Waveguide Holography: Towards True 3D Holographic Glasses – Supplementary Material

Changwon Jang[1,+,*], Kiseung Bang[1,+], Minseok Chae[2], Byoungho Lee[2], and Douglas Lanman[1]

[1]Reality Labs Research, Meta, Redmond, 98052, WA, United States
[2]Seoul National University, Seoul, 08826, Republic of Korea
[*]changwon.jang@fb.com
[+]these authors contributed equally to this work

S1 Numerical simulation model of wave propagation in pupil replicating waveguide

Diffractive EPE waveguide can be designed with different types of gratings such as SRG[1–3], volume Bragg gratings[4,5], and polarization dependent gratings[6–9]. Each type may be implemented as reflection type or transmission type, or often both modes can be utilized at the same time. Physically rigorous simulation method of light interaction in waveguide varies depending on the types and configurations of the gratings and typically they are computationally prohibited for large scale simulations. Often, approximation to geometrical optics regime can be useful for the system level design of waveguide displays[10,11], however, coherent interaction must be considered in our study. Thus, we formulate simplified numerical simulation model in wave optics regime by making some assumptions. First, thickness of grating is assumed to be negligible as typical grating has micrometer scale thickness. Second, only the first order interaction is considered meaning that multi-order diffraction or non-TIR Fresnel reflection of light is neglected. Lastly, polarization of light is not considered since SLM is active to only a single polarization and we put a polarizer at the output side. For simplicity, we assume the waveguide consists of three SRG gratings; input coupler grating, EPE grating, and output coupler grating. Although, the simulation method can be applied to different designs with different shapes, numbers, and arrangement of gratings with a simple modification. Let the input wavefront $u_{in}(x,y)$ be initially in-coupled into the waveguide and guided as total internal reflection mode u_g:

$$u_{g|m=0} = u_{in}\eta_{in}A_{in}e^{k_{gx1}x+k_{gy1}y}, \tag{1}$$

where m is the number of pair of TIR reflection that wavefront has undergone, A_{in} is the aperture function of the input coupler grating, k_{gx1} and k_{gy1} are the momentum of the input grating, and η_{in} represents the coupling efficiency simplified as a single scalar value. Note that the diffraction efficiency can be represented as a function of angular component of the wavefront to simulate more accurate diffraction efficiency. The coupled wavefront bounces back to the grating surface every two total internal reflection mode propagation:

$$u_{g|m=1} = u_{g|m=0} \otimes h_{ASM}, \tag{2}$$

where h_{ASM} is a wave propagation kernel of distance $2t$ in the substrate. Note that TIR does not invert the phase of the wavefront. Without losing generality, the m-th TIR propagation from input coupler can be described as follows (m>1):

$$u_{g|m+1} = u_{g|m} \otimes h_{ASM} - u_{g2|m}e^{-j(k_{gx2}x+k_{gy2}y)} - u_{recoupled|m}e^{j(k_{gx1}x+k_{gy1}y)}, \tag{3}$$

where

$$u_{g2|m} = u_{g|m} \otimes h_{ASM}\eta_{epe}M_{epe}e^{j(k_{gx2}x+k_{gy2}y)}, \tag{4}$$

$$u_{recoupled|m} = u_{g|m} \otimes h_{ASM}\eta'_{in}M_{in}e^{-j(k_{gx1}x+k_{gy1}y)}. \tag{5}$$

$u_{g2|m}$ is the diffracted light from EPE grating that is guided to different direction and propagate towards the output coupler grating. Depending on the layout of waveguide, some portion of light $u_{recoupled|m}$ may be re-coupled as -1 order diffraction of the input grating and exit the waveguide with diffraction efficiency η'_{in}. In practical, the coupling efficiency for -1 order can be designed to be smaller than 1^{st} order. For simplicity, we neglect $u_{recoupled}$ and assume EPE mask to be homogeneous, then we can re-write u_g as:

$$\begin{aligned} u_{g|m+1} &= u_{g|m} \otimes h_{ASM}(1-\eta_{epe}) \\ &= u_{g|m} \otimes h_{epe}. \end{aligned} \tag{6}$$

Then the input wavefront of the EPE grating can be rewritten as:

$$\begin{aligned} u_{in2} &= u_{g|1} + u_{g|2} + u_{g|3} + u_{g|4} + \cdots \\ &= u_{in}\eta_{in}e^{j(k_{gx1}x+k_{gy1}y)} \otimes (h_{epe} + h_{epe}^{\otimes 2} + h_{epe}^{\otimes 3} + h_{epe}^{\otimes 4} + \cdots) \\ &= u_{in}\eta_{in}e^{j(k_{gx1}x+k_{gy1}y)} \otimes h'_{epe}, \end{aligned} \tag{7}$$

where $h_{epe} = h_{ASM}(1-\eta_{epe})$ and $^{\otimes n}$ indicates $n-1$ successive convolution of itself. Similarly, we can write input wavefront of the output coupler grating as:

