Dual-coded compressive hyperspectral imaging

Xing Lin,[1,2] Gordon Wetzstein,[3] Yebin Liu,[1,2] and Qionghai Dai[1,2,*]

[1]*Department of Automation, Tsinghua University, Beijing 100084, China*

[2]*Beijing Key Laboratory of Multi-dimension & Multi-scale Computational Photography (MMCP), Tsinghua University, Beijing 100084, China*

[3]*Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

**Corresponding author: qionghaidai@mail.tsinghua.edu.cn*

Received January 16, 2014; accepted February 20, 2014;
posted February 26, 2014 (Doc. ID 204679); published March 26, 2014

This Letter presents a new snapshot approach to hyperspectral imaging via dual-optical coding and compressive computational reconstruction. We demonstrate that two high-speed spatial light modulators, located conjugate to the image and spectral plane, respectively, can code the hyperspectral datacube into a single sensor image such that the high-resolution signal can be recovered in postprocessing. We show various applications by designing different optical modulation functions, including programmable spatially varying color filtering, multiplexed hyperspectral imaging, and high-resolution compressive hyperspectral imaging. © 2014 Optical Society of America

OCIS codes: (120.6200) Spectrometers and spectroscopic instrumentation; (110.1758) Computational imaging; (110.4234) Multispectral and hyperspectral imaging.
http://dx.doi.org/10.1364/OL.39.002044

Hyperspectral (HS) imaging is concerned with capturing a 3D datacube with 2D spatial and 1D spectral variation. This kind of data has important applications in a wide range of fields, including remote sensing, scientific imaging, surveillance, and spectroscopy. The most widely used approaches for spectral imaging use mechanical or temporal scanning techniques that record one or a few data points at a time. Compared with scanning methods, snapshot approaches capture the full 3D datacube in a single image, which is a distinct advantage for capturing dynamic scenes or aerial imaging. Snapshot approaches can be implemented by multiplexing a high-dimensional signal onto a 2D sensor, thereby sacrificing image resolution. Examples of that approach include the 4D imaging spectrometer (4DIS) [1], the snapshot image mapping spectrometer (IMS) [2], and also computed tomography imaging spectrometer (CTIS) [3].

Recently, a coded aperture snapshot spectral imager (CASSI) [4,5] was proposed that applies compressive computational reconstructions to encode optical signals. This computational imaging approach was demonstrated to overcome previous tradeoffs between spatial and spectral image resolution. The CASSI system can be improved by acquiring multiple shots that are recorded from a coded mask that shifts on a piezostage [6]. A more flexible alternative is the digital micromirror device-based (DMD) multishot spectral imaging system (DMD-SSI) proposed in [7]. The trend toward computational imaging systems that optically code recorded data and recover it via compressive computation is obvious [8]. However, all approaches to compressive spectral imaging *code the color spectrum in a spatially uniform* manner, which places a fundamental limit on the quality that can be expected from compressive sparsity constrained compressive reconstruction algorithms.

In this Letter, we present a new snapshot approach to compressive hyperspectral imaging that we call dual-coded hyperspectral imaging (DCSI). The proposed approach separately codes both spatial and spectral dimensions within a single exposure, achieving an independent spectral code for each sensor pixel. By combining spectral and spatial modulation, DCSI facilitates extremely flexible capture modes customized for different applications. For example, DCSI can be used to implement programmable spatially varying color filtering (such as Bayer filtering, Fig. 2) or multiplexed hyperspectral imaging (Fig. 3). More importantly, from a single shot, DCSI allows for compressive HS imaging with high-quality and high-resolution reconstructions for both the spatial and spectral dimensions. Compared with CASSI and DMD-SSI approaches, DCSI provides a higher degree of randomness in the measured projections, which is beneficial for compressive reconstruction algorithms, as described by the restricted isometry property (RIP) and the mutual incoherence property (MIP) [9,10]. Basically, a more flexible amount of signal mixing is achieved when a high-dimensional signal is optically projected to a lower-dimensional space. We present HS reconstruction results with physical experiments and evaluate the proposed techniques in simulation.

