Lensfree auto-focusing imaging with coarse-to-fine tuning method: supplemental document

The supplemental document is organized as follows:
(1) The filter selection of RoG metric is given.
(2) The natural image dataset used for training is presented.
(3) The auto-focusing results of RoG on random amplitude mask are added.
(4) Extra auto-focusing results of multi-exposure measurement are added.
(5) The definition and diagram of A_c and R_c are given.
(6) Extra auto-focusing results using different coarse-to-fine tuning strategies are given.

1. Filter selection of RoG metric

In our RoG metric, filter denoising is a key step to control the accuracy and robustness of the sharpness curve. Thus, it is necessary to find an optimal filter for auto-focusing. Here median filter, Gaussian filter, Wiener filter, wavelet filter, guided filter [1], bilateral filter, BM3D filter [2], and denoising convolutional neural network (DnCNN) [3] are used to find an optimal denoiser for RoG metric. We choose two samples, including stained intestine and ovary of fish, to accomplish this tuning task and arrange the corresponding results in Fig. S1 and Fig. S2. The kernel sizes of median filter, Gaussian filter, and Wiener filter are set as 7, and the kernel sizes of guided filter and bilateral filter are 5. The global variance of the bilateral filter is 3 and the local variance is 0.01, and the regularization parameter of guided filter is 0.01.

The auto-focusing sharpness curves of RoG with the above filters are shown in Fig. S1(a) and Fig. S2(a), where the estimated distances pointed from the peak values are listed in Tab. S1 and Tab. S2. As shown in Fig. S1(a), the curves of Gaussian and Wiener filters are fluctuated, which introduces multiple peak values on the sharpness curves. The curves of BM3D and DnCNN progressively increase and the peak value cannot be remarkably found. Different from the above filters, median, wavelet, guided, and bilateral filters can acquire the unimodal distribution around the real distance and its detail has been expanded in Fig. S1(b). With the use of the estimated

distances from Tab. S1, the amplitude images of *ovary of fish* are retrieved in Figs. S1(c-f). It is noted that the distance sets of Group I output a clear image and other images show remarkable defocused blur. This means that four filters, including median, wavelet, guided, and bilateral filters, accomplish the auto-focusing task for *ovary of fish*.

