Dispersed Structured Light for Hyperspectral 3D Imaging

Suhyun Shin[1] Seokjun Choi[1] Felix Heide[2] Seung-Hwan Baek[1]

[1] POSTECH [2] Princeton University

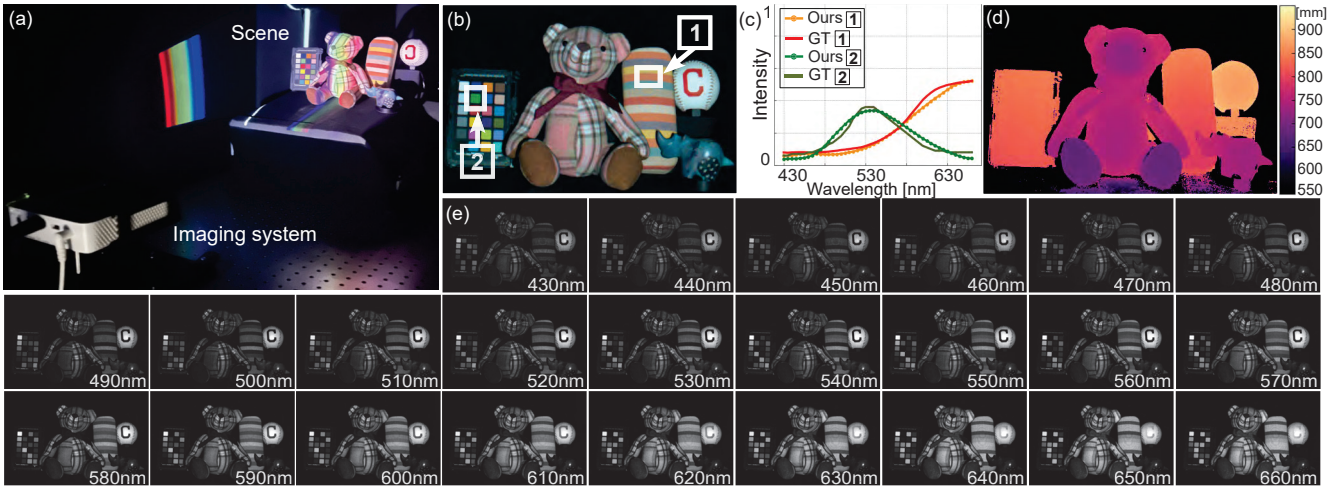


Figure 1. We introduce dispersed structured light, a low-cost high-quality hyperspectral 3D imaging method. By placing a diffraction grating film on a conventional camera-projector setup, we disperse structured-light patterns. Analyzing the images captured under the dispersed structured light enables accurate hyperspectral 3D reconstruction. (a) Capture configuration, (b) estimated hyperspectral image in sRGB, (c) comparison with spectroradiometer measurements, (d) estimated depth map, (e) estimated hyperspectral image.

Abstract

Hyperspectral 3D imaging aims to acquire both depth and spectral information of a scene. However, existing methods are either prohibitively expensive and bulky or compromise on spectral and depth accuracy. In this paper, we present Dispersed Structured Light (DSL), a cost-effective and compact method for accurate hyperspectral 3D imaging. DSL modifies a traditional projector-camera system by placing a sub-millimeter thick diffraction grating film front of the projector. This configuration enables dispersing structured light based on light wavelength. To utilize the dispersed structured light, we devise a model for dispersive projection image formation and a per-pixel hyperspectral 3D reconstruction method. We validate DSL by instantiating a compact experimental prototype. DSL achieves spectral accuracy of 18.8 nm full-width half-maximum (FWHM) and depth error of 1 mm, outperforming prior work on practical hyperspectral 3D imaging. DSL promises accurate and practical hyperspectral 3D imaging for diverse application domains, including computer vision and graphics, cultural heritage, geology, and biology.

1. Introduction

Hyperspectral 3D imaging aims to capture both depth and spectrum per pixel. Allowing for geometric-spectral analysis of a scene [4, 21–23, 44], this imaging modality has potential applications across diverse domains, including food ripeness detection [41], mineral detection [42], art authentication [36], classification [26], and cultural heritage preservation [19].

Despite of the potential in capturing rich geometric and spectral information, existing hyperspectral 3D imaging methods are often impractical, due to the high instrumentation costs, large form factor, and low accuracy. Specifically, high-end systems such as using coded-aperture snapshot spectral imaging systems (CASSI) and liquid-crystal tunable filters provide accurate depth and spectral information, however mandate high cost and large form factor [7, 14, 19, 37, 44–46, 48]. In contrast, affordable hyperspectral 3D imaging systems suffer from low accuracy of depth and spectrum [4, 15, 20–23, 29, 35].

In this paper, we exploit dispersion, a phenomenon where light rays are spatially redirected according to their wavelength either by refraction or diffraction. Using the optical dispersion, we present dispersed structured light

aiming for practical and accurate hyperspectral 3D imaging. DSL augments a conventional projector-camera system with diffraction grating film, which adds a negligible cost and size of around 10 USD with a sub-millimeter thickness. The diffraction grating in front of the projector spatially disperses broadband projector light based on its wavelength, resulting in a dispersed pattern consisting of several diffraction orders [13].

Specifically, zero-order diffraction permits light to pass through the diffraction grating as if the grating is not present, enabling us to exploit it for 3D imaging. First-order diffraction changes the direction of a light ray based on wavelength, resulting in dispersed structured light patterns that aid hyperspectral reconstruction. We develop a model for dispersive projection image formation and a per-pixel hyperspectral 3D reconstruction method using multi-order diffractions.

DSL achieves an average depth error of 1 mm and spectral FWHM of 18.8 nm in the visible spectrum, outperforming existing practical hyperspectral 3D imaging method [21] that has a spectral FWHM of 83 nm.

In summary, we make the following contributions.

- We present dispersed structured light that enables compact, low-cost, and high-quality hyperspectral 3D imaging by augmenting a projector-camera setup with a diffraction grating film.
- We develop a model for disperisve projection image formation and per-pixel hyperspectral 3D reconstruction method that considers both zero-order and first-order diffractions.
- We perform extensive evaluations and show that DSL achieves an average depth error of 1 mm and spectral FWHM of 18.8 nm, outperforming the state-of-the-art affordable hyperspectral 3D imaging methods.

2. Related Work

Hyperspectral 3D Imaging Existing work on hyperspectral 3D imaging typically combines the field of hyperspectral imaging [5, 24, 25, 27, 47] with depth imaging [15, 38, 44]. Specifically, researchers have paired CASSI [43] with structured light [7, 19], time-of-flight camera [37], light-field camera [45], and stereo [44]. An alternative to CASSI is the use of spectral bandpass filters with structured light [4, 14] or a light field camera [48]. Xu et al.[46] present a hyperspectral projector that incorporates collimation optics, diffraction grating, and digital micro-mirror device. Although these systems acquire accurate hyperspectral 3D information, their increased complexity due to relay lenses, narrow-band spectral filters, and plate-based dispersion optics results in a high building cost and large form factor.