$$\begin{aligned} u_{in3} &= u_{in2}\eta_{epe}e^{j(k_{gx2}x+k_{gy2}y)} \otimes (h_{out} + h_{out}^{\otimes 2} + h_{out}^{\otimes 3} + h_{out}^{\otimes 4} + \cdots) \\ &= u_{in2}\eta_{epe}e^{j(k_{gx2}x+k_{gy2}y)} \otimes h'_{out}, \end{aligned} \tag{8}$$

where $h_{out} = h_{ASM}(1-\eta_{out})$. And finally, the output wavefront is represented as:

$$u_{out} = u_{in3}\eta_{out}e^{j(k_{gx3}x+k_{gy3}y)} \tag{9}$$

where k_{gx3} and k_{gy3} are the momentum of the output coupler grating. In the typical design of EPE waveguide $k_{gx1}+k_{gx2}+k_{gx3} = 0$ and $k_{gy1}+k_{gy2}+k_{gy3} = 0$, which leads to:

$$u_{out} = u_{in} \otimes h_{wg} \tag{10}$$

where

$$h_{wg} = \mathscr{F}^{-1}(\eta_{in}\eta_{epe}\eta_{out}H'_{epe}(f_x - k_{gx1}/2\pi, f_y - k_{gy1}/2\pi) \times H'_{out}(f_x - k_{gx2}/2\pi, f_y - k_{gy2}/2\pi)). \tag{11}$$

Equation 11 implies that the coherent interaction in the waveguide can be simply represented as a complex valued kernel h_{wg} by neglecting multi-order diffraction and the clipping effect at the grating boundaries. This result offers a physical intuition that the waveguide can be closely approximated as an LSI system. Also, the result can be utilized for useful numerical simulations presented in the design space analysis section since the computation can be significantly reduced compared with full tracking of light path that requires numerous convolution process. Although, more precise simulation can be performed based on Eq. 1-5 without applying assumptions. To validate the single kernel approximation, we have trained the model with simulated wavefront dataset without approximation and compared the fidelity. In the simulation, we added some non-idealities of the system such as DC noise and slight aberration. With the single kernel, 24.7 dB of c-PSNR was achieved, while 9-channel model could achieve 27.1 dB. We use the result of Eq.11 for design space and scalability analysis presented in Fig. 9 in the manuscript.

S2 Analytic derivation of gradient in the waveguide model

We provide the analytic derivation of the waveguide model that can be used for model training. The estimated output of the waveguide model u' can be represented as:

$$u' = \mathscr{A}(u_{in}) = \left[\sum_i \{(u_{in} \times Q_i) \otimes h_i\} \times R_i\right] + DC, \tag{12}$$

where $\mathscr{A}$ is the waveguide model , u_{in} is the input field. u_{in} and Q_i are $N \times N$ complex valued matrices, u',R_i, and DC are $M \times M$ complex-valued matrices, and h_i is $(N+M) \times (N+M)$ complex valued matrix. $\otimes$ denotes a non-circular convolution operation defined in the input and output domain of $N^2, (N+M)^2 \rightarrow M^2$. Generally, discrete convolution operations can be calculated using fast Fourier transform (FFT), which has lower order of time complexity. However, input and output domain

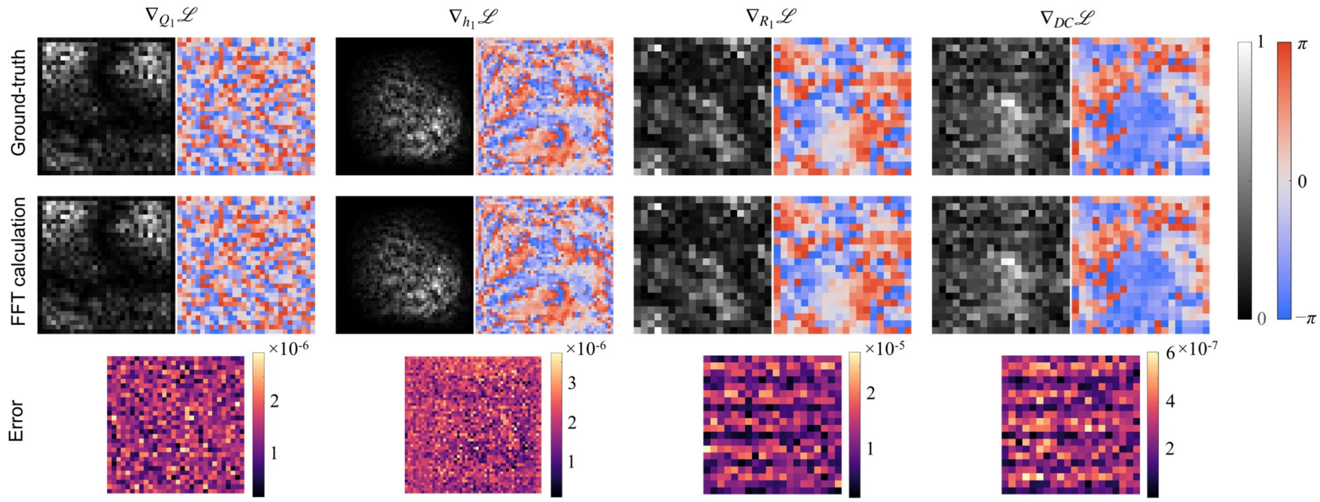

Figure S1. Example of gradients for each modeling parameter for validation of Eq. 15-18. The amplitude of the gradients was normalized to unity. The top row shows the ground-truth gradients, the middle row shows the gradient calculated using FFT as shown in Eq. 15-18, and the bottom row shows the norm of the complex error between the ground-truth and the FFT calculation. The ground-truth was obtained with a small-step change for each pixel. The FFT calculation result matches the ground-truth correctly with only a negligible difference, which is caused by the finite step change in ground-truth calculation. All the illustrated parameters are down-sampled from the original size for visualization.