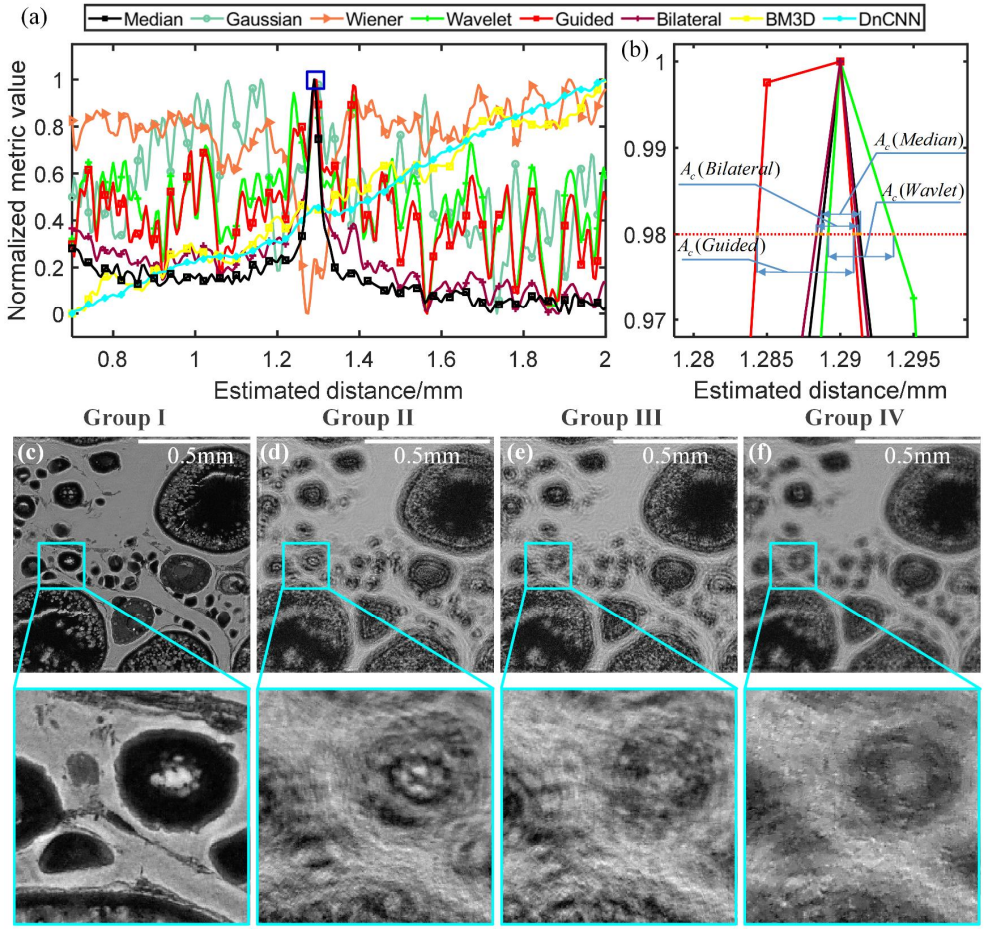

Fig. S1. Auto-focusing performance of *ovary of fish* by using RoG metric with different filters. (a) sharpness curve. (b) is the zoomed image cropped from the blue boxes in (a). (c-f) are retrieved images with four groups of the estimated distances.

Similar to Fig. S1, the sharpness curves of intestine tissue are shown in Fig. S2(a) and Fig. S2(b). With the use of the estimated distances from Tab. S2, the amplitude images of intestine tissue are retrieved in Figs. S2(c-g). It is noted that only median and bilateral filters could output a clear sample image. These two tests of Fig. S1 and Fig. S2 have shown that the robustness and

unimodality performance of sharpness curves highly rely on the filter for RoG metric.

Tab. S1. Estimated distances by using different filters for *ovary of fish*.

	Z_1/mm	Z_2/mm	Z_3/mm	Z_4/mm	Z_5/mm	Filter
Group I	1.290	1.390	1.490	1.590	1.690	Median, Wavelet, Guided, Bilateral
Group II	1.160	1.210	1.235	0.890	0.770	Gaussian
Group III	1.910	1.810	0.800	0.870	1.020	Wiener
Group IV	2.000	1.995	1.995	1.995	2.000	BM3D, DnCNN

Tab. S2. Estimated distances by using different filters for *intestine tissue*.

	Z_1/mm	Z_2/mm	Z_3/mm	Z_4/mm	Z_5/mm	Filter
Group I	1.340	1.440	1.540	1.640	1.740	Median, Bilateral
Group II	0.700	0.700	0.715	0.700	0.940	Gaussian
Group III	1.825	0.760	1.545	0.785	1.740	Wiener
Group IV	1.300	1.395	1.995	1.590	1.400	Wavelet, Guided
Group V	1.915	1.440	1.545	1.640	1.740	BM3D, DnCNN

Tab. S3. The accuracy criterion values (A_c) of Fig. S1(a) and Fig. S2(a) for different filters

Sample type	Accuracy criterion (unit:μm)							
	Median	Gaussian	Wiener	Wavelet	Guided	Bilateral	BM3D	DnCNN
Ovary of fish	**4.9**	×	×	5.0	5.5	5.0	×	×
Intestine	**5.1**	×	×	×	×	9.1	×	×

Tab. S4. The resolution criterion values (R_c) of Fig. S1(a) and Fig. S2(a) for different filters

Sample type	Resolution criterion (unit:μm)							
	Median	Gaussian	Wiener	Wavelet	Guided	Bilateral	BM3D	DnCNN

Ovary of fish	**243.3**	×	×	473.2	321.5	267.1	×	×
Intestine	**168.3**	×	×	×	×	343.6	×	×

To further show the filter performance, the accuracy and resolution criterion values using different filters are displayed in Tab. S3 and Tab. S4. For these two criterions, we can find that the sharpness curves of median filter own the smallest value. The running time of the RoG metric using different filters is also listed in Tab. S5, in which the sharpness curves of median filter take the least amount of time for distance estimation. Therefore, we select the median filter as a denoiser in the RoG metric for auto-focusing imaging.

