

# Dense Dispersed Structured Light for Hyperspectral 3D Imaging of Dynamic Scenes

Suhyun Shin  
POSTECH

Seungwoo Yoon  
POSTECH

Ryota Maeda  
POSTECH, University of Hyogo

Seung-Hwan Baek  
POSTECH

In this supplemental document, we provide additional results and details in support of our findings in the main manuscript.

## Contents

<b>1. Experimental Prototype</b>	<b>2</b>
1.1. Geometric Calibration . . . . .	2
1.2. Initial Radiometric Calibration of Spectral/Emission Function . . . . .	2
1.2.1 . Camera Response Function . . . . .	2
1.2.2 . Projector Emission Function . . . . .	2
1.3. Refinement of Spectral/Emission Function . . . . .	2
1.4. Diffraction Grating Efficiency . . . . .	3
1.5. Fast Capture Synchronization . . . . .	3
<b>2. Comparison</b>	<b>3</b>
2.1. Ours vs DSL . . . . .	3
2.2. Quantitative Comparison with Other Systems . . . . .	4
2.3. Discussion on Other Methods . . . . .	4
<b>3. Additional Discussion</b>	<b>4</b>
3.1. Failure Cases . . . . .	4
3.2. Future Directions . . . . .	4
3.3. Depth Accuracy . . . . .	4
3.4. Real-time Processing . . . . .	5
<b>4. Hyperspectral 3D Reconstruction</b>	<b>6</b>
4.1. Active-stereo Depth Estimation . . . . .	6
4.2. Matrix-vector Multiplication . . . . .	6
4.2.1 . Intensity Matrix . . . . .	6
4.2.2 . System Matrix . . . . .	6
4.3. Hyperspectral Reconstruction . . . . .	6
4.4. Details on Backward Model . . . . .	7
4.4.1 . Data Acquisition Method . . . . .	7
4.4.2 . Fitting Method . . . . .	7
<b>5. Additional Details on Black Optical Flow</b>	<b>8</b>
<b>6. Results on Depth Imaging</b>	<b>9</b>
<b>7. Results on Hyperspectral Imaging</b>	<b>10</b>
7.1. Static Scenes . . . . .	10
7.2. Dynamic Scenes . . . . .	13

## 1. Experimental Prototype

We list all parts used to build the experimental prototype system in Table 1.

Item #	Part description	Quantity	Model name
1	RGB Camera	2	FLIR GS3-U3-32S4C-137 C
2	RGB Projector	1	Epson CO-FH02
3	Objective lens	2	Edmund #33-303
4	Diffraction grating sheet	1	Edmund 158 #54-509
5	Holder	1	3D printed

Table 1. Part list of our imaging system.

### 1.1. Geometric Calibration

The geometric calibration of the camera and projector is preformed without using diffraction grating [7]. Similarly, the stereo-camera system undergoes geometric calibration based on the approach proposed in [9]. Both calibration procedures achieve an average sub-pixel reprojection error of 0.1 pixels, ensuring precise geometric alignment with high accuracy. Using the calibrated geometric parameters of camera and projector, we establish the correspondence between a camera pixel  $p$  and a projector pixel  $q$  through the operations  $\text{unproject}(\cdot)$  and  $\text{project}(\cdot)$ , as defined below:

$$q = \text{project}(\text{unproject}(p, \mathbf{p}_z)), \quad (1)$$

where  $\mathbf{p}_z$  represents the depth value of scene point  $\mathbf{p}$  corresponding to camera pixel  $p$ .

### 1.2. Initial Radiometric Calibration of Spectral/Emission Function

#### 1.2.1. Camera Response Function

This section outlines the radiometric calibration process for determining the spectral response and emission functions. To compute the initial spectral response function of the camera, we employ a Spectralon target with a uniform reflectance of 99% across all visible wavelengths. The target is illuminated by an LED light source, whose spectral power distribution is precisely measured using a spectroradiometer (JETI Specbos 1211) to provide the ground truth reference for the calibration.

The camera’s spectral response is defined over the wavelength range of 440 nm to 660 nm, sampled at 10 nm intervals. Narrow band spectral filters, corresponding to each interval, are sequentially placed in front of the camera. Each spectral images are captured under the same LED illumination for a total of 23 spectral bands. By analyzing the intensity values across these bandpass-filtered images, we compute the camera’s spectral response function, as expressed below:

$$I_\lambda(p, c) = \sum_{\lambda \in \Lambda} \Omega_{c,\lambda}^{\text{cam}} H_{\text{Spectralon}}(p, \lambda) LED(\lambda), \quad (2)$$

where  $I_\lambda(p, c)$  represents the intensity of the captured image at pixel  $p$  within the region of interest (RoI) for spectral band  $\lambda$  and RGB channel  $c$ . The term  $H_{\text{Spectralon}}(p, \lambda)$  denotes the hyperspectral reflectance of the Spectralon target which is 99%. The LED spectral power distribution, denoted as  $LED(\lambda)$ , serves as the ground truth for the light source, and  $\Omega_{c,\lambda}^{\text{cam}}$  represents the spectral response function of the camera. We show the initial radiometric parameters in Figure 1(a).

#### 1.2.2. Projector Emission Function

Next, we calibrate the projector’s emission function, denoted as  $\Omega_{c,\lambda}^{\text{proj}}$ . This is achieved by projecting RGB patches onto the Spectralon target. The reflected spectral radiance of each projected patch is then measured using a spectroradiometer, allowing us to obtain the spectral emission curves for each of the RGB channels. The the initial radiometric parameter of projector is depicted in Figure 1(a).

### 1.3. Refinement of Spectral/Emission Function

We refine the spectral response and emission functions using multiple white scanline patterns, similar to the approach proposed by Shin et al. [6]. The refinement process minimizes the difference between the measured intensity  $\mathbf{I}$ , and the simulated intensity under scanline patterns, as expressed in Equation 3. This refinement is based on the simulated pixel intensity graph for 21 central points  $p \in \mathcal{P}$  of the Classic ColorChecker, using the known ground truth hyperspectral reflectance  $H(p, \lambda)$ .

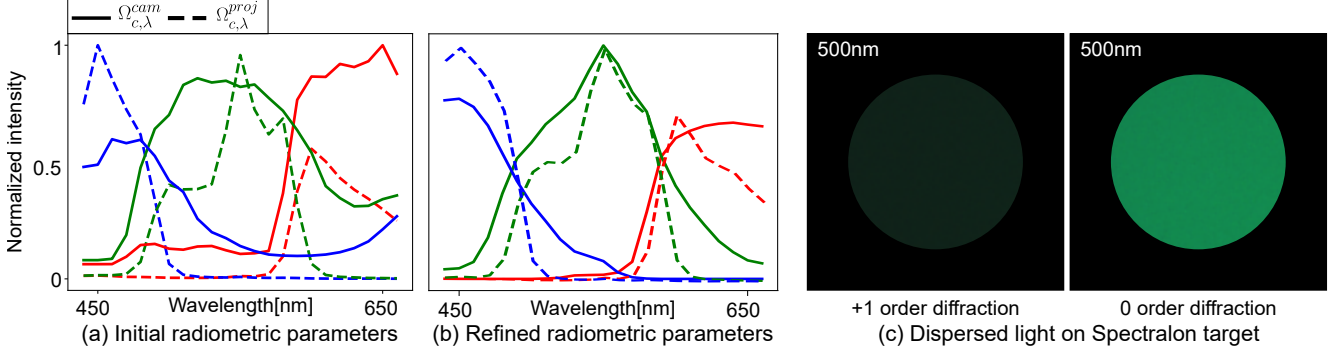


Figure 1. **Radiometric Calibration.** (a) Initial spectral and emission functions before optimization. (b) Refined radiometric parameters after optimization. (c) Captured dispersed light on the Spectralon target at a specific wavelength (500 nm) used for diffraction grating efficiency calibration.

$$\underset{\Omega's}{\operatorname{argmin}} \sum_{p \in \mathcal{P}} \|\mathbf{I}_k - \Omega_{c,\lambda}^{\text{cam}} H_k(p, \lambda) \eta_\lambda L(q_\lambda, \lambda)\|_2^2 + w(\|\nabla_\lambda \Omega_{c,\lambda}^{\text{cam}}\|_2^2 + \|\nabla_\lambda \Omega_{c',\lambda}^{\text{proj}}\|_2^2). \quad (3)$$

In this formulation,  $\nabla_\lambda$  represents the gradient operator along the spectral axis, which ensures the smoothness of the spectral response functions. The regularization term, weighted by  $w = 0.008$ , penalizes variations in the spectral gradients of the both camera's spectral response function  $\Omega_{c,\lambda}^{\text{cam}}$ , and the projector's emission function  $\Omega_{c',\lambda}^{\text{proj}}$ . The results of the optimized spectral response and emission functions are presented in Figure 1(b), and this refinement leads to a more accurate representation of the radiometric properties, enabling improved spectral and spatial reconstruction.

