Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Design and Implementation of an Image-Based Autofocus Control System for Optical Microscopy

Wei Jiang
Hengxing University of Science and Technology,
Hongyan Liu
Hengxing University of Science and Technology,
Jinsheng Guo
Hengxing University of Science and Technology,
Waseem ur Rahman
engr.waseem778@gmail.com

Hengxing University of Science and Technology,

Article

Keywords: Autofocus, Focus evaluation function, Hill-climbing search, Optical microscopy, Image processing

Posted Date: September 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10773575/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Design and Implementation of an Image-Based Autofocus Control System for Optical Microscopy

Wei Jiang[1]*, Hongyan Liu [1], Jinsheng Guo[1], Waseem ur rahman[1]*
[1]School of Intelligent Manufacturing, Hengxing University of Science and Technology, Qingdao 266100, China.
Corresponding Authors: Dr. Waseem ur Rahman & Wei Jiang
Email: engr.waseem778@gmail.com &706080037@qq.com

Abstract

Autofocusing is an essential function of optical microscopy for industrial inspection and scientific research applications. This paper presents the design and implementation of an image-based autofocus control system for an optical microscope, integrating a stepper motor-driven focusing stage with a Charge-Coupled Device (CCD) camera for real-time image acquisition. Four focus evaluation functions improved average gradient, Roberts operator, Sobel operator, and Laplacian operator were systematically evaluated using silicon wafer specimens with distinct surface textures. An interactive region-of-interest (ROI) selection approach was introduced to reduce computational Complexity and minimize background interference, while a hill-climbing search strategy was employed for efficient focal position detection. Experimental results demonstrate that the improved average gradient method provides the best overall performance in terms of single-peak characteristics, monotonicity, and sensitivity, achieving a focusing time of approximately 4s. The minimum focusing step size of the proposed system is 1.25μm, corresponding to the minimum resolution of the stepper motor. These results demonstrate the effectiveness of the proposed system for automated microscopic inspection tasks requiring high-precision focusing.

Keywords: Autofocus; Focus evaluation function; Hill-climbing search; Optical microscopy; Image processing

1. Introduction

Optical microscopy remains one of the most widely employed tools for observing microscopic structures in materials science, semiconductor manufacturing, biological research, and industrial quality control [1, 2]. An optical microscope's primary function is to magnify and resolve specimen features via a precisely engineered optical system, allowing researchers and engineers to observe fine structural details beyond the resolution of the unaided eye. However, the effectiveness of microscopic observation depends critically on accurate focus; poor focus blurs details, reduces contrast, and causes loss of information, often leading to erroneous analysis or inspection outcomes [3].

Modern microscopy applications, manual focusing presents several significant limitations. First, it imposes considerable demands on the operator’s experience and judgment, as assessing image sharpness is inherently subjective and varies among individuals. Second, prolonged manual adjustment leads to operator fatigue, which further degrades focusing consistency and reliability. Third, in industrial settings where high-throughput inspection is required, manual focusing becomes a significant bottleneck, limiting overall system efficiency. These challenges have motivated extensive research into automatic focusing technologies that

can achieve rapid, accurate, and reproducible focal position detection without human intervention [4].

The automatic focusing techniques reported in the literature can be broadly classified into two categories: active methods and passive methods [5]. Active autofocus systems typically employ external sensors such as ultrasonic, infrared, or laser range finders to measure the distance between the objective lens and the specimen surface. Although these methods can provide rapid distance measurement, they require additional hardware components and may be susceptible to interference from the operating environment. In contrast, passive autofocus methods determine the focal position by analyzing image content, eliminating the need for additional sensors. Passive methods have gained significant popularity due to their lower hardware complexity, lower cost, and non-intrusive nature.

Among passive autofocus approaches, those based on digital image processing have attracted considerable research attention. The fundamental principle underlying image-based autofocus is that a properly focused image contains more high-frequency components and sharper edges compared to a defocused image. Consequently, a variety of focus evaluation functions, also referred to as sharpness evaluation functions or digital focus indices, have been developed to quantify the sharpness of captured images [6]. These functions typically operate on gradient information, statistical properties, or frequency-domain characteristics of the image, and their performance is evaluated based on criteria such as single-peak property, monotonicity, sensitivity, and computational efficiency. Commonly employed focus evaluation functions include gradient-based operators (e.g., average gradient, Roberts operator, Sobel operator), Laplacian-based operators, wavelet-based methods, and frequency-domain measures such as the discrete cosine transform and fast Fourier transform.

In addition to selecting an appropriate focus evaluation function, the search strategy used to locate the optimal focal position is equally critical. Popular search algorithms include the hill-climbing (or mountain-climbing) method, the Fibonacci search, the golden section search, and rule-based adaptive strategies. The hill-climbing method is particularly favored in practical implementations due to its simplicity, low computational overhead, and satisfactory performance in unimodal focus curves [7]. In this approach, the system iteratively compares the focus values of successive images and adjusts the lens position accordingly until the maximum focus value is detected.

The application of automatic focusing technology to optical microscopy has been investigated by numerous research groups. Guo et al. [8] conducted a comprehensive study of image processing techniques for optical microscopes, demonstrating the feasibility of digital-image-based autofocus systems. Liang et al. [9] investigated various autofocusing algorithms and their implementation for microscopic imaging systems. Xu et al. [10] proposed a robust automatic focus algorithm for low-contrast images using a novel contrast measure, addressing the challenge of focusing on specimens with weak texture variations. Tsai and Lin [11] developed an autofocus method based on focal-window searching and tracking for digital imaging applications. Li et al. [12] designed and implemented a complete autofocus system for

optical microscopes, providing valuable insights into system integration and performance optimization.

