SPATIO-SPECTRAL MODULATION USING A BINARY PHOTOMASK FOR COMPRESSIVE CHROMOTOMOGRAPHY

Xuesong Zhang[*], *Jing Jiang*[†], *Anlong Ming*[*], *Xuejing Kang*[*], *and Gonzalo R. Arce*[‡]

[*]Beijing University of Posts and Telecommunications, Beijing, 100876, China
[†] Department of Communication Engineering, Beijing Union University, Beijing, 100101, China
[‡]Department of ECE, University of Delaware, Newark, Delaware, 19716, USA

ABSTRACT

Recent advances in compressive spectral imagers have demonstrated the potential of spatio-spectral modulation (SSM) for improved reconstruction performance. Existing SSM techniques, however, use either a color filter array or a complex optical arrangement, both of which can only provide limited modulation bandwidth in the spectral dimension. This paper proposes a practical SSM method to help address the "missing cone" problem of chromo-tomography. A high-resolution binary coded aperture is used to modulate the dispersed images, which in the Fourier domain fulfills a 3D convolution of the probed spectrum with the aperture's wide spectrum. This spectrum spreading process facilitates the compressed sensing strategy determined by the Fourier Slice Theorem and we demonstrate the advantages of the proposed approach with numerical experiments.

Index Terms— Hyperspectral imaging, coded aperture, spatio-spectral modulation, chromo-tomography, compressive sensing

1. INTRODUCTION

Most imaging spectrometers employ a 2D focal plane array (FPA) to capture a 3D representation of the investigated scene, i.e., the optical wavelength extension of a two spatial dimension. Such sensing constraint results in a scanning process of conventional sensors, either spatially or spectrally, to sequentially collect the entire data cube, which demands long acquisition time and therefore is not applicable to measuring dynamic scenes. In the last two decades, various types of spectral imagers have been developed for higher light collection efficiency and fast sensing [1], which is unachievable for classical spectroscopic imaging geometries.

The advent of compressive sensing (CS) theory has significantly promoted the progress of snapshot sensing [2-5]. Following the CS theory [6], coded aperture (CA) based compressive spectral imaging uses a mask to modulate the hyperspectral data cube such that the resulting measurement matrix is incoherent with the sparsifying basis of hyperspectral imagery as much as possible. Depending on the optical implementation and layout, CAs can fulfill 2D spatial modulation, such as the coded aperture snapshot spectral imager (CASSI) [7], and 3D spatio-spectral modulation, such as the colored CASSI (CCASSI) [8] and the spatial-spectral encoded compressive hyperspectral imager (SSCSI) [9]. Since 3D modulation can provide more flexible coding schemes for probing the underlying signal features, they exhibited superior reconstruction performance over the 2D versions in the numerical comparisons [3], but their physical implementations still have to solve two important issues, optical throughput and modulation efficiency, so that 3D modulation can play a more practical role in real spectroscopic applications.

Specifically, color filter array (CFA) based 3D modulation usually attenuate light by at least more than 50% to realize the color-sensitive modulation [8, 10]. Also the use of 3 or 4 types of optical filters in CFA can just provide a low-resolution modulation in the spectral dimension. The optical throughput gets much more limited for SSCSI because it must use slow optics (F# > 10) to approximately form a spectral plane on which rays with the same wavelength converge to a single point and then get modulated by a binary mask near the sensor [9]. Low optical efficiency means longer sensor integration time for sufficient signal-to-noise ratio and therefore significantly discounts the snapshot advantage.

In this work, we propose a new compressive spectral imaging method, called coded aperture compressive chromo-tomography (CACCT). Two state-of-the-art computational spectral imaging architectures, chromo-tomographic hyperspectral imaging spectrometer (CTHIS) [11] and CASSI [7, 12], can be regarded as a special case of CACCT. In the following, we show this point via a Fourier domain analysis and verify the advantages of CACCT in terms of

Partially supported by the National Natural Science Foundation of China (61871055, 61701036), and partially by the Fundamental Research Funds for the Central Universities (2018RC54).

