Pro-Cam SSfM: projector–camera system for structure and spectral reflectance from motion

Chunyu Li · Yusuke Monno · Masatoshi Okutomi

1 Department of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology, Tokyo, Japan

Abstract

Image-based reconstruction of an object’s 3D shape having the wavelength-by-wavelength spectral reflectance property enables higher-fidelity object 3D modeling compared with typical RGB-based modeling. In this paper, we propose a novel projector–camera system for practical and low-cost acquisition of a dense object 3D model with the spectral reflectance property. Different from existing spectral 3D data acquisition systems that use a dedicated multispectral camera or light, we use a standard RGB camera and an off-the-shelf projector as active illumination for both the 3D reconstruction and the spectral reflectance estimation. We first propose a calibration-free multi-view structured-light method to reconstruct the 3D points while estimating the intrinsic parameters and the poses of both the camera and the projector, which are alternately moved around the object during our image acquisition procedure. We then exploit the projector for multispectral imaging and estimate the spectral reflectance of each 3D point based on a novel spectral reflectance estimation method considering the geometric relationship between the reconstructed 3D points and the estimated projector positions. Experimental results on both synthetic and real data demonstrate that our system can precisely acquire a dense spectral 3D model using off-the-shelf devices.

Keywords Projector–camera system · Structured light · Structure from motion · Spectral reflectance estimation

1 Introduction

The nature of an object is typically represented by two properties: geometric and photometric properties. The geometric property is determined by the 3D structure of the object, while the photometric property is determined by how the incident light is reflected at each 3D point of the object surface. Among various photometric parameters, spectral reflectance is one of the most fundamental physical quantities, which defines the amount of reflected light over that of the incident light at each wavelength. In this work, our aim is to acquire the spectral 3D information of an object from multi-view images captured using low-cost off-the-shelf devices (see Fig. 1). Practical and low-cost acquisition of the spectral 3D data has many potential applications requiring high-fidelity object 3D modeling, such as cultural heritage [11, 33], plant modeling [7, 39], spectral rendering [14], and multimedia [41].

Image-based 3D reconstruction is a very active research area in computer vision. A popular approach is a structure-from-motion (SfM) [52, 55] and multi-view stereo (MVS) [19, 53] pipeline, which is based on feature correspondences among multiple images taken from different viewpoints. Another common approach is a structured-light method [18, 20, 21, 26, 51], which uses a projector–camera system to densely reconstruct the 3D points even for texture-less
(i.e., feature-less) objects by actively projecting structured-light feature patterns. Although these approaches have extensively been studied, they usually focus on the geometric 3D reconstruction. Some recent methods combine the geometric and the photometric 3D reconstruction \[31,42,43\]. However, they still focus on the estimation of RGB albedos. Different from the spectral reflectance, the RGB albedos are not inherent to the object, since the RGB values depend on device-dependent RGB camera sensitivity, meaning that the RGB albedos have essentially no physical meaning in the sense of “object” modeling.

Multispectral imaging is another active research area. Various hardware-based systems \[6,9,12,23,44,48,57\] and software-based methods \[2,5,15,28,46,54\] have been proposed for recovering scene’s spectral reflectance. However, they usually assume a single-viewpoint input image and do not consider the geometric relationship between the object surface and the light source, only achieving scene- and viewpoint-dependent spectral recovery, where the shading and the shadows that are apparent in the scene are “baked-in” as the estimated spectral reflectance.

Some systems have also been proposed for spectral 3D data acquisition \[25,27,32,34,47,58\]. However, they rely on a dedicated setup using a multispectral camera \[32,47,58\] or a multispectral light source \[25,27,34\]. Also, cumbersome geometric calibration of the camera and the light source is often required \[32,34,47\]. These limitations make the system expensive and impractical for most users.

In this paper, we propose a novel projector–camera system, named Pro-Cam SSfM, for structure and spectral reflectance from motion. In Pro-Cam SSfM, we use a standard RGB camera and leverage an off-the-shelf projector for two roles: structured light and multispectral imaging. As shown in Fig. 2, for our data acquisition, structured light patterns (for geometric observations) and uniform color illuminations (for multispectral observations) are sequentially projected onto the object surface while alternately moving the camera and the projector to arbitrary positions around the object. Using the captured multi-view structured light data, we first reconstruct the dense 3D points while estimating the poses of all moved cameras and projectors in a self-calibration manner based on SfM using the tracked feature correspondences among all the projector and the camera viewpoints.

Using the captured multi-view multispectral data, we then estimate the spectral reflectance of each 3D point consid-

Fig. 2 The overview of Pro-Cam SSfM. a The projector projects a sequence of structured light patterns (gray code) and uniform color illuminations to acquire geometric and multispectral observations. b For image capturing, the camera and the projector are alternately moved around the object, so that the structured-light codes can be connected among all the camera and the projector positions. Then, the 3D points, the projector poses, and the camera poses are jointly reconstructed using the connected structured-light codes (i.e., feature points) in a self-calibration manner based on SfM. Finally, the spectral reflectance for each 3D point is estimated using the images captured with the uniform color illuminations while considering the geometric relationship between the estimated 3D points and projector positions.
ering the geometric relationship between the reconstructed 3D points and the estimated projector positions. Owing to the proposed self-calibrating 3D reconstruction, the reconstructed 3D points and the estimated projector positions can be used to construct a lighting model to eliminate the baked-in effect of the shading and the shadows from the estimated spectral reflectance without requiring any geometric calibration.

Technical contributions of this work are summarized as follows.

1. We propose a novel data acquisition procedure and self-calibrating 3D reconstruction method based on multi-view structured-light and SfM techniques. Our method provides dense 3D points as well as camera and projector poses, even for texture-less objects by tracking the structured light patterns for all camera and projector viewpoints. The estimated projector positions are further exploited for the following spectral reflectance estimation.

2. We propose a novel spectral reflectance estimation method by incorporating the geometric relationship between the reconstructed 3D points and the estimated projector positions into the cost optimization. Our method leads accurate estimation of the inherent spectral reflectance of each 3D point while eliminating the baked-in effect of the shading and the shadows.

3. By integrating the above two key techniques into one system, we propose Pro-Cam SSfM, a novel projector–camera system for practical and low-cost spectral 3D acquisition. We experimentally demonstrate that Pro-Cam SSfM can precisely reconstruct a dense object 3D model with the spectral reflectance property. To the best of our knowledge, Pro-Cam SSfM is the first spectral 3D acquisition system using off-the-shelf devices.

