DEPTH FROM DEFOCUS TECHNIQUE BY MUTUAL REFOCUSING

Kazumi TAKEMURA and Toshiyuki YOSHIDA

Dept. of Information Science, Graduate school of Eng., University of Fukui, Fukui 910-8507, Japan

Abstract - This paper proposes a new DFD technique based on mutual refocusing. Our technique exploits the fact that two blurred images coincide if they are refocused (blurred) by the other's focus setting. It requires no lens parameters under the use of an object-side telecentric lens. The simulation results given in this paper demonstrate the validity of our approach.

Keywords : Depth estimation, Depth from defocus(DFD), Multi-focus images, Refocusing, Error reduction

I. INTRODUCTION

"Depth from Focus (DFF)" and "Depth from Defocus (DFD)" have been known as one of the 3-D shape reconstruction techniques [1–6]. The techniques commonly estimate a depth profile to the surface of a target object by using a set of multi-focus images, which are captured by varying the focus setting to the target object.

One of the common DFF/DFD techniques utilizes a focus measure that evaluates the degree of focus for each block. In Ref. [5], the authors have proposed a DFD technique that estimates the depth for each block by finding the perfectly focused position based on a focus measure. Although the technique requires no lens parameters, e.g., the focal length and the aperture radius, unlike Refs. [3, 4, 6], it suffers from an inherent error due to the passband width of the bandpass filter (BPF) utilized in evaluating the focus measure, as reviewed in the next section.

In Ref. [6], the authors have proposed a unified approach of the DFD and stereo matching, where multi-view and multi-focus images are utilized to improve the accuracy in depth estimation. As also reviewed in the next section, the technique utilizes the all-focused image recovered from multi-focus images. In the case that the multi-focus images are heavily unfocused and blurred, however, the recovered all-focused image suffers from artifacts due to quantization noise, which leads to a large depth estimation error.

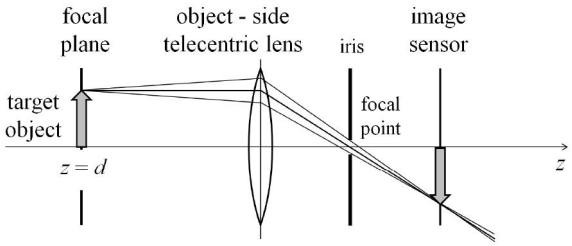
In this paper, a new DFD technique is proposed by exploiting the fact that two multi-focus images coincide if they are blurred with their point spread functions exchanged. Since the technique requires neither of the all-focused image nor the BPF in evaluating the focus measure, the estimation errors in Refs. [5, 6] can be reduced without any lens parameters.

In what follows, Sect. II briefly reviews the DFD techniques in Refs. [5, 6] and summarizes their advantages and disadvantages. Sect. III elaborates the proposed technique, and Sect. IV gives the experimental results. Finally, Sect. V concludes this paper.

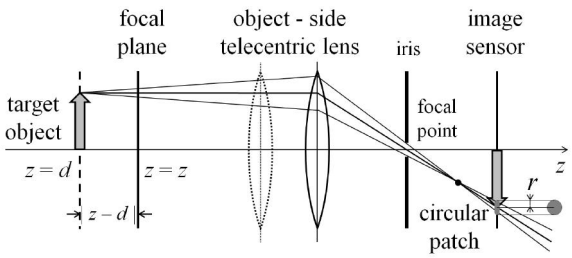
II. RELATED WORKS

A. Telecentric lens system and point spread fuction

Fig. 1 (a) shows a DFD system that captures a set of multi-focus images of the target object placed at $z = d$. The focus setting can be varied by shifting the whole acquisition system along with the z axis wile maintaining the distance between the lens and the image sensor. To avoid a variation of the magnification factor due to the variation of the lens-object distance, the object-side telecentric lens is employed in the system.



(a) focal plane on the target object



(b) focal plane off the target object

Fig. 1 : DFF/DFD system.

Fig. 1 (a) shows the case the target object is perfectly focused, i.e., the acquisition system is placed at which its focal plane is just on the target object, where no blurring is observed on the image sensor. If the system is shifted so that its focal plane is placed at $z = z$, a single point entering into the lens is blurred to a circular patch with radius r on the image sensor, as shown in Fig. 1 (b). Under the assumption that the object-side telecentric lens is employed, the blur radius r is proportional to the distance between the focal plane and the target object, which can be represented as

$$r = k' \left| z - d \right|, \tag{1}$$

where k' is a constant.