#### 1.4. Diffraction Grating Efficiency

A diffraction grating disperses incident light into multiple diffraction orders. The zero-order diffraction preserves the original path of the incident light, maintaining the same intensity across all wavelengths. In contrast, the first-order diffraction, which is utilized in our system, separates the incident light into different angles based on wavelength. Unlike the zero-order light, the first-order diffraction exhibits varying efficiency across different wavelengths which is the diffraction grating efficiency  $\eta_\lambda$ .

The diffraction grating efficiency  $\eta_\lambda$  is calibrated by measuring the intensity of the zeroth order and positive first order light. To achieve this, we place spectral bandpass filters in front of the camera at 10nm intervals, wavelengths from 440nm to 660nm. For each wavelength  $\lambda$ , the corresponding dispersed light projected onto a Spectralon target is captured. Figure 1(d) depicts an example of the captured Spectralon target for the 500nm wavelength, showing both the positive first-order diffracted light and the zero-order light. We compute the final diffraction efficiency by deriving the intensity ratio of each positive first-order wavelength over zero-order intensity.

#### 1.5. Fast Capture Synchronization

We capture dynamic scenes by synchronizing the stereo cameras and the projector. Using this setup, we capture one group of  $M$  DDSL patterns with a single black pattern, at a frame rate of 6.6 fps. Synchronization between the projector and stereo cameras is managed through the OpenGL library. The next projector pattern is preloaded in advance, and the buffer is swapped immediately after the future tasks for image capture are completed. The synchronization process is outlined in the pseudo code below.

## 2. Comparison

### 2.1. Ours vs DSL

Our method differs from DSL [18] in five key aspects: (1) a Gaussian-fitted image formation model, (2) a spectrally multiplexed DDSL pattern design, (3) an active stereo system with stereo cameras and a dispersed projector, (4) motion-aware hyperspectral reconstruction using black optical flow, and (5) single-shot depth estimation using structured-light cues. These differences enable (a) per-frame depth imaging with accuracy comparable to the multi-frame structured light approach of DSL [18], and (b) a 4,000 $\times$  increase in image acquisition speed while maintaining superior spectral accuracy.

---

**Algorithm 1:** Pseudo-code for Fast Multi-Camera and Projector

---

**Input:** Stereo camera serial numbers  $serial_1$ ,  $serial_2$ .

**Output:** Stereo camera captured image lists.

```
1 Initialization
2   Initialize stereo cameras using serial numbers  $serial_1$  and  $serial_2$ ;
3   Configure cameras;
4   Enable software trigger mode for both cameras;
5   Initialize fullscreen display;
6   Create group of  $M$  DDSL and single black pattern  $Pattern\_group$ ;
7 Start Image Acquisition
8   Start image acquisition for both cameras;
9   Display the first image on the screen;
10  foreach image  $i$  in  $Pattern\_group$  do
11    Swap display buffers;
12    Execute software trigger for both cameras;
13    Launch threads to capture images from both cameras;
14    if  $i < len(Pattern\_group) - 1$  then
15      Preload the next image into the display buffer;
```

---

## 2.2. Quantitative Comparison with Other Systems

In Figure 2(a) we show quantitative simulation results comparing our method with [18, 12] and CASSI method [11] in terms of PSNR, SAM, and capture speed on the simulated spectral filter scene similar to Fig.6 in the main paper. Our method achieves the best SAM, PSNR, and capture speed. Note that simulating other complex systems [1, 2, 9] was infeasible due to their reliance on complex wave optics and precise electronics modeling.

## 2.3. Discussion on Other Methods

We additionally compare our method with other compact imaging approaches shown as Figure 2(b): Baek et al.[4] employs a diffractive optical element (DOE) with a single-camera setup and utilizes a global learning-based reconstruction method. This approach mainly suffer from low-frequency spectral point spread function (PSFs). Li et al. [4] uses a conventional projector-camera system for depth and spectral reconstruction, which illuminates broad-band spectral illumination providing ambiguous spectral cues, limiting to achieve high accurate hyperspectral information. However as shown in Figure 2(b) our method projects narrow-band spectral illumination which provides dense cues for accurate hyperspectral reconstruction. Note that neither paper provides public code, which limits direct comparison between our methods.

## 3. Additional Discussion

### 3.1. Failure Cases

Fast-moving objects can degrade the accuracy of black optical flow due to frame-dependent blur, leading to spectral and depth inaccuracies. Hardware projector-camera synchronization could enable using a shorter exposure time to better handle fast motion.

### 3.2. Future Directions

Differentiable joint reconstruction of spectral and depth information with learning-based pattern design is an interesting direction. However, key challenges include developing a differentiable reconstruction pipeline and establishing a large-scale hyperspectral 3D dataset for learning-based approaches. Also, integrating optical flow into our single-shot active stereo depth could enhance depth accuracy through temporal denoising.

### 3.3. Depth Accuracy

Although DSL [18] achieves accurate depth reconstruction (MAE 1 mm) using 40 binary code patterns it is limited to static scenes. However, our DDSL leverages DDSL patterns as stereo-matching cues for depth reconstruction, allowing per-frame



dynamic depth imaging with a comparable 4 mm depth MAE.

### 3.4. Real-time Processing

While our method achieves a 4,000 $\times$  increase in acquisition speed over DSL [18], reconstruction is performed offline, taking approximately 3 minutes. Future real-time reconstruction may be possible through an SVD-based optimization in C++ and reformulating the problem as a least-squares optimization without L1 terms.

