Depth from Blur by Zooming

Naoki Asada Masashi Baba Ai Oda

Department of Intelligent Systems
Hiroshima City University
Hiroshima 731-3194, Japan
E-mail:{asada,baba,ai}@cv.its.hiroshima-cu.ac.jp

Abstract

This paper presents a novel approach to depth recovery from image blur by zooming. We first discuss the optical properties of a zoom lens system; the image blur varies according to zoom as well as focus and iris settings. Then a thin lens based camera model is proposed that is characterized by the effective focal length and lens diameter. Using this model, we have proved that the depth information is recovered from the image blur at any setting of zoom, focus and iris parameters, and have developed a method to estimate the depth from a set of multiple images taken at different zoom settings. Experimental results using real images have shown that the depth information is reliable around the focused distance and is stably recovered from multiple zoom images. Finally, we have demonstrated an application of the depth recovery by zooming to image synthesis; that is, the virtual objects embedded in a real image are defocused appropriately based on the recovered 3D geometry.

1. Introduction

Depth recovery is an essential problem in computer vision to understand 3D scenes from 2D images. Among various cues for depth estimation such as stereopsis and motion parallax, blurring phenomena due to defocusing have been studied to realize monocular range sensing. Since the image blur is observed when an object is out of focus of a lens that has an aperture opening, the defocusing is controlled by focus as well as iris settings of the lens. Based on this principle, most of previous work has achieved the depth from blur by focusing[1, 2, 3] or irising[4, 5, 6].

In addition to the focus and iris, zoom lens systems have another parameter *zoom*, and the image blur also varies according to the zoom setting. Figure 1 shows an example where the blurring effect by zooming is different from that of enlarged image by interpolation. This observation suggests that depth could be estimated from blur by zooming.

In literature, some methods for depth recovery by zooming have been proposed[7, 8]. However, since the zoom lens is modeled as a set of pinhole cameras each of which has different focal length, the blurring phenomena caused by zooming have not considered.

In this paper, we will show theoretically and experimentally that the depth information is obtained from image blur by zooming as well as focusing and irising. In the following section, we first discuss the optical properties of zoom lens parameters, then present a novel camera model that describes the relationship between zoom, focus and iris parameters. Unlike the previous modeling of zoom lens[9], our model allows to analyze the image blur due to defocusing. Using this camera model with calibration parameters, we will show some experimental results and accuracy evaluation of the depth from blur by zooming.

2. Zoom Lens Camera Model

2.1. Zoom, focus and iris parameters

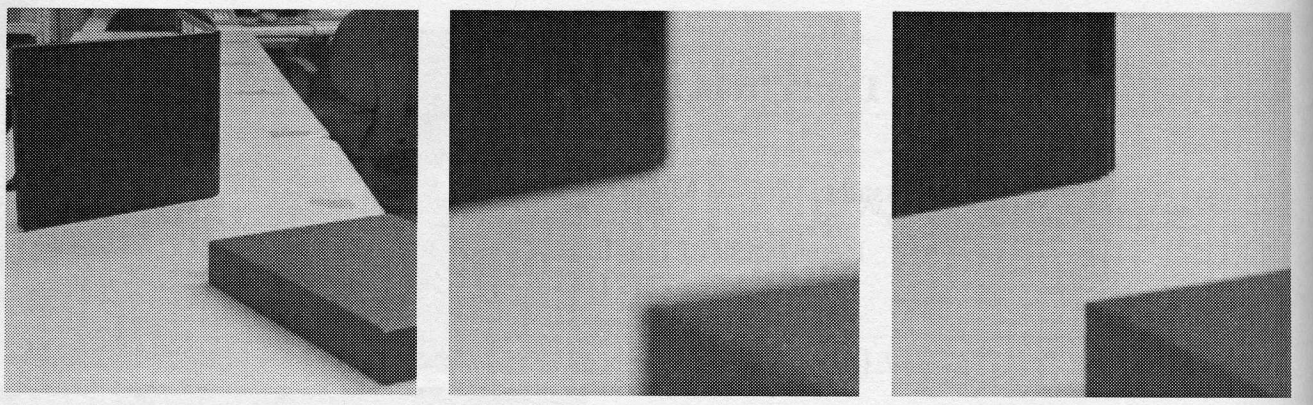
In most zoom lens systems, zoom, focus and iris parameters have the following optical properties.

Zoom allows to take images with arbitrary magnification of a scene without changing the focused distance and image intensity.

