University of Kent

Kent Academic Repository

Huang, Yizhi, Hossain, Md. Moinul, Liu, Yan, Sun, Kai, Zhang, Biao and Xu, Chuanlong (2024) ***Data rectification and decoding of a microlens array-based multi-spectral light field imaging system.*** **Optics and Lasers in Engineering, 180 . ISSN 0143-8166.**

Downloaded from

https://kar.kent.ac.uk/106066/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.1016/j.optlaseng.2024.108327

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

Additional information

For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

Title:

Data rectification and decoding of a microlens array-based multi-spectral light field imaging system

The type of article: Full-length Paper

Authors and affiliations:

Yizhi Huang[1], Md. Moinul Hossain[2], Yan Liu[3, 4], Kai Sun[3], Biao Zhang[1],

Chuanlong Xu[1, *]

1. National Engineering Research Center of Power Generation Control and Safety, School of Energy and Environment, Southeast University, Nanjing, 210096, China
2. School of Engineering, University of Kent, Canterbury, Kent, CT2 7NT, UK
3. Jiangsu Zhongke Research Center for Clean Energy and Power, Lianyungang 222000，China
4. The Institute of Engineering Thermophysics, Chinese Academy of Sciences，Beijing 100190, China

E-mail address:

230189043@seu.edu.cn (Y. Huang)
M.Hossain@kent.ac.uk (M. M. Hossain)
liuyan@iet.cn (Y. Liu)
sunkai@jscep.ac.cn (K. Sun)
zhangbiao@seu.edu.cn (B. Zhang)
chuanlongxu@seu.edu.cn (C. Xu)

* Corresponding author：
Name: Chuanglong Xu
Address: School of Energy and Environment, Southeast University, Nanjing, 210096, China
Tel&Fax: +86-025-83794395
E-mail address: chuanlongxu@seu.edu.cn

Data rectification and decoding of a microlens array-based multi-spectral light field imaging system

Abstract: Spectral light field imaging systems have gained emerging attention for their capacity to capture multidimensional spatial-spectral data in a single snapshot. However, the spectral data poses significant challenges to image decoding methods for existing light field imaging due to the aberrations and spectral aliasing. To resolve these challenges, in this study, a microlens array-based light field spectral imaging system is designed and assembled. A novel joint decoding method incorporating aberration correction for spatial-spectral data is proposed. The proposed method is implemented through a novel set of aberrations assessment and correction procedures for both local and global non-uniform aberrations. An improved light field digital refocusing method is used to overcome the spectral aliasing in refocusing the spectral light field images, enabling the reconstruction of depth and spectral intensity distribution of a target scene. Experimental verification is conducted on a uniform light source and real-world imagery. Results demonstrate that the proposed method can provide accurate aberration correction and reconstruction of the spectral intensity distribution of the targeted scene under different depths, thus showing the effectiveness of the proposed light field spectrometer and spatial-spectral joint decoding method for real-world imagery.

Keywords: Spectral light field imaging, Aberration correction, Data decoding, Depth reconstruction, Spectral intensity reconstruction.

1. Introduction

Spectral imaging technology (SIT) significantly enhances the dimensionality of image data obtained by multiple narrow-band wavelengths. Its application is apparent across diverse industrial processes such as medicine pill uniformity, packaging cleanliness and non-destructive semiconductor wafer quality testing. The SIT, therefore, has enhanced the quality control standards significantly in these industrial processes [1-6]. Traditional SIT such as the grating dispersion spectrometer and the variable filter spectrometer requires line-by-line scanning or wavelength-by-wavelength to generate a complete three-dimensional (3-D) spectral data cube (i.e., 2-D spatial light intensity and 1-D spectrum distribution) [7, 8]. However, this scanning procedure is time-consuming, which makes the traditional SIT unsuitable for complex industrial applications such as diagnosing turbulent flame combustion or performing temperature measurements of engine turbine blade surface. In recent years, snapshot-based SITs such as encompass coded aperture [9], computed tomography [10] and spatial sampling spectral imaging [11, 12] have made considerable progress and are applied in complex measurement applications [9-13]. These technologies disperse the spectral and spatial information (or spectral interference information) of the target to different sensor positions simultaneously and facilitate the high-resolution spectral imaging of complex scenes.

Spectral light field imaging (SLFI) is a single snapshot-based spatial sampling technology and recently various progress has been made in this technology [8, 14-16]. This technology incorporates a spectral filter array (SFA) into the optical path and each light field (LF) sub-image contains both the angular and spectral information of light rays, thus, facilitating the five-dimensional (i.e., 2-D light intensity spatial distribution, 2-D light ray angular information and 1-D spectral distribution information) spectral LF information in a single exposure [17-22]. The angular information of light rays can be utilized to achieve depths and 3-D measurements of the target scene. This technology provides a reliable system structure, straightforward spectral acquisition without any kind of interpolation and high data reliability.

A joint decoding technique (spatial-spectral data) is necessary to distillate the 3-D spectral data cube from the 5-D SLFI data [22, 25]. However, the whole decoding procedure requires several steps [23-25] such as (1) pixel targeting

to determine the relative position of each pixel within the sub-image for obtaining the corresponding angle and wavelength information, (2) pixel rearrangement based on angle information for creating a sub-aperture image sequence and (3) reconstruct the spatial and spectral intensity distribution of the target from the sub-aperture image sequence. The spatial-spectral data joint decoding lies on the pixel targeting. The existing pixel targeting methods depend on the LF ray projection model [23-25] which uses geometric optics to elucidate the light propagation process and determines the propagation path of each pixel through inverse ray tracing. However, the precision of these methods depends on the center sub-image recognition accuracy. Optical system aberrations such as vignetting and distortion affect the recognition accuracy. The vignetting and distortion occur by primary and relay lenses along with tilt, and rotational errors due to microlens array (MLA) installation. Additionally, in SLFI, each wavelength filter only covers a smaller portion of the main aperture. This small aperture reduces the effective imaging area compared to the actual image sensor size. Thus, the image quality will be degraded such as blur and lower dynamic range.

In recent, data-driven and imaging-based correction methods have been employed to rectify the aberrations of the SLFI. The data-driven aberration correction models use a large amount of light field imaging (LFI) data and enable rapid calculations. However, these methods elevate the hardware costs due to the excessive low-aberration images required for training [26, 27]. In contrast, imaging-based methods use the LFI forward model to estimate errors [28, 29]. For instance, Jin et al. [29] employed an imaging-based method to correct the image distortion caused by misalignment of the MLA. This correction is conducted by fitting and estimating the rotation angles of the MLA row by row based on the center points of the microlenses. However, existing algorithms provide a lower accuracy for microlense center point recognition when applied to the sub-images with coupled aberrations thus, subsequently affecting the precision of aberration correction. Additionally, these methods correct the overall image without considering the local variations between the sub-images and potentially introduce additional biases in the spatial information. Besides, various aberrations can be seen during the imaging process and these aberrations can be represented as the product of different aberration matrices. These aberration matrices can be calculated by using matrix optics principles [30, 31] sequentially. The aberrations can be global [i.e., vignetting and distortion] and local [i.e., introduced by the SFA]. Therefore, an efficient correction method is necessary for full-frame and non-uniform aberrations including vignetting and distortion corrections to resolve the image blurring and improve the dynamic range induced by the SFA. It is also possible to overcome the limitations of existing LF aberration correction methods by categorizing the aberrations systematically and correcting them one by one. This will allow for accommodating the local distortion and improving the precision of spatial-spectral data joint decoding.

In LFI, the light intensity of defocused objects is scattered within the resulting diffuse spots. Hence, it is necessary to reconstruct the light intensity for extracting the spectral intensity from the diffuse spots. The LFI employs digital refocusing approaches to accomplish the intensity reconstruction. However, for SLFI, this refocusing calculation provides mixed spectral information on the refocused image. Therefore, a common practice for the SLFI is to position all the target objects only on the focal plane [8, 21]. In this practice, all targets are in focus, thereby eliminating the need for intensity and depth reconstruction. However, sacrifices the acquired angular information concurrently and ignores the ability to distinguish targets at different depths. Moreover, in LFI applications such as natural scene imaging [32], pedestrian recognition on roads [33] and fast-moving vehicle identification [34], the target objects are typically located within a range of different depths. Therefore, to effectively leverage the depth resolution and accurately quantify the 3-D spectral intensity distribution of the target scene, it is crucial to develop a joint decoding method for space-spectral data. The rectified 5-D spectral LF data can be decoded into 4-D LF data (2-D spatial and 2-D angular information) for each wavelength by integrating aberration correction. The decoded 4-D LF data can overcome the spectral mixing problem encountered in intensity reconstruction, thereby enabling the accurate extraction of depth information and spectral intensity distribution at various depths within the target scene.

This study presents the development of a microlens array-based SLFI system to investigate the aberration and data decoding issues involved in the SLFI. A novel spatial-spectral joint decoding method incorporating aberration correction is proposed and discussed. The performance of the proposed method is verified. The proposed method eliminates the aberrations and decodes the 3-D spectral intensity distribution of SLFI data. The image quality of the SLFI system is evaluated through Structure Similarity Index Measure (SSIM) based on LF sub-images. A step-by-step process is considered for correcting the aberrations. An improved digital refocusing is employed to the spectral aliasing in refocusing of SLFI data. Defocus analysis is carried out along with the focus stack algorithm to achieve the spatial-spectral joint decoding, thus enabling the reconstruction of depth and spectral intensity distribution of the target scene. Experimental evaluations are carried out by setting up a real-world imagery case based on the spectral channels of the LFS. Results obtained from this study are presented and discussed.

2. Principles of spectral light field imaging

2.1 Imaging model

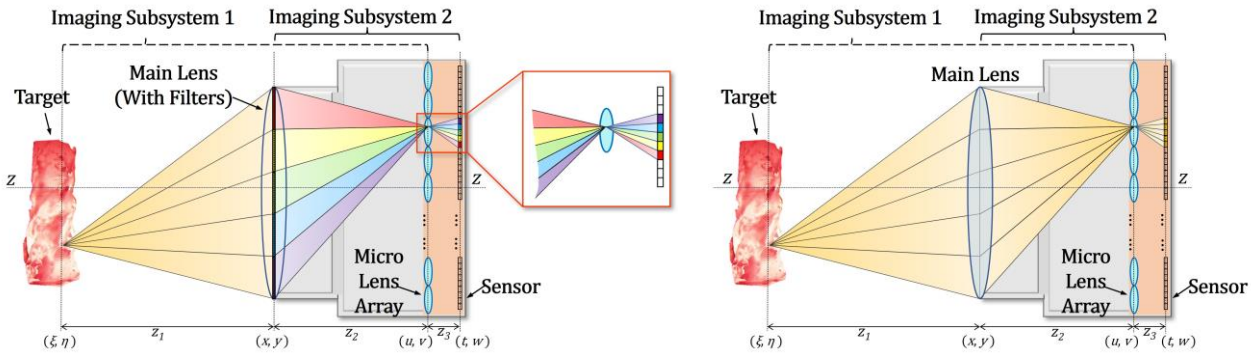
Figure 1 illustrates the principle of an unfocused light field camera (LFC) and a LFS. The main lens of the LFC is simplified as an ideal thin lens on the (x, y) plane. The target, located on the (ζ, η) plane, is projected onto the MLA on the (u, v) plane. The MLA discretizes the image and re-images it onto the camera sensor plane (t, w). After re-imaging, each image obtained under a microlens is referred to as a sub-image, and the image formed by extracting the same pixel under each sub-image is called a sub-aperture image [23]. In Figure 1, $z_1$is the distance between the target and the main lens planes, z_2 is the distance between the MLA and the main lens planes, and z_3 is the distance between MLA and the camera sensor. The focal lengths of the main lens and the microlens are denoted as f_1 and f_2, respectively. For an unfocused LFC, $z_3 = f_2$ [23].

