

RESEARCH

Open Access



Snapshot depth–spectral imaging based on image mapping and light field

Xiaoming Ding¹, Liang Hu², Shubo Zhou³, Xiaocheng Wang¹, Yupeng Li¹, Tingting Han¹, Dunqiang Lu¹ and Guowei Che^{1*}

*Correspondence:
gwche@tjnu.edu.cn

¹ Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China

² CAS Key Laboratory of Computational Optics Imaging Technology, Aerospace - Information Research Institution, Beijing 100094, China

³ Institute of Information Science and Technology, Donghua University, Shanghai 201620, China

Abstract

Depth–spectral imaging (DSI) is an emerging technology which can obtain and reconstruct the spatial, spectral and depth information of a scene simultaneously. Conventionally, DSI system usually relies on scanning process, multi-sensors or compressed sensing framework to modulate and acquire the entire information. This paper proposes a novel snapshot DSI architecture based on image mapping and light field framework by using a single format detector. Specifically, we acquire the depth – spectral information in two steps. Firstly, an image mapper is utilized to slice and reflect the first image to different directions which is a spatial modulation processing. The modulated light wave is then dispersed by a direct vision prism. After re-collection, the sliced dispersed light wave is recorded by a light field sensor. Complimentary, we also propose a reconstruction strategy to recover the spatial depth – spectral hypercube effectively. We establish a mathematical model to describe the light wave distribution on every optical facet. Through simulations, we generate the aliasing raw spectral light field data. Under the reconstruction strategy, we design an algorithm to recover the hypercube accurately. Also, we make an analysis about the spatial and spectral resolution of the reconstructed data, the evaluation results conform the expectation.

Keywords: Depth estimation, Light field, Image mapper, Spectral imaging

1 Introduction

Multidimensional imaging technique can acquire plenty of optical information of scenes as much as nine dimensions $(x, y, z, \theta, \phi, \psi, \chi, \lambda, t)$, including three-dimensional (3D) spatial intensity distribution (x, y, z) , propagation polar angles (θ, ϕ) , propagations (ψ, χ) , wavelength (λ) for spectral intensity and time (t) [1]. The multidimensional imaging has a variety of applications in astrophysics, remote sensing, security, and biochemistry et al. [2–6]. Especially, 3D spatial distribution and one-dimensional (1D) spectral intensity is significant in target detection, recognition, tracking, scene classification and other computer vision fields et al. [7–10].

The 4D imaging processing (3D spatial and 1D spectral information) produces a mass of data, however, the detector that records the 4D data is usually a single, two or several format sensors. So that, in order to collect the whole 4D data, the input light distribution should be modulated onto the sensors to form the raw data, and then, particular

algorithms will be performed to recover the 4D data from the raw data. There are two categories to measure the depth information of targets. One is the active imaging strategy including structured light and Time-of-Flight (ToF) approaches, and the other is the passive imaging strategy including binocular vision and light field approaches [11, 12]. To measure the spectral characteristics of each spatial point in real time, the snapshot spectral imaging techniques emerges in recent years including direct measurement strategy and computational imaging strategy. The former one includes the approaches of image-division [13], aperture-division [14] and optical-path-division [15] formations, and the latter one includes the approaches based on computed tomography [16], compressed sensing (CS) [17] and Fourier transform [18].

In this paper, we propose a Snapshot Depth – Spectral Imager based on Image mapping and Light field (SDSIIL), in which the image mapper, dispersion element and light field sensor are used to modulated the input optical information and record the spatial – spectral light field simultaneously. Thereafter, a reconstruction strategy is introduced to recover the depth – spectral hypercube effectively. Three main contributions of this work can be summarized as: (1) A novel snapshot depth – spectral image framework is proposed. We design a compact optical structure to realize this framework with less optical element and fixed joint sensor; (2) A relative comprehensive mathematical model describing the imaging process of this optical system is established, and a simulation platform is performed in order to generate plenty and justified raw data. (3) An effective reconstruction method is proposed to realize the recovery of depth – spectral hypercube of the input scene, which verifies the feasibility of SDSIIL.

The remainder of this paper is arranged as follows. In Sect. 2, we introduce the related works about the depth–spectral imaging technique in recent years. In Sect. 3, the general principle of SDSIIL is discussed, and a mathematical model describing the distribution of the light wave on every optical facet is derived in detail. In Sect. 4, the reconstruction approach to recover the 4D depth-spectral hypercube data from the raw data is described. In Sect. 5, we perform simulations to generate the raw data record by detector according to the mathematical model, and reconstruct the spatial – spectral datacube at different depth, and also estimate a depth map of the input scene. In Sect. 6 and Sect. 7, we evaluate the spatial and spectral resolution respectively of this optical system by simulations, the results reveal that the resolution in spatial domain and spectral domain are both in accordance with theoretical expectations.

2 Related works

To obtain the 4D information of scenes in real time, several synthesized approaches with the combination of 3D imager and spectral imager are proposed. According to the depth imaging approaches, there are mainly three measurement strategies: binocular-vision formation, ToF formation and light field formation.

The binocular-vision-based technique refers to the strategy that use two imaging channels to calculate the depth map of object. One or each of the channels is spectral imaging, such as the 3D imaging spectroscopy proposed by Kim et al. [19], the cross-modal stereo system proposed by Wang et al. [20] and the spectral–depth imaging system based on deep learning reconstruction proposed by Yao et al. [21]. This binocular-vision

strategy usually needs two or more imaging channels and sensors, which always introduces non-synchronous problems between these sensors, especially for dynamic scenes.

The ToF-based technique refers to the strategy that use the ToF as the depth estimation channel combined with a spectral imaging channel, such as the snapshot compressive ToF + spectral imaging system proposed by Rueda-Chacon et al. [22]. The use of ToF complicates the entire system, and results in limitations in outdoor applications.

The light field-based technique refers to the strategy that uses light field camera usually combined with coded aperture spectral imager or other snapshot spectral imaging approach to record the angular information of the monochromatic light rays to calculate the depth map of objects at different wavelengths, such as the compressive spectral light field imager proposed by Marquez et al. [23], 3D compressive spectral integral imager proposed by Feng et al. [24] and the compressed spectral light field imager proposed by Liu [25]. Combining the image mapping spectrometer (IMS) and light field, Cai et al. proposed a hyperspectral light field imaging based on image mapping spectrometry by making the light field camera as the fore optics of IMS [26]. The light field of the scene is sampled, sliced and dispersed by IMS to record the entire spatial, angular and spectral information simultaneously. However, placing light field system before IMS makes the light field distribution of targets is sliced and separated by the strip mirrors. According to the previous researches, the image mapper has some intrinsic system errors such as “edge cutting” [27] and sliced image tilts [28], besides, the prism also introduces nonlinear dispersion [29]. All these issues make the calibration of microlens center projection on the sensor difficult and less accurate, which further influences the precision of depth estimation.

To overcome these problems, SDSIIL proposed in this paper uses a microlens array fixed joint with the sensor which will make the calibration more accurate. At the same time, the intermediate objective lens is unused in SDSIIL to make less optical elements needed, since the light field module is moved to the end of the system, which makes the structure more compact.

