Designing Phase Masks for Under-Display Cameras

Anqi Yang[†], Eunhee Kang[‡], Hyong-Euk Lee[‡], Aswin C. Sankaranarayanan[†]
[‡]Samsung Advanced Institute of Technology Suwon-si, South Korea
[†]Carnegie Mellon University, Pittsburgh, USA

Abstract

Diffractive blur and low light levels are two fundamental challenges in producing high-quality photographs in under-display cameras (UDCs). In this paper, we incorporate phase masks on display panels to tackle both challenges. Our design inserts two phase masks, specifically two microlens arrays, in front of and behind a display panel. The first phase mask concentrates light on the locations where the display is transparent so that more light passes through the display, and the second phase mask reverts the effect of the first phase mask. We further optimize the folding height of each microlens to improve the quality of PSFs and suppress chromatic aberration. We evaluate our design using a physically-accurate simulator based on Fourier optics. The proposed design is able to double the light throughput while improving the invertibility of the PSFs. Lastly, we discuss the effect of our design on the display quality and show that implementation with polarization-dependent phase masks can leave the display quality uncompromised.

1. Introduction

Diffraction is perhaps the fundamental limiter of performance in most imaging systems, and nowhere is this more readily apparent than in an under-display camera (UDC). UDCs obtain photographs from the light that passes through the space between display pixels; since the spacing is small, often in tens of microns, the resulting blur due to diffraction spans hundreds of pixels in the captured images and is hard to remove in post-processing. This deblurring is further complicated by the display blocking a large portion of light which results in the captured photograph having an unusually low signal-to-noise ratio (SNR).

Much of the prior work in recent UDC imaging has been in designing deep neural networks to deblur [27, 12, 5, 26, 11]. Learning-based restoration largely outperforms conventional methods, such as Wiener deconvolution or regularized linear inverse techniques [3]. However, the performance of all techniques, whether regularized by analytical or data-driven priors, is affected strongly by the conditioning of the imaging model and SNR of the captured images.

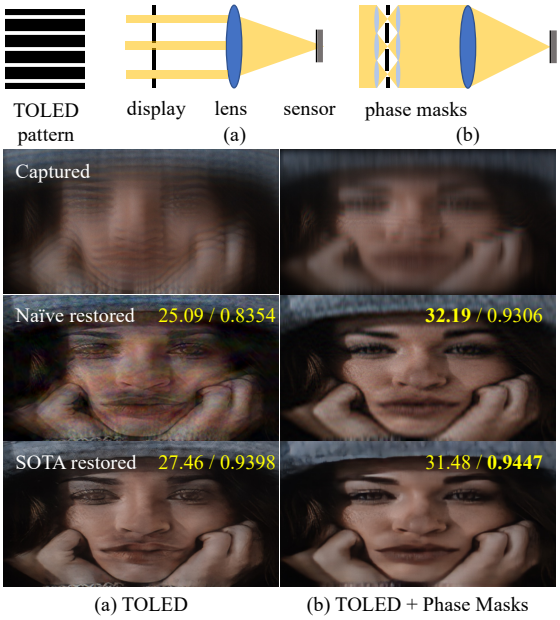


Figure 1: **A comparison between a UDC under a transparent-OLED display (a) without and (b) with the proposed phase masks.** (From top to bottom) Rows show the setup, images captured under each, restored images using a naive iterative solver, and using a state-of-the-art deep network [5]. We show PSNR(↑) in dB and SSIM(↑) for restored images. Both UDCs have a pixel density of 600 DPI.

Inspired by a large body of work that enhances capability of imaging systems with phase masks [22, 17, 10, 20, 8, 16, 19], we propose to design phase masks to suppress diffractive blur and increase light throughput for UDCs. We first show using basic Fourier optics [7] that inserting a thin phase mask at the display is ineffective in improving UDCs. To overcome the limitation of a single phase mask, we propose to use two phase masks—specifically two microlens arrays– placed *in front of* and *behind* the display; we do this for the specific case of transparent-OLEDs (TOLED), a display model commonly used in today's cellphones. The

first phase mask distributes light to locations where the display is transparent, and the second phase mask recovers the original waveform. Within some limits, this allows the incident light to pass through without being blocked by the display or effectively renders the display fully transparent!

In order to prevent the microlens arrays from hindering the display quality, we propose to implement them as thin polarization-dependent phase masks. A naive implementation of microlens arrays as thin optics is to fold them at a fixed height. However, this results in severe chromatic aberrations. We instead choose a different height for each microlens through optimization, so that diffractive blur is suppressed equally at all wavelengths. Using simulations, we show that the proposed phase mask significantly increases the image quality of a UDC (see Figure 1). The code for this work is publicly available [23].

In summary, we make the following contributions:

- We show that a phase mask placed tightly against a display is inadequate to improve the image quality of UDCs.
- We propose to insert two microlens arrays in front of and behind a TOLED, which effectively allows more light to reach the camera and produces more invertible PSFs.
- We implement the proposed microlens arrays as thin polarization-dependent phase masks, a design that ensures light emitting from the display is not modulated and therefore guarantees high display quality.
- When implementing microlenses as thin optics, we optimize the folding height of each microlens to minimize chromatic aberrations.
- We conduct simulation based on wave optics and physically-accurate camera pipeline and demonstrate that the proposed setup outperforms the conventional UDC.

Limitations. The proposed method has two limitations. First, we show phase correction can suppress the diffractive blur of TOLED display, whose pattern is separable along x and y directions. Extending this to 2D displays is hard due to the high computational cost of simulating 2D short-distance propagation. Second, the field of view of the resulting UDC can be constrained by the use of phase masks. This is determined by the ratio of the focal length f of microlens arrays and the size of the pixel opening. Our choice of f produces a field of view of around 14°.

