High spectral resolution imaging based on dual-camera system with filter wheel (November 2021)

Jiakai Yu1, Jufeng Zhao1,2, Junjie Zhu1 and Guangmang Cui1,2
1School of Electronics and Information, Hangzhou Dianzi University, Hangzhou 310018, China
2 Zhejiang Provincial Key Lab of Equipment Electronics, Hangzhou Dianzi University, Hangzhou 310018, China
Corresponding author: Jufeng Zhao (e-mail: dabaozj@hdu.edu.cn).

This work was supported in part by Zhejiang Provincial Natural Science Foundation of China under Grant No. LY22F050002 and No.LGF20F050003, in part by Zhejiang association for science and technology YuCai project No.SKX201901, and in part by National Natural Science Foundation of China under No.61805063.

ABSTRACT High spectral and high spatial resolution imaging is necessary in optical imaging. However, the filter wheel imaging system which is composed of band pass filters obtains discrete spectral information, with high spatial resolution, but low spectral resolution. This paper proposes a novel high spectral resolution imaging approach based on dual-camera system with filter wheel, to reconstruct high spectral and high spatial resolution images from a limited number of spectral images. The proposed architecture comprises three key components: design of dual-camera system with filter wheel (DCS-FW), establish the mapping relationship between RGB image and spectral images, achieve spectrum reconstruction by interpolation. The proposed system combines useful information from both cameras. And the conservation of energy is used to achieve the mapping relationship between RGB image and spectral images, which we used to establish interpolating compensation computational reconstruction. Both simulation and experiments are adopted to verify the validness of our approach, the results show that our method achieves an accurate and robust high spectral resolution reconstruction while maintaining spatial resolution.

INDEX TERMS High spectral resolution, spatial resolution, filter wheel, band pass filters, dual-camera, energy conservation, interpolation

I. INTRODUCTION
Humans can observe the colorful world, because the cone cells in the human visual system can convert light in the visible spectrum. Conventional cameras imitate the human eyes by using RGB color filters for recording RGB measurements on the sensor, which leads to a great loss in spectra details. Spectral imaging technology can record the reflection response of natural light in different spectral bands, and record the physical and chemical properties of substances that are difficult to be obtained in RGB images, such as the material, morphology, surface smoothness of the objects. These attributes are widely applied in many fields, including remote sensing [1], [2], military reconnaissance [3], medical diagnosis [4], agriculture [5], object detection and tracking [6], [7], etc. Meanwhile, high spatial resolution images can provide a lot of spatial details, which cannot be ignored.

Spectral imaging techniques can be divided into dispersion type, interference type, and filter type based on the spectroscopic principle. The dispersive components used in the dispersive imaging spectrometer include prisms and gratings. The dispersive spectrometer has a simple structure, but it is difficult to obtain both high resolution and high luminous flux. To improve the spectral resolution, the luminous flux has to be sacrificed. Different from the dispersive spectrometer, the interference spectrometer [8] has the characteristics of high luminous flux, however, because of the scanning mirror structure, it also has poor structural stability, which means a small vibration may cause a large imaging error. There are many implementation schemes for filter type spectrometers [9], [15], among which the tunable filter method and the physical filter wheel method are widely used. The specific form of the physical filter wheel method is...
the filter wheel structure, the spectral images of different wavelengths, which have high spatial resolution and retain a large amount of spatial information, can be obtained by controlling the rotation of the filter wheel. But there also exist great limitations. On the one hand, the spectral information of the obtained spectral image is relatively limited, only discrete spectral reflectance curves can be obtained, which leads to low spectral resolution. On the other hand, time loss brought by the rotation of the filter wheel can’t be ignored.

In this paper, we propose a novel computational high spectral resolution imaging method named High Spectral Resolution Imaging using Interpolating Compensation based on Filter Wheel Dual-Camera (HSRIIC-FWDC) to obtain high spectral resolution images without using learning-based methods. Specifically, we construct the mapping relationship between RGB channels and spectral data based on energy conservation. The basic idea is that the RGB images contain the information of the three-channel spectral sensitivity curve of RGB camera, which reflects the same physical quantity as the spectral reflectance curve. Thus, we can establish the corresponding relationship. Then, to achieve high resolution spectral reflectance curve, we adopt interpolating compensation method. More importantly, the time efficiency is further improved by optimizing the filter wheel control system.

Fig. 1 demonstrates the overview of the proposed method which contains three steps. First, the RGB image and spectral images are captured by our system. Then, with the known camera spectral sensitivity, we can establish the mapping relationship between RGB channels and spectral data. Finally, the spectral reflectance is recovered by interpolating compensation, thus, we can achieve high spectral resolution imaging. Additionally, the spatial resolution of the imaging system comes from RGB camera.

In particular, we summarize the major features of our method as follows:

- It completes the reconstruction by using the captured images, instead of using large numbers of spectral images to form spectral datasets, saving resources on dataset acquisition.
- The corresponding relationship between RGB image and spectral images is calculated according to energy conservation, ensuring high accuracy.
- The recovery of spectral reflectance based on interpolating compensation method, which has low computational complexity and is based on original data captured by our imaging system, thus, it is efficient and reliable.
- The dual camera structure enables our imaging system to achieve high spectral resolution imaging and maintain high spatial resolution.

