

## Research Article

# High Zoom Ratio Foveated Snapshot Hyperspectral Imaging for Fruit Pest Monitoring

Yaoyao Hu,<sup>1</sup> Jun Chang ,<sup>1</sup> Yiting Li,<sup>1</sup> Wenchao Zhang,<sup>1</sup> Xiaoxiao Lai,<sup>1</sup> and Quanquan Mu<sup>2</sup>

<sup>1</sup>School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

<sup>2</sup>Key Laboratory of Optical System Advanced Manufacturing Technology, China Changchun Institute of Optics, Fine Mechanics and Physics, Changchun, Jilin, China

Correspondence should be addressed to Jun Chang; [optics\\_chang@126.com](mailto:optics_chang@126.com)

Received 10 November 2022; Revised 14 December 2022; Accepted 15 December 2022; Published 3 January 2023

Academic Editor: Jose S. Camara

Copyright © 2023 Yaoyao Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Snapshot hyperspectral imaging technology is increasingly used in agricultural product monitoring. In this study, we present a 9× local zoom snapshot hyperspectral imaging system. Using commercial spectral sensors with spectrally resolved detector arrays, we achieved snapshot hyperspectral imaging with 14 wavelength bands and a spectral bandwidth of 10–15 nm. An experimental demonstration was performed by acquiring spatial and spectral information about the fruit and *Drosophila*. The results show that the system can identify *Drosophila* and distinguish well between different types of fruits. The results of this study have great potential for online fruit classification and pest identification.

## 1. Introduction

Nowadays, computer vision technology inspired by the human visual system is widely applied to plant protection and agricultural management and has made great achievements in pest detection and recognition. Numerous methods of pest detection and recognition have been investigated in recent years, such as computer-assisted estimation [1], k-means clustering [2], support vector machines (SVMs) [3], cognitive vision [4], and optimal deep residual learning [5–7]. Most of these technologies perform subsequent image processing based on the results collected by an RGB camera in an indoor environment. Among the spectral systems mentioned in these studies, some are spectrometers, which cannot acquire two-dimensional images, and some are hyperspectral imaging systems, which focus mainly on fruit classification algorithms; these optical systems have a fixed focal length.

Multi- and hyperspectral imaging, which refers to the recording of two-dimensionally resolved spectral information, has been rapidly applied in the agricultural food sector [8–11]. Different spectral imaging techniques have been developed in the past few years to meet the individual

requirements of each specific application, including the “snapshot” spectral imaging used in real-time monitoring. The snapshot imaging system allows the recording of spatial and spectral information in one acquisition, and it creates a complete three-dimensional hyperspectral cube in one step. To this end, the spectral imaging data were multiplexed onto a single frame of the image sensor.

Snapshot spectral imaging technology has been proposed since 1978; according to different implementations, it can be divided into the following types: Image reproduction spectral imaging (IRIS) technology uses image arrays and filter arrays to achieve imaging of different bands. Spectroscopic components include spectroscopic prisms, light pipes [12], Fabry–Perot interference filters [13], lens arrays [14, 15] and Wollaston prisms [16–18]. This method requires a complex design and sophisticated equipment. The image mapping spectrometer (IMS) and integrated field technology use array elements to divide images and then reintegrate and superimpose the subimages to obtain the final result. The most common array elements include fiber arrays [19], microlens arrays [20], and linear gradient filters [21–23]. They need to properly process the pupils and finally reimagine them on a two-dimensional detector. The

disadvantages of these technologies include the need for a high-precision image slicer and the complexity of the subsequent optical path to ensure proper pupil processing and spectral decomposition. Based on the compressed sensing theory, compressed sensing snapshot spectral imaging technology [24, 25] collects snapshots and performs spectral image restoration. However, the restoration process of this method is complex, and the key devices of the system have weak environmental adaptability. With the rapid development of device technology in recent years, some spectral sensors have been manufactured for snapshot spectroscopy, including quantum dot spectrometers [26, 27] and spectral sensors based on photonic crystal plates [28]. This technology is compact and can be coupled to various optical lenses. However, these technologies are difficult to apply because of their complex manufacturing processes and high costs.

Spectrally resolved detector arrays (SRDA) have been widely used in agricultural monitoring [29]. The spectral filters of the SRDA are deposited either in front of a glass substrate [28] or directly on top of the sensor array [30]. It is possible to acquire video-rate spectral images when spectral filters are deposited on the sensor in a pixel-wise mosaic pattern [30]. SRDA is compact and robust, has the potential to be produced at a very low cost, and has been commercialized. However, the number of wavelength channels supported by the SRDA is currently limited by the number of CMOS pixels. Increasing the wavelength channels bins CMOS pixels into a single hyperspectral imaging pixel and inevitably reduces the imaging spatial resolution [29, 30].

In early research, a foveated imaging technique based on bionic human eye vision, which has the characteristics of a large field and local high-resolution imaging, was proposed [31–33]. This technology can compensate for the problem of low spatial resolution caused by the limitation of the CMOS pixels of the SRDA. Depending on their implementation, foveated imaging techniques can be divided into three major categories: local aberration-corrected foveated imaging [34–37], multichannel fused foveated imaging [38, 39], and local magnification foveated imaging [40]. Local magnification foveal imaging technology realizes the local magnification of the region of interest through the foveated channel to improve the object space resolution of the system and can be widely coupled with different detectors.

Based on the foveated magnification foveated imaging technology, we followed the methods of Shen et al. 2018 [41] and used SRDA to build a  $9\times$  foveated-zoom snapshot hyperspectral imaging (LHSI) system for fruit pest monitoring. We achieved a 14-bands (460 nm–630 nm) using a  $256\times 512$ -pixel foveated zoom hyperspectral imaging system, which has two imaging channels with the same image plane. One channel is designed as foveated zoom imaging for pest recognition, whereas the other channel is peripheral imaging for searching pests in a wide FOV. Both imaging channels contain 14-band spectral information. The experimental results verified that the use of foveated imaging theory can improve the spatial resolution of SRDA-based spectral imaging systems. This system is widely used for imaging *Drosophila* on fruits. The zoomed *Drosophila*

images in a  $2\times$  to  $9\times$  foveated region of interest (LRI) and the unzoomed *Drosophila* images in a wide field of view (FOV) were captured by the established setup. The foveated zoomed spectral images can help pest high-resolution recognition, whereas the peripheral unzoomed spectral images play an important role in expanding the observation scope of the system. We also performed image color recovery, fruit contrast enhancement, and fruit and pest classification using spectral analysis. This study provides a potential solution for hyperspectral imaging applications in other fields.

## 2. LHSI System Design Strategy

**2.1. Principle of LHSI System.** The foveated-zoom snapshot hyperspectral imaging system comprises a fixed imaging lens, a laterally moving foveated zoom lens, a fixed relay lens, and an SRDA, as shown in Figure 1.  $f_i$  and  $f_r$  are the fixed focal lengths of the imaging and relay lenses, respectively. The foveated zoom lens was located between the imaging lens and the relay lens. The foveated zoom lens consisted of a variator for foveated zooming and a compensator to eliminate the axial shift of the image plane of the LRI. The focal lengths of the variator and compensator are  $f_v$  and  $f_c$ , respectively. In Figure 1(a), the peripheral imaging channel is marked with light blue, which is imaged by the imaging lens and the relay lens at the image plane without the foveated zoom lens and detected by the SRDA. The LRI of the target scene to be imaged is shown in Figure 1(b), which is magnified at different magnifications of  $2\times$  to  $9\times$ . The light paths of the magnifications are marked in yellow, green, red, and orange in Figure 1(a). The LRI passes through the imaging lens, foveated zoom lens, and relay lens. The light path is a foveated zoom imaging channel. This channel magnifies the LRI and improves its spatial resolution. The foveated scene was zoomed at different magnifications by nonlinearly shifting the positions of the variator and compensator.

