MetaHDR: Supplementary Material

This document accompanies the main paper for additional details and information.

1. DETAILS OF THE CALIBRATION PROCESS

Geometric alignment. Consider the coordinate system in Fig. 5 in the main paper. A 3D point with the coordinate $\mathbf{X} = [X, Y, Z]^T$ is projected to nine corresponding points in the nine sub-images. The projection model of each sub-image I_i is:

$$\mathbf{x}_i = R_{\text{meta}} D_i K R_{\text{meta}}^{-1} \mathbf{X}, \tag{S1}$$

where the 2D point $\mathbf{x}_i$ with homogeneous coordinate $[x_i, y_i, 1]^T$ is the projected point on sub-image I_i. The matrix $R_{\text{meta}} \in \mathbb{R}^{3\times3}$ is the rotation matrix from the metasurface normal vector $\mathbf{n}$ to the z-direction $\mathbf{z}$. The matrix D_i models the deflection of the ith beam by the metasurface to form sub-image I_i:

$$D_i = \begin{bmatrix} 1 & 0 & \Delta x_i \\ 0 & 1 & \Delta y_i \\ 0 & 0 & 1 \end{bmatrix}. \tag{S2}$$

The matrix K is the intrinsic matrix of the camera. Thus, the correspondence between two 2D points $\mathbf{x}_i = [x_i, y_i, 1]$ and $\mathbf{x}_j = [x_j, y_j, 1]$ in sub-image I_i and I_j can be modeled by the homography H_{ij}:

$$\mathbf{x}_i = H_{ij}\mathbf{x}_j, \quad H_{ij} \in \mathbb{R}^{3x3}, \tag{S3}$$

where

$$H_{ij} = R_{\text{meta}} D_i D_j^{-1} R_{\text{meta}}^{-1}. \tag{S4}$$

We can use the standard homography calibration methods to estimate H_{ij} using at least four corresponding points $\mathbf{x}_i$ and $\mathbf{x}_j$ for every pair of sub-images I_i and I_j [1]. However, as all sub-images have different dynamic ranges, it is possible that a key point $\mathbf{x_i}$ can be clearly determined from the visual features of I_i while the corresponding point $\mathbf{x_j}$ cannot be located in sub-image I_j as the region in I_j is under or over-exposed. To tackle this problem, we capture multiple full images with different exposures of the same scene and select each sub-image I_i from the full image with the most proper exposure. We use a calibration target with a regular grid of disks and the Hough circle detection to locate corresponding key points between sub-images, as shown in Fig. 6a-b in the main paper.

Contrast registration. After estimating the unnormalized power ratios $\tilde{\alpha}_i$ as described in Sec. 5.2 in the main paper, we normalize them via:

$$\alpha_i = \tilde{\alpha}_i \cdot \frac{\sum_{x,y} \left(I_2(x,y) - \tilde{B}_2(x,y)\right)/t_2}{\sum_{x,y} I_i(x,y)/t_i}, \tag{S5}$$

where t_i is the exposure time that we use to take a properly exposed sub-image $I_i(x,y)$, and $\tilde{B}_2(x,y)$ is the background intensity in the sub-image $I_2(x,y)$, which is estimated using the average intensities of a portion of image pixels outside the field of view.

2. DETAILS OF LIGHT EFFICIENCY ANALYSIS

We consider exposure mosaicking methods [2–4] that use mosaicked pixels with reduced light efficiency as a comparison to the proposed MetaHDR method. Ideally, the light efficiency of a V-multiplex exposure mosaicking imager, with power ratios of α between adjacent multiplex, is:

$$\eta_{\text{mosaic}}(V;\alpha) = \frac{1-\alpha^V}{V(1-\alpha)}. \tag{S6}$$

For 9-multiplex and a power ratio of 0.5, the ideal light efficiency of exposure mosaicking is $\eta_{\text{mosaic}}(9;0.5) = 22.2\%$. For MetaHDR, we estimate its light efficiency based on the normalized power ratios $\{\alpha_i, i = 1, \cdots, V\}$ and the overall light efficiency of the metasurface nanocells:

