

## Article

# Multi-Target Adaptive Focusing Algorithm Based on Image Phase Difference

Ninghui Pan and Feihong Yu \*

College of Optical Science and Engineering, Zhejiang University, Hangzhou 310007, China; 22230084@zju.edu.cn

\* Correspondence: feihong@zju.edu.cn

**Abstract:** In the field of image detection, when focusing on multiple targets of interest with different depths, in addition to optimizing the focusing speed of a single target, a comprehensive consideration of the interrelationships among multiple targets can further enhance the overall focusing efficiency. In this context, this paper presents a multi-target adaptive focusing algorithm based on image phase difference. The algorithm improves the accuracy of phase difference calculation through multi-scale image matching and adaptively adjusts the defocus estimation model based on the phase difference. Additionally, the algorithm orders multiple target regions by defocus estimation to reduce the overall focus time and combines the defocus estimate with image clarity to achieve precise focus on each region. The experimental results show that the method not only effectively improves the accuracy of defocus estimation and does not need to rely on lens calibration, but also significantly improves the speed and efficiency of multi-target focusing compared with traditional focusing algorithms.

**Keywords:** autofocus; defocus estimation; image clarity; multi-target



**Citation:** Pan, N.; Yu, F. Multi-Target Adaptive Focusing Algorithm Based on Image Phase Difference. *Electronics* **2024**, *13*, 4917. <https://doi.org/10.3390/electronics13244917>

Academic Editor: Beiwen Li

Received: 23 October 2024

Revised: 6 December 2024

Accepted: 11 December 2024

Published: 13 December 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Autofocus technology is one of the key technologies in imaging systems [1]. With the development of technology and the richness of application scenarios, people have higher demand for the speed and reliability of autofocus technology. Currently, the mainstream focusing algorithms [2,3] are mainly divided into two kinds: contrast detection autofocus (CDAF) and phase detection autofocus (PDAF).

CDAF operates by measuring image clarity [4] at various positions, determining the focus direction and position based on these values. The point of maximum clarity value represents the theoretical focus position [5]. In recent years, researchers have actively been exploring more sensitive image clarity evaluation functions and more efficient focus search algorithms to improve the performance of contrast detection focusing [6–9]. Commonly used search algorithms include the hill-climbing search algorithm, the golden-section search algorithm, and the function fitting search algorithm. While CDAF offers high accuracy, it struggles in low-contrast environments where it is difficult to estimate lens movement direction, resulting in longer search times and slower focusing.

PDAF [10–12] achieves autofocus by detecting phase offsets in the images captured by photoreceptors. When a phase difference is observed between the output signals of two photoreceptors, the necessary focus position adjustment is calculated based on this difference. Compared to CDAF, PDAF provides faster focusing, but at the cost of lower accuracy, particularly in low-light conditions. To improve phase detection accuracy, Jang et al. [13] proposed a multiresolution feature extraction algorithm that can more rapidly and robustly estimate phase differences in defocus images under low illumination. Additionally, they incorporated a hybrid focusing algorithm [14] to combine the strengths of both CDAF and PDAF, further improving focusing speed. Chan et al. [15] introduced the skewed normal distribution model for phase shift estimation and applied Bayesian theorem-based techniques for more precise defocus distance estimation, enhancing the accuracy of focus adjustments.

The term multi-target in this paper refers to the presence of areas of different depths in the field of view, which can be different feature areas of the same object or different types of objects. Multi-target-focused image acquisition involves focusing on and capturing images of multiple regions of interest within the field of view separately. This technique is commonly applied in industrial and medical image detection [16], multi-target image fusion, and other related scenarios [17–19]. The autofocus algorithms mentioned above focus on the time required for single-target focusing, and thus are not optimal for the continuous multi-target focusing photography. These algorithms also face limitations in adapting to changes in lens parameters when employing calibration-based lens models [20] for autofocus.

In this paper, defocus estimation is conducted for different target regions to optimize the focus search path. This approach reduces the distance the motor must travel repeatedly, thereby accelerating the overall focusing speed of the multi-target focusing algorithm. Additionally, by refining the phase difference calculation method, the autofocus algorithm provides more accurate defocus estimates, which further enhances the speed of focusing on individual targets. The adaptive adjustment of the defocus estimation model during the focusing process offers the advantage of not being constrained by lens calibration, in contrast to traditional methods.

## 2. Principle

Contrast detection autofocus is performed based on the image clarity evaluation function [21–23], and autofocus is realized by finding the position with the largest clarity evaluation value [22]. Due to the real-time requirement of autofocus for image clarity, a spatial domain evaluation function is usually chosen to evaluate the image clarity of the current region by calculating the gray gradient value of the focusing region. Among them, the commonly used Tenengrad function calculates the gray gradient of the image based on the Sobel operator [24] and calculates the clarity value by adding the square of the gradient amplitude:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (1)$$

$$G_x(x, y) = f(x, y) * S_x, G_y(x, y) = f(x, y) * S_y \quad (2)$$

$$G_{Sobel}(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (3)$$

$$C = \frac{1}{M \times N} \sum_x \sum_y G_{Sobel}(x, y) \quad (4)$$

where  $G_x(x, y)$  and  $G_y(x, y)$  are the horizontal and vertical gradient values of the grayscale image  $f(x, y)$  convolved with the Sobel algorithm kernels  $S_x$  and  $S_y$ , respectively,  $*$  is the convolution symbol, and  $G_{Sobel}(x, y)$  is the gradient image with the image size of  $M \times N$ .  $C$  is the computed clarity factor, and the larger  $C$  is, the clearer the image.

Phase detection autofocus relies on shield pixels on the sensor [25]. Figure 1a shows the imaging principle of a sensor with shield pixels. A mask in the form of a metal shield is placed on the left or right pixel. When placed above the left (or right) pixel, it blocks the light from the right (or left) side so that the light cannot reach the photodiode below the mask. When in focus, the intensity of light arriving at the left pixel and arriving at the right pixel is the same; when out of focus, the intensity of light deviates. The photodiode converts the received light intensity into a voltage signal or a digital signal, which generates the corresponding gray value. The gray pixel values obtained on the same row of masked pixel points when the sensor is imaged are shown in Figure 1b, where the horizontal coordinate is the horizontal position of the pixel point and the vertical coordinate is the gray value. When out of focus, there is a certain pixel displacement between signals, i.e., a

phase difference  $PD$  between the response signals of the rows of left pixels and those of the rows of right pixels.

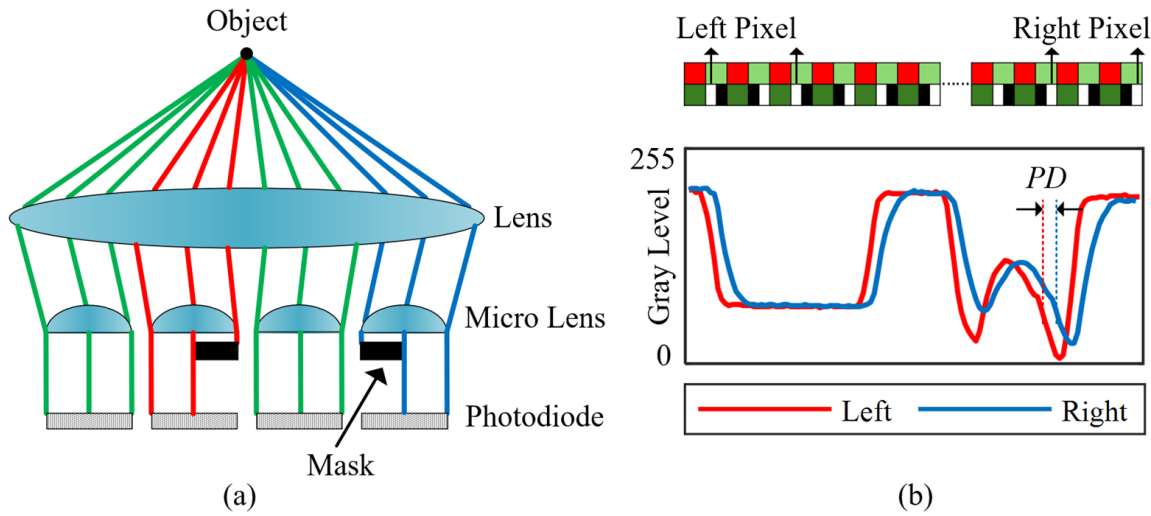


**Figure 1.** Sensor-based defocus phase difference generation principle: (a) sensor with shield pixels; (b) phase difference between left and right signals.

