IPSJ Transactions on Computer Vision and Applications Vol.6 88–92 (July 2014)

[DOI: 10.2197/ipsjtcva.6.88]

Express Paper

Depth from Projector's Defocus Based on Multiple Focus Pattern Projection

Hitoshi Masuyama[1,a) Hiroshi Kawasaki[1,b) Ryo Furukawa[2,c)

Received: March 14, 2014, Accepted: April 24, 2014, Released: July 25, 2014

Abstract: For 3D active measurement methods using video projector, there is the implicit limitation that the projected patterns must be in focus on the target object. Such limitation set a severe constraints on possible range of the depth for reconstruction. In order to overcome the problem, Depth from Defocus (DfD) method using multiple patterns with different in-focus depth is proposed to expand the depth range in the paper. With the method, not only the range of the depth is extended, but also the shape can be recovered even if there is an obstacle between the projector and the target, because of the large aperture of the projector. Furthermore, thanks to the advantage of DfD which does not require baseline between the cameras and the projector, occlusion does not occur with the method. In order to verify the effectiveness of the method, several experiments using the actual system was conducted to estimate the depth of several objects.

Keywords: Depth from Defocus (DfD), multiple projection patterns, projector-camera system

1. Introduction

Various active 3D shape measurement methods have been researched to measure 3D shape with single shot which is suitable for capturing moving objects. As a representative method of active 3D measurement method, two dimensional pattern projector and a camera based system is known. This method projects unique pattern to the object, then captures the scene with a camera to finally obtain correspondences. The correspondences are then analyzed for depth estimation using triangulation method. In order to obtain wide measurement range of depth, methods using laser have been proposed [5].

Since the narrow and straight nature of a laser beam, it makes possible to achieve wider range of depth on measurement. However, if there are obstacles between the target object and the laser source, it is easily affected by them, thus resulting in unstable and sparse measurement. On the other hand, since a projector based 3D scanning system (i.e., the projector with a point light source), which is also a major active 3D measurement method [11], is usually equipped with a large aperture. Therefore, the emitted light from the system diffracts around the obstacles and enables to achieve measurements without being affected by presence of obstacles. However, the measurement range of the system is narrow and has less depth of field, because of the large aperture.

In order to increase the depth of field on projector based system, Depth from Defocus (DfD) based method, which estimates the depth from defocus blur on the projected pattern, has been proposed [2]. This method estimates the depth value by using the point light source with a coded aperture installed in the projector, and analyzes the size and shape of the blur which is projected by the projector. However, the expansion of the measurement range of this method is limited, because possible blur size for depth estimation is not large.

To overcome the above problem, we propose a DfD based method to expand the measurement range of depth by using multiple patterns for projector with various in-focus distances. The proposed method makes use of in-focus distances of each projected pattern which is generated by transparent glass with multiple types of line pattern placed between the light source and the lens. The depth value of the projected pattern is then estimated by identifying the in-focus pattern on the target object in the captured image. Since the large aperture can be used to this type of projector, 3D measurement even in presence of an obstacle between the projector and the target is possible. In addition, by using multiple patterns, 3D measurement range of depth becomes longer than a general projector. In this paper, we discuss the proposed method in detail as well as showing its effectiveness to obtain depth value in the real experiments.

2. Related Work

DfD is generally known as a method based on defocus blur of *camera*, which estimates the depth value from single image [9]. However, in order to achieve successful DfD result, it requires high-frequency texture on the measurement target thus limits applicable scenes. Nayar et al. proposed a DfD method based on defocus blur of the *projected light pattern* [8]. It projects a number of points and DfD is done from the blur of them, which does not require texture on the target. Their research does not focus on acquiring the depth value, but rather on refocusing the captured image using the result obtained from DfD as a reference map. For this reason, light points are positioned sparsely, where each of them is separated distantly, making it difficult to directly be used for shape measurement.

