optica

Aberration-corrected three-dimensional positioning with a single-shot metalens array: supplement

WENWEI LIU,[1,†] DINA MA,[1,†] ZHANCHENG LI,[1] HUA CHENG,[1,5] DUK-YONG CHOI,[2] JIANGUO TIAN,[1] AND SHUQI CHEN[1,3,4,*]

[1]*The Key Laboratory of Weak Light Nonlinear Photonics, Ministry of Education, School of Physics and TEDA Institute of Applied Physics,Nankai University, Tianjin 300071, China*
[2]*Laser Physics Centre, Research School of Physics, Australian National University, Canberra, ACT 2601, Australia*
[3]*The Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China*
[4]*Collaborative Innovation Center of Light Manipulations and Applications, Shandong Normal University, Jinan 250358, China*
[5]*e-mail: hcheng@nankai.edu.cn*
[*]*Corresponding author: schen@nankai.edu.cn*
[†]*These authors contributed equally to this work.*

SUPPLEMENTARY MATERIAL FOR

Aberration-corrected three-dimensional positioning with single-shot metalenses array

WENWEI LIU[1,†], DINA MA[1,†], ZHANCHENG LI[1], HUA CHENG[1,*], DUK-YONG CHOI[2], JIANGUO TIAN[1], AND SHUQI CHEN[1,3,4]*

[1]*The Key Laboratory of Weak Light Nonlinear Photonics, Ministry of Education, School of Physics and TEDA Institute of Applied Physics, Nankai University, Tianjin 300071, China.*

[2]*Laser Physics Centre, Research School of Physics and Engineering, Australian National University, Canberra, ACT 2601, Australia.*

[3]*The Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China.*

[4]*Collaborative Innovation Center of Light Manipulations and Applications, Shandong Normal University, Jinan 250358, China.*

[†]*These authors contributed equally to this paper.*

email: hcheng@nankai.edu.cn

Corresponding author, email: schen@nankai.edu.cn

1. SAMPLE FABRICATION

The TiO_2 metasurface was fabricated by using the spin-coating technique to deposit a 550-nm-thick electron-beam resist (ZEP 520A from Zeon, Japan) on a cleaned glass substrate, which as baked on a hot plate at 180 °C for 1 min. To prevent charging during electron-beam writing, a thin layer of the E-spacer 300Z (Showa Denko) was coated on the resist. Then, the nanostructures were defined by electron-beam lithography (Raith150) at 30 kV with 20-pA current, followed by development in n-amyl acetate solvent. Conformal TiO_2 layers with a thickness of approximately 70 nm were deposited by an atomic layer deposition system (Picosun) using titanium tetrachloride and H_2O as precursors at a reactor temperature of 130 °C. Subsequently, CHF_3 plasma in an inductively coupled plasma-reactive ion etching (ICP-RIE, Oxford system 100) was performed to blank the TiO_2 layer until the ZEP 520A resist was exposed. Here the etching conditions were 30 sccm of CHF_3 at 50 W bias power/500 W induction power at an operating pressure of 10 mTorr, which resulted in a TiO_2 etching rate of ~20 nm per minute. Finally, O_2 plasma was used to fully remove the remaining resist.

2. CROSS-CORRELATION-BASED GRADIENT DESCENT (CCGD) ALGORITHM

We employed a CCGD algorithm to correct the imaging aberrations, achieve positioning, and obtain the reconstructed images. Four parameters characterizing the properties of imaging were utilized in the algorithm, that is, k_1, k_2, t_x, and t_y, where k_1

and k_2 are the parameters to correct the distortions of the image, and t_x, t_y are translation parameters between different image parts. The algorithm can be summarized in the following steps:

(1) Generating the image matrix to minimize the impact of the gray information in the images, we first use the function "adaptiveThreshold()" from the Open Source Computer Vision (OpenCV) library [1, 2] to binarize the image. This function divides the image into small blocks and performs the thresholding calculation.

(2) Initial distortion correction. The radial transformations of the image can be corrected using Eqs. S1 and S2 by two parameters k_1 and k_2, which will be optimized during the iteration in step (4).

$$\Delta x = x_r \left(k_1 r^2 + k_2 r^4 \right), \tag{S1}$$

$$\Delta y = y_r \left(k_1 r^2 + k_2 r^4 \right), \tag{S2}$$

where (Δx, Δy) is the correction of the distorted image, (x_r, y_r) are the coordinates of a pixel in the image with the original point located at the distortion center, and r is the distance between the pixel and the distortion center. Then, we divide the image into three parts, Part I, II, and III, corresponding to three matrices P_1, P_2, and P_3.

(3) A cross-correlation method to determine the affine transformation. P_1, P_2, and P_3 describe the same image with different imaging information. Taking P_1 and P_3 as an example, the affine transformation between the coordinates of P_1 and P_1' can be described by

$$\begin{pmatrix} x_1^{'} \\ y_1^{'} \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}. \tag{S3}$$

Here $P_1^{'}$ is a new image part that should overlap with P_3. Because $P_1^{'}$ and P_3 contain different imaging information, the optimization method introduced in Step (4) is utilized to optimize the overlapping.