II. RELATED WORK

A. SPECTROMETERS FOR SPECTRAL IMAGING

Spectral imaging has long interested researchers for its great potential in various applications, and an imminent task is capturing three-dimensional spectral datasets using a two-dimensional sensor. Traditional spectral imaging systems involve spatial dimension scanning approaches [10], [11] and spectral dimension filtering approaches [12], [13]. Spatial dimension scanning approaches include whisk-broom and push-broom. Whisk-broom spectrometers, such as Air Visible/Infrared Imaging Spectrometer (AVIRIS) [10], capture the spectrum of a single scene point at each time and move the scanning system to record the spectrum of all scene points. Push-broom spectrometers, such as Hyperspectral Digital Imagery Collection Experiment (HYDICE) [11], records the spectrum of a slit at each time, which improves the scanning efficiency. Spectral dimension filtering spectrometer uses multiple color band-pass filters to record the spectrum of different spectral bands. For example, a spatial variant color filter is distributed over the sensor which requires multiple exposures to record the spectrum of whole scene points [14], or a tunable spectral filter [12] or rotating filter wheel [15] is placed in front of the lens of the camera to change the spectral bands of the capture images.

The above-mentioned conventional spectral imaging methods are difficult to obtain high temporal and spectral resolution at the same time. Compared with traditional spectral imaging, computational spectral imaging can obtain the entire high-dimensional spectral data in a single exposure time. The spectral images can be tomographically
reconstructed by Computed Tomography Imaging Spectrometer (CTIS) [16], [17] which projects the high-dimensional data onto the sensor with the computer-generated hologram disperser, but the CTIS has the inevitable problem of undersampling. Coded Aperture Snapshot Spectral Imaging (CASSI) [18], [19] based on compressive sensing method, replacing the aperture with a two-dimensional random encoding device. The basic assumption is that the natural scene spectrum has multi-scale inherent sparseness. The CASSI system can be improved by acquiring multiple snapshots that are recorded from a coded mask shifting on a piezo stage [20]. Meanwhile, the CASSI also has some obvious disadvantages. First, the CASSI is based on the premise assumption that the natural scene is inherently sparse, so there are inevitable errors in the reconstruction. Second, the reconstruction algorithm of the CASSI has a high computational complexity, which leads to the inability to reconstruct high-dimensional spectral data matrix in real time. Prism-mask Multispectral Video Imaging System (PMVIS) [21], using spatial sampling methods, such as a mask or a microlens array, combined with the traditional dispersion method, captures the scattered spectrum by the high-resolution sensor. Aiming at the limitation of the low spatial resolution of the PMVIS system, the method of capturing high spatial resolution spectral images through a dual-channel system [22], [23] is proposed, and this method propagate the multispectral data to the other pixels according to color similarity and spatial proximity. The light-field spectrometer couples spectral information to light angle information [24], which is easy to integrate, but the improvement of its spectral resolution is built on the sacrifice of spatial resolution.

B. RECONSTRUCTION OF SPECTRAL REFLECTANCE

The spectrometers base on deep learning is widely used in recent years, whose core technical point is to establish the mapping relationship between the RGB values and the spectral high-dimensional data, and reconstruct the spectral reflectance. Several learning-based methods to get the mapping relationship have been widely researched, including Radial Basis Functions (RBF) [27], sparse coding [28]-[30], and Convolutional Neural Networks (CNN) [32]-[36]. To solve the problem of low spectral resolution, a quantity of datasets that contain tristimulus values from the surface reflectance spectrum is used to acquire the principal components of the spectrum [25]. Abed et al. [26] proposed a method that uses local linear interpolating compensation to estimate the reflectance curve with tristimulus values. Nguyen et al. [27] used a radial basis function network to map the RGB values to the reflectance spectra of the scene. Hyperspectral prior is collected and pre-processed only once using tools from the sparse representation literature to recover hyperspectral signatures from RGB measurements [29]. Fu et al. [30] presented an approach that learns multiple non-negative sparse coding dictionaries from the training spectral datasets in terms of clustering results. Otsu et al. [31] presented an approach that reformulates conversion of tristimulus colors to spectra via principal component analysis, and proposes a greedy clustering algorithm that minimizes reconstruction error. Xiong et al. [34] proposed a deep learning framework, i.e., HSCNN, which is one of the first CNN-based methods for hyperspectral recovery from a single RGB image. However, the upsampling operation in HSCNN requires the knowledge of an explicit spectral response function, it thus restricts the applicability of HSCNN when the spectral response function is unknown or difficult to obtain. Z. Shi et al. [35] improved HSCNN by removing the hand-engineered upsampling step and using a deep residual network and densely connected network to achieve more accurate results. Fubara et al. [32] proposed a CNN-based strategy for learning RGB to hyperspectral cube mapping by learning a set of basis functions and weights in a combined manner and using them both to reconstruct the hyperspectral signatures of RGB data.

The above spectral reconstruction methods by mapping RGB values and the spectral high-dimensional data need large prior datasets. Considering that the prior datasets are sensitive to scene and light, this paper proposes a mapping method that is suitable for any real-world scene. We use the same model of cameras to capture RGB images and spectral images and get the mapping relationship by matching the energy of RGB images and spectral images based on conservation of energy.

