Research Article | Vol. 59, No. 26 / 10 September 2020 / Applied Optics | 7853

applied optics

Design of binary-phase diffusers for a compressed sensing snapshot spectral imaging system with two cameras

JONATHAN HAUSER,[1,*] AMIR AVERBUCH,[2] MENACHEM NATHAN,[1] VALERY A. ZHELUDEV,[2] MICHAEL KAGAN,[1] AND MICHAEL A. GOLUB[1,†]

[1]School of Electrical Engineering, Faculty of Engineering, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel
[2]School of Computer Science, Faculty of Exact Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel
*Corresponding author: hauserjo@post.tau.ac.il

Received 17 April 2020; revised 6 July 2020; accepted 30 July 2020; posted 31 July 2020 (Doc. ID 395541); published 8 September 2020

We propose designs of pupil-domain optical diffusers for a snapshot spectral imaging system using binary-phase encoding. The suggested designs enable the creation of point-spread functions with defined optical response, having profiles that are dependent on incident wavefront wavelength. This efficient combination of dispersive and diffusive optical responses enables us to perform snapshot spectral imaging using compressed sensing algorithms while keeping a high optical throughput alongside a simple fabrication process. Experimental results are reported.

https://doi.org/10.1364/AO.395541

1. INTRODUCTION

Spectral imaging (SI) systems are designed to capture the spatial and spectral information of an object or a scene, forming a "spectral cube." Previous work by our research team has demonstrated the feasibility of performing snapshot spectral imaging (SSI) together with color imaging using a simple, high-light-throughput architecture. The SSI system used a single standard monochromatic digital camera with an added pupil-domain 14-level phase diffuser [1,2]. The diffuser featured 400 randomly permutated strips with 14 unique heights and was fabricated using four-layer photolithography and dry etching. The diffuser design combined approaches of compressed sensing (CS) theory and classical spectroscopy, which require highly randomized response along with strong chromatic dispersion. However, while the random point-spread function (PSF) response was different for each wavelength of interest, adjacent wavelengths had similar optical responses with strong spatial overlapping on the image sensor (i.e., weak dispersion). Moreover, the system's performance was limited due to discrepancies between the actual optical model and the sensing matrix. The addition of a regular RGB camera side-by-side with the monochromatic camera with the diffuser contributed to the reconstruction quality of the spectral cube, using CS-based algorithms [3].

The focus of this work is on the introduction of new diffuser designs with improved chromatic dispersion that keep diffusion properties. The efficient combination of dispersion and diffusion is achieved by binary-phase encoding [4], which greatly simplifies the fabrication process through the use of one-layer photolithography and dry etching. In addition, large spatial resolution is achieved thanks to the one-dimensional (1D) symmetry of the diffuser, which enables us to exploit the full dimension of the sensor along the nondispersed and diffused (DD) image axis. Accordingly, reconstruction of spectral cubes of pixel size $1944 \times 256 \times 27$ in the [420, 680] nm spectral range are reported.

2. MATHEMATICAL MODEL OF THE OPTICAL SYSTEM

The binary diffuser is designed as a thin phase optical element that is placed at the entrance pupil of the monochromatic camera with lateral coordinates (u_M, v_M) and consists of vertical straight-line strips that are parallel to the v_M axis. Each strip's depth and respective phase are constant within the strip's width and quantized to a binary level. Since the diffuser is placed at the entrance pupil of the imaging lens, it sets a scaled complex pupil function on the exit pupil $P(u'_M, v'_M; \lambda_l)$ for each λ_l wavelength of interest:

$$P_M(u'_M, v'_M; \lambda_l) = \left|P_M\left(u'_M, v'_M; \lambda_l\right)\right| e^{[i\varphi_M(u'_M, v'_M; \lambda_l)]}, \quad (1)$$

where (u'_M, v'_M) are the lateral coordinates of the exit pupil plane, $|P_M(u'_M, v'_M; \lambda_l)|$ is the field amplitude transmittance, and $\varphi_M(u'_M, v'_M; \lambda_l)$ is the added phase at the exit pupil. Assuming that the optical system is ideal in the absence of the diffuser, the added phase $\varphi_M(u'_M, v'_M; \lambda_l)$ is solely determined by the diffuser, and the field amplitude transmittance is determined by the clear aperture's shape and size at the exit pupil plane. The incoherent PSF $h_{I,M}$ for the λ_l wavelength can be calculated as

$$h_{I,M}\left(x'_M, y'_M; \lambda_l\right) = \left|\mathrm{FT}_{2D}^{-1}\left\{P_M\left(u'_M, v'_M; \lambda_l\right)\right\}\right|^2, \quad (2)$$

where $\mathrm{FT}_{2D}^{-1}\{\cdot\}$ denotes the inverse 2D Fourier transform and $|\cdot|^2$ denotes the absolute-square operator.

The binary diffuser is intended to create both dispersion and diffusion of light on the image sensor by setting the PSF of the monochromatic camera. 1D binary-phase encoding is performed by hard clipping [4]:

$$\varphi_M\left(u'_M; \lambda_{des}\right) = \begin{cases} \pi, & \text{if } \cos\left[2\pi f_{enc}\cdot u'_M + \varphi_T\left(u'_M; \lambda_{des}\right)\right] > \sqrt{1 - \frac{|T(u'_M;\lambda_{des})|^2}{T_{max}^2}}, \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where u'_M is the horizontal coordinate of the exit pupil plane, λ_{des} is the design wavelength, f_{enc} is the encoding spatial frequency that sets the dispersion extent, $\varphi_T(u'_M; \lambda_{des})$ is the encoded phase function that sets the diffusion extent and shape, $T(u'_M; \lambda_{des})$ is the complex field transmittance function of the diffuser, and T_{max} is the maximal field transmittance. Assuming the diffuser has ideal field amplitude transmittance, $|T(u'_M; \lambda_{des})| = T_{max} = 1$. Accordingly, Eq. (3) can be expressed as

$$\varphi_M\left(u'_M; \lambda_{des}\right) = \begin{cases} \pi, & \text{if } \cos\left[2\pi f_{enc}\cdot u'_M + \varphi_T\left(u'_M; \lambda_{des}\right)\right] > 0. \\ 2\pi, & \text{otherwise} \end{cases} \quad (4)$$

Assuming that an incident wavefront of wavelength λ hits the diffuser with an angle of incidence θ_{inc} with respect to the plane of the diffuser, the angular separation between the jth diffraction order angle θ_j and θ_{inc} is given by

$$\sin\left(\theta_j\right) - \sin(\theta_{inc}) = j\lambda f_{enc}. \quad (5)$$

Since the diffuser is placed at the entrance pupil plane of the monochromatic camera's imaging lens, the distance between the jth diffraction order and the zeroth order at the image plane can be calculated as

$$\Delta_j(\lambda) = j\lambda R\frac{D_{EnP}}{D_{ExP}} f_{enc}, \quad (6)$$

where D_{ExP} and D_{EnP} are the diameters of the exit pupil and the entrance pupil, respectively, and R denotes the distance between the exit pupil plane and the image plane.

Selection for the encoding spatial frequency f_{enc} and the encoded function $\varphi_T(u'_M; \lambda_{des})$ is subject to inherent trade-offs. Selecting a large value for f_{enc} contributes to the spectral reconstruction, as it generates stronger chromatic dispersion by extending the polychromatic signal over a larger number of sensor pixels. However, strong dispersion of the signal requires large sensor arrays and may harm the spatial reconstruction quality. The selection for the encoded phase function $\varphi_T(u'_M; \lambda_{des})$, which sets the diffusion shape and extent, is also subject to its own trade-offs. While CS theory indicates that the signal should be randomly projected on the sensor, rapid alternations of the PSF maybe challenging to characterize experimentally. Therefore, we designed the diffusers to provide three types of diffusion: for low alternations, we chose $\varphi_T(u'_M; \lambda_{des})$ to perform quasi-uniform diffusion, denoted as top-hat diffusion; for moderate alternations, we chose $\varphi_T(u'_M; \lambda_{des})$ to be cubic phase [5]; and for high alternations, we chose $\varphi_T(u'_M; \lambda_{des})$ to be random.

A. Top-Hat Diffusion Function

To derive the encoded phase function $\varphi_T(u'_M)$ for top-hat diffusion, we assume that the optical system without the diffuser is ideal and in focus. Accordingly, the added phase is calculated as

$$\varphi_T\left(u'_M; \lambda_{des}\right) = \frac{2\pi}{\lambda_{des}}\left(S\left(u'_M\right) - S_{Lens}\left(u'_M\right)\right), \quad (7)$$

where $S_{Lens}(u'_M)$ is the eikonal function of the lens at the exit pupil plane and $S(u'_M)$ is the eikonal function after the diffuser. $S_{Lens}(u'_M)$ is calculated as

$$S_{Lens}\left(u'_M\right) = \sqrt{u'^2_M + R^2} + R, \quad (8)$$

where R is the distance between the exit pupil plane and the image plane. $S(u_M)$ is calculated as

$$S\left(u'_M\right) = \sqrt{u'^2_M + \frac{R^2 D'^2_M}{\left(a - D'_M\right)^2}} - \frac{R D'_M}{a - D'_M}, \quad (9)$$

where R is the distance between the exit pupil plane and the image plane, D'_M is the radius of the exit pupil, and a is the nominal half-width of the top-hat profile at the image plane.