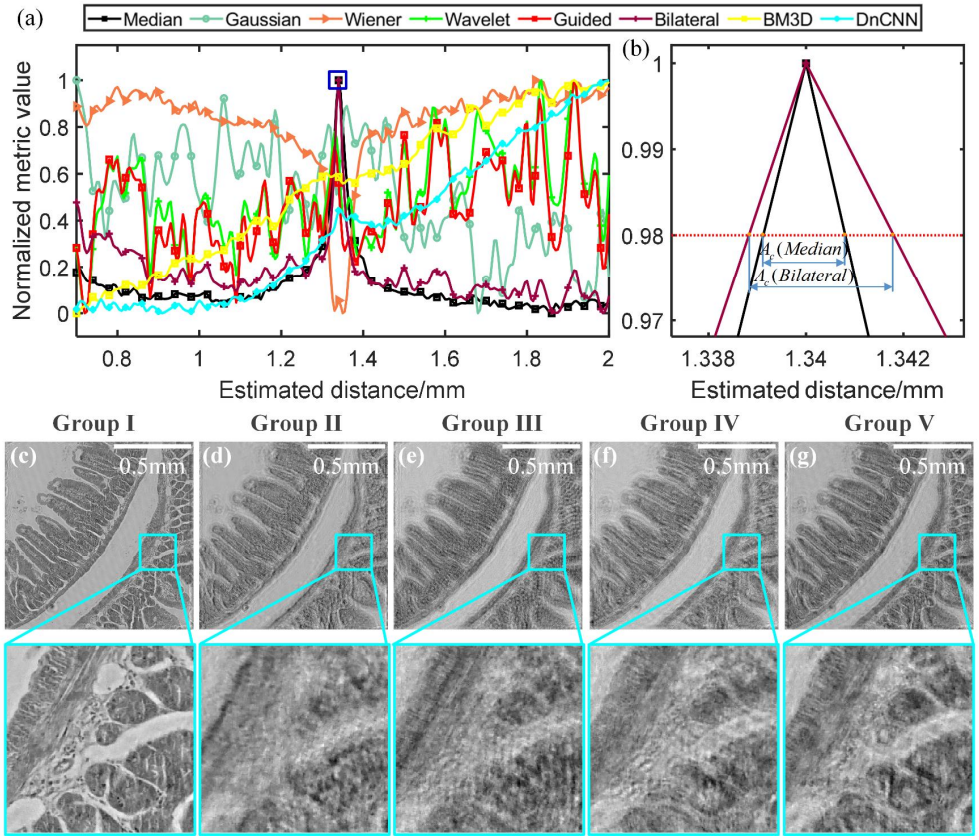

Fig. S2. Auto-focusing performance of *intestine* by using RoG metric with different filters. (a) sharpness curve. (b) is the zoomed image cropped from the blue boxes in (a). (c-g) are retrieved images with five groups of the estimated distances.

Tab. S5. The running time using different filters.

Filter type	Median	Gaussian	Wiener	Wavelet	Guided	Bilateral	BM3D	DnCNN
Running time	66.1s	20.5s	37.1s	140.6s	85.3s	1140.5s	25729s	13399s