Focus enables to take focused images of an object at any distance with varying the image magnification without changing image intensity.

Iris is capable of controlling the image intensity with changing the depth of field.

Since these properties are based on the lens effects, we have built a camera model by using a thin lens imaging system, as shown in Fig. 2. Note that while the thick lens



(a) A slightly defocused image of a scene

(b) Enlarged image by zooming

(c) Enlarged image by interpolation

Figure 1. Blurring effect by zooming is different from that of enlarged image by interpolation.

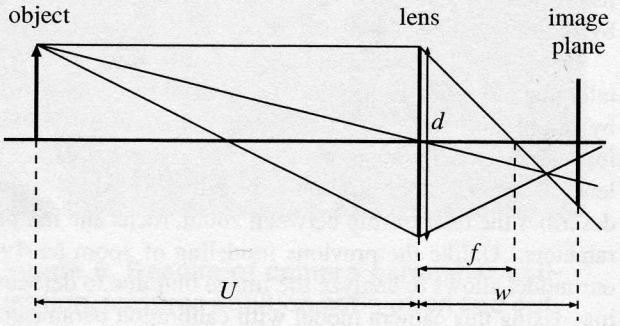


Figure 2. Zoom lens camera model.

imaging system is necessary for modeling zoom lens system with pose and position calibration[11], this paper discusses not external but internal parameter calibration of a stationary zoom lens system. This is why such a simple camera model based on thin lens is sufficient for modeling a zoom lens system in our study.

2.2. Properties of the camera model

To describe the optical properties of the camera model, we define some notations of real, lens and model parameters. The real parameters denote the actual value of zoom, focus, and iris settings such as a voltage of servo control. The lens parameters represent the optical characteristics of the lens such as the focal length and F-number. The model parameters are the effective values that describes the relationship between zoom, focus, and iris.

Real parameter : Z, F, and I.

Z is a value of zoom setting from 25(wide) to 75(telephoto), F is a value of focus setting from 0(infinite) to 50(nearest), and I is a value of iris setting from 25(open) to 75(close).

Lens parameter : f, U, and A.

f is the focal length that is a function of Z, U is the focused distance from the lens that is a function of F, and A is the F-number of the lens that is a function of I.

Model parameter : w, d, and B.

w is an effective focal length that is defined as a distance between the lens and the image plane, see Fig. 2. d is an effective lens diameter, and B denotes an effective F-number that is defined by w/d which is proportional to A.

Table 1 shows the relationship and dependency between real, lens, and model parameters. Note that all parameters but U are related to internal conditions of the camera model. Thus, we define the effective focal length w by using not U but the focus setting F.

Property 1: The effective focal length w is represented by a polynomial function of the focus setting F, that is,

$$w = f + \sum_{n=1}^{N} \alpha_n(f) F^n \qquad (1)$$

where $\alpha_n(f)$ denotes polynomial functions of the focal length f. Note that w satisfies the condition $w = f$ when $F = 0$. This means that w becomes f when the object at infinite distance is in focus.

Eq. (1) represents the imaging effects by zoom and focus parameters; that is, the image magnification by zoom Z and focus F is a function of $w(f(Z), F)$. Note

Table 1. Relationship between real, lens, and model parameters.

real parameter	zoom Z	focus F	iris I
lens parameter	focal length $f(Z)$	focused distance $U(F)$	F-number $A(I)$
model parameter	effective focal length $w(f, F)$		effective F-number $B(A)$
	effective lens diameter $d(w, B)$		

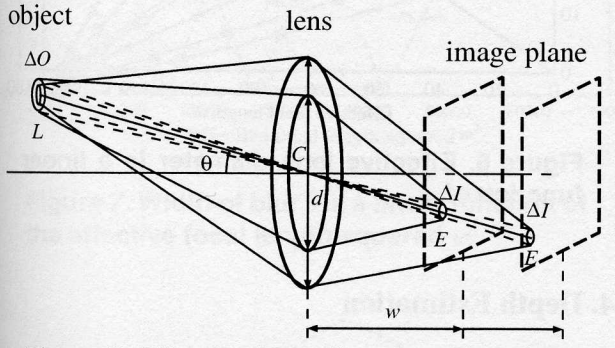


Figure 3. Effective lens diameter d is a function of the effective focal length w.

that zoom makes no effect on the focusing condition because U is defined solely by F irrespectively of zoom Z.

