



Edge and Depth from Focus*

NAOKI ASADA[†]

*Department of Intelligent Systems, Hiroshima City University, 3-4-1 Ozuka-Higashi, Asa-Minami,
Hiroshima 731-31, Japan*

asada@its.hiroshima-cu.ac.jp

HISANAGA FUJIWARA

Industrial Research Center of Okayama, 5301 Haga, Okayama 701-12, Japan

fujiiwara@okakogi.go.jp

TAKASHI MATSUYAMA[†]

Department of Electronics and Communication, Kyoto University, Kyoto 606-01, Japan

tm@kuee.kyoto-u.ac.jp

Received September 4, 1994; Revised October 12, 1996; Accepted June 9, 1997

Abstract. This paper proposes a novel method to obtain the reliable edge and depth information by integrating a set of *multi-focus images*, i.e., a sequence of images taken by systematically varying a camera parameter *focus*. In previous work on depth measurement using focusing or defocusing, the accuracy depends upon the size and location of local windows where the amount of blur is measured. In contrast, no windowing is needed in our method; the blur is evaluated from the intensity change along corresponding pixels in the multi-focus images. Such a blur analysis enables us not only to detect the edge points without using spatial differentiation but also to estimate the depth with high accuracy. In addition, the analysis result is stable because the proposed method involves integral computations such as summation and least-square model fitting. This paper first discusses the fundamental properties of multi-focus images based on a step edge model. Then, two algorithms are presented: edge detection using *an accumulated defocus image* which represents the spatial distribution of blur, and depth estimation using *a spatio-focal image* which represents the intensity distribution along focus axis. The experimental results demonstrate that the highly precise measurement has been achieved: 0.5 pixel position fluctuation in edge detection and 0.2% error at 2.4 m in depth estimation.

Keywords: information integration, multi-focus images, depth from focus/defocus, edge detection, depth estimation

*This research was partially supported by the Grant-in-Aid for Scientific Research No. 05680298, No. 08680410 and No. 07558047 of The Ministry of Education and by the Grant for Special Academic Research No. A419 of The Hiroshima City University.

[†]N. Asada and T. Matsuyama were with Okayama University when this research was initiated.

1. Introduction

Integration of multiple images taken by systematically varying a camera parameter allows us to extract the reliable scene information which cannot or hardly be extracted from a single image. When we control the lens aperture, for example, accurate chromaticity and brightness can be computed from a set of multi-iris images (Asada and Matsuyama, 1992). In this paper, we propose an integration method to obtain the reliable edge and depth information from *multi-focus images*, i.e., a sequence of images taken by varying the focus parameter.

Among various cues for depth recovery such as stereopsis and motion parallax, blurring phenomena due to defocusing have received much attention to realize “Depth from Focus” (DFF in short) or “Depth from Defocus” (DFD in short). DFF is a search-based method (Jarvis, 1983; Krotkov, 1987; Darrell and Wohn, 1988; Nair and Stewart, 1992; Nayar, 1992); the depth of a scene point is measured by searching the “optimal” focus position where the image of the point is just in focus. On the other hand, DFD is a model-based method (Grossmann, 1987; Pentland, 1987; Subbarao, 1988; Lai et al., 1992; Subbarao and Wei, 1992; Ens and Lawrence, 1993); the depth map of a scene is computed by solving the inverse problem of blurring process based on camera and edge models¹. Much efforts have been made to improve focus criterion functions (Lighthart and Groen, 1982; Jarvis, 1983; Krotkov, 1987) and search algorithms (Krotkov, 1987; Xiong and Shafer, 1993) in DFF, as well as image blurring models (Subbarao, 1988) and computation procedures (Ens and Lawrence, 1993) in DFD.

Although the computational schemes are different between DFF and DFD, almost all algorithms proposed so far employ local windows to measure the amount of blur. Such a window-based method poses difficult technical problems. First, since the smaller window allows the more precise but less stable depth recovery, how can we determine the appropriate window size? Second, since the window should be located to include a single feature of depth, e.g., a blurred edge, how can we find it before the depth recovery? Since images generally contain a variety of edges with different degrees of blur, simple edge detector using spatial differentiation cannot always extract such edges reliably. If the window size is inappropriate and/or the window is misplaced, the amount of blur would be measured falsely due to the presence of other blurred edges within or around the window.

To avoid these windowing problems and to obtain the reliable depth information, we have developed a novel method called “Edge and Depth from Focus”; that is, the depth estimation is accomplished together with the edge detection by measuring the blur from the intensity change along corresponding pixels in the multi-focus images. Our method processes a sequence of images taken at different focus positions like DFF, and employs an image blurring model to estimate the depth like DFD. However, neither search procedure nor strict blurring model is needed. Instead, we assume a step edge model as a primitive cue of scene features. Such a simple edge model enables us to develop integration-based algorithms which yield stable and reliable results.

In the following sections, we first discuss some important properties of multi-focus images based on an ideal edge model. Then, we describe two algorithms: edge detection using an *accumulated defocus image* which represents the spatial distribution of blur, and depth estimation using a *spatio-focal image* which represents the intensity distribution along focus axis. Finally, we demonstrate the effectiveness of our method with experimental results using natural scenes and show the quantitative accuracy evaluation.

