

OpticFusion: Multi-Modal Neural Implicit 3D Reconstruction of Microstructures by Fusing White Light Interferometry and Optical Microscopy

Supplementary Material

A. Network Details

The input spatial coordinates are initially encoded using the multi-resolution hash technique [30]. This encoding involves 16 levels of hash feature grids, each with feature dimensions of 2. The coarsest grid starts with a resolution of 16, and each subsequent level increases by a scale factor of 1.447. The SDF network consists of an MLP with one hidden layer of size 64, utilizing ReLU activation. For the global and residual color networks, each is structured as an MLP with two hidden layers of size 64, also employing ReLU activation. The output from these color networks is normalized to a range of 0 to 1 using the sigmoid function.

B. Computation of Roughness Parameters

As demonstrated in Sec. 4.3 of our main paper, we provide two roughness parameters for microscale surfaces: the arithmetical mean height of the surface (Sa) and the root mean square height of the surface (Sq). For a selected surface area, we first fit the surface to a plane and then rotate and translate it to align with the XY plane. Then, on this transformed plane, N points are uniformly sampled at equal intervals along the XY coordinates, and the Z-axis heights z of these points are recorded. The arithmetic mean of the heights \bar{z} is calculated as follows:

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i. \quad (\text{B1})$$

Sa is defined as the mean absolute height difference between the surface points and the mean plane:

$$\text{Sa} = \frac{1}{N} \sum_{i=1}^N |z_i - \bar{z}|. \quad (\text{B2})$$

Sq represents the standard deviation of the point heights:

$$\text{Sq} = \sqrt{\frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2}. \quad (\text{B3})$$

C. Generation of Synthetic WLI Data

Here, We demonstrate how we generate synthetic WLI data in our synthetic experiment. Real WLI data has high measurement accuracy and can be regarded as noise-free depth maps in the simulated environment. Moreover, WLI data often contains voids because the CCD camera fails to capture interference fringes with sufficient contrast on some

surfaces. This issue typically arises when the sample's tilt angle α (the angle between the surface normal and the Z-axis) is too large, preventing the white light from reflecting back into the objective lens. As discussed in Sec. 4.4, for a perfectly specular micro surface, this tilt angle limit can be calculated from the numerical aperture (NA) of the objective lens, NA is 0.13 in our real experiments. On the other hand, rough surfaces with larger tilt angles can still have measurements. We summarize the characteristics of WLI data generation as follows:

- When the tilt angle is less than $\theta = \arcsin(\text{NA})$, measurements can be obtained on any surface.
- When the tilt angle exceeds θ , specular surfaces are less likely to yield measurements, while diffuse surfaces are more likely to do so.

Our objective is to ensure that the synthetic WLI data replicates these characteristics of real WLI data. To achieve this, we render a depth map d , normal map n , diffuse color c_d , and specular color c_s from each viewpoint in Blender. Using n and the view direction, we calculate the surface tilt angle α for each pixel. We then define $\beta = \frac{c_s}{c_d + c_s}$ as the reflectance of the model surface. Based on the first rule, we design a piecewise function f to calculate the probability p that the depth value of each pixel is valid:

$$p = f(\alpha, \beta) = \begin{cases} 1, & \text{for } \alpha < \theta, \\ g, & \text{for } \alpha \geq \theta. \end{cases} \quad (\text{C4})$$

Following the second rule, we define the function g . We set two intermediate variables $x = \frac{\alpha - \theta}{90 - \theta}$, and $t = 2\beta - 1$:

$$g(x, t) = \begin{cases} 1 - e^{10(1-x)t} \cdot x, & \text{for } t < 0, \\ e^{-10xt} \cdot (1 - x), & \text{for } t \geq 0. \end{cases} \quad (\text{C5})$$

As illustrated in Fig. C1, the f curve closely aligns the characteristics of real WLI data. We then randomly remove values in d according to the calculated p of each pixel and then obtain synthetic WLI data that contain voids.