Several methods have explored hyperspectral 3D imaging with compact setups. Baek et al. [4] propose a diffractive optical element (DOE) producing a point spread function (PSF) that varies with scene spectrum and depth, facilitating single-shot hyperspectral 3D imaging through blur analysis. However, the depth and spectral reconstruction accuracy are limited due to the low-frequency characteristics of the PSF. Li et al. [21, 23] employ a projector-camera setup to capture a scene with varying trichromatic projector primaries with known spectra, allowing for hyperspectral reconstruction. Despite its practicality, the spectral accuracy is limited by large spectral bandwidth of the projector primary spectra, resulting in 83 nm FWHM. In comparison, the proposed DSL enables accurate hyperspectral 3D imaging with 18.8 nm FWHM and average depth error of 1 mm by dispersing projector light at a cost and form factor on-par with conventional structured light systems.

Structured Light Structured light techniques project illumination patterns onto a scene and analyze the reflected light using a camera [8]. The projected patterns enable establishing correspondence between camera and projector pixels for 3D imaging. DOEs, combined with narrow-band coherent laser, have often been employed as cost-effective components for generating structured light [2, 18, 32]. Employing a conventional projector instead of the DOE-based laser illumination allows for using multiple patterns in a programmable manner, significantly enhancing depth accuracy [9, 38]. Various structured light patterns have been proposed to facilitate 3D imaging robust to global illumination [11, 31], light transport analysis [10, 33], and energy-efficient 3D imaging [34]. Introducing polarizing optics to a conventional projector further enhances its capability, making 3D imaging of translucent objects and polarimetric light transport analysis feasible [1, 17]. The proposed DSL augments structured light by only placing a diffraction grating film in front of a projector, enabling accurate hyperspectral 3D imaging with a compact setup.

Dispersive Optics for Cameras and Projectors Dispersive optics, such as prism and diffraction grating, have been often employed in the design of cameras and displays, in particular for hyperspectral imaging. CASSI employs dispersive optics and a coded mask to obtain masked spectral images with a wavelength-dependent translation [43]. Using a prism and a coded mask without relay lenses enables video-rate hyperspectral imaging [6]. Computed tomography imaging spectrometers leverage multi-order dispersion from a diffraction grating for hyperspectral imaging [30]. Hostettler et al.[16] use a prism and a mask to implement a trichromatic color projector based on dispersion. Mohan et al. [28] employ a diffraction grating and an attenuation mask to control the spectral power distribution of projector light. Recently, Sheinin et al.[39, 40] use a diffraction grating and a line camera to track fast-moving sparse scene

points. Our DSL augments a projector-camera system only with a diffraction grating film, allowing to keep the form factor compact and cost low. Instead of directly masking the spectrum emitted from the projector using additional components, we analyze and exploit how structured light patterns are dispersed and illuminate a scene.

3. Dispersive Projection Image Formation

Imaging Setup We devise the proposed DSL imaging system with the goals of compactness and affordability. To this end, we combine a conventional trichromatic camera (FLIR GS3-U3-32S4C-C) with a trichromatic projector (LG PH30N) and a diffraction-grating film (Edmund #54-509) mounted in front. This configuration makes light from the projector undergo dispersion, and, as such, patterns emitted from the projector are spatially dispersed depending on wavelength. Figure 2 shows our experimental prototype and a captured image with a projector pattern composed of four squares. For zero-order diffraction, light does not experience wavelength-dependent changes, and the projected squares retain their original shapes. For first-order diffractions, each square undergoes wavelength-dependent shift.

3.1. Image Formation

Background on Diffraction Grating Diffraction grating consists of repetitive, even-spaced micron-scale grooves characterized by the groove density g. The interaction of light with the grating incurs wavelength-dependent diffraction [13], which is modeled in the geometric-optics perspective as a redirection of the light path:

$$\mathtt{dg}(\mathbf{v}, m, \lambda) = (\underbrace{-mg\lambda + v_x}_{d_x}, \underbrace{v_y}_{d_y}, \sqrt{1 - d_x^2 - d_y^2}), \quad (1)$$

where m is the diffraction order, $\mathbf{v} = (v_x, v_y, v_z)$ is the incident light vector. x-axis here is aligned with the groove direction and λ is the wavelength. $\mathtt{dg}(\mathbf{v}, m, \lambda)$ is the ray direction after diffraction. In addition to the direction change, the diffraction grating introduces a change in light intensity depending on diffraction order and wavelength, represented as diffraction efficiency $\eta_{m,\lambda}$ [13].

We use a diffraction grating which creates positive and negative first-order diffractions appearing to the left and right sides with respect to the zero-order component as shown in Figure 2(c). Note that higher-order diffractions are not detected inside of our camera FoV.

Forward Model For a projector pattern $P(q, c)$, where q is a projector pixel and c is a color channel (R, G, or B), the light intensity $L(q, \lambda)$ emitted for each wavelength λ is

$$L(q, \lambda) = \sum_c \Omega^{\text{proj}}_{c,\lambda} P(q, c), \quad (2)$$

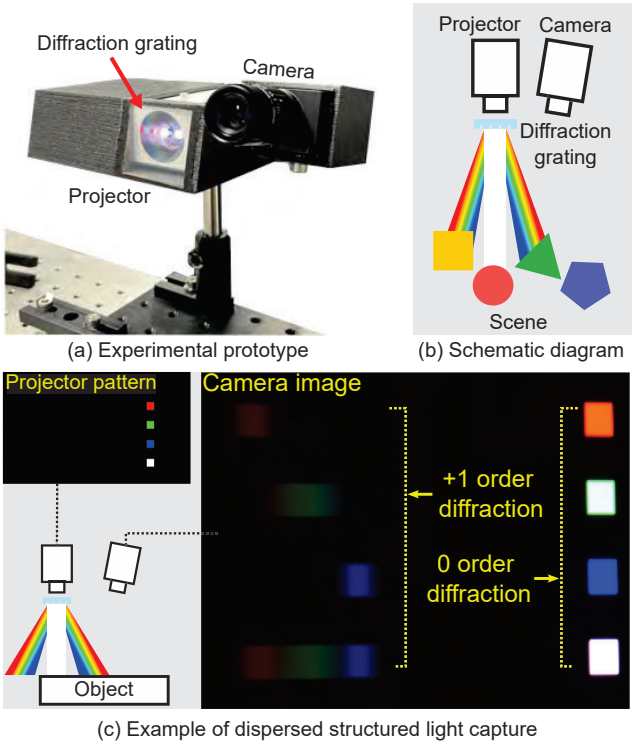


Figure 2. **Experimental prototype.** (a) & (b) Our prototype consists of an RGB projector equipped with a diffraction grating film, and an RGB camera. (c) An example projector pattern and its corresponding captured image, exhibiting clear first-order diffraction.

where $\Omega^{\text{proj}}_{c,\lambda}$ is the projector emission function shown in Figure 3(a).