Preliminary versions of this paper appeared in [38] and [36]. This journal version unifies [38] and [36] and provides a more comprehensive and thorough description of our proposed Pro-Cam SSfM. More specifically, we expand on the description about the data acquisition (Sect. 3.1) and the pattern detecting and tracking with respect to self-calibrating 3D reconstruction (Sect. 3.2). Furthermore, we include additional simulation experiments using newly created synthetic ground-truth spectral 3D data to quantitatively evaluate the 3D shape quality and the spectral reflectance accuracy in comparison with existing methods (Sect. 4.2).

2 Related work

Structured light systems Structured light is a well-adopted technique to densely reconstruct the 3D points irrespective of surface textures by projecting structured light patterns [4,21,26,56], as summarized in detail in [10,24,45,51]. In the structured-light system, a common approach [16,17] to obtaining the whole shape of the object is to align the multiple scans (point clouds) from different viewpoints into a single global coordinate using the iterative closest point (ICP) method [8]. However, the ICP alignment is not trivial because it does not perform well when the object has repetitive structures or no distinctive 3D structures.

While most of structured-light methods are based on a pre-calibrated projector–camera system, some multi-view structured light methods [18,20] have realized self-calibrating reconstruction of the 3D points. The key of these methods is that the structured light patterns projected by one fixed projector is captured from more than two camera viewpoints having viewing angle overlaps, so that feature matching and tracking can be made to connect all projectors and cameras. This setup has been realized by simultaneously using multiple projectors and cameras [18,20].

Different from existing multi-view methods [18,20], our Pro-Cam SSfM realizes self-calibrating 3D reconstruction using only one camera and one projector by proposing a practical data acquisition manner. We also exploit the estimated projector positions for modeling the spectral reflectance, while the existing methods [18,20] only focus on the geometric reconstruction.

Multispectral imaging systems Existing hardware-based [6,9,12,23,44,48,57] or software-based [2,5,15,28,46,54] multispectral imaging systems commonly apply a single-viewpoint image-based spectral reflectance estimation method ignoring scene’s or object’s geometric information. This means that they only achieve scene-dependent spectral reflectance estimation, where the viewpoint- and scene-dependent shading or shadow is baked-in as the estimated spectral reflectance.

In Pro-Cam SSfM, we use an off-the-shelf RGB camera and projector for multispectral imaging. Although this setup is similar to the one used in [23], in which an digital light processing (DLP) projector and a high-speed camera are used to realize video-rate spectral image acquisition, we propose a novel spectral reflectance estimation method for recovering the object’s inherent spectral reflectance considering the geometric information that can be obtained at the self-calibrating 3D reconstruction step.

Spectral 3D acquisition systems Existing systems for spectral 3D acquisition are roughly classified into two categories: single-view and multi-view systems. Single-view systems are based on photometric stereo. Kitahara et al. proposed a system for simultaneous estimation of spectral reflectance and surface normal using a multispectral light stage [34]. Ozawa et al. proposed a single-shot multispectral photometric stereo system using a multispectral camera and spatially
and spectrally designed illumination \[47\]. Although these systems can acquire dense surface normals, they have two main limitations: (i) they assume that the light positions are calibrated and (ii) they only provides a spectral 3D model from a single viewpoint, and thus, the separated models at each viewpoint should be merged to generate a whole model.

Multi-view systems are based on SfM and MVS. Zia et al. proposed a 3D reconstruction method from multispectral images that effectively combines the reconstructed models at different wavelengths \[58\]. Ito et al. proposed an MVS system using the images obtained using a narrow-band multispectral light source \[27\]. These SfM- and MVS-based systems are difficult to reconstruct the 3D points on textureless surfaces. Kim et al. proposed a 3D imaging spectroscopy system impractical and expensive.

The closest system to Pro-Cam SSfM is the one by Hirai et al. \[25\]. They also used a projector–camera setup and applied the structured light to acquire the 3D points. However, their system has several limitations: First, their system is based on precise geometric calibration using a calibration target. Second, their system requires a multispectral projector to generate so-called basis light. Finally, their system is a single-view system and not designed to recover the spectral 3D data of a whole object. These limitations make their systems expensive.

Our Pro-Cam SSfM overcomes the limitations of existing systems because (i) it uses a low-cost off-the-shelf camera and projector, (ii) it does not require geometric calibration, and (iii) it generates a dense 3D model based on the structured light.

### 3 Proposed Pro-Cam SSfM

As shown in Fig. 2a, the projector is used to project a sequence of structured light patterns and uniform color illuminations to acquire geometric and photometric observations. As the structured light patterns, we use the binary gray code \[21,24\] because of its robustness. The proposed structured-light system generates \(\log_2 O\) bits of gray code to assign the independent value for each pixel where \(O\) is the number of projector rows (or columns). For instance, Table 1 shows the generated \(\log_2 8 = 3\) different strip codes for 8 columns.

| Gray codes | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
|------------|---|---|---|---|---|---|---|
| 0 0 1 1 1 1 0 0 |
| 0 1 1 0 0 1 1 0 |

To improve the accuracy of decoding, the gray code patterns insert the original and its inverse. For example, a sequence of gray code patterns consists of 42 frames for a projector with the resolution of \(1024 \times 768\) (including \(2 \times \lceil \log_2 1024 \rceil = 20\) patterns representing the columns, \(2 \times \lceil \log_2 768 \rceil = 20\) patterns representing the rows, and one pair of white and black images for identifying shadow regions). As the uniform color illuminations, we use the seven illuminations: red, green, blue, cyan, magenta, yellow, and white illuminations, which are generated using the binary combinations of the RGB primaries as \((R,G,B) = (1,0,0), (0,1,0), (0,0,1), (0,1,1), (1,0,1), (1,1,0), (1,1,1),\) respectively. The sequence of active projections is effectively exploited in the 3D reconstruction and the spectral reflectance estimation.

The data acquisition starts with initial projector and camera positions (e.g., position 1 in Fig. 2b). After capturing the data with the initial positions, the position of the projector is fixed and the camera is moved to another viewpoint to capture the patterns (e.g., camera position 2 in Fig. 2b). Next, the projector is moved to a different position (e.g., projector position 2 in Fig. 2b) while fixing the position of the camera, so that two sequences of patterns projected from different projector positions are captured from the same camera viewpoint. This connects the codes projected by different projector positions. The above procedures are repeated until the whole surface of the target object is scanned. This acquisition procedure enables to connect the structured light codes (i.e., feature points) between successive projectors and cameras. All connected feature points using all projector and camera positions are used as correspondences for the SfM pipeline, which enables self-calibrating reconstruction of the 3D points while estimating the poses of all moved projectors and cameras.