Compared with the unfocused LFC, the LFS incorporates a SFA onto the main aperture. This SFA consists of $N \times M$ narrow-band filters with different wavelengths. As depicted in Figure 1(a), the intensity of a sub-image pixel formed by a narrow-band filter $\lambda_i(x_i, y_i)$ and can be denoted as $I(\lambda_i(x_i, y_i), u, v)$. By combining all the pixels corresponding to the narrow-band filter $\lambda_i(x_i, y_i)$, the spectral image of the target at the wavelength λ_i can be obtained. It can be represented as $\sum_u \sum_v I(\lambda_i(x_i, y_i), u, v)$. Consequently, by recombining all pixels of each sub-image, a total of $N \times M$ spectral images of the target at different wavelengths can be obtained.

In a LFC, the propagation of light rays can be formed as matrices [30, 31] and the ray from a point on the target plane (ζ, η) can be represented as $\boldsymbol{L}(\zeta, \eta, \varphi, \theta)$,

$$\boldsymbol{L_{CCD}}\,(t, w, \theta, \varphi) = \boldsymbol{A_{LF}} \cdot \boldsymbol{L}(\zeta, \eta, \theta, \varphi) \tag{1}$$

where θ is the angle between the projection of the light ray onto the $t - z$ plane and the z-axis, φ is the angle between the projection of the light ray onto the $w - z$ plane and the z-axis (where the z-axis is the main optical axis). $\boldsymbol{L_{CCD}}\,(t, w, \theta, \varphi)$ represents the collected light information on the camera sensor.



(a) (b)

Fig. 1. Imaging principles of (a) the LFS and (b) the original LFC.

The ray propagates to the LFC undergoes refraction through the main lens and the MLA and reaches the camera sensor plane (t, w). This process can be represented using a propagation matrix $\boldsymbol{A_{LF}}$ and define as:

$$\boldsymbol{A_{LF}} = \boldsymbol{T}(f_2) \cdot \boldsymbol{R}(f_2) \cdot \boldsymbol{S}(u_c, v_c) \cdot \boldsymbol{T}(z_2) \cdot \boldsymbol{R}(f_1) \cdot \boldsymbol{T}(z_1) \tag{2}$$

where $\boldsymbol{T}(z_1)$ represents the light propagation from the object to the main lens, $\boldsymbol{T}(z_2)$ represents the light propagation from the main lens to the MLA, and $\boldsymbol{T}(f_2)$ represents the light propagation from the MLA to the sensor plane. $\boldsymbol{R}(f_1)$ represents the refraction effect of the main lens (simplified as an ideal thin lens) and $\boldsymbol{R}(f_2)$ represents the refraction effect of the microlens. $\boldsymbol{T}(l)$ represents the light propagation matrix, where l denotes the propagation distance of the light ray,

$$\boldsymbol{T}(l) = \begin{bmatrix} 1 & 0 & l & 0 & 0 \\ 0 & 1 & 0 & l & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$\boldsymbol{R}(f)$ is the refraction matrix and can be defined in Eq. (4), where f represents the focal length of the lens.

$$\boldsymbol{R}(f) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1/-f & 0 & 1 & 0 & 0 \\ 0 & 1/-f & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$\boldsymbol{S}(u_c, v_c)$ represents the discrete matrix of the microlens and calculates the distance from the intersection point of the light ray with the microlens surface to the center of that microlens. It can be expressed as,

$$\boldsymbol{S}(u_c, v_c) = \begin{bmatrix} 1 & 0 & 0 & 0 & -u_c \\ 0 & 1 & 0 & 0 & -v_c \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

where (u_c, v_c) represents the center of the microlens where the light ray enters.

In LFI, the LF information is represented by the position and direction of the sensor plane, $\boldsymbol{L_{CCD}}(t, w, \theta, \varphi)$, is

typically transformed into dual-plane coordinates $\boldsymbol{LF}(u, v, x, y)$ [i.e., main lens aperture plane (x, y) and MLA plane (u, v)]. Based on geometrical optics, this transformation process involves tracing the light rays backwards from the sensor plane to the MLA plane. This process can be expressed using a ray-tracing matrix $\boldsymbol{A_{RT}}$,

$$\boldsymbol{LF}(u, v, x, y) = \boldsymbol{W}(z_2) \cdot \boldsymbol{A_{RT}} \cdot \boldsymbol{L_{CCD}}\ (t, w, \theta, \varphi) \tag{6}$$

$$\boldsymbol{A_{RT}} = \boldsymbol{S}(-u_c, -v_c) \cdot \boldsymbol{R}(-f_2) \cdot \boldsymbol{T}(-f_2) \tag{7}$$

where $\boldsymbol{W}(z_2)$ represents the coordinate transformation matrix which is used to convert the coordinates of the LF information $[\boldsymbol{A_{RT}} \cdot \boldsymbol{L_{CCD}}\ (t, w, \theta, \varphi)]$ obtained from backward ray tracing into the dual-plane coordinates. The transformation is carried out as follows:

$$\boldsymbol{W}(z_2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & -z_2 & 0 & 0 \\ 0 & 1 & 0 & -z_2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

The LF image $\boldsymbol{IMG}(u, v)$ can be obtained by integrating the LF information $\boldsymbol{LF}(u, v, x, y)$, as follows:

$$\boldsymbol{IMG}(u, v) = \frac{1}{{z_2}^2} \iint \boldsymbol{LF}(u, v, x, y) dx dy \tag{9}$$

In LFI, the SFA filters the incident light rays passing through the main lens. However, the imaging model discussed in Eqs. (6) and (9) focus on the directional changes of light rays and do not explicitly consider their intensity and wavelength. Since the differences in refractive index across visible wavelengths (400 nm -780nm) are negligible, the above model, therefore, remains applicable for SLFI within this range.

2.2 Digital refocusing

Traditional spectral imaging (TSI) systems are developed based on linear systems with translational invariance [2, 7, 31]. To capture the spectral images of objects simultaneously at different depths within the field of view (FOV), it is necessary to focus on each depth individually and then capture the corresponding spectral images for subsequent synthesis. For LFS, this requirement for multiple focusing can be replaced by the LF refocusing process. However, it is crucial to demonstrate that the process of LF refocusing aligns with the assumption of a linear system. In other words, it is necessary to establish an equivalence relationship between the LF refocusing image and the image obtained by a TSI system.

The LF refocusing converts the captured LF information $\boldsymbol{LF}(u, v, x, y)$ into the refocused image plane and the main lens aperture plane coordinates, $\boldsymbol{LF_\alpha}(u', v', x, y)$. This can be expressed as follows:

$$\boldsymbol{LF_\alpha}(u', v', x, y) = \boldsymbol{A_{rf}}\ (\alpha) \cdot \boldsymbol{LF}(u, v, x, y) \tag{9}$$

$$\boldsymbol{A_{rf}}(\alpha)=\begin{bmatrix}\alpha & 0 & 1-\alpha & 0 & 0\\ 0 & \alpha & 0 & 1-\alpha & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix} \tag{10}$$

$$\alpha = {z_2}'/z_2 \tag{11}$$

where $\boldsymbol{A_{rf}}(\alpha)$ represents the refocusing matrix; α is the refocusing factor that determines the position of the refocused image plane and ${z_2}'$ is the distance between the refocused image plane and the main lens.

The refocused image $\boldsymbol{IMG_{\alpha}}(u',v')$ can be obtained by integrating the $\boldsymbol{LF_{\alpha}}(u',v',x,y)$ and expressed as follows:

$$\boldsymbol{IMG_{\alpha}}(u',v')=\frac{1}{\alpha^2{z_2}^2}\iint \boldsymbol{LF_{\alpha}}(u',v',x,y)dxdy \tag{12}$$

Considering that the TSI and SLFI have consistent responses for wavelength λ_1 and neglects the light intensity and wavelength variations during the imaging process. To investigate this, the imaging models described in Section 2.1 can be applied. It is worth noting that the focal length of the main lens in the TSI is denoted as f_1.

By combining Eqs. (1), (6) and (9), the process of light transmission to the refocused image plane can be represented as $\boldsymbol{A_{RF}}$,

$$\boldsymbol{LF_{\alpha}}(u',v',x,y)=\boldsymbol{A_{RF}}\cdot\boldsymbol{L}(\zeta,\eta,\theta,\varphi) \tag{13}$$

$$\boldsymbol{A_{RF}}=\boldsymbol{A_{rf}}(\alpha)\cdot\boldsymbol{W}(z_2)\cdot\boldsymbol{A_{RT}}\cdot\boldsymbol{A_{LF}} \tag{14}$$

The light ray propagation process of a TSI can be simplified into three steps such as (1) the transmission of light from the target object to the main lens, (2) the refraction of the light by the main lens and (3) the transmission of the refracted light to the image plane (u',v'). Therefore, the light transmission process $\boldsymbol{A_{ConvS}}$ can be represented as $\boldsymbol{A_{ConvS}}$,

$$\boldsymbol{L_{\alpha}}(u',v',x,y)=\boldsymbol{A_{ConvS}}\cdot\boldsymbol{L}(\zeta,\eta,\theta,\varphi) \tag{15}$$

$$\boldsymbol{A_{ConvS}}=\boldsymbol{W}(z_2')\cdot\boldsymbol{T}(z_2')\cdot\boldsymbol{R}(f_1)\cdot\boldsymbol{T}(z_1) \tag{16}$$

where $\boldsymbol{T}(z_2')$ represents the transmission of light from the main lens to the image plane. $\boldsymbol{W}(z_2')$ is the coordinate transformation matrix that converts the propagated light into the main lens aperture (x,y) and the image plane (u',v').

By combining Eqs. (3)-(5), (8), and (10),

$$\boldsymbol{A_{ConvC}}=\boldsymbol{A_{RF}}=\begin{bmatrix}1-\frac{\alpha z_2}{f_1} & 0 & \alpha z_2+s_1\left(1-\frac{\alpha z_2}{f_1}\right) & 0 & 0\\ 0 & 1-\frac{\alpha z_2}{f_1} & 0 & \alpha z_2+z_1\left(1-\frac{\alpha z_2}{f_1}\right) & 0\\ 1 & 0 & z_1 & 0 & 0\\ 0 & 1 & 0 & z_1 & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix} \tag{17}$$

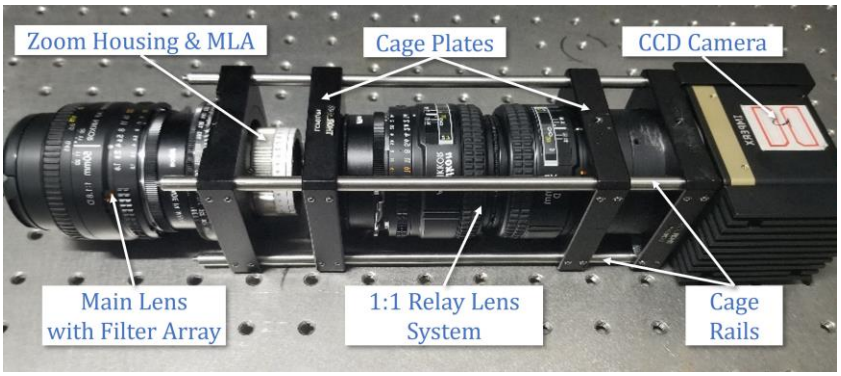
Eq. (17) shows that the light propagation of LF refocusing and the TSI are equivalent. This implies that the multispectral images obtained through refocusing the SLFI have the translational invariance of a linear system, like that obtained through a TSI. These refocused multispectral images, therefore, can be utilized to analyze the spectral intensity distribution of objects that contain multiple depths.