2. Background

Under-Display Cameras. Let us denote the effective aperture of a UDC as $a(x, y) \in \{0, 1\}$ where zeros represent display regions with RGB pixels and circuits and block light, and ones represent fully transparent regions. If the focal length of the camera lens is f, the blur kernel at a wavelength λ is given as

$$k(x, y; \lambda) \propto \left| A\Big(\frac{x}{\lambda f}, \frac{y}{\lambda f}\Big) \right|^2 \tag{1}$$

where $A(\cdot)$ denotes the Fourier transform of $a(\cdot)$. An image captured with a UDC can be formed as,

$$I_b(x, y) = \int_\lambda [I_s(x, y; \lambda) * k(x, y; \lambda)] \cdot s(\lambda) d\lambda + n(x, y), \tag{2}$$

where I_s is the high-quality image, I_b is the captured photograph, $s(\lambda)$ is the sensor spectral response, and n is noise.

OLED Displays. Organic LED (OLED) displays have become prevalent in recent years in smartphones, tablets, monitor screens, and televisions. In this paper, we focus on the simplest layout TOLED, whose opening pattern is shown in the upper left of Figure 1 and is a separable along $x-$ and $y-$directions. Along the x-direction, every display pixel has an opening of around $23.8\,\%$; in y-direction, the aperture is fully open. The PSF produced by TOLED is also separable. Therefore, we focus on the design of one-dimensional phase masks for x-direction.

Image Restoration for UDCs. Recently there has been a large body of work focusing on restoring high-quality images from those captured under UDCs [27, 12, 5, 26, 21, 14, 25, 15, 6, 11]. Zhou et al. [26] designed a variant of UNet [18] to deblur and denoise images captured with a UDC. Kwon et al. [12] proposed a CNN that takes the degraded images, noise level, and spatially-varying blur kernels as input, and reconstructs sharp images. Feng et al. [5] take into account high dynamic range and saturation, and propose a Dynamic Skip Connection Network to remove diffraction artifacts. These techniques achieve impressive image reconstruction performance for UDCs.

UDC Design. To tackle the fundamental limitation of diffraction blur in UDCs, a promising direction is to re-design the hardware of a UDC. Yang and Sankaranarayanan [24] showed that the diffraction blur in UDCs is determined by the shape of display openings. They optimize the shape of display openings as well as suggest the use of a random display tiling. The resulting PSF is more robust to inversion in the presence of noise and the image quality of the UDC is significantly improved. However, this method required a non-trivial effort to re-design the display. In contrast, we propose a complementary design by inserting carefully designed phase masks, that do not require any modification to the display fabrication.

3. Phase Mask Design for UDCs

We explore the design space of phase masks in UDCs and look into two scenarios — first, a single phase mask placed tightly against the display and second, two phase modulations in front of and behind the display.

3.1. Inadequacy of Single Phase Masks

Phase masks modulate the phase of an incident wavefront and can potentially correct the wavefront to form a PSF that is easily invertible. We first examine the most common setup of placing a single phase mask at the aperture plane [22, 17, 10, 20, 8, 16, 19]. Unfortunately, using basic Fourier optics, we show that such a phase mask is insufficient to reduce diffraction blur in UDCs.

Lemma (Inadequacy of a single phase mask)**.** *A single-sided phase mask can not improve the invertibility of the point spread function of a UDC.*

Proof. Let $a(x)$ be the aperture of a UDC, and $h(x)$ be the height map of a single-sided phase mask that is placed tightly against the display panel. We assume that the aperture and the phase mask are on the same plane, and the overall aperture function $b(x)$ can be written as

$$b(x) = a(x)e^{j\frac{2\pi}{\lambda}(n-1)h(x)} \tag{3}$$

where n is the refractive index of the phase mask and λ is the wavelength of the incident wavefront. The invertibility of PSF $k_b(x)$ can be measured by its amplitude spectrum $|K_b(u)|$, where values close to zero are hard to invert, and large values are robust to noise in inversion. From (1), $k_b(u)$ is the (scaled) power spectral density of the aperture, its Fourier transform $K_b(u) = \mathcal{AC}_b(\tau)$, where $\mathcal{AC}_b(\tau)$ is the (scaled) autocorrelation function of the aperture. We compute the autocorrelation of the overall aperture function,

$$\mathcal{AC}_b(\tau) = \int_{-\infty}^{\infty} a(x)a(x+\tau)e^{j\Delta\Phi_\tau(x)}dx, \tag{4}$$

and $\Delta\Phi_\tau(x) = \frac{2\pi}{\lambda}(n-1)(h(x)-h(x+\tau))$. We then compute the intensity of $\mathcal{AC}_b(\tau)$, and by triangle inequality,

$$|\mathcal{AC}_b(\tau)| = \Big|\int_{-\infty}^{\infty} a(x)a(x+\tau)e^{j\Delta\Phi_\tau(x)}dx\Big| \tag{5}$$

$$\leq \int_{-\infty}^{\infty} |a(x)a(x+\tau)|dx. \tag{6}$$

Since aperture function $a(x)$ is non-negative, we can further simplify the above equation,

$$|\mathcal{AC}_b(\tau)| \leq \int_{-\infty}^{\infty} a(x)a(x+\tau)dx = |\mathcal{AC}_a(\tau)|. \tag{7}$$

We can see that $|\mathcal{AC}_b(\tau)| \leq |\mathcal{AC}_a(\tau)|$ for all τ and

$$|K_b(u)| \leq |K_a(u)|, \tag{8}$$

implying that PSF produced by a display panel with a single-sided phase mask is always worse in terms of invertibility than that produced by a pure display panel. ■

3.2. Double Phase Masks

If inserting a thin phase mask at the display plane is ineffective to improve the image quality of a UDC, would inserting multiple phase masks help? Diffraction blur in a UDC is produced by the small openings on the display pixels that have sizes comparable to the wavelength of incident light. Smaller opening results in a more severe diffraction blur [24]. Would it be possible to *optically* expand the size of display openings, i.e. let a larger portion of light pass through display openings?

