

3D compressive spectral integral imaging

WEIYI FENG,^{1,2} HOOVER RUEDA,² CHEN FU,² GONZALO R. ARCE,²
WEIJI HE,¹ AND QIAN CHEN^{1,*}

¹Jiangsu Key Laboratory of Spectral Imaging and Intelligence Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, China

²Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA
*chenq@163.com

Abstract: A novel compressive 3D imaging spectrometer based on the coded aperture snapshot spectral imager (CASSI) is proposed. By inserting a microlens array (MLA) into the CASSI system, one can capture spectral data of 3D objects in a single snapshot without requiring 3D scanning. The 3D spatio-spectral sensing phenomena is modelled by computational integral imaging in tandem with compressive coded aperture spectral imaging. A set of focal stack images is reconstructed from a single compressive measurement, and presented as images focused on different depth planes where the objects are located. The proposed optical system is demonstrated with simulations and experimental results.

© 2016 Optical Society of America

OCIS codes: (110.0110) Imaging systems; (110.4234) Multispectral and hyperspectral imaging; (110.1758) Computational imaging; (170.1630) Coded aperture imaging.

References and links

1. M. H. Kim, T. A. Harvey, D. S. Kittle, H. Rushmeier, J. Dorsey, Richard O. Prum, and B. J. Javidi, "3D Imaging Spectroscopy for Measuring Hyperspectral Patterns on Solid Objects." *ACM Trans. Graphics* **31**(4), 13–15 (2012).
2. P. Latorre-Carmona, E. Sanchez-Ortiga, X. Xiao, F. Pla, M. Martínez-Corral, H. Navarro, G. Saavedra, and B. Javidi, "Multispectral integral imaging acquisition and processing using a monochrome camera and a liquid crystal tunable filter." *Opt. Express* **20**(23), 25960–25969 (2012).
3. A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging" *Appl. Opt.* **47**(10), B44–B51 (2008).
4. A. Plaza and C.-I. Chang, "Preface to the special issue on high performance computing for hyperspectral imaging," *International Journal of High Performance Computing Applications*, **22**(4), 363–365 (2008).
5. M. Cho and B. Javidi, "Three-dimensional visualization of objects in turbid water using integral imaging," *J. Disp. Technol.* **6**(10), 544–547 (2010).
6. I. Quinzán Suárez, P. Latorre Carmona, P. García Sevilla, E. Boldo, F. Pla, V. García Jimenez, R. Lozoya, and G. Pérez de Lucía, "Int. Conf. on Pat. Rec. Applic. and Methods, 386–393 (2012).
7. G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "An introduction to compressive coded aperture spectral imaging," *IEEE Signal Processing Magazine* **31**(1), 105–115 (2014).
8. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, "Video rate spectral imaging using a coded aperture snapshot spectral imager," *Opt. Express* **17**(8), 6368–6388 (2009).
9. P. Llull, X. Yuan, L. Carin, and D. J. Brady, "Image translation for single-shot focal tomography," *Optica* **2**(9), 822–825 (2015).
10. T. H. Tsai, P. Llull, X. Yuan, L. Carin, and David J. Brady, "Spectral-temporal compressive imaging," *Opt. Lett.* **40**(17), 4054–4057 (2015).
11. X. Xiao, B. Javidi, M. Martínez-Corral, and A. Stern, "Advances in three-dimensional integral imaging: sensing, display, and applications [Invited]," *Appl. Opt.* **52**(4), 546–560 (2013).
12. J. -S. Jang and B. Javidi, "Three-dimensional projection integral imaging using micro-convex-mirror arrays," *Opt. Express* **12**(6), 1077–1083 (2004).
13. S. -H. Hong, J. -S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Opt. Express* **12**(3), 483–491 (2004).
14. R. Horisaki, X. Xiao, J. Tanida, and B. Javidi, "Feasibility study for compressive multi-dimensional integral imaging," *Opt. Express* **21**(4), 4263–4279 (2013).
15. W. Feng, H. Rueda, C. Fu, Q. Chen, and G. R. Arce, "Compressive spectral integral imaging using a microlens array," *Proc. SPIE* **9857**, 985706 (2016).
16. X. Lin, J. Suo, G. Wetzstein, Q. Dai, and R. Raskar, "Coded focal stack photography," *IEEE International Conference on Computational Photography*, 1–9 (2013).
17. M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," *ACM Trans. Graphics* **25**(3), 924–934 (2006).

18. D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.* **49**(36), 6824–6833 (2010).
19. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, "Spectral image estimation for coded aperture snapshot spectral imagers," *Proc. SPIE* **7076**, 707602 (2008).
20. J. Tan, Y. Ma, H. Rueda, D. Baron, and G. R. Arce, "Compressive hyperspectral imaging via approximate message passing," *IEEE J. of Selected Topics in Signal Processing*, **10**(2), 389–401 (2016).
21. X. Yuan, T. H. Tsai, R. Zhu, P. Llull, D. Brady, and L. Carin, "Compressive hyperspectral imaging with side information," *IEEE J. Sel. Top. Sig. Proc.* **9**(6), 964–976 (2015).
22. H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," *J. Opt. Soc. Am. B* **28**(11), 2400–2413 (2011).
23. F. Jin, J. -S. Jang, and B. Javidi, "Effects of device resolution on three-dimensional integral imaging," *Opt. Lett.* **29**(12), 1345–1347 (2004).
24. M. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Sig. Proc.* **1**(4), 586–597 (2007).
25. J. -S. Jang and B. Javidi, "Depth and lateral size control of three-dimensional images in projection integral imaging," *Opt. Express* **12**(16), 3778–3790 (2004).
26. H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Trans. Image Processing* **23**(4), 1896–1908 (2014).
27. H. Rueda, D. Lau, and G. R. Arce, "Multi-spectral compressive snapshot imaging using RGB image sensors," *Opt. Express* **23**(9), 12207–12221 (2015).
28. H. Rueda, H. Arguello, and G. R. Arce, "DMD-based implementation of patterned optical filter arrays for compressive spectral imaging," *J. Opt. Soc. Am. B* **32**(1), 80–89 (2015).
29. Y. Wu, I. O. Mirza, G. R. Arce, and D. W. Prather, "Development of a digital-micromirror-device-based multishot snapshot spectral imaging system," *Opt. Lett.* **36**(14), 2692–2694 (2011).
30. X. Lin, G. Wetzstein, Y. Liu, and Q. Dai, "Dual-coded compressive hyperspectral imaging," *Opt. Lett.* **39**(7), 2044–2047 (2014).