size in FFT should be same and it gives circular convolution result. Therefore, to calculate the convolution $\otimes$ with FFT, it should be applied in modified way using additional zero-padding and crop operators.

$$u' = \mathscr{A}(u_{in}) = \left[\sum_i \mathscr{P}_{(N+M)^2\to M^2}\left[\mathscr{F}^{-1}\left\{\mathscr{F}\left(\mathscr{P}_{N^2\to(N+M)^2}(u_{in}\times Q_i)\right)\times\mathscr{F}(h_i)\right\}\right]\times R_i\right] + DC \tag{13}$$

$\mathscr{F}$ denotes the 2D FFT defined in the input and output domain of $(N+M)^2, (N+M)^2 \to (N+M)^2$. $\mathscr{P}_{A\to B}$ is the zero-padding and crop operator where A and B is the size of input and output. When $A < B$, it operates as a zero-padding operator, and when $A > B$, it operates as a crop operator around the center.

The cost function for model accuracy is defined as follows:

$$\mathscr{L} = \|u^{'} - u\|^2, \tag{14}$$

where u is the captured output $M\times M$ complex wavefront matrix. While optimizing the model, we update Q_i, h_i, R_i, and DC, so we need the gradient of the cost function for each parameter. Each gradient can be calculated using the FFT as follows:

$$\nabla_{h_i}\mathscr{L} = -2\mathscr{F}^{-1}\left[\overline{\mathscr{F}\left\{\mathscr{P}_{N^2\to(N+M)^2}(u_{in}\times Q_i)\right\}}\times\mathscr{F}\left\{\mathscr{P}_{M^2\to(N+M)^2}\left(\overline{R_i}\times(u'-u)\right)\right\}\right], \tag{15}$$

$$\nabla_{Q_i}\mathscr{L} = -2\mathscr{P}_{(N+M)^2\to N^2}\mathscr{F}^{-1}\left[\overline{\mathscr{F}\{h_i\}}\times\mathscr{F}\left\{\mathscr{P}_{M^2\to(N+M)^2}\left(\overline{R_i}\times\left(u'-u\right)\right)\right\}\right]\times\overline{u_{in}}, \tag{16}$$

$$\nabla_{R_i}\mathscr{L} = -2\overline{\mathscr{P}_{(N+M)^2\to M^2}\mathscr{F}^{-1}\left[\mathscr{F}\{h_i\}\times\mathscr{F}\left\{\mathscr{P}_{N^2\times(N+M)^2}(u_{in}\times Q_i)\right\}\right]}\times\left(u'-u\right), \tag{17}$$

$$\nabla_{DC}\mathscr{L} = -2(u'-u), \tag{18}$$

where the gradient of a real-valued function g in the complex domain is defined as $\nabla_{a+ib}g = \nabla_a g + i\nabla_b g$, when a and b are real numbers, and $\overline{(\cdot)}$ is the complex conjugation operator. The correctness of Eq. 15-18 is demonstrated in Fig. S1 with the comparison to the ground-truth result. Direct access to gradient calculation of Eq. 15-18 could enables the faster implementation in the future, but we used auto-gradient functionality of Pytorch[12] for utility of the modeling.

S3 Wavefront camera calibration algorithm

The wavefront camera is built based on Mach-Zehnder type interferometer that consists of a beam splitter, a piezo actuator (PZ-38), a neutral density filter, and a CMOS sensor. The sensor captures the interference pattern generated by the signal wavefront and the plane reference wavefront, with shifted phase controlled by the piezo actuator. In order to compensate the phase fluctuation caused by airflow and vibration, the piezo actuator is updated using phase locking algorithm and minimize the phase error to get correct interference pattern with target phase shift. The pre-calibration of the interferometer system is useful for achieving the precise phase locking during the wavefront data acquisition. The details of the pre-calibration and phase locking algorithm is described in the following paragraph.

