Snapshot hyperspectral imaging via spectral basis multiplexing in Fourier domain

CHAO DENG,[1,2,3] XUEMEI HU,[1,3] JINLI SUO,[1,*] YUANLONG ZHANG,[1] ZHILI ZHANG,[2] AND QIONGHAI DAI[1]

[1] *Department of Automation, Tsinghua University, Beijing, 100084, China*
[2] *High-Tech Institute of Xi'an, Xi'an, 710025, China*
[3] *These authors contribute equally to this work*
[*] jlsuo@tsinghua.edu.cn

Abstract: Hyperspectral imaging is an important tool having been applied in various fields, but still limited in observation of dynamic scenes. In this paper, we propose a snapshot hyperspectral imaging technique which exploits both spectral and spatial sparsity of natural scenes. Under the computational imaging scheme, we conduct spectral dimension reduction and spatial frequency truncation to the hyperspectral data cube and snapshot it in a low cost manner. Specifically, we modulate the spectral variations by several broadband spectral filters, and then map these modulated images into different regions in the Fourier domain. The encoded image compressed in both spectral and spatial are finally collected by a monochrome detector. Correspondingly, the reconstruction is essentially a Fourier domain extraction and spectral dimensional back projection with low computational load. This Fourier-spectral multiplexing in a 2D sensor simplifies both the encoding and decoding process, and makes hyperspectral data captured in a low cost manner. We demonstrate the high performance of our method by quantitative evaluation on simulation data and build a prototype system experimentally for further validation.

1. Introduction

Hyperspectral imaging aims to capture the details of the spectra of natural scene . It is playing great roles in both scientific research and engineering applications, such as military security [1,2], environmental monitoring [3], biological science [4,5], medical diagnosis [6,7], scientific observation [8,9], and many other fields [10–12].

Since only 1D and 2D commercial imaging sensor are available, primary hyperspectral imaging are implemented in a scanning mode, either along spatial or spectral [13,14] dimension. Although the advanced pixel count and sensitivity of array detectors boost the spatial and spectral resolution of scanning-based hyperspectral imaging systems, the requirement of steady scanning limits the speed and robustness of scanning based methods. Hence, there is a general scientific interest in investigating spectral imaging of dynamic samples such as combustion processes [15] or fluorescent probes used in biological and biomedical imaging [16,17], and snapshot hyperspectral imaging is one of the main research focuses in this area.

The typical way to perform snapshot hyperspectral imaging is mapping different spectral bands (either single narrow band or multiplexed ones) to different positions and then collecting them by one or more detectors. So far, various methods have been proposed to implement such band mapping. The integral field spectroscopy (IFS) [18, 19], multispectral beamsplitting (MSBS) [20], image replicating image spectrometer (IRIS) [21, 22] and multispectral sagnac interferometer (MSI) [23] are the paradigms directly using stacked optical components to split spectral channels, which complicate the system and impose strict limitation on the light path building. Another drawback of the direct mapping method is limited spectral band number. For example, the spectral band number of MSBS is dependent on the beam splitter which is upper limited by 4, and the spectral resolution of IRIS is limited to 16 wavelength bands due

to lack of large-format sufficient-birefringence Wollaston polarizers. In comparison, computed tomography imaging spectrometer (CTIS) [24], multi-aperture filtered camera (MAFC) [25], image mapping spectrometry (IMS) [26] and snapshot hyperspectral imaging Fourier transform (SHIFT) [27] are more compact but need customized components, which is usually expensive and of high fabrication difficulty. Introducing new reconstruction algorithms might inspire new imaging schemes and raise the final performance. The coded aperture snapshot spectral imager (CASSI) [28] is the first imager exploiting the compressive sensing theory to recover hyperspectrum. Getting rid of simple combination of optical elements, this method demands careful calibration and heavy computational load, which limits its broad applications in the scenarios where online reconstruction is needed. In sum, a snap shot spectral imager, with simple implementation, low computation load, and high reconstruction performance are worth studying.

One inspiring work in this direction is the frequency recognition algorithm for multiple exposures (FRAME) [29–31], which uses multiplexed sinusoidal illumination to encode several images into a single one simultaneously. Making use of both the spatial [32, 33] and spectral sparsity of the natural hyperspectral data, we propose a snapshot Fourier-Spectral-Multiplexing (FSM) method for hyperspectral imaging using a monochrome camera. Dividing the image sensor into a few subfields is a simple way to record various channels simultaneously. This method either shrinks the field of view (FOV) or linearly decrease the pixel resolution. Differently, considering that the Fourier spectrum of nature scenes concentrate at the central low frequency region, FSM can effectively eliminate severe resolution loss and FOV shrinkage. Our snapshot hyperspectral imaging method can achieve high spatial-spectral resolution and high frame rate, with low cost setup and computational workload. The qualitative performance comparison of the different state-of-the-arts is shown in Table 1. In summary, we make the following contributions:

Table 1. **The performance comparison of different hyperspectral imaging methods**

	Spatial Res.	Spectral Res.	Computational Load	Cost
IFS	low	low	low	high
MSBS	high	low	low	high
IRIS	low	low	low	high
MSI	high	low	low	high
CTIS	low	low	low	high
MAFC	low	high	low	high
IMS	low	high	low	high
SHIFT	low	high	low	high
CASSI	high	high	high	low
FRAME	high	low	low	low
FSM	**high**	**high**	**low**	**low**

- We *combine* spectral dimension reduction and Fourier-Spectral-Multiplexing (FSM) strategy as a new way of low cost hyperspectral imaging.

- We *validate* our method in simulation both quantitatively and qualitatively.

- We *build* a snapshot hyperspectral imaging prototype to verify this approach and *demonstrate* its practical benefits.

The structure overview of our paper is given as below. In section 2, we introduce our imaging scheme and formulation. In section 3, we set the parameters and quantitatively evaluate the effectiveness in simulation. In section 4, we demonstrate our method with physical experiments. In section 5, we discuss the advantages and limitations of our method, and the future work of our imaging system.