1. Introduction

Imaging spectroscopy has long been used as an important tool of analysis in many applications, such as environmental monitoring, geological exploration, biological chemistry, and so on [1–3]. Although its performance and resolution have been improved over the past decades, spectral imagers are challenging in some applications [4], especially in imaging microfluidics or combustion reactions where the objects are fast moving or changing in color. Traditional spectral imaging techniques like push-broom and Fourier imaging spectrometers are too slow to track the moving specimen or to detect the changes in spectrum due to their scanning mechanism [3]. Also, they cannot locate the reflectance data for 3D objects. The pixel-wise geometrical structure and spectral information can be useful to create 3D models with the inclusion of multispectral information, like underwater 3D visualization [5] or skin cancer detection [6].

The coded aperture snapshot spectral imager (CASSI) system, based on the principles of compressive sensing (CS), is a remarkable imaging architecture that allows capturing a spectral cube with just a single shot 2D measurement [3, 7, 8]. Advances in this area are growing rapidly and it facilitates flexible capture modes for different applications. For instance, a 3D imaging spectroscopy system for measuring hyperspectral patterns on solid objects was proposed in [1]. There, the authors integrated a CASSI system with a 3D range scanning system covering from the UV to NIR spectrum, in order to measure the radiometric characteristics of the entire surface of a 3D object. However, the laser radar scanner limits the applications of dynamic scenes due to the intrinsic slow scanning mechanism. Other multidimensional detection system using CS has also been developed [9, 10]. In this paper, a fast acquisition system based on a modification of CASSI is presented to obtain multidimensional spectral data.

Integral imaging is a passive multi-perspective imaging technique [11–13], which records multiple two-dimensional images of a scene from different perspectives. The concept is based on capturing many perspectives of a 3D scene by means of a microlens array. Then, 3D reconstruction can be achieved computationally by simulating the optical back projection of the multiple 2D images. The individual 2D images are usually called elemental images (EIs). In [14], a feasibility generalized framework for compressive multi-dimensional integral imaging was

studied. A pixel-wise color filtering element was simulated for 3D scenes in three spectral channels. However, the reconstruction quality is reduced as the number of spectral bands increases. Motivated by 3D integral imaging [11], the proposed system introduces a microlens array (MLA) in the CASSI system. This paper shows that the combination of integral imaging with CASSI can realize spectral imaging for 3D scenes using a dispersive element.

As a first attempt, a simple joint reconstruction procedure was explored in [15] to obtain spectral information of 3D scenes. In the previous system, elemental images are obtained by a pinhole camera array. The reconstruction process is a combination of spectral reconstruction and spatial reconstruction. In this paper, a prototype of the compressive 3D integral imaging spectrometer is presented. A different imaging sensing model is developed so as to accurately characterize the imaging phenomena and to develop an inverse algorithm. This imager can compress spatial and spectral dimensions of a dynamic 3D scene onto a monochrome detector. The obtained stack of images are arranged in plane-by-plane depth slices, like a focal stack [16, 17] which contains clear or blurred contributions by features in on, or off, the planes where they are focused, where each spatial pixel location contains its spectral response.

2. Optical sensing model

The proposed integral imager within the CASSI architecture is shown in Fig. 1. It consists of a microlens array, a coded aperture, a relay lens, a prism and a detector plane. The spatial information of the 3D scene is represented by a focal stack, a sequence of images each focused on a different plane with spectral information. Each slice in a focal stack is defined as a depth slice. Sensing of the 3D scene is carried out by obtaining elemental images optically through the microlens array to capture different perspectives of the scene. Then, the spectral information of elemental images is projected by the CASSI architecture onto a 2D compressive coded image on the detector plane.

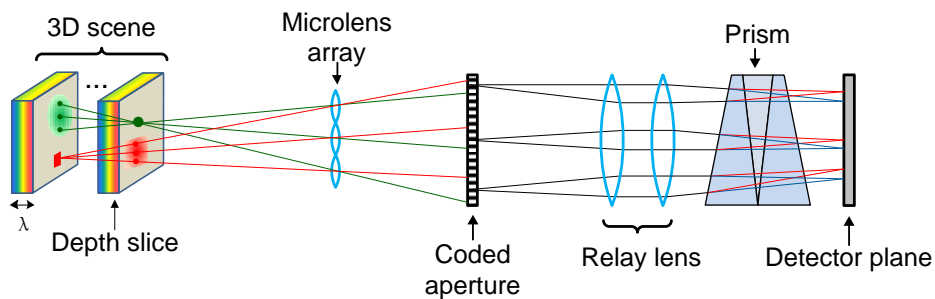


Fig. 1. Scheme of the 3D compressive spectral integral imaging system. It consists of a microlens array, a coded aperture, a relay lens, a prism and a detector plane. The focal stack of 3D spectral images is captured by the MLA and the CASSI system.

In the proposed method, integral imaging in tandem with CASSI realize a depth-variant and spectral impulse response. The imaging process of a single depth data cube for a single measurement shot is depicted in Fig. 2. There, the optical elements are represented by their effect on the discretized data cube. N_x' and N_y' are the numbers of elements of the scene along the x-axis and y-axis, respectively. N_x and N_y are the numbers of elements of the coded aperture, respectively. L is the number of spectral channels. $V = N_x \cdot (N_y + L - 1)$. For simplicity, it only depicts the projections by one row of micro elemental optics for 5 spectral bands. First, a depth slice from a 3D scene is projected with the parallax translation by each microlens. Each

projected elemental image with its own perspective is zoomed out on the coded aperture plane according to the elemental optics. Secondly, the projected elemental image data cube is coded in amplitude by the coded aperture. Third, a dispersive element plays a role of a spatial shifting of each spectral band. Finally, the coded and dispersed information of each EI is integrated along the spectral axis onto the focal plane array (FPA) detector.