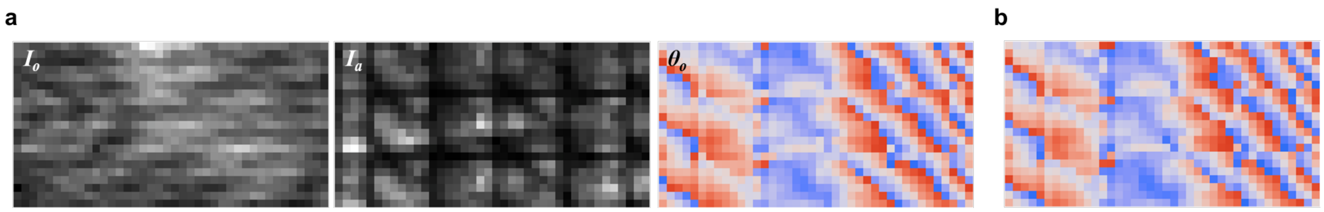

Figure S2. Example of pre-calibration result of Eq.19. **a** Illustration of I_o, I_a and θ_o after the pooling process. **b** A sample phase of output wavefront after the pooling process captured during the dataset acquisition, which shows a very similar phase profile as θ_o in **a**.

With the fixed signal wavefront and phase-shifted reference wavefront, the intensity of the interference pattern follows a sinusoidal function of the phase shift ϕ_s.

$$I(x,y;\phi_s) = I_o(x,y) + I_a(x,y)\cos\left(\theta_o(x,y) + \phi_s\right) \tag{19}$$

In the pre-calibration stage, we capture multiple interferograms with unknown phase shift ϕ_s varying the piezo-actuator and find the optimal values of the pre-calibration parameters such as intensity offset I_o, intensity amplitude I_a, and phase offset θ_o that closely approximate Eq. 19. With pre-calibration parameters, the unknown phase shift ϕ_o in the data acquisition stage can be estimated with a fairly high accuracy when arbitrary interferogram I_{cap} is acquired. We can quantify the accuracy of the estimation by the cost function g and use it to find the estimated phase shift ϕ_e.

$$g = \|I_{cap} - I(x,y;\phi_s)\| \tag{20}$$

$$\phi_e = \mathrm{argmin}_{\phi_s} g \tag{21}$$

With the knowledge of the cost function g being a sinusoidal function of ϕ_s, a simple binary search can result in a quick and accurate estimation. Using the estimated phase, we update the piezo actuator to be matched to the target phase shift. We also used the value of cost function to evaluate the perturbation in the system, which helped to get high-quality dataset for the waveguide modeling. Note that above phase locking algorithm is assuming that the signal wavefront remains the same in both the calibration and estimation stage. In practice, the SLM pattern needs to be updated as multiple random phases during the data acquisition process. In order to provide the robustness to the algorithm, average-pooling process is applied to the captured interference pattern with a pooling area of 100×100, which eliminates the effect of the high frequency components by the random phase input. This allows the system to calibrate for arbitrary fully random phase inputs after I_o, I_a, and θ_o are acquired. Figure S2 illustrates the pre-calibration results of I_o, I_a, and θ_o. After obtaining $N (N > 3)$ phase shifted interferograms, a complex wavefront can be reconstructed as follows:

$$u = \frac{1}{N}\sum_{k=1}^{N} I_k \exp\left(i\frac{2\pi k}{N}\right) \tag{22}$$

We built two wavefront cameras in the benchtop; one for capturing the relayed SLM image and the other for capturing waveguide output wavefront. Captured data of relayed SLM wavefront is used to calibrate the physical propagation module that can be set as an initialization for waveguide model training.

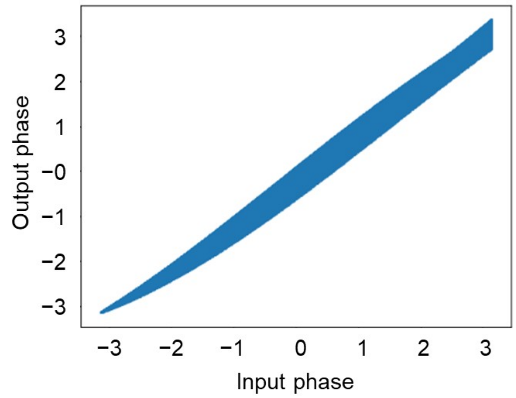

Figure S3. Example of calibration result of gamma function of the SLM.

S4 Details of physical propagation module

Gamma function of the spatial light modulator

Ideally, the phase response of all SLM pixels is linear to the gray level (or voltage) and covers the full modulation range of $-\pi$ to $+\pi$. However practically, the phase response is not linear, nor consistent over the SLM area. To calibrate the phase response of the SLM, we inserted a gamma function layer as follows.

$$\phi_{out}(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3}\sum_{k=0}^{3} c_{ijk} \times x^i \times y^j \times \phi_{in}(x,y)^k, \tag{23}$$

where x and y are the position of the pixel, and ϕ_{in} and ϕ_{out} are the input phase (gray level) and the output phase. c_{ijk} is the set of coefficients that is adjusted for calibration. This formula allows to model a smoothly-varying nonlinear curve over the spatial positions. Figure S3 is the overlap of phase response curves of all SLM pixels after calibration. A similar calibration result has been reported recently[13].

Crosstalk of the spatial light modulator

It is known that the high resolution liquid crystal on silicon panel suffers from the field fringe effect.[14] When the voltage is applied to a LC cell in a pixel, the effective electric field is disturbed by the leaked electric field from the adjacent pixels. To include such field fringe effect, we inserted a 3×3 convolution layer after the gamma function layer. The convolution is applied in a real domain of the phase after upsampling layer with a factor of two, before converted to the complex wavefront. The upsampling is performed because the field-fringe effect noise includes a higher frequency components that cannot be captured by the maximum spatial frequency supported by the SLM pixel pitch.

