BiDi Screen: A Thin, Depth-Sensing LCD for 3D Interaction using Light Fields
Supplementary Appendix

Matthew Hirsch[1] Douglas Lanman[2] Henry Holtzman[1] Ramesh Raskar[1]
[1]MIT Media Lab [2]Brown University

A. Optimizing the Mask Properties

In this appendix we describe how to optimize the sensor-mask (or diffuser-mask) separation, for pinhole arrays and tiled-broadband codes. As with other light field cameras, the total number of samples (given by the product of the spatial and angular resolutions) cannot exceed the number of camera pixels. In our system the LCD discretization further limits the mask resolution, restricting the total number of light field samples to be approximately equal to the total number of pixels in the display (i.e., 1680×1050 pixels). Thus, as described in Section 6.1, we achieve a spatial resolution of 73×55 samples and an angular resolution of 19×19 samples with a pinhole or MURA tile spacing of d_p = 4.92 mm and a mask separation of $d_i = 2.5$ cm. However, by adjusting the spacing and separation, the spatio-angular resolution trade-off can be adjusted.

A.1. Pinhole Array Mask Configuration

As shown in Figure 3, each pinhole must be separated by $d_p = 2d_i \tan(\alpha/2)$ if diffraction is negligible (otherwise Equation 2 must be used). Thus, the pinhole array separation d_i is given by

$$d_i = \frac{d_p}{2\tan(\alpha/2)}. \qquad \text{(A.1)}$$

The field of view α, shown in Figure 4, is determined by the vignetting for each sensor pixel (e.g., that due to the diffuser and camera's field of view) or by an angle-limiting film. Wider fields of view may be desirable for some applications. However, for a fixed field of view, one must choose the mask separation d_i to optimize the effective spatial resolution in front of the display. Thus, Equation 4 can be used to maximize $N_{spatial}$ as a function of d_i. In our design we assume an average object distance of $d_o = 25$ cm. Duplicating Figure 6, we plot the effective spatial resolution as a function of the mask separation d_i (see Figure A.1). Note that the selected distance $d_i = 2.5$ cm is close to the maximum, allowing slightly higher angular resolution (via Equation 3) without a significant reduction in spatial resolution.

A.2. Tiled-Broadband Mask Configuration

The placement and design of tiled-broadband masks was described in [Lanman et al. 2008]. However, their design was for a transmission-mode system with a uniform array of LEDs placed a fixed distance in front of the sensor. Our reflection-mode system requires the mask be placed at a different distance from the sensor to allow light field capture. In this section the notation and derivation mirrors that paper. We describe 2D light fields and 1D sensor arrays, however the extension to 4D light fields and 2D sensor arrays arrives at a similar mask separation d_m.

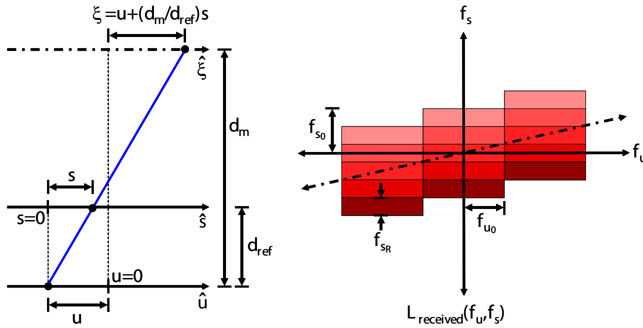

Figure A.2: *Geometric derivation of tiled-broadband mask separation. (Left) The two-plane parameterization (u,s) of the optical ray shown in blue. The ray intersects the mask at $\xi = u + (d_m/d_{ref})s$. (Right) The received light field spectrum contains multiple spectral replicas shown in shades of red. The shield field spectrum must lie along the dashed line (i.e., $f_s/f_u = (d_m/d_{ref}) = f_{s_R}/(2f_{u_0})$).*

As shown in Figure A.2, consider the two-plane parameterization, where u denotes the horizontal coordinate (along the sensor or diffuser) and s denotes the horizontal position of intersection (in the local frame of u) of an incident ray with a plane that is a distance $d_{ref} = 1$ cm away from, and parallel to, the first plane. A mask separated by d_m from the sensor creates a *shield field* that acts as a volumetric occlusion function $o(u,s) = m(u + (d_m/d_{ref})s)$. Thus, each ray is attenuated by the occlusion function $m(\xi)$ evaluated at $\xi = u + (d_m/d_{ref})s$. Taking the 2D Fourier transform yields the mask's shield field spectrum $O(f_u, f_s)$, given by

$$O(f_u, f_s) = M(f_u)\delta(f_s - (d_m/d_{ref})f_u), \qquad \text{(A.2)}$$

where $M(f_\xi)$ is the 1D Fourier transform of $m(\xi)$. As described in [Lanman et al. 2008], the effect of the mask can be modeled by convolving the incident light field spectrum $L_{incident}(f_u, f_s)$ with the shield field spectrum $O(f_u, f_s)$. This implies that the mask spectrum must be composed of a series of impulses (i.e., a Dirac comb), with the tiled-MURA mask being one such valid pattern when the tile dimensions are equal to the pinhole spacing d_p.

The mask separation d_m must be adjusted so the received image can be decoded to recover the incident light field. Assume that $L_{incident}(f_u, f_s)$ is bandlimited to f_{u_0} and f_{s_0}, as shown in Figure A.2. Since Equation A.2 implies that the mask spectrum lies along the line $f_s = (d_m/d_{ref})f_u$, we conclude that

$$d_m = \frac{d_{ref} f_{s_R}}{2 f_{u_0}} = \frac{d_p}{2\tan(\alpha/2)}, \qquad \text{(A.3)}$$

where $f_{s_R} = 1/(2d_{ref}\tan(\alpha/2))$ and $f_{u_0} = 1/(2d_p)$. Note that Equations A.1 and A.3 imply that the pinhole array and tiled-broadband codes must be the same distance away from the sensor.

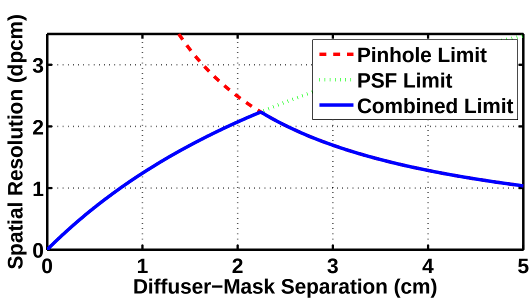

Figure A.1: *Effective spatial resolution as a function of diffuser-mask separation d_i for a pinhole array, as given by Equation 4. System parameters correspond with the prototype in Section 6.1.*

References

LANMAN, D., RASKAR, R., AGRAWAL, A., AND TAUBIN, G. 2008. Shield fields: Modeling and capturing 3d occluders. *ACM Trans. Graph. 27*, 5.