The blurring profile of the circular patch is modeled by a point spread function (PSF): this paper assumes the Gaussian PSF given as

$$p_{\sigma^2}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{2}$$

978-1-5386-2615-3/18/$31.00

where σ^2 denotes the blurring parameter. Since σ is proportional to the blur radius r,

$$\sigma = k\,|z - d|\,, \tag{3}$$

where k is a constant.

As the acuired image in Fig. 1 (b) is a blurred version of the all-focused image $f(x, y)$ for the target object, it can be represented as

$$f(x, y) **p_{\sigma^2}(x, y), \tag{4}$$

where $**$ stands for a convolution.

B. DFD technique in Ref. [5]

The technique in Ref. [5] utilizes three multi-focus images

$$f_n(x, y) = f(x, y) **p_{\sigma_n^2}(x, y) \tag{5}$$

acquired for $n = 0, 1, 2$ with the target object and the focal plane placed at $z = d$ (unknown) and $z = z_n$ $(n = 0, 1, 2)$, respectively, as shown in Fig. 2. The blurring parameter σ_n $(n = 0, 1, 2)$ in Eq. (5) is given by

$$\sigma_n = k\,|z_n - d| \quad (n = 0, 1, 2). \tag{6}$$

To evaluate the block-wise focus measure, f_ns are first filtered with a 2-D circularly symmetric BPF $h(x, y)$ as

$$g_n(x, y) = f_n(x, y) **h(x, y) \tag{7}$$

and then

$$m(n) = \sum_{(x,y)\in B} g_n^2(x, y) \tag{8}$$

is calculated as the focus measure for each block B on f_n $(n = 0, 1, 2)$.

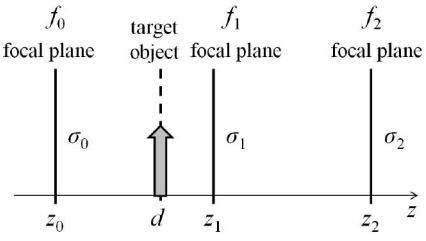


Fig. 2 : DFD technique in Ref. [5].

Eq. (8) can be equivalently expressed in the frequency domain as

$$m(n) = \iint_{-\pi}^{\pi} |F_n|^2\,|H|^2\,e^{-\sigma_n^2(\omega_x^2+\omega_y^2)}\,d\omega_x d\omega_y, \tag{9}$$

where F_n and H are the Fourier transforms of f_n and h, respectively. If we can assume that the passband width of the 2-D circularly symmetric BPF is quite narrow to extract the frequency components on the ring with radius ω_r, Eq. (9) can be simplified into

$$m(n) = e^{-\sigma_n^2\,\omega_r^2} \iint_{\omega_x^2+\omega_y^2=\omega_r^2} |F_n|^2\,d\omega_x d\omega_y, \tag{10}$$

where the Gaussian factor has been excluded in the integration because it is a constant in the integration on the ring $\omega_x^2 + \omega_y^2 = \omega_r^2$.

Eq. (10) indicates that the focus measure $m(n)$ varies along with the Gaussian function $e^{-\sigma_n^2\,\omega_r^2}$ under the assumptions that the telecentric lens is employed and that the passband width of the BPF is quite narrow. The unknown position of the target object d can therefore be estimated by approximating the discrete points $m(n)$ with the Gaussian curve and by finding its maximum point. Note that no lens parameters are necessary in this technique.

However, actual BPFs inevitably have a finite passband width, which makes the theoretical curve of $m(n)$ deviate from the Gaussian function, resulting in a large error in depth estimation. Such a difficulty is notable in particular in the case where the acquisition interval for multi-focus images is large, as demonstrated in Sect. IV.

C. DFD technique in Ref. [6]

Ref. [6] has proposed a unified approach of the DFD and stereo matching. In this section, the technique is briefly reviewd from a DFD viewpoint.

