

Supplementary Materials for *Fovea Stacking: Imaging with Dynamic Localized Aberration Correction*

SHI MAO, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

YOGESHWAR NATH MISHRA, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

WOLFGANG HEIDRICH, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

1 DPP SURFACE MODELING WITH ZERNIKE POLYNOMIALS

The optical path difference (OPD) introduced by the deformable phase plate (DPP) is parameterized as Zernike polynomials.

$$D(\rho, \varphi) = \sum_k^K w_k Z_k(\rho, \varphi), \quad (1)$$

which are defined as

$$Z_n^m(\rho, \varphi) = \begin{cases} c(n, m) R_n^m(\rho) \cos(m\varphi) & m \geq 0 \\ c(n, m) R_n^m(\rho) \sin(-m\varphi) & m < 0 \end{cases}, \quad (2)$$

where n is the radial order and m is the azimuth order, with $n \geq |m| \geq 0$. According to OSA standard indices:

$$k = \frac{n(n+1) + m}{2} \quad (3)$$

The radial polynomials R_n^m are defined as:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-|m|}{2}} (-1)^k \binom{n-k}{k} \binom{\frac{n-|m|}{2}}{\frac{n-|m|}{2}-k} \rho^{n-2k} \quad (4)$$

The coefficients $c(n, m)$ is a normalization factor to ensure the variance on the unit circle area equals 1, and is calculated as:

$$c(n, m) = \begin{cases} \sqrt{n+1} & m = 0 \\ \sqrt{2(n+1)} & \text{otherwise} \end{cases} \quad (5)$$

As the Zernike polynomials is defined in a normalized polar coordinate, to relate it to the cartesian coordinate with physical size, the parameters is converted:

$$\begin{cases} \rho = \frac{\sqrt{x^2 + y^2}}{R} = \frac{r}{R} \\ \varphi = \arctan\left(\frac{y}{x}\right) \end{cases} \quad (6)$$

where R is the radius of the DPP aperture, and we denotes $r = \sqrt{x^2 + y^2}$.

1.1 Derivative of Zernike Polynomials

To refract the ray according to the introduced optical path difference, the derivatives of the OPD with respect to the Cartesian coordinates are required:

$$\frac{\partial D}{\partial x} = \frac{\partial D}{\partial \rho} \frac{\partial \rho}{\partial x} + \frac{\partial D}{\partial \varphi} \frac{\partial \varphi}{\partial x} \quad (7)$$

$$\frac{\partial D}{\partial y} = \frac{\partial D}{\partial \rho} \frac{\partial \rho}{\partial y} + \frac{\partial D}{\partial \varphi} \frac{\partial \varphi}{\partial y} \quad (8)$$

$$(9)$$

Where the Jacobian for transforming the cartesian coordinate to the normalized polar coordinates is calculated as:

$$\begin{bmatrix} \frac{\partial \rho}{\partial x} & \frac{\partial \varphi}{\partial x} \\ \frac{\partial \rho}{\partial y} & \frac{\partial \varphi}{\partial y} \end{bmatrix} = \frac{1}{R} \begin{bmatrix} \frac{x}{r} & -\frac{y}{r^2} \\ \frac{y}{r} & \frac{x}{r^2} \end{bmatrix} \quad (10)$$

The derivative of the OPD D w.r.t. normalized polar coordinates is a linear combination of its component Zernike polynomials:

$$\frac{D(\rho, \varphi)}{\partial \rho} = \sum_k^K w_k \frac{Z_n^m(\rho, \varphi)}{\partial \rho} \quad (11)$$

$$\frac{D(\rho, \varphi)}{\partial \varphi} = \sum_k^K w_k \frac{Z_n^m(\rho, \varphi)}{\partial \varphi} \quad (12)$$

The derivative of Zernike Polynomials is well separated for ρ and φ and given by:

$$\frac{Z_n^m(\rho, \varphi)}{\partial \rho} = \begin{cases} c(n, m) \frac{\partial R_n^m(\rho)}{\partial \rho} \cos(m\varphi) & m \geq 0 \\ c(n, m) \frac{\partial R_n^m(\rho)}{\partial \rho} \sin(-m\varphi) & m < 0 \end{cases} \quad (13)$$

$$\frac{Z_n^m(\rho, \varphi)}{\partial \varphi} = \begin{cases} -c(n, m) R_n^m(\rho) m \sin(m\varphi) & m \geq 0 \\ -c(n, m) R_n^m(\rho) m \cos(m\varphi) & m < 0 \end{cases} \quad (14)$$

$$(15)$$

Where the partial derivative of φ for radial polynomials is:

$$\frac{\partial R_n^m(\rho)}{\partial \rho} = \sum_{k=0}^{\frac{n-|m|}{2}} (-1)^k \binom{n-k}{k} \binom{\frac{n-|m|}{2}}{\frac{n-|m|}{2}-k} (n-2k) \rho^{n-2k-1} \quad (16)$$

