ENHANCEMENT OF PHASE DETECTION FOR AUTOFOCUS

Chin-Cheng Chan, Shao-Kang Huang, and Homer H. Chen

National Taiwan University, Taipei, Taiwan

ABSTRACT

Phase detection autofocus (PDAF) is a technique that uses sensors on left and right pixels to determine the relative position between the object and the focal plane. When an image is out of focus, a shift between the assembled left and right pixels is resulted, which can be used to determine the lens movement during the autofocus process. However, the presence of noise and blur often affects the accuracy of shift estimation and hence the performance of autofocus. In this paper, we propose a method to enhance phase detection for such sensors. A Gaussian filter is applied to overcome the adverse effect of image noise and blur. Experiments are conducted to show that the proposed method leads to a more accurate estimate of lens movement than the traditional phase correlation.

Index Terms— Autofocus, phase detection autofocus, phase correlation.

1. INTRODUCTION

Autofocus (AF) enables a digital camera to capture sharp images without human intervention. An AF algorithm usually consists of two components, one responsible for focus measurement and the other for lens movement decision (aka search strategy). The focus measurement component computes the sharpness of a captured image or equivalently an estimate of lens distance to the in-focus position, and the search strategy component guides the lens to the in-focus position as fast as possible based on the information provided by the focus measurement component. Therefore, the quality of focus measurement is critical to an AF algorithm. The goal of this work is to investigate how to obtain reliable measurement from phase detectors.

Methods using image contrast as focus measurement for autofocus are called contrast detection autofocus (CDAF), for which the image contrast is measured in either the image domain [1], [2] or the frequency domain [3], [4] and the search strategy can be global or local. A global search strategy guarantees the optimal lens position, because the lens traverses through every possible position, at the expense of efficiency [5]. In contrast, a local search strategy, which uses only the current and a small number of previous focus measurements for lens movement decision, cannot guarantee the optimal lens position especially when the images used for focus measurement are blurry or the lighting condition is poor [6]. In such cases, the local search strategy is often unable to accurately predict the direction and the distance of lens movement.

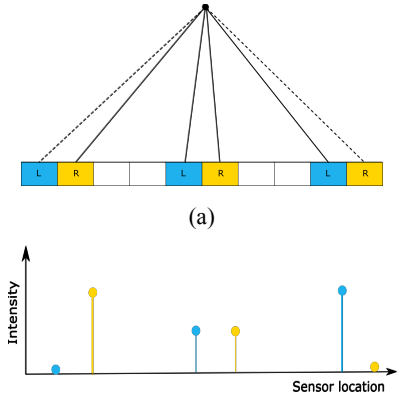
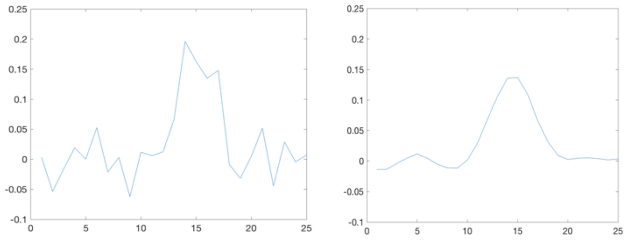


Fig. 1. (a) An arrangement of left and right pixels on an image sensor (one-dimensional view). Solid lines represent light rays that can be detected by the sensor, and dashed line represent light rays that are blocked by the masks placed on the sensor. (b) The responses to the left and right example pixels shown in (a) to an impulse.

Methods using phase difference as focus measurement for autofocus are called phase detection autofocus (PDAF) [7], [8]. Such methods alleviate the weakness of CDAF in out-of-focus regions but requires some modification of the image sensor. Masks are placed above a number of photodiodes at selected pixels of the sensor to permit the light coming from either the left side (left pixel) or the right side (right pixel) of a pixel to be detected by the photodiode underneath. The shift between the left image, formed by all the left pixels, and the right image, formed by all the right pixels, is used as the focus measurement. When an object is in-focus, the left image and the right image are almost identical, resulting in zero shift. In contrast, when an object is out-of-focus, there is a shift between the two images, and the sign of the shift (i.e., whether the object is in front of or behind the focal plane) represents the position of the object relative to the focal plane [9]. The shift can be obtained by calculating the phase correlation [10] of the left and right images. Since PDAF is able to distinguish the relative position between the object and the focal plane, it provides a



(a) (b)

Fig. 2. (a) A typical row of the phase correlation matrix. (b) Result of Gaussian filtering.

more accurate estimate of lens movement direction than CDAF for out-of-focus images.