Let $f_n(x, y)$ $(n = 0, 1)$ be two multi-focus images of the target object acquired as in Fig. 2. The technique in Ref. [6] first recovers the all-focused image $f(x, y)$ by assuming the unknown object position as $z = d'$. Actually, the Fourie transform of the all-focused image

$$F = \frac{\sum_{n=0}^{1} \overline{P}_n\,F_n}{\sum_{n=0}^{1} |P_n|^2 + c^2} \tag{11}$$

is calculated, where P_n is the Fourier transform of the psf $p_n(x, y)$ for n, and c is a constant to prevent the denominator from being zero. $f(x, y)$ can be recovered by taking the inverse Fourier transformation of Eq. (11). Then, the recovered $f(x, y)$ is again blurred by the PSFs $p_n(x, y)$ $(n = 0, 1)$ using the assumed object position d'. Note that the actually acquired image $f_n(x, y)$ and the reblurred version $f'_n(x, y)$ completely coincide if the assumed d' is the true value, i.e., $d' = d$. The technique in Ref. [6] thus varies d' in a prescribed range in each block to find the optimal value that makes f_n and f'_n identical.

The technique, however, suffers from the problem that the all-focused image cannot be accurately recovered in the case the multi-focus images are heavily unfocused and blurred: in such a case, since the denominator of Eq. (11) gets close to zero, the quantization noise in the multi-focus image F_n is amplified, which leads to a large estimation error.

III. PROPOSED TECHNIQUE

This section proposes a new DFD technique that overcomes the difficulties in Refs. [5, 6].

A. Fundamental idea

Fig. 3 illustrates the fundamental idea of the proposed DFD technique. Let $f(x, y)$ and $f_n(x, y)$ $(n = 0, 1)$ be the all-focused and the acquired images for the target object as in

Eqs. (5) and (6). Then as in Fig. 3, f_0 and f_1 are reblurred by the PSF $p_{\sigma'^2_1}$ and $p_{\sigma'^2_0}$, respectively, where σ'_1 and σ'_0 are given as

$$\sigma'_n = k' \left| z_n - d' \right| \quad (n = 0, 1) \tag{12}$$

for the assumed object position d' and the constant k'. Such reblurring yields

$$\begin{aligned} f_{01}(x,y) &= f_0(x,y) **p_{\sigma'^2_1}(x,y) \\ &= f(x,y) **p_{\sigma^2_0}(x,y) **p_{\sigma'^2_1}(x,y) \end{aligned} \tag{13}$$

$$\begin{aligned} f_{10}(x,y) &= f_1(x,y) **p_{\sigma'^2_0}(x,y) \\ &= f(x,y) **p_{\sigma^2_1}(x,y) **p_{\sigma'^2_0}(x,y). \end{aligned} \tag{14}$$

Note that f_{01} and f_{10} are strictly identical if the assumed object position d' and the constant k' are the true values, i.e., $d' = d$ and $k' = k$.

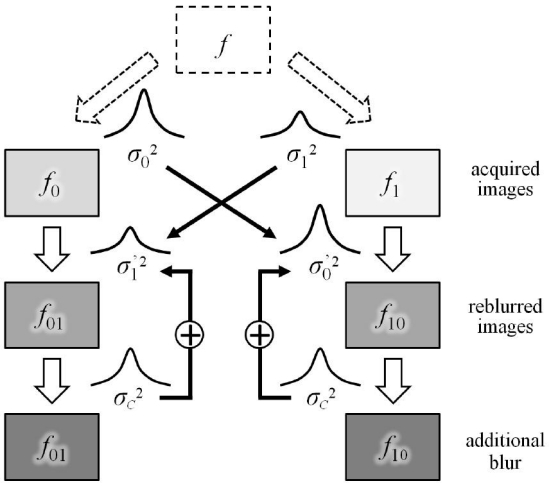


Fig. 3 : Proposed DFD technique

Based on this fact, the true object position d and the constant k can be estimated by minimizing the error criterion

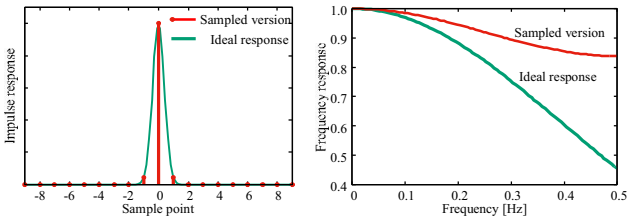
$$E(d', k') = \sum_{x,y} \left(f_{01}(x,y) - f_{10}(x,y) \right)^2 \tag{15}$$

with respect to d' and k'. Actually, a nonlinear minimization technique such as the Gauss-Newton technique can be employed to find the d' and k' that minimizes $E(d', k')$.