3 General principle and mathematical model

The system layout is shown in Fig. 1a. The fore optics consists of the pupil aperture and L1, which is a telecentric lens in imaging space to ensure the chief rays onto image mapper parallel with the optical axis. The image mapper slices, separates and reflects the input light to different directions as the slit mirrors containing different tilt angles (illustrated in Fig. 1b). Each direction relates to a sub spectrometer, and correlative slit mirrors on image mapper is equivalent to the input “slits” of each sub spectrometer. The prism is placed behind the collimating lens L2 to disperse the light. A reimaging lens array L3 is used to collect the dispersed light. The microlens array L4 is placed on the focal plane of L3. The sensor chip is just fixed on the focal plane of L4. As a result of this combination, the spatial–spectral light field of the input scene is recorded simultaneously, namely, both the spatial and angular information of each spectral channel is detected at the same time.

To describe the imaging formation mathematically, an imaging model based on the light propagation theory is established as followed. As shown in Fig. 2, the target is assumed to be a 3D object with the coordinate as (x_o, y_o, z_o) . An arbitrary point on the

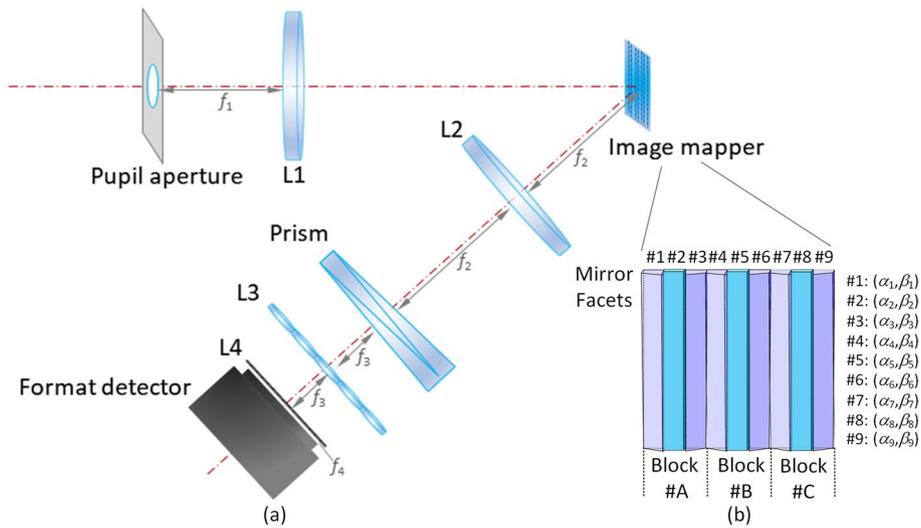


Fig. 1 The layout of SDSIIL. **a** is a schematic diagram of the optical path and structure, the pupil aperture and L1 form the telecentric fore optics to make the first image plane on image mapper; **b** is the structure schematic diagram of the image mapper, which only contains 3 blocks and 9 facets for simplicity. L2 is the collimating lens, the dispersion element usually should be a direct vision prism, such as Amici Prism. L3 is a reimaging lens array, and L4 is a microlens array combined with the format detector to form a light field sensor

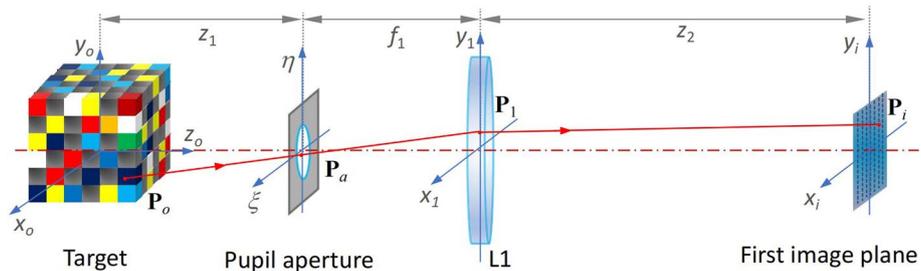


Fig. 2 The optical path and structure of the fore optics. P_o means an object point on the hypercube target. A random light ray from P_o intersects the pupil aperture and L1 at P_a and P_1 respectively. The light ray from P_1 intersects the image mapper at P_i

target is represented as $P_o(x_o, y_o, z_o)$ and the ideal object distance is assumed to be z_1 , namely, the distance between the ideal object plane and entrance pupil. The global original location is settled at the ideal object plane, and the propagation of light is the positive direction. The pupil aperture is at the front focal plane of L1, which means that the pupil aperture is the entrance pupil of the system. The focal length of L1 is f_1 . The coordinate of the pupil aperture plane is (ξ, η) .

The ideal object plane is conjugated with the first image plane meaning that the distance noted by z_2 between L1 and the first image plane should satisfy the Gauss Formula [30], i.e., $z_2 = (f_1 + z_1)f_1/z_1$. A random input ray from P_o can be defined as, $L(x_o, y_o, z_o, \xi, \eta)$, which means that the light L propagates through point $P_o(x_o, y_o, z_o)$ and point $P_a(\xi, \eta)$. According to the geometric principle [31], once the original position, propagation direction and distance are known, the end position is determined.

As propagation, $L(x_o, y_o, z_o, \xi, \eta)$ intersects L1 at $\mathbf{P}_1(x_1, y_1)$, and the \mathbf{P}_1 is considered as a vector, which can be calculated by,

$$\mathbf{P}_1 = \frac{f_1}{z_{R_o}} \hat{\mathbf{R}}_o + \mathbf{P}_a, \tag{1}$$

where $\hat{\mathbf{R}}_o$ represents the unit vector of $L(x_o, y_o, z_o, \xi, \eta)$, which can be calculated by $\hat{\mathbf{R}}_o = (\mathbf{P}_a - \mathbf{P}_o) / |\mathbf{P}_a - \mathbf{P}_o|$, z_{R_o} is the z value of $\hat{\mathbf{R}}_o$. We just calculate the paraxial rays and do not consider the aberration introduced by a real lens. So that, based on the Gauss Optics principle [30], if we know position and direction of the input light, we can calculate the direction of output light. In addition, the intersection joint noted by \mathbf{P}_i between the light and the first image plane is determined,

$$\mathbf{P}_i = \frac{z_2}{z_{R_1}} \hat{\mathbf{R}}_1 + \mathbf{P}_1, \tag{2}$$

where $\hat{\mathbf{R}}_1$ is the unit vector of the light through L1, which is given by,

$$\hat{\mathbf{R}}_1 = \frac{(f_1/z_{R_o})\hat{\mathbf{R}}_o + (0,0,f_1+z_1) - \mathbf{P}_1}{|(f_1/z_{R_o})\hat{\mathbf{R}}_o + (0,0,f_1+z_1) - \mathbf{P}_1|} \tag{3}$$

To acquire the ideal spatial–spectral light field distribution of the target, we tentatively put aside the influence of entire tilt of the image mapper to avoid some undesired phenomena, such as incline of the slit images, the “edge cutting” and the nonlinear space between adjacent slit images et. al. These problems need more deep research to be corrected by accurate calibrations. No consideration of the tilt angle means that the image mapper and each slit mirror is presumed to be just on the first image plane. As we know, based on the classical geometry optics theory, the mirror has no influence on the optical path difference (OPD), so that the optical axis after image mapper can be considered as a straight line together with the optical axis of the fore optics as shown in Fig. 3.

