Supplementary Material: The Diffractive Achromat Full Spectrum Computational Imaging with Diffractive Optics

Yifan Peng[2,1] Qiang Fu[1] Felix Heide[2,1] Wolfgang Heidrich[1,2]
[1] King Abdullah University of Science and Technology [2] The University of British Columbia

1 Generating Regular Fresnel Lens

As has been discussed in main text, we use a designed initial guess derived from a mixture of a series of Fresnel lenses' profiles for different wavelengths. Without loss of generality, we take an amplitude-type Fresnel zone-plate (FZP) as the instance for principle illustration. Leveraging scalar diffraction theory, the complex amplitude of a point at the image plane is the sum-up of all complex amplitudes derived from all sub-waves emitted from ring-zones of a FZP. To obtain constructive interference at the image plane, the zones should switch from the opaque to transparent at radius

$$R_n = \sqrt{n\lambda f + \frac{n^2\lambda^2}{4}}, \tag{1}$$

where n is the integer for the $n^{\mathbf{th}}$ ring, λ is the wavelength of the light the zone plate is meant to focus, and f is the distance from the center of the zone plate to the focal point. In this case, each two neighbouring ring-zones holds the optical path difference of half wavelength (corresponding to the modulation of 2π), such that they contribute positive and negative amplitude to the spatial point, respectively. Those negative components are blocked.

With respect to the phase modulation, we expect the incident light to travel through the substrate medium with different thickness, such that create the required optical path difference without blocking the light. That is to say, we can use the same math described above to derive the radius for each ring zone, and then assign them with different physical heights, corresponding to the phase delay they modulate on the incident light rays. These height change can be designed continuous or discrete subject to the fabrication method one intend to apply.

2 Fabrication Details

As has been mentioned in main text, the DOEs are fabricated by repeatedly applying the lithography technique followed by a reactive ion etching (RIE) step. Each round of fabrication results in a binary profile on the substrate. Therefore, by repeating the process for N times, a 2^N-level phase plate could be fabricated.

In our work, we use 4 inch UV-grade fused silica wafers as the substrates. Four iterations of lithography and etching are repeated to produce 16-level phase zone plates. The fabrication process is shown in Fig. 1.

In our designs, the principle wavelength is $\lambda_0 = 550nm$. A total 2π phase delay at this wavelength corresponds to a depth of $d = \lambda/(n_\lambda - 1) = 1195nm$, where $n_\lambda = 1.46008$ is the refractive index of the wafer at the principle wavelength. Therefore, each level has a depth of $75nm$ in our designs. A final depth error of $\pm 10nm$ has been achieved. The microscope images of the fabricated DOE plates are shown in Fig. 2. We see that a Fresnel lens has more regular periodical ring distribution, while a diffractive achromat has more non-periodical ring distribution, corresponding to the optimized surface profile.

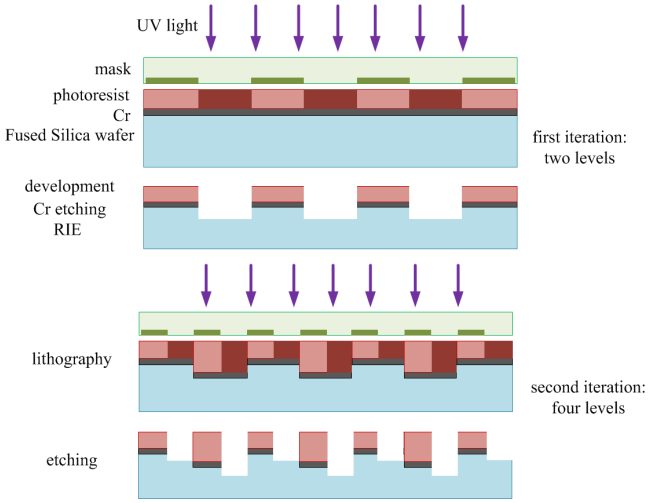

Figure 1: *Fabrication of DOEs consists of two major steps— lithography and etching. In the lithography step, patterns are transferred from the mask to photoresist on the wafer through exposure to UV light. In the etching step, the transferred patterns are converted to a binary profile on the wafer by reactive ion beam bombardment.*

3 Reconstruction Algorithms

As has been mentioned in main text, we seek to apply a two-step fast deconvolution scheme to recovery the images. The reconstruction pipeline is illustrated in Fig. 3.