2. Method

2.1. Imaging scheme
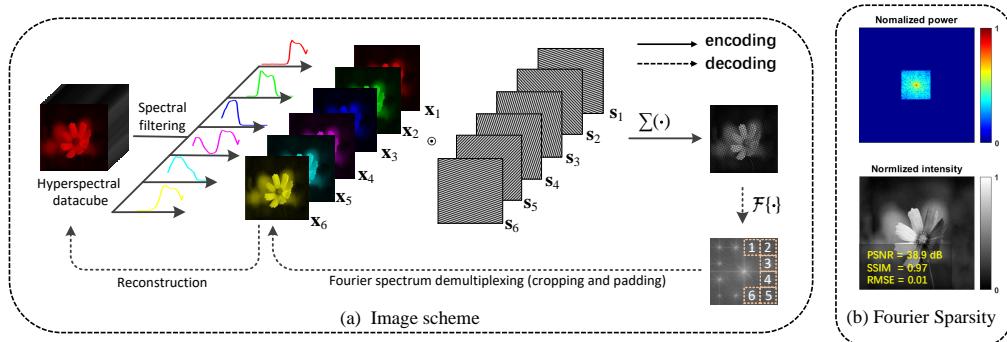
Fig. 1. (a) The scheme of the proposed hyperspectral imaging system. The hyperspectral data is spectrally filtered and projected into six images: $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_6$, and then codified by six sinusoidal patterns $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_6$, respectively. The six sinusoidal patterns shift the Fourier distribution of six projected images away from the origin into six different regions, and the gray scale camera captures the modulated images in an add-up way. The hyperspectral images are reconstructed through a two-stage method, including: Fourier spectrum demultiplexing and linear system reconstruction. (b) The centralized spreads of the nature images' Fourier coefficients. The reconstruction (bottom image) from 6.25% (0.25^2) Fourier coefficients locating at the centroid region (upper image) is quite clear.

The architecture for the snapshot hyperspecral imaging is shown in Fig. 1(a), including both encoding and decoding modules. During encoding, the 3D hyperspectral data cube is sequentially projected to a low dimensional space after going through a series of wide-band spectral filters, denoting as $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_J$. Since the spectrum of natural scenes are of low intrinsic dimension, the hyperspectral datacube can be denoted as the linear summation of a few spectral basis. To ensure that the reconstruction problem is well posed, the number of color filters is equal to that of spectral basis. Statistics tells that six bases are sufficient for high fidelity hyperspectral data representation (see [34]), so here we set J as 6. Then the projections are modulated by different sinusoidal patterns $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_J$, respectively. Mathematically, the sinusoidal modulation is represented as point-wise product "$\odot$", and would shift the Fourier spectrum of each projection into a specific region in the Fourier domain. The sinusoidal patterns are designed to ensure few overlap among the Fourier distributions of $\{\mathbf{x}_i\}$. Finally, the fast coded image is recorded by a gray scale camera in an add-up way as shown in the Fig. 1(a). Corresponding to the coding procedure, the decoding is straightforward. We first transform the captured image to the Fourier domain, and extracting the Fourier spectrum of each projection $\mathbf{x}_i$ according to the shifting effect of the corresponding sinusoidal modulation $\mathbf{s}_i$. Then we transform the separated Fourier spectrum back to the spatial domain. Finally, we back project the reconstructions to the high dimensional

data cube by solving a linear system. As Fig. 1(b) shows, the Fourier coefficients of natural scenes mainly concentrate at the low frequency region, and the reconstruction from a small proportion of the center coefficients (padding zeros for the surrounding ones) can reconstruct the image with high quality. Therefore, we propose to encode multiple spectral projections into a single image in the Fourier domain. Correspondingly, the Fourier spectrum demultiplexing includes two processes: Fourier spectrum cropping and zero padding to the original resolution. In Sec. 2.2, we introduce the formulation of the whole encoding and decoding process in details.

2.2. Formulation

Though going through the ith spectral filter, the spectrum of the target scene is filtered with the corresponding spectral transmission, which can be represented as

$$\mathbf{x}_i = \int_{\lambda} \mathbf{I}_i(\lambda)\mathbf{r}(\lambda)d\lambda, \tag{1}$$

in which $\mathbf{I}_i(\lambda)$ is the spectrum transmission of the ith spectral filter and $\mathbf{r}(\lambda)$ denoting the spectrum reflectance/transmission. Here we omit 2D spatial coordinate for brevity. Statistically, the spectra of the natural materials can be represented as a linear summation of a few (e.g., $J=6$) characteristic spectral basis [35], i.e.

$$\mathbf{r}(\lambda) = \sum_{j=1}^{J} \alpha_j \mathbf{b}_j(\lambda), \tag{2}$$

where $\mathbf{b}_j$ is the jth spectral basis, and α_j is the corresponding coefficients. Substituting Eq. (2) into Eq. (1) [36, 37], we could get

$$\mathbf{x}_i = \sum_{j=1}^{J} \alpha_j \int_{\lambda} \mathbf{I}_i(\lambda)\mathbf{b}_j(\lambda)d\lambda. \tag{3}$$

In Eq. (3), $\mathbf{I}_i(\lambda)$ is precalibrated and $\mathbf{b}_j(\lambda)$ is trained from hyperspectral database [35]. So the hyperspectrum of the surface can be represented by J parameters α_j.

The J sinusoidal patterns $\{\mathbf{s}_i\}$ for Fourier domain multiplexing of J projected images $\{\mathbf{x}_i\}$ are designed to have as minimal overlap in the Fourier domain as possible. Generally, the sinusoidal pattern can be written as

$$\mathbf{s}_i = 1 + \cos(\mathbf{p} \cdot \boldsymbol{\omega}_i), \tag{4}$$

where $\mathbf{p}$ is the 2D spatial coordinate, $\boldsymbol{\omega}_i$ ($i = 1, 2, \cdots, J$) represents the spatial frequency, and the Fourier transform of Eq. (4) includes three delta functions $\delta(\omega + \omega_i)$, $\delta(\omega)$, $\delta(\omega - \omega_i)$ in the Fourier domain. After applying such sinusoidal modulation to the target scene, its spatial spectrum is duplicated into three replicas centered at $\omega = -\omega_i$, $\mathbf{0}$, ω_i, respectively. Research in the field of natural image statistics suggests that the Fourier domain of natural images is sparse [38] and enables efficient coding and sampling of natural images [33]. Taking advantage of this prior information, if the shifted distance $\|\omega_i\|$ is set properly, these three replicas can be separated to eliminate aliasing effectively. To set the sinusoidal patterns statistically, we first analyze the Fourier spectrum of natural scenes and then estimate the proper shift distance and the corresponding cropping size accordingly (details are discussed in Sec. 3).