Thus, the phase shift between the signal on the left pixel and the signal on the right pixel can be modeled by a one-dimensional translation:

$$R(x) = L(x + PD) \quad (5)$$

where  $R$  is the signal on the right pixel of a row and  $L$  is the signal on the left pixel of a row. According to the correlation principle, the current signal offset, i.e., the phase difference  $PD$ , is calculated:

$$Corr(R, L_{\Delta x}) = \frac{\sum (R - \bar{R})(L_{\Delta x} - \bar{L}_{\Delta x})}{\sqrt{\sum (R - \bar{R})^2 \sum (L_{\Delta x} - \bar{L}_{\Delta x})^2}} \quad (6)$$

$$PD = \arg \max_{\Delta x} Corr(R, L_{\Delta x}) \quad (7)$$

where  $L_{\Delta x}$  is the signal offset  $\Delta x$  from the original  $L$ ,  $\bar{R}$  is the average value of  $R$ ,  $\bar{L}_{\Delta x}$  is the average value of  $L_{\Delta x}$ , and  $Corr$  is the correlation calculation function.

Focusing, in principle, involves controlling the focus motor to change the distance from the lens to the sensor so that the imaging position is on the focal plane. As shown in Figure 2, the clarity value and phase difference of the image in the region are calculated when the lens is in different positions. For CDAF, the position when the image clarity is maximal is the focus position; for PDAF, the phase difference of the image in focus is close to 0, and the larger the absolute value, the further the lens is from the focus position of this target [26]. Therefore, when defocusing, it is possible to estimate the current step size  $S$  needed to move the lens to the focus position based on the current phase difference  $PD$  and the proportionality relationship  $\theta$  between the phase difference and the lens position [27]:

$$S = -PD/\theta \quad (8)$$

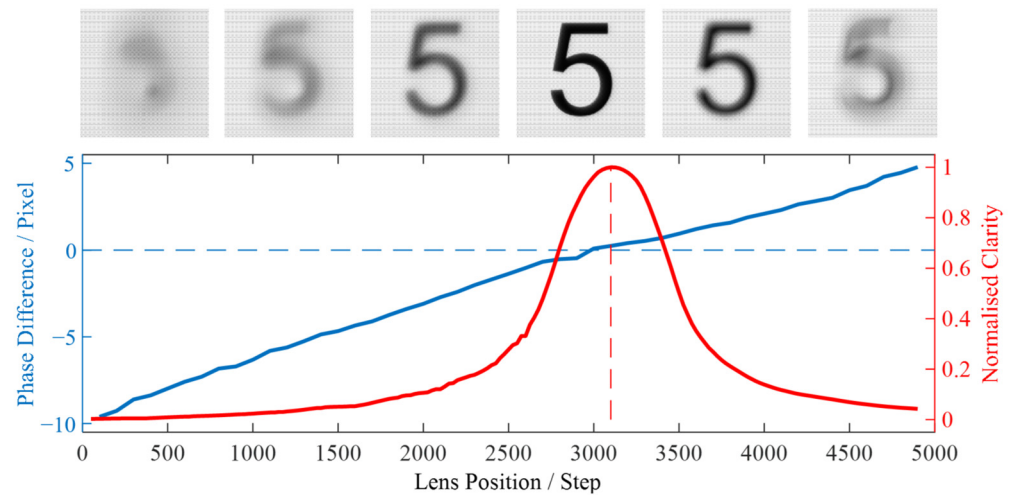


**Figure 2.** Image phase difference and image clarity versus lens position.

### 3. Materials and Methods

#### 3.1. Method of Calculating Phase Difference

The accuracy of the defocus estimation is directly related to the phase difference computation, and this paper proposes a phase difference computation method based on multi-scale two-dimensional image matching (MS2D Match) to improve the accuracy of the phase difference computation. Since a single line of pixel signals may lack the feature information of the object in focus, the present study splits the matching of the left image  $I_l$  and the right image  $I_r$  from the original Raw format image  $I$ :

$$I = I_r + I_l + I' \quad (9)$$

where  $I'$  is the image made up of non-shield pixels. When increasing the field of view angle, the total light intensity that can be received by the left and right pixels will be different, as shown in Figure 3a. By normalizing the left and right images to obtain  $I_R$  and  $I_L$  to reduce the effect of the difference in illumination on the matching result, this reduces the matching error in the fringe field of view. Image pyramids are constructed for the normalized left and right images, respectively [28]: a single image is continuously downsampled to obtain images of different resolutions, as shown in Figure 3b, which shows a three-layer pyramid ( $I_{L1}$ ,  $I_{L2}$ ,  $I_{L3}$ ) of the left image and a three-layer pyramid ( $I_{R1}$ ,  $I_{R2}$ ,  $I_{R3}$ ) of the right image. The left and right images of different scales are matched, respectively:

$$NCC_i(\Delta x) = \text{Corr}(I_{Ri}, I_{Li}(\Delta x)) \quad (10)$$

where  $I_{Li}(\Delta x)$  denotes an image offset by  $\Delta x$  pixels horizontally compared to the original image  $I_{Li}$ , and  $NCC_i(\Delta x)$  is the normalized correlation between the left and right images at the current scale at the  $\Delta x$  offset, as shown in Figure 3c. After the correlation on each scale is calculated, the results of the lower scales are scaled up to the original resolution for fusion using interpolation:

$$\overline{NCC} = \frac{1}{N} \sum_{i=1}^N NCC_i \quad (11)$$

where  $N$  is the number of layers of the image pyramid. The phase difference  $PD$  is the subpixel level  $\Delta x$  at which the correlation is highest, obtained by secondary interpolation:

$$PD = \Delta x_{\max} + \frac{\overline{NCC}(\Delta x_{\max} - 1) - \overline{NCC}(\Delta x_{\max} + 1)}{2 \times (\overline{NCC}(\Delta x_{\max} - 1) - 2 \times \overline{NCC}(\Delta x_{\max}) + \overline{NCC}(\Delta x_{\max} + 1))} \quad (12)$$

where  $\Delta x_{\max} = \arg \max(\overline{NCC})$ .