B. Cubic-Phase Diffusion Function

The inspiration for using a cubic-phase modulation function was derived from the work of Dowski and Cathey [5], albeit for a different usage. While Dowski and Cathey demonstrated that a cubic-phase modulation mask extends the depth of field by making the optical transfer function (OTF) independent of misfocus, we used the cubic-phase modulation function for generating a quasi-ramped shaped diffusive pattern with moderate spatial alternations. The cubic-phase diffusive function is calculated as

$$\varphi_T\left(u'_M; \lambda_{des}\right) = \alpha u'^3_M, \quad (10)$$

where the parameter α is proportional to the spatial extent of diffusion at the image plane.

C. Random Diffusion Function

The realization of a random diffusion function is done in a similar way to our formerly reported random diffuser [1–3]. We start with generating a periodic discrete binary vector

$$\varphi_{\mathrm{Bin}}[k] = \begin{cases} 2\pi, & k \text{ is even} \\ \pi, & k \text{ is odd} \end{cases}, \quad k = \overline{1, N_d}, \tag{11}$$

where k is an integer and N_d is the length of the vector $\varphi_{\mathrm{Bin}}[k]$. We then generate a random binary vector

$$\varphi_{\mathrm{Rand}}[k] = \mathrm{randperm}\{\varphi_{\mathrm{Bin}}[k]\}, \quad k = \overline{1, N_d}, \tag{12}$$

where the operator randperm $\{V\}$ returns a random permutation of the vector V. We derive the phase function $\varphi_T(u'_M)$ by expanding the vector $\varphi_{\mathrm{Rand}}[k]$ to continuous coordinates:

$$\varphi_T\left(u'_M; \lambda_{\mathrm{des}}\right) = \sum_{k=0}^{N_d-1} \varphi_{\mathrm{Rand}}[k] \,\mathrm{rect}\left(\frac{u'_M - (k+0.5)\,\Delta u'_M}{\Delta u'_M}\right), \tag{13}$$

where the value of the function rect$\{\xi\}$ is 1 for $|\xi| < 0.5$ and 0 at other points, and $\Delta u'_M = \frac{D'_{ExP}}{N_d}$ is the width of a phase strip at the exit pupil plane. The parameter N_d, which equals to the number of phase strips, is proportional to the spatial extent of diffusion at the image plane. The diffuser's design height profile h_{des} is proportional to the design phase and can be calculated in the paraxial case as [6]

$$h_{\mathrm{des}}\left(u'_M; \lambda_{\mathrm{des}}\right) = \frac{\varphi_M\left(u'_M; \lambda_{\mathrm{des}}\right)}{2\pi} \frac{\lambda_{\mathrm{des}}}{\left(n\left(\lambda_{\mathrm{des}}\right) - 1\right)}, \tag{14}$$

where λ_{des} is the design wavelength and $n(\lambda_{\mathrm{des}})$ is the refractive index of the diffuser's material for λ_{des}. Although the binary diffuser is designed for wavelength λ_{des}, other wavelengths in the entire spectral range are incident on it. The spectral cube uses a finite number L of spectral bands with central wavelengths λ_l, $l = \overline{1, L}$, out of the entire continuous wavelength range. At a certain central wavelength λ_l, the groove depths $h(u'_M; \lambda_{\mathrm{des}})$ remain unchanged while the phase changes as [1,6]

$$\varphi_M\left(u'_M; \lambda_l\right) = \varphi_M\left(u'_M; \lambda_{\mathrm{des}}\right) \frac{\lambda_{\mathrm{des}}}{\lambda_l} \frac{n(\lambda_l) - 1}{n(\lambda_{\mathrm{des}}) - 1}. \tag{15}$$

When added to a regular digital camera, the binary diffuser converts the original image into a dispersed and diffused (DD) image with programmed blur. Since the binary diffuser has a 1D diffraction pattern, the intensity $I'_{M,l} \equiv I'_M(x'_M, y'_M; \lambda_l)$ of the DD image in the presence of the diffuser at the λ_l wavelength may be expressed by a 1D convolution $I'_{M,l} = h_{I,M,l} \otimes I_l$ of the ideal ("nondispersed") image, $I_l \equiv I(x, y; \lambda_l)$, with the 1D representation of the incoherent PSF of the monochromatic camera $h_{I,M,l} \equiv h_{I,M}(x'_M; \lambda_l)$, for each y coordinate of the object. The contribution of polychromatic light to the DD image can be expressed as a weighted sum of the intensities of monochromatic DD images $I'_{M,l}$, determined by the spectral sensitivity of the optical system. Similar analysis for the RGB branch is considered as described in Ref. [3].

3. DUAL-CAMERA ARRANGEMENT FOR SNAPSHOT SPECTRAL IMAGING WITH A DISPERSIVE BINARY-PHASED DIFFUSER

A schematic layout of the dual-camera optical system is shown in Fig. 1.

As already reported in Ref. [3], our dual-camera system arrangement includes two camera branches: one camera branch includes a monochromatic digital sensor, an imaging lens complete with a wide band-pass filter (BPF), and a transparent, binary-phased optical diffuser made of quartz, positioned at the entrance pupil of the imaging lens. The usage of a pupil-domain, binary-phase encoding diffuser enables us to create a PSF response that consists of two strong diffusive patterns around the first positive and negative diffraction orders of each wavelength of interest. Since the position of the diffraction orders is linearly dependent with wavelength, it is possible to set the chromatic dispersion extent by proper selection of the encoding frequency.

The other camera branch includes a regular RGB image sensor with an imaging lens and a wide BPF and is intended to provide additional precise spatial and coarse spectral information about the scene. A digital processor of the dual-camera system is digitally coupled to the cameras and processes the combination of the monochromatic DD and sharp RGB images to reconstruct the spectral cube using CS-based algorithms. We utilize the generalized matrix equation model for the spectral imaging system that combines the responses of both cameras to the same spectral cube of the object [3]

$$AX = Y; \quad A = \begin{bmatrix} A_M \\ A_{\mathrm{RGB}} \end{bmatrix}, \quad Y = \begin{bmatrix} Y_M \\ Y_{\mathrm{RGB}} \end{bmatrix}, \tag{16}$$

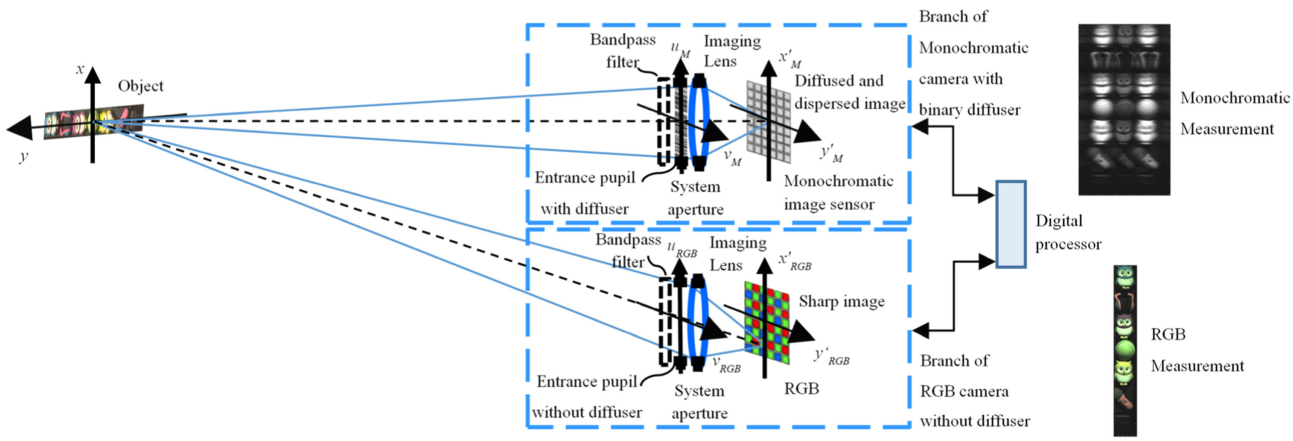

Fig. 1. Schematic layout of the optical system.

where the sensing matrix A models the dispersion and diffusion originating from the diffuser, the spectral sensitivity of the monochromatic camera (embedded in the submatrix A_M), the diffraction-limit blur, and the spectral sensitivity of the RGB camera (embedded in the submatrix A_{RGB}). X is the matrix form of the cascaded spectral cube; Y is a cascaded matrix of the recorded DD monochromatic image Y_M and the sharp demosaiced RGB image Y_{RGB}. The 1D symmetry of the diffuser, which disperses and diffuses the signal along each sensor row separately, enables us to increase the number of vertical pixels of the spectral cube without increasing the dimensions of the sensing matrix. To solve the ill-posed Eq. (16), we utilize the framelet-wavelet transform to set the sparsity along both the 2D-spatial and 1D-spectral dimensions of the spectral cube. We then use a singular value decomposition (SVD) pre-step followed by a Split Bregman iterative process (SBI) for spectral cube reconstruction [3, 7–9].