Property 2: The irradiance E of an infinitesimal area $\triangle I$ on the image plane is given by

$$E = L\frac{\pi}{4}\left(\frac{d}{w}\right)^2 \cos^4\theta \qquad (2)$$

where L denotes the radiance of an infinitesimal area $\triangle O$ on the object surface, and θ is an angle of incident light on the lens[10], as shown in Fig. 3.

Eq. (2) tells that image intensity depends upon the effective focal length w which is a function of the focal length f that relies on the zoom setting Z. However, zoom does not affect the image intensity as stated in section 2.1. Thus, the effective F-number B, that is defined by w/d, should be independent of the zoom setting. This means that the effective lens diameter d is proportional to the effective focal length w. Figure 3 depicts the that d varies according to w to keep the irradiance E constant.

3. Camera Calibration

We used a CCD video camera SONY XC-007 with a zoom lens CANON J16×9.5B4RAS to take real images. Table 2 shows the optical characteristics of the zoom lens. The

Table 2. Optical characteristics of the zoom lens.

focal length (f)	9.5mm – 152mm
focused distance (U)	1.0m – ∞
F-number (A)	1.8 – close

spatial and intensity resolutions of images were 512×480 pixels and 256 gray levels for each RGB signal.

Using a calibration target that consists of a structured pattern, we obtained the following relationship between real and lens parameters.
focal length:

$$f(Z) = 4.73 \times 10^{-2}Z^2 - 2.37Z + 4.27 \times 10 \qquad (3)$$

focused distance:

$$U(F) = 4.25 \times 10F^{-1} + 5.25 \times 10^{-2} \qquad (4)$$

F-number:

$$A(I) = 1.8\sqrt{2}^{(1.75\times10^{-1}I - 4.39)} \qquad (5)$$

3.1. Effective focal length w

To determine w of Eq. (1), we evaluated the image magnification due to focus F. The focus magnification M_{focus} is defined by the ratio of w to f, that is

$$M_{focus} = 1 + \frac{1}{f}\sum_{n=1}^{N}\alpha_n(f)F^n \qquad (6)$$

The focus magnification measured at the zoom setting $Z = 25, 30, 35, 40, 45, 50, 60$ is plotted in Fig. 4. Using the least squares method, we have found that the focus magnification is represented by a quadratic function of F whose regression curves are also shown in Fig. 4.

From the coefficients of M_{focus} in terms of F, we determined $\alpha_1(f)$ and $\alpha_2(f)$ then obtained w as follows.

$$w = f + (5.88 \times 10^{-5}f^2 + 2.38 \times 10^{-3}f)F + (1.48 \times 10^{-6}f^2 - 1.62 \times 10^{-5}f)F^2 \qquad (7)$$

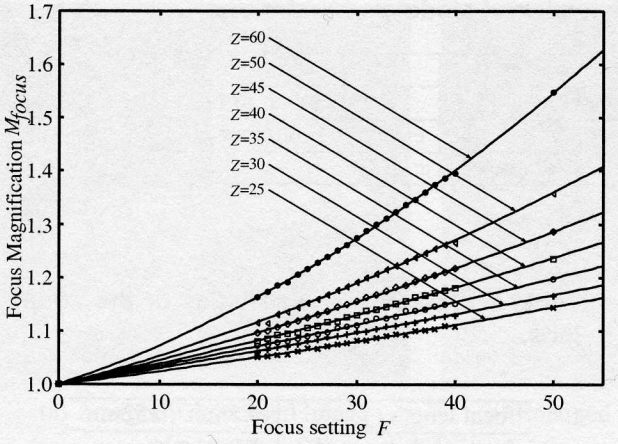


Figure 4. Focus magnification and regression curves.

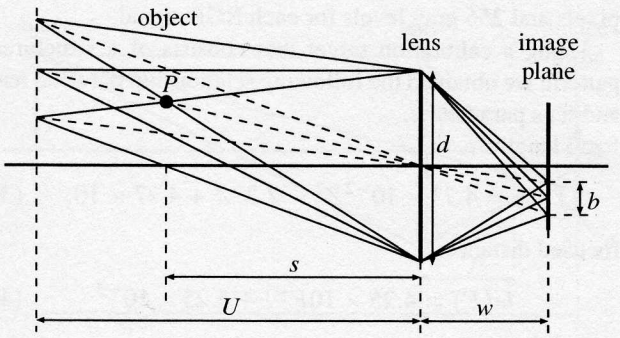


Figure 5. Width of blur b is a function of w and d.