We employ a cross-correlation method to determine the value of t_x and t_y during each iteration, which combines the overall information of an image instead of finding featured points. Consider two images related to $P_b = P_a\left(x - t_x, y - t_y\right)$. The cross-correlation between P_a and P_b in real space is defined by $R_{ccf}^{(r)}\left(x, y\right) = P_a\left(x, y\right) * P_b\left(-x, -y\right)$, where '*' signifies the convolution operator. Accordingly, the cross-correlation function in Fourier space is

$$R_{ccf}^{(f)}\left(k_x, k_y\right) = \mathcal{F}\left(P_a\right)\mathcal{F}^*\left(P_a\right)\exp[2\pi j\left(k_x t_x + k_y t_y\right)], \tag{S4}$$

where $\mathcal{F}$ represents the Fourier transform operator. The inverse Fourier transform of Eq. (S4) can be calculated by $R_{ccf}^{(r)}\left(x, y\right) = G\left(x, y\right) * \delta\left(x - t_x, y - t_y\right)$, where $G(x, y)$ is the inverse Fourier transform of $\mathcal{F}\left(P_a\right)\mathcal{F}^*\left(P_a\right)$. The cross-correlation function in real space can be further simplified as

$$R_{ccf}^{(r)}\left(x, y\right) = G\left(x - t_x, y - t_y\right). \tag{S5}$$

The maximum value of $\left|R_{ccf}^{(r)}\right|$ is located at (t_x, t_y), which characterizes the translation relationship between P_a and P_b.

(4)   Calculate the loss function and optimize. The fitness function can be calculated by the mean distances between different image parts:

$$J_{31} = \frac{\sum\limits_{i=1}^{m_1} \sum\limits_{j=1}^{n_1} \left| P_1'(i,j) - P_3(i,j) \right|^2}{m_1 \times n_1} , \qquad (S6)$$

$$J_{23} = \frac{\sum\limits_{i=1}^{m_2} \sum\limits_{j=1}^{n2} \left| P_2'(i,j) - P_3(i,j) \right|^2}{m_2 \times n_2} . \qquad (S7)$$

By updating and optimizing the parameters in $-\nabla J$, that is, the gradient of the objective function J, the objective function J can be minimized to obtain the aberration-corrected image and the positioning parameters. To improve the convergence efficiency, we also employ the adaptive moment estimation (Adam) optimizer [3] to update the parameters.

(5) Continue the iteration until the fitness functions converge. Generally, convergence can be achieved within 150 iterative cycles in our study.

3. SENSITIVITY OF POSITIONING WITH THE METALENS ARRAY

The proposed principle can realize three-dimensional positioning of the target object by analyzing the two-dimensional information of the aberration-corrected images. Compared with positioning via mechanical platforms, this method relies on the optical intensity of the images, which is only fundamentally limited by the Abbe diffraction limit. As demonstrated in the main text, the object distance and the horizontal displacement can be calculated as

$$S_\perp = f D / (D - d), \qquad (S8)$$

$$\delta = \Delta_1 (D - d) / d, \qquad (S9)$$

where we replaced Δ with Δ_1 compared with that in the main text to distinguish from

the differential operator Δ. Accordingly, the relative sensitivity of positioning can be further calculated by

$$\left|\frac{\Delta S_\perp}{S_\perp}\right| = \frac{d}{D-d}\left|\frac{\Delta D}{D}\right|, \tag{S10}$$

$$\left|\frac{\Delta \delta}{\delta}\right| = \left|\frac{\Delta \Delta_1}{\Delta_1} + \frac{\Delta D}{D-d}\right|. \tag{S11}$$

It can be seen that the resolution of positioning is determined by the accuracy of the measured geometric length in the image plane. By employing improved algorithms, utilizing other metalens designs to directly revise the aberrations, or by adopting a high-accuracy translation stage, the positioning accuracy can be further improved. The relative sensitivity of positioning can also be described by the metalens design and the geometric parameters in the object plane:

$$\left|\frac{\Delta S_\perp}{S_\perp}\right| = \frac{\dfrac{S_\perp}{f} + \dfrac{f}{S_\perp} - 2}{d}\left|\Delta D\right|, \tag{S12}$$

$$\left|\frac{\Delta \delta}{\delta}\right| = \left|\frac{|S_\perp - f|}{f}\left|\frac{\Delta \Delta_1}{\delta} + \frac{\Delta D}{d}\right|, \tag{S13}$$

where $\Delta\Delta_1$ is the measured accuracy of the images in the image plane, and can be estimated by the accuracy of the periodicity of the images ΔD (also measured in the image plane). Thus, Eq. S10 can be further written as

$$\left|\frac{\Delta \delta}{\delta}\right| = \frac{(d+\delta)|S_\perp - f|}{fd\delta}\left|\Delta D\right|. \tag{S14}$$

Because the measured geometric accuracy in the image plane is settled for a metalens array, the vertical and horizontal positioning accuracy both decrease for large object distances, which is a universal result for positioning systems [4].

6

4. NUMERICAL SIMULATIONS

The focusing properties of the metalens were determined using the finite-difference time-domain method with a Gaussian beam incidence for different incident angles. The optical indices of the TiO_2 nanopillars and fused silica substrate were obtained from Palik's handbook [5]. We simulated the electric distribution at a plane at a distance of 2 µm above the metalens and employed far-field extraction to obtain the light distribution at the focal plane. To further reduce the simulation occupancy of resources, we simulated a smaller metalens with a diameter of 20 µm, but with the same numerical aperture NA = 0.50 as that of the fabricated metalens. The size of the smaller metalens is still dozens of times larger than the operating wavelength, which means that the near-field effects are negligible and the diffraction effects dominate the focusing properties (the same as the metalens introduced in our scheme).