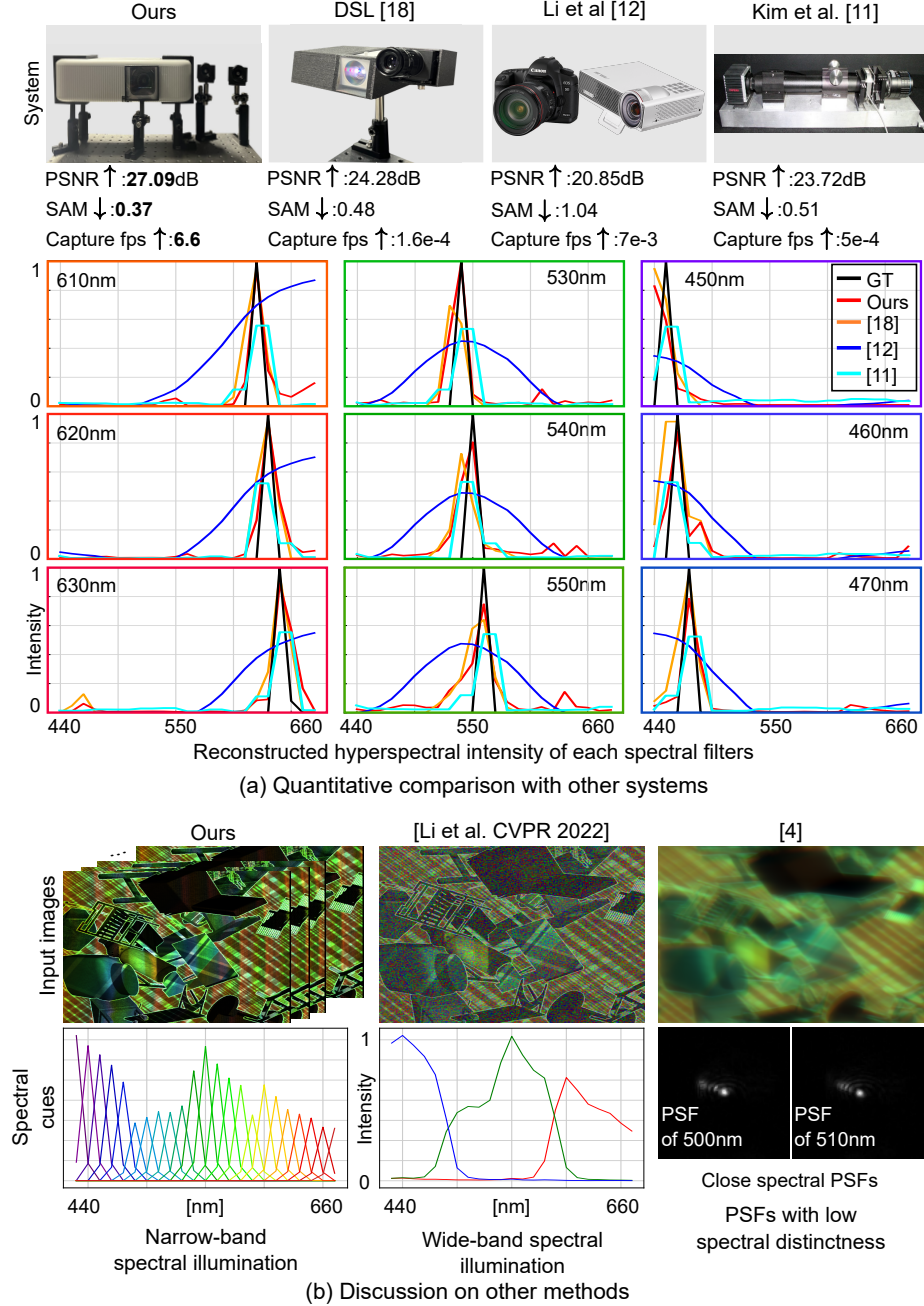


Figure 2. **Comparison with other methods.** (a) Quantitative comparison with other systems (b) Discussion on other methods.

## 4. Hyperspectral 3D Reconstruction

### 4.1. Active-stereo Depth Estimation

We employ an active stereo imaging system to achieve accurate single-shot depth estimation using the pretrained RAFT-stereo network [8]. The network takes two rectified stereo images  $\text{rectify}(I_k^i)$  as input and outputs a disparity map corresponding to the rectified camera view.

To integrate RAFT-Stereo [5] into our system, we first utilize the precalibrated geometric parameters to rectify the stereo image pairs before feeding them into the network. The output disparity map from RAFT-Stereo, initially aligned with the rectified view, is then transformed back to the original camera view through an inverse rectification  $\text{unrectify}(\cdot)$  process. Finally, using the disparity map and the geometric calibration parameters, we compute the depth map for the target camera view. This process ensures accurate depth estimation with high fidelity, using both the state-of-the-art RAFT-Stereo framework and the precise geometric calibration of our system. Figure 3 provides an overview of this pipeline.

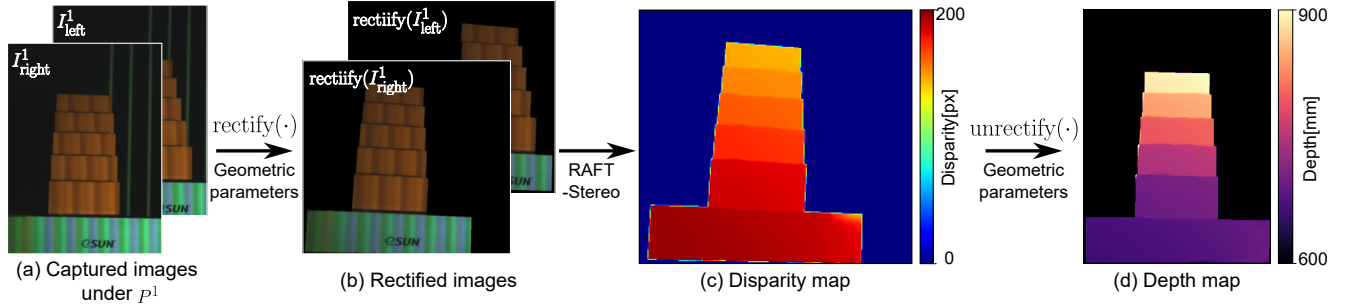


Figure 3. **Raft Stereo.** (a) Captured images under  $P^1$ . (b) Rectified images based on calibrated geometric parameters of stereo cameras. (c) Output of disparity map of RAFT-Stereo [5]. (d) Depth map.

### 4.2. Matrix-vector Multiplication

In this section we detail the matrix-vector multiplication form as shown in Equation 4 of our image formation. Following describes each intensity matrix  $\mathbf{I}_k$  and system matrix  $\mathbf{L}$ .

$$\mathbf{I}_k = \mathbf{L}\mathbf{H}_k. \quad (4)$$

#### 4.2.1. Intensity Matrix

To construct the intensity matrix  $\mathbf{I}_k$ , images are captured under  $M$  DDSL patterns, denoted as  $\{I_k^i(p_k^i, c)\}_{i=1}^M$ . To get rid of the effects of residual light and enhance reconstruction accuracy, the intensity of a black pattern image  $I_k^B$  is subtracted from each captured image. By minimizing these residual effects, this ensures more accurate intensity measurements, contributing to improved reconstruction performance.

#### 4.2.2. System Matrix

With the calibrated and refined radiometric parameters, we define the system matrix  $\mathbf{L}$  as expressed as below:

$$\mathbf{L} = [\mathbf{L}_R; \mathbf{L}_G; \mathbf{L}_B]^\top \in \mathbb{R}^{3M \times N}, \quad (5)$$

$$\mathbf{L}_c(i, j) = \Omega_{c, \lambda_j}^{\text{cam}} \eta_{\lambda_j} (L^i * G) (q_{\lambda_j}^i, \lambda_j), \quad (6)$$

where  $L_c(i, j)$  represents the system matrix element for channel  $c \in \{R, G, B\}$ , with  $i \in \{1, \dots, M\}$  corresponding to the  $i$ -th DDSL pattern and  $j \in \{1, \dots, N\}$  representing the target wavelength index. Here,  $N$  denotes the total number of target wavelengths. The term  $L^i$  represents the light intensity of the  $i$ -th pattern, which is convolved with a Gaussian kernel  $G$  with kernel size 7 and standard deviation 3 to account for optical blurring effects. This formulation ensures accurate modeling and reconstruction.

### 4.3. Hyperspectral Reconstruction

Here we outline the details of our hyperspectral optimization. Our goal is to reconstruct a highly accurate hyperspectral image  $\mathbf{H}_k$ , using our proposed optimization method. The optimization process is performed over 1000 epochs, employing

the Adam optimizer [1]. The learning rate is initialized at 0.05 and is reduced according to a decay schedule with a step size of 400 epochs and a decay factor of 0.5. This schedule ensures stable convergence by gradually lowering the learning rate as the optimization progresses. The proposed method is designed to refine the spectral reconstruction iteratively, achieving high accuracy across the spectral range.

#### 4.4. Details on Backward Model

We describe the process for obtaining data samples and fitting our backward model as defined in Equation 7.

$$q_\lambda = \psi(\mathbf{p}, \lambda). \quad (7)$$

Here,  $q_\lambda \in \mathbb{S}_q$  represents the projector pixel sample corresponding to a scene point  $\mathbf{p} \in \mathbb{S}_p$  and a wavelength sample  $\lambda \in \mathbb{S}_\lambda$ . The backward model  $\psi$  establishes the relationship between the projector pixel coordinates and the scene geometry dependent on wavelength  $\lambda$ . The following subsections detail the data acquisition process used to obtain the required samples for model fitting.

##### 4.4.1. Data Acquisition Method

To acquire the necessary data samples, we position a Spectralon target at five different depth positions. Using a white scanline projection, following the method proposed in [6]. The scene point samples  $\mathbf{p}$  are then determined using structured light triangulation method [2], ensuring precise 3D coordinates.

For each depth position, spectral data is collected by placing narrow band spectral filters in front of the camera, spanning the wavelength range of interest. Each filter captures specific wavelength  $\lambda$ , enabling the capture of wavelength-specific images for each scanline projection as Figure 4. To determine the projector pixel  $q_\lambda$  corresponding to a given wavelength  $\lambda$  and scene point  $\mathbf{p}$ , we analyze the pixel intensity graph. These pixel intensity graph for specific wavelength  $\lambda$  and pixel  $\mathbf{p}$  are then modeled using a Gaussian fitting function. Previous method took the maximum value of projector coordinate whereas ours take the mean of Gaussian fitted function for sub-accurate sampling of projector coordinate  $q_\lambda$ . This step refines the projector coordinates  $q_\lambda$  to sub-pixel accuracy, providing precise data samples for fitting the backward model.