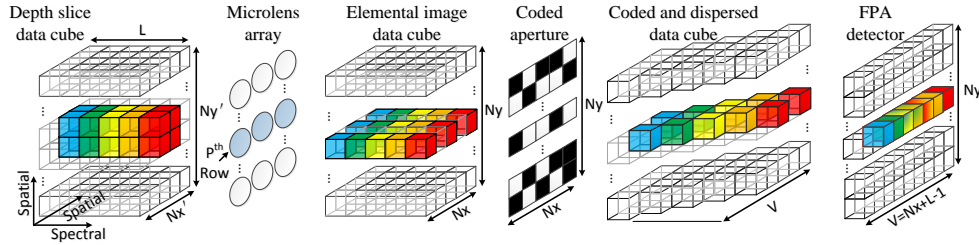


Fig. 2. Illustration of the spatio-spectral optical flow in the proposed scheme. The optical elements are represented by their effect on the discretized data cube. A region in a depth slice is projected with the parallax translation by each microlens. The elemental image array is coded by a row of the coded aperture and dispersed by the prism. The detector integrates the intensity of the coded and dispersed field.

Mathematically, the spatio-spectral density source entering into the system is represented as $f_0(x, y, z, \lambda)$ where (x, y, z) are the spatial coordinates and λ is the wavelength. Suppose that there are p rows and q columns of elemental images projected by the MLA on the coded aperture plane. Each elemental image has the same size. The imaging process of a microlens array in computational integral imaging can be described as follows [14]

$$f_k(x, y, \lambda) = \int f_0(x - \Gamma_k(z'), y - \Gamma_k(z'), z', \lambda) dz', \quad \text{for } k = 1, 2, \dots, p \cdot q \quad (1)$$

where $f_k(x, y, \lambda)$ is the projection by the k^{th} microlens optics, and $\Gamma_k(z)$ represents a translation by the parallax in the k^{th} microlens optics, respectively. Equation (1) can be written in matrix form as

$$\mathbf{f}_k = \mathbf{T}_k \mathbf{f}_0 \quad (2)$$

where $\mathbf{f}_k \in \mathbb{R}^{N_e \cdot N_e \cdot L}$ is the vector form of $f_k(x, y, \lambda)$, $\mathbf{T}_k \in \mathbb{R}^{(N_e \cdot N_e \cdot L) \times (N_x' \cdot N_y' \cdot N_z \cdot L)}$ is a translation matrix indicating the parallax translation produced by the k^{th} microlens optics, and $\mathbf{f}_0 \in \mathbb{R}^{N_x' \cdot N_y' \cdot N_z \cdot L}$ is the vector form of $f_0(x, y, z, \lambda)$, respectively. $\mathbb{R}^{a \times b}$ is real valued $a \times b$ matrix. $N_e \cdot N_e$ is the size of one elemental image in the k^{th} microlens optics. N_z is the number of depth slices.

There are p rows and q columns of elemental images through the coded aperture. Thus, the entire signal before entering the CASSI can be written as

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_k \end{bmatrix} = \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \\ \vdots \\ \mathbf{T}_k \end{bmatrix} \mathbf{f}_0 = \mathbf{T} \mathbf{f}_0, \quad \text{for } k = 1, 2, \dots, p \cdot q \quad (3)$$

where $\mathbf{f} \in \mathbb{R}^{N_x \cdot N_y \cdot L}$ is the vector form of elemental images captured by the entire microlens optics, and $\mathbf{T} \in \mathbb{R}^{(N_x \cdot N_y \cdot L) \times (N_x' \cdot N_y' \cdot N_z \cdot L)}$ is the entire translation matrix. Here, $N_x = N_e \cdot p$ and $N_y = N_e \cdot q$. Due to the zoom effect by the microlens optics, the depth slice in the object space is zoomed out onto the image on the coded aperture plane. That is, $N_x' > N_x$ and $N_y' > N_y$.

\mathbf{T} is a shifted matrix which describes the parallax translation and the zoom effect by the microlens optics in the range of depth. \mathbf{T} also accounts for the compression in spatial domain. In this paper, \mathbf{T} is constructed using computational integral imaging method [13]. Figure 3 shows an example of the matrix \mathbf{T} for $N_x = 8$, $N_y = 8$, $N_z = 2$ and $L = 2$, where $N_e = 4$, $p = 2$, $q = 2$, $N_x' = 16$, and $N_y' = 16$. The white squares show the locations of the '1's in \mathbf{T} , representing the corresponding sensing procedure from the depth slice data cube to the elemental image data cubes. The ratio of the distance between the microlens array and the depth slice to the focal length of the microlens array is denoted as magnification factor. The magnification factors of the two depth slices are set to be 2 and 3, respectively. That is, one element on the focal plane of the microlens array will be mapped to 2×2 and 3×3 elements on the corresponding depth slices.