5. MEASUREMENT PROCEDURE

The light source in the experiments was a Mercury-Xenon lamp (Thorlabs, SLS402) with a liquid light guide (Thorlabs, LLG5-4Z) as the output. After a 10 nm-bandpass filter centered at 532 nm, the polarization of the incident light was controlled by a linear polarizer and a quarter-wave plate pair to obtain the LCP incidence (see Supplementary Fig. 5 for the measurement setups). The diffuser can reduce the light speckles in the field-of-view (FOV) and can mimic passive lighting from a rough object. An objective (Obj1, Sigmakoki EPL-5, 5×, NA = 0.13, WD = 11.6 mm) was

employed to collect the incident light onto the target object S1. In our setup, we created a drawing of a bee on a thin Cr film as S1 with the laser direct writing method [6], and the target sample did not have specific limitations in our design. The image of S1 was zoomed out by another objective (Obj2, Sigmakoki EPLE-50, 50×, NA = 0.55, WD = 8.2 mm) and then imaged by the metalens array. In our setup, the actual object, which is positioned by the metalens array (S2), is the zoomed-out image behind Obj2. The filter, P1, QW1, diffuser, Obj1, S1, and Obj2 were mounted on an electric translation stage (TS1) with a spatial resolution of 100 nm to control the movement of the target sample. A pair of QW2 and P2 were utilized to acquire the RCP output components of the sample, which were mounted on another translation stage (TS2 with a spatial resolution of 10 μm) with an objective (Obj3, Sigmakoki EPLE-20, 20×, NA = 0.4, WD = 11.1 mm), a tube lens (TL, Thorlabs, ITL200), and a CCD camera (Hamamatsu, C13440-20CU) to form a scanning system and to reduce the background signals.

The optical positioning setup can be adjusted through the following steps: First, the metalens sample S2 was imaged by the scanning system. By adjusting TS1 along the vertical direction (z-axis), the scanning system on TS2 can simultaneously image the target object and S2 in the FOV (see Supplementary Fig. 6 for details of the captured image), which is the initial state to realize positioning. At this step, sample S2 is located at both the working planes of Obj2 and Obj3, and the object distance of S2 is zero. Then, by moving TS2 along the $+z$ direction and moving TS1 along the $-z$ direction, one can image the target object again when satisfying the Gaussian lens

8

formula, $1/S_0 + 1/S_1 = 1/f_{\text{metalens}}$. At this step, the captured image is the image of S2. In realistic implementations, the optical setup can be simplified in the following aspects. In Fig. 4, we demonstrated that P1 and QW1 could be removed, maintaining comparable imaging/positioning performances. The filter could be integrated on S2, such as coating a bandpass film. The lamp, diffuser, and S1 could be replaced by an object with diffuse reflection. Objective Obj1 was used to collect the incident light and to increase the signal-to-noise ratio, which is also not fundamentally necessary. On the other hand, if one does not need to directly measure the object distance, Obj2 and TS1 can also be removed for larger metalens with large focal length and long object distance.

SUPPLEMENTARY FIGURES
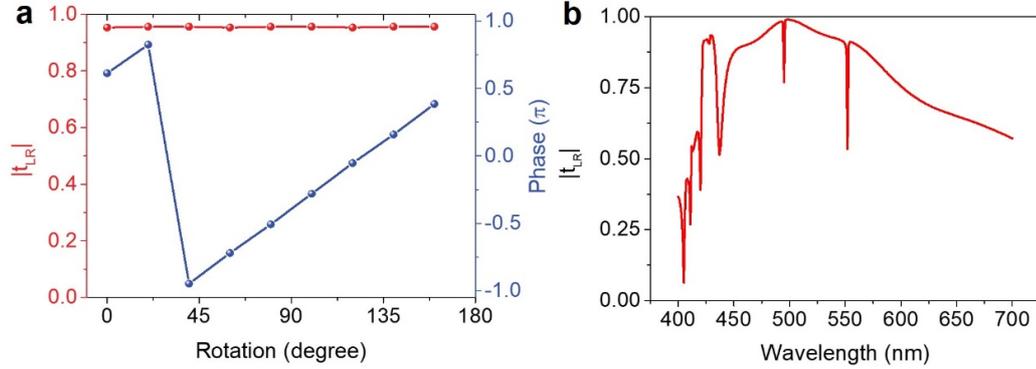
Fig. S1. Simulated transmission for the designed nanopillars. (a) Simulated transmission coefficient $|t_{LR}|$ of the RCP light for LCP incidence, and the Pancharatnam-Berry phase as a function of the orientation angle θ of nanopillars. (b) The simulated transmission coefficient for different incident wavelengths.
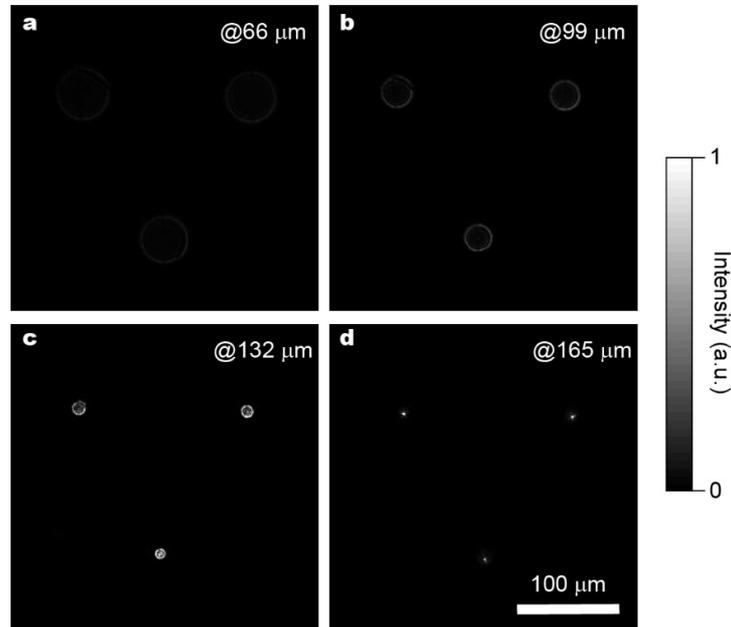