Homography mapping function

To match the coordinates between the SLM pixels and the camera pixels and also compensate for the distortion, we re-map the homography of the wavefront before applying the optical propagation. We used up to 2nd order coefficients of homography transformation using 12 parameters. Similar functionality can be achieved by piece-wise affine transform[15].

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{20} & a_{11} & a_{02} & a_{10} & a_{01} & a_{00} \\ b_{20} & b_{11} & b_{02} & b_{10} & b_{01} & b_{00} \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \\ x \\ y \\ 1 \end{bmatrix} \tag{24}$$

Optical propagation and tilt

Band-limited angular spectrum method[16] is used for numerical propagation in the model. The angle alignment error between the SLM and waveguide can be also considered in the model by using 3D tilt propagation[17], however the computation involves re-sampling process which adds a slight error. Therefore we did not include it in the final result of the benchtop prototype calibration. The propagation parameter is pre-calibrated using a wavefront camera that images a relayed SLM and fixed during the waveguide model training stage.

S5 Additional Results

Figure S4 demonstrates the experimental results of the waveguide holography when images are displayed at infinity conjugate. Additional results captured in different pupil locations are provided in Figs. S5–S7. Relative pupil (x, y, z) locations of pupil 1–6 are $(-4,3,0)$, $(1,1,0)$, $(5,4,0)$, $(-3,-2,2)$, $(-5.5,-4.5,-5)$, and $(3,-2,5)$ in millimeter unit, respectively. The center of the eyebox is set as 25 mm distance (normal direction) from the center of the out-coupler grating. Figure S8 illustrates the effect of pupil offset, simulating the eyetracking error during the software-pupil steering. The upper row (cat image) is displayed at 1 D and the bottom row (tower image) is displayed at 0 D.

Figure S4. Experimental results of waveguide holography when the images are displayed at infinity (0 D). The insets correspond to 1.3 degree of field of view. Our model improves the image quality significantly even when the image is displayed at the infinity depth, where there is no explicit presence of ghost noise. It demonstrates that our method can overcome the non-idealities in the waveguide display system such as surface aberration and beam clipping effect.

References

1. Levola, T. 7.1: Invited paper: Novel diffractive optical components for near to eye displays. In *SID Symposium Digest of Technical Papers*, vol. 37, 64–67 (Wiley Online Library, 2006).
2. Kress, B. & Shin, M. Diffractive and holographic optics as optical combiners in head mounted displays. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 1479–1482 (2013).
3. Kress, B. C. & Cummings, W. J. 11-1: invited paper: towards the ultimate mixed reality experience: Hololens display architecture choices. In *SID symposium digest of technical papers*, vol. 48, 127–131 (Wiley Online Library, 2017).
4. Mukawa, H. *et al.* A full-color eyewear display using planar waveguides with reflection volume holograms. *J. society for information display* **17**, 185–193 (2009).
5. Han, J., Liu, J., Yao, X. & Wang, Y. Portable waveguide display system with a large field of view by integrating freeform elements and volume holograms. *Opt. express* **23**, 3534–3549 (2015).
6. Weng, Y. *et al.* Liquid-crystal-based polarization volume grating applied for full-color waveguide displays. *Opt. Lett.* **43**, 5773–5776 (2018).
7. Yoo, C., Bang, K., Chae, M. & Lee, B. Extended-viewing-angle waveguide near-eye display with a polarization-dependent steering combiner. *Opt. Lett.* **45**, 2870–2873 (2020).
8. Yoo, C. *et al.* Dual-focal waveguide see-through near-eye display with polarization-dependent lenses. *Opt. Lett.* **44**, 1920–1923 (2019).
9. Afra, T., Salehi, M. & Abiri, E. Design of two compact waveguide display systems utilizing metasurface gratings as couplers. *Appl. Opt.* **60**, 8756–8765 (2021).
10. Wyrowski, F. & Bryngdahl, O. Digital holography as part of diffractive optics. *Reports on Prog. Phys.* **54**, 1481 (1991).
11. Lee, S. *et al.* Viewing angle enhancement of an integral imaging display using bragg mismatched reconstruction of holographic optical elements. *Appl. Opt.* **55**, A95–A103 (2016).
12. Pytorch. https://pytorch.org/.
13. Liebmann, M., Valverde, J. & Kerbstadt, F. Wavefront compensation for spatial light modulators based on Twyman-Green interferometry. In Lee, J.-H., Wang, Q.-H. & Yoon, T.-H. (eds.) *Advances in Display Technologies XI*, vol. 11708, 117080X, 10.1117/12.2583077. International Society for Optics and Photonics (SPIE, 2021).
14. Fan-Chiang, K.-H., Wu, S.-T. & Chen, S.-H. Fringing-field effects on high-resolution liquid crystal microdisplays. *J. Disp. Technol.* **1**, 304 (2005).
15. Piecewise affine transformation. https://scikit-image.org/docs/stable/auto_examples/transform/plot_piecewise_affine.html.
16. Matsushima, K. & Shimobaba, T. Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields. *Opt. Express* **17**, 19662–19673, 10.1364/OE.17.019662 (2009).
17. Matsushima, K., Schimmel, H. & Wyrowski, F. Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves. *J. Opt. Soc. Am. A* **20**, 1755–1762, 10.1364/JOSAA.20.001755 (2003).