2. Multi-Focus Images

2.1. Camera Model

A simple camera model consisting of a thin lens and an image plane is used to derive some fundamental characteristics of focusing based on geometrical optics. Figure 1 illustrates a focusing mechanism to bring the image plane into the position in focus. When an object is placed at distance u from the lens, the sharply focused image of the object is taken if the distance

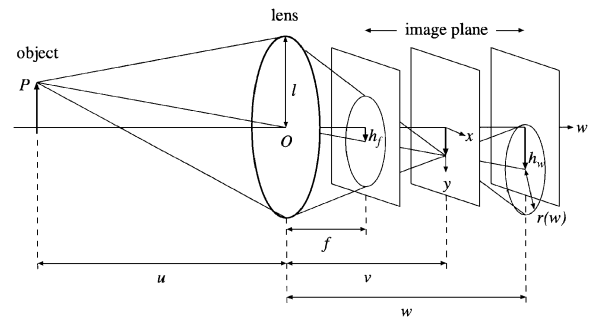


Figure 1. Camera model and the focusing mechanism.

between the image plane and the lens is v , where u , v and the focal length f of the lens satisfy the thin lens formula,

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}. \quad (1)$$

A set of multi-focus images is modeled as a sequence of images taken at different focus positions of the image plane. Such an image sequence defines a three-dimensional geometric space represented by two-dimensional image coordinates x - y and one-dimensional focus axis w , where the origin is given at the lens center O as shown in Fig. 1. We will consider a focusing range between $f \leq w \leq w_{\max}$ which corresponds to the depth between infinite distance and a certain nearest distance from the lens. In principle, the range of focus position is sampled randomly to generate the multi-focus images. Note that such an image sequence does not necessarily include the focused images of objects.

2.2. Image Blurring Model

The blurred image $B(x, y, w)$ taken at a focus position w is mathematically modeled by the convolution of a non-blurred image $I(x, y, w)$ with a blurring kernel $K(x, y, w)$ as follows².

$$B(x, y, w) = I(x, y, w) * K(x, y, w), \quad (2)$$

where $*$ denotes the convolution operation. In general, $K(x, y, w)$ is a rotationally symmetric function like a two-dimensional Gaussian (Subbarao et al., 1994) which satisfies

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x, y, w) dx dy = 1. \quad (3)$$

The radius of $K(x, y, w)$ is given by

$$r(w) = \frac{l |w - v|}{v}, \quad (4)$$

where l is the radius of the lens. We assume for simplicity that $K(x, y, w)$ is zero outside of the blurring circle defined by $r(w)$.

2.3. Pixel Correspondence

Focusing varies not only the degree of blur but also the size of image. In Fig. 1, the image of an arrow object is

enlarged as the focus position w increases. We define a magnification factor m as a ratio of length h_w of the arrow image at a focus position w to the corresponding length h_f at the focus position f . The ratio of h_w to h_f is equivalent to that of w to f , so that we have

$$m = \frac{h_w}{h_f} = \frac{w}{f}. \quad (5)$$

Using the magnification factor, we can establish a pixel correspondence in the multi-focus images; that is, a pixel at (x, y, f) corresponds to the pixel at (mx, my, w) . A set of such corresponding pixels forms a straight line, i.e., a line passing through P and the lens center O , which we call a *pixel correspondence line* of P .

2.4. Edge Model

For the analysis of multi-focus images, we use the following edge model.

- A straight and isolated edge exists on a plane placed perpendicularly to the optical axis of the lens.
- Each side of the edge has uniform brightness whose intensity is denoted by L_1 and L_2 ($L_1 > L_2$).

The intensity profile across the edge is represented by an ideal step function. Figure 2 shows the edge model aligned in parallel with the x axis and three images taken at different focus positions.

The change of blur around the pixel correspondence line of an edge point P is best characterized on the plane ϕ which is determined by the pixel correspondence line and a line perpendicular to the edge on the plane. To analyze the change of blur with respect to the focus position w , we project the intensity distribution of ϕ orthogonally onto the plane μ which includes

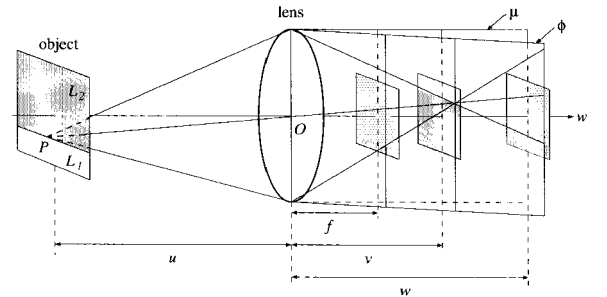


Figure 2. Step edge model and an example of multi-focus images. The plane ϕ includes the pixel correspondence line of P and μ is defined on the focus axis.

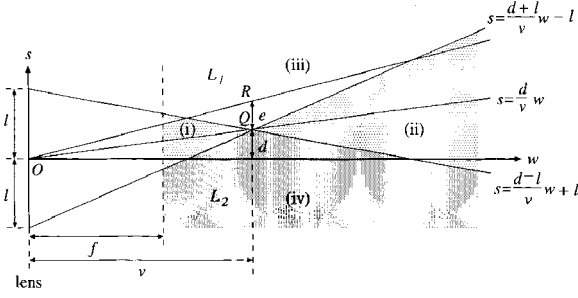


Figure 3. Intensity distribution in a spatio-focal image. $Q(v, d)$ denotes the focused image of the edge point P .

the focus axis and a lens diameter determined by ϕ . Figure 3 shows the intensity distribution of μ where the abscissa is the focus axis w and the ordinate is the spatial axis s , so that we call this intensity distribution a *spatio-focal image*.

2.5. Spatio-Focal Image

In Fig. 3, the focused image of P is given by $Q(v, d)$ and the pixel correspondence line of P is represented by $s = \frac{d}{v}w$. The upper side of this line corresponds to the L_1 region, and the lower to the L_2 . The left side of the focused position v corresponds to the blurred images of P when the farther distance is in focus, and the right side corresponds to the nearer.

Since $r(w)$ is represented as a linear function of the focus position w , the spatio-focal image is partitioned into four regions by the following two lines,

$$s = \frac{d+l}{v}w - l, \quad s = \frac{d-l}{v}w + l. \quad (6)$$

The blurred regions are labeled (i) and (ii) where the intensity varies between L_1 and L_2 . On the other hand, the constant intensities L_1 and L_2 are observed in the regions (iii) and (iv), respectively.

