

<https://doi.org/10.1038/s44172-024-00288-z>

Fast autofocusing based on single-pixel moment detection

Check for updates

Huiling Chen^{1,2}, Dongfeng Shi^{1,2,3} ✉, Zijun Guo^{1,2}, Runbo Jiang^{1,2}, Linbin Zha^{1,2}, Yingjian Wang^{1,2,3} & Jan Flusser⁴

Traditional image processing-based autofocusing techniques require the acquisition, storage, and processing of large amounts of image sequences, constraining focusing speed and cost. Here we propose an autofocusing technique, which directly and exactly acquires the geometric moments of the target object in real time at different locations by means of a proper image modulation and detection by a single-pixel detector. An autofocusing criterion is then formulated using the central moments, and the fast acquisition of the focal point is achieved by searching for the position that minimizes the criterion. Theoretical analysis and experimental validation of the method are performed and the results show that the method can achieve fast and accurate autofocusing. The proposed method requires only three single-pixel detections for each focusing position of the target object to evaluate the focusing criterion without imaging the target object. The method does not require any active object-to-camera distance measurement. Comparing to local differential methods such as contrast or gradient measurement, our method is more stable to noise and requires very little data compared with the traditional image processing methods. It may find a wide range of potential applications and prospects, particularly in low-light imaging and near-infra imaging, where the level of noise is typically high.

Rapid and accurate autofocusing is an important issue in imaging technologies^{1,2}. There has been increasing research on passive autofocusing techniques (i.e., techniques that do not require any distance measurement by laser range-finders or similar sensors) in recent years. These techniques can be classified into the following categories: traditional methods using image information for autofocusing, such as contrast-based autofocusing and phase autofocusing^{3–6} or evaluating advanced local differential focus measures⁷ and the use of artificial intelligence approaches, such as the deep learning autofocusing techniques^{8–10}. However, the AI-based methods require a large number of image sequences, being computationally intensive, and being applicable only to specific scenes^{11,12}. Conventional imaging techniques rely on surface array detectors; however, surface array detectors are expensive and sometimes unresponsive in special wavelength bands such as infrared and terahertz.

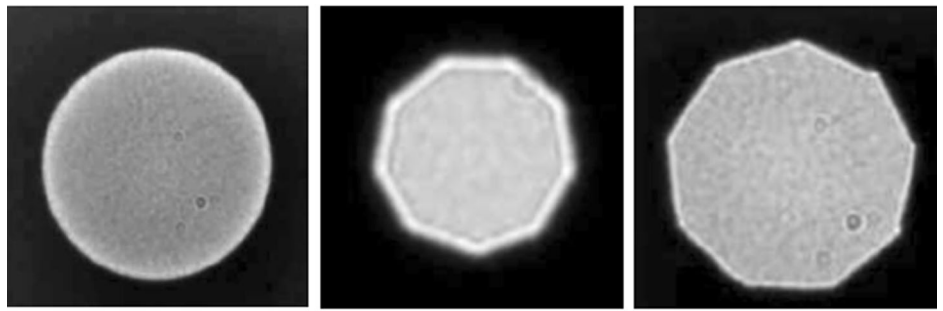
Single-pixel imaging (SPI) is a novel imaging technique¹³, which was invented in connection with compressed sensing. In SPI, optical information is modulated by a spatial-optical modulator and projected onto a single-point photosensitive detector. SPI originally appeared in connection with

compressive sensing, where a complete image of the target is constructed using an algorithm by varying the modulation patterns of the spatial-optical modulator and obtaining measurements. Here we employ SPI in an original way for a real-time evaluation of a focusing criterion.

The typical advantages of SPI over conventional imaging techniques¹⁴ are as follows. First, SPI has high sensitivity. Conventional imaging techniques typically use an array of pixels to capture an image, in which each pixel is responsible for recording the intensity of light at a specific location. However, SPI takes a different approach and captures image information by measuring the total light intensity in the entire scene. This global measurement approach makes SPI highly sensitive and capable of capturing extremely faint light signals¹⁵. Second, SPI responds over a wide spectral range. Different wavelengths of light have different properties and information; therefore, the ability to acquire images over a wide spectral range is critical for many applications. SPI uses a single-point detector with a much wider response spectrum, making it valuable for important applications in bands such as infrared and terahertz¹⁶. Since SPI actually performs a spatial integration of the incoming light, it is very robust to noise.

¹School of Environmental Science and Optoelectronic Technology, University of Science and Technology of China, Hefei, 230026, China. ²Key Laboratory of Atmospheric Optics, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, 230031, China. ³Advanced Laser Technology Laboratory of Anhui Province, Hefei, 230037, China. ⁴Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czech Republic. ✉e-mail: dfshi@aiofm.ac.cn

Fig. 1 | Examples of real point spread functions (PSF's) that cause out-of-focus image blur. Circular PSF of a fully open aperture (left) and polygonal PSF's of partially open aperture. The shape is given by diaphragm blades. The images of the PSF's were obtained as photos of a single bright point far from the focus distance.



The use of non-imaging methods to achieve fast focusing is of great value for many systems, such as conventional surface arrays and new SPI. Recently, Shaoting Qi et al. studied dual-modulation image-free active self-focusing and its application to Fourier SPI¹⁷. However, this technique requires the addition of auxiliary grating devices, which increases the complexity of the imaging systems. Zilin Deng et al. proposed an autofocus technique that employs Fourier coefficients as a criterion¹⁸. Ten Fourier coefficients are acquired at each position and combined as a criterion for autofocus; the number of detections required at each position is 30 when using a three-step phase-shift method. However, universal rules for selecting quantitative spectral values for a large number of rich images are lacking.

In this paper, we aim at a development of a novel autofocusing technique that should work in real time, should be robust to noise in image data, should be implemented directly in the camera hardware and should not require any additional range measurement. To meet all these requirements, we propose to use focusing criterion which is based on image moments. To achieve the desired speed, we propose a real-time technique for moment calculation by means of SPI.

We provide a detailed theoretical analysis of the universal autofocusing judgment rules. The moment values of the target image at different positions of the imaging lens are obtained by direct detection, and the central moments are used as a criterion to search for the position of the imaging lens with the smallest central moment to achieve autofocusing. In comparison with traditional image processing autofocusing methods, our method is a non-imaging autofocusing method with extremely fast data acquisition, transmission, and processing. It manifests higher autofocusing speed with only three detection intensity values at different positions¹⁸ and does not require additional devices, such as gratings¹⁷. Moreover, our method can be applied to diverse real-world images.

Methods

If the image is not perfectly focused, it appears blurred, which means its high frequencies are suppressed. If the scene is flat, the blurred image $g(x, y)$ can be modeled as a convolution of a focused image $f(x, y)$ and a point spread function $h(x, y)$ ¹⁹; the convolution symbol is denoted by $*$, where (x, y) are the coordinates of the two-dimensional (2D) image

$$g(x, y) = (f * h)(x, y). \tag{1}$$

In case of images of 3D scenes, the convolution model holds only locally and/or approximately.

Common assumptions are the positivity of the image

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy > 0 \tag{2}$$

and overall conservation of energy in the imaging system

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) dx dy = 1. \tag{3}$$

In case of out-of-focus blur, $h(x, y)$ is a constant function inside a polygon which has the same shape as the aperture of the camera. If the aperture is fully open, $h(x, y)$ is a circle (see Fig. 1). For a detailed explanation of out-of-focus blur see²⁰, Chapter 6.

Now we explain how image moments are changed under wrong focus. Geometric moment of order $(p + q)$ of image f is defined as

$$m_{pq}^{(f)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy. \tag{4}$$

Since geometric moments are not shift-invariant, we work with central moments $u_{pq}^{(f)}$ that are independent on the position of the object in the field of view of the camera

$$u_{pq}^{(f)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_c)^p (y - y_c)^q f(x, y) dx dy, \tag{5}$$

where (x_c, y_c) are the coordinates of the centroid of $f(x, y)$

$$x_c = \frac{m_{10}}{m_{00}}, y_c = \frac{m_{01}}{m_{00}}. \tag{6}$$

The moments of the blurred image $g(x, y)$ can be expressed in terms of moments of the original image $f(x, y)$ and the point spread function $h(x, y)$ as follows²¹:

$$m_{pq}^{(g)} = \sum_{i=0}^p \sum_{j=0}^q \binom{p}{i} \binom{q}{j} m_{ij}^{(h)} m_{p-i, q-j}^{(f)} \tag{7}$$

$$u_{pq}^{(g)} = \sum_{i=0}^p \sum_{j=0}^q \binom{p}{i} \binom{q}{j} u_{ij}^{(h)} u_{p-i, q-j}^{(f)}. \tag{8}$$

To derive a moment-based focus measure, we employ the second-order moments. First-order moments do not carry enough information about the focus and higher-order moments do not bring any advantage over the second order. So, from (8) we obtain

$$M_1 = u_{20}^{(g)} = u_{00}^{(h)} u_{20}^{(f)} + 2u_{10}^{(h)} u_{10}^{(f)} + u_{20}^{(h)} u_{00}^{(f)} \tag{9}$$

$$M_2 = u_{02}^{(g)} = u_{00}^{(h)} u_{02}^{(f)} + 2u_{01}^{(h)} u_{01}^{(f)} + u_{02}^{(h)} u_{00}^{(f)}. \tag{10}$$

Since the imaging system is assumed to be energy-preserving, we have $u_{00}^{(h)} = 1$. Because $u_{10} = u_{01} = 0$ for arbitrary function, Eqs. (9) and (10) obtain the form

$$M_1 = u_{20}^{(g)} = u_{20}^{(f)} + u_{20}^{(h)} u_{00}^{(f)}, \tag{11}$$

$$M_2 = u_{02}^{(g)} = u_{02}^{(f)} + u_{02}^{(h)} u_{00}^{(f)}. \tag{12}$$

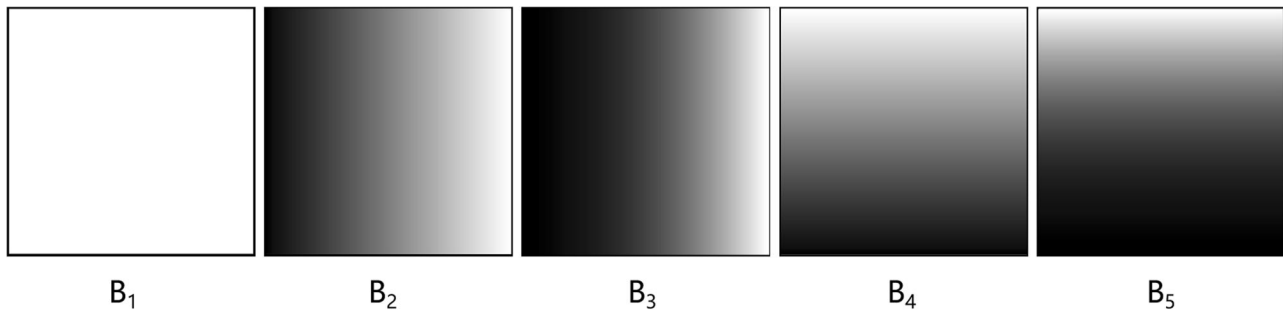


Fig. 2 | The modulation patterns for geometric moment calculation. $B_1, B_2, B_3, B_4,$ and B_5 are used to obtain $m_{00}, m_{10}, m_{20}, m_{01},$ and m_{02} values. (Note that the image coordinates use a Cartesian coordinate system with the origin at the lower left corner.)