However, similar to the case of CDAF, image noise and blur profoundly impact the accuracy of phase correlation (i.e., shift estimation) and, in turn, the performance of PDAF especially under low-light conditions. The structure of the phase detectors on the image sensor plane also affects the result of phase correlation. In this paper, we propose an algorithm to enhance the shift estimation.

The remainder of this paper is organized as follows. Section 2 describes a simplified model of PDAF, and Section 3 describes the proposed phase processing algorithm. Experiment results are presented in Section 4, followed by the conclusion in Section 5.

2. A SIMPLFIED PDAF MODEL

2.1. Model Description

Consider a single row of photodiodes of a camera (Figure 1(a)). Left pixels, marked in blue, and right pixels, marked in yellow, are placed on the sensor at a constant interval. The response of the camera to an impulse forms an image on the sensor plane—the point spread function (PSF) of the camera. The intensity distribution of a PSF is usually assumed to be uniform over a region for regular pixels (those without phase detection) [11]. The PSF for the left pixels, however, is not uniform since left pixels only sense the light coming from the left-hand side (the same for the right pixels). Figure 1(b) provides an illustration of the PSFs for the left pixels and the right pixels. These two PSFs, though not identical to each other when the image is out-of-focus, exist a spatial shift that can be used for focus measurement.

2.2. Image Formation

For a general one-dimensional scene, the images captured by the left and right pixels, respectively, can be expressed as the convolution of the scene with the left and right PSFs:

$$l(x) = s(x) * h_L(x) \tag{1}$$

$$r(x) = s(x) * h_R(x), \tag{2}$$

where x denotes pixel location, $s(\cdot)$ denotes the scene, $l(\cdot)$ and $h_L(\cdot)$, respectively, denote the image and the PSF of left pixels, $r(\cdot)$ and $h_R(\cdot)$, respectively, denote the image and the PSF of right pixels, and "$*$" denotes convolution.

Although the above discussion is for one-dimensional images, it can be extended to two-dimensional images by treating each row and each column of the images as a one-dimensional image and process them separately.

3. PROPOSED PHASE DETECTION ALGORITHM

Based on the model described in Section 2.1, a phase detection algorithm based on phase-correlation is described in this section.

3.1. Phase Correlation

Let $l(x, y)$ be the left image and $r(x, y)$ the right image. Also denote the Fourier transforms of $l(x, y)$ and $r(x, y)$ by L and R, respectively.

We formulate the relation between the left image and the right image as a pure spatial shift in the x-direction. Then $r(x, y)$ could be expressed in terms of $l(x, y)$ by

$$r(x,y) = l(x + \Delta x, y), \tag{3}$$

where Δx is dependent on the size of the PSF.

To calculate Δx of the pair of left and right images, we first calculate the phase correlation matrix $p(x, y)$ of the two images by

$$p(x,y) = \mathcal{F}^{-1}\{\frac{L \circ \bar{R}}{|L \circ \bar{R}|}\}, \tag{4}$$

where $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse two-dimensional Fourier transform, "$\circ$" denotes the Hadamard product, which performs entry-wise multiplication, and "$\bar{}$" denotes complex conjugate.

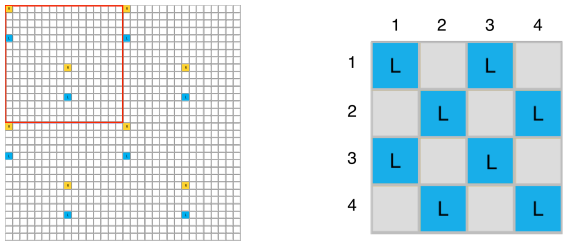
Notice that the shift is one-dimensional, one could calculate a phase correlation matrix for each row of the image using one-dimensional Fourier transform. However, we choose to use the whole image to account for the global shift.