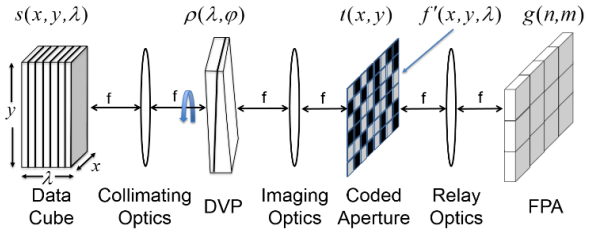

Fig. 1. Schematic of CACCT. A fixed coded aperture serves as a spatial-spectral modulator and allows for efficient sensing with the rotation of DVP.

information collection capability and reconstruction performance.

2. PROPOSED METHOD

2.1. System model

A schematic of CACCT is shown in Fig.1, where we implant into the original CTHIS architecture a binary photomask at a plane conjugated to both the object and the FPA. The spectrally resolvable image $f(x,y,\lambda)$ is first collimated and then dispersed by a direct-vision prism (DVP). The DVP rotates around the optical axis and accordingly defines the dispersion direction angle φ on the $x-y$ plane.

Let λ_c denote the undeviated wavelength after passing through the DVP, and k the dispersion coefficient of the DVP. The deviated rays of wavelength λ are reimaged onto the plane of the CA, with a spatial shift,

$$\rho(\lambda,\varphi)=\left(\delta_x^2+\delta_y^2\right)^{1/2},$$

where $\delta_x=k\left(\lambda-\lambda_c\right)\cos\varphi$ and $\delta_y=k\left(\lambda-\lambda_c\right)\sin\varphi$ are the wavelength dependent shifts in x and y directions, respectively. The dispersion and projection geometry of the data cube is shown in Fig. 2.

One can see that every point (x,y,λ) is shifted to $(x+\delta_x,y+\delta_y,\lambda)$, forming the dispersion line, and that every point $(x-\delta_x,y-\delta_y,\lambda)$ is shifted to (x,y,λ), defining the projection line with the angles pair (α,φ), where $\alpha=\cos^{-1}\left[1\Big/\left(1+k^2\right)^{1/2}\right]$ is the deviation angle from the λ axis.

The spectral density right after the CA is given by

$$f'(x,y,\lambda)=f(x-\delta_x,y-\delta_y,\lambda)t(x,y), \tag{1}$$

where $t(x,y)$ is the spatial transmission function imposed by the CA. Defining a (λ,φ) dependent transmission function

$$t'_\varphi(x,y,\lambda)=t(x+\delta_x,y+\delta_y), \tag{2}$$

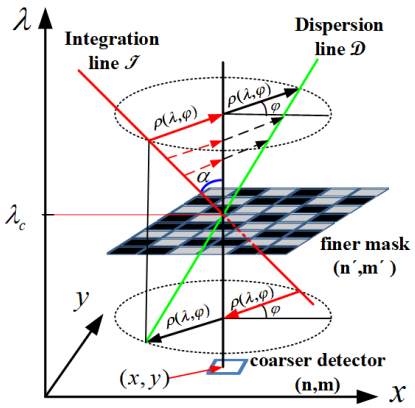

Fig. 2. Projection model of CACCT. Every point on the projection line will be spatially shifted to (x, y), modulated by $t(x, y)$ and integrated by the detector (n, m).

we reveal the fact that at each DVP's rotation angle the original spectral image $f(x,y,\lambda)$ is modulated by a spectrally variant mask $t'_\varphi(x,y,\lambda)$,

$$\begin{aligned} g_\varphi(x,y)&=\int rect\left(\frac{\lambda-\lambda_c}{\Lambda}\right)f\left(x-\delta_x,y-\delta_y,\lambda\right)\\ &\quad\times t'_\varphi\left(x-\delta_x,y-\delta_y,\lambda\right)d\lambda \\ &\triangleq\int_\Lambda S^\lambda_{(\varphi,\alpha)}\left[f(x,y,\lambda)t'_\varphi\left(x_x,y,\lambda\right)\right]d\lambda, \end{aligned} \tag{3}$$

where we define a shift operator $S^\lambda_{(\varphi,\alpha)}[\cdot]$ that shifts each spectral channel by $\rho(\lambda,\varphi)$, and the continuous image $g_\varphi(x,y)$ on the FPA is the tomographic measurement of $f(x,y,\lambda)$ modulated by $t'_\varphi(x_x,y,\lambda)$.