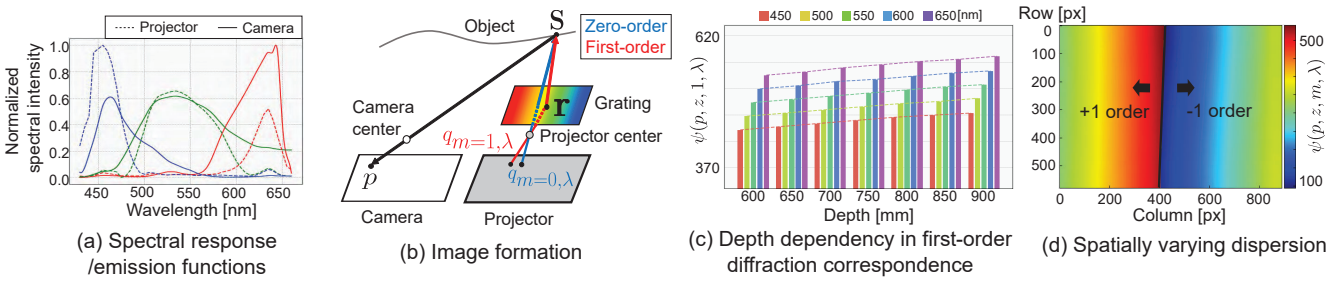
The light then passes through the diffraction grating and splits into multiple light rays with varying propagation directions $\mathtt{dg}(\mathbf{v}, m, \lambda)$, depending on diffraction order m and wavelength λ. Here, we focus on a ray corresponding to wavelength λ and diffraction order m. The ray will propagate and eventually illuminate a scene point $\mathbf{S}$ at depth z. The reflected light from the scene point will be captured by a camera pixel p. We then establish a correspondence function ψ that maps the camera pixel p, depth z, diffraction order m, and wavelength λ to the corresponding projector pixel $q_{m,\lambda}$ that emitted the light ray illuminating the scene point:

$$q_{m,\lambda} = \psi\left(p, z, m, \lambda\right). \quad (3)$$

We proceed to describe our full image formation before describing how we model the correspondence function ψ.

The intensity at the camera pixel p and color channel c then can be modeled as

$$I(p, c) = \sum_{\lambda \in \Lambda} \Omega^{\text{cam}}_{c,\lambda} H(p, \lambda) \underbrace{\frac{1}{d(p)^2} \sum_{m=-1}^{1} \eta_{m,\lambda} L(q_{m,\lambda}, \lambda)}_{\text{Dispersed structured light}}, \quad (4)$$



(a) Spectral response /emission functions (b) Image formation (c) Depth dependency in first-order diffraction correspondence (d) Spatially varying dispersion

Figure 3. **Image formation.** (a) Camera response function and projector emission function. (b) Schematic diagram of image formation. (c) Depth dependency of the correspondence function ψ for $m = 1$ and a camera pixel p. (d) Spatially-varying correspondence map for depth 700 mm and wavelength 430 nm for the first-order diffractions $m = -1, 1$.

where $\Omega^{\text{cam}}_{c,\lambda}$ is the camera response function shown in Figure 3(a), $\Lambda = [\lambda_1, \cdots, \lambda_{47}]$ is the set of 47 wavelengths sampled from 430 nm to 660 nm with 5 nm interval. $H(p, \lambda)$ is the hyperspectral intensity, which we aim to estimate. $d(p)$ is the propagation distance from the projector to the scene point $\mathbf{S}$, as such, the term $1/d(p)^2$ describes inverse-square law.

The dispersed structured light term in Equation (4) describes the light intensity projected to the scene point $\mathbf{S}$ by aggregating zero- and first-order light energy $L(q_{m,\lambda}, \lambda)$ weighted by the corresponding diffraction efficiency $\eta_{m,\lambda}$. In the following, we model the correspondence function ψ introduced in Equation (3).

3.2. Correspondence under Dispersion

Zero-order Diffraction For zero-order diffracted light ($m = 0$), light transport can be analyzed as if no diffraction grating exists [13]: given a camera pixel p and its depth z, we obtain the corresponding projector pixel $q_{0,\lambda}$ for any visible wavelength λ by applying perspective unprojection and projection [12]:

$$q_{0,\lambda} = \psi\left(p, z, 0, \lambda\right) \tag{5}$$

$$= \texttt{project}\left(\texttt{unproject}\left(p, z\right)\right), \tag{6}$$

where $\texttt{unproject}(\cdot)$ is the perspective unprojection from a camera pixel p to a 3D point $\mathbf{S}$ using the depth z. $\texttt{project}(\cdot)$ is the perspective projection from the 3D point $\mathbf{S}$ to the projector.

First-order Diffractions Correspondence for first-order diffractions ($m = -1$ or 1) is more challenging to model. Due to the direction change by diffraction grating, direct perspective projection from the scene point $\mathbf{S}$ to the projector is no longer valid. Instead, we need to identify a point $\mathbf{r}$ on the diffraction grating where the light ray from the projector center $\mathbf{t}_{\text{proj}}$ diffracts at $\mathbf{r}$ and reaches to the scene point $\mathbf{S}$. This is illustrated in Figure 3(b) and by using Equation (1) it can be formulated as

$$\underset{\mathbf{r}}{\text{minimize}} \; \|\texttt{dg}(\overrightarrow{\mathbf{t}_{\text{proj}}\mathbf{r}}, m, \lambda) - \overrightarrow{\mathbf{r}\mathbf{S}}\|_2, \tag{7}$$

where $\dot{\overrightarrow{x}}$ represents the normalized vector of $\overrightarrow{x}$. While the formulation is intuitive, solving Equation (7) over the entire 3D volume, for each wavelength λ and diffraction order m, is prone to calibration errors and is computationally demanding when approached with iterative numerical methods like the Newton-Raphson method or root-finding algorithms.

Hence, instead of directly solving Equation (7), we develop a data-driven model for efficient and accurate correspondence mapping of the first-order diffractions. We prepare samples: first-order corresponding projector pixel positions $q'_{m,\lambda}$ for wavelengths $\lambda' \in \mathbb{S}_\lambda$, grid-sampled camera pixels $p' \in \mathbb{S}_p$, and depths $z' \in \mathbb{S}_z$. Section 5 describes the sample acquisition. We then fit an exponential function to the samples for modeling depth dependency of the correspondence function ψ and perform linear interpolation for spatial and spectral dimensions. As the first-order diffraction in our prototype occurs along the horizontal axis, and our hyperspectral reconstruction only uses horizontal correspondence, we only model the horizontal coordinate of the correspondence function, i.e., the column index. More details on our data-driven correspondence model can be found in the Supplemental Document.

Our data-driven first-order correspondence model obtains a mean reprojection error of 1px and can be efficiently evaluated with tabulation in $O(1)$. Figure 3(c) shows the depth dependency of the modeled correspondence function ψ. Figure 3(d) shows the spatially-varying correspondence map for depth 700 mm, wavelength 430 nm, and the diffraction orders $m = 1$ and -1.