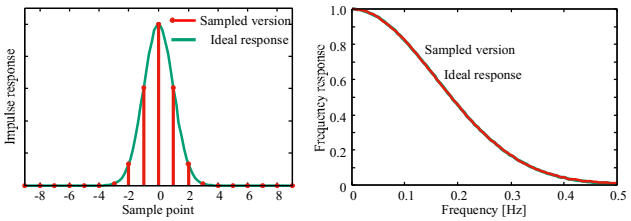
Unfortunately, however, a direct minimization of Eq. (15) suffes from two problems. The following subsections elaborate the problems, and give the revised alogorithm.

B. Compensation for PSF

Since the acquired images f_0 and f_1 are discrete ones, the corresponding PSFs $p_{\sigma'^2_1}$ and $p_{\sigma'^2_0}$ should be sampled versions of Eq. (2). Fig. 4 illustrates the sampled versions of Eq. (2) with $\sigma = 0.4$ and 1.0 and their corresponding frequency responses. 1-D responses are shown in Fig. 4 for simplicity. The frequency response for $\sigma = 0.4$ (a) deviates from its ideal response $e^{-\sigma^2\omega^2/2}$ while that for $\sigma = 1.0$ coincides with the ideal one in (b). Such a deviation can be attributed to the sampling process for very narrow Gaussian function as in Fig. 4 (a). Since the deviation causes a large estimation error for d' and k', it should be compensated.



(a) $\sigma = 0.4$



(b) $\sigma = 1.0$

Fig. 4 : Sampled PSFs and their frequency responses

We now focuse on the fact that no such deviation is observed for relatively large σ in Fig. 4 (b). Therefore, additional blur with the blurring parameter σ_c is introduced at the bottom of Fig. 3. Then, σ_c is combined with σ'_1 and σ'_0, respectively, which leads to a modification of the PSF in Eq. (2) into

$$P_{\sigma^2+\sigma_c^2}(x,y) = \frac{1}{2\pi(\sigma^2+\sigma_c^2)} e^{-\frac{x^2+y^2}{2(\sigma^2+\sigma_c^2)}}. \tag{16}$$

Since the blurring parameter is increased from σ^2 to $\sigma^2+\sigma_c^2$, the deviation observed in Fig. 4 (a) can be avoided, leading to a reduction of the estimation error. σ_c should be selected as small as possible while reducing the deviation sufficently. $\sigma_c = 0.86$ is utilized in this paper although the optimization process is omitted due to limited space.

C. Extention to three-image approach – revised algorithm

Fig. 3 shows that the error criterion (15) takes its minimum value at which d' and k' satisfies $\sigma_0^2 + \sigma'^2_1 = \sigma_1^2 + \sigma'^2_0$. A substitution of Eqs. (6) and (12) into the condition leads to

$$\begin{aligned} & k^2(z_0 - d)^2 + k'^2(z_1 - d')^2 \\ &= k^2(z_1 - d)^2 + k'^2(z_0 - d')^2. \end{aligned} \tag{17}$$

Obviously, the optimal d' and k' cannot be uniquely detemined with the single condition.

To avoid such a difficulty, the proposed technique is extended to use three multi-focus images defined as in Fig. 2 and Eq. (5) for $n = 0, 1, 2$. First, the following six reblurred images

$$\begin{aligned} f_{01}(x,y) &= f_0(x,y) **p_{\sigma'^2_1}(x,y) \\ f_{10}(x,y) &= f_1(x,y) **p_{\sigma'^2_0}(x,y) \\ f_{12}(x,y) &= f_1(x,y) **p_{\sigma'^2_2}(x,y) \\ f_{21}(x,y) &= f_2(x,y) **p_{\sigma'^2_1}(x,y) \\ f_{02}(x,y) &= f_0(x,y) **p_{\sigma'^2_2}(x,y) \\ f_{20}(x,y) &= f_2(x,y) **p_{\sigma'^2_0}(x,y), \end{aligned}$$

are generated, and then the error criterion

$$E(d',k') = \sum_{x,y}\left(f_{01}(x,y)-f_{10}(x,y)\right)^2 + \sum_{x,y}\left(f_{12}(x,y)-f_{21}(x,y)\right)^2 + \sum_{x,y}\left(f_{02}(x,y)-f_{20}(x,y)\right)^2 \quad (18)$$

is minimized with respect to d' and k'. The three error terms in Eq. (18) guarantee a unique determination of the optimal d' and k'.