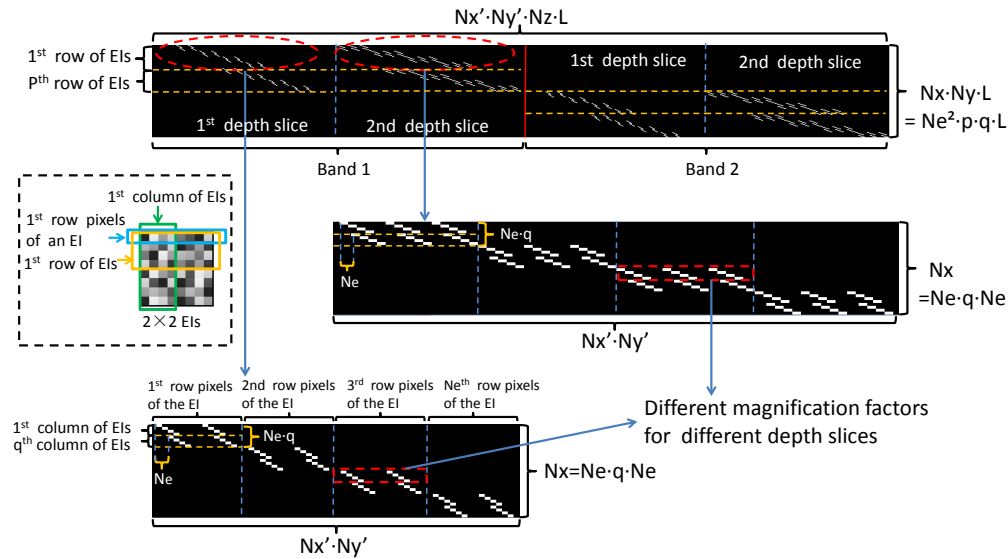


Fig. 3. Illustrative example of the translation matrix \mathbf{T} for $N_x = 8$, $N_y = 8$, $N_z = 2$ and $L = 2$. The structure of the matrix accounts for the parallax translation and the zoom out effect by the microlens optics. It depicts the sensing procedure of 2×2 EIs from the depth slice data cube to the elemental image data cubes. The white squares show the locations of the '1's in \mathbf{T} . Due to the zoom out effect by the microlens, the magnification ratio of the two depth slices to the focal plane of the microlens array are 2 and 3, respectively.

In the single shot matrix model for CASSI [3, 7, 18], the FPA measurement is a linear combination of the coded and shifted spectral image planes. The corresponding forward model is given by

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \omega, \quad (4)$$

where $\mathbf{g} \in \mathbb{R}^{V \times 1}$ is the vector of the captured data on the detector plane, $\mathbf{H} \in \mathbb{R}^{V \times (N_x \cdot N_y \cdot L)}$ accounts for the coded apertures and the dispersive element effects, and ω is the measurement noise, respectively. Figure 4 shows an example of the matrix \mathbf{H} for $N_x = 8$, $N_y = 8$ and $L = 2$.

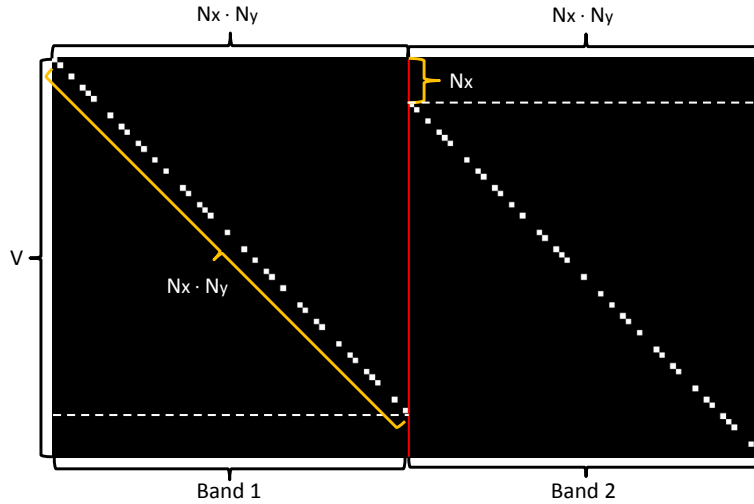


Fig. 4. Illustrative example of the matrix \mathbf{H} for $N_x = 8$, $N_y = 8$ and $L = 2$. The structure of the matrix accounts for the effects of the coded aperture and the dispersion by off setting the diagonal structure from band to band. White squares represent the '1's by unlock elements of the coded aperture. Here, $V = N_x \cdot (N_y + L - 1)$.

By replacing Eq. (3) in (4), the compressive coded measurement of the proposed system can be written as

$$\mathbf{g} = \mathbf{HT}\mathbf{f}_0 + \omega = \mathbf{A}\mathbf{f}_0 + \omega, \quad (5)$$

where $\mathbf{A} = \mathbf{HT}$ represents the sensing matrix of the proposed system. To solve this under-determined linear system, the Two-step iterative shrinkage/thresholding (TwIST) [19] reconstruction algorithm is used although a number of other method can be used [20, 21]. The signal recovery is obtained \mathbf{f}_0 as the solution of

$$\underset{\mathbf{f}_0}{\operatorname{argmin}} \|\mathbf{g} - \mathbf{A}\mathbf{f}_0\|_{l_2} + \tau R(\mathbf{f}_0), \quad (6)$$

where $\|\cdot\|_{l_2}$ denotes the l_2 norm, τ is a regularization parameter, and R represents a total variation (TV) regularizer given by

$$R(\mathbf{f}_0) = \sum_m \sum_n \sum_{i,j} [(f_{(i+1)jnm} - f_{ijnm})^2 + (f_{i(j+1)nm} - f_{ijnm})^2]^{1/2}, \quad (7)$$

where f_{ijnm} is a rearranged element of \mathbf{f}_0 , representing one element of the original multidimensional spatio-spectral data $f_0(x, y, z, \lambda)$ in discrete form.

3. Simulations and experimental results

In this section, the proposed method will be firstly verified by simulations of two data sets. In addition, a prototype has been constructed in the laboratory to test the proposed system. To improve the contrast of the 3D scenes, only the reconstructed depth slices where the objects are focused are shown. A random pattern with transmittance of 50% was used for the coded aperture in all the simulations and experiments. The transmittance of 50% guarantees a good trade-off between the compression ratio per pixel and the light throughput of the system. Code aperture optimization for the proposed imaging device is also an important tool that can improve the imaging result but it is left for future research [22].

3.1. Simulations

In order to evaluate the proposed method, a test data cube of size of $N_x' \cdot N_y' \cdot N_z \cdot L = 256 \times 256 \times 2 \times 8$, as shown in Fig. 5, is first used to simulate the system model. The eight spectral channels are 500, 510, 530, 550, 580, 600, 620, and 640 nm, respectively. The data is divided into two, as shown in Fig. 5(a). Fig. 5(b) shows the simulated 3D scene consisting of two depth slices. The image in the left image is assumed to be closer to the system than the image in the right. Then, a distance z between the two image planes is set so that the left image is zoomed out twice on the coded aperture plane and the right one is zoomed out triple by the microlens optics.