Zhang et al. analyzed the defocus blur of the projected image

1 Faculty of Engineering, Kagoshima University, Kagoshima 890–0065, Japan
2 Faculty of Information Sciences, Hiroshima City University, Hiroshima 731–3194, Japan
a) sc109064@ibe.kagoshima-u.ac.jp
b) kawasaki@ibe.kagoshima-u.ac.jp
c) ryo-f@hiroshima-cu.ac.jp

from a projector, which measures depth information on the projected surface [12]. They used half-mirrors and placed a projector and a camera on the same axis, allowing having successful depth measurement. By using coaxial projector-camera system, it allows to calculate depth value for each pixel without any missing parts. However, this is ineffective for capturing a moving object because of time consuming process for scan, since it requires 24 images to illuminate the entire scene by shifting the pattern.

Recently, in computational photography field, various new methods analyzing camera blur and light reflectance and their practical applications have been researched [4], [7], [9], [10]. For example, by changing the shape of the aperture of the camera from circular to unique shape and analyzing the degree of blur, the methods to realize efficient deblurring or depth estimate have been proposed. However, there still are only few researches that utilize computational photography technique on the projector. Grosse et al. proposed a system that uses programmable coded aperture on a video projector [1]. Since it is impossible to make entire projected image to be in-focus when projecting the pattern to ordinary scenes, the paper proposed the method to minimize the blur in the areas that is out of focus by dynamically adjusting the coded aperture as well as deliberately designing the shape of aperture; note that such condition is often used for AR system when patterns are projected over the scene.

Horita et al. proposed a method that measures accurate and dense depth value by using coded aperture on a projector [3]. Our proposed method resembles the method of Horita et al. in respect to acquiring the depth value using defocus blur of the projector. However, it differs in terms of optical design and algorithm for reconstruction.

3. Overview of the System

3.1 Configuration of the Optical System

The configuration of the proposed system is shown in **Fig. 1** (a). Since it is based on DfD, it is constructed to realize coaxial system on a projector and a camera by using a half mirror [6]. By using half mirror, although there are several drawbacks such as the amount of light drastically decreases and it requires precise adjustment on optics, we used the half mirror since it does not cause distortion even when projecting multiple patterns with different focal distance.

The projected light from the light source passes through the multiple patterns placed in front, then through the half mirror and reflected back from the measurement target and lens (Fig. 1 (a)). In order to capture the projected pattern on the target object without any distortion, the optical axis of the camera and the projector were matched to each other. Furthermore, in order to avoid the leakage light from the opposite side of the half-mirror, black curtain was placed behind. The actual system looks like Fig. 1 (b).

3.2 Configuration of Multiple Patterns

If multiple transparent glasses which are drawn with line patterns are placed between the lens and the light source, each projected pattern forms an in-focus image at different distance form the optical system depending on the position of the glass respectively as shown in **Fig. 2**. The proposed method analyzes the shape and the degree of blur on its appearance on captured image and determines the depth by comparing with the original pattern.

In terms of relation between pattern position and its in-focal position, the pattern positioned at the farthest from the lens (pattern in a red box in Fig. 2) forms a in-focus pattern closest to the lens. In case of constructing the measuring system, the pattern position and its number is set according to the system requirement on accuracy and the range of measurement depth.

3.3 Data Acquisition Process

The light emitted from the light source is captured by a camera by the following procedure.

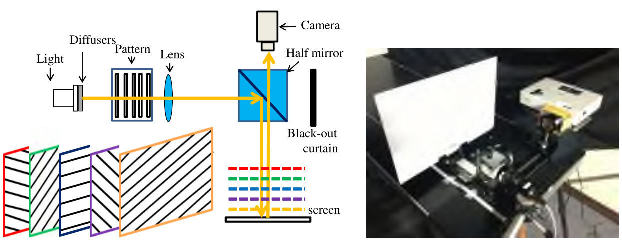
(1) The pattern is projected by passing through the pattern, the lens, and the half mirror, respectively.
(2) Then, the patterns are reflected on the target object.
(3) Finally, the reflected patterns are then coming back through the half mirror to be captured by a camera.

