


Shape and Reflectance Reconstruction Using Concentric Multi-Spectral Light Field

Mingyuan Zhou, Yuqi Ding, Yu Ji, S. Susan Young, *Senior Member, IEEE*, Jingyi Yu, and Jinwei Ye 

Abstract—Recovering the shape and reflectance of non-Lambertian surfaces remains a challenging problem in computer vision since the view-dependent appearance invalidates traditional photo-consistency constraint. In this paper, we introduce a novel concentric multi-spectral light field (CMSLF) design that is able to recover the shape and reflectance of surfaces of various materials in one shot. Our CMSLF system consists of an array of cameras arranged on concentric circles where each ring captures a specific spectrum. Coupled with a multi-spectral ring light, we are able to sample viewpoint and lighting variations in a single shot via spectral multiplexing. We further show that our concentric camera and light source setting results in a unique single-peak pattern in specularity variations across viewpoints. This property enables robust depth estimation for specular points. To estimate depth and multi-spectral reflectance map, we formulate a physics-based reflectance model for the CMSLF under the surface camera (S-Cam) representation. Extensive synthetic and real experiments show that our method outperforms the state-of-the-art shape reconstruction methods, especially for non-Lambertian surfaces.

Index Terms—Shape reconstruction, surface reflectance, multi-spectral, light field

1 INTRODUCTION

SURFACE shape and reflectance reconstruction from images is a fundamental problem in computer vision that can benefit numerous applications ranging from graphics rendering to scene understanding. Well-established solutions based on multi-view stereo [1], [2], [3] or photometric stereo [4], [5], [6], [7] often assume Lambertian surfaces, from which the light is equally reflected towards all directions. However, most real world objects have more complex reflectances that exhibit view-dependent characteristics (such as specular highlights). These surface violate the Lambertian assumption and result in erroneous depth and reflectance estimation.

In recent years, light field has emerged as a powerful tool in computer vision and graphics for 3D-related applications. A light field camera can be essentially viewed as a multi-view device. Notable examples include the hand-held light field camera [8] and the light field camera array [9]: the former combines a lenticular lens array and a single high-resolution sensor with each lenslet emulating a pinhole

camera, while the latter uses multiple cameras in order to allow for wide baseline and large Field-of-View (FoV) acquisition. In both settings, the viewpoints are arranged on rectangular grids. Earlier uses of light field focused on refocused rendering [8], [10] and view interpolation [11]. More recent approaches have employed light field for 3D reconstruction [12], [13], [14]. To handle non-Lambertian reflectance, focus cues [15], angular coherence [16] and BRDF-invariants [17], [18] are proposed on light field data. However, additional surface priors such as smoothness or polynomial shape need to be imposed.

In this paper, we introduce a novel concentric multi-spectral light field (CMSLF) array that arranges the cameras on concentric circles. As shown in Fig. 1, each circle of cameras captures images at the same specific spectrum. We show that by coupling with a multi-spectral ring light, our CMSLF is advantageous in surface shape and reflectance reconstruction. This is mainly because 1) through spectral multiplexing, we are able to simultaneously sample multiple viewpoints under varying lighting directions without interference; and 2) our concentric viewpoint arrangement results in a unique single-peak pattern in specularity variations across views that can be utilized for specularity analysis and thus enables robust depth estimation for non-Lambertian points.

We develop tailored algorithms based on our CMSLF design for surface shape and reflectance estimation. Specifically, we first formulate a dichromatic Phong reflectance model under CMSLF using the surface camera (S-Cam) representation [19]. An S-Cam models angular reflectance distribution with respect to a 3D scene point. It can be formed by tracing rays originated from the scene point back to the captured light field. By analyzing the reflectance model, we show that diffuse and specular surface points

- M. Zhou is with the DGene Digital Technology, Baton Rouge, LA 70820. E-mail: mingyuan.zhou@dgene.com.
- Y. Ding is with the Louisiana State University, Baton Rouge, LA 70803. E-mail: yding18@lsu.edu.
- Y. Ji is with DGene Digital Technology, Baton Rouge, LA 70820. E-mail: yu.ji@dgene.com.
- S. S. Young is with the US Army Research Laboratory, Adelphi, MD 20783. E-mail: susanyoung00@yahoo.com.
- J. Yu is with the Shanghai Tech University, Shanghai 201210, China. E-mail: yujingyi@shanghaitech.edu.cn.
- J. Ye is with the Division of Computer Science and Engineering, Louisiana State University, Baton Rouge, LA 70803. E-mail: jye@csc.lsu.edu.

Manuscript received 30 Mar. 2020; accepted 4 Apr. 2020. Date of publication 13 Apr. 2020; date of current version 3 June 2020.

(Corresponding author: Jinwei Ye.)

Recommended for acceptance by A. Chakrabarti.

Digital Object Identifier no. 10.1109/TPAMI.2020.2986764

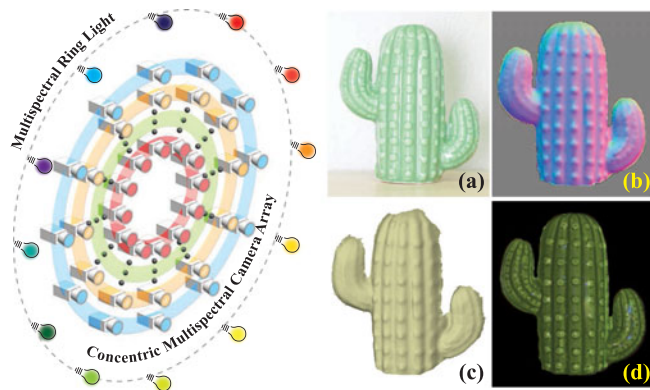


Fig. 1. Left: our concentric multi-spectral light field (CMSLF) array. We arrange the cameras on concentric circles, where each circle has the same number of cameras that capture at the same specific spectrum. A multi-spectral ring light surrounding the cameras provides direction-varying illumination for each camera circle. Right: our reconstruction results. (a) scene photograph; (b) recovered normal map; (c) recovered 3D surface; and (d) recovered reflectance map.

exhibit very different characteristics in S-Cam under our concentric viewpoint arrangement. Specifically, the specularity variations across views form a unique single-peak pattern which can enable robust depth estimation for specular points. For depth and reflectance estimation, we first initialize a depth map using multi-view stereo on S-Cams. We separate the diffuse and specular points by thresholding the intensity variance (because the diffuse points have constant intensity across views). We then rely on the single-peak pattern to refine the depths of specular points. Through our specularity analysis, we can subtract the specular components from the captured intensities to form specular-free S-Cams. We finally apply the photometric cues on the specular-free S-Cams to jointly estimate the surface normal and reflectance. We iterate these steps for refinement. We conduct extensive synthetic and real experiments to show that our approach is accurate and robust. We also show that our method outperforms state-of-the-art multiview-based and photometric stereo-based methods in shape and reflectance reconstruction, especially for non-Lambertian surfaces.

2 RELATED WORK

Our work is closely related to reflectance modeling and image-based surface shape and reflectance reconstruction.

Modeling surface reflectance is important to computer vision and graphics. The classical method in computer graphics uses the Lambertian model to characterize diffuse reflection and the Phong model for specular reflection. Although this method is not theoretically correct, it is still widely used and indispensable in computer graphics due to its simplicity in mathematical modeling. To characterize complex surface reflectance, the bidirectional reflectance distribution function (BRDF) [20] that measures the ratio between incident irradiance and exit radiance at a surface point is commonly used. The full BRDF model of a point requires a large parameter space as it exhausts all combinations of incident and exit lighting directions. A special case of the BRDF model is the dichromatic reflectance model, which was originally proposed by Shafer [20] to model dielectrics. It assumes that the BRDF of a surface can be

decomposed into two additive components: the interface (specular) reflectance and the body (diffuse) reflectance. Since all wavelength variations can be factorized from the two components, it is well suited for modeling multi-spectral reflectance. In our multi-spectral specular analysis, we combine the dichromatic reflectance model with the classical Phong model to characterize the surface reflectance sampled by our CMSLF.

Recovering the surface shape and reflectance from images is a fundamental problem in computer vision. The most popular two classes of methods are multi-view photogrammetry [2], [21], [22], [23], [24], [25], [26] and photometric stereo [4], [5], [6], [7], [27], [28], [29]. The former ones recover the 3D shape by triangulating rays from multiple viewpoints while the latter perform reconstruction under a fixed viewpoint but with various lighting directions. Although great success has been achieved on diffuse surfaces, specular highlights pose great challenges as they violate the color consistency assumption. Some methods [15], [30], [31] consider specular highlights as outliers and try to remove them. Some [32], [33], [34] rely on geometric and color distribution priors to compensate for specular regions. A recent work by Mecca *et al.* [35], [36] separates specular points from pure Lambertian reflection and treat them differently. However, their approach needs to take many images (around ten) as input. Oxholm and Nishino [37] recover the shape and reflectance of the homogeneous surface from a single image captured under uncontrolled illumination. Fyffe *et al.* [6] use spectral multiplexing to perform single-shot photometric stereo. Ikehata *et al.* [38] and Wu *et al.* [39] use the sparse representations to solve photometric stereo to compensate the corruptions caused by specularities. Zuo *et al.* [40] estimate surface geometry and albedo from RGB-D videos. Chandraker [41], [42] explores the motion cue for recovering shape and reflectance of a homogeneous object under a single directional light source. Wang *et al.* [17] extend the similar motion cue to spatial-varying BRDF using light field. Li *et al.* [18] improve the optimization framework for shape estimation with BRDF-invariant features. In this work, we present a novel concentric multi-spectral light field that is advantageous for handling non-Lambertian scenes because the concentric setting results in a unique pattern in specularity variation.