In this approximate case, the direction of the reflect light from image mapper just has the linear relationship with the tilt angle of each mirror. So that according to the Snell’s law [31], the reflected light unit vector from point $\mathbf{P}_i(x_p, y_i)$ is

$$\hat{\mathbf{R}}_i = \mathbf{M}_\alpha \times \mathbf{M}_\beta \times \hat{\mathbf{R}}_1, \tag{4}$$

where

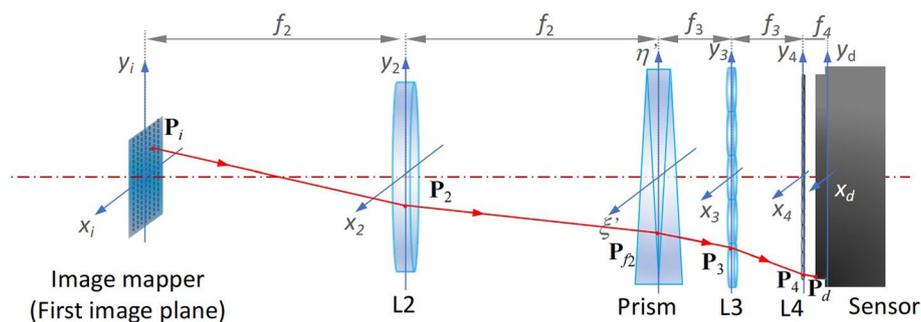


Fig. 3 The optical path and structure of the spectrometer and light field sensor. The light ray reflected from image mapper intersects L2, the prism, L3, L4 and the sensor at $\mathbf{P}_2, \mathbf{P}_{2'}, \mathbf{P}_3, \mathbf{P}_4$ and \mathbf{P}_d respectively

$$M_\alpha = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha(P_i) & -\sin \alpha(P_i) \\ 0 & \sin \alpha(P_i) & \cos \alpha(P_i) \end{bmatrix} \tag{5}$$

and

$$M_\beta = \begin{bmatrix} \cos \beta(\mathbf{P}_i) & 0 & \sin \beta(\mathbf{P}_i) \\ 0 & 1 & 0 \\ -\sin \beta(\mathbf{P}_i) & 0 & \cos \beta(\mathbf{P}_i) \end{bmatrix}. \tag{6}$$

In the above Eq. (5) and (6), the tilt angles (α, β) of the slit mirrors on image mapper are determined by the position of the point $\mathbf{P}_i(x_i, y_i)$. The light exiting from image mapper intersects the collimating lens L2 at $\mathbf{P}_2(x_2, y_2)$. Based on the same theory used in Eq. (1) and Eq. (2), \mathbf{P}_2 is given by,

$$\mathbf{P}_2 = \frac{f_2}{z_{R_i}} \hat{\mathbf{R}}_i + \mathbf{P}_i. \tag{7}$$

According to Gauss Optics principle, the intersection point noted by \mathbf{P}_{f_2} of light and the focal plane of L2 can be calculated as the $\hat{\mathbf{R}}_i$ and \mathbf{P}_2 is known, which is $\mathbf{P}_{f_2} = (f_2/z_{R_i}) \hat{\mathbf{R}}_i + (0, 0, f_1 + z_1 + z_2 + 2f_2)$. In addition, the unit vector from \mathbf{P}_2 to \mathbf{P}_{f_2} is noted by $\hat{\mathbf{R}}_2 = (\mathbf{P}_{f_2} - \mathbf{P}_2) / |\mathbf{P}_{f_2} - \mathbf{P}_2|$.

As the reflection and slicing function of the image mapper, a sub pupil array is formed at the back focal plane of L2 which is conjugate with the pupil aperture. The prism is placed on this plane to disperse the light. As mentioned before, any nonideal issues caused by the prism such as nonlinear dispersion, field distortion and scattering et al. are not considered in this model. We just focus the research on the feasibility of the principle in ideal conditions. The prism introduces an angular dispersion according to the wavelength. The angular dispersion is assumed to be linear with the wavelength λ which is represented by $\phi(\lambda)$. In an actual system, the $\phi(\lambda)$ is usually more complicated and nonlinear with wavelength λ . In this paper, we just discuss the ideal condition. We choose a center wavelength λ_c , and the light at λ_c has no change in propagation direction through the prism. As described above, λ is given by

$$\theta(\lambda) = D(\lambda - \lambda_c), \tag{8}$$

where D is a constant which is defined as $D = d\theta/d\lambda$.

After the dispersion by prism, the light propagation direction changes. In mathematics, the dispersion process actually performs an angle rotation on the unit vector of the rays in one direction in the 3D coordinates, which can be described as,

$$\hat{\mathbf{R}}_p = \mathbf{M}_p \times \hat{\mathbf{R}}_2, \tag{9}$$

where \mathbf{M}_p is the rotation transfer matrix. If we assume the light is dispersed along x direction, then the rotation axis is y direction. \mathbf{M}_p is given by

$$\mathbf{M}_p = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \tag{10}$$

After the prism, a reimaging lens array L3 is used to collect the dispersed light and reimage on the image plane of each sub spectrometer. The distance between the prism and L3 is f_3 which is the focal length of each sub lens of L3 ensuring the telecentric in imaging space of each sub spectrometer. The intersect point of the light with L3 is represented by \mathbf{P}_3 ,

$$\mathbf{P}_3 = \frac{f_3}{z_{R_p}} \hat{\mathbf{R}}_p + \mathbf{P}_{f_2} \tag{11}$$

Since the L3 contains several sub lenses, the sub lenses are placed one by one regularly in 2D directions with the center distance (the same with diameter of each sub lens) d_{l_3} between adjacent sub lenses. We assume that \mathbf{P}_3 is located on the N_o . (m, n) sub lens, which means that sub lens (m, n) collects this light ray. The intersect point on its focal plane is noted by $\mathbf{P}_{f_3}(m, n)$,

$$\mathbf{P}_{f_3}(m,n) = \frac{f_3}{z_{R_p}} \hat{\mathbf{R}}_p + (c_m, c_n, f_1 + z_1 + z_2 + 2f_2 + 2f_3) \tag{12}$$

where (c_m, c_n) is the center location of N_o . (m, n) sub lens. The unit vector of from \mathbf{P}_3 to $\mathbf{P}_{f_3}(m, n)$ is determined by the intersect location and the related sub lens, which is noted by $\hat{\mathbf{R}}_{3m,n} = (\mathbf{P}_{f_3}(m, n) - \mathbf{P}_3) / (|\mathbf{P}_{f_3}(m, n) - \mathbf{P}_3|)$.