III. OUR METHOD

In this section, we elaborate on the working principle of DCS-FW, and then demonstrate the basic principle of HSRIC-FWDC by mathematical formulating its mapping relationship between the RGB values and the spectral high-dimensional data. In addition, the spectrum of each scene point can be reconstructed efficiently.

Compared with the CASSI and PMVIS, our method has the characteristics of high spatial resolution. In our system, spatial information provided by RGB image can achieve high spatial resolution imaging.
A. DUAL CAMERA SYSTEM WITH FILTER WHEEL

Dual-camera system with filter wheel is composed of filter wheel, beam splitter and two cameras of the same basic parameters, one of which is used to capture RGB images and the other to capture spectral images as Fig. 1 shows. To ensure that the dual cameras image the same scene, we design the optical path of the dual-camera system precisely. The angle between the plane of the beam splitter and the plane of the filter is controlled to 45°, and the angle between the plane of the RGB camera objective and the plane of the filter is controlled to 90°. Further, the dual cameras are controlled to have same optical range to the plane of the beam splitter. The light containing spectrum is separated into two rays which have the same information and same energy according to the spectroscopic characteristic of the beam splitter, and the rays captured by dual cameras are theoretically equal according to the light propagation principle. By controlling the rotation speed and the acquisition rate of cameras, we can obtain the images efficiently, guaranteeing high time resolution.

B. MAPPING RELATIONSHIP BASED ON CONSERVATION OF ENERGY

Let \( h(x, y, \lambda) \) denotes the 3D images and \( x, y \) is the 2D spatial coordinates, \( \lambda \) is the spectral dimension. In conventional photography, RGB camera captures a 2D image \( i(x, y) \) without spectral dimension, we use \( i \) replace \( i(x, y) \). The RGB camera has the characteristic of obtaining RGB values from the camera spectral sensitivity, and the relationship between RGB values and the camera spectral sensitivity can be formulated as
\[
i^i_m = \int s^i_\lambda(\lambda) r^i_\lambda(\lambda) d\lambda, \tag{1}
\]
where \( i^i_m \) is the RGB values at pixel index \( m \), for the RGB camera, \( n = \{R, G, B\} \) and \( i^i_m \in [0, 255] \), and \( s^i_\lambda(\lambda) \) is the camera spectral sensitivity of RGB channels, \( r^i_\lambda(\lambda) \) is the spectral reflectance.

In practice, the spectral information is discretized across wavelength and (1) can be rewritten as
\[
i^i_m = \sum_{\lambda_b=1}^{B} s^i_{\lambda_b}(\lambda_b) r^i_{\lambda_b}(\lambda_b), \tag{2}
\]
where \( \lambda_b \) (\( b = 1, 2, \cdots, B \)) is the discrete representation of wavelength \( \lambda \) and \( B \) is the number of spectral bands.

Our DCS-FW can capture limited number of spectral images of different bands, in other words, the number of spectral bands \( B \) is relatively small, which leads to low spectral resolution.

We further simplify (2) in matrix form as
\[
I = SR, \tag{3}
\]
where \( I \in \mathbb{R} ^{3 \times M} \) denotes the testing RGB values \( \{ i^i_m \} \), \( S \in \mathbb{R} ^{3 \times B} \) denotes the camera spectral sensitivity of RGB channels \( \{ s^i_\lambda \} \), \( R \in \mathbb{R} ^{B \times M} \) denotes the spectral reflectance \( \{ r^i_\lambda \} \), \( M \) is the number of pixel in the RGB image.

Our task is to reconstruct the spectral reflectance \( R \) from the limited number of spectral images, and the estimation function can be formulated as
\[
\min_{R} || I - SR ||_2^2. \tag{4}
\]

where \( || \cdot ||_2 \) is L2-norm.

Here, we propose a method that uses weight coefficients of the camera spectral sensitivity of RGB channels to reconstruct the spectral reflectance based on conservation of energy. For our system, only narrow bands of light can pass through the filters, which leads to a large energy loss, while the RGB camera does not. The energy matching is essential to the subsequent steps.

Known several spectral images of different bands which is captured by our spectral camera, we can obtain discrete spectral reflectance curve. According to previous studies [37], the spectral reflectance at each pixel can be approximately represented by using a small basis as
\[
r^i_\lambda(\lambda_b) = \sum_{j=1}^{J} \alpha^i_\lambda d^j(\lambda_b), \tag{5}
\]
where \( d^j \) denotes the basis function for spectral reflectance, and \( \alpha^i_\lambda \) denotes the corresponding coefficient, and \( J \) denotes the number of basis used.

In our terms, we use the camera spectral sensitivity of RGB channels \( \{ s^i_\lambda \} \) to act as \( d^j \) in order to meet the requirement of energy matching, because \( \{ s^i_\lambda \} \) has the characteristic of spectral reflectance. Furthermore, because the RGB camera has three channels, we assign 3 to \( J \). Thus, (5) can be rewritten as
\[
r^i_\lambda(\lambda_b) = \sum_{j=1}^{J} \alpha^i_\lambda s^j_\lambda(\lambda_b). \tag{6}
\]

The unknown \( \{ \alpha^i_\lambda \} \) can be formulated by solving an optimization problem:
\[
\min_{\alpha^i_\lambda} \| r^i_\lambda - \sum_{j=1}^{3} \alpha^i_\lambda s^j_\lambda(\lambda_b) \|_2^2, \text{ s.t. } \forall j = 1, 2, 3, \alpha^i_\lambda \geq 0. \tag{7}
\]

In practice, the number of spectral bands \( B \) equals the number of filters in our system. In a hence, the mapping relationship between RGB image and spectral images is established. Then we use an efficient interpolating reconstruction method to pursue high spectral resolution imaging.

