Diffractive lensless imaging with optimized Voronoi-Fresnel phase: supplemental document

This document provides supplementary information to "Diffractive lensless imaging with optimized Voronoi-Fresnel phase". Technical details of the Point Spread Function, Modulation Transfer Function, Centroidal Voronoi Tesselation, Voronoi-Fresnel phase optimization, image formation and reconstruction, prototype fabrication, and additional results are provided to support the findings in the main paper.

1. POINT SPREAD FUNCTION

Here we derive the mathematical expressions for the point spread functions (PSFs) in the main texts. First, we inspect a single Voronoi-Fresnel cell V_i. The complex optical field after propagating from the phase element to the image sensor plane is

$$
\begin{aligned}
P_i(x,y,\lambda) &= \mathscr{F}_z\left\{A_i(\xi-\xi_i,\eta-\eta_i)\exp\left(-j\frac{2\pi}{\lambda}\frac{(\xi-\xi_i)^2+(\eta-\eta_i)^2}{2z}\right)\right\}\\
&= \frac{1}{j\lambda z}\iint A_i(\xi',\eta')\exp\left(-j\frac{\pi}{\lambda z}\left(\xi'^2+\eta'^2\right)\right)\exp\left(j\frac{\pi}{\lambda z}\left((x-\xi_i-\xi')^2+(y-\eta_i-\eta')^2\right)\right)\mathrm{d}\xi'\mathrm{d}\eta'\\
&= \frac{\exp\left(\frac{\pi}{\lambda z}\left((x-\xi_i)^2+(y-\eta_i)^2\right)\right)}{j\lambda z}\iint A_i(\xi',\eta')\exp\left(-j\frac{2\pi}{\lambda z}\left((x-\xi_i)\xi'+(y-\eta_i)\eta'\right)\right)\mathrm{d}\xi'\mathrm{d}\eta'\\
&= \frac{\exp\left(\frac{\pi}{\lambda z}\left((x-\xi_i)^2+(y-\eta_i)^2\right)\right)}{j\lambda z}\mathcal{F}\left\{A_i(\xi',\eta')\right\}\Big|_{\left[\frac{x-\xi_i}{\lambda z},\frac{y-\eta_i}{\lambda z}\right]}\\
&= \frac{C(\xi_i,\eta_i)}{j\lambda z}\mathcal{F}\left\{A_i(\xi',\eta')\right\}\Big|_{\left[\frac{x-\xi_i}{\lambda z},\frac{y-\eta_i}{\lambda z}\right]},
\end{aligned}
\tag{S1}
$$

where we have replaced the variables by $\xi'=\xi-\xi_i$ and $\eta'=\eta-\eta_i$ in the second line. The integral in the third line is exactly the Fourier transform of the aperture function A_i evaluated at spatial frequencies $[(x-\xi_i)/\lambda z,(y-\eta_i)/\lambda z]$. The corresponding PSF is then

$$
\begin{aligned}
\mathbf{PSF}_i(x,y,\lambda)(x,y,\lambda) &= |P_i(x,y,\lambda)|^2\\
&\propto \left|\mathcal{F}\left\{A_i(\xi',\eta')\right\}\Big|_{\left[\frac{x-\xi_i}{\lambda z},\frac{y-\eta_i}{\lambda z}\right]}\right|^2\\
&= \mathbf{PSF}_i^0(x-\xi_i,y-\eta_i,\lambda),
\end{aligned}
\tag{S2}
$$

where we denote $\mathbf{PSF}_i^0(x,y,\lambda)$ as the centered PSF as if the aperture were located at the origin,

$$
\begin{aligned}
\mathbf{PSF}_i^0(x,y,\lambda) &\propto \left|\mathcal{F}\left\{A_i(\xi',\eta')\right\}\right|^2\\
&= \left|\iint A_i(\xi',\eta')\exp\left(-j\frac{2\pi}{\lambda z}(x\xi'+y\eta')\right)\mathrm{d}\xi'\mathrm{d}\eta'\right|^2.
\end{aligned}
\tag{S3}
$$

This implies that the PSF for the i-th Voronoi-Fresnel cell is a shifted version of the centered PSF. It is worth noting that the shape and distribution of the centered PSF depend on the geometry of the aperture functions. These apertures are finite-edge polygons, so the centered PSFs are actually compact yet highly directional filters. It is important to diversify such directional filtering of the PSFs to achieve optimal performance.

The effective PSF of the Voronoi-Fresnel lensless camera is obtained in the same way by taking the whole phase into account,

$$
\begin{aligned}
\mathbf{PSF}(x,y,\lambda) &= \left|\mathscr{F}_z\left\{\Phi(\xi,\eta,\lambda)\right\}\right|^2 \\
&= \left|\mathscr{F}_z\left\{\sum_{i=1}^{K} A_i(\xi-\xi_i,\eta-\eta_i)\cdot\exp\left(-j\frac{2\pi}{\lambda}\cdot\frac{(\xi-\xi_i)^2+(\eta-\eta_i)^2}{2z}\right)\right\}\right|^2 \\
&= \left|\sum_{i=1}^{K} P_i(x,y,\lambda)\right|^2 \\
&= \left(\sum_{i=1}^{K} P_i^*(x,y,\lambda)\right)\left(\sum_{i=1}^{K} P_i(x,y,\lambda)\right) \\
&= \sum_{i=1}^{K}\left|P_i(x,y,\lambda)\right|^2 + \sum_{\substack{j=1\\ i\neq j}}^{K}\sum_{i=1}^{K} P_i^*(x,y,\lambda)\,P_j(x,y,\lambda)\,.
\end{aligned}
\tag{S4}
$$

Although the Voronoi cells share no intersections, $V_i \cap V_j = \emptyset$, $\forall i \neq j$, and the individual aperture functions have no overlapped areas, the diffraction patterns $P_i(x,y,\lambda)$ and $P_j(x,y,\lambda)$ in general would interfere with each other, so the cross terms in Eq. (S4) are not necessarily zero.

We investigate the cross terms in two situations. One is a random distribution of adjacent Fresnel centers, and the other is where the centers are at the centroids. A simple example is a rectangular space tessellated with two adjacent cells, as shown in Fig. S1. The optical parameters are the same as in Fig. 3 in the main paper. When the centers of two adjacent cells are not located in the centroids, and very close to each other near the boundary, the sum of individual PSFs are different from the real PSF predicted by the Fresnel propagation. However, when the centers are at the centroids of the cells, the sum of individual PSFs is approximately the same as the accurate PSF predicted by the Fresnel diffraction. The maximum error is less than 1%.