On the image plane of the spectrometer, a light field sensor consisting of a microlens array and a large format detector is placed to record the angular information of the spectral rays. As a result, any ray L determined by a spatial coordinate $(x_o, y_o, z_o, \xi, \eta)$ and a spectral coordinate (λ) can be measured by the detector, including both the spectral intensity and direction information. The microlens array L4 is placed on the image plane after L3 and the format detector is fixed behind the microlens array with a distance of f_4 which is the focal length of each microlens. The joint on L4 is given by,

$$\mathbf{P}_4 = \frac{f_3}{z_{R_{3m,n}}} \hat{\mathbf{R}}_{3m,n} + \hat{\mathbf{P}}_3 \tag{13}$$

Actually, the value of \mathbf{P}_4 is the same with $\mathbf{P}_{f_3}(m,n)$, however, we derive the formula for the integrity of the entire imaging model. The unit vector after L4 is,

$$\hat{\mathbf{R}}_{4s,t} = \frac{\mathbf{P}_{f_4}(s,t) - \mathbf{P}_4}{|\mathbf{P}_{f_4}(s,t) - \mathbf{P}_4|} \tag{14}$$

where $\mathbf{P}_{f_4}(s,t) = (f_4/z_{R_{3m,n}}) \hat{\mathbf{R}}_{3m,n} + (c_s, c_t, f_1 + z_1 + z_2 + 2f_2 + 2f_3 + f_4)$, and (c_s, c_t) is the center location of N_o . (s, t) microlens. Propagating through the microlens, the final intersect point of the light on the detector is represent by $\mathbf{P}_d(x_d, y_d, z_d)$,

$$\mathbf{P}_d = \frac{f_4}{z_{R_{4s,t}}} \hat{\mathbf{R}}_{4s,t} + \mathbf{P}_4 \tag{15}$$

which means that the light $L(x_o, y_o, z_o, \xi, \eta, \lambda)$ passing through the whole system shoots on the format detector at the location $\mathbf{P}_d(x_d, y_d, z_d)$. As discretely sampled by the pixels on the detector, the pixel index (i_d, j_d) of $\mathbf{P}_d(x_d, y_d, z_d)$ is given by

$$(i_d, j_d) = [(x_d, y_d)/d_p] + (N_{dx} - 1, N_{dy} - 1)/2 \tag{16}$$

where d_p is the size of a single pixel, and the (N_{dx}, N_{dy}) is the number of pixels of the detector in the 2D directions. In this formula, (N_{dx}, N_{dy}) are both considered to be odd. As a result, the gray level $I(i_d, j_d)$ of the pixel (i_d, j_d) should consider the rays from the entire object field and through the entire pupil aperture, which is given by,

$$I(i_d, j_d) = \iiint \iiint L(x_d, y_d, z_d, \xi, \eta, \lambda) dx_d dy_d dz_d d\xi d\eta d\lambda. \tag{17}$$

4 Reconstruction method for SDSIIL data

The collected raw data on detector records the 2D spatial, 2D angular and 1D spectral information, so that the digital refocus method for light field data should be considered and improved in this processing according to the spectral distribution. Based on the digital refocus principle in light field camera, we should establish the reconstruction approach in the SDSIIL situation considering the influence of dispersion element on the rays' direction.

When considering the dispersion, as shown in Fig. 4, the rays should be calculated in two directions separately. Based on the model established above, the light is dispersed along the x direction and the ideal image plane is on the microlens array (L4). So that, the digital refocusing is performed on each single slit spectral light field image separately. The ray L' intersects the refocused plane S' at point $\mathbf{P}'_d(s', t')$ and intersects L4 at point $\mathbf{P}_d(s, t)$. The distance between S' and L4 is Δl , so that the gray value at $\mathbf{P}'_d(s', t')$ is given by

$$I'(y'_d) = \int L'(u, y'_d) du \tag{18}$$

Namely, $I'(y'_d) = \int L(u, y_d) du$, where $L(u, y_d)$ is a ray recorded by the light field sensor. According to the similar triangle theorem, $y_d = y'_d/a + u(1-1/a)$, and Eq. (18) is transformed to,

$$I'(y'_d) = \int L[u, y'_d/a + u(1-1/a)] du \tag{19}$$

For the dispersion direction, we can consider the different spectral light as different field light in spatial domain. As shown in Fig. 4, according to the prism model (Eq. (9)), the relative spatial location x_d (relative to the center wavelength) is related with the

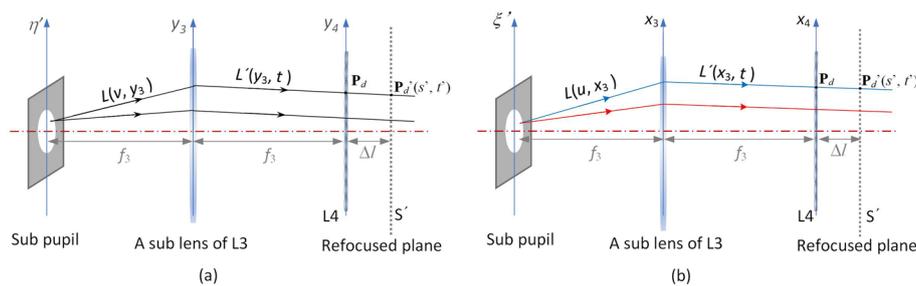


Fig. 4 The optical layout for digital refocusing of each slit light field image. **a** is illustration in y direction, and **b** is the illustration in x direction, i.e., the dispersion direction. The image is assumed to be refocused on S' plane, and the ray L' intersects plane S' at point \mathbf{P}'_d

wavelength λ . So that $x_d = (\lambda - \lambda_c) Df_3 + x_{dc}$, where x_{dc} is the location in x direction of the center wavelength. The intensity at P'_d point for the x direction is

$$I'(x'_d) = \int L(u, x'_d) du \tag{20}$$

Namely, $I'(x'_d) = \int L(u, x_d) du$, where $L(u, x_d)$ is a ray recorded by the light field sensor. According to the similar triangle theorem, $x_d = x'_d/a + u(1 - 1/a)$, then,

$$l = [x'_d/a + u(1 - 1/a) + \lambda_c Df_3 - x_{dc}] / (Df_3) \tag{21}$$

and Eq. (20) is transformed to

$$I'(x'_d) = \int L[u, x'_d/a + u(1 - 1/a)] du \tag{22}$$

Through the method mentioned above, we can refocus and recovery any spectral images from the raw data at a chosen depth. Combining with the definition evaluation for refocused images, we can scan along the depth direction in object space in order to estimate the depth map of the target. This is a classical and direct method to estimate the depth information, more accurate and advanced strategies to reconstruct the 3D spatial information through light field data have been proposed and reported [32, 33]. The estimation precision is determined by the angular resolution for each microlens. With the same value of the relative diameter (or the same NA), more pixels covered by the microlens usually means more distinguishable angular and higher depth resolution.