3.2. Effective lens diameter d

The effective lens diameter d and the pixel size β are determined from the blurred images of a step edge. Figure 5 illustrates that the width b of blurred image of a point object P is a function of d, w, U, and object distance s, that is

$$b = wd\left|\frac{1}{s} - \frac{1}{U}\right| = \frac{w^2}{B}\left|\frac{1}{s} - \frac{1}{U}\right|, \tag{8}$$

where $B = w/d$.

From the number of pixels of blurred edge taken at $F = 10, 15, 20, 25, 30$ and $Z = 55, 60, 65, 70$ with $A = 1.8$, we determined $\beta = 1.40 \times 10^{-2}$mm and d as a linear function of w; that is, $d = 0.37w$ in Fig. 6. Thus, we obtained d,

$$d = wB^{-1} = 6.66 \times 10^{-1} wA^{-1}. \tag{9}$$

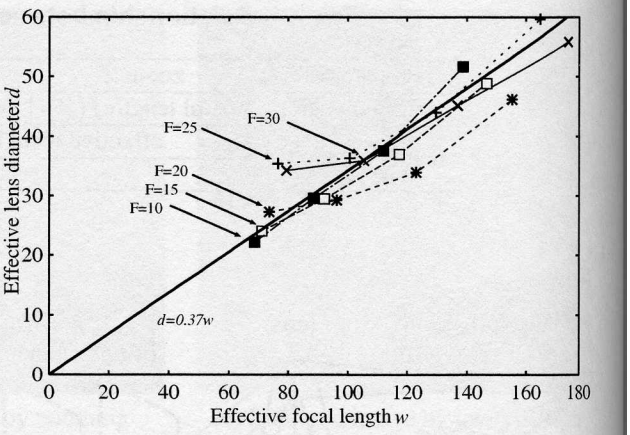


Figure 6. Effective lens diameter is a linear function of w.

4. Depth Estimation

Using Eq. (8), we have the depth s of an edge from its width of blur b, that is,

$$s = \begin{cases} \dfrac{w^2 U}{w^2 + bUB} & (s < U) \\ \dfrac{w^2 U}{w^2 - bUB} & (s > U) \end{cases} \tag{10}$$

where U and B depend upon the focus and iris parameters, respectively, and w is determined by focus and zoom settings. Thus, the depth s is computed from the width of blur b at any setting of zoom, focus, and iris parameters.

From a set of multiple images taken at different zoom settings, we also have the depth information. Since Eq. (8) tells that the width of blur b is proportional to the effective focal length squared w^2, the coefficient g of w^2, that is

$$g = \frac{1}{B}\left|\frac{1}{s} - \frac{1}{U}\right|, \tag{11}$$

gives the depth s as follows.

$$s = \begin{cases} \dfrac{U}{1 + gBU} & (s < U) \\ \dfrac{U}{1 - gBU} & (s > U) \end{cases} \tag{12}$$

where U and B remain constant if the focus and iris parameters are unchanged. Thus, the depth s is determined from the coefficient g of w^2 on the regression line of b given from the multi-zoom images. Note that the pixel correspondence between different zoom images is established by the coordinates of image center and image magnification ratio due to zooming given by the camera calibration.

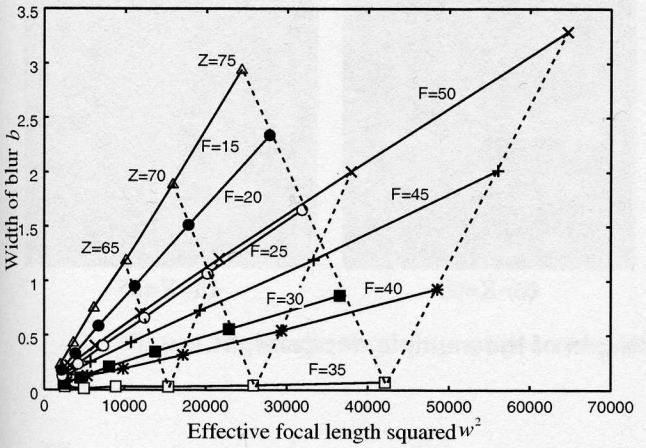


Figure 7. Width of blur b is a linear function of the effective focal length squared w^2.