2.5.1. Intensity Change along Pixel Correspondence Line.

Suppose R is a point in the L_1 side slightly displaced from the focused edge point Q . When we increase w , the intensity change along pixel correspondence line of R forms a peak; that is, the intensity first rises up to L_1 in (i), then keeps L_1 in (iii), and finally falls down from L_1 in (ii). On the contrary, the intensity change forms a valley if R exists in the L_2 side.

This consideration suggests that we may determine which side of the edge, L_1 or L_2 , the point belongs to.

In other words, we can detect the edge points by the following procedure:

1. Measure a certain feature that represents whether the intensity change along pixel correspondence line has a peak or valley.
2. Synthesize an image that represents the spatial distribution of the measured feature.
3. Detect those points that are located between “peak” and “valley” features in the synthesized image.

Based on this idea, we have developed the following edge detection algorithm in which the cumulative difference of intensities is used as a stable feature to characterize the intensity change along pixel corresponding line.

Algorithm 1: Edge Detection

- *Step 1.* Apply the following steps 2 and 3 for every pixel (x_0, y_0) in the image $B(x, y, f)$.
- *Step 2.* Compute a base line intensity function $B_{\text{base}(x_0, y_0)}(w)$ from two intensity values $B(x_0, y_0, f)$ and $B(mx_0, my_0, w_{\text{max}})$. We consider that $B_{\text{base}(x, y)}(w)$ accounts for the intensity reduction due to image magnification, lens vignetting, and so on.

$$\begin{aligned} B_{\text{base}(x_0, y_0)}(w) &= \frac{B(mx_0, my_0, w_{\text{max}}) - B(x_0, y_0, f)}{w_{\text{max}} - f}(w - f) \\ &\quad + B(x_0, y_0, f) \end{aligned} \quad (7)$$

- *Step 3.* Compute the intensity difference between $B(mx_0, my_0, w)$ and $B_{\text{base}(x_0, y_0)}(w)$ for each focus position w , then sum up all difference values.

$$\begin{aligned} B_{\text{accumulate}}(x_0, y_0) &= \sum_{w=f}^{w_{\text{max}}} (B(mx_0, my_0, w) - B_{\text{base}(x_0, y_0)}(w)) \end{aligned} \quad (8)$$

- *Step 4.* Synthesize an image $B_{\text{accumulate}}(x, y)$ where each pixel (x_0, y_0) has $B_{\text{accumulate}}(x_0, y_0)$ as its value. We call the image $B_{\text{accumulate}}(x, y)$ an *accumulated defocus image*.
- *Step 5.* Detect zero-crossing points in the accumulated defocus image. The $B_{\text{accumulate}}(x, y)$ is positive if pixel (x, y) is located in the L_1 region, negative if in the L_2 , and zero if just on the edge.

2.5.2. Isointensity Lines in Blurred Regions. Depth estimation is performed based on the following idea.

1. Compute the spatio-focal image for each detected edge point.
2. Find a pair of straight lines delimiting four regions labeled (i), (ii), (iii) and (iv) in Fig. 3.
3. Compute the depth of the edge point from the focus position w of intersection point of those delimiting lines.

Since it is difficult to find the intensity boundaries of four regions accurately, we used a pair of isointensity lines instead of the delimiting lines. In what follows, we first examine the properties of isointensity lines and then describe the depth estimation algorithm.

Suppose a point C is located in the L_1 side and its displacement ρ_0 from the edge is smaller than the radius $r(w_0)$ of the blurring kernel $K(x, y, w_0)$. In such a situation, the blurring kernel centered at C covers a part of the L_2 side. Thus, the intensity L_C of the point C is represented by

$$L_C = \tau L_1 + (1 - \tau) L_2, \quad (9)$$

where

$$\tau = \int_{-\rho_0}^{r(w_0)} \int_{-r(w_0)}^{r(w_0)} K(x, y, w_0) dx dy.$$

The intensity L_C is a function of τ which is determined by the ratio of ρ_0 to $r(w_0)$ because $K(x, y, w_0)$ is rotationally symmetric. Therefore, the observed intensity keeps constant as long as $\frac{\rho(w)}{r(w)} = \frac{\rho_0}{r(w_0)}$ is satisfied even if the focus position w varies. In other words, $\rho(w) = \lambda_1 r(w)$ is an isointensity condition in the blurred region of the L_1 side, where λ_1 denotes a constant value between $0 < \lambda_1 < 1$.

Using the isointensity condition, we obtain the following two lines in the blurred regions of the L_1 side.

$$s = \begin{cases} \frac{d - \lambda_1 l}{v} w + \lambda_1 l & f \leq w < v \\ \frac{d + \lambda_1 l}{v} w - \lambda_1 l & v < w \leq w_{\max} \end{cases} \quad (10)$$

These two lines meet at the focused edge point $Q(v, d)$ and the isointensity lines form a “V” shape which is depicted by bold lines in Fig. 4. Similarly, we have a pair of isointensity lines in the blurred regions of the

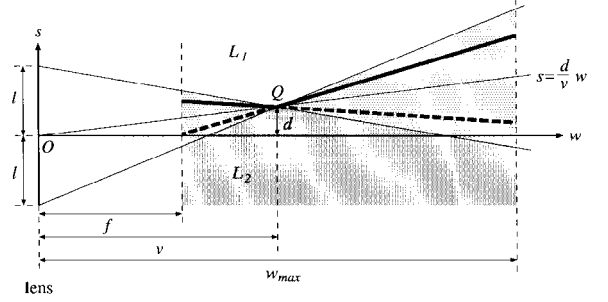


Figure 4. Isointensity lines drawn in the blurred regions.