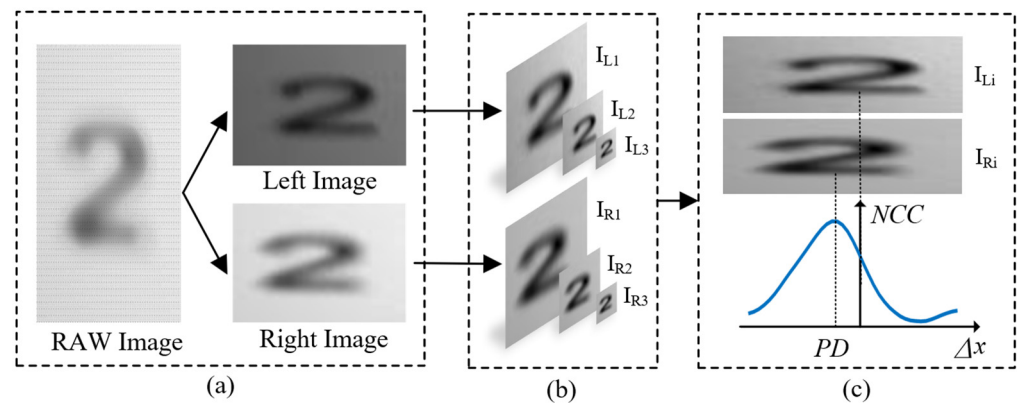


**Figure 3.** Phase difference calculation process based on MS2D Match: (a) separate left and right images; (b) constructing multi-scale images; (c) calculating the defocused image phase difference based on the correlation coefficient.

### 3.2. Adaptive Defocus Estimation Model

The defocus estimation model [29] concerns to the relationship between the phase difference and the lens defocus distance. As shown in Figure 4, when defocused, the two light rays originally converging at the same point receive the same response on the left and right pixels of the sensor, respectively. The pixel distance between them on the sensor is the phase difference  $PD$ , and the corresponding defocus distance is  $d$ . The ideal imaging situation at different object distances is shown in the figure. Figure 4a shows a closer object, resulting in a negative phase difference, indicating that the lens needs to move closer to the direction of the sensor; Figure 4b shows a farther object, resulting in a positive phase difference, indicating that the lens needs to move away from the direction of the sensor. The  $PD$  generated by the objects between Object\_1 and Object\_2 satisfies  $PD_1 < PD < PD_2$ , so that the relative distances of the objects can be estimated based on the phase difference and the desired lens-focusing situation can be obtained.

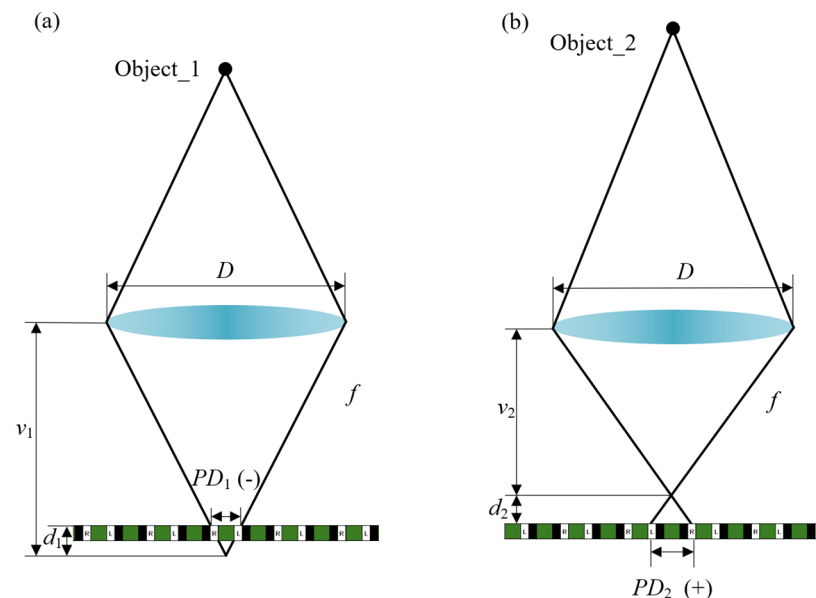


**Figure 4.** Lens-focusing model. (a) Imaging situation that produces negative phase differences; (b) imaging situation that produces positive phase differences.

According to the imaging principle, the relationship between the phase difference and the defocus distance can be estimated:

$$\frac{PD}{d} = \frac{D}{v} \xrightarrow{F\# = \frac{f}{D}} \frac{PD}{d} = \frac{f}{vF\#} \quad (13)$$

where  $v$  is the image distance,  $D$  is effective lens aperture,  $f$  is the focal length of the lens, and  $F\#$  is the aperture number of the lens. An estimate of the lens defocus distance for the lens can be constructed as follows:

$$d = PD \times \theta \quad (14)$$

where  $\theta$  is the conversion coefficient between the phase difference  $PD$  and the lens defocus distance  $d$ , which is related to the lens aperture, focal length, focus distance, etc. According to the above principle, the lens module needs to be calibrated in advance to convert the image phase difference into the defocus distance through the defocus estimation model, and the conversion coefficient  $\theta$  under different parameters needs to be calibrated.

This paper uses an adaptive defocus estimation model to adjust the conversion coefficient calculation under different lens parameters. When the lens is at a certain defocus position  $Z_i$ , the phase difference  $PD_i$  of the image in the current motor position and its corresponding left and right image correlation coefficients  $\omega_i$  can be calculated according to the principle in Section 3.1, i.e.,  $\omega_i = \max(\overline{NCC})$ . Using the initial two sets of data— $(Z_0, PD_0)$  and  $(Z_1, PD_1)$ —the initial conversion coefficient is estimated:

$$\theta_1 = \frac{PD_1 - PD_0}{Z_1 - Z_0} \quad (15)$$

During the autofocus process, the conversion coefficient is adjusted by calculating  $(Z_i, PD_i)$  at each position and using the weighted least squares method [30] linear fitting:

$$S = \sum_{j=1}^i \omega_j (PD_j - (\theta_i Z_j + c)) \quad (16)$$

$$\frac{\partial S}{\partial \theta_i} = 0, \frac{\partial S}{\partial c} = 0 \rightarrow \theta_i = \frac{\sum_{j=1}^i \omega_j Z_j (PD_j - \overline{PD})}{\sum_{j=1}^i \omega_j Z_j (Z_j - \overline{Z})} \quad (17)$$

where  $c$  is the intercept of the fitting line, and  $Z_j$ ,  $PD_j$ , and  $\omega_j$  represent the data calculated in the previous focusing process; they can be used to estimate the defocus model.

$\overline{PD} = \frac{\sum_{j=1}^i \omega_j PD_j}{\sum_{j=1}^i \omega_j}$ ;  $\overline{Z} = \frac{\sum_{j=1}^i \omega_j Z_j}{\sum_{j=1}^i \omega_j}$ . In addition, when the lens changes the aperture or focal length, it is necessary to re-estimate the conversion coefficient based on the phase difference  $PD_i$  before the change and the phase difference  $PD'_i$  after the change:

$$\theta'_i = \theta_i \times \frac{PD_i}{PD'_i} \quad (18)$$

### 3.3. Multi-Target Focusing Algorithm

When focusing on a single target, the autofocus algorithm based on contrast detection can obtain a more accurate focusing result, but the number of searches is much greater than that of the focusing algorithm based on phase detection. In addition, when focusing on multiple selected targets, since the focus positions of each target are different, focusing in the order that the targets are selected will cause the motor to move back and forth frequently, and the overall focusing time will be longer, as shown in Order1 in Figure 5. Focusing in a certain order can shorten the focusing process, as shown in Order2 in Figure 5.