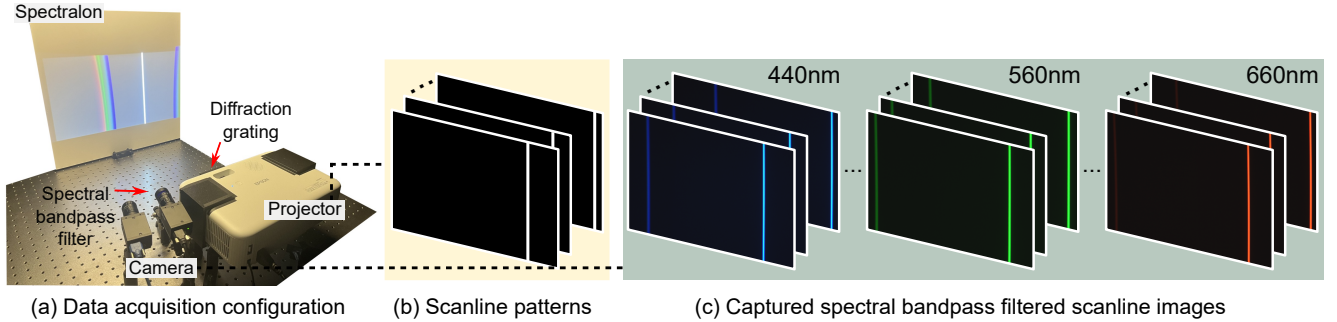


Figure 4. **Details on backward model.** (a) Data acquisition configuration system. (b) Projected white scanline patterns. (c) Captured each spectral bandpass filtered images under white scanline patterns at specific depth position.

##### 4.4.2. Fitting Method

For each scene point  $\mathbf{p}$  with depth coordinate  $\mathbf{p}_z$ , the corresponding projector pixel  $q_\lambda$  is modeled using the following parametric equation:

$$q_\lambda = \alpha \mathbf{p}_z^\beta + \gamma, \quad (8)$$

where  $\alpha$ ,  $\beta$ , and  $\gamma$  are the parameters defining the model. These parameters are estimated by fitting the collected data samples  $\mathbf{p}_z$  and  $q_\lambda$  to the equation using MATLAB's non-linear power function. Additionally, linear interpolation is employed to estimate values for intermediate wavelength samples  $\lambda$  and spatial samples  $\mathbf{p}$ , ensuring continuity and accuracy across the sampled range. This process forms a sub pixel accurate backward model effectively mapping each projector pixel  $q_\lambda$ , wavelength  $\lambda$  and scene point  $\mathbf{p}$ .

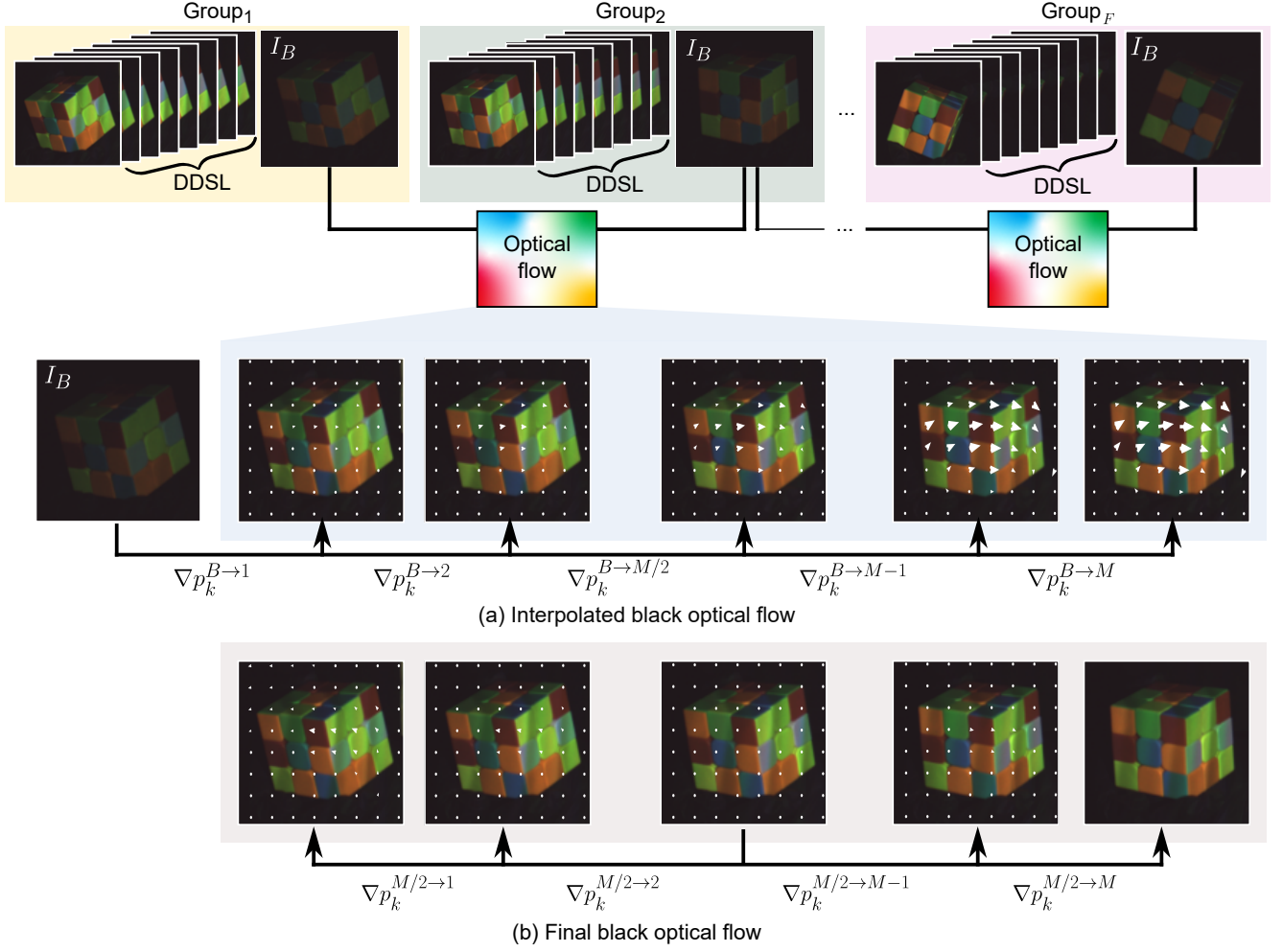


Figure 5. **Black Optical Flow.** We show the overall pipeline of black optical flow. (a) Interpolated black optical flow (b) Final black optical flow.

## 5. Additional Details on Black Optical Flow

In this section we detail the black optical flow of our method. To estimate each optical flow in between black captured images  $I_B$  we used pretrained network RAFT [8]. The overall black optical flow is shown in Figure 5.

We first estimate all optical flows in between black captured images in  $\{\text{Group}_i\}_{i=1}^F$  and all black images are multiplied with constant 3 for bright intensity cue for optical flow. As shown in Figure 5(a) with the interpolated black optical flow of a specific group, we interpolate this using cubic interpolation method and obtain the optical flow from black captured image  $I_B$  to each image captured under  $i$ -th DDSL pattern  $\nabla p_k^{B \rightarrow i}$ . Since our goal is to obtain the target optical flow  $\nabla p_k^{M/2 \rightarrow i}$  we subtract each optical flow to obtain this as the Equation below and the Figure 5(b):

$$\nabla p_k^{M/2 \rightarrow i} = \nabla p_k^{B \rightarrow i} - \nabla p_k^{B \rightarrow M/2} \quad (9)$$

## 6. Results on Depth Imaging

We evaluate the consistency of the reconstructed depth across  $M$  DDSL patterns and compare it with the ground truth obtained using the structured light method [2]. The results of the reconstructed depth under  $M$  DDSL patterns are presented in Figure 6, demonstrating the accuracy of our reconstruction result.

