IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 3, 2017, accepted March 28, 2017, date of publication April 4, 2017, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2690455

Depth Map Reconstruction for Underwater Kinect Camera Using Inpainting and Local Image Mode Filtering

HUIMIN LU[1], YIN ZHANG[2], YUJIE LI[3], QUAN ZHOU[4,5], (Member, IEEE), RYUNOSUKE TADOH[1], TOMOKI UEMURA[1], HYOUNGSEOP KIM[1], AND SEIICHI SERIKAWA[1]

[1]Kyushu Institute of Technology, Kitakyushu 8048550, Japan
[2]Zhongnan University of Economics and Law, Wuhan 430073, China
[3]Yangzhou University, Yangzhou 225127, China
[4]Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[5]Nanjing University of Information Science and Technology, Nanjing 210044, China

Corresponding author: Huimin Lu (dr.huimin.lu@ieee.org)

This work was supported by Leading Initiative for Excellent Young Researcher (LEADER) of Ministry of Education, Culture, Sports, Science and Technology-Japan (16809746), Grants-in-Aid for Scientific Research of JSPS (17K14694), Research Fund of CAS (MGE2015KG02), Research Fund of SKL of Marine Geology in Tongji University (MGK1608), Research Fund of SKL of Ocean Engineering in Shanghai Jiaotong University (1315;1510), Research Fund of The Telecommunications Advancement Foundation and Fundamental Research Developing Association for Shipbuilding and Offshore, NSFC (61401228), China Postdoctoral Science Foundation (2015M581841), Postdoctoral Science Foundation of Jiangsu Province (1501019A), PAPD of Jiangsu Higher Education Institutions, and Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology.

ABSTRACT Underwater optical cameras are widely used for security monitoring in ocean, such as earthquake prediction and tsunami alarming. Optical cameras recognize objects for autonomous underwater vehicles and provide security protection for sea-floor networks. However, there are many issues for underwater optical imaging, such as forward and backward scattering, light absorption, and sea snow. Many underwater image processing techniques have been proposed to overcome these issues. Among these techniques, the depth map gives important information for many applications of the post-processing. In this paper, we propose a Kinect-based underwater depth map estimation method that uses a captured coarse depth map by Kinect with the loss of depth information. To overcome the drawbacks of low accuracy of coarse depth maps, we propose a corresponding reconstruction architecture that uses the underwater dual channels prior dehazing model, weighted enhanced image mode filtering, and inpainting. Our proposed method considers the influence of mud sediments in water and performs better than the traditional methods. The experimental results demonstrated that, after inpainting, dehazing, and interpolation, our proposed method can create high-accuracy depth maps.

INDEX TERMS Underwater Kinect camera, depth map, inpainting, local mode filtering.

I. INTRODUCTION

The ocean contains a large amount of resources, which will be vital for the future of human life. Hydrothermal deposits contain elements such as gold, silver, zinc, and lead, and exist at about 200 locations around Japan. These deposits have been estimated to be 750 million tons, with an estimated 450 million tons in Japan's Exclusive Economic Zone. There is estimated to be about 50,000 km^2 of cobalt-rich crust, amounting to 2.4 billion tons. Furthermore, there is estimated 50,000 km^2 of methane hydrates with a volume of 12.6×10^{12} m^3. This amount of minerals could last approximately 100 years in Japan [1].

All of these minerals need to be developed and ocean mining machine monitoring systems are very important. Monitoring system need to track the position of terrain around mining machines and resources on the seabed. Sonar, laser, and optical cameras are typically used to monitor underwater terrain. In recent years, sonar has been widely used to map the terrain of the ocean. Sonar imaging [2] has many benefits, such as allowing for long-range imaging and being robust for turbid water. However, it is not suitable for short-range imaging because it has low resolution and limited color information. Laser imaging [3] is also widely used in underwater scientific research. However, it relies heavily on system configuration and is hard to equip in mining machines. Therefore, optical cameras [4] may be the only suitable option for underwater short-range imaging equipment. However, it is difficult to use the camera for security monitoring during

the machine mining systems operation [53], [54]. We briefly introduce the above three types of underwater imaging in the following.

A. SONAR IMAGING SYSTEM

Sonograms [2] are formed by sonar imaging devices, which emit beam pulses to the sea floor. Sonar beams are narrow at the sending device and are wide when they reach their targets. Sonar imaging devices receive acoustic reflections from the ocean floor. As the sonar beam moves, reflections depict a series of lines in the direction of movement. Then the device stitches beam lines to form a single image of the targeted objects. This sonar imaging process is time consuming and only distance information is used for imaging, which causes the loss of 3D information of the object.

B. LASER IMAGING SYSTEM

Range-gated or time-gated laser imaging is a method used to improve image quality and visibility in turbid conditions. In range-gated laser imaging systems, the camera is adjacent to the light source, whereas the underwater target is behind the scattering medium [5]. Range-gated systems select reflected light from the object arriving at the camera and block optical back-scatter light [6]. Range-gated systems include a broad-beam pulse illumination source, a high speed camera, and a synchronization gate duration control [6]. Tan *et al.* [5] presented a sample plot of range-gated imaging timing. Laser imaging systems capture grayscale images only, as the color information is missed. Also, lasers are affected by strong scattering and absorption in highly turbid water.