C. HIGH SPECTRAL RESOLUTION IMAGING BASED ON INTERPOLATING RECONSTRUCTION

The spectral images captured by our system can analysis out discrete spectral reflectance curve as Fig. 1 shows, and the values of this curve can make up \( \{ r^i_\lambda(\lambda_b) \} \). We use the calculated \( \{ \alpha^i_\lambda \} \) combines with the camera spectral sensitivity of RGB channels \( S \) to estimate the spectral reflectance \( R \), which also restricted by (4), specifically, adjust \( \{ \alpha^i_\lambda \} \) adaptively.

The values of the spectral reflectance curve can be estimated as
\[ r_n(\lambda) = \sum_{j=1}^{3} \alpha_j s_j(\lambda) \quad \text{s.t.} \quad \|i_n - \int s_n(\lambda) r_n(\lambda) d\lambda\|_2^2 \leq \varepsilon, \quad (8) \]

where \( \varepsilon \) is the residual error.

IV. EXPERIMENTS

In this section, the implementation details and analysis are provided. First, we describe the images captured by our spectral system and the image preprocessing we implement. Then, we use the proposed approach to recover the spectral reflectance.

A. SYSTEM SETTINGS AND IMAGES ACQUISITION

We capture the RGB images and spectral images by DCS-FW, and the prototype is demonstrated in Fig. 2. The sensor of RGB camera and grayscale camera we use are the FL2-2084 series which have maximum resolution of 1624\,×\,1224 at 15fps and pixel size of 4.4\,µm, and the objective lens we use have focal length of 35mm and aperture value of F1.4. We use the filter wheel control system to control the rotation speed of the filter wheel to 2 circles/s. Meanwhile, considering the frame rate of the grayscale camera, we set the image sampling interval of the grayscale camera to 12 fps, and we are thus able to capture two sets of images per second, maintaining time resolution. The average time for our method to reconstruct spectral reflectance from a whole image is 0.256s, which is much less than the acquisition time of a set of images. Therefore, in our imaging system, the acquisition time of the spectral cube depends mainly on the acquisition time of the images, in other words, the acquisition time of the spectral cube is 2\,cubes/s.

In this section, we use narrowband filters with center wavelengths of 450nm, 470nm, 540nm, 560nm, 630nm, 680nm to obtain spectral images, and control the spatial dimension of RGB images and spectral images to 1024\,×\,1024. To investigate the impact of illumination on the reconstruction performance, we do our experiment under the illuminations of CIE D65 and sunlight, the images we capture are shown in Fig. 3, and the top three rows are captured in the indoor environment under the illumination of CIE D65, the bottom three rows are captured in the outdoor environment under the illumination of sunlight. We select the representative points of the scenes as sampling points for testing, and we show the discrete spectral reflectance curves in Fig. 4.

B. RECOVERY OF SPECTRAL REFLECTANCE

Here, we set the residual error \( \varepsilon \) to 1, which is an empirical value and we recover the spectral reflectance of sampling points in the region of interest that are meaningful for the algorithm demonstration. The results of reconstruction of spectral reflectance are showed in Fig. 5.

V. ANALYSIS AND DISCUSSION

In this section, we use objective evaluation metrics to analyze the accuracy and robustness of our method objectively, and use comparative experiments to verify the effect of our method.

We use three objective evaluation metrics to evaluate the quantitative performance for the spectral reflectance reconstruction, including root mean square error (RMSE) as in (9) and mean relative absolute error (MRAE) as in (10) to measure the error and spectral angle mapper (SAM) as in (11) to measure the similarity, and smaller values of RMSE and MRAE mean better performance, otherwise, the bigger values of SAM mean better performance, and SAM is in the range of 0 to 1.

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (R_i - R_i^g)^2}, \quad (9)
\]
\[
\text{MRAE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{R_i - R_i^g}{R_i} \right|, \quad (10)
\]
\[
\text{SAM} = \frac{\sum_{i=1}^{n} R_i R_i^g}{\left( \sum_{i=1}^{n} R_i^2 \right) \left( \sum_{i=1}^{n} R_i^g^2 \right)^{1/2}}, \quad (11)
\]

where \( n \) denotes the total number of data points in the spectral cube, \( R_i \) and \( R_i^g \) denote the reconstructed and ground-truth spectra.