Since the proposed method uses a half mirror, if the luminance of light source is too strong, the projected light will be directly reflected back to the camera through the half mirror. To solve this problem, we first capture the background image without the reflection from the target object, then subtract the background image from the captured image to produce the image unaffected by the self-reflection of half mirror. Here we show the actual captured image, the background image and the reflection-free image by the background subtraction in **Fig. 3** (a)–(c) as for example. It is confirmed that the contrast of the pattern is drastically improved by the process. Rest of the processes are assumed to use the reflection-free image.

4. Shape Reconstruction Algorithm

4.1 Design of the Projected Pattern

Multiple patterns are created by rotating the parallel-line pat-



(a) (b)

Fig. 1 Configuration of the system: (a) optical configuration of the system with half-mirror and (b) experimental environment.

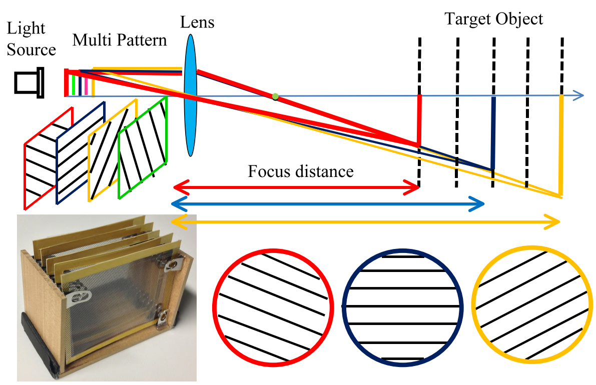


Fig. 2 Estimation of depth value using multiple patterns.

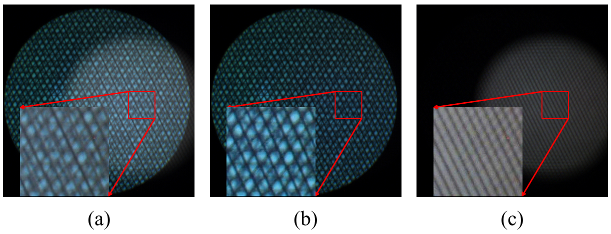


Fig. 3 Result of removal of the background effect: (a) captured image, (b) background image, (c) image without the background effect.

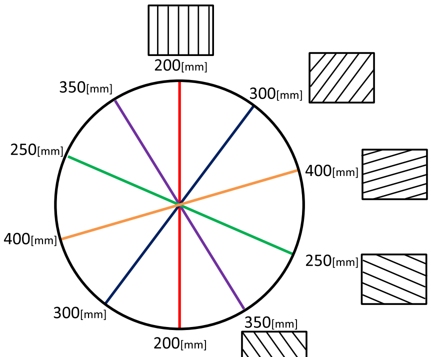


Fig. 4 Arrange pattern.

tern by changing the rotation angles as shown in Fig. 2. Each pattern is placed between the light source and the lens and positions of them are spaced so that interference of projected patterns becomes small. In addition, to decrease the interference of the pattern, we put the largely different angle of line direction for adjacent depth pair to be as shown in **Fig. 4**. This is because two patterns of adjacent depth are easily overlapped on the surface of the target object and it is difficult to decompose them if those two angles are close each other.

As for the interval between parallel lines, even though it is beneficial to extract spectrum information robustly and accurately using the pattern with high frequency, we used sparse patterns in our method, since the amount of light passing through the pattern is drastically reduced by dense pattern. Combination of dense and sparse pattern is our future research topic.

4.2 Spectrum Analysis on Projected Pattern

The proposed method estimates the depth value of each pixel position by extracting the in-focus patterns obtained from the image captured by the camera. In our technique, pattern detection is done through spectrum analysis of the captured image. This is done by calculating 2D Fourier transformation for the image to retrieve power spectrum of the image, and extracting the quantity of the spectrum for specific direction by the following procedure. First, for an input image such as **Fig. 5** (left), we define a search area in which a search window is shifted by one pixel along x and y directions to estimate the depth for all pixels. For each search window image (Fig. 5 (middle)), 2D Fourier transformation is calculated and its power spectrum image (Fig. 5 (right)) is obtained. Then, we determine the direction in which strongest spectrum appears. To calculate the amount of the power spectrum of the certain direction from the image, we used the sum of luminosity value within the narrow rectangular region placed