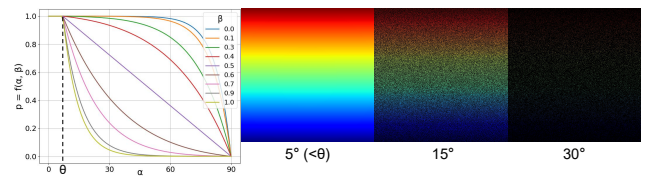


Figure C1. The curve of the function f , and synthetic WLI data for planes with different tilt angles.

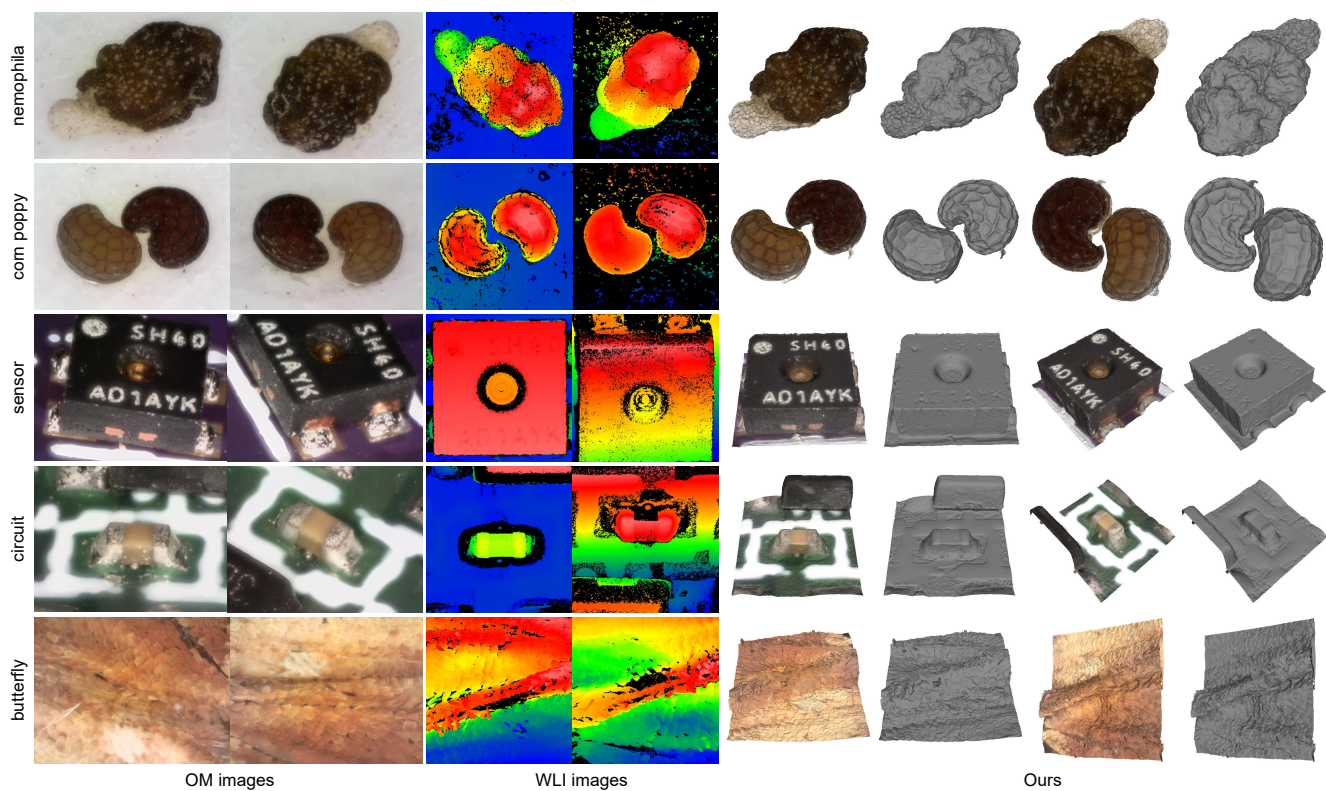


Figure C2. More reconstruction results on the real-world WLI-OM dataset.