### 3.1 Data acquisition

Figure 2 illustrates the data acquisition procedure of Pro-Cam SSfM. An off-the-shelf projector is used as active illumination and a standard RGB camera is used as the imaging device to capture the object illuminated by the projector.
3.2 Self-calibrating 3D reconstruction

Given the structured-light encoded images, we first perform self-calibrating reconstruction of the 3D points while estimating the intrinsic parameters and the poses of all moved cameras and projectors.

3.2.1 Assigning gray codes to image pixels

By projecting gray code patterns as shown in Fig. 2a, the projector can add “features” on object surfaces. Those features have different codes whose number is the same as the projector resolution. The decoding algorithm in [24] is followed which outputs a list of \( \{ CODE_a, X_b \} \) of all camera viewpoints. \( CODE_a \) denotes the unique identifications (gray codes) in the pattern projected by projector position \( a \). \( X_b \in \mathbb{R}^{2 \times N_f} \) denotes the positions of extracted gray codes in the image captured by camera viewpoint \( b \), where \( N_f \) is the number of extracted gray codes. If the same gray code is assigned to multiple pixels, the average of pixel coordinates is computed as the final position in the image.

3.2.2 Matching image pixels using gray-code patterns

The next step is to match the pixels across projector and camera viewpoints using the codes. Figure 3 shows the examples of the feature correspondences and tracks. The feature correspondences for camera–projector pairs and camera–camera pairs at the common projector positions (e.g., camera1–projector1–camera2 or camera2–projector2–camera3 tracks in Fig. 3), which share the same projector code (\( CODE_1 \) or \( CODE_2 \)), are easily obtained by searching the same code.

To represent the camera poses and the 3D structures in a global coordinate system, it is necessary to find the projector rays intersecting at the same surface. This is achieved by connecting the projector patterns (gray codes) generated from different projector positions using the camera viewpoint that captures two sequences of projector patterns (e.g., camera2 for projector1 and projector2 in Fig. 3). The codes located at the same pixel position are linked. More specifically, for each code \( \{ CODE_1, X_2 \} \) projected by projector position 1, if the nearest code \( \{ CODE_2, X_2 \} \) of projector position 2 is less than 0.5 pixels (e.g., the yellow circles in Fig. 3), the codes of \( CODE_1 \) and \( CODE_2 \) are connected. In this case, the code positions of \( CODE_1 \) and \( CODE_2 \) are merged into one point (midpoint) and that point is used as the feature point for SfM. Notice that the corresponding points across different projector positions appear only in the overlapping area of projected patterns on the object surface.

3.2.3 3D point and projector–camera pose estimation

The set of all obtained correspondences is then fed into a standard SfM pipeline [52,55] to estimate the 3D points, the projector poses, and the camera poses. In the SfM pipeline, we modify the bundle adjustment formulation [40] by setting different weights to the reprojection errors on camera and projector views as

\[
E = \sum_i \sum_k w_i \|x_{k,i} - H_i(p_k)\|^2, \tag{1}
\]

where \( p_k \) is the 3D coordinate of point \( k \), \( x_{k,i} \) is the corresponding pixel coordinates in \( i \)-th viewpoint (camera or projector), and \( H_i(p) \) is a function that projects the 3D point to \( i \)-th viewpoint (camera or projector) using intrinsic and extrinsic parameters for each projector and each camera. Because the feature positions of cameras are detected from observed images, they may contain non-negligible errors. In contrast, the feature positions of projectors are more reliable, because they directly come from the projector patterns which have almost no errors. Therefore, in Eq. (1), we set a larger weight to impose higher penalties for the reprojection errors of the projector as

\[
w_i = \begin{cases} 
1, & \text{if viewpoint } i \text{ is a camera} \\
w_p, & \text{if viewpoint } i \text{ is a projector} 
\end{cases} \tag{2}
\]
3.3 Spectral reflectance estimation

Given the estimated 3D points, projector positions, and camera poses, we next estimate the spectral reflectance of each 3D point. For this purpose, we use multispectral images captured under the uniform color illuminations. In what follows, we first introduce the rendering model adopted in our method and then explain cost optimization to estimate the spectral reflectance using multi-view multispectral images.

3.3.1 Rendering model

We here introduce the rendering model for each 3D point using a single projector–camera pair, as illustrated in Fig. 4. Suppose the object surface is modeled by Lambertian reflectance and the camera response is linear, the camera’s pixel intensity for the spectral reflectance is expressed by eight basis functions. The shading factor is calculated from the geometric relationship between the projector position and the 3D point position.

\[
y_k = s_k C^T L r_k, \tag{4}
\]

where \( r_k \in \mathbb{R}^{N_b} \) represents the spectral reflectance, \( L = [L_1; \ldots; L_{N_t}] \in \mathbb{R}^{N_t \times N_b} \) is the illumination matrix, where \( L_n \in \mathbb{R}^{N_b \times N_b} \) is the n-th diagonal illumination matrix, and \( C^T = \text{BlockDiag}(C_{rgb}^T, \ldots, C_{rgb}^T) \in \mathbb{R}^{3 \times N_t \times N_b} \) is the block diagonal matrix, where \( C_{rgb} \in \mathbb{R}^{3 \times N_b} \) is the camera sensitivity matrix. In this work, we assume that the spectral power distributions of the projected illuminations and the camera sensitivity (i.e., \( C^T \) and \( L \)) are known or preliminarily estimated (e.g., by [29,57]).