The next step is to calculate the shift by finding the peak in the phase correlation matrix. However, PSFs of the left and right pixels are not identical to each other when the lens is at an out-of-focus position, and thus the relation between the left and right images is not a pure shift. As a result, the phase correlation matrix is noisy and has multiple peaks (see Figure 2(a)), making it difficult to accurately determine the shift. Therefore, we propose to smooth the phase correlation matrix using a one-dimensional Gaussian kernel $g(x)$ with a pre-determined variance, which yields

$$p_f(x) = p_1(x) * g(x), \tag{5}$$

where $p_1(x)$ denotes a row of the correlation matrix, $p_f(x)$ denotes the filtered result. An example result is shown in Figure 2(b).

Finally, Δx can be obtained by finding the maximum value of $p_f(x)$,



(a) (b)

Fig. 3. (a) A portion of the image sensor considered in this work with right pixels marked in yellow and left pixels in blue. The red rectangle represents the basic pixel pattern that repeats over the sensor. (b) Even pixels and odd pixels of an assembled image are not vertically aligned. An assembled left image is shown here as an illustration.

$$\Delta x = \arg\max_{x} p_f(x). \tag{6}$$

3.2. Subpixel Accuracy

Only integral spatial shift can be obtained from (6) since the index x is an integer. However, subpixel accuracy is required so that shift has the precision needed in the regular image domain [12]. We adopt the method proposed by Tian et al. [13] to compute subpixel spatial shift,

$$\Delta x_{sub} = \Delta x - \frac{p_f^- - p_f^+}{2(p_f^+ - 2p_f + p_f^-)}, \tag{7}$$

where $p_f^- = p_f(\Delta x - 1)$ and $p_f^+ = p_f(\Delta x + 1)$.

4. EXPERIMENT

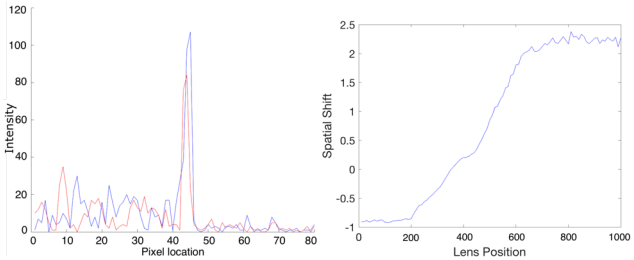
4.1. PDAF Platform

The PDAF platform used in the experiment has a sensor with 3280×2464 pixels, and the range of lens positions is [0, 1024]. Figure 3(a) shows a portion of the sensor and the distribution of left and right pixels. Since the even pixels and the odd pixels of the assembled left image shown in Figure 3(b) are not vertically aligned, we further divide the left image into two sub-images—one consists of odd pixels and the other of even pixels—and work on the two sub-images separately. The average of the shifts of the two sub-images are then taken as the shift of the entire left image. The same procedure is applied to the right image.

4.2. Impulse Response

We verify the simplified model described in Section 2.1 by analyzing the impulse response of the system. However, it is difficult to generate an impulse in practice. Instead, we analyze the step response and mathematically derive the impulse response from the step response.

The step input is generated by capturing an edge with black on the left side and white on the right side. Consider the direction perpendicular to the edge, and let $i_L(x)$ be the observed left image and $h_L(x)$ the blur kernel for left pixels. We have



(a) (b)

Fig. 4. (a) Impulse responses of left pixels (blue) and right pixels (red) in an example case. (b) The corresponding shift profile plotted against the lens position. The actual in-focus lens position is at 350. Each spatial shift is computed by taking the average of 100 tests.

$$i_L(x) = s(x) * h_L(x). \tag{8}$$

Mathematically, the impulse response $j_L(x)$ of the system can be expressed in terms of the step input by

$$j_L(x) = (s(x) - s(x-1)) * h_L(x). \tag{9}$$

Then impulse response is related to the step response by

$$j_L(x) = \mathcal{F}^{-1}\{(1 - e^{-j2\pi mx/N}) I_L(m)\}, \tag{10}$$

where $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse discrete Fourier transform, N is the length of $i_L(x)$, and $I_L(m)$ is the discrete Fourier transform of $i_L(x)$.

When an image is out of focus, there exists a shift between the impulse response of the left image and that of the right image, and the shift can be obtained by calculating the distance between the two peaks (Figure 4(a)).