2. Natural image dataset for training

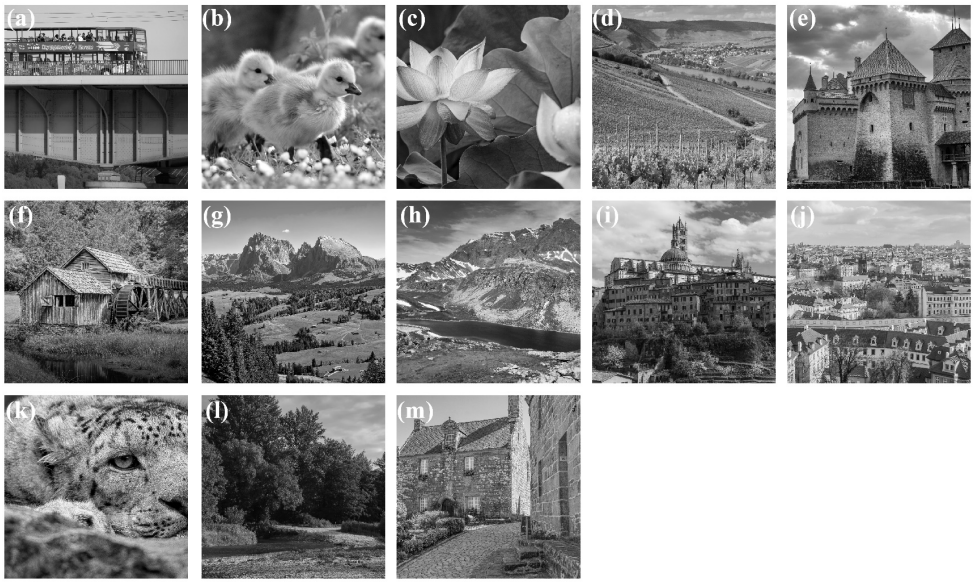

Fig. S3. The dataset contains 13 natural images with a size of 3000×3000 from Pixabay (under CC0 Creative Commons) [4].

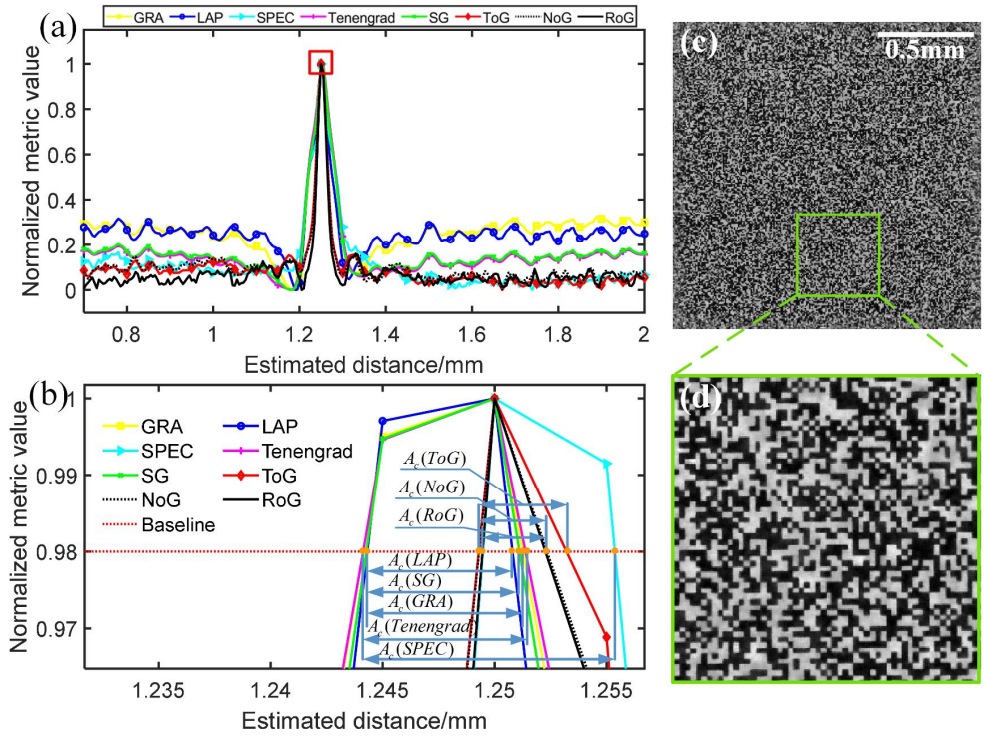

Fig. S4. The auto-focusing results of random amplitude mask. (a) The sharpness curve at the first height. (b) is the magnified curve from the red box in (a). (c) is the reconstructed result and its detail is expanded in (d).