4. EVALUATION AND CALIBRATION OF THE EXPERIMENTAL SYSTEM

The SSI arrangement was assembled from the same hardware as reported previously in Ref. [3], except for the new binary diffusers. Several binary diffusers were fabricated at the Tel Aviv University Nano-Center facilities with a standard binary staircase technology on a 2.4 mm thick, 127 mm chrome mask made of quartz. Direct writing over the chrome mask was done using a Heidelberg DWL66 Laser Writer. The writing resolution was 0.8 µm, and the critical dimension of the diffusers was 2 µm. The chrome mask was then developed to create a transparent and opaque binary pattern matching the binary-phase design. A design wavelength λ_{des} was set to 550 nm, which was approximately in the middle region of the visible spectrum of interest. This was followed by a single reactive ion etching process. To obtain phase levels of π and 2π for λ_{des}, the nominal etching depth d_{des} was set according to Eq. (14):

$$d_{des} = \frac{2\pi - \pi}{2\pi} \frac{\lambda_{des}}{(n(\lambda_{des}) - 1)} = \frac{1}{2} \frac{550}{(1.4599 - 1)} = 598 \text{ nm}, \tag{17}$$

where $n(\lambda_{des}) = 1.4599$ is the literature value of the refractive index of fused quartz at 550 nm [10]. The mask included several different binary diffuser designs, each with a 3.2 mm clear aperture that matched the entrance pupil diameter of the monochromatic camera. This binary fabrication process is significantly simpler than our former fabrication process of the 14-level phase diffuser [1–3], being composed of a single development and etching iteration instead of four development and etching iterations with three alignment procedures in between iterations. Moreover, it provides better production accuracy, being less prone to spatial errors and depth errors. Post-fabrication validation of the diffusers confirmed that an excellent spatial accuracy was achieved, whereas the actual etching depth was approximately 570 nm, considered to be accurate enough.

SSI experiments were done separately for three types of binary-encoded diffusers: a top-hat encoded diffuser, a cubic-phase encoded diffuser, and a random-phase encoded diffuser. For the top-hat encoded diffuser, we used Eqs. (7)–(9) to set the nominal half-width of the top-hat profile a to be 22 µm, equal to 10 sensor pixels; for the cubic-phase encoded diffuser, we used Eq. (10) and set $\alpha = 90$; for the random-phase encoded diffuser, we set $N_d = 160$ according to Eqs. (11)–(13). Considering the physical parameters of our optical system and the diffuser fabrication process limitations, we chose f_{enc} to be 100 [mm^{-1}], which corresponded to dispersion of ∼4.4 pixels per band (4.4 pixels per 10 nm) at the monochromatic sensor plane. To analyze the performance of the binary-encoding approach, we fabricated two additional binary diffusers: a 1D periodic phase grating and a 1D random phase grating. The additional analysis was used to examine whether the suggested design for generating dispersion and diffusion by binary-phase encoding is advantageous over solely binary dispersion (as provided by the 1D periodic grating) or diffusion (as provided by the 1D random grating). Accordingly, the analytical expression for the phase of the 1D periodic grating is given by

$$\varphi_M(u'_M; \lambda_{des}) = \begin{cases} \pi, & \text{if } \cos\left[2\pi f_{enc,PG} \cdot u'_M\right] > 0, \\ 2\pi, & \text{otherwise} \end{cases} \tag{18}$$

where $f_{enc,PG} = 62.5$ [mm^{-1}]. The analytical expression for the non-encoded 1D random binary diffuser is given in Eqs. (11)–(13) with $N_d = 160$ strips. While our previously reported 1D diffuser featured 400 randomly permutated strips with 14 unique heights [1–3], the new non-encoded 1D random binary diffuser sets a similar optical response. Hence, it was also used to compare our new system configuration with binary-encoded diffusers to a system configuration that has similar optical properties to our previously reported dual-camera system [3].

All five diffusers were fabricated on the same quartz mask and with the same fabrication process. For each diffuser, modeling of the sensing matrix A of the dual-camera SSI optical system was done by direct measurements of the 1D PSFs and spectral efficiency for each wavelength of interest, separately for the monochromatic and the RGB branches, as detailed in Ref. [3]. Figure 2 shows exemplary measured 1D PSF profiles at 420, 550, and 670 nm for the fabricated diffusers, with marked dispersion extent Δ_{disp} and diffusion extent Δ_{diff} at the first diffraction order. The intensity variation between wavelengths is due to the dependence of the diffraction efficiency on wavelengths. Specifically, the diffraction efficiency of the mid-range wavelengths (500–590 nm) is higher on average than that of the blue (420–490 nm) and red (600–670 nm) wavelengths, due to their spectral proximity to the design wavelength λ_{des}.

After characterization of the relative spectral sensitivity, all the wavelengths' 1D PSFs were normalized to have their sum of elements equal to the relative spectral sensitivity for each corresponding wavelength, sensor, and color channel. The normalized PSF was integrated into the corresponding block Toeplitz subsensing matrix as explained in Refs. [1,3]. Figure 3 shows the optically measured sensing matrices A for each system configuration, which correspond to the PSF calibration measurements and the resolution of the spectral cube according to Eq. (16).

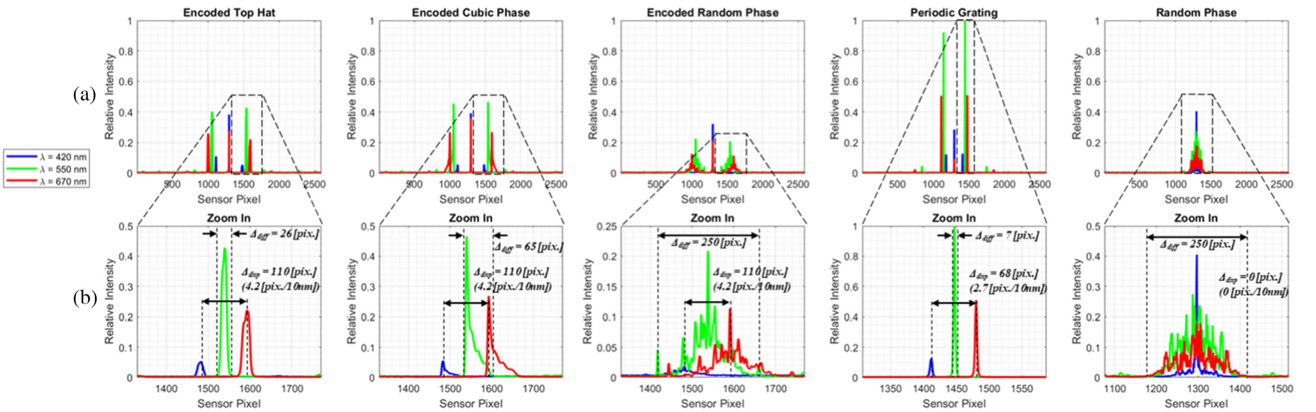

Fig. 2. (a) Experimentally measured 1D PSF profiles for the encoded top-hat, encoded cubic-phase, encoded random-phase, periodic grating, and random-phase diffusers; (b) Zoom-in at the diffusion patterns located in the first diffraction order (zoom-in at the zeroth order for the random phase), with dispersion extent Δ_{disp} and diffusion extent Δ_{diff}.

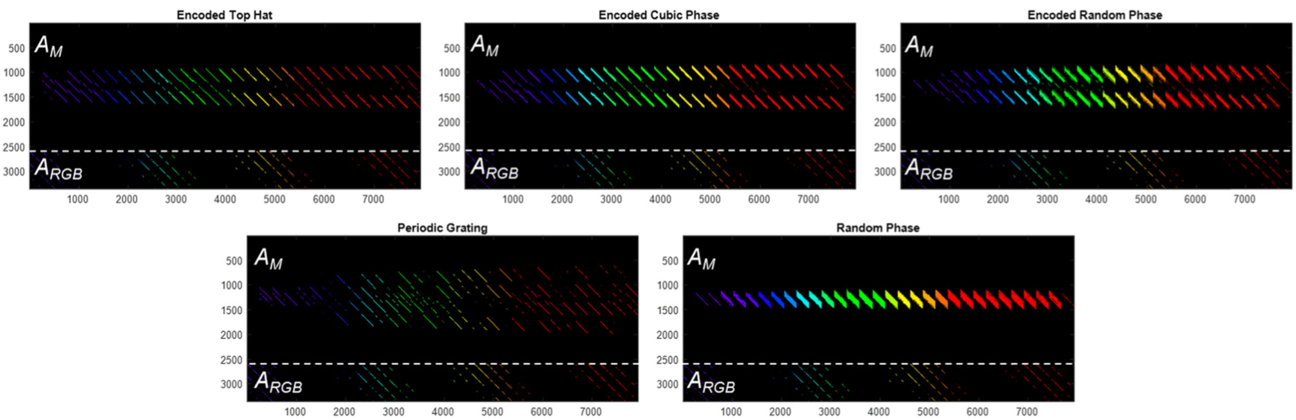

Fig. 3. Sensing matrices A, built from calibration measurements for encoded top-hat, encoded cubic-phase, encoded random-phase, periodic grating, and random-phase diffusers in accordance to Eq. (16). The submatrices A_M, A_{RGB} are annotated in white. For illustration purposes, each block Toeplitz submatrix was colored with correspondence to its matching wavelength as detailed in Ref. [11]. In addition, the matrix values are normalized to the range of [0,1] and encoded in the range of [0,0.0025] for enriched display.