L_2 side by using $\rho(w) = \lambda_2 r(w)$.

$$s = \begin{cases} \frac{d + \lambda_2 l}{v} w - \lambda_2 l & f \leq w < v \\ \frac{d - \lambda_2 l}{v} w + \lambda_2 l & v < w \leq w_{\max} \end{cases} \quad (11)$$

The isointensity lines in the L_2 side form a “A” shape bent at the focused edge point $Q(v, d)$, which are depicted by broken bold lines in Fig. 4.

We suppose two points C_1 and C_2 located at a focus position w in the blurred regions of the L_1 and L_2 sides, respectively. If the displacement ρ_1 of C_1 from the edge is the same as that ρ_2 of C_2 , the value of τ is identical between C_1 and C_2 . Therefore, we have the following equations.

- $|L_1 - L_{C_1}| = |L_2 - L_{C_2}|$, where L_{C_1} and L_{C_2} denote intensities of C_1 and C_2 , respectively.
- $\lambda_1 = \lambda_2$ because $\rho_1 = \rho_2$ that implies $\lambda_1 r(w) = \lambda_2 r(w)$.

This means that the first equation of Eq. (10) coincides with the second one of Eq. (11) and the second one of Eq. (10) is identical with the first one of Eq. (11). In other words, the V-shaped isointensity lines of $L_1 - \delta$ in the L_1 side and the A-shaped ones of $L_2 + \delta$ in the L_2 side are merged into a pair of straight lines crossing at the focused edge point $Q(v, d)$. Therefore, the focus position v of the focused edge point $Q(v, d)$ is determined from the intersection point of such a pair of isointensity lines. Based on this idea, we have developed the following algorithm.

Algorithm 2: Depth Estimation

- *Step 1.* Generate a spatio-focal image for every detected edge point.

- *Step 2.* Find maximum and minimum intensities in the spatio-focal image. They are denoted by L_1 and L_2 , respectively.
- *Step 3.* Partition the image into halves by the pixel correspondence line of the current edge point.
- *Step 4.* Let δ be a value between $0 < \delta < \frac{L_1 - L_2}{2}$. Apply binarization to the spatio-focal image using $L_1 - \delta$ as a threshold value for the L_1 side and $L_2 + \delta$ for the L_2 . Then, detect boundary points of thresholded regions.
- *Step 5.* Generate a pair of straight lines by applying the least square fitting method to those boundary points, and find their intersection point.
- *Step 6.* Repeat the Steps 4 and 5 using several different δ s, and compute the average v of focus positions of the intersection points.
- *Step 7.* Compute the depth u from v by using Eq. (1).

This algorithm enables us to generate a depth map of edge points.

Here it is necessary to address a size problem of spatio-focal images. The smaller size is desirable to apply the step edge model as an approximation of various kinds of actual edges. The smaller size, however, poses the less stable result in the depth estimation even though the algorithm involves integral computations such as model fitting and averaging. This is analogous to the window problem discussed in Section 1. Nevertheless, we can find an appropriate size of spatio-focal images; that is, the width is determined from the focus range within which a single peak or valley is included in the intensity change along pixel correspondence lines, and the height is determined from the spatial range between the nearest peak and valley from the edge point in the accumulated defocus image.

3. Experimental Results and Performance Evaluation

We performed experiments in a laboratory room by using a video camera (SONY DXC-325) with a zoom lens (CANON VCL-810BX). The lens parameters were selected to enhance the depth resolution; the largest aperture 1.4 and the longest focal length 80 mm were used. The focus position was varied manually step by step to take multi-focus images. In order to reduce random noise, we applied a temporal averaging to 64 frames taken at the same focus position. The spatial and intensity resolutions of images were 512×480 pixels and

256 gray levels, respectively. We performed a camera calibration to determine some intrinsic parameters of our camera system.

3.1. Edge Detection

Figure 5 shows a scene consisting of two planes, each of which includes black and white rectangular regions. The left plane was placed at 2.4 m from the camera lens, and the right at 5.9 m. There are five edge segments labeled a number from 1 to 5 as shown in Fig. 5. Note that edge segments 1, 2, and 3 are occluding edges, whereas those labeled 4 and 5 are surface edges. Thirty one images were taken by varying the focus parameter in the range from infinite distance to 1.5 m. Figure 6 shows five images of them; the focused distance decreases from the left image to the right one. The farther plane is almost in focus in the second image and the nearer is almost in focus in the fourth.

Figure 7 depicts two examples of intensity change along pixel correspondence line; the upper and lower curves show the intensity changes in the white and black sides of the edge segment 4, respectively. The broken straight lines represent the base line intensity function computed by Eq. (7). This figure demonstrates the importance of intensity compensation using

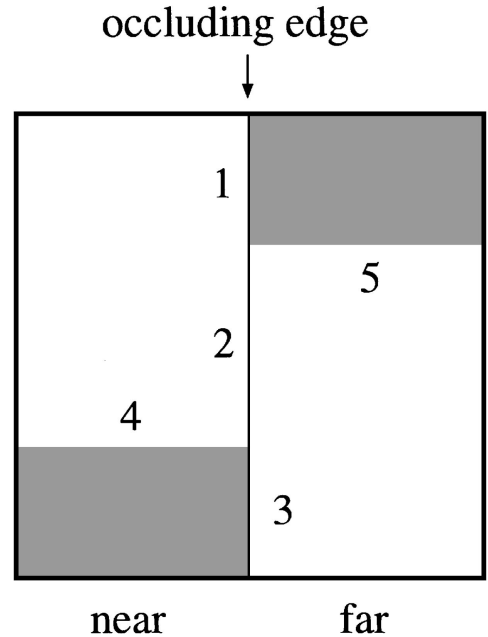


Figure 5. Experimental scene consisted of two planes placed at different depths. Edge segments labeled 1, 2 and 3 are occluding edges and those labeled 4 and 5 are surface ones.