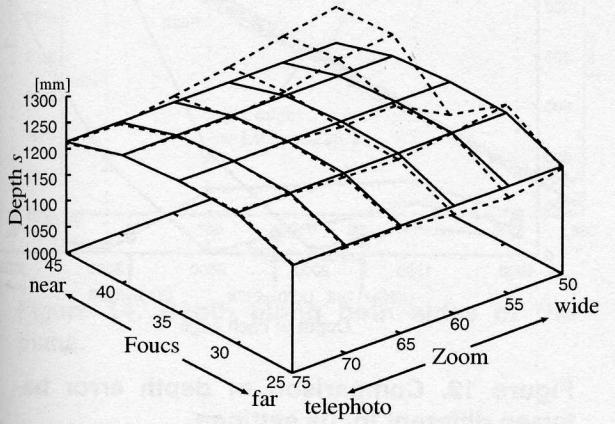


Figure 8. Depth from each single image and multi-zoom images.

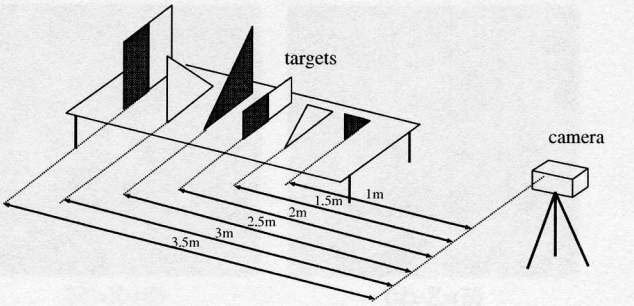


Figure 9. Experimental configuration of multiple depth target.

5. Experiments and Evaluation

5.1. Single depth target

We used the video camera with zoom lens whose calibration parameters are obtained in the section 3. A planar target having a step edge was placed at 1250mm distant from the camera position. The width b of the blurred edge was measured from each image taken at $Z = 50, 55, 60, 65, 70, 75$ and $F = 15, 20, 25, 30, 35, 40, 45, 50$ with $I = 35$. Among forty-eight images, we had almost focused image at $F = 35$; that is, the depth was about $U(35) = 1260$mm.

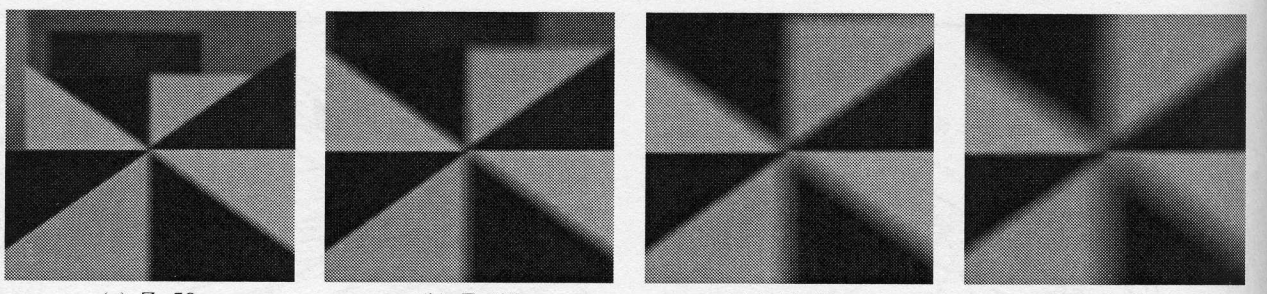
We have plotted the width of blur b against w^2 for each F when varying w by Z on Fig. 7, and we have found that b is proportional to w^2 for every F. This observation is consistent with the theoretical consideration described in section 4. Note that the change of b for each Z when varying F forms "V" shape which is a familiar observation in the depth from focus method; that is, the width of blur decreases then increases when varying the focus setting around the focused position $F = 35$ as shown by broken lines in Fig. 7.

Figure 8 shows the depth s from blur b at each Z and F by broken lines and s from coefficient g of w^2 in the multi-zoom images by solid lines. The average of depth obtained from single zoom images was 1236mm with standard deviation s.d.=27.4mm and that computed from the multi-zoom images was 1233mm with s.d.=19.8mm. From these results, we have the following observations.

- The depth determined from a set of multi-zoom images is more stable than that from a single image.
- The depth from single image is computed more stably when using the telephoto setting of the zoom. This means that we can obtain more reliable b from images with higher spatial resolution.