5. RELATION BETWEEN THE COMPRESSION RATIO AND THE DIFFRACTION EFFICIENCY

The compression ratio ρ of a system can be considered the ratio between the number of rows and columns of the sensing matrix. Hence, the compression ratios of each camera branch and the combined dual camera system are given by

$$\rho_M = \frac{N_{x,s,M}}{L N_{x,i}}, \qquad \rho_{\text{RGB}} = \frac{3N_{x,i}}{L N_{x,i}} = \frac{3}{L},$$

$$\rho_{\text{DC}} = \frac{N_{x,s,M} + 3N_{x,i}}{L N_{x,i}}, \tag{19}$$

where ρ_M, ρ_{RGB}, and ρ_{DC} are the compression ratios of the monochromatic camera branch with the diffuser, the RGB branch, and the dual-camera configuration, respectively. $N_{x,s,M}$ is the number of horizontal pixels in the monochromatic sensor, $N_{x,i}$ is the number of horizontal pixels in the spectral cube, and L is the number of spectral bands. However, the dispersed and diffused optical signal may not impinge the entire monochromatic sensor, or, on the other hand, exceed it. As explained in Section 4, the binary-encoded diffuser creates two strong diffusive patterns that are located around the first positive and negative diffraction orders of each wavelength of interest. In practice, the binary-phase encoding method creates an infinite number of diffraction orders, which weaken as the diffraction order number increases. The finite size of the monochromatic sensor creates the restriction that only a limited number of diffraction orders will impinge the sensor plane. The extent of the dispersed and diffused signal between the $-j$th and $+j$th diffraction orders on the monochromatic sensor plane can be calculated as

$$\Delta_{\text{DD}}^{j} = \frac{2}{\delta_p} \frac{|j|\,\lambda_{\max} R f_{\text{enc}} D_{\text{EnP}}}{D_{\text{ExP}}} + \Delta_{\text{diff}} + N_{x,i}, \tag{20}$$

where δ_p is the size of a sensor pixel, $\lambda_{\max}$ is the longest wavelength that enters the optical system, R is the distance between the exit pupil and the sensor plane, f_{enc} is the spatial encoding frequency of the diffuser, D_{EnP} and D_{ExP} are the diameters of the entrance pupil and exit pupil, respectively, and Δ_{diff} is the local diffusion extent around the jth diffraction order. Therefore, the effective compression ratio of the optical system can be

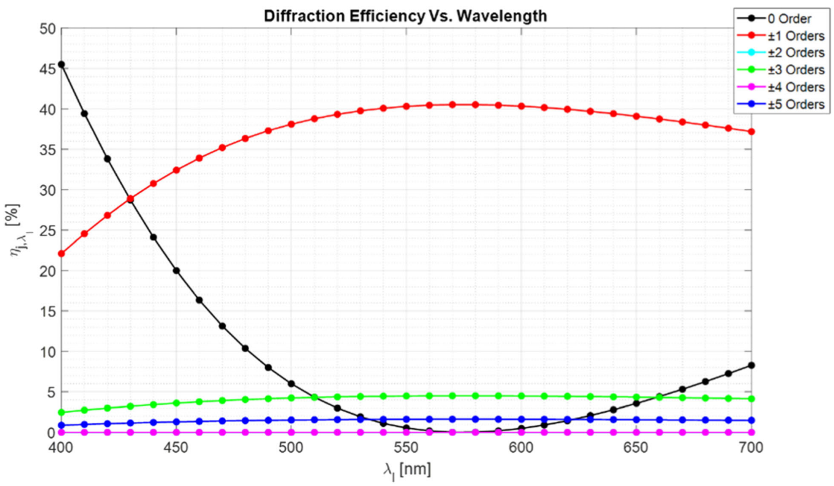

Fig. 4. Analytical calculation for the diffraction efficiency η_{j,λ_l} of a periodic binary phase grating as function of wavelength, for 11 diffraction orders between -5 and $+5$. The efficiencies of the even diffraction orders ± 2 and ± 4 are negligible.

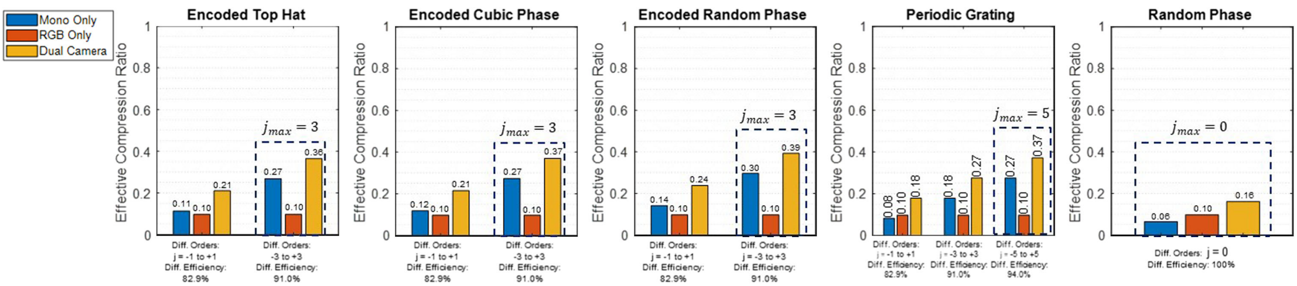

Fig. 5. Relation between the effective compression ratio, the range of diffraction orders, and the total diffraction efficiency for either a single-camera or dual-camera configuration with each of the tested diffusers. The analysis for the non-encoded random-phase diffuser was treated as a binary-staircase periodic DOE with zero frequency.

calculated as

$$\rho_{M,\mathrm{eff}} = \frac{\min\{\Delta_{\mathrm{DD}}^{j_{\max}}, N_{x,s,M}\}}{L N_{x,i}}, \qquad \rho_{\mathrm{RGB}} = \frac{3N_{x,i}}{L N_{x,i}} = \frac{3}{L},$$
$$\rho_{\mathrm{DC,eff}} = \frac{\min\{\Delta_{\mathrm{DD}}^{j_{\max}}, N_{x,s,M}\} + 3N_{x,i}}{L N_{x,i}}, \tag{21}$$

where $j_{\max}$ is the maximal diffraction order that impinges on the monochromatic sensor.

To evaluate the diffraction efficiency of the jth order of the suggested binary diffusers, we considered an ideallytrans-mitting, binary-staircase, periodic diffractive optical element (DOE) with nominal phase modulation of π around λ_{des} and a duty cycle of 50%. Being a periodic element, the field transmittance of the jth diffraction order is calculated as the jth Fourier coefficient of its Fourier series representation:

$$T_{j,\lambda_l} = \int_0^1 \exp\left[i Q_{\lambda_l}(s)\right]\exp[-i2\pi s]\,\mathrm{d}s, \tag{22}$$

where

$$Q_{\lambda_l}(s) = \begin{cases} 0, & 0 \le s < 0.5 \\ c_{h,\lambda_l}\pi, & 0.5 \le s < 1, \end{cases} \tag{23}$$

where the parameter c_{h,λ_l} is a groove-depth-dependent mismatch combined with a wavelength-dependent mismatch. c_{h,λ_l} can be calculated as

$$c_{h,\lambda_l} = \frac{\hat{h}_{\max}}{h_{\max}} \frac{\lambda_{\mathrm{des}}}{\lambda_l} \frac{n(\lambda_l) - 1}{n(\lambda_{\mathrm{des}}) - 1}, \tag{24}$$

where $h_{\max}$ is the maximal groove depth of an ideal binary-staircase periodic DOE and $\hat{h}_{\max}$ is the maximal groove depth of the fabricated DOE with actual height profile. It can be shown that the power efficiency of the jth diffraction order and the wavelength λ_l is given by

$$\eta_{j,\lambda_l}\,[\%] = 100 \cdot \left|T_{j,\lambda_l}\right|^2 = 100 \cdot \mathrm{sinc}^2\left(\frac{j}{2}\right) \frac{\mathrm{sinc}^2\left[c_{h,\lambda_l} - j\right]}{\mathrm{sinc}^2\left[\frac{c_{h,\lambda_l} - j}{2}\right]}. \tag{25}$$

Figure 4 shows analytical calculations for the diffraction efficiency η_{j,λ_l} of the fabricated binary diffuser as function of wavelength, for 11 diffraction orders between -5 and $+5$, according to Eqs. (24) and (25). As indicated by the diffraction efficiency curves, most of the optical energy is concentrated in the -1, 0, and $+1$ diffraction orders.