4. Hyperspectral 3D Reconstruction

With the image formation model from above in hand, we estimate a depth map using binary-code structured light and then reconstruct the hyperspectral image using scanline structured light.

4.1. Depth Reconstruction

For depth estimation, we use binary-code structured light patterns shown in Figure 4 and represented as

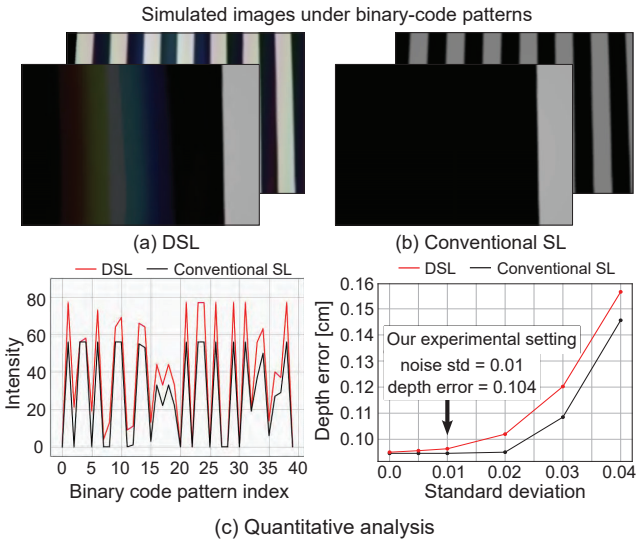


Figure 4. **Binary decoding under dispersion.** Simulated images under the binary-code patterns for (a) DSL and (b) conventional SL. (c) Intensity of a camera pixel with DSL and conventional SL. (d) Depth error of binary-code decoding for DSL and conventional SL with varying Gaussian noise.

$[P_1, \dots, P_{K_b}]$: $P_i(q, \forall c) = \texttt{bit}(q, i)$, where $\texttt{bit}(q, i)$ returns the i-th binary bit of the decimal horizontal and vertical location of the projector pixel q, encoding the projector pixel location as a binarized code [9]. $K_b = 40$ is the number of patterns. Then, each camera pixel p observes the fluctuating intensity values under these binary-code patterns, producing an observation vector $[I_1(p, \forall c), \dots, I_{K_b}(p, \forall c)]$ shown in Figure 4(c).

Binary-code Decoding From the observation vector, we aim to estimate the zero-order correspondence $q_{0,\lambda}$, which will facilitate the estimation of depth z through triangulation using the zero-order correspondence function $\psi(p, z, 0, \lambda)$. Note that zero-order correspondence has no dependency on wavelength λ. For the i-th projector pattern P_i, the conventional decoding method binarizes the captured intensity I_i into 0 or 1 based on its intensity level by using RGB-to-gray conversion and thresholding with a constant τ. The binary code is then decoded to a decimal number, which is the location of the corresponding projector pixel $q_{0,\lambda}$.

Validation of Binary-code Decoding for DSL An important question is whether the conventional binary decoding works for DSL in the presence of dispersed light. We provide both mathematical and experimental verification. In theory, this can be confirmed if the intensity lit by the zero-order diffracted light exceed the intensity when not receiving the zero-order diffracted light. We derive the minimum captured intensity $I^{\text{on}}_{\text{min}}$ when a scene point is only illuminated by the zero-order structured light. Also, we model the maximum captured intensity $I^{\text{off}}_{\text{max}}$ when the zero-order structured light does not illuminate the point and the first-order light illuminates. The quantities are written as

$$I^{\text{on}}_{\text{min}}(p, c) = \sum_{\lambda} \Omega^{\text{cam}}_{c,\lambda} H(p, \lambda) \eta_{0,\lambda} \sum_{c'} \Omega^{\text{proj}}_{c',\lambda}, \tag{8}$$

$$I^{\text{off}}_{\text{max}}(p, c) = \sum_{\lambda} \Omega^{\text{cam}}_{c,\lambda} H(p, \lambda) \sum_{m=1,-1} \eta_{m,\lambda} \sum_{c'} \Omega^{\text{proj}}_{c',\lambda}. \tag{9}$$

We then compare both quantities, giving an observation that the inequality $I^{\text{off}}_{\text{max}} < I^{\text{on}}_{\text{min}}$ holds in our setup configuration, because for all wavelength λ the following inequality holds $\eta_{0,\lambda} \sum_{c'} \Omega^{\text{proj}}_{c',\lambda} > \sum_{m=1,-1} \eta_{m,\lambda} \sum_{c'} \Omega^{\text{proj}}_{c',\lambda}$. This validates the use of conventional binary-code decoding for DSL in theory. Refer to the Supplemental Document for the derivation details.

Next, we further validate the applicability of binary decoding on 10^6 simulated samples with random hyperspectral reflectance for a planar object with Gaussian measurement noise. Example observation vectors with and without considering first-order diffractions are shown in Figure 4(c). We set the Gaussian standard deviation 0.01 corresponding to the real-world noise in our hardware measurements. The average depth error after decoding, shown in Figure 4(d), is 1.04 mm in the presence of first-order diffractions. This experiment further validates the use of binary-code decoding for DSL.

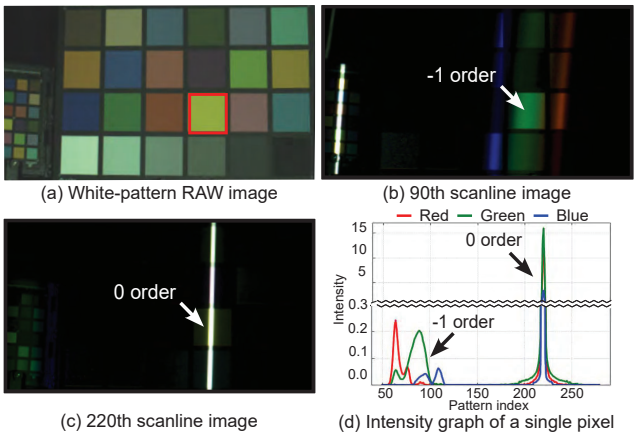
4.2. Hyperspectral Reconstruction

Once depth z is obtained, we proceed to estimate per-pixel hyperspectral intensity $\mathbf{H} = [H(p, \lambda_1), \cdots, H(p, \lambda_N)] \in \mathbb{R}^{N \times 1}$. To this end, we use the scanline structured light patterns $[P_1, , ..., P_{K_s}]$ that scans through the whole column with a line width w. $K_s = 318$ is the number of scanline patterns. Figure 5 shows that the first-order diffracted light from the i-th scanline pattern produces a narrow-band illumination spread across multiple columns. Consequently, scanning columns from left to right enables illuminating each scene point with every narrow-band light, facilitating high-quality hyperspectral reconstruction.