5.2. Multiple depth target

We have examined the accuracy of simultaneous recovery of different depths in a scene. Figure 9 shows the experimental configuration that consists of step edges of six different depths, and each edge label was numbered as illustrated in Fig. 11. Twelve images were taken at $Z = 50, 55, 60, 65$ with each $F = 15, 20, 25$ and $I = 35$, four of which at $F = 20$ are displayed in Fig. 10. The depth of each edge computed from multi-zoom images was plotted in Fig. 12, and the accuracy evaluation of average depth, error, and standard deviation for each edge are shown in Table 3. Note that since the blur of occluding edge is theoretically identical to that of the nearer surface edge[12], the edge label 1 and 2, 4 and 5 yield the same depth. In Fig. 12, the scale



(a) Z=50 (b) Z=55 (c) Z=60 (d) Z=65

Figure 10. Image sequence of zooming-in of the multiple depth target.

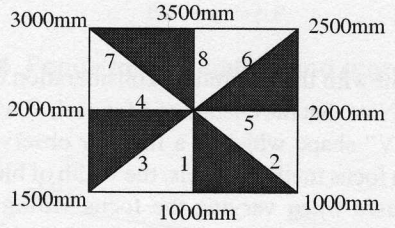


Figure 11. View of multiple depth target and edge labels.

Table 3. Accuracy evaluation of multiple depth recovery at $F = 20$.

edge	truth*	average*	error*	s.d.*
1,2	1000	881.4	118.6	6.7
3	1500	1294.5	205.5	3.4
4,5	2000	1887.0	113.0	10.1
6	2500	2496.4	3.6	12.1
7	3000	3364.4	364.4	33.6
8	3500	4545.9	1045.9	42.4

*unit=[mm] (s.d.=standard deviation)

of the vertical axis is marked by the depth of each edge (ground truth) and that of the horizontal one gives the depth error. The dots with arrows on each line denote the focused distance at $F = 15, 20, 25$, respectively. From Fig. 12 and Table 3, we have the following findings.

- The depth error was very small around the focused distance for each focus setting.
- The error was relatively small in the nearer distance, whereas the depth accuracy deteriorated rapidly in the farther one.
- The standard deviation increased in the farther distance.

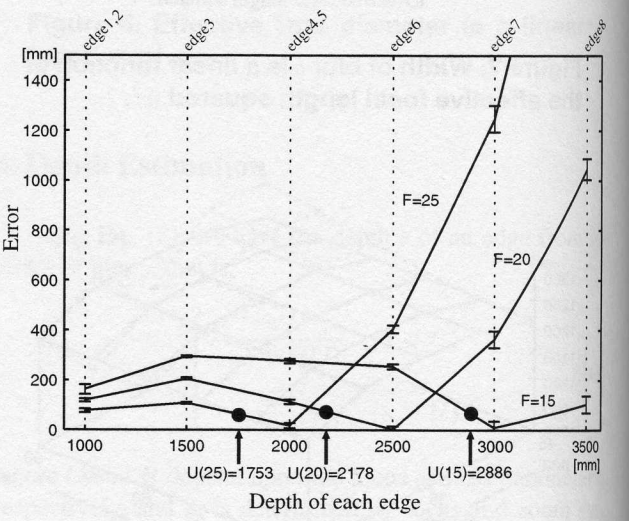


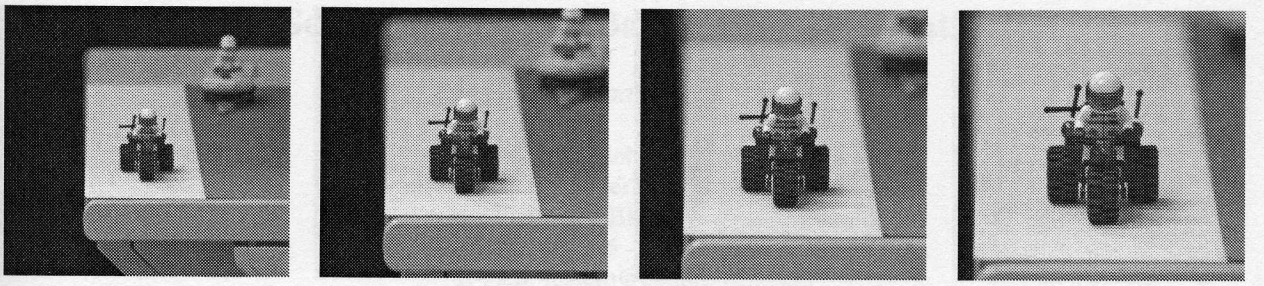
Figure 12. Comparison of depth error between different focus settings.

These phenomena are observed commonly in the method of depth from image blur; that is, the over-blurred edges and images with lower spatial resolution yield larger errors and less stable results. This suggests that we should apply the depth from zoom method to the objects almost in focus to obtain the reliable information.