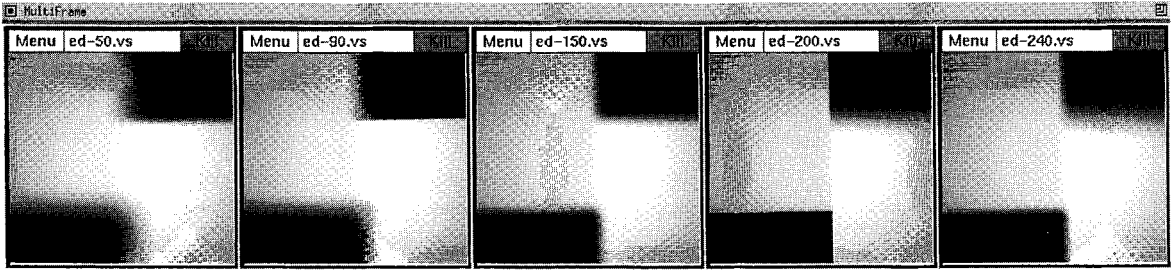


Figure 6. Five images selected from a sequence of 31 multi-focus images of Fig. 5. The focused distance decreases from the left image to the right one; the farther plane is almost in focus in the second image and the nearer is almost in focus in the fourth.

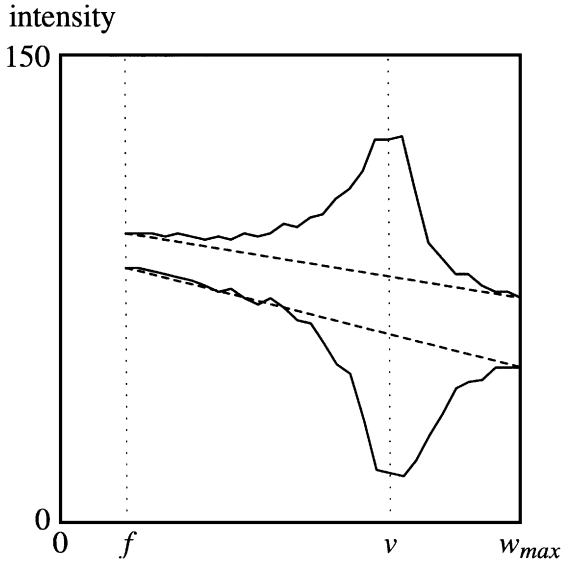


Figure 7. Intensity changes along pixel correspondence line of two points around the edge segment 4.

$B_{\text{base}(x,y)}(w)$ when comparing the intensities observed at different focus positions.

Figure 8 shows the accumulated defocus image (left) and the edge image (right) obtained by the Algorithm 1. The edge segments 1, 3, 4 and 5 were detected successfully, but the segment 2 was difficult to find. This reason is as follows: the intensity difference across the edge segment 2 was very small, and thus zero-crossing points located between small positive and negative values in the accumulated defocus image. Such edge point candidates were eliminated as a noise by thresholding.

Note that the blurring phenomenon at the left end of the edge segment 5 was complicated. While the segment 5 was in focus in the second image of Fig. 6, its left end was unclear because of the extension of the blurred segments 1 and 2. Obviously, the segment 5 itself was out of focus in other images. Therefore, no sharp image of the left end of segment 5 was included in the multi-focus images. We detected,

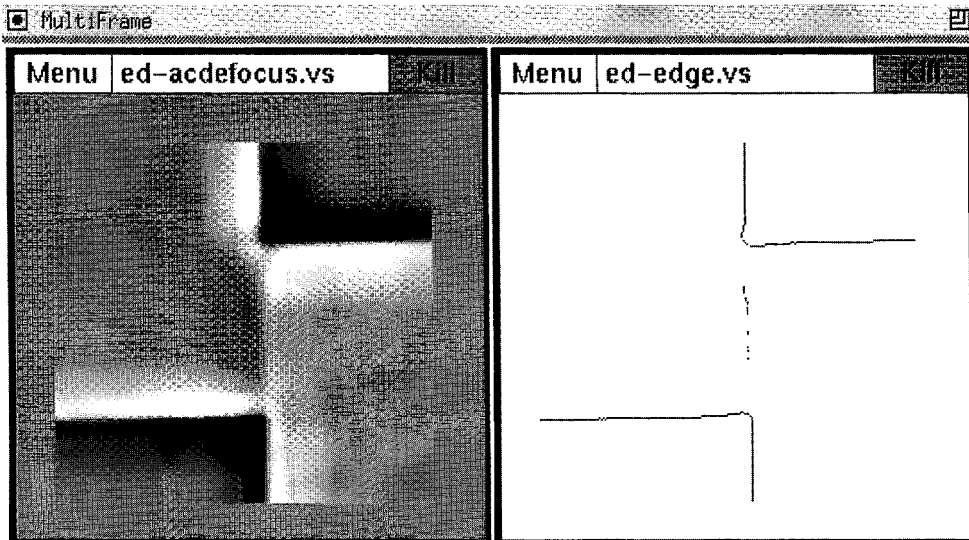


Figure 8. Accumulated defocus image (left) and edge image (right) computed from the entire image sequence.

however, the whole of segment 5 as shown in the edge image. This demonstrates that our integration-based algorithm works successfully even in such a complicated situation.

3.2. Depth Estimation

Figure 9 shows the spatio-focal image in which a pair of isointensity lines is superimposed. We used three different δs , i.e., 16%, 23%, 30% of $\frac{L_1-L_2}{2}$, for thresholding in the Algorithm 2. The depth of each edge point was computed from the average focus position of three intersection points. Based on the number of frames and the intensity distribution of accumulated defocus image, we determined the size of spatio-focal images, i.e., 31 (width) \times 16 (height) as shown in Fig. 9.