Consider now a system with two phase masks, on either sides of a display. The first surface with a height profile $h_1(x)$ modulates light incident on the display so that, after propagating for some distance z m, most of the intensity of the wavefront is concentrated at the display openings. The second surface $h_2(x)$ modulates the diffused wavefront so as to revert the effect of the first phase mask. If successful, the display panel would be rendered invisible.

Mathematically, this can be modeled as follows: Given a wavefront $p_\theta(x;\lambda)$ that incidents from angle θ and has a wavelength of λ. The incident wavefront passes through a phase mask, a display panel, followed by another phase mask, and becomes

$$p'_\theta(x;\lambda) = \underbrace{(\Phi_2 \circ \mathcal{Q}_z \circ a \circ \mathcal{Q}_z \circ \Phi_1)}_{a_\Phi}(p_\theta(x;\lambda)) \tag{9}$$

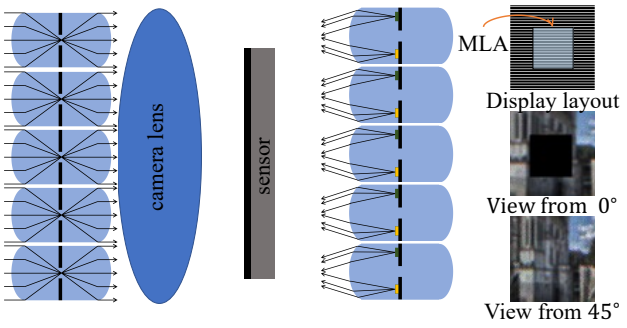
where $a(x)$ describes the display openings, $\Phi_i(x) = \exp\{\frac{2\pi}{\lambda}(n-1)h_i(x)\}, i = 1, 2$ are phase modulations of the first and second height maps, and $\mathcal{Q}_z(\cdot)$ is the operator corresponding to wave propagation of z m.

Our goal is to design height maps $h_1^*(x)$, $h_2^*(x)$ and distance z^* such that the resulting aperture $a_\Phi^*(x)$ is approximately a fully-open aperture, $a_\Phi^*(x) \approx 1$.

3.3. Proposed Design: Double Microlens Arrays

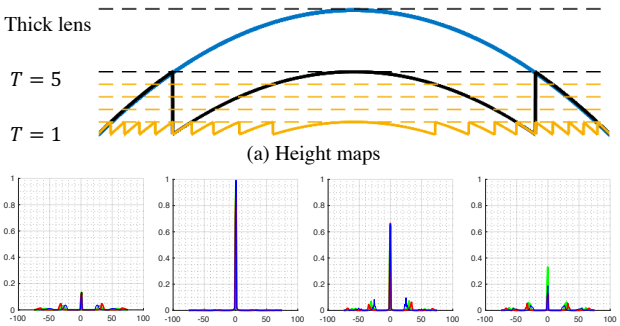
In theory, height maps and thickness of an optimal double-sided phase mask $h_1^*(x)$, $h_2^*(x)$, z^* can be solved through an optimization problem. However, propagating incoherent wavefronts with a physically accurate model at each iteration is an expensive procedure, and gradient descent only allows solving for the height range that corresponds to the range of 2π modulation.

Our design is to place two microlens arrays (MLA) with equal focal lengths on either sides of the display such that the display panel lies in the focal plane of both MLAs,



(a) UDC imaging arm (b) OLED display arm

Figure 2: **Proposed microlens arrays for UDCs.** In the right column, we render OLED display with MLA from two viewpoints. MLA is placed under the UDC aperture (center square). When viewed from the front (i.e, 0°), the display appears dark.



(a) Height maps

(b) PSFs produced by light incident along optical axis

Figure 3: **Thick lens versus phase masks** wrapped at the dash lines $T = 1, 5$ and their PSFs.

as shown in Figure 2(a). Light incident from the scene is concentrated by each microlens, passes through the display opening, and diverges to a parallel beam by the second set of microlenses. Compared to UDCs with a pure display, the proposed setup allows a larger portion of light to reach to camera main lens, and therefores improve the conditioning of incident wavefront and SNR.

However, the microlens array in front of the display also modulates light emitting from the display pixels. An illustration is shown in Figure 2(b). Since the display subpixels are misaligned with the optical axis of each microlens, light emitting from subpixels is rarely refracted to the direction along the optical axis. This implies the display would appear dark when users view it from orthogonal viewpoint.

3.4. Folding MLAs to Thin Plates

One approach to prevent microlenses from affecting the display is to implement them as polarization-dependent optics and place a pair of orthogonal linear polarizers on both sides of the display panel.

The microlens arrays only modulate the phase of light along p-polarization state. First, we examine the camera point of view. Light incident from the scene is a mixture of both states. Phase mask only modulates the p-state (shaded lines) and leaves s-state (solid color) unchanged. The polarizer behind the display selects p-states and filters out the rest. Thus the camera works in the same principle as we described in the previous section. This polarization-dependent implementation reduces the light throughput by half, which is taken into consideration in all simulations. Second, we look into the effect of phase masks on the display quality. Due to the presence of s-polarization filter, the display RGB subpixel emits light along s-state. As our phase mask only modulates light along p-polarization state, light emitting from the display is left untouched.