5.3. Depth varying target

The third experiment was performed in more realistic situation. A toy was placed on a desk and an image sequence of zooming-in was taken with fixed focus and iris settings. Figure 13 shows four images taken at $Z = 55, 60, 65, 70$ with $F = 24$ and $I = 35$. The front wheel of the toy was in focus and its distance was given from the focus setting; that is $U(24) = 1823$mm.

The depth was computed from the blur along the left and



(a) Z=55 (b) Z=60 (c) Z=65 (d) Z=70

Figure 13. Image sequence of zooming-in of a scene including depth varying edges.

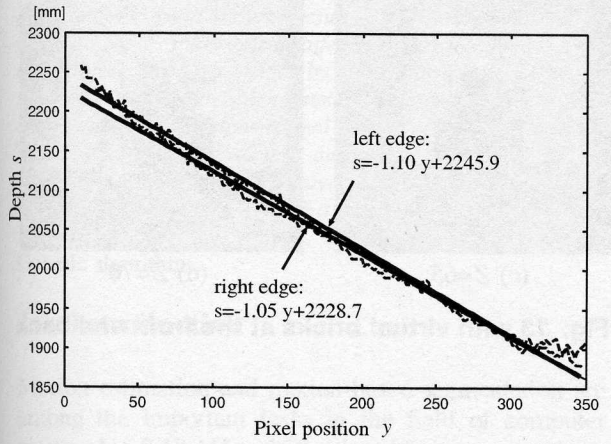


Figure 14. Depth along both sides of the plane.

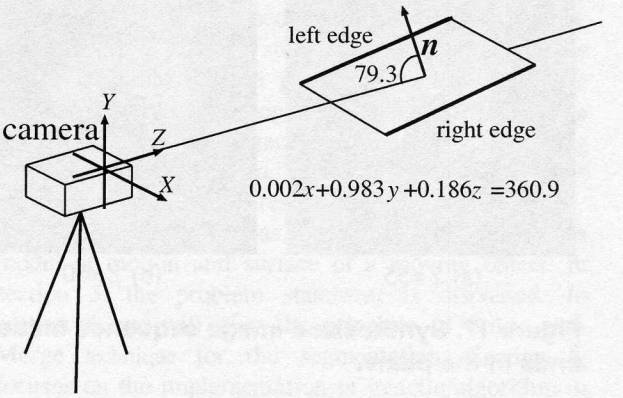


Figure 15. Recovered 3D geometry of the plane.

right edges of the bright plane under the toy. Figure 14 shows the depth from each edge point with regression lines whose correlation coefficients were both over 0.99. The results indicate that the depth varies almost linearly along the edges, which is consistent with the scene configuration.

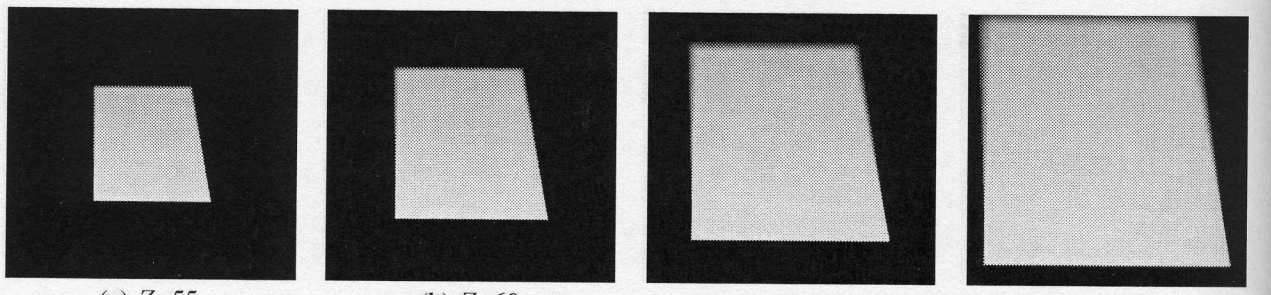
Using the depth of the left and right edge lines, we have recovered the 3D geometry of the plane represented in the camera-centered coordinates, as shown in Fig. 15. Once we have 3D information of an object in the real images, we can reconstruct it with the same pose and position in a virtual space , as shown in Fig. 16. Moreover, virtual objects can be embedded in the real image with keeping geometric and photometric consistency. Figure 17, for example, displays a synthesized image sequence made from Fig. 13 with virtual bricks at the front and back ends of the bright plane. Note that the virtual objects are defocused appropriately based on the camera model and 3D geometry.