Suppose the CA is pixelated with square elements of size Δ_M, i.e.,

$$t(x,y)=\sum_{m'n'}t_{n'm'}rect\left(\frac{x}{\Delta_M}-m',\frac{y}{\Delta_M}-n'\right). \tag{4}$$

Then the spectral images $f'(x,y,\lambda)$ relayed to the detector plane will be integrated in a range $\Delta_D\times\Delta_D\times\Lambda$, with $\Delta_D>\Delta_M$ the pixel pitch and Λ the spectral response range of the detectors. The detector measurements under rotation angle φ are given by

$$\begin{aligned} g^{(\varphi)}_{nm}&=\sum_{m'n'}t_{n'm'}\iint rect\left(\frac{x}{\Delta_D}-m,\frac{y}{\Delta_D}-n\right)rect\left(\frac{x}{\Delta_M}-m',\frac{y}{\Delta_M}-n'\right)\\ &\quad\int rect\left(\frac{\lambda-\lambda_c}{\Lambda}\right)f\left(x-\delta_x,y-\delta_y,\lambda\right)dxdyd\lambda. \end{aligned} \tag{5}$$

If we discretize $f(x,y,\lambda)$ into $M'\times N'\times L$ voxels of size $\Delta_M\times\Delta_M\times\Delta_\lambda$ and take into account the imaging noise item ω_{nm}, (5) can be written as

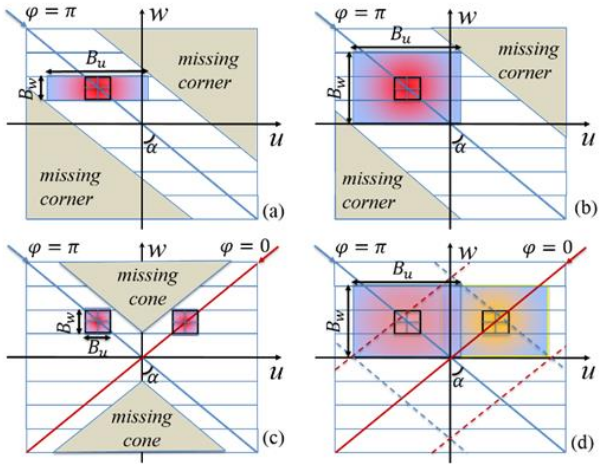

Fig. 3. Frequency sampling strategy comparison. (a) SD-CASSI. (b) DD-CASSI. (c) CTHIS. (d) CACCT. The plane $v=0$ in the Fourier cube is shown, where the missing region is determined by both the filter's bandwidth and the projection angle pair (α,φ): CACCT can fully cover the whole 3D spectrum using at least two modulated chromo-tomographic projections.

$$g_{nm}^{(\varphi)}=\sum_{l=1}^{L}\sum_{n'm'}\tilde{t}_{n'm'}f\left(m'\Delta_M-\delta_x,n\Delta_M-\delta_y,l\right)+\omega_{nm}, \quad (6)$$

where

$$\tilde{t}_{n'm'}=t_{n'm'}\iint rect\left(\frac{x}{\Delta_D}-m,\frac{y}{\Delta_D}-n\right)rect\left(\frac{x}{\Delta_M}-m',\frac{y}{\Delta_M}-n'\right)dxdy \quad (7)$$

is the footprint of the photomask element (n',m') on the pixel (n,m) of the FPA.