A. ACCURACY ANALYSIS

Table I shows the MRSE, RMAE, and SAM between the reconstructed spectral reflectance and the original discrete curve. We can see that the spectral reflectance reconstructed by our method bases on the original values, and maintains the feature of original values. Combining Fig. 5 and Table I, we can conclude that the error between the results of our method and the original values is small, indicating that our method has good fidelity.

| TABLE I RMSE, MRAE AND SAM RESULTS BETWEEN THE RECONSTRUCTED SPECTRAL REFLECTANCE AND THE ORIGINAL DISCRETE CURVE |
|---|---|---|---|
| Illumination | Scene | RMSE | MRAE | SAM |
| CIE D65 | Scene I | 0.0798 | 0.2841 | 0.9556 |
| | Scene II | 0.0058 | 0.1692 | 0.9896 |
| | Scene III | 0.1314 | 0.1780 | 0.9728 |
| | Scene IV | 0.0811 | 0.2950 | 0.9725 |
| Daylight | Scene V | 0.0700 | 0.2796 | 0.9795 |
| | Scene VI | 0.0544 | 0.0712 | 0.9966 |
| Average | | 0.0704 | 0.2129 | 0.9778 |

To further analyze the accuracy of our method, we use three comparison methods, including principal component analysis based on a greedy clustering algorithm which is proposed by Otsu et al. [31], the sparse representation-based method (SR) [29], and a learning-based method, i.e., HSCNN-D [35]. Because the methods based on dictionary learning require large training set, we use CAVE [38] spectral datasets as the training set for comparison methods. CAVE spectral datasets contain 32 real-world objects, and each object contains an RGB image and 31 spectral images, whose wavelength is across 400nm to 700nm at 10nm intervals and spatial...
Scene I: Wooden object in indoor environment

Scene II: Alloy object in indoor environment

Scene III: Real apple in indoor environment

Scene IV: Plant in outdoor environment

Scene V: Wall in outdoor environment

Scene VI: Plastic in outdoor environment

(a) RGB (b) 450nm (c) 470nm (d) 540nm (e) 560nm (f) 630nm (g) 680nm

FIGURE 3. RGB images and spectral images we captured to test and verify the proposed approach. The top three rows are captured in indoor environment under illumination of CIE D65, and the bottom three rows are captured in outdoor environment under illumination of sunlight, and the observed objects are wooden object, alloy object, real apple, plant, wall and plastic from top to bottom. (a) RGB images. (b)-(g) Spectral images of different wavelengths.
FIGURE 4. The original discrete curves of the sampling points of the scenes.

FIGURE 5. Spectral reflectance of the sampling points of the scenes reconstructed by our method.
FIGURE 6. Spectral reflectance of the sampling points of the scenes reconstructed by comparison methods.

FIGURE 7. The reconstruction results of colorchecker’s reflectance under indoor illumination of CIE D65.
are captured under the illumination of CIE D65. Taking the performance under different illuminations, which indicates effect of the comparison reconstruction methods. For our obtained by PMVIS serve as ground truth for spectral spectral reflectance curves reconstructed by our method are default premise is that the spectral reflectance obtained by showed in Fig. 6.

To analyze the robustness of our method, we also compare the actual reconstruction results of eight patches in the color checker in Fig. 7, and the image is taken under indoor illumination of CIE D65. The quantitative errors of these patches are also shown in Table III. Specifically, in Fig. 7(b) and Fig. 7(h) we can see that the reconstruction results of our method, Otsu, and HSCNN-D have roughly the same curve trend with the ground truth, however, SR method does not perform well in this aspect. Compared with Otsu and HSCNN-D, our method has more detailed curve information, e.g., in Fig. 7(b), same as the ground truth, the curve reconstructed by our method has a peak near 620nm, which is not found in the curves reconstructed by Otsu and HSCNN-D. In Fig. 7(h), we can see that the curve reconstructed by our method has peaks near 600nm and 670nm, which are the same as the ground truth and are not found in the comparison methods. The details described above prove that our method performs better in obtaining detailed features. From items (b) and (h) in Table III, we can see that our method has good results on all three objective evaluation metrics in both items, better than the results of SR, Otsu, and HSCNN-D overall, in other words, the reconstruction results of our method have lower error. Combining the results of Fig. 7 and Table III, we can also get the conclusion that the reconstruction results of our method 

### Table II

| Ill. | Scene | RMSE | MRAE | SAM |
|------|-------|------|------|-----|
| CIE D65 | Scene I | 0.0578 | 0.0583 | 0.0771 | 0.0889 | 0.2180 | 0.2128 | 0.3083 | 0.2656 | 0.9752 | 0.9743 | 0.9507 | 0.9627 |
|      | Scene II | 0.0092 | 0.0161 | 0.0106 | 0.0308 | 0.1927 | 0.3632 | 0.1853 | 0.2081 | 0.9698 | 0.9002 | 0.9586 | 0.9573 |
|      | Scene III | 0.0687 | 0.0954 | 0.0721 | 0.0174 | 0.2796 | 0.3234 | 0.1893 | 0.1585 | 0.9775 | 0.9500 | 0.9763 | 0.9823 |
|      | Average | 0.0452 | 0.0566 | 0.0533 | 0.0457 | 0.2301 | 0.2998 | 0.2276 | 0.2107 | 0.9742 | 0.9415 | 0.9619 | 0.9674 |
| Daylight | Scene IV | 0.0671 | 0.1293 | 0.0738 | 0.1759 | 0.2907 | 0.7840 | 0.3052 | 0.3661 | 0.9628 | 0.8414 | 0.9410 | 0.9069 |
|      | Scene V | 0.0510 | 0.1178 | 0.0719 | 0.0968 | 0.1969 | 0.2769 | 0.3637 | 0.4876 | 0.9779 | 0.9737 | 0.9802 | 0.9336 |
|      | Scene VI | 0.0445 | 0.1290 | 0.0544 | 0.1527 | 0.1110 | 0.3529 | 0.1438 | 0.2200 | 0.9940 | 0.9174 | 0.9924 | 0.9711 |
|      | Average | 0.0497 | 0.0910 | 0.0600 | 0.0938 | 0.2148 | 0.3855 | 0.2448 | 0.2843 | 0.9762 | 0.9262 | 0.9665 | 0.9523 |