Fig. S2. Measured intensity distribution of focusing for different *x-y* cut-planes captured by the scanning system. The cut-planes are (a-d) from 66 μm to 165 μm behind the metalens array.
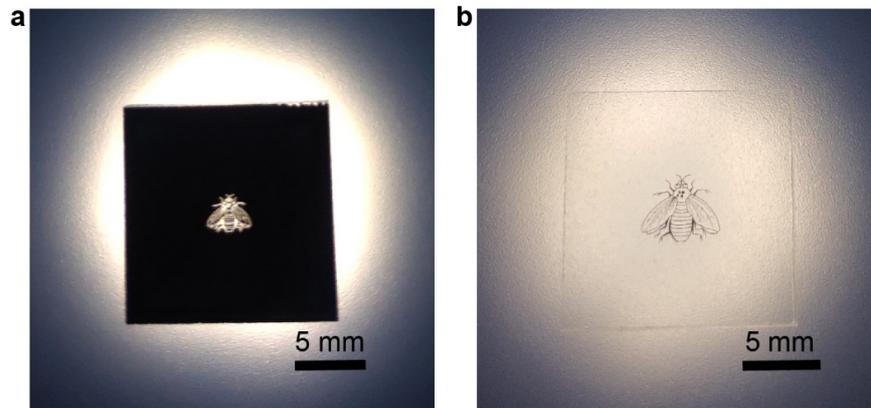
Fig. S3. Optical images of the target object. The object was created on a fused silica substrate with laser direct writing. The opaque areas are covered by a film of Cr, forming (a) a negative pattern with a total size of 3 mm and (b) a positive pattern with a total size of 5 mm.
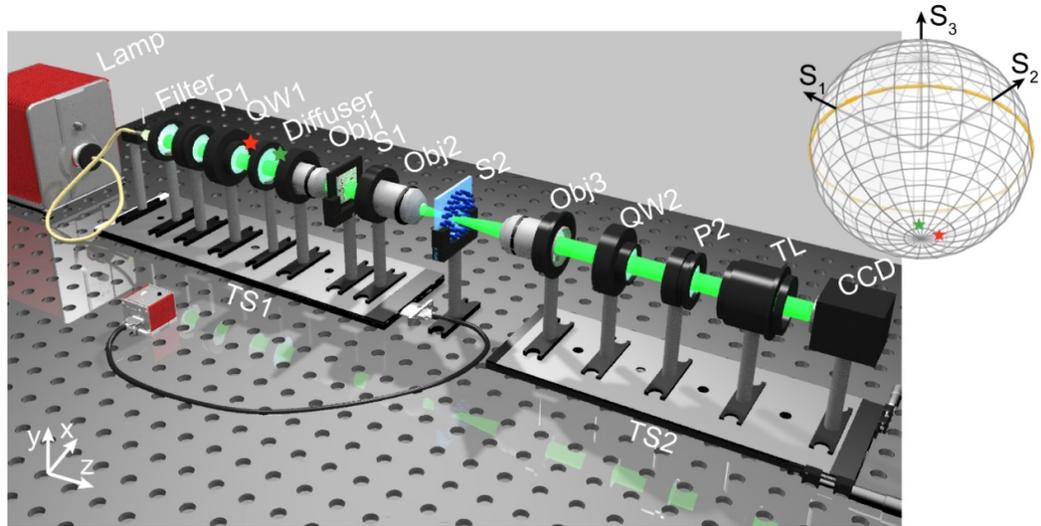
Fig. S4. Measurement setup to realize imaging/positioning and to measure the accurate distance to the object. Note the configuration before the metalens can be simplified in practical implementations because the real object distance does not need to be measured. The incident polarization is controlled by a linear polarizer (P1) and a quarter-wave