1. Overview

In the supplementary material, we first provide a comprehensive account of imaging model in section 2, elucidating the variances in the PSF based on different parameters. Then we expound on the test set employed in the simulation experiments of this paper in section 3. Additionally, detailed discussion and comparison of simulation experiments at lower exposures are presented in Section 4. Ultimately, additional results of ADIS’s reconstruction from the real acquisitions are illustrated in section 5.

2. Detailed process of imaging model

 Imaging forward model. We now consider a multi-slit aperture comprising $N \times N$ parallel rectangular apertures, with each rectangular aperture having a width of $a$, a length of $b$, and a center-to-center distance between adjacent slits of $d$. The Fraunhofer diffraction formula, a fundamental calculation method in optics, is utilized to characterize the diffraction phenomenon when light passes through an aperture. When light passes through a finite-sized aperture, it generates a series of interference and diffraction patterns within the far-field region. The Huygens-Fresnel principle states that each point on a wave surface can be treated as a new secondary wave source, and the wave surface can be considered as a superposition of spherical waves emitted by an infinite number of point sources.

Linear systems possess an essential characteristic known as the principle of superposition, which asserts that the resulting output of a linear system, when multiple input signals are applied, is the linear superposition of these input signals. This principle is applicable to the phenomenon of aperture mask diffraction, wherein each aperture behaves as a point source. The waves emanating from every point source combine constructively and destructively to produce the output wave of the entire aperture. This process of wave superposition is linear, implying that the spatial distribution of the output wave corresponds to the superposition of waves generated by all point sources as the number of apertures increases. Consequently, each rectangular aperture in the aperture mask can be viewed as a point source, and the waves produced by all point sources can be added coherently to yield the diffraction pattern across the entire aperture mask.

The initial form of Fraunhofer diffraction is:

$$E_p = c \int_A e^{ikr}dA$$

(1)

Considering a single square hole mask with a width of $a$ and a length of $b$, we can write its imaging distribution on the Fourier surface as:

$$E_p = c \int_0^b \int_0^a e^{ik(r_0 + x \sin \phi + y \sin \theta)} dx dy$$

(2)

Therefore, for a multi-slit mask comprising $N \times N$ parallel rectangular apertures, we can write the diffraction formula as follows:

$$E_p = c e^{ikr_0} \left[ \int_0^b e^{iky \sin \theta_1} dy + \cdots + \int_0^{(N-1)d+b} e^{iky \sin \theta_1} dy \right]$$

$$\times \left[ \int_0^a e^{ikx \sin \theta_2} dx + \cdots + \int_0^{(N-1)d+a} e^{ikx \sin \theta_2} dx \right]$$

(3)

Calculated to get:

$$E_p = c e^{ikr_0} \frac{e^{ikb \sin \theta_1} - 1}{ik \sin \theta_1} \times \frac{1 - e^{iN \delta \sin \theta_1}}{1 - e^{ikd \sin \theta_1}}$$

$$\times \frac{e^{ika \sin \theta_2} - 1}{ik \sin \theta_2} \times \frac{1 - e^{i\delta \sin \theta_2}}{1 - e^{ikd \sin \theta_2}}$$

(4)
To simplify the parameters, let:

\[
\begin{align*}
\beta_1 &= \frac{1}{2} kb \sin \theta_1, \quad \beta_2 = \frac{1}{2} ka \sin \theta_2, \\
\gamma_1 &= \frac{1}{2} kd \sin \theta_1, \quad \gamma_2 = \frac{1}{2} kd \sin \theta_2
\end{align*}
\]

(5)

Then we get:

\[
E_p = \frac{e^{ikr_0}}{ab} \frac{\sin \frac{2i\beta_1}{\beta_1} - 1}{2i\beta_1} \times \frac{1 - e^{2iN\gamma_1}}{1 - e^{2i\gamma_1}} \times \frac{\sin \frac{2i\beta_2}{\beta_2} - 1}{2i\beta_2} \times \frac{1 - e^{2iN\gamma_2}}{1 - e^{2i\gamma_2}}
\]

(6)

Since focusing is performed under paraxial conditions, the formula \(\sin \theta_1 \approx \tan \theta_1 = \frac{x_m}{f_2}\), \(\sin \theta_2 \approx \tan \theta_2 = \frac{y_m}{f_2}\):

\[
E_p = E_0 \frac{\sin \frac{\beta_1}{\beta_1}}{\sin \frac{\gamma_1}{\beta_1}} \frac{\sin \frac{\beta_2}{\beta_2}}{\sin \frac{\gamma_2}{\beta_2}}
\]

(7)

\[
I(x_m, y_m, \lambda) = I_0 \cdot D(x_m, y_m, \lambda) \cdot P(x_m, y_m, \lambda)
\]