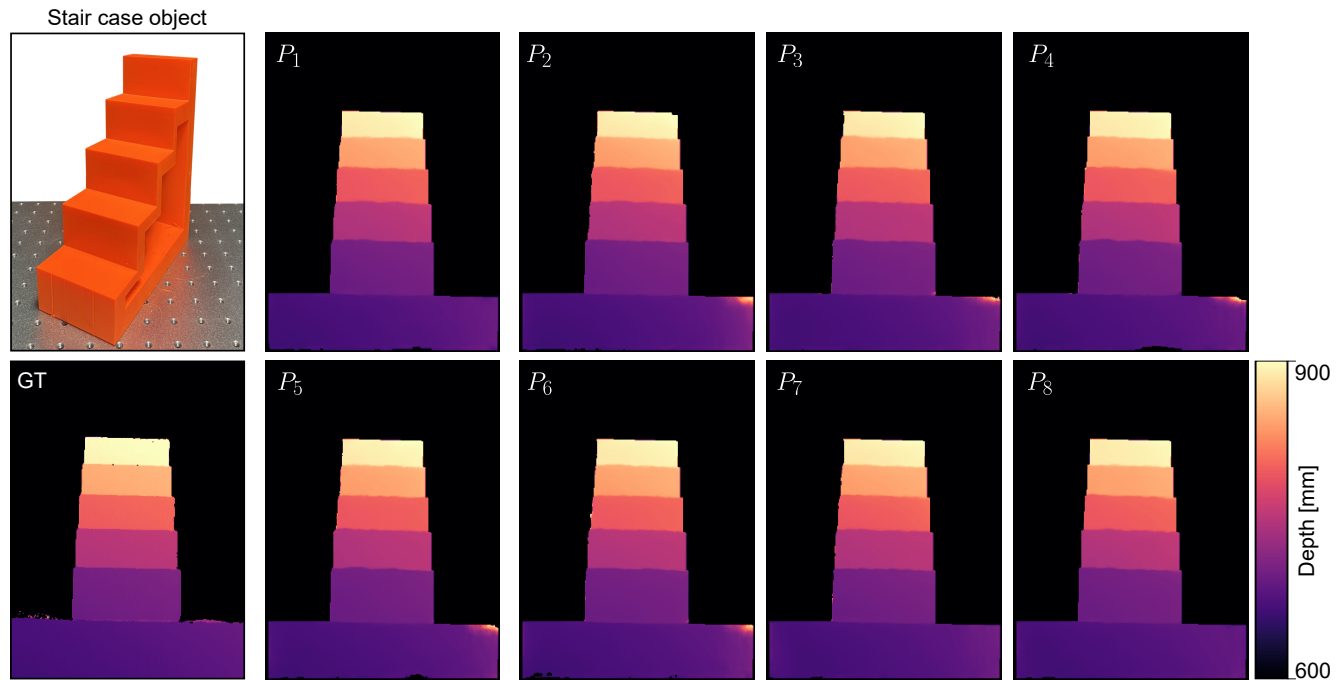


Figure 6. **Depth evaluation.** We show the depth reconstructed results for each  $M$  DDSL patterns and compare it with ground truth depth earned by Structured light method [2].

## 7. Results on Hyperspectral Imaging

### 7.1. Static Scenes

We present detailed results for metamerism, ColorChecker, and high-frequency spectral curves. The reconstructed hyperspectral images, corresponding sRGB images, and spectral curves are shown for a wavelength range of 440 nm to 660 nm, sampled at 10 nm intervals. We validate the accuracy of the reconstructions with the ground truth spectral curves which were measured using a spectroradiometer. The comparison focuses on spectral curves at specific points within each scene, as illustrated in Figures 7, 8 and 9 highlighting the accuracy of our method.

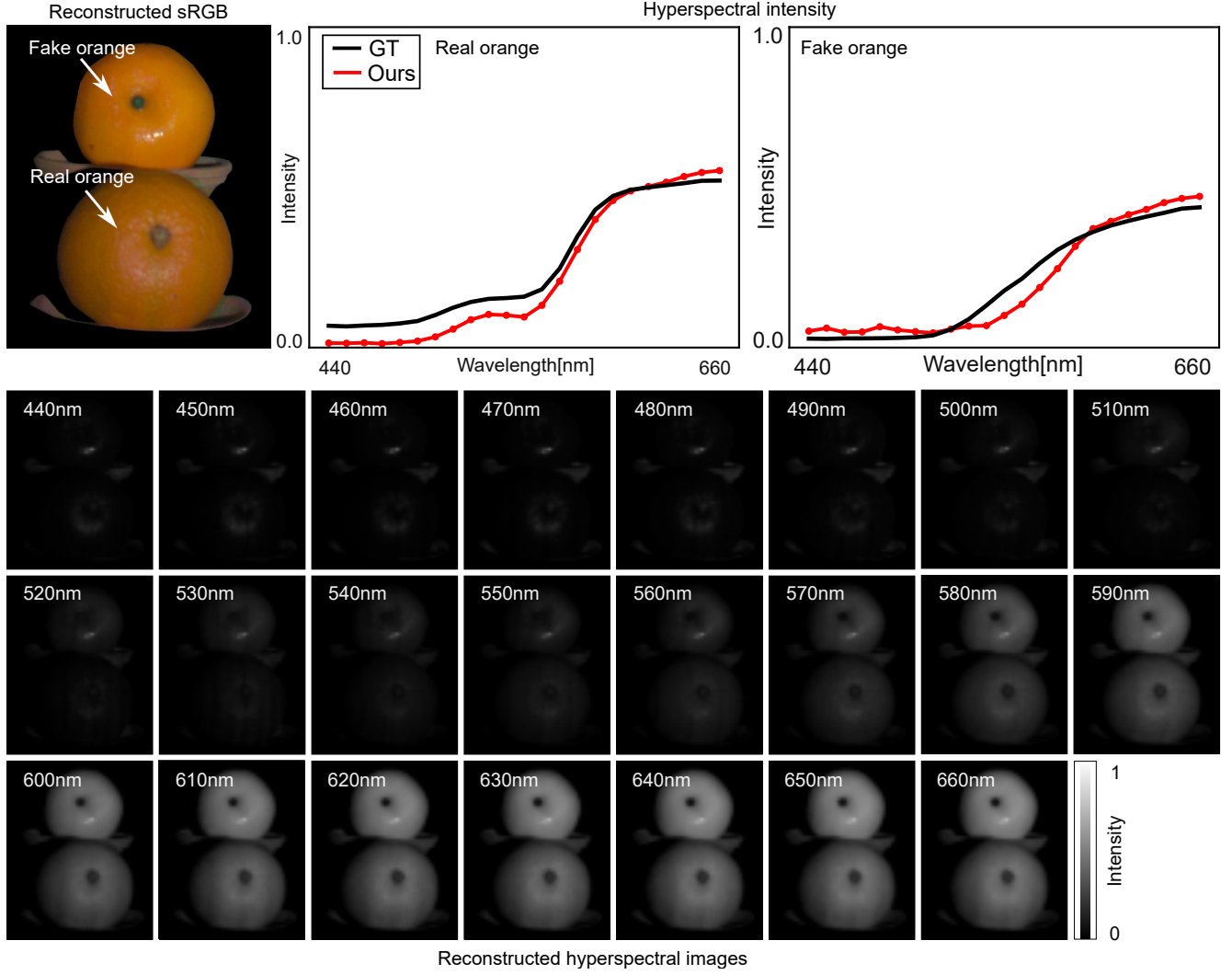


Figure 7. **Metamerism.** We show the metamerism result of a fake and real fruit (orange). Reconstructed sRGB, hyperspectral intensity with comparison with spectrometer measurements and hyperspectral images from 440nm to 660nm for 10 nm interval are depicted.