plate (QW1). A diffuser was used to reduce the lamp speckles. The objective (Obj1) was employed to focus the beam onto the target sample (S1), and Obj2 was used to zoom out the image of S1 in front of the metalens sample (S2). The vertical position of the object plane was controlled by a translation stage (TS1), and the image plane was detected by a scanning system mounted on a translation stage (TS2). Inset: The polarization state of the incident beam on the Poincaré sphere before (red star) and behind (green star) the diffuser, demonstrating that the diffuser does not introduce significant changes of the polarization states of the incident beams.
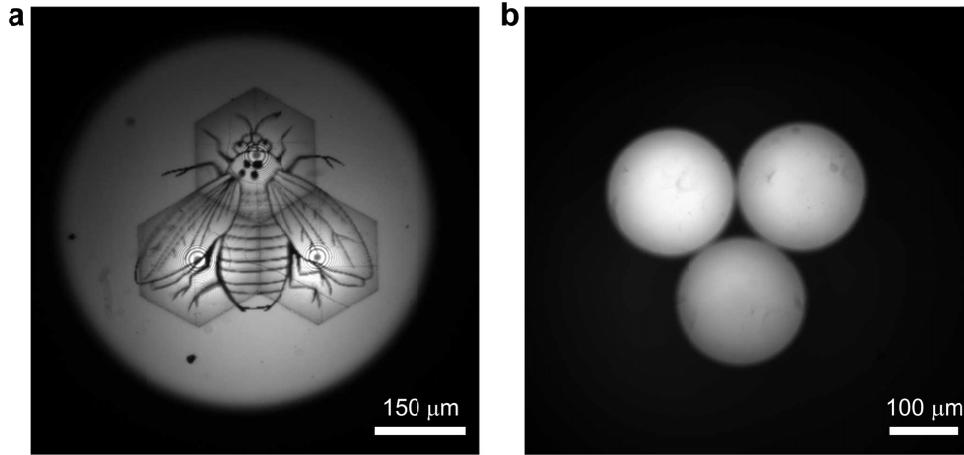
Fig. S5. Captured images for calibration of the measurement. (a) The initial state of positioning when the metalens array is simultaneously located at the working planes of Obj2 and Obj3, showing both of