6. Conclusion

This paper has proposed a zoom lens camera model that describes the relationship between zoom, focus and iris parameters, and presented a novel approach to depth recovery from image blur by zooming. We performed experiments using real images and the depth was successfully computed from a set of images taken at different zoom settings. We also showed an application of the depth recovery to realistic image synthesis. These results demonstrated the validity of the camera model as well as the effectiveness of the depth from zoom method.

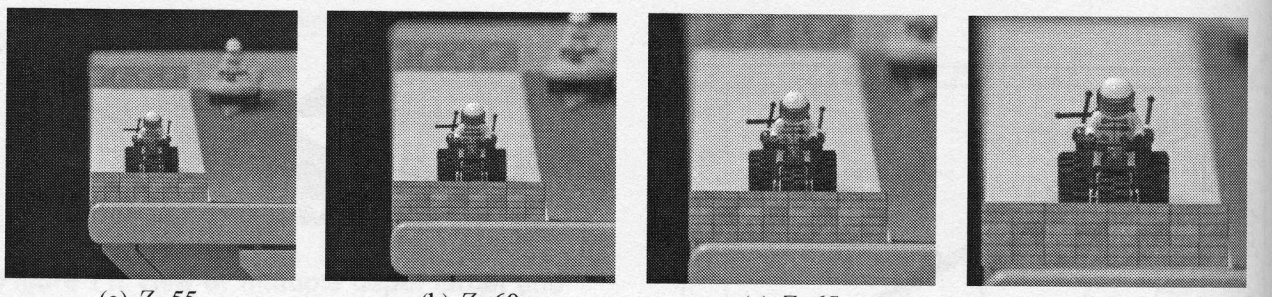
Since the reliable depth information are obtained in the range around focused distance, the proposed method is expected to be used with the depth from focus like the third experiment. Note that we often encounter such a situation where an object centered in the image was in focus and the image sequence of zooming-in was taken by a fixed pose and position camera.

In future, it is necessary to verify the camera model by examining the applicability to many types of zoom lens. We also need to evaluate the performance of depth recovery in



(a) Z=55 (b) Z=60 (c) Z=65 (d) Z=70

Figure 16. Generated image sequence of the plane placed in the same 3D geometry as Fig. 13.



(a) Z=55 (b) Z=60 (c) Z=65 (d) Z=70

Figure 17. Synthesized image sequence made from Fig. 13 with virtual bricks at the front and back ends of the plane.

many situations.

Acknowledgments

This research was partially supported by the Grant-in-Aid for Scientific Research (C) No.12680357 of Japan Society for the Promotion of Science and by the Grant for Special Academic Research No.0018 of The Hiroshima City University.

References

[1] E. Krotkov, "Focusing", *IJCV*, pp.223-237, 1987.

[2] M. Subbarao, "Parallel depth recovery by changing camera parameters", *Proc. ICCV*, pp.149-155, 1988.

[3] N. Asada, H. Fujiwara and T. Matsuyama, "Edge and depth from focus", *IJCV*, Vol.26, No.2, pp.153-163, 1998.

[4] A.P. Pentland, "A new sense for depth of field", *IEEE Trans. PAMI*, Vol.9, No.4, pp.523-531, 1987.

[5] J. Ens and P. Lawrence, "An investigation of methods for determining depth from focus", *IEEE Trans. PAMI*, Vol.15, No.2, pp.97-108, 1993.

[6] G. Surya and M. Subbarao, "Depth from defocus by changing camera aperture: A spatial domain approach", *Proc. CVPR*, pp.61-67, 1993.

[7] J. Ma and S.I. Olsen, "Depth from zooming", *J. Opt. Soc. Am. A*, Vol.7, No.10, pp.1883-1890, 1990.

[8] J.M. Lavest, G. Rives and M. Dhome, "Three-dimensional reconstruction by zooming", *IEEE Trans. RA*, Vol.9, No.2, pp.196-207, 1993.

[9] K. Tarabanis, R.Y. Tsai and D.S. Goodman, "Calibration of a computer controlled robotic vision sensor with a zoom lens", *CVGIP*, Vol.59, No.2, pp.226-241, 1994.

[10] B.K.P. Horn, *Robot Vision*, MIT Press, 1986.

[11] R. Kingslake, *Optical System Design*, Academic Press, 1983.

[12] N. Asada, H. Fujiwara and T. Matsuyama, "Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model", *IEEE Trans. PAMI*, Vol.20, No.2, pp.155-167, 1998.