5 Simulations

The simulations were performed using MATLAB 2020a software. The input scene for the simulations should be a 4D data with high resolution in 3D spatial domain and 1D spectral domain. To our best knowledge, no such standard database can be used directly. So that we make use of the hyperspectral database ICVL [34] to generate the depth distribution manually. As shown in Fig. 5a, this is a 3D datacube in ICVL after proper image resizer and interpolation to form a size of $825 \times 825 \times 61$. However, no

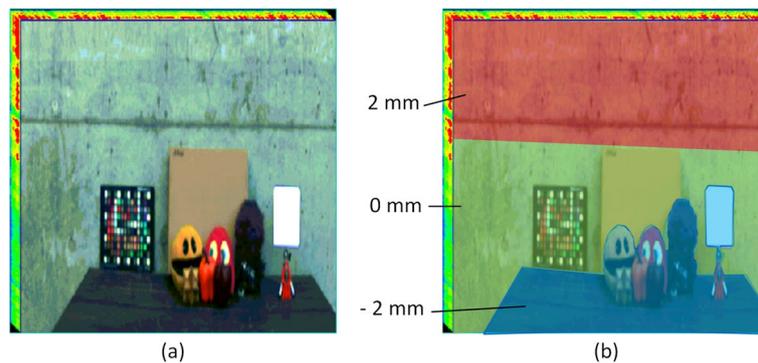


Fig. 5 The input datacube, **a** the datacube with size of $825 \times 825 \times 61$, **b** the depth value of each region, the blue region is -2 mm depth, the green one is 0 mm depth and the red one is 2 mm depth

depth data. The depth range that a light field camera can estimate is limited by the microlens array parameters, which is given by

$$\left[\frac{(d_p z - d_4 f_4) f_3}{d_p (f_3 - z) + d_4 f_4}, \frac{(d_p z + d_4 f_4) f_3}{d_p (f_3 - z) - d_4 f_4} \right] \tag{23}$$

where z is the distance between the L2 and the microlens array ($z = f_2 + f_3$), d_4 is the diameter of each microlens. In this paper, the depth range is about from -2 mm to 2 mm. So that we marked some region (shown in Fig. 5b), Blue region) with the depth of -2 mm, some (Green region) is marked 0 mm and the rest (Red region) is 2 mm. When simulating, the region with 0 mm depth is placed at the ideal object plane. The system parameters are shown in Table 1.

The imaging simulation result is shown in Fig. 6, which is the raw data collected by the detector. We find that 5×5 sub images distribute regularly according to the $M \times N$ tilt angles of mirrors. In each sub image, 11 dispersed light field slit images are recorded as shown in Fig. 6b, and each microlens covers 11×11 pixels (shown in Fig. 6f), which means 11×11 rays with different angular directions are measured for each spatial point.

Refocus these light field slit images, and combine them together by the order of object field distribution to get the datacube at a certain depth. As we know that the input scene has 3 marked regions with different depths, so we refocus them in the depth of -2 mm, 0 mm and 2 mm in object space, respectively. In each depth, 21 spectral images are recovered and form a datacube with 2D spatial and 1D spectral information, as shown in Fig. 7.

From the reconstructed datacube, we can find that when refocusing on depth -2 mm, 0 mm, 2 mm, the region 1, 2, 3 are the most distinct area, respectively, which is consistent with the input data. Using the method proposed by Tao [35], we estimate the depth map from the raw data, the result is shown in Fig. 8. We can find that although some singular values exist in the estimated depth map for the reason of spectral mixing and stereo matching errors, the result mostly represents the depth distribution of the input sense.

To verify the measurement ability of spectral information, we choose some representative points to plot the spectral curves, as shown in Fig. 9. To avoid the spatial

Table 1 The system parameter for simulations

| Parameter | Value | Parameter | Value |
|-----------|---|------------------------|------------------------|
| z_1 | 200 mm | $M \times N$ | 5×5 |
| f_1 | 200 mm | b | 0.19 mm |
| f_2 | 20 mm | d_p | 0.0045 mm |
| f_3 | 10 mm | $N_{dx} \times N_{dy}$ | $22,000 \times 22,000$ |
| D_{ape} | 55 mm | D | 0.3465 rad/ μ m |
| d_3 | 19.8 mm | Spectral range | 450 nm–750 nm |
| d_4 | 0.0495 mm | Bands | 21 |
| α | $0, 0, 0, \pm 22.76, \pm 22.37, \pm 23.60, \pm 28.81, \pm 32.93, \pm 32.05, \pm 31.62, \pm 32.97, \pm 38.20, \pm 41.79$ | | |
| β | $0, \pm 22.36, \pm 31.60, \pm 10.39, \pm 18.04, \pm 3.52, \pm 13.59, \pm 19.55, \pm 12.84, \pm 18.54, \pm 7.84, \pm 3.32, \pm 6.07$ | | |

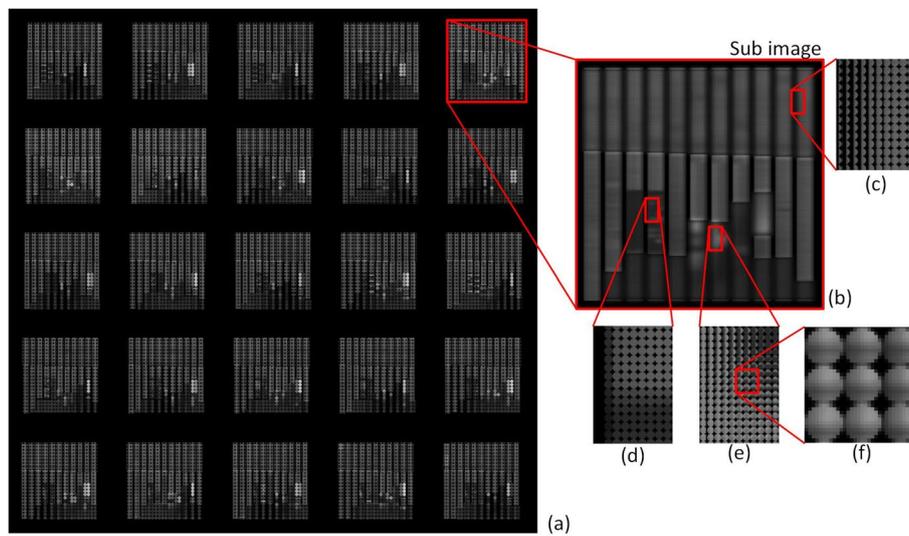


Fig. 6 The simulation raw data collected by the detector of the SDSIIL, **a** is the entire raw data, **b** is one of the sub images, **c** is the enlarged view of the -2 mm depth region, **d** is the enlarged view of the 0 mm depth region, **e** is the enlarged view of the 2 mm depth region, **f** is the enlarged view of sub pupil image behind some microlens

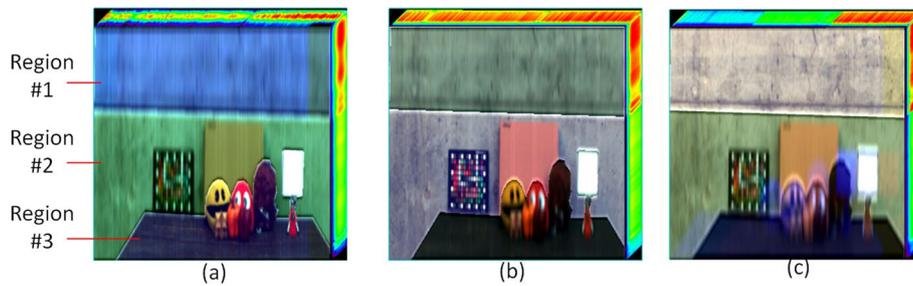


Fig. 7 The reconstructed datacube, **a** is the refocused datacube at depth of -2 mm, **b** is the refocused datacube at depth of 0 mm, **c** is the refocused datacube at depth of 2 mm