Despite significant progress in optical microscopic autofocus technology, significant limitations remain in practical engineering applications. The selection of the optimal focus evaluation function is highly dependent on specimen characteristics, and different sharpness operators exhibit considerable performance discrepancies for images with diverse texture features [13]. In addition, an inherent trade-off between focusing accuracy and computational speed remains a critical challenge for autofocus systems. The step size of the focusing motor directly determines the overall focusing performance. A smaller step size achieves higher focusing accuracy but requires more iterations and increases computational time, whereas a larger step size improves efficiency but tends to skip the optimal focal plane, failing to meet the requirements of real-time detection. These drawbacks are particularly prominent in semiconductor lithography wafer inspection. Most existing lightweight focusing algorithms reduce computational consumption by directly abandoning the squared weighting term. Although such strategies improve computational speed, they significantly degrade the response sensitivity to low-contrast lithographic textures on wafers. In contrast, the standard average gradient algorithm delivers excellent focusing performance owing to its ability to capture subtle edge features, yet its pixel-wise floating-point square-root operations impose excessive computational overhead, limiting its applicability for online, real-time wafer detection. Currently, few studies have focused on developing lightweight gradient algorithms that retain bidirectional squared gradient terms while eliminating square-root operations specifically for lithography wafer scenarios, and systematic comparative evaluations of classical sharpness operators are still insufficient. To address these issues, a physical inspection platform consisting of an optical microscope, a wafer stage, and a Z-axis focusing motor is established in this work. MATLAB is used for image acquisition, comparison of various sharpness operators, calculation of quantitative indices, and performance analysis. Meanwhile, motion commands are issued through MATLAB to drive Z-axis motor displacement and implement complete physical autofocus tests. The experimental results verify that the proposed lightweight evaluation function effectively balances focusing accuracy and computational speed, achieving superior comprehensive performance compared with traditional focusing operators.

The objective of this study is to develop and experimentally evaluate a microscopic autofocus system for the inspection of silicon wafer specimens. In this work, an improved average gradient sharpness evaluation operator is proposed, and comparative experiments with the Laplacian, Robert, and Sobel operators are carried out. The experimental results demonstrate that the proposed algorithm can simultaneously satisfy the requirements of focusing real-time performance and imaging accuracy. Comprehensive performance tests are conducted on a self-built microscopic platform equipped with a stepper-motor lifting mechanism, and the system achieves a minimum focusing precision of 1.25 μm. After introducing the region-of-interest extraction strategy, the proposed improved average gradient operator takes

only 4s per focusing process and outperforms the other three conventional evaluation operators in overall performance.

2 Materials and Methods

2.1 Experimental Setup

Figure 1 shows the architecture of the proposed automatic focusing control system. The system comprises four main components: a personal computer (PC) for data processing, a microcontroller-based motor controller, a stepper motor coupled to the microscope focusing stage, and a charge-coupled device (CCD) camera mounted on the optical microscope for image acquisition. The PC communicates with the motor controller via a serial port (RS-232) interface. Upon receiving a command from the PC, the controller drives the stepper motor to rotate by a specified angle, which is mechanically converted into linear displacement of the objective lens barrel (focusing stage). The CCD camera continuously captures video frames of the specimen, and the digitized images are transmitted back to the PC as feedback signals. The PC executes the focus evaluation algorithms to calculate the sharpness value for each captured image and determines the next control command using a hill-climbing search strategy. This closed-loop control process continues until the optimal focal position is identified.

The test specimen is a silicon wafer with clearly visible surface texture patterns. Silicon wafers are ideal test objects for evaluating autofocus systems, because they exhibit well-defined surface features, including parallel stripes and crystallographic patterns, that produce distinct focus evaluation curves. The objective lens was adjusted along the optical axis (z-axis) by the stepper motor, which has a minimum resolution of 1.25μm per step.

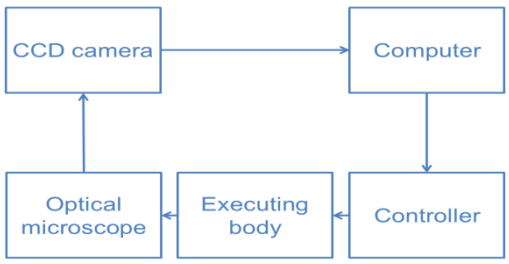


Figure 1. System architecture of the proposed autofocus control system for optical microscopy

2.2 Autofocus Evaluation Algorithm

For a given imaging system, the degree of defocus is directly related to the perceived sharpness of the captured image. When an image is in sharp focus, its fine details are well preserved, resulting in pronounced intensity variations at edges and textured regions. Correspondingly, the high-frequency components in the spatial frequency spectrum of a focused image are significantly stronger than those of a defocused image. This fundamental characteristic enables the construction of various focus evaluation functions that quantify image sharpness.

An effective focus evaluation function should satisfy three essential properties: (1) Unimodality (single-peak property): the function should exhibit exactly one global maximum at the true focal position, with no spurious local maxima that could mislead the search algorithm. (2) Monotonicity: the function value should monotonically increase as the lens approaches the focal plane from either

direction, ensuring that the search direction can be reliably determined. (3) High sensitivity: the slope of the function in the vicinity of the peak should be steep, enabling the system to discriminate between nearly focused and optimally focused positions with high precision [14].