3. Random amplitude mask

We also use a random amplitude mask (pixel size: 5μm) as the sample and arrange the auto-focusing results in Fig. S4. The sharpness curve at the first height is displayed in Fig. S4(a) and its highlighted curve is magnified in Fig. S4(b). As shown in Fig. S4(a) and Fig. S4(b), all metrics point to a common distance. For other heights, the estimated distances are also identical for all metrics. With the estimated distances, the amplitude image of the random mask is shown in Fig. S4(c) and Fig. S4(d), where the sharp edge of the mask has been acquired. Fig. S4(b) shows that the A_c values are listed as follows: 7.1μm (GRA), 6.5μm (LAP), 5.9μm (SPEC), 7.2μm (Tenengrad), 6.8μm (SG), 5.4μm (ToG), 5μm (NoG) and 4.9μm (RoG). Similarly, the values of R_c are listed as follows: 368.7μm (GRA), 347.1μm (LAP), 151.7μm (SPEC), 263.4μm (Tenengrad), 550μm (SG), 154.8μm (ToG), 190.6μm (NoG) and 149μm (RoG). The comparison between local and global criteria demonstrates

that our RoG metric surpasses other metrics in terms of both accuracy and unimodality.

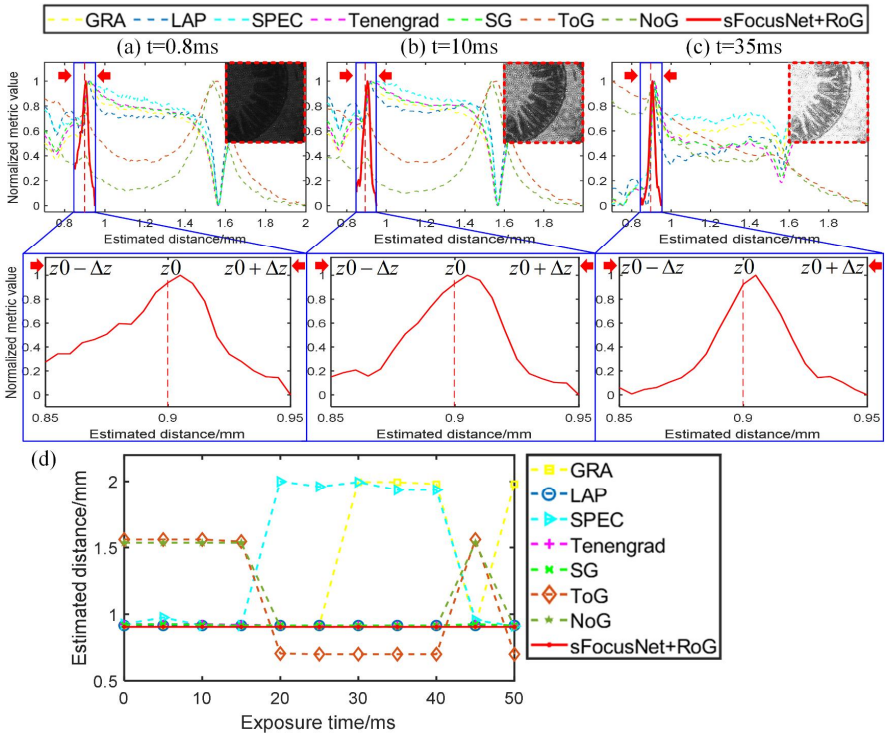

Fig. S5 Auto-focusing results of the homemade intestine tissue with different exposure time. (a-c) Auto-focusing curves of different methods. (d) The estimated distances of different methods under different exposure time.

4. Extra auto-focusing results for multi-exposure dataset

The auto-focusing results of the homemade intestine tissue are given in Fig. S5. Our method still outperforms other methods with high robustness and stability.

