Learned Display Radiance Fields with Lensless Cameras

ZIYANG CHEN, University College London, United Kingdom
YUTA ITOH, University of Tokyo, Japan
KAAN AKŞIT, University College London, United Kingdom

1 HARDWARE SETUP

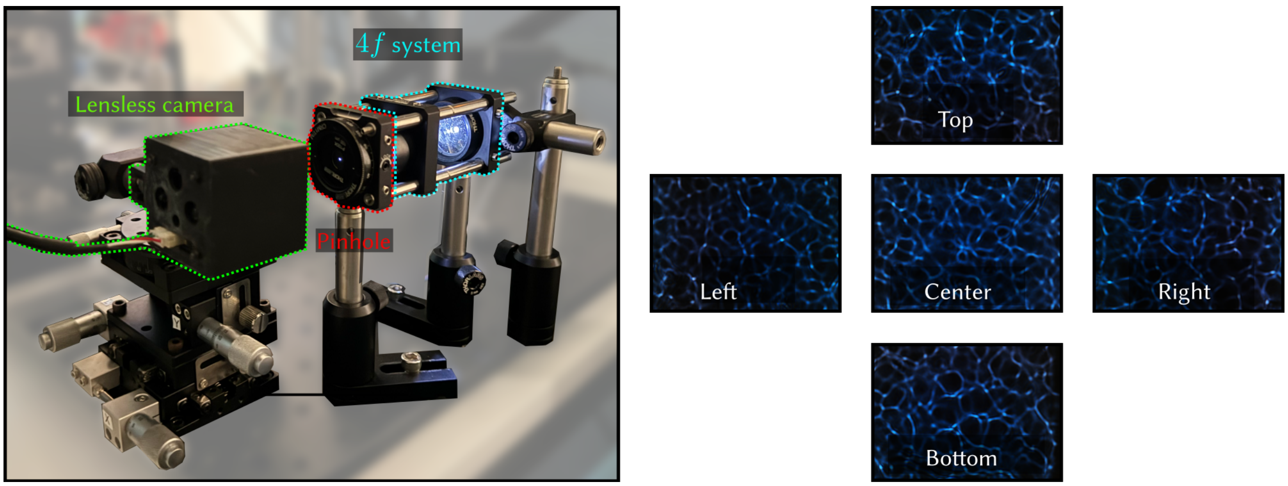

Fig. 1. The setup for capturing the Point Spread Function (PSF) (left). The captured PSF from each aperture (right).

Point Spread Function Capture. We place the lensless camera in front of a white LED source and acquire PSFs at multiple incident angles as shown on the left of Fig. 1. A $4f$ system widens the angular spread and separates spatial-frequency components at the Fourier plane. A pinhole at the Fourier plane enables estimation of the display pixel pitch. This setup lets the recorded PSFs reflect the camera response to light from single display pixels. During acquisition, a linear stage translates the camera laterally to sample apertures across angles. The resulting PSFs are shown on the right of Fig. 1.

Pixel Light Field Capture. Facing the panel, the lensless camera records light fields while we sequentially illuminate individual pixels with steps of 51 rows and 57 columns. The reconstruction pipeline first aligns each capture to the base PSF via cross-correlation to remove shifts and rotations. We then downsample the calibrated images and crop the Region Of Interest (ROI) to 2160×2800 pixels to match the base PSF. We compute the incident angle as $\alpha = 2\tan^{-1}(n/m)$, where n denotes the aperture size and m is the sensor–aperture distance.

2 IMPLICIT NEURAL REPRESENTATION

Positional encoding augments the Multi-Layer Perceptron (MLP) input coordinates to model fine-scale variation in pixel emission. We analyze the frequency levels for each coordinate group and report their effect on novel-pixel prediction in Fig. 2. Next, the training objective considers several losses. Specifically, Fig. 3(a) visualizes novel-pixel reconstructions under L_1, L_2, and $\mathcal{R}$ alone and in combination. The Implicit Neural Representation (INR) synthesizes display behavior at pixel resolution, as illustrated in Fig. 3(b).