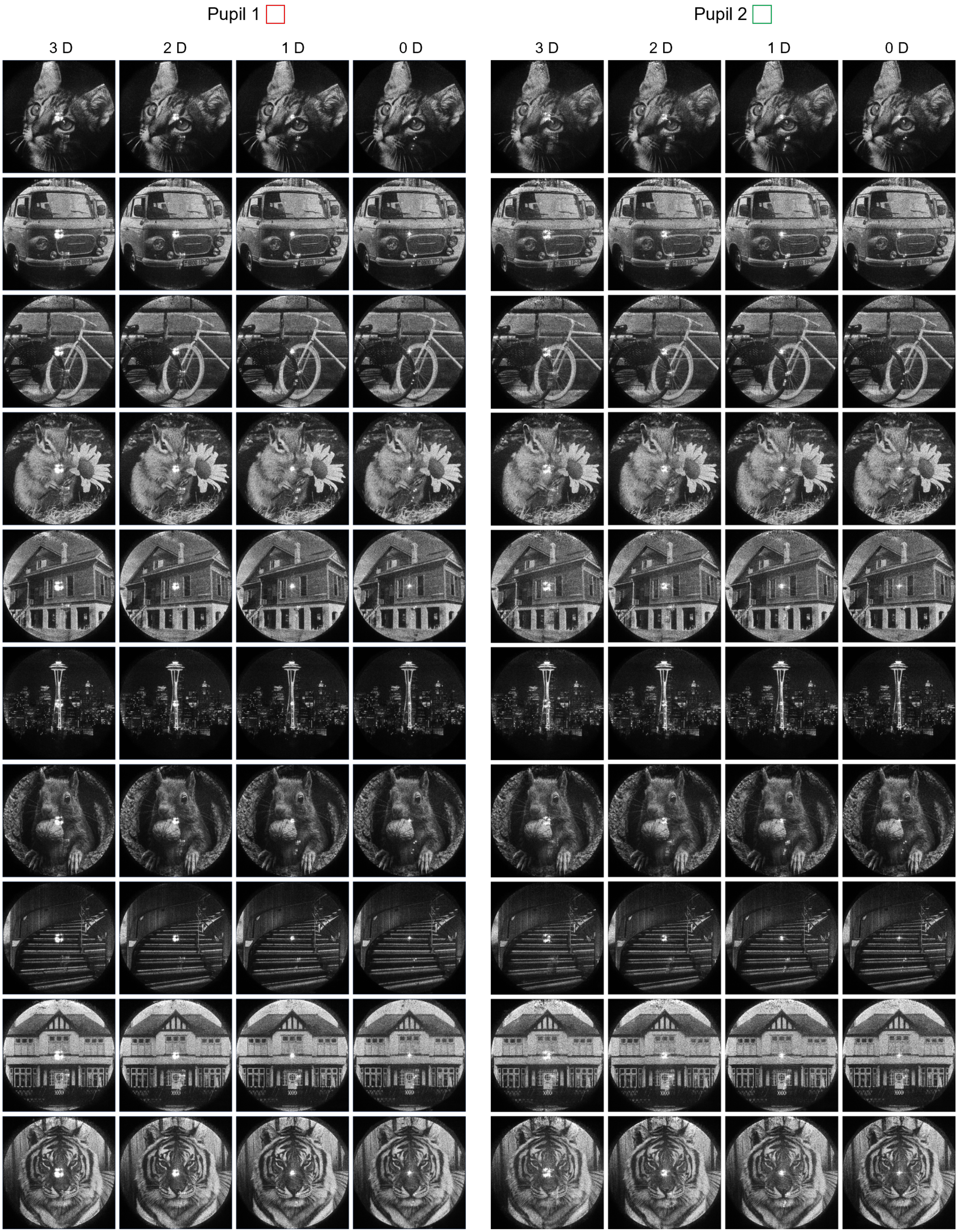

Figure S5. Display results at different pupil locations (pupil 1 and pupil 2), captured with imaging camera.