3. Aberration correction and data decoding

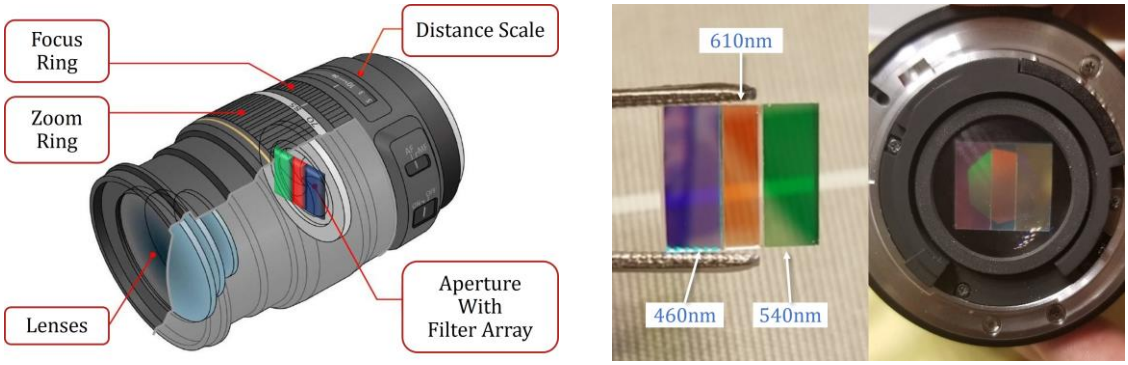
As a 5-D dataset encompassing light intensity, ray angles and spectral information, LFS images cannot be directly used for spectral analysis. Hence, a decoding process is essential to extract the 3-D spectral data cube from this 5-D dataset. Before decoding, it is crucial to perform aberration correction to reduce the decoding errors. Thus, a joint decoding method incorporating aberration correction is proposed and details are given below.

3.1 Microlens array-based caged-type light field spectrometer

An adjustable caged-type SLFI system based on the MLA has been employed. Unlike LFCs using pinhole arrays, the use of a MLA allows for better light flux and improves the image quality [8, 21]. As shown in Figure 2(a), the SLFI consists of two head-to-head lenses (Nikon 50mm f/1.4D) as relay lenses that project the intermediate image of the MLA onto the image sensor at a ratio of 1:1. The apertures of the relay lenses are fixed at f/1.4, and the focal length is set to infinity to obtain high-quality relay images. The use of the relay lenses allows for convenient assembly of the SLFI and preserves the ability to replace the different MLA with different parameters. The employed MLA (RPC photonics S100-f/4) is mounted using a high-precision zoom sleeve (Thorlabs SM1ZM) with a maximum travel distance of 4.1mm. As shown in Figures 2(b) and (c), the main objective lens (Nikon 50mm f/1.8D) of the system has been modified with a custom narrow-band SFA containing three different wavelengths: 460nm, 540nm and 610nm, respectively. The full width at half maximum (FWHM) of the narrow-band filters is 10nm. The modified lens has a fixed aperture but retains full focusing capability. All optical components are assembled by a cage plate (Thorlabs LCP06/M) and cage rods (Thorlabs ER10) to ensure their coaxial and perpendicularity to the principal axis. The SLFI allows flexible adjustment of the MLA position through the rotation of the zoom housing, enabling the trade-off between the spatial resolution and directional resolution. In future, this SLFI system can easily be tailored based on different application requirements [35, 36]. The SLFI is fabricated with maximum directional resolution to achieve high-precision measurement of spectral intensity distribution and facilitate the extraction of spectral information [37].



(a)



(b) (c)

Fig. 2. The employed Caged-Typed LFS:
(a) Structure of the LFS; (b) Schematic diagram of SFA installation; (c) Photos of the SFA and multispectral main lens.

3.2 Aberrations effect

Figure 3 illustrates an image captured by the LFC using a planar uniform light source positioned perpendicular to the main optical axis. Ideally, the captured image should exhibit consistent brightness across its entire surface. However, the brightness of the image gradually decreases from the center towards the edges and the boundaries of the sub-image become blurred. This indicates a significant presence of vignetting phenomena caused by the multiple lenses within the imaging system due to the main lens, relay lenses and the MLA. Structures such as lenses, lens barrels and lens frames obscure the off-axis light rays (rays deviating from the main optical axis). Consequently, the brightness of off-axis object points in the image is diminished. The brightness reduces severely for the maximum FOV and creates a more prominent vignetting effect [30].

The SLF image quality also degrades by the distortion effects. The SLFI system employed in this study uses two identical relay lenses with an axis-symmetric structure to ensure no distortion. A precise alignment of the main lens, relay lenses and the MLA with the main optical axis is carried out to eliminate the presence of tangential distortion. However, the radial distortion is observed due to the main lens [30]. The radial distortion effect is characterized by analyzing the discrepancy between the ideal and measured center coordinates of the sub-image. To quantify the radial distortion, a relative radial distortion coefficient R is calculated by,

$$R = \frac{\sqrt{(x_d - x_u)^2 + (y_d - y_u)^2}}{\sqrt{(x_u - x_c)^2 + (y_u - y_c)^2}} \times 100\% \tag{18}$$

Where $(x_c,\ y_c)$ denotes the center coordinates of an image and $(x_u,\ y_u)$ and $(x_d,\ y_d)$ represent the actual and ideal coordinates of the sub-image center, respectively. The ideal coordinates of the sub-image are calculated based on the MLA arrangement. Table 1 summarizes the distortion coefficients of the LF image at various measurement points as depicted in Figure 3. The maximum relative radial distortion value of 0.53% is found at measurement point c. Although the distortion does not alter the order of sub-image arrangement or the arrangement of pixels within each sub-image, noticeable distortion may interfere with the accurate center positioning of the microlens and lead to decoding errors.

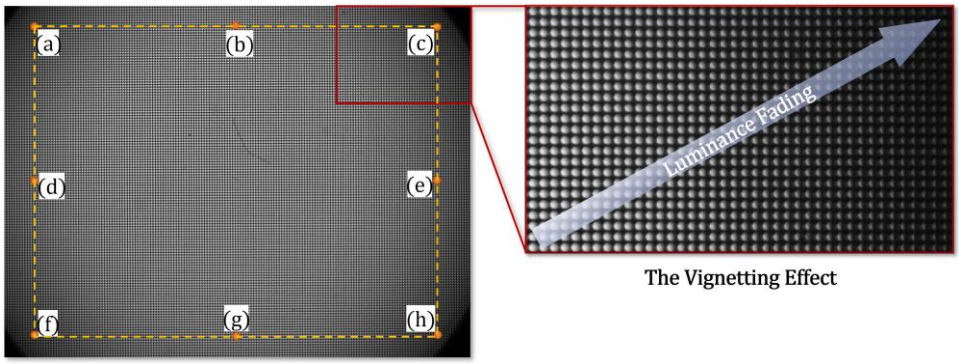
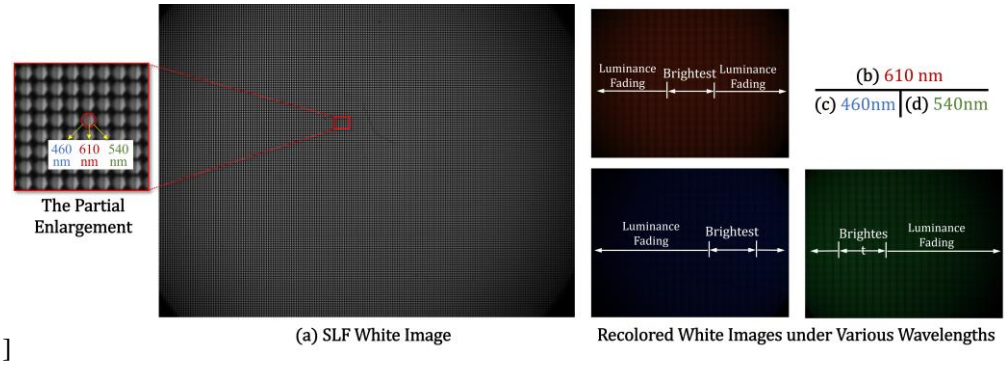


Fig. 3. Example of the vignetting effect of LF image and distortion measurement points

Table 1. Relative distortion values at each measurement point

Measurement points	Relative distortion value (%)
a	0.47
b	0.21
c	0.53
d	0.33
e	0.39
f	0.49
g	0.28
h	0.51

Figure 4(a) illustrates an example spectral light field (SLF) white image obtained by the SLFI system. Each spectral LF sub-image comprises three distinct regions characterized by varying grayscale intensity and they mainly correspond to the three wavelengths such as 460nm 610nm and 540nm (from left to right) respectively. Figure 4(b)-(d) illustrates the three LF images extracted from the SLFI white image, corresponding to these three wavelengths. Non-uniform brightness distribution can be seen in these LF images. The image part with normalized average brightness above 0.95 is labeled as 'Brightest'. At 540nm, the left region exhibits the highest brightness and the brightness gradually decreases towards the right. Conversely, at 460nm, the 'Brightest' region appears on the right side and decreases towards the left. At the 610nm image, the 'Brightest' region is in the center. Therefore, it is revealed that the images obtained by the SLFI system exhibit different brightness regions but the locations of these regions are varied under different wavelengths due to the SFA. Because the SFA divides the main aperture into multiple regions, allowing light only from a specific wavelength to pass through each region. Consequently, on the main aperture, the energy of light becomes clustered in the region covered by the corresponding wavelength filter. When this clustered energy is imaged onto the sensor, uneven brightness distribution is then formed in the image.



]

Fig. 4. Example of (a) SLF white image and extracted white images under various wavelengths: (b)610nm; (c) 460nm; (d) 540nm. Note that these images are enhanced and recolored.

Additionally, the SLF sub-images which are located far from the main optical axis exhibit shape distortion. Figure 5 shows a comparison of the SLF sub-images at three distinct positions obtained at 540nm. The right-side sub-images show progressive shape distortion and the number of pixels is decreased compared to the left and center-side sub-images. This indicates that off-axis light beams are partially or completely blocked, particularly those far from the optical axis. To facilitate the assembly and replacement of different MLAs, a relay lens system is employed. The aperture of this lens system is limited and intercepts off-axis light rays. The aperture of the microlenses is also limited and restricts the ability to capture off-axis light rays. The microlenses that are significantly off-axis no longer have a full FOV to cover the main aperture, thus leading to the shape distortion of off-axis sub-images.