3 OUR APPROACH

In this section, we present our approach for surface shape and reflectance reconstruction using our concentric multi-spectral light field (CMSLF). We first present the system configuration (Section 3.1). We then formulate a physics-based reflectance model that characterize the CMSLF using the surface camera (S-Cam) (Section 3.2). Next, we perform a specularity analysis on our reflectance model and show that the specularity variations across views exhibit a unique single-peak pattern. We show that it is useful for the depth estimation for specular points (Section 3.3). Finally, we describe our shape and reflectance reconstruction algorithm for non-Lambertian surfaces (Section 3.4).

3.1 Concentric Multi-Spectral Light Field

As shown in Fig. 1, our CMSLF acquisition system consists of multi-spectral cameras and light sources that are arranged on coplanar and concentric circles. Each circle has

the same number of cameras which are uniformly spaced and capture the same specific spectrum. The surrounding multi-spectral ring light provides direction-varying illumination for each camera ring. It's worth noting that we use narrowband spectral filters for all cameras and light sources. We can, therefore, simultaneously sample multiple viewpoints under varying lighting directions without interference via spectral multiplexing.

To parameterize the CMSLF, we adopt the classical two-plane parametrization (2PP) [11] light field representation. Since our cameras are on coplanar circles, we set the center-of-projection (CoP) plane as the st plane at $z = 0$ and the image plane as the uv plane at $z = 1$. We use st for camera indices and uv for pixels.

Assuming we have m concentric camera rings in total and n cameras on each ring in a CMSLF, the camera position on the st plane can be written as $(s(i, j), t(i, j)) = (r_j \cos \phi_i, r_j \sin \phi_i)$, where $i \in \{1, \dots, n\}$ is the camera index in each concentric ring; $j \in \{1, \dots, m\}$ is the ring index, m also is the total number of spectral samples; r_j is the radius of the j th camera ring; $\phi_i = (i - 1)\tilde{\phi}$ is the spanned angle between the i th camera spoke and the x -axis ($\tilde{\phi} = 2\pi/n$ is the interval angle between neighboring camera spokes). The j th camera ring captures wavelength λ_j .

On the illumination side, since the lighting spectra match the ones sampled by the camera array, the number of point light sources equals to the number of camera rings (i.e., m). Assuming the light source ring is on a circle with radius r_l , the position of the j th light source in 3D can be written as $P_j = [r_l \cos \theta_j, r_l \sin \theta_j, 0]$ where $\theta_j = \theta_1 + (j - 1)\tilde{\theta}$ (θ_1 is the angular position of the first light source and $\tilde{\theta} = 2\pi/m$ is the angular interval between neighboring light sources).

We use the vector $\mathbf{P} = [P_1; \dots; P_m]$ to represent the set of all light source positions. Since we use narrowband spectral filters, the spectral illumination emitted from the j th point light source can only be received by the j th camera ring.

3.2 CMSLF Reflectance Model

We then formulate a physics-based reflectance model that characterize the CMSLF using the surface camera (S-Cam) representation [19].

3.2.1 Dichromatic Phong Model

We adopt the Dichromatic Reflectance Model (DRM) [43] to model reflectance. DRM separates the surface reflectance into body reflectance and interface reflectance, both terms account for geometry and color. DRM is suitable for modeling inhomogeneous materials.

Given a light source with the spectral distribution $E(\lambda)$ where λ refers to the wavelength, and a camera with spectral response function $Q(\lambda)$, the observed image intensity I under DRM at pixel p can be formulated as:

$$I(p) = w_d(p) \int_{\lambda_1}^{\lambda_N} R(p, \lambda) E(p, \lambda) Q(\lambda) d\lambda + w_s(p) \int_{\lambda_1}^{\lambda_N} E(p, \lambda) Q(\lambda) d\lambda, \quad (1)$$

where $[\lambda_1, \lambda_N]$ is the range of sampled wavelengths; $R(p, \lambda)$ is the surface reflectance; $w_d(p)$ and $w_s(p)$ are geometry-

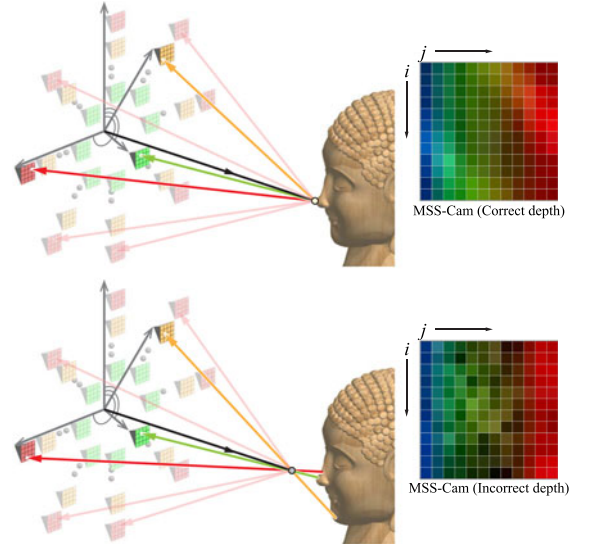


Fig. 2. Multi-spectral Surface Camera (MSS-Cam) sampling. Top: MSS-Cam sampled at the correct depth; Bottom: MSS-Cam for the same point but sampled at an incorrect depth.

related scale factors. The first term in Eq. (1) represents body reflectance that models light reflection after interacting with the surface reflectance. The second term represents interface reflectance that models light immediately reflected from the surface and thus causing specularities.

We apply numerical integration with step $\tilde{\lambda}$ on Eq. (1) which, after dropping pixel p becomes:

$$I = w_d \mathbf{R} \mathbf{E} \mathbf{Q} + w_s \mathbf{J} \mathbf{E} \mathbf{Q} \quad (2)$$

where \mathbf{J} is a row vector with all ones, $\mathbf{R} = [R(\lambda_1), R(\lambda_1 + \tilde{\lambda}), \dots, R(\lambda_N)]$, $\mathbf{E} = \text{diag}(E(\lambda_1), E(\lambda_1 + \tilde{\lambda}), \dots, E(\lambda_N))$ which is a diagonal matrix, and $\mathbf{Q} = [Q(\lambda_1), Q(\lambda_1 + \tilde{\lambda}), \dots, Q(\lambda_N)]^T$.

To take the scene geometry into consideration, we formulate the Phong dichromatic model that combines the classical Phong model with the DRM and assumes the near point lighting (NPL). Specifically, the diffuse and specular factors w_d and w_s are modeled with lighting position, viewing direction, surface normal and roughness. The image intensity I can then be written as:

$$I = k_d \left(\frac{L \cdot N}{\|P - X\|^2} \right) \mathbf{R} \mathbf{E} \mathbf{Q} + k_s \left(\frac{(D \cdot V)^\alpha}{\|P - X\|^2} \right) \mathbf{J} \mathbf{E} \mathbf{Q}, \quad (3)$$

where N is the surface normal at a 3D point X ; P is the position of light source; $L = (P - X) / \|P - X\|$ is the normalized lighting direction; V is the viewing direction; $D = 2(L \cdot N) - L$ is the reflection direction; α is the shininess parameter that models the surface roughness; k_d and k_s refer to the diffuse and specular surface reflectivity.

3.2.2 Multi-Spectral Surface Camera (MSS-Cam)

Next, we apply the Phong dichromatic reflectance model on our CMSLF using the Surface Camera (S-Cam) [19]. S-Cam characterizes the angular sampling characteristics of a light field. Given a 3D scene point, its S-Cam can be synthesized by tracing rays originated from the scene point into the light field to fetch color (see Fig. 2).

After applying the S-Cam on our CMSLF, we obtain the multi-spectral S-Cam (MSS-Cam). We now present the MSS-Cam image formation using our dichromatic Phong model. Given a pixel (u, v) in the center camera view with position $(s, t) = (0, 0)$, and assuming its corresponding 3D scene point is $X(u, v, z) = (x, y, z)$, we can synthesize its MSS-Cam M_X from the captured multi-spectral light field images. The pixels in a column of M_X are taken from cameras on the same ring that is sampled under a specific spectrum according to our concentric camera/light source arrangement. Moreover, each column captures the specular variation with respect to a single light source for non-Lambertian points. The pixels in the same row of M_X are taken from cameras on different rings but at the same camera spanned angle. To obtain MSS-Cam, we trace rays from the point X to each camera in the CMSLF. For a pixel (i, j) in the M_X , it samples the ray from the camera at $(s(i, j), t(i, j))$. By applying Eq. (3), we can write the intensity at an MSS-Cam pixel $M_X(i, j)$ as:

$$M_X(i, j) = k_d \left(\frac{L_j \cdot N}{\|P_j - X\|^2} \right) \mathbf{c} \mathbf{B}_j \mathbf{E}_j \mathbf{Q}_j + k_s \left(\frac{(D_j \cdot V_{i,j})^\alpha}{\|P_j - X\|^2} \right) \mathbf{J} \mathbf{E}_j \mathbf{Q}_j, \quad (4)$$

where $V_{i,j}$ is the viewing direction from the X to the camera $(s(i, j), t(i, j))$; D_j is the reflection direction of L_j ; $\mathbf{c} \mathbf{B}_j = \mathbf{R}$ indicates the linear decomposition of the reflectance spectra \mathbf{R} , with $\mathbf{c} = [c_1, \dots, c_w]$ as the reflectance coefficient vector and B_j as a $w \times h$ linear reflectance basis matrix under the spectral range $[\lambda_j - \frac{(h-1)}{2} \tilde{\lambda}, \lambda_j + \frac{(h-1)}{2} \tilde{\lambda}]$. E_j and Q_j have the same spectral range and are with sizes $h \times h$ and $h \times 1$ respectively.