### Table III

| Scene | RMSE | MRAE | SAM |
|-------|------|------|-----|
| (a)   | 0.0371 | 0.0501 | 0.0332 | 0.0368 | 0.2883 | 0.5054 | 0.2436 | 0.3529 | 0.9631 | 0.9372 | 0.9571 | 0.9491 |
| (b)   | 0.0334 | 0.3622 | 0.0373 | 0.0590 | 0.2306 | 2.1171 | 0.2408 | 0.4394 | 0.9648 | 0.5658 | 0.9646 | 0.8853 |
| (c)   | 0.0494 | 0.0890 | 0.0352 | 0.0743 | 0.2339 | 0.5017 | 0.2028 | 0.3889 | 0.9639 | 0.8849 | 0.9837 | 0.9220 |
| (d)   | 0.0227 | 0.0760 | 0.0510 | 0.0684 | 0.2237 | 0.7109 | 0.3661 | 0.4866 | 0.9822 | 0.8187 | 0.9564 | 0.8646 |
| (e)   | 0.0382 | 0.0432 | 0.0383 | 0.0618 | 0.3028 | 0.3581 | 0.2593 | 0.4771 | 0.9628 | 0.9848 | 0.9762 | 0.8928 |
| (f)   | 0.0424 | 0.0591 | 0.0324 | 0.0417 | 0.4657 | 0.6462 | 0.3804 | 0.4044 | 0.8931 | 0.8211 | 0.9569 | 0.9131 |
| (g)   | 0.0244 | 0.0474 | 0.0522 | 0.0512 | 0.2084 | 0.9446 | 0.5632 | 0.5312 | 0.9775 | 0.9064 | 0.9139 | 0.8772 |
| (h)   | 0.0427 | 0.1488 | 0.0562 | 0.0353 | 0.1459 | 0.3615 | 0.2361 | 0.1570 | 0.9856 | 0.7981 | 0.9833 | 0.9894 |
| Average | 0.0363 | 0.1095 | 0.0420 | 0.0536 | 0.2625 | 0.7682 | 0.3115 | 0.4047 | 0.9616 | 0.8351 | 0.9615 | 0.9117 |

dimension is 512×512. In addition, CAVE spectral datasets are captured under the illumination of CIE D65. Taking the RGB images we captured as a testing dataset to reflect the effect of the comparison reconstruction methods. For our experiments demonstrate, the spectral reflectance curves obtained by PMVIS serve as ground truth for spectral reflectance recovery.

To show the performance of each method clearly, we put recovery results of spectral reflectance of each method in the same scene into the same coordinate system, which is showed in Fig. 6.

We also put the spectral reflectance obtained by PMVIS into the coordinate system as the reference. In this paper, our default premise is that the spectral reflectance obtained by PMVIS has a small error from the ground truth, which has been proved in previous research [22].

From Fig. 6 we can see that in each scene, the reconstructed spectral reflectance curves have roughly the same trend as the ground truth, which indicates that our method and the comparison methods are feasible. In addition, spectral reflectance curves reconstructed by our method are much closer to the ground truth in most scenes.

Table II shows RMSE, MRAE, and SAM results of recovered spectral reflectance for different methods under different illuminations respectively. We can see from Table II that our method performs better than other comparison methods in most cases. Comparing the average values of each measurement, we can see that our method has a better performance under different illuminations, which indicates that our method has better capability in terms of accuracy and fidelity.

### B. ROBUSTNESS ANALYSIS

To analyze the robustness of our method, we also compare the actual reconstruction results of eight patches in the color checker in Fig. 7, and the image is taken under indoor illumination of CIE D65. The quantitative errors of these patches are also shown in Table III. Specifically, in Fig. 7(b) and Fig. 7(h) we can see that the reconstruction results of our method, Otsu, and HSCNN-D have roughly the same curve trend with the ground truth, however, SR method does not perform well in this aspect. Compared with Otsu and HSCNN-D, our method has more detailed curve information, e.g., in Fig. 7(b), same as the ground truth, the curve reconstructed by our method has a peak near 620nm, which is not found in the curves reconstructed by Otsu and HSCNN-D. In Fig. 7(h), we can see that the curve reconstructed by our method has peaks near 600nm and 670nm, which are the same as the ground truth and are not found in the comparison methods. The details described above prove that our method performs better in obtaining detailed features. From items (b) and (h) in Table III, we can see that our method has good results on all three objective evaluation metrics in both items, better than the results of SR, Otsu, and HSCNN-D overall, in other words, the reconstruction results of our method have lower error. Combining the results of Fig. 7 and Table III, we can also get the conclusion that the reconstruction results of our method
have lower error and higher shape fidelity, in other words, our method has a better performance.