3.3.2 Spectral reflectance model

It is known that the spectral reflectance of natural objects is well represented by a small number of basis functions [49]. Based on this observation, we adopt a widely used basis model [23,48,57], where the spectral reflectance is modeled as

\[
r_k = B \alpha_k, \tag{5}
\]

where \( B \in \mathbb{R}^{N_b \times N_b} \) is the basis matrix, where \( N_b \) is the number of basis functions, and \( \alpha_k \in \mathbb{R}^{N_b} \) is the coefficient vector. The basis model can reduce the number of parameters (since \( N_b < N_k \)) for spectral reflectance estimation. Using the basis model, Eq. (4) is rewritten as

\[
y_k = s_k C^T L \alpha_k. \tag{6}
\]
3.3.3 Shading model

Different from common single-view image-based methods (e.g., [23,25,48,57]), we take the shading factor into account for spectral reflectance estimation, which results in more accurate model and estimation. Figure 4 illustrates the relationship between the projector position \( p_{\text{pro}} \), \( k \)-th 3D point \( p_k \), and the point normal \( n_k \), where the point normal is obtained by averaging the normals of adjacent meshes reconstructed by a surface reconstruction method (e.g., Screened Poisson Surface Reconstruction [30] for our experiments). We define the shading factor as

\[
s_k = \frac{I_k}{l},
\]

(7)

where \( I_k \) is the irradiance of the incident light at \( k \)-th point and \( l \) is the intensity of the emitted light from the projector. For the above calculation, we rewrite as \( l(\lambda) = l \), since the shading factor is wavelength independent and defined only from the geometric relationship between the projector position and the 3D point. The irradiance \( I_k \) is represented as

\[
I_k = \frac{l}{\|p_{\text{pro}} - p_k\|^2} \times \frac{p_{\text{pro}} - p_k}{\|p_{\text{pro}} - p_k\|} \cdot n_k,
\]

(8)

where the first term represents the light intensity at the 3D point, which is inversely proportional to the square of the distance from the projector to the 3D point, and the second term represents the inner product of the normalized lighting vector and the point normal (see Fig. 4).

Assuming that the ambient light is negligible and omit interreflection from the model, the shading factor is calculated from Eqs. (7) and (8) as

\[
s_k = \frac{p_{\text{pro}} - p_k}{\|p_{\text{pro}} - p_k\|^2} \cdot n_k.
\]

(9)

Based on the defined model, the shading factor can be calculated from the 3D point, the point normal, and the projector position that we have already obtained by the self-calibrating 3D reconstruction. The final rendering model is derived by substituting Eq. (9) into Eq. (6).

3.3.4 3D mesh reconstruction and visibility calculation

In our method, the object surface is used to calculate the visibility of each 3D point, which is required to estimate the spectral reflectance using multi-view input images. Also, the surface representation is useful for some rendering applications such as spectral relighting. Thus, we reconstruct the object surface from the obtained point cloud (we used Screened Poisson Surface Reconstruction [30] for this purpose because of its easy availability in Meshlab [13]).

After 3D meshes are reconstructed by the surface reconstruction, the visibility of each mesh vertex from each camera and projector is checked using the estimated camera and projector poses. If the vertex \( k \) is reprojected into the view frustum of a camera (or a projector) and its sight ray is not occluded by any other mesh triangle, this vertex is considered as visible from this camera (or projector). For each camera–projector pair, the vertexes that are invisible from either the camera or the projector are not considered in the spectral reflectance estimation. Especially, the invisible vertexes for the projector are regard as cast shadow areas, so that we can discount the effects of the cast shadows for the spectral reflectance estimation.

3.3.5 Cost optimization

Using the rendering model of Eq. (6), we solve an optimization problem to estimate the spectral reflectance of each 3D point from multi-view images obtained from all projector–camera pairs. The cost function is defined as

\[
\text{arg min}_{\alpha_k} E_{\text{ren}}(\alpha_k) + \gamma E_{\text{ssm}}(\alpha_k),
\]

(10)

where \( E_{\text{ren}} \) is the rendering term and expressed as

\[
E_{\text{ren}}(\alpha_k) = \sum_{c \in \mathcal{V}(k)} \frac{\|y_{k,c}^{\text{obs}} - y_{k,c}(\alpha_k)\|^2}{|\mathcal{V}(k)|},
\]

(11)

where \( y_{k,c}^{\text{obs}} \in \mathbb{R}^{3N_c} \) is the observed multispectral intensity vector obtained from \( c \)-th projector–camera pair, \( y_{k,c}(\alpha_k) \) is the estimated intensity vector based on the rendering model, and \( \mathcal{V}(k) \) is the visible set for \( k \)-th point. This term evaluates the data fidelity between the observed and the rendered intensities. \( E_{\text{ssm}} \) is a commonly used spectral smoothness term [23,48,57], which is defined by

\[
E_{\text{ssm}}(\alpha_k) = D B \alpha_k,
\]

(12)

where \( D \in \mathbb{R}^{N_s \times N_c} \) is the operation matrix to calculate the second-order derivative [48]. This term evaluates the smoothness of the estimated spectral reflectance. The balance of \( E_{\text{ren}} \) and \( E_{\text{ssm}} \) is determined by the parameter \( \gamma \).

3.4 Setup details

We used an ASUS P3B projector and a Canon EOS 5D Mark-II digital camera. The sequence of the structured light patterns was captured using a video format with 1920×1080 resolution, while the color illuminations were captured using a RAW format, which has a linear camera response, with a higher resolution. The RAW images were then resized to have the same resolution with the video format. As shown...
in Fig. 5, the camera spectral sensitivity of Canon EOS 5D Mark-II was obtained from the camera sensitivity database [29] and the spectral power distributions of the color illuminations were measured using a StellarNet BlueWave-VIS Spectrometer.

For the 3D reconstruction, we used Colmap [52] to run the SFM pipeline and Screened Poisson surface reconstruction [30], which is integrated with Meshlab [13], to visualize the obtained 3D model. For the spectral reflectance estimation, we set the target wavelength range as 410 nm to 670 nm with every 10nm intervals because the used projector illuminations only have the spectral power within this range. The spectral basis functions were calculated using the spectral reflectance data of 1269 Munsell color chips [49] by principal component analysis. We used eight basis functions based on the observation that they are able to represent more than 99% of the total variance of common reflectance data [57].

The spectral smoothness weight in Eq. (10) was determined by an empirical manner and set as $\gamma = 0.06$ for the intensity range [0,1]. The C++ Ceres solver [3] was used to solve the nonlinear optimization problem of Eq. (10).

### 4 Experimental results

#### 4.1 Evaluation of spectral reflectance estimation using colorchart

We firstly evaluate the performance of our spectral reflectance estimation method using a standard colorchart with the 24 patches. We first show the effect of the number of spectral bands. In our experiment, seven color illuminations and RGB camera channels, as shown in Fig. 5, were used, resulting in a total of 21-band measurements. To select the best band set, we evaluated all possible band sets for each number of spectral bands. Figure 6 shows RMSE for the 24 patches of the colorchart when using the selected best band set for each number of spectral bands. We can observe that RMSE is reduced by using multispectral information and becomes very close when more than six bands are used. The six-band set of (light, camera) = ($L$green, $C$blue), ($L$blue, $C$green), ($L$blue, $C$blue), ($L$cyan, $C$red), ($L$magenta, $C$red), ($L$yellow, $C$green) provides the minimum RMSE among the evaluated all possible band sets. Even though six spectral bands are fewer than the dimensionality of the spectral basis functions ($N_b = 8$ in our experiments), our optimization problem is well-posed because we regularize the optimization problem through the spectral smoothness term $E_{ssm}$.