3.3 Specularity Analysis

We now perform a specularity analysis on MSS-Cam and show that the specularity variation across views exhibits a unique single-peak pattern that is useful for specular region depth estimation.

We first separate the diffuse and specular points by thresholding the intensity variance because the diffuse points have constant reflectance across views. Fig. 2 shows that the MSS-Cam sampled at the correct scene depth exhibits photo consistency for ideal diffuse points (same intensities for pixels at the same column) and smooth intensity variations for specular points. If MSS-Cam is sampled at an incorrect depth, rays are integrated from different surface points. This results in random fluctuation of intensities in the MSS-Cam. The photo consistency for diffuse points in an MSS-Cam image M_X can be formulated as:

$$C(M_X) = \frac{1}{m} \sum_{j=1}^m std(M_X(1, j), \dots, M_X(n, j)), \quad (5)$$

where $std(\cdot)$ is the standard deviation function. This function indicates that the standard deviation for pixels taken from the same column should be very close to 0 if the point is diffuse and the MSS-Cam M_X is sampled at the correct depth. We therefore set a threshold on $C(M_X)$ to separate diffuse and specular points.

For specular points, we show that the specularity variation across views exhibits a unique single-peak pattern because of the concentric configuration of cameras. Intuitively, as shown in Fig. 3a, the cone-shaped lighting directions result in a reflection cone that is symmetric to the normal. Since the light field camera sampling for each spectrum is on a circle, the intensities from each column of the MSS-Cam will be changing like a sinusoidal curve with the camera's sampling angle ϕ from 0 to 2π in the circle (see Fig. 3b). Below we formalize this property as a proposition and provide its proof.

Proposition 1. *The specularity variation along the column of a MSS-Cam always forms a single-peak function.*

Proof. For simplification, we consider a specular point in 3D, $X = [0, 0, z]$, with the surface normal N . Its MSS-Cam M_X sampled at the correct scene depth can be formulated by applying Eq. (4). Now we analyze the intensity variation along the j th column of its MSS-Cam under the spectral light source j at $P_j = [r_j \cos \theta_j, r_j \sin \theta_j, 0]$. Assuming the normalized lighting direction is L , and the reflection direction w.r.t. the surface normal is $D = [d_x, d_y, d_z]$. The camera position from the camera ring is $[r_j \cos \phi, r_j \sin \phi, 0]$. The viewing direction can be computed as $V = [r_j \cos \phi, r_j \sin \phi, -z] / \sqrt{r_j^2 + z^2}$. The diffuse component and the term $k_s \mathbf{J} \mathbf{E}_j \mathbf{Q}_j / \|P_j - X\|^2$ in the specular component in Eq. (4) are constants c_d and c_s along the j th column of the MSS-Cam, therefore, the intensity along the j th column of the MSS-Cam can be rewritten as:

$$\begin{aligned} I(\phi) &= c_d + c_s (D \cdot V)^\alpha \\ &= c_d + \frac{c_s}{(r_j^2 + z^2)^{\frac{\alpha}{2}}} (r_j d_x \cos \phi + r_j d_y \sin \phi - z d_z)^\alpha \\ &= c_d + \frac{c_s}{(r_j^2 + z^2)^{\frac{\alpha}{2}}} (c_v \sin(\phi + \tilde{\phi}) - z d_z)^\alpha, \end{aligned} \quad (6)$$

where

$$c_v = \sqrt{(r_j d_x)^2 + (r_j d_y)^2} \\ \sin \tilde{\phi} = \frac{d_x}{\sqrt{d_x^2 + d_y^2}}, \quad \cos \tilde{\phi} = \frac{d_y}{\sqrt{d_x^2 + d_y^2}}.$$

The terms c_d and $c_s / (r_j^2 + z^2)^{\frac{\alpha}{2}}$ are constants along each column of the MSS-Cam. Therefore we can see that the intensity variation along a column of the MSS-Cam is an exponential sinusoidal curve w.r.t. the camera's angle ϕ in the ring, which is a single-peak function in the interval of $[0, 2\pi]$. \square

Notice that this proposition will not be affected by the arrangement of multi-spectral light sources. The spectral arrangement results in different shape of the curve because it changes the column ordering, but curve always remains one period of a sinusoid. Further, since the $\sin^n(\cdot)$ can be expanded as the Fourier series, we can approximate Eq. (6) as a Fourier series w.r.t. the camera's angle ϕ in the ring as:

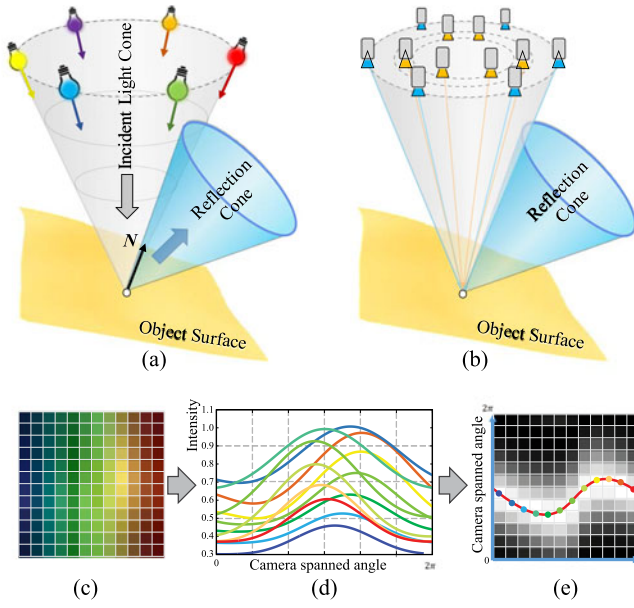


Fig. 3. The specularity variations in MSS-Cam exhibit unique pattern in our CMSLF. (a) The cone-shaped lighting directions result in a reflection cone that is symmetric to the normal. (b) Because of concentric camera setting on circles, the intensities from each column of the MSS-Cam will be changing on a sinusoidal curve (single-peak in the interval $[0, 2\pi]$). (c) An MSS-Cam with specularity. (d) We plot the pixel intensities from the same MSS-Cam column and show that they form single-peak patterns. (e) The camera spanned angles corresponding to the peaks of each curve in (d) form a sinusoidal curve.

$$I(\phi) \approx F(\phi) = a_0 + \sum_p a_p \cos p\phi + \sum_p b_p \sin p\phi. \quad (7)$$

We therefore fit a set of Fourier series F_1, \dots, F_m to model intensity variation of specular point per column in the M_X . On the other hand, as is shown in Figs. 3b and 3e, by taking the camera spanned angles corresponding to the peaks from each specularity curve, we can form a similar sinusoidal curve. We then fit another Fourier series F_0 to represent this curve of peak camera angles. Therefore, if a point is specular, its intensity in MSS-Cam should follow the consistency measurement below:

$$S(M_X) = \frac{1}{m} \sum_{j=1}^m \|\mathbf{U}(j) - \mathbf{F}(j)\| + \|\Phi - \mathbf{F}(0)\|, \quad (8)$$

where $\mathbf{U}(j) = [M_X(1, j), \dots, M_X(n, j)]$ are MSS-Cam pixels from the same column, $\mathbf{F}(j) = [F_j(\phi_1), \dots, F_j(\phi_n)]$ is the fitted Fourier series for this column. $\Phi = [\phi_s^{(1)}, \dots, \phi_s^{(m)}]$ are camera spanned angles corresponding to the peaks from each column, and $\mathbf{F}(0) = [F_0(1), \dots, F_0(m)]$ is the Fourier series fitted for them.

3.4 Shape and Reflectance Reconstruction

Finally, we show how to apply the CMSLF reflectance model and the specularity analysis on MSS-Cam for robust surface shape and reflectance reconstruction. Our reconstruction pipeline is shown in Fig. 4.

Depth Initialization. Given a pixel (u, v) and its corresponding 3D point $X(u, v, z)$ in the virtual center view at $[0, 0, 0]$. We first apply our photo-consistency measurement

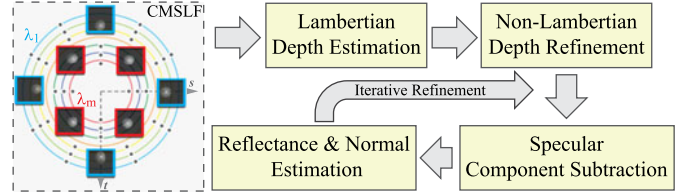


Fig. 4. The algorithmic pipeline for shape and reflectance reconstruction.

on the MSS-Cams with every hypothetical depth z of X to initialize its depth as:

$$z' = \arg \min_z C(M_X). \quad (9)$$

We classify this point as diffuse or specular point by thresholding the intensity variance across views. If a point is non-Lambertian, we use Eq. (8) refine its depth as:

$$z' = \arg \min_z S(M_X). \quad (10)$$

Note that we only perform Fourier series fitting for non-Lambertian points because this process is time consuming and applying it for all pixels would be inefficient.