Synthesizing the results of the experiments we did, we can get the conclusion that our method is robust, it has a good reconstruction effect no matter under different illuminations or for different objects.

C. FURTHER DISCUSSION

Our method combines the captured discrete spectral reflectance curve and RGB spectral sensitivity with detailed features, which guarantees the constructed spectral reflectance with more detailed features, meanwhile, get better reconstruction results. The comparison methods rely on the training dataset, and the quality of training dataset affects the reconstruction results directly. We can see from Table II that under the illumination of daylight, the comparison methods have worse performance than under the illumination of CIE D65 in most cases, because the training datasets are captured under the illumination of CIE D65, and the reconstruction performance of spectral reflectance is sensitive to the environments and illuminations of training hyperspectral image datasets. If the training datasets contain different illuminations, the performance would be better. Since the work of collecting hyperspectral image datasets is laborious, the effect of learning-based methods is limited.

Known the spectral reflectance of an object is determined by its material composition and the spectral properties of the illumination, the illumination is very influential, our method captures directly under different illuminations, which avoid the errors caused by illumination.

VI. CONCLUSION

In this paper, we present an effective spectral reconstruction method based on the mapping relationship between spectral images and RGB image. Spectral images and RGB image are first captured by dual-camera system with filter wheel. To establish high spectral resolution imaging, the camera spectral sensitivity and conservation of energy are used to establish the mapping relationship which is the key to spectral reconstruction. Compared with previous works, our method does not require hyperspectral prior, which is based on hyperspectral image datasets. From the experiment results under different illuminations, we show that the spectral reflectance can be recovered with low RMSE, low MRAE, and high SAM, which proves that our method has good accuracy and robustness. We believe that the proposed approach can facilitate a wide range of spectral imaging applications.

A limitation of our work is that the time resolution of our method is closely related to the acquisition speed of spectral images. In the future, the method to improve time resolution should be further optimized. Besides, another limitation would be the capability of cameras that may cause additional errors, the cameras in this paper are close to cut-off when approaching 700nm, so we are aiming at solving this problem from the perspective of algorithm and equipment.

Since the range of three-channel spectral sensitivity curve of RGB camera is in the visible light band, our method can reconstruct visible spectrum. The reconstruction of the spectral that beyond the wavelength of the RGB will be further researched by using cameras with wider spectral range or combining with learning-based methods, and how to integrate these technologies is the key point of future research.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” IEEE Geosci. Remote Sens. Lett., vol. 1, no. 2, pp. 6-36, 2013.
[2] Z. Wang, C. Zou, and W. Cai, “Small sample classification of hyperspectral remote sensing images based on sequential joint deep learning model,” IEEE Access, vol. 8, pp. 71353-71363, 2020.
[3] C. R. Schwartz, M. T. Eismann, J. N. Cederquist, and R. O. Johnson, “Thermal multispectral detection of military vehicles in vegetated and desert backgrounds,” in Proc. of SPIE, vol. 2742, 1996, pp. 404-408.
[4] M. A. Calin, S. V. Parasca, D. Savastru, and D. Manea, “Hyperspectral imaging in the medical field: present and future,” Appl. Spectrosc. Rev., vol. 49, no. 6, pp. 435-447, 2014.
[5] K. L. Ang and J. K. P. Seng, “Big Data and Machine Learning With Hyperspectral Information in Agriculture,” IEEE Access, vol. 9, pp. 36699-36718, 2021.
[6] V. Ferraris, N. Dobigeon, Q. Wei, and M. Chabert, “Robust fusion of multiband images with different spatial and spectral resolutions for change detection,” IEEE Trans. Comput. Imaging, vol. 3, no. 2, pp. 175-186, Jun. 2017.
[7] Z. Chen, and X. Huang, “Pedestrian detection for autonomous vehicle using multi-spectral cameras,” IEEE Trans. Intell. Veh., vol. 4, no. 2, pp. 211-219, Jun. 2019.
[8] R. G. Sellar, and G. D. Boreman, “Classification of imaging spectrometers for remote sensing applications,” Opt. Eng., vol. 44, no. 1, pp. 013602, Jan. 2005.
[9] V. Sapkota, and K. Youcef-Toumi, “Design of a mechanical-tunable filter spectrometer for noninvasive glucose measurement,” Appl. Opt., vol. 43, no. 13, pp. 2680-2688, 2004.
[10] R. O. Green, M. L. Eastwood, C. M. Sarture, et al, “Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS),” Remote Sens. Environ., vol. 65, no. 3, pp. 227-248, 1998.
[11] P. A. Mitchell, “Hyperspectral digital imagery collection experiment (HYDICE),” in Proc. of Geographic Information Systems, Photogrammetry, and Geological/Geophysical Remote Sensing, vol. 2587, 1995, pp. 70-95.
[12] N. Gai, “Imaging spectroscopy using tunable filters: a review,” in Proc. of Wavelet Applications VII, vol.4056, 2000, pp. 50-64.
[13] H. R. Morris, C. C. Hoyt, and P. J. Treado, “Imaging spectrometers for fluorescence and Raman microscopy: acousto-optic and liquid crystal tunable filters,” Appl. Spectrosc., vol. 48, no. 7, pp. 857-866, 1994.
[14] Y. Y. Schechner, and S. K. Nayar, “Generalized mosaicing: wide field of view multispectral imaging,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 10, pp. 1354-1348, Oct. 2002.
[15] M. Yamaguchi, H. Haneshi, H. Fukuda, J. Kishimoto, H. Kanazawa, M. Tsuchida, R. Iwama, and N. Ohyama, “High-fidelity video and still-image communication based on spectral information: Natural vision system and its applications,” in Proc. of Spectral Imaging: Eighth International Symposium on Multispectral Color Science, vol. 6062, 2006.
[16] M. Descourand, and E. Dereniak, “Computed-tomography imaging spectrometer: experimental calibration and reconstruction results,” Appl. Opt., vol. 34, no. 22, pp. 4817-4826, 1995.
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3166946, IEEE Access