The moments $u_{20}^{(h)}$ and $u_{02}^{(h)}$ are in fact variances of the PSF in x and y direction, respectively. The variance σ of the PSF indicates the degree of spatial dispersion of light energy; a smaller variance implies a more concentrated distribution of the PSF, leading to sharper images, whereas a larger variance corresponds to a more dispersed PSF distribution and thus, blurrier images²². The larger blur, the higher these moments. In the optimal focus position we have $u_{20}^{(h)} = u_{02}^{(h)} = 0$ and both M_1 and M_2 reach their minimum (note that the other moments in Eqs. (11) and (12) are always positive and independent on the focus setting).

Therefore, M_1 and M_2 can be used to measure the degree of image blurring whatever the particular PSF is. Taking $M_1 + M_2$ as the criterion is even more convenient choice—this quantity measures the blur in both directions and is invariant to image rotation. To calculate the central moments, we use the relationship between the center of mass and the geometric moments. The central moments of the blurred image $g(x,y)$ can be expressed as

$$M_1 = u_{20}^{(g)} = m_{20}^{(g)} - \frac{m_{10}^{(g)^2}{m_{00}^{(g)}} \tag{13}$$

$$M_2 = u_{02}^{(g)} = m_{02}^{(g)} - \frac{m_{01}^{(g)^2}{m_{00}^{(g)}} \tag{14}$$

Hence, only an algorithm for calculating geometric moments is sufficient to evaluate our focus measure.

Traditional moment calculation from the definition replaces the integral in Eq. (4) by a sum over all image pixels. This is expensive as it requires $O(N^2)$ operations for an $N \times N$ image. To reduce this complexity to $O(1)$, we take the advantage of SPI and optical modulation unit called digital micromirror device (DMD). We spatially modulate the image by a polynomial, which is equivalent to multiplication. Then the entire modulated image is captured by a single-pixel detector, which is equivalent to integration of the modulated light over the field of view. The obtained scalar value is the image moment, which corresponds to the polynomial used for modulation. By changing the pattern on DMD, we can calculate any set of moments in real time.

To describe this process formally, we multiply the scene information using a series of modulation patterns $B_n(x, y)$, and use a single-pixel detector to record the total amount of light I_n reflected or transmitted from the scene, where n denotes the index of the modulation pattern

$$I_n = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) B_n(x, y) dx dy. \tag{15}$$

If

$$B_n(x, y) = x^p y^q \tag{16}$$

Then obviously $I_n = m_{pq}$.

Using the detected value as the geometric moments, fast acquisition and calculation of the central moments can be achieved. Being a non-imaging method, the proposed method does not require image acquisition of the target and requires only a small amount of computation and data, which enables fast and accurate autofocusing. Compared to existing methods^{17,18}, this method does not require additional hardware devices, and the modulation sampling intensity values are lower with higher efficiency. At the same time, the integration of the light into a single value ensures the robustness to noise. The generated modulation patterns are shown in Fig. 2. When the scene information is modulated using $B_1, B_2, B_3, B_4,$ and B_5 , the detected intensity values correspond to geometric moments $m_{00}, m_{10}, m_{20}, m_{01},$ and m_{02} of the scene, respectively. Notice, to enhance the clarity of the modulation pattern boundaries in the plotting, a black border has been added at each boundary of the modulation modes for better visibility.

Results

Simulations results

In the first series of experiments, we used computer simulations only. No actual DMD and SPI were used. The aim of these experiments is to demonstrate that $M_1, M_2,$ and $M_1 + M_2$ can actually measure the level of blur in images and to identify slight differences between them. We performed this study using 256×256 pixel binary and grayscale images, as shown in Fig. 3. The images were convolved with different types of point spread functions to simulate various out-of-focus blur types. The absolute difference in distance between the imaging lens at different positions and the reference focused position corresponds to the degree of blurring. Thus, we can simulate the entire image focusing process, from defocus to focus and back to defocus, by changing the position of the imaging lens.

Through theoretical analysis, it is observed that M_1 represents the image variance in the horizontal direction, M_2 represents the variance in the vertical direction, and $M_1 + M_2$ includes variances in both horizontal and vertical directions. Using Eqs. (13) and (14), we obtained the central moments M_1 and M_2 of the simulated images with different defocus levels. To normalize them, both image central moments M_1 and M_2 were divided by the central moments of the original focused image.

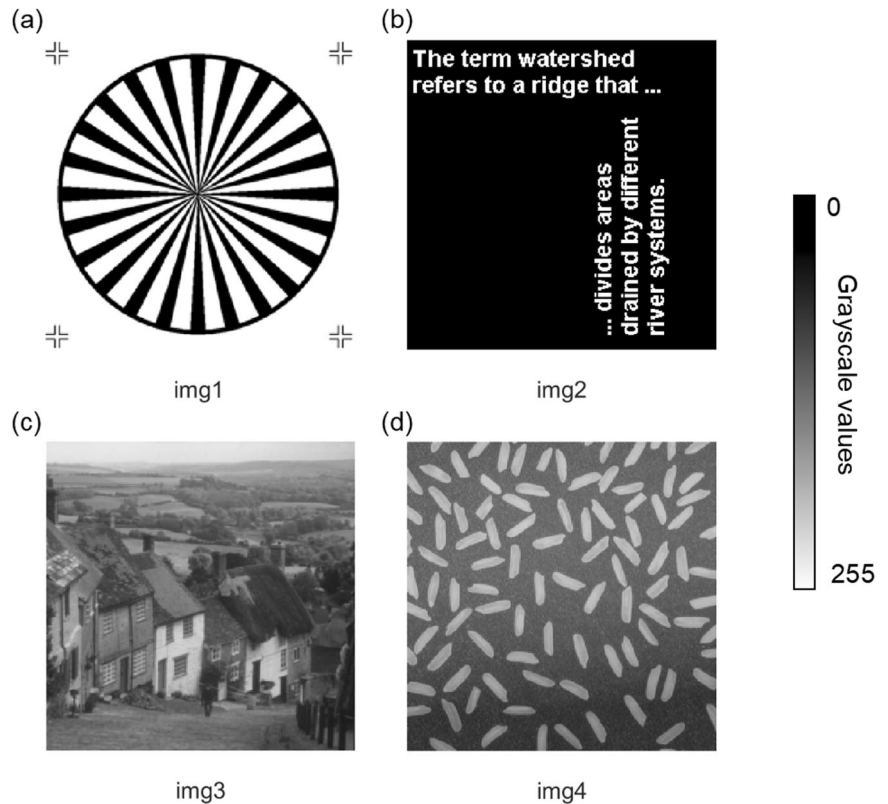
Gaussian point spread function. For the first set of simulation experiments, the commonly used isotropic 2D truncated Gaussian function was chosen as the point spread function²³

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp - \frac{1}{2} \left(\frac{x^2}{\sigma^2} + \frac{y^2}{\sigma^2} \right) \tag{17}$$

The sequence of blurred images with different blur size σ ($\sigma = 0.5-5$, intervals of 0.5) is shown in Fig. 4.

The central moments M_1 and M_2 of the images in Fig. 3 were calculated using Eqs. (13) and (14), respectively. The parameter σ in Fig. 5 corresponds to the position z of the imaging lens set. For the focused image, the imaging lens position is $z = 0$. Ten different values of σ were considered. When $z > 0$, $\sigma = z/2$, simulating back defocus; when $z < 0$, $\sigma = -z/2$, simulating front

Fig. 3 | Images for simulating experiments. Binary images img1 and img2 and gray-level images img3 and img4 used for the simulation experiments.



defocus. The value of σ represents the degree of blurring in the defocused image. The larger the σ value, the greater the blurring, while the smaller the σ value, the lesser the blurring. The vertical coordinates in Fig. 5 show the normalized central moments M_1 and M_2 of the simulated images at different degrees of defocus divided by the central moments of the focused images. Thus, when the imaging lens is in the focus position, the central moments M_1 and M_2 of the focused image are equal to 1.

The central moments M_1 and M_2 vary with variance σ , and the results are shown in Fig. 5. The magnitudes of the central moments of the image progressively decrease as the images approach proper focus ($z = 0$). For isotropic images, like img1, when the PSF has equal variance σ in both the x and y directions, the resulting blurred image's M_1 and M_2 moments are identical. For anisotropic images, such as img2, img3, and img4, despite having a PSF with equal variance σ in x and y directions, the M_1 and M_2 moments of their blurred versions are distinct. However, both approach their minima at the focal point ($z = 0$). In this case, any of the three moments (M_1 , M_2 or $(M_1 + M_2)/2$) can be used as a focus criterion.

For the next set of simulation experiments, considering the possibility of different variations in the point spread function along the x and y axes, we employ an elliptic 2D Gaussian function

$$h(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp - \frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \quad (18)$$

was chosen as the point spread function to simulate anisotropic optical systems. The sequence of blurred images with different variances σ_x and σ_y ($\sigma_x = 0.5-5$, intervals of 0.5; $\sigma_y = 1-10$, intervals of 1) is shown in Fig. 6.