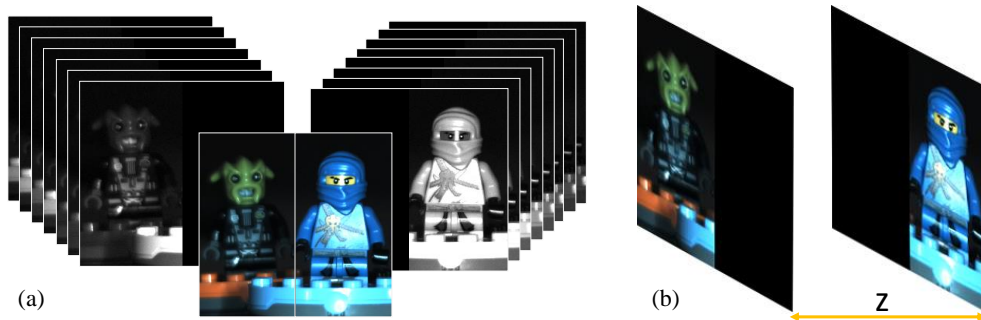


Fig. 5. Test data with the size of $256 \times 256 \times 2 \times 8(N_x' \times N_y' \times N_z \times L)$. (a) Original data which is divided into two parts. (b) Simulated 3D scene with a separation z .

With the simulated 3D scene, 4×4 elemental images with 8 spectral channels are obtained using Eq. (2), as shown in the first two rows of Fig. 6. Each elemental image is zoomed and parallax translated by simulating the microlens array. Note that, more elemental images will improve the reconstruction quality but increase the computational requirement [11]. The other four rows of the figure show two reconstructed depth slices with 8 spectral channels using the TwIST algorithm. It can be seen that the two objects are recovered clearly at their original locations like two focal surfaces. The compression ratio is defined as the size of the estimated data over the size of the measurement, which is $(N_x' \cdot N_y' \cdot N_z \cdot L) / [(N_x + L - 1) \cdot N_y]$. The size of the reconstructed data is $256 \times 256 \times 2 \times 8$. The size of the captured image is 256×263 . Thus, the compression ratio is about 15.6. The averaged peak signal-to-noise ratio (PSNR) of the reconstruction is calculated by comparing the reconstructed images with the original data along the spectral channels. The averaged PSNR of the central two rows is 29.6 dB and the bottom two rows is 26.8 dB. Note that, the averaged PSNR at the longer distance is reduced by 3 dB. The reason for getting a lower PSNR is due to the higher compression used to sense elemental images at longer distances, since the EIs will shrink more when the objects are placed at longer distances. However, the reconstructed depth is limited. The depth of focus of the reconstructed 3D scene is proportional to the number of pixels of the EIs and the focal length of the MLA [23].

To evaluate the spectral reconstruction accuracy, two points on the objects are used for comparison, as shown in Fig. 7. The first column and the third column are the comparison of RGB images between the reconstructed result and the original data. The full spectral bands are used for simulating RGB vision. From the second column, it can be found that the reconstructed spectral curves of the two points match very well with the original ones, which suggests that the proposed method is able to correctly measure the spectral information of 3D scenes. The intensities of spectral curves were normalized to the scale of 0 to 1.