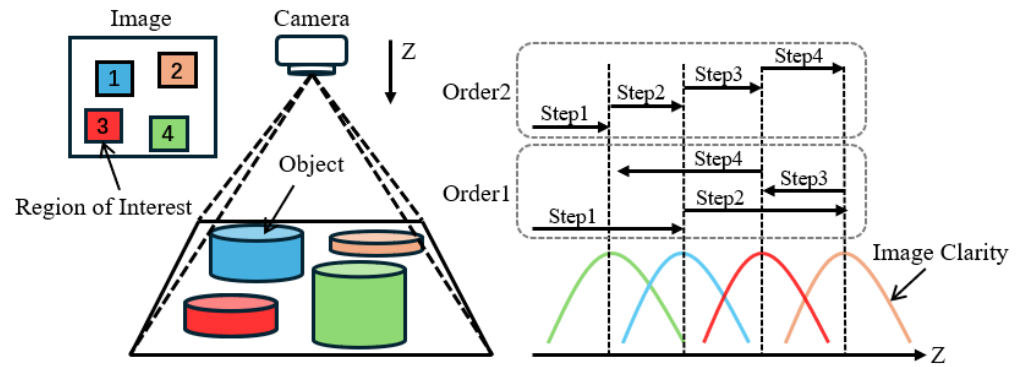


**Figure 5.** Schematic of multi-target focusing and its focus paths.

This paper proposes a multi-target focusing algorithm based on the above phase difference calculation method and adaptive defocus estimation model, sorts multiple targets based on defocus estimation, and implements a multi-target focusing algorithm by mixing a phase detection autofocus method and a contrast detection autofocus method. The steps are shown in Figure 6:

- Focus initialization. For the selected  $n$  regions  $ROI[k]$  ( $1 \leq k \leq n$ ), calculate the phase differences  $PD_0[k]$  and  $PD_1[k]$ , respectively, at two focus positions,  $Z_0$  and  $Z_1$ , that are distanced from  $S_{min}$ .  $S_{min}$  is the minimum search step of the hill-climbing search algorithm.
- Multi-target sorting. The conversion coefficient  $\theta_1$  of the initial defocus model is calculated according to formula (15), and the defocus distance  $d_1[k]$  of each focus region is estimated using the phase difference  $PD_1[k]$ . All focus regions are sorted from near to far according to the defocus distance to obtain  $ROI_{sort}[k]$ . Subsequent autofocus in the sorted order can reduce the repeated movement of the motor.
- Autofocus on the region of interest. To autofocus on the current  $ROI_{sort}[k]$ , the focus algorithm is divided into two parts: PDAF steps and CDAF steps. First, according to the defocus distance  $d_1$ , the defocus step length  $S_d = |d_1|$  required for the lens to move to the position where the theoretical phase difference is 0 can be calculated. In order to reduce the number of repeated movements of the motor, the search interval is reserved during PDAF. The step length  $S_{PD}$  of autofocus is as follows:

$$\begin{cases} S_{PD} = |S_d - 4S_{min}| & D = 1, d > 0 \\ S_{PD} = |S_d + 4S_{min}| & D = -1, d < 0 \end{cases} \quad (19)$$

where  $D$  is the direction of lens movement. Calculate the image phase difference data  $(Z_i, PD_i, \omega_i)$  after each movement and use the data to update the conversion coefficient  $\theta_i$  of the defocus estimation to obtain a new defocus estimation distance  $d_i$ . Repeat this step until  $|d_i| < 4S_{min}$  and perform CDAF by completing a hill-climbing search focus based on the image clarity evaluation function. The judgment condition for successful focusing using the hill-climbing search algorithm is as follows: when the motor moves with a step length  $S_{min}$ , the image clarity values  $C$  corresponding to five consecutive positions show two increases and two decreases. At the same time, in order to prevent the peak from being missed due to calculation errors or image instability, add the judgment condition: when the image clarity decreases three times in a row, turn the focus direction.

- Adaptive adjustment of the defocus estimation model. The defocus estimation model is divided into two steps in the focusing process: model self-correction and model prediction. Model self-correction is when it is focusing on the current region in step (c). If there are certain data  $((Z_j, PD_j, \omega_j), 0 < j < i)$  within the range of  $Z_i - S_{min} < Z_i < Z_i + S_{min}$ , these data will be eliminated, and at most the five latest sets of  $(Z_j, PD_j, \omega_j)$  data will be retained during model self-correction to update the current defocus estimation conversion coefficient  $\theta_i$ . Model prediction means that in

the focusing process, in addition to calculating the phase difference and clarity of the currently focused  $ROI_{sort}[k]$ , the image phase difference data of the  $ROI_{sort}[k+1]$  is also calculated at each focusing position. These data are used to correct the defocus estimation model in the next position, so that when starting to focus on the next region, a more accurate conversion coefficient  $\theta_1$  corresponding to the  $ROI_{sort}[k+1]$  can be obtained.

- (e) Multiple target focus. Repeat steps (c) and (d) in the order of  $ROI_{sort}[k]$  until all regions are in focus.

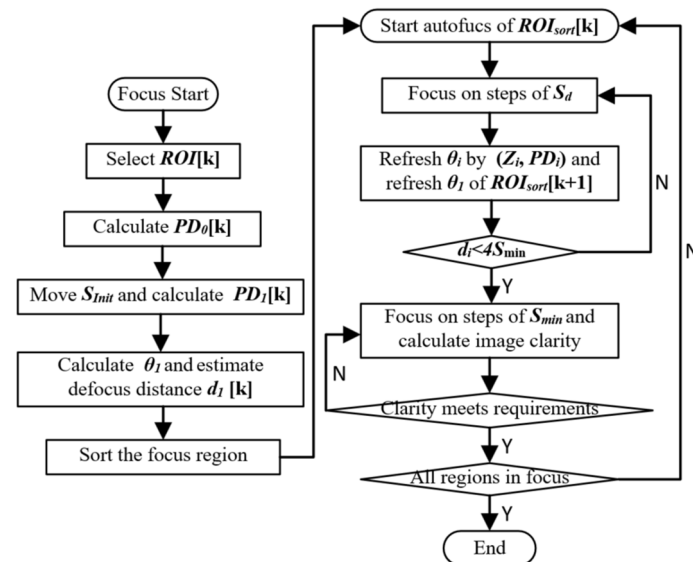


Figure 6. Multi-target focusing algorithm.

## 4. Experiments and Analyses

### 4.1. Experiment Environment

The sensor model of the camera used for the experiment was a Sony IMX410 with the parameters shown in Table 1.

Table 1. Image sensor parameters.

| Parameter Name    | Parameter Value                                        |
|-------------------|--------------------------------------------------------|
| Image sensor size | 43.3 mm (2.7 inch)                                     |
| Pixel size        | 5.94 $\mu\text{m}$ (H) $\times$ 5.94 $\mu\text{m}$ (V) |
| Resolution        | 6064 (H) $\times$ 4040 (V)                             |

H represents the horizontal direction and V represents the vertical direction. In full-resolution mode, there is 1 line of phase focus detection pixels for every 12 lines. The camera's maximum frame rate at full resolution is 12 fps. The lenses used for the experiment were manual zoom and motorized focus lenses. The camera sends commands to the lens through electronic contacts to control the movement of the lens to achieve focusing, and the total moving range of the focusing group, which is the travel of the focusing motor, is 0–7000 steps, and the current position of the motor is represented by the variable Z (steps). The focal length range of the lens is 24 mm–70 mm and the maximum aperture is  $F\# = 4.0$ . The experiment was supplemented with ambient lighting or a ring-light source.

### 4.2. Experiment of Phase Difference Calculation

This section verifies the effectiveness of the phase difference calculation method based on image matching.