Table 1 provides a detailed description of the relationship between the position z of the imaging lens system and the variance of the Gaussian function. The first row of Table 1 shows the position z of the imaging lens system, while the second row displays the variations in the isotropic Gaussian function variance σ . The third and fourth rows show the variations of the variances σ_x and σ_y , respectively.

We calculated the moments M_1 and M_2 of the image using Eqs. (13) and (14), respectively. The parameter σ in Fig. 7 corresponds to the position z of the imaging lens. When $z > 0$, we set $\sigma_x = z/2$, $\sigma_y = z$, simulating back defocus; when $z < 0$, $\sigma_x = -z/2$, $\sigma_y = -z$, that simulates front defocus. The vertical coordinates in Fig. 7 are the central moments of the blurred images and are normalized by dividing by the central moment of the original focused image. Due to the variance of the elliptical 2D Gaussian function differs in the x and y directions (that is, $\sigma_x \neq \sigma_y$), for isotropic images such as img1, the central moments M_1 and M_2 of the image will also be different.

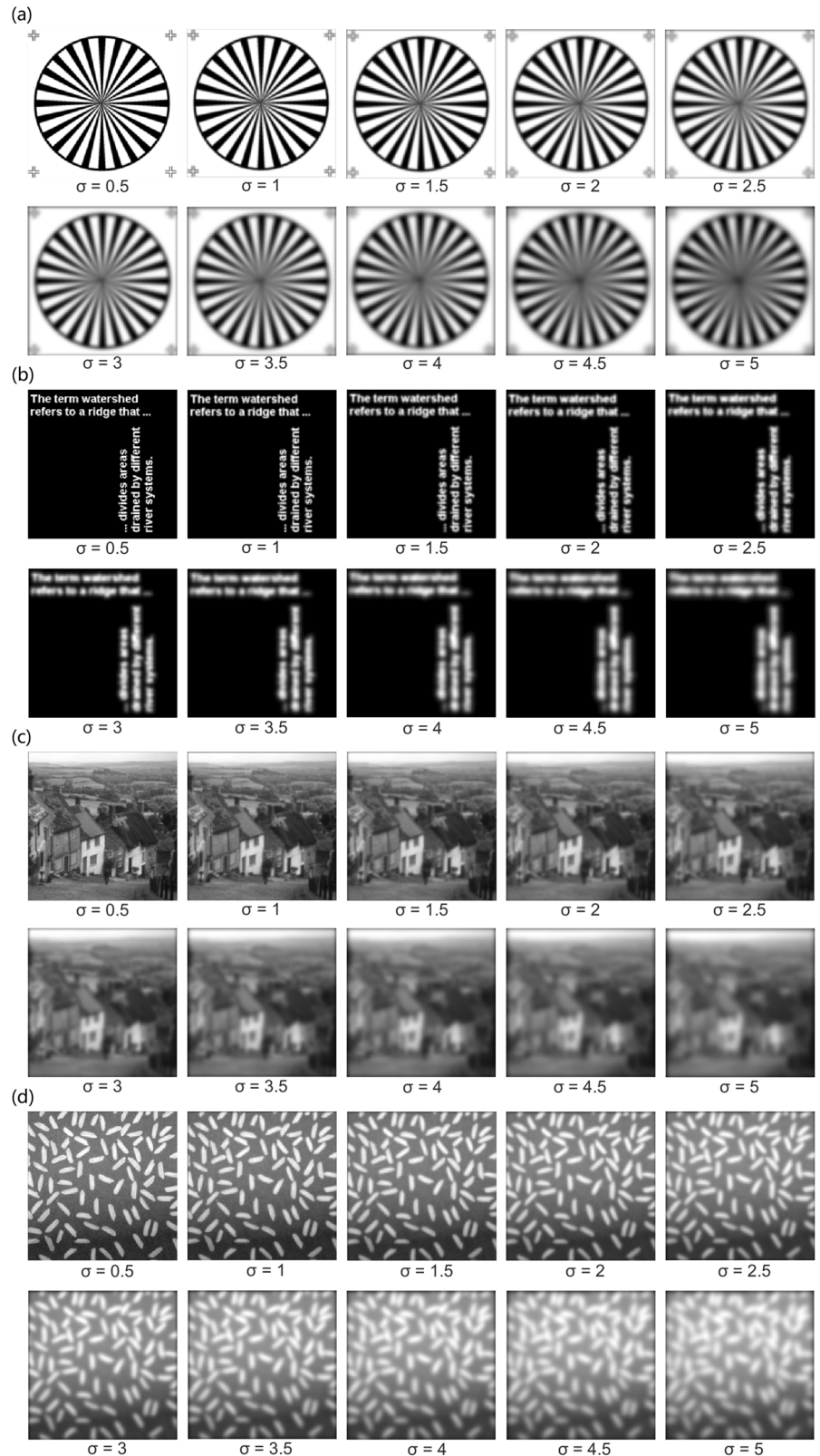
According to the results in Figs. 5 and 7, it can be observed that regardless of whether the point spread function is an isotropic Gaussian function or an anisotropic Gaussian function, the central moments of the image effectively reflect the degree of image blurring. The central moment values serve as a reliable criterion for determining focus in both of these cases.

Performance analysis. In this section, we delve into the impact of noise on the accuracy of autofocusing and meticulously evaluate the robustness and reliability of the proposed autofocusing criterion when confronted with typical noise disturbances. To comprehensively understand the potential implications of noise on the focus determination mechanism, we simulate the introduction of prevalent noise types found in imaging systems onto test images: Gaussian noise. This analytical step aims to reveal the performance of the method under complex real-world conditions, ensuring that the proposed approach is not only effective under ideal circumstances but also remains sturdy when faced with practical challenges.

Let $g(x,y)$ represent the noise-free blurry image, and $n(x,y)$ denote the additive noise. The resulting noisy blurry image, $g'(x,y)$, is then given by

$$g'(x,y) = g(x,y) + n(x,y) \quad (19)$$

Fig. 4 | Defocused images of objects with different levels of blurring. a–d A sequence of blurred images of img1, img2, img3, and img4, respectively. As σ increases, the images become more blurred.



The moments M' of the noisy image can be expressed as

$$M' = u_{pq}^{(g')} = u_{pq}^{(g)} + u_{pq}^{(n)} = M + M_n \quad (20)$$

Where M is the moment of the image, and M_n is the moment of the noise.

Suppose the additive noise follows a Gaussian distribution with mean 0 and variance $\text{Var}(N)$, For zero-mean noise, assuming the noise function is

$n(x,y)$, then $E[n(x,y)] = 0$, and thus we have:

$$E[M_n] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E[x^p y^q n(x,y)] dx dy = 0 \quad (21)$$

This means that the contribution of the noise to the expected value of image moments is zero, $E[M'] = E[M]$. Zero-mean noise, due to its property

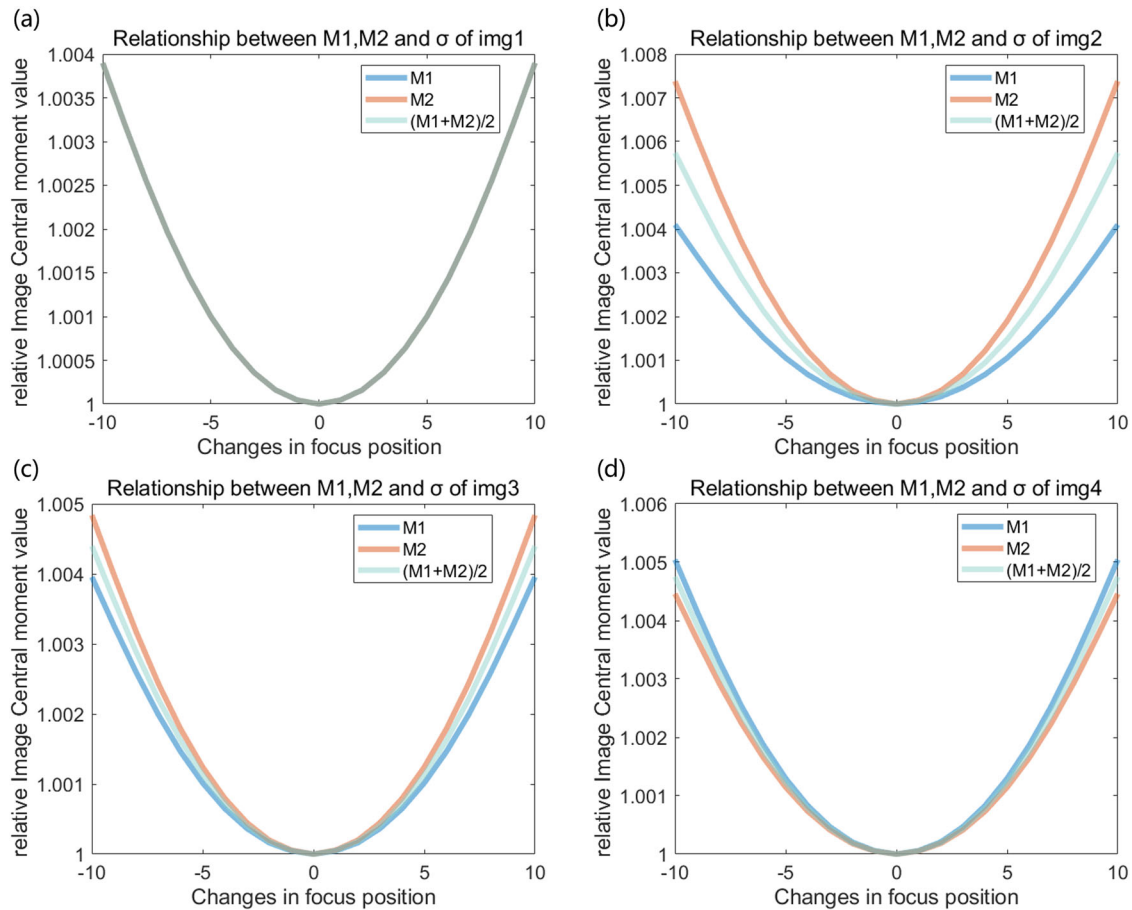


Fig. 5 | The variation of central moments during the process of isotropic blurring of the images. Relationship between central moments (M_1 , M_2 , and $M_1 + M_2$) and variance (σ) of the 2D Gaussian point spread function for img1 (a), img2 (b), img3 (c), and img4 (d).

of having a zero expectation value, is averaged out during the calculation of image moments, making its effect on them mild, thus preserving the stability of image moments. For specific instances of noise appearance, the expected value of the moments may be non-zero but still very low. This is because particular forms of noise might introduce some offset, causing the expected value of the moments to deviate slightly from zero. Nevertheless, such deviations are usually very small and can be negligible for most practical applications. The central moment values of a noisy image primarily depend on the central moment values of the original image. Therefore, during the blurring process, changes in the central moments are still predominantly influenced by the point spread function.