1.2 Phase Plate Model V.S. Refractive Surfaces Model

We validate that the DPP can be sufficiently modeled as a thin plate that introduce phase change, compare to a more complex refractive modeling that perform ray tracing through refractive surfaces. As shown in Fig 1, the physical model of DPP is composed of a optical fluidic chamber, filled with tridecane ($n = 1.4256$), and a fused silica substrate [Rajaeipour et al. 2021]. A refractive surfaces modeling (Fig 1b) is compared with the implemented phase plate modeling

Authors' addresses: Shi Mao, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, shi.mao@kaust.edu.sa; Yogeshwar Nath Mishra, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, mishrayn@iitj.ac.in; Wolfgang Heidrich, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, wolfgang.heidrich@kaust.edu.sa.

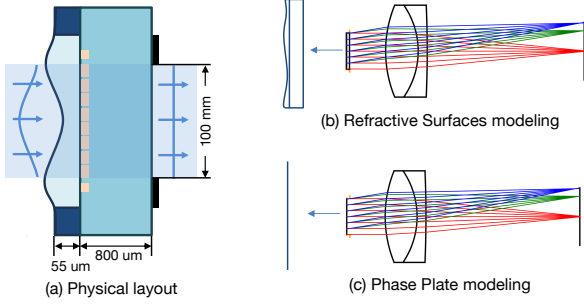


Fig. 1. Phase plate model v.s. refractive surfaces Model. (a) The Physical Structure of DPP [Rajaeipour et al. 2021]. The optical fluidic chamber is $55\mu\text{m}$ thick, while the fused-silica substrate is $800\mu\text{m}$ thick. (b) The refractive modeling of DPP, corresponding to the phase structure. (c) The phase modeling of DPP, which is modeled as a thin plate that introduce optical path difference.

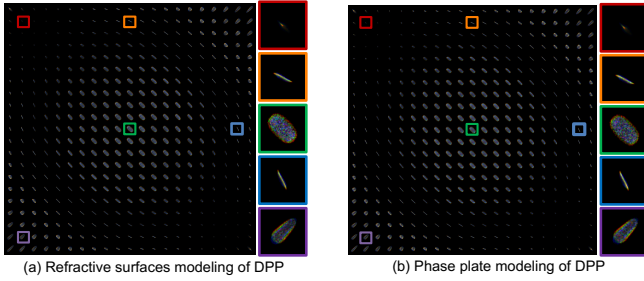


Fig. 2. Ray tracing comparison of refractive surfaces Model (a) and phase plate model (b).

(Fig 1c). Notice the optical fluidic surface thickness H is related to the introduced OPD D by: $D = (n - 1)H$.

With the same optimized DPP deformation pattern, we perform ray-tracing through different models and visualized the PSF in Fig. 2. As there is no identifiable difference between this two simulation results, we regard the phase plate model as a good approximation of the underlying refraction process, and choose the phase plate model for its simplicity and computational efficiency.

2 LOCALIZED ABERRATION CORRECTION

2.1 Additional Results

For a single DPP deformation optimization at oblique angle ϕ , sagittal and tangential MTF were analyzed using ray tracing for different wavelengths shown in Fig. 3. The sampled angles align with Fig. 4 in the main paper and provide further analysis. Red, green, and blue channels were traced using wavelengths of 640 nm, 525 nm, and 470 nm, respectively, matching the prototype system's sensor color response [FILR 2017]. Compared to the unoptimized initial state, defocus-only optimization improved sagittal MTF while tangential MTF remained low. Full Zernike optimization corrected higher-order aberrations, leaving only chromatic aberration, which single refractive elements cannot effectively address.