We next demonstrate the effectiveness of our spectral reflectance estimation method considering the geometric information. As shown in Fig. 7a, we laid the colorchart on a table and captured the structured light and multispectral data by four projector–camera pairs according to the data acquisition procedure of Pro-Cam SSfM. The estimated projector positions, camera positions, and 3D points of the colorchart are shown in Fig. 7b. The example captured images (under white illumination) by using all projector–camera pairs are shown in Fig. 7c. Figure 7d compares the estimated spectral reflectance results for the 24 patches, where the blue line is the ground truth, the red line is our result (average within each patch) using all projector–camera pairs, and the yellow and purple dashed lines are the results of two existing single-view image-based methods [5,23] (average within each patch) only using the projector–camera pair 4. Figure 7e shows the corresponding RMSE comparison for each patch, where we can confirm that our method achieves much lower RMSE than the existing methods.

As can be seen in Fig. 7d and e, the single-view methods fail to correctly estimate the spectral reflectance including the relative scales between the patches. For example, the results of patches 4 and 8 have lower scales, whereas the results of patches 21–23 have higher scales, compared to ground-truth spectral reflectances. This is due to the shading effect appeared in the colorchart setup, as shown in Fig. 7c. Pair 4, which was used as the input for the single-view methods, where we can observe that patches 4 and 8 are under deeper shading than patches 21–23. In contrast, our method can provide accurate estimation results with correct relative scales. The benefit of our method is to estimate the spectral reflectance while considering the shading effect, which is
ignored in the single-view methods. With this essential difference, our method is especially beneficial when the shading exists in the scene. If the shading does not exist, the accuracy of our method could be similar to that of the existing methods. However, such no-shading condition is very special and possible only under fully controlled illumination.

4.2 Evaluation using synthetic images

4.2.1 Synthetic dataset creation

We next quantitatively evaluate the 3D shape quality and the spectral reflectance accuracy. For this purpose, we performed simulation experiments using two common CG models (Armadillo and Dragon) obtained from Stanford 3D Scanning Repository [22]. Using the CG model, we created a ground-truth spectral-3D model, which has ground-truth spectral reflectance data for every 3D point. As ground-truth spectral reflectance data, we used the spectral reflectance data of the colorchart’s 24 patches. According to the texture as shown in the left column of Fig. 8, one of the 24 patch’s spectral reflectance data was assigned to each 3D point. Using the created ground-truth spectral-3D model, multi-view input images were generated using 18 camera and projector positions, as shown in the middle column of Fig. 8. For each camera viewpoint, two sequences of images under structured light patterns and uniform color illuminations were rendered according to our image acquisition setup described in Sect. 3.4. Examples of the rendered multi-view images are shown in the right column of Fig. 8.

4.2.2 3D reconstruction results

We first evaluate the 3D reconstruction results. We compared our method with a standard SfM + MVS pipeline using Colmap [52] and an existing structured-light method with ICP alignment (SL [26] + ICP [8]).

For the SfM + MVS pipeline, we rendered another input image set under a passive light (an ambient light) condition, as the SfM + MVS pipeline assumes general images as inputs.
The viewpoints for the rendering follow the setup of Pro-Cam SSfM shown in Fig. 8, where we used all 36 viewpoints, including the projector viewpoints, as camera viewpoints to guarantee that the SfM + MVS pipeline has the same number of viewpoints as Pro-Cam SSfM.

As an existing structured-light approach, we compared a standard procedure that combines a calibrated structured-light method and the ICP alignment (SL [26] + ICP [8]), where the ICP alignment is performed to integrate multiple point clouds generated from each projector–camera viewpoint. Specifically, we performed the following steps: (i) Point clouds for each camera–projector viewpoint are reconstructed by conducting the triangulation using detected code patterns [26], where the ground-truth intrinsic and extrinsic parameters of the projector–camera pair are given for the SL + ICP method, though Pro-Cam SSfM estimates them in the proposed self-calibrating manner. (ii) Then, a global registration is performed using extract characteristic points from the point clouds using fast point feature histograms [50]. (iii) Then, the ICP alignment [8] is conducted to further refine the registration. (iv) Finally, to compare the registered point cloud and the ground truth under a unified coordinate system, we align the position of one of the camera–projector pairs with the ground-truth position, so that the entire coordinate system is aligned with the ground truth.

To evaluate the 3D shape quality, we used two common metrics, i.e., accuracy and completeness, which are used in some 3D reconstruction benchmark papers [1,35]. The accuracy is the distance from each estimated 3D point to its nearest ground-truth 3D point. The completeness is the distance from each ground-truth 3D point to its nearest estimated 3D point.

Figure 9 shows the visual comparison of the reconstructed 3D point clouds and their accuracy (distance to the closest ground truth point) for Armadillo (a) and Dragon (b). Compared to the SfM + MVS pipeline which uses a local feature extraction algorithm based on a passive illumination, our method reconstructs more complete point clouds than the SfM + MVS pipeline because the structured light can provide more abundant feature points. Also, the structured light results in more precise correspondences, so that our method achieves better accuracy than the SfM + MVS pipeline. As the SL + ICP method, ICP may lead to accumulated errors when superimposing the multiple point clouds sequentially. We can observe that there are many error points generated around the surface. In contrast, our method estimates the 3D points by globally optimizing all the camera and the projector poses with bundle adjustment, which leads to more accurate point clouds.

Figure 10 shows the comparison of the average accuracy and the average completeness errors for the two models and their visualization for every point by color map error representation. We can see that our method achieves the best accuracy and completeness for the both models, even though the SL + ICP method provides the largest number of the reconstructed points.