Figure 7 shows the phase relationship obtained by image matching using different methods when the defocus distance is large. Notably, the 1D Match shown in Figure 7a uses a single-row image for matching, the 2D Match shown in Figure 7b uses a two-dimensional image of the focus region for matching, and the MS2D Match shown in Figure 7c performs image matching based on multi-scale transformation. The curves represent the correlation

between the two images when the left and right images are offset by X-Offset pixels and when the focus motor is at  $Z = 100$ ,  $Z = 200$ ,  $Z = 300$ ,  $Z = 400$ , and  $Z = 500$ . The theoretical focus position is at  $Z = 2400$ . The closer  $Z$  is to 2400, the smaller the defocus distance, and the smaller the absolute value of the theoretical phase difference. In other words, as  $Z$  approaches the focused position at 2400, the X-Offset corresponding to the maximum correlation coefficient, that is, the  $PD$ , gradually tends to 0. The specific  $PD$  values and the relationship between  $PD$  and the current position of  $Z$  are shown in Figure 7d. The experiments show that as  $Z$  increases, the overall trend in the phase difference calculation results obtained using the 1D Match and 2D Match is an upward fluctuation in the value of  $PD$ . The use of MS2D Match results in a smoother correlation curve and a reduced fluctuation in the calculated phase difference (red line in Figure 7d), which aligns with the theoretical upward trend. This is attributed to the consideration of the disparate scales of the images and their two-dimensional nature.

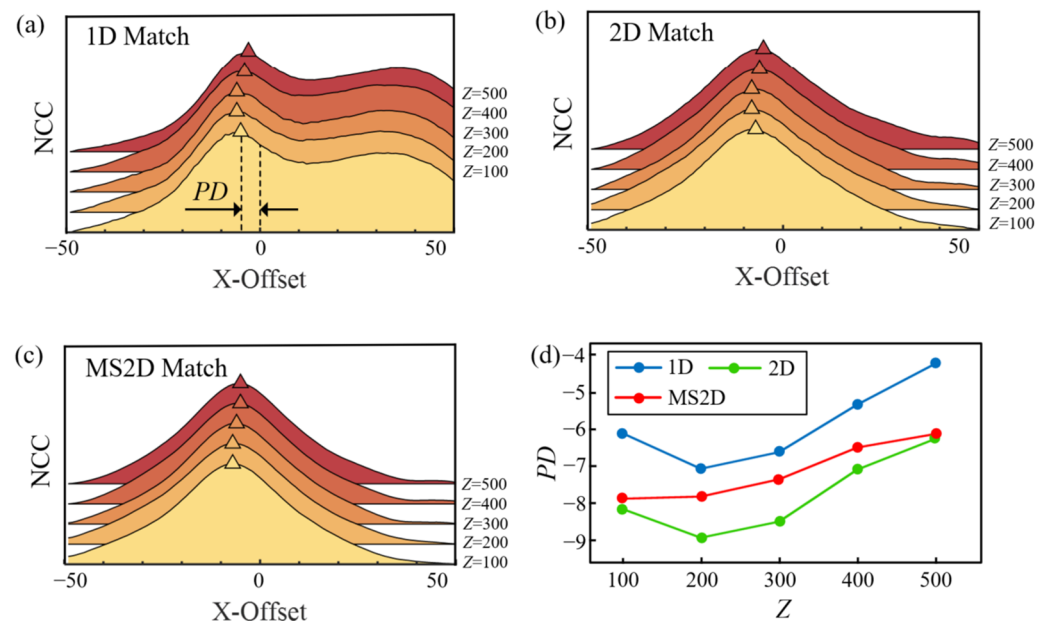


**Figure 7.** Correlation curves and phase differences obtained by using different image matching methods: (a) correlation curve of 1D Match method; (b) correlation curve of 2D Match method; (c) correlation curve of MS2D Match method; (d)  $PD$  obtained from the correlation curves shown in (a–c).

The phase difference is calculated for three different scenes, as shown in Figure 8a–c, corresponding to a resolution board with patterns of different frequencies and contrasts; a circuit board with complex geometries, irregular textures, and multi-depth features, representing real industrial inspection scenarios; and a keyboard with a low-contrast area, respectively. The three curves in Figure 8d–f represent the relationship between the phase difference and defocus distance calculated by the matching methods of 1D Match, 2D Match, and MS2D Match for the target region. The ordinate is the phase difference, and the abscissa is the current position of the lens motor.

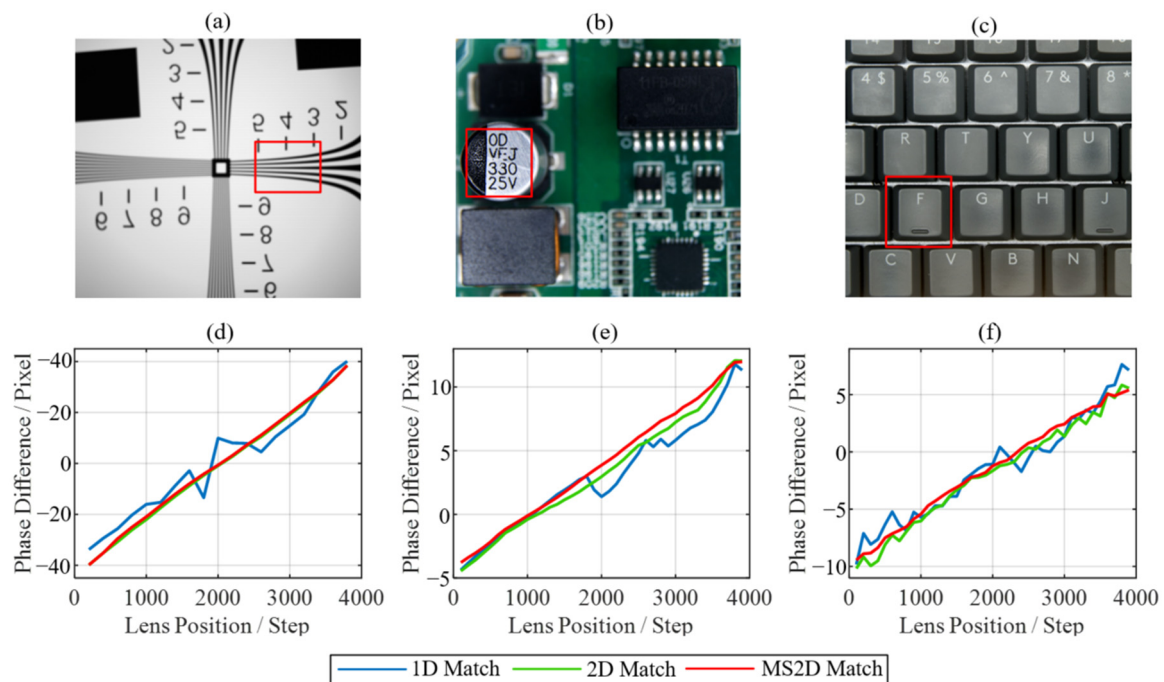


**Figure 8.** Calculated phase difference vs. lens position in different scenes: (a) focus on a certain area of a resolution board; (b) focus on a certain area of a circuit board; (c) focus on a certain area of a keyboard; (d) phase difference results for (a); (e) phase difference results for (b); (f) phase difference results for (c).

The experiments show that compared with 1D Match, the curves of the other two methods have smaller fluctuations, and the defocus distance and phase difference relationship curve calculated using MS2D Match has the smallest fluctuation.