C. OPTICAL CAMERA SYSTEM

Optical imaging sensors [4] can provide a great deal of high speed information. They are commonly used in many terrestrial and airborne robotic applications. However, the interaction between electromagnetic waves and water means that optical imaging and vision systems need to be specifically designed for use in underwater environment.

Underwater images have specific characteristics that should be considered during collection and processing. Light attenuation and scattering, non-uniform lighting, shadows, color distortion, and suspended particles or abundance of marine life near the target are frequently found in typical underwater scenes.

One effect of the inherent optical properties of the ocean is that it becomes darker with depth because as water depth increases, the sunlight is absorbed and scattered. For example, euphotic depth is less than 200 meters in clean ocean water. In addition, the spectral composition of sunlight also changes with water depth. Long wavelengths (red color) have larger absorption than short (blue or green color) wavelengths. Therefore, most underwater images taken under natural light (sunlight) will appear blue or green. Consequently, additional illumination is required in deep or turbid water.

Above all, optical cameras have various features and advantages. However, it is difficult to recover time variations adverse to visibility in the presence of floating turbidity sediments. In this paper, we propose a Kinect-based underwater optical imaging method. The underwater dark channel used prior to descattering. At the same time, we use inpainting to refine depth maps. Finally, depth maps are super-resolved using weight enhanced image mode filtering.

II. RELATED WORK

The challenge of acquiring high-quality depth information in real-world environment is an important computer vision task. Depth information or depth map estimation methods can be categorized as active or passive. Passive methods use cameras to capture images and learning algorithms to recover visual depth information [7]. Passive methods suffer from some issues for practical applications. For example, they require strict image reflection and are inefficient for textureless objects. On the other hand, active approaches, such as structured light sensors and time-of-flight (ToF) cameras are subject to errors, such as noise, ambiguity, scattering, and motion blurring [8]. In the following, we give a survey of the most related work.

A. PASSIVE METHODS

1) SINGLE MONOCULAR IMAGE-BASED DEPTH ESTIMATION

Traditional single image-based methods use Markov random field (MRF) techniques to learn depth information. Saxena *et al.* [9] used discriminatively trained global and local image features to model individual depth points and the relationship between depths and neighbors. Liu *et al.* [10] performed a semantic segmentation of images and used semantic labels to determine the depth. Girshick *et al.* [11] presented a regions with convolutional neural network (R-CNN) features method for estimating depth information. Liu *et al.* [12] proposed a deep continuous conditional random fields (CRF) model to combine adjacent superpixels. Inspired by [12], Liu *et al.* [13] presented a deep convolutional neural field model for estimating the depth map by considering the capacities of CNN and CRF. He *et al.* [14] analyzed numerous natural sky images and determined that most color images contain a dark channel. Based on this finding, they proposed using a dark channel for depth estimation. However, most single image-based depth estimation methods are time consuming.

2) STEREO IMAGE-BASED DEPTH ESTIMATION

Many stereo image-based depth estimation methods have been proposed over the past few decades. Stereo image-based depth estimation algorithms can be generally grouped into either global or local approaches. Global-based approaches treat depth assignment as a problem of minimizing the global energy function for all disparity values. Salmen *et al.* [15] introduced a modified dynamic programing algorithm for calculating depth maps from stereo images. Perez and Sanchez [16] proposed an extended belief propagation algorithm that had good immersive feeling. Wang *et al.* [17]

proposed a hierarchical bilateral disparity structure algorithm for improving graph cuts without any loss to depth maps.

In contrast, local-based approaches are also known as window-based approaches. They calculate depth at a given pixel within a predefined support window. These methods include matching cost computation, cost aggregation, disparity selection, and disparity refinement steps. Winner take all (WTA) optimization is usually used for disparity selection. Lee *et al.* [18] used a three-mode cross census for WTA optimization. Matsuo *et al.* [19] proposed a local-based approach that used absolute differences and the Sobel operator in matching cost, and then used box filtering with WTA and joint bilateral filters for refinement. Werner *et al.* [20] used a normalized cross correlation algorithm to compute matching cost, WTA optimization for disparity selection, and cross bilateral filters for refinement.

3) MULTIPLE IMAGES-BASED DEPTH ESTIMATION

Multiple-image-based depth estimation is widely used in photometric stereo or other applications. Tsiotsios *et al.* [21] proposed an occlusion-aware depth estimation algorithm for light-field camera-based multiple image depth estimation. This method can capture sharp object boundaries correctly. Wang *et al.* [22] proposed a taking a three-lighted image from turbid water and recovering the scene as well as removing scattering.

B. ACTIVE METHODS

There are generally three types of depth cameras: LiDAR cameras, structured-light based cameras (e.g., Kinect), and ToF cameras. We review these cameras and related depth refinement methods.