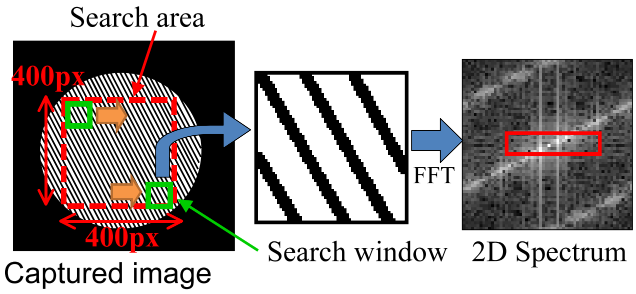


Fig. 5 Sliding search window and its 2D Fourier transformation.

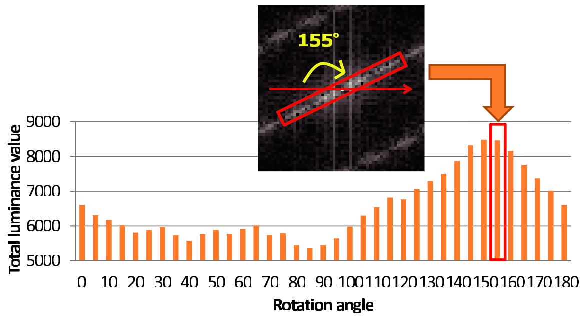


Fig. 6 Histogram of the total luminosity value for each angle.

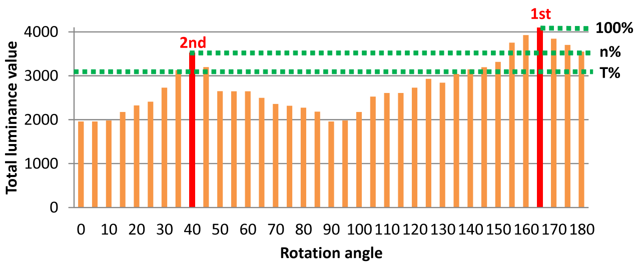


Fig. 7 Middle Estimate.

at the center of the image. The actual size of rectangular region used for the calculation as shown in Fig. 5 (right) red box is 3×11. By rotating the rectangular region to the certain degree and calculating the sum of luminosity value, the amount of the power spectrum for the direction is obtained.

4.3 Depth Estimation from Spectrum Analysis

The depth is determined by extracting an angle which is the maximum direction of the power spectrum image. The actual process for depth estimation is as follows. First, the rectangular region is rotated by five degrees and the power spectrum is calculated each time to create the histogram as shown in **Fig. 6**. Since the angle that peaks in the histogram is the dominant direction of the spectrum image, the focal length of the pattern that has the same direction of the peak is considered depth of that position.

Since the single peak detection approach can estimate only a limited number of discrete depths, solution is required to increase the resolution of depth estimation. In this paper, we use the second peak from the histogram and the ratio of the first and second peaks as shown in **Fig. 7**. Since there is a bias on power spectrum, if the second peak is smaller than the threshold, we consider there is no second peak and depth is only estimated by single peak. If the second peak is larger than threshold, the depth is calculated by

$$\frac{d_{first} \times (100 - T) + d_{second} \times (n - T)}{(100 - T + n - T)}, \qquad (1)$$

where T is the threshold and n is the height of second peak normalized by the height of the first peak, respectively.

5. Experiment

The experiments were conducted by using the optical system shown in Fig. 1 (b). Five patterns were used with two sets of line directions. The first set consists of two patterns with ±45 degrees and two patterns with ±22.5 degrees from perpendicular direction and one pattern of horizontal direction. The second set consists of directions with the same interval degree between them, i.e., 36 degrees.