The experimental results are analyzed quantitatively below. According to Equation (14), in a phase detection autofocus system, the defocus steps of the focus motor are calculated by the product of the phase difference and a preset conversion factor. To determine the conversion factor, a straight line  $\hat{f}(Z)$  is fitted to the curve of the defocus steps versus lens position (the original curve  $f(Z)$ ). The coefficient of determination  $R^2$  is taken as an evaluation index of the correlation between the original curve and the fitted curve:

$$R^2 = 1 - \frac{RSS}{TSS} \quad (20)$$

where the residual sum of squares  $RSS = \sum_Z \left( f(Z) - \hat{f}(Z) \right)^2$  and the total sum of squares  $TSS = \sum_Z \left( f(Z) - \bar{f}(Z) \right)^2$ .  $\bar{f}(Z)$  is the mean of  $f(Z)$ . The closer the  $R^2$  value is to 1, the smaller the deviation is. Table 2 shows the  $R^2$  of the curve with its fitted straight line.

**Table 2.** Goodness of fit of the curve to its fitted straight line.

|            | 8(d)   | 8(e)   | 8(f)   |
|------------|--------|--------|--------|
| 1D Match   | 0.7971 | 0.9681 | 0.9628 |
| 2D Match   | 0.9994 | 0.9927 | 0.9893 |
| MS2D Match | 0.9994 | 0.9989 | 0.9956 |

The smaller the deviation, the more accurate the results that can be obtained when using this method to calculate the phase difference for conversion factor calibration and defocus estimation. It can be seen from the results shown in Table 2 that using two-

dimensional images for matching (2D Match and MS2D Match) can significantly improve the accuracy of the phase difference calculation. On this basis, when using MS2D Match, the image phase difference can be calculated more accurately, especially for scenes with low contrast. As shown in Figure 8c, the accurate improvement effect of the phase difference calculation is obvious.

#### 4.3. Focusing Algorithm Experiment

##### 4.3.1. Adaptive Defocus Estimation

In the traditional lens calibration method, when the lens is at the center of the motor range, an object is placed at an object distance that can be clearly focused, and the relationship between the phase difference and the motor position is calibrated. According to this method, for a lens with a travel range of 0~7000 steps that is focused at the center of the travel range (3500), a calibrated conversion coefficient of  $\theta_{CAL} = 0.010255$  can be obtained as a reference. The results of the defocus estimation according to the calibration parameters and the results of the defocus estimation according to the adaptive method are compared. As shown in Figure 9, (A) is the overall defocused image, and five regions (B)~(F) are selected for defocus estimation.

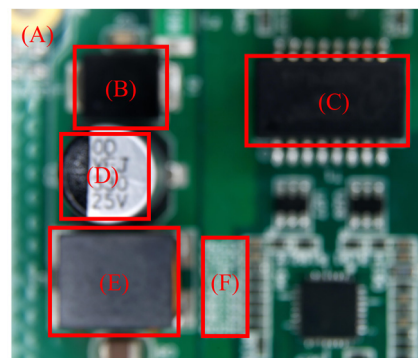


**Figure 9.** Multi-target focus region selection.

Table 3 shows the comparison of the defocus estimation results when the motor focus clear range is within the range of 3250–3750, that is, close to the calibration center position of 3500. The experimental results show that the focus region shown in the figure has different fields of view and object distances, and the calibration results do not consider the impact of different fields of view and actual focus position on the defocus model. Therefore, in most cases, the adaptive defocus estimation performs better.

**Table 3.** Estimation of defocus steps near the calibration center.

| Target Region | Actual Defocus Steps | Estimated by Calibrated Defocus | Estimated by Adaptive Model |
|---------------|----------------------|---------------------------------|-----------------------------|
| (B)           | 1230                 | <u>1222</u> <sup>1</sup>        | <u>1222</u>                 |
| (C)           | 1562                 | 1450                            | <u>1538</u>                 |
| (D)           | 1890                 | 1955                            | <u>1866</u>                 |
| (E)           | 1382                 | <u>1412</u>                     | 1458                        |
| (F)           | 1112                 | 1204                            | <u>1070</u>                 |

<sup>1</sup> The underlined data indicate that the estimated value obtained by this method is closer to the actual value.

Table 4 shows the comparison of the defocus estimation results when the motor focus clear range is within the range of 2000–2500, that is, not near the calibration center position. The experimental results show that when the clear focus position is far away from the actual calibration center position, the error of the defocus estimation based on the calibration result will increase, and the result of the defocus estimation using the adaptive method is more accurate.

**Table 4.** Estimation of defocus steps away from calibration center.

| Target Region | Actual Defocus Steps | Estimated by Calibrated Defocus | Estimated by Adaptive Model |
|---------------|----------------------|---------------------------------|-----------------------------|
| (B)           | 2016                 | 2131                            | <u>2015</u>                 |
| (C)           | 1536                 | 1749                            | <u>1567</u>                 |
| (D)           | 1124                 | 1329                            | <u>1149</u>                 |
| (E)           | 1754                 | 1917                            | <u>1839</u>                 |
| (F)           | 2184                 | 2429                            | <u>2153</u>                 |

#### 4.3.2. Autofocus Speed

This paper uses a high-resolution camera with a frame rate of 12 frames per second to test the autofocus search algorithm. Since the autofocus algorithm needs to determine the next movement based on the evaluation of the current image, that is, each time a motor movement command is sent, it is necessary to ensure that the image has been refreshed, which requires an interval of about 83 ms at least. Therefore, under the condition of a certain frame rate, the fewer the number of movements, the shorter the time required for the focus search.

Figure 10 compares the focusing processes of multi-target focusing, considering the multi-target adaptive focusing algorithm in this paper and traditional methods, including the golden-section search algorithm, the Fibonacci search algorithm, and the hill-climbing search algorithm combining coarse-fine search steps. The horizontal coordinate in the figure represents the search process, the vertical coordinate represents the current lens position, and each point represents a focus movement. The parts represented by letters B~F in Figure 10 correspond to the focusing process of the areas (B)~(F) in Figure 9. The traditional methods focus according to the selection order of the focusing area (B)~(F), while the multi-target adaptive focusing algorithm focuses according to its algorithmic estimation in the order of (F)-(B)-(E)-(C)-(D). Based on the degree of fluctuation in the graph, when using the golden-section algorithm and the Fibonacci algorithm for focusing, the motor repeatedly moves over a longer distance, and therefore the total focusing stroke is longer. When using the hill-climbing algorithm, the density of the points is greater, which means that the number of focus searches is higher.

The focusing performance of the four methods is quantitatively analyzed in Table 5. Overall, the multi-target adaptive focusing algorithm proposed in this paper has advantages in terms of the number of searches, search trips, and focusing time compared with the other three algorithms.