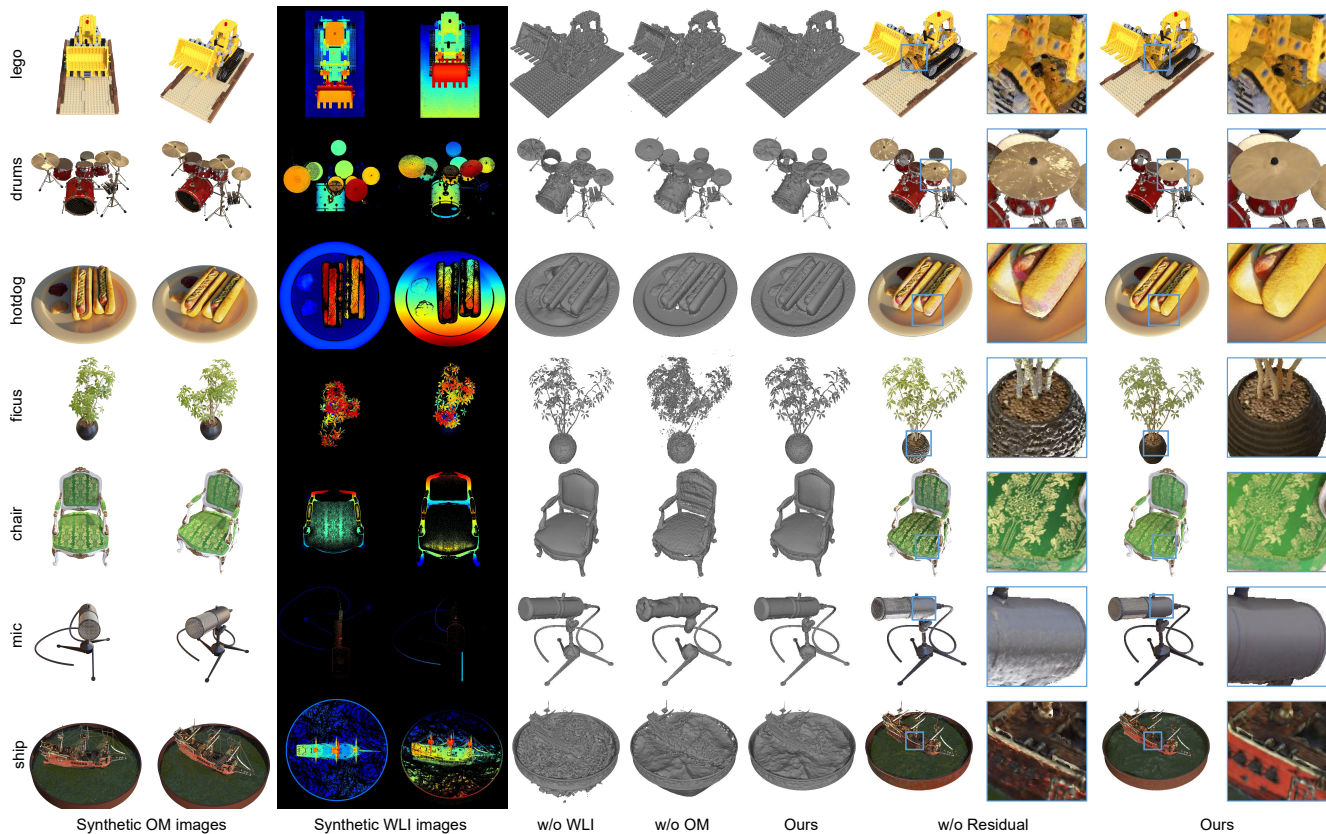


Figure C3. More reconstruction results on the synthetic multi-view WLI and OM dataset.