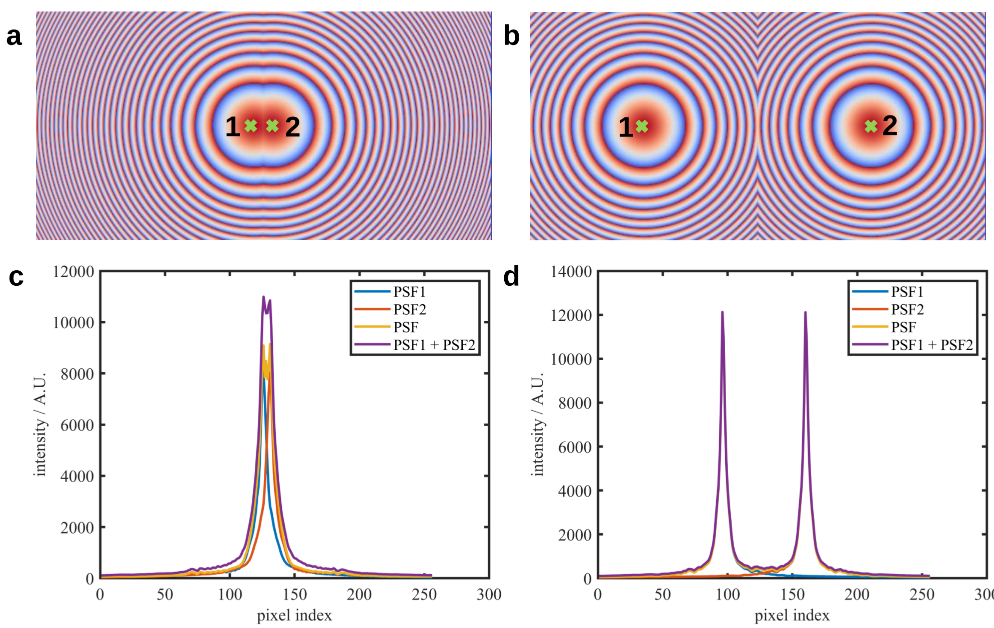

Fig. S1. Cross terms in two adjacent PSFs. (a) Two cells with very close centers near the boundary. (b) Two cells with centers at the centroids. (c) Cross-section of the individual PSFs, total PSFs and the sum of individual PSFs in (a). (d) Cross-section of the individual PSFs, total PSFs and the sum of individual PSFs in (b). Maximum error is less than 1% when the the centers are at the centroids of the two cells.

Hence, for the Centroidal Voronoi case, the entire PSF can be approximated by omitting the cross terms, just for the purpose of analysis,

$$\mathbf{PSF}\left(x,y,\lambda\right) \approx \sum_{i=1}^{K}\left|P_i\left(x,y,\lambda\right)\right|^2 = \sum_{i=1}^{K}\mathbf{PSF}_i^0\left(x-\xi_i,y-\eta_i,\lambda\right), \tag{S5}$$

Note that in the simulation and optimization, we do not simulate individual PSFs and superposition them, but use the entire constructed phase function to obtain the panchromatic PSF.

The above PSF is for monochromatic light. To get the panchromatic PSF, we simply integrate all the spectral PSFs over the interested spectral range,

$$\mathbf{PSF}\left(x,y\right) = \int_{\lambda_1}^{\lambda_2}\mathbf{PSF}\left(x,y,\lambda\right)\mathrm{d}\lambda. \tag{S6}$$

2. MODULATION TRANSFER FUNCTION

Here we provide a detailed derivation and analysis for the Modulation Transfer Function (MTF). MTF is defined as the magnitude of the Optical Transfer Function (OTF), which is the Fourier transform of the PSF for incoherent imaging systems,

$$\mathbf{MTF} = \left|\mathbf{OTF}\right| = \left|\mathcal{F}\left\{\mathbf{PSF}\right\}\right|, \tag{S7}$$

where $0 \leq \mathbf{MTF} \leq 1$. We can obtain the MTF by taking the Fourier transform of the above PSF,

$$\begin{aligned}\mathbf{MTF}\left(f_X,f_Y\right) &= \left|\mathcal{F}\left\{\mathbf{PSF}\left(x,y\right)\right\}\right| \\ &\approx \left|\sum_{i=1}^{K}\mathcal{F}\left\{\mathbf{PSF}_i^0\left(x-\xi_i,y-\eta_i\right)\right\}\right| \\ &= \left|\sum_{i=1}^{K}\mathcal{F}\left\{\mathbf{PSF}_i^0\left(x,y\right)\right\}\exp\left(-jf_X\xi_i,-jf_Y\eta_i\right)\right| \\ &= \left|\sum_{i=1}^{K}\mathbf{OTF}_i^0\left(f_X,f_Y\right)\exp\left(-jf_X\xi_i,-jf_Y\eta_i\right)\right|,\end{aligned} \tag{S8}$$

where f_X and f_Y are the Fourier domain frequencies, and we have applied the translation property of Fourier transform. The individual OTFs in the complex form are

$$\mathbf{OTF}_i^0\left(f_X,f_Y\right) = \mathcal{M}_i\left(f_X,f_Y\right)\exp\left(-j\mathcal{P}_i\left(f_X,f_Y\right)\right). \tag{S9}$$

The remaining phase delay terms are simplified as

$$\exp\left(-j\Psi_i\left(f_X,f_Y\right)\right) = \exp\left(-jf_X\xi_i,-jf_Y\eta_i\right), \tag{S10}$$

so the MTF can now be rewritten as

$$\mathbf{MTF}\left(f_X,f_Y\right) = \left|\sum_{i=1}^{K}\mathcal{M}_i\left(f_X,f_Y\right)\exp\left(-j\mathcal{P}_i\left(f_X,f_Y\right)\right)\exp\left(-j\Psi_i\left(f_X,f_Y\right)\right)\right|. \tag{S11}$$

This equation reveals three factors that affect the MTF, the diffraction by each similar apertures that determines the $\mathcal{M}\left(f_X,f_Y\right)$ and $\mathcal{P}_i\left(f_X,f_Y\right)$ terms; the additional phase delay terms $\Psi_i\left(f_X,f_Y\right)$ that is introduced by the amount of spatial shifts from the centered PSFs; and the total number of Voronoi-Fresnel cells K.