In this study, four representative focus evaluation functions were selected for comparative evaluation: the improved average gradient, the Roberts operator, the Sobel operator, and the Laplacian operator. These operators were chosen because they represent different mathematical approaches to edge detection and sharpness quantification, and they are widely used in the literature on image-based autofocus.

2.2.1. Improved Average Gradient

The improved average gradient is a first-derivative-based focus measure that computes the mean magnitude of the intensity gradient across the image. The underlying principle is that the gradient magnitude is proportional to the rate of change in gray-level intensity; therefore, images with sharper edges and finer details exhibit larger gradient values. Let F(i, j) denote the gray-level intensity at pixel location (i, j) in an image of size M×N pixels. The improved average gradient g is defined as:

$$g = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} [\{F(i,j) - F(i+1,j)\}^2 + \{F(i,j) - F(i,j+1)\}^2]$$

A larger value of g indicates a sharper image with more pronounced edge information. When g attains its global maximum, the corresponding image is considered to be at the optimal focal position.

2.2.2. Roberts Operator

The Roberts operator is one of the earliest edge detection operators. For an input image f(x, y), the Roberts operator is defined as:

$$g(x,y) = \{\left(\sqrt{f(x,y)} - \sqrt{f(x+1,y+1)}\right)^2 + (\sqrt{f(x+1,y)} - \sqrt{f(x,y+1)})^2\}^{\frac{1}{2}}$$

The term g(x, y) represents the gradient magnitude at coordinates (x, y), which characterizes the local edge energy. To obtain a global sharpness measure for the entire image, the Roberts-based focus value J is computed as the mean of the gradient magnitudes over all pixel positions:

$$J = \sum_{N} \sum_{M} g(x,y)$$

2.2.3. Sobel Operator

In each image, each pixel can be expressed as the convolution of two nuclei: one nuclear mainly affects the quality of vertical edges, and the other mainly affects the quality of horizontal edges. And the output value g(x, y) of the point is represented by the maximum value of two convolutions. The expression is as follows:

$$J = \sum_{N} \sum_{M} g(x,y)$$

The corresponding calculation template of sobel operator is as follows.

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

2.2.4. Laplacian Operator

The Laplacian operator is used to calculate the second-order differential of the image, which can be used as a final evaluation value. The expression is as follows.

$$f(n) = \sum_x \sum_y [S(x,y)]^2$$

Where S(x, y) is the result of the image processing by Laplacian operator, the template is as follows:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

2.3 Autofocus Search Strategy

2.3.1 Region-of-Interest Selection

Prior to executing the autofocus search procedure, it is essential to select an appropriate region of interest (ROI) within the captured image. The ROI selection serves two important purposes. First, the smaller ROI can help PC restrict the computation, the overall processing time can be significantly reduced, which is particularly important for real-time applications. Second, computing the focus function over the full image with flat background regions yields suboptimal results. Background interference conceals the specimen's sharpness fluctuations and leads to erroneous focal plane positioning.

In this work, an interactive ROI selection method was implemented. The operator uses a computer mouse to draw a rectangular box on the displayed image, selecting the region that contains the specimen of interest. The system subsequently computes the focus evaluation function exclusively within the selected rectangular region, thereby reducing the computational burden while improving the reliability of the focus detection.

2.3.2 Hill-Climbing Search Algorithm

The autofocus search procedure implemented in this study is based on the hill-climbing search strategy, which is a widely used approach for unimodal optimization problems in autofocus applications. The search flowchart is illustrated in Figure 2.

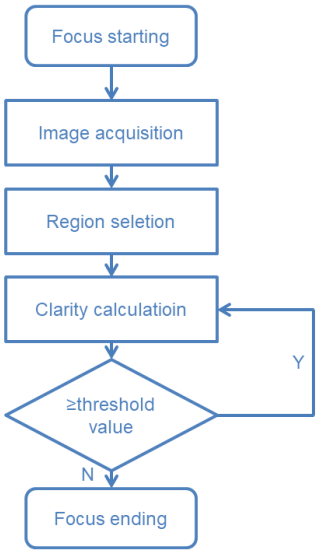


Figure 2. Flowchart of the hill-climbing autofocus search strategy

The search procedure operates as follows. First, the CCD camera captures an initial image, and the

PC selects the ROI and computes the initial focus value. The stepper motor then advances the objective lens by one step, and a second image is captured. The focus metrics of two consecutive images are compared. When the focus score of the second image exceeds the first, this value is updated as the current maximum threshold, and the motor maintains its moving direction. This process is repeated iteratively: at each step, the current focus value is compared with the threshold, and the threshold is updated whenever a larger value is encountered. When the focus value decreases below the current threshold, indicating that the peak has been passed, the stepper motor is commanded to reverse direction and return to the position corresponding to the maximum focus value. At this point, the autofocus procedure is completed, and the system is in sharp focus.

The key advantage of the hill-climbing strategy is its simplicity and computational efficiency. The algorithm requires only pairwise comparisons of successive focus values and does not need to evaluate the entire focus curve before making a decision. This makes it well suited for real-time autofocus applications where rapid response is essential.