Next, we analyze the characteristics of central moment variations in blurred images under noisy conditions using the four images depicted in Fig. 3. To simulate blur, we convolve the images with isotropic two-dimensional Gaussian point spread functions of varying variances ($\sigma = 0.5-5$, intervals of 0.5), followed by the addition of Gaussian noise with a mean of 0 and a variance of 0.01, as illustrated in Fig. 8. After noise addition, the signal-to-noise ratios (SNRs) for Images 1 through 4 are measured to be 20.02 dB, 10.73 dB, 15.07 dB, 14.01 dB, respectively.

Similarly, we used Eqs. (13) and (14) to calculate the central moments M_1 and M_2 for images with varying degrees of blur, including the presence of noise, as depicted in Fig. 9. It is evident that all three focus measures perfectly pinpoint the focus, demonstrating the robustness of the method despite the introduced noise.

Taking the analysis a step further, we added Gaussian noise at varying variance to the images and computed the changes in the central moments of the images with added noise. Specifically, for img1, we applied Gaussian noise with variance of 0.01, 0.02, 0.05, and 0.1. The outcomes of this process are illustrated in Fig. 10 below. Images subjected to these noise variance

exhibited respective signal-to-noise ratios (SNRs) of 20.02 dB, 17.01 dB, 13.18 dB, and 10.23 dB. Likewise, the effects of different noise variance on the central moments of blurred images are showcased in Fig. 11. From these results, it becomes evident that even at an SNR as low as 10 dB, using central moments as a metric for determining focus remains robust and effective.

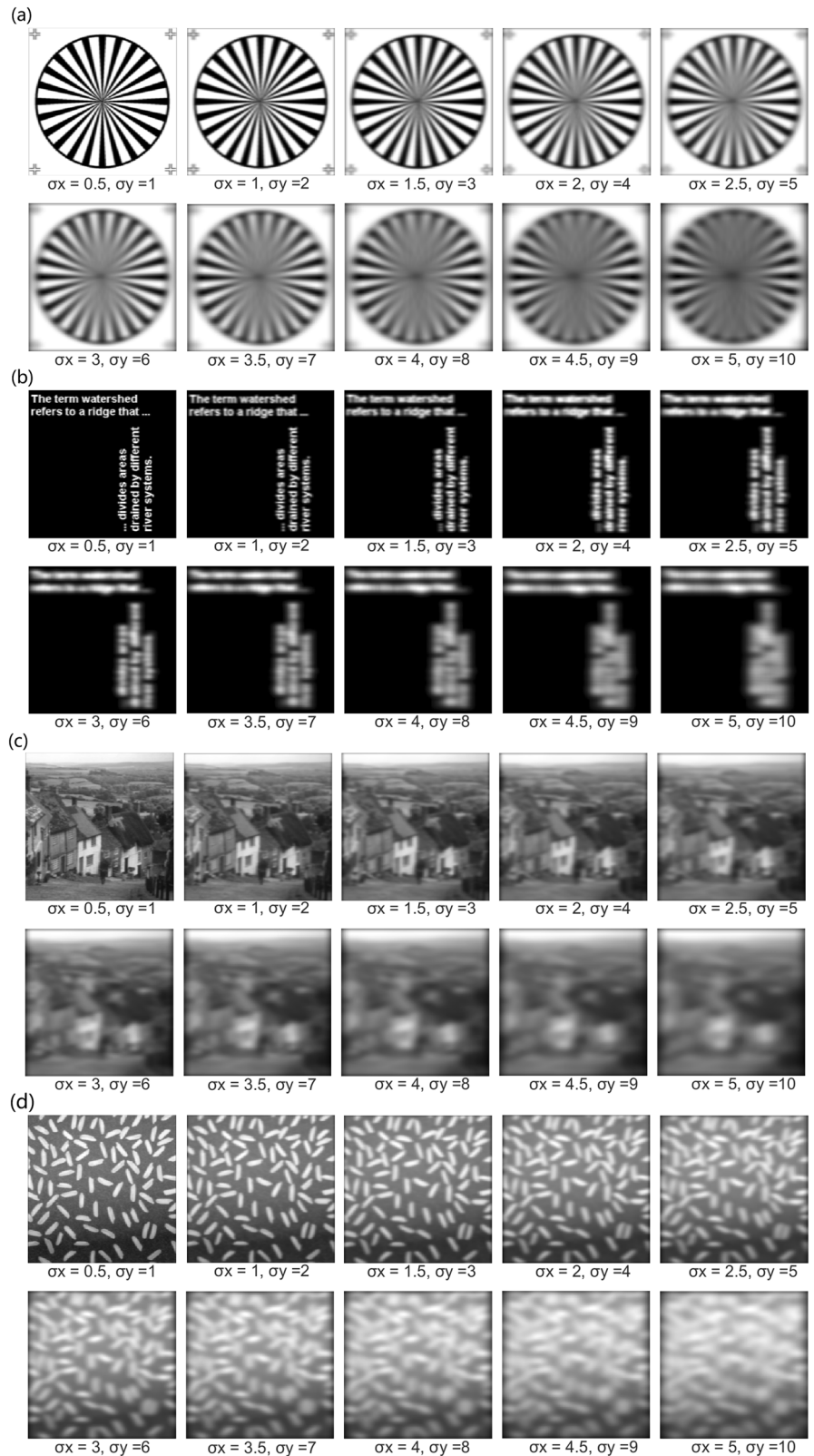
Our simulation conclusions reveal that under conditions of varying noise variance, a proportional relationship exists between the degree of image blur and its central moments. Therefore, it is possible to perform autofocus by searching for the minimum value of one of the three examined focus measures, thereby affirming the feasibility and robustness of utilizing central moments as a metric for assessing and optimizing image focus in a wide range of scenarios and challenges.

Generally, it is recommended to use preferably $M_1 + M_2$, because this criterion is rotation invariant and isotropic. For our constructed optical verification system, the point spread function is symmetric along the horizontal and vertical directions on the optical axis. Therefore, the central moment M_1 can be taken as the standard parameter for obtaining the position of the focal point. In this case, three modulation modes (B_1 , B_2 , and B_3) are used to obtain the values of geometric moments (m_{00} , m_{10} , and m_{20}), and then the central moment M_1 is calculated. Leaving out M_2 speeds-up the process without sacrificing the accuracy of the focus point detection.

Real experiments results

Experimental hardware design. To verify the proposed autofocusing technique in practice, we designed an experimental hardware system (see Fig. 12). A light source is used to illuminate the object. The light passes through the lens and reaches a digital micromirror device (DMD).

Fig. 6 | Defocused images of objects with different levels of blurring. a–d A sequence of blurred images of img1, img2, img3, and img4, respectively. σ_x and σ_y are the variance of the 2D Gaussian point spread function in the horizontal and vertical directions, respectively. As σ_x and σ_y increases, the images become more blurred.



DMD displays a 2D modulation pattern generated by a computer. In our case, low-order polynomials are used to modulate the image (see Fig. 1). The patterns are loaded onto the DMD using the spatial dithering method²⁴. The modulated optical signal is captured by a single-pixel detector. The SPI value, which is in fact an optically calculated image geometric moment, is passed to a computer to calculate central moments and to evaluate the focus measure. This process is iterated over a certain range of the lens positions

and the position providing the optimal focus is determined as the minimum of the criterion function.

A light-emitting diode (LED) (M530L4-C1, Thorlabs) with a maximum power of 200 mW, wavelength of 530 nm, and bandwidth of 35 nm was used as the light source. A combined lens set was used as the imaging lens set and was placed on a motorized displacement stage (Zolix TSA100-B) with a total travel distance of 100 mm. For the same traveling distance,

four experiments with different focusing speed were conducted to verify the effectiveness of the method proposed in this paper by controlling the moving speed of the motorized displacement stage. The moving speed was 10, 15, 20, and 30 mm s⁻¹, respectively. The total distance traveled by the imaging lens was 60 mm. Focusing was performed by gradually moving closer to and then further away from the focal point. The light modulator (DMD) was a Texas Instruments Discovery V7000 with 1024 × 768 micromirrors (the effective usage area of DMD is the middle 768 × 768 micromirrors). The modulated light from the DMD was reflected using a silver-plated mirror, the reflected light was collected using a convex lens with a focal length of 31.5 mm, and the modulated light signal was captured by a single-pixel detector (PDA100A2 Thorlabs). The detector output signals were sampled and acquired using a high-speed digitizer (ADLINK PCI-9816H).

Experimental results. Autofocusing experiments were conducted using the system described above. Two objects were manufactured of a metal plate with hollow patterns, as shown in Fig. 13. During the experiment, the imaging lens moved horizontally at a constant speed to change the image of the object on the DMD. The modulation time of the three

Table 1 | Relationship between focus position x and Gaussian function variance

position z	0	1	2	3	4	5	6	7	8	9	10
σ	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
σ_x	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
σ_y	0	1	2	3	4	5	6	7	8	9	10

patterns modulated by the DMD was set to 1000 μ s, and the sampling rate of the acquisition card was 100 kHz. Equation (13) was used to calculate the central moment of the object at different imaging lens positions.