Since polarization-dependent optics are only available in thin optics, either as phase spatial light modulators (SLM) or thin optical elements [9, 13]. It is necessary to fold each microlens into a thin phase plate at maximum height d_0,

$$\hat{h}(x) = \mathrm{mod}\Big(-\frac{x^2}{2(n-1)f_0}, d_0\Big). \tag{10}$$

Figure 3 shows an example. Larger d_0 produces a phase mask that contains few phase wrappings and performs almost equally across all wavelengths; and small d_0 leads to much more phase wrappings and the resulting performance is strongly wavelength dependent. Phase plate wrapped at d_0 has preferable performance for light of a certain wavelength λ_0 over those of other wavelengths. This is because d_0 can be viewed as $\frac{T\lambda_0}{n-1}$, where T is an arbitary positive integer that coarsely controls the thickness of a phase mask and λ_0 is a wavelength that decides the exact thickness. A thick microlens can be written as $h(x) = \hat{h}(x) + c(x)\frac{T\lambda_0}{n-1}$, where $c(x) \in \mathbb{Z}$, and produces a phase modulation of $\exp\{j\frac{2\pi}{\lambda}(n-1)\hat{h}(x)\}\exp\{j\frac{2\pi}{\lambda}c(x)\frac{T\lambda_0}{n-1}\}$. For incident light of wavelength $\lambda = \lambda_0$, the modulation of thick lens is the same as that of the phase plate. For incident light of other wavelengths, the phase plate produces wrapping artifacts and thinner plates have more chromatic aberration.

A typical phase SLM is able to achieve phase modulations within a range of 2π or equivalently $T = 1$; other liquid crystal-based non-programmable optics can be implemented with larger phase retardation. Therefore, we design phase masks for two thicknesses: $T = 1$ for the phase SLM and $T = 5$ for thicker retarders.

3.5. Phask Masks Optimization

As mentioned in the previous section, a thin phase plate of uniform height d favors a corresponding wavelength λ

and produces wrapping artifacts for other wavelengths, resulting in chromatic aberration. We propose to optimize a different height $d[l]$ for each microlens l such that the optimized inevitability of the overall PSF is the same across RGB channels to eliminate chromatic aberration.

Given a UDC with L microlenses with corresponding heights $\{d_l | l = 1, ..., L\}$, each of which takes the value in a set of heights i.e., $d_l \in \{h_j | j = 1, ..., N\}$. The set of discrete heights is created by uniformly sampling N wavelengths from $400\,\text{nm}$ to $700\,\text{nm}$. The goal is to find the number of d_l with the same height for each h_j.[1] Thus we define $m_j = \sum \mathbb{I}(d_l = h_j), \forall l = 1, .., L$ and a vector $\mathbf{m} = [m_1, ..., m_N]^\top$ for all N heights.

We calculate the invertibility of a system with different heights as a weighted combination of that of constant ones. The invertibility is measured by $v^j(\lambda)$, the region under modulation transfer function for microlens of height h_j and a specific wavelength λ, and higher scores are better. Specifically, we form form a matrix $\mathbf{V} \in \mathbb{R}^{N \times N}$, where $\mathbf{V}_{j,k} = v^j(\lambda_k)$ is the system invertibility for height h_j and wavelength λ_k. The invertibility of a new system with mixed heights $\mathbf{m}$ can therefore be computed by $\mathbf{V}_k^\top \mathbf{m}$.

Different wavelengths contribute to the performance of RGB channels differently, for example, wavelengths close to $470\,\text{nm}$, $530\,\text{nm}$, $610\,\text{nm}$ matter more to the overall performance than other wavelengths, and the importance is characterized by the sensor spectral response function. We thus discretize the function into a matrix $\mathbf{S} = [\mathbf{s}_R^\top, \mathbf{s}_G^\top, \mathbf{s}_B^\top]$ and $\mathbf{S}^\top \mathbf{V}^\top \mathbf{x}$ computes the RGB performance under $\mathbf{m}$. We optimize the following problem,

$$\min_{\mathbf{m}} \|\mathbf{S}^\top \mathbf{V}^\top \mathbf{m} - \mathbf{1}\|_2^2 + \alpha \|\mathbf{m}\|_1 \tag{11}$$

$$\text{s.t.} \quad m_i \geq 0, \quad i = 1, \ldots, N. \tag{12}$$

The first term guarantees that performance of RGB channels is equally high, and the second term is a regularization. In optimization, $\mathbf{m} \in \mathbb{R}^N$ is a continuous variable, and in evaluation it is rounded up to integers. m_i is non-negative since it represents a count. We use the log-barrier approach to solve this constrained optimization problem. The detailed algorithm is provided in the supplementary. Figure 4 shows an example of optimized phase masks. Compared to a fixed height that favors the green sensor channel and produces severe chromatic aberration, the optimized heights perform equally across RGB channels. Compared to uniformly varying heights, the optimized profile produces sharper PSFs.

[1]We show that the *ordering* of d_l has negligible effects on the performance in supplementary. Therefore, we only optimize for the counts.

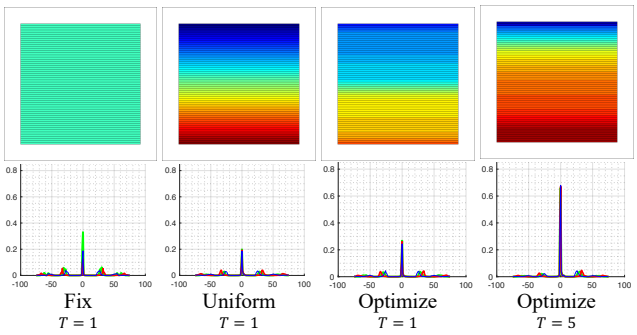


Figure 4: **Choice of d_0s at different locations.** The center part indicates the screen under the UDC aperture, and the white edges indicate the normal screen. The display is at 600 DPI. Colors from dark blue to red indicate d_0s determined by wavelengths from $400\,\text{nm}$ to $700\,\text{nm}$.

4. Imaging Model and Its Characteristics

In this section, we describe the image formation model of the proposed setup and analyze its characteristics.