3 Results

3.1. Focus Evaluation Function Comparison

The silicon wafer specimen that is used in the experiments exhibits well-defined surface texture patterns, including parallel stripes that are served as excellent features for evaluating autofocus performance. In Figure 3, image (a) shows a defocused image of the silicon wafer specimen, where the surface stripes are barely distinguishable, whereas image (b) presents a focused image in which the stripe patterns are clearly resolved. The difference in image quality between these two states underscores the importance of accurate focusing for microscopic inspection tasks. Prior to autofocus operation, the operator can select the ROI by drawing a rectangular window on the displayed image, as shown in image (c). The system computes the focus evaluation function exclusively within the selected region, thereby significantly reducing computational time and suppressing the influence of background regions.

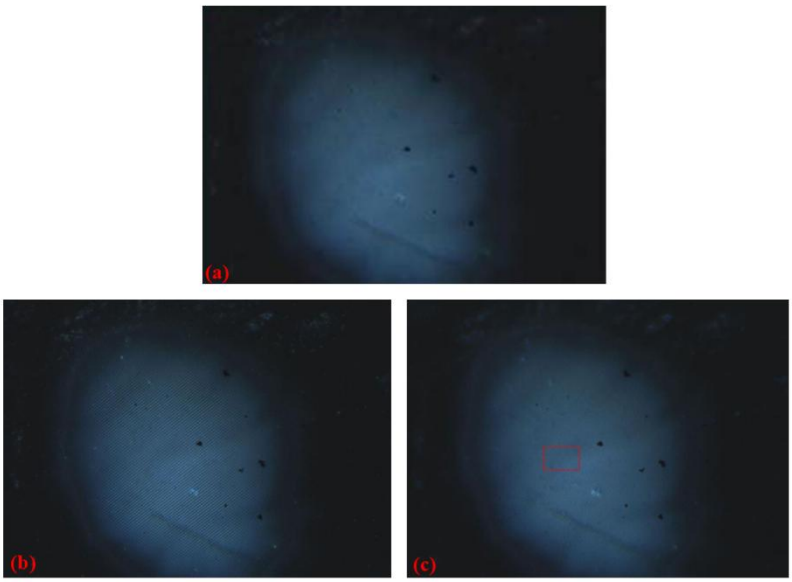


Figure 3. Silicon wafer image

After selecting the ROI, the system captures a sequence of images at regular z-axis intervals and

computes the focus value for each image by using the four evaluation functions. Figure 4 presents the normalized focus evaluation curves obtained using the improved average gradient, Roberts, Sobel, and Laplacian operators. The horizontal axis represents the image sequence number, and the vertical axis represents the normalized focus function value. Images were captured at intervals of 6.25μm. The 17th image in the sequence corresponds to the sharpest focal position, at which all functions attain their respective maxima.

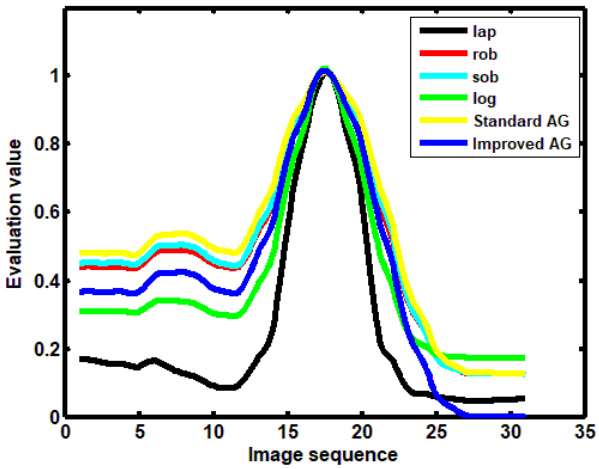


Figure 4. Normalized focus evaluation curves for the four operators

The experimental results reveal noticeable differences in the performance of the four focus evaluation methods. The improved average gradient operator produces a smooth, well-defined unimodal curve with a sharp peak and good monotonicity on both sides of the focal position. The Roberts operator also exhibits good unimodal behavior and sensitivity comparable to that of the improved average gradient operator. The Laplacian operator shows satisfactory single-peak characteristics with moderate sensitivity. By comparison, the improved average gradient operator achieves a slightly faster focusing speed than the Sobel operator around the focal region.

The focus evaluation and curve plotting are implemented in MATLAB. Using the improved average gradient operator, the autofocus procedure completes in approximately 4 seconds, satisfying the real-time requirements of most microscopic inspection applications.

Table 1 provides a comprehensive comparison of the four focus evaluation functions across five performance criteria.

Table 1. Comparison of focus evaluation functions.

Focus Function	Single-Peak Property	Monotonicity	Sensitivity	Computation Time	Overall Performance
Improved gradient	Good	Good	High	4 s	Best
Roberts operator	Good	Good	High	7.1 s	Good
Sobel operator	Moderate	Moderate	Low	8.5 s	Poor
Laplacian operator	Good	Good	Moderate	6.1 s	Good

The results in Table 1 indicate that the improved average gradient operator achieves the best overall performance among the four evaluated functions. It meets the three key characteristics required for an ideal focus evaluation function, namely unimodality, monotonicity, and high sensitivity, and keeps the

computational expense at a moderate level. The Roberts operator performs comparably well, while the Laplacian operator exhibits slightly lower sensitivity but still produces reliable results. The Sobel operator is widely used in edge detection applications, but it performs poorly in this autofocus application due to its low sensitivity to the texture of the silicon wafer specimen.

3.2. Focusing Accuracy Analysis

The focusing resolution of an autofocus system is fundamentally limited by the positioning actuator's step size. If the step size is too large, the system may overshoot the optimal focal plane or fail to resolve small focus variations, resulting in reduced accuracy. To systematically investigate the achievable focusing accuracy of the proposed system, experiments are conducted using the Laplacian operator at three different step sizes: 6.25 μm, 3.75 μm, and 1.25 μm. The improved average gradient operator is selected for this analysis because it provides a good balance between performance and computational cost.