**Table 5.** Focusing algorithm comparison.

| Algorithm             | Number of Searches |          |          |           |          | Total Number of Searches | Search Trips (Steps) | Focusing Time (s) |
|-----------------------|--------------------|----------|----------|-----------|----------|--------------------------|----------------------|-------------------|
|                       | (B)                | (C)      | (D)      | (E)       | (F)      |                          |                      |                   |
| Golden-section        | 11                 | 11       | 12       | <u>11</u> | 11       | 56                       | 76,194               | 18.9              |
| Fibonacci             | 13                 | 14       | 13       | <u>11</u> | 17       | 68                       | 77,942               | 20.4              |
| Hill-climbing         | 27                 | 17       | 15       | 20        | 19       | 98                       | 7684                 | 11.2              |
| Multi-target adaptive | <u>8</u>           | <u>7</u> | <u>6</u> | <u>11</u> | <u>7</u> | <u>39</u>                | <u>3457</u>          | <u>4.9</u>        |

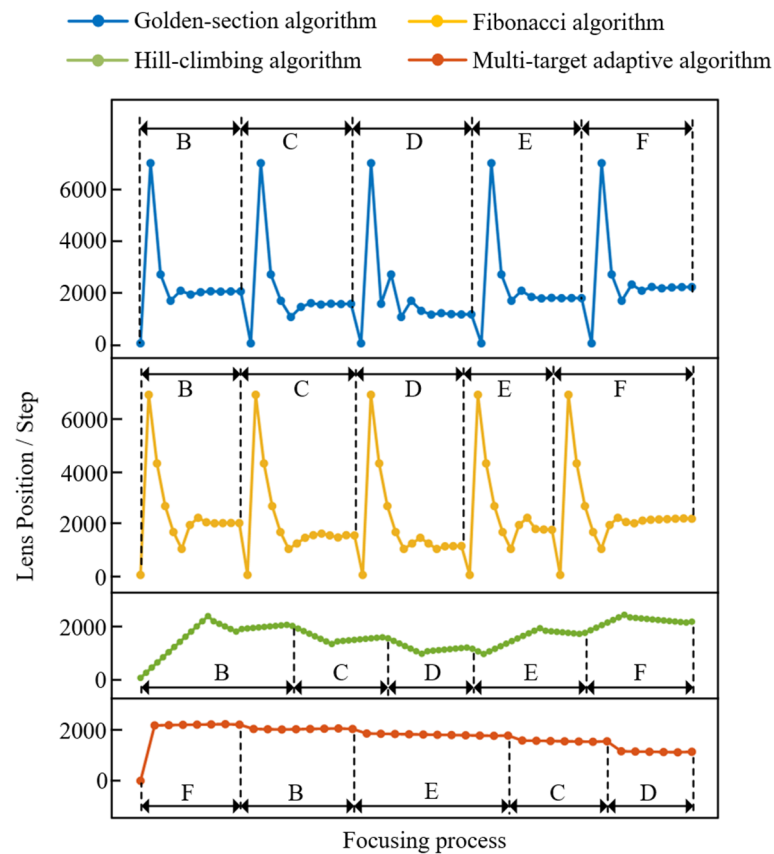


**Figure 10.** Multi-target focusing process.

The final result is shown in Figure 11, where (A) is the initial defocused image and (B)~(F) are the focus results images of each focus target region. The red boxes in the figure indicate the selected focus regions.

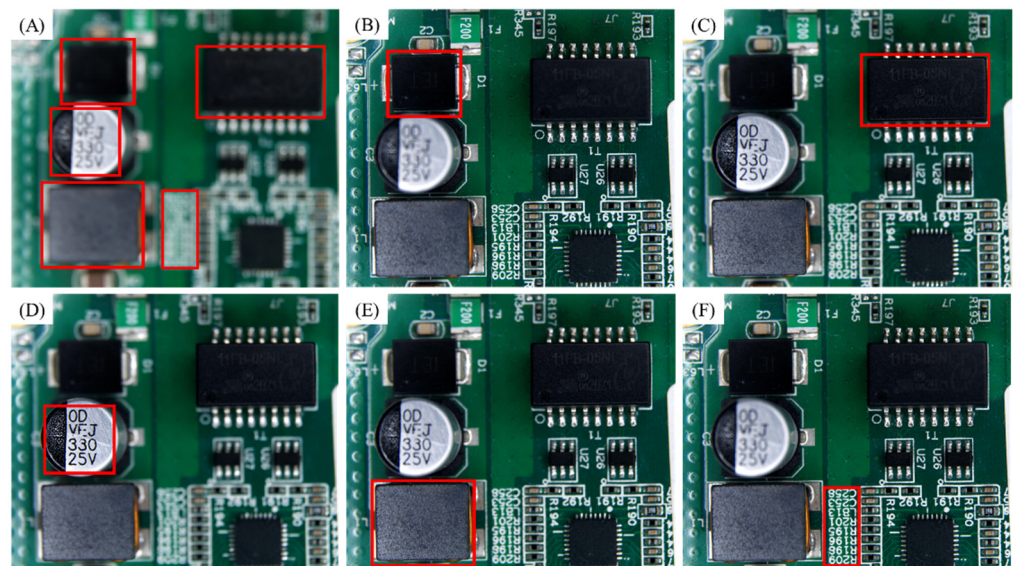


**Figure 11.** Multi-target focusing results: (A) initial defocused image and regions to be focused; (B)~(F) focus results of each region in (A).

The all-in-focus image obtained by merging all the focused regions is shown in Figure 12.

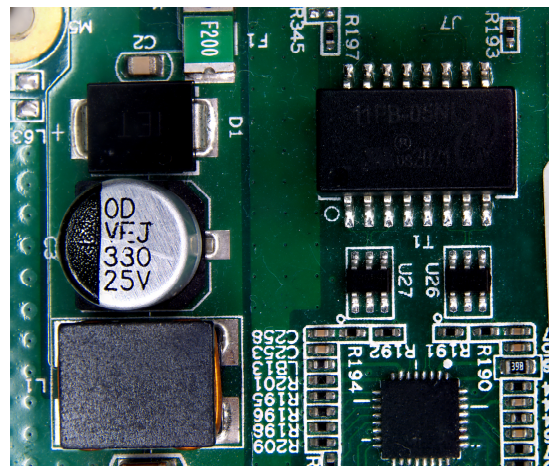


Figure 12. All-in-focus image.

## 5. Conclusions

The multi-target adaptive focusing algorithm proposed in this paper uses multi-scale image matching to calculate the phase difference of the defocused image, improves the accuracy of the phase difference calculation, and uses the image phase difference data for adaptive defocus estimation. Then, the defocus distance estimation of targets in different regions is realized by combining the phase difference and the defocus model, which is used for the optimization of the multi-target focus path and the design of the autofocus step size of the target. The advantage of this method is that through the optimization of the focus path and the adjustment of the adaptive model, the speed of autofocusing on multiple targets in the field of view is accelerated, thereby accelerating the acquisition process of multi-focus images. With the introduction of the adaptive defocus estimation model, this technology breaks away from the limitations of traditional lens calibration and can flexibly respond to changes in lens parameters. It is particularly suitable for scenes with frequent adjustment of magnification and depth, such as medical microscopes and industrial microscopic detection systems. It can provide innovative technical solutions for fields such as biology, medicine, materials science, and industrial detection, and has important application value. Since this method needs to be implemented using phase difference data, and the calculation of phase difference data has specific requirements for image sensors, it is hoped that other methods such as deep learning models can be introduced in the future to replace traditional phase difference calculations, so that the algorithm can be applied to more types of imaging systems.