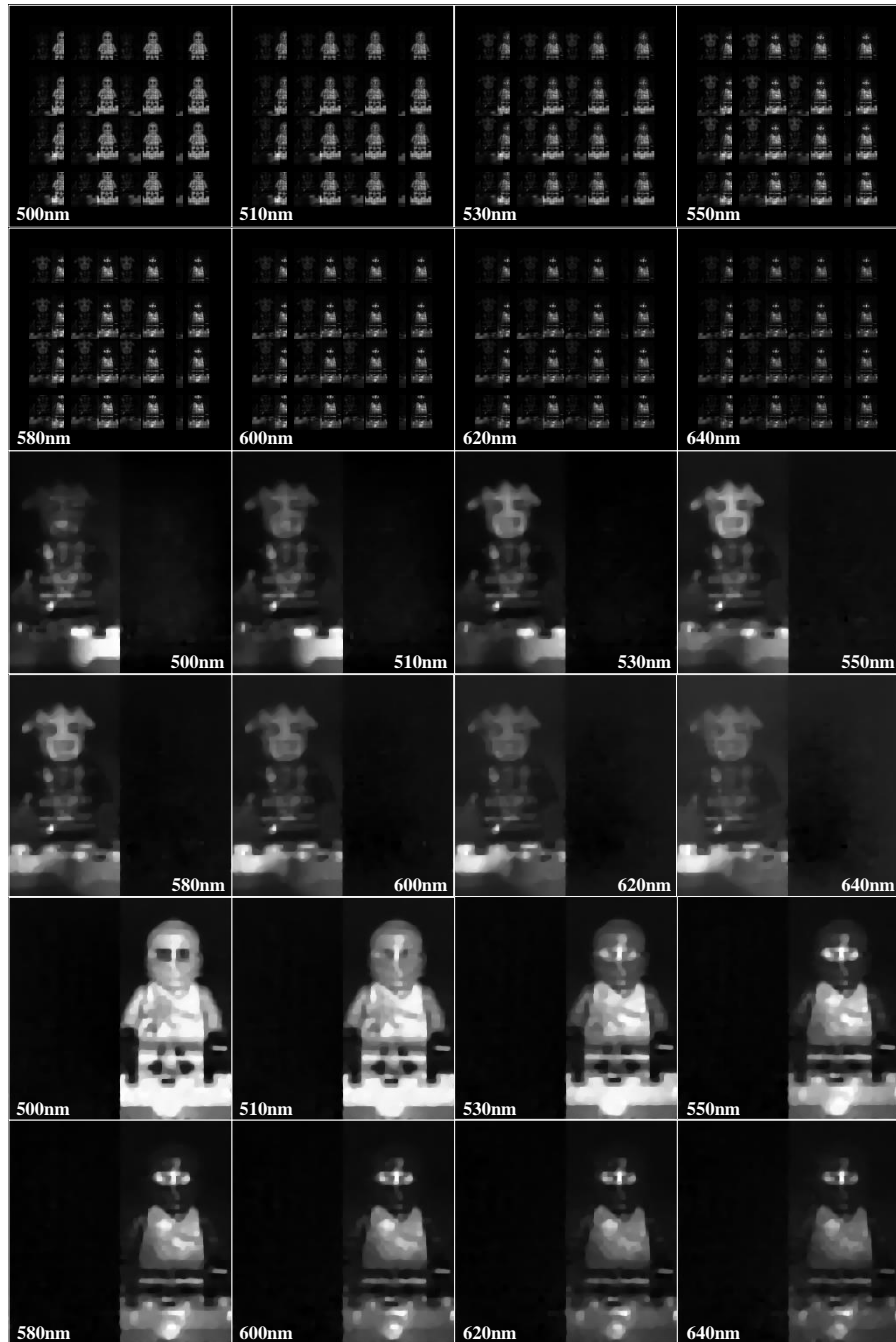


Fig. 6. Simulated results with 8 spectral channels. The first two rows show the elemental image array formed on the coded aperture plane by the simulated 4×4 microlens array. The central two rows of the figure show the reconstructed result of the object in front. The bottom two rows show the reconstructed result of the second object. The eight spectral channels are 500, 510, 530, 550, 580, 600, 620, and 640 nm, respectively. The averaged PSNR of the central two rows is 29.6 dB and the bottom two rows is 26.8 dB.

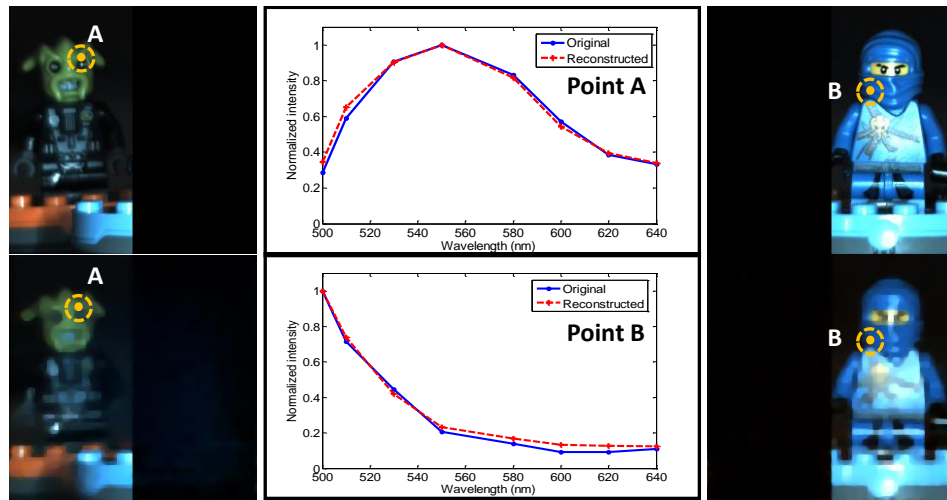


Fig. 7. Spectral intensity of two pixels located at two objects. In the left and right of the figure, the top and bottom RGB images are the comparison between the original data and the reconstruction. The full spectral bands are used for RGB display. Two points A and B on the objects were selected for spectral comparison. The central column shows the spectral curves of the original and reconstructed pixels.

The second set of simulated data is captured by a shifted pinhole CCD camera. It is a scenario close to a real experiment. In Fig. 8(a), two objects separated by a distance of 70 mm are used to form the 3D scene. The distance between the pinhole and the CCD camera is 50 mm. The first object is located at a distance of $z = 210$ mm from the pinhole. So the magnification factors are 4.2 and 5.6 for the object in front and the one at the back, respectively. Figure 8(b) shows 6×10 elemental images which were recorded by shifting the pinhole CCD camera with equal separations of 2.5 mm in both horizontal and vertical directions. Each captured elemental image has 32×32 pixels with 11 spectral bands ranging from 450nm to 650nm. The captured elemental images are used to verify the feasibility of the proposed optical system.

Figure 9 shows the reconstructed results. The top row shows two depth slices at the distance of $z = 210$ mm where the first object is located and of $z = 280$ mm where the second one is located. It can be noticed that different objects at different distances were reconstructed quite well, which like a focal stack, the sequence of images each focused on a different plane. That is, the spatial and depth information of the 3D scene is obtained. The size of the reconstructed data is $300 \times 450 \times 2 \times 11$. The size of the captured image is 192×330 . Thus, the compression ratio is about 46.9. Also, to evaluate the spectral fidelity of the reconstructions, the intensity variations of the 11 spectral channels were measured at two pixel locations, including point A on the first object and point B on the second object, as indicated in Fig. 8(a). A commercial spectrometer (Ocean Optics, USB2000+) was used to generate reference spectral curves at those locations. From the bottom row of Fig. 9, it can be seen that the spectral curves of the two points match well with the original ones. The curves here have larger errors than those in Fig. 7 because the EIs here are obtained by real data with optical aberrations instead of simulated data.

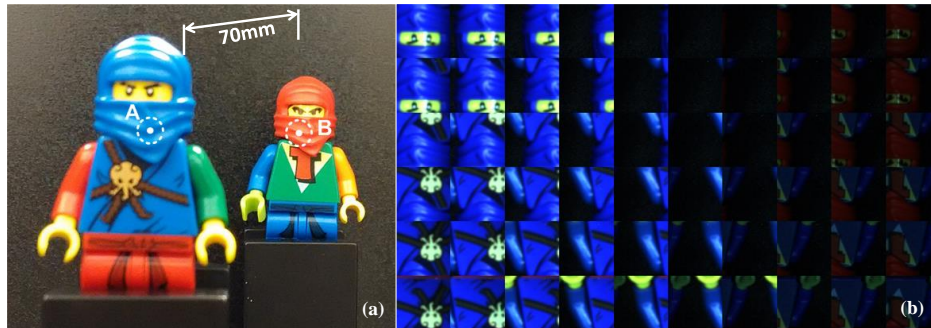


Fig. 8. Captured 3D scene by a shifted pinhole camera. (a) Simulated 3D scene with two toys. The distance between the two objects is 70 mm. Two points are selected for spectral comparison. (b) 6×10 elemental images with different perspectives. Each elemental image is upside down due the pinhole optics.