Figure 5 shows the focus evaluation curve obtained with a step size of 6.25μm using the improved average gradient operator. The curve exhibits excellent unimodality with no bias, and there are no visible artifacts or spurious fluctuations near the peak. The system demonstrates good resolution at this step size, and the achievable focusing accuracy is determined to be 6.25 μm.

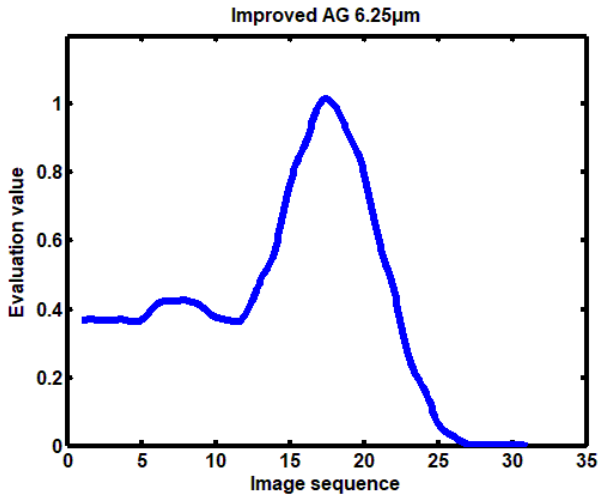


Figure 5. Focus evaluation curve at a step size of 6.25μm

Figure 6 presents the focus evaluation curve obtained with a step size of 3.75μm. The single-peak characteristic remains clearly visible, and the curve is notably smoother compared to the 6.25μm case due to the increased sampling density along the z-axis. The system achieves a focusing accuracy of 3.75μm at this step size.

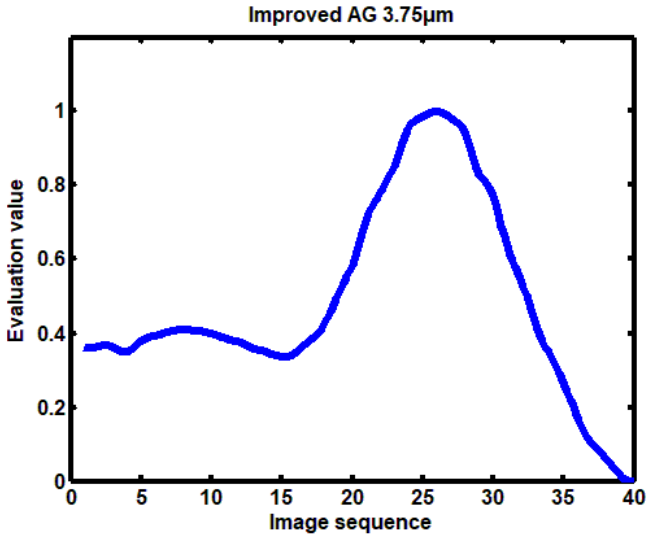


Figure 6. Focus evaluation curve at a step size of 3.75μm

Figure 7 shows the focus evaluation curve obtained with the smallest step size of 1.25μm. Although minor fluctuations appear near the focal peak due to measurement noise and discrete image sampling, the overall trend of the curve remains steep and well-defined. The focusing function curve continues to satisfy the essential properties of a focus evaluation function, and the system achieves a focusing accuracy of 1.25μm, which corresponds to the minimum resolution of the stepper motor employed in this study.

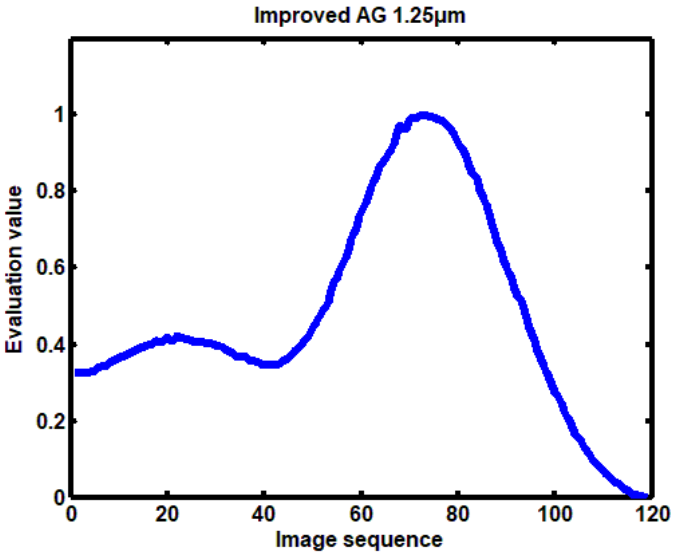


Figure 7. Focus evaluation curve at a step size of 1.25μm

Table 2 summarizes the focusing performance at the three tested step sizes.

Table 2. Focusing accuracy at different step sizes using the improved average gradient operator.

Step Size (μm)	Single-Peak Quality	Curve Smoothness	Achievable Accuracy	Resolution Capability
6.25	Good	Good	6.25μm	Good
3.75	Good	Smooth	3.75μm	Good
1.25	Good	Steep	1.25μm	Excellent

Table 3 summarizes the overall system performance parameters.

Table 3. System performance summary.