The encoded image is recorded by the gray scale camera as

$$\mathbf{y} = \sum_{i=1}^{J} \mathbf{x}_i \odot \mathbf{s}_i. \tag{5}$$

In terms of reconstruction, we first transform the captured image $\mathbf{y}$ to Fourier domain and demultiplexing each image by Fourier spectrum cropping and padding according to each

sinusoidal pattern. After inverse Fourier transform to spatial domain, we obtain estimations of J spectral filtered images $\{\hat{\mathbf{x}}_i\}$. To further remove the aliasing, we use generalized alternating projection (GAP) to improve the reconstruction using these estimations as the initial value. GAP is basically an extended alternating projection algorithm which solves the compressive sensing problem in the transformed domain, i.e. discrete cosine transformation or wavelet domain [39, 40]. The optimization problem is formulated as

$$\min_{\mathbf{W}} \|\mathbf{W}\|_{\ell_{2,1}^{\mathcal{G}\,\beta}}, \quad \text{subject to} \quad \mathbf{Y} = \mathbf{S}\mathbf{X} \quad \text{and} \quad \mathbf{X} = \mathbf{T}\mathbf{W}, \tag{6}$$

where the capital notation $\mathbf{Y}$ and $\mathbf{X}$ are vectorized forms of $\mathbf{y}$ and $\mathbf{x}$ [e.g., $\mathbf{Y} = \text{vec}(\mathbf{y})$, $\mathbf{X} = \text{vec}(\mathbf{x})$], respectively. $\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_J]$ with $\mathbf{S}_i$ being diagonalization of vectorized form of $\mathbf{s}_i$ [e.g., $\mathbf{S}_i = \text{diag}(\text{vec}(\mathbf{s}_i))$]. $\mathbf{T}$ is wavelet transformation matrix, and $\mathbf{W}$ is corresponding coefficient in the transformed domain.

The $\|\cdot\|_{\ell_{2,1}^{\mathcal{G}\,\beta}}$ is weighted group-$\ell_{2,1}$ norm, calculated as

$$\|\mathbf{W}\|_{\ell_{2,1}^{\mathcal{G}\,\beta}} = \sum_{k=1}^{m} \beta_k \|\mathbf{W}_{\mathcal{G}_k}\|_2, \tag{7}$$

where m is group number, $\mathbf{W}_{\mathcal{G}_k}$ is a subvector of $\mathbf{W}$ from kth group, with $\mathcal{G}_k$ being components index and β_k being the group weight coefficient. The Eq. (6) would converge with desired accuracy after a number of iterations. After de-aliasing, we obtain estimations of six spectral-filtered images $\{\hat{\mathbf{x}}_i\}$ with high quality.

To obtain the hyperspectral images, we inverse the linear system defined in Eq. (3) and get the following constrained optimization:

$$\arg\min_{\boldsymbol{\alpha}} \|\hat{\mathbf{x}} - \mathbf{F}\boldsymbol{\alpha}\|_2^2 + \eta \left\|\frac{\partial^2 \mathbf{r}(\lambda)}{\partial \lambda^2}\right\|_2^2 \quad \text{s.t.} \quad \mathbf{r}(\lambda) \geq 0. \tag{8}$$

Here $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \cdots, \hat{\mathbf{x}}_J]^T$, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_J]^T$, and $\mathbf{F}_{ij} = \int_\lambda \mathbf{I}_i(\lambda)\mathbf{b}_j(\lambda)d\lambda$. The parameter η weights the spectrum smoothness and we set $\eta = 200$ in implementation empirically. This constrained optimization problem is solved with the quadratic programming solver in Matlab, which is based on the Lagrangian multiplier method.

## 3. Parameter setting and quantitative evaluation
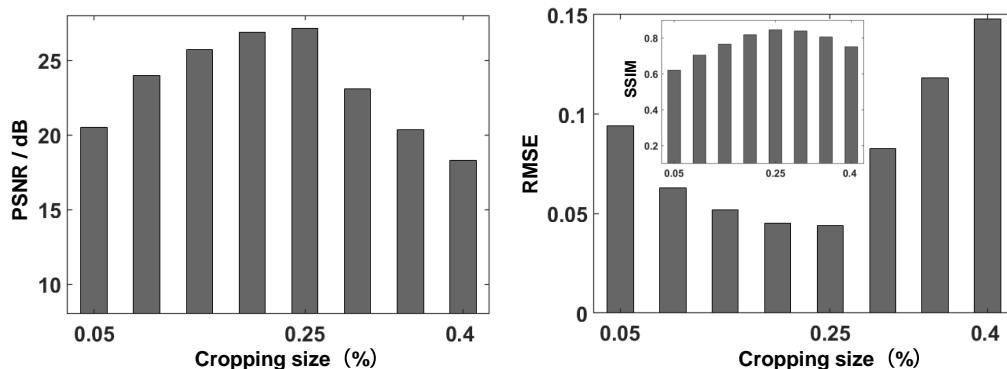
Fig. 2. The PSNR, RMSE and SSIM scores of different cropping sizes, ranging from 0.05 to 0.4 of image width.