When the imaging lens moves to different positions within the range, the change of the central moment of the object “USTC” is shown in Fig. 14. Figure 14a–d shows the results for focusing speeds of 10, 15, 20, and 30 mm s⁻¹, respectively. The central moment was calculated using the detections under pattern modulation of modes B₁, B₂, and B₃ at different focusing positions (the patterns were pre-computed and stored off-line). The blue curves are the calculated data for the central moment. The horizontal axis represents distance in millimeters, and the central moment is plotted along the vertical axis. To eliminate the impact of noise, fluctuations in the detector signals, and other factors, the calculated central moment data curves were fitted using the curve fitting method. This process ensured a smooth data curve, shown in orange. As shown in Fig. 14, the four experiments initially approached the focal point and then moved away. As the lens moved, the imaging pattern transitioned from blurry to clear and back to blurry. The orange smooth-fitting curves in Fig. 14a–d show that the central moment of the objects in all four experiments reached its minimum at the focal point ($z = 30$ mm). Because of the time interval between the signal acquisition by the acquisition card and the start of lens movement, the central moment curve initially exhibits a period of flatness. From Fig. 14a–d, it can be observed that the central moment is minimum at the same position ($z \approx 29.4$ mm) for each of the four focusing speeds, indicating that the corresponding optical system has the same focal point.

The image of the object on the DMD during the 10 mm s⁻¹ speed experiment was for illustration captured using a Xiaomi 13 mobile phone, and the results are shown in the images indicated by the black dashed arrows in Fig. 14a. The seven images in Fig. 14(i–vii) are images of the objects

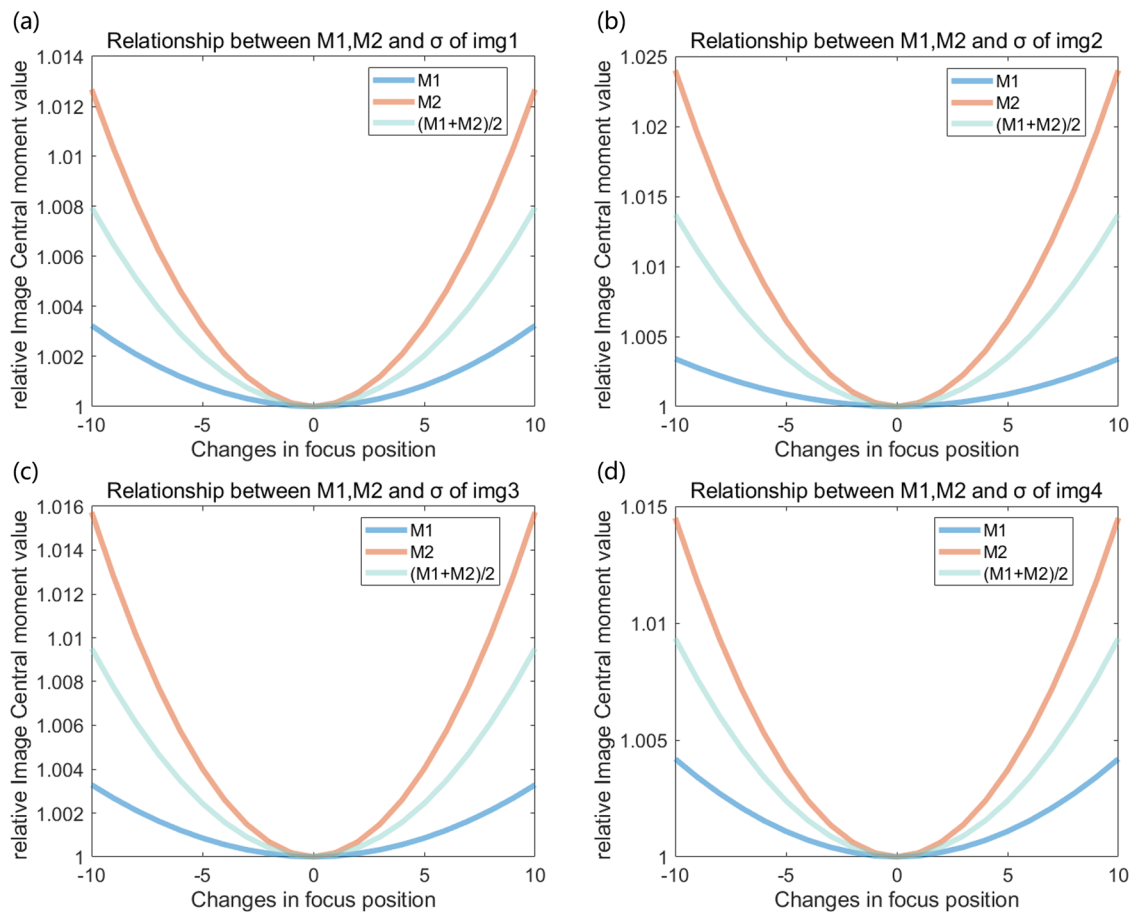
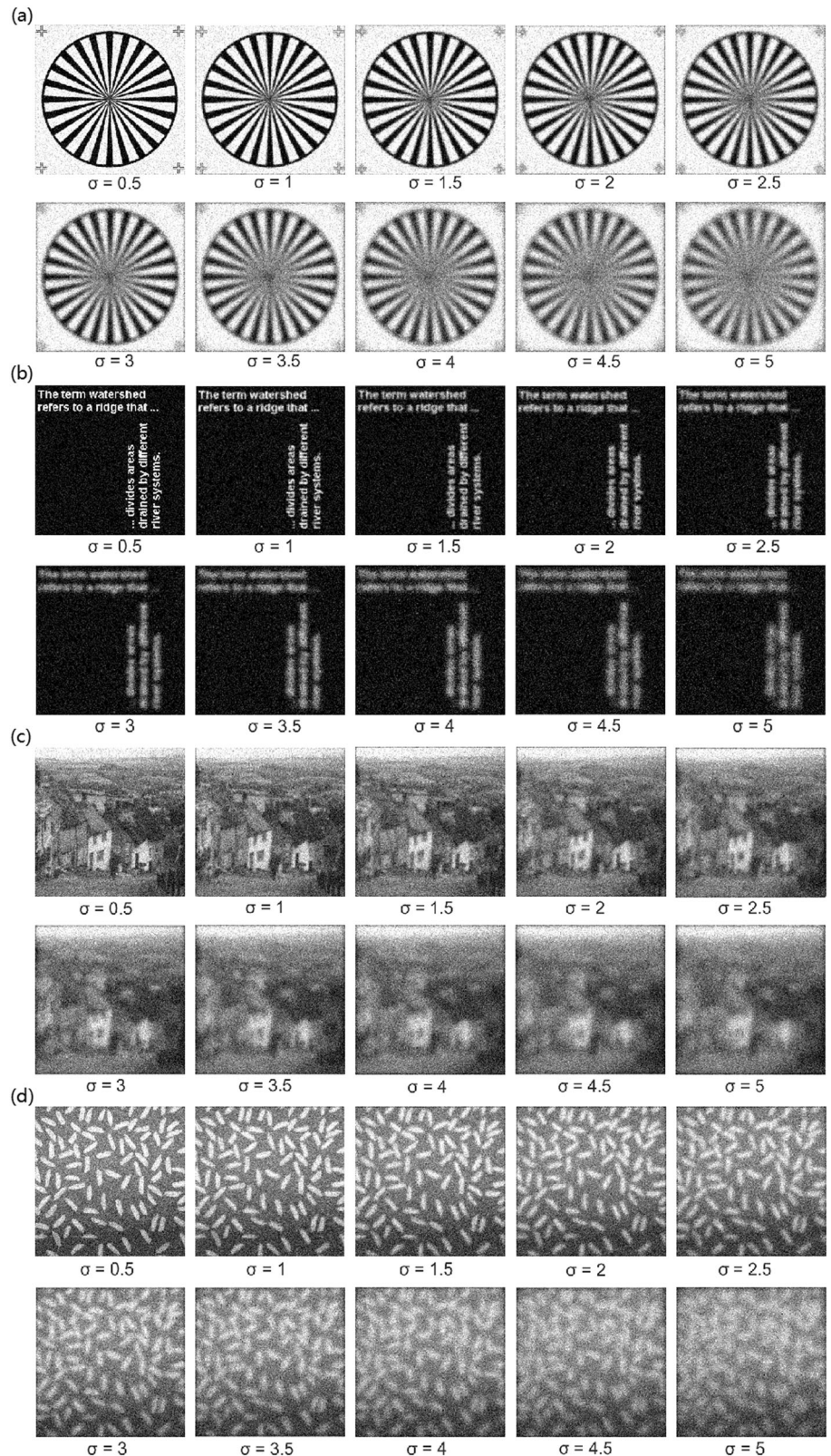


Fig. 7 | The variation of central moments during the process of anisotropic blurring of the images. Relationship between central moments (M_1 , M_2 and $M_1 + M_2$) and variance (σ_x and σ_y) of the elliptic 2D Gaussian point spread function for img1 (a), img2 (b), img3 (c), and img4 (d).

Fig. 8 | Noisy defocused images of objects with different levels of blurring. a–d A sequence of Gaussian noise blurred images of img1, img2, img3, and img4, respectively. As σ increases, the images become more blurred.



captured when the imaging lens moved approximately to positions 0, 10, 20, 30, 40, 50, and 60 mm, respectively. In the supplement (see “Supplementary Movie 1”), some distortion can be observed in the images due to the shooting angle of the mobile phone. The angle between the mobile phone lens and the DMD target surface is $\sim 45^\circ$.

In addition to video, taking by the mobile phone, we performed another test that enables visual assessment. In addition to polynomial

modulation, which was used for moment calculation, we applied modulation by Hadamard basis functions.