1) LiDAR

Furukawa and Ponce [23] proposed an algorithm for calibrating multi-view stereopsis and outputting high-quality depth maps. Strecha *et al.* [24] proposed a partial differential equation-based formulation to refine high- and low-resolution LiDAR depth information. Chen *et al.* [25] introduced a self-adaptive method to upsample LiDAR point-projected sparse depth maps. Bevilacqua *et al.* [26] used a primal-dual optimization method to estimate a complete depth map using image gradients and visibility indicators.

2) KINECT DEPTH CAMERAS

A depth map captured by Kinect has random depth information missing because of occlusions. Berdnikov and Vatolin [27] proposed a deepest neighbor method and a spatial interpolation method to solve the problem of missing depth information due to occlusions. Maimone and Fuchs [28] investigated a two-pass median filter with a non-fixed window to fill missing depth. Matyunin *et al.* [29] proposed a simple temporal filtering algorithm for depth restoration. Dakkak and Husain [30] devised an iterative diffusion method to merge the RGBD segmented regions of the depth map. Camplani and Salgado [31] proposed using a joint bilateral filter to remove noise and recover depth values in the temporal domain. Yu *et al.* [32] introduced a shape-from-shading method to refine noisy depth maps. Lee and Ho [33] proposed an image inpainting method to fill missing depth map information. Yang *et al.* [34] used an adaptive autoregressive model to recover colored depth maps. Li *et al.* [35] proposed a fast global smoothing technique for global depth map interpolation.

3) ToF DEPTH CAMERAS

ToF camera-based depth map refinement methods can be divided into two categories: MRF-based methods and advanced filtering methods. Diebel and Thrun [36] used a two-layer MRF method to model the correlation of range measurements and refine the depth map using the conjugate gradient method. Huhle *et al.* [37] proposed a third-layer MRF to improve the depth map. Garro *et al.* [38] proposed a graph-based segmentation method to interpolate missing depth information. Yang *et al.* [39] proposed a hierarchical joint bilateral filtering method for depth map upsampling. Zhu *et al.* [40] designed a dynamic MRF for improving depth map spatial and temporal accuracy. He *et al.* [41] proposed a guided image filter to smooth depth information with a reference image. Lu *et al.* [42] incorporated amplitude values of a ToF camera and designed a data term to fit the characteristics of a depth map. Park *et al.* [43] considered the use of a non-local term and weighted color image filtering to recover depth information. Aodha *et al.* [44] used a generic database of local patches to increase depth map resolution. Min *et al.* [45] proposed a weighted mode filtering method to modify a depth map histogram. Lu *et al.* [46] proposed a local multipoint regression method to estimate the depth information of each point. Ferstl *et al.* [47] used an anisotropic diffusion tensor obtained from a reference image. Liu *et al.* [48] investigated geodesic distance to compute the filtering coefficient of similar pixels.

III. DEPTH MAP ENHANCEMENT ARCHITECTURE

In this paper, we propose the following pipeline for underwater depth map refinement shown in Figure 1. We remove color image and inpainting processing scattering from the depth maps. So, we use the processed color image and depth map to create an unsampling depth map. We discuss the steps of our proposed method in detail in the following.

A. GUIDED IMAGE DESCATTERING

Traditional guided-image depth map refinement requires high-quality images, but underwater images are highly distorted by turbidity. To this end, we propose an underwater imaging model that follows a standard attenuation model to accommodate wavelength attenuation coefficients. In this paper, we adopted a modified Jaffe-McGlamery model, [49] which has been estimated as a description of water's absorption effects on an observer. For underwater imaging, observed irradiance is linear and is attenuated in the route of sight and by scattered ambient light as depicted in Figure 1.

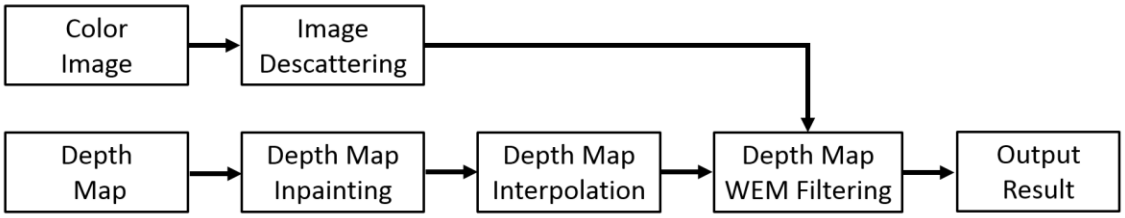


FIGURE 1. Pipeline of proposed approach for obtaining a noise-free depth map.

Therefore, a modified Jaffe-McGlamery model suits underwater lighting conditions.

The modified Jaffe-McGlamery model can be written as

$$Y(i) = X(i)e^{-kd(i)} + \rho \cdot X(i)(1 - e^{-kd(i)}), \tag{1}$$

where $X(i)$ is the real scene at pixel i of the related captured image Y, ρ is the normalized radiance of a scene point, d is the distance from the scene point to the camera, and k is the total beam attenuation coefficient, which is nonlinear and wavelength-dependent. In most highly turbid water, the value of k is usually chosen to be 0.78.

During our previous experiments, we found that the lowest pixel value of RGB channels in turbid water is not always the red color channel. Sometimes the blue color channel is the lowest. Consequently, we only need to take the dual-channel (red and blue) to compute a coarse depth map.