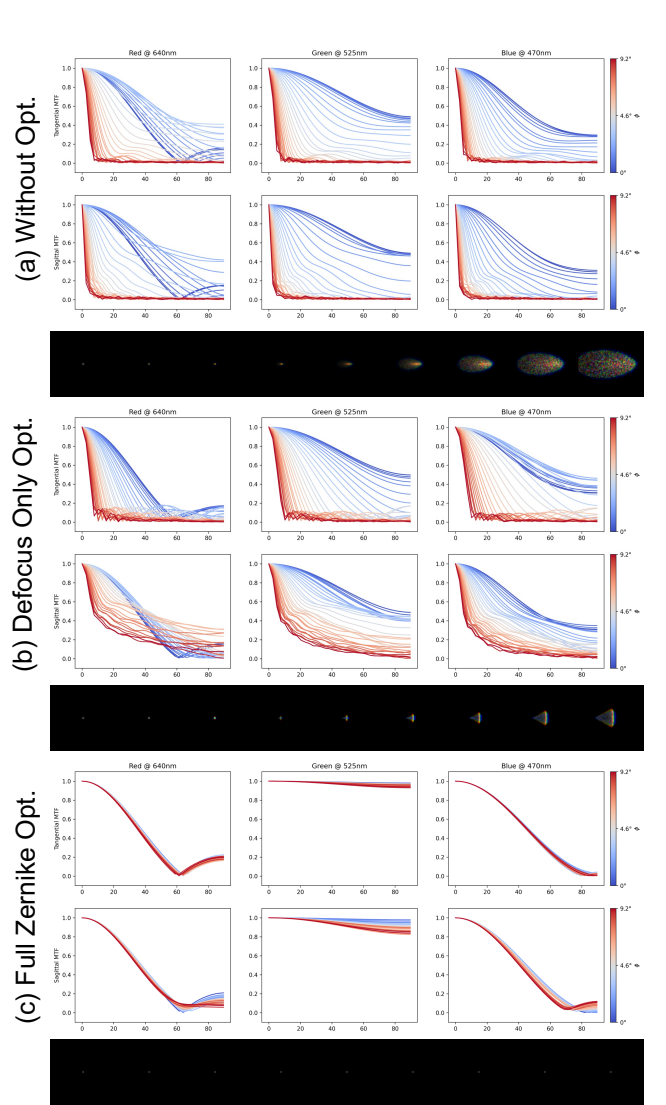


Fig. 3. Detailed analysis on MTF and PSF for various oblique angle ϕ . (a) Initial system without optimization. (b) optimization of the defocus Zernike parameter only. (c) Full Zernike parameter optimization up to 4-th order.

The fovea in the radial-angular diagram is approximated as a rectangular region (Fig 4). To determine the images needed for full coverage, we iterate radially over the sampled 32 optimized patterns, identifying the next largest overlapping pattern index above a sharpness threshold. The replication count for each selected pattern is then estimated by dividing 2π by its angular size. When $\rho > \sqrt{2}/2$, we estimate corner coverage by dividing $\pi/2 - 2 \arccos(1/\rho\sqrt{2})$ by the angular size and multiplying by 4 to obtain the full coverage replication count, assuming no radial symmetry for extreme ρ values.

With a $5.5\mu\text{m}$ pixel size, the sensor has approximately 182 pixels/mm, resulting in a Nyquist frequency of 91 lp/mm. Given the

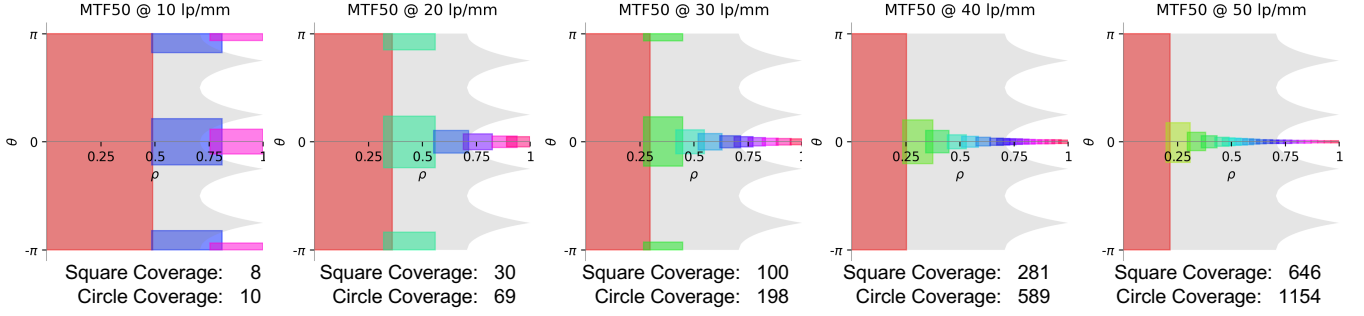


Fig. 4. Fovea coverage above different sharpness thresholds, ranging from 10 lp/mm to 50 lp/mm.