the images of S1 and S2. In this case the object plane of S2 is zero. The accurate object distance is obtained by controlling TS1. (b) FOV of the system captured by the CCD camera when S1 is transparent.
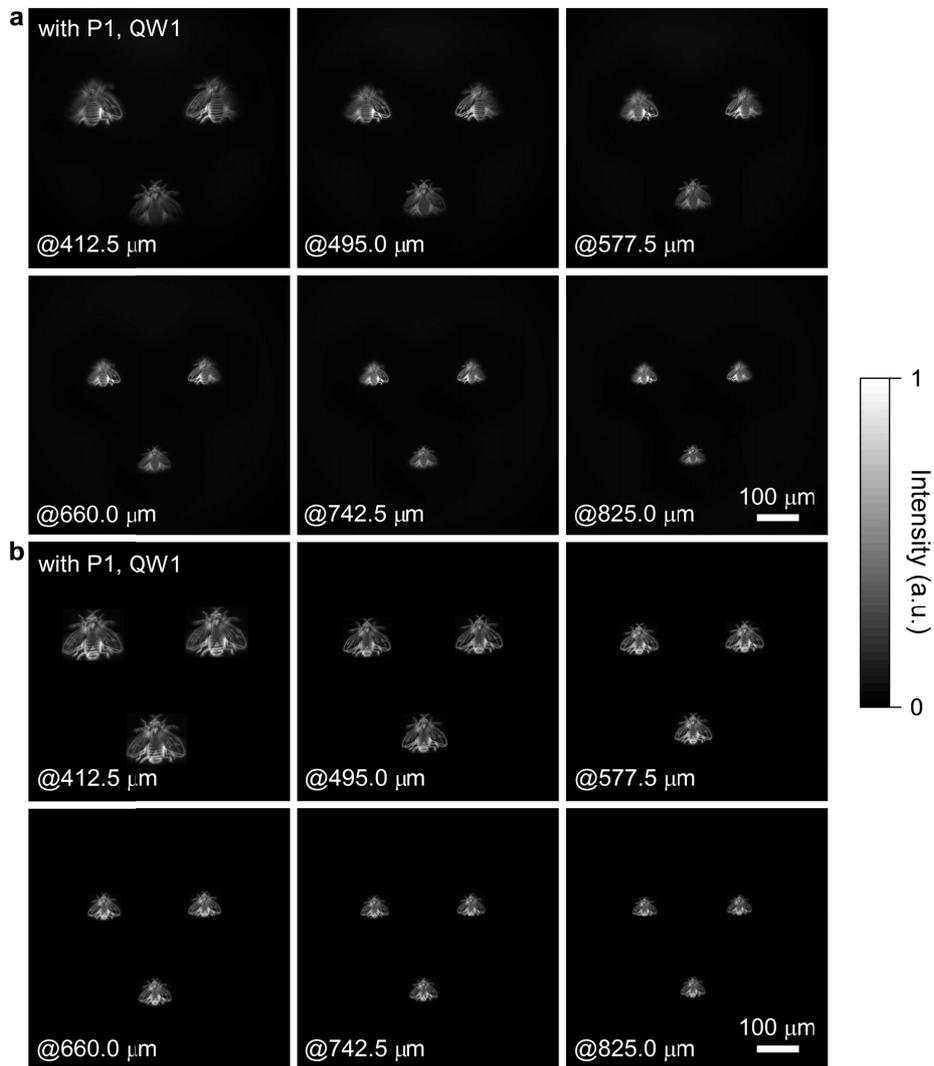
Fig. S6. Image distributions with P1 and QW1. (a) Raw images captured by the CCD camera for different object distances, showing obvious aberrations near the edge of the FOV. (b) The reconstructed images following the CCGD algorithm to correct the imaging aberrations.
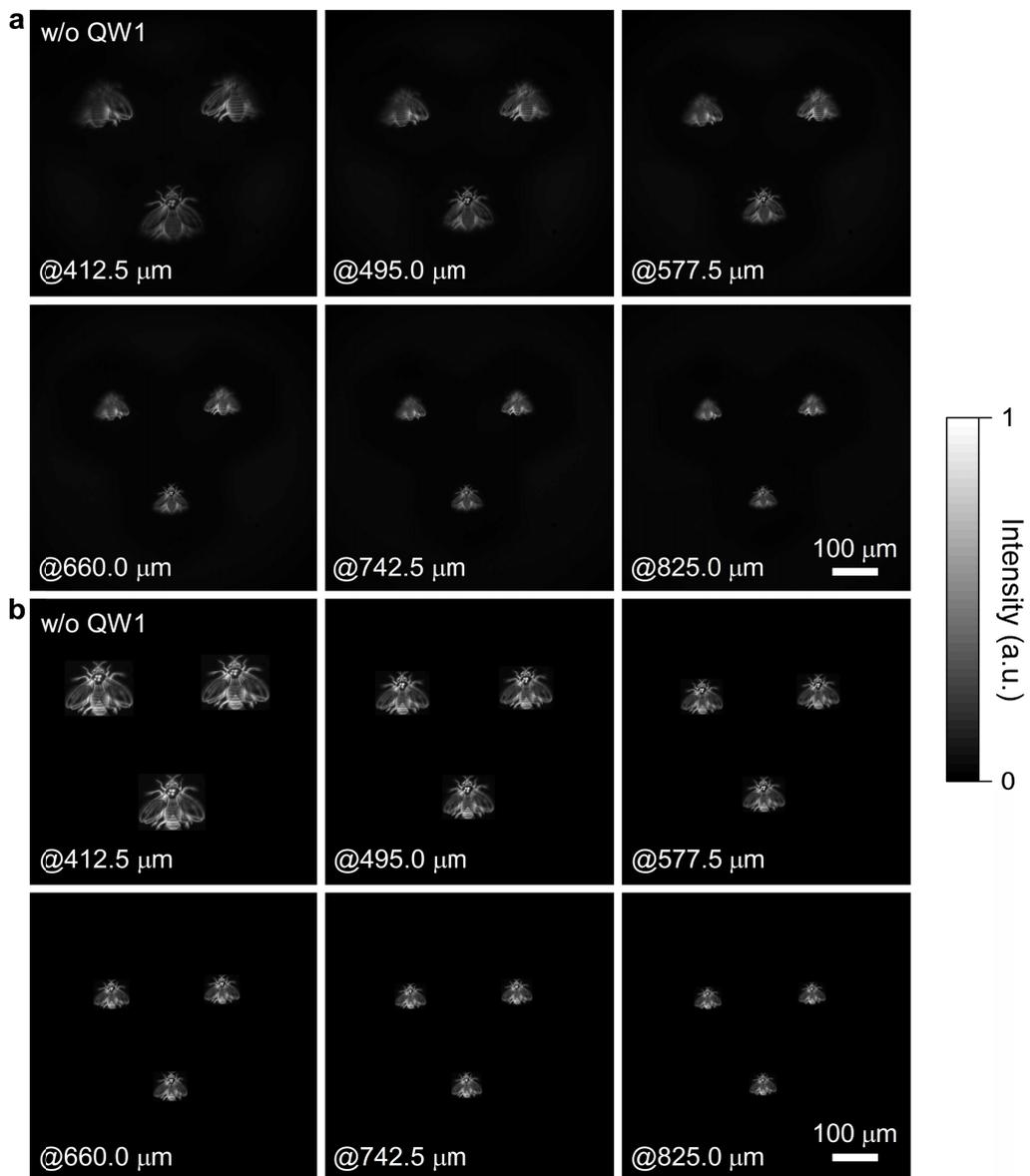
Fig. S7. Image distributions with P1 and without QW1. (a) Raw images captured by the CCD camera for different object distances. (b) Reconstructed images following the CCGD algorithm to correct the imaging aberrations. The images and positioning accuracy are comparable with those in Fig. S6.
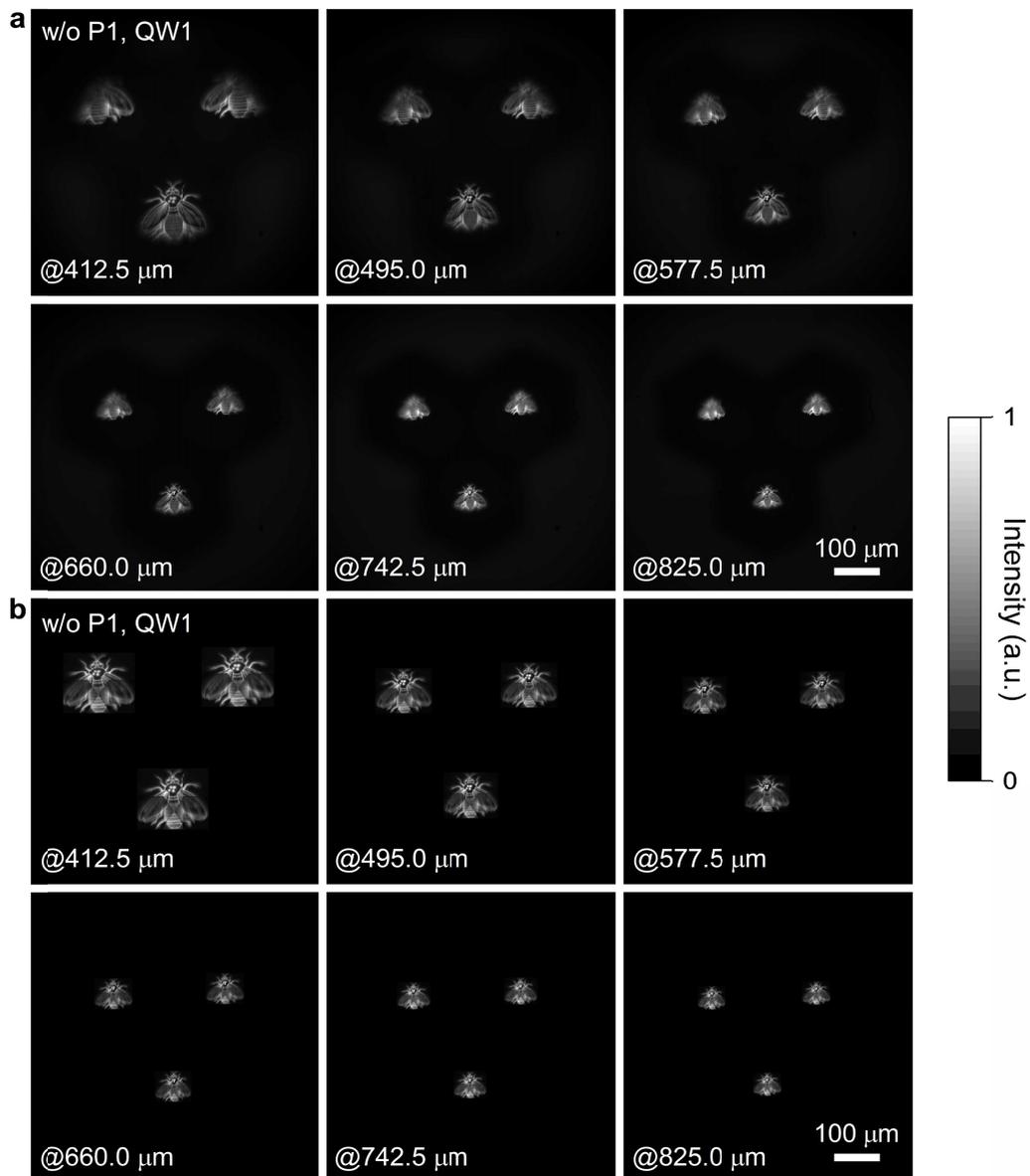
a — w/o P1, QW1