Our method is based on [50], where the depth map is initialized using an underwater median dual channel prior (UMDCP). As mentioned before, we can estimate the coarse depth map $\tilde{d}(x)$ as:

$$\tilde{d}(i) = \underset{\Omega(m,n)}{median}\left(\min_{c\in\{r,b\}} \frac{Y_c(i)}{\rho X(i)}\right) \tag{2}$$

where Ω is a 7×7 pixels sliding window. For each pixel located at (m, n) of the sliding window Ω, the lower value from the red and blue color channels is chosen. The proposed method can prevent a halo effect around occlusion boundaries.

Accordingly, the refined coarse depth map is obtained by

$$d(i) = 1 - \tau\tilde{d}(i) \tag{3}$$

where $\tau = 0.98$ for most scenes. Then, an enhanced image can be obtained by:

$$X(i) = \frac{Y(i) - \rho X(i)}{\max\left(d(i), d_0\right)} + \rho X(i) \tag{4}$$

where d_0 usually is equal to 0.1.

B. COARSE DEPTH MAP INPAINTING

During our experiments, we found that there are some occurrences of missing depth information in images obtained by Kinect. Image inpainting [51] as a powerful tool for filling the lost information. In this section, we explain the mathematical model of inpainting method in this article. Given a small enough missing area ε, the first order approximation $d_q(p)$ of the inpainted coarse depth map in pixel p, the coarse depth map $d(q)$, and gradient $\nabla d(q)$ values of pixel q, the inpainting model can be expressed as:

$$d_q(p) = d(q) + \nabla d(q)(p - q) \tag{5}$$

Next, we inpaint the pixel p as a function of all pixels q in the neighborhood $B_\varepsilon(p)$ by summing the estimates of all pixels q. We calculated the inpainted coarse depth map by adding a normalized weighting function $\omega(p, q)$ as:

$$d(p) = \frac{\sum_{q\in B_\varepsilon(p)} \omega(p, q)[d(q) + \nabla d(q)(p - q)]}{\sum_{q\in B_\varepsilon(p)} \omega(p, q)} \tag{6}$$

The weighted function $\omega(p, q)$ is defined as:

$$dir(p, q) = \frac{p - q}{||p - q||} \cdot N(p) \tag{7}$$

$$dst(p, q) = \frac{d_0^2}{||p - q||^2} \tag{8}$$

$$lev(p, q) = \frac{T_0}{1 + ||T(p) - T(q)||} \tag{9}$$

$$\omega(p, q) = dir(p, q) \cdot dst(p, q) \cdot lev(p, q) \tag{10}$$

where T is the distance map of the pixels in the region to be inpainted. $N(p)$ is the gradient T at the pixel p. The terms dst and lev measure reference distances d_0 and T_0. In practice, d_0 and T_0 are usually set to 1. To iteratively apply Eq. (6) to all the losing pixels in the coarse depth map.

C. FINE DEPTH MAP ESTIAMTION

The coarse depth map is repaired after inpainting; however, the small-scaled coarse depth map needs to be enlarged to match the high-resolution color image. In this paper, we propose a weighted enhanced image mode filter to refine the enlarged depth map after interpolation. The relaxed histogram $H(p,j)$ at reference pixel p and the j-th bin can be defined as:

$$H(p, j) = \sum_{q\in M(p)} G_r(j - d(q)) \tag{11}$$

where G_r is a Gaussian function and $M(p)$ is the neighboring pixels around the pixel p. The local histogram $H_L(p,j)$ is defined as:

$$H_L(p, j) = \sum_{q\in M(p)} G_r(j - d(q))G_S(p - q) \tag{12}$$

FIGURE 2. Side-view of the experimental setting.

(a)

(b)

FIGURE 3. Experimental result of scatter removal. (a) Input image; (b) Image after descattering.

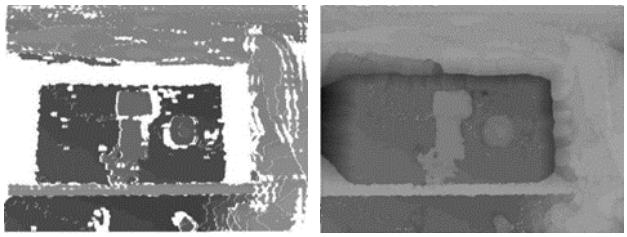
where G_S is the spatial Gaussian function. Finally, the global mode of the histogram $H_G(p,j)$ for the weighted enhanced image mode is generated by

$$H_G(p,j) = \sum_{q \in M(p)} G_I(X(p) - X(q)) G_S(p-q) G_r(j - d(q)) \tag{13}$$

where G_I is a weighted function of enhanced image $X(p)$ different from the coarse interpolated depth map $d(p)$ to be filtered.