Figure 2(a) shows a schematic of the imaging of the SRDA in one step. The yellow, green, red, and orange circles are images with LRI magnifications of  $2\times$  to  $9\times$ . The outside circles represent unzoomed images of the peripheral scene. The SRDA used in this study is a commercial product of IMEC from Belgium ( $2/3$  CMOS,  $5.5\times 5.5\ \mu\text{m}$  CMOS pixel size before and after the dielectric coating, USB interface, 120 fps maximum imaging speed). The  $1024\times 2048$  CMOS pixels on the imaging chip are divided into  $256\times 512$  hyperspectral imaging pixels. Each hyperspectral imaging pixel contained  $4\times 4$  CMOS pixels, representing 16 wavelength channels. For each CMOS pixel, a dielectric thin film Fabry–Perot (FP) cavity filter was monolithically fabricated on the surface. Figure 2(b) shows a cross-sectional view of a hyperspectral pixel (16 CMOS pixels). The incident light is imaged on the CMOS after passing through the FP filter. In our device, 16 different filter thicknesses were allocated, allowing 16 channels of different wavelengths from 460 to 630 nm to pass. The spectral response of the SRDA is shown in Figure 2(c). After spectral calibration, 14 effective corresponding bands were finally obtained, which are 480, 483, 492, 494, 508, 520, 533, 546, 571, 583, 595, 607, 619, and

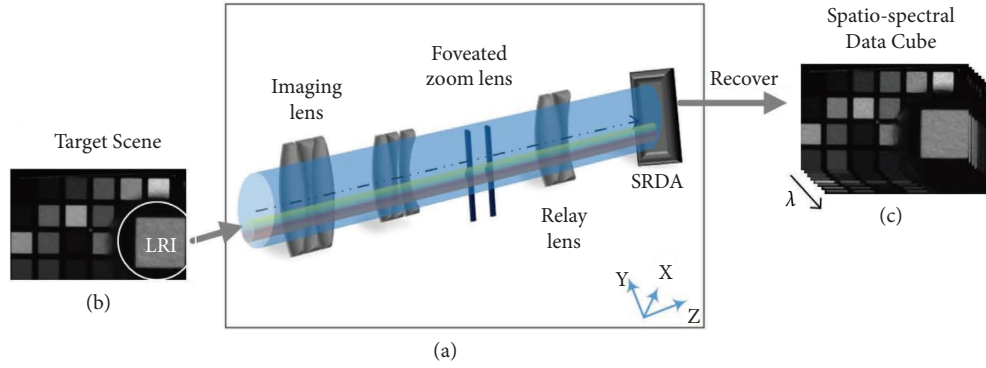


FIGURE 1: Principle layout of the LSHI. (a) System structure. (b) Target scene. (c) Spatio-spectral data cube of the reconstruction scene.

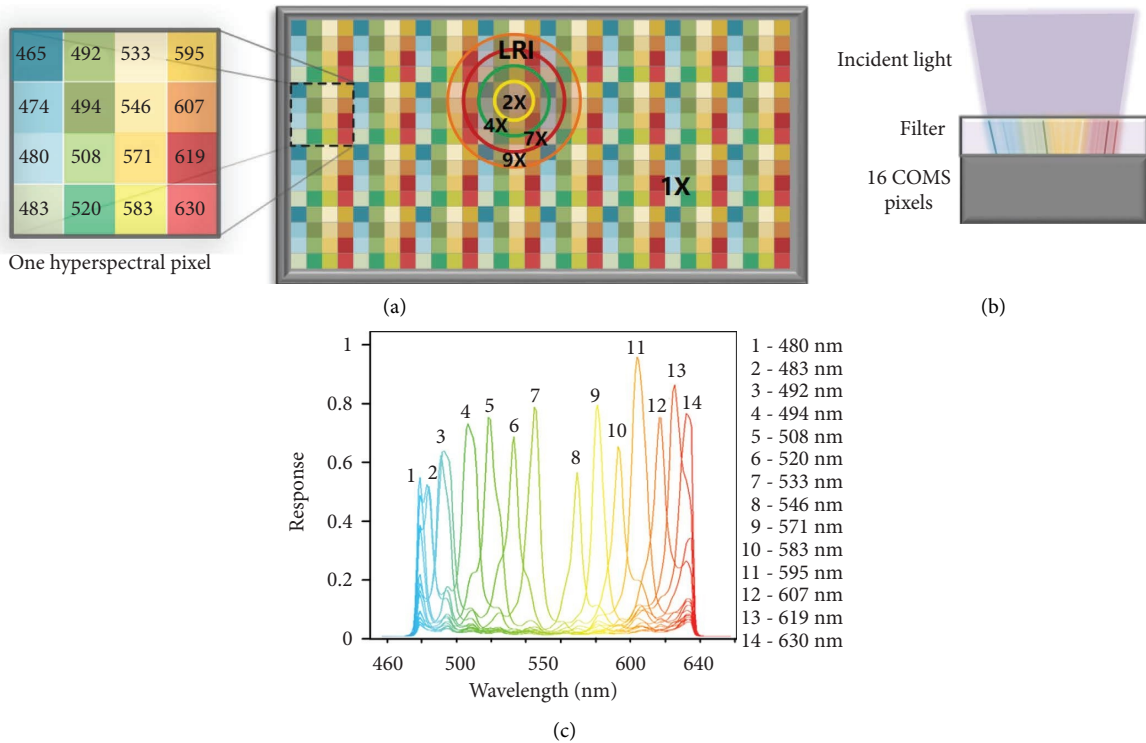


FIGURE 2: (a) Illustration of spectral images on mosaic CMOS and SRDA hyperspectral imaging pixel. (b) Illustration of one hyperspectral imaging pixel (16 CMOS pixels) with an FP filter. (c) Spectral responses of the 14 wavelength channels.

630 nm. The spectral bandwidth is 10–15 nm. Among them, 595 nm was the highest, and 480 nm and 483 nm were the lowest wavelengths. The zoomed and unzoomed channels are simultaneously imaged once exposed on the 14-band SRDA to obtain the spatio-spectral data cube, as shown in Figure 1(c).

The foveated magnification of the foveated zoom imaging channel was determined by the axial positions of the variator and compensator. To achieve different foveated zoom ratios, the variator and compensator moved along the z-axis, parallel to the optical axis. By scanning the foveated zoom lens group in the x-y plane perpendicular to the optical axis, the LRIs of different lateral positions on the object are magnified, and its spatial resolution imaging is improved.

The goal of our design is to achieve a single system with two spatial channels in 14 bands. One channel has a constant peripheral focal length and wide FOV to search for pests and monitor complex environments. The other has a different foveated focal length and a higher spatial resolution, which is used to finely identify pests in the LRI. The two channels have the same image plane. Table 1 lists the specifications of the 9× foveated-zoom system.

The change in distance between the variator and compensator causes a change in the foveated focal length and magnification. When the system was zoomed locally, the focal length of the peripheral imaging channel remained unchanged. In our previous research, the foveated magnification was written as follows [6]:

TABLE 1: Specifications of a 9× LSHI.

Items	Specifications
Peripheral focal length	16 mm
Foveated focal lengths	16–144 mm
Foveated-zoom ratio	1×–9×
Peripheral channels F#	1.8
Peripheral channels FOV	31° × 23.4°
Peripheral spatial resolution (lp/mm)	23 (image space) 1.8 (object space)
Foveated spatial resolution (lp/mm)	23 (image space) 1.8–16 (object space)
Spectral resolution	10–15 nm

$$M = \frac{f_v (f_c - l_{ci})}{f_c (f_v + l_{vi})}, \quad (1)$$

where  $f_v$  and  $f_c$  are the focal lengths of the variator and the compensator, respectively.  $l_{vi}$  and  $l_{ci}$  are the object distance of the zoom and the image distance of the compensator, respectively, under different foveated focal lengths. The value of the foveated magnification is positive to ensure that the partially magnified image is not inverted in the image plane. The paraxial analysis solution was implemented using paraxial lenses in this system. The parameters of the 9× foveated zoom system were calculated according to the specifications in Table 1. The spatial resolution was a specification of the LSHI system tested at an object distance of 200 mm.

**2.2. Optical System and Performance Evaluation.** A foveated-zoom snapshot hyperspectral imaging system was designed as shown in Figure 3. The components of the system are a 16 mm focal length imaging lens, a 6 mm focal length variation, a -6 mm focal length compensator, and a relay lens with an object-to-image ratio of 1:1 and NA 2.5. The effective observation FOV was 31° × 23.4°. The peripheral channel F number is 1.8.

By controlling the three-axis translation stage, the variator and compensator were moved along the optical axis to change the foveated focal length and magnification. Figure 4 shows the distances of the variator from the relay lens, the imaging lens from the compensator, and the compensator from the variator at different magnifications. This movement distance was used in the subsequent experiments described in Sections 3 and 4. Simultaneously, the foveated zoom lens is scanned in the XOY plane, perpendicular to the optical axis (Z-axis), to obtain a foveated zoom with different FOVs.