Fig. 8 The estimated depth map of the reconstructed data

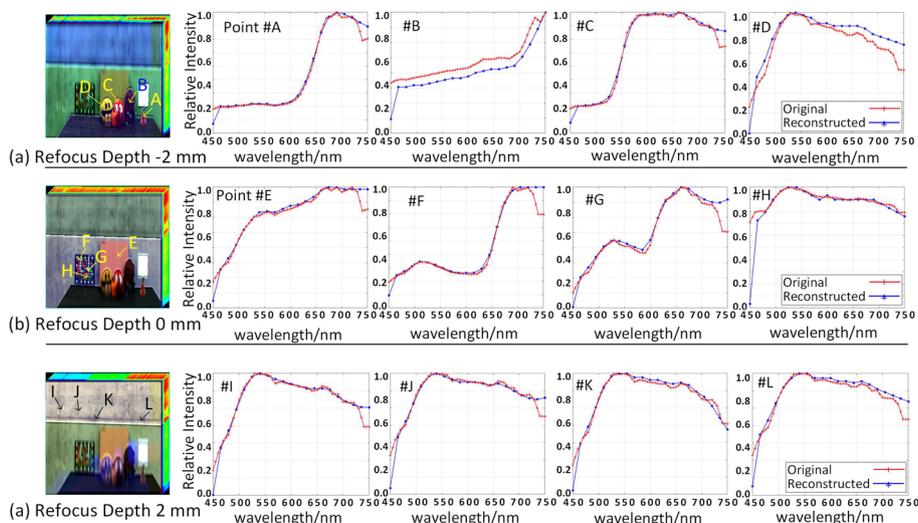


Fig. 9 The spectral curves of some representative points. (a), (b) and (c) are the reconstructed datacube and spectral curves at refocused depth -2 mm, 0 mm and 2 mm, respectively

Table 2 The evaluations for the spectral curves

| Points | SA | RQE |
|--------|--------|--------|
| A | 0.0547 | 0.1188 |
| B | 0.0750 | 0.2102 |
| C | 0.0475 | 0.0952 |
| D | 0.0531 | 0.1195 |
| E | 0.0503 | 0.0988 |
| F | 0.0844 | 0.1858 |
| G | 0.0781 | 0.1411 |
| H | 0.0727 | 0.1528 |
| I | 0.0492 | 0.0897 |
| J | 0.0562 | 0.1011 |
| K | 0.0418 | 0.0758 |
| L | 0.0542 | 0.1033 |

and spectral mixing between adjacent pixels, we just choose the refocused regions to plot the spectral curves.

Comparing with the ground true in the database, we calculate the Spectral Angle (SA) and Relative Spectral Quadratic Error (RQE) [36] of each spectral curve. The values are listed in Table 2. The SA and RQE are both the evaluation functions that judge the similarity of two spectral curves. The SA value and RQE value being smaller to 0 mean that the two curves are more similar. The evaluation results verify that the reconstructed spectral curve is just similar to the ground true and the spectral information measured is accurate. The results mentioned above all verify that the principle of the system proposed in this paper is feasible. The refocused and recovery method performed is effective.

6 Analysis of spatial resolution

The spatial resolution in image space of SDSIIL is determined by both the width (b') of slit mirror image on L4 and the microlens diameter (d_4), which is given by

$$R_s = \frac{1}{2\max(b',d_4)} \tag{24}$$

In this paper, the width of slit mirror image is the same with the diameter of microlens, so that the spatial resolution R_s is about $1/(2d_4) = 1/(2 \times 0.0495) = 10.1 \text{ lp/mm}$ in theory. We conducted simulations for evaluation. The USAF1951 image is used as the input sense. Since the object field is $13.6 \text{ mm} \times 13.6 \text{ mm}$, the input image should contain the group 3 element 2–4 in USAF1951 image. We scan the depth range from -18 to 18 mm with step 1 mm at wavelength 600 nm to generate the raw data. The monochromatic reconstructed images are shown in Fig. 10a, and the resolvable lines are checked from the cross view of the images.

We can find that the spatial resolution decreases with the depth of input sense departure from the zero plane. Although the light field data can be refocused at different depth, the definition of the spatial details decreases when the absolute value of the target depth increases. Especially, beyond the -2 mm and 2 mm , resolution decreases rapidly. Synthesizing the spatial resolution (Fig. 10b) curve and the average gradient curve (Fig. 10c), the depth range $[-2 \text{ mm}, 2 \text{ mm}]$ in the object space is the valid depths which can be recovered with high spatial accuracy.

7 Analysis of spectral resolution

In order to acquire the spectral resolution at each spectral band, we need to calculate the spectral response function (SRF) of SDSIIL. Since the raw data of this system is not the direct measurement of the scene, the traditional measurement method for SRF is not suitable for SDSIIL. After scanning the wavelength range to get the response at a certain

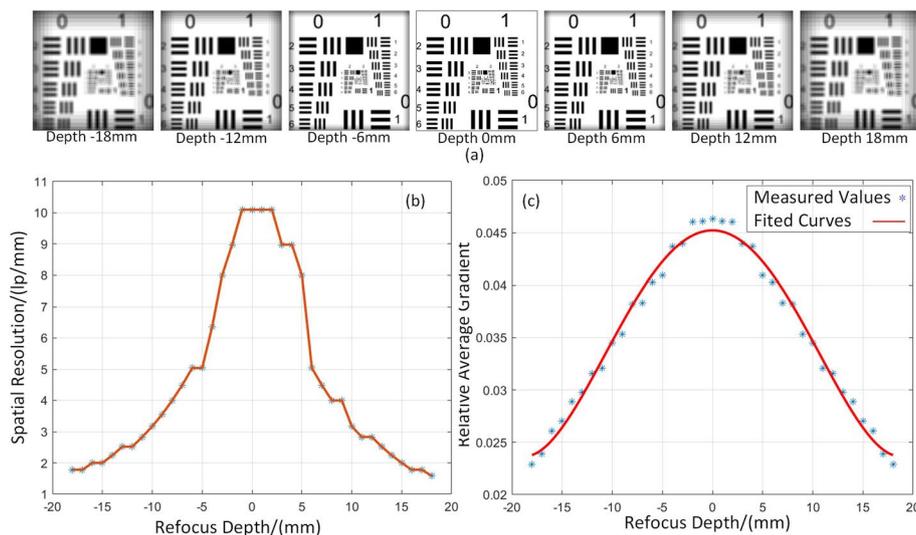


Fig. 10 The evaluations of refocused images at different depths, **a** some of the refocused images at different depths, **b** the spatial resolution measured by the USAF1951 of the refocused images, **c** the relative average gradient of the refocused images and the fitted curves