In addition, we show how MTFv is related to the Strehl ratio. Strehl ratio is defined as the peak intensity ratio between the aberrated PSF and the diffraction limited PSF [1],

$$\mathrm{SR} = \frac{I_{\mathrm{ab}}(0,0)}{I_{\mathrm{dl}}(0,0)}, \tag{S12}$$

where $I_{\mathrm{ab}}(0,0)$ is the peak intensity of the aberrated PSF in the origin, and $I_{\mathrm{dl}}(0,0)$ is the corresponding intensity of the diffraction limited PSF. In practice, it is difficult to use Strehl ratio in this form as an optimization metric, as it is challenging to optimize the peak intensity of the PSF, and there may be multiple peaks in a composite system like ours. According to the definition of Fourier transform, we can rewrite the above equation in the Frequency domain,

$$\mathrm{SR} = \frac{\iint\mathbf{OTF}_{ab}\left(f_X,f_Y\right)df_Xdf_Y}{\iint\mathbf{OTF}_{dl}\left(f_X,f_Y\right)df_Xdf_Y}. \tag{S13}$$

Now it becomes a quantity that can be calculated more easily with OTFs. We also note that, since the diffraction limited OTF is in the denominator, and is often a fixed value for a given system. We only need the term in the numerator. In addition, OTF consists of complex values. For stable numerical computation, we can replace OTF with MTF in the above equation. Since MTF is positive definite, the integral of MTF is always no less than the integral of OTF [2],

$$\iint \mathbf{OTF}_{ab}\left(f_X, f_Y\right) df_X df_Y \leq \iint \mathbf{MTF}_{ab}\left(f_X, f_Y\right) df_X df_Y. \tag{S14}$$

So finally we have

$$\mathbf{MTFv} = \iint \mathbf{MTF}_{ab}\left(f_X, f_Y\right) df_X df_Y. \tag{S15}$$

We note that, strictly speaking, although MTFv is related to the Strehl ratio, they are not identical. MTFv is a more generalized term that can be calculated easily from the MTF. It is also a tractable quantity that measures the information collected in an optical system. Therefore, we choose MTFv as a figure-of-merit for the optimization of our lensless imaging system.

3. CENTROIDAL VORONOI TESSELLATION

Finding the optimal tessellation of the 2D design space is basically a sampling problem. Among various sampling methods, blue noise sampling [3, 4] offers minimal low-frequency components and no concentrated spikes of energy, which is the required properties for our application. An effective way to achieve blue noise sampling is by Centroidal Voronoi Tessellation (CVT) [5]. The CVT is a special Voronoi diagram where the sites coincide with the mass centroids of the corresponding Voronoi regions. The mass centroid of a Voronoi region V_i is defined as

$$\mathbf{c}_i = \frac{\int_{V_i} \mathbf{p}\tau\left(\mathbf{p}\right) \mathrm{d}\mathbf{p}}{\int_{V_i} \tau\left(\mathbf{p}\right) \mathrm{d}\mathbf{p}}, \tag{S16}$$

where $\mathbf{p}$ is a point in the Cartesian coordinates, and τ is a given density function. We can assume a constant density across the 2D plane for simplicity, i.e., $\tau \equiv 1$. CVT is a critical point of the energy function defined by

$$E_{\text{CVT}}\left(P\right) = \sum_{i=1}^{K} \int_{V_i} \tau\left(\mathbf{p}\right) \left\|\mathbf{p} - \mathbf{p}_i\right\|^2 \mathrm{d}\mathbf{p}. \tag{S17}$$

There are various algorithms to optimize the above energy function and generate optimal CVT [5, 6]. A classic method is to use Lloyd iterations [7] to update the Voronoi sites by their centroids until convergence. In each iteration, the mass centroids are computed for the current Voronoi regions. Then these generated sites are replaced by the calculated centroids, and a new Voronoi tessellation is constructed. The process is repeated until a convergence criterion is met.

4. SCALABILITY

We find that the optimal number of Voronoi-Fresnel cells is linearly proportional to the design area. Here we analyze and validate this assumption for different design areas with various aspect ratios. For all the experimental designs below, we assume the substrate is fused silica, and design the phase at 550 nm. The distance between the phase and sensor is 2 mm. Phase pixel size is 1 μm. In Fig. S2a-e we show the optimization curves for these experimental designs of aspect ratios of 1:1, 4:3, 3:2, 16:9, and 21:9, respectively. These aspect ratios allow us to account for common sensor shapes. For each aspect ratio, we start from a small area to optimize the Voronoi-Fresnel phase for various number of cells. Then we double the size in both dimensions (scaling to a quadruple area). We repeat this procedure 4 times for each aspect ratio. The best number of Voronoi-Fresnel cells in each design is obtained by fitting the data into a 5th order polynomial, and evaluate the cell number when the MTFv reaches the peak. These data are used in Fig. 5 in the main paper.

5. IMAGE RECONSTRUCTION

A color image recorded on the image sensor is an integral of spectrally-blurred images weighted by the color response of the sensor. This process can be expressed as

$$g_c\left(x, y\right) = \int_{\lambda_1}^{\lambda_2} q_c\left(\lambda\right) \left(f\left(x, y, \lambda\right) * h\left(x, y, \lambda\right)\right) \mathrm{d}\lambda, \tag{S18}$$