Specular Component Subtraction. For non-Lambertian point, given its initial depth, we retrieve its MSS-Cam and then subtract the specular component to obtain a specular-free MSS-Cam. Specifically, we first compute the vertical gradients of the MSS-Cam to remove the diffuse component in Eq. (4) as:

$$\begin{aligned} \nabla M_X(i, j) &= (M_X(i+1, j) - M_X(i, j)) \\ &= k_s((D_j \cdot V_{i+1, j})^\alpha - (D_j \cdot V_{i, j})^\alpha) \frac{\mathbf{J}\mathbf{E}_j\mathbf{Q}_j}{\|P_j - X\|^2}. \end{aligned} \quad (11)$$

By rearranging Eq. (11), we have:

$$G_X(i, j) = \nabla M_X(i, j) \frac{\|P_j - X\|^2}{\mathbf{J}\mathbf{E}_j\mathbf{Q}_j}, \quad (12)$$

where $G_X(i, j) = k_s((D_j \cdot V_{i+1, j})^\alpha - (D_j \cdot V_{i, j})^\alpha)$.

Given the pre-calibrated term $\mathbf{J}\mathbf{E}_j\mathbf{Q}_j$ and the lighting position P_j (the calibration process is described in Section 4.2.2), and the calculated gradients, we compute the observed \tilde{G}_X to optimize the surface normal N , specular reflectivity k_s and surface roughness α simultaneously through the following objective function:

$$\arg \min_{N, \alpha, k_s} \sum_{i, j} \|\tilde{G}_X(i, j) - k_s((D_j \cdot V_{i+1, j})^\alpha - (D_j \cdot V_{i, j})^\alpha)\|, \quad (13)$$

where $D_j = 2(L_j \cdot N)N - L_j$. As this optimization is non-linear, we can apply the Levenberg-Marquardt method to solve the specular parameters that satisfy the specularity gradient variations.

Given these specular parameters, we can then subtract the specular components from our MSS-Cam and form a specular-free MSS-Cam A_X as:

$$A_X(i, j) = M_X(i, j) - k_s \left(\frac{(D_j \cdot V_{i,j})^\alpha}{\|P_j - X\|^2} \right) \mathbf{J} \mathbf{E}_j \mathbf{Q}_j. \quad (14)$$

We form a row vector $\mathbf{A}_X = [\text{median}(A_X(:, 1)), \dots, \text{median}(A_X(:, m))]$ by taking the median values from each column of A_X .

Normal and Reflectance Refinement. Finally, we apply multi-spectral photometric stereo on our specular-free MSS-Cam to obtain more accurate surface normal and reflectance. Our objective function is formulated as:

$$\arg \min_{N, c} ((\mathbf{c} \mathbf{W}) \circ (\mathbf{L} \mathbf{N})^T - \mathbf{A}_X), \quad (15)$$

where \circ is Hadamard product (element-wise multiplication), $\mathbf{L} = [L_1; \dots; L_m]$. The term $\mathbf{W} = [\mathbf{W}_1, \dots, \mathbf{W}_m]$ where $\mathbf{W}_j = \mathbf{B}_j \mathbf{E}_j \mathbf{Q}_j$, which is composed of the pre-calibrated camera and light source spectral responses. When the number of sampled spectra is greater than or equal to the dimensions of the reflectance spectra, that is $w \times 3 \leq m$, the above optimization can be formulated as an over-determined linear system. We can write the following linear system by unrolling the matrices in Eq. (15):

$$b = H \setminus \mathbf{A}_X^T, \quad (16)$$

where

$$H = \begin{bmatrix} \mathbf{W}(1,1)\mathbf{L}(1,1) & \mathbf{W}(1,1)\mathbf{L}(1,2) & \mathbf{W}(1,1)\mathbf{L}(1,3) & \dots & \mathbf{W}(w,1)\mathbf{L}(1,3) \\ \mathbf{W}(1,2)\mathbf{L}(2,1) & \mathbf{W}(1,2)\mathbf{L}(2,2) & \mathbf{W}(1,2)\mathbf{L}(2,3) & \dots & \mathbf{W}(w,2)\mathbf{L}(2,3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{W}(1,m)\mathbf{L}(m,1) & \mathbf{W}(1,m)\mathbf{L}(m,2) & \mathbf{W}(1,m)\mathbf{L}(m,3) & \dots & \mathbf{W}(w,m)\mathbf{L}(m,3) \end{bmatrix}$$

$$b = [c_1 n_x, c_1 n_y, c_1 n_z, \dots, c_w n_x, c_w n_y, c_w n_z]^T.$$

The surface normal and reflectance coefficients can be directly derived from b . Otherwise, when $w \times 3 > m$, we apply the Levenberg-Marquardt algorithm to solve Eq. (15).

After recovering the surface normal, we update the depth estimation through normal integration. We then recompute all MSS-Cams with the updated depths and repeat the above steps to iteratively refine the shape and reflectance estimation.

For multi-spectral reflectance estimation, we fit a dense multi-spectral reflectance $\mathbf{R} = \mathbf{c}' \mathbf{B}$ using the estimated reflectance coefficients \mathbf{c}' . \mathbf{B} are the dense spectral sampling reflectance basis functions determined on the Munsell color sets [44], [45].

4 EXPERIMENTS

In this section, we present synthetic and real experimental results to evaluate our approach. All experiments are performed on a desktop computer with Intel i7 7820 CPU (2.9GHz Quad-core) and 32G memory. Our algorithm is implemented in Matlab. Although we are able to recover multi-spectral reflectance, we still show the reflectance maps in RGB for the ease of visualization.

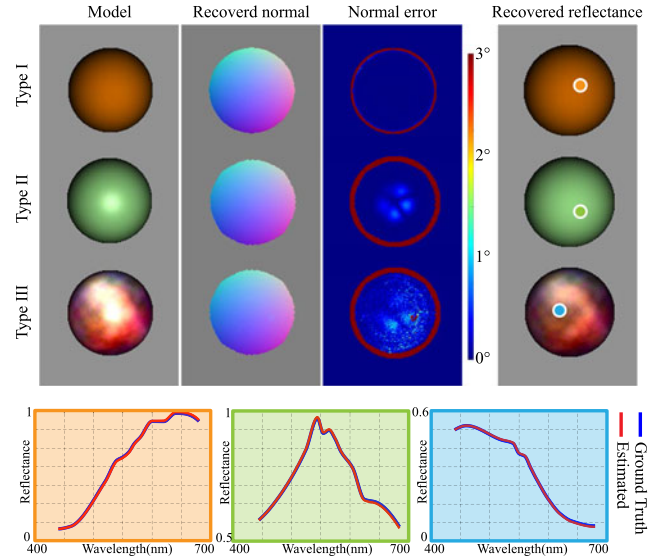


Fig. 5. Synthetic results of a sphere object with three types of materials. Top: we show the input image of the sphere models, the recovered normal map, the error map of normal estimation in degree, and the recovered reflectance (from left to right); Bottom: we show the multi-spectral reflectance curve (red) of a randomly sampled point on each sphere in comparison with the ground truth reflectance curve (blue).

4.1 Synthetic Experiments

To simulate our CMSLF images, we implement a multi-spectral renderer to generate the input images. We consider the same camera and light source setup as used in our real experiments (described in Section 4.2.1). To simulate multi-spectral reflectance from RGB textures, we use high-dimensional reflectance linear basis to fit the missing spectra.

We first work on a simple sphere object with three different types of reflectance: ideal diffuse (Type I), specular (Type II) and specular with texture (Type III) (see Fig. 5). Through this experiment, we show that our approach can be applied to a variety of surface materials. In the material setting, the diffuse coefficients are all set to 0.7; the specular coefficients for each material are set to 0, 0.3, and 0.5 for Types I to III respectively; and the roughness for Type II/III is set to 10. The radius of the sphere is set to 20 units and the distance between the CMSLF and sphere is 120 units. As for our CMSLF, the radius of our ring light source is 80 units and contains 12 spectral light sources. We use a 12×8 concentric camera array (i.e., 12 circles with 8 cameras on each circle). The radius of the concentric camera circles ranges from 29 to 40 units with interval 1 unit for each circle in between. We consider spectral samples between wavelength 440 nm to 660 nm with interval 20 nm for both cameras and light sources. For each camera, we simulate image of the object at resolution 320×320 . For depth estimation, we discretize the depth values from 108 to 125 units with step 0.2. Our sphere has 76 depth layers. We apply our algorithm described in Section 3.4 to recover the shape and reflectance for the three types of sphere. Our reconstruction results are shown in Fig. 5. To evaluate our shape reconstruction, we compare our recovered normal maps with the ground truth ones and compute the error maps. The maximum normal error is less than 2° for all three cases. To evaluate the reflectance reconstruction, we plot the multi-spectral reflectance curve for a randomly sampled point on