The total diffraction efficiency for all wavelengths between the $\pm j'$ orders can be calculated as

$$\eta_{j',\mathrm{tot}} = \frac{1}{L}\sum_{l=1}^{L}\left[\sum_{j=-j'}^{j'}\eta_{j,\lambda_l}\right]. \quad (26)$$

Figure 5 summarizes the relation between the effective compression ratio, the range of diffraction orders, and the total diffraction efficiency for either a single-camera or dual-camera configuration with each of the tested diffusers.

6. OPTICAL EXPERIMENTS FOR SSI WITH DUAL-CAMERA AND BINARY DIFFUSERS

Each system test with a binary diffuser was done by precisely placing its active area at the front aperture plane of the monochromatic camera's lens, which consolidates with its entrance pupil. A column of six colored volumetric toys with approximate size of 30 mm × 30 mm each was illuminated by two 12 W white LEDs and placed 50 cm in front of the prototype. Reference spectral cubes of the object were acquired in a separate procedure by the monochromatic camera without the diffuser, using a set of $L = 31$ narrow-band-pass filters at the spectral range of 400–700 nm in equal spacing of 10 nm as described in Ref. [3]. The full resolution of the monochromatic camera was $N_{y,s} \times N_{x,s} = 1944 \times 2592$ pixels (same as the RGB camera). Due to poor optical signal in the remote bands, only 27 bands were evaluated in the spectral range of 420–680 nm, and hence the resolution of the spectral cube was $N_{y,i} \times N_{x,i} \times L = 1944 \times 256 \times 27$. For each system setting with a binary diffuser, a single snapshot of the scene was taken from each camera as shown in Fig. 6. This provided the measurement Y as described in Eq. (16). Table 1 specifies the integration time of the cameras for each system configuration.

After snapshot acquisition, we used the SVD pre-step followed by the SBI process for reconstruction of the spectral cube [3,8,9]. Specifically, we adapted our wavelet and framelet decomposition reconstruction algorithm to fit large spectral cubes of 1944 × 256 spatial pixels, rather than 256 × 256 pixels. Quality evaluation of the reconstructed spectral cubes was done by comparison with reflectance reference spectral cubes measured directly and in advance with the set of $L = 27$ narrow-band-pass filters. Reflectance measurements of the "Toys" scene included an acquisition of a spectral cube of a diffusive white paper target, followed by normalization of the raw spectral cube of the "Toys" scene by the spectral cube of the white paper target.

Table 1. Camera Integration Time per System Configuration

Diffuser Type	Integration Time [ms]
Encoded top hat	51.4
Encoded cubic phase	46.8
Encoded random phase	42.1
Periodical grating	53.1
Random phase	30.4

Figure 7 shows the experimental results for the "Toys" scene. Reconstruction results for the spectral cube were achieved after using preliminary SVD and followed by three SBI-type iterations. The SVD and each SBI-type iteration run time took 113.5 s and 35.4 s to complete, respectively, using Matlab software, which was installed on a PC with an Intel i7-8665U, 1.9 GHz processor, 32 G RAM, and Windows 10 operating system. For comparison, the run times for the SVD and an SBI-type iteration for a spectral cube of size 256 × 256 × 27 take 15.5 s and 5.2 s to complete on the same computer, and hence the runtime is linearly proportional to the overall dimensions of the spectral cube. Figure 7(a) shows the RGB representation of the reference spectral cube. The conversion of the spectral coordinates to RGB values was done in accordance with the CIE standard observer color matching functions as detailed in Ref. [11]. Figure 7(b) shows zoom in images to eight regions of interest (ROIs) for the spectral reconstruction analysis. Figure 7(c) shows reference and reconstructed spectra for eight spatial points marked by a red dot in the corresponding row in Fig. 7(b), for each system configuration with a binary diffuser. Figures 7(d) and 7(e) show average values of peak signal-to-noise ratio (PSNR) and spectral angle mapping (SAM) [3,12], respectively, for 1024 spatial pixels marked by black rectangles in the corresponding zoom in images in Fig. 5(b). The system configurations are abbreviated as ETH (encoded top-hat binary diffuser), ECP (encoded cubic-phase binary diffuser), ERP (encoded random-phase binary diffuser), PG (periodic grating binary diffuser), and RP (random-phase binary diffuser).

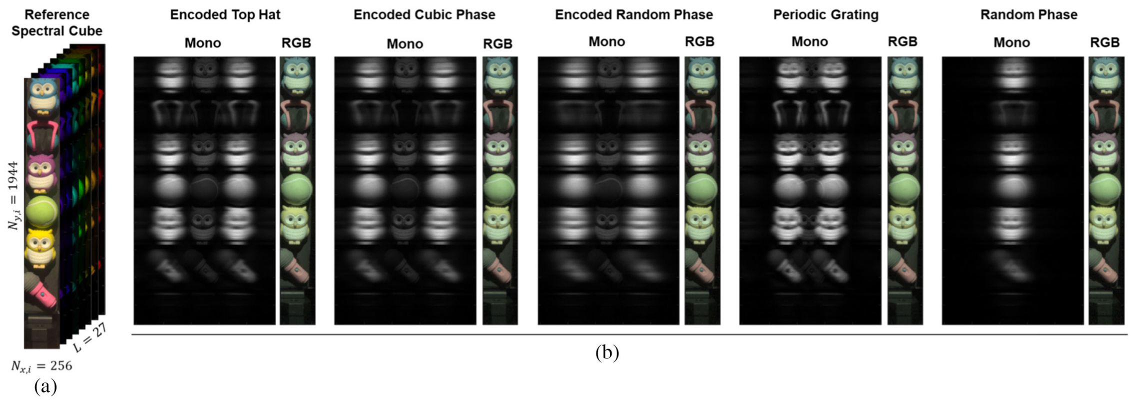

Fig. 6. (a) RGB representation and dimensions of the reference spectral cube for the "Toys" scene; (b) experimental sensor measurements from the dual-camera prototype with each type of binary diffuser. For clarity, the displays of the monochromatic snapshots do not exceed the ±1 diffraction orders.

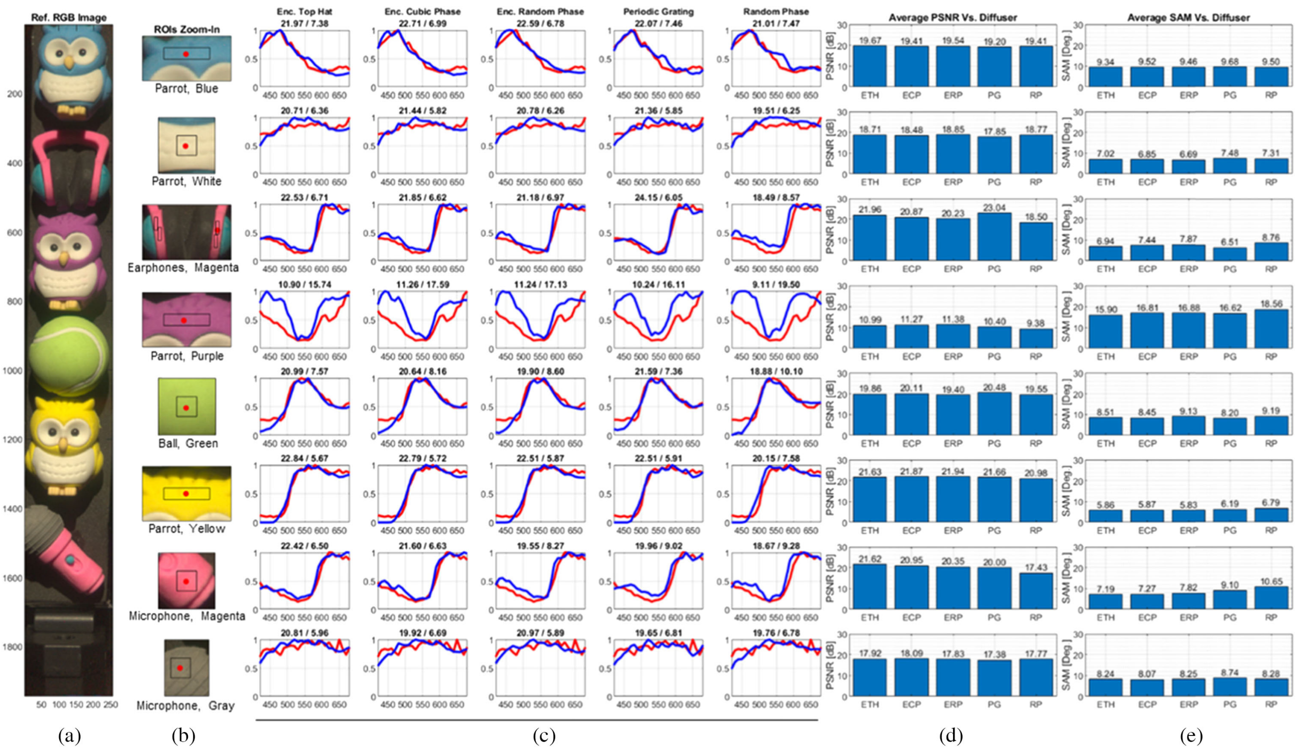

Fig. 7. (a) RGB representation of the reference spectral cube for the "Toys" scene; (b) zoom-in to eight regions of interest; (c) reference (red) and reconstructed (blue) spectra for the marked red dot in the corresponding zoomed-in images in Fig. 5(b) for each system configuration, with PSNR/SAM values marked in the title, respectively; (d) average PSNR and (e) SAM values for 1024 spatial pixels marked by black rectangles in the corresponding zoomed-in images in Fig. 5(b).