The MTF (Modulation Transfer Function), representing the optical imaging quality of the peripheral and foveated channels at different magnifications, is shown in Figure 5. The cut-off frequency was 23 lp/mm based on the hyperspectral pixel size. Figure 5(a) shows the MTF of the full field of view of the peripheral channels, all above 0.7. Figures 5(b)–5(f) show the MTFs of the foveated channels at 2×, 3×, 5×, 7×, and 9× magnification, respectively, with foveated FOVs of 1.4°, 1.1°, 0.7°, 0.4°, and 0.3°. We adjusted the weight of the peripheral and foveated channels into

balanced and proper values to maintain their MTF values higher than 0.5 at 23 lp/mm. Hence, the image of this system can maintain good quality.

### 3. Experimental Setup and Performance Analysis

**3.1. LSHI Setup.** In Figure 6, a proof-of-concept experimental system is built to demonstrate the effectiveness of the proposed 9× foveated zoom snapshot hyperspectral imaging system. The system consists of a bandpass filter, an imaging lens with a focal length of 16 mm, a foveated zoom lens including a variator and a compensator, a three-axis translation stage, a relay lens with a magnification of 1×, and a 14-band SRDA. The bandpass filter comprises a short-pass filter (OD4-650 nm, Edmund Optics) and a long-pass filter (OD4-475, Edmund Optics), which limit the wavelength range to 475–630 nm in front of the imaging lens. The diameter of the foveated zoom lens was 3 mm. The variator and compensator were connected by a cage-type coaxial structure, and the distance between the two could be finely moved to achieve zoom imaging. The three-axis translation stage is used to scan the foveated zoom lens laterally in a plane perpendicular to the optical axis and to move the variator and compensator along the z-axis, axially parallel to the optical axis. The moving distance was determined according to the data shown in Figure 4. The scanning of the foveated zoom lens is designed to dynamically modulate the focal length of any foveated FOV. During the scan, the variator and the compensator remained coaxial.

**3.2. Spatial Resolution and Magnification.** The spatial resolution of the LSHI system is determined by imaging the target (GCG-0206, Daheng Optics) of the United States Air Force (USAF) test chart in 1951, as shown in Figure 7(a). In general, we use image space resolution to evaluate the spatial resolution of an imaging system. Therefore, we converted the measured spatial resolution of the object space into the spatial resolution of the image space through vertical axis magnification. The LSHI system imaged the USAF test chart at an object distance of 200 mm. In this case, the vertical-axis magnification of the 1× LSHI system was 1/12.7. The pixel-number-limited image resolution was approximately 22 μm in the peripheral channel. Therefore, the Nyquist frequency was calculated as  $\nu_{ny} = 1/(2\delta_{px})$  using the pixel pitch  $\delta_{px}$  of the image sensor [42]. A Nyquist frequency of approximately 23 lp/mm should correspond to an object space spatial resolution of 1.8 lp/mm at an object distance of 200 mm. On the USAF test chart, the 0th group, 6th element corresponds to a resolution of 1.8 lp/mm. The imaging result of the peripheral channel of the LSHI system without a foveated zoom lens group in the USAF test chart is shown in Figure 7(b). The 1st group 1st element (position ① in Figures 7(a)–7(b)) is the largest resolvable fringe, and its corresponding object space spatial resolution is 2 lp/mm, which meets the design value of 1.8 lp/mm. When the foveated zoom lens group is working, the LSHI system achieves 9× foveated magnification at ② in (a), and the

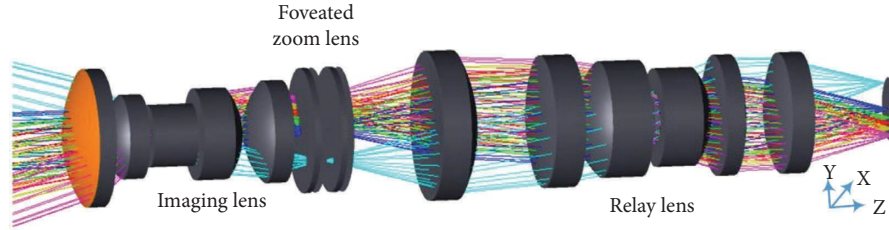


FIGURE 3: Optical layouts of LSHI.

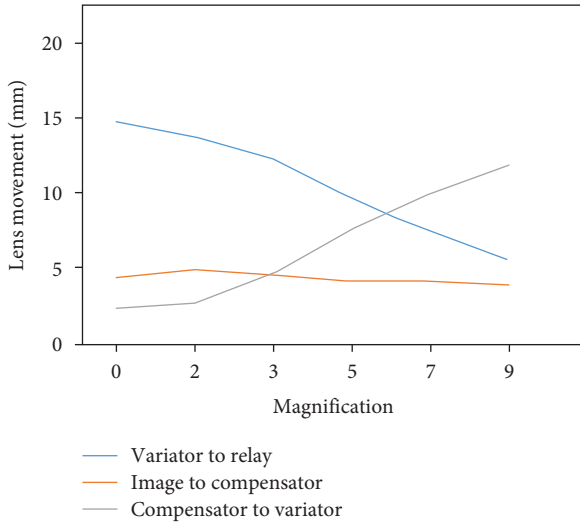


FIGURE 4: Movement distance of the variator and compensator.

imaging result is shown in Figure 7(c). When only the peripheral channel is working, the result shows that the ② in Figure 7(b) cannot be distinguished. When the foveated zoom lens group is working, the stripes in box ② in Figure 7(b) can distinguish the 4th group's 1st element (shown in the red box in Figure 7(c)), and the corresponding object space spatial resolution is 16 lp/mm. In the  $9\times$  LSHI system, because the vertical axis magnification is  $1/1.41$  ( $9\times 1/12.7$ ), the image space spatial resolution of the  $9\times$  LSHI system is 22.56 lp/mm, which almost meets the design value of 23 lp/mm. The resolution in the object space is 2–16 lp/mm, which also meets the design value of 1.8–16 lp/mm.

To calculate the foveated magnification, we selected a grid calibration chart (GCG-020606, Daheng Optics) as the object to be imaged. Figure 8 shows foveated magnified images. The peripheral channel of the LSHI system without foveated zoom lens imaging results from the grid calibration plate is shown in Figure 8(a), where each grid (in the red box) occupies five pixels. When the foveated zoom lens is working, the LSHI system realizes foveated magnification at the LRI, as shown in Figure 8(a), and the imaging results are shown in Figures 8(b)–8(h). Foveated magnification is the ratio of the grid in the foveated zoom area to the grid in the peripheral area. Foveated magnification can be easily obtained by comparing the grids in the two areas. By analyzing the imaging results, a grid in the red box in Figures 8(b)–8(h) occupies 10, 13, 18, 22, 30, 35, and 47 pixels, respectively.

Therefore, the foveated magnifications of Figures 8(b)–8(h) are, respectively  $2\times$ ,  $2.6\times$ ,  $3.6\times$ ,  $4.4\times$ ,  $6\times$ ,  $7\times$ ,  $9.4\times$ . When the distance between the variator and compensator moves according to the curve in Figure 4, continuous  $1\times$ – $9.4\times$  foveated zoom imaging can be obtained.

**3.3. Spectral Imaging Performance and Spectral Angle Mapping Classification.** A standard color checker (X-rite) was used as the target scene to test the spectral resolution capability of the LSHI system, which was fixed at 24 color blocks. To eliminate background noise, a completely dark pattern was projected onto the target, and then the reflected light captured by the SRDA was averaged and subtracted in the following measurement. Figure 9 shows the X-rite imaging results and the analysis. Figures 9(a)–9(h) show selected images corresponding to the eight wavelengths from the raw hyperspectral stack. The results show that the intensity of each color block under different wavebands has differences that can be distinguished. The recovered color image and representative spectrum are shown in Figures 9(i) and 9(k), respectively. The spectrum shows the reflection characteristics of the different colors. Red (typically regions 3, 10, 11, and 13) reflects more at long wavelengths. Blue (typically regions 2 and 12) reflects shorter wavelengths. Yellow and green (typically regions 1, 5, and 7) have a higher reflectance at mid-wavelengths, whereas black (typically region 8) has a flat spectrum. The white circle in (i) represents the imaging result of the foveated zoom channel. Figures 9(a)–9(h) and 9(k) show that the reflectivity of the color card in the white circle at 607–630 nm is high, which is consistent with the actual situation of the red color block. Therefore, we verified that both the peripheral and foveated zoom channels could obtain clear spectral imaging.