Then the SPI values can be used for image reconstruction. This process is known as the Hadamard single-pixel imaging and was originally employed in compressed sensing. We employed the differential Hadamard single-pixel imaging technique and utilized 32,768 Hadamard patterns to reconstruct the image on a 128×128 pixel grid (note that for a complete

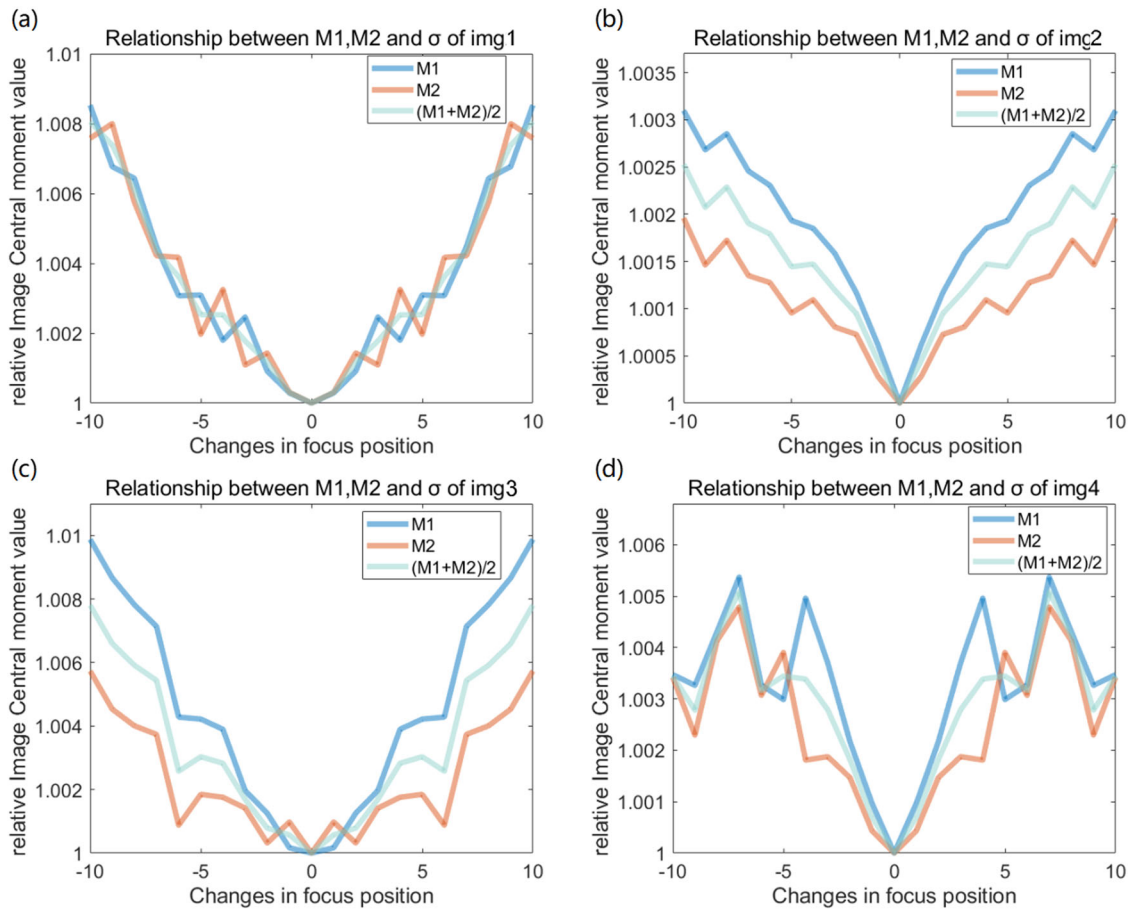
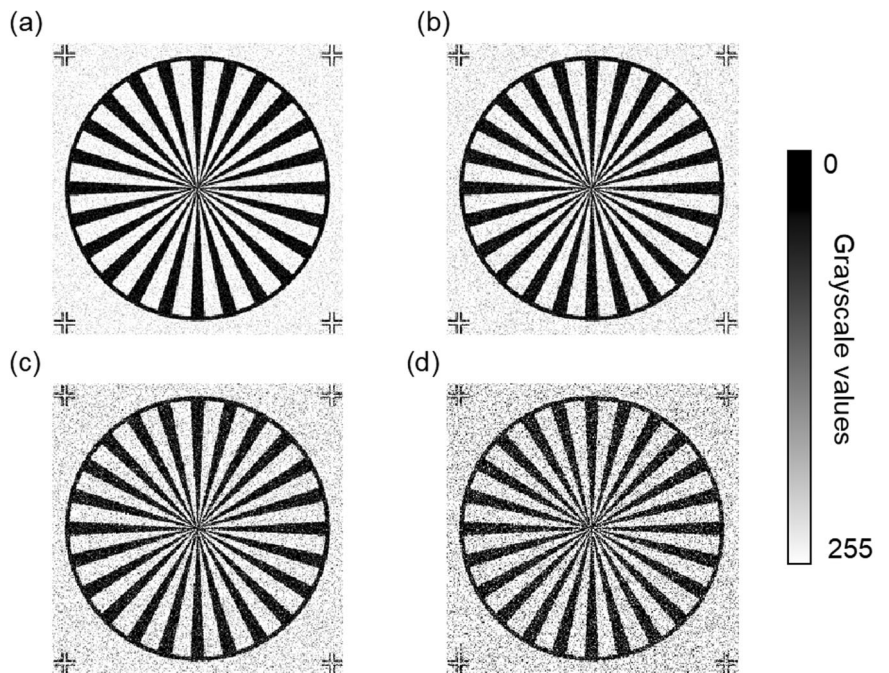


Fig. 9 | The variation of central moments during the isotropic blurring process of the noisy images. Relationship between central moments (M1, M2, and M1 + M2) and variance (σ) of the 2D Gaussian point spread function for img1 (a), img2 (b), img3 (c), and img4 (d), under Gaussian noise conditions.

Fig. 10 | Noise images for simulating experiments. Img1 with Gaussian noise of different variance, where the noise variance for a, b, c, and d are 0.01, 0.02, 0.05, and 0.1, respectively.



reconstruction, this technique requires two times more SPI values than the number of reconstructed pixels)²⁵⁻²⁹.

The reconstructions are shown in Fig. 14e–g. Figure 14f shows the image at the focused position, whereas Fig. 14e, g show the imaging results at 10 mm to the left of the approximate focus position (green

dots on the curves in Fig. 14b–d) and 10 mm to the right (yellow dots on the curves in Fig. 14b–d), respectively. The object image appears the sharpest at the calculated focal point, which supports the results measured by the mobile and illustrates the effectiveness of the proposed method.

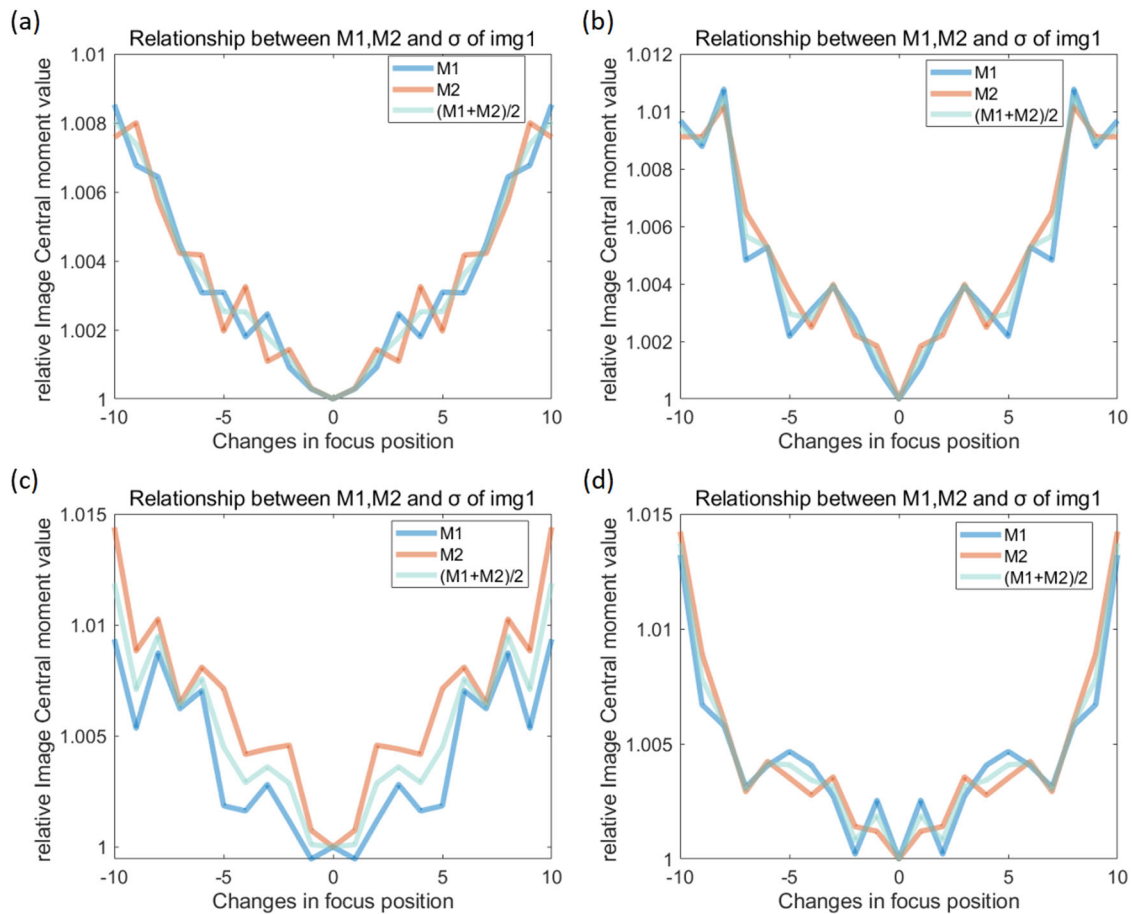


Fig. 11 | The variation of central moments during the isotropic blurring process of noisy image1. The relationship between the central moments of img1 with Gaussian noise of different variance and the variance (σ) of a 2D Gaussian point spread function, where the noise variance for **a**, **b**, **c**, and **d** are 0.01, 0.02, 0.05, and 0.1, respectively.

Fig. 12 | The experimental hardware for testing the proposed autofocus technique. The light illuminates the object, passes through the lens, is modulated on the Digital Micromirror Device (DMD) and continues to the single-pixel detector. The experiment is monitored by a mobile phone.