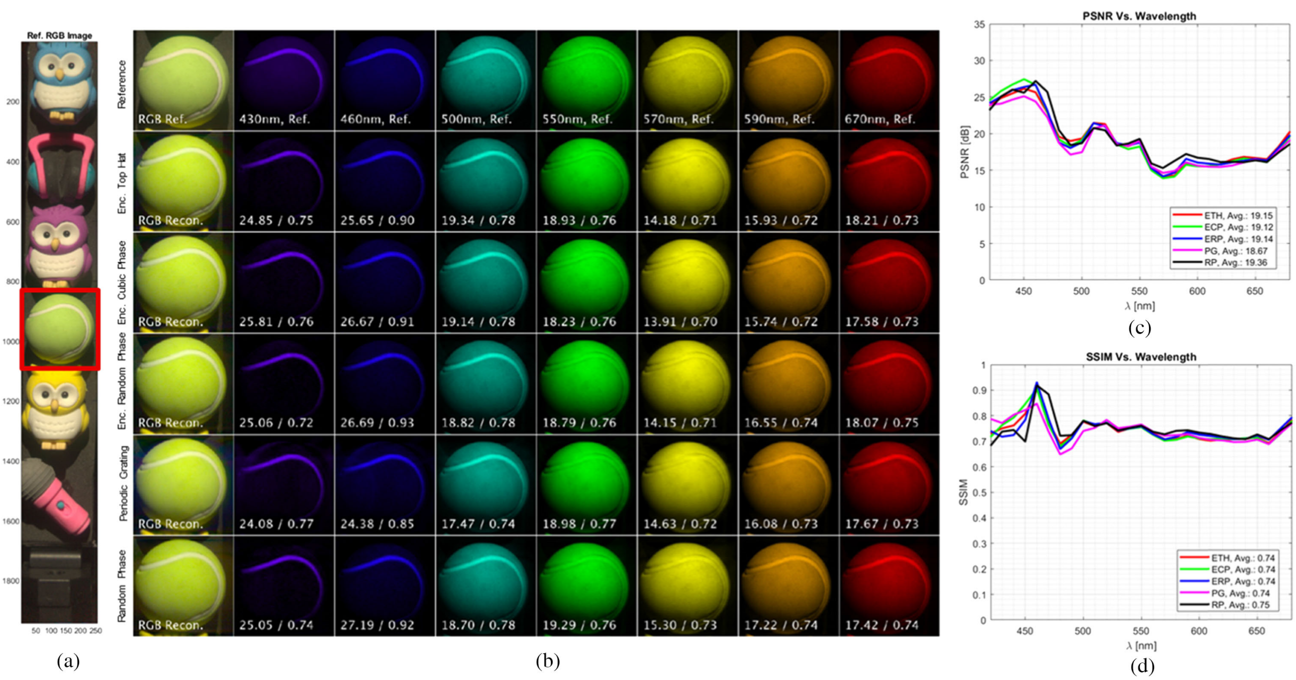

Fig. 8. (a) RGB representation of the reference spectral cube for the "Toys" scene, with the "Tennis Ball" object marked by a red square; (b) reference and reconstructed single-wavelength images at selected wavelengths, with reconstructions' PSNR/SSIM values, respectively; (d) PSNR and (e) SSIM values as functions of wavelength for experimentally reconstructed spectral cubes of the "Tennis Ball" object from the "Toys" scene.

Figure 8(b) shows selected 7 out of 31 reference and reconstructed single-wavelength images for each system configuration, for the "Tennis Ball" object, marked by a red square in Fig. 8(a). The size of the "Tennis Ball" ROI is 256 × 256 spatial pixels. Block-matching and 3D denoising (BM3D) [13] was used to subtract spatial noise from each reference and

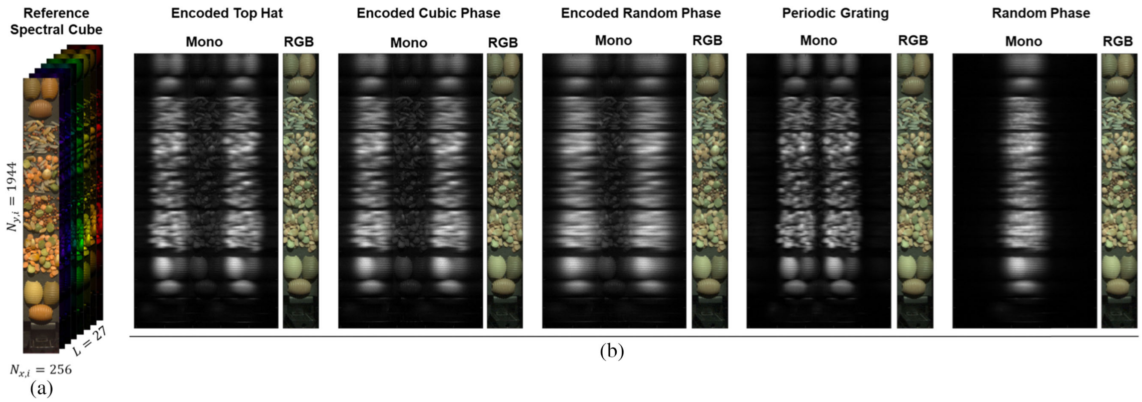

Fig. 9. (a) RGB representation and dimensions of the reference spectral cube for the "Seeds and Pasta" scene; (b) experimental sensor measurements from the dual-camera prototype with each type of binary diffuser. For clarity, the displays of the monochromatic snapshots do not exceed the ±1 diffraction orders.

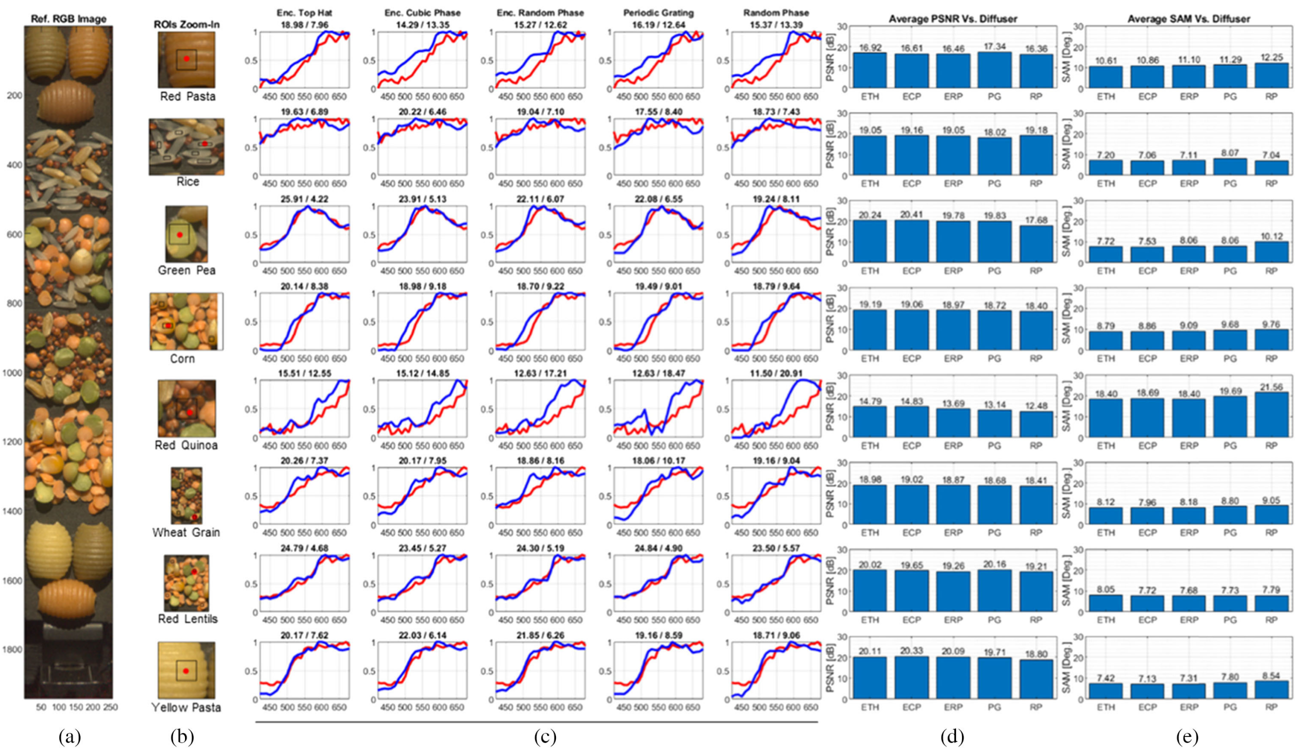

Fig. 10. (a) RGB representation of the reference spectral cube for the "Seeds and Pasta" scene; (b) zoom-in to eight regions of interest; (c) reference (red) and reconstructed (blue) spectra for the marked red dot in the corresponding zoomed-in images in Fig. 8(b) for each system configuration, with PSNR/SAM values marked in the title, respectively; (d) average PSNR and (e) SAM values for 1024 spatial pixels marked by black rectangles in the corresponding zoomed-in images in Fig. 8(b).

reconstructed single-wavelength image. PSNR and structural similarity index matching (SSIM) [14] values are marked for each reconstructed single-wavelength image. Figures 8(c) and 8(d) shows the PSNR and SSIM values as function of wavelength, for the reconstructed single-wavelength "Tennis Ball" object.