The spectrum of a single pixel can be described as a vector in an  $n$ -dimensional space, where  $n$  is the number of spectral bands. Each vector has a specific length that represents the brightness of the pixel and a direction that represents the spectral shape of the pixel. The more similar the spectra, the better their correlation and the smaller the spectral angle between them. This method is called spectral angle mapping (SAM) [43]. In this study, we used the SAM for classification. The spectral angle between the pixels of the image and the pixels of the selected area was calculated. When the pixel difference was less than the set maximum angle, they were assigned to the same category. We labeled and extracted the pixels of region 1–13 in Figure 9(i). Figure 9(j) is obtained after the SAM classification. The results show that this classification method can accurately

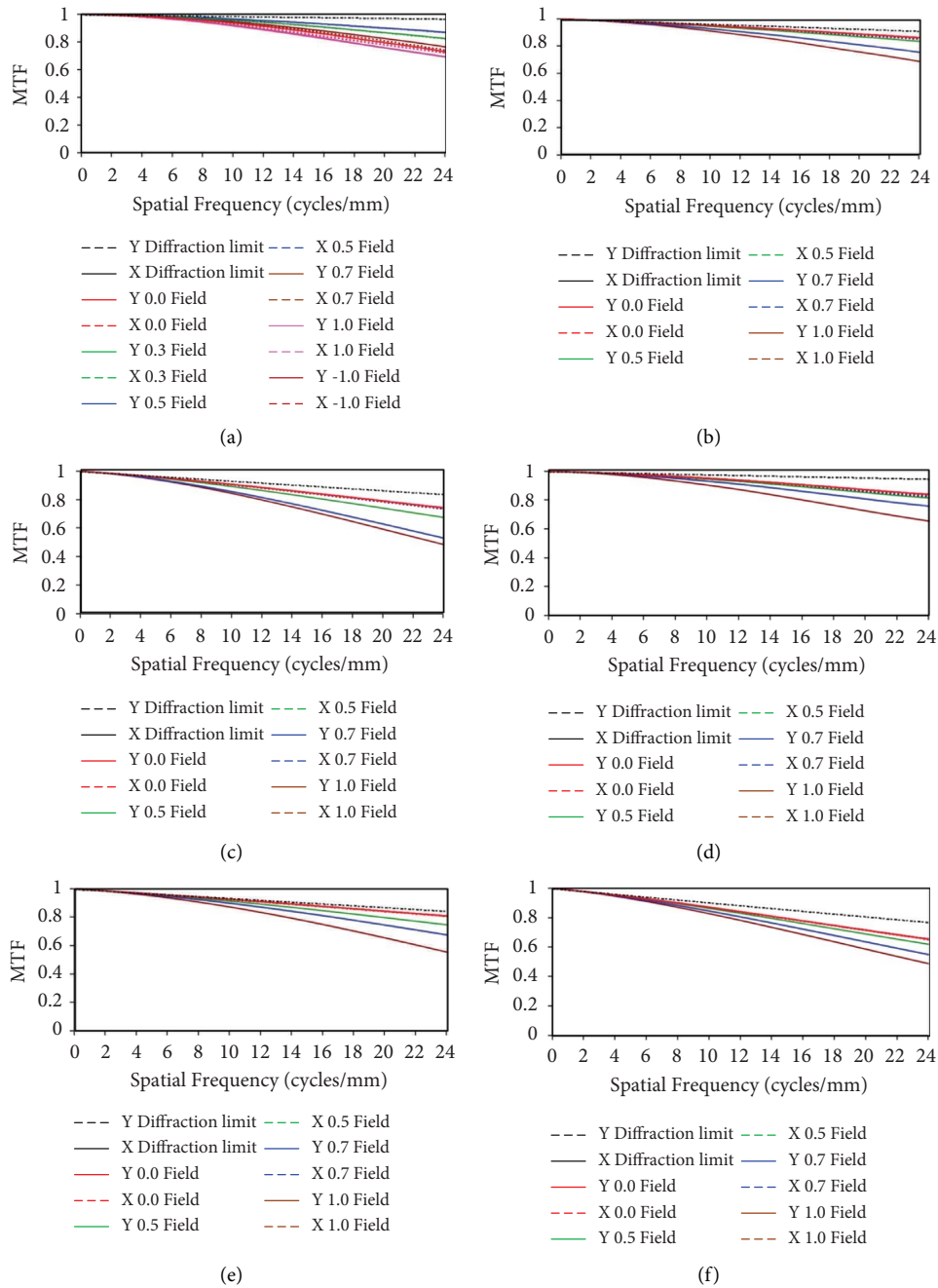


FIGURE 5: MTF of the periphery and foveated channels at different magnifications. (a) MTF of the peripheral channel in full field. (b)–(f) MTF of foveated channels in the foveated field at 2 $\times$ , 3 $\times$ , 5 $\times$ , 7 $\times$ , and 9 $\times$  magnifications, respectively.

classify the color blocks corresponding to the selected 13 regions. This will be of significant help in our next fruit classification experiment.

#### 4. Fruit Pest Monitoring Application

**4.1. *Drosophila* Foveated Zoom Hyperspectral Fruit Imaging Classification.** The proof-of-concept experiment system built as shown in Figure 6 continues to be used to image fruits and *Drosophila* on fruits. In this experiment, we used full-spectrum halogen lamps as the illumination sources to

simulate the outdoor daylight environment. Two 25W halogen lamps were placed on both sides of the scene to obtain uniform illumination. The target scene consisted of mixed red and green grapes, cherry tomatoes, and two *Drosophila* specimens. The foveated zoom channel performs 9 $\times$  foveated zoom imaging on one of the *Drosophila* species, and the peripheral channel images the entire scene. Figures 10(a)–10(h) show selected spectral images corresponding to the eight wavelengths from the raw hyperspectral stack. The recovered color image and representative spectrum are shown in Figures 10(i) and 10(j), respectively.



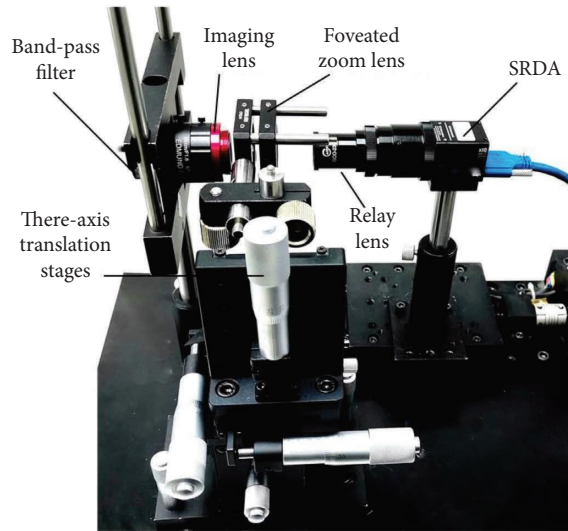


FIGURE 6: Foveated-zoom snapshot hyperspectral imaging system experimental setup.

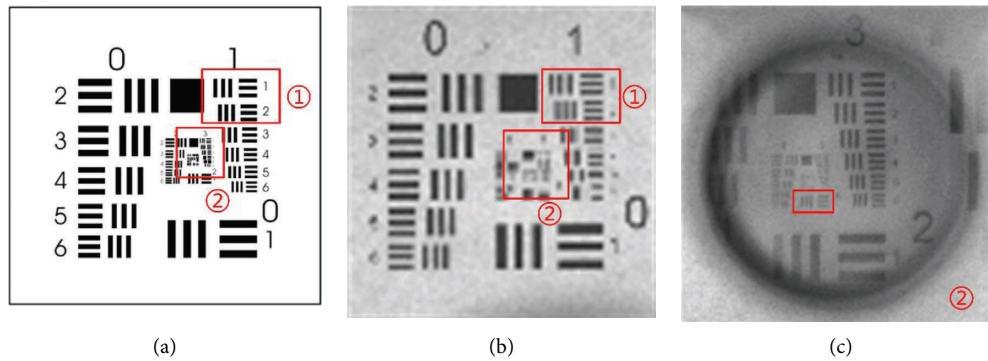


FIGURE 7: USAF test chart imaging results. (a) USAF test chart in 1951. (b) Imaging result of the peripheral channel. (c) The imaging result of 9× foveated zoom channel.