Authors' Contact Information: Ziyang Chen, University College London, London, United Kingdom, ziyang.chen.22@ucl.ac.uk; Yuta Itoh, University of Tokyo, Tokyo, Japan, yuta.itoh@iii.u-tokyo.ac.jp; Kaan Akşit, University College London, London, United Kingdom, k.aksit@ucl.ac.uk.

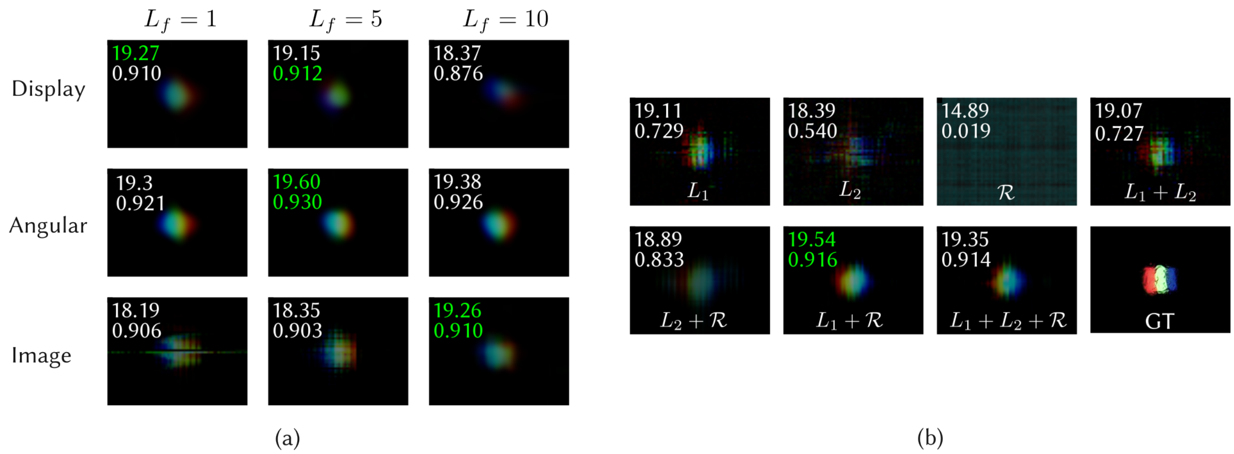

Fig. 2. (a) The novel pixels generated with positional encoding frequency levels from 1 to 10 for each coordinate groups. (b) The novel pixel view generated with different loss functions. The $\mathcal{R}$ denotes the range loss that regulates the model to keep the predicted light field within the range of $[0, 1]$. We show the PSNR and SSIM scores in the top left corner of each image.

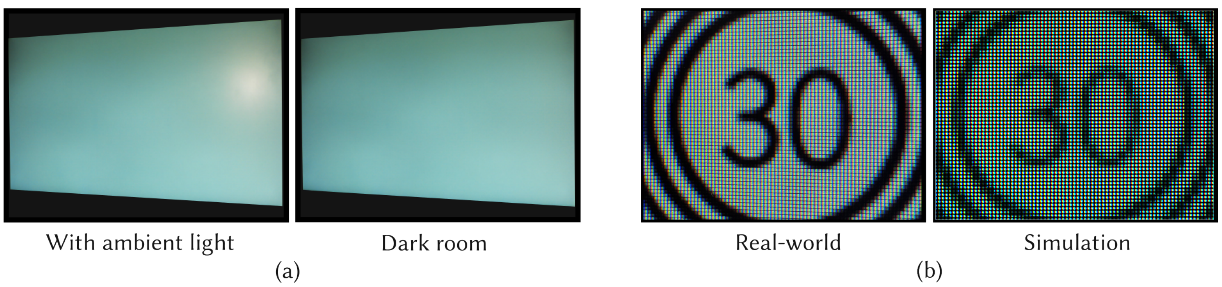

Fig. 3. (a) The display captured with ambient lights (left). The display captured in a dark room (right). (b) The zoom-in view of the display captured with a conventional camera (left). The simulated display with our INR model (right).