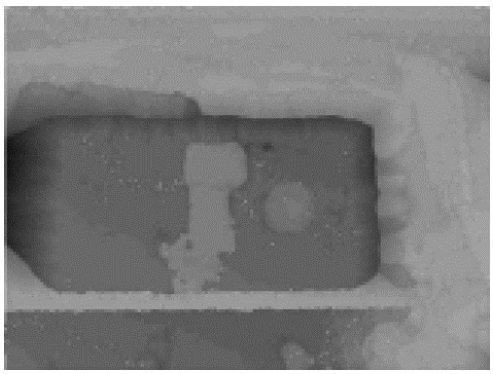
IV. EXPERIMENTAL RESULTS

In this paper, we used Kinect for Windows [52] and took images of an object in a water tank. The equipment setting

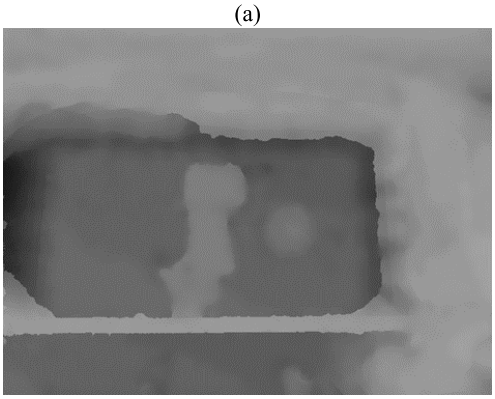


(a) (b)

FIGURE 4. Experimental result of inpainting. (a) Captured depth map; (b) After inpainting.



(a)



(b)

FIGURE 5. Comparison of the experimental results. (a) Interpolation; (b) Proposed method.

is shown in Figure 2. We set the Kinect on the top of the water tank. There were two bricks, a milk bottle and a Chinese-style cup lid on the water tank. In the experiment, we added a lot of deep-sea soil to the water tank. As shown in Figure 3(a), the captured color image contains many noise and scatter. The result of after descattering is shown in Figure 3(b), and the inpainting result is shown in Figure 4(b). From Figure 4(a), we found that the captured depth map contains depth information missing. The boundaries of two bricks and Chinese-style cup are missing. Hence, inpainting was adopted into these coarse depth map to complete the depth information. Next, the weighted enhanced image mode filtering was used to upsample the final depth map in Figure 5. Other depth reconstruction experiments are shown in Figure 6. We can conclude that the proposed method performs well to recover the color image and the depth map.