Figure 1 shows a schematic of the proposed optical system and a photograph of the prototype device. DCSI is presented as a two-arm system, including a spatial modulation arm and a spectral modulation arm. In the spatial modulation arm, an objective lens (focal length 150 mm, diameter 50.8 mm) is used to optically form images onto a DMD. A DMD is employed as a high-resolution spatial light modulator which can provide high contrast and light-efficient binary modulation for pixel-wise coded exposures. The DMD used in our setup is a Texas Instruments (TI) 4100, with a resolution of 1920×1080 micromirrors and a pixel pitch of 10.8 μm. In the spectral modulation arm, a diffraction grating is applied to disperse the spatially modulated light into its spectrum [11], and a liquid crystal on silicon (LCOS) display is added as the spectral modulator. An off-the-shelf transmission diffraction grating (Thorlabs GT50-06V with 600 grooves/mm and 28.7 deg blaze angle) is used and the LCOS is taken from a Newsmy PHO5C projector with a resolution of 1024×768 pixels. A bandpass filter with transmission window 400–820 nm is used to filter out unwanted spectral bands from the system. The sensor is a PointGray GRAS-50S5M-C grayscale camera with resolution 2448×2048 and pixel pitch 3.45 μm. In our

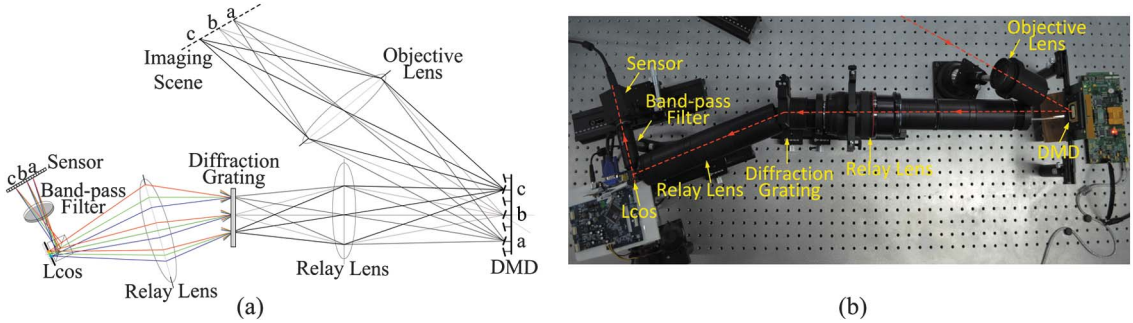

Fig. 1. (a) Schematic of the DCSI system and (b) a photograph showing the prototype system.

experiments, each signal pixel is represented by a 4×4 window of a DMD micromirrors subset; hence the maximum modulation resolution is 480×270, and each DMD pixel subset is imaged onto an area of the sensor with a size of 5×5 pixels. During the exposure time of a single image, we dynamically modulated the spectrum over a range of bands while selectively coding the exposure pattern of each pixel.

In DCSI, the coded sensor image $i(\mathbf{x})$ is formed by a spatio-spectral modulation $m(\mathbf{x}, \lambda)$ on the HS datacube $h(\mathbf{x}, \lambda)$, followed by a projection along the spectral dimension over the domain Ω_λ:

$$i(\mathbf{x}) = \int_{\Omega_\lambda} h(\mathbf{x}, \lambda) m(\mathbf{x}, \lambda) \mathrm{d}\lambda, \tag{1}$$

where $\mathbf{x} = \{x, y\}$ is the 2D spatial coordinate on the sensor and λ is the wavelength. Note that m is a modulation function imposed by the pixel states of both SLMs. The spectral sensitivity and other sensor-specific effects are ignored in this formulation, because, in preprocessing, they can be calibrated and corrected.

In practice, Eq. (1) is discretized as

$$\mathbf{i} = \Phi \boldsymbol{h}, \tag{2}$$

where $\mathbf{i} \in \mathbb{R}^M$ and $\boldsymbol{h} \in \mathbb{R}^N$ are the vectorized pixel measurements and the vectorized target HS datacube, respectively, and $\Phi \in \mathbb{R}^{M \times N}$ is the modulation matrix. By decoupling and synchronizing the spectral and spatial coding, DCSI facilitates various modulation functions that can be designed for different applications with the goal of achieving flexible HS imaging modalities.