1st step We implement the first step of deconvolution at a downsampled scale, say half size of the width and height, to solve Eq.13 in the paper, which is

$$\mathbf{i_d} = \mathbf{argmin_i}\, \mu\|\mathbf{b} - \mathbf{Ki}\|_2^2 + \beta\|\mathbf{Di}\|_2^2, \tag{2}$$

where $\mathbf{D}$ here is the first-order derivative filters matrix, μ, β are the weights for each term. As has been mentioned, Eq.2 is a closed-form optimization such that we have

$$\begin{aligned}
&\mathbf{i}_{\text{opt}} = \mathbf{argmin_i}\, \mu\|\mathbf{b} - \mathbf{Ki}\|_2^2 + \beta\|\mathbf{Di}\|_2^2 \\
&\Leftrightarrow 2\mu\left(\mathbf{K}^T\mathbf{K}\mathbf{i}_{\text{opt}} - \mathbf{K}^T\mathbf{b}\right) + 2\beta\mathbf{D}^T\mathbf{D}\mathbf{i}_{\text{opt}} \stackrel{!}{=} 0 \\
&\Leftrightarrow \mu\mathbf{K}^T\mathbf{K}\mathbf{i}_{\text{opt}} + \beta\mathbf{D}^T\mathbf{D}\mathbf{i}_{\text{opt}} = \mu\mathbf{K}^T\mathbf{b} \\
&\Leftrightarrow \left(\mu\mathbf{K}^T\mathbf{K} + \beta\mathbf{D}^T\mathbf{D}\right)\mathbf{i}_{\text{opt}} = \mu\mathbf{K}^T\mathbf{b} \\
&\Leftrightarrow \mathbf{i}_{\text{opt}} = \mathcal{F}^{-1}\left(\frac{\mu\mathcal{F}(\mathbf{K})^*\mathcal{F}(\mathbf{b})}{\mu\mathcal{F}(\mathbf{K})^*\mathcal{F}(\mathbf{K}) + \beta\mathcal{F}(\mathbf{D})^*\mathcal{F}(\mathbf{D})}\right)
\end{aligned} \tag{3}$$

where $\mathcal{F}(\cdot)$ represents Fourier transform while $\mathcal{F}^{-1}(\cdot)$ represents