Parameter	Description
Minimum focusing accuracy	1.25μm
Focusing time (improved average gradient operator)	4 seconds
Search strategy	Hill-climbing
Focus area selection	Interactive ROI
Test specimen	Silicon wafer
Motor type	Stepper motor
Motor minimum resolution	1.25μm/step
Communication interface	RS-232 serial port
Image sensor	CCD camera

The results demonstrate that the proposed system achieves a minimum focusing resolution of 1.25μm, which is sufficient for a wide range of microscopic inspection applications in semiconductor manufacturing and materials science. The focusing time of approximately 4 seconds represents a significant improvement over manual focusing and meets the requirements for automated inspection workflows.

4 Discussion

The experimental results presented in Section 3 reveal significant and informative differences in the performance characteristics of the four evaluated focus evaluation methods. Understanding the underlying causes of these differences is essential for selecting the most appropriate focus measure for a given application and for guiding the design of future autofocus systems.

The improved average gradient operator achieves the best overall performance among the four tested methods, producing a smooth, unimodal curve with good monotonicity and high sensitivity near the focal position. The superior performance of the improved average gradient operator can be attributed to several factors [15]. First, the silicon wafer specimen used in the experiments has distinct directional surface textures, and the parallel stripes shown in Figure 4 produce strong first-derivative responses. The improved average gradient operator computes gray-level variation in both horizontal and vertical directions, making it highly suitable for the directional textures of silicon wafers. Second, the gradient magnitude is computed from first-order differences between adjacent pixels, which provides a direct measure of edge strength. For specimens with well-defined edge structures, this direct measurement approach yields higher sensitivity than second-derivative operators or more complex edge-detection schemes. Third, since the improved average gradient operator only requires simple arithmetic operations on pixel intensities, it runs quickly and completes one autofocus cycle in about 4 seconds, which makes it well suited for real-time applications.

The Roberts operator exhibited performance comparable to that of the improved average gradient operator, with good unimodality and high sensitivity [16]. The Roberts operator is also a first-derivative-based measure, but it adopts diagonal differences instead of axial differences. Consequently, the Roberts operator and the improved average gradient operator achieve similar focusing performance. This consistency indicates that the silicon wafer's surface texture yields strong gradient responses in both the axial and diagonal directions. However, the Roberts operator uses a smaller kernel, which makes it more sensitive to noise and may result in less stable performance on specimens with weaker texture patterns.

The Sobel operator produced the poorest performance among the four evaluated functions, characterized by a relatively flat response curve with low sensitivity near the focal position [17]. This result is somewhat counterintuitive because the Sobel operator is widely used and generally effective in edge detection applications. Several factors may explain this apparent discrepancy. First, the Sobel operator uses weighted kernels that place greater emphasis on the center pixel, which can suppress responses to fine texture details when the image is slightly defocused. Second, the Sobel operator's output is dominated by strong edges, and if the specimen lacks pronounced edges or if the edge contrast is relatively uniform across the defocus range, the resulting focus curve will exhibit low discrimination. Third, the silicon wafer surface texture consists primarily of fine, closely spaced parallel stripes rather than isolated strong edges; this type of texture may not produce the strong bilateral gradient responses that the Sobel operator is designed to detect. This finding is consistent with the observation reported by Furnari et al. [18] that gradient-based focus measures generally outperform Sobel-based measures for specimens with fine directional textures.

The Laplacian operator demonstrates satisfactory performance with good unimodality and moderate sensitivity. As a second-derivative operator, the Laplacian responds to the rate of change of the gradient rather than the gradient itself, making it particularly sensitive to fine details and rapid intensity transitions. The moderate sensitivity observed in the experiments suggests that while the Laplacian effectively captures the fine texture variations of the silicon wafer, its response to gradual defocus changes is less pronounced than that of first-derivative operators. This behavior is consistent with the theoretical properties of second-derivative operators, which tend to amplify high-frequency noise and may produce less stable focus curves in the presence of image noise. Nevertheless, the Laplacian operator remains a viable choice for applications where fine detail preservation is critical.

The experimental results clearly demonstrate that the step size of the positioning actuator is a critical parameter that directly determines the achievable focusing accuracy of the autofocus system [19]. The relationship between step size and focusing accuracy can be understood from both theoretical and practical perspectives.

From a theoretical standpoint, the achievable focusing accuracy is fundamentally limited by the motor's step size. In practice, the actual accuracy may be further degraded by factors such as mechanical backlash in the focusing stage, image sensor noise, and quantization errors in the computation of focus values.

The experimental results obtained at the three tested step sizes (6.25 μm, 3.75 μm, and 1.25 μm) are consistent with this theoretical analysis. At the largest step size of 6.25 μm, the focus curve is smooth

and well-defined, but the accuracy is limited to 6.25 μm. Reducing the step size to 3.75 μm yields smoother focus curves with improved sampling density, and the accuracy improves correspondingly to 3.75μm. At the finest step size of 1.25 μm, the system achieves its maximum theoretical accuracy of 1.25 μm, although minor fluctuations become visible near the peak of the focus curve.

These fluctuations at the smallest step size can be attributed to several factors. First, as the step size decreases, the image-to-image variation in focus value decreases, and the influence of image noise and sensor quantization becomes relatively more pronounced than the signal variation. Second, the discrete nature of digital image sampling means that small changes in the z-position may not always produce monotonic changes in the computed focus value, particularly when the specimen texture has periodic structures that interact with the pixel sampling grid. Third, mechanical vibrations and stepper motor resonance may introduce small positional errors that become significant at fine step sizes. Despite these challenges, the focus curve with a 1.25 μm step maintains good quality for reliable peak detection, verifying that the system can run at the full resolution of the motor.