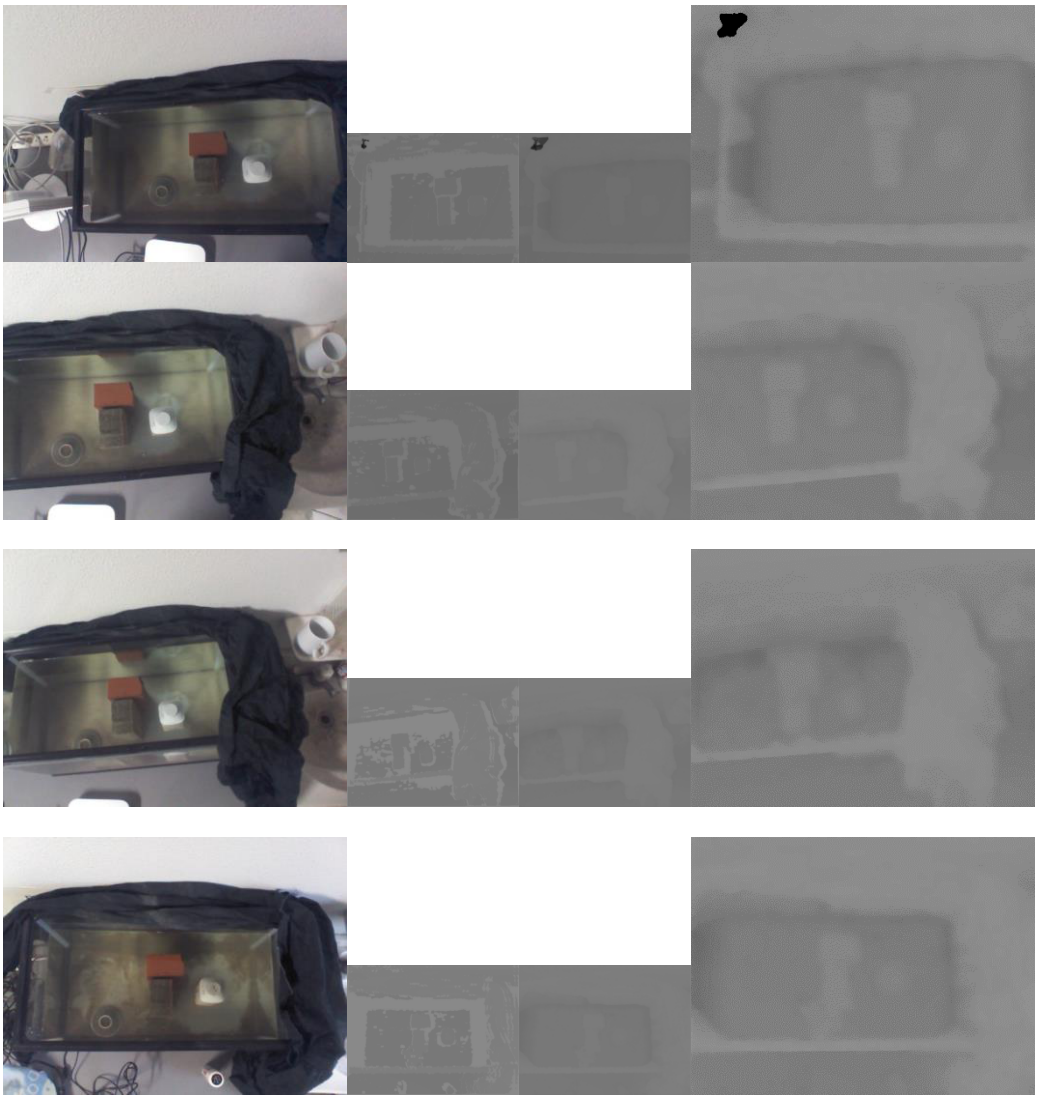


FIGURE 6. Experimental results of underwater depth map reconstruction.

V. CONCLUSIONS

In this paper, we have solved the issue of the occlusion underwater depth map using Kinect. The issue is caused by relative displacement of the projector and camera. We considered an approach different from traditional depth recovery methods to recover the high-resolution depth map using images captured in turbid water. Experimental results in real-world scenes show that our method outperforms existing conventional methods, and is suitable for restoring underwater image depth information. In the future, we will do a three-dimensional restoration of the sea floor as well as descattering of heavily scattered images.

REFERENCES

[1] *Basic Plan on Ocean Policy*, accessed on Jan. 20, 2017. [Online]. Available: http://www.kantei.go.jp/jp/singi/kaiyou/kihonkeikaku/130426kihonkeikaku_e.pdf

[2] J. F. Fish, C. S. Johnson, and D. K. Ljungblad, "Sonar target discrimination by instrumented human divers," *J. Acoust. Soc. Amer.*, vol. 59, no. 3, pp. 602–606, 1976.

[3] D. M. Kocak and F. M. Caimi, "The current art of underwater imaging—With a glimpse of the past and vision of the future," *Marine Technol. Soc. J.*, vol. 39, no. 3, pp. 5–26, 2005.

[4] F. Bonin, A. Burguera, and G. Oliver, "Imaging systems for advanced underwater vehicles," *J. Maritime Res.*, vol. 8, no. 1, pp. 65–86, 2011.

[5] C. S. Tan, A. L. Sluzek, G. Seet, and T. Y. Jiang, "Range gated imaging system for underwater robotic vehicle," in *Proc. IEEE Int. Symp. Appl. Ferroelectr.*, May 2007, pp. 1–6.

[6] C. Tan, G. Seet, G. Sluzek, and D. He, "A novel application of range-gated underwater laser imaging system (ULIS) in near-target turbid medium," *Opt. Laser Eng.*, vol. 43, no. 9, pp. 995–1009, Sep. 2005.

[7] B. Bodkin, *Real-Time Mobile Stereo Vision*. Knoxville, TN, USA: Univ. Tennessee, 2012.

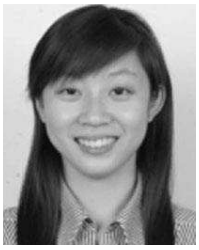
[8] R. A. Hamzah and H. Ibrahim, "Literature survey on stereo vision disparity map algorithms," *J. Sensors*, vol. 2016, 2016, Art. no. 8742920.

[9] A. Saxena, S. Chung, and A. Ng, "Learning depth from single monocular images," in *Proc. NIPS*, 2005, pp. 1161–1168.