The spectra show the absorption signatures of different fruits and *Drosophila* species. The reflectance of green grapes is higher in the 520–583 nm band, and it absorbs more at 482–508 nm. The spectral curve trends of the red grapes and cherry tomatoes were roughly similar. In particular, the reflectivity of red grapes was higher in the 479–620 nm band and lower in the 620–630 nm band. *Drosophila* specimens exhibited the highest reflectivity at 546 nm. Therefore, it is better to monitor and identify *Drosophila* using this band.

Figure 11 shows the imaging results of *Drosophila* with six different focal lengths in the 546 nm band. The result of the peripheral channel without the foveated zoom lens of the LSHI system imaging the target scene is shown in Figure 11(a). *Drosophila* occupies five pixels and can be detected but cannot be identified. However, when the foveated zoom lens group works, the LSHI system realizes the foveated magnification of *Drosophila* in Figure 11(a), and the imaging results are shown in Figures 11(b)–11(f). *Drosophila* in Figures 11(b)–11(f) account for 13, 20, 27, 33, and 45 pixels, respectively. Therefore, the foveated magnifications of Figures 11(b)–11(h) were 2.6×, 4×, 5.4×, 6.6×, and 9×, respectively. When the foveated magnification is

2.6×, *Drosophila* can be identified, as shown in Figure 11(b). When the foveated magnification is 4× or higher, *Drosophila* can not only be recognized, but its details (wings and tentacles) can be distinguished, as shown in Figures 11(c)–11(h).

**4.2. Image Color Contrast Enhancement and Fruit Classification.** To better distinguish different fruits during the color image restoration process. We enhanced the contrast of the different fruits during the color image restoration process. We used two-color image restoration methods. One is to create artificial color images and apply the Commission Internationale de L' Eclairage (CIE) curve to spectral data. The results are shown in Figure 12(a). The other is called false color. A specific single band of the target is assigned to the red, green, and blue color channels to obtain a high-contrast rendered color image. We chose different bands as reference wavelengths for the red, green, and blue channels to enhance the weight of the selected band. The color restoration results when the red, green, and blue channels chose 607, 533, and 482 nm as reference wavelengths, respectively, are shown in Figure 12(b). It can

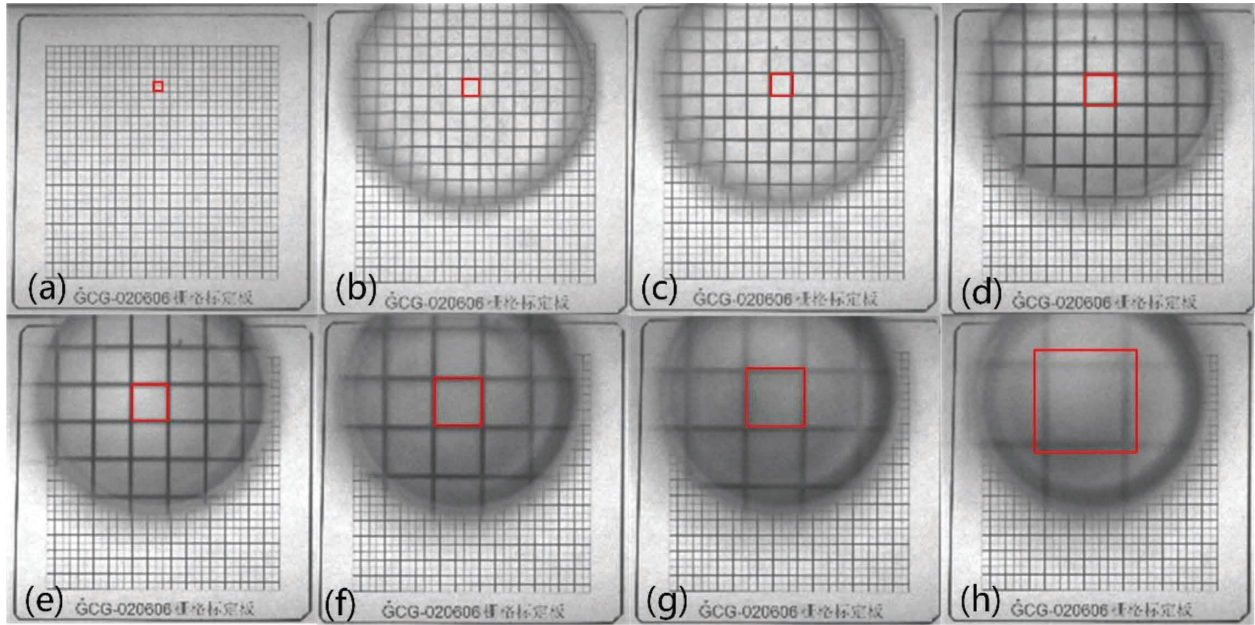


FIGURE 8: Grid calibration chart imaging results. (a) 16 mm focal length peripheral channel imaging without foveated zoom lens. (b)–(h) 2 $\times$ , 2.6 $\times$ , 3.6 $\times$ , 4.4 $\times$ , 6 $\times$ , 7 $\times$ , 9.4 $\times$  foveated zoom imaging with a foveated zoom lens.

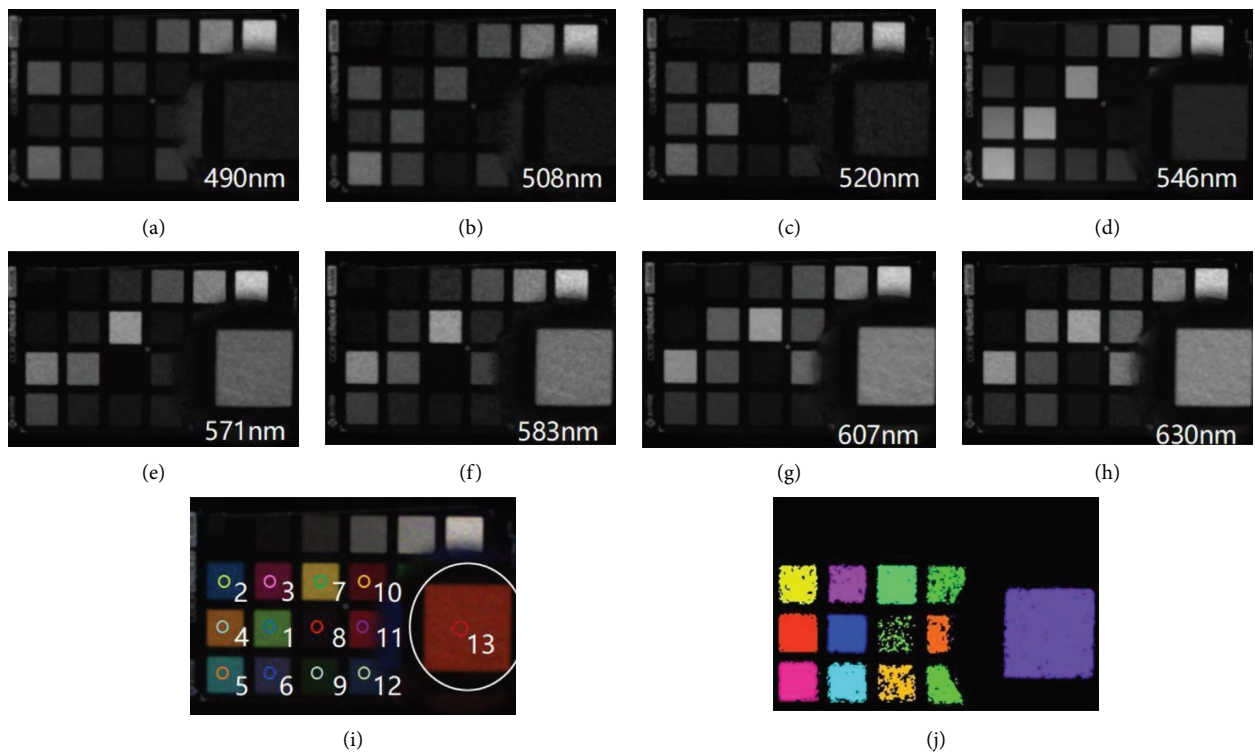
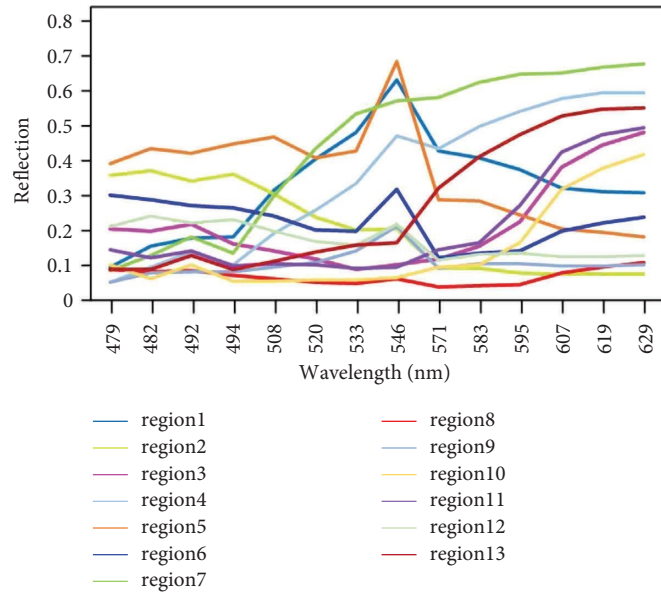