5.1 Evaluation of the Method

The first set of pattern was used for the evaluation. Images were captured by shifting the target screen placed on a motorized stage. The range of the screen position is 280 mm–420 mm from the half mirror, and the image capturing interval was 2.5 mm. Under the condition, a total of 57 images ware captured. The depth value was estimated using the proposed method, and **Fig. 8** shows the result. Estimated depth was calculated by averaging 100 points of estimated depth of each frame. In the graph, we can find several accurate depths followed by a number of inaccurate depths estimation. Accurate depths are estimated near the in-focus depth of each pattern and inaccurate depths are estimated between them. We consider that the inaccuracy comes from the overlapping two blurry patterns.

Figure 9 (a)–(j) show the five selected images corresponding to the red circle in Fig. 8. We can confirm that depths are robustly estimated near in-focus depth except 290 mm depth. In the 290 mm depth image, since the error has occurred with the different pattern which has similar direction of the pattern of 362.5 mm depth, more precise estimation of angle is required to solve the problem and this is our important future research.

In the Fig. 9 (a)–(j), the decrease of the amount of light intensity is clearly observed when the distance between light source and object increases. However, there were no large differences in accuracy. This is most likely due to the fact that the proposed method is based on spectral analysis and not affected by intensity; this is one of the advantage of our method.

5.2 Depth Estimation of Slanted Plane

Next, we measured the slanted plane using the second set of pattern. We tilted the plane with 45 degrees so that the closest screen position to be 280 mm and furthest position to be 325 mm from the half mirror. Actual set up of the system and the captured image are shown in **Figs. 10** (a) and (b), respectively. Since the depth of the slanted plane from the half mirror is continuously changing, we can observe that the projected patterns are also changing continuously. The result of the depth estimation with our method is shown in **Fig. 11** (a) and its profile of the depth along the yellow line in (a) is shown in Fig. 11 (b). From the both figures, we can confirm that the depth is correctly estimated near in-focus distance. We can also confirm that the depths between in-focus distances are also estimated to some extent.

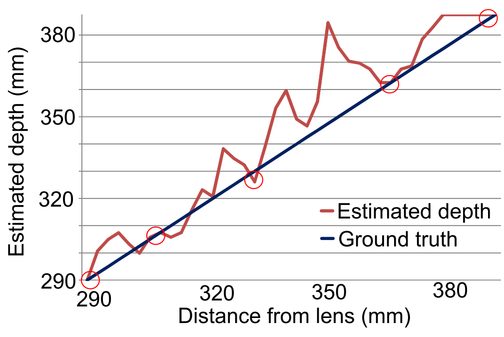


Fig. 8 Depth estimation result.