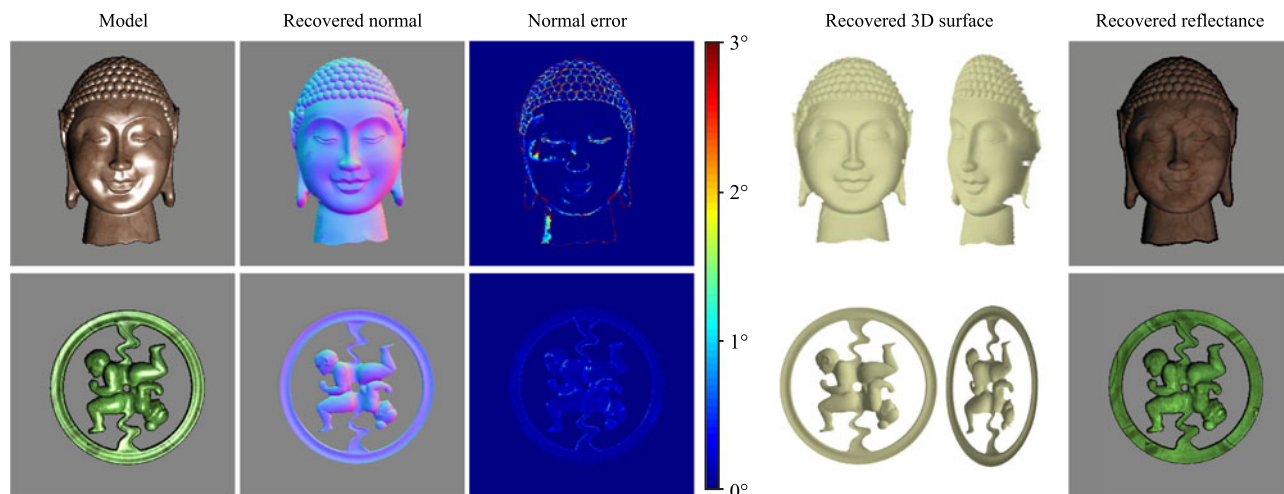


Fig. 6. Synthetic results for the buddha and plate. From left to right, we show the input model images, recovered normal maps, the recovered normal maps, the error maps of normal estimation in degree, the recovered 3D surfaces and the recovered reflectance images.

the sphere. We compare the curve with the ground truth reflectance we set for the object. We can see that the two curves are highly close for all three cases, which indicates that our method recovers accurate multi-spectral reflectance for various materials.

To evaluate our performance on more complex material models, we perform experiments using the MERL dataset [46]. We use the same sphere and CMSLF setting, except that we use the BRDF functions in the MERL dataset for material modeling. We evaluate the reconstruction error for all materials. The average normal error is 3.8° . The complete set of results are included in the supplementary material, which can be found on the Computer Society Digital Library at <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2020.2986764>.

Next, we perform synthetic experiments on two objects with more complex geometries, a Buddha head (or buddha) and an antique pendant plate (or plate). For these experiments, our light source setting and spectral sampling scheme are the same as the sphere object experiment. For the CMSLF, the radius of the camera circles range from 4 to 2.9 with step 0.1. Each camera renders images at a higher resolution (500×500) to capture the fine details of the objects. For surface material, we set the specular coefficients to 0.6 (buddha) and 0.4 (plate). The roughness is set to 20 for both. The distances between objects and camera are 41 (buddha) and 34 (plate). We discretize the depth values in the range from 35 to 42 with step 0.1 for the buddha and from 33 to 35 with step 0.1 for the plate. Our shape and reflectance reconstruction results are shown in Fig. 6. We evaluate our shape reconstruction in terms of normal errors and we can see that the maximum normal error is less than 3° for both objects. We also show our recovered 3D surfaces. We can see that the geometric details are well preserved. We re-render the diffuse images to show that our recovered reflectance is highly accurate.

We compare our shape reconstruction with two state-of-the-art light field-based depth estimation methods [30] and [18]. Tao *et al.* [30] use specular properties in the epipolar plane images (EPI) to recover glossy surfaces with light field. Li *et al.* [18] develop an optimization framework to

recover the shape and reflectance for surfaces with spatial varying BRDF. Both methods are expected to handle non-Lambertian surfaces well. We render 12×12 grid-based light field images and use them as input to the two methods. We use the source codes provided by the authors for testing. The comparison results on buddha is shown in Fig. 7. We can see that our recovered 3D surfaces are with higher quality. This is mainly because of two reasons: 1) our concentric circular light field setting provide additional photometric cues for non-Lambertian points depth estimation; and 2) our multi-spectral sampling scheme provides more diverse lighting and viewpoint samples, which allows for more robust shape and reflectance reconstruction.

Finally, we analyze the performance of our approach with respect to different numbers of cameras, light sources and image noises. Theoretically, the minimal setup of CMSLF requires 5 multi-spectral light sources and 5×3 concentric camera array (i.e., five camera rings and three cameras on each ring) for recovering the surface shape and reflectance. In practice, more multi-spectral light sources as well as cameras can help improve the reconstruction

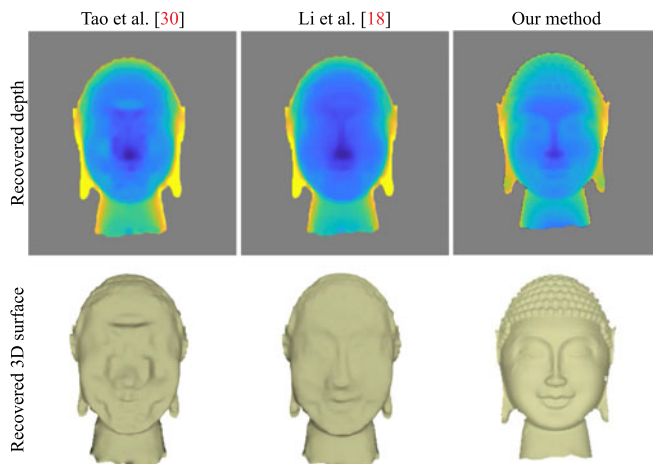


Fig. 7. Comparison results on synthetic data. We compare the recovered depth map and 3D surfaces with two state-of-the-art light field-based methods [30] and [18].

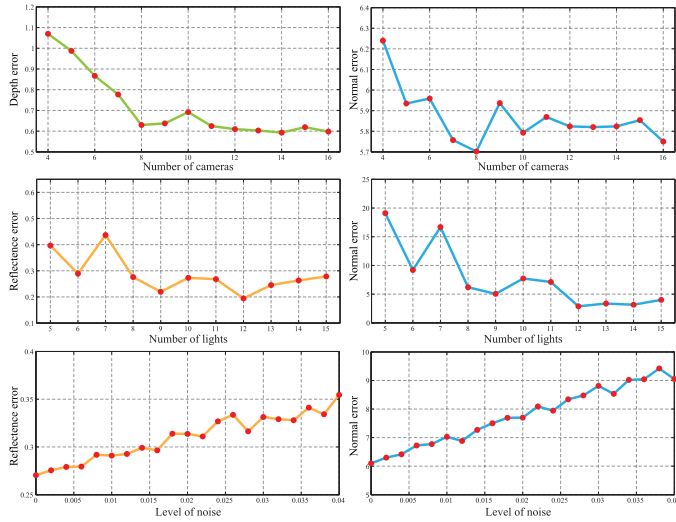


Fig. 8. Performance analysis with respect to the number of cameras, light sources and image noise.

accuracy. Therefore, we perform experiments by changing CMSLF settings to understand the impact of number of cameras and light sources. We use the sphere model with Type III material as the reconstruction target. First, we fix the number of multi-spectral light sources to 12 and change the number of cameras on each circle from 4 to 16 to evaluate the system performance. The reconstruction error line plots are shown in the first row of Fig. 8. We can see that although both the depth and normal errors decreases as the number of cameras increases, the improvement becomes marginal when the number of cameras is greater than 8 on each circle. We then fix the number of cameras to 8 and vary the number of multi-spectral light sources from 5 to 15. The reconstruction error line plots are shown in the second row of Fig. 8. We can see that the accuracy improvement becomes marginal when the number of light sources is greater than 12. Therefore we conclude that the optimal configuration for CMSLF is to use 12 multi-spectral light sources and 8 cameras in each circle. We use this configuration to build our real prototype system. We also evaluate the performance with respect to image noise. We add zero-mean Gaussian noise with different variances (0 ~ 0.4) to the input images. The reconstruction error line plots are shown in the third row of Fig. 8. We can see that both the reflectance and normal errors increases as the noise level goes up. The normal estimation is more sensitive to noise than the reflectance.

4.2 Real Experiments

We perform experiments on various real scenes using a CMSLF system prototype.

4.2.1 System Configuration

We construct a CMSLF using customized hardware. Our prototype system is shown in Fig. 9. To emulate the concentric multi-spectral light field camera array with a single camera, we mount a monochrome camera (Point Grey GS3-U3-51S5M-C) with a 50mm lens on a translation stage to move the camera to specific positions on concentric circles

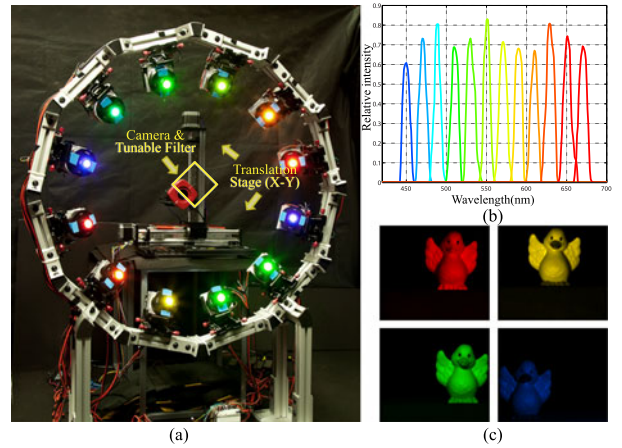


Fig. 9. (a) Our prototype concentric multi-spectral light field (CMSLF) system. (b) The illumination spectra of our multi-spectral ring light source. (c) Sample multi-spectral images captured with our CMSLF.

on a 2D plane (i.e., the st plane). Specifically, we emulate a 12×8 concentric camera array (i.e., 12 circles with 8 cameras on each circle). We mount a tunable liquid crystal spectral filter (KURIOS-WL1) in front of the camera to capture the scene under specified wavelengths. The horizontal field-of-view of the camera is 13° and the resolution of the captured image is 2448×2048 . According to the dimension of our system, the optimal acquisition distance is to place the target around 100cm away from our system.