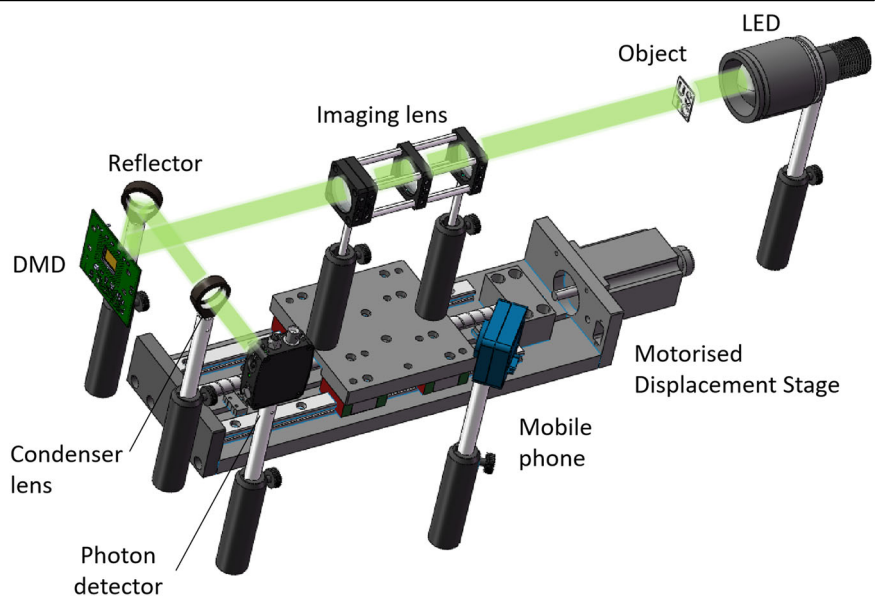


Fig. 13 | Objects used for the experiments. **a** The letters “USTC” and **b** part of the USAF1951 resolution plate.

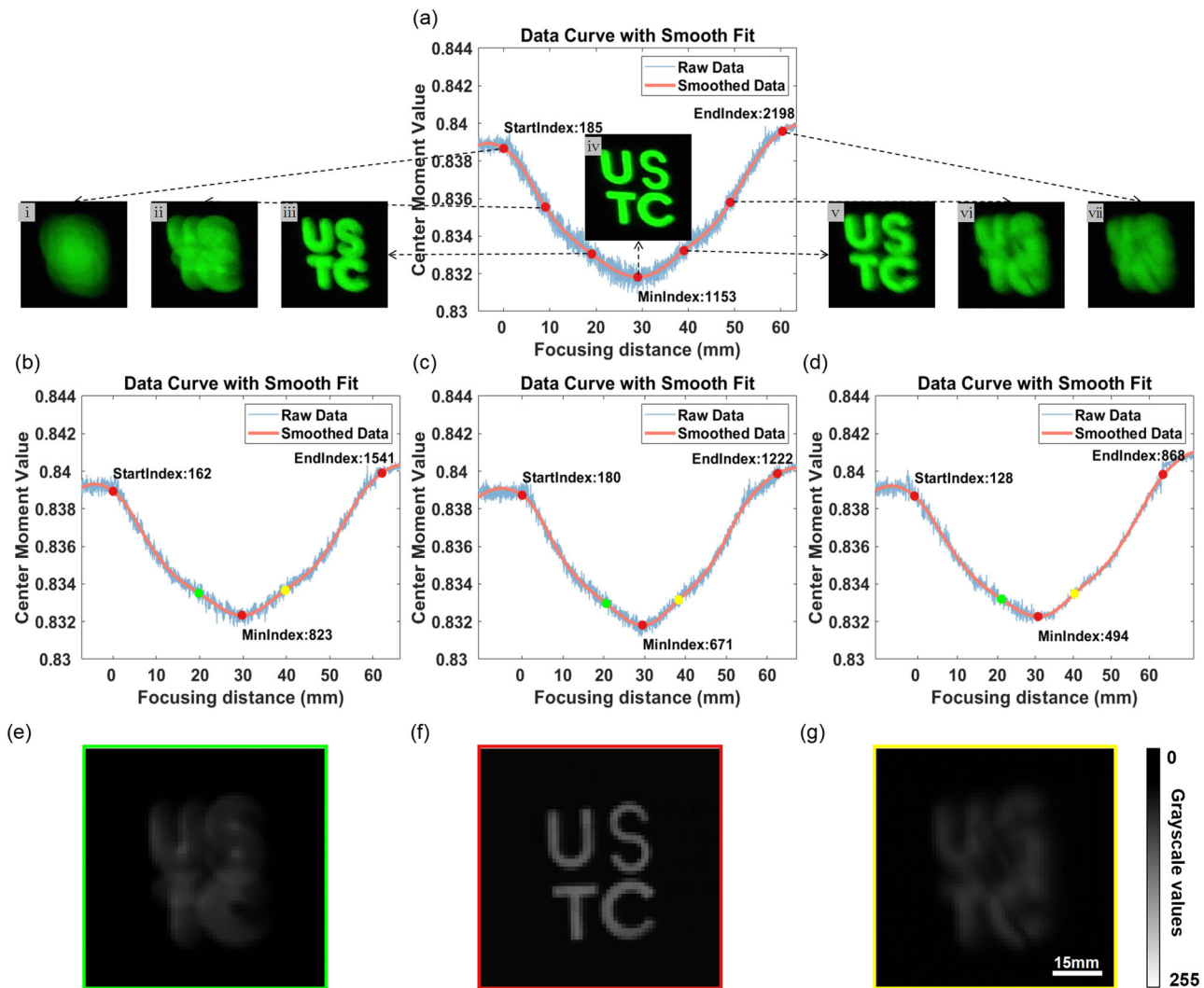
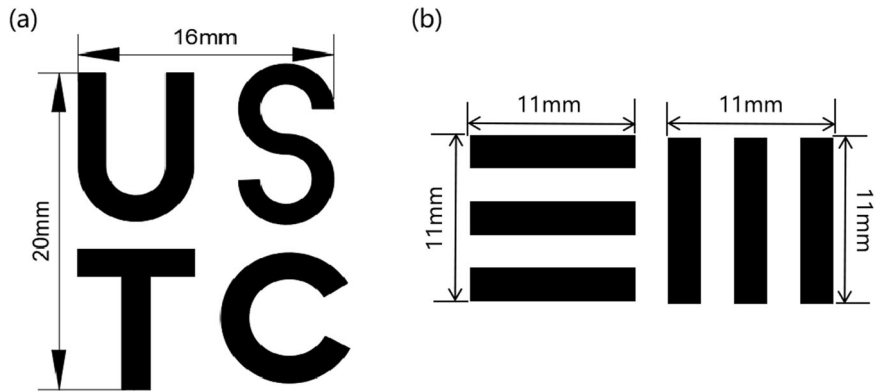


Fig. 14 | Experimental results for the USTC object. **a–d** The results under moving speeds of 10, 15, 20, and 30 mm s⁻¹, respectively. **i–vii** are the photographed images using a Xiaomi 13 mobile phone when the imaging lens moved at the positions of 0, 10, 20, 30, 40, 50, 60 mm, respectively. **e–g** The Hadamard single-pixel reconstructions of the object 10 mm left of the focal point, at the focal point of the central moment minimum, and 10 mm to the right of the focal point, respectively (see Supplementary Movie 1 for more details).

Figure 15 shows the results of an identical experiment conducted on the resolution plate shown in Fig. 13b. The blue curves in Fig. 15a–d represent the original central moment data of the object obtained by the imaging lens at various positions using the detections under pattern modulation of modes B₁, B₂, and B₃ at different focusing positions. The orange

curves show the results of fitting the original blue data curves to achieve a smoothed representation. The black dashed arrows pointing to **i–vii** in Fig. 15a are the images of the object captured with a Xiaomi 13 mobile phone when the imaging lens moved about 0, 10, 20, 30, 40, 50, and 60 mm, respectively. In the supplemental document (see “Supplementary Movie 2”),

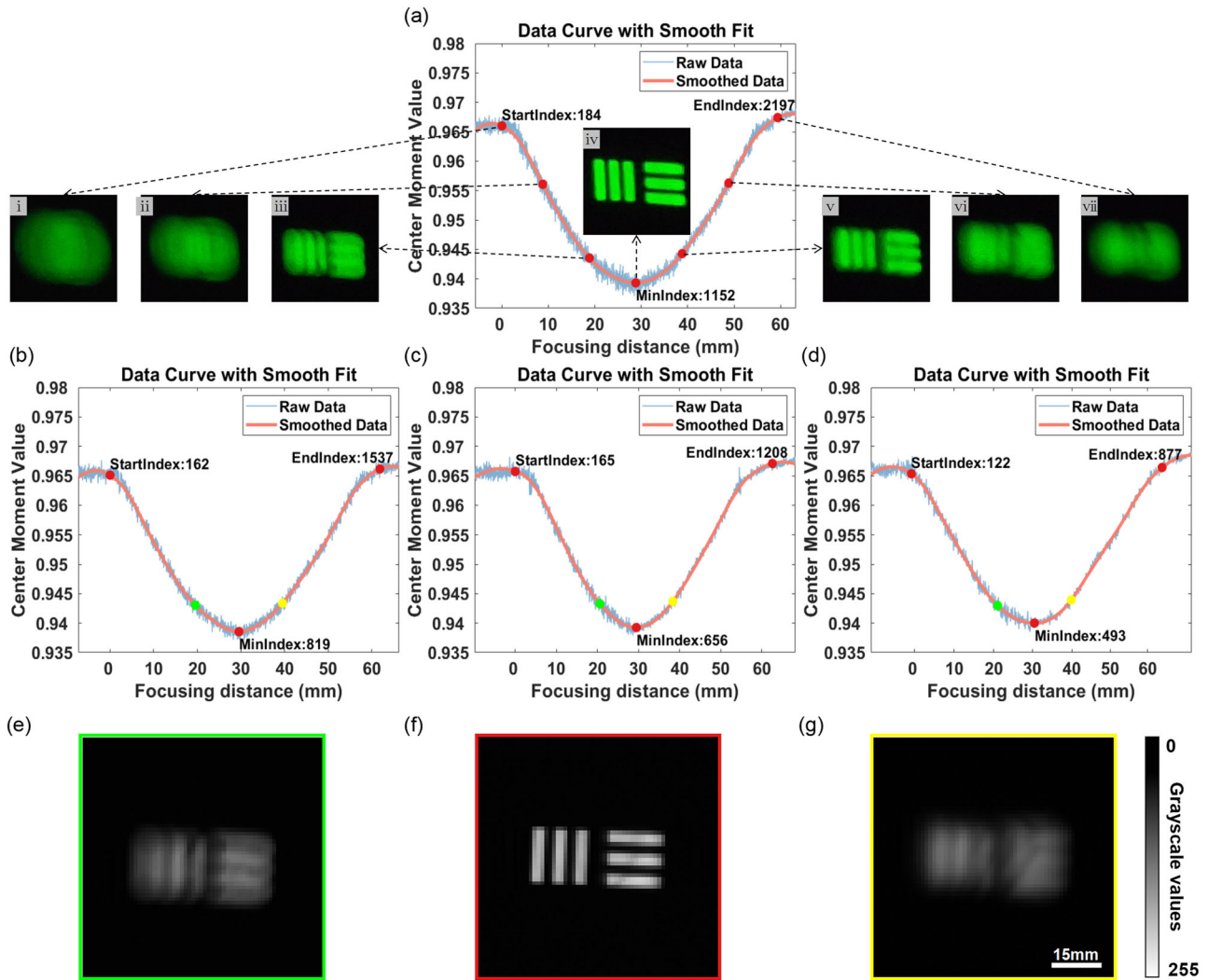


Fig. 15 | Experimental results for the USAF1951 resolution plate. a–d The results under moving speeds of 10, 15, 20, and 30 mm s⁻¹, respectively. i–vii are the photographed images using a Xiaomi 13 mobile phone when the imaging lens moved at the positions of 0, 10, 20, 30, 40, 50, 60 mm, respectively. e–g The Hadamard single-