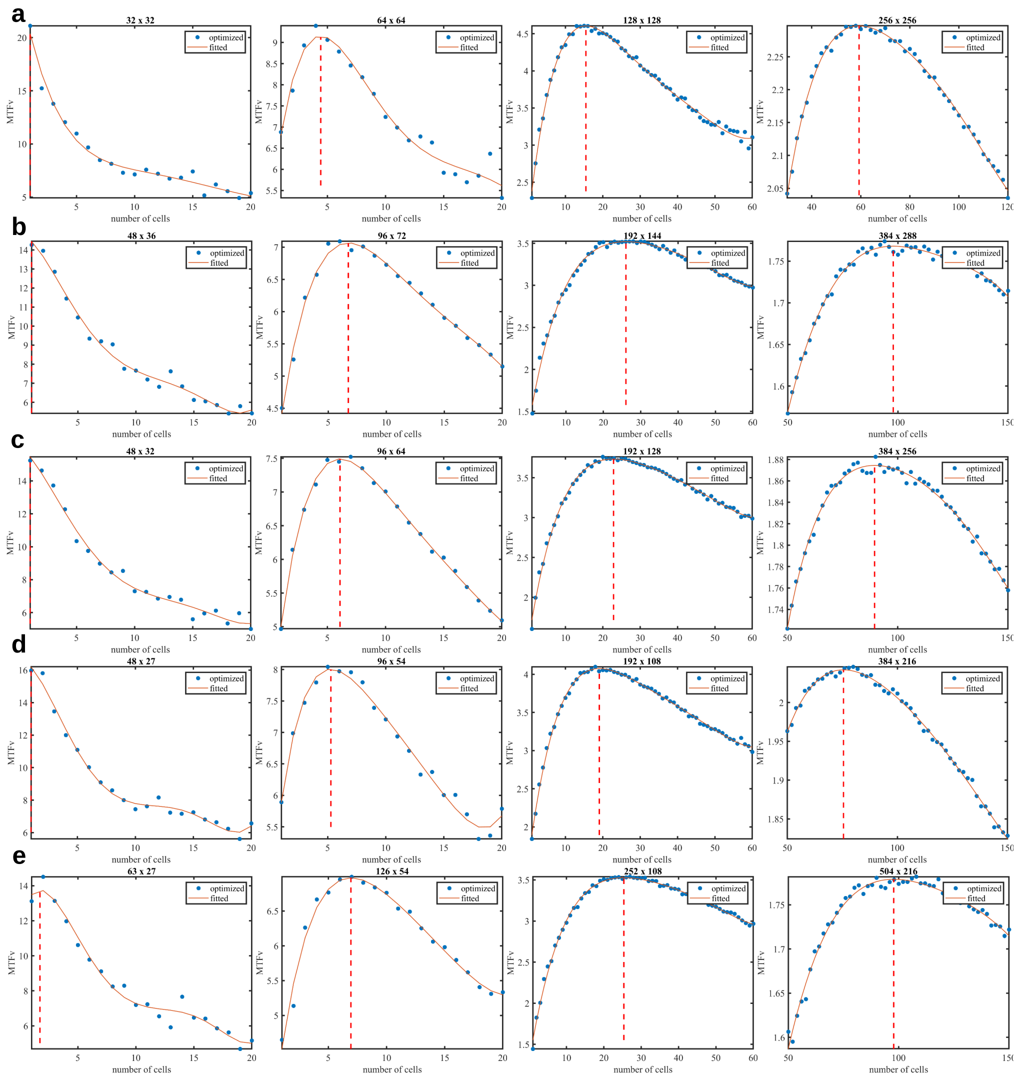

Fig. S2. Linear scalability analysis for various aspect ratios. (a) Aspect ratio 1:1. (b) Aspect ratio 4:3. (c) Aspect ratio 3:2. (d) Aspect ratio 16:9. (e) Aspect ratio 21:9. For each aspect ratio, we quadruple the previous design area from left to right, as indicated in the figure titles (unit: μm^2).

where $f(x, y, \lambda)$ is the latent spectral image at wavelength λ, $h(x, y, \lambda)$ is the spectral PSF, and $*$ denotes spatial convolution. $q_c(\lambda)$ is the color response function of the sensor, and $g_c(x, y)$ is the captured color image in color channel c (for sensors with Bayer filters, $c = r, g, b$).

If the imaging system is perfect, we assume the spectral PSFs are all identically Dirac delta functions, i.e., $h(x, y, \lambda) = \delta(x, y, \lambda) = \delta(x, y)$. The ground-truth sharp image would be

$$f_c(x, y) = \int_{\lambda_1}^{\lambda_2} q_c(\lambda) f(x, y, \lambda) \, \mathrm{d}\lambda. \tag{S19}$$

On the other hand, if the image is a spectrally-uniform ideal point source, $f(x, y, \lambda) = \delta(x, y, \lambda) = \delta(x, y)$, the captured image would be a color PSF

$$h_c(x, y) = \int_{\lambda_1}^{\lambda_2} q_c(\lambda) h(x, y, \lambda) \, \mathrm{d}\lambda. \tag{S20}$$

Note that if the imaging system is approximately achromatic, $h(x, y, \lambda) \approx h(x, y)$, Eq. (S18) can

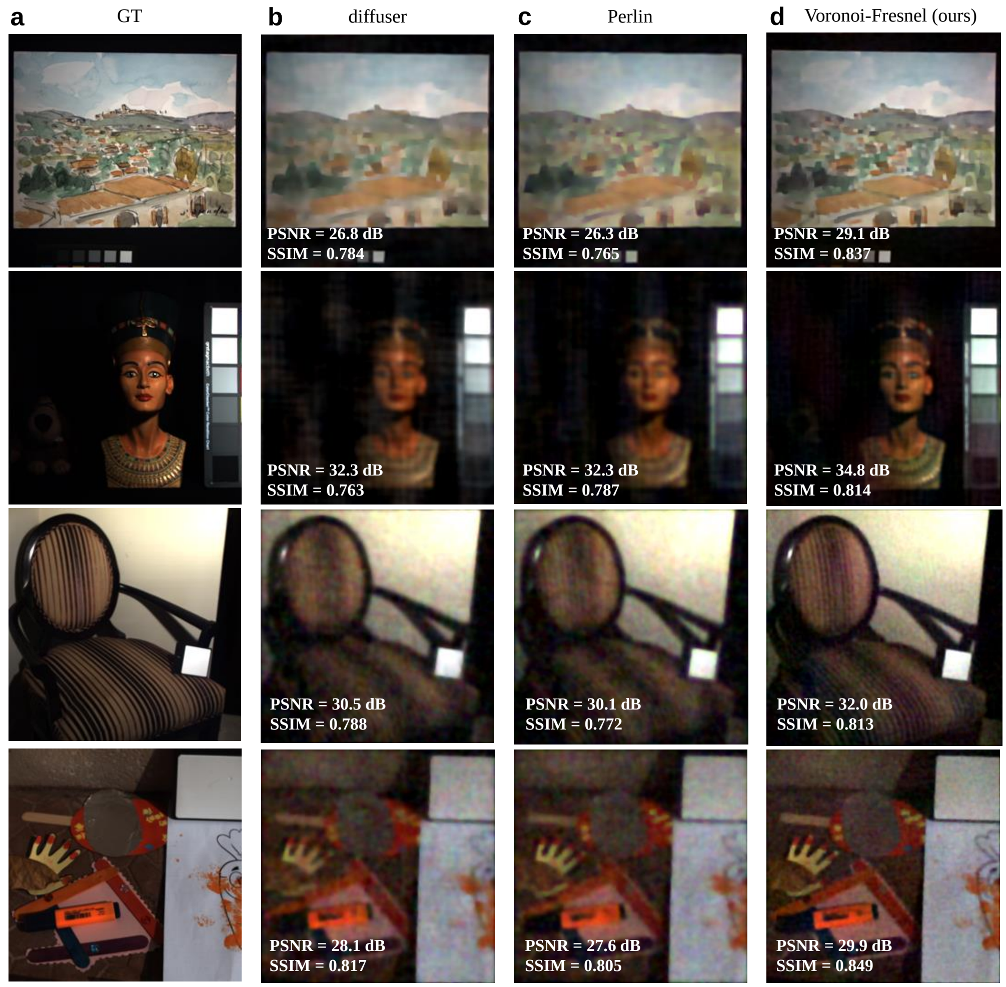

Fig. S3. Additional comparison results. (a) Reference ground truth image. (b) Reconstructed image by the diffuser PSF. (c) Reconstructed image by the Perlin PSF. (d) Reconstructed image by the Voronoi-Fresnel PSF.