Programmable Spatially-varying Color Filtering: DCSI provides various imaging modes. The most intuitive one achieves programmable spatially varying color filtering, for example, Bayer filtering where red, green, and blue samples of the datacube are captured in an interleaved grid (see Fig. 2). Assuming that response functions for the Bayer color channels are $f_0(\lambda), f_1(\lambda), f_2(\lambda)$, the modulation function can be formulated as $m(\mathbf{x}_k, \lambda) = f_k(\lambda)$ for pixels in channel $\mathbf{x}_k, k = 0, 1, 2$. This modulation is achieved by dividing the sensor into a grid of interleaved segments - one for each color channel. The spatial sampling pattern $\mathbf{x}_k$ is selected by the DMD and the corresponding spectral filter function f_k is imposed by the LCOS. Based on this intuitive, all-optical technique, we can either apply color-demosaicing algorithms to a multiplexed sensor image or capture the high-resolution signal in multiple shots.

Multiplexed HS imaging: Bayer-type multiplexing can be easily extended to include different spectral coding, as well as different spatial layouts. All of these approaches trade spatial resolution for spectral resolution if captured in a single exposure. As shown in Fig. 3, we can group different spectral codes into super-pixels, with each pixel inside forming a single channel, *e.g.*, $\mathbf{x}_k$ ($k = 0, 1, \ldots, 16$ in Fig. 3). In this example, we sequentially expose channel $\mathbf{x}_k$ with spectral filter f_k using a Hadamard multiplexed scheme to achieve a higher light efficiency. Assuming that all pixels in any super-pixels share the same spectral profile, a low spatial resolution HS datacube can be recovered by applying the inverse Hadamard matrix to each super-pixel.

Compressive HS imaging: By combining optical spatio-spectral modulation and the sparsity-constrained reconstruction algorithm, a high-resolution HS data cube

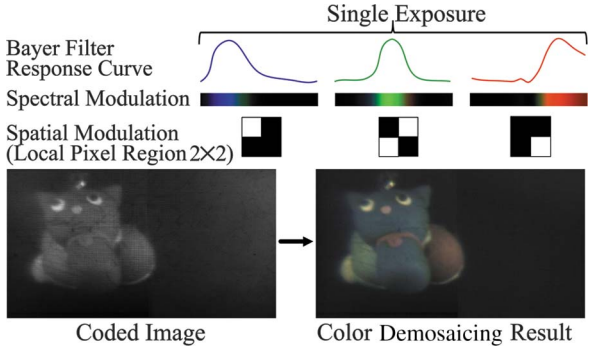

Fig. 2. Programmable spatially varying color filtering. A traditional Bayer color filter array, implemented with the proposed prototype, is shown in this example (Media 1).

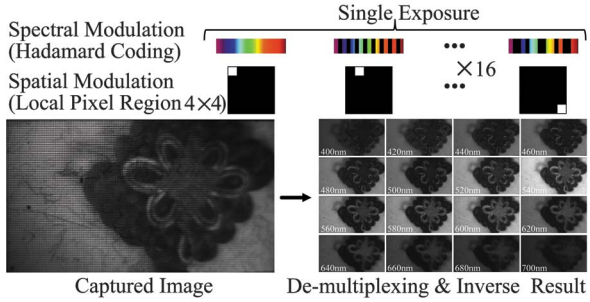

Fig. 3. Multiplexed hyperspectral imaging. In this example, an HS data cube with 16 spectral channels is reconstructed via demosaicing.