4.1. Image Formation Model

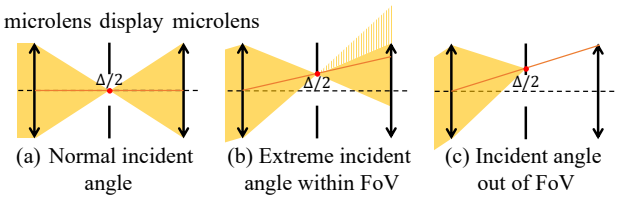
In UDCs, diffraction is usually non-negligible due to the small size of the display openings, thus we resort to wave optics in simulation. The height profiles of the first and second phase masks $h_1(x), h_2(x)$ are specified as microlens arrays as in Equation 10. Given a set of plane waves $p_\theta(x; \lambda)$ with unit irradiance. We can plug the $h_1(x), h_1(x), z = f_0$ into Equation 9, and obtain the modulated wavefront $p'_\theta(x; \lambda)$ under our design. The modulated wavefront $p'_\theta(x; \lambda)$ is then focused by the camera main lens and forms a set of blur kernels $k_\theta(x; \lambda)$, as specified in Equation 1. The blur kernel produced by a wide spectrum light source coming from angle θ can be computed as an integral of blur kernels with wavelength λs weighted by sensor spectral sensitivity $s(\lambda)$, $k_\theta(x) \approx \int_\lambda k_\theta(x; \lambda) s(\lambda) d\lambda$. We simulate 300 wavelengths from $400\,\text{nm}$ to $700\,\text{nm}$.

Since TOLED is fully-open in y-direction and it produces a blur kernel of approximately a Dirac Delta function, the captured image can be written as

$$\mathbf{I}^{\text{UDC}} = \mathbf{K}_x \mathbf{I} + \mathbf{n} \tag{13}$$

where $\mathbf{K}_x$ is a concatenation of 1D blur kernels, $\mathbf{I} \in \mathbb{R}^{1024 \times 2048}$ is a high quality image, and $\mathbf{n}$ is noise. We simulate blur kernels produced by 1,024 incoming directions that correspond to sensor pixel locations along x-direction.

Reconstruction. We first apply BM3D denoiser [4] to captured images. And then we minimize the least square error between the captured $\mathbf{I}^{\text{UDC}}$ and estimated blurry image $\mathbf{K}_x \mathbf{I}$, and regularize the estimated $\mathbf{I}$ with Tikknov priors.



(a) Normal incident angle (b) Extreme incident angle within FoV (c) Incident angle out of FoV

Figure 5: **Field of view of our design.**

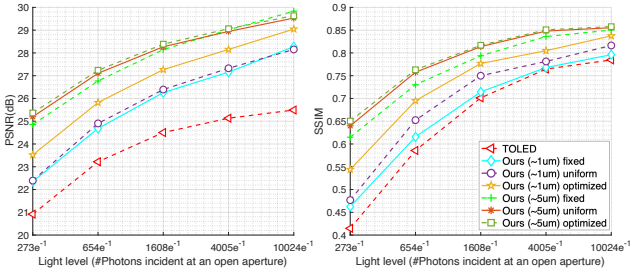
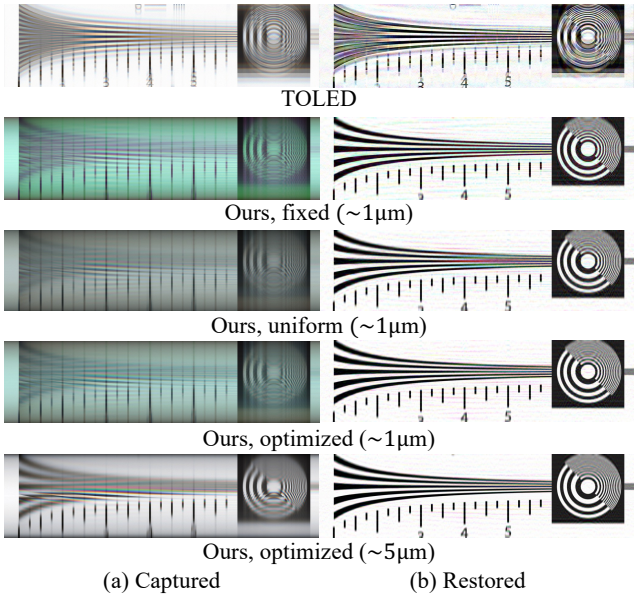


Figure 6: **Comparison of our setups with TOLED.**



(a) Captured (b) Restored

Figure 7: **Effect of phase mask optimization.**

We solve the target function using a naive iterative solver with a 'full' boundary condition and then crop the estimated $\mathbf{I}$ to have the same shape as $\mathbf{I}^{\text{UDC}}$.

4.2. Characterisics of Our Design

Field of View (FoV). Figure 5 illustrates light from different directions incident on one pair of microlenses and display pixel. Let the display pixel has an opening of Δ and the microlenses have a focal length of $\kappa\Delta$ where κ is a design choice. We choose $\kappa\Delta$ to equal display pitch, the smallest focal length if assuming spherical lenses. Normal incident light is focused to a point in the center of the display opening. As the incident angle increases, the focus point also shifts away from the center, until it reaches extreme angle. Any incident angle larger than the extreme angle is blocked by the display. The FoV of the system is

$$\text{FoV} = 2\tan^{-1}\left(\frac{\Delta/2}{\kappa\Delta}\right) \approx \frac{1}{\kappa}. \tag{14}$$

Light Transmission Ratio (LTR). Normal incident light passes through our setup without being blocked. As the incident angle increases, a larger portion of the light is out of the range of the second microlens. At the largest angle within the FoV, the LTR of our system is

$$\text{LTR}_{\text{min}} \approx 1 - \frac{1}{\kappa}. \tag{15}$$

Due to the polarization-dependent implementation of our system, the LTR is reduced by half.

5. Experiments

We design phase masks for UDCs under TOLED of pixel densities ranging from 150 to 600 DPI. All displays have an open ratio of 23.8 %. We choose the focal length to equal display pixel pitch, and thus $\kappa = 4.2$. The resulting FoV is around 14°and the LTR is between 38 % and 50 %.