The cropping size is designed from following aspects: (i) The Fourier distributions of six filtered versions are slightly different but similar, so we set the cropping size uniformly; (ii) To minimize the crosstalk, the six modulation frequencies should be located away from the origin and adjacent version; (iii) The cropped regions should occupy as much area in the Fourier domain to avoid losing many details in the final reconstruction. To compare different cropping sizes, we simulate the performance in terms of peak signal-to-noise ratio (PSNR), root mean square error (RMSE) and structural similarity (SSIM) [41–43], the result is shown in Fig. 2. Specifically, the cropping size ranges from 0.05 to 0.4 of image width, and the PSNR, RSME and SSIM values are the average of six Fourier de-multiplexing channels, respectively. From the simulation, we can see that the optimum performance is achieved when the cropping size is set to 0.25 of image width. Moreover, we conduct a statistical analysis over the multi-spectral database built by the CAVE laboratory at Columbia University [44] to evaluate the 0.25 Fourier cropping, in terms of PSNR, RMSE and SSIM, these three metrics maintain at 32.0 dB, 0.03 and 0.94 on average. We use the same modulation scheme for different samples in order to develop a general imager. The scheme is applicable for most nature scenes, because their spectral spreads are similar and applying GAP de-aliasing can handle the slight variance among samples.
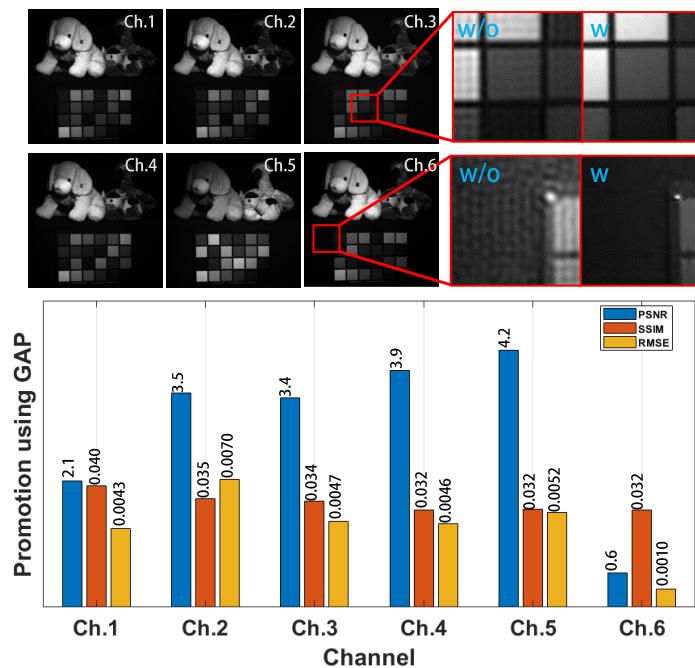
Fig. 3. Performance promotion by GAP. **Upper part**: six channels reconstructed after GAP optimization. Ch.1~Ch.6 represent the six projections, and the close-ups compare the de-aliasing before and after using GAP. w/o: without GAP optimization; w: with GAP optimization. **Bottom part**: PSNR, SSIM and RMSE *promotion* through GAP optimization.

During de-multiplexing of the spatial projections from the recorded encoded measurement, as aforementioned, we use GAP algorithm after Fourier extracting to further suppress aliasing. The simulation on the CAVE multi-spectrum database reveals 3.05 dB improvement on average by introducing GAP optimization. Fig. 3 displays the result on an example image with and without GAP optimization. We can clearly see the improvement introduced by GAP, in terms of PSNR, RMSE, and SSIM.

To evaluate the performance of our hyperspectral imaging system quantitatively, we test the
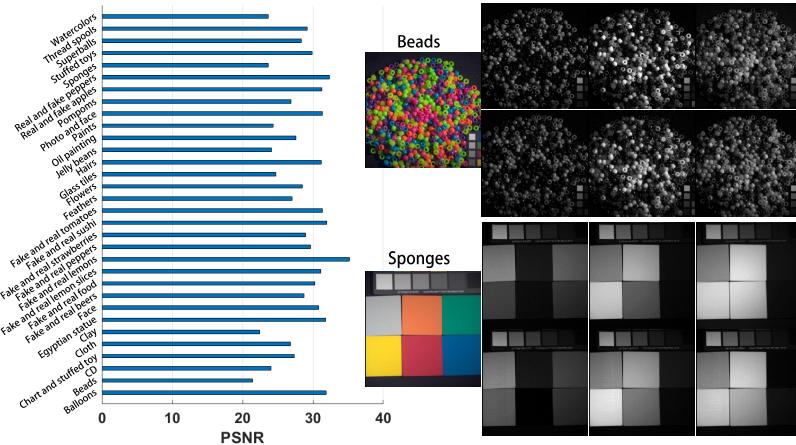
Fig. 4. The PSNR of the reconstruction and two examples from the CAVE multi-spectral database. In each example, the upper row is the ground truth in 500 nm, 600 nm, and 700 nm, respectively and the lower row is the corresponding reconstruction.

imaging accuracy on the CAVE's multi-spectral image database. For each example, we simulate the encoded image based on the six sinusoidal patterns $\{s_i\}$ and CW's spectral responses $\{I_i\}$ demonstrated in Fig. 5(c). The reconstruction results are quite promising, with PSNR averagely over 28.4 dB, as shown in the left subfigure of Fig. 4. For a clearer demonstration, we also display two scenes with the lowest PSNR in the right part: Beads and Sponges, each with the ground truth (upper row) and corresponding reconstruction (lower row) displayed in comparison. We can see high fidelity reconstruction are achieved.

4. Imaging prototype and experiment on captured data

We build a prototype to verify our imaging scheme as shown in Fig. 5. The collimated broadband light is first modulated by a synchronized rotating color wheel (CW). The off-the-shelf six segment color wheel is driven by a high speed motor (a 60k rpm Walkera Super CP motor) and placed on the Fourier plane of the 4f system composed of L1 and L2 for spectral modulation. The spot on the color wheel is small enough so that the large temporary overlap between two segments can be neglected. As the correlation among the spectral responses of the off-the shelf segments is high, we conduct a slight modification. Specifically, we choose three segments (red, green and blue) from the off-the-shelf color wheel, and attach three other broadband color papers to the transparent segment. The three new spectral filters are selected according to two criteria. First, for effective reconstruction, the spectrum of these filters should cover the target spectrum. Second, from an available spectral filters set, we traverse all the possible combinations to choose the optimum filter group with minimum correlation by following optimization

$$\min_{\substack{1 \leq i,j \leq 6, \\ i \neq j}} \max corr(I_i, I_j), \tag{9}$$

in which I_i, I_j are two spectral filters in each group, and the notation $corr$ represents correlation of two vectors. Except for large ratio between bandwidth and the number of reconstructed wavelength [45], broadband spectral filters has higher signal-to-noise ratio than narrowband ones. Then a gray scale sinusoidal pattern module (GSM) modulates the incoming light. Specifically, we use a DMD to modulate the binary approximation of gray scale sinusoidal patterns to achieve full modulation rate, and apply pinhole filtering on the Fourier plane to obtain ideal sinusoidal patterns. As for binarization, we choose dithering algorithm considering its low approximation