To build the multi-spectral ring light, we mount twelve 30 Watt LED chips onto a dodecagon frame, the distance between each LED chip and the center of the dodecagon is 50cm. We then place 12 narrow-band spectral filters with wavelengths ranging from 450 nm to 670 nm with step 20 nm in front of the LED chips to emit multi-spectral illumination. We arrange the spectral filters such that the shading variation along the illumination ring for a fluctuated curve with respect to the wavelength, as shown in Fig. 10. This is because we need to reduce the ambiguity on reflectance and shading variation separation in Eq. (15). When the multi-spectral lights are arranged in this way, the surface normal can be more easily converged to the global optimum.

4.2.2 System Calibration

We need to perform geometric and photometric calibration on our system. Specifically, we calibrate the camera intrinsics and extrinsics, the light sources' 3D positions and spectral responses. The reflectance basis functions and camera spectral response are also pre-computed in the calibration process.

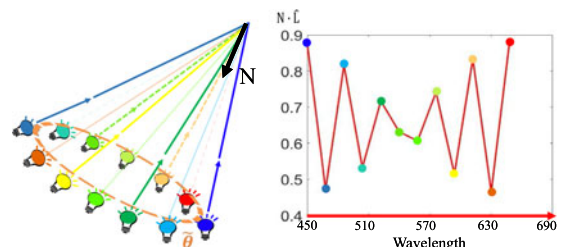


Fig. 10. Our desired multi-spectral light source arrangement. The shading variation of the light sources with respect to the wavelength form a fluctuated curve.

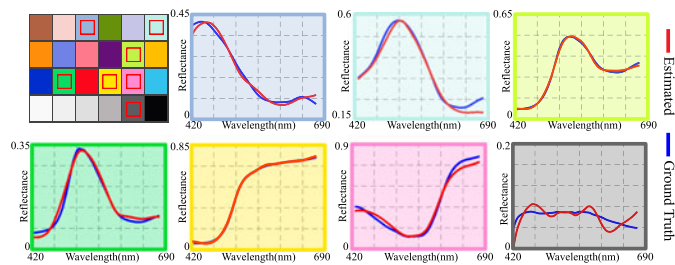


Fig. 11. Evaluation of spectral calibration results. We perform spectral calibration using standard colorchecker chart. We evaluation our calibration results by comparing our estimated spectral response for the color checker with the ground truth ones.

Camera Calibration. We first calibrate the camera intrinsic and extrinsic parameters using traditional camera calibration method [47]. The camera is then moved to specific positions on concentric circles with a high-precision 2D translation stage.

Light Source Calibration. To calibrate the light source positions, we first move the camera to the center of the concentric circle. For each light source, we capture a sequence of images with a chrome ball at different locations. In each image, we detect the specular spot on the chrome ball as an incident point. We then compute the incident ray direction using the reflection ray direction (i.e., the camera ray) and the surface normal. The light source position can then be determined by backtracing all incident rays from the image sequence to form an intersection point. Intuitively, this procedure needs to be repeated for every light source. In practice, we capture the images for all light source in one pass through spectral-multiplexing.

Spectral Calibration. In this step, we calibrate the mixed spectral response $\mathbf{W}_j = \mathbf{B}_j \mathbf{E}_j \mathbf{Q}_j$ (where \mathbf{E}_j is the light source's spectral response and \mathbf{Q}_j is the camera's) by capturing images of the standard colorchecker chart for each spectral light. As we know the ground truth reflectance response of the color swatches, we can apply PCA to extract the coefficient vector for each color and combine them together to a coefficients matrix $\mathbf{C} = [\mathbf{c}_1; \dots; \mathbf{c}_{24}]$ with size $24 \times w$ (where 24 is the total number of color swatches), we also can get the dense spectral reflectance basis \mathbf{B} . For the j th spectral light source, we adjust our tunable spectral filter in front of the camera to capture the color checker only under its specific spectrum. With the known checker position and the light position, we can eliminate the shading term $(L_j \cdot N) / \|P_j - X\|^2$. We then average the intensities without shading components for each color swatch to form a vector I'_j with size is 24×1 . We can solve \mathbf{W}_j by optimizing the following objective function:

$$\arg \min_{\mathbf{W}_j} (\mathbf{C} \mathbf{W}_j - I'_j). \quad (17)$$

We perform this optimization for all spectra and obtain $\mathbf{W} = [\mathbf{W}_1, \dots, \mathbf{W}_m]$. The spectral response $\mathbf{E}_j \mathbf{Q}_j$ can then be computed from the calculated \mathbf{W}_j with the known reflectance responses.

We evaluate our spectral calibration by comparing our estimated spectral responses for the colorchecker chart with the ground truth ones. The results are shown in Fig. 11.

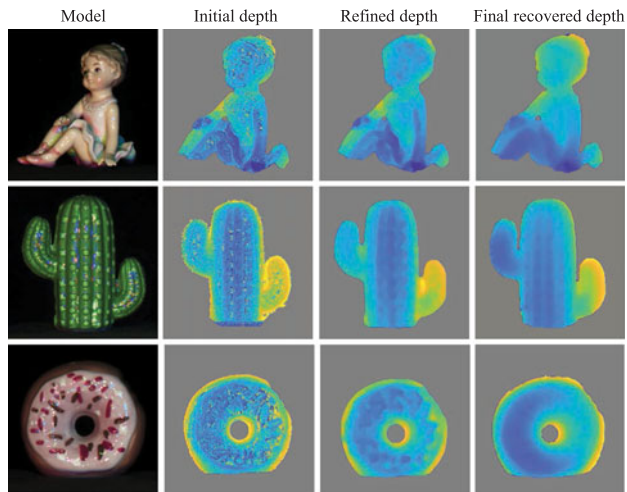


Fig. 12. Depth maps from each step of our algorithm.

From our sample results, we can see that our spectral estimations are highly accurate. The estimation accuracy is lower for the dark gray and black checkers because their captured intensities are low resulting in low signal-noise ratio in the capture images.

4.2.3 Results

Here we present real scene results using our CMSLF system. We use our prototype CMSLF system to capture objects with diverse reflectances, ranging from diffuse to highly specular. Our shape and reflectance reconstruction results are shown in Fig. 13. We can see that our approach works well for specular objects, such as ceramic figurine and plastic toys. We also compare our recovered 3D surfaces with state-of-the-arts light field-based methods [18], [30] and photometric stereo-based methods [6], [38], [39], [48]. The visual comparison results of the recovered 3D surfaces are shown in Fig. 14. Specifically, both light field-based methods are expected to handle specularly well. Among the photometric stereo-based methods, [48] and [6] are not designed to recover non-Lambertian surfaces; [38] and [39] compensate the corruption caused specularly with sparse solvers. We can see that our approach outperforms all these methods. This shows that our concentric multi-spectral light field setting is advantageous.

To better understand our algorithm, we show the intermediate depth estimation results from three examples in Fig. 12. The “initial depth” is the depth map initialized with the photo-consistency constraint. As this constraint only works for diffuse points, we can notice erroneous depth estimations for specular points. The “refined depth” refers to the estimation that refines the depths for specular points using the single-peak curve. We can notice obvious improvement on specular points. The “final recovered depth” is our final result after iterative refinement.

We also quantitatively evaluation our shape and multi-spectral reflectance reconstruction using a standard colorchecker chart. We compare our method with the multi-spectral photometric stereo [6]. The results are shown in Table 1. We can see that both the normal errors and reflectance errors of our method are significantly smaller.