IV. EXPERIMENTS

The proposed technique was applied to a set of multi-focus images to evaluate the estimation accuracy. Fig. 5 shows the target object (slanted plane) and the focal planes for f_0, f_1, and f_2. The Gaussian noise with $\mu = 128$ and $\sigma = 20$ was mapped as a texture at the surface of the object plane, and the multi-focus images were obtained in 312×208 resolution at the focal planes $z_0 = 0$, $z_1 = 1$, and $z_2 = 2$ in Fig. 5 with $k = 1.6$. Each image was then divided into 13×13 blocks, and the depth in each block is estimated by the proposed techniques. The optimal d' and k' were obtained by a brute-force search with respect to d' and k' in this section. The technique in Ref. [5] was also applied for comparison.

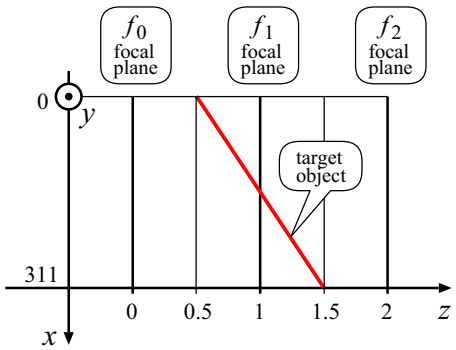


Fig. 5 : Target object and focal planes in the experiment

Table 1 lists the root mean squared error (RMSE) between the true and estimated depths under three conditions with respect to the 8 bit quantization and acquisition noise (Gaussian with $\sigma = 0.5$). Fig. 6 shows the estimated depth distributions with both of the 8bit quantization and the acquisition noise activated. Table 1 and Fig. 6 illustrates that the proposed technique gives much better result compared with Ref. [5], from which we can confirm the validity of the proposed apporach.

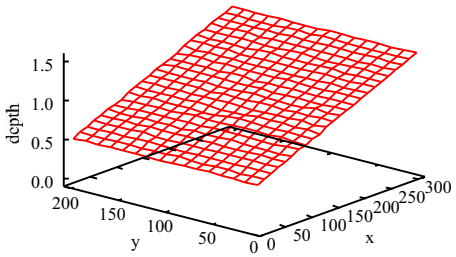
V. CONCLUSIONS

This paper has proposed a new DFD technique based on a mutual refocusing approach. The technique requires no lens parameters while it gives a lower estimation error in comparison with the focus-measure-based approach. Since our technique allows a large acquisition interval in capturing multi-focus images, the number of required images can be reduced in a piecewise approach for objects with a wide depth distribution.

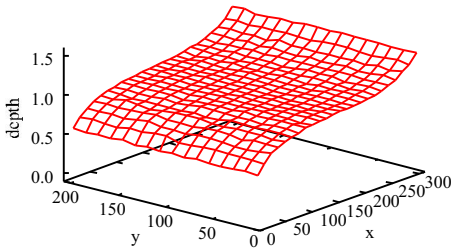
In this paper, however, the error criterion was minimized by a brute-force search, resulting in an increase of the computation cost. An efficient minimization technique such as the Gauss-Newton technique is necessary to reduce the computation cost, which is left for future work.

Table 1: RMSE of the estimated depths.

8bit Quant.	Acq. Noise	Ref. [5]	Proposed
no	no	0.110	0.00263
yes	no	0.118	0.00403
yes	yes	0.133	0.00665



(a) Proposed



(b) Ref. [5]

Fig. 6 : Estimated depth distributions.

REFERENCES

[1] Y. Xiong and S. A. Shafer, "Depth from focusing and defocusing", Proc. of CVPR '93, pp.68-73, June 1993.

[2] S. K. Nayer, "Shape from focus system", Proc. of CVPR '92, pp.302-308, June 1992.

[3] S. Chaudhuri and A. N. Rajagopalan, Depth from defocus: a real aperture imaging approach, Springer-Verlag, New York, 1999.

[4] M. Watanabe and S. K. Nayer, "Rational filters for passive depth from defocus", Int'l Journal of Computer Vision, vol. 27, no. 3, pp.203-225, 1998.

[5] T. Yoshida, "3-D shape measurement with shape from focus/defocus technique and its applications", IEICE technical Report, SIS2015-40, pp. 61-66, Dec. 2015 (in Japanese).

[6] Y. Takeda, S. Hiura, and K. Sato, "Integration of depth from defocus and stereo using coded aperture", Trans. IFICE-D, vol.J96-D, no.8, pp.1688-1700, Aug. 2013 (in Japanese).