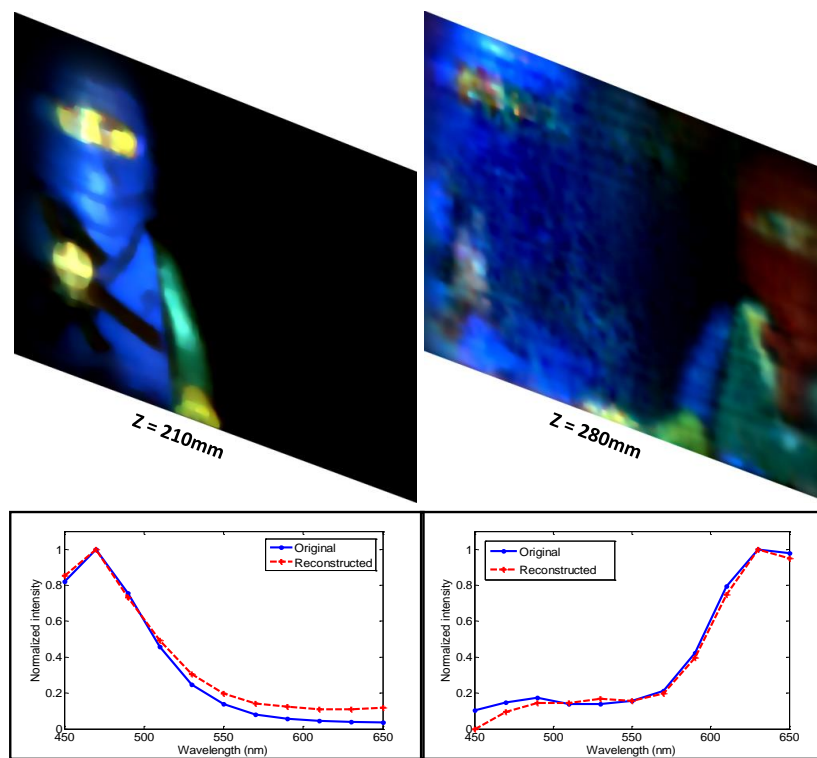


Fig. 9. Reconstructed results using the elemental images array in Fig. 8(b). The top row shows a focal stack consisting of two focal plane where the objects are located (one at $z = 210\text{ mm}$ and the other at $z = 280\text{ mm}$). The full visible spectral channel are used for RGB display. The bottom row shows the reconstructed spectral curves of the two points compared to the original ones.

3.2. Experimental results

For the experiments, a prototype was constructed in the laboratory to test the proposed concept. Figure 10 shows the experimental setup. A custom designed microlens array is placed in front of a binary litographic coded aperture to form the elemental image array. This MLA(Thorlabs Inc.) has 20×20 square refractive lenses in a 10 mm^2 area with a focal length of 13.8 mm. The coded aperture(Photo Sciences Inc.) is a 256×256 random binary pattern with the smallest code feature of 19.8 um (2×2 CCD pixels). In this setup, the elemental images are formed on the coded aperture plane. Afterwards, the elemental images go through the CASSI system. A pair of visible achromatic lenses was used as the relay lens, which relays the image from the plane of the coded aperture to the CCD. A double Amici prism(Shanghai Optics Inc.) made of glasses SF4 and SK2 was installed in between the exit aperture of the relay lens and the monochromatic CCD camera(AVT Marlin, Allied Vision Technologies). Therefore, the coded and dispersed elemental image array of the 3D scene is mapped onto the detector plane. System calibration is a critical process to ensure the accuracy of the reconstruction. The calibration of the CASSI system is described in [19]. Before calibrating the spectral response of the proposed system, make sure that the MLA is placed in front of the coded aperture.

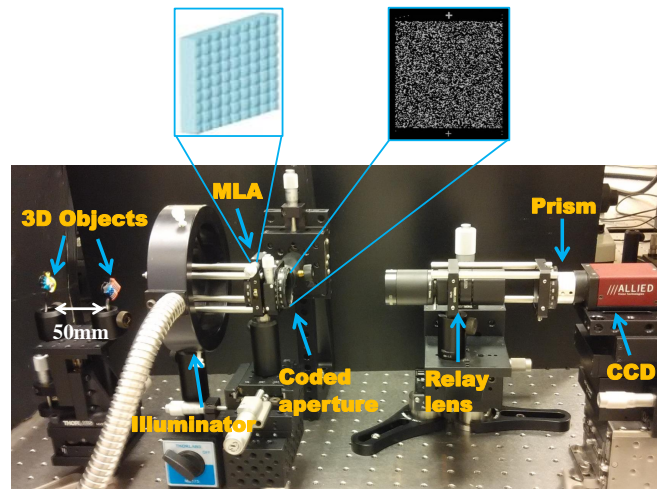


Fig. 10. Prototype of the proposed system. The 3D scene is illuminated by an incoherent broadband lightsource. The microlens array is placed in front of the CASSI system to form the elemental images. The zoomed pictures show the patterns of the microlens array and the coded aperture.

In the experiment, only one snapshot was used for reconstruction as well. Figure 11 shows the experimental result using the prototype. In Fig. 11(a), the scene contains two cards on which two letters “U” and “D” are printed. The objects in the scene are illuminated by an incoherent broadband lightsource. The distance between the first object and the microlens array is 150 mm, and the distance between the second object and the microlens array is 200 mm. The dotted region is the imaging scene of the system due to the FOV of the MLA. Two points are chosen for spectral analysis afterwards. Figure 11(b) shows a compressive measurement of 4 columns and 8 rows of elemental images captured by the CCD camera. Each elemental image has 30×30 pixels. It is spatially coded by the coded aperture and spectrally dispersed by the prism. Figure 11(c) and Figure 11(e) show the two reconstructed depth slices using the TwIST algorithm. Figure 11(d) and Figure 11(f) show the two reconstructed depth slices using the Gradient Projection for Sparse Reconstruction(GPSR) algorithm [24]. The GPSR is also an algorithm used for