(a) Optical diagram

(b) Imaging setup

(c) Transmissive response of CW

(d) GSM
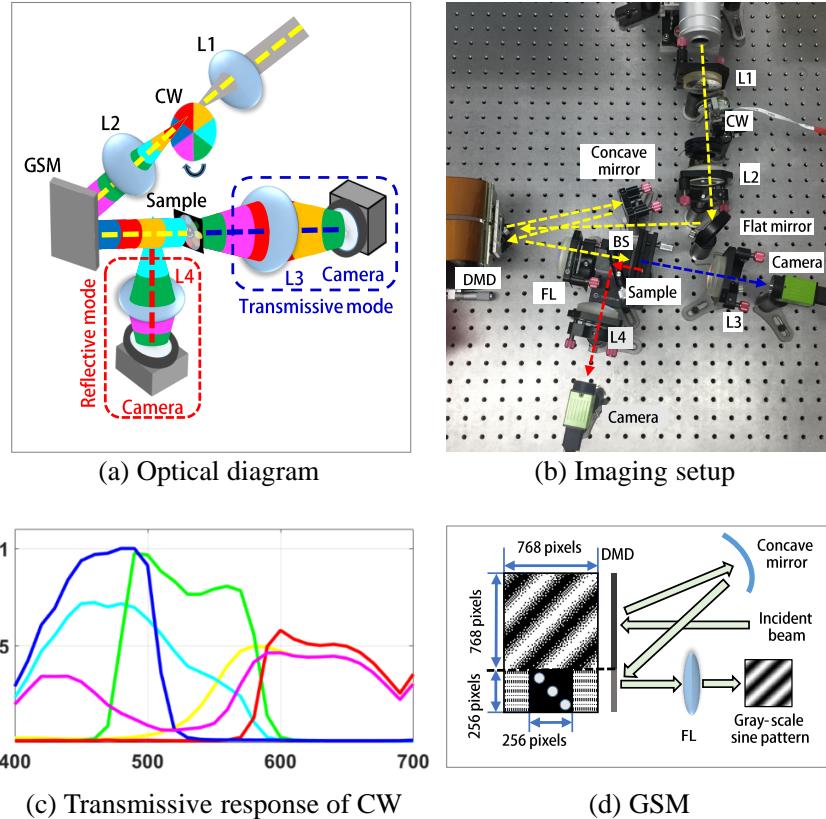
Fig. 5. The imaging scheme of proposed method. (a) The optical diagram including transmissive and reflective mode (the beam splitter is omitted in reflective mode). L1 and L2: converging lens composing a 4f system. CW: color wheel. GSM: gray scale sinusoidal modulation. L3 and L4: Converging lens. (b) The imaging setup including both transmissive and reflective mode. DMD: digital micromirror device; FL: Fourier lens. (c) The transmissive response of CW. (d) The GSM module implementing fast gray scale sinusoidal modulation.

error. In implementation, we use a single DMD to conduct binary patterning and Fourier-domain filtering simultaneously. The left part of the DMD modulates the incident light beam with a dithering sinusoidal pattern, and the right half of the DMD optically filters the spatial spectrum to remove the approximation error introduced by dithering algorithm [Fig. 5(d)]. The filter selects three dominant frequencies of an ideal sinusoidal pattern, i.e., -1, 0, and +1. Here the Fourier transform is taken by a concave mirror, which is well designed so that the right region of DMD is the Fourier plane of the left counterpart. This modulation can achieve 20 kHz gray scale sinusoidal patterning and is sufficient for our imaging scheme. To properly configure the optical elements in the narrow space, we slightly rotate the DMD plane so that its incident and outgoing beams are of a small angle. Moreover, to achieve compact optical design, we can also exploit short-focus lenses with small aperture and customized optical bracket. We use the GSM modulation for three reasons. First, unlike ideal gray scale sinusoidal modulation, digital representation could introduce unwanted frequency components in the Fourier domain and thus deteriorate the final reconstruction. Second, the GSM arrangement can also block the diffraction orders introduced by DMD. Third, as the refresh rate of gray scale DMD modulation is upper limited to 253 Hz (8 bit), we adopt the same method proposed in [46] to implement fast sinusoidal
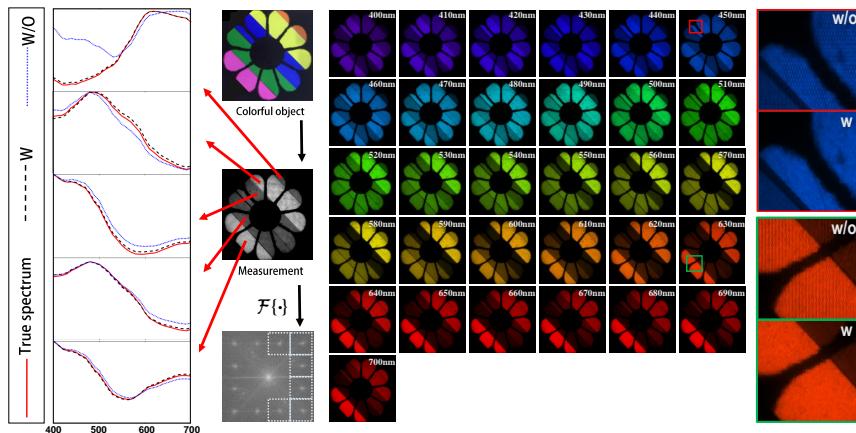
Fig. 6. The experimental reconstruction of a flower film with five strips of different transparent color paper. The RGB image of the object, the measurement together with its Fourier space and the reconstructed five spectrum of transparent color paper are displayed in the left part. The solid red, dashed black and dotted blue curve are the true spectrum, reconstruction with (w) GAP and without (w/o) GAP, respectively. The reconstructed hyperspectral images are displayed in the middle. The spectral range is 400 nm ~ 700 nm. We highlight the distinction of reconstruction without and with GAP optimization in the right part.

modulation for scalability, with the scheme illustrated in Fig. 5(d). After modulation, the light beam is focused by the converging lens onto the sample. Finally, the encoded sample information is collected by a gray scale camera (GO-5000C-USB, JAI) after converged by a lens.