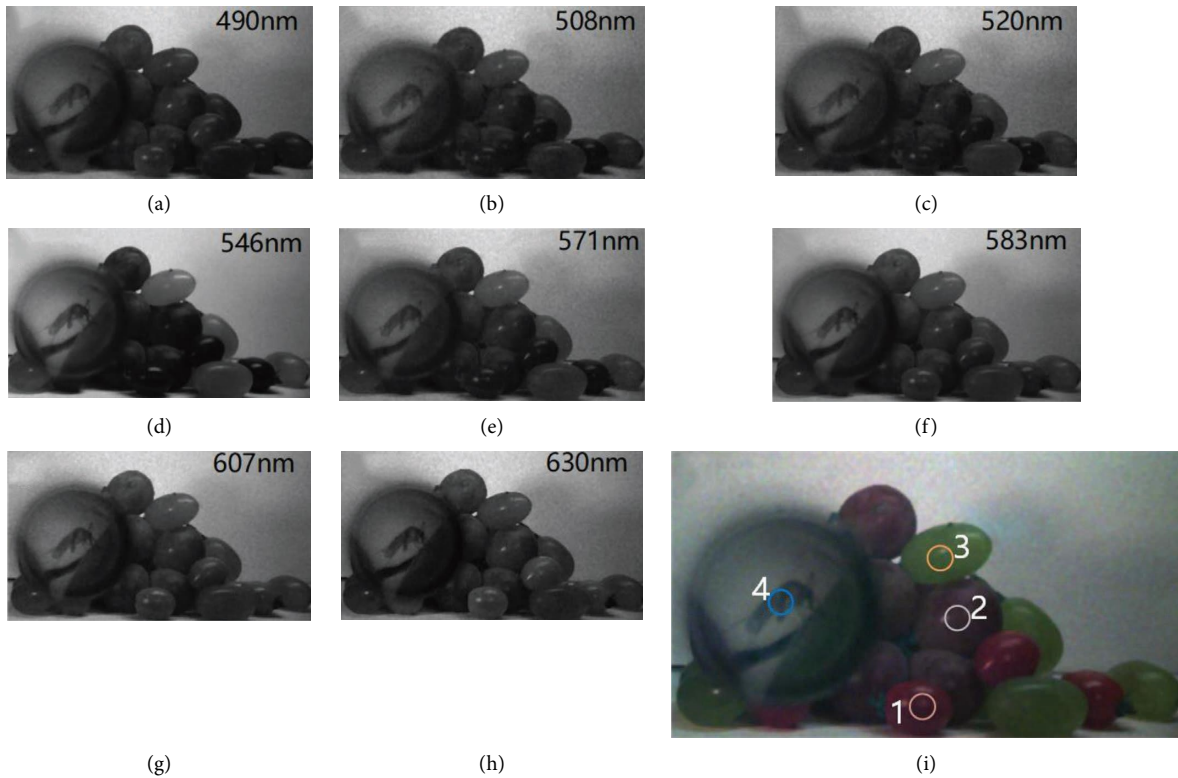
FIGURE 9: Continued.





(k)

FIGURE 9: X-rite color card hyperspectral imaging results and analysis. (a)–(h) X-rite images from 8 selected spectral channels. (i) Color image recovered from 16 spectral channels. (j) SAM classification image. (k) Representative spectra from the 13 regions labeled in (i).



(g) (h) (i)  
FIGURE 10: Continued.

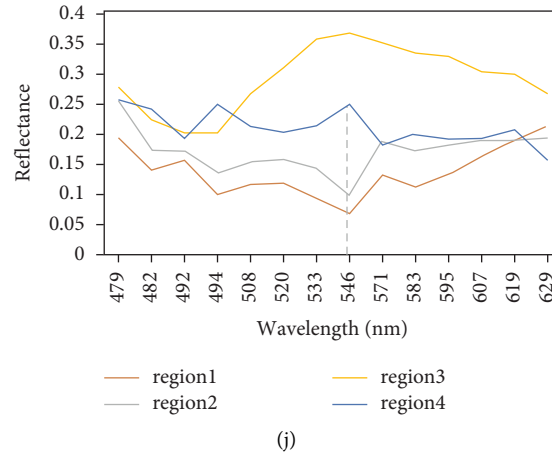


FIGURE 10: *Drosophila* and fruit hyperspectral imaging results and analysis. (a)–(h) *Drosophila* and fruit images from 8 selected spectral channels. (i) Color image recovered from 16 spectral channels. (j) Representative spectra from the 4 regions labeled in (i).

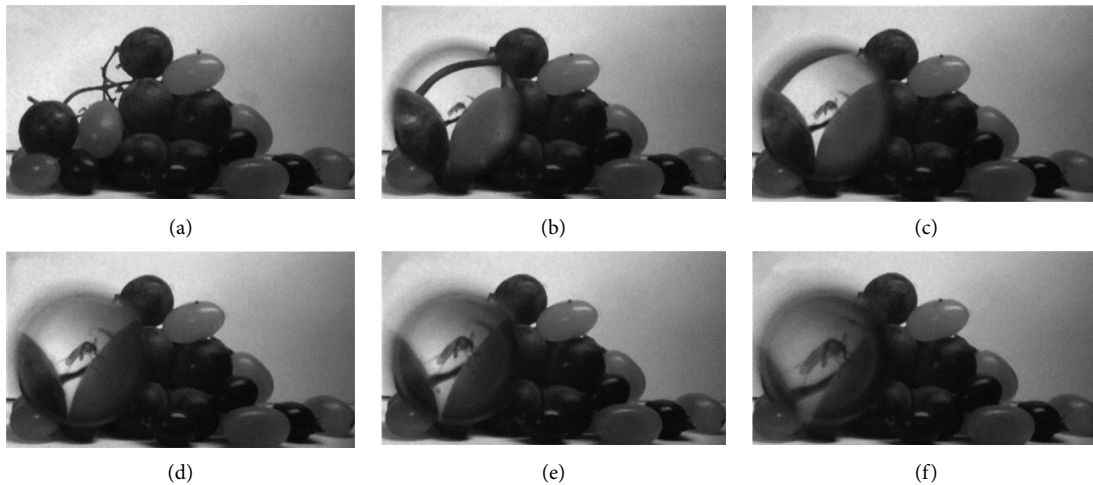


FIGURE 11: *Drosophila* imaging at different foveated focal lengths at 546 nm. (a) 16 mm focal length peripheral channel imaging without a foveated zoom lens group. (b)–(f) 2.6 $\times$ , 4 $\times$ , 5.4 $\times$ , 6.6 $\times$ , 9 $\times$ foveated zoom imaging with foveated zoom lens group.

be observed from Figures 10(h) and 10(j) that the reflectivity of cherry tomatoes at 629 nm is relatively high. To enhance the contrast of the cherry tomatoes, 629 nm was selected as the reference wavelength for the red channel. The results are shown in Figure 12(c). Similarly, we chose 546 nm as the reference wavelength for the green channel to enhance the contrast of the green grapes. The result is shown in Figure 12(d). Compared with the color image in Figures 12(b)–12(d), which show a higher color contrast of cherry tomatoes and green grapes for better visualization. Comparing the two-color restoration methods, the first restoration is a continuous waveband, which results in less image noise. The second restoration method can select different weighted wavelengths of red, green, and blue according to the requirements, which provides an effective way to enhance contrast. Note that by optimizing the color arrangement or full-spectrum analysis, the contrast of objects of interest in complex scenes can be further increased,

where multiple wavelengths can form a color channel together. Furthermore, this method plays an important role in real-time monitoring.