$$\eta_{\text{meta}}(V) = \text{mean}_{m,n}\left(T(m,n)^2\right) \cdot \left(\sum_{i=1}^{V} \alpha_i\right), \tag{S7}$$

where the term $\text{mean}_{m,n}\left(T(m,n)^2\right)$ estimates the overall power that is transmitted through the metasurface. The estimated light efficiency of the MetaHDR prototype is $\eta_{\text{meta}}(V) = 53.7\%$ under a 10nm bandwidth centered at the working wavelength 650nm. Therefore, the MetaHDR prototype could achieve 2.42 times higher light efficiency than an ideal exposure mosaicking imager when the application only requires a narrow band of operating wavelengths.

We then analyze the effect of light efficiency on the quality of HDR reconstruction for the two methods. We capture two full images of an HDR scene, $I_1(x, y, t_{\text{mosaic}})$ and $I_2(x, y, t_{\text{meta}})$ using MetaHDR with the exposure times $t_{\text{meta}} = 2.42 t_{\text{mosaic}}$. The two full images I_1, I_2 correspond to the estimated light efficiency of exposure mosaicking and MetaHDR. We reconstruct the spatial gradient of the irradiance map from each of the full images using the gradient-based HDR fusion method described in this paper. Table 3 in the main paper reports the quantitative accuracy of the reconstruction compared to the ground-truth HDR measurements captured with a long exposure time and low gain. We repeat the experiment for 3 different scenes in total and report the average number. Fig. S1 shows the cross-section of sample HDR reconstruction results corresponding to the two light efficiencies.

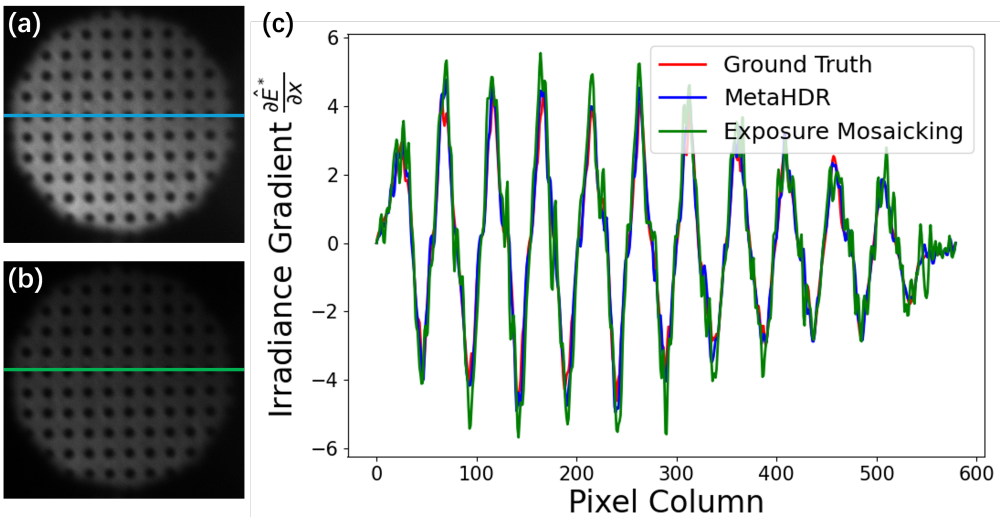

Fig. S1. HDR reconstruction for light efficiency comparison. (a) Sub-image of a textured plane captured at exposure t_{meta}. (b) The same sub-image of the same scene captured at exposure t_{mosaic}. (c) Cross-section of the reconstructed spatial gradient of the irradiance map. The intensity profile that corresponds to the light efficiency of exposure mosaicking is noisier.

3. ADDITIONAL RESULTS

A. HDR Video

We include 2 HDR videos captured using MetaHDR, with information listed below:

- HDRvideo_flame.gif. Without the objective lens, we used MetaHDR to record an HDR video of a waving flame (same scene as Fig.7(c) in the main paper).
- HDRvideo_circuit board.gif. With the objective lens, we used MetaHDR to scan the surface of a circuit board (same scene as Fig.7(e) in the main paper).