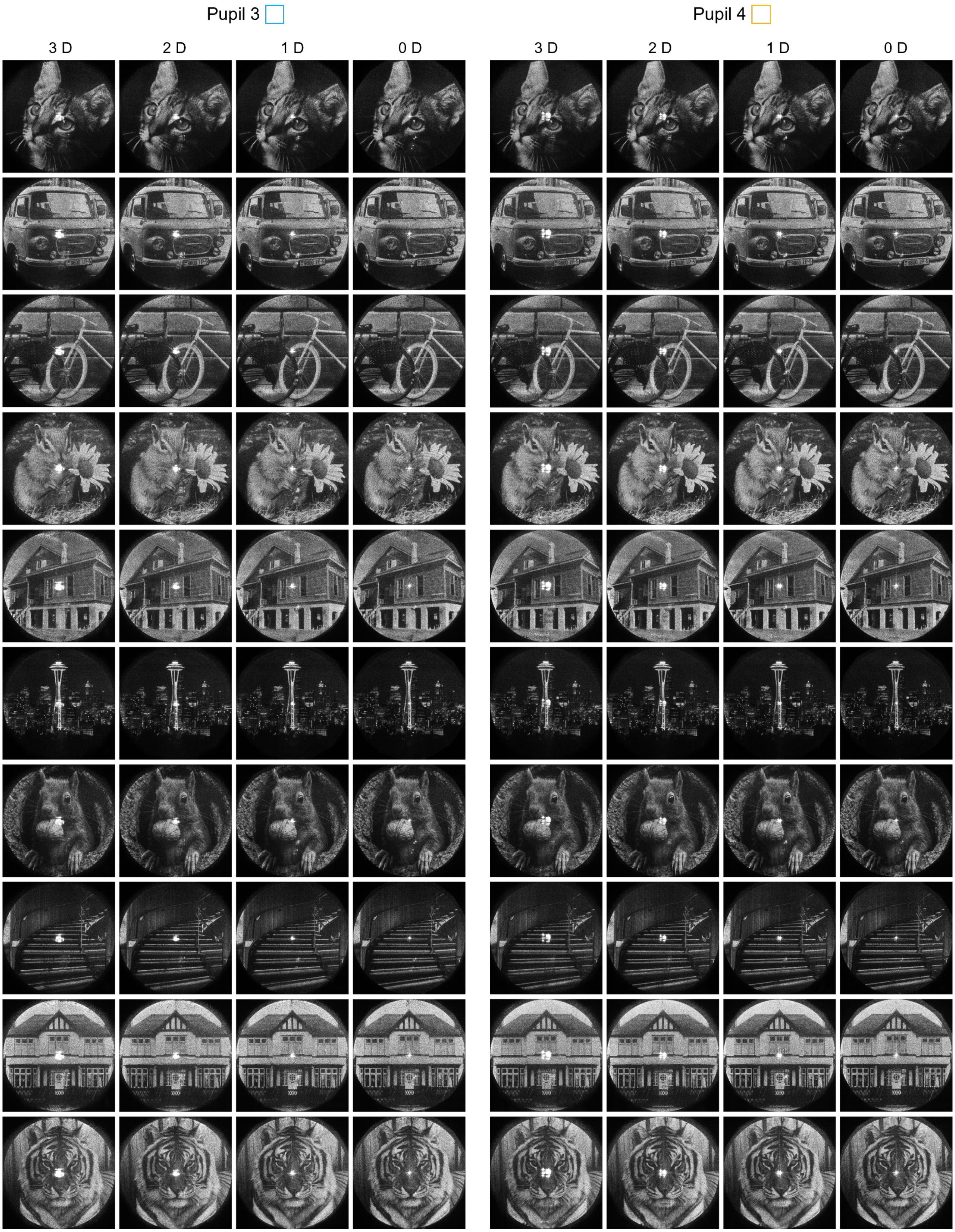

Figure S6. Display results at different pupil locations (pupil 3 and pupil 4), captured with imaging camera.