Setup. We compare the proposed setup with conventional UDCs under TOLED. All displays have a pixel density of 600 DPI, equivalently a pixel pitch of 42 μm. We simulate a smartphone front camera with an aperture size of 2.3 mm and focal length of 4.67 mm. To simulate captured images, we apply spatially-varying blur kernels to ground-truth sharp images, and then add noise according to a physically-accurate noise model, and quantized to 12-bit. We emulate a sensor that has a full well capacity of 15,506 electrons and a standard deviation of 4.87 electrons, which are commonly seen in smartphone camera sensors. We set the gain to be inversely proportional to the LTR of each setup so that the captured image has consistent intensities across setups. We vary the light level by changing the number of the photons incident on an open aperture on the display from 250 to 10,000 photons. All setups are evaluated on a test set containing thirty images and using PSNR and SSIM as evaluation metrics.

Effect of Phase Masks. We compare TOLED without and with two sets of proposed phase masks that have thickness of around 1 μm and 5 μm. For each thickness, we compare three choices of wrapping heights — a fixed height determined by $\lambda_0 = 530$ nm, different heights determined by wavelengths uniformly sampled from 400 nm to 700 nm, and optimized heights. Figure 6 shows that the proposed se-

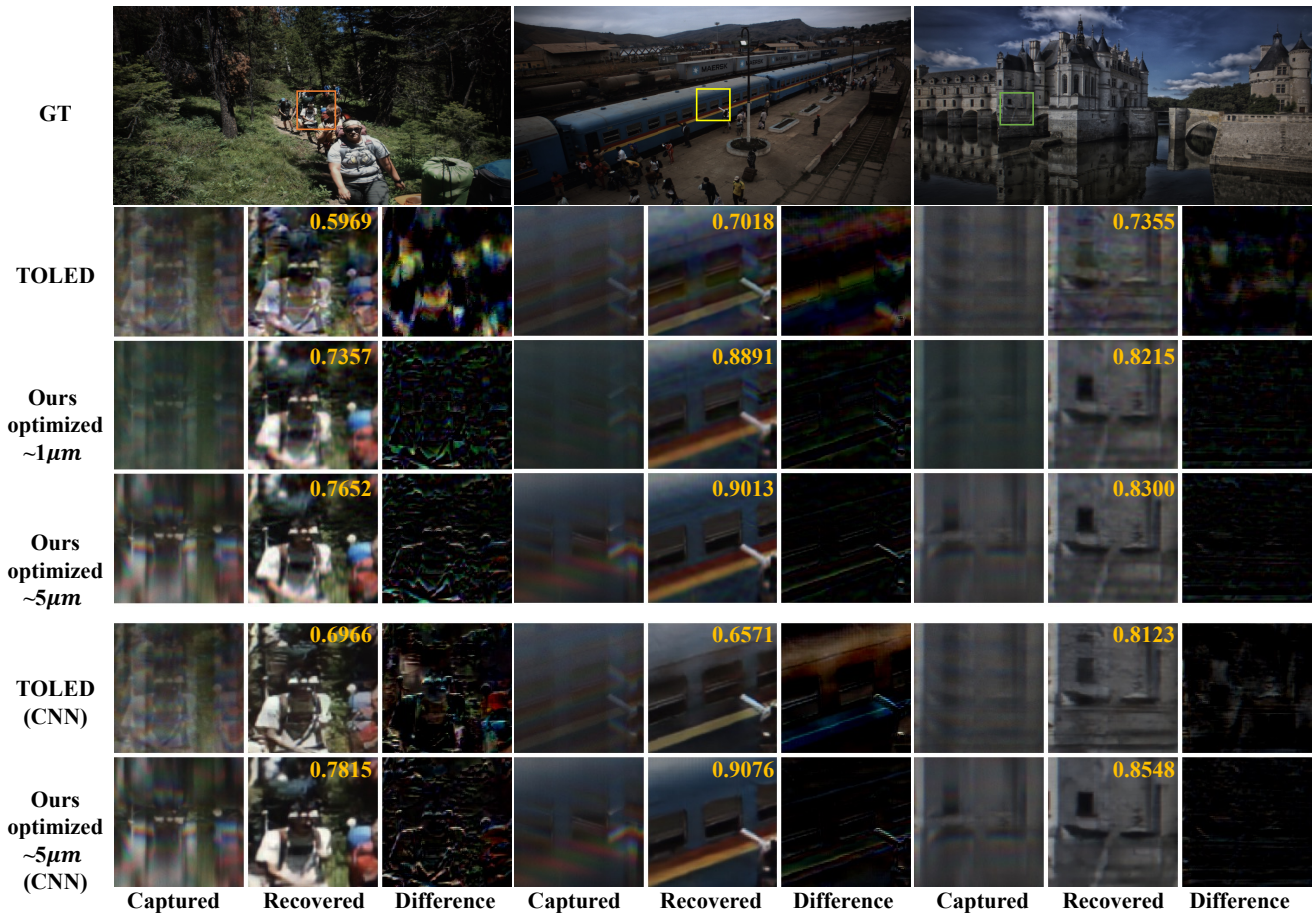


Figure 8: **Qualitative results from UDC under TOLED and our setups.** For each scene, columns from left to right show captured, recovered, and difference maps between the restored and ground truth. The intensities of difference maps are magnified by 2 times. We show SSIM for each restored image. Higher score means better quality.

tups outperform TOLED at all light levels. At $1\,\mu\text{m}$, the optimized height map largely outperforms the fixed one; while at $5\,\mu\text{m}$, different designs perform similarly. Because thinner phase masks have more phase wrappings and are more sensitive to the selection of d_0. At $5\,\mu\text{m}$, the phase mask is quite similar to a thick lens and the system performance is more consistent across different choices of d_0. Comparisons on validation set is in the supplementary.