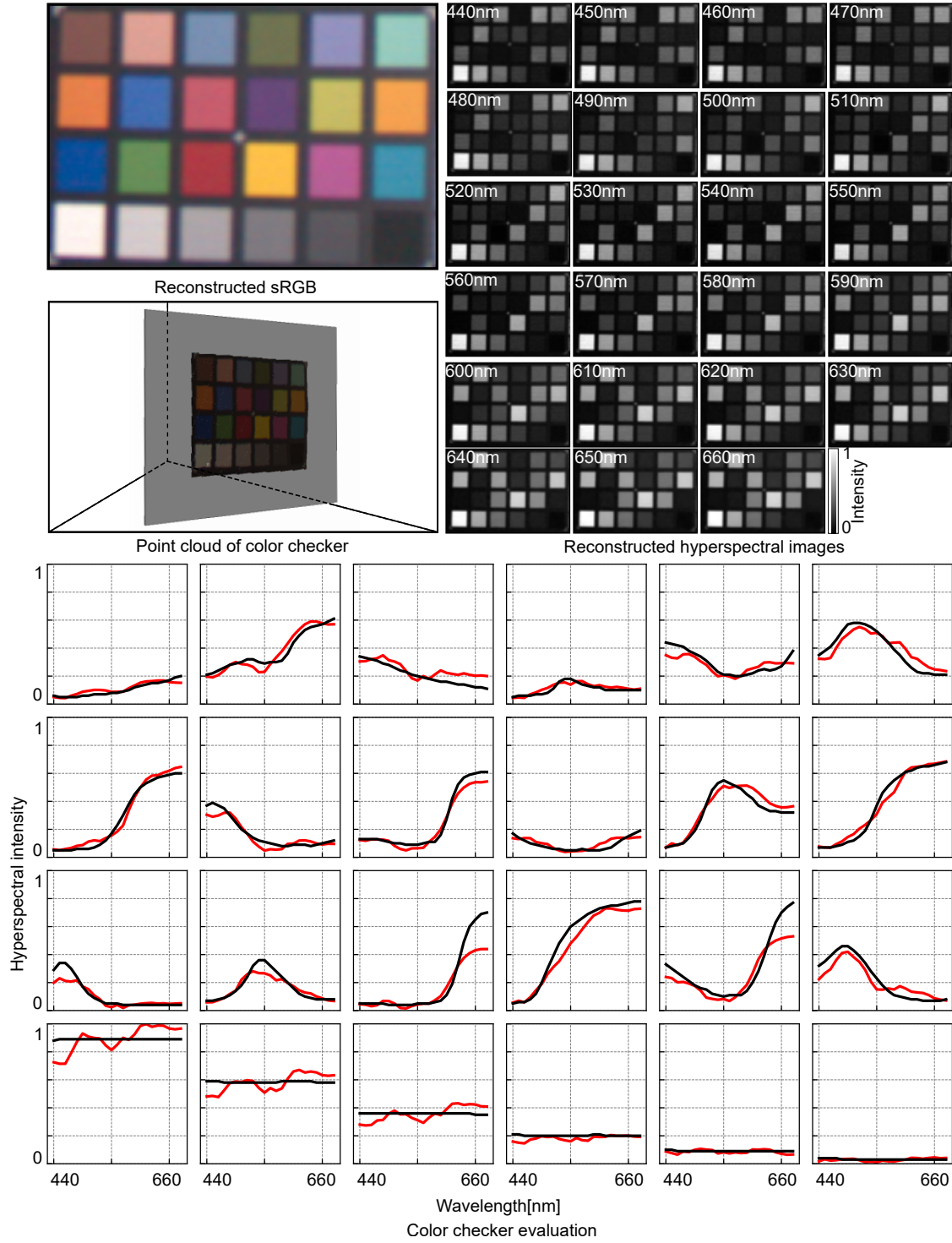


Figure 8. **Hyperspectral 3D imaging.** We show the reconstructed hyperspectral image in sRGB and reconstructed 3d point cloud, and detailed hyperspectral images from 440nm to 660nm at 10 nm interval. We show the Color Checker hyperspectral intensity comparing with ground truth spectral curves.

**High-frequency Spectral Curves** We compare the performance of high-frequency spectral curve reconstruction using our DDSL method with Li et al.[3] and Shin et al.[6]. Figure 9 presents the reconstructed hyperspectral images produced by all three methods, covering wavelengths from 440 nm to 660 nm at 10 nm intervals. Additionally, we illustrate the acquisition time for each method, demonstrating that our DDSL method achieves a frame rate of 6.6 fps, surpassing the performance of the other approaches. Our DDSL method successfully reconstructs high-frequency hyperspectral spectral curves across all nine bandpass filters with high accuracy. In contrast, the method by Li et al. [3] struggles to resolve adjacent spectral curves, highlighting the superior reconstruction capabilities of our approach.

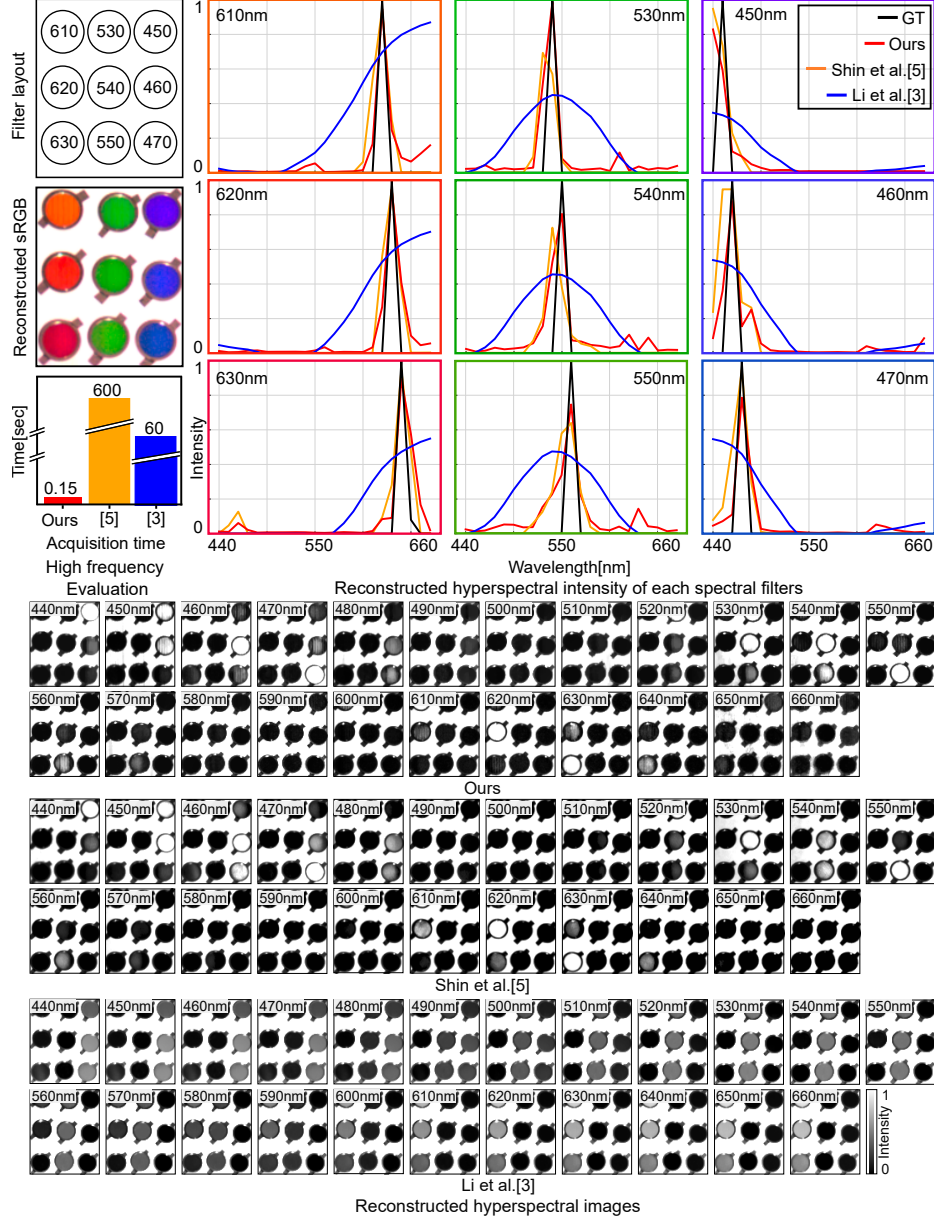


Figure 9. **High frequency evaluation.** We show the reconstructed hyperspectral images in sRGB and acquisition time of high frequency band pass filter scene. We compare the reconstructed hyperspectral intensity and images with Shin et al. [6] and Li et al. [3].



## 7.2. Dynamic Scenes

We present additional results for dynamic scenes in Figures 10, 11, 12, and 13. These results include reconstructed hyperspectral images in sRGB, corresponding depth maps, and detailed reconstructed hyperspectral images at 10 nm intervals. Additionally, we provide detailed spectral curves for Figures 10 and 11. Face scanned result is shown in 13, depicts the hyperspectral images from 440nm to 660nm at 10nm interval. We show dynamic videos in Supplemental Video.