5. The definition and diagram of A_c and R_c

We prepare objective and subjective assessment criterions to judge the auto-focusing performance of the sharpness metrics. For subjective assessment, we use visual observation to judge whether or not the retrieved images are in-focus. For objective assessment, we use accuracy (A_c) and resolution criterions (R_c) to judge the auto-focusing performance of sharpness curves. The definition of A_c and R_c are illustrated in Fig. S6. As shown in Fig. S6(a), the accuracy criterion (A_c) is a parameter used to describe the width of a sharpness curve. A smaller A_c denotes that the sharpness metric is more sensitive to defocus blur. The resolution criterion (R_c) is a parameter used to

describe the unimodality and fluctuation of a sharpness curve, which can be calculated as $R_c = \frac{1}{\|Y\|_2}\sqrt{\left|\sum_{r=1}^{R}\left(e_r - z_{gt}\right)^2 |Y(e_r)|^2\right|}$. As shown in Fig. S6(b), a bigger R_c denotes that the sharpness curve is heavily fluctuated or the sharpness curve has multiple peaks. In an ideal condition, we hope that the sharpness curve has only one narrow peak and the defocused blur can be quantitatively represented. Thus, the accuracy (A_c) and resolution criterions (R_c) are usually used to judge the performance of the auto-focusing sharpness curve.

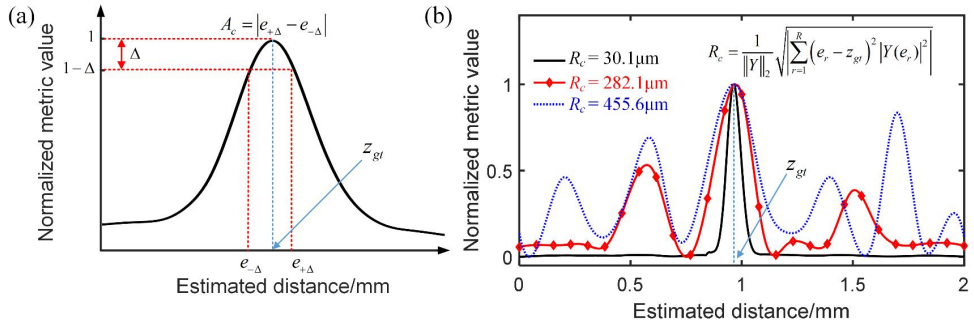

Fig. S6. The diagram of criterion definition for A_c and R_c.

6. Extra auto-focusing results using different coarse-to-fine tuning strategies

This metric-based coarse-to-tine strategy also fails to realize auto-focusing for over-exposure datasets. To show this, we provided the auto-focusing result of resolution target with the exposure time of 25ms in Fig. S7. Fig. S7(a) is the recorded intensity image. Fig. S7(b) is the auto-focusing curve of different sharpness metrics using the coarse-tuning. The coarse positions of peak values are marked with red, blue, and orange boxes, where the initial distances are specified as 0.7mm, 0.8mm, and 2mm. The coarse position acquired by our sFocusNet is 1mm, which is labelled with a red-dashed line.