Pixel Intensity under Scanline Patterns For the i-th scanline pattern, a camera pixel p receives either zero-order diffraction, first-order diffractions, or no illumination at all. This simplifies our image formation as

$$I_i(p, c) = \begin{cases} \sum_{\lambda} \Omega^{\text{cam}}_{c,\lambda} H(p, \lambda) \eta_{0,\lambda} L(q_{0,\lambda}, \lambda) & \text{for zero order,} \\ \Omega^{\text{cam}}_{c,\lambda} H(p, \lambda) \eta_{m,\lambda} L(q_{m,\lambda}, \lambda) & \text{for first orders,} \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Figure 5(c) shows the intensity graph of a camera pixel p for varying index of the scanline patterns:



(a) White-pattern RAW image (b) 90th scanline image (c) 220th scanline image (d) Intensity graph of a single pixel

Figure 5. **Hyperspectral imaging with scanline illuminations.** Captured images under (a) white pattern and (b)&(c) scanline patterns. Narrow-band illumination over multiple columns is shown in (b), which originates from the first-order diffraction. (d) Pixel intensity with respect to varying scanline pattern index.

$[I_1(p,c), \cdots, I_{K_s}(p,c)]$. For the first-order cases ($m = -1$ or 1), the intensity stems from a specific wavelength λ, hence narrow-band illumination.

Index Mapping for First-order Diffraction We exploit first-order diffractions for accurate hyperspectral reconstruction. We find the scanline pattern index i of which m-order diffracted light with wavelength λ illuminates the camera pixel p. In fact, we already obtained this index mapping in the form of the correspondence function ψ:

$$\delta_{m,i} = \texttt{px2index}\left(\psi\left(p, z, m, \lambda_i\right)\right), \tag{11}$$

where $\delta_{m,i}$ is the corresponding scanline pattern index and `px2index` is the conversion function that simply returns the scanline pattern index that lights up at the input projector pixel location.

We then collect the intensity captured under narrow-band illumination for all target wavelengths in $[\lambda_1, \cdots, \lambda_N]$ for each diffraction order m:

$$\mathbf{I}_m = [I_{\delta_{m,1}}(p,c), \cdots, I_{\delta_{m,N}}(p,c)] \in \mathbb{R}^{N\times 3}. \tag{12}$$

Optimization We formulate a per-pixel optimization problem for accurate hyperspectral reconstruction:

$$\underset{\mathbf{H}}{\operatorname{argmin}} \underbrace{\sum_{m=-1}^{1} \kappa_m \|\mathbf{A}_m\mathbf{H} - \mathbf{I}_m\|_2^2}_{\text{Data term}} + \underbrace{\kappa_\lambda \|\nabla_\lambda \mathbf{H}\|_2^2}_{\text{Regularization term}}, \tag{13}$$

where κ_m is the spatially-varying per-pixel balancing weight for the diffraction order m. Refer to the Supplemental Document how we set the balancing weight. $\mathbf{A}_m$ is the system matrix defined as

$$\mathbf{A}_m = \begin{cases} \sum_\lambda \Omega_{c,\lambda}^{\text{cam}} \eta_{m,\lambda} L(q_{m,\lambda}, \lambda) & \text{for zero order,} \\ \Omega_{c,\lambda}^{\text{cam}} \eta_{m,\lambda} L(q_{m,\lambda}, \lambda) & \text{for first orders.} \end{cases} \tag{14}$$

∇_λ is the gradient operator along the spectral axis. The first term in Equation (13) accounts for the reconstruction loss across the multiple diffraction orders. The second term is the spectral smoothness prior [3]. We use gradient descent for the optimization, which takes three minutes to converge. Our hyperspectral reconstruction method operates on a per-pixel basis and exploits first-order diffractions, enabling accurate hyperspectral image reconstruction.

5. Calibration

We calibrate the image formation parameters of the projector, camera, and diffraction grating, as briefly described in the following. A detailed description of the calibration procedure can be found in the Supplemental Document.

Diffraction Efficiency To calibrate the diffraction efficiency $\eta_{m,\lambda}$, we measure the intensity of m-order diffracted light at each wavelength λ projected onto a Spectralon sample. We use spectral bandpass filters at 10 nm intervals, from 430 nm to 660 nm. Diffraction efficiency is then computed as the intensity ratio of each first-order wavelength measurement over the zero-order intensity.

Spectral Response and Emission Functions The projector spectral emission function, $\Omega_{\lambda,c}^{\text{proj}}$, was obtained by projecting red, green, and blue dots onto a Spectralon target, measuring the reflected radiance with a spectroradiometer (JETI Specbos 1211), and normalizing the results with Spectralon reflectance. For the camera response function, $\Omega_{\lambda,c}^{\text{cam}}$, we use the data provided by the camera manufacturer. For both emission and response functions, we perform refinements of which details can be found in the Supplemental Document. Figure 3(a) shows the calibrated functions.

First-order Correspondence Model To calibrate the first-order correspondence model described in Section 3.2, we acquired images of flat Spectralon surfaces at five different depth positions with scanline illumination patterns present. These images were captured using multiple bandpass filters. For sampled camera pixels, denoted p', we identified the corresponding projector pixel $q'_{m,\lambda}$ from the captured images. Using the samples, we obtain the data-driven corresponding model ψ.

6. Assessment

Hyperspectral 3D Reconstruction DSL enables accurate hyperspectral 3D imaging. We estimate a hyperspectral

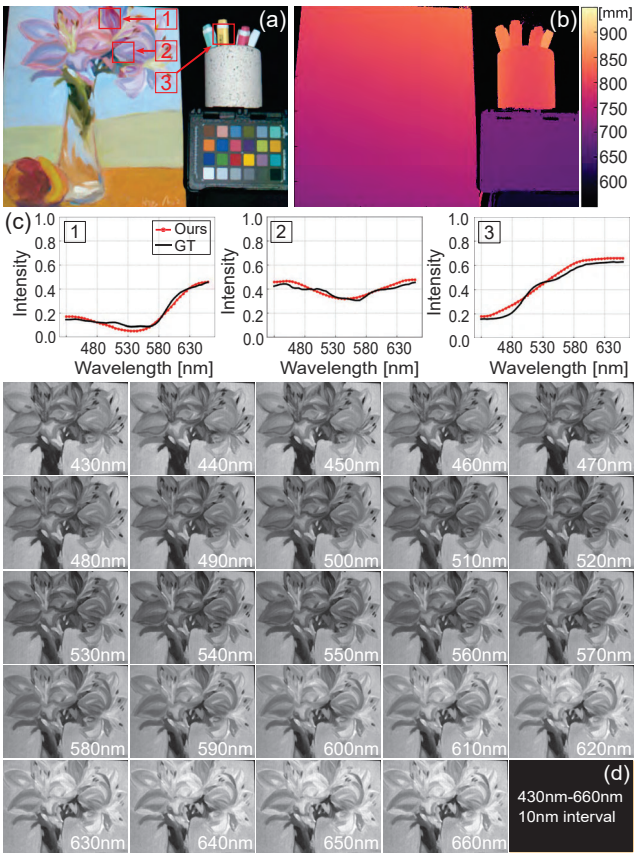


Figure 6. **Hyperspectral 3D imaging.** (a) Reconstructed hyperspectral image visualized in sRGB, (b) reconstructed depth map, (c) estimated hyperspectral intensity for three different points compared with spectroradiometer measurements, (d) estimated hyperspectral image.

image with 46 channels, from 430 nm to 660 nm at 5 nm intervals, along with a depth map. Figures 1 and 6 shows reconstruction results for two real-world scenes.