Table 1 shows the average and standard deviation of depth and edge position. Note that the accuracy of edge position was evaluated from the consistency between three intersection points in the spatio-focal image.

From Table 1, we have the following observations. First, the precise measurement was achieved at the most part of the segments 1, 3, 4 and 5: 0.2% error at 2.4 m and 0.3% at 5.8 m in the depth estimation and 0.5 pixel position fluctuation in the edge detection. These results suggest that our method would show a high performance in some situations where the real edges are observed as a step edge. Note that the position stability of most part of segments 1 and 3 is very high, whereas those average depths do not coincide with each other.

Table 1. Quantitative evaluation of the accuracy of depth estimation and edge detection.

| Edge segment | Number of edge points | Depth (mm) | | Position (pixel) | |
|--------------------|-----------------------|------------|------------|------------------|------------|
| | | \bar{u} | σ_u | \bar{s} | σ_s |
| 1 (center + upper) | 153 | 2412 | 5.4 | -0.36 | 0.59 |
| 1 (lower) | 51 | 2443 | 7.2 | 1.39 | 0.48 |
| 2 (all) | 42 | 2466 | 46.2 | 0.73 | 1.05 |
| 3 (upper) | 34 | 2432 | 6.6 | -2.02 | 0.43 |
| 3 (center + lower) | 145 | 2442 | 3.4 | -0.77 | 0.68 |
| 4 (center + left) | 425 | 2431 | 5.2 | 0.10 | 0.71 |
| 4 (right) | 19 | 2458 | 11.6 | 2.01 | 0.16 |
| 5 (center + right) | 282 | 5825 | 20.2 | -0.35 | 0.56 |
| 5 (left) | 69 | 5717 | 56.3 | -1.64 | 0.70 |

This means that the occluding edge was almost perpendicular to the optical axis of the lens but it was slightly slanted. To confirm it, we computed the average depth and standard deviation of eight sub-segments of the edge segments 1 and 3. Figure 10 shows an ascending straight line computed by least square fitting to the average depth values of the sub-segments. This shows that the occluding edge leaned a little toward the camera.

Second, accuracy degradation appears at the end of segments 1, 3 and 4 where the meet of multiple edge segments forms “T-junctions”. Our simple edge model does not hold for such a T-junction and complicated intensity change would be found on the pixel

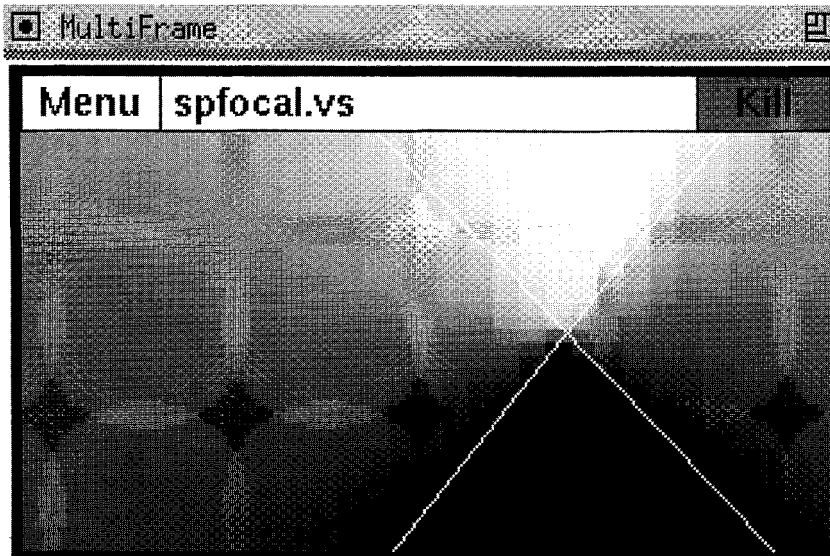


Figure 9. Depth estimation using a pair of isointensity lines in the spatio-focal image.

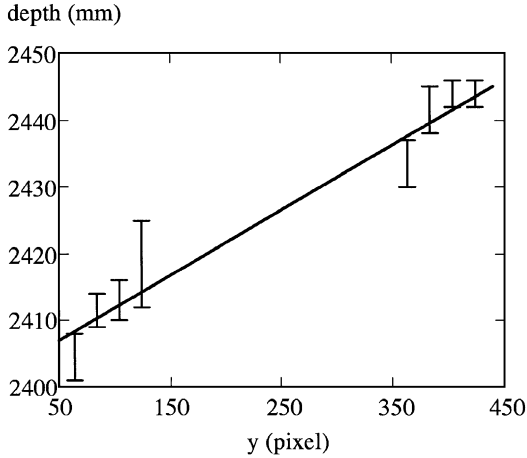


Figure 10. Depth change along the occluding edge.

correspondence lines as well as in the spatio-focal images. However, approximate depth and edge position were obtained even around such a multiple edge complex. This suggests that the finer analysis of the pixel correspondence lines would improve the precision of edge detection and depth estimation results.

Lastly, poor accuracy was found in the segments 2 and 5 (left). This comes from the low contrast between the adjacent regions for the segment 2 and the disturbance of blurred edge segments 1 and 2 for the segment 5 (left), as described above.