B. Single-shot Surface Reflectance Calibration

Fig. S2 shows additional results of surface reflectance calibration on two glossy materials.

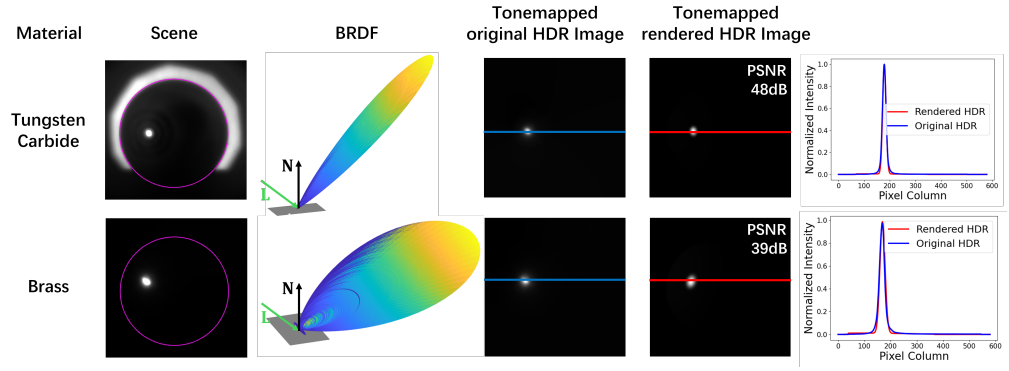

Fig. S2. Single-shot surface reflectance calibration using MetaHDR. Similar to Fig. 9 in the main paper, we calibrate the surface reflectance of two metal materials using precision balls, visualize the reflectance as the bidirectional radiance distribution functions (BRDFs), and render the image of the precision ball under directional lighting to validate the accuracy of the calibration.

C. Single-shot surface curvature estimation

Using the reflectance models shown in Fig. S2, we can estimate the surface curvature of unknown shapes made of the two materials, as shown in Fig. S3.

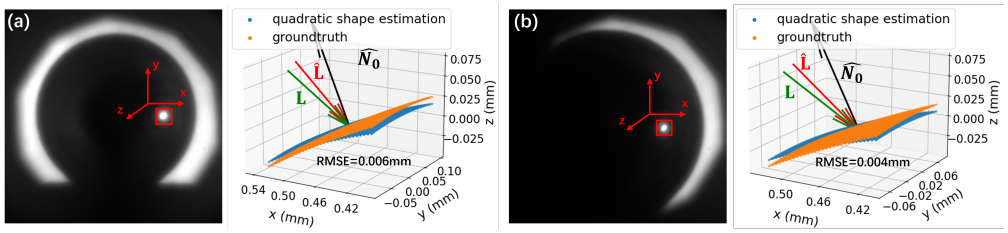

Fig. S3. Single-shot surface curvature estimation. We use the surface reflectance measured in advance from two balls in Fig. S2 to reconstruct the shape of unknown surfaces of the same materials at the specularities (red patches) and the lighting direction. The materials of objects are Tungsten Carbide (a) and Brass (b). We plot the reconstructed shapes and overlay them with ground truths, and visualize the estimated lighting direction $\widehat{\mathbf{L}}$, the estimated surface normal of the brightest pixel $\widehat{\mathbf{N}_0}$, and the ground truth lighting direction $\mathbf{L}$. The average error of the estimated lighting direction $\widehat{\mathbf{L}}$ is 9.2°.

REFERENCES

1. R. Hartley and A. Zisserman, *Multiple view geometry in computer vision* (Cambridge university press, 2003).
2. S. Nayar and T. Mitsunaga, "High dynamic range imaging: spatially varying pixel exposures," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, vol. 1 (2000), pp. 472–479 vol.1.
3. M. Xie, M. Chan, and C. Metzler, "Snapshot high dynamic range imaging with a polarization camera," arXiv preprint arXiv:2308.08094 (2023).
4. H. Cho, S. J. Kim, and S. Lee, "Single-shot high dynamic range imaging using coded electronic shutter," Comput. Graph. Forum **33**, 329–338 (2014).