be simplified by

$$\begin{aligned} g_c(x,y) &\approx \left(\int_{\lambda_1}^{\lambda_2} q_c(\lambda) f(x,y,\lambda)\, \mathrm{d}\lambda\right) * h(x,y) \\ &= f_c(x,y) * \int_{\lambda_1}^{\lambda_2} q_c(\lambda) h(x,y)\, \mathrm{d}\lambda \\ &= f_c(x,y) * h_c(x,y), \end{aligned} \tag{S21}$$

where the sensor response is normalized such that $\int_{\lambda_1}^{\lambda_2} q_c(\lambda)\, \mathrm{d}\lambda = 1$. To solve a minimization problem in the ADMM framework, we introduce a slack variable $\mathbf{z} = \mathbf{D}\mathbf{x}$, and apply the augmented Lagrangian multiplier,

$$\arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \mu\|\mathbf{z}\|_1 + \frac{\rho}{2}\|\mathbf{D}\mathbf{x}-\mathbf{z}\|_2^2, \tag{S22}$$

where ρ is the weight for the slack variable. The ADMM iterations consist of three steps,

$$\begin{cases} \mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \frac{\rho}{2}\|\mathbf{D}\mathbf{x}-\mathbf{z}+\mathbf{u}\|_2^2 \\ \mathbf{z}^{k+1} = \arg\min_{\mathbf{z}} \mu\|\mathbf{D}\mathbf{x}-\mathbf{z}\|_1 + \frac{\rho}{2}\|\mathbf{D}\mathbf{x}-\mathbf{z}+\mathbf{u}\|_2^2 \\ \mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{D}\mathbf{x}^{k+1} - \mathbf{z}^{k+1}, \end{cases} \tag{S23}$$

where $\mathbf{u}$ is the scaled dual variable. The $\mathbf{x}$-problem can be efficiently solved in the Fourier domain. The analytical solution is

$$\mathbf{x}^{k+1} = \left(\mathbf{A}^T\mathbf{A} + \rho\mathbf{D}^T\mathbf{D}\right)^{-1}\left(\mathbf{A}^T\mathbf{A} + \rho\mathbf{D}^T\mathbf{D}\right). \tag{S24}$$

The $\mathbf{z}$-problem has a closed-form solution,

$$\mathbf{z}^{k+1} = \mathcal{S}_{\mu/\rho}\left(\mathbf{D}\mathbf{x}^{k+1} + \mathbf{u}^k\right), \tag{S25}$$

where $\mathcal{S}_\kappa(v) = (1 - \kappa/|v|)_+ v$ is an element-wise soft thresholding operator. Finally $\mathbf{u}$ is updated with the new $\mathbf{x}$ and $\mathbf{z}$.

We show more comparison results in Fig. S3 for the three candidate designs, the diffuser PSF [8], the Perlin PSF [9], and our Voronoi-Fresnel PSF. The top two rows are from dataset [10] with lab setup scenes, and the bottom two rows are from the dataset [11] with natural indoor scenes. Our design outperforms the other two in both PSNR and SSIM.

6. COMPARISON WITH AUTO-CORRELATION

The auto-correlation of the PSF is a concept related to MTF, and could in principle be used for optimization of optical phase elements, such as in Miniscope3D [12]. Since auto-correlation is not a single-number, it cannot be used directly as an optimization metric. A diffraction limited MTF must be taken as the reference. A major difference in the proposed MTFv concept is that, spatial and spectral information in the PSF are distilled into a single number that can be used for numerical optimization. Our MTFv metric requires no reference value, and evaluates the amount of useful information by itself.

In addition, we present an example to show that, auto-correlation may not be consistent in certain cases from the perspective of image quality, whereas our MTFv metric is more directly related to the final performance. Here we use four PSFs that show different auto-correlation properties. The first three are those we use in Fig. 4 in the main paper. Their auto-correlation functions are shown in Fig. S4a-c. The fourth one is a white Gaussian noise pattern, which has an extremely sharp Dirac-like auto-correlation function (Fig. S4d). As a comparison, the rectangular-grid and hexagonal-grid PSFs exhibit very large support, whereas the optimized Voronoi-Fresnel PSF has a moderate shape. We also compute their corresponding MTFv values, 1.14 (rectangular), 1.19 (hexagonal), 2.29 (Voronoi-Fresnel), and 0.83 (Gaussian). To evaluate the final image quality for these PSFs, we reconstruct the sharp images using the same parameters. The reconstructed images are shown in Fig. S4e-h, with the PSNR and SSIM values shown below. Our Voronoi-Fresnel PSF yields the best PSNR (32.4 dB) and SSIM (0.765), significantly better than the other three.

This example indicates that, although the white Gaussian noise pattern has a strong peak in the auto-correlation, it fails in reconstructing a reasonable image. The rectangular and hexagonal PSFs show very large support in the auto-correlation, but not as good as the optimized Voronoi-Fresnel PSF. The result also emphasizes the importance of a proper reference quantity for the success of the auto-correlation metric. As a comparison, the proposed MTFv metric is consistently related with the image quality, and naturally rules out the white Gaussian noise PSF, so MTFv is a more robust metric for this design problem.

7. FABRICATION

The experimental samples are fabricated by a combination of photolithography and reactive-ion etching (RIE) techniques. The substrate is a 4 inch fused silica wafer with a thickness of 0.5 mm. We binarize the optimized Voronoi-Fresnel phase profile into $2^4 = 16$ levels, and repeat 4 iterations of the basic photolithography with different masks and then RIE with doubled etching time. The masks are fabricated on soda lime substrates by laser direct writing on a Heidelberg μPG 501 mask maker. In each iteration, the wafer is first cleaned in Piranha solution at 115°C for 10 min, and dried with N_2. A 150-nm-thick Chromium (Cr) layer is deposited by sputtering on one side of the wafer. A 0.6-μm-thick layer of photoresist AZ1505 is then spin-coated on top of Cr after HMDS (Hexamethyldisilazane) vapor priming. To transfer the mask patterns to the wafer, we align the wafer with the mask on a contact aligner EVG6200 ∞ in the hard+vacuum mode, and then apply UV exposure with a dose of 9 mJ/cm^2. The photoresist is developed in AZ726MIF developer (2.38% TMAH in H_2O) for 20 sec before De-Ionized water clean and N_2 drying. To

