Quantization-aware Deep Optics
for Diffractive Snapshot Hyperspectral Imaging
Supplementary Material

Lingen Li Lizhi Wang* Weitao Song Lei Zhang
Beijing Institute of Technology
{lingenli, wanglizhi, swt, leizhang}@bit.edu.cn

Zhiwei Xiong
University of Science and Technology of China
zwxiong@ustc.edu.cn

Hua Huang
Beijing Normal University
huahuang@bnu.edu.cn

Due to the space limit, some results and details could not be included in the main paper. Here we provide them for a better demonstration of the performance of our methods. This supplementary material is organized as follows: Sec. 1 provides an in-depth visualization of the DOE quantization results. Sec. 2 supplements some simulation details. Sec. 3 presents complete experiment results. Sec. 4 discusses the limitation of our methods.

1. Quantization Results Visualization

The QDO+A model can adjust the physical heights of the quantized levels with the learnable weights $W_{l\in\{2,\ldots,L-1\}}$ to reduce the deviation caused by the uniform quantization of QDO. We visualize the distribution of the height maps optimized by the full-precision DO, 4-level QDO, and 4-level QDO+A in Fig. 1. Clearly, the height map distribution of the optimized full-precision DOE is not uniform, indicating that the evenly divided levels are not the optimal quantization strategy. Correspondingly, QDO+A can adaptively follow the distribution of the full-precision DOE.

2. Synthetic Simulation Details

To reduce the GPU memory consumption of simulation and the difficulty of fabrication, we use a rotationally symmetric parameterization which calculates the full-precision height map H by rotating a 1-D vector of radius r length:

$$H(x,y) = H_{radius}(\sqrt{x^2+y^2}). \quad (1)$$

where H_{radius} represents the 1-D vector weight of length r. This parameterization idea is similar to the method of [2], but our implementation does not include the interpolation operation because any interpolation will introduce extra levels, which is not appropriate for quantization-aware mod-

*Corresponding author: Lizhi Wang.

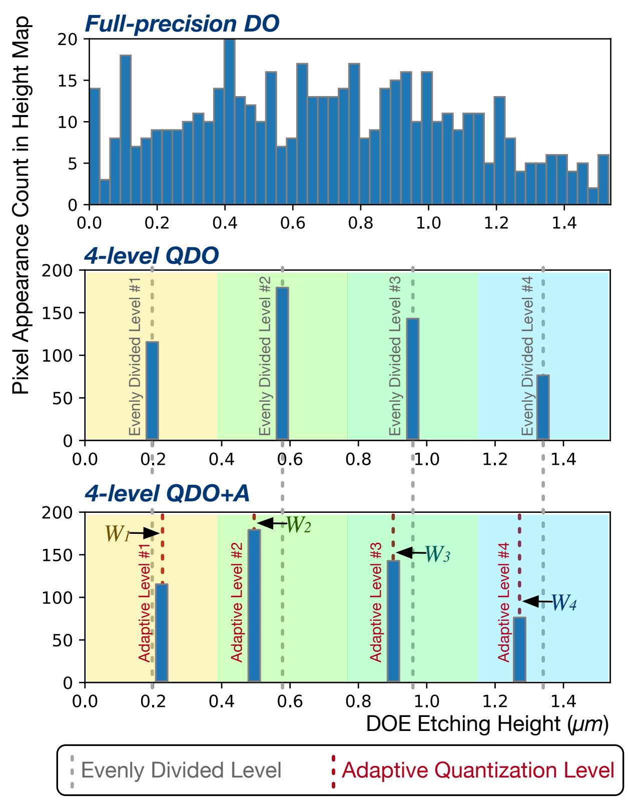

Figure 1. The height distribution of DOEs optimized by the full-precision DO model, the 4-level QDO model, and the 4-level QDO+A model. The base plane thickness $2mm$ is excluded in this figure. The learnable weight $W_{l\in\{2,\ldots,L-1\}}$ indicates the adaptive adjustment from the evenly divided quantization level.

els. In our simulation and experiments, the DOE height maximum h_{max} is set to produce an 2π phase delay at $700nm$ according to the refractive index of $SK1300$ material. Moreover, white Gaussian noise with standard deviation of $20nm$ is added to the height maps during training, to simulate the random fabrication noise and thus make the proposed methods more robust.

3. Physical Experiment Details

In physical experiments, we adopt the network associated with the 4-level QDO+A model trained on ICVL [1] dataset to reconstruct hyperspectral images. All captured scenes are located 1 meter away from the prototype camera, which is consistent with simulation experiments. To avoid color processing inside the imaging pipeline of the camera, we disable gamma correction and only use its linear demosaicking method to convert RAW images to RGB images.