pixel reconstructions of the object 10 mm left of the focal point, at the focal point of the central moment minimum, and 10 mm to the right of the focal point, respectively (see Supplementary Movie 2 for more details).

some distortion can be observed in the images due to the shooting angle of the mobile phone.

The central moment values of the object in all four experiments shown in Fig. 15a–d varied from large to small and then to large, and the image reached the focal point when the central moment value was the smallest. Regardless of the focusing speed, the minimum value of the central moment was reached at the same position ($z \approx 29.35$ mm).

The Hadamard single-pixel reconstruction was used again to visualize the results (see Fig. 15e–g).

It is worth noting that the proposed technique actually works in real time. The modulation frequency of the DMD was up to 22.2 kHz, and each calculation required three modulation modes. Therefore, the achievable frequency for obtaining the focus parameters is 7.4 kHz. In other words, calculation of focus parameters, including signal modulation and single-pixel detection, requires negligible time comparing to the time of traveling of the lens from one distance to another.

Conclusion

In this study, a moment-detection fast autofocusing technique was proposed, and the effectiveness of the method was verified through theoretical and experimental analysis. The main advantages of the proposed technique are as follows: First, it assesses image sharpness and focus quality in real time

using a non-imaging hardware-implemented method and achieves accurate focusing results. Unlike local differential focus criteria, the moment-based measure is integral and hence robust to noise. Even in the presence of noise and with a low signal-to-noise ratio, the method still performs well. On the other hand, there exist a limitation of our method which is implied by the global nature of the moments. To be precise, our focus measure requires objects on a black background along the image borders. If this is not the case, we face so-called boundary effect, which means that the moment values of the blurred image are influenced by the objects laying outside the visual field. Fortunately, if the blur is small comparing to the image size, the boundary effect does not violate the performance of the method significantly. The single-pixel moment-detection fast autofocusing technology reduces the cost and complexity. Comparing to active techniques, our method does not require any distance measurement using a rangefinder or a similar device and our method does not need any timing circuits. In comparison with traditional image processing-based autofocusing techniques, where the entire image is captured in each focus position and the focus measure is calculated numerically, our technology uses one light detection unit to measure the light intensity. Both approaches of course require scanning along the optical axis, so they need to move the lens. This simplified structure renders the technology more practically and conveniently for various applications. Furthermore, our technique utilizes a single-pixel

detector, which performs well even under low-light conditions. It has advantages in acquiring optical signals within specific wavelength bands, such as the infrared and terahertz ranges. Moreover, this single-pixel moment detection fast autofocusing method requires only three single-pixel detection values at each focus position, which results in less computational data and faster focusing. In summary, the proposed single-pixel moment-detection fast autofocusing technique provides an innovative solution for achieving fast and accurate autofocusing. This technique has significant advantages in the field of autofocusing, including high speed, accurate focusing, low cost, and low complexity, making it ideal for application in various imaging and projection fields.

Data availability

The data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Code availability

The programming code used to compute all the results in this paper was implemented using Matlab Version R2022b. The code can be obtained by contacting the author upon reasonable request.

Received: 16 April 2024; Accepted: 30 September 2024;

Published online: 09 October 2024

References

1. Kong-feng, Z. & Wei, J. New kind of clarity-evaluation-function of image. *Infrared Laser Eng.* **34**, 464–468 (2005).
2. Subbarao, M. & Tyan, J.-K. Selecting the optimal focus measure for autofocusing and depth-from-focus. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 864–870 (1998).
3. Kwon, O.-J., Choi, S., Jang, D. & Pang, H.-S. All-in-focus imaging using average filter-based relative focus measure. *Digit. Signal Process.* **60**, 200–210 (2017).
4. Li, H., Li, L. & Zhang, J. Multi-focus image fusion based on sparse feature matrix decomposition and morphological filtering. *Opt. Commun.* **342**, 1–11 (2015).
5. Zan, G. An auto-adaptive algorithm to auto-focusing. *Acta Opt. Sin.* **26**, 1474–1478 (2006).
6. Strobl, K. H. & Lingenauber, M. Stepwise calibration of focused plenoptic cameras. *Comput. Vis. Image Underst.* **145**, 140–147 (2016).
7. Kautsky, J., Flusser, J., Zitova, B. & Simberova, S. A new wavelet-based measure of image focus. *Pattern Recognit. Lett.* **23**, 1785–1794 (2002).
8. Liu, Y. et al. Multi-focus image fusion with a deep convolutional neural network. *Inf. Fusion* **36**, 191–207 (2017).
9. Rizvi, S., Cao, J. & Hao, Q. Deep learning based projector defocus compensation in single-pixel imaging. *Opt. Express* **28**, 25134–25148 (2020).
10. Qin, X. X., Zou, H. X., Yu, W. & Wang, P. Superpixel-oriented classification of PolSAR images using complex-valued convolutional neural network driven by hybrid data. *IEEE Trans. Geosci. Remote Sens.* **59**, 10094–10111 (2021).
11. Jinxing, Li. et al. DRPL: deep regression pair learning for multi-focus image fusion. *IEEE Trans. Image Process.* **29**, 4816–4831 (2020).
12. Ma, B. Y. et al. SESF-Fuse: an unsupervised deep model for multi-focus image fusion. *Neural Comput. Appl.* **33**, 5793–5804 (2021).
13. Edgar, M. P., Gibson, G. M. & Padgett, M. J. Principles and prospects for single-pixel imaging. *Nat. Photonics* **13**, 13–20 (2018).
14. Lu, T. A. et al. Comprehensive comparison of single-pixel imaging methods. *Opt. Lasers Eng.* **134**, 106301 (2020).
15. Gibson, G. M. et al. Single-pixel imaging 12 years on: a review. *Opt. Express* **28**, 28190–28208 (2020).
16. Phillips, D. B. et al. Adaptive foveated single-pixel imaging with dynamic supersampling. *Sci. Adv.* **3**, e1601782 (2017).
17. Qi, S. T. et al. Image-free active autofocusing with dual modulation and its application to Fourier single-pixel imaging. *Opt. Lett.* **48**, 1970–1973 (2023).
18. Deng, Z., Qi, S. T., Zhang, Z. & Zhong, J. Autofocus Fourier single-pixel microscopy. *Opt. Lett.* **48**, 6076–6079 (2023).
19. Zhang, Y., Zhang, Y. & Wen, C. A new focus measure method using moments. *Image Vis. Comput.* **18**, 959–965 (2000).
20. Flusser, J., Suk, T. & Zitova, B. *2D and 3D Image Analysis by Moments* (Wiley, 2016).
21. Flusser, J. & Suk, T. Degraded image analysis: an invariant approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 590–603 (1998).
22. Goodman, J. W. *Fourier Optics*, 3rd edn (Roberts and Company Publishers, 2005).
23. Kostkova, J. et al. Handling Gaussian blur without deconvolution. *Pattern Recognit.* **103**, 107264 (2020).
24. Zha, L. et al. Single-pixel tracking of fast-moving object using geometric moment detection. *Opt. Express* **29**, 30327–30336 (2021).
25. Zhang, Z. B. et al. Hadamard single-pixel imaging versus Fourier single-pixel imaging. *Opt. Express* **25**, 19619–19639 (2017).
26. Yu, W.-K. Super sub-Nyquist single-pixel imaging by means of cake-cutting Hadamard basis sort. *Sensors* **19**, 4122 (2019).
27. Sun, S. et al. DCT single-pixel detecting for wavefront measurement. *Opt. Laser Technol.* **163**, 109326 (2023).
28. Wang, Z. et al. DQN based single-pixel imaging. *Opt. Express* **29**, 15463–15477 (2021).
29. Ma, M. et al. Direct noise-resistant edge detection with edge-sensitive single-pixel imaging modulation. *Intell. Comput.* **2**, 0050 (2023).

Acknowledgements

This study was supported by the Youth Innovation Promotion Association of the Chinese Academy of Sciences (Grant No. 2020438), HFIPS Director's Fund (Grant No. YZJJ202404-CX), and Czech Science Foundation (Grant No. GA21-03921S).

Author contributions

Dongfeng Shi proposed the idea of applying single-pixel moment detection to autofocusing and supervised all research for the paper. Jan Flusser provided theoretical analysis support regarding moments. Huiling Chen designed the method and conducted the experiments and computational analysis. Zijun Guo, Runbo Jiang, and Linbin Zha contributed to the data analysis, interpretation, and evaluation. Huiling Chen wrote the manuscript. Yingjian Wang and Jan Flusser performed the final review of the manuscript. All authors discussed the progress of the research and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44172-024-00288-z>.

Correspondence and requests for materials should be addressed to Dongfeng Shi.

Peer review information *Communications Engineering* thanks Richard Bowman, Yusuke Saita, and the other, anonymous, reviewer for their contribution to the peer review of this work. Primary Handling Editors: Anastasiia Vasylychenkova.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024