2.2. Spatio-spectral modulation analysis in Fourier domain

Neglecting the limited size of data cube and taking the Fourier transform (FT) w.r.t. (x,y) on both sides of (3), we have

$$\begin{aligned}\hat{g}_\varphi(u,v)&=\int d\lambda\,\exp\left[-i2\pi k(u\cos\varphi+v\sin\varphi)(\lambda-\lambda_c)\right]\\&\quad\times\left[\left(\hat{t}'_\varphi\right)_{uv}*\left(\hat{f}\right)_{uv}\right](u,v,\lambda)\\&=\left[\left(\hat{t}'_\varphi\right)_{uvw}*\left(\hat{f}\right)_{uvw}\right](u,v,k(u\cos\varphi+v\sin\varphi)),\end{aligned} \quad (8)$$

where the hat on a function represents its FT, * is the convolution operation, and the subscripts uv and uvw are the frequency variants involved in FT. Eq. (8) is the modulated version of the Fourier slice theorem: CACCT samples the convolved spectrum of f and t'_φ on a plane passing through the origin, with the third frequency dimension:

$$w=k(u\cos\varphi+v\sin\varphi). \quad (9)$$

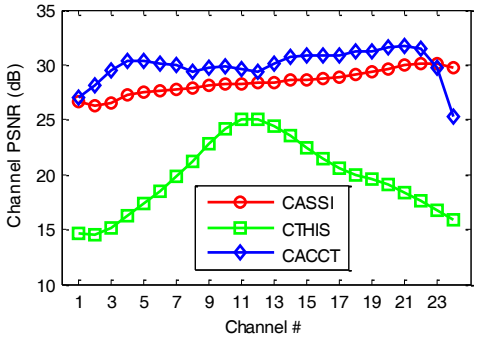

Fig. 4. PSNRs of each reconstructed spectral bands. The regularization parameters of CACCT are $\tau=$ 1e-5 and $\gamma=5\text{e}-5$.

The formulation in (8) facilitates the analysis of sensing efficiency for different architectures. In Fig. 3 we illustrate the sensing advantage of CACCT over single disperser (SD-) CASSI [7], dual-disperser (DD-) CASSI [12] and CTHIS, where the vertical plane $v=0$ in the Fourier data cube is displayed and the slices correspond to the lines passing through the origin. Because of the dispersion and integration process in each scheme, their FPA measurements all contain the frequency information acquired on the plane of (9). However, they differ in the transmission filter t'_φ, determining the volume involved in the convolution, as well as the angle φ, defining the orientation of the sampling plane.

Suppose the source cube is of size $A\times B\times\Lambda$. SD-CASSI (Fig. 3(a)) uses a common spatial filter as defined in (4) to modulate all spectral images, and the spectral filter is just $rect((\lambda-\lambda_c)/\Lambda)$. Therefore, the resulting effective spectral bandwidth is $B_w\propto\Lambda^{-1}$. As the MTF of CA sweeps across the slice corresponding to $\varphi=\pi$, much more frequency information is acquired via the convolution but a "missing corner" cannot be sensed. Both DD-CASSI (Fig. 3(b)) and CACCT (Fig. 3(d)) employ a 3D filter, resulting in a much wider bandwidth B_w in the chromatic frequency dimension, compared with Λ^{-1}. However, CACCT has the freedom to change the slice orientation, leading to a complete sensing coverage of the source spectrum with multiple DVP rotations. In contrast, the CASSIs just can alleviate the undersampling problem with multiple shots, but the "missing corner" remains. CTHIS in Fig.3(c) just can provide a minute extent of the spectrum spreading, as a result of its plain field stop.

2.3. Efficient reconstruction algorithm in spatial domain

We reconstruct the underlying data cube $\mathbf{f}$ from the compressed samples $\mathbf{g}$ in (6) by solving the following sparsity promoting problem

$$\hat{\mathbf{f}}=\arg\min_{\mathbf{f}}\frac{1}{2}\left\|\mathbf{PSf}-\mathbf{g}\right\|_2^2+\tau\left\|\mathbf{\Psi}^H\mathbf{f}\right\|_1+\gamma\left\|\mathbf{f}\right\|_{\text{TV}}, \quad (10)$$