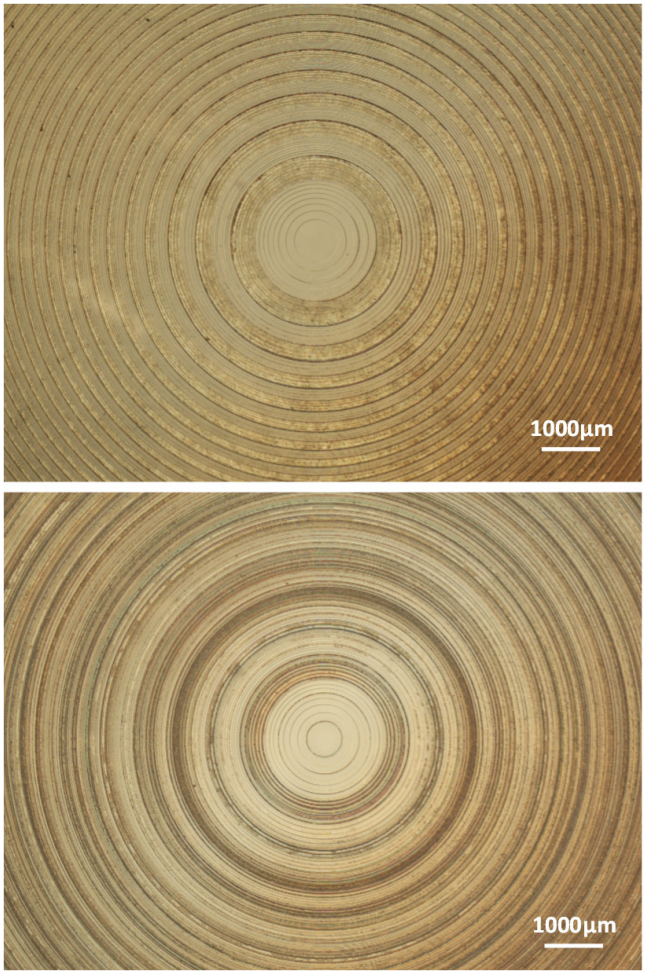

Figure 2: *Illustration of 5X microscope images of the fabricated Fresnel lens (top), and diffractive achromat (bottom). Note that the bright and dark changes indicate the physical height distribution.*

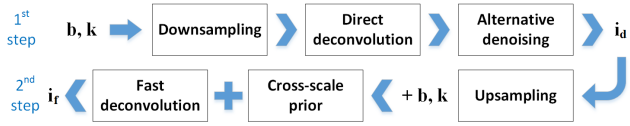

Figure 3: *Reconstruction of the images consists of two major steps: In the 1st step, the blurred image is downscaled, then deblurred and denoised at a downsampled scale; In the 2nd step, the processed image from the 1st step is upsampled to the full scale, and used in a cross-scale prior to assist the deblurring of the original image.*

the inverse Fourier transform. Here the superscript $*$ indicates the complex conjugate operation.

2nd step We implement the second step of deconvolution at the full scale, to solve the Eq.15 in the paper, which is

$$\mathbf{i_f} = \mathbf{argmin_i} \frac{\mu}{2}\|\mathbf{b}-\mathbf{Ki}\|_2^2 + \beta\|\mathbf{Di}\|_1 + \gamma\|\mathbf{Di}-\mathbf{Di_s}\|_1, \quad (4)$$

where $\mathbf{D}$ is the same derivative filters matrix as above, and $\mathbf{i}_s$ is the image upsampled from the one deconvolved in the 1st step. Adding the cross-scale prior results in a non-linear optimization. An alternative solution is to introduce slack variable for l_1 term. Specially, we formulate the proximal operators like the ones in [Boyd et al. 2011; Chambolle and Pock 2011], thus the l_1 deconvolution is handled as shrinkage operators, we use $\mathbf{p} = \mathbf{Di}$ as variable

$$\begin{aligned}\mathbf{prox}_{\theta\|\cdot\|_1}(\mathbf{p}) &= \max\left(1-\frac{\theta}{|\mathbf{p}|},0\right)\odot\mathbf{p}\\ \mathbf{prox}_{\theta\|\cdot-\alpha\|_1}(\mathbf{p}) &= \max\left(1-\frac{\theta}{|\mathbf{p}-\alpha|},0\right)\odot\mathbf{p}+\alpha\end{aligned} \quad (5)$$

where $\alpha = \mathbf{Di_s}$. The proximal operators for convex conjugates can be derived from [Boyd et al. 2011]. Next, we use a similar half-quadratic penalty scheme in [Krishnan and Fergus 2009].

Specifically, the Eq. 4 can be rewritten by introducing the variable $\mathbf{p} = \mathbf{Di}$ that has been denoted. Then, we have

$$\mathbf{i}_{\text{opt}} = \mathbf{argmin_i} \frac{\mu}{2}\|\mathbf{b}-\mathbf{Ki}\|_2^2+\beta\|\mathbf{p}\|_1+\gamma\|\mathbf{p}-\mathbf{Di_s}\|_1+\frac{\rho}{2}\|\mathbf{Di}-\mathbf{p}\|_2^2, \quad (6)$$

where ρ is a weight that shall be gradually increased during the optimization, such that once the weight ρ is becoming very large, the modified objective Eq. 6 is equivalent to the original objective we have defined in Eq. 4. We set the values namely integer powers of $\sqrt{2}$ between 1 and 256 in all our implementation.