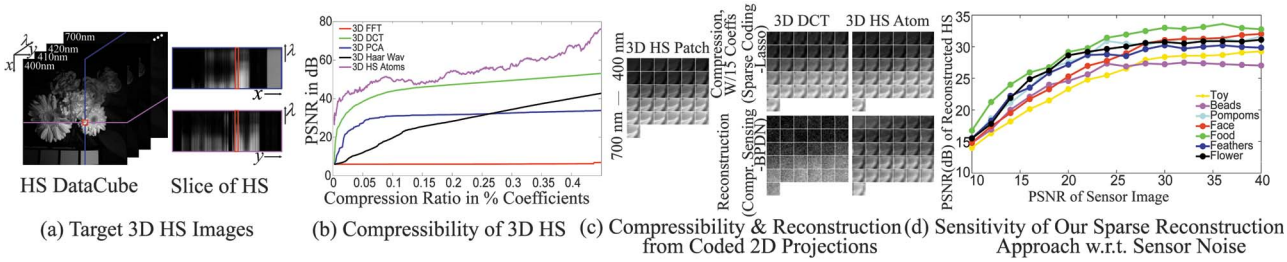

Fig. 4. (a), (b) We quantitatively evaluate compressibility by transforming the HS datacube into different bases as well as the proposed over-complete dictionary. (c) Comparison of 3D HS images reconstructed from a single-coded 2D projection. (d) Sensitivity of the proposed sparse reconstruction algorithm w.r.t. to sensor noise.

can be computed from a single DCSI snapshot measurement. For this purpose, we solve for $\boldsymbol{h}$ in the highly under-determined linear system [Eq. (2)] with sparsity constraint on the underlying signal.

As shown in Fig. 4(a), the individual slices in 3D HS datacube are highly redundant. Based on compressive sensing theory [12,13], $\boldsymbol{h}$ can be faithfully resolved if it has sparse representation α, for instance using a dictionary $\mathbf{D}$ as

$$\boldsymbol{h} = \mathbf{D}\boldsymbol{\alpha} = \mathbf{d}_1\alpha_1 + \cdots + \mathbf{d}_k\alpha_k, \tag{3}$$

where $\mathbf{d}_1, \ldots, \mathbf{d}_k \in \mathbb{R}^N$ are the atoms (columns) of the dictionary $\mathbf{D} \in \mathbb{R}^{N\times K}$, and $\alpha = [\alpha_1, \ldots, \alpha k]^{\mathbf{T}}$ are sparse coefficients. As shown in Fig. 4(b), when approximating a hyperspectral datacube with only a few coefficients, dictionary atoms provide better quantitative compression for this example than other basis representations. We also qualitatively compare the compressibility of a single, small 3D patch using the 3D discrete cosine transform (3D DCT) and the proposed atoms. Atoms provide more faithful reconstructions (Fig. 4(c) first row). In addition, we compare sparse reconstruction from a simulated snapshot coded projection with both DCT and the HS atoms. Again, 3D HS atoms can significantly improve the reconstruction quality (Fig. 4(c), second row).

The over-complete dictionary $\mathbf{D}$ is learned from a collection of training samples that are small 3D spatio-spectral patches, each with a resolution of $10 \times 10 \times 31$ pixels. These samples can be obtained by randomly choosing a predefined number of patches from public hyperspectral data sets (e.g., [14,15]). The dictionary learning is formulated as an optimization problem:

$$\min_{\{\mathbf{D},\mathcal{A}\}} \|\mathbf{T} - \mathbf{D}\mathcal{A}\|_2^2; \qquad s.t. \quad \forall\, i, \|\boldsymbol{\alpha}_i\|_0 \le k, \tag{4}$$

where $\mathbf{T} \in \mathbb{R}^{M\times O}$ is a training set composed of O small patches, and $\mathcal{A} \in \mathbb{R}^{K\times O}$ is a matrix containing the k-sparse vectors α_i in its columns. We use the K-SVD algorithm [16] to solve Eq. (4). An example of such a dictionary is visualized in Fig. 5. HS atoms represent the essential building blocks of natural HS images—most HS images can be represented by weighted sum of very few of these atoms.

With a dictionary-based sparse representation, the random modulated image can be formulated as

$$\mathbf{i} = \Phi\boldsymbol{h} = \Phi\mathbf{D}\boldsymbol{\alpha}. \tag{5}$$

We randomly modulate both the spectral and spatial dimensions during the exposure time, which provides a theoretical light efficiency of 25%. These codes are encoded as the projection matrix Φ. We apply the SPGL1 algorithm [17] for robust recovery of the sparse unknown vector $\boldsymbol{\alpha}$ and solve the basis pursuit denoise problem (BPDN) as

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \; s.t. \; \|\mathbf{i} - \Phi\mathbf{D}\boldsymbol{\alpha}\|_2^2 \le \varepsilon. \tag{6}$$

The optimizations are solved on all 3D patches independently. The recovered HS patches are merged to recover the desired 3D HS datacube.