In our experiment, the pixel resolution of both sinusoidal patterns displayed in DMD and the detector are 768×768, and the Fourier cropping size is 192 pixels ($768 \times 25.0\%$). The spectral resolution is mainly determined by three factors: First, the number of PCA bases, as mentioned in Chapter 4. The spectral resolution increases with the number of PCA bases (e.g., eight in [35]). Second, the selection of the transmissive spectrums of the color wheel segments. We can obtain higher spectral resolution by reducing the correlation among the transmissive spectrums. Third, the regularization parameter. An improper regularization parameter η in Eq. 8 would degenerate the spectral resolution: a too small η cannot suppress noise effectively, while a too large η would smooth out the spectral curve. Empirically, we set $\eta = 200$ in implementation. We use external trigger mode for synchronization: the trigger out signal per circle of the color wheel triggers the DMD for displaying the sinusoidal patterns, and again the trigger out signal of DMD triggers the exposure of the camera. The final imaging speed is limited by the lowest speed of three elements – the rotation of color wheel (1000 Hz), camera's frame rate (~ 100 Hz), and DMD refresh rate (20000 Hz). Since the rotating frequency of color wheel is equal to the camera frame rate, which is upper limited to hundreds Hz, the noise introduced by such low speed rotation and vibration is negligible, as widely used in a commercial projector. In our implementation, we set the camera frame rate as 24 Hz, which can handle daily moving scenes and a higher frame rate camera can be used for faster movements.

To test the reconstruction accuracy of the hyperspectral images, we use a static color scene to evaluate our imaging system quantitatively. The test scene is strip-wisely uniform and we can calibrate the spectrum with a spectrometer. From the results in Fig. 6, we can see that the reconstructed spectrum of each patch after GAP optimization is of great consistency with the ground truth spectrum. We highlight the distinction after GAP optimization for clearer observation in the right part.

To demonstrate the dynamic acquisition capability of our method, we conduct two hyperspectral
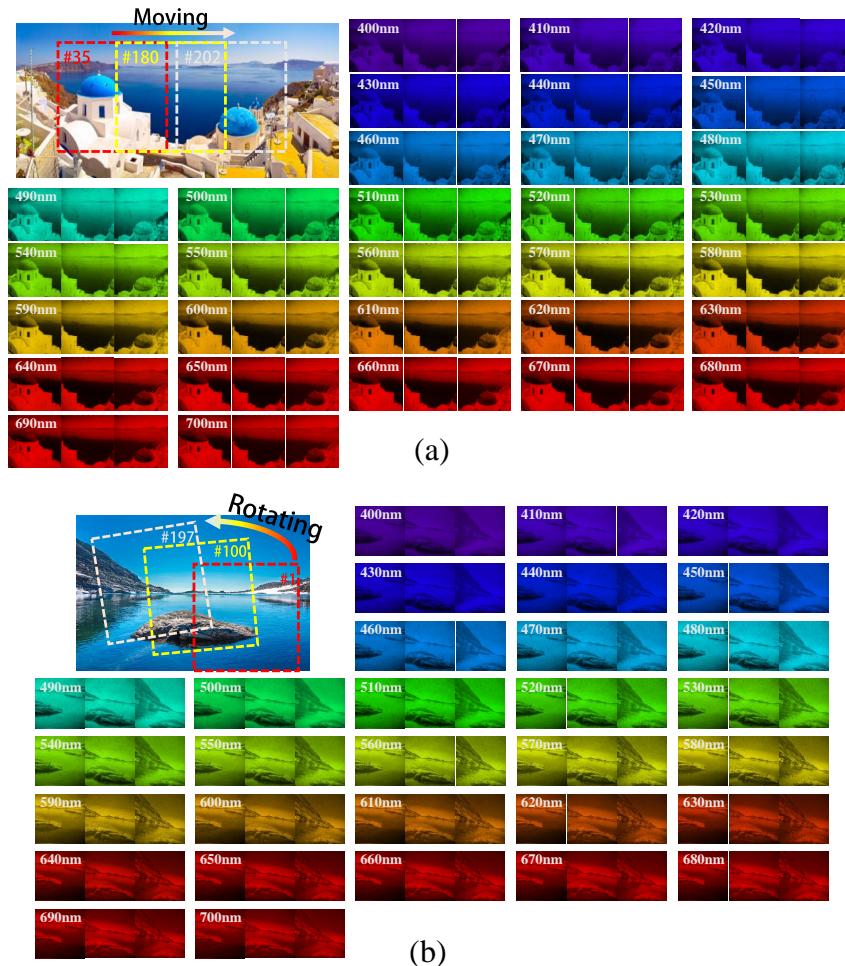
Fig. 7. Hyperspectral reconstruction of regular motion: (a): a color film of Santorini Island mounted on a translation stage moving at a constant speed of 3.7 mm/s; (b): a color film of a landscape fixed on a rotary translation stage with angular speed of 0.023 rad/s.

reconstructions of dynamic scenes. The first one is composed of two transmissive scenes with regular motions at a constant speed, as shown in Fig. 7. The translation scene is a color film of Santorini Island mounted on a translation stage moving at a constant speed of 3.7 mm/s. The gray scale camera works at 24 frames per second. We display three frames in Fig. 7(a), and the hyperspectral vedio of total 274 frames is available online (Visualization 1). The rotation one is a colorful rotating film [Fig. 7(b)] fixed on a rotary translation stage. The angular speed is 0.023 rad/s. The hyeperspectral video showing the whole process is also available online (197 frames in total, see Visualization 2).

The second experiment is the diffusion process of two color pigments poured into a glass of water, captured in the reflective mode. We first pour the light blue pigment and then the orange one. The color of water gradually changes from light blue to orange, and eventually to the mixture of blue and orange. The spectrum distribution is changing over the diffusion process. We can also see the sinking and mixing process of these two color pigments in the water. We display three frames in Fig. 8 to show the dynamics. The video of total 77 frames is available online (Visualization 3). From the two reconstructions, we can clearly see that the proposed method can
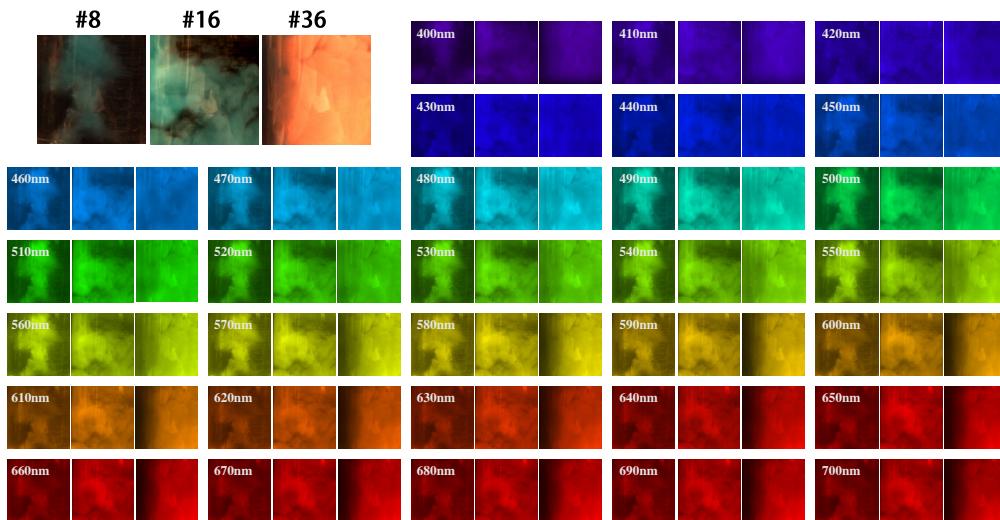
Fig. 8. Hyperspectral reconstruction of the diffusion process of color pigments poured into clean water. Three out of 77 frames are displayed.