Reconstruction of High-frequency Spectral Curves Figure 7 shows the results of our DSL in comparison with Li et al. [21], state-of-the-art practical hyperspectral 3D imaging method. We captured a scene containing nine bandpass spectral filters, each with a bandwidth of 10 nm. DSL accurately identifies the center wavelengths and achieves an average FWHM of 18.8 nm. In contrast, Li et al. [21] rely on broadband RGB illuminations of projector, resulting in a significantly broader FWHM of 83 nm, due to the limited capability in differentiating high-frequency spectral features. This performance gap mainly originates from our use of first-order diffractions. We test the DSL without using the first-order term in Equation (13). As expected, the resuting spectra is overly smooth with 50 nm FWHM and cannot detect the high-peak spectral features. This is aligned with the results of Li et al.[21], demonstrating the importance of using first-order diffractions.

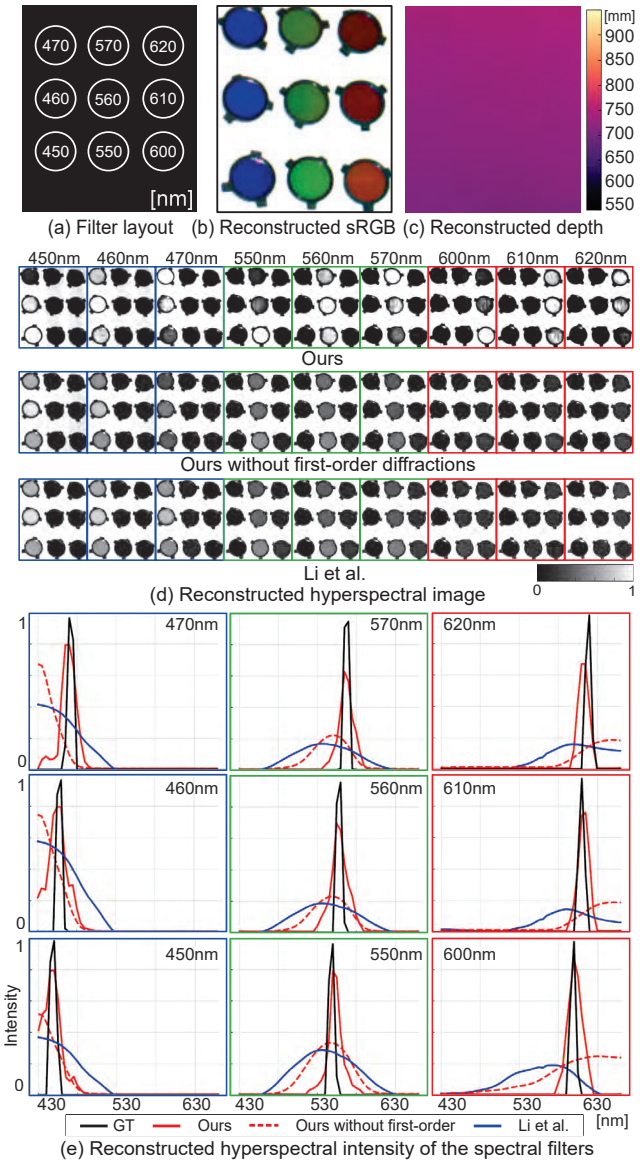


Figure 7. **High-frequency spectral imaging.** (a) Filter layout, (b) reconstructed hyperspectral image in sRGB, (c) reconstructed depth, (d) reconstructed hyperspectral images, (e) spectral curves for the nine bandpass filters.

Colorchecker and Metamerism Figure 8 shows the spectral accuracy of our DSL measured on a ColorChecker and metameric fake and real leaves. For smooth spectral curves of color patches, DSL successfully reconstructs the hyperspectral intensity. Also, DSL enables telling clear difference between fake and real leaves. We obtained the ground-truth intensity using a spectroradiometer.

Depth Accuracy To assess the depth accuracy of our experimental prototype, we assess both relative and absolute depth errors. Figure 9(a) show that DSL accurately estimates the distance between two boxes, with a marginal error

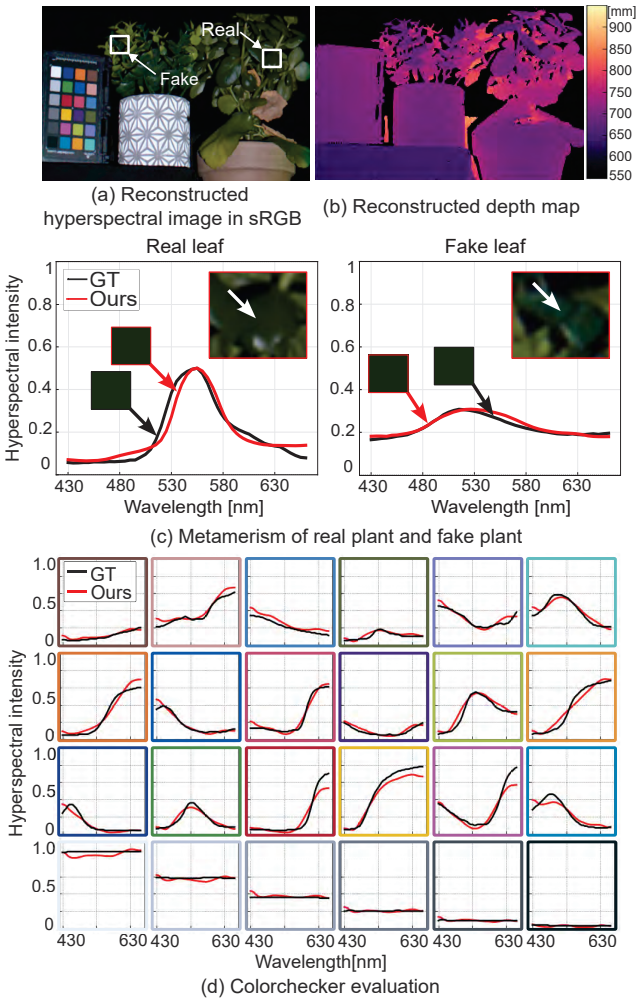


Figure 8. **Evaluation on a ColorChecker and metameric samples.** (a) Reconstructed hyperspectral image in sRGB, (b) reconstructed depth map, (c) spectra of metameric samples (real and fake leaves), (d) reconstructed spectra of each color patch.

of only 1 mm. Figure 9(b) evaluates the absolute depth error by capturing a planar object mounted on a linear translation stage (Thorlabs #LTS150C). Across the working range of the translation stage with a 10 mm step size, DSL achieves an average depth error of 1.35 mm. Note that these experimental results are aligned with our synthetic experiments shown in Figure 4.