One can then divide the objective into two subproblems and apply alternating methods to solve them iteratively. For the two quadratic terms, by fixing $\mathbf{p}$ we have a similar closed-form solution as in Eq. 7 can be derived as

$$\begin{aligned}\mathbf{i}_{\text{opt}} &= \mathbf{argmin_i} \frac{\mu}{2}\|\mathbf{b}-\mathbf{Ki}\|_2^2+\frac{\rho}{2}\|\mathbf{Di}-\mathbf{p}\|_2^2\\ \Leftrightarrow \mathbf{i}_{\text{opt}} &= \mathcal{F}^{-1}\left(\frac{\mu\mathcal{F}(\mathbf{K})^*\mathcal{F}(\mathbf{b})+\rho F\mathbf{D}^*\mathcal{F}(\mathbf{p})}{\mu\mathcal{F}(\mathbf{K})^*\mathcal{F}(\mathbf{K})+\rho\mathcal{F}(\mathbf{D})^*\mathcal{F}(\mathbf{D})}\right).\end{aligned} \quad (7)$$

Next we fix i, and optimize the two l_1 norm terms that contain $\mathbf{p}$, the proximal operators that have been defined in Eq. 5 turn them to be pixel-wisely shrinkage problem.

The quantitative comparisons tested on 50 hyperspectral dataset images using 5 different reconstruction schemes are presented in Fig. 4. From which we see that, generally speaking, our 2-step cross-scale deconvolution method provides competitive results with individual method.

4 Additional Results

We present full resolution images of the blurred and deblurred results presented in main text. See the captions of Fig. 6, Fig. 7 for details. As a comparison, we present the synthetic results using a diffractive achromat and and a regular Fresnel lens with the same focal power at the principle wavelength, shown in Fig. 5. Averaging all 50 hyperspectral tested images yields a PSNR 26.2 dB using our diffractive achromat, while 22.9 dB using a regular Fresnel lens. Besides, benefiting from the spectral invariant kernel behaviour, the reconstructed results from our method exhibit more robustness with smaller variance of PSNR values of full dataset. This means that less artifacts occur when capturing a full spectrum image.