Figure 9 shows experimental measurements obtained with the dual-camera prototype for a "Seeds and Pasta" scene, which was assembled from mixtures of real dried colored pasta, rice, red quinoa, wheat grains, green peas, red lentils, and corn.

Figure 10 shows the experimental results for the "Seeds and Pasta" scene that was illuminated by the white LEDs. Reconstruction results for the spectral cube were also achieved after using preliminary SVD and SBI-type iterations, with equal run times as the "Toys" scene's reconstruction. Figure 10(a) shows the RGB representation of the reference spectral cube. Figure 10(b) shows zoomed-in images to eight ROIs for spectral reconstruction analysis. Figure 10(c) shows reference and reconstructed spectra for eight spatial points marked by the red dot in the corresponding row in Fig. 10(b), for each system configuration with a binary diffuser. Figures 10(d) and 10(e)

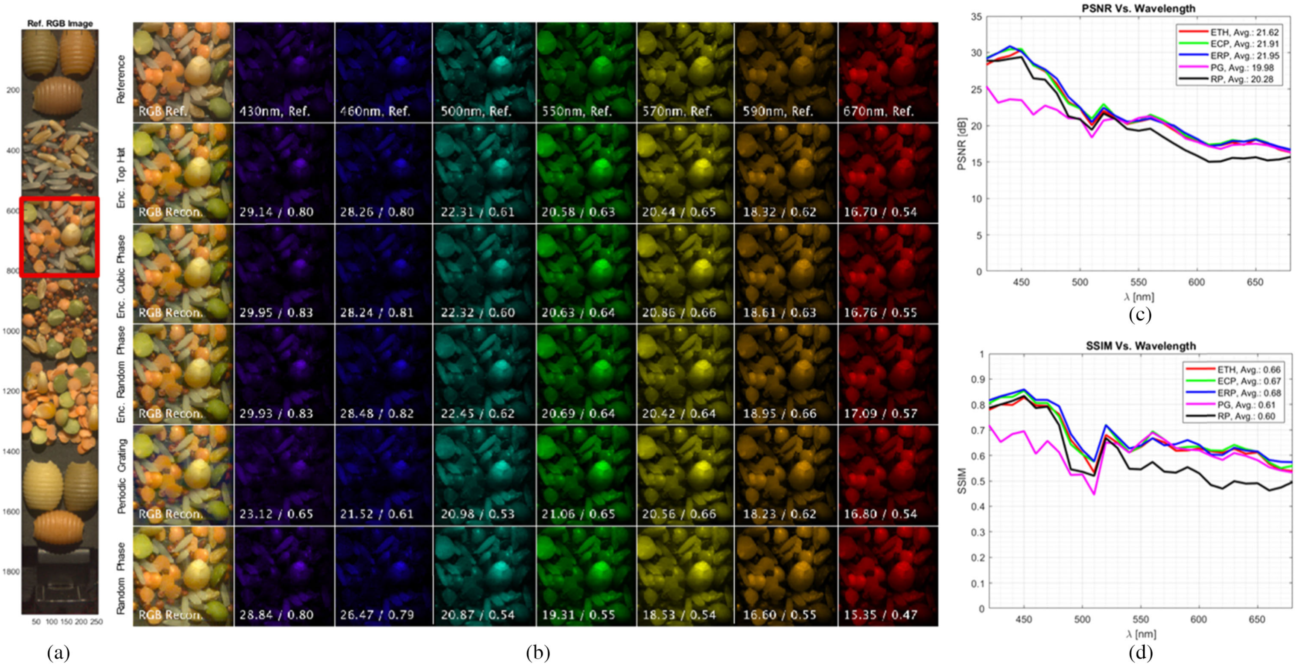

Fig. 11. (a) RGB representation of the reference spectral cube for the "Seeds and Pasta" scene, with the "Seeds Mixture" marked by a red square; (b) reference and reconstructed single-wavelength images at selected wavelengths, with reconstructions' PSNR/SSIM values, respectively; (d) PSNR and (e) SSIM values as functions of wavelength for experimentally reconstructed spectral cubes of the "Seeds Mixture" from the "Seeds and Pasta" scene.

show average values of PSNR and SAM, respectively, for 1024 spatial pixels marked by black rectangles in the corresponding zoom in images in Fig. 10(b).

Figure 11(b) shows selected 7 out of 31 reference and reconstructed single-wavelength images for each system configuration, for the "Seeds Mixture" marked by a red square in Fig. 11(a). The size of the "Seeds Mixture" ROI is 256×256 spatial pixels and contains rice, corn, red quinoa, red lentils, green peas, and wheat grains. BM3D denoising was used to subtract spatial noise from each reference and reconstructed single-wavelength image. PSNR and SSIM values are marked for each reconstructed single-wavelength image. Figures 11(c) and 11(d) shows the PSNR and SSIM values as function of wavelength for the reconstructed single-wavelength "Seeds Mixture" objects.

Figures 7, 8, 10, and11 indicate that very good reconstruction results are achieved for both scenes and for each system configuration. The differences in performance for each system configuration are more prominent in terms of numerical figures of merit. Tables 2 and 3 summarize the average PSNR, SAM, and SSIM values for spectral and spatial reconstruction, for each scene and for each system configuration. To assess the contribution of each camera branch, we have also performed signal reconstruction from each of the camera branches separately. The best performance for each scene is marked by bold text.

The spectra reconstruction summary in Table 2 indicates that each combination of a monochromatic camera with a binary diffuser and an RGB camera outperforms the performance achieved by each camera branch separately. This observation coincides with our previous work [3], which has demonstrated that adding a regular RGB camera branch side-by-side with the monochromatic branch with a 14-level diffuser improves the quality of the reconstruction. While the RGB camera has an important role for the reconstruction quality, our current analysis also shows that the dual-camera arrangement achieves better spectral reconstruction with respect to the RGB camera branch by itself. This is explained by the fact that a dual-camera system model is better presented with respect to either a monochromatic system model or an RGB system model. Considering the dual-camera arrangements, Table 2 indicates that the suggested phase encoding designs, which combine both dispersion and diffusion, achieve better results than either a periodic grating (i.e., PSF with strong dispersion but negligible diffusion) or a random phase (PSF with negligible dispersion but with strong diffusion). Specifically, the encoded top-hat and encoded cubic-phase designs achieve overall better results in terms of PSNR and SAM figures of merit. This can be explained by the stronger dispersion, which is introduced by the encoded phase diffusers ($\Delta_{disp} = [4.4$ Pixels/10 nm] versus $\Delta_{disp} = [2.7$ Pixels/10 nm]), with the combination of the local diffusion. Specifically, the empirical results indicate that a limited level of diffusion leads to better spectral reconstruction on average.