7. Conclusion

In this paper, we introduced DSL, an accurate, low-cost, and compact hyperspectral 3D imaging method. Our dispersion-aware image formation, per-pixel hyperspectral 3D reconstruction, and calibration enables accurate hyperspectral 3D imaging. DSL can be implemented in an affordable experimental prototype by using a diffraction grating adding sub-millimeter thickness at a cost of 10 USD. Our experimental prototype achieves depth error of 1 mm

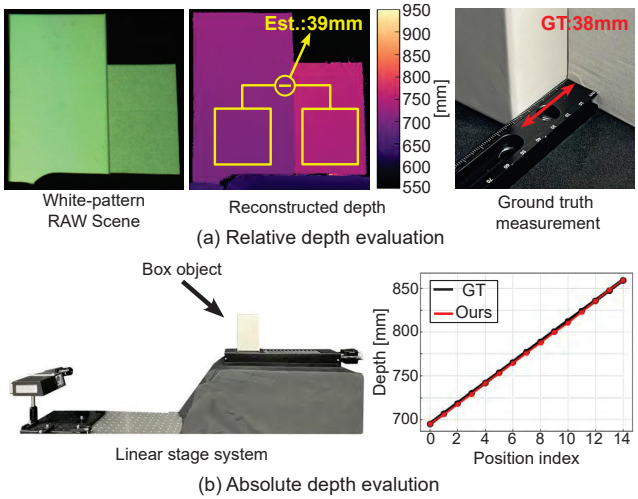


Figure 9. **Depth accuracy.** We estimate (a) the distance between two boxes and (b) the depth of the box mounted on a translation stage. We obtain the depth errors of 1 mm for the relative-depth experiment and 1.35 mm for the absolute depth experiment.

and spectral FWHM of 18.8 nm, outperforming prior work on practical hyperspectral 3D imaging. We find that DSL makes a step towards practical hyperspectral 3D imaging for applications beyond computer vision and graphics.

Limitations and Future Work While accurate and low-cost, our current experimental prototype takes around 10 minutes to capture a scene, restricting it to static scenes. Also, the low intensity of first-order diffracted light limits the working depth range to be less than a meter.

For handling dynamic scenes and increasing depth range, we leave developing a light-efficient capture system and joint depth-spectrum reconstruction method as an interesting direction for future research. Another unexplored direction is to find the optimal design of the diffraction pattern for efficient and accurate hyperspectral 3D imaging. Recent differentiable optimization of imaging systems could be employed to achieve the goal. Lastly, the core principle of DSL can be applied to other spectral ranges beyond visible spectrum.

Acknowledgements Seung-Hwan Baek was supported by Korea NRF (RS-2023-00211658, 2022R1A6A1A03052954), Korea IITP MSIT (No.2019-0-01906, Artificial Intelligence Graduate School Program-POSTECH), and POSCO. Felix Heide was supported by an NSF CAREER Award (2047359), a Packard Foundation Fellowship, a Sloan Research Fellowship, a Sony Young Faculty Award, a Project X Innovation Award, and an Amazon Science Research Award.

References

[1] Seung-Hwan Baek and Felix Heide. Polarimetric spatio-temporal light transport probing. *ACM Trans. Graph.*, 40 (6):1–18, 2021. 2

[2] Seung-Hwan Baek and Felix Heide. Polka lines: Learning structured illumination and reconstruction for active stereo. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 5757–5767, 2021. 2

[3] Seung-Hwan Baek, Incheol Kim, Diego Gutierrez, and Min H Kim. Compact single-shot hyperspectral imaging using a prism. *ACM Trans. Graph.*, 36(6):1–12, 2017. 6

[4] Seung-Hwan Baek, Hayato Ikoma, Daniel S Jeon, Yuqi Li, Wolfgang Heidrich, Gordon Wetzstein, and Min H Kim. Single-shot hyperspectral-depth imaging with learned diffractive optics. In *Int. Conf. Comput. Vis.*, pages 2651–2660, 2021. 1, 2

[5] Yuanhao Cai, Jing Lin, Xiaowan Hu, Haoqian Wang, Xin Yuan, Yulun Zhang, Radu Timofte, and Luc Van Gool. Coarse-to-fine sparse transformer for hyperspectral image reconstruction. In *Eur. Conf. Comput. Vis.*, pages 686–704. Springer, 2022. 2

[6] Xun Cao, Xin Tong, Qionghai Dai, and Stephen Lin. High resolution multispectral video capture with a hybrid camera system. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 297–304. IEEE, 2011. 2

[7] Elkin Díaz, Jaime Meneses, and Henry Arguello. Hyperspectral+ depth imaging using compressive sensing and structured light. In *3D Image Acquisition and Display: Technology, Perception and Applications*, pages 3M3G–6. Optica Publishing Group, 2018. 1, 2

[8] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Adv. Opt. Photon.*, 3(2):128–160, 2011. 2

[9] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Adv. Opt. Photon.*, 3(2):128–160, 2011. 2, 5

[10] Mohit Gupta, Yuandong Tian, Srinivasa G Narasimhan, and Li Zhang. (de) focusing on global light transport for active scene recovery. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2969–2976. IEEE, 2009. 2

[11] Mohit Gupta, Qi Yin, and Shree K Nayar. Structured light in sunlight. In *Int. Conf. Comput. Vis.*, pages 545–552, 2013. 2

[12] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. 4

[13] Eugene Hecht. *Optics (Fourth Edition)*. Addison-Wesley, 2001. 2, 3, 4

[14] Stefan Heist, Chen Zhang, Karl Reichwald, Peter Kühmstedt, Gunther Notni, and Andreas Tünnermann. 5d hyperspectral imaging: fast and accurate measurement of surface shape and spectral characteristics using structured light. *Optics express*, 26(18):23366–23379, 2018. 1, 2

[15] Keita Hirai, Ryosuke Nakahata, and Takahiko Horiuchi. Measuring spectral reflectance and 3d shape using multi-primary image projector. In *Image and Signal Processing: 7th International Conference, ICISP 2016, Trois-Rivières, QC, Canada, May 30-June 1, 2016, Proceedings 7*, pages 137–147. Springer, 2016. 1, 2

[16] Rafael Hostettler, Ralf Habel, Markus Gross, and Wojciech Jarosz. Dispersion-based color projection using masked prisms. *Comput. Graph. Forum*, 34(7), 2015. 2

[17] Daniel S Jeon, Andréas Meuleman, Seung-Hwan Baek, and Min H Kim. Polarimetric itof: Measuring high-fidelity depth through scattering media. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 12353–12362, 2023. 2

[18] Gyeongtae Kim, Yeseul Kim, Jooyeong Yun, Seong-Won Moon, Seokwoo Kim, Jaekyung Kim, Junkyeong Park, Trevon Badloe, Inki Kim, and Junsuk Rho. Metasurface-driven full-space structured light for three-dimensional imaging. *Nature Communications*, 13(1):5920, 2022. 2