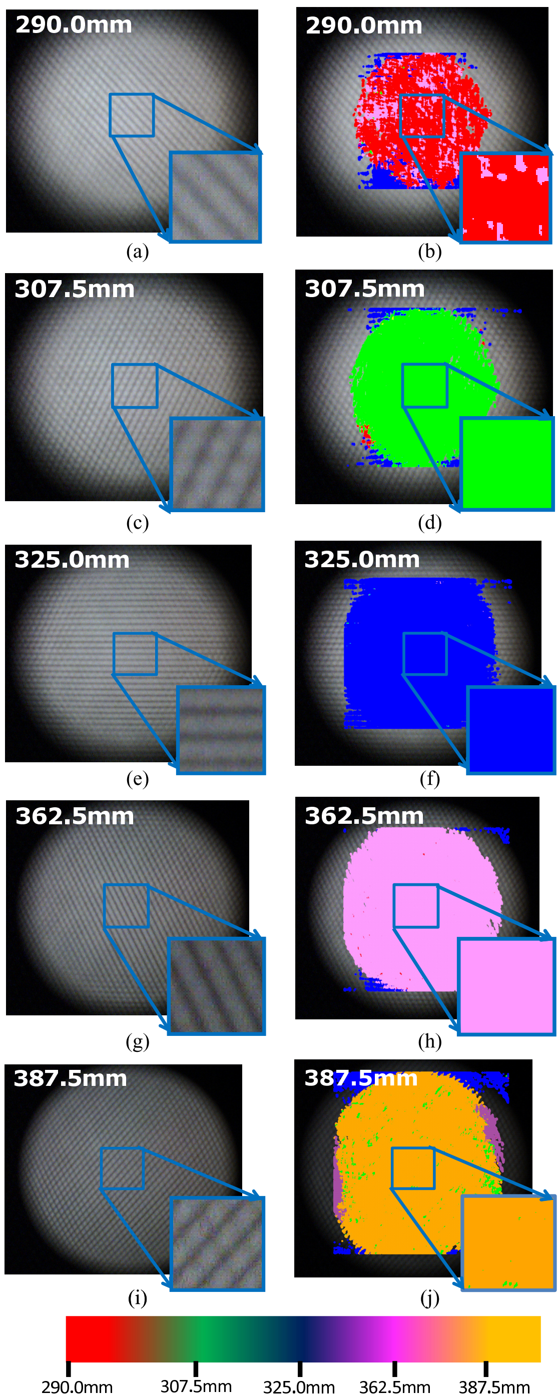
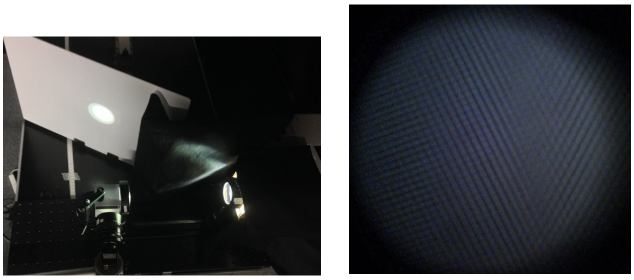
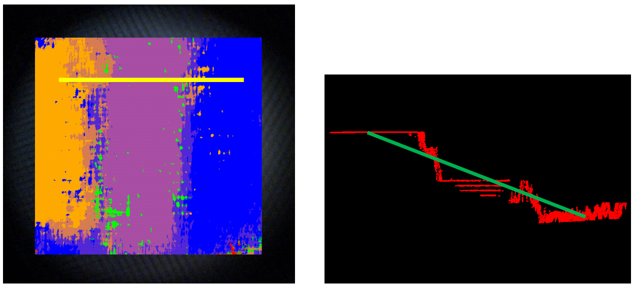


Fig. 9 Left column: captured image. Right column: estimated peak pattern for each pixel. Color represent the each pattern type closer to further order; i.e., red, green, blue, pink and yellow.



(a) (b)

Fig. 10 (a) Experimental set up and (b) Captured image.



(a) (b)

Fig. 11 (a) Estimation depth of tilted plane and (b) its profile along the yellow line in (a). Green line is the ground truth.

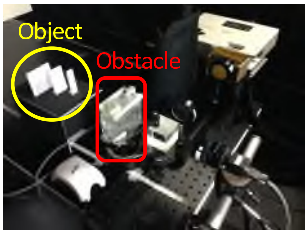


Fig. 12 Actual environment with obstacles.

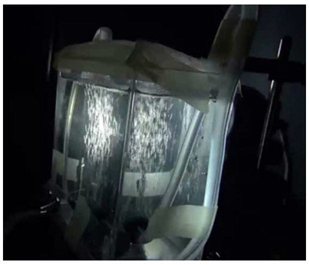


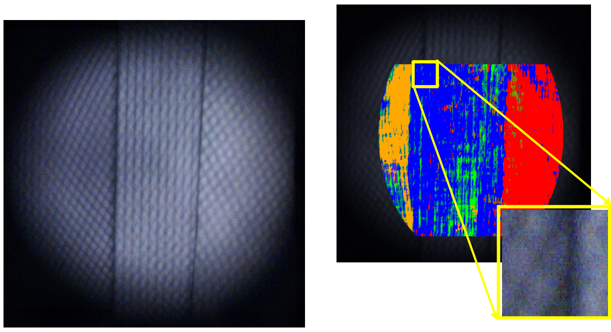
Fig. 13 Obstacle: water tank with bubble.