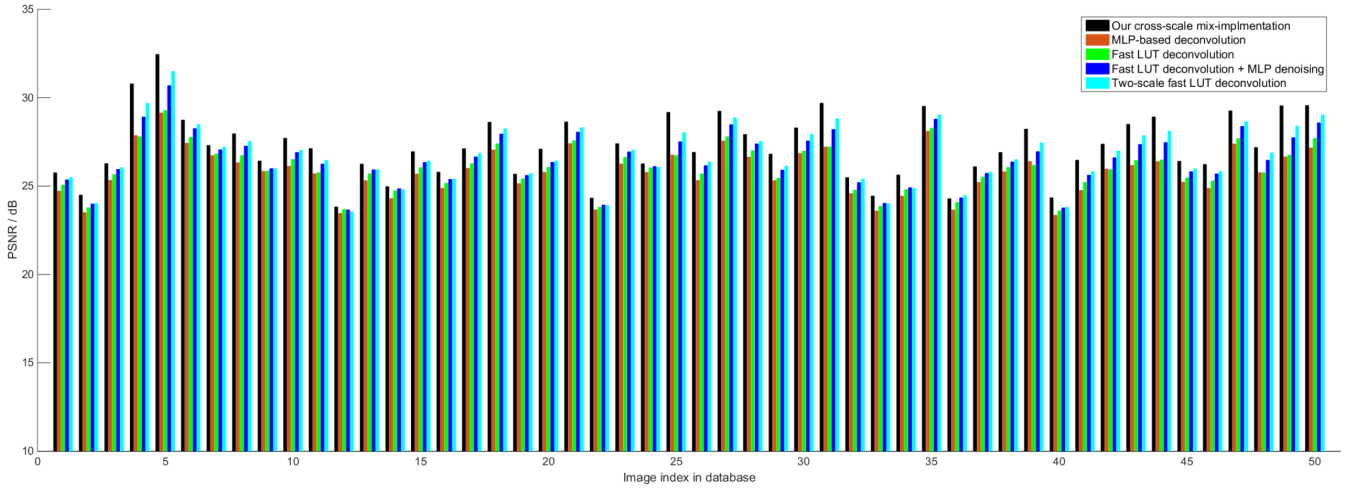

Figure 4: *PSNR comparison of results reconstructed from 5 different reconstruction schemes, tested on 50 hyperspectral images from dataset [Chakrabarti and Zickler 2011], with $\sigma = 0.005$ Gaussian white noise added. The color bars indicate the implementation with different reconstruction schemes, as has been illustrated in the legends.*

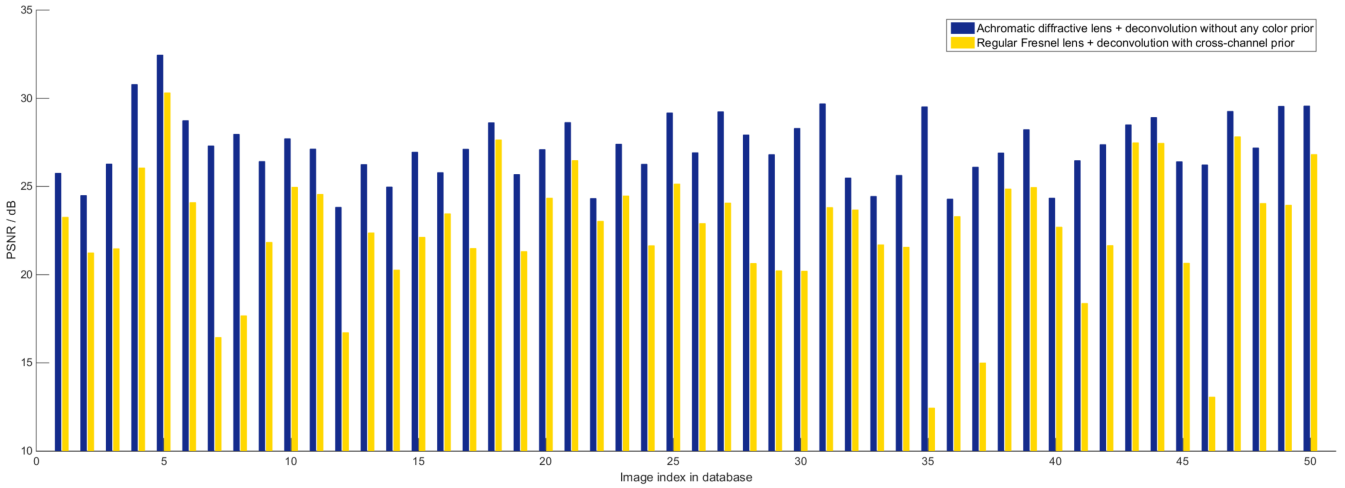

Figure 5: *PSNR comparison of reconstructed images captured using a Fresnel lens and our diffractive achromat, tested on 50 hyper-spectral images from dataset [Chakrabarti and Zickler 2011], with $\sigma = 0.005$ Gaussian white noise added. The yellow bars indicate the results captured by a Fresnel lens, while the blue bar indicate the results captured by our diffractive achromat. Note that a cross-channel based deconvolution like the ones in [Peng et al. 2015] is applied to obtain the results of Fresnel lens' case.*

We present another comparison of experimental results in Fig. 8. From which we see that even after a preliminary color calibration, the reconstructed result from a Fresnel phase plate suffers from a variety of color artifacts, especially at those regions where green color component intensity is very low in real scene (red and blue color regions), the cross-channel prior and color calibration inaccurately enforce the gradient sharing between channels. Consequently, for those regions where the real objects lack sufficient full color information, they are reconstructed either with considerable color artifacts or inaccurate edge enhancement. On the contrary, ours exhibits a higher color fidelity over all image, meanwhile presenting an acceptable spatial resolution.