(8)

\[
D(x_m, y_m, \lambda) = \sin \left( \frac{\pi b}{\lambda f_2} x_m \right) \sin \left( \frac{\pi a}{\lambda f_2} y_m \right)
\]

(9)

\[
P(x_m, y_m, \lambda) = \left[ \frac{\sin \left( \frac{N \pi d}{\lambda f_2} x_m \right)}{\sin \left( \frac{\pi d}{\lambda f_2} x_m \right)} \right]^2 \times \left[ \frac{\sin \left( \frac{N \pi d}{\lambda f_2} y_m \right)}{\sin \left( \frac{\pi d}{\lambda f_2} y_m \right)} \right]^2
\]

(10)

Where the formula \(D(x_m, y_m, \lambda)\) is the diffraction factor describes the diffraction effect of each rectangular square hole. \(P(x_m, y_m, \lambda)\) is the interference factor describes the effect of multi-slit interference. \((x_m, y_m)\) denotes the spatial coordinates on the receiving screen, while \(f_2\) denotes the distance between the diffraction array and the sensor.

Finally, we get the formula under orthogonal aperture diffraction:

\[
I = I_0 \left( \frac{\sin \frac{\beta_1}{\beta_1}}{\sin \frac{\gamma_1}{\gamma_1}} \right)^2 \left( \frac{\sin \frac{N \gamma_1}{\beta_1}}{\sin \frac{N \gamma_2}{\beta_1}} \right)^2 \left( \frac{\sin \frac{\beta_2}{\beta_2}}{\sin \frac{\gamma_2}{\gamma_2}} \right)^2
\]

(11)

4. Test set of simulation experiment

Here we show our test set of 10 scenes selected from the KAIST [11] dataset as depicted in Figure 2. The 256*256
4. Simulation Experiments (Low Exposure)

Unlike the main text’s simulation experiments conducted with regular exposure, here, the measurements’ amplitude is scaled down to approximately one-fourth of the original to simulate varying exposure conditions. Similar to [8, 12, 13, 14, 15, 16, 17], 28 wavelengths are selected from 450nm to 650nm and derived by spectral interpolation manipulation for the HSI data.

Simulation Dataset. We adopt two datasets, i.e., CAVE-1024 [8] and KAIST [11] for simulation experiments. The CAVE-1024 consists of 205 HSIs with spatial size 1024x1024 obtained by interpolating and splicing from the CAVE [18] dataset. The KAIST dataset contains 30 HSIs of spatial size 2704×3376. 10 scenes from the KAIST dataset are selected for testing, while the CAVE-1024 dataset and another 20 scenes from the KAIST dataset are selected for training.

Implementation Details. The dispersion step of the primary diffraction is 0.5 spatial pixels, while the simulation experiment is deployed in the range of 400nm to 670nm, which means that 586x586x28 data cubes are needed to generate 256x256 resolution measurements for conducting experiments while preserving the tertiary diffraction. We...
implement CSST by Pytorch. All CSST models are trained with Adam [19] optimizer ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) using Cosine Annealing scheme [20] for 300 epochs on an RTX 3090 GPU. The initial learning rate is $4 \times 10^{-4}$.

**Quantitative Analysis.** Table 1 compares the results of CSST and 10 methods including one baseline method(Unet [1]), six reconstruction methods(lambda-Net [6], HDNet [3], BIRNAT [4], TSA-Net [8], HSCNN+ [2] and MST++ [9]), three Super-resolution algorithms (Restorner [10], MPRNet [7], MIRNet[5]) on 10 simulation scenes at low exposure. CSST shows the best experimental results on the ADIS spectral reconstruction task, i.e., $32.16\text{dB}$ in PSNR and $0.944$ in SSIM. CSST-9stg significantly outperforms two recent SOTA methods Restorner and MST++ by $1.84\text{dB}$ and $2.46\text{dB}$, demonstrating stronger reconstruction performance compared to previous methods under low exposure conditions and robustness against exposure variations.

**Qualitative Analysis.** Figure 4 illustrates the comparative performance of our CSST and other methods in the HSI reconstruction of ADIS on the same scene at low exposure. Visual inspection of the image reveals that the CSST-9stg method provides more intricate details, sharper textures, and well-defined structures. Conversely, the previous approaches produce either overly smooth results that compromise the underlying structure or introduce color artifacts and speckled textures. Moreover, the lower left corner of the figure presents the spectral profile of the intensity-wavelength corresponding to the fuchsia square.

5. Additional real reconstruction results

Here we further show the reconstruction outcomes of various scenes captured by ADIS in the figure 5. These results exhibit distinct textures and well-structured edges, thereby corroborating the efficacy of ADIS in snapshot super-pixel resolution spectral imaging.

**References**