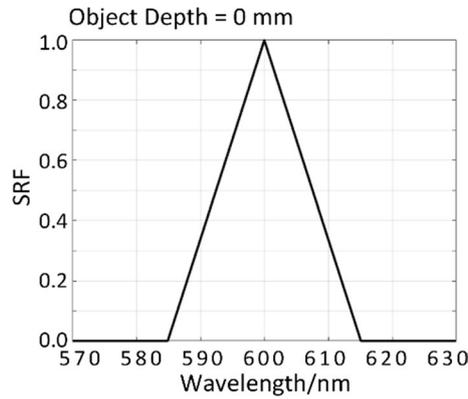


Fig. 11 The theoretical curve of SRF with object depth 0 mm

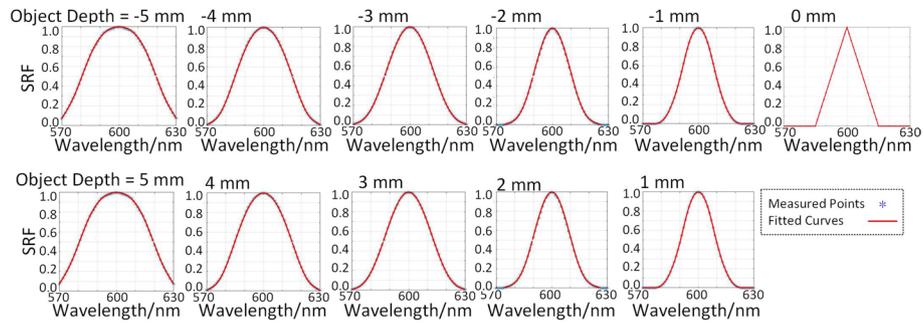


Fig. 12 The SRFs of refocused images with different object depths at center wavelength 600 nm. The object depth of each figure is noted on the top. The object depth of the figures in first line are -5 mm, -4 mm, -3 mm, -2 mm, -1 mm and 0 mm, respectively. The object depth of the figures in second line are 5 mm, 4 mm, 3 mm, 2 mm and 1 mm, respectively

spectral band, we should reconstruct the spectral slit image firstly, and then calculate the SRF at the spectral band. Since the model we established above does not consider the nonlinear dispersion or the field distortion et. al., the SRF should be consistent along the spectral bands. So that we choose the center wavelength (600 nm) at center object field for consideration. In theory, the ideal SRF is determined by the resolving power of the dispersive element (represented by R , $R = D \times f_3$), slit width (b), the magnification of the spectrometer (M_{spec}), the detector element width (d_p) and the f-number at the exit pupil plane ($f / \#$). If we only consider the geometric SRF without the influence of diffraction, the geometric SRF is given by

$$g(\lambda) = \text{rect} \left[\frac{R(\lambda - \lambda_c)}{M_{spec} b} \right] * \text{rect} \left[\frac{R(\lambda - \lambda_c)}{d_p} \right] \tag{25}$$

We substitute the parameters in Table 1 to Eq. (27) to calculate the theoretical SRF shown in Fig. 11. FWHM of this theoretical SRF is 15 nm.

The simulations are conducted under different unfocused situations, including -5 mm to 5 mm with step of 1 mm along z axis in the object space. We refocus the raw data, and generate the SRFs, as shown in Fig. 12.

We can find that the curve above noted by object depth 0 mm is similar with the theoretical curve in Fig. 11, which verifies the mathematical model. The FWHM of the SRF

in Fig. 12 becomes larger when the depth of the object increases. In addition, we calculate the FWHM of the SRF at different object depths from -5 mm to 5 mm with dense sampling. Figure 13 shows the simulation results.

The figure above indicates that the SRF becomes diffusive with the enlargement of the object depth. When the object depth is larger than 2 mm, the FWHM of SRF becomes more than 20 nm, which makes the spectral resolution become lower than 30%. The variation of the object depth has influence on the SRF after reconstruction. Even though the reconstruction processing refocuses the light field image on the certain image plane to be most distinct, the SRF still becomes diffusive compared with that the object depth is 0 mm. The phenomenon may be caused by the angular light sampled by the adjacent microlens when the target is not on the ideal object plane. The above analysis further indicates that the depth range $[-2 \text{ mm}, 2 \text{ mm}]$ in the object space is the valid depths which can be recovered with relatively high spectral accuracy.

8 Conclusion

This paper proposed a SDSIIL system to capture the depth–spectral information of input scene within a single snapshot via image mapping and light field framework. SDSIIL performs the image slicing through an image mapper to reflect different parts of the first image to relevant directions. A direct vision prism is used to disperse the mapped and collimated light. To measure the angular information at different wavelengths, a microlens array is placed before the detector to form a light field sensor. Under the digital refocused framework, we can make depth estimate for different spectral images after remapping the spectral sliced light field. As potential advantages in contrast with state-of-the-art systems that may rely on multi-sensors, coded mask or filter array, SDSIIL benefits from the high light throughput without encoding element or filters, broad applicability for sparse or non-sparse targets, high reliability and compact structure with a single imaging sensor. The mathematical model was established to describe the light propagation process through the whole system and the light intensity distribution on the image plane. Based on the model, the simulations were conducted to acquire the raw

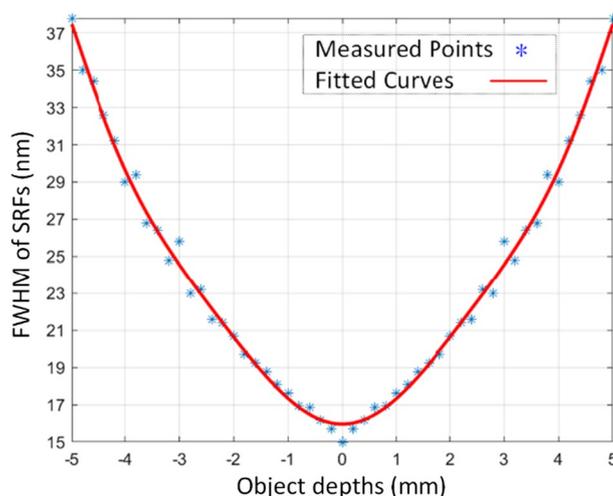


Fig. 13 FWHM of SRF at different object depths

data of spectral light field, and the reconstruction method were performed to recover the spectral images with depth estimation. We used the SA, RQE to evaluate the reconstructed spectral curves, and generated the depth map to compare with the original data. The results indicated that accurate reconstructed depth–spectral information can be recovered from the raw data, which confirmed that our proposed system SDSIIL represents an effective and efficient approach to obtain depth–spectral images of targets in a single snapshot of the sensor.

In addition, the spatial and spectral resolution were analyzed from theory and simulations. The results showed that even though the images were reconstructed on the refocused plane, the spatial and spectral resolution decreased, as the input scene deviated from the zero-depth plane. The variation of the object depth has influence on the spatial and spectral resolution after reconstruction. The analysis further indicates that the depth range $[-2\text{ mm}, 2\text{ mm}]$ in the object space is the valid depths which can be recovered with relatively high spatial and spectral accuracy.

In future work, we plan to make a more accurate mathematical model to describe the aberrations introduced by practical elements in order to help to develop more applicable and efficient reconstruction strategies. A prototype will be established to test the effectiveness of the theory model and the reconstruction method.

Abbreviations

| | |
|--------|---|
| DSI | Depth-spectral imaging |
| 3D | Three-dimensional |
| 1D | One-dimensional |
| ToF | Time-of-flight |
| CS | Compressed sensing |
| IMS | Image mapping spectrometer |
| SDSIIL | Snapshot depth-spectral imager based on image mapping and light field |
| OPD | Optical path difference |
| SA | Spectral angle |
| RQE | Relative spectral quadratic error |
| SRF | Spectral response function |
| FWHM | Full width at half maxima |

Acknowledgements

Not applicable.

Author contributions

XD proposed the framework of the whole ideal, structure of the model and the algorithm; LH and SZ helped to perform the simulations and conduct the analysis of the results. XW, YL and TH provided the relative data, participated in the conception, and helped to revise the manuscript; CG provide the framework and application background of this project. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) (62001328, 62001327, 61901301), the Natural Science Foundation of Tianjin Municipality (20JCYBJC00300), the Scientific Research Project of Tianjin Educational Committee (2021KJ182).