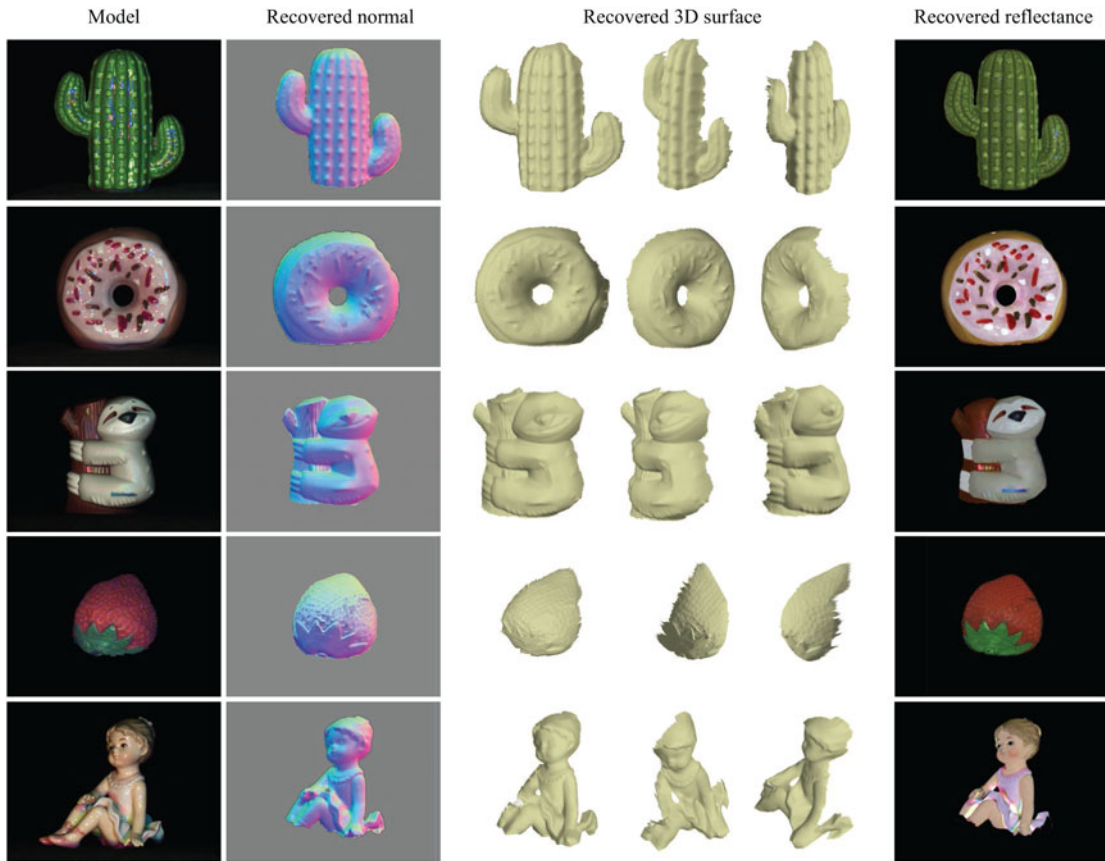


Fig. 13. Real experiment results. From left to right, we show the input model images, recovered normal maps, the recovered normal maps, the recovered 3D surfaces and the recovered reflectance images.

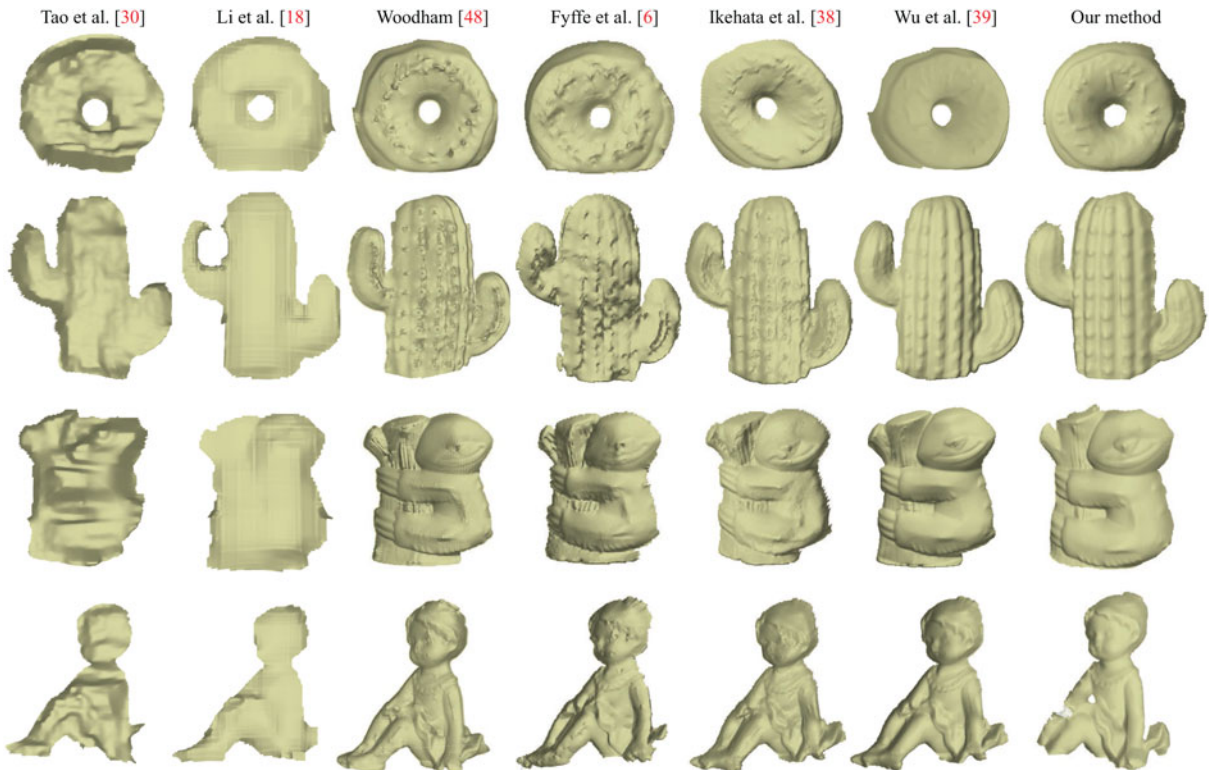


Fig. 14. Visual comparison of recovered 3D surface with state-of-the-arts light field-based methods [18], [30] and photometric stereo-based methods [6], [38], [39], [48].

TABLE 1

Quantitative Evaluation of Multi-Spectral Reflectance Estimation Using the Standard Colorchecker Chart in Comparison With [6]

		dark skin	light skin	blue sky	foliage	blue flower	bluish green	orange	purple red	moderate red	purple	yellow green	orange yellow
Normal RMSE	[6]	18.7°	21.3°	16°	27.4°	24.8°	28.3°	15.7°	20.8°	16.2°	9.54°	18.9°	30.4°
	Ours	3.79°	1.89°	2.74°	6.37°	2.16°	5.48°	4.29°	2.12°	6.61°	2.89°	8.69°	3.72°
Reflectance RMSE	[6]	0.328	0.097	0.129	0.340	0.261	0.365	0.152	0.179	0.130	0.212	0.280	0.293
	Ours	0.0173	0.0187	0.0099	0.0188	0.0150	0.0212	0.0389	0.0299	0.0320	0.0183	0.0205	0.0416
		blue	green	red	yellow	magenta	cyan	white	neutral 8	neutral 65	neutral 5	neutral 35	black
Normal RMSE	[6]	17.9°	41.2°	30.2°	15.9°	14.7°	20.6°	35.4°	22.4°	12.5°	15.6°	31.9°	41.6°
	Ours	1.76°	2.37°	2.08°	5.18°	1.40°	5.07°	8.32°	5.14°	1.32°	1.85°	2.56°	10.67°
Reflectance RMSE	[6]	0.204	0.562	0.300	0.269	0.227	0.324	0.454	0.273	0.197	0.203	0.291	0.375
	Ours	0.0144	0.0092	0.0330	0.0134	0.0318	0.0216	0.0219	0.0229	0.0208	0.0066	0.0040	0.0047

5 CONCLUSION AND DISCUSSIONS

In summary, we have presented a concentric multi-spectral light field (CMSLF) based method for surface shape and reflectance reconstruction. We've shown that our concentric camera and light source setting results in a unique single-peak pattern in specular variations across views. The pattern allows for robust depth estimation for specular points. Through comprehensive synthetic and real experiments, we show that our method can achieve highly accurate and robust results for non-Lambertian surfaces.

However, our method has limitations in handling occlusions and shadows. Our shape and reflectance reconstruction accuracy will be degraded if a point is in shadow (see occluded regions in the ceramic girl figurine scene in Fig. 13). One promising future direction is to exploit characteristics of occluding boundaries in our CMSLF and develop reconstruction algorithms that are robust to occlusions. Our current prototype system is bulky and requires moving the camera with a translational stage to simulate the concentric light field. In the future, we plan to build a CMSLF system using an array of cameras with spectral filters to realize single-shot acquisition.

ACKNOWLEDGMENTS

This work was supported in part by the Louisiana Board of Regents under Grant LEQSF(2018-21)-RD-A-10.