One improvement of our proposed method is achieved by introducing the weight of Eq. (1) for bundle adjustment, which poses larger penalties to the projector’s reprojection errors. Figure 11 shows the effect of changing the projector weight $w_p$ on quantitative evaluation for Armadillo model. We can confirm that a larger weight leads to a better result, though the value exceeding 100 does not make a significant difference. Thus, we set as $w_p = 100$ for all the experiments in this paper.
4.2.3 Spectral reflectance estimation results

We next evaluate the spectral reflectance estimation accuracy for the CG models. Since there is no comparable method and public code for the spectral 3D reconstruction, to the best of our knowledge, we compared our multi-view-based spectral reflectance estimation method with an existing single-image-based spectral reflectance estimation method [23]. In the compared method, we applied the image-based spectral reflectance estimation method [23] to every camera viewpoint using the same six-band images as the ones we have selected in Section 4.2. Then, we calculated the spectral reflectance of each 3D point by averaging the spectral reflectance results estimated at all the pixels corresponding to the 3D point projection to all visible cameras. Here, we employed ground-truth 3D points and camera poses for discounting the effect of geometric 3D reconstruction errors.
from the evaluation of the spectral reflectance estimation methods.

The top row of Fig. 12 shows the visual comparison of sRGB, which was converted from the estimated spectral reflectances. We can observe that the effects of shading and shadows remain in the results of the image-based spectral reflectance estimation method (the first column for each CG model), especially around the complicated surfaces such as the hands of Armadillo and the head of Dragon. In contrast, our sRGB result (the second column for each CG model) is close to the ground truth and represents the object’s inherent property less affected by the shading and the shadows by considering them in the spectral reflectance estimation.

The bottom row of Fig. 12 shows the error map for estimated spectral reflectances, where RMSE for all wavelengths is visualized for each 3D point. We can confirm that our method achieves lower RMSE errors compared with the existing method, especially around the complicated surfaces where the shading and shadows are likely to appear. The average result of all wavelengths and all points are shown under the error map. From those results, we can confirm that our method generally outperforms the existing method. The third column for each CG model shows the results by using the estimated 3D points by our method instead of using the ground-truth 3D points. From the comparison of the second and the third columns, we can confirm that, for our spectral reflectance estimation method, there is only small differences on the spectral reflectance accuracy even with the estimated 3D points. This indicates that Pro-Cam SSfM can provide the 3D points accurate enough for the spectral reflectance estimation.

4.3 Evaluation using real images

We next evaluate our method for real images. Figure 13 shows the spectral 3D acquisition results of a clay sculpture with roughly 30 centimeter height. To schematically show the layout of the moved projector and camera positions, we show a synthesized top-view image of Fig. 13a. The estimated projected and camera positions by our method are overlaid as the red and the green triangular pyramids using a manually aligned scale. It is demonstrated that the projector and the camera positions are correctly estimated by our method. Figure 13b shows the final reconstructed surface result for our method, where the detail structures of the sculpture are precisely reconstructed.

Figure 13c shows the spectral reflectance results for some 3D points. It is demonstrated that our method can accurately estimate the spectral reflectance compared with the ground truth measured by a StellarNet BlueWave-VIS Spectrometer. Figure 13c–e compares the sRGB results converted from the estimated spectral reflectances by our method and the single-view method [23]. We can observe that our spectral reflectance estimation method considering the geometric information can effectively remove the baked-in effect of the shading and the shadow, which is apparent in the sRGB result of the single-view method.

Since Pro-Cam SSfM can accurately estimate both the 3D points and the spectral reflectance, it is possible to perform the spectral 3D relighting of the object for synthesizing the appearance illuminated under an arbitrary light orientation.

![Fig. 11](image)

Quantitative evaluation for Armadillo as a function of $w_p$.

![Fig. 12](image)

Spectral reflectance results for synthetic models. Top: the sRGB results converted from the estimated spectral reflectance results; Bottom: visualized spectral reflectance error (RMSE for all wavelengths) for each 3D point, and the average of all points and all wavelengths.
and spectral distribution. Figure 13f shows the result of spectral relighting under the projector-cyan illumination, where we can confirm that our relighting result is close to the reference actual image taken in the same illumination orientation and spectral distribution. The differences at the concave skirt regions are due to the effect of interreflections, which is not considered in our current model. Figure 13g shows the more complex relighting result under two mixed light sources (projector cyan and halogen lamp), which are located at different sides of the object. Figure 13h shows the 3D relighting results under different illumination orientations. As shown in those results, we can effectively perform the spectral 3D relighting based on the estimated 3D model and spectral reflectance.

Figure 14a–c shows the spectral reflectance and the sRGB results, the 3D shape results, and the spectral patterns at each wavelength for a stuffed toy. Those results demonstrate the potential of Pro-Cam SSfM for accurate spectral 3D scanning and rendering. Figure 15 shows the results of the spectral 3D acquisition on a Bizen ware, which is a Japanese traditional craft. It can be confirmed that the detailed surface of Bizen ware can be recovered. In recent years, the distribution of forgery of art and traditional crafts has become a major issue. By creating a database of detailed 3D shapes and spectral reflectance information of valuable arts and traditional crafts obtained by the proposed system, it is expected to be used for digital archives and authenticity judgment.

Additional spectral 3D acquisition results by our proposed system can be seen in the supplemental video or our project page (www.ok.sc.e.titech.ac.jp/res/PCSSfM/).

5 Concluding remarks

In this paper, we have proposed Pro-Cam SSfM, the first spectral 3D acquisition system using an off-the-shelf pro-
jector and camera. By effectively exploiting the projector as active lighting for both the geometric and the photometric observations, Pro-Cam SSfM can accurately reconstruct a dense object 3D model with the spectral reflectance for each 3D point. We first have proposed a practical data acquisition manner and 3D reconstruction method that enables self-calibrating 3D reconstruction using only one camera and one projector. We then have proposed a novel spectral reflectance estimation method by incorporating the geometric relationship between the estimated 3D points and projector positions into the cost optimization.

Both of the proposed components have been validated by the experiments using the synthetic images, which have demonstrated that our Pro-Cam SSfM can precisely reconstruct the 3D points and estimate the spectral reflectance by eliminating the effects of the shading and the shadows. We also have demonstrated the potential of Pro-Cam SSfM through the spectral 3D acquisition results on several real objects. The reconstructed model by Pro-Cam SSfM can be further used as an initial model for joint geometric and photometric refinement, as we have reported in our parallel study of [37].

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00371-022-02434-0.

Acknowledgements This work was partly supported by JSPS KAKENHI Grant Number 17H00744 and 21K17762.

Declarations

Conflict of interest The authors declared that they have no conflicts of interest to this work.