Availability of data and materials

Please contact the authors for data requests.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 20 November 2022 Accepted: 23 January 2023

Published online: 08 February 2023

References

1. L. Gao, L.V. Wang, A review of snapshot multidimensional optical imaging: measuring photon tags in parallel. *Phys. Rep.* **616**, 1–37 (2016)
2. A.F.H. Goetz, G. Vane, J.E. Solomon, B.N. Rock, Imaging spectrometry for earth remote sensing. *Science* **228**(4704), 1147–1153 (1985)
3. J. Braga, Coded aperture imaging in high-energy astrophysics. *Publ. Astron. Soc. Pacific* **132**(1007), 12001 (2020)
4. R.R. Iyer et al., Full-field spectral-domain optical interferometry for snapshot three-dimensional microscopy. *Biomed. Opt. Express* **11**(10), 5903 (2020)
5. J. Huang, K. Liu, M. Xu, M. Perc, X. Li, Background purification framework with extended morphological attribute profile for hyperspectral anomaly detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **10**(14), 8113–8124 (2021)
6. K. Liu, Z. Jiang, M. Xu, M. Perc, X. Li, Tilt correction toward building detection of remote sensing images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **14**, 5854–5866 (2021)
7. G. Cheng, P. Zhou, J. Han, Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **54**(12), 7405–7415 (2016)
8. Van Nguyen, H., Banerjee, A. and Chellappa, R., Tracking via object reflectance using a hyperspectral video camera. 2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. - Work. CVPRW 2010 44–51 (2010).
9. J.M. Ramirez, H. Arguello, spectral image classification from multi-sensor compressive measurements. *IEEE Trans. Geosci. Remote Sens.* **58**(1), 626–636 (2020)
10. F.J. Rodríguez-Pulido, B. Gordillo, F.J. Heredia, M.L. González-Miret, CIELAB – spectral image MATCHING: An app for merging colorimetric and spectral images for grapes and derivatives. *Food Control* **125**, 108038 (2021)
11. F. Liu et al., Binocular light-field: imaging theory and occlusion-robust depth perception application. *IEEE Trans. Image Process.* **29**, 1628–1640 (2020)
12. Ding, X. et al., Snapshot compressive spectral - depth imaging based on light field. *EURASIP J. Adv. Signal Process.* **2022**(1), (2022).
13. M.E. Pawlowski, J.G. Dwight, T.-U. Nguyen, T.S. Tkaczyk, High performance image mapping spectrometer (IMS) for snapshot hyperspectral imaging applications. *Opt. Express* **27**(2), 1597 (2019)
14. C. Yu et al., Microlens array snapshot hyperspectral microscopy system for the biomedical domain. *Appl. Opt.* **60**(7), 1896 (2021)
15. S.E. Headland, H.R. Jones, A.S.V. D'Sa, M. Perretti, L.V. Norling, Cutting-edge analysis of extracellular microparticles using mainstream imaging flow cytometry. *Sci. Rep.* **4**(1), 1–10 (2014)
16. F.S. Oktem, F. Kamalabadi, J.M. Davila, A parametric estimation approach to instantaneous spectral imaging. *IEEE Trans. Image Process.* **23**(12), 5707–5721 (2014)
17. Meng, Z., Yu, Z., Xu, K. and Yuan, X., Self-supervised neural networks for spectral snapshot compressive imaging. *Proc. IEEE Int. Conf. Comput. Vis.* 2602–2611 (2021).
18. L.C. Petre, V. Damian, Snapshot interferometric multispectral imaging using deconvolution and colorimetric fit. *Opt. Laser Technol.* **111**, 100–109 (2019)
19. M.H. Kim et al., 3D imaging spectroscopy for measuring hyperspectral patterns on solid objects. *ACM Trans. Graph.* **31**(4), 1–11 (2012)
20. L. Wang, Z. Xiong, G. Shi, W. Zeng, F. Wu, Simultaneous depth and spectral imaging with a cross-modal stereo system. *IEEE Trans. Circuits Syst. Video Technol.* **28**(3), 812–817 (2018)
21. M. Yao, Z. Xiong, L. Wang, D. Liu, X. Chen, Spectral-depth imaging with deep learning-based reconstruction. *Opt. Express* **27**(26), 38312 (2019)
22. H. Rueda-Chacon, J.F. Florez-Ospina, D.L. Lau, G.R. Arce, Snapshot compressive ToF+spectral imaging via optimized color-coded apertures. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**(10), 2346–2360 (2020)
23. M. Marquez, H. Rueda-Chacon, H. Arguello, Compressive spectral light field image reconstruction via online tensor representation. *IEEE Trans. Image Process.* **29**, 3558–3568 (2020)
24. W. Feng et al., 3D compressive spectral integral imaging. *Opt. Express* **24**(22), 24859 (2016)
25. Liu, X. et al. Multi-information fusion depth estimation of compressed spectral light field images. *Imag. Appl. Opt. Congress, OSA Technical Digest, paper DW1A.2.*, (2020).
26. Q. Cui, J. Park, R.T. Smith, L. Gao, Snapshot hyperspectral light field imaging using image mapping spectrometry. *Opt. Lett.* **45**(3), 772–775 (2020)
27. R.T. Kester, L. Gao, T.S. Tkaczyk, Development of image mappers for hyperspectral biomedical imaging applications. *Appl. Opt.* **49**(10), 1886–1899 (2010)
28. A. Liu, L. Su, Y. Yuan, X. Ding, Accurate ray tracing model of an imaging system based on image mapper. *Opt. Express* **28**(2), 2251 (2020)
29. L. Gao, Correction of vignetting and distortion errors induced by two-axis light beam steering. *Opt. Eng.* **51**(4), 043203 (2012)
30. W.J. Smith, *The Design of Optical Systems: General* (Modern Optical Engineering, McGraw-Hill, USA, 2000)
31. M. Born and E. Wolf, *Principles of Optics* (Pergamon, 1980), chap. 3.
32. S. Zhu, A. Lai, K. Eaton, P. Jin, L. Gao, On the fundamental comparison between unfocused and focused light field cameras. *Appl. Opt.* **57**(1), A1 (2018)
33. Y. Li, Q. Wang, L. Zhang, G. Lafruit, A lightweight depth estimation network for wide-baseline light fields. *IEEE Trans. Image Process.* **30**, 2288–2300 (2021)

34. Alvarez-Gila, A., Van De Weijer, J. and Garrote, E., Adversarial networks for spatial context-aware spectral image reconstruction from RGB. Proc. - 2017 IEEE Int. Conf. Comput. Vis. Work. ICCVW 2017 2018-Janua, 480–490 (2017).
35. Tao, M. W., Hadap, S., Malik, J. and Ramamoorthi, R., Depth from combining defocus and correspondence using light-field cameras. Proc. IEEE Int. Conf. Comput. Vis. 673–680 (2013).
36. B. Aiazzi et al., Tradeoff between radiometric and spectral distortion in lossy compression of hyperspectral imagery. Math. Data/Image Coding, Compress., Encrypt. VI, Appl. **5208**, 141–152 (2004)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
