



Compressive Multi-Mission Electro-Optical Sensor System

Le Li¹, Yi Yang¹, Juefei (Jeff) Zhou¹, Yongmei Li¹, Kevin F. Kelly², Anthony Giljum², Sanjeev Agarwal³, Son Nguyen³ ¹Kent Optronics, Inc. 40 Corporate Park Dr., Hopewell Junction, NY 12533 ²William Marsh Rice University, 6100 Main Street, Houston, TX 77005 ³U.S. Army CCDC C5ISR Center, 10221 Burbeck Road Fort Belvoir, VA 22060 UNITED STATES

ABSTRACT

We present the design, preliminary implementation, and initial validation results for a compressive multispectral and hyperspectral sensor (CMHS) system with spatial and spectral foveal imaging. The proposed compressive sensor system design is agile enough, enabling both multispectral (MS) sensing for wide area situational awareness and hyperspectral (HS) sensing for target recognition and identification. The agility of the sensor system enables configuration of the operational parameters such as spatial, spectral and temporal resolutions based on mission requirements. Such a multi-mission sensor provides the warfighters with the much needed capabilities to effectively employ the same sensor in dynamically challenging multiple-threats environments and obviates the need to carry multiple sensors for different threats and specific mission tasks. A breadboard compressive sensing (CS) multi-mission sensor has been assembled to demonstrate the feasibility via SLM patterning, data acquisition and reconstruction using the associated algorithm and software. The performance characteristics are presented along with the procedures involving data recording and reconstruction.

Keywords: hyperspectral video, compressed sensing, foveated imaging.

1.0 INTRODUCTION

Hyperspectral imaging is quickly growing in importance for a variety of security situations. However, the practical use of hyperspectral imaging is yet limited in several aspects. As a digital imaging technique, conventional hyperspectral microscopy is also in the regime of the Shannon-Nyquist sampling theorem. The bandwidth of the information that can be captured by the sensor, be stored and represented by digital data is determined by the number of sampling points. Because of the fact that hyperspectral data have 3-dimensions, 2 spatial dimensions and one additional dimension in spectral domain, in order to achieve a detailed spatial-spectral structure of the object, both the spatial resolution and the spectral resolution are required to be sufficiently high. Further in the case of moving object and/or moving sensor, we often require high frame rate video to effectively capture temporal features/characteristics of the scene. This 4-dimentinal data capture for hyperspectral video requires a very high sampling rate for the scene. As an illustration a modest 1Kx1K spatial resolution at standard frame rate of 30 Hz for 1024 band hyperspectral video would require a data rate of 30 Giga samples per second.

There is no off-the-shelf "hyperspectral sensor" to directly capture the 3D hyperspectral videos. In conventional methods, a one-dimensional or two-dimensional detector array needs to scan either in the spatial dimensions or in the spectral dimension to get the 3D hyperspectral data cube with differing, but significant compromise in temporal resolution. Another limitation exists in the sensitivity of the sensors. With the conventional scanning methods, each element of the detector array can only receive the signal photons from one pixel of the image and within one spectral band. The signal to noise ratio reduces significantly with increasing image resolution which requires longer integration time or a high sensitivity detector array in order to compensate. The first approach results in compromise in video frame rate while latter adds to the cost. This makes the whole data acquisition process time consuming and inefficient and



limiting in terms of size, weight, power and cost (SWaP-C).

Even though the data sampling requirement for hyperspectral video capture are large, it is important to note that this data is highly correlated (redundant) in space, time and spectrum. The true data requirement for the representation of this hyperspectral video data is much lower as is evidenced by the high compressibility of such video post-acquisition. The challenge is then to sample the scene more efficiently so that the redundancy in the measurements is minimized. Traditionally this has been achieved by reducing the sampling rate in space (spatial resolution), time (frame rate) or spectrum (band selection) depending on the need of the application or mission tasks. The sensors are thus developed for a specific mission profile such as a large format sensor for wide area reconnaissance, hyper spectral imager for HME detection, and high speed camera for target tracking applications. Compressive sensing provides an opportunity to design EO/IR sensors that hold out the promise to allow this trade between spectral, spatial, temporal resolution on the fly during data collection (sensing). When such a selection/adaptation can be made for selected region of interest (ROI) determined from the content of the scene during data collection, it results in a powerful sensor with multi-mission capabilities. The objective of this paper is to explore and evaluate such compressive hyperspectral video sensor design.

2.0 SENSOR DESIGN AND IMPLEMENTATION

The sensor design in this specific approach aims at the following key features:

- Flexible switching between MS and HS sensing modes from a single sensor system;
- Fast sampling rate and imagery display refresh at video rate or greater;
- Spatial and spectral fovea of region of interest (ROI) driven by the detection result(s);
- Reduced data volume and data transmission bandwidth;
- Small foot-print profile in size, weight, and power consumption as well as cost (SWaP-C).

Figure 1 schematically illustrates the primary design of the compressive multispectral and hyperspectral sensor (CMHS) system, where the dashed line enclosed area represents a spectrometer. In the figure, total internal reflection (TIR) prism couples the incoming light to and from the digital mirror device (DMD). The objective lens collects and focuses broadband scene light to the DMD for spatial coding. The coded light beam is then coupled into the entrance of the optical fiber in the spectrometer. The grating in the spectrometer disperses the coded broadband light into spectrally distributed sub-beams which are then detected by the one dimensional (1D) linear array. The detector output pins are registered to the corresponding wavelengths (λ s). The data from the 1D detector pins are then reconstructed to generate 4D HS imagery set (x, y, λ , t). Here, the optional complementary 2-D camera is for wide area monitoring and recording of the events during the experiment and is not used in data reconstruction.