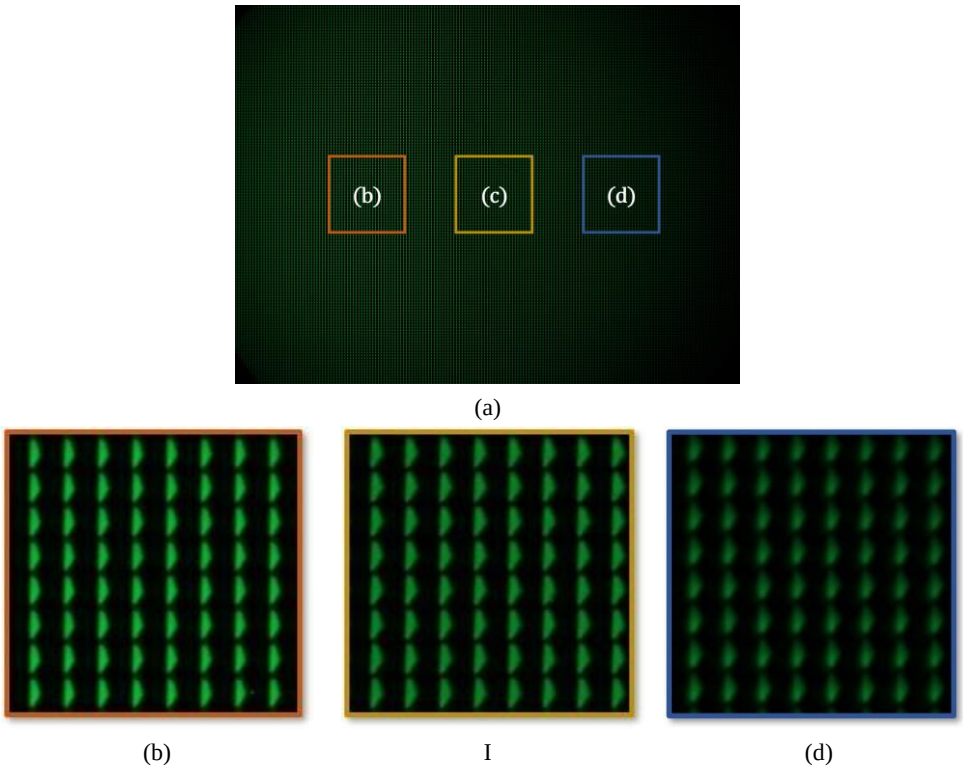


Fig. 5. The shape distortion of SLFI sub-images under various positions at 540nm: (a) extracted SLFI white image; (b) left-side sub-images; (c) center sub-images; (d) right-side sub-images. Note that these images are enhanced and re-colored.

3.3 Aberration assessment and correction

It is necessary to investigate the aberrations for correction and decode them accurately to assess the quality of SLF images. As discussed in Sections 2.1 and 3.2, an LFC identical to the employed SLFI system without the SFA as the LFC suffers the same degree of vignetting and distortion. Thus, the white image acquired by the LFC is used for the aberration evaluation. The results obtained from the evaluation are then used for the vignetting and distortion corrections and then employed on the SLF images. The correction of uneven spectral brightness is employed to correct the vignetting of SLF images.

The vignetting and distortion occur mainly in particular regions of the image that are far from the image center (main optical axis) and the center sub-images are aberration-free as discussed in Section 3.2. The aberrated sub-images are differences in shape and grayscale distribution compared to aberration-free sub-images. In this study, SSIM is employed to quantify the aberrated sub-images [38]. For reference sub-images, a set of center sub-images (16 in this

study) are considered and their average value is calculated. The SSIM value is calculated between each sub-image and the reference sub-image. The higher SSIM value (~1) indicates the smaller aberrations. The region of the image that has a lower SSIM value is significantly distorted and removed from that region from the image. The remaining regions of the image that have higher SSIM are kept for subsequent correction of vignetting and spectral brightness distribution. A step-by-step process is considered for correcting the aberrations. This process is initially started with global aberrations such as vignetting and followed by local aberrations introduced by the SFA.

The vignetting correction is carried out on the sub-image by sub-image. A gain coefficient is assigned to each pixel of the sub-image to compensate for the brightness loss caused by the vignetting. For example, the unfocused LFC image $\boldsymbol{L}$ consists of $U \times V$ sub-images, each having $a \times a$ pixels. The u-th, v-th sub-image is denoted as $\boldsymbol{SubL}_{u,v}(lp_{11}, \dots, lp_{ij}, \dots, lp_{aa})$, where lp_{ij} represents the i-th, j-th pixel in the sub-image. The reference sub-image is denoted as $\boldsymbol{SubLS}_{reference}(lpr_{11}, \dots, lpr_{ij}, \dots, lpr_{aa})$, where lpk_{ij} represents the i-th, j-th pixel in the sub-image. The gain coefficient for pixel lp_{ij} in sub-image $\boldsymbol{SubL}_{u,v}$ is taken as c_vig_{ij} and expressed as:

$$c_vig_{ij} = I_{lpr_{ij}} / I_{lp_{ij}} \tag{19}$$

Thus, the u-th & v-th vignetting-corrected sub-image $\boldsymbol{SubL}_{rvig_u,v}$ is obtained by:

$$\boldsymbol{SubL}_{c_vig_u,v} = (c_vig_{11} \cdot I_{lp_{11}}, \dots, c_vig_{ij} \cdot I_{lp_{ij}}, \dots, c_vig_{aa} \cdot I_{lp_{aa}}) \tag{20}$$

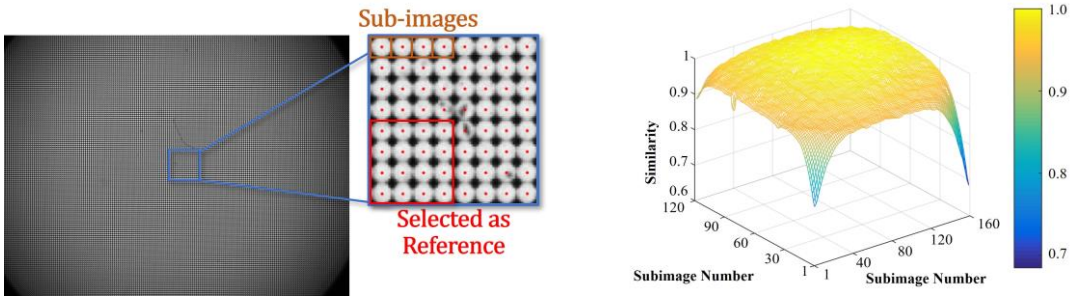
The gain coefficient is directly applied to the SLF image for vignetting correction. The SLFI system employs a $N \times M$ narrowband SFA, thus the SLF image $\boldsymbol{SL}$, composed of $U \times V$ sub-images and each sub-image has $a \times a$ pixels. The u-th and v-th sub-image is denoted as $\boldsymbol{SubSL}_{u,v}(sp_{11}, \dots, sp_{ij}, \dots, sp_{aa})$, where sp_{ij} represents the i-th, j-th pixel in the sub-image. Thus, the u-th & v-th vignetting-corrected SLF sub-image $\boldsymbol{SubSL}_{C_vig_u,v}$ is obtained by:

$$\boldsymbol{SubSL}_{c_vig_u,v} = (c_vig_{11} \cdot I_{sp_{11}}, \dots, c_vig_{ij} \cdot I_{sp_{ij}}, \dots, c_vig_{aa} \cdot I_{sp_{aa}}) \tag{21}$$

The uneven spectral brightness of the SLF image is corrected by using a set of sub-images that have the best dynamic range and the dynamic range of the remaining sub-images is then rectified. In total $N \times M$ monochromatic LF images ($\boldsymbol{MonoSL}_{c_vig}{}^{w_i}$ for wavelength w_i) could be extracted from vignetting corrected SLF image $\boldsymbol{SL}_{c_vig}$ and they correspond to the $N \times M$ bands. For reference sub-images at three wavelengths, the best dynamic range of sub-images $\boldsymbol{MonoSL}_{c_vig}{}^{w_i}$ are chosen. A gain coefficient (i.e., is obtained by averaging the pixel brightness ratio of reference sub-images at three wavelengths, respectively) is assigned to each pixel of the sub-image of $\boldsymbol{SL}_{c_vig}{}^{w_i}$ to compensate for the spectral brightness loss.

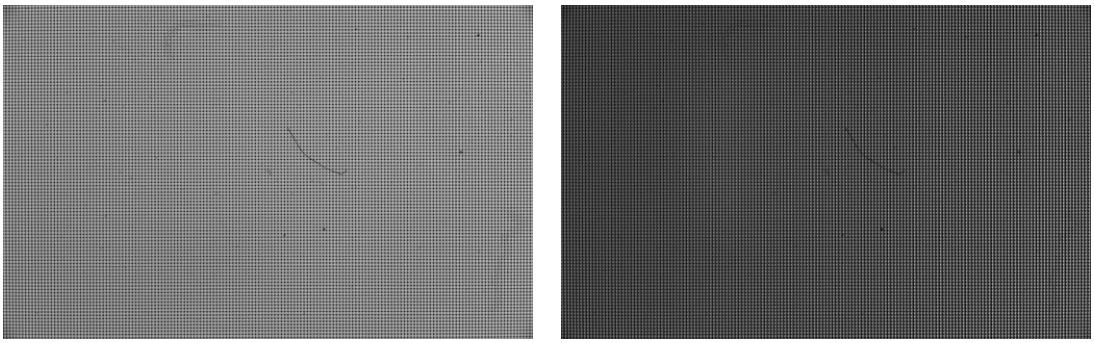
Figure 6 shows the preprocessing of the SLF image which is used for obtaining the reference sub-image and assessing the aberrations. The aberration-free central sub-images are considered to verify the paraxial approximation. As shown in Figure 6(a), 16 central sub-images are selected to obtain the reference sub-image. Each pixel intensity of the reference sub-image is the arithmetic mean of the intensity of the corresponding pixel of the selected sub-images. The SSIM values between the sub-images and the reference sub-image are plotted in Figure 6(b). In the central region of the image, the SSIM values are stable and around 0.95, showing small aberrations. However, at the edges of the image (especially in the four image corners), the SSIM values drop to 0.6, thus a significant difference can be seen in brightness and shape compared to the reference sub-image. Also shows a noticeable distortion and vignetting effect. The sub-images with SSIM values below 0.91 are removed due to the vignetting and distortions. Therefore, 164 ×

103 sub-images are retained in the central portion of the image from the full frame of 180 × 135 sub-images. The cropped image shown in Figure 7 is then used in further aberration correction and decoding processes.