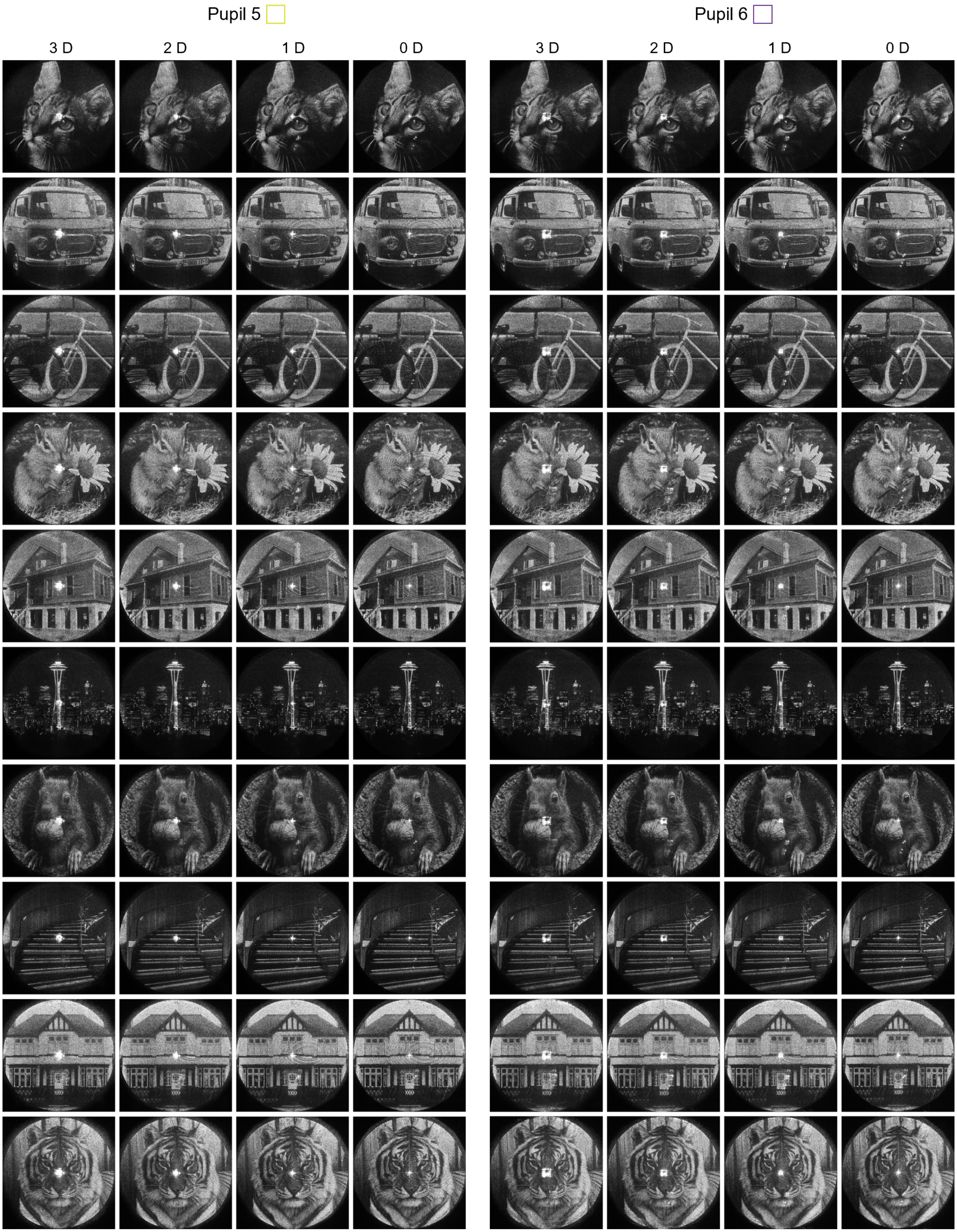

Figure S7. Display results at different pupil locations (pupil 5 and pupil 6), captured with imaging camera.

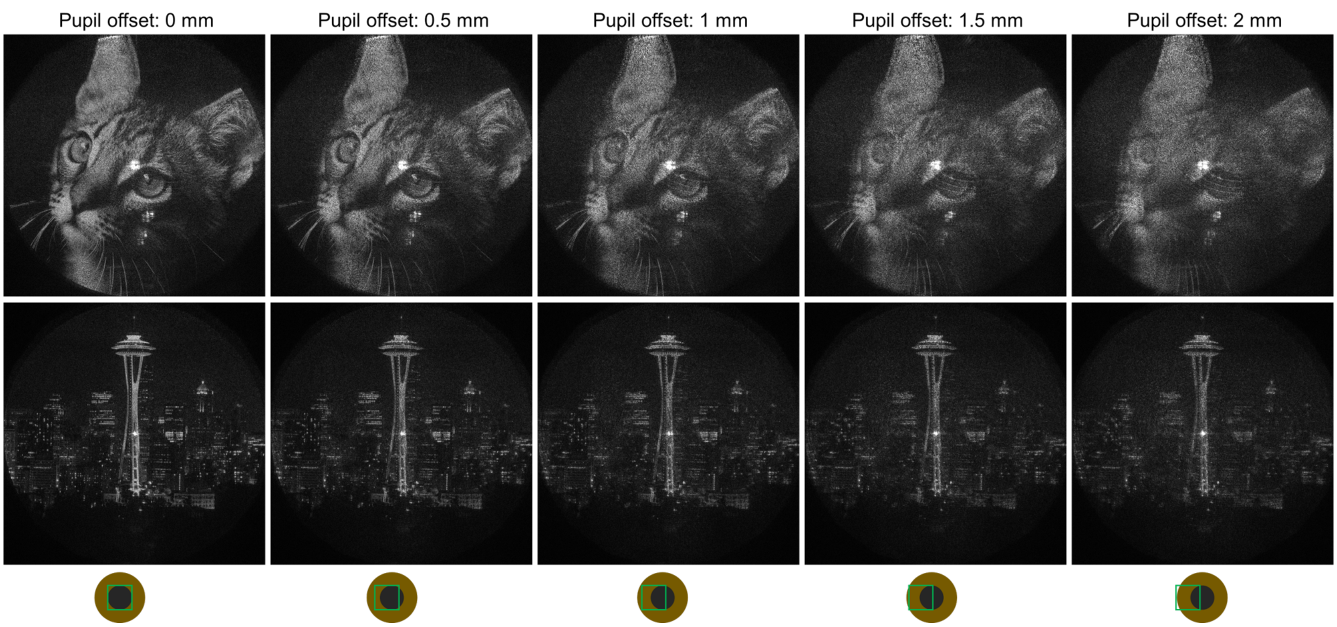

Figure S8. Effect of pupil offset. Captured wavefront is cropped with different offset values and numerically propagated to simulate the image degradation with presence of eyetracking error. The optimized pupil size could be set larger than user's actual pupil size to provide a margin for eyetracking error.