References

1. Aanæs, H., Jensen, R.R., Vogiatzis, G., Tola, E., Dahl, A.B.: Large-scale data for multiple-view stereopsis. Int. J. Comput. Vis. 120(2), 153–168 (2016)
2. Aeschbacher, J., Wu, J., Timofte, R.: In defense of shallow learned spectral reconstruction from RGB images. In: Proc. of IEEE Int. Conf. on Computer Vision (ICCV) Workshops pp. 471–479 (2017)
3. Agarwal, S., Mierle, K.: Ceres solver. http://ceres-solver.org
4. Aliaga, D.G., Xu, Y.: Photogeometric structured light: a self-calibrating and multi-viewpoint framework for accurate 3D modeling. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) pp. 1–8 (2008)
5. Arad, B., Ben-Shahar, O.: Sparse recovery of hyperspectral signal from natural RGB images. In: Proc. of European Conf. on Computer Vision (ECCV) pp. 19–34 (2016)
6. Baek, S.H., Kim, I., Gutierrez, D., Kim, M.H.: Compact single-shot hyperspectral imaging using a prism. ACM Trans. Graph. 36(6), 1–12 (2017)
7. Behmann, J., Mahlein, A.K., Paulus, S., Dupuis, J., Kuhlmann, H., Oerke, E.C., Plümer, L.: Generation and application of hyperspectral 3D plant models: methods and challenges. Mach. Vis. Appl. 27(5), 611–624 (2016)
8. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14(2), 239–256 (1992)
9. Cao, X., Du, H., Tong, X., Dai, Q., Lin, S.: A prism-mask system for multispectral video acquisition. IEEE Trans. Pattern Anal. Mach. Intell. 33(12), 2423–2435 (2011)
10. Caspi, D., Kiryati, N., Shamir, J.: Range imaging with adaptive color structured light. IEEE Trans. Pattern Anal. Mach. Intell. 20(5), 470–480 (1998)
11. Chane, C.S., Mansouri, A., Marzani, F., Boosch, F.: Integration of 3D and multispectral data for cultural heritage applications: survey and perspectives. Image Vis. Comput. 31(1), 91–102 (2013)
12. Chi, C., Yoo, H., Ben-Ezra, M.: Multi-spectral imaging by optimized wide band illumination. Int. J. Comput. Vis. 86, 140–151 (2010)
13. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G.: Meshlab: an open-source mesh processing tool. In: Proc. of Eurographics Italian Chapter Conference pp. 129–136 (2008)
14. Devlin, K., Chalmers, A., Wilkie, A., Purgathofer, W.: Tone reproduction and physically based spectral rendering. Eurographics State of the Art Report pp. 1–23 (2002)
15. Fu, Y., Zheng, Y., Zhang, L., Huang, H.: Spectral reflectance recovery from a single RGB image. IEEE Trans. Comput. Imaging 4(3), 382–394 (2018)
16. Furukawa, R., Inose, K., Kawasaki, H.: Multi-view reconstruction for projector camera systems based on bundle adjustment. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops pp. 69–76 (2009)
17. Furukawa, R., Nagamatsu, G., Kawasaki, H.: Simultaneous shape registration and active stereo shape reconstruction using modified bundle adjustment. In: Int. Conf. on 3D Vision (3DV), pp. 453–462 (2019)
18. Furukawa, R., Sagawa, R., Kawasaki, H., Sakashita, K., Yagi, Y., Asada, N.: One-shot entire shape acquisition method using multiple projectors and cameras. Proc. of Pacific-Rim Symposium on Image and Video Technology (PSIVT) pp. 107–114 (2010)
19. Furukawa, Y., Ponce, J.: Accurate, dense, and robust multi-view stereopsis. IEEE Trans. Pattern Anal. Mach. Intell. 32(8), 1362–1376 (2010)
20. Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F.J., Marín-Jiménez, M.J.: Simultaneous reconstruction and calibration for multi-view structured light scanning. J. Vis. Commun. Image Represent. 39, 120–131 (2016)
21. Geng, J.: Structured-light 3D surface imaging: a tutorial. Adv. Opt. Photonics 3(2), 128–160 (2011)
22. Greg, T., Marc, L.: The stanford 3d scanning repository. http://graphics.stanford.edu/data/3Dscanrep/
23. Han, S., Sato, I., Okabe, T., Sato, Y.: Fast spectral reflectance recovery using DLP projector. Int. J. Comput. Vis. 110(2), 172–184 (2014)

24. Herakleous, K., Poullis, C.: 3DUNDERWORLD-SLS: an open-source structured-light scanning system for rapid geometry acquisition. CoRR abs/1406.6595 (2014)

25. Hirai, K., Nakahata, R., Horiiuchi, T.: Measuring spectral reflectance and 3D shape using multi-primary image projector. In: Proc. of Int. Conf. on Image and Signal Processing (ICISP) pp. 137–147 (2016)

26. Inokuchi, S.: Range imaging system for 3-D object recognition. ICPR 1984, 806–808 (1984)

27. Ito, S., Ito, K., Aoki, T., Tsuchida, M.: A 3D reconstruction method with color reproduction from multi-band and multi-view images. In: Proc. of Asian Conf. on Computer Vision (ACCV) pp. 236–247 (2016)

28. Jia, Y., Zheng, Y., Gu, L., Subpa-Asa, A., Lam, A., Sato, Y., Sato, I.: From RGB to spectrum for natural scenes via manifold-based mapping. In: Proc. of IEEE Int. Conf. on Computer Vision (ICCV) pp. 4705–4713 (2017)

29. Jiang, J., Liu, D., Gu, J., Süsstrunk, S.: What is the space of spectral sensitivity functions for digital color cameras?. In: Proc. of Workshop on Applications of Computer Vision (WACV) pp. 168–179 (2013)

30. Kazhdan, M., Hoppe, H.: Screened Poisson surface reconstruction. ACM Trans. Graph. 32(3), 1–13 (2013)

31. Kim, K., Torii, A., Okutomi, M.: Multi-view inverse rendering under arbitrary illumination and albedo. In: Proc. of European Conf. on Computer Vision (ECCV) pp. 750–767 (2016)

32. Kim, M.H., Harvey, T.A., Kittle, D.S., Rushmeier, H., Dorsey, J., Prum, R.O., Brady, D.J.: 3D imaging spectroscopy for measuring hyperspectral patterns on solid objects. ACM Trans. Graph. 31(4), 1–11 (2012)

33. Kim, M.H., Rushmeier, H., Ffrench, J., Passeri, I., Tidmarsh, D.: Hyper 3D: 3D graphics software for examining cultural artifacts. ACM J. Comput. Cult. Herit. 7(3), 1–19 (2014)