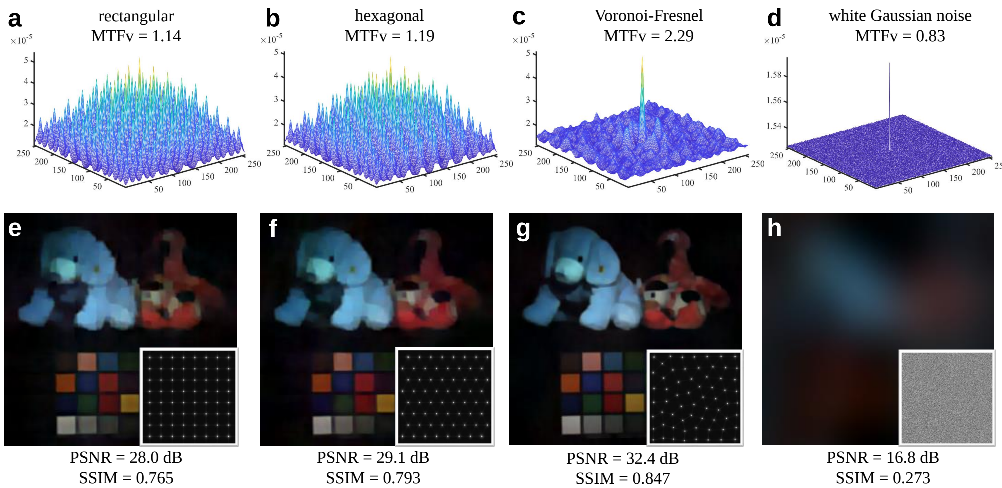

Fig. S4. Image performance comparison with the auto-correlation metric. From (a) to (d) are the auto-correlation functions of the rectangular-grid PSF, hexagonal-grid PSF, optimized Voronoi-Fresnel PSF, and a white Gaussian noise PSF. From (e) to (h) are the reconstructed images for the PSFs shown above. The image quality is evaluated by PSNR and SSIM. The respective PSFs are shown in the bottom right corner insets.

transfer the patterns from the photoresist to the Cr layer, we wet-etch the wafer with Cr etchant (mixtures of $HClO_4$ and $(NH_4)_2[Ce(NO_3)_6]$) for 1 min and remove the residual photoresist with Acetone. In the RIE step, SiO_2 in the wafer is etched by plasma of 15 sccm CHF_3 and 5 sccm O_2 at 10 °C. The etching depths are 75 nm, 150 nm, 300 nm and 600 nm respectively for a design wavelength at 550 nm. An additional Cr layer is deposited and etched to serve as the aperture to preserve the shift-invariance of PSFs.

8. PROTOTYPE RESULTS

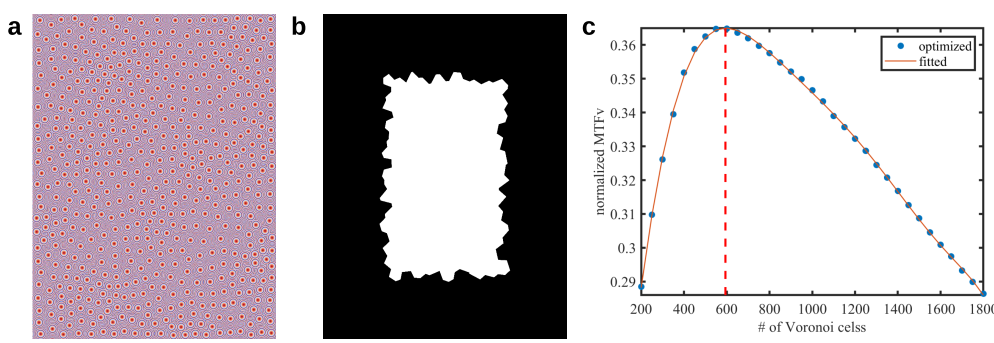

Fig. S5. Design of the prototype Voronoi-Fresnel lensless camera. (a) Optimized full-area phase profile. (b) An aperture to maintain the PSF shift-invariance within $\pm 20^\circ$ field-of-view. (c) Best number of Voronoi-Fresnel cells in the parameter sweep step. The optimal number of cells is 594 without the aperture. Effective cells after applying the aperture is 424.

The prototype Voronoi-Fresnel phase is designed at 550 nm for 2π modulation. Considering the fabrication resolution and sensor pixel size, we fix the upsampling ratio of $3\times$ for the optical element, i.e., 1.15 μm, which is well controlled by our fabrication method. The sensor has 1440×1080 pixels, and the Voronoi-Fresnel phase has 4320×3240 pixels. The optimized full-area phase profile is shown in Fig. S5a. Our sensor has a field-of-view of $\pm 20^\circ$ in the horizontal direction, and $\pm 15^\circ$ in the vertical direction, so we design an aperture (Fig. S5b) that excludes the cells outside of the field-of-view. The optimal number of Voronoi-Fresnel cells from the optimization is 594, as shown in Fig. S5c after the parameter sweep step. The effective cells within the field-of-view is 424, which is about 71.4% of the total number.

The final presentation of the image requires some necessary pre-processing and post-processing for better image presentation. We first demosaic the raw sensor data into color image data according to the sensor Bayer layout. Before image reconstruction, we normalize the blurred data by its norm in each color channel. For all the experimental results, the weights are all set to $\mu = 1e^{-7}$, and $\rho = 1e^{-5}$. In the post-processing, we use a simple gray world algorithm in MATLAB (`chromadapt`) for automatic white balancing in the linear color space. The illuminant is estimated by excluding 10 percentile of pixels. Finally gamma correction ($\gamma = 1.25$) is applied.

We present additional characteristic test results for the prototype in Fig. S6. First, we evaluate the geometry distortion using a checkerboard target. As shown in Fig. S6a, the geometry is restored very well across the $\pm 20^\circ \times \pm 15^\circ$ field-of-view. On the border regions, there are residual chromatic artifacts, however. This may arise from two factors. One is the off-axis aberrations of the base Fresnel phase, and the other is the difference in PSFs from on-axis to off-axis leading to the drop in reconstruction quality. Second, we evaluate the color fidelity using a color checker target in Fig. S6b. Despite the residual color artifacts in the border regions, the overall color fidelity is retrieved from the raw data. Potential improvement could be a more advanced white balancing algorithm other than the simple gray world algorithm used here. Last, we evaluate the spatial resolution variation using a Siemens star target in Fig. S6c. A uniform spatial resolution change is observed from the reconstructed image, demonstrating again that the MTF optimization is effective with uniform response.