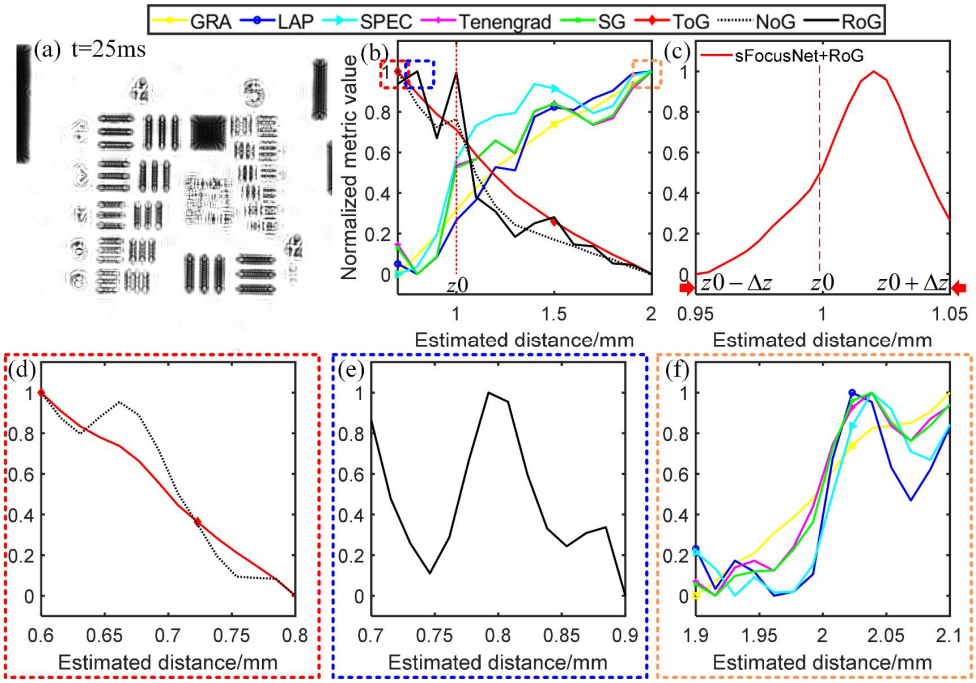

Fig. S7. The auto-focusing curves of resolution target with different coarse-to-fine tuning strategies. (a) Intensity image under over-exposure measurement. (b) The auto-focusing curve of sharpness metrics using coarse tuning. (c) The coarse-to-fine tuning curve of our method. (d-f) are fine-tuning results using the coarse positions from red, blue, and orange boxes of (b).

Tab. S6. Fine-tuned estimated distances with different sharpness metrics for resolution target.

	Z/mm	Metric
Group I	2.100	GRA
Group II	2.025	LAP
Group III	2.035	SPEC, Tenengrad, SG
Group IV	0.600	NoG, ToG
Group V	0.800	RoG
Group VI	1.020	sFocusNet+RoG

With these four initial positions, the fine-tuned auto-focusing curves are plotted in Fig. S7(c), Fig. S7(d), Fig. S7(e), and Fig. S7(f), in which the final diffractive distances are listed in Tab. S6. As the distances in Tab. S6 are performed, the amplitude images of the back-propagated results are shown in Fig. S8. It is noted that only our method outputs a clear target and other metrics fail to acquire the initial position.

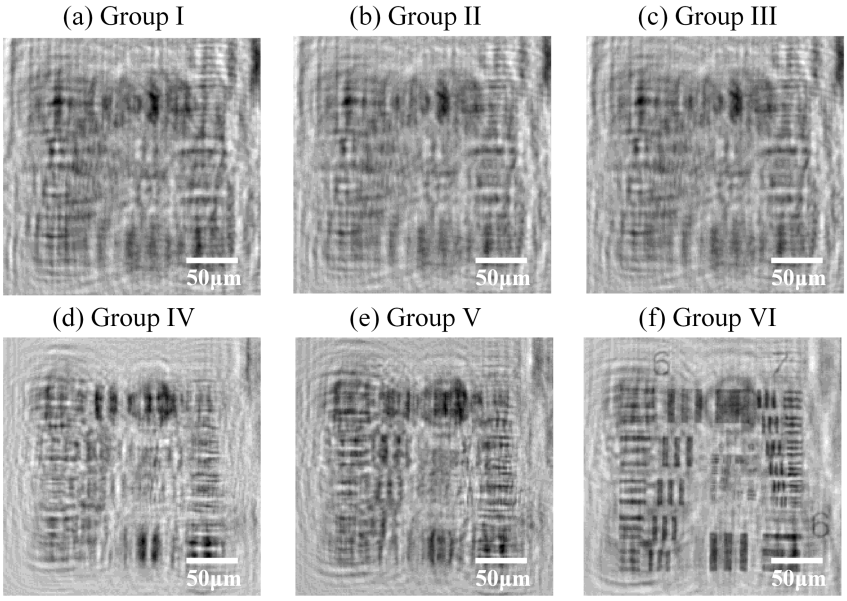

Fig. S8. The retrieved amplitude images of resolution target by using different groups of estimated distances from Tab. S6.

References

1. He K, Sun J and Tang X. Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. 2012; 35(6): 1397-1409.
2. Dabov K, Foi A, Katkovnik V, and Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process 2007; 16(8): 2080-2095.
3. Zhang K, Zuo W, Chen Y, Meng D, and Zhang L. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Trans. Image Process 2017; 26(7): 3142-3155.
4. https://pixabay.com/images/search/nature/.