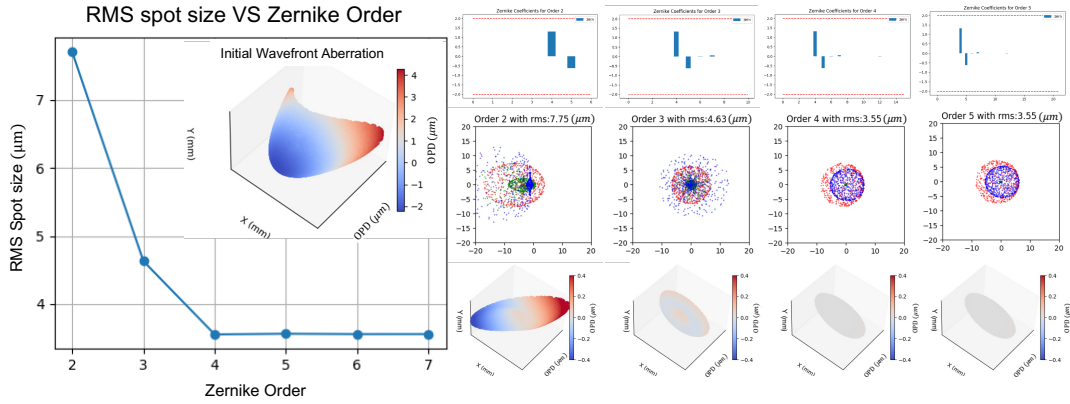


Fig. 5. Required Zernike orders to correct for off-axis aberrations. Left: The optimized PSF RMS spot size in μm compared to the number of allowable Zernike orders on DPP, superimposed by the initial wavefront aberration at an oblique angle of 6.5° . Right: from top to bottom are zernike coefficients, spot diagram, and the remaining wavefront aberrations on the green channel after correction with corresponding order-limited DPP deformation.

Bayer pattern, sharp imaging requires 45 lp/mm resolution. The lowest 10 lp/mm in Fig. 4 corresponds to a PSF spanning 9×9 raw sensor pixels. Notice this threshold serves as the lowest (instead of an average) quality over the entire image. Increasing image quality rapidly increases the number of required images to cover the full sensor area.

2.2 Analysis on the Required Zernike Orders

The optical system determines the wavefront aberrations at the exit pupil for different sensor regions. In this work, using a simple achromatic main lens (Thorlab’s AC254-050A), aberrations are confined to lower radial orders. To identify the necessary Zernike orders for off-axis correction, we optimized DPP patterns limited to different radial orders at a 6.5° oblique angle under hyperfocal conditions. As shown in Fig. 5, off-axis aberrations are effectively corrected with Zernike terms above 4th order, and a device supporting up to 7th order is sufficient. This conclusion is system-dependent; other optical systems may require more complex deformation.

3 DATASET COLLECTION FOR DPP CONTROL MODELS

To collect control voltage - Zernike coefficient pairs, random sampling on the “voltage space” rarely generates large amplitude deformation, which is useful for large off-axis aberration correction. This is because large amplitude deformation requires a joint activation of a local spatial region of electrodes (jointly reaching its maximum voltage to pull down the membrane with larger electrostatics force), which is rarely sampled if electrodes voltages are sampled independently. To enable the sampling of such large deformation, we perform the sampling in the “Zernike coefficient space” within the operational range and obtain the control voltage through constrained optimization using Eq. 8 in the main paper. The precise deformation is then measured by the wavefront sensor.

4 INTERPOLATION FOR DPP CONTROL

For applications like fovea tracking, a fast change of the control signal is required. In this section, we discuss how to pre-compute optimized patterns specialize on a grid of locations across the image, and acquire control signal for any locations through interpolation.