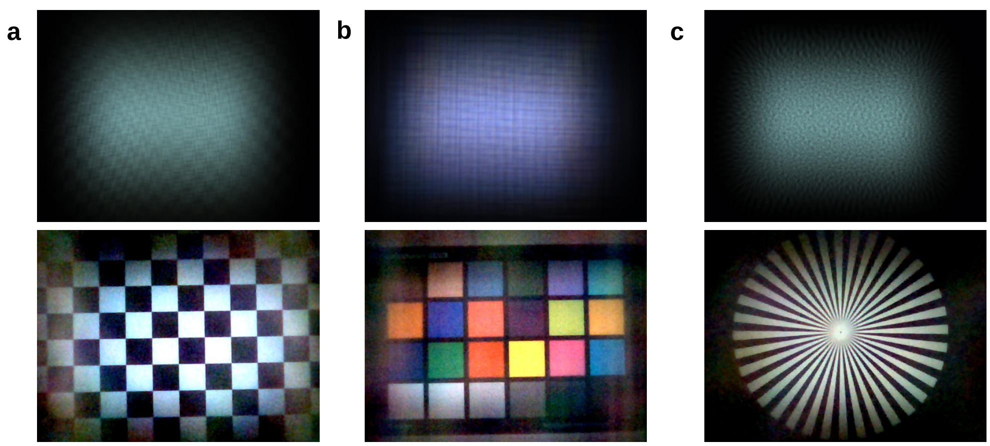

Fig. S6. Prototype characteristic test results. The reconstruction of a checkerboard image in (a) shows little geometry distortion. The color checker in (b) indicates good color recovery, although residual color artifacts exist in the image border. (c) A Siemens star image shows uniform resolution preservation in all directions.

Color reproduction remains a challenge in the current prototype. To analyze the color fidelity, we extract the color patches from the reconstructed image of the color checker (Fig. S6a), and tile them side by side according to their original orders in the color checker to synthesize an image in Fig. S7a. Each extracted patch has 50×50 pixels. As a reference, we create a reference color patch image from the corresponding true RGB values with the same size, as shown in Fig. S7b, with their indices labeled. The color fidelity is measured as the color difference dE using the CIEDE2000 standard [13]. We can visualize the pixel-wise color difference with the dE map in Fig. S7c. Since the dE map varies within each color patch due to the noisy reconstruction results, we take the average value in each color patch as the reconstructed color values. The color difference dE values are then calculated and plotted in Fig. S7d. The maximum dE (largest color difference) is 31.6 at index 19, which is the "White" patch, and the the minimum dE (smallest color difference) is 3.7 at index 10, which is the "Purple" patch. This also agrees with the visual perception.

We attribute the quality of color fidelity in our prototype mainly to the residual chromatic aberrations, and the simple white balancing algorithm we currently use. Since the base Fresnel

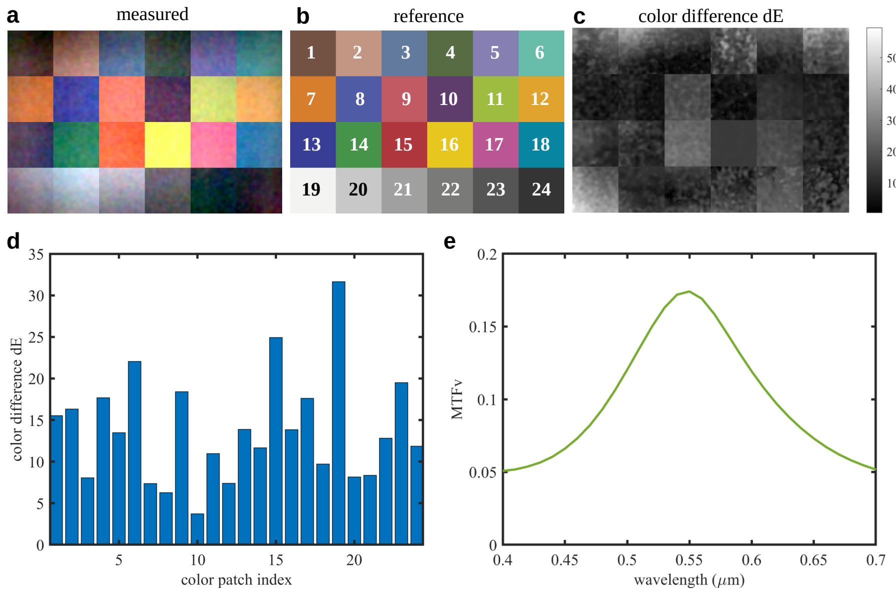

Fig. S7. Color difference analysis. (a) Extracted color patches from the reconstructed color checker image. (b) Reference color patches with true RGB values. Indices are labeled for each patch. (c) Color different dE map calculated from (a) and (b) using the CIEDE2000 standard. (d) Color difference dE values calculated with averaged RGB values in each color patch. (e) MTFv varies with wavelength in the visible band for the Voronoi-Fresnel phase used in the prototype.

phase is static at one single wavelength (550 nm in our prototype), the PSF geometry changes if the illumination wavelength is different, and hence the MTFv is also a function of wavelength. We optimize for the spectral integral of the MTF, but not the cross differences between wavelengths. We evaluate this chromatic effect by plotting the MTFv with respect to wavelength across the visible band from 400 nm to 700 nm, as shown in Fig. S7e. The MTFv drops around 71% for both the short and long ends of the wavelength range, compared with the design wavelength at 550 nm. This could be mitigated by using a base phase function that is optimized achromatic, instead of the static Fresnel phase we currently use. Further improvement can be made to use more advanced white balance algorithms to improve the color fidelity.

Although we use the panchromatic PSF in the visible band in our design, there are still residual chromatic aberrations in the current result. We attribute mainly two important factors for the residual chromatic aberrations. First, the fixed base Fresnel phase in each cell is inherently dispersive, which is a fundamental property of all the diffractive optical elements. It could be possible to use an optimized achromatic base phase in each cell, which would require additional efforts. Second, the simple image reconstruction algorithm we adopt here does not account for chromatic aberration correction. The total variation regularization only takes care of spatial structures in the image. It would then be more advantageous to employ neural networks for the reconstruction to alleviate chromatic aberrations.

To evaluate the resolution of the prototype, we capture a cross-hair target, and reconstruct the final image. The raw data is shown in Fig. S8a, and the reconstructed image is shown in Fig. S8b. We can plot the cross-sections in the horizontal and vertical directions. The spot diameter is measured around the lines as the intensity first reaches zero. Since we know the pixel size, we can calculate the physical diameters. In the horizontal direction, the diameters are measured as 17.25 μm, 17.25 μm, 20.7 μm for the red, green and blue channels respectively. The average diameter is 18.4 μm. In the vertical direction, the diameters are 20.7 μm, 17.25 μm, 24.15 μm for the red, green and blue channels respectively. The average diameter is 20.7 μm. To compare with the theoretical value, we calculate the effective diameter of all the Voronoi-Fresnel cells. The ideal spot diameter is 15.7 μm, so the resolution of the prototype is indeed close to the theoretical value.

Additionally, we present more example results in Fig. S9. Again, the results in Fig. S9a and S9b are captured from self-illuminating images displayed on a monitor, and Fig. S9c-d are real objects with ambient illumination. Details in the objects are well preserved in both cases.