Diffractive-refractive hybrid lens design As has been mentioned in main text, one promising application of our achromatic diffractive imaging method is to benefit the design of flat-field lens. In those application scenarios where high quality color imaging with large field of view but compact optics structure is expected, one can incorporate our diffractive achromat with refractive lenses. As a proof of concept, the captured results of prototyping a diffractive-refractive hybrid lens are presented in Fig. 9. Compared to a pure refractive lens with the same optical power, the thickness and material volume (over 35% reduced in the prototype case), as well as off-axis geometric distortion of our hybrid design have been reduced drastically.

Comments

We recommend the readers to refer to the project web-page named *http://www.cs.ubc.ca/labs/imager/tr/2016/DiffractiveAchromatImaging/* for the visualization of more experimental results.

Figure 6: *Blurred and deblurred results of selected synthetic scenarios from hyperspectral dataset [Chakrabarti and Zickler 2011; Skauli and Farrell 2013], with $\sigma = 0.005$ Gaussian white noise added.*

References

BOYD, S., PARIKH, N., CHU, E., PELEATO, B., AND ECKSTEIN, J. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning 3*, 1, 1–122.

CHAKRABARTI, A., AND ZICKLER, T. 2011. Statistics of real-world hyperspectral images. In *Proc. CVPR*, IEEE, 193–200.

CHAMBOLLE, A., AND POCK, T. 2011. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis. 40*, 1, 120–145.

KRISHNAN, D., AND FERGUS, R. 2009. Fast image deconvolution using hyper-laplacian priors. In *Advances in Neural Information Processing Systems*, NIPS, 1033–1041.

PENG, Y., FU, Q., AMATA, H., SU, S., HEIDE, F., AND HEIDRICH, W. 2015. Computational imaging using lightweight diffractive-refractive optics. *Opt. Express 23*, 24, 31393–31407.

SKAULI, T., AND FARRELL, J. 2013. A collection of hyperspectral images for imaging systems research. In *Proc. IS&T/SPIE Electronic Imaging*, SPIE, 86600C–86600C.