The experimental results demonstrate the effectiveness of the proposed autofocus system; several limitations should be acknowledged. First, the system was evaluated using a single specimen type (silicon wafer), and the performance of different focus evaluation functions may vary significantly when applied to specimens with different texture characteristics, such as biological tissues, metallic surfaces, or transparent materials. Future work will involve comprehensive tests on various specimens to formulate general guidelines for selecting focus evaluation functions [20].

Second, the hill-climbing search strategy employed in this study, while simple and efficient, is susceptible to local maxima in the focus curve. Although the evaluated focus functions all exhibited good unimodality on the silicon wafer specimen, more complex specimens or imaging conditions with partial occlusions may produce multi-modal focus curves that could mislead the hill-climbing algorithm. The integration of more robust search strategies, such as the Fibonacci search or rule-based adaptive step-size methods, could improve the system's reliability in challenging scenarios.

Third, the current system requires manual selection of the ROI prior to each focusing operation [21]. While this approach provides flexibility and allows the operator to target specific regions of interest, it introduces a manual intervention step that limits full automation. Future implementations could incorporate automatic ROI detection algorithms based on image segmentation or saliency detection to achieve a fully automated focusing workflow [22].

Fourth, the focusing time of approximately 4 seconds, while acceptable for many applications, may be insufficient for high-speed industrial inspection tasks. Several approaches could be explored to reduce the focusing time, including: (1) the use of parallel processing techniques such as GPU acceleration to compute focus values more rapidly; (2) the implementation of predictive search strategies that estimate the focal position based on partial curve data; and (3) the optimization of the serial communication protocol between the PC and the motor controller to reduce latency.

Finally, the system currently operates as an open-loop control system in which the stepper motor position is commanded but not directly verified through feedback. The integration of an encoder or linear displacement sensor could provide closed-loop position verification, further improving the reliability and repeatability of the focusing process.

Despite these limitations, the proposed system offers several practical advantages that make it well-suited for deployment in industrial and laboratory settings. The use of a stepper motor-driven focusing mechanism provides a cost-effective solution with sufficient accuracy for most microscopic inspection tasks. The interactive ROI selection allows operators to target specific features of interest, providing valuable flexibility in quality control applications where different inspection criteria may apply to different regions of the specimen. The integration of four focus evaluation functions provides the user with options to select the most appropriate function based on the specimen characteristics.

The experimental findings of this study provide practical guidance for engineers and researchers developing autofocus systems for optical microscopy [23]. The demonstrated superiority of the improved average gradient operator for silicon wafer specimens suggests that first-derivative-based measures should be preferentially considered for specimens with directional surface textures. The clear relationship between step size and achievable accuracy provides a quantitative basis for selecting the appropriate motor resolution based on the application requirements [24]. These insights contribute to the informed design of autofocus systems that balance cost, accuracy, and speed for specific use cases [25].

5 Conclusion

This paper has presented the design, implementation, and experimental evaluation of an automatic focusing control system for optical microscopy. The system integrates a stepper motor-driven focusing stage, a CCD camera for image acquisition, and a PC-based control algorithm to achieve fully automated focal position detection. The principal findings and contributions of this work are summarized as follows.

First, a systematic comparative evaluation of four focus evaluation functions was conducted using silicon wafer specimens. The experimental results demonstrate that the improved average gradient operator achieves the best overall performance, satisfying all three essential properties of an ideal focus function (unimodality, monotonicity, and high sensitivity) while maintaining a computational efficiency that enables autofocus completion in approximately 4 seconds. The Roberts and Laplacian operators also produced satisfactory results, whereas the Sobel operator showed low sensitivity for the tested specimen type. Second, the relationship between motor step size and achievable focusing accuracy was quantitatively characterized through experiments at three step sizes: 6.25 μm, 3.75 μm, and 1.25 μm. The results confirm that the focusing accuracy is directly determined by the step size of the positioning actuator, with a minimum achievable accuracy of 1.25 μm corresponding to the stepper motor's minimum resolution. Third, the implementation of an interactive ROI selection mechanism effectively reduced computational overhead and suppressed background interference, thereby improving processing speed and enhancing the reliability of focus detection. The proposed system achieves a minimum focusing accuracy of 1.25 μm with a processing time of approximately 4 seconds, which is sufficient for a wide range of automated microscopic inspection applications in semiconductor manufacturing, materials characterization, and quality control.

Future work will focus on several directions: (1) extending the evaluation to a broader range of specimen types to establish general guidelines for focus function selection; (2) implementing adaptive search strategies with variable step sizes to improve both speed and accuracy; (3) exploring hardware

acceleration techniques such as GPU-based parallel processing to reduce the autofocus time; and (4) developing automatic ROI detection algorithms to eliminate the need for manual intervention and achieve a fully automated focusing workflow. Additionally, integrating machine learning techniques for intelligent focus prediction and incorporating closed-loop position feedback mechanisms are promising directions for further enhancing system performance.

Author Contributions: Wei Jiang: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing-original draft preparation, writing review and editing; Hongyan Liu: Software, writing review and editing; Jincheng Guo: formal analysis, resources; Waseem Ur Rahman: supervision, writing review and editing.

Funding: This research received no external funding.

Data Availability: The data are available from the corresponding author upon reasonable request.