We first sampled $(2^n + 1) \times (2^n + 1)$ corner-aligned grids across the image. Then the interpolation is performed in ideal Zernike space,

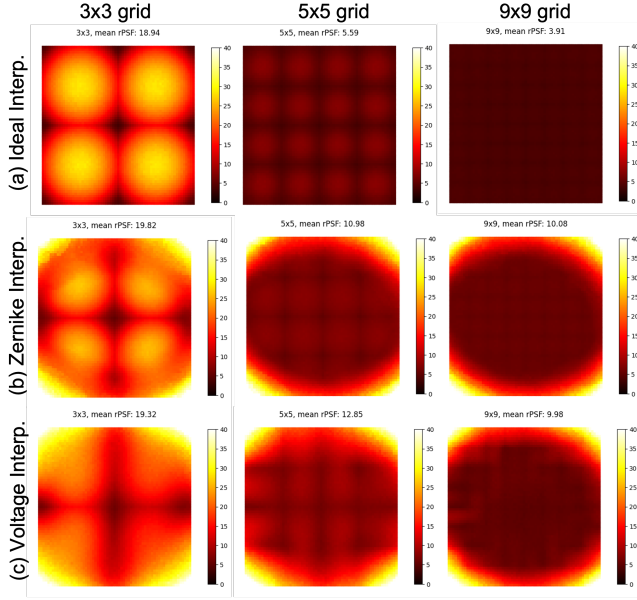


Fig. 6. The imaging quality using interpolated control signals in different spaces. (a) Ideally, the interpolation is performed in Zernike parameters. (b) When hardware response is considered, the interpolated Zernike parameters is applied to the device and the predicted response from the proxy model is used. (c) The interpolation is performed in the voltage space.

physical Zernike space, and voltage space respectively. For each interpolated DPP deformation pattern, the PSF of its corresponding location is calculated through ray tracing, as shown in Fig 6. For ideal Zernike interpolation, no underlying physics is considered, and the DPP's Zernike parameter is interpolated linearly. While the physical Zernike interpolation considered the DPP's hardware response by apply the control strategy discussed in Sec. 4.4 in the main paper, and use the decoded Zernike parameters to perform ray-tracing. For physical voltage interpolation, the voltage for each pre-computed grid points are computed and interpolated, a decoder is applied to obtain the recovered Zernike parameters for ray-tracing. When the DPP control model is considered, the image quality on the edges is limited, this is considered as a limitation of the device.

Fig. 7 shows that finer grids reduce the average RMS spot size across the image, but improvements become marginal beyond a 9×9 grid for all interpolation methods, even for ideal interpolation. Therefore, optimizing a 9×9 grid is considered sufficient for linear interpolation based control. For real-world applications, the voltage interpolation is preferred over Zernike interpolation because no additional computation is required before sending the signal to the DPP device, thus improve the real-time performance.

5 IMAGE FUSION METHODS

The sharpness-based fusion algorithm is presented in Algorithm 1. Since the DPP's Zernike vertical and horizontal tilt terms are constrained to zero during optimization, only minor misalignment is expected in the foveated image stack. Consequently, we perform fusion without prior alignment. Additionally, severe blurring in the image

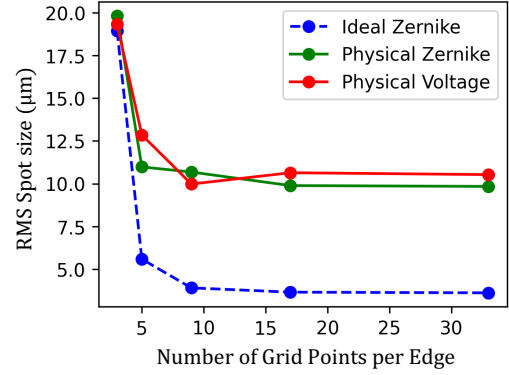


Fig. 7. Image quality under different levels of interpolation.

ALGORITHM 1: Fusion algorithm for Fovea Stacking

Input: Images I_1, I_2, \dots, I_N of size $H \times W \times 3$

Output: Fusion image O of size $H \times W \times 3$

for $k \leftarrow 1$ **to** N **do**

$G_k \leftarrow$ Convert I_k to grayscale;

$B_k \leftarrow$ GaussianBlur(G_k , kernel size k_{gauss});

$L_k \leftarrow |\text{Laplacian}(B_k, \text{kernel size } k_{\text{lap}})|$;

$S_k \leftarrow$ GaussianBlur(L_k , kernel size k_{blur});

 Store S_k in sharpness_list;

end

Stack all S_k into a 3D array S of shape $N \times H \times W$;

Stack all I_k into a 4D array I of shape $N \times H \times W \times 3$;

for each pixel location (i, j) **do**

 Enhance sharpness:

$S[k, i, j] \leftarrow (S[k, i, j] / \max_{l=1}^N (I[l, i, j]))^{k_{\text{sharp}}}$;

 Normalize weights: $w_k \leftarrow S[k, i, j] / \sum_{l=1}^N S[l, i, j]$;