@412.5 μm   @495.0 μm   @577.5 μm

@660.0 μm   @742.5 μm   @825.0 μm   100 μm

b — w/o P1, QW1

@412.5 μm   @495.0 μm   @577.5 μm

@660.0 μm   @742.5 μm   @825.0 μm   100 μm

Intensity (a.u.)   1 — 0

Fig. S8. Image distributions without P1 and QW1 mimicking the natural light with arbitrary incident polarization. (a) Raw images captured by the CCD camera at different object distances. (b) Reconstructed images following the CCGD algorithm to correct the imaging aberrations. The images and positioning accuracy are comparable with those in Figs. S6 and S7.
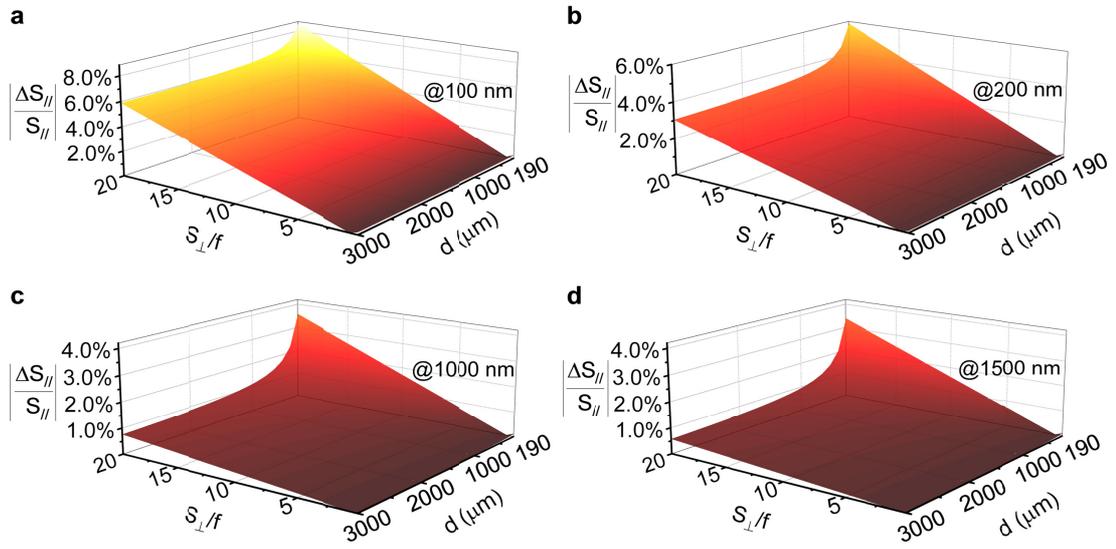
Fig. S9. Calculated relative positioning accuracy in the horizontal direction for different $S_{//}$. The relative positioning accuracy $|\Delta S_{//} / S_{//}|$ was calculated based on the measurement of ΔD for different $S_{//}$ preset values of (a) 100 nm, (b) 200 nm, (c) 1000 nm, and (d) 1500 nm. The results show that the horizontal relative positioning accuracy increases when increasing the length of $S_{//}$.
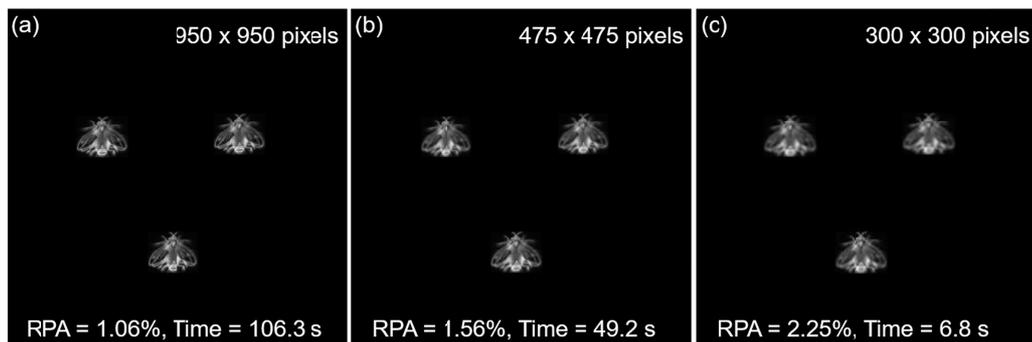