Furthermore, the shift can be calculated for each lens position and plotted against the lens location to obtain a shift profile, as shown in Figure 4(b). It can be seen that the shift is approximately zero at the in-focus lens position 350 and that the middle section of the curve is approximately linear. Also, on both ends of the curve, the shift stays almost constant. This is reasonable because the nearest focus distance of the PDAF platform used in this experiment is at lens position 750. Images taken within this distance appear equally blur. Likewise, at position 150, the lens is focused at a distant point behind the object. Therefore, the shift profile stays flat (and the image appears equally blur) as the lens is focused at a farther distance.

4.3. Evaluation Metrics

We use two metrics to evaluate the performance of the enhanced phase detection algorithm. The first one is the shape of the shift profile. Since the shift between the left and right

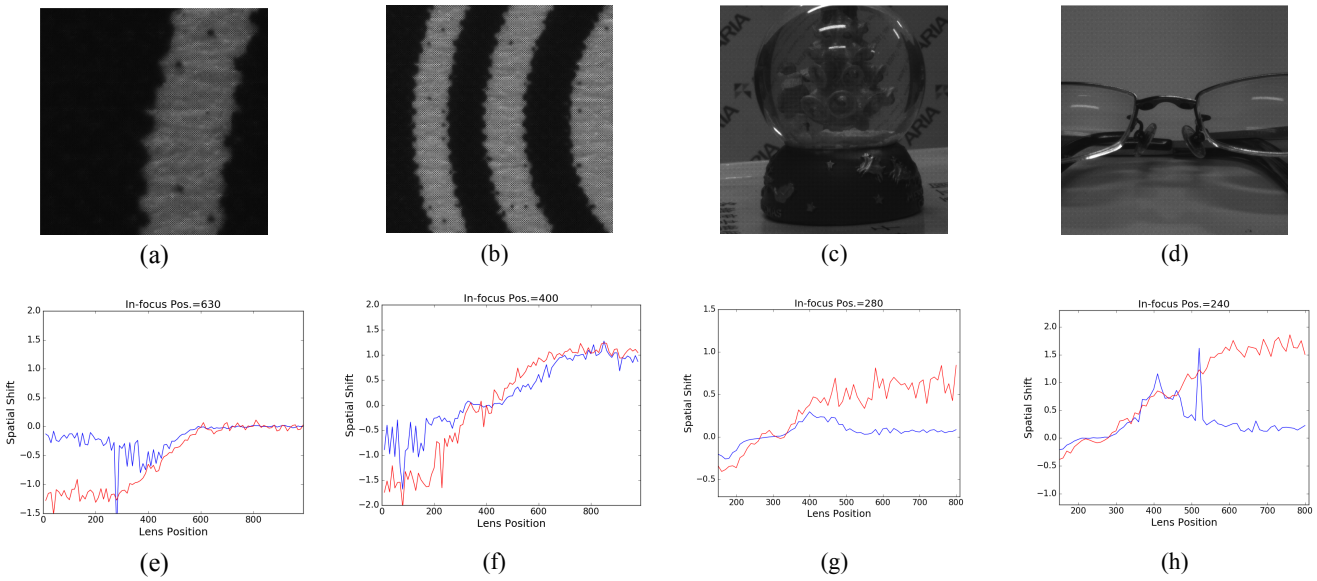


Fig. 5. Comparisons of the proposed method with traditional phase correlation. (a)-(d) The focus windows of Scenes 1, 2, 3 and 4 in order with the in-focus lens positions at 630, 400, 280, and 240 respectively. The focus window size is 208×208 pixels. Images are scaled for display. (e)-(h) The shift profiles obtained by the proposed method (red curves) and traditional phase correlation (blue curves).

Table 1. The number of lens movements needed to reach the near-focus region: proposed method/traditional method.

	Initial lens position					
	200	300	400	500	600	700
Scene 1	5/9	4/6	4/4	3/4	0/0	1/1
Scene 2	1/5	1/3	0/0	1/3	2/4	2/5
Scene 3	1/2	0/0	2/2	3/7	4/12	5/17
Scene 4	0/0	1/1	2/2	3/4	3/7	4/8

images depends only on the distance between the object and the focal plane, the shift profile should be linear in the middle section and flat on both ends, just like the shift profile of the impulse response.