**Author Contributions:** Conceptualization, N.P. and F.Y.; methodology, N.P.; validation, N.P.; resources, F.Y.; data curation, N.P.; writing—original draft preparation, N.P.; writing—review and editing, F.Y. and N.P.; project administration, F.Y. and N.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** All data included in this study are available upon request to the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Zhu, Y.; Ling, Z.; Zhang, Y. Research progress and prospect of machine vision technology. *J. Graph.* **2020**, *41*, 871–890.
2. Zhang, Y.; Liu, L.; Gong, W.; Yu, H.; Wang, W.; Zhao, C.; Wang, P.; Ueda, T. Autofocus System and Evaluation Methodologies: A Literature Review. *Sens. Mater.* **2018**, *30*, 1165–1174.
3. Xia, H.; Yu, F. Auto Focusing Algorithm of Digital Microscope. *Laser Optoelectron. Prog.* **2021**, *58*, 21–28.
4. Zhu, M.; Yu, L.; Wang, Z.; Ke, Z.; Zhi, C. Review: A Survey on Objective Evaluation of Image Sharpness. *Appl. Sci.* **2023**, *13*, 2652. [[CrossRef](#)]

5. Yang, H.; Feng, X.Z.; Liu, J.Y.; Yang, B. Microscope Autofocus Method Based on Image Evaluation. *Laser Optoelectron. Prog.* **2023**, *60*, 305–315.
6. Guan, Y.; Cui, Z.; Zhu, J.; Zhou, W. Accurate and Fast Autofocusing in Off-Axis Digital Holography Based on Step Reduction Search and Particle Swarm Optimization. *Opt. Lasers Eng.* **2024**, *181*, 108421. [[CrossRef](#)]
7. Ding, Z.; Zheng, S.; Zhang, F.; Li, Q.; Guo, C. Lensfree Auto-Focusing Imaging with Coarse-to-Fine Tuning Method. *Opt. Lasers Eng.* **2024**, *181*, 108366. [[CrossRef](#)]
8. Lin, S.; Zhang, Y.; Yu, L.; Zhou, B.; Luo, X.; Pan, J. Autofocus for Event Cameras. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 16344–16353.
9. Jia, D.; Zhang, C.; Wu, N.; Zhou, J.; Guo, Z. Autofocus Algorithm Using Optimized Laplace Evaluation Function and Enhanced Mountain Climbing Search Algorithm. *Multimed. Tools Appl.* **2022**, *81*, 10299–10311. [[CrossRef](#)]
10. Hamada, M. Imaging Device Including Phase Detection Pixels Arranged to Perform Capturing and to Detect Phase Difference. US20130088621A1, 11 April 2013.
11. Galor Gluskin, M.; Velarde, R.M.; Lee, J. Phase Detection Autofocus Noise Reduction. 20170094149, 30 March 2017.
12. Dey, S.; Roy, R. Autofocus Technologies in Digital Cameras: A Comparative Study. In Proceedings of the Inspiria Excellence Series Multidisciplinary Conference on Future Paradigms in Industry (ICFPI-2018), Siliguri, IN, Indian, 1–2 November 2018.
13. Jang, J.; Yoo, Y.; Kim, J.; Paik, J. Sensor-Based Auto-Focusing System Using Multi-Scale Feature Extraction and Phase Correlation Matching. *Sensors* **2015**, *1*, 5747–5762. [[CrossRef](#)] [[PubMed](#)]
14. Jang, J.; Park, S.; Jo, J.; Kim, J.; Paik, J. Hybrid Auto-Focusing System Using Dual Pixel-Type CMOS Sensor with Contrast Detection Algorithm. In Proceedings of the Imaging Systems and Applications 2016 (Optica Publishing Group, 2016), Heidelberg, Germany, 25–28 July 2016. paper IW3F2.
15. Chan, C.-C. Modeling Phase Shift Data of Phase-Detection Autofocus by Skew-Normal Distribution. *J. Electron. Imaging* **2019**, *28*, 1. [[CrossRef](#)]
16. Zhang, C.; Jia, D.; Wu, N.; Guo, Z.; Ge, H. Autofocus Method Based on Multi Regions of Interest Window for Cervical Smear Images. *Multimed. Tools Appl.* **2022**, *81*, 18783–18805. [[CrossRef](#)]
17. Hou, L.; Zou, J.; Zhang, W.; Chen, Y.; Shao, W.; Li, Y.; Chen, S. An Improved Shape from Focus Method for Measurement of Three-Dimensional Features of Fuel Nozzles. *Sensors* **2022**, *23*, 265. [[CrossRef](#)] [[PubMed](#)]
18. Ren, J.; Fu, X.; Wang, M.; Zhao, T.; Wang, Z.; Feng, K.; Liang, Y.; Wang, S.; Lei, M. Advances in Rapid Three-Dimensional Wide Field Microscopy. *Chin. J. Lasers* **2023**, *50*, 0307104–0307116.
19. Tang, L.; Ge, M.; Dong, W. Design and research of fully automatic push-broom hyperspectral microscopic imaging system. *Chin. Opt.* **2021**, *14*, 1486–1494.
20. Xin, S.; Wadhwa, N.; Xue, T.; Barron, J.T.; Srinivasan, P.P.; Chen, J.; Gkioulekas, I.; Garg, R. Defocus Map Estimation and Deblurring from a Single Dual-Pixel Image. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021; pp. 2208–2218.
21. Her, L.; Yang, X. Research of Image Sharpness Assessment Algorithm for Autofocus. In Proceedings of the 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), Xiamen, China, 5–7 July 2019; pp. 93–98.
22. Dimeo, P.; Sun, L.; Du, X. Fast and Accurate Autofocus Control Using Gaussian Standard Deviation and Gradient-Based Binning. *Opt. Express* **2021**, *29*, 19862–19878. [[CrossRef](#)] [[PubMed](#)]
23. Groen, F.C.A.; Young, I.T.; Lighthart, G. A Comparison of Different Focus Functions for Use in Autofocus Algorithms. *Cytometry* **1985**, *6*, 81–91. [[CrossRef](#)] [[PubMed](#)]
24. Alfaraj, M.; Wang, Y.; Luo, Y. Enhanced isotropic gradient operator. *Geophys. Prospect.* **2014**, *62*, 507–517. [[CrossRef](#)]
25. Kobayashi, M.; Ohmura, M.; Takahashi, H.; Shirai, T.; Sakurai, K.; Ichikawa, T.; Yuzurihara, H.; Inoue, S. High-Definition and High-Sensitivity CMOS Image Sensor with All-Pixel Image Plane Phase-Difference Detection Autofocus. *Jpn. J. Appl. Phys.* **2018**, *57*, 1002B5. [[CrossRef](#)]
26. Yamasaki, T.; Nakamura, T.; Funatsu, R.; Shimamoto, H. Hybrid Autofocus System by Using a Combination of the Sensor-Based Phase-Difference Detection and Focus-Aid Signal. In Proceedings of the 2018 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 12–14 January 2018; pp. 1–2.
27. Shim, E.S.; Lee, K.; Pyo, J.; Choi, W.; Yun, J.; Jung, T.; Lee, K.; Kim, S.; Lee, C.; Baek, S.; et al. All-Directional Dual Pixel Auto Focus Technology in CMOS Image Sensors. In Proceedings of the 2021 Symposium on VLSI Circuits, Kyoto, Japan, 13–19 June 2021; pp. 1–2.
28. Mochizuki, Y.; Imiya, A. Pyramid Transform and Scale-Space Analysis in Image Analysis. In *Outdoor and Large-Scale Real-World Scene Analysis*; Dellaert, F., Frahm, J.-M., Pollefeys, M., Leal-Taixé, L., Rosenhahn, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 78–109.
29. Jang, J.; Park, S.; Jo, J.; Paik, J. Depth Map Generation Using a Single Image Sensor with Phase Masks. *Opt Express* **2016**, *24*, 12868–12878. [[CrossRef](#)] [[PubMed](#)]
30. Širca, S. Method of Least Squares. In *Probability for Physicists*; Springer International Publishing: Cham, Switzerland, 2016; pp. 227–258.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.