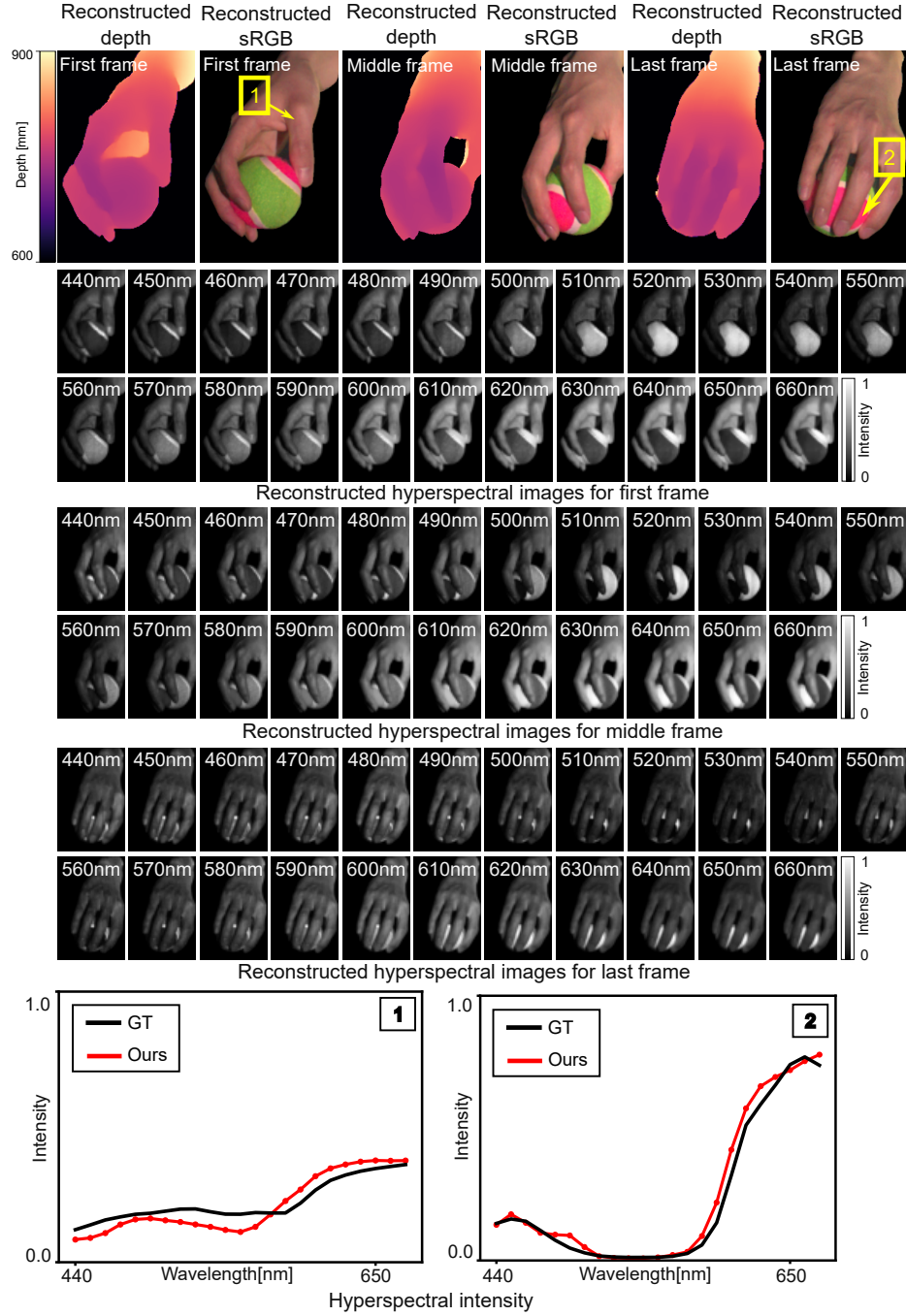


Figure 10. **Hyperspectral 3D imaging for dynamic scene.** We show the reconstructed hyperspectral image in sRGB, reconstructed depth map, detailed reconstructed hyperspectral images from 440nm to 660nm at 10nm interval and the detailed spectral curves for specific points.

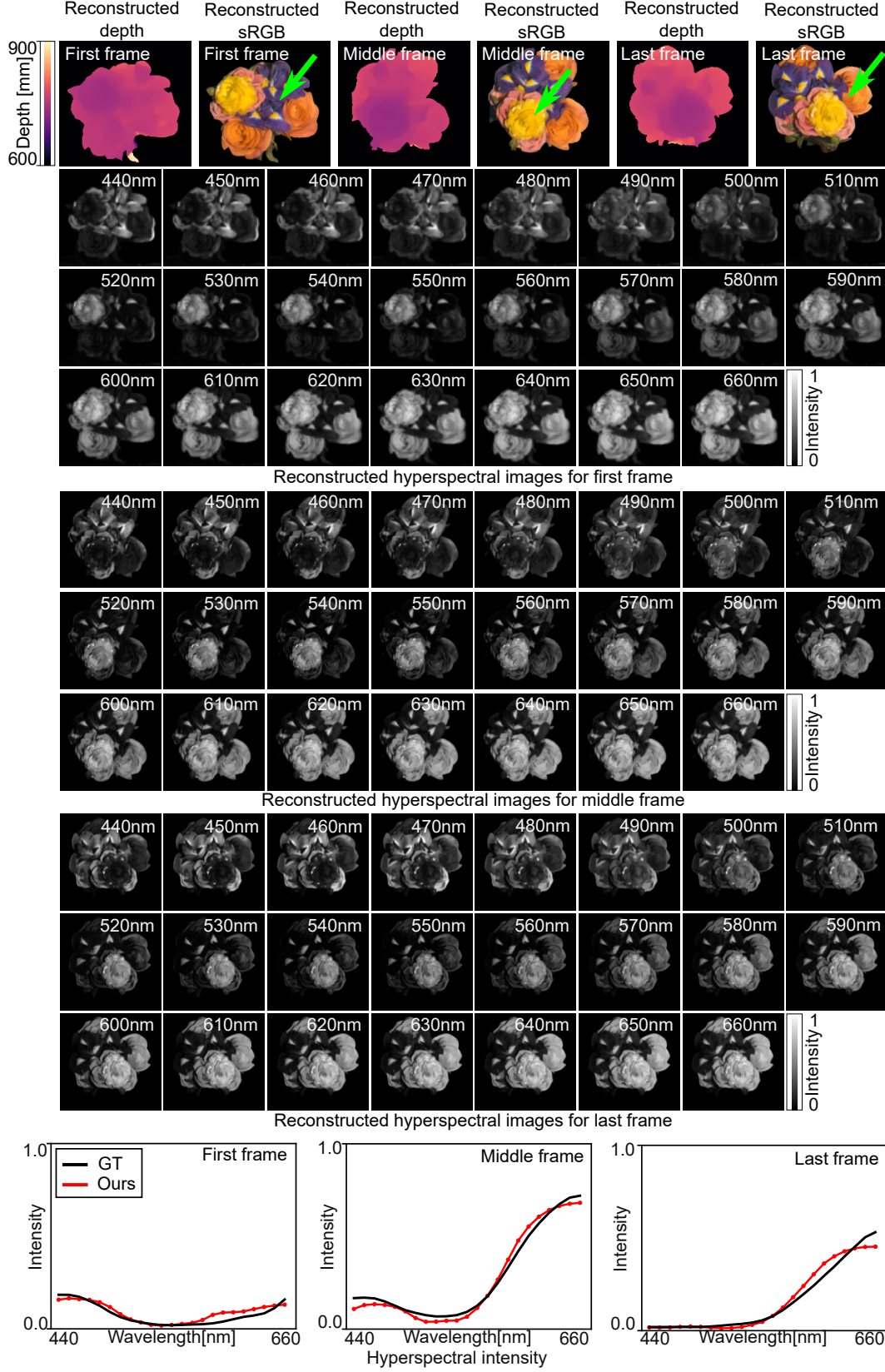


Figure 11. **Hyperspectral 3D imaging for dynamic scene.** We show the reconstructed hyperspectral image in sRGB, reconstructed depth map, and detailed reconstructed hyperspectral images from 440nm to 660nm at 10nm interval. The detailed spectral curves are shown for specific points.