(a) (b)

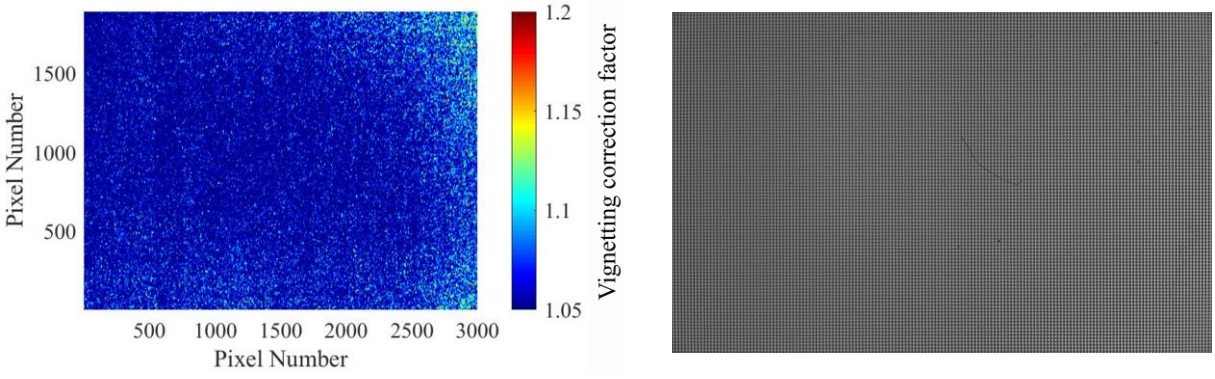
Fig. 6. The preprocessing of aberration correction procedure:(a) process of obtaining the reference sub-image and (b) aberration assessment results through SSIM.



(a) (b)

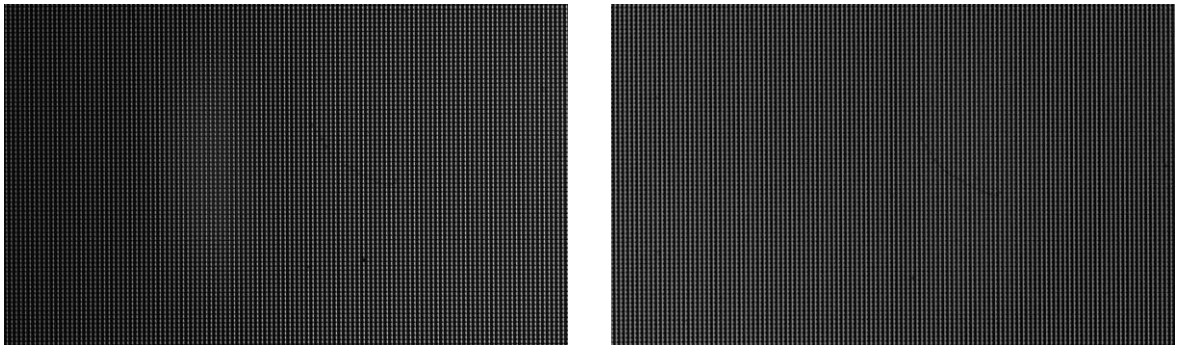
Fig. 7. Images after cropping the regions with significant aberrations: (a) LF white image and (b) SLF white image.

The distribution of the vignetting correction coefficient and the corrected images are shown in Figure 8. The uncorrected image shows significant vignetting at its edges which requires larger vignetting correction coefficients as seen in Figure 8(a). However, the corrected images [Figures 8(b)] showed improved brightness uniformity across the entire image frame. The LF images at 460nm, 540nm and 610nm are extracted from the vignetting corrected LFS white image. The correction of spectral brightness is conducted on these LF images, respectively. Figure 9 illustrates a comparison of the LF image at 460nm before and after the spectral brightness correction. The corrected image shows a more uniform brightness performance. This confirms the efficacy of the proposed method described in Section 3.3.



(a) (b)

Fig. 8. Vignetting correction results: (a) calculated vignetting correction coefficient distribution and (b) vignetting corrected SLF white image.



(a) (b)

Fig. 9. Comparison of LF white image at 460nm: (a) before and (b) after spectral brightness correction.

3.4 Proposed Joint Decoding Method

In this study, a joint decoding method is proposed along with the aberration correction for SLF images. The proposed method decodes the SLF images to obtain a complete spectral distribution of the images. The process of the proposed method is depicted in Figure 10.

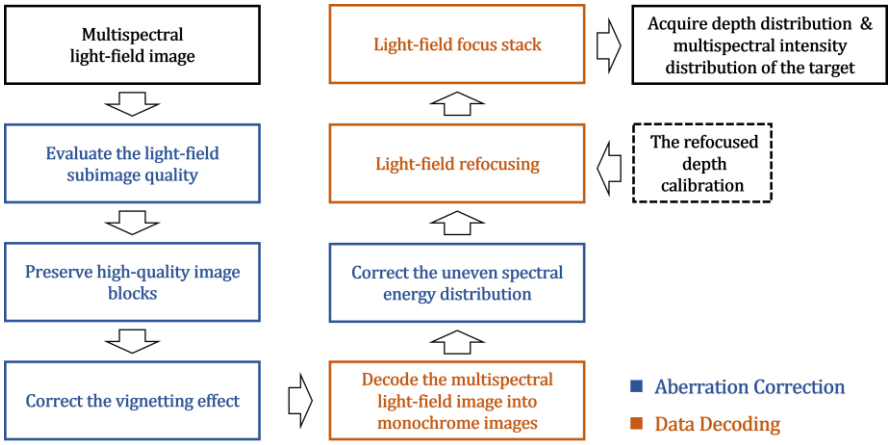


Fig. 10. Technical strategy of the proposed joint decoding method for SLF image with aberration corrections.

Initially, the aberrated corrected LF image is used as input for the LF refocusing calculation and obtains the spectral intensity distribution of images on each object plane. To avoid refocusing errors from the frequency domain, the Shift-and-add Refocusing algorithm [described in Section 2.2] [23] is utilized rather than the Fourier Slice Refocusing algorithm [24]. The $N \times M$ refocused image sequences are obtained through the refocusing calculation. Each refocused image corresponds to a specific focal plane in space. Thus, the correlation between the target objects and the focal plane of the refocused images is then analyzed by the defocus response analysis. For each object, the refocused image with the highest defocus response value presents the sharpest image. Thus, the sharpest image represents the spectral intensity of the object. The decoding process is then completed by synthesizing the sharpest images of each object using focus stacking, resulting in all-focused images at each specific wavelength. The all-focused images represent the spectral intensity distribution of the test scene.

The analysis of defocus response and focus stacking are achieved based on the Tenengrad function. The Tenengrad is a no-reference image sharpness evaluation function based on image gradients, which indicates that edges in a sharp image have larger gradient values [38-40]. It exhibits good unimodality and unbiasedness in the presence of Gaussian noise [39, 40] and it uses the Sobel operator to extract the gradients in the horizontal and vertical directions, as described in the following process:

For the point (t, w) in the image, the gradient value $G(t, w)$ at that point can be calculated as follows:

$$G(t,w) = \sqrt[2]{{G_t}^2(t,w) + {G_w}^2(t,w)} \tag{20}$$

where $G_t(t, w)$ and $G_w(t, w)$ are the gradient values in the horizontal and vertical directions of the pixel point, respectively and defined as:

$$G_t(t,w) = I(t,w) \otimes g_t \tag{21}$$

$$G_w(t,w) = I(t,w) \otimes g_w \tag{22}$$

where $\otimes$ denotes the convolution operation, and g_t and g_w are the Sobel operator templates in the horizontal and vertical directions, respectively, given by:

$$g_t = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{23}$$

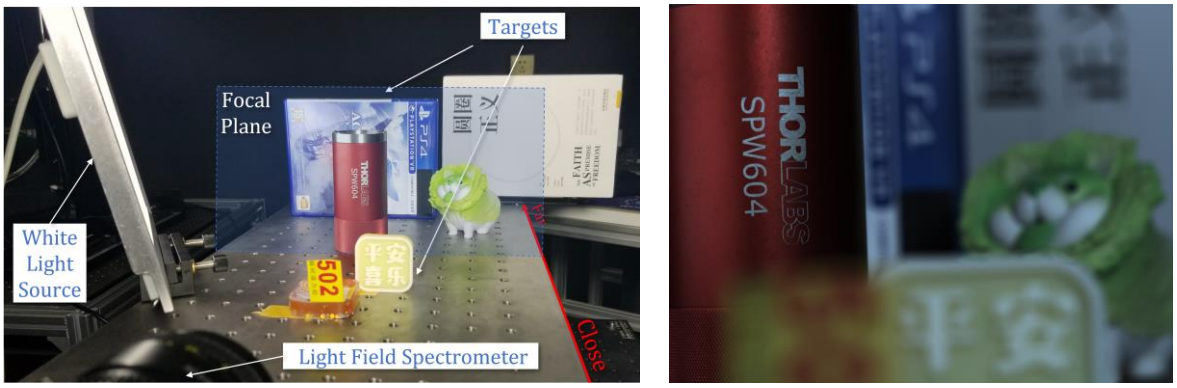
$$g_w = \begin{bmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{24}$$

The synthesis of all-focused SLF images is based on the results of the sharpness evaluation, where the highest pixel gradients in the sequence of SLF images are combined.

4. Experimental Validation

4.1 Test case setup

To validate the proposed method, a test case is considered as illustrated in Figure 11. The SLFI system is placed at one end of the test case and a uniform white light is used to illuminate the scene. The test case consists of various color objects positioned at different distances from the SLFI system, however, they are all within the 90 cm range. The approximate distances between each object and the SLFI system are shown in Table 2. The color objects in the test case are selected as red, blue and green to match the wavelengths of the SLFI system. The proposed method and SLFI system are verified by measuring the spatial information and spectral intensity at wavelengths of 460nm, 540nm and 610nm, respectively. A red socket wrench is placed at the center of the test case in an axial direction and the focal plane is adjusted to focus on its surface. The proposed method is applied to extract the spectral intensity of the whole test case.



(a) (b)

Fig. 11. Example images of the experimental test case: (a) showing relative positions of the objects and (b) image captured by the traditional CCD camera.

Table 2. Overview of distances between each object and the SLFI system and color features of the objects

Test case objects	Color features	Distances from the SLFI system/cm
Paper standing sign	Yellow & Red	12.4
Resin standing sign	Light yellow	16.2
Socket wrench	Red	37.9
Toy dog	Green	42.6
Game disc case	Blue	58.4

4.3 .2 Aberration correction

An example spectral image of the test case captured by the SLFI system is shown in Figure 12(a). The raw captured image has a resolution of 3340 (H) × 2488(V) and the total sub-images are 24300 [180 (H) × 135 (V)]. Each sub-image is square-shaped and occupies 19 × 19 pixels. Within each sub-image, three regions are defined which correspond to the spectral intensity of the three wavelengths. The aberration assessment and correction including the gain coefficients described in Section 3.3 are applied to the captured SLF images [Figure 12(a)]. Figure 12 (b)-(f) shows comparative images with and without the aberration correction. Based on the aberration assessment, 164 × 103 sub-images located in the central part of the image are retained as these are without aberrations thus no need for correction. The rest sub-images have significant aberrations [Figure 12(b)] and need subsequent image correction and decoding processes. The correction of the vignetting effect is also considered. Figure 12(b) shows a significant improvement in brightness uniformity across the entire image compared to the uncorrected original image [Figure 12(a)].

The vignetting corrected SLF image [Figure 12(b)] is subsequently decomposed into LF images at the three wavelengths. Each wavelength LF image then undergoes individual spectral brightness correction. As an example, the corrected image at 540nm is shown in Figure 12(e). Compared to Figure 12(d), the left part of the corrected image at 540nm [i.e., the part of the socket wrench image] shows an improved brightness. The spectral brightness correction has also been conducted for the 460nm and 610nm, respectively. The corrected LF images for these three wavelengths are combined to obtain a corrected SLF image as depicted in Figure 12(f). The corrected image exhibits better uniform and brighter spectral distributions. This demonstrates the efficacy of the proposed method described in Section 3.3.