work well for hyperspectral imaging of dynamic scenes.

5. Summary and discussions

In summary, we propose a snapshot hyperspectral imaging technique jointly utilizing the spectral and spatial redundancy of hyperspectral data cube. Specifically, we conduct spectral dimension reduction and spatial frequency multiplexing under the computational imaging scheme. For reconstruction, we can resolve spectral transmission/reflectance with low computational load in realtime. Benefiting from a recently proposed fast sinusoidal modulation [47] working at up to 20 kHz and a rotator working at 1 k revolution per second (rps), our imaging speed is mainly limited by the frame rate of the adopted camera and thus can work well for the common dynamic scenes. Moreover, utilizing the temporal redundancy of natural scenes, imaging speed can be further improved by introducing random modulation based on compressive sensing [48–52]. The spatial multiplexing scheme would reduce the reconstruction quality slightly in our scheme. To compensate the loss of high frequency details, we can introduce super resolution techniques, either single image based algorithms or sequence based techniques. In short, our technique is quite promising to be extended to a high performance hyperspectral imaging system.

It is worth mentioning that the proposed approach might be of sectioning ability for thicker transparent/translucent samples. The sectioning ability depends on the imaging mode: for fluorescent imaging, only the focusing plane is excited and thus we can capture the hyerspectral data of a specific layer; for non-fluorescent imaging, the result would be a summation of different z-planes, including one focusing plane and some out-of-focus ones. Besides, since we focus on developing a general hyperspectral imaging, six fixed modulation frequencies are used for all the natural scenes. Scene adaptive modulation may be more efficient since samples have different frequency spreads in the Fourier domain, which is worth further studying.

Funding

References

1. M. T. Eismann, C. R. Schwartz, J. N. Cederquist, J. A. Hackwell, and R. J. Huppi, "Comparison of infrared imaging hyperspectral sensors for military target detection applications," Proc. SPIE **2819**, 91–102(1996).
2. J. P. Ardouin, J. Lévesque, and T. A. Rea, "A demonstration of hyperspectral image exploitation for military applications," in *Proceedings of IEEE Conference on Information Fusion* (IEEE, 2007), pp. 1–8.
3. S. Delalieux, A. Auwerkerken, W. W. Verstraeten, B. Somers, R. Valcke, S. Lhermitte, J. Keulemans, and P. Coppin, "Hyperspectral reflectance and fluorescence imaging to detect scab induced stress in apple leaves," Remote Sens. **1**, 858–874 (2009).
4. V. Backman, M. B. Wallace, L. Perelman, J. Arendt, R. Gurjar, M. Müller, Q. Zhang, G. Zonios, E. Kline, T. McGillican, S. Shapshay, T. Valdez, K. Badizadegan, J. M. Crawford, M. Fitzmaurice, S. Kabani, H. S. Levin, M. Seiler, R. R. Dasari, I. Itzkan, J. Van Dam and M. S. Feld, "Detection of preinvasive cancer cells," Nature **406**, 35–36 (2000).
5. G. Zavattini, S. Vecchi, G. Mitchell, U. Weisser, R. M. Leahy, B. J. Pichler, D. J. Smith, and S. R. Cherry, "A hyperspectral fluorescence system for 3D in vivo optical imaging," Phys. Med. & Biol. **51**, 2029 (2006).
6. R. T. Kester, N. Bedard, L. Gao, and T. S. Tkaczyk, "Real-time snapshot hyperspectral imaging endoscope," J. Biomed. Opt. **16**, 056005 (2011).
7. T. Vo-Dinh, "A hyperspectral imaging system for in vivo optical diagnostics," IEEE Eng. Med. Biol. Mag. **23**, 40–49 (2004).
8. M. J. Barnsley, J. J. Settle, M. A. Cutter, D. R. Lobb, and F. Teston, "The proba/chris mission: A low-cost smallsat for hyperspectral multiangle observations of the earth surface and atmosphere," IEEE Tran.on Geosci. Remote Sens. **42**, 1512–1520 (2004).
9. J. P. Bibring, Y. Langevin, A. Gendrin, B. Gondet, F. Poulet, M. Berthé, A. Soufflot, R. Arvidson, N. Mangold, J. Mustard, P. Drossart, and the OMEGA team, "Mars surface diversity as revealed by the omega/mars express observations," Science **307**, 1576–1581 (2005).
10. Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," IEEE Trans. on Pattern Anal. Mach. Intell. **25**, 1552–1560 (2003).
11. D. J. Brady, *Optical imaging and spectroscopy* (John Wiley & Sons, 2009).
12. G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," IEEE Signal Process. Mag. **31**, 105–115 (2014).
13. A. F. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," Science **228**, 1147–1153 (1985).
14. R. G. Sellar and G. D. Boreman, "Comparison of relative signal-to-noise ratios of different classes of imaging spectrometer," Appl. Opt. **44**, 1614–1624 (2005).
15. A. J. Reiter and S. C. Kong, "Combustion and emissions characteristics of compression-ignition engine using dual ammonia-diesel fuel," Fuel **90**, 87–97 (2011).
16. P. Kauranen, S. Andersson-Engels, and S. Svanberg, "Spatial mapping of flame radical emission using a spectroscopic multi-colour imaging system," Appl. Phys. B **53**, 260–264 (1991).
17. C. E. Volin, B. K. Ford, M. R. Descour, J. P. Garcia, D. W. Wilson, P. D. Maker, and G. H. Bearman, "High-speed spectral imager for imaging transient fluorescence phenomena," Appl. Opt. **37**, 8112–8119 (1998).
18. L. Weitzel, A. Krabbe, H. Kroker, N. Thatte, L. Tacconi-Garman, M. Cameron, and R. Genzel, "3d: The next generation near-infrared imaging spectrometer," Astron. Astrophys. Suppl. Ser. **119**, 531–546 (1996).
19. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," Opt. Eng. **52**, 090901 (2013).
20. J. D. Matchett, R. I. Billmers, E. J. Billmers, and M. E. Ludwig, "Volume holographic beam splitter for hyperspectral imaging applications," Proc. SPIE **6668**, 66680K (2007).
21. A. Gorman, D. W. Fletcher-Holmes, and A. R. Harvey, "Generalization of the lyot filter and its application to snapshot spectral imaging," Opt. Express **18**, 5602–5608 (2010).
22. A. R. Harvey and D. W. Fletcher-Holmes, "High-throughput snapshot spectral imaging in two dimensions," Proc. SPIE **4959**, 4959136 (2003).
23. M. W. Kudenov, M. E. Jungwirth, E. L. Dereniak, and G. R. Gerhart, "White-light sagnac interferometer for snapshot multispectral imaging," Appl. Opt. **49**, 4067–4076 (2010).
24. T. Okamoto and I. Yamaguchi, "Simultaneous acquisition of spectral image information," Opt. Lett. **16**, 1277–1279 (1991).
25. A. Hirai, T. Inoue, K. Itoh, and Y. Ichioka, "Application of multiple-image fourier transform spectral imaging to measurement of fast phenomena," Opt. Rev. **1**, 205–207 (1994).
26. L. Gao, R. T. Kester, and T. S. Tkaczyk, "Compact image slicing spectrometer (iss) for hyperspectral fluorescence microscopy," Opt. Express **17**, 12293–12308 (2009).
27. M. W. Kudenov and E. L. Dereniak, "Compact snapshot birefringent imaging fourier transform spectrometer," Proc. SPIE **7812**, 6–12 (2010).
28. M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," Opt. Express **15**, 14013–14027 (2007).
29. K. Dorozynska and E. Kristensson, "Implementation of a multiplexed structured illumination method to achieve snapshot multispectral imaging," Opt. Express **25**, 17211–17226 (2017).
30. E. Kristensson, Z. Li, E. Berrocal, M. Richter, and M. Aldén, "Instantaneous 3D imaging of flame species using