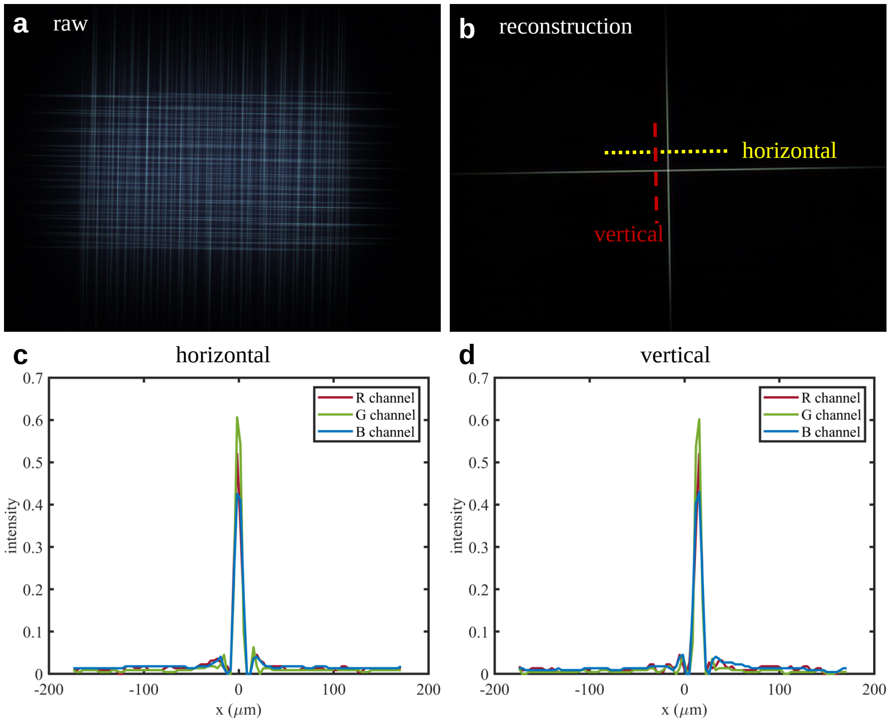

Fig. S8. Resolution measurement. (a) Raw capture of the cross-hair target. (b) Reconstructed image of the cross-hair target. The yellow dotted line indicates where the horizontal cross-section is taken, and the red dash line indicates where the vertical cross-section is taken. (c) Horizontal cross-section of (b). (d) Vertical cross-section of (b). The spot diameter is measured around the lines as the intensity first reaches zero.

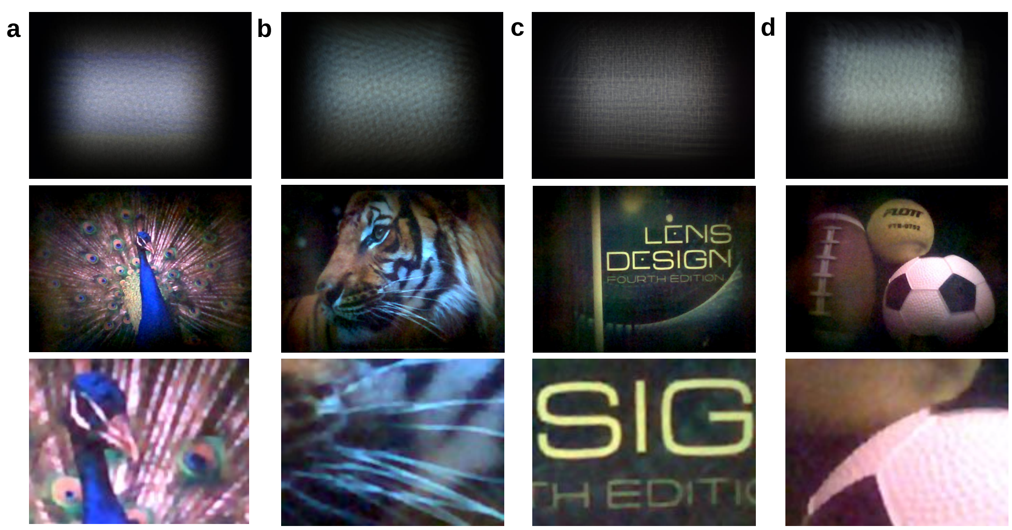

Fig. S9. Additional prototype results. (a) and (b) show results for self-illuminating images displayed on a computer monitor. (c) and (d) show results for real objects with ambient illumination. Top row are the captured raw data; middle row are reconstructed images; and bottom row are zoom-in details.

REFERENCES

1. R. T. Brigantic, M. C. Roggemann, K. W. Bauer, and B. M. Welsh, “Image-quality metrics for characterizing adaptive optics system performance,” Appl. Opt. **36**, 6583–6593 (1997).

2. L. C. Roberts and C. R. Neyman, “Characterization of the aeos adaptive optics system1,” Publ. Astron. Soc. Pac. P. **114**, 1260 (2002).
3. D.-M. Yan, J.-W. Guo, B. Wang, X.-P. Zhang, and P. Wonka, “A survey of blue-noise sampling and its applications,” J Comput. Sci. Technol. **30**, 439–452 (2015).
4. F. De Goes, K. Breeden, V. Ostromoukhov, and M. Desbrun, “Blue noise through optimal transport,” ACM Trans. Graph. **31**, 1–11 (2012).
5. Q. Du, V. Faber, and M. Gunzburger, “Centroidal Voronoi tessellations: Applications and algorithms,” SIAM Rev. Soc. Ind. Appl. Math. **41**, 637–676 (1999).
6. D.-M. Yan, K. Wang, B. Lévy, and L. Alonso, “Computing 2D periodic centroidal voronoi tessellation,” in *2011 Eighth International Symposium on Voronoi Diagrams in Science and Engineering,* (IEEE, 2011), pp. 177–184.
7. S. Lloyd, “Least squares quantization in PCM,” IEEE Trans. Inf. Theory **28**, 129–137 (1982).
8. N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, and L. Waller, “DiffuserCam: lensless single-exposure 3D imaging,” Optica **5**, 1–9 (2018).
9. V. Boominathan, J. Adams, J. Robinson, and A. Veeraraghavan, “Phlatcam: Designed phase-mask based thin lensless camera,” IEEE Trans. Pattern Anal. Mach. Intell. (2020).
10. F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar, “Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum,” Tech. rep. (2008).
11. Y. Li, Q. Fu, and W. Heidrich, “Multispectral illumination estimation using deep unrolling network,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision,* (2021), pp. 2672–2681.
12. K. Yanny, N. Antipa, W. Liberti, S. Dehaeck, K. Monakhova, F. L. Liu, K. Shen, R. Ng, and L. Waller, “Miniscope3D: optimized single-shot miniature 3d fluorescence microscopy,” Light. Sci. Appl. **9**, 1–13 (2020).
13. G. Sharma, W. Wu, and E. N. Dalal, “The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations,” Color. Res. Appl. **30**, 21–30 (2005).