3.3. Experiment Using Natural Objects

To evaluate the applicability of the proposed method to natural scenes, we prepared two objects of a monster animal “godzilla” placed at about one meter apart each other, as shown in Fig. 11. Figure 12 shows the

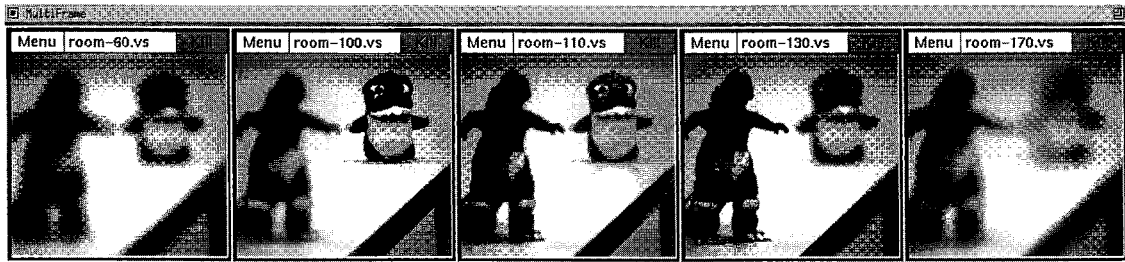


Figure 11. Five images selected from a sequence of 21 multi-focus images of two objects on a table.

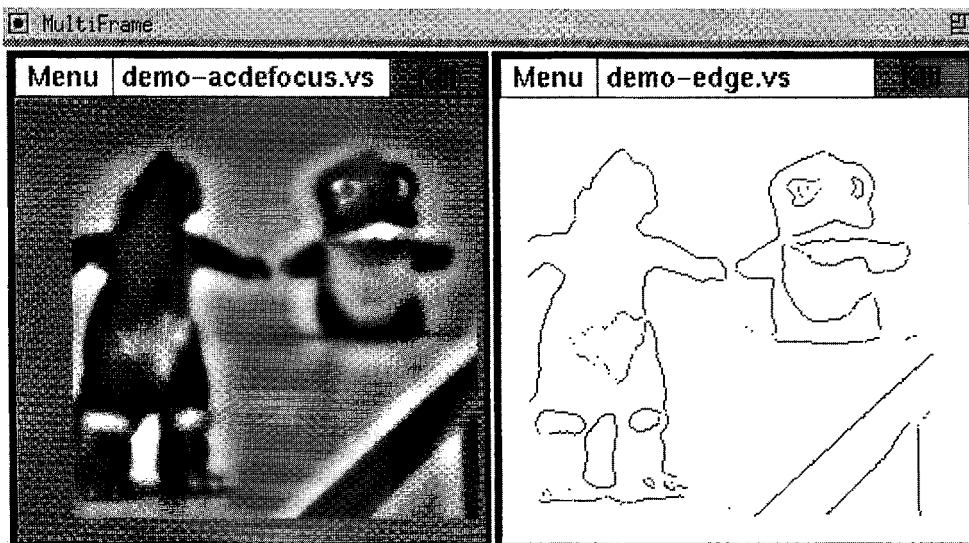


Figure 12. Accumulated defocus image (left) and edge image (right) computed from the entire image sequence.

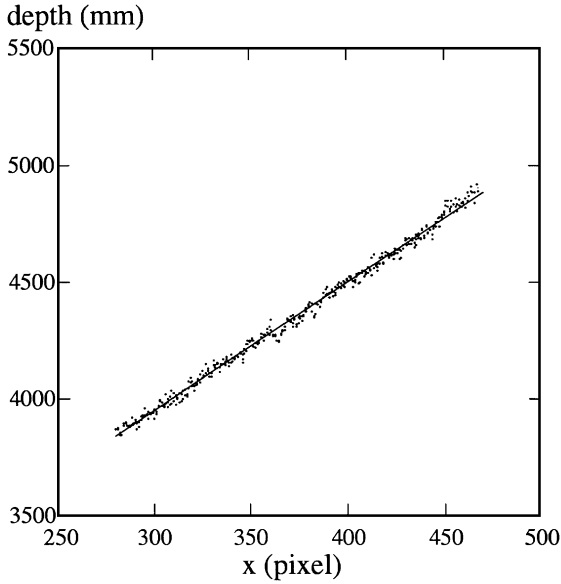


Figure 13. Depth change along the table edge computed by the proposed method.

accumulated defocus image (left) and the edge image (right). Although the scene included curved edges that were different from the simple edge model, most of the edge points were detected stably without any modification of the proposed algorithms.

The standard deviation of depth of the nearer object was within 100 mm which was consistent with the actual depth of the object. However, it was difficult to evaluate quantitatively the depth accuracy of each individual edge point because the shape of the object was not simple. Instead of such complex objects, we examined the table edge whose depth varied linearly. Note that the blur of depth-varying edge was different from that of the simple edge model. Figure 13 shows the estimated depth and a line generated by the least square fitting, whose correlation coefficient and residual error sigma were 0.99 and 3.95 mm, respectively. This demonstrated that the error of depth estimation at 4 m distance was less than 0.1%.

For further evaluation of our method, we applied conventional DFF and DFD methods to the same set of the multi-focus images. Figure 14 shows a DFF result of estimated depth along the table edge. The Tenengrad operator (Krotkov, 1987) worked successfully to select the focused frame, but the depth resolution was quite low. This means that DFF method requires a greater number of frames than ours to improve the depth accuracy. On the other hand, based on the Grossmann's

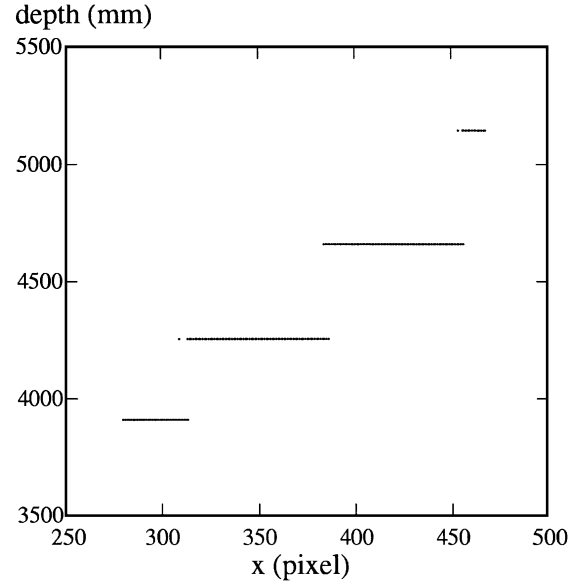


Figure 14. Depth change along the table edge computed by a DFF method.