coded laser illumination," Proc. Combust. Inst. **36**, 4585–4591 (2017).

31. A. Ehn, J. Bood, Z. Li, E. Berrocal, M. Aldén, and E. Kristensson, "FRAME: femtosecond videography for atomic and molecular dynamics," Light. Sci. Appl. **6**, 17045 (2017).

32. A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2011), pp. 193–200.

33. L. Bian, J. Suo, X. Hu, F. Chen, and Q. Dai, "Efficient single pixel imaging in fourier space," J. Opt. **18**, 085704 (2016).

34. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," IEEE Trans. Image Process. **19**, 2241–2253 (2008).

35. J. P. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of munsell colors," J. Opt. Soc. Am. A **6**, 318–322 (1989).

36. J. I. Park, M. H. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral imaging using multiplexed illumination," in *Proceedings of IEEE Conference on Computer Vision* (IEEE, 2007), pp. 1–8.

37. S. Han, I. Sato, T. Okabe, and Y. Sato, "Fast spectral reflectance recovery using dlp projector," in *Proceedings of Asian Conference on Computer Vision,* (Springer, 2010), pp. 323–335.

38. M. Rabbani, "JPEG2000: Image compression fundamentals, standards and practice," J. Electron. Imaging **11**, 286 (2002).

39. X. Liao, H. Li, and L. Carin, "Generalized alternating projection for weighted-2,1 minimization with applications to model-based compressive sensing," SIAM J. on Imaging Sci. **7**, 797–823 (2014).

40. X. Yuan, "Generalized alternating projection based total variation minimization for compressive sensing," in *Proceedings of IEEE International Conference on Image Processing* (IEEE, 2016), pp. 2539–2543.

41. 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).

42. R. Dosselmann, and X. Yang, "A comprehensive assessment of the structural similarity index," Signal, Image and Video Process. **5**, 81–91 (2009).

43. Z. Wang, and Q. Li, "Information Content Weighting for Perceptual Image Quality Assessment," IEEE Trans. Image Process. **20** 1185–1198 (2009)

44. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," IEEE Trans. Image Process. **19**, 2241–2253 (2010).

45. P. Wang, and R. Menon, "Computational spectrometer based on a broadband diffractive optic," Opt. express **22**, 14575–14587 (2014).

46. Y. Zhang, J. Suo, Y. Wang, and Q. Dai, "Doubling the pixel count limitation of single-pixel imaging via sinusoidal amplitude modulation," Opt. Express **26**, 6929–6942 (2018).

47. Z. Zhang, S. Liu, J. Peng, M. Yao, G. Zheng, and J. Zhong, "Simultaneous spatial, spectral, and 3d compressive imaging via efficient fourier single-pixel measurements," Optica **5**, 315–319 (2018).

48. Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Video from a single coded exposure photograph using a learned over-complete dictionary," in *Proceedings of IEEE International Conference on Computer Vision* (IEEE, 2011), pp. 287–294.

49. J. Holloway, A. C. Sankaranarayanan, A. Veeraraghavan, and S. Tambe, "Flutter shutter video camera for compressive sensing of videos," in *Proceedings of IEEE International Conference on Computational Photography,* (IEEE, 2012), pp. 1–9.

50. P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D. J. Brady, "Coded aperture compressive temporal imaging," Opt. Express **21**, 10526–10545 (2013).

51. D. Reddy, A. Veeraraghavan, and R. Chellappa, "P2c2: Programmable pixel compressive camera for high speed imaging," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2011), pp. 329–336.

52. X. Yuan, P. Llull, X. Liao, J. Yang, D. J. Brady, G. Sapiro, and L. Carin, "Low-cost compressive sensing for color video and depth," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2014), pp. 3318–3325.