J. Yu et al.: High Spectral Resolution Imaging Based on Dual-Camera System with Filter Wheel (November 2021)

[17] W. R. Johnson, D. W. Wilson, G. H. Bearman, and J. Backlund, “An all-reflective computed tomography imaging spectrometer,” in Proc. of SPIE 5660, Instruments, Science, and Methods for Geospace and Planetary Remote Sensing, vol. 5660, 2004, pp. 88-97.

[18] A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, “Video rate spectral imaging using a coded aperture snapshot spectral imager,” Opt. Express, vol. 17, no. 8, pp. 6368-6388, 2009.

[19] R. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, “Higher-order computational model for coded aperture spectral imaging,” Appl. Opt., vol. 52, no. 10, pp. D12-D21, 2013.

[20] D. Kittler, K. Choi, A. Wagadarikar, and D. J. Brady, “Multiframe image estimation for coded aperture snapshot spectral imagers,” Appl. Opt., vol. 49, no. 36, pp. 6824-6833, 2010.

[21] H. Du, Y. Tong, X. Cao, and X. Lin, “A prism-based system for multispectral video acquisition,” in 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 175-182.

[22] X. Cao, H. Du, X. Tong, Q. Dai, and S. Lin, “A prism-mask system for multispectral video acquisition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 12, pp. 2423-2435, 2011.

[23] J. Feng, X. Fang, X. Cao, C. Ma, Q. Dai, H. Zhu, and Y. Wang, “Advanced hyperspectral video imaging system using Amici prism,” Opt. Express, vol. 22, no. 16, pp. 19348-19356, 2014.

[24] L. Su, Z. Zhou, Y. Yuan, L. Hu, and S. Zhang, “A snapshot light field imaging spectrometer,” Optik, vol. 126, no. 9-10, pp. 877-881, 2015.

[25] F. Ayala, J. F. Echávarri, P. Renet, and A. I. Negueruela, “Use of three tristimulus values from surface reflectance spectra to calculate the principal components for reconstructing these spectra by using only three eigenvectors,” J. Opt. Soc. Am. A (JOSA A), vol. 23, no. 8, pp. 2020-2026, 2006.

[26] F. M. Abed, S. H. Amirshahi, and M. R. M. Abed, “Reconstruction of reflectance data using an interpolation technique,” J. Opt. Soc. Am. A (JOSA A), vol. 26, no. 3, pp. 613-624, 2009.

[27] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, “Training-based spectral reconstruction from a single RGB image,” in Proc. of European Conference on Computer Vision (ECCV), 2014, pp. 186-201.

[28] S. Lansel, M. Parmar, and B. A. Wandell, “Dictionaries for sparse representation and recovery of reflectances,” in Proc. of Computational Imaging VII, vol. 7246, 2009.

[29] B. Arad, and O. Ben-Shahar, “Sparse recovery of hyperspectral signal from natural RGB images,” in Proc. of European Conference on Computer Vision (ECCV), 2016, pp. 19-34.

[30] Y. Fu, Y. Zheng, L. Zhang, and H. Huang, “Spectral reflectance recovery from a single RGB image,” IEEE Trans. Comput. Imaging, vol. 4, no. 3, pp. 382-394, 2018.

[31] H. Otsu, M. Yamamoto, and T. Hachisuka, “Reproducing spectral reflectances from tristimulus colours,” in Proc. of Computer Graphics Forum, 2018, pp. 370-381.

[32] B. J. Fubara, M. Sedky, and D. Dyke, “RGB to spectral reconstruction via learned basis functions and weights,” in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 480-481.

[33] S. Galliani, C. Lanaras, D. Marmanis, E. Baltasvias, and K. Schindler, “Learned spectral super-resolution,” arXiv preprint arXiv:1703.09470, 2017.

[34] Z. Xiong, Z. Shi, H. Li, L. Wang, D. Liu, and F. Wu, “Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections,” in Proc. of the IEEE International Conference on Computer Vision Workshops (ICCV), 2017, pp. 518-525.

[35] Z. Shi, C. Chen, Z. Xiong, and D. Liu, “Hscnn+: Advanced CNN-based hyperspectral recovery from RGB images,” in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 939-947.

[36] T. Stiebel, S. Koppers, P. Seltsam, D. Merhof, “Reconstructing spectral images from rgb-images using a convolutional neural network” in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 948-953.

[37] L. T. Maloney, “Evaluation of linear models of surface spectral reflectance with small numbers of parameters,” J. Opt. Soc. Am. A (JOSA A), vol. 3, no. 10, 1673-1683, 1986.