spectral image estimation. The spectral data cubes contain 13 spectral channels separated by the dispersive prism. They are 450, 460, 470, 480, 492, 505, 515, 532, 550, 568, 589, 620 and 650 nm, respectively. The full visible spectral channel are used for approximating the RGB vision system. It can be seen that the reconstruction result is quite good. The outline of the card “U” is clear at the distance of $z = 150$ mm while the letter “D” at the distance of $z = 200$ mm is also clear. Thus, the reconstructed plane in Fig. 11(c) focuses on the card “U” and the reconstructed one in Fig. 11(e) focuses on the card “D”. The size of the reconstructed data is $300 \times 400 \times 2 \times 13$. The size of the captured image is 120×252 . Thus, the compression ratio is about 103. Compared the two algorithms, it can be found that the edge of reconstructed objects using GPSR algorithm is clearer than that using TwIST algorithm. With the choice of the total variation regularizer, the TsIST algorithm yields spatially smoother estimates during the iterative reconstruction process. However, TwIST iterations are much faster than GPSR iterations. In addition, the spatial resolution of the reconstructed 3D scene is determined by various factors, such as the size of the microlens array, the resolution of the elemental images, and the magnification factor [23, 25].

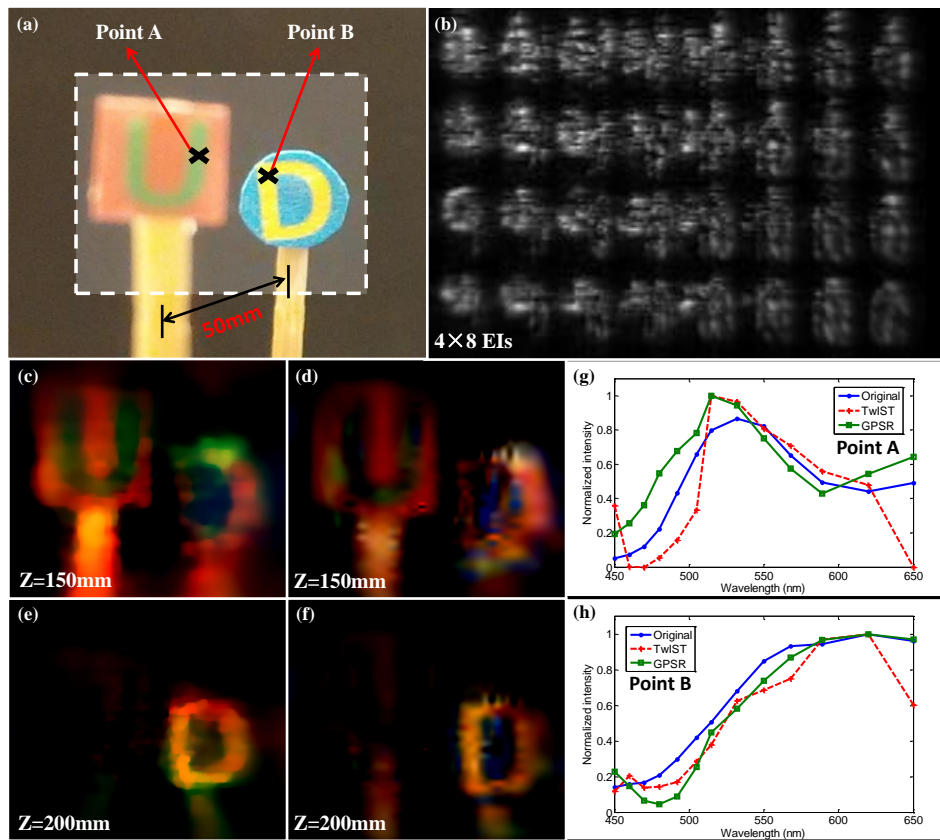


Fig. 11. Experimental result with the captured scene. (a) A 3D scene consisting of two letters “U” and “D”. The distance is 50 mm. (b) A compressive measurement of 4 columns and 8 rows of elemental images captured by the CCD camera. (c) and (e) show the reconstructed depth slices using the TwIST algorithm. (d) and (f) show the reconstructed depth slices using the GPSR algorithm. (g) and (h) show the spectral comparison of the two points in Fig. 11(a).

The experimental and reference spectral curves of the reconstructed 3D scene at two pixel locations in Fig. 11(a) are compared in Fig. 11(g) and Fig. 11(h). It can be seen that the reconstructed spectral data using one snapshot provides a proper spectral accuracy. The results using GPSR algorithm are better than that using TwIST algorithm. The reconstructed spectral profiles are matched but there are some offsets. This is due to the difference between the simulations and the real experiment with MLA which provokes aberrations. In addition, the size of an EI projected by a microlens is very small, which will also effect the reconstructed quality. Note that, only one measurement was used in the experiment.

4. Conclusion

A new compressive spectral integral snapshot imaging system for measuring 3D objects was proposed based on the ideas of compressive spectral imaging and computational integral imaging. It can simultaneously sense a 3D scene at different depths and obtain the spectral responses of the spatial pixel locations. 3D objects with different perspectives are optically recorded by a microlens array as elemental images. Then through a CASSI system, the spatio-spectral data is captured in a single compressive 2D measurement image on the detector plane. The optical model and the imaging process of the proposed system were described. The simulations and experiment were conducted to test and verify the performance of the proposed setup. The results show that 3D spatial information and the spectral intensities can be successfully reconstructed as depth slices. The proposed imager provides the flexibility for improving the spatial reconstruction quality by simply applying a coded aperture with higher resolution which can contain more EIs. Future work will focus on performance improvement of the proposed system by designing better coded apertures [22] or using color coded apertures [26–28]. The reconstructed quality can also be enhanced by high-resolution compressive spectral imagers like a digital-micromirror-device (DMD)-based multishot snapshot spectral imaging (DMD-SSI) system [29] or a dual coded hyperspectral imaging (DCSI) [30].

Funding

China Scholarship Council (CSC) (201406840021); National Science Foundation (NSF) (Intel/NSF VEC 1538950); Jiangsu College graduate Research and Innovation Program (KYZZ0112); Key Laboratory of low-light-level night vision technology (J20130501).

Acknowledgments

Weiyi Feng acknowledges the support from the University of Delaware. Hoover Rueda is supported by a Colciencias Fulbright Scholarship.