[10] B. Liu, S. Gould, and D. Koller, "Single image depth estimation from predicted semantic labels," in *Proc. CVPR*, Jun. 2010, pp. 1253–1260.
[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. CVPR*, 2014, pp. 580–587.
[12] M. Liu, M. Salzmann, and X. He, "Discrete-continuous depth estimation from a single image," in *Proc. CVPR*, Jun. 2014, pp. 716–723.
[13] F. Liu, C. Shen, G. Lin, and I. Reid, "Learning depth from single monocular images using deep convolutional neural fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 10, pp. 2024–2039, Oct. 2016.
[14] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
[15] J. Salmen, M. Schlipsing, J. Edelbrunner, S. Hegemann, and S. Lüke, "Real-time stereo vision: Making more out of dynamic programming," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, 2009, pp. 1096–1103.
[16] J. M. Pérez and P. Sánchez, "Real-time stereo matching using memory-efficient Belief Propagation for high-definition 3D telepresence systems," *Pattern Recognit. Lett.*, vol. 32, no. 16, pp. 2250–2253, Dec. 2011.
[17] Y.-C. Wang, C.-P. Tung, and P.-C. Chung, "Efficient disparity estimation using hierarchical bilateral disparity structure based graph cut algorithm with a foreground boundary refinement mechanism," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 5, pp. 784–801, May 2013.
[18] Z. Lee, J. Juang, and T. Q. Nguyen, "Local disparity estimation with three-moded cross census and advanced support weight," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 1855–1864, Dec. 2013.
[19] T. Matsuo, S. Fujita, N. Fukushima, and Y. Ishibashi, "Efficient edge-awareness propagation via single-map filtering for edge-preserving stereo matching," *Proc. SPIE*, vol. 9393, pp. 93930S-1–93930S-10, Mar. 2015.
[20] M. Werner, B. Stabernack, and C. Riechert, "Hardware implementation of a full HD real-time disparity estimation algorithm," *IEEE Trans. Consum. Electron.*, vol. 60, no. 1, pp. 66–73, Feb. 2014.
[21] C. Tsiotsios, M. E. Angelopoulou, T.-K. Kim, and A. J. Davison, "Backscatter compensated photometric stereo with 3 sources," in *Proc. CVPR*, 2014, pp. 2251–2258.
[22] T.-C. Wang, A. A. Efros, and R. Ramamoorthi, "Depth estimation with occlusion modeling using light-field cameras," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2170–2181, Nov. 2016.
[23] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multi-view stereopsis," in *Proc. CVPR*, Jun. 2007, pp. 1–8.
[24] C. Strecha, R. Fransens, and L. Van Gool, "Combined depth and outlier estimation in multi-view stereo," in *Proc. CVPR*, Jun. 2006, pp. 2394–2401.
[25] L. Chen, Y. He, J. Chen, Q. Li, and Q. Zou, "Transforming a 3-D LiDAR point cloud into a 2-D dense depth map through a parameter self-adaptive framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 165–176, Jan. 2017.
[26] M. Bevilacqua, J.-F. Aujol, M. Brédif, and A. Bugeau, "Visibility estimation and joint inpainting of LiDAR depth maps," in *Proc. ICIP*, Sep. 2016, pp. 1–5.
[27] Y. Berdnikov and D. Vatolin, "Real-time depth map occlusion filling and scene background restoration for projected-pattern based depth cameras," in *Proc. IETP Graph. Conf.*, 2011, pp. 121–126.
[28] A. Maimone and H. Fuchs, "Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras," in *Proc. 10th IEEE ISMAR*, Oct. 2011, pp. 137–146.
[29] S. Matyunin, D. Vatolin, Y. Berdnikov, and M. Smirnov, "Temporal filtering for depth maps generated by Kinect depth camera," in *Proc. 3DTV*, May 2011, pp. 1–4.
[30] A. Dakkak and A. Husain, "Recovering missing depth information from Microsoft's Kinect," in *Proc. Embedded Vis. Alliance*, 2012, pp. 1–9.
[31] M. Camplani and L. Salgado, "Adaptive spatio-temporal filter for low-cost camera depth maps," in *Proc. Int. Conf. Emergence Signal Process. Appl.*, Jan. 2012, pp. 33–36.
[32] L.-F. Yu, S.-K. Yeung, Y.-W. Tai, and S. Lin, "Shading-based shape refinement of RGB-D images," in *Proc. CVPR*, 2013, pp. 1415–1422.
[33] S.-B. Lee and Y.-S. Ho, "Joint multilateral filtering for stereo image generation using depth camera," *Era Interact. Media*, pp. 373–383, 2013.
[34] J. Yang, X. Ye, K. Li, C. Hou, and Y. Wang, "Color-guided depth recovery from RGB-D data using an adaptive autoregressive model," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3443–3458, Aug. 2014.
[35] Y. Li, D. Min, M. N. Do, and J. Lu, "Fast guided global interpolation for depth and motion," in *Proc. ECC*, 2016, pp. 717–733.
[36] J. Diebel and S. Thrun, "An application of Markov random fields to range sensing," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 18. 2005, p. 291.
[37] B. Huhle, S. Fleck, and A. Schilling, "Integrating 3D time-of-flight camera data and high resolution images for 3DTV applications," in *Proc. 3DTV*, May 2007, pp. 1–4.
[38] V. Garro, P. Zanuttigh, and G. M. Cortelazzo, "A new super resolution technique for range data," in *Proc. Associazione Gruppo Telecomun. Tecnol. Informazione*, 2009, pp. 1–7.
[39] Q. Yang, K.-H. Tan, B. Culbertson, and J. Apostolopoulos, "Fusion of active and passive sensors for fast 3D capture," in *Proc. IEEE Int. Workshop Multimedia Signal Process.*, Oct. 2010, pp. 69–74.
[40] J. Zhu, L. Wang, J. Gao, and R. Yang, "Spatial-temporal fusion for high accuracy depth maps using dynamic MRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 899–909, May 2010.
[41] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. ECCV*, 2010, pp. 1–14.
[42] J. Lu, D. Min, R. Pahwa, and M. N. Do, "A revisit to MRF-based depth map super-resolution and enhancement," in *Proc. IEEE ICASSP*, May 2011, pp. 985–988.
[43] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon, "High quality depth map upsampling for 3D-TOF cameras," in *Proc. ICCV*, Nov. 2011, pp. 1623–1630.
[44] O. M. Aodha, N. D. F. Campbell, A. Nair, and G. J. Brostow, "Patch based synthesis for single depth image super-resolution," in *Proc. ECCV*, 2012, pp. 71–84.
[45] D. Min, J. Lu, and M. N. Do, "Depth video enhancement based on weighted mode filtering," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1176–1190, Mar. 2012.
[46] J. Lu, K. Shi, D. Min, L. Lin, and M. N. Do, "Cross-based local multipoint filtering," in *Proc. CVPR*, Jun. 2012, pp. 430–437.
[47] D. Ferstl, C. Reinbacher, R. Ranftl, M. Ruether, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation," in *Proc. ICCV*, 2013, pp. 993–1000.
[48] M.-Y. Liu, O. Tuzel, and Y. Taguchi, "Joint geodesic upsampling of depth images," in *Proc. CVPR*, 2013, pp. 169–176.
[49] H. Lu, Y. Li, L. Zhang, and S. Serikawa, "Contrast enhancement for images in turbid water," *J. Opt. Soc. Amer.*, vol. 32, no. 5, pp. 886–893, 2015.
[50] H. Lu, Y. Li, and S. Serikawa, "Single underwater image descattering and color correction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2015, pp. 1623–1627.
[51] A. Telea, "An image inpainting technique based on the fast marching method," *J. Graph. Tools*, vol. 9, no. 1, pp. 23–34, 2004.
[52] *Kinect for Windows*, accessed on Jan. 20, 2017. [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows
[53] H. Abbas, M. Q. Mahmoodzadeh, F. A. Khan, and M. Pasha, "Identifying an OpenID anti-phishing scheme for cyberspace," *Secur. Commun. Netw.*, vol. 9, no. 6, pp. 481–491, 2016.
[54] H. Abbas, C. Magnusson, L. Yngström, and A. Hemani, "Addressing dynamic issues in information security management," *Inf. Manage. Comput. Secur.*, vol. 19, no. 1, pp. 5–24, 2011.