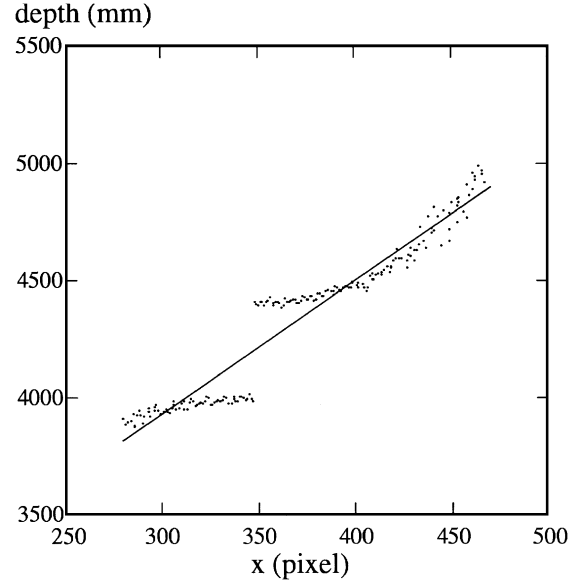


Figure 15. Depth change along the table edge computed by a DFD method.

method (Grossmann, 1987), we computed the depth along the table edge from the width of blur measured in the twelfth frame. Figure 15 shows the estimated depth of each edge point and a line generated by the least square fitting, whose residual error sigma was 14.82 mm. While we have to consider that the depth

was recovered from a single image, the result shows a serious drawback involved in DFD method; the depth accuracy around the focused edge point ($x = 350$ in Fig. 15) was deteriorated due to the intrinsic blur of the optical systems. Our method overcomes this problem by integrating multiple images. From these comparative studies, it is expected that the proposed method would show a higher performance than those DFF and DFD methods.

4. Conclusion

This paper has described an integration-based method to extract reliable edge and depth information from multi-focus images. Although the proposed method requires multiple images and the strict edge model, the computation scheme involving accumulation and model fitting has demonstrated that high stability and precision are achieved in the edge detection and depth estimation.

Further investigations should be done on the following topics:

- *Detailed Analysis of the Intensity Change Along Pixel Correspondence line:* The method proposed in this paper is based on a single straight step edge model. Natural scenes, however, contain a variety of edges of different properties as well as textured regions. More detailed analysis of the intensity change along pixel correspondence line will yield useful information to characterize such properties.
- *Development of a Practical Camera System:* To implement our method in a camera system, we should analyze how many and how different focus positions are necessary to make the practical system.

Notes

1. Some researchers obtained blurred images by controlling the lens aperture instead of the focus in order to avoid the pixel correspondence problem discussed in Section 2.3 (Pentland, 1987; Subbarao and Wei, 1992; Ens and Lawrence, 1993).
2. The blurring phenomenon around an occluding edge is complicated because two different blurring kernels should be considered.

It is shown, however, that the blurred image of an occluding edge is identical with that computed by convolution using a blurring kernel corresponding to the nearer surface if the light intensity of the occluded surface is uniform around the occluding edge (Asada et al., 1995).

References

- Asada, N. and Matsuyama, T. 1992. Color image analysis by varying camera aperture. In *Proc. International Conference on Pattern Recognition*, pp. A466–A469.
- Asada, N., Fujiwara, H., and Matsuyama, T. 1995. Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model. In *Proc. International Conference on Computer Vision*, pp. 150–155.
- Darrell, T. and Worn, K. 1988. Pyramid based depth from focus. In *Proc. Computer Vision and Pattern Recognition*, pp. 504–509.
- Ens, J. and Lawrence, P. 1993. An investigation of methods for determining depth from focus. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(2):97–108.
- Grossmann, P. 1987. Depth from focus. *Pattern Recognition Letters*, 5:63–69.
- Jarvis, R.A. 1983. A perspective on range finding technique for computer vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5(2):122–139.
- Krotkov, E. 1987. Focusing. *International Journal of Computer Vision*, 1(3):223–237.
- Lai, S.H., Fu, C.W., and Chang, S. 1992. A generalized depth estimation algorithm with a single image. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(4):405–411.
- Lighthart, G. and Groen, C.A. 1982. A comparison of different autofocus algorithms. In *Proc. International Conference on Pattern Recognition*, pp. 597–600.
- Nair, H.N. and Stewart, C.V. 1992. Robust focus ranging. In *Proc. Computer Vision and Pattern Recognition*, pp. 309–314.
- Nayar, S.K. 1992. Shape from focus system for rough surface. In *Proc. Image Understanding Workshop*, pp. 539–606.
- Pentland, A.P. 1987. A new sense for depth of field. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(4):523–531.
- Subbarao, M. 1988. Parallel depth recovery by changing camera parameters. In *Proc. International Conference on Computer Vision*, pp. 149–155.
- Subbarao, M. and Wei, T.C. 1992. Depth from defocus and rapid autofocus: A practical approach. In *Proc. Computer Vision and Pattern Recognition*, pp. 773–776.
- Subbarao, M., Wei, T.C., and Surya, G. 1994. Focused image recovery from two defocused images recorded with different camera settings. In *Proc. Computer Vision and Pattern Recognition*, pp. 786–791.
- Xiong, Y. and Shafer, S.A. 1993. Depth from focusing and defocusing. In *Proc. Computer Vision and Pattern Recognition*, pp. 68–73.