REFERENCES

- [1] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, "Multi-view stereo for community photo collections," in *Proc. 11th Int. Conf. Comput. Vis.*, 2007, pp. 1–8.
- [2] G. Oxholm and K. Nishino, "Multiview shape and reflectance from natural illumination," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2155–2162.
- [3] P. F. Gotardo, T. Simon, Y. Sheikh, and I. Matthews, "Photogeometric scene flow for high-detail dynamic 3D reconstruction," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 846–854.
- [4] G. Vogiatzis, P. Favaro, and R. Cipolla, "Using frontier points to recover shape, reflectance and illumination," in *Proc. 10th IEEE Int. Conf. Comput. Vis.*, 2005, pp. 228–235.
- [5] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1060–1071, Jun. 2010.
- [6] G. Fyffe, X. Yu, and P. Debevec, "Single-shot photometric stereo by spectral multiplexing," in *Proc. IEEE Int. Conf. Comput. Photography*, 2011, pp. 1–6.
- [7] J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 8, pp. 1670–1687, Aug. 2015.
- [8] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Comput. Sci. Tech. Report*, vol. 2, no. 11, pp. 1–11, 2005.
- [9] "The (new) stanford light field archive," 2008. [Online]. Available: <http://lightfield.stanford.edu/>
- [10] M. Levoy, "Light fields and computational imaging," *Computer*, vol. 39, no. 8, pp. 46–55, Aug. 2006.
- [11] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. 23rd Annu. Conf. Comput. Graph. Interactive Techn.*, 1996, pp. 31–42.
- [12] S. Wanner and B. Goldluecke, "Globally consistent depth labeling of 4D light fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 41–48.
- [13] Z. Yu, X. Guo, H. Lin, A. Lumsdaine, and J. Yu, "Line assisted light field triangulation and stereo matching," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 2792–2799.
- [14] H.-G. Jeon *et al.*, "Accurate depth map estimation from a lenslet light field camera," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1547–1555.
- [15] M. W. Tao, T.-C. Wang, J. Malik, and R. Ramamoorthi, "Depth estimation for glossy surfaces with light-field cameras," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 533–547.
- [16] M. W. Tao, P. P. Srinivasan, J. Malik, S. Rusinkiewicz, and R. Ramamoorthi, "Depth from shading, defocus, and correspondence using light-field angular coherence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1940–1948.
- [17] T.-C. Wang, M. Chandraker, A. A. Efros, and R. Ramamoorthi, "SVBRDF-invariant shape and reflectance estimation from light-field cameras," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 5451–5459.
- [18] Z. Li, Z. Xu, R. Ramamoorthi, and M. Chandraker, "Robust energy minimization for BRDF-invariant shape from light fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5571–5579.
- [19] J. Yu, L. McMillan, and S. Gortler, "Scam light field rendering," in *Proc. 10th Pacific Conf. Comput. Graph. Appl.*, 2002, pp. 137–144.
- [20] S. A. Shafer, "Using color to separate reflection components," *Color Res. Appl.*, vol. 10, no. 4, pp. 210–218, 1985.
- [21] H. Jin, D. Cremers, A. J. Yezzi, and S. Soatto, "Shedding light on stereoscopic segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2004, pp. I–I.
- [22] H. Jin, S. Soatto, and A. J. Yezzi, "Multi-view stereo reconstruction of dense shape and complex appearance," *Int. J. Comput. Vis.*, vol. 63, no. 3, pp. 175–189, 2005.
- [23] C. Wu, Y. Liu, Q. Dai, and B. Wilburn, "Fusing multiview and photometric stereo for 3D reconstruction under uncalibrated illumination," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 8, pp. 1082–1095, Aug. 2011.
- [24] A. Ghosh, G. Fyffe, B. Tunwattanapong, J. Busch, X. Yu, and P. Debevec, "Multiview face capture using polarized spherical gradient illumination," *ACM Trans. Graph.*, vol. 30, 2011, Art. no. 129.
- [25] J. Park, S. N. Sinha, Y. Matsushita, Y.-W. Tai, and I. S. Kweon, "Robust multiview photometric stereo using planar mesh parameterization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 8, pp. 1591–1604, Aug. 2017.
- [26] G. Fyffe, P. Graham, B. Tunwattanapong, A. Ghosh, and P. E. Debevec, "Near-instant capture of high-resolution facial geometry and reflectance," *Comput. Graph. Forum*, vol. 35, pp. 353–363, 2016.
- [27] C. Hernández, G. Vogiatzis, G. J. Brostow, B. Stenger, and R. Cipolla, "Non-rigid photometric stereo with colored lights," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, 2007, pp. 1–8.
- [28] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008.
- [29] R. Anderson, B. Stenger, and R. Cipolla, "Color photometric stereo for multicolored surfaces," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 2182–2189.
- [30] M. W. Tao, J.-C. Su, T.-C. Wang, J. Malik, and R. Ramamoorthi, "Depth estimation and specular removal for glossy surfaces using point and line consistency with light-field cameras," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 6, pp. 1155–1169, Jun. 2016.

- [31] S. Lin, Y. Li, S. B. Kang, X. Tong, and H.-Y. Shum, "Diffuse-specular separation and depth recovery from image sequences," in *Proc. Eur. Conf. Comput. Vis.*, 2002, pp. 210–224.
- [32] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur, "Beyond lambert: Reconstructing specular surfaces using color," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 619–626.
- [33] Y. Sato and K. Ikeuchi, "Temporal-color space analysis of reflection," *J. Optical Soc. America A*, vol. 11, no. 11, pp. 2990–3002, 1994.
- [34] K. Nishino, Z. Zhang, and K. Ikeuchi, "Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis," in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, 2001, pp. 599–606.
- [35] R. Mecca and Y. Quéau, "Unifying diffuse and specular reflections for the photometric stereo problem," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2016, pp. 1–9.
- [36] S. Tozza, R. Mecca, M. Duocastella, and A. Del Bue, "Direct differential photometric stereo shape recovery of diffuse and specular surfaces," *J. Math. Imaging Vis.*, vol. 56, no. 1, pp. 57–76, 2016.
- [37] G. Oxholm and K. Nishino, "Shape and reflectance from natural illumination," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 528–541.
- [38] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 318–325.
- [39] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *Proc. Asian Conf. Comput. Vis.*, 2010, pp. 703–717.
- [40] X. Zuo, S. Wang, J. Zheng, and R. Yang, "Detailed surface geometry and albedo recovery from RGB-D video under natural illumination," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 3133–3142.
- [41] M. Chandraker, "On shape and material recovery from motion," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 202–217.
- [42] M. Chandraker, "What camera motion reveals about shape with unknown BRDF," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2179–2186.
- [43] S. A. Shafer, "Using color to separate reflection components," *Color Res. Appl.*, vol. 10, no. 4, pp. 210–218, 1985. [Online]. Available: <http://dx.doi.org/10.1002/col.5080100409>
- [44] J. P. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of munsell colors," *J. Optical Soc. America A*, vol. 6, no. 2, pp. 318–322, 1989.
- [45] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *J. Optical Soc. America A*, vol. 3, no. 10, pp. 1673–1683, 1986.
- [46] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 759–769, Jul. 2003.
- [47] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
- [48] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Eng.*, vol. 19, no. 1, 1980, Art. no 191139.



Mingyuan Zhou received the BE degree in intelligence science and technology from the Beijing Information Science and Technology University, China, in 2011, the ME degree in computer engineering from the Stevens Institute of Technology, in 2014, and the PhD degree in computer science from the University of Delaware, in 2019. He is now a research scientist at DGene. His research interests include computer vision and computational photography.



Yuqi Ding received the BS and MS degrees in geographical information system from Wuhan University, in 2013 and 2009. He is currently working toward the PhD degree at the Division of Computer Science and Engineering, Louisiana State University. His research interests include computational photography, computer vision, and computer graphics.



Yu Ji received the bachelor's degree in electrical engineering from the Huazhong University of Science and Technology, in 2009, the MSc degree in digital media from Nanyang Technological University, in 2011, and the PhD degree in computer science from the University of Delaware, in 2014. He is currently a principal scientist at DGene Digital Media. His research interests include computational photography, computer vision, and computer graphics.



S. Susan Young (Senior Member, IEEE) received the BS and MS degrees in electrical and computer engineering from the South China University of Technology, Guangzhou, China, in 1984 and 1987, respectively, and the PhD degree from the State University of New York at Buffalo, Amherst, in 1995. She joined as a research associate with the Department of Radiation Medicine, Roswell Park Cancer Institute, Buffalo, NY, in 1995, where she was involved in research on medical imaging problems for radiation therapy. From 1996 to 2002, she was a senior research scientist with the Health Imaging Research Laboratory, Eastman Kodak Company, Rochester, NY, where she was engaged in multi-resolution object recognition and image analysis on medical images, and medical image compression. Since 2002, she has been with the U.S. Army Research Laboratory, Adelphi, MD, researching automatic target recognition, image processing technology for various army weapon applications, thermal modeling and perception, and biometric recognition. She has co-authored a text book *Signal Processing and Performance Analysis for Imaging Systems* (Artech House, 2008). She has authored or coauthored more than 70 papers in reputable journals and conferences. She holds eleven patents for inventions on medical diagnostic imaging, image compression, image enhancement, pattern analysis, image super-resolution and biometric recognition. She is an associate editor of *Applied Optics*. She is a recipient of 2007 and 2012 U.S. Army Research and Development Achievement Awards.



Jingyi Yu received the BS degree from Caltech, in 2000, and the PhD degree from MIT, in 2005. He is currently a professor and the executive dean of the School of Information Science and Technology, ShanghaiTech University. He is also affiliated with the Department of Computer and Information Sciences, University of Delaware. He has published more than 120 papers at highly refereed conferences and journals, and holds more than ten international patents on computational imaging. His research interests span a range of topics in computer vision and computer graphics, especially on computational photography and non-conventional optics and camera designs. He is a recipient of the NSF CAREER Award and the AFOSR YIP Award, and has served as an area chair of many international conferences including CVPR, ICCV, ECCV, ICCP, and NIPS. He is currently an associate editor of the *IEEE Transactions on Pattern Analysis and Machine Intelligence*, the *IEEE Transactions on Image Processing*, and Elsevier CVIU, and will be program chair of ICPR 2020 and IEEE CVPR 2021.



Jinwei Ye received the BEng degree in electrical engineering from the Huazhong University of Science and Technology, in 2009, and the PhD degree from the University of Delaware, in 2014. She is currently an assistant professor with the Division of Computer Science and Engineering, Louisiana State University (LSU). Before joining LSU, she worked as a researcher in the US Army Research Laboratory (ARL) and Canon U.S.A., Inc. Her research interests include computational photography, computer vision, and computer graphics.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.