[19] Min H Kim, Todd Alan Harvey, David S Kittle, Holly Rushmeier, Julie Dorsey, Richard O Prum, and David J Brady. 3d imaging spectroscopy for measuring hyperspectral patterns on solid objects. *ACM Trans. Graph.*, 31(4):1–11, 2012. 1, 2

[20] Masahiro Kitahara, Takahiro Okabe, Christian Fuchs, and Hendrik PA Lensch. Simultaneous estimation of spectral reflectance and normal from a small number of images. In *VISAPP (1)*, pages 303–313, 2015. 1

[21] Chunyu Li, Yusuke Monno, Hironori Hidaka, and Masatoshi Okutomi. Pro-cam ssfm: Projector-camera system for structure and spectral reflectance from motion. In *Int. Conf. Comput. Vis.*, pages 2414–2423, 2019. 1, 2, 7

[22] Chunyu Li, Yusuke Monno, and Masatoshi Okutomi. Spectral mvir: Joint reconstruction of 3d shape and spectral reflectance. In *IEEE Int. Conf. Comput. Photo.*, pages 1–12. IEEE, 2021.

[23] Chunyu Li, Yusuke Monno, and Masatoshi Okutomi. Deep hyperspectral-depth reconstruction using single color-dot projection. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 19770–19779, 2022. 1, 2

[24] Lingen Li, Lizhi Wang, Weitao Song, Lei Zhang, Zhiwei Xiong, and Hua Huang. Quantization-aware deep optics for diffractive snapshot hyperspectral imaging. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 19780–19789, 2022. 2

[25] Miaoyu Li, Ying Fu, Ji Liu, and Yulun Zhang. Pixel adaptive deep unfolding transformer for hyperspectral image reconstruction. In *Int. Conf. Comput. Vis.*, pages 12959–12968, 2023. 2

[26] S Mahesh, DS Jayas, J Paliwal, and NDG White. Hyperspectral imaging to classify and monitor quality of agricultural materials. *Journal of Stored Products Research*, 61:17–26, 2015. 1

[27] Maksim Makarenko, Arturo Burguete-Lopez, Qizhou Wang, Fedor Getman, Silvio Giancola, Bernard Ghanem, and Andrea Fratalocchi. Real-time hyperspectral imaging in hardware via trained metasurface encoders. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 12692–12702, 2022. 2

[28] Ankit Mohan, Ramesh Raskar, and Jack Tumblin. Agile spectrum imaging: Programmable wavelength modulation for cameras and projectors. In *Comput. Graph. Forum*, pages 709–717. Wiley Online Library, 2008. 2

[29] Giljoo Nam and Min H Kim. Multispectral photometric stereo for acquiring high-fidelity surface normals. *IEEE computer graphics and applications*, 34(6):57–68, 2014. 1

[30] F. Narea-Jiménez, J. Castro-Ramos, J. J. Sánchez-Escobar, and A. Mu noz Morales. Assessment of a computed tomography imaging spectrometer using an optimized expectation-maximization algorithm. *Appl. Opt.*, 61(20):6076–6085, 2022. 2

[31] Shree K Nayar, Gurunandan Krishnan, Michael D Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. In *ACM Trans. Graph.*, pages 935–944. ACM New York, NY, USA, 2006. 2

[32] Yibo Ni, Sai Chen, Yujie Wang, Qiaofeng Tan, Shumin Xiao, and Yuanmu Yang. Metasurface for structured light projection over 120 field of view. *Nano Letters*, 20(9):6719–6724, 2020. 2

[33] Matthew O'Toole, John Mather, and Kiriakos N Kutulakos. 3d shape and indirect appearance by structured light transport. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 3246–3253, 2014. 2

[34] Matthew O'Toole, Supreeth Achar, Srinivasa G Narasimhan, and Kiriakos N Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM Trans. Graph.*, 34 (4):1–13, 2015. 2

[35] Keisuke Ozawa, Imari Sato, and Masahiro Yamaguchi. Hyperspectral photometric stereo for a single capture. *JOSA A*, 34(3):384–394, 2017. 1

[36] Adam Polak, Timothy Kelman, Paul Murray, Stephen Marshall, David JM Stothard, Nicholas Eastaugh, and Francis Eastaugh. Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. *Journal of Cultural Heritage*, 26:1–11, 2017. 1

[37] Hoover Rueda-Chacon, Juan F Florez-Ospina, Daniel L Lau, and Gonzalo R Arce. Snapshot compressive tof+ spectral imaging via optimized color-coded apertures. *IEEE Trans. Pattern Anal. Mach. Intell.*, 42(10):2346–2360, 2019. 1, 2

[38] Simon Schreiberhuber, Jean-Baptiste Weibel, Timothy Patten, and Markus Vincze. Gigadepth: Learning depth from structured light with branching neural networks. In *Eur. Conf. Comput. Vis.*, 2022. 2

[39] Mark Sheinin, Dinesh N Reddy, Matthew O'Toole, and Srinivasa G Narasimhan. Diffraction line imaging. In *Eur. Conf. Comput. Vis.*, pages 1–16. Springer, 2020. 2

[40] Mark Sheinin, Matthew O'Toole, and Srinivasa G Narasimhan. Deconvolving diffraction for fast imaging of sparse scenes. In *IEEE Int. Conf. Comput. Photo.*, pages 1–10. IEEE, 2021. 2

[41] Da-Wen Sun. *Hyperspectral imaging for food quality analysis and control*. Elsevier, 2010. 1

[42] Mahesh Kumar Tripathi and H Govil. Evaluation of aviris-ng hyperspectral images for mineral identification and mapping. *Heliyon*, 5(11), 2019. 1

[43] Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47(10):B44–B51, 2008. 2

[44] Lizhi Wang, Zhiwei Xiong, Guangming Shi, Wenjun Zeng, and Feng Wu. Simultaneous depth and spectral imaging with a cross-modal stereo system. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(3):812–817, 2016. 1, 2

[45] Zhiwei Xiong, Lizhi Wang, Huiqun Li, Dong Liu, and Feng Wu. Snapshot hyperspectral light field imaging. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 3270–3278, 2017. 2

[46] Yibo Xu, Anthony Giljum, and Kevin F Kelly. A hyperspectral projector for simultaneous 3d spatial and hyperspectral imaging via structured illumination. *Optics Express*, 28(20): 29740–29755, 2020. 1, 2

[47] Xuanyu Zhang, Yongbing Zhang, Ruiqin Xiong, Qilin Sun, and Jian Zhang. Herosnet: Hyperspectral explicable reconstruction and optimal sampling deep network for snapshot compressive imaging. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 17532–17541, 2022. 2

[48] Kang Zhu, Yujia Xue, Qiang Fu, Sing Bing Kang, Xilin Chen, and Jingyi Yu. Hyperspectral light field stereo matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 41(5):1131–1143, 2018. 1, 2