Fig. S10. Reconstructed images, associated RPAs, and the postprocessing time for different sizes of the original image. The entire calculation was performed using an Intel i5-9400F CPU in a commercial laptop.
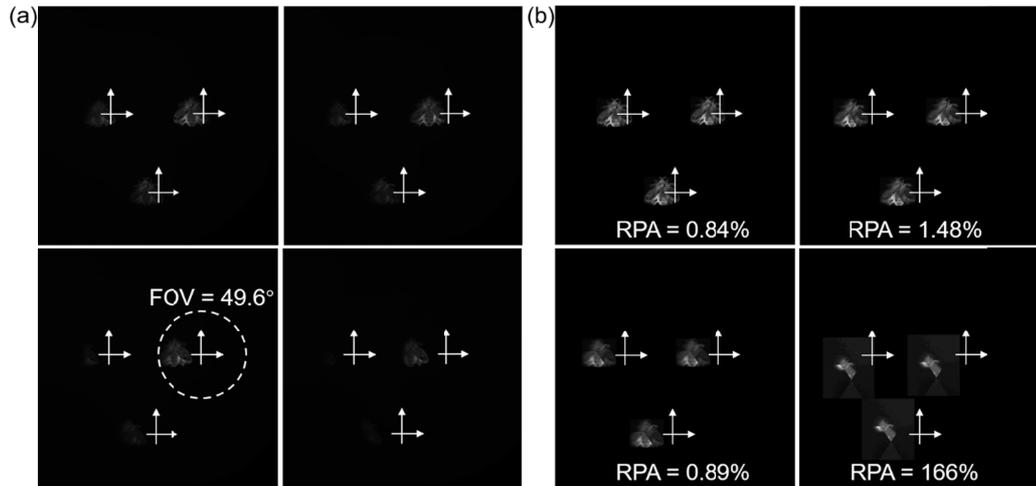
Fig. S11. Original images (a) and reconstructed images (b) for different setups of off-axis imaging. The white arrows indicate the center of each metalens. The white dashed circle in (a) indicates the FOV of 49.6°, which reaches the limit of the prototype of the metalens array. For a larger off-axis setup the RPA rapidly increases to 166%.
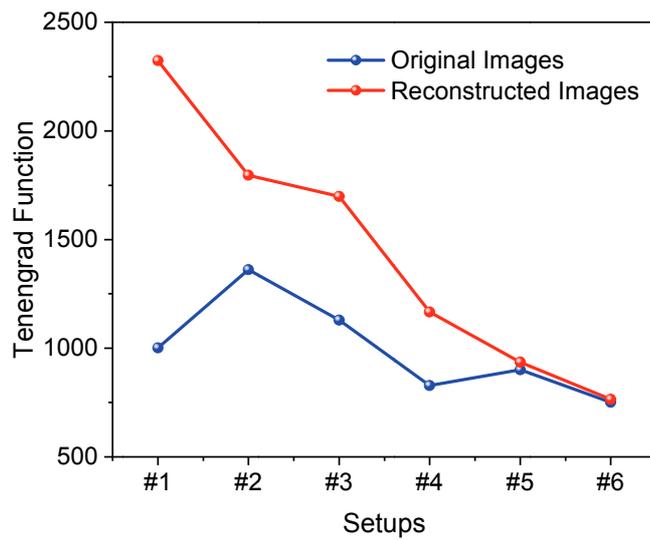

Fig. S12. Calculated Tenengrad function of the original and reconstructed images for different setups.

SUPPLEMENTARY VIDEO LEGENDS

Movie S1. Evolution of captured images. The movie shows the evolution of the raw images for the negative target object for different object distances.

Movie S2. Evolution of the captured images for the positive target object.

SUPPLEMENTARY REFERENCES

1. I. Culjak, D. Abram, T. Pribanic, H. Dzapo, and M. Cifrek, "A brief introduction to OpenCV," Proceedings of the 35th International Convention MIPRO, 1725-1730 (2012).

2. R. Laganiere, *OpenCV Computer Vision Application Programming Cookbook* (Packt Publishing, 2014).

3. D. P. Kingma and J. L. Ba, "ADAM: A Method for stochastic optimization," arXiv:1412.6980 (2014).

4. N. Zeller, F. Quint, and U. Stilla, "Depth estimation and camera calibration of a focused plenoptic camerafor visual odometry," ISPRS J. Photogramm **118**, 83 (2016).

5. E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, New York, 1998).

6. Y. Gao, Q. Li, R. Wu, J. Sha, Y. Lu, and F. Xuan, "Laser direct writing of ultrahigh sensitive SiC-based strain sensor arrays on elastomer toward electronic skins," Adv. Funct. Mater. **29**, 1806786 (2019).