The LSHI system can achieve partial zooming of any LRI in an entire target scene. Region 1 in the restored color image in Figure 13(a) is locally zoomed, as shown in Figures 10 and 11. Therefore, we implemented a foveated zoom for region 2 through the three-axis translation stage, and the result of the color recovery image is shown in Figure 13(b). We classify the three types of fruit in Figure 13(a), and the results are shown in Figure 13(c). Both the green and red grapes in the target scene were classified and identified. However, only two cherry tomatoes were classified and labeled because of their lower spectral response of cherry tomatoes. When the light source is unevenly illuminated, the spectral angle of some cherry tomatoes will be greater than the threshold, causing these tomatoes to not be assigned to the same category. In addition, the two *Drosophila* species occupy too

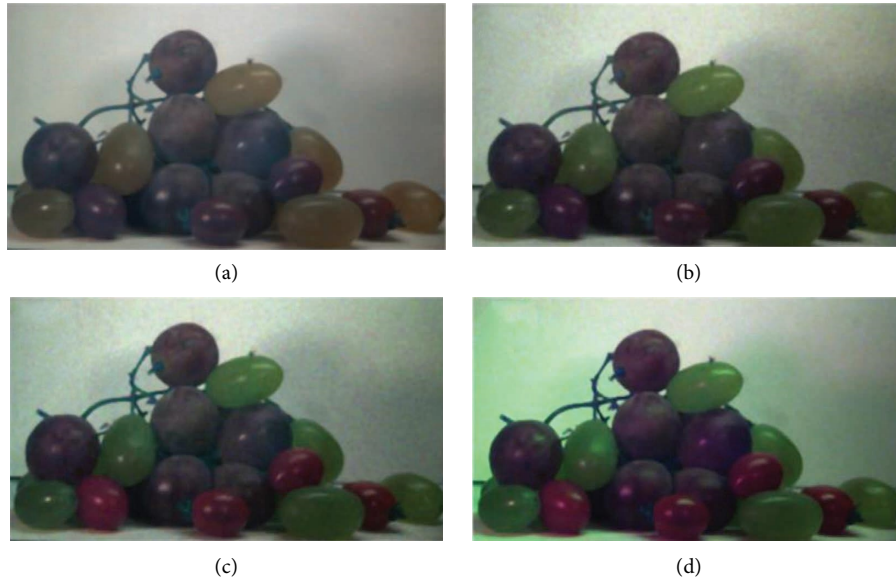


FIGURE 12: Image color recovery and fruits contrast enhancement. (a) Artificial color image using CIE curve. (b) False color with reference wavelengths of 607, 533, and 482 nm. (c) False color with reference wavelengths of 629, 533, and 482 nm. (d) False color with reference wavelengths of 629, 546, and 482 nm.

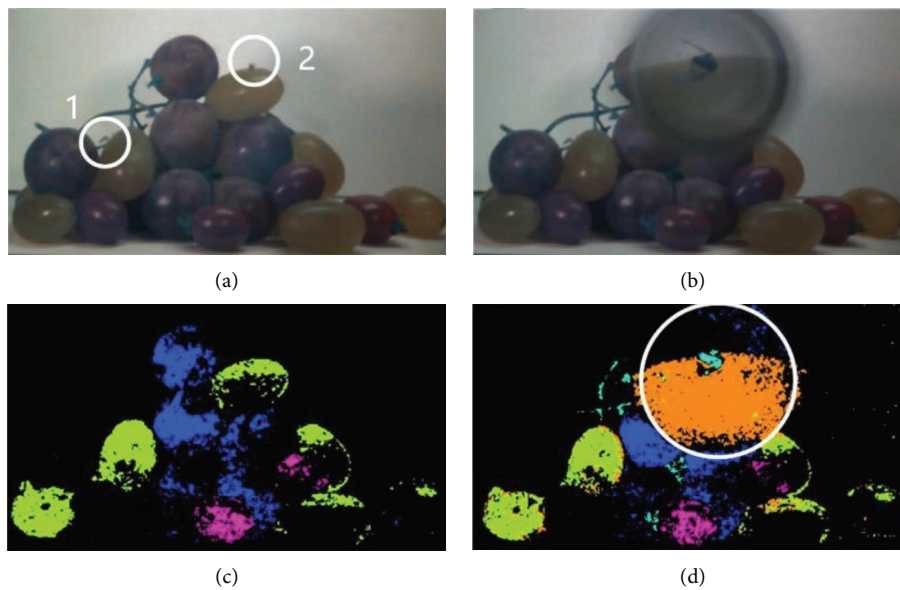


FIGURE 13: *Drosophila* and fruit classification. (a) Artificial color image without a foveated zoom lens group. (b) Foveated magnified color image of *Drosophila* in region 2. (c) Classification image of (a). (d) Classification image of (d).

few pixels, resulting in them not being classified and marked in Figure 13(c). However, when *Drosophila* is foveatedly magnified, they can be classified and labeled, as shown in Figure 13(d). Figure 13(d) is the classification diagram of Figure 13(b). The white circle shows that the *Drosophila* (aquamarine) and the fruit (orange) are classified and labeled. The *Drosophila* in region 1 that has not been locally magnified cannot be classified due to its low spatial resolution. Therefore, foveated zoom channels play an important role in the spectrum's classification and recognition.

## 5. Conclusion

In this study, we proposed a  $9\times$  foveated zoom snapshot hyperspectral imaging system using commercial SRDA to monitor and identify fruit pests and fruit classification. We achieved a  $9\times$  foveated zoom by controlling the three-axis translation stage to change the distance between the variator and the compensator. The spatial resolution of the LSHI system has been increased from 2 lp/mm to 16 lp/mm using foveated zoom channels, which significantly compensates

for the low spatial resolution of the SRDA. Using the X-rite color card as the imaging target, we verified that the LSHI system can simultaneously obtain 14-band spectral images of the peripheral and foveated channels from 479 nm to 630 nm with a spectral bandwidth of 10–15 nm through a single exposure. We also performed 9× foveated zoom hyperspectral imaging of certain fruits and *Drosophila* species. We successfully obtained 14 spectral bands and spectral response curves for red grapes, green grapes, cherry tomatoes, and *Drosophila*. We also showed two different color restoration results. Color-enhanced fruit images presented better contrast than those without using specific reference bands. Finally, SAM was used for the classification. The results show that the use of foveated zoom channels is conducive to the accurate classification of targets. This study is suitable for online fruit classification and pest identification. In addition, if the three-axis translation stage is replaced with a lighter and smaller precision controller in the future, the demonstrator will be particularly suitable for remote sensing applications with drones, such as agricultural and environmental monitoring, because the system is compact, lightweight, and real-time.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

Key Laboratory of Optical System Advanced Manufacturing Technology, Chinese Academy of Sciences (2022KLOMT02-01).