Spatial reconstruction analysis was limited to two subscenes: the rich-textured "Seeds Mixture" and the "Tennis Ball" object, which has small number of textures and was slightly less prone to perspective shifts than the other toys (the perspective shifts are caused by the side-by-side camera arrangement). Table 3 indicates that the dual-camera arrangements also achieve better spatial reconstruction with respect to the monochromatic branch by itself. However, the RGB camera branch may achieve better spatial reconstruction quality when considered

Table 2. Spectra Reconstruction: Average PSNR and SAM for Eight Selected ROIs

Scene	System Configuration	Figure of Merit	Encoded Top Hat	Encoded Cubic Phase	Encoded Random Phase	Periodic Grating	Random Phase
Toys	Monochromatic branch only	PSNR [dB]	7.70	7.87	8.69	8.19	8.36
		SAM [deg]	36.74	36.12	32.85	34.20	33.79
	RGB branch only	PSNR [dB]			16.27		
		SAM [deg]			11.06		
	Dual camera	PSNR [dB]	**19.04**	18.88	18.69	18.75	17.72
		SAM [deg]	8.62	8.78	8.99	9.06	9.88
Seeds and Pasta	Monochromatic branch only	PSNR [dB]	7.31	7.73	8.40	8.60	9.55
		SAM [deg]	41.42	39.07	35.44	34.50	30.23
	RGB branch only	PSNR [dB]			15.93		
		SAM [deg]			13.33		
	Dual camera	PSNR [dB]	**18.66**	18.63	18.27	18.20	17.56
		SAM [deg]	9.54	**9.48**	9.62	10.14	10.76

Table 3. Single-Wavelength Reconstruction: Average PSNR and SSIM for Selected Subscenes

Scene	System Configuration	Figure of Merit	Encoded Top Hat	Encoded Cubic Phase	Encoded Random Phase	Periodic Grating	Random Phase
Toys, Tennis Ball	Monochromatic branch only	PSNR [dB]	16.96	18.56	**20.01**	15.32	16.61
		SSIM	0.67	0.72	0.74	0.6	0.52
	RGB branch only	PSNR [dB]			19.02		
		SSIM			**0.75**		
	Dual camera	PSNR [dB]	19.15	19.12	19.14	18.67	19.36
		SSIM	0.74	0.74	0.74	0.74	**0.75**
Seeds and Pasta, Seeds Mixture	Monochromatic branch only	PSNR [dB]	17.32	17.49	19.03	13.64	19.14
		SSIM	0.49	0.48	0.51	0.34	0.5
	RGB branch only	PSNR [dB]			**22.24**		
		SSIM			**0.69**		
	Dual camera	PSNR [dB]	21.62	21.91	21.95	19.98	20.28
		SSIM	0.66	0.67	0.68	0.61	0.60

by itself. This is explained by the spatial information loss in the monochromatic branch, which is resulted by the diffuser, and hence the information combination may slightly decrease the spatial figures of merit. Furthermore, it is evident that the RGB camera achieves better quality for the "Seeds Mixture" region of interest with respect to the "Tennis Ball" region of interest. This can be explained by the larger depth profile of the Tennis Ball with respect to the seeds mixture's profile, which may yield perspective errors in the RGB camera branch. Nevertheless, very good visual reconstructions are achieved for each dual-camera arrangement for both scenes as can be seen in Figs. 8 and 11. Considering the dual-camera arrangements, Table 3 indicates that the random-phase binary diffuser achieves slightly better results for the "Tennis Ball" object; however the encoded random phase achieves best results for the more complex "Seeds Mixture" scene. Specifically, it outperforms the non-encoded random phase profile by nearly 1.7 [dB] and 0.08 SSIM units.

7. DISCUSSION AND CONCLUSIONS

In this work we present the design, fabrication, and experiments with a set of new 1D binary-phase diffusers for performing SSI using two regular digital cameras and by resorting to CS algorithms. The investigated dual-camera arrangement is an expansion of the previously reported SSI system, which consists of one or two regular cameras with a 14-level random diffuser, which was fabricated in a standard four-layer photolithography process [1–3]. The suggested binary diffusers are not only fabricated in a much simpler one-layer photolithography process but also maintain high optical throughput along with improved optical response by combining strong dispersion and diffusion properties. The latter is achieved efficiently by binary-phase encoding, which creates a 1D PSF profile where the region of the diffusive patterns is primarily located at the -1 and 1 diffractive orders, and hence its location is linearly dependent with wavelength.

Three different types of binary-encoded phase diffusers, a binary periodic grating diffuser and a binary non-encoded random diffuser, were fabricated and evaluated. The analysis indicates that the binary-encoded phase diffusers exhibit overall better results of spectral reconstruction. Specifically, the system configuration with the binary-encoded top-hat profile yields the best results on average. Spatial reconstruction is found to be similar between all system configurations in cases which the scene is relatively uniform (e.g., the tennis ball in the "Toys" scene), with slight advantage to the non-encoded random-phase diffuser. However, for more complex scenes (e.g., the seeds mixture in the "Seeds and Pasta" scene), the system configuration with the binary-encoded random-phase diffuser achieves the best results, where also the other binary-encoded diffusers outperform the binary periodic grating or the non-encoded random grating. Nevertheless, better spatial reconstructions may be obtained if only the RGB branch is considered. An approximate comparison between the suggested binary-encoded phase diffusers and our previously reported 14-level random-phase diffuser [1–3] is done through the system tests with the non-encoded binary random-phase diffuser, which has similar optical properties. To conclude, the suggested combination of dispersion and diffusion by binary-phase encoding could be especially efficient for spatial and spectral reconstruction of scenes that include rich or moderate levels of spatial details and moderate spectral details. Hence, we expect that the reconstruction performance of our system will be superior for natural scenes that include smooth reflectance spectra. The latter is contributed to the CS-based algorithm that is utilized by our suggested system, which assumes sparsity along both spectral and spatial domains. While the monochromatic camera branch provides important spatial and spectral information about the scene, the secondary RGB camera provides important precise spatial and coarse spectral information about the scene, which contributes to the overall quality of the reconstruction for each system configuration. An analysis for the added value of the RGB camera is also covered in our previous publication [3].

In this work we also introduce enhanced system capabilities and fully exploit the nondispersed and diffused vertical dimension of the sensor. This enables to reconstruct spectral cubes with spatial dimensions of 1944×256 pixels in a single snapshot, compared with the previously reported 256×256 spatial pixels [1,3]. While our reconstruction algorithm outputs 31 spectral bands in the 400–700 nm interval, we limited our analysis to 27 bands in the 420–680 nm region due to weak signal at the far regions of the spectrum. The snapshot acquisition times of either 1944×256 scenes or 256×256 scenes are equivalent, however the reconstruction time for the larger cubes increases with linear proportion to the overall dimensions of the cube. Nevertheless, the 1D symmetry of the diffuser enables us to use the same sensing matrix for large 1944×256 scenes or smaller 256×256 scenes and thus contributes to reasonable run times albeit with the significant enlargement of the spectral cube.

The results of this work indicate the progress towards applications of snapshot spectral imagers in variety of fields, specifically in cases where light throughput, acquisition time, size, and cost are cardinal.

Funding. Ministry of Science, Technology and Space (3-10416, 3-13601); Blavatnik Family Foundation.

Acknowledgment. The authors dedicate this research article to the late co-author, Professor Michael A. Golub, who passed away in December 2019. Professor Golub was a senior OSA member and a worldwide expert in the field of diffractive optics, micro-optics, optical design, and computational imaging. Professor Golub was a true professional, mentor, and friend, and our research group cherishes his work and remembrance.

Disclosures. The authors declare no conflicts of interest.

† Deceased.

REFERENCES

1. M. A. Golub, A. Averbuch, M. Nathan, V. A. Zheludev, J. Hauser, S. Gurevitch, R. Malinsky, and A. Kagan, "Compressed sensing snapshot spectral imaging by a regular digital camera with an added optical diffuser," Appl. Opt. **55**, 432–443 (2016).
2. J. Hauser, M. A. Golub, A. Averbuch, M. Nathan, V. A. Zheludev, O. Inbar, and S. Gurevitch, "High-photon-throughput snapshot colour imaging using a monochromatic digital camera and a pupil-domain diffuser," J. Mod. Opt. **66**, 710–725 (2019).
3. J. Hauser, M. A. Golub, A. Averbuch, M. Nathan, V. A. Zheludev, and M. Kagan, "Dual-camera snapshot spectral imaging with pupil-domain optical diffuser and compressed sensing algorithms," Appl. Opt. **59**, 1058–1070 (2020).
4. S. Shwartz, M. A. Golub, and S. Ruschin, "Computer-generated holograms for fiber optical communication with spatial-division multiplexing," Appl. Opt. **56**, A31–A40 (2017).
5. E. R. Dowski and W. Cathey, "Extended depth of field through wavefront coding," Appl. Opt. **34**, 1859–1866 (1995).
6. M. A. Golub, "Generalized conversion from the phase function to the blazed surface-relief profile of diffractive optical elements," J. Opt. Soc. Am. A **16**, 1194–1201 (1999).
7. Z. Shen, "Wavelet frames and image restorations," in *International Congress of Mathematicians*, R. Bhatia, ed. (Hindustan Book Agency, 2010), Vol. **IV**, pp. 2834–2863.
8. T. Goldstein and S. Osher, "The split Bregman method for L1 regularized problems," SIAM J. Imaging Sci. **2**, 323–343 (2009).
9. J. F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," Multiscale Model. Simul. **8**, 337–369 (2009).
10. I. H. Malitson, "Interspecimen comparison of the refractive index of fused silica," J. Opt. Soc. Am. **55**, 1205–1208 (1965).
11. D. H. Foster, 2018 https://personalpages.manchester.ac.uk/staff/david.foster/Tutorial_HSI2RGB/Tutorial_HSI2RGB.html.
12. F. A. Kruse, A. B. Lefkoff, J. B. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data," Remote Sens. Environ. **44**, 145–163 (1993).
13. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," Proc. SPIE **6**, 606414 (2006).
14. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).