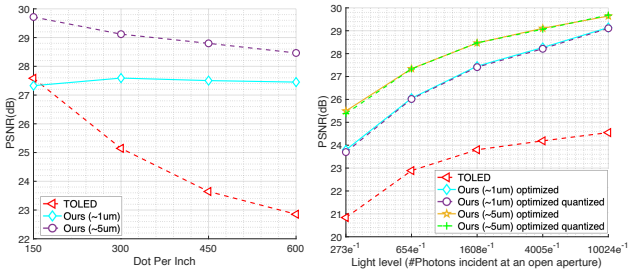
Effect of Optimization. Figure 7 compares qualitative results of three choices of wrapping heights on ISO 12233 resolution chart [1]. The light level is about $10{,}000$ photons. The restored image of TOLED contains a significant amount of ringing artifacts. The phase mask designed at a fixed height and with $1\,\mu\text{m}$ thickness produces apparent chromatic aberration. The image captured under uniformly sampled heights appears less greenish. Phase masks with optimized heights further suppress chromatic artifacts and retain more details, and the one at $5\,\mu\text{m}$ performs even better than the one at $1\,\mu\text{m}$. For example, it recovers more high-frequency details on the circle.

Qualitative Results. Figure 8 shows the qualitative results of TOLED and those of our setups. Light level is around $10{,}000$ photons. The upper rows show results from a naive iterative solver, and the lower two rows are from the cutting-edge CNN for UDCs [5]. Ours are consistently better than TOLED in SSIM. It is worth noting that TOLED results, even with CNN, contain apparent ringing artifacts. For example, in the first scene, ghosting artifacts appear on the blue and red bags, and in the third scene, there is an extra copy of the window left in the restored image.

Comparisons with Other OLED Displays. We compare our design with two display layouts commonly used in smartphone screens, TOLED and POLED [27, 26], and two displays layouts designed specifically for UDCs [24, 5]. POLED contains a poly-amide substrate, which causes extremely low light throughput of around $8\,\%$ and produces

	Display Changes	LTR%	PSNR / SSIM
TOLED	–	23.8	22.42 dB / 0.59
POLED	–	8.3	26.22 dB / 0.67
Ours	–	47.6	28.01 dB / 0.75
Yang et al.	Modify layout	22.6	32.93 dB / 0.88
ZTE Axon	Low DPI	~ 75	38.24 dB / 0.96

Table 1: **Comparisons with other OLED displays.** TOLED, POLED, and ours do not require a change to the display openings, while Yang et al. and ZTE Axon require significant modifications to the display layout. We list averaged PSNR(↑) and SSIM(↑) across scenes from typical indoor to outdoor light levels, from 250 to 10,000 photons.



(a) Effect of pixel densities (b) Effect of phase mask quantization

Figure 9: **Effect of (a) setups with varying display pixel densities and (b) quantization of phase masks.**

a yellowish color shift in the captured images. The display designed by Yang and Sankaranarayanan [24] modifies the display openings, and subsequently requires significant engineering effort to accommodate display RGB subpixels and circuits. ZTE Axon 20 phone largely reduces the display pixel density to make room for transparent regions for light to pass through. We evaluate the performance of ZTE using the PSF provided by Feng et al. [5] and an estimated LTR of around 75 %. Note that reducing the pixel density results in apparent artifacts on the display.

Table 1 summarizes the design, LTR, and imaging performance of UDCs under various OLED displays. Ours falls into the category of requiring no change of the display openings and outperforms the other two common displays, TOLED and POLED. While Yang et al. [24] and ZTE Axon have higher imaging quality, the modifications of the display have non-trivial negative effects on the display quality. Detailed performance at different light levels and qualitative results are in the supplementary.

Effect of Pixel Density. Figure 9(a) evaluates the performance of UDCs at various pixel densities. Displays at 150 DPI are commonly used for desktop monitors and laptops; 600 DPI for high-quality cellphone displays and tablets. Light level is around 1,600 photons. At 5 µm, optimized phase masks outperform TOLED at all four pixel densities, and at 1 µm ours outperform TOLED with pixel densities larger than 300 DPI. Because microlens arrays for larger pixel pitch have larger radii, and results in phase warpping artifacts when implemented as thin plates. Additional SSIM plots are shown in the supplementary.

CNN-based Restoration. We adopt DISCNet [5], one of the best UDC restoration networks. We utilize the 240 high-quality images in UDC dataset [26] and simulate the captured images using the pipeline described earlier. Figure 1 and Figure 8 showcase restored images for TOLED and ours. Compared to the naive iterative solver, CNN-based restoration largely improves imaging quality. The proposed setup consistently outperforms TOLED.

Effect of phase mask quantization. In fabrication, phase masks are often quantized into discrete height maps with a step of 200 nm, for example in two-photon lithogaphy [2]. Figure 9(b) shows that quantized phase masks perform similarly as ones before quantization.

6. Discussions

In this paper, we design phase masks to improve the image quality of UDCs. First, we show that inserting one phase mask behind the display is ineffective. Second, we propose to place two MLAs in front of and behind the display. The first MLA concentrates light to locations where the display is open, and the second recovers the original wavefront. The proposed design allows more light to reach the camera main lens and shapes the wavefront to a better condition. To ensure the display quality uncompromised, we implement microlens arrays as polarization-dependent phase masks and optimize their heights to suppress chromatic aberration. The proposed design largely improves the imaging quality of UDCs under TOLED display.

Scene at different depths. The effect of the proposed phase masks is nearly constant across scenes at different depths in the working range of selfie cameras. This is due to the small focal lengths of the proposed microlens arrays, which are at the scale of hundreds of microns.

Diffraction blur. A byproduct of inserting phase masks a short distance away from the display is that the captured images lose some details towards the edge (see Figure 1(b)). Similar to diffractive grating, the diffraction becomes apparent as the angle of incident light increases. In contrast, TOLED retains those details, however, with widespread ringing artifacts that are difficult to remove even with SOTA deep neural networks (Figure 1(a)). Consequently, our method yields much more visually appealing reconstructions and higher benchmark scores.