Declaration of competing interests: The authors declare no conflicts of interest.

Reference:

[1] W. Lee, H. S. Nam, Y. G. Kim, Y. J. Kim, J. H. Lee, and H. Yoo, "Robust autofocusing for scanning electron microscopy based on a dual deep learning network," *Scientific Reports,* vol. 11, no. 1, p. 20933, 2021.

[2] J. Zhang, Y. F. Fu, H. Shen, Q. Liu, L. N. Sun, and L. G. Chen, "Precision autofocus in optical microscopy with liquid lenses controlled by deep reinforcement learning," *Microsystems & Nanoengineering,* vol. 10, no. 1, 2024.

[3] M. B. Sanz, F. M. Sánchez, and S. Borromeo, "An algorithm selection methodology for automated focusing in optical microscopy," *Microscopy research and technique,* no. 5, p. 85, 2022.

[4] X. Qu, C. Ma, W. Hu, H. Den, and S. Yang, "A robust autofocusing method for microscopic imaging based on an event camera," *Optics and lasers in enginering,* vol. 175, no. 000, p. 11, 2024.

[5] A. Masini, M. Diani, A. Porta, and M. Romagnoli, "Automatic focusing techniques for infrared sensors," *Proceedings of SPIE - The International Society for Optical Engineering,* vol. 6737, 2007.

[6] Y. Yao, B. Abidi, N. Doggaz, and M. Abidi, "Evaluation of sharpness measures and search algorithms for the auto focusing of high-magnification images," *International Society for Optics and Photonics,* 2006.

[7] P. Bettinger, "Hill-Climbing Search," *Springer, Cham,* 2025.

[8] K. Guo, J. Liao, Z. Bian, X. Heng, and G. Zheng, "InstantScope: a low-cost whole slide imaging system with instant focal plane detection," *Biomedical Optics Express,* vol. 6, no. 9, pp. 3210–3216, 2015.

[9] Y. Liang, M. Yan, Z. Tang, Z. He, and J. Liu, "Learning to Autofocus Based on Gradient Boosting Machine for Optical Microscopy," *Optik - International Journal for Light and Electron Optics,* vol. 198, 2019.

[10] X. Xu, Y. Wang, J. Tang, X. Zhang, and X. Liu, "Robust Automatic Focus Algorithm for Low Contrast Images Using a New Contrast Measure," *Sensors,* vol. 11, no. 9, pp. 8281–8294, 2011.

[11] T. H. Tsai and C. Y. Lin, "A new auto-focus method based on focal window searching and

tracking approach for digital camera," *IEEE,* 2008.

[12] Li and Qingxiang, "Autofocus system for microscope," *Optical Engineering,* vol. 41, no. 6, pp. 1289–1294, 2002.

[13] S. Zhang, J. H. Liu, S. Li, G. Jin, and T. An, "The Research of Mixed Programming Auto-Focus Based on Image Processing," *DBLP,* 2010.

[14] W. J. Lee, H. S. Nam, Y. G. Kim, Y. J. Kim, J. H. Lee, and H. Yoo, "Robust autofocusing for scanning electron microscopy based on a dual deep learning network," *Scientific Reports,* vol. 11, 2021.

[15] W. Jianwei, "An Improved Method of Color Image Edge Detection Based on One Order Gradient Operator," *International Journal of Hybrid Information Technology,* vol. 6, no. 5, pp. 151–162, 2013.

[16] W. Fangchao, Z. Min, and G. Limei, "Image Edge Detection Algorithm of Roberts Operator," *Journal of Detection & Control,* 2016.

[17] L. Z. Qi and L. Cheng, "Edge Thinning Based on Sobel Operator," *JOURNAL OF IMAGE AND GRAPHICS,* 2000.

[18] A. Furnari, G. M. Farinella, A. R. Bruna, and S. Battiato, "Distortion adaptive Sobel filters for the gradient estimation of wide angle images ☆," *Journal of Visual Communication and Image Representation,* vol. 46, no. jul., pp. 165 – 175, 2017.

[19] C. Luo, Y. Liang, J. Gao, G. Lei, and C. Han, "A Scene-Adaptive Phase Gradient Autofocus Algorithm," *Remote Sensing,* vol. 17, no. 2, 2025.

[20] J. Li *et al.*, "Focusing and leveling in dual stage lithographic system," *5th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Design, Manufacturing, and Testing of Micro- and Nano-Optical Devices and Systems,* vol. 7657, p. 76571H, 2010.

[21] C. Zhang, D. Jia, N. Wu, Z. Guo, and H. Ge, "Autofocus method based on multi regions of interest window for cervical smear images," *Multimedia tools and applications,* no. 13, p. 81, 2022.

[22] Y. Huang *et al.*, "A bifocal compound liquid lens with continuous zoom based on selective wettability," *Optics letters,* vol. 47, no. 15, pp. 3824–3827, 2022.

[23] L. Silvestri, M. C. Müllenbroich, I. Costantini, A. P. D. Giovanna, and F. S. Pavone, "Universal autofocus for quantitative volumetric microscopy of whole mouse brains," *Nature Methods*.

[24] Y. Luo, L. Huang, Y. Rivenson, and A. Ozcan, "Single-shot autofocusing of microscopy images using deep learning," 2020.

[25] H. Tobón-Maya *et al.*, "Autofocusing method for active Hadamard single-pixel microscopy using gradient descent algorithms," *Optics and Lasers in Engineering,* vol. 185, no. 000, 2025.