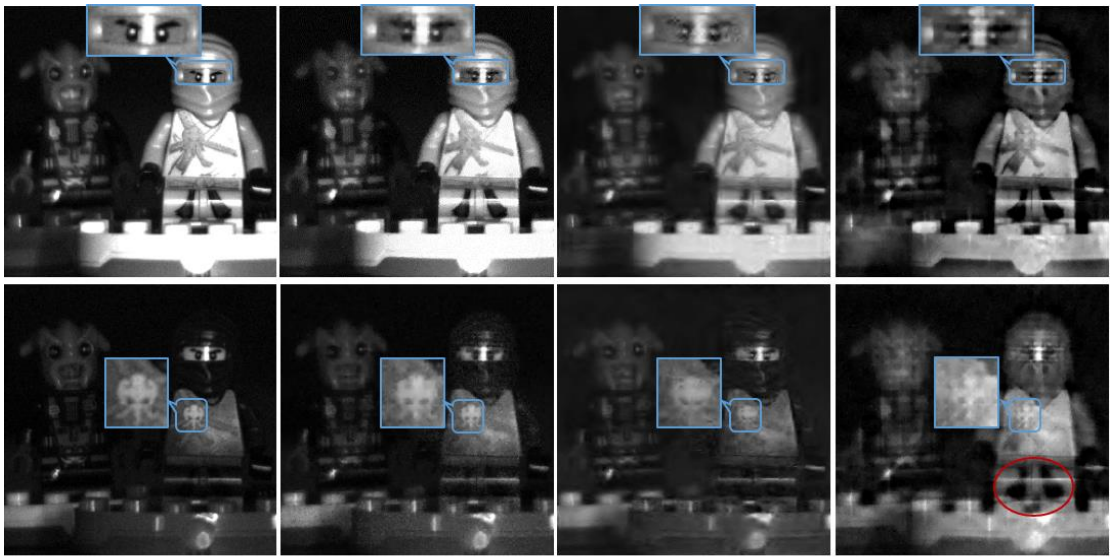

Fig. 5. Reconstructions of the 4th (top row) and the 24th (bottom row) band. (a) Original. (b) CACCT. (c) CASSI. (d) CTHIS. The 4th and 24th band correspond to the best and the worst reconstruction channel of CACCT, respectively. A closeup is presented in each panel and the red ellipse in the bottom right panel indicates the spectral crosstalk in the CTHIS reconstruction.

where $\mathbf{P} = \underbrace{\left[\mathbf{D}_{MN\times M'N'}\tilde{\mathbf{T}},\cdots,\mathbf{D}_{MN\times M'N'}\tilde{\mathbf{T}}\right]}_{L}$ stands for the process of modulation, decimation and summation of all L spectral channels, $\tilde{\mathbf{T}} = diag\left[\left\{\tilde{t}_{n'm'}\right\}_{n'm'=11}^{N'M'}\right]$ is the modulation matrix from (7), $\mathbf{S} = blockdiag\left[\mathbf{S}_1,\cdots,\mathbf{S}_L\right]$ implements the spatial shifts of channels with $\mathbf{S}_l = \mathbf{F}^H diag\left[\left\{e^{-j2\pi\left(m'\delta_x^{(l)}/M'+n'\delta_y^{(l)}/N'\right)}\right\}_{n'm'=11}^{N'M'}\right]\mathbf{F}$, $\mathbf{\Psi}$ is the 3D wavelet transform matrix, τ and γ are the regularization parameters balancing data fidelity and sparsity, and the TV norm is defined as $\|\mathbf{f}\|_{\mathrm{TV}} = \sum_{l=1}^{L}\sum_{i=1}^{M'N'}\|D_i f_l\|_2$ with D_i the first order finite difference operator for the i-th pixel in f_l.

Extension of (10) to the scenario of multi-frame measurements under different DVP rotation angles is straightforward. And this TV-L1-L2 minimization problem can be efficiently solved based on an iterative operator-splitting framework [13].

3. SIMULATIONS AND RESULTS

We numerically implemented CACCT to verify its imaging concept. First, a Xenon lamp and a visible monochromator were used to illuminate two targets and the spectral images were collected by a CCD camera. The resulting spectral data cube has a spatial resolution of 256×256 and 24 spectral bands. The spectral range is from 448nm to 664nm and the center wavelength $\lambda_c = 550$ nm. Then we simulated the chromo-projections according to the measurement model in (6). The code of CA is a pattern of a periodic uniformly redundant arrays [14], which has a wide and plain frequency spectrum. The dispersion coefficient of DVP was set to $k = \Delta_D/\Delta_\lambda = 2\Delta_M/\Delta_\lambda$, corresponding to a spatial shift of 1 pixel (also 2 mask elements) per spectral channel. Eight frames of FPA data were generated with the DVP rotation angles φ uniformly taken in the interval $[0, 2\pi]$.