Acknowledgements. This work was supported by Global Research Outreach program of Samsung Advanced Institute of Technology and the NSF CAREER award CCF-1652569.

References

[1] ISO 12233 photography. https://www.graphics.cornell.edu/~westin/misc/ISO_12233-reschart.pdf.

[2] Nanoscribe. https://www.nanoscribe.com/en/.

[3] Jan Biemond, Reginald L Lagendijk, and Russell M Mersereau. Iterative methods for image deblurring. *Proceedings of the IEEE*, 78(5):856–883, 1990.

[4] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[5] Ruicheng Feng, Chongyi Li, Huaijin Chen, Shuai Li, Chen Change Loy, and Jinwei Gu. Removing diffraction image artifacts in under-display camera via dynamic skip connection network. In *CVPR*, 2021.

[6] KeMing Gao, Meng Chang, Kunjun Jiang, Yaxu Wang, Zhihai Xu, Huajun Feng, Qi Li, Zengxin Hu, and YueTing Chen. Image restoration for real-world under-display imaging. *Optics Express*, 29(23):37820–37834, 2021.

[7] Joseph W Goodman. *Introduction to Fourier optics*. Roberts & Co., 2005.

[8] Felix Heide, Qiang Fu, Yifan Peng, and Wolfgang Heidrich. Encoded diffractive optics for full-spectrum computational imaging. *Scientific reports*, 6(1):1–10, 2016.

[9] Jie Hu, Sankhyabrata Bandyopadhyay, Yu-hui Liu, and Liyang Shao. A review on metasurface: from principle to smart metadevices. *Frontiers in Physics*, 8:586087, 2021.

[10] Daniel S Jeon, Seung-Hwan Baek, Shinyoung Yi, Qiang Fu, Xiong Dun, Wolfgang Heidrich, and Min H Kim. Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Transactions on Graphics*, 2019.

[11] Jaihyun Koh, Jangho Lee, and Sungroh Yoon. BNUDC: A two-branched deep neural network for restoring images from under-display cameras. In *CVPR*, 2022.

[12] Kinam Kwon, Eunhee Kang, Sangwon Lee, Su-Jin Lee, Hyong-Euk Lee, ByungIn Yoo, and Jae-Joon Han. Controllable image restoration for under-display camera in smartphones. In *CVPR*, 2021.

[13] Sida Li, Yueda Liu, Yan Li, Shuxin Liu, Shuyi Chen, and Yikai Su. Fast-response pancharatnam-berry phase optical elements based on polymer-stabilized liquid crystal. *Optics Express*, 27(16):22522–22531, 2019.

[14] Shichao Nie, Chengconghui Ma, Dafan Chen, Shuting Yin, Haoran Wang, LiCheng Jiao, and Fang Liu. A dual residual network with channel attention for image restoration. In *ECCV*, 2020.

[15] Youngjin Oh, Gu Yong Park, Haesoo Chung, Sunwoo Cho, and Nam Ik Cho. Residual dilated u-net with spatially adaptive normalization for the restoration of under display camera images. In *2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2021.

[16] Yifan Peng, Qiang Fu, Felix Heide, and Wolfgang Heidrich. The diffractive achromat full spectrum computational imaging with diffractive optics. In *ACM Transactions on Graphics*, pages 1–2. 2016.

[17] Yifan Peng, Qilin Sun, Xiong Dun, Gordon Wetzstein, Wolfgang Heidrich, and Felix Heide. Learned large field-of-view imaging with thin-plate optics. *ACM Transactions on Graphics*, 38(6):219–1, 2019.

[18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015.

[19] Zheng Shi, Yuval Bahat, Seung-Hwan Baek, Qiang Fu, Hadi Amata, Xiao Li, Praneeth Chakravarthula, Wolfgang Heidrich, and Felix Heide. Seeing through obstructions with diffractive cloaking. *ACM Transactions on Graphics*, 41(4):1–15, 2022.

[20] Vincent Sitzmann, Steven Diamond, Yifan Peng, Xiong Dun, Stephen Boyd, Wolfgang Heidrich, Felix Heide, and Gordon Wetzstein. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Transactions on Graphics*, 37(4):1–13, 2018.

[21] Varun Sundar, Sumanth Hegde, Divya Kothandaraman, and Kaushik Mitra. Deep atrous guided filter for image restoration in under display cameras. In *ECCV*, 2020.

[22] Yicheng Wu, Vivek Boominathan, Huaijin Chen, Aswin Sankaranarayanan, and Ashok Veeraraghavan. PhaseCam3D–learning phase masks for passive single view depth estimation. In *IEEE International Conference on Computational Photography*, 2019.

[23] Anqi Yang, Eunhee Kang, Hyong-Euk Lee, and Aswin C. Sankaranarayanan. Design phase masks for under-display cameras: Software. https://github.com/Image-Science-Lab-cmu/UDC_Phase_ICCV23, 2023.

[24] Anqi Yang and Aswin C Sankaranarayanan. Designing display pixel layouts for under-panel cameras. *IEEE TPAMI*, 43(7):2245–2256, 2021.

[25] Qirui Yang, Yihao Liu, Jigang Tang, and Tao Ku. Residual and dense unet for under-display camera restoration. In *ECCV*, 2020.

[26] Yuqian Zhou, Michael Kwan, Kyle Tolentino, Neil Emerton, Sehoon Lim, Tim Large, Lijiang Fu, Zhihong Pan, Baopu Li, Qirui Yang, et al. Udc 2020 challenge on image restoration of under-display camera: Methods and results. In *ECCV*, 2020.

[27] Yuqian Zhou, David Ren, Neil Emerton, Sehoon Lim, and Timothy Large. Image restoration for under-display camera. In *CVPR*, 2021.