The second metric is the number of lens movements needed to reach the near-focus region. If the distance of the lens to the in-focus position is less than 50, it is considered to be in the near-focus region. Because CDAF is more accurate than PDAF in this region, we want PDAF to bring the lens to this region as fast as possible and let CDAF take over the remaining task afterwards. The direction of the lens movement is determined by the sign of the shift, and the travel distance of the lens in each movement is determined by the following rule:

$$\text{movement} = \begin{cases} 200, & \text{if } |\text{shift}| > 1.5 \\ 150, & \text{if } 1.5 > |\text{shift}| > 1.0 \\ 100, & \text{if } 1.0 > |\text{shift}| > 0.5 \\ 50, & \text{if } 0.5 > |\text{shift}| > 0.25 \\ 20, & \text{otherwise} \end{cases} . \qquad (11)$$

4.4. Performance Evaluation

We compare the proposed algorithm with the traditional phase correlation method (with subpixel accuracy) on four test scenes (see Figure 5) by the two evaluation metrics. Each scene has a different depth of focus. We set the variance of Gaussian kernels to 2 in all tests.

We first plot the shift profiles for the four scenes using the proposed method and traditional phase correlation. The top row of Figure 5 shows the focus windows. The bottom row shows the shift profiles obtained by the two methods (blue for the traditional phase correlation method and red for the proposed method). It can be seen that the shift profiles of the proposed method are more similar to the shift profile of the impulse response shown in Figure 4(b).

Table 1 shows the number of lens movements required by the two methods to reach the near-focus region from different initial lens positions. The proposed method needs at most 5 movements, while the traditional phase correlation method may require 17 movements. Overall, the proposed method needs fewer lens movements when the lens starts off from an out-of-focus position. The results clearly show that the proposed method is more reliable than the traditional phase correlation method and that the reliability leads to efficiency advantage for autofocus.

5. CONCLUSION

In this paper, we have described an enhanced shift estimation method to alleviate the effect of image noise and blur and the effect of phase sensor structure on the performance of PDAF. The ideal shape of the shift profile is obtained by analyzing the impulse response of the PDAF system and exploited in the design and test of the proposed method. We have also shown that that the proposed method for shift estimation enables a PDAF system to bring the lens to the near-focus region faster than the traditional phase correlation method.

6. REFERENCES

[1] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 8, pp. 824–831, 1994.

[2] M. Kristan, J. Pers, M. Perse, and S. Kovacic, "A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform," *Pattern Recognit. Lett.*, vol. 27, no. 13, pp. 1431–1439, 2006.

[3] K. S. Choi, J. S. Lee, and S. J. Ko, "New autofocusing technique using the frequency selective weighted median filter for video cameras," *IEEE Trans. Consum. Electron.*, vol. 45, no. 3, pp. 820–827, Aug. 1999.

[4] J. T. Huang, C. H. Shen, S. M. Phoong, and H. H. Chen, "Robust measure of image focus in the wavelet domain," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, 2005, p. 157–160.

[5] E. P. Krotkov, *Active Computer Vision by Cooperative Focus and Stereo*. New York: Springer-Verlag, 1989.

[6] M. Subbarao and J.-K. Tyan, "Selecting the optimal focus measure for autofocusing and depth-from-focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 864–870, Aug. 1998.

[7] M. Hamada, "Imaging device including phase detection pixels arranged to perform capturing and to detect phase difference," US20130088621, 2013.

[8] M. G. Gluskin, R. M. Velarde, and J. Lee, "Phase detection autofocus noise reduction," US9420164, 2016.

[9] M. Levoy, "Autofocus: phase detection.," 2010. [Online]. Available:http://courses.cs.washington.edu/courses/cse131/13sp/applets/autofocusPD.html.

[10] C. D. Kuglin, "The Phase Correlation Image Alignment Method," in *Proc. Int. Conf. Cybernetics and Society*, 1975, pp. 163–165.

[11] M. Cannon, "Blind deconvolution of spatially invariant image blurs with phase," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 1, pp. 58–63, Feb. 1976.

[12] H. Foroosh, J.B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Process.*, vol. 11, no. 3, pp. 188–200, Mar. 2002.

[13] Q. Tian and M. N. Huhns, "Algorithms for subpixel registration," *Comput. Vision, Graph. Image Process.*, vol. 35, no. 2, pp. 220–233, Aug. 1986.