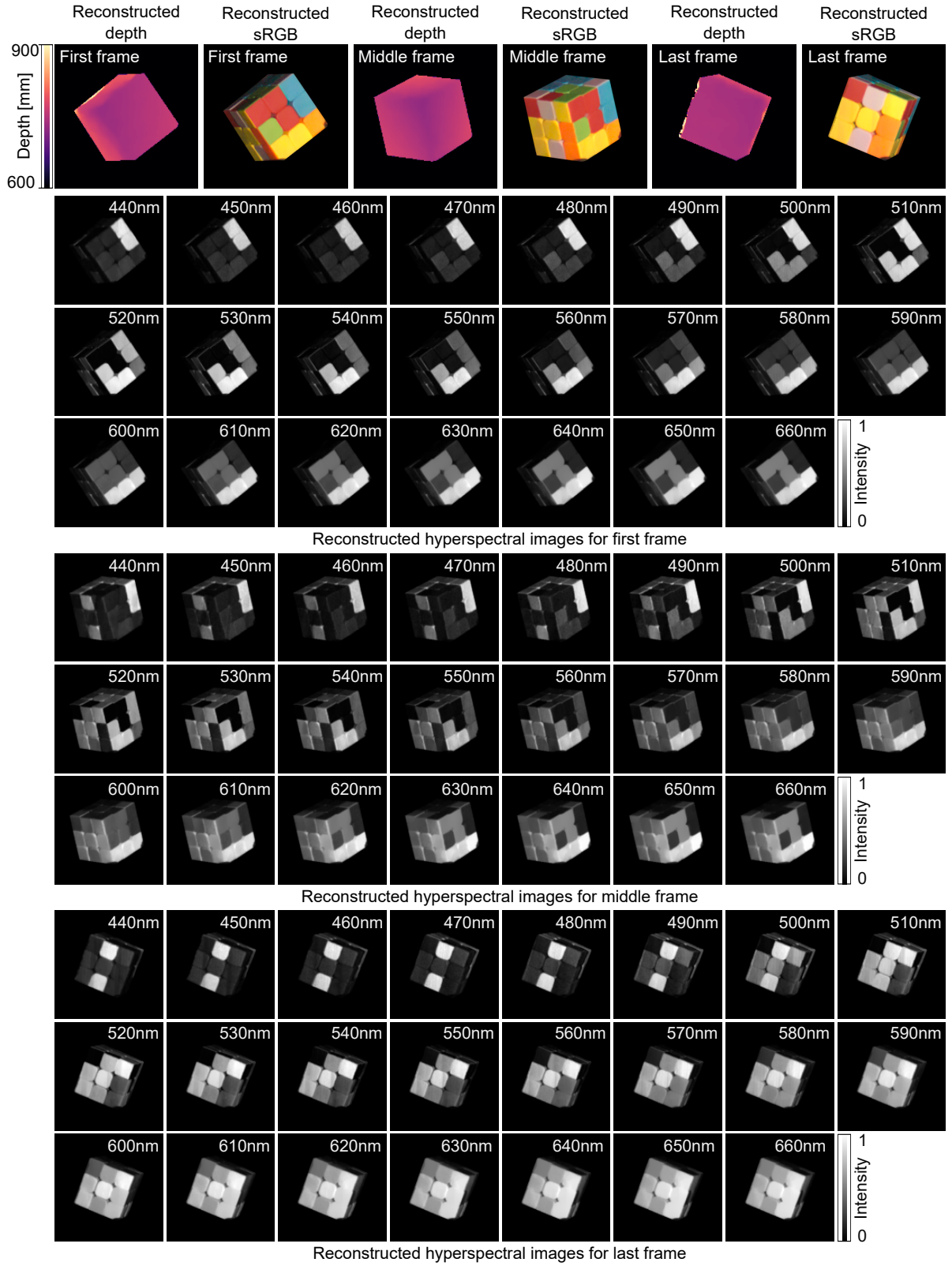


Figure 12. **Hyperspectral 3D imaging for dynamic scene.** We show the reconstructed hyperspectral image in sRGB, reconstructed depth map, and detailed reconstructed hyperspectral images from 440nm to 660nm at 10nm interval.

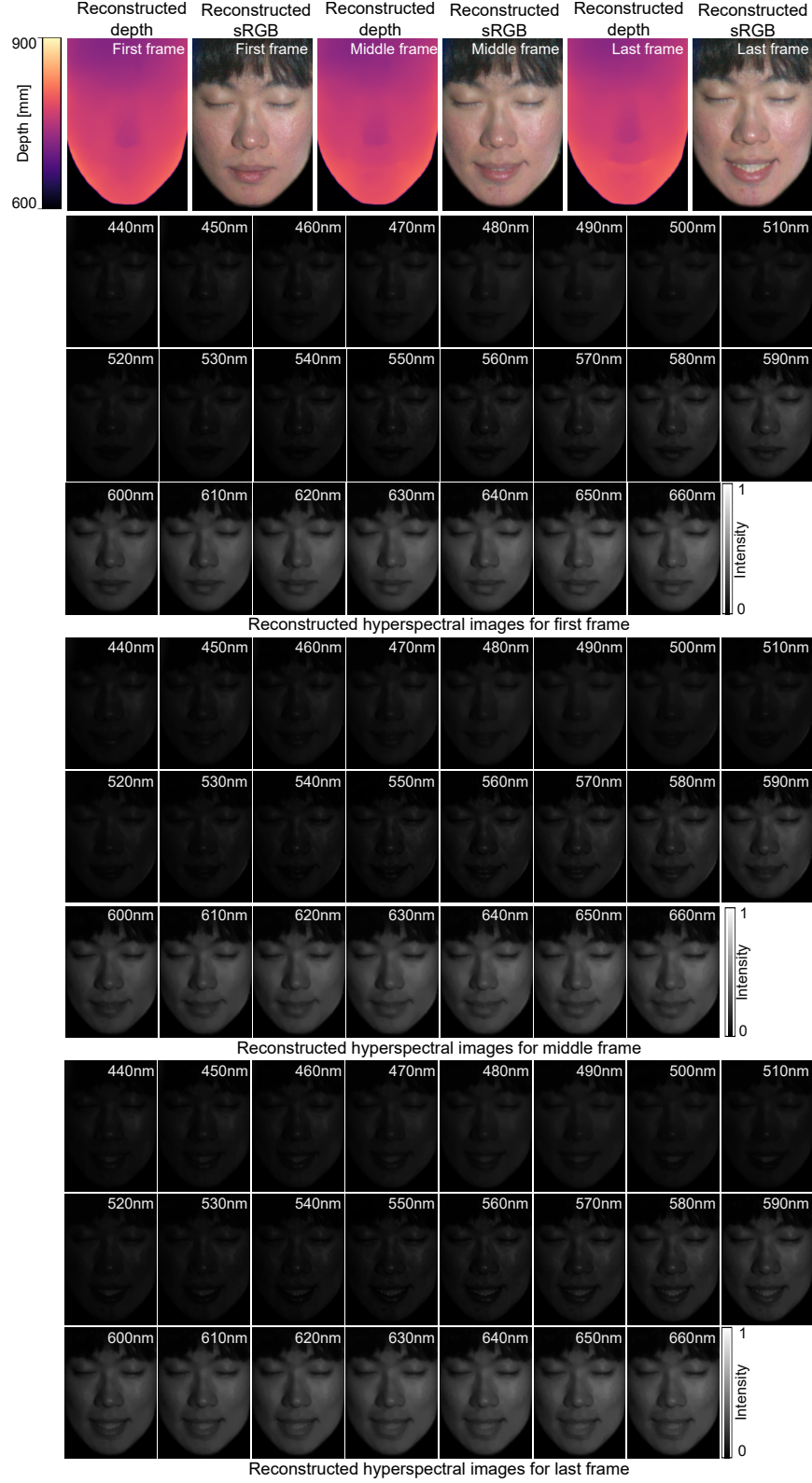


Figure 13. **Hyperspectral 3D imaging for dynamic scene.** We show the result of face scanned reconstructed hyperspectral image in sRGB, reconstructed depth map, and detailed reconstructed hyperspectral images from 440nm to 660nm at 10nm interval.

## References

- [1] P Kingma Diederik. Adam: A method for stochastic optimization. (*No Title*), 2014. 7
- [2] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Adv. Opt. Photon.*, 3(2):128–160, Jun 2011. 7, 9
- [3] Chunyu Li, Yusuke Monno, Hironori Hidaka, and Masatoshi Okutomi. Pro-cam ssfm: Projector-camera system for structure and spectral reflectance from motion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2414–2423, 2019. 12
- [4] Chunyu Li, Yusuke Monno, and Masatoshi Okutomi. Deep hyperspectral-depth reconstruction using single color-dot projection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19770–19779, 2022. 4
- [5] Lahav Lipson, Zachary Teed, and Jia Deng. Raft-stereo: Multilevel recurrent field transforms for stereo matching. In *2021 International Conference on 3D Vision (3DV)*, pages 218–227. IEEE, 2021. 6
- [6] Suhyun Shin, Seokjun Choi, Felix Heide, and Seung-Hwan Baek. Dispersed structured light for hyperspectral 3d imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24997–25006, 2024. 2, 7, 12
- [7] Gabriel Taubin, Daniel Moreno, and Douglas Lanman. 3d scanning for personal 3d printing: build your own desktop 3d scanner. In *ACM SIGGRAPH 2014 Studio*, pages 1–66. ACM, 2014. 2
- [8] Zachary Teed and Jia Deng. Raft: Recurrent all-pairs field transforms for optical flow. In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16*, pages 402–419. Springer, 2020. 6, 8
- [9] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000. 2