$O[i, j, :] \leftarrow \sum_{k=1}^N w_k \cdot I[k, i, j, :]$;

end

return O ;

stack hinders the identification of sufficient corresponding points for alignment. For all experiments with captured 2048×2048 images, the fusion algorithm uses parameters $k_{\text{gauss}} = k_{\text{lap}} = 19$, $k_{\text{blur}} = 127$, and $k_{\text{sharp}} = 10$. The parameter k_{sharp} balances mean/max fusion. With $k_{\text{sharp}} = 1$, fusion softly blends images via sharpness-weighted averaging. As $k_{\text{sharp}} \rightarrow \infty$, fusion hard blends by selecting the maximum sharpness value across the stack.

Fig. 8 visualizes the sharpness of individual and fused images obtained via sharpness-based fusion. Images 3, 7, and 11 focus on near, middle, and far distances, respectively. The calculated sharpness effectively identifies the sharp regions in each image, and the maximum sharpness index reflects the depth ordering, with smaller values indicating closer depths.

REFERENCES

- FILR. 2017. Imaging performance specification for FLIR Grasshopper3 cameras. <https://www.apostar.com.tw/data/10000/upload/ProductFiles/TeledyneFLIRIIS/GS3-U3/GS3-U3-EMVA-Imaging.pdf> Accessed: May 19, 2025.

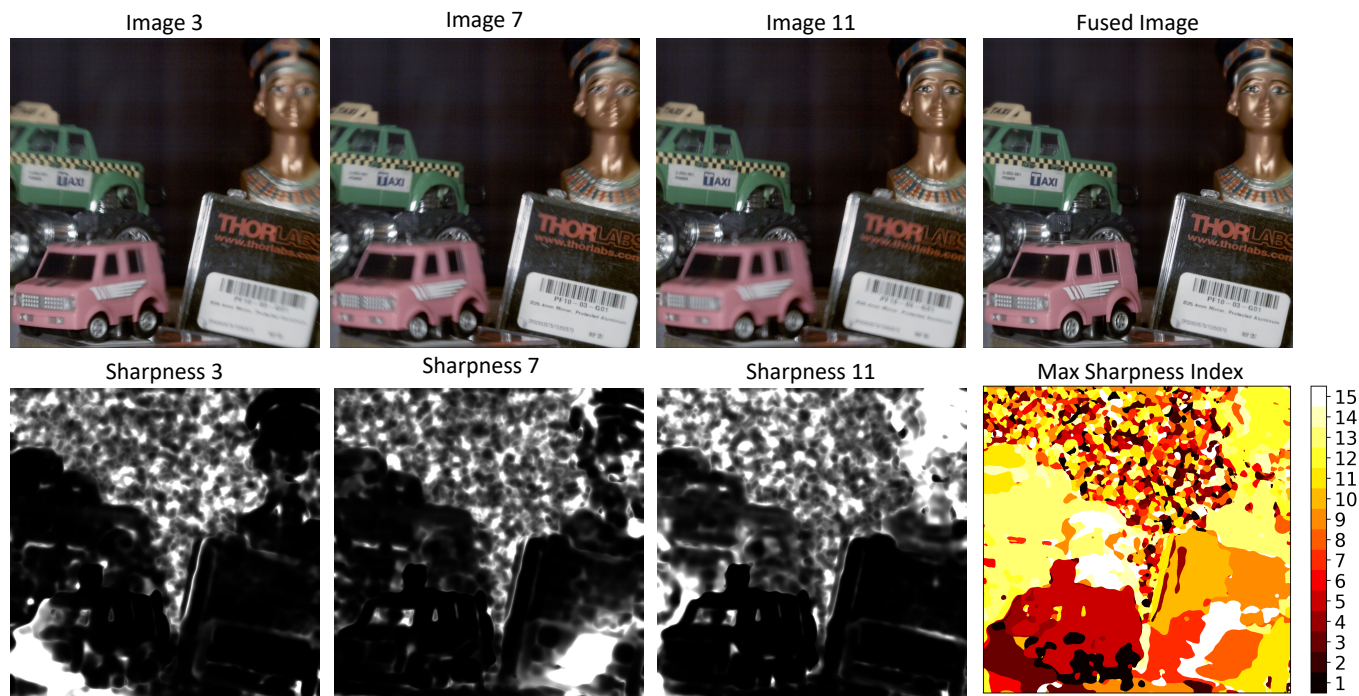


Fig. 8. Sharpness based fusion on extended depth of field application.