Fig. 2 shows full spectral images from $400nm$ to $700nm$ which are not completely shown in the main paper due to the space limit.

4. Limitation

Effective Scope of the Quantization-aware Model. Our quantization-aware model outperforms the conventional model in situations where the fabrication technique can etch only 32 or fewer levels. When the quantization precision in fabrication becomes higher, the quantization error is no longer one of the main factors affecting the reconstruction performance. In that case, the improvement brought by the quantization-aware method will not be apparent.

Since more levels in DOEs require more complex etching processing of lithography, which also leads to a high cost, the quantization-aware model is a reasonable trade-off between the optimization degree of freedom and the physical limitation of fabrication. This quantization-aware deep optics method is always meaningful before the DOE lithography technology supports etching hundreds of levels or even more precise structures at a very low cost.

Approximation of the Optical Simulation Model. Our optical simulation model is constructed based on the paraxial approximation, *i.e.*, the PSF is spatial-invariant. In order to meet the requirements of paraxial approximation, our system is designed to have a relatively large f-number (*i.e.* $\approx$ 12.2), which leads to a small valid field of view in our imaging system.

Using the spatial-variant PSF model can solve this problem. However, this model has much more computational complexity than the current spatial-invariant model, and it also requires much more GPU memory. Finding an efficient modeling approach of spatial-invariant PSF would be meaningful future work.

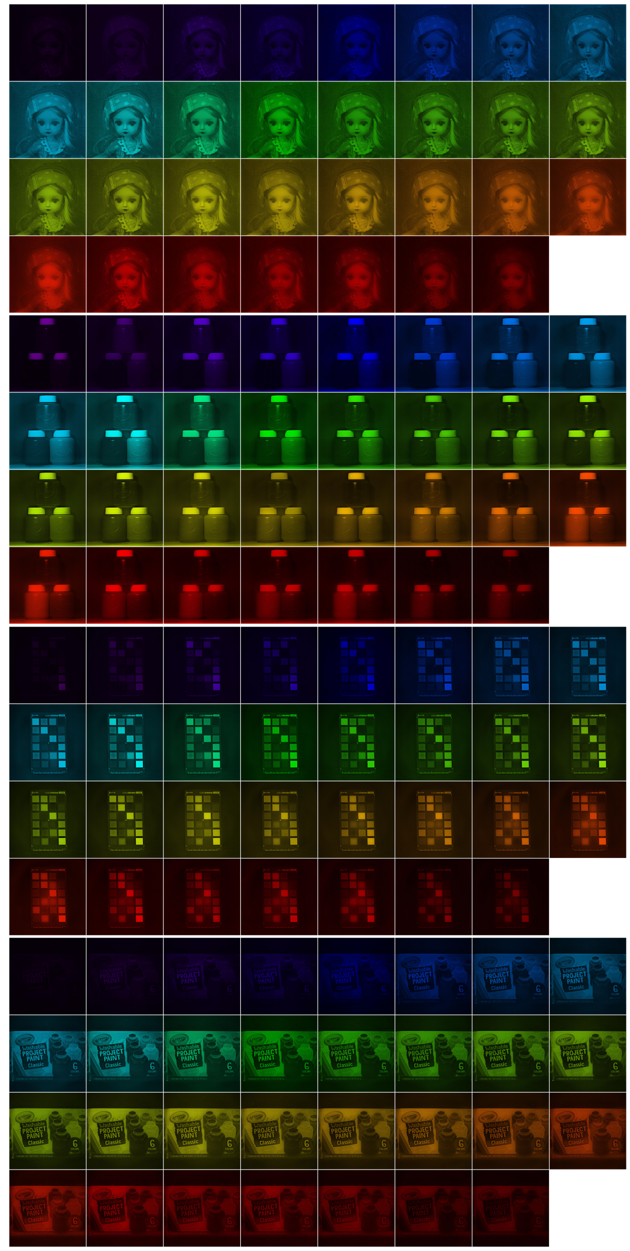

Figure 2. The full spectral images (from $400nm$ to $700nm$ with $10nm$ intervals) reconstructed from the images captured by our prototype camera.

References

[1] Boaz Arad and Ohad Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *European Conference on Computer Vision*, pages 19–3corresponding author4, 2016. 2

[2] Xiong Dun, Hayato Ikoma, Gordon Wetzstein, Zhanshan Wang, Xinbin Cheng, and Yifan Peng. Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging. *Optica*, 7(8):913–922, 2020. 1