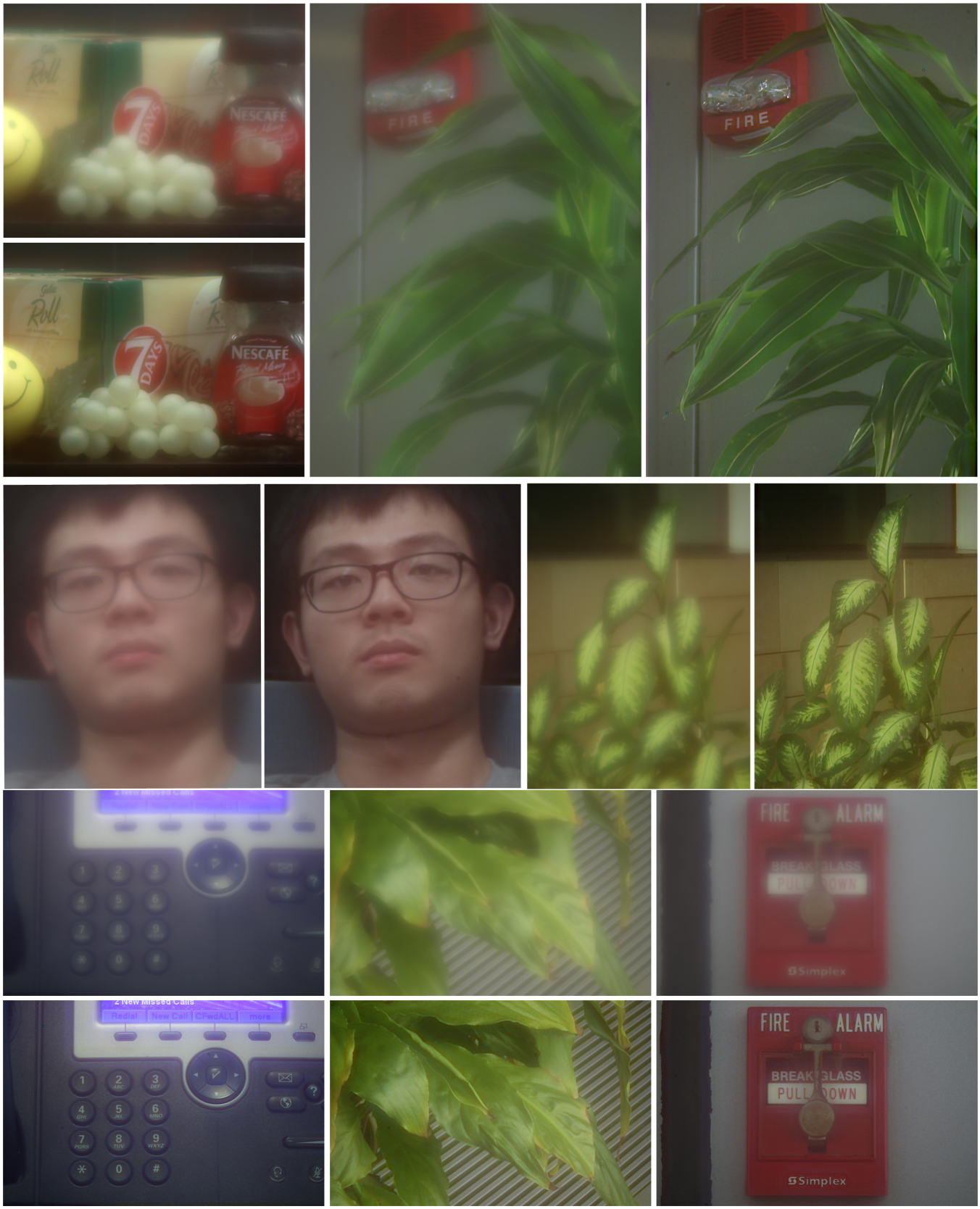

Figure 7: *Blurred and deblurred pair results of real captured scenarios. All scenes are captured at different depths with single exposure, using a single 0.5 mm ultrathin diffractive achromat we fabricate. Note that we roughly calibrate single PSF at one depth (2m), and use it for all deconvolutions.*

Figure 8: *Comparison of a Fresnel lens (left column pair) and a diffractive achromat (right column pair), with their blurred and deblurred results of natural scenarios. Both are captured at same depth with single exposure under same indoor artificial and mixed illumination, using a single 0.5mm ultrathin Fresnel phase plate and a diffractive achromat we have fabricated, respectively. Note that for the Fresnel one, a preliminary color calibration is added using Adobe Photoshop application. From which we see that although blur has been mostly removed for the Fresnel lens one, considerable color artifacts are observed, explicitly illustrated by the failure of recover blue and red colors. On the contrary, the recovery from our diffractive achromat successfully removes the blur, as well as preserve the color fidelity.*