Figure 6 shows the reconstruction of a resolution chart with the proposed dual-coded compressive HS imaging approach. In this example, a 31 waveband HS image with a resolution of 470×260 pixels is recovered by applying overlapping patches reconstruction. This example shows the resolution performance of our system. The utilization of random dual-coding decreases the captured photon count by a factor of four. Thus, we also demonstrate the sensitivity of sparse reconstructions w.r.t. to noise

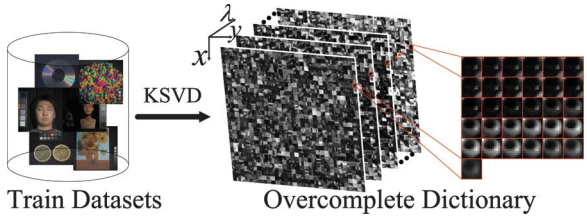

Fig. 5. Visualization of learned hyperspectral atoms in an over-completed dictionary.

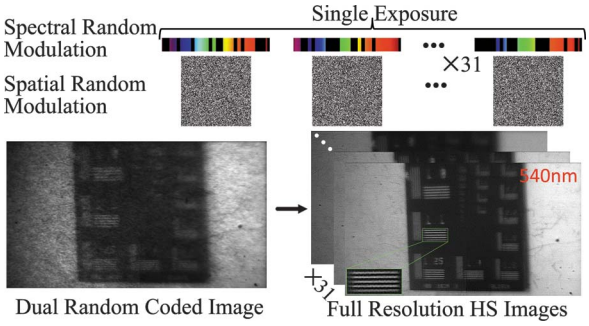

Fig. 6. Compressive hyperspectral imaging. An HS data cube containing a resolution chart with 31 spectral channels is reconstructed in this example.

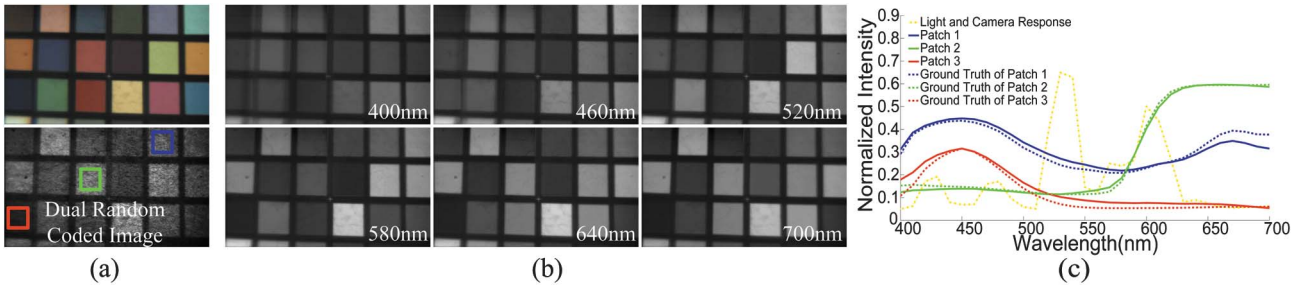

Fig. 7. Compressive HS imaging of a color checker scene. [(a), bottom row] and (b). From a randomly coded image, a high-resolution HS image with 31 waveband is reconstructed (Media 1). [(a), top]. A color-coded image of the reconstruction is shown. (c). We plot the retrieved spectrum of individual regions and compare them to ground-truth.