### References

- [1] B. Škaloudová, V. Křivan, and R. Zemek, “Computer-assisted estimation of leaf damage caused by spider mites,” *Computers and Electronics in Agriculture*, vol. 53, no. 2, pp. 81–91, 2006.
- [2] F. Fina, P. Birch, R. Young, J. Obu, B. Faithpraise, and C. Chatwin, “Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters,” *International Journal of Advanced Biotechnology and Research*, vol. 4, pp. 189–199, 2013.
- [3] M. A. Ebrahimi, M.-H. Khoshtaghaza, S. Minaei, and B. Jamshidi, “Vision-based pest detection based on SVM classification method,” *Computers and Electronics in Agriculture*, vol. 137, pp. 52–58, 2017.
- [4] P. Boissard, V. Martin, and S. Moisan, “A cognitive vision approach to early pest detection in greenhouse crops,” *Computers and Electronics in Agriculture*, vol. 62, no. 2, pp. 81–93, 2008.
- [5] J. Zhu, A. Wu, X. Wang, and H. Zhang, “Identification of grape diseases using image analysis and BP neural networks,” *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 14539–14551, 2020.
- [6] Z. Gao, Y. Shao, G. Xuan, Y. Wang, Y. Liu, and X. Han, “Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning,” *Artificial Intelligence in Agriculture*, vol. 4, pp. 31–38, 2020.
- [7] N. Ekramirad, A. Y. Khaled, L. E. Doyle et al., “Nondestructive detection of codling moth infestation in apples using pixel-based NIR hyperspectral imaging with machine learning and feature selection,” *Foods*, vol. 11, no. 1, p. 8, 2021.
- [8] C. A. Sánchez-Ríos, A. M. Cardona, A. Echavarría, and J. Herrera-Ramírez, “Detection of effects of three different plant pests for chrysanthemum flowers by using image processing and spectral imaging,” *Journal de Physique: Conf. Ser.* vol. 1547, no. 1, Article ID 012015, 2020.
- [9] T. Adão, J. Hruška, L. Pádua et al., “Hyperspectral imaging: a review on UAV-based sensors, data processing and applications for agriculture and forestry,” *Remote Sensing*, vol. 9, no. 11, p. 1110, 2017.
- [10] D. Caballero, R. Calvini, and J. M. Amigo, “Hyperspectral imaging in crop fields: precision agriculture,” in *Data Handling in Science and Technology*, pp. 453–473, Elsevier, Amsterdam, Netherlands, 2020.
- [11] P. Moghadam, D. Ward, E. Goan, S. Jayawardena, P. Sikka, and E. Hernandez, “Plant disease detection using hyperspectral imaging,” in *Proceedings of the 2017 International Conference on Digital Image Computing: Techniques and Applications*, pp. 1–8, DICTA), Sydney, NSW, Australia, November 2017.
- [12] S. Pacheco and R. Liang, “Snapshot, reconfigurable multi-spectral and multi-polarization telecentric imaging system,” *Optics Express*, vol. 22, no. 13, pp. 16377–16385, 2014.
- [13] Y. Oiknine, I. August, and A. Stern, “Multi-aperture snapshot compressive hyperspectral camera,” *Optics Letters*, vol. 43, no. 20, pp. 5042–5045, 2018.
- [14] Y. Zhao, T. Yue, L. Chen et al., “Heterogeneous camera array for multispectral light field imaging,” *Optics Express*, vol. 25, no. 13, pp. 14008–14022, 2017.
- [15] J. Wu, B. Xiong, X. Lin, J. He, J. Suo, and Q. Dai, “Snapshot hyperspectral volumetric microscopy,” *Scientific Reports*, vol. 6, pp. 24624–24710, 2016.
- [16] M. W. Kudenov, M. Miskiewicz, N. Sanders, and M. J. Escuti, “Achromatic Wollaston prism beam splitter using polarization gratings,” *Optics Letters*, vol. 41, no. 19, pp. 4461–4463, 2016.
- [17] J. Chamoun and S. Farhadi, “Hyperspectral imaging using a Wollaston prism,” in *Proceedings of the Optics and Photonics for Sensing the Environment*, pp. EM2C–5, Optical Society of America, Washington, DC, US, January 2020.
- [18] B. F. Bathel and J. M. Weisberger, “Compact, self-aligned focusing schlieren system,” *Optics Letters*, vol. 46, no. 14, pp. 3328–3331, 2021.
- [19] Y. Wang, M. E. Pawlowski, and T. S. Tkaczyk, “High spatial sampling light-guide snapshot spectrometer,” *Optical Engineering*, vol. 56, no. 8, Article ID 081803, 2017.
- [20] J. G. Dwight and T. S. Tkaczyk, “Lenslet array tunable snapshot imaging spectrometer (LATIS) for hyperspectral fluorescence microscopy,” *Biomedical Optics Express*, vol. 8, no. 3, pp. 1950–1964, 2017.
- [21] M. Hubold, E. Montag, R. Berlich, R. Brunner, and R. Brünig, “Multi-aperture system approach for snapshot multispectral imaging applications,” *Optics Express*, vol. 29, no. 5, pp. 7361–7378, 2021.
- [22] T. Mu, F. Han, H. Li et al., “Snapshot hyperspectral imaging polarimetry with full spectropolarimetric resolution,” *Optics and Lasers in Engineering*, vol. 148, Article ID 106767, 2022.

- [23] T. Mu, F. Han, D. Bao, C. Zhang, and R. Liang, "Compact snapshot optically replicating and remapping imaging spectrometer (ORRIS) using a focal plane continuous variable filter," *Optics Letters*, vol. 44, no. 5, pp. 1281–1284, 2019.
- [24] W. Feng, H. Rueda, C. Fu, G. R. Arce, W. He, and Q. Chen, "3D compressive spectral integral imaging," *Optics Express*, vol. 24, no. 22, pp. 24859–24871, 2016.
- [25] Z. He, N. Williamson, C. D. Smith, M. Gragston, and Z. Zhang, "Compressed single-shot hyperspectral imaging for combustion diagnostics," *Applied Optics*, vol. 59, no. 17, pp. 5226–5233, 2020.
- [26] J. Bao and M. G. Bawendi, "A colloidal quantum dot spectrometer," *Nature*, vol. 523, no. 7558, pp. 67–70, 2015.
- [27] X. Zhu, L. Bian, H. Fu et al., "Broadband perovskite quantum dot spectrometer beyond human visual resolution," *Light: Science & Applications*, vol. 9, no. 1, p. 73, 2020.
- [28] Z. Wang, S. Yi, A. Chen et al., "Single-shot on-chip spectral sensors based on photonic crystal slabs," *Nature Communications*, vol. 10, pp. 1020–1026, 2019.
- [29] A. S. Luthman, S. Dumitru, I. Quiros-Gonzalez, J. Joseph, and S. E. Bohndiek, "Fluorescence hyperspectral imaging (fHSI) using a spectrally resolved detector array," *Journal of Biophotonics*, vol. 10, no. 6-7, pp. 840–853, 2017.
- [30] H. Li, W. Liu, B. Dong, J. V. Kaluzny, A. A. Fawzi, and H. F. Zhang, "Snapshot hyperspectral retinal imaging using compact spectral resolving detector array," *Journal of Biophotonics*, vol. 10, no. 6-7, pp. 830–839, 2017.
- [31] X. Liu, J. Chang, W. L. Chen et al., "A dynamic foveated infrared imager for surveillance," *Optics and Lasers in Engineering*, vol. 124, Article ID 105825, 2020.
- [32] B. Shen, J. Chang, Y. Niu, W. Chen, and Z. Ji, "Deformable mirror-based optical design of dynamic local athermal longwave infrared optical systems," *Optics and Lasers in Engineering*, vol. 106, pp. 1–9, 2018.
- [33] X. Wang, J. Chang, Y. Niu et al., "Design and demonstration of a foveated imaging system with reflective spatial light modulator," *Frontiers of Optoelectronics*, vol. 10, no. 1, pp. 89–94, 2017.
- [34] G. Curatu and J. E. Harvey, "Lens design and system optimization for foveated imaging," in *Current Developments in Lens Design and Optical Engineering IX* pp. 170–178, SPIE, San Diego, CA, USA, 2008.
- [35] X. Dallaire and S. Thibault, "Wide-angle lens miniaturization through foveated imaging," in *Optical Systems Design 2015: Optical Design and Engineering VI* pp. 323–328, SPIE, Bellingham, Washington, USA, 2015, <https://doi.org/10.1117/12.2191359>.
- [36] N. Hagen and T. S. Tkaczyk, "Foveated endoscopic lens," *Journal of Biomedical Optics*, vol. 17, no. 02, 2012.
- [37] K. Kuniyoshi, N. Kita, K. Sugimoto, S. Nakamura, and T. Suehiro, "A foveated wide angle lens for active vision," in *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, pp. 2982–2988, IEEE, Nagoya, Japan, May 1995.
- [38] C. Xu, D. Cheng, J. Chen, and Y. Wang, "Design of all-reflective dual-channel foveated imaging systems based on freeform optics," *Applied Optics*, vol. 55, no. 9, pp. 2353–2362, 2016.
- [39] J. I. Katz, S. y. Lee, and H. Hua, "Improved multi-resolution foveated laparoscope with real-time digital transverse chromatic correction," *Applied Optics*, vol. 59, no. 22, p. G79, 2020.
- [40] J. Xiong, C. Yoo, S. Moon et al., "Foveated display system based on a doublet geometric phase lens," *Optics Express*, vol. 28, no. 16, pp. 23690–23702, 2020.
- [41] B. Shen, J. Chang, C. Wu et al., "Local zoom system for agricultural pest detection and recognition," *Applied Physics B*, vol. 124, no. 11, p. 219, 2018.
- [42] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Prentice-Hall, New York, USA, 3rd Edition, 2006.
- [43] O. A. De Carvalho and P. R. Meneses, "Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM)," in *Proceedings of the Summaries of the 9th JPL Airborne Earth Science Workshop*, JPL publication Pasadena, Pasadena, California, January 2000.