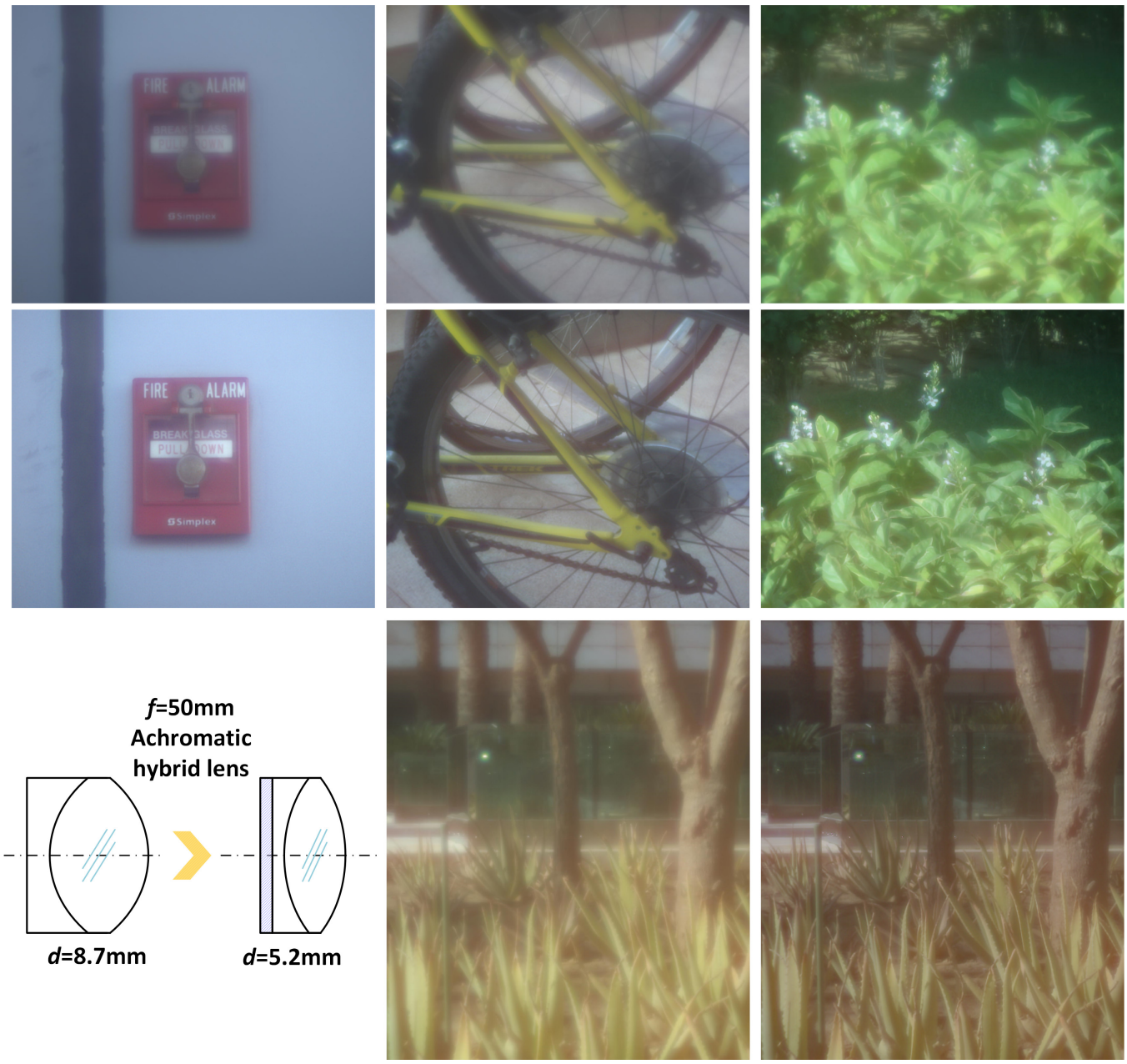

Figure 9: *Blurred and deblurred pair results of natural scenes captured using a diffractive-refractive hybrid lens. All these scenes are casually captured at different depths with single exposure and no gain under outdoor sunlight illumination, using a diffractive achromat attached with an achromatic double element refractive lens from Thorlabs. The equivalent focal length is 50 mm, with a total optics thickness 5.2 mm (For comparison purpose, a pure refractive achromatic lens with same focal length and same material has the optics thickness 8.7 mm). Note that we roughly calibrate single PSF at one depth (around 2 m) indoor, and use it for all deconvolutions.*