34. Kitahara, M., Okabe, T., Fuchs, C., Lensch, H.P.A.: Simultaneous estimation of spectral reflectance and normal from a small number of images. In: Proc. of Int. Conf. on Computer Vision Theory and Applications (VISAPP) pp. 303–315 (2015)

35. Ley, A., Hänsch, R., Hellwich, O.: Sybr3: a realistic synthetic benchmark for 3D reconstruction from images. In: Proc. of European Conf. on Computer Vision (ECCV), pp. 236–251 (2016)

36. Li, C., Monno, Y., Hidaka, H., Okutomi, M.: Pro-cam ssfm: projector-camera system for structure and spectral reflectance from motion. In: Proc. of the IEEE/CVF Int. Conf. on Computer Vision (ICCV), pp. 2414–2423 (2019)

37. Li, C., Monno, Y., Okutomi, M.: Spectral MVIR: joint reconstruction of 3D shape and spectral reflectance. In: IEEE Int. Conf. on Computational Photography (ICCP), pp. 1–12 (2021)

38. Li, C., Torii, A., Okutomi, M.: Robust, precise, and calibration-free shape acquisition with an off-the-shelf camera and projector. In: Proc. of IEEE Conf. on Consumer Electronics (ICCE) pp. 1–6 (2018)

39. Liang, J., Zia, A., Zhou, J., Sirault, X.: 3D plant modelling via hyperspectral imaging. In: Proc. of IEEE Int. Conf. on Computer Vision Workshops (ICCVW) pp. 172–177 (2013)

40. Lourakis, M., Argyros, A.A.: SBA: a software package for generic sparse bundle adjustment. ACM Trans. Math. Softw. 36(1), 1–30 (2009)

41. Mansouri, A., Lathuiliere, A., Marzani, F., Voisin, Y., Gouton, P.: Toward a 3D multispectral scanner: an application to multimedia. IEEE Multimed. 14(1), 40–47 (2007)

42. Maurer, D., Ju, Y.C., Breuß, M., Bruhn, A.: Combining shape from shading and stereo: a variational approach for the joint estimation of depth, illumination and albedo. In: Proc. of British Machine Vision Conference (BMVC) pp. 76–1–14 (2016)

43. Mérou, J., Quéau, Y., Durou, J., Castan, F., Cremers, D.: Variational reflectance estimation from multi-view images. J. Math. Imaging Vis. 60(9), 1527–1546 (2018)

44. Monno, Y., Kikuchi, S., Tanaka, M., Okutomi, M.: A practical one-shot multispectral imaging system using a single image sensor. IEEE Trans. Image Process. 24(10), 3048–3059 (2015)

45. Nguyen, N., Nguyen, D., Wang, Z., Kieu, H., Le, M.: Real-time, high-accuracy 3D imaging and shape measurement. Appl. Opt. 54(1), A9–A17 (2015)

46. Nguyen, R.H.M., Prasad, D.K., Brown, M.S.: Training-based spectral reconstruction from a single RGB image. In: Proc. of European Conf. on Computer Vision (ECCV) pp. 186–201 (2014)

47. Ozawa, K., Sato, I., Yamaguchi, M.: Hyperspectral photometric stereo for a single capture. J. Opt. Soc. Am. A 34(3), 384–394 (2018)

48. Park, J.I., Lee, M.H., Grossberg, M.D., Nayar, S.K.: Multispectral reconstruction from a single RGB image. In: Proc. of European Conf. on Computer Vision (ICCV) pp. 1–8 (2007)

49. Parkkinen, J.P., Hallikainen, J., Jaaskelainen, T.: Characteristic spectra of munsell colors. JOSA A 6(2), 318–322 (1989)

50. Rusu, R.B., Blodow, N., Beetz, M.: Fast point feature histograms (FPFH) for 3D registration. In: Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 3212–3217 (2009)

51. Salvi, J., Fernandez, S., Pribanic, T., Lldado, X.: A state of the art in structured light patterns for surface profilometry. Pattern Recognit. 43(8), 2666–2680 (2010)

52. Schonberger, J.L., Frahm, J.M.: Structure-from-motion revisited. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) pp. 4104–4113 (2016)

53. Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) pp. 519–528 (2006)

54. Shi, Z., Chen, C., Xiong, Z., Liu, D., Wu, F.: HSCNN+: Advanced CNN-based hyperspectral recovery from RGB images. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops pp. 939–947 (2018)

55. Snavely, N., Curless, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3D. ACM Trans. Graph. 25(3), 835–846 (2006)

56. Weinmann, M., Schwartz, C., Ruiters, R., Klein, R.: A multicamera, multi-projector super-resolution framework for structured light. In: Proc. of Int. Conf. on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT) pp. 397–404 (2011)

57. Wug Oh, S., Brown, M.S., Pollefeys, M., Joo Kim, S.: Do it yourself hyperspectral imaging with everyday digital cameras. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) pp. 2461–2469 (2016)

58. Zia, A., Liang, J., Zhou, J., Gao, Y.: 3D reconstruction from hyperspectral images. In: Proc. of IEEE Winter Conf. on Applications of Computer Vision (WACV) pp. 318–325 (2015)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Chunyu Li received the Bachelor of Engineering degree from the Software Institute, Nanjing University, in 2016, and the Master of Engineering degree from the Department of Systems and Control Engineering, Tokyo Institute of Technology, in 2019. Currently he is a PhD candidate in the Department of Systems and Control Engineering, Tokyo Institute of Technology. His research interests include 3D reconstruction (structured light, structure from motion) and hyperspectral imaging.

Yusuke Monno received the B.E., M.E., and Ph.D degrees from Tokyo Institute of Technology, Tokyo, Japan, in 2010, 2011, and 2014, respectively. From November 2013 to March 2014, he joined the Image and Visual Representation Group at École Polytechnique Fédérale de Lausanne as a research internship student. He is currently an assistant professor with the Department of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology. His research interests are in both theoretical and practical aspects of image processing, computer vision, and biomedical engineering.

Masatoshi Okutomi received the B.Eng. degree from the Department of Mathematical Engineering and Information Physics, the University of Tokyo, Tokyo, Japan, in 1981, and the M.Eng. degree from the Department of Control Engineering, Tokyo Institute of Technology, Tokyo, in 1983. He joined the Canon Research Center, Canon Inc., Tokyo, in 1983. He received the Dr.Eng. degree from Tokyo Institute of Technology, in 1993, for his research on stereo vision. Since 1994, he has been with Tokyo Institute of Technology, where he is currently a Professor with the Department of Systems and Control Engineering, the School of Engineering.