4.3 Spectral digital refocusing

To extract the spectral distribution of each object at different depths, the aberration corrected SLF image [Figure 12(f)] is refocused [i.e., shifting the original focal plane (surface of the red socket wrench) to other target objects outside the original depth of field]. The LF images are extracted for the three wavelengths [Figure 12(f)] and refocused respectively by using Eq. (12). These monochrome LF images lack complete LF information due to pixels with valid signals occupying only a portion of the sub-image. However, this limitation does not preclude refocusing calculations, as the process involves shifting and adding sub-aperture images.

In this study, the refocused image $\boldsymbol{IMG_{\alpha}}$ can also be expressed as follows [6],

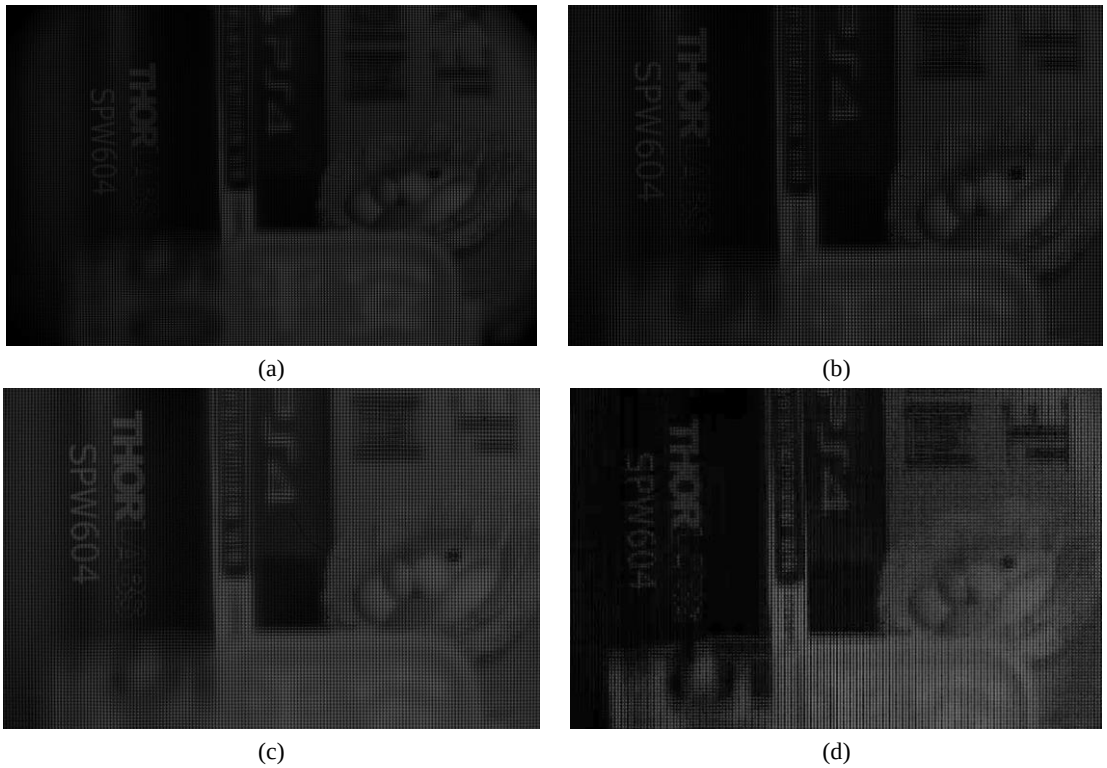
$$\boldsymbol{IMG_{\alpha}}(u', v') = \frac{1}{\alpha^2 {z_2}^2} \iint \boldsymbol{LF}^{(x,y)} \left(x\left(1 - \frac{1}{\alpha}\right) + \frac{u'}{\alpha}, y\left(1 - \frac{1}{\alpha}\right) + \frac{v'}{\alpha} \right) dxdy \tag{25}$$

where $\boldsymbol{LF}^{(x,y)} \left(x\left(1 - \frac{1}{\alpha}\right) + \frac{u'}{\alpha}, y\left(1 - \frac{1}{\alpha}\right) + \frac{v'}{\alpha} \right)$ is simply the sub-aperture image $\boldsymbol{LF}^{(x,y)}$, dilated by α and shifted by a factor of $(x(1 - 1/\alpha), y(1 - 1/\alpha))$. In other words, digital refocusing can be implemented by shifting and

adding the LF sub-aperture images. The shift, $(x(1-1/\alpha), y(1-1/\alpha))$, of each sub-aperture image in Equation 3 increases with the distance of the sub-aperture from the center of the lens (x, y). In monochrome LF images, pixels with valid signals only occupy a part of the sub-image. Therefore, compared to the images obtained by the original LFC, monochrome images could form fewer sub-aperture images. However, from the above derivation process, it can be seen that Eq. (25) still applies for these fewer sub-aperture images from monochrome images.

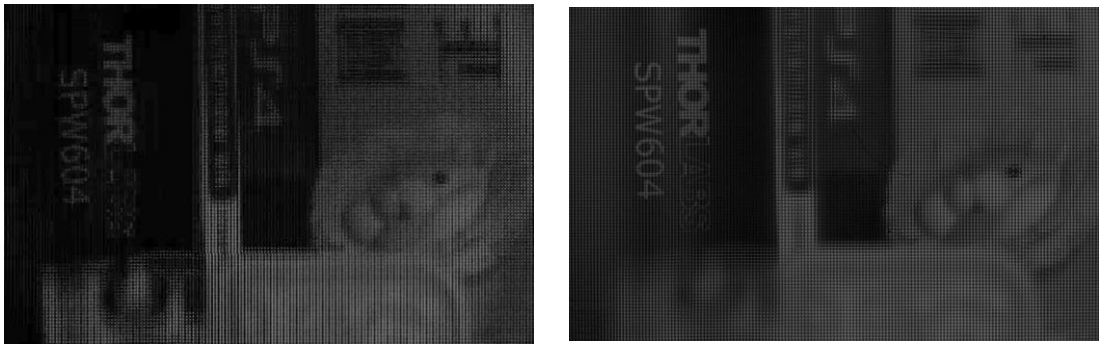
A total of 60 refocused images are obtained for each wavelength at different depths. These refocused images are then used for depth and spectral reconstruction. The resolution of the refocused image is 164×103 which is consistent with the number of sub-images in the horizontal and vertical directions of the SLF image. The refocusing method restructures the entire LF image into sub-aperture images (i.e., extract the pixels at the same position from each sub-image and rearrange them according to the order of the sub-images) and subsequently computes the weighted average sub-aperture images as refocused images [39]. Figure 15 illustrates an example of a sub-aperture image sequence of the SLF image [Figure 12(f)]. This sub-aperture image is categorized into the three wavelengths corresponding to 540nm, 610nm and 460nm, respectively.

However, the refocusing calculation is conducted wavelength by wavelength. Thus, for each wavelength, the sub-aperture images related to other wavelengths are considered as a null signal. For example, at 540nm: the sub-aperture images are only the leftmost region as shown in Figure 13, and the sub-aperture images of 610nm and 460nm contain null signals. As these null sub-aperture images are involved in the averaging process of digital refocusing [24], the refocused image brightness is diluted. It has been observed that the digital refocused image intensity is notably lower than the actual intensity due to the null signals of other wavelengths as shown in Figure 14. Thus, it is necessary to correct the intensity after the refocusing calculation. Based on the refocusing principle [Eq. (9)], the intensity after digital refocusing is proportional to the area covered by the 540nm filter on the main aperture. Thus, the intensity correction factor of each specific sub-aperture image can be calculated. The correction factors achieved for all sub-aperture images are shown in Table 3. An example of digital refocused images before and after the intensity correction is shown in Figure 14.



(a) (b)

(c) (d)



(e) (f)

Fig. 12. Comparison of test case SLF images before and after correction: (a)Raw SLF image; (b) SLF image after cropping out regions with significant aberrations; (c) vignetting corrected SLF image; (d) Vignetting corrected 540nm LF image; (e) Spectral brightness corrected 540nm LF image; (f) Spectral brightness corrected SLF image.

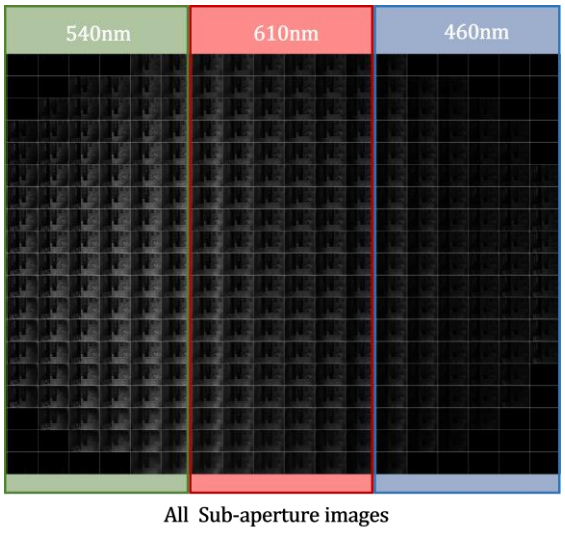


Fig. 13. SLF sub-aperture images sequence and their corresponding wavelengths.

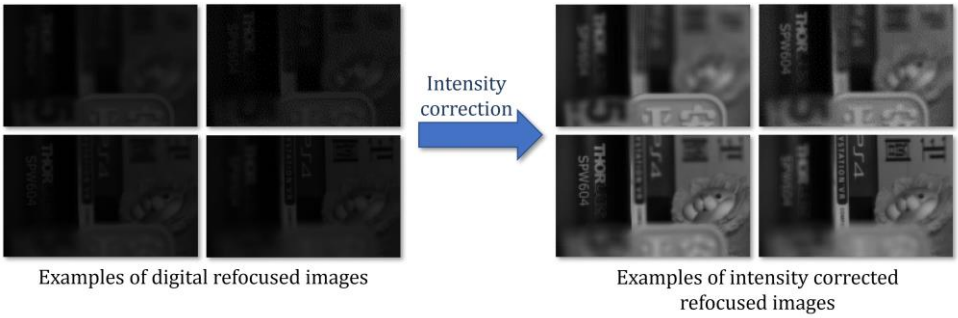


Fig. 14. Comparison of the refocused image under 540nm at different depths before and after intensity correction.

Table 3. Refocusing intensity correction coefficients at different wavelengths

Wavelength/nm	Effective sub-aperture images	Proportion of filter in main aperture / %	Intensity correction factor
460	90	30.23	3.40
540	100	32.63	3.06
610	116	37.14	2.64

4.4 Intensity and depth reconstruction

To determine the depth of each object with defocus response analysis, a relationship between the refocusing coefficient α and the depth is established through calibration. A tilted ruler with scales is used to capture an image that contains distance information and clear distinctions in sharpness among different parts, as depicted in Figure 15. The ruler image is divided into *n* parts along the scale direction, and the sharpness of each part is calculated to obtain a sharpness curve using a quadratic function, as illustrated in Figure 16. The highest sharpness part corresponds to the refocused focal plane, and the scale of the refocused focal plane is then converted to the horizontal distance, thereby completing the calibration of refocusing positions and intervals [36].