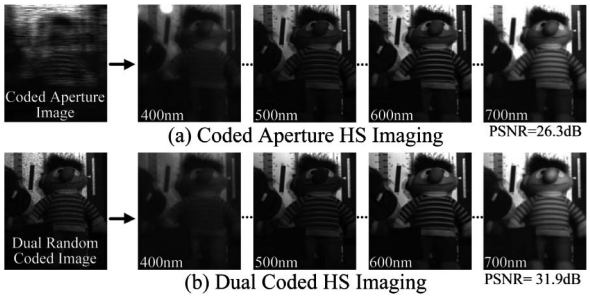

Fig. 8. Comparison between coded aperture method (CASSI, top row) and proposed approach (bottom row).

for different, simulated experiments in Fig. 4(d). In this implementation, we add zero-mean independent and identically distributed Gaussian noise to the modulated image and calculate the peak signal-to-noise ratio (PSNR) of the reconstruction. Fig. 4(d) shows that the proposed dictionary-based method works well, even with high noise levels.

We also evaluate the accuracy of the proposed compressive algorithm in Fig. 7. The reconstructed results in Fig. 7(b) are calibrated for lighting and camera response, as shown by the yellow curve in Fig. 7(c). The light and camera response is precalibrated using the color checker. The normalized intensity variation of 31 spectral channels for three of the color checker patches marked in red, green, and blue, respectively, are shown in Fig. 7(c) and compared to ground truth. In this case, the sum of squares due to error (SSE) between reconstructed results and ground truth are 0.0204, 0.0040, and 0.0197, respectively. Finally, we synthetically compare DCSI with CASSI [4] in Fig. 8 by using the dataset from [15]. The proposed method provides better accuracy (31.9 dB) for this example than the coded aperture method (26.3 dB).

In summary, the proposed dual-coded compressive hyperspectral imaging approach provides flexible capture modes, but requires two spatial light modulators. Whereas this is more expensive and requires additional calibration efforts, the DMD and relay optics used in the presented experiments allow for pixel-wise coded exposures. In future sensor implementations, these could be directly fabricated in silicon. Compared with spectral scanning methods, the proposed approach provides significant improvements in sensor readout time and facilitates dynamic scene capture.

Yet, our current prototype has several limitations. First, the optical relay imposes a tradeoff between f-number (photon count) and sensitivity in the spectral plane [11]. Second, we assume that photographed scenes are static during the exposure time. In the future, we would like to experiment with static coded masks in front of the sensor that would make the setup significantly more inexpensive and improve the form factor.

This work was supported by the Project of NSFC (Nos. 61327902, 61120106003, and 61035002). Gordon Wetzstein was supported by the NSERC PDF.

References

1. N. Gat, G. Scriven, J. Garman, M. D. Li, and J. Zhang, Proc. SPIE **6302**, 63020M (2006).
2. L. Gao, R. T. Kester, N. Hagen, and T. S. Tkaczyk, Opt. Express **18**, 14330 (2010).
3. B. K. Ford, M. R. Descour, and R. M. Lynch, Opt. Express **9**, 444 (2001).
4. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, Opt. Express **17**, 6368 (2009).
5. H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, Appl. Opt. **52**, D12 (2013).
6. D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, Appl. Opt. **49**, 6824 (2010).
7. Y. Wu, I. O. Mirza, G. R. Arce, and D. W. Prather, Opt. Lett. **36**, 2692 (2011).
8. Y. August, C. Vachman, Y. Rivenson, and A. Stern, Appl. Opt. **52**, D46 (2013).
9. E. J. Candes and T. Tao, IEEE Trans. Inf. Theory **51**, 4203 (2005).
10. D. L. Donoho and X. Huo, IEEE Trans. Inf. Theory **47**, 2845 (2001).
11. A. Mohan, R. Raskar, and J. Tumblin, Comput. Graph. Forum **27**, 709 (2008).
12. D. L. Donoho, IEEE Trans. Inf. Theory **52**, 1289 (2006).
13. E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, Appl. Comput. Harmon. Anal. **31**, 59 (2011).
14. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, IEEE Trans. Image Process. **19**, 2241 (2010).
15. www.cs.columbia.edu/CAVE/databases/multispectral/.
16. M. Aharon, M. Elad, and A. Bruckstein, IEEE Trans. Signal Process. **54**, 4311 (2006).
17. E. van den Berg and M. P. Friedlander, SIAM J. Sci. Comput. **31**, 890 (2008).