5.3 Depth Estimation for Scenes with Obstacles

Finally, we measured the target object in an environment with obstacles. **Figure 12** shows the measuring environment. We placed a water tank in which bubbles are emerging continuously as the obstacle in front of the target object as shown in **Fig. 13**. When the patterns were projected to the target object with the proposed system, clear patterns are observed due to a wide aperture of the projector even in the presence of obstacles as shown in **Fig. 14** (a). The result of the depth estimation using the captured image is shown in Fig. 14 (b). Here, we can confirm that the depth was correctly estimated for three objects with different depth. However, error occurred near the boundary of the objects. This is due to the fact that the window size used for the spectrum analysis was large, and was affected by occluding boundary, which is clearly shown in zoom-up view in Fig. 14 (b); note that vertical strong line is observed in the zoom-up view, whereas patterns of both depths are inclined. Further research utilizing adaptive window is necessary for solution.

6. Conclusion

We propose a DfD method that projects multiple patterns with different focal length to extend the range of depth and decrease the effect of obstacles in front of the object. In the proposed



(a) (b)

Fig. 14 (a) Capture image and (b) estimation result with close-up view near the boundary of the object.

method, we place multiple patterns at different distances from the lens to actualize multiple focal length, and analyzing the corresponding patterns being projected for depth estimation. Furthermore, since it is based on DfD, baseline is not necessary which makes it possible to measure the shape without occlusion. In the future, better method on angle detection to improve the depth resolution as well as efficient interpolation method between two patterns is important.

Acknowledgments This work was supported in part by Funding Program for Next Generation World-Leading Researchers No.LR030, Grant-in-Aid for Scientific Research No.24500120 and Strategic Information and Communications R&D Promotion Programme in Japan.

References

[1] Grosse, M. and Bimber, O.: Coded aperture projection, *SIGGRAPH Posters*, p.60 (2008).

[2] Kawasaki, H., Horita, Y., Masuyama, H., Ono, S., Kimura, M. and Takane, Y.: Optimized aperture for estimating depth from projector's defocus, *International Conference on 3D Vision* (*3DV*), pp.135–142 (2013).

[3] Kawasaki, H., Horita, Y., Morinaga, H., Matugano, Y., Ono, S., Kimura, M. and Takane, Y.: Structured light with coded aperture for wide range 3D measurement, *IEEE Conference on Image Processing* (*ICIP*), pp.2777–2780 (2012).

[4] Levin, A., Fergus, R., Durand, F. and Freeman, W.T.: Image and depth from a conventional camera with a coded aperture, *ACM SIGGRAPH 2007 papers*, *SIGGRAPH '07*, Vol.26, No.3, p.70, ACM (2007).

[5] Microsoft: Xbox 360 Kinect (2010), available from ⟨http://www.xbox.com/en-US/kinect⟩.

[6] Moreno-Noguer, F., Belhumeur, P.N. and Nayar, S.K.: Active refocusing of images and videos, *ACM Trans. Graph.*, Vol.26, No.3 (2007).

[7] Nagahara, H., Kuthirummal, S., Zhou, C. and Nayar, S.K.: Flexible Depth of Field Photography, *ECCV*, pp.60–73 (2008).

[8] Nayar, S., Watanabe, M. and Noguchi, M.: Real-Time Focus Range Sensor, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.18, No.12, pp.1186–1198 (1996).

[9] Pentland, A.: A New Sense for Depth of Field, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.9, No.4, pp.523–531 (1987).

[10] Raskar, R., Agrawal, A. and Tumblin, J.: Coded Exposure Photography: Motion Deblurring using Fluttered Shutter, *SIGGRAPH*, pp.795–804 (2006).

[11] Sagawa, R., Sakashita, K., Kasuya, N., Kawasaki, H., Furukawa, R. and Yagi, Y.: Grid-based Active Stereo with Single-colored Wave Pattern for Dense One-shot 3D Scan, *Proc. 2012 2nd Joint 3DIM/3DPVT Conference*, Zurich, Switzerland, pp.363–370 (2012).

[12] Zhang, L. and Nayar, S.K.: Projection Defocus Analysis for Scene Capture and Image Display, *ACM Trans. Graphics* (*also Proc. ACM SIGGRAPH*), Vol.25, No.3, pp.907–915 (2006).

(Communicated by *Hajime Nagahara*)