The defocus response analysis is carried out on the refocused image group to obtain the index map of the sharpest image for each object. As depicted in Figure 17(a), the index value of each pixel represents the number sequence of the sharpest image of that pixel. This number sequence corresponds to a unique α and spatial depth. Therefore, by combining the specific α *of* the sharpest image of each object with the curve in Figure 16, the spatial depth distribution of the target scene is determined accurately, as shown in Figure 17(b).

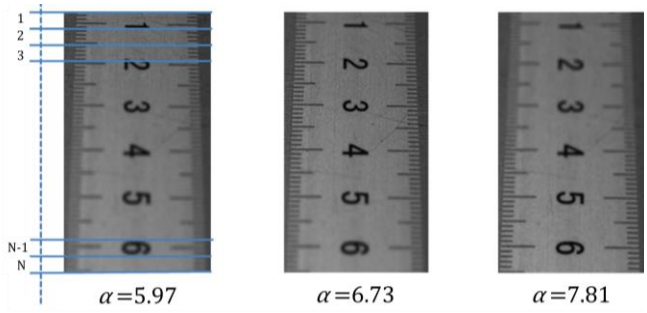


Fig. 15. Depth calibration at different refocusing coefficients α.

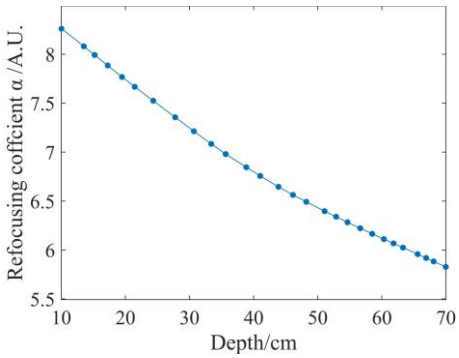
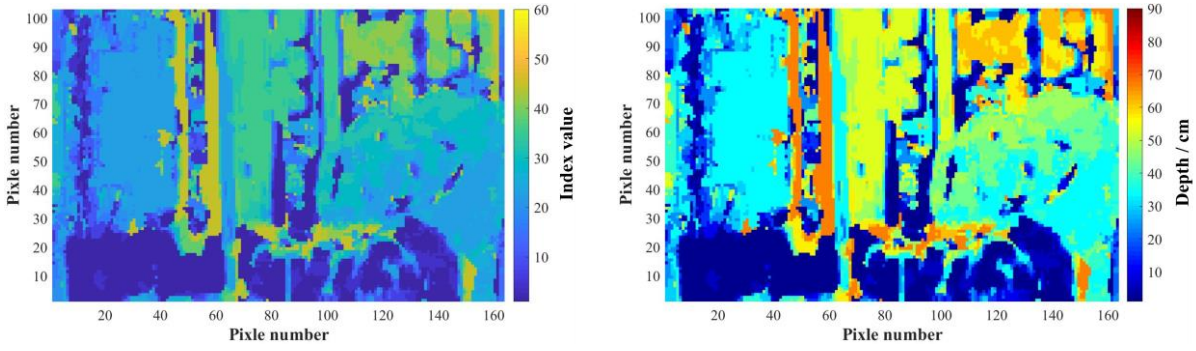


Fig. 16. Variation of refocusing coefficients with depth



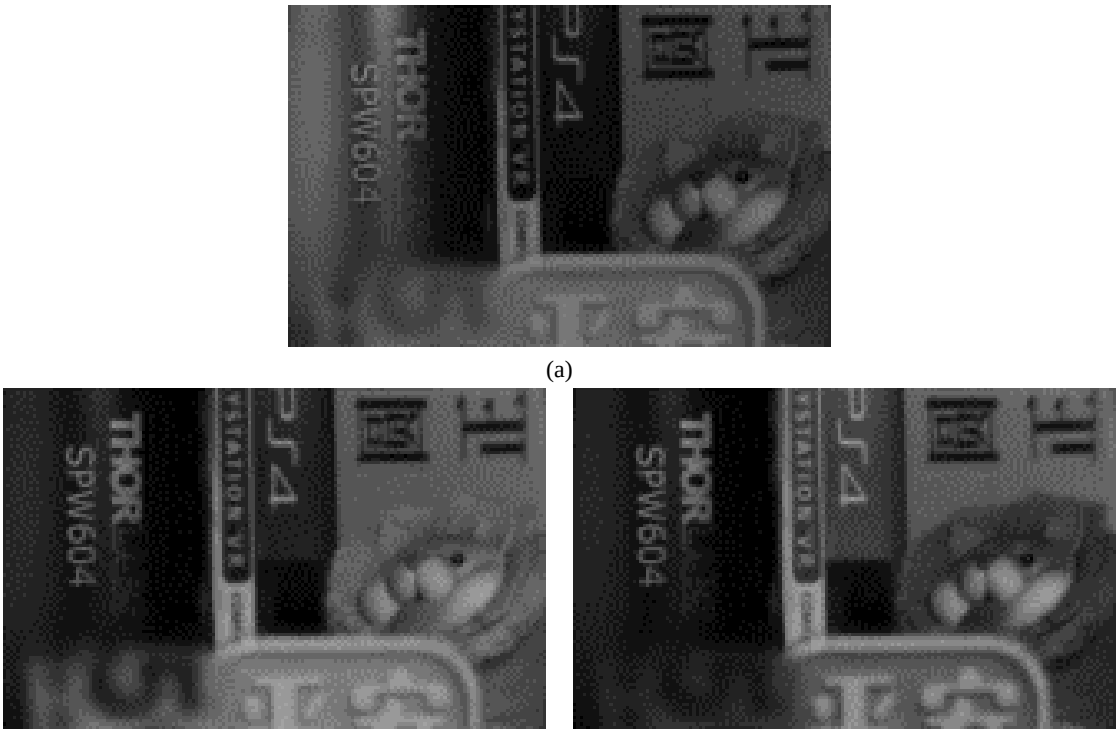
(a) (b)

Fig. 17. Defocus response analysis of the SLF test case image: (a) index distribution of defocus response and (b) spatial depth distribution.

The Tenengrad used for the defocus response analysis builds on the fact that, on average sharper images will produce larger gradient magnitudes when compared with blurry images. However, the aberrations in the SLFI system reduce the image brightness, thereby reducing the gradient magnitude of the image. This may hinder the identification of the clearest view of the target, reducing the accuracy of depth reconstruction. Thus, the aberration correction procedure is a necessary step for the subsequent intensity and depth reconstruction. The depth maps presented in this study (Figure 17) are not disturbed by aberration because it was calculated based on the aberration-corrected images.

Figure 18(a)-(c) shows the focused spectral images of the target scene obtained by the focus stack [41-44] calculation. These images exhibit a clear focus on foreground, middle and background objects and represent the relative spectral intensity distributions of the target objects. To evaluate the accuracy of the spectral distributions of the target objects, the focused spectral images are fused to a color image by the Bayer fusion algorithm as shown in Figure 19(a). The fused image is then evaluated by SSIM with the generated focus stack images (four) captured by a traditional CCD camera at different depths [Figure 19(b)]. The resolution of the traditional camera images is downsampled to 164×103 pixels to ensure comparability with Figure 19(a).

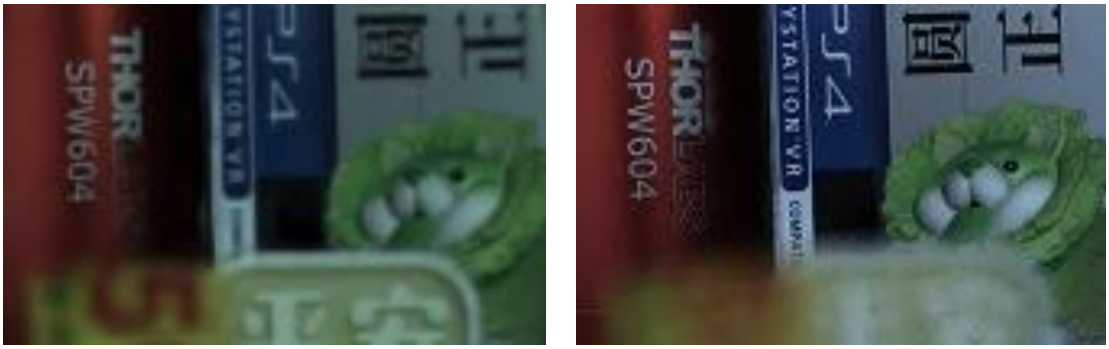
The SSIM value between the all-focused color images generated from (a) SLF images and (b) CCD camera images is 0.948, and the similarity of image structure, contrast, and brightness is 0.913, 0.975 and 0.951, respectively. The high SSIM value confirms the high degree of similarity between these two images. The contrast and brightness values are higher than 0.95, indicating a strong agreement in the overall image brightness and dynamic range. The image structure similarity is slightly lower than the contrast and luminance due to the blur of the foreground, which is significantly limited by the conventional camera's depth of field and focusing range. It demonstrates that the SLFI system benefits from an extended depth of field, resulting in a clearer foreground in the refocused images. Therefore, the evaluation further verifies the reconstruction accuracy of the spectral distribution of the target objects, the efficacy of the proposed SLFI system and the decoding and correction approaches.



(a)

(b) (c)

Fig. 18. Spectral distribution of the test scene at various wavelengths: (a) 610nm; (b) 540nm; (c) 460nm.



(a) (b)

Fig. 19: Comparison of all-focused color images generated from (a) SLF image and (b) CCD camera image.

5. Conclusion

This study presents a microlens array-based spectral light field imaging (SLFI) to investigate the data decoding and aberration corrections of spectral light field images. The aberrations of the proposed LFS are investigated and discussed, specifically vignetting and uneven spectral intensity distribution problems. A novel joint decoding method incorporating aberration correction for SLFI is proposed and implemented. The proposed method utilizes the LF sub-images for both aberration assessment and correction and shows a versatile approach to local and global aberration correction. An improved digital refocusing method is employed for the decoding of SLF images. The proposed SLFI system and the joint decoding method are validated experimentally. The spectral intensity and depth distribution of a test case are reconstructed through the proposed method. The concluding remarks obtained from this study are summarized as follows:

- It has been observed that vignetting has significant effects on the intensity distribution of SLF images, causing almost 30% of sub-images to be unsuitable for further decoding. Thus, improving the relay lens system to reduce vignetting is the key challenge of the SLFI system.
- Obvious differences have been noted in the spectral intensity distribution under various wavelengths and such differences are fundamentally determined by the filter arrangement in the SLFI system.
- Spectral aliasing in digital refocusing could be eliminated by decomposing a 5-D SLF image into 4-D LF images for different wavelengths. However, the image decomposition deteriorates the intensity of LF images. Therefore, a further intensity rectification of refocused images is necessary.
- The comparative study shows that the reconstructed spectral intensity and depth distribution have a good match with the test case and demonstrate the effectiveness of the proposed SLFI system and joint decoding method.
- Though the structures of SLFI systems are continually evolving and varied with their research objectives, the concept of the imaging model is the same including that in this study. The proposed aberration correction and data decoding methods are developed based on the characteristic of LF images being composed of a series of sub-images that can be used in different structures of SLFI systems.