HUIMIN LU received the B.S. degree in electronics information science and technology from Yangzhou University in 2008, the M.S. degrees in electrical engineering from the Kyushu Institute of Technology and Yangzhou University in 2011, and the Ph.D. degree in electrical engineering from the Kyushu Institute of Technology in 2014. From 2013 to 2016, he was a JSPS Research Fellow with the Kyushu Institute of Technology, where he is currently an Assistant Professor. His research interests include computer vision, robotics, artificial intelligence, and ocean observing.

YIN ZHANG was a Post-Doctoral Fellow with the School of Computer Science and Technology, Huazhong University of Science and Technology, China. He is currently an Assistant Professor with the School of Information and Safety Engineering, Zhongnan University of Economics and Law, China. He has published over 30 prestigious conference and journal papers. His research interests include data analysis, data mining, healthcare big data, and social network. He is the Vice-Chair of the IEEE Computer Society Big Data STC. He also served as the TPC Co-Chair of CloudComp 2015, and the Local Chair of TRIDENTCOM 2014. He serves as a Guest Editor of the IEEE Sensors Journal and *New Review of Hypermedia and Multimedia*.



YUJIE LI received the B.S. degree in computer science and technology from Yangzhou University in 2009, the M.S. degrees in electrical engineering from the Kyushu Institute of Technology and Yangzhou University in 2012, and the Ph.D. degree from the Kyushu Institute of Technology in 2015. She is currently a Lecturer with Yangzhou University. Her research interests include computer vision, sensors, and image segmentation.

QUAN ZHOU (M'13) received the M.S. and Ph.D. degrees in communication and information system from the Huazhong University of Science and Technology, China, in 2006 and 2013, respectively. He is currently an Associate Professor with the Nanjing University of Posts and Telecommunications. He has published over 20 research papers in SCI journals, such as the IEEE Transactions on Image Processing, the IEEE Transactions on Multimedia, and *Pattern Recognition*, and conferences, such as ICIP, ICASSP, ACCV, and ICPR, in image processing and computer vision. His research interests include computer vision and pattern recognition. He now serves as a TPC Member or the Chair of many international conferences and a Reviewer for a series of SCI journals, including the IEEE Transactions on Image Processing, the IEEE Transactions on Multimedia, the IEEE Transactions on Circuits System for Video Technology, *Pattern Recognition*, and *Neurocomputing*.

RYUNOSUKE TADOH received the B.S. degree in electronics and electrical engineering from the Kyushu Institute of Technology in 2017, where he is currently pursuing the master's degree. His research interests include computer vision, robotics, and artificial intelligence.

TOMOKI UEMURA received the B.S. degree in mechanical and control engineering from the Kyushu Institute of Technology in 2016, where he is currently pursuing the master's degree. His research interests include computer vision, medical imaging, and artificial intelligence.

HYOUNGSEOP KIM received the B.A. degree in electrical engineering and the master's and Ph.D. degrees from the Kyushu Institute of Technology in 1994, 1996, and 2001, respectively. He is currently a Professor with the Department of Control Engineering, Kyushu Institute of Technology. His research interests are focused on medical application of image analysis.

SEIICHI SERIKAWA received the B.S. and M.S. degrees from Kumamoto University in 1984 and 1986, respectively, and the Ph.D. degree from the Kyushu Institute of Technology in 1994, all in electronic engineering. He is currently a Vice President with the Kyushu Institute of Technology and also serves as a Professor with the Center for Socio-Robotic Synthesis and the Department of Electrical and Electronic Engineering. His current research interests include computer vision, sensors, and robotics.