The PSNRs of each recovered band are plotted in Fig. 4, where Gaussian noise with a SNR of 30dB was added. For comparison, eight CASSI frames were generated with a different random CA for each shot. As to CTHIS, we set a complete open aperture and obtain the results using the same algorithm as CACCT because the original CTHIS [11] only works with an overdetermined system. Every algorithm was terminated when the relative improvement was less than 1e-5. We also display the 4th and the 24th bands in Fig. 5 for visual inspection. One can notice the apparent blurry effect of CASSI due to its "missing corner" in the Fourier cube, which corresponds to high frequency details. While for CTHIS there exists severe cross-talk among spectral channels.

4. CONCLUSION

A binary photomask has been introduced into the chromotomography architecture to easily realize a high throughput spatio-spectral modulation. Via the Fourier domain analysis, we elaborated on the sampling advantage of this scheme over state-of-the-art geometries. The experimental results revealed the promising potential of CACCT for fast and accurate hyperspectral imaging applications.

5. REFERENCES

[1] N. Hagen, R. T. Kester, L. Gao, and T. S. Tkaczyk, "Snapshot advantage: a review of the light collection improvement for parallel high-dimensional measurement systems," *Optical Engineering,* vol. 51, no. 11, pp. 111702 2012.

[2] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive Coded Aperture Spectral Imaging: An Introduction," *IEEE Signal Processing Magazine,* vol. 31, no. 1, pp. 105-115, 2014.

[3] X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational Snapshot Multispectral Cameras: Toward dynamic capture of the spectral world," *IEEE Signal Processing Magazine,* vol. 33, no. 5, pp. 95-108, 2016.

[4] J. Tan, Y. Ma, H. Rueda, D. Baron, and G. R. Arce, "Compressive Hyperspectral Imaging via Approximate Message Passing," *IEEE Journal of Selected Topics in Signal Processing,* vol. 10, no. 2, pp. 389-401, 2016.

[5] I. Choi, D. S. Jeon, G. Nam, D. Gutierrez, and M. H. Kim, "High-quality hyperspectral reconstruction using a spectral prior," *ACM Trans. Graph.,* vol. 36, no. 6, pp. 1-13, 2017.

[6] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory,* vol. 52, no. 2, pp. 489-509, 2006.

[7] A. Wagadarikar, R. John, R. Willett, and D. J. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics,* vol. 47, no. 10, pp. B44-B51, 2008.

[8] H. Arguello, and G. R. Arce, "Colored Coded Aperture Design by Concentration of Measure in Compressive Spectral Imaging," *IEEE Transactions on Image Processing,* vol. 23, no. 4, pp. 1896-1908, 2014.

[9] X. Lin, Y. Liu, J. Wu, and Q. Dai, "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Trans. Graph.,* vol. 33, no. 6, pp. 1-11, 2014.

[10] H. Rueda, D. Lau, and G. R. Arce, "Multi-spectral compressive snapshot imaging using RGB image sensors," *Optics Express,* vol. 23, no. 9, pp. 12207-12221, 2015/05/04, 2015.

[11] J. M. Mooney, V. E. Vickers, M. An, and A. K. Brodzik, "High-throughput hyperspectral infrared camera," *J. Opt. Soc. Am. A,* vol. 14, no. 11, pp. 2951 - 2961, 1997.

[12] M. E. Gehm, R. John, R. Willett, T. Schultz, and D. Brady, "Single-shot compressive spectral imaging with a dual disperser architecture," *Opt. Express,* vol. 15, no. 21, pp. 14013-14027, 2007.

[13] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM J. Imag. Sci.,* vol. 1, no. 4, pp. 248-272, 2008.

[14] E. E. Fenimore, "Coded aperture imaging: predicted performance of uniformly redundant arrays," *Applied Optics,* vol. 17, no. 22, pp. 3562-3570, 1978.