It demonstrates that the proposed LFS and joint decoding method can unleash its potential to capture the complete spectral distribution of a 3-D scene in a single snapshot. As our ongoing research, both the SLFI system and the proposed method will be used to investigate flame chemiluminescence and particle image velocimetry.

Acknowledgements

The authors wish to express their gratitude to the National Natural Science Foundation of China (No. 52376158), Innovation Capacity Building of Lianyungang City Science and Technology Plan Project (Research and Development of Efficient and Low Carbon Gas Turbine Test Facility) (No. CX2207) and Engineering & Physical Sciences Research Council (EPSRC) Project Reference: EP/X020789/1.

References

[1].V. Backman, M. Wallace, L. Perelman, J. Arendt, M. Müller, Q. Zhang, G. Zonios, E. Kline, T. Mcgillican., Detection of preinvasive cancer cells, Nature 406(6791) (2000) 35 -36. https://doi.org/10.1038/35017638

[2]. W. Debskia, P. Walczykowskia, A. Klewskia, M. Zyznowskib, Analysis of usage of multispectral video technique for distinguishing objects in real-time. 20th ISPRS Congress. Istanbul: ISPRS 201(323) 2004.

[3]. S. Delalieux, A. Auwerkerken, W. Verstraeten, B Somers, R Valcke, S Lhermitte, J Keulemans, P Coppin, Hyperspectral reflectance and fluorescence imaging to detect scab induced stress in apple leaves, Remote Sensing 1 (4) (2009) 858 -874. https://doi.org/10.3390/rs1040858

[4]. R. Kester, N. Bedard, L. Gao, T. Tkaczyk, Real-time snapshot hyperspectral imaging endoscope. Journal of Biomedical Optics,16(5) (2011) 056005. https://doi.org/10.1117/1.3574756

[5]. M. Kim, T. Harvey, D. Kittle, H. Rushmeier, D. Brady, 3D imaging spectroscopy for measuring hyperspectral patterns on solid objects. ACM Transactions on Graphics, 31(4) (2012) 38. https://doi.org/10.1145/2185520.2185534

[6]. G. Wong, Snapshot hyperspectral imaging and practical applications, Journal of Physics: Conference Series (178) (2009) 012048. https://doi.org/10.1088/1742-6596/178/1/012048

[7]. H. Liu, S. Zheng, H. Zhou, C. Qi, Measurement of distributions of temperature and wavelength-dependent emissivity of a laminar diffusion flame using hyper-spectral imaging technique, Measurement Science and Technology 27(2) (2016) 025201. https://doi.org/10.1088/0957-0233/27/2/025201

[8]. L.J. Su, Z.L. Zhou, Y. Yuan, L. Hu, S.Y. Zhang, A snapshot light field imaging spectrometer, Optik. 126 (2015) 877–881. https://doi.org/10.1016/j.ijleo.2015.01.034

[9]. M. Descour, E. Dereniak, Computed-tomography imaging spectrometer: experimental calibration and reconstruction results, Applied Optics 34(22) (1995) 4817-4826. https://doi.org/10.1364/AO.34.004817

[10]. G.R. Arce, D.J. Brady, L. Carin, H. Arguello, Compressive coded aperture spectral imaging: an introduction, IEEE Signal Processing Magazine 31(1) (2014) 105-115. https://doi.org/10.1109/MSP.2013.2278763

[11]. X. Cao, H. Du, X. Tong, Q.H. Dai, S. Lin, A Prism-Mask System for Multispectral Video Acquisition, IEEE Transactions on Pattern Analysis and Machine Intelligence 33(12) (2011) 2423-2435. https://doi.org/10.1109/TPAMI.2011.80

[12]. X. Cao, X. Tong, Q.H. Dai, S. Lin, High resolution multispectral video capture with a hybrid camera system, IEEE Conference on Computer Vision and Pattern Recognition 1 (2011) 297-304. https://doi.org/10.1109/CVPR.2011.5995418

[13]. C. Ma, C. Xun, T. Xin, Q. Dai, S. Lin, Acquisition of High Spatial and Spectral Resolution Video with a Hybrid Camera System, International Journal of Computer Vision 110(2) (2014) 141-155. https://doi.org/10.1007/s11263-013-0690-4

[14]. L. Hu, Y. Yan, L.J. Su, M. Huang, Y. Li, Spectral resolution of light field spectral imaging system under diffraction model. Optics Communications 405 (2017) 421-425. https://doi.org/10.1016/j.optcom.2017.08.058

[15]. D.L. Kelly, D.E. Scarborough, B.S. Thurow, A novel multi-band plenoptic pyrometer for high-temperature applications, Meas. Sci. Technol. 32 (2021) 105901. https://doi.org/10.1088/1361-6501/ac0465

[16]. Y. Luan, D. Mei, S. Shi, Light-field multi-spectral radiation thermometry, Optics Letters 46(1) (2021) 9-12. https://doi.org/10.1364/OL.408437

[17]. M. Levoy, P. Hanrahan, Light field rendering, Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (1996) 31–42. https://doi.org/10.1145/237170.237199

[18]. “Lytro,” 2017. [Online]. Available: https://www.lytro.com/

[19]. “RayTrix. 3D light field camera technology.” 2017. [Online]. Available: http://www.raytrix.de/

[20]. G.C. Wu, B. Masia, A. Jarabo, Y.C. Zhang, L.Y. Wang, Q.H. Dai, T.Y. Chai, Y.B. Liu. Light field image processing: An overview, IEEE Journal of Selected Topics in Signal Processing, 11 (7) (2017) 926-954. https://doi.org/10.1109/JSTSP.2017.2747126

[21]. R. Horstmeyer, G. Euliss, R. Athale, M. Levoy, Flexible multimodal camera using a light field architecture. IEEE International Conference on Computational Photography. (2010) https://doi.org/10.1109/ICCPHOT.2009.5559016

[22]. Z. Zhou, Y. Yan, X.L. Bin, Light Field Imaging Spectrometer: Conceptual Design and Simulated Performance, Frontiers in Optics (2010) Laser Science XXVI, OSA Technical Digest (CD) FThM3. https://doi.org/10.1364/FIO.2010.FThM3

[23]. R. Ng, M. Levoy, M. Br´edif, G. Duval, M. Horowitz, P. Hanrahan, Light field photography with a hand-held plenoptic camera, Stanford University Computer Science Tech Report CSTR (2005) 2005-02.

[24]. R. Ng, Fourier slice photography, Proceedings of Special Interest Group for Computer Graphics and Interactive Techniques (2005) 735–744. https://doi.org/10.1145/1073204.1073256

[25]. D.G. Dansereau, O. Pizarro, S.B. Williams. Decoding, calibration and rectification for lenselet-based plenoptic cameras, Proceedings of the IEEE conference on computer vision and pattern recognition (2013) 1027-1034. https://doi.org/10.1109/CVPR.2013.137

[26]. R. Raghavendra, B. Yang, K.B. Raja, C. Busch, A new perspective — Face recognition with light-field camera, International Conference on Biometrics (2013) https://doi.org/1~8. 10.1109/ICB.2013.6612980

[27]. Wagner, N., Beuttenmueller, F., Norlin, N. et al. Deep learning-enhanced light-field imaging with continuous validation. Nature Methods 18 (2021) 557–563. https://doi.org/10.1038/s41592-021-01136-0

[28]. D. Cho, M. Lee, S. Kim, Y.W. Tai, Modeling the calibration pipeline of the Lytro camera for high quality light-field image reconstruction, IEEE International Conference on Computer Vision (2013) 3280-3287. https://doi.org/10.1109/ICCV.2013.407

[29]. J. Jin, Y.W. Cao, W.J. Cai, W.L. Zheng, P. Zhou, An effective rectification method for lenselet-based plenoptic cameras, Conference on Optoelectronic Imaging and Multimedia Technology IV (2016) https://doi.org/10020. 10.1117/12.2245540

[30]. Gross H, Zügge H, Peschka M, et al. Handbook of Optical Systems, Volume 3: Aberration Theory and Correction of Optical Systems. Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2007.

[31]. Mahajan V N. Optical Imaging and Aberrations, Part I: Ray Geometrical Optics. Washington: SPIE, 1998.

[32]. M.A. Taylor, T. Nöbauer, A. Pernia-Andrade, F. Schlumm, A. Vaziri, Brain-wide 3D light-field imaging of neuronal activity with speckle-enhanced resolution, Optica 5 (2018) 345-353. https://doi.org/10.1364/OPTICA.5.000345

[33]. C. Jia, F. Shi, Y. Zhao, M. Zhao, W. Zhe, S. Chen, Identification of pedestrians from confused planar objects using light field imaging, IEEE Access 6 (2018) 39375-39384. https://doi.org/10.1109/ACCESS.2018.2855723

[34]. M. Ren, R. Liu, H. Hong, J.J. Ren, G.B. Xiao, Fast Object Detection in Light Field Imaging by Integrating Deep Learning with Defocusing, Applied Sciences 7(12) (2017) 1309. https://doi.org/10.3390/app7121309

[35]. X. Song, M. Gu, L. Cao, Z. Tang, C. Xu A microparticle image velocimetry based on light field imaging. IEEE Sensors Journal 19 (2019) 06–17. https://doi.org/10.1109/JSEN.2019.2927414

[36]. Y. Liu, M. Zhu, T. Wang, G. Lei, M.M. Hossain, B. Zhang, J. Li, C.L. Xu, Spatial resolution of light field sectioning pyrometry for flame temperature measurement. Optics and Lasers in Engineering 140(2) (2021) 106545. https://doi.org/10.1016/j.optlaseng.2021.106545

[37]. S. Zhu, A. Lai, K. Eaton, P. Jin, L. Gao, On the fundamental comparison between unfocused and focused light field cameras, Applied Optics 57 (2018) 1-11. https://doi.org/10.1364/AO.57.0000A1

[38]. W. Zhou, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing 13 (4) (2004) 600–612. https://doi.org/10.1109/TIP.2003.819861

[39]. N. Kanopoulos, N. Vasanthavada, R. L. Baker, Design of an image edge detection filter using the Sobel operator, IEEE Journal of solid-state circuits 23(2) (1988) 358-367. https://doi.org/10.1109/4.996

[40]. P. Fan, R.G. Zhou, W. Hu, N. Jing, Quantum image edge extraction based on classical Sobel operator for NEQR, Quantum Information Processing 18(1) (2019) 1-23. https://doi.org/10.1007/s11128-018-2131-3

[41]. T. Broad, M. Grierson, Light field completion using focal stack propagation, ACM Special Interest Group for Computer GRAPHICS Posters. (2016) 54. https://doi.org/10.1145/2945078.2945132

[42]. A. Mousnier, E. Vural, C. Guillemot, Partial light field tomographic reconstruction from a fixed-camera focal stack, Computer Science (2015). https://doi.org/10.48550/arXiv.1503.01903

[43]. M. W. Tao, S. Hadap, J. Malik, R. Ramamoorthi, Depth from combining defocus and correspondence using light-field cameras, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2013) 673–680. https://doi.org/10.1109/ICCV.2013.89

[44]. M.W. Tao, P.P. Srinivasan, J. Malik, S. Rusinkiewicz, R. Ramamoorthi, Depth from shading, defocus, and correspondence using light-field angular coherence, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2015) 1940–1948. https://doi.org/10.1109/CVPR.2015.7298804