In this design the DMD is used only to provide spatial coding of the scene. The output from DMD is collated and projected through the dispersion element (i.e., grating) to spread the spectrum over the 2048 pixels of the 1D detector array to effectively create 2048 single pixel cameras to capture 2048 band hyperspectral data cube.





Figure 1. Schematic illustration of the 1st design of the compressive multispectral and hyperspectral sensor (CMHS) system.

Figure 2 shows the photo picture of the breadboard CMHS system based on the design in Figure 1. The objective lens is the Asahi Pentax Super Takumar 50mm F1.4 lens. The DMD is Vialux V9501 from Digital Light Innovations. It has a resolution of 1920×1080 , which extends the capability of displaying STOne patterns with 1024×1024 pixels. The DMD package including the driver weight is 0.5 lb and power consumption is typically 0.2W. Other specifications for the DMD are shown in Table 1. The spectrometer is Ocean-FX from Ocean Optics, whose parameters are shown in Table 2.



Figure 2. Photo picture of the breadboard CMHS system based on the design in Figure 1.

Table 3 shows the design specs for the CMHS sensor. It is important to note that under this design, the light intensity impingent on each detector pixel is greatly reduced resulting in inherently lowered photon sensitivity and lower signal to noise ratio (SNR) which further limits the sensor frame rate and image quality. For example, as shown in later section, the upper limit frequency of the DMD is limited to 1,000 Hz due to the chosen spectrometer and the required integration time. This data rate limits the effective frame rate of the



hyperspectral data capture. Notwithstanding the photon sensitivity, the frame rate would be limited by the scan rate of the spectrometer (4500Hz). If a faster spectrometer is available, the scan rate in limited by the DMD frame rate of 17,850 Hz. A full resolution of 1024x1024 hyperspectral data with current setup would require 17.5 minutes per data cube at operating frequency of 1000Hz and 59 seconds for highest possible frame rate limited by DMD of 17,850Hz. Due to this limitation, the maximum spatial resolution for all results in this paper is limited to 128x128.

Table 1. Key specifications of the selected DMD (V9501)

V9501	Key Features	V9501
1920 x 1080	Layout	XY Grid
0.95	On Board Memory	8 GB DDR3
10.8	Binary Pattern Storage	31,086
17,857	Controller Interface	USB 3.0
+/-12	Window Coating	VIS/UV
	V9501 1920 x 1080 0.95 10.8 17,857 +/-12	V9501Key Features1920 x 1080Layout0.95On Board Memory10.8Binary Pattern Storage17,857Controller Interface+/-12Window Coating

Table 2. Key specifications of the Ocean-FX spectrometer

Item	Parameter				
Spectral coverage (nm)	350 - 850				
Detector format	1×2048				
Scan rate (Hz)	4,500				
Slit (µm)	5	50	100	200	600 (no slit)
Spectral resolution (nm)	0.8	2.2	4.1	8.1	24.4

Parameter	Target	Achieved	Remarks
Spectrum (nm)	400 - 850	400 - 850	Limited by Ocean-FX Spectrometer
Spatial resolution	1×1 to 1024×1024	1×1 to 1024×1024	Spatial foveal; limited by DMD memory & computation power of PC
Spectral resolution (nm)	0.8 - 450 nm	0.8 - 450 nm	Spectral foveal; limited by slit width and sensitivity of silicon detector
Frame rate (fps)	30	<< 30	
Field of view (deg)	12.6 × 12.6	12.6 × 12.6	Not optimized
Size (Inch)	12" × 8" ×8"	12"×8"×8"	Excluding PC
Weight (Lb)	~2	~2	Excluding PC
Power consumption (W)	≤ 55	≤ 55	Excluding PC

Table 3. Design specifications of the CMHS system



3.0 ALGORITHM AND SOFTWARE

By combining a spectrometer with a DMD in our hardware design, we have developed a highly agile hyperspectral imaging platform where spatial resolution, spectral sensitivity, and video rate can be tuned using a variety of patterns on the micromirrors in combination with differing reconstruction algorithms. The user can choose to apply traditional raster scanning, L2-based patterns such as Hadamard, random/permuted patterns for compressive imaging as in the original single-pixel camera, ROI specific dedicated foveation patterns, or compressive video specific patterns such as the Sum-To-One (STOne) transform.¹

By combining with the spectrometer as shown in Figure 1 each spectral bin becomes its own single-pixel camera that can be reconstructed independently or in various combinations of the other spectral channels. If all the bins are summed in total this reduces to grayscale imaging while other binning combinations can span from RGB, to multispectral, to the full hyperspectral resolution of the detector. The tradeoff of going to more spectral channels is a decrease in SNR within each spectral band and an increase in reconstruction time for the total data cube.

Given the flexibility of this platform, the spatially foveated image can be obtained from the approach as reported by D.B.Philips *et al*² who applied a spatially variable [log(z)] pattern to code DMD during data acquisition and to reconstruct the data to directly generate foveal image. However this approach is not optimal in regular L1 compressive video reconstruction and can in no way generate fast reconstructed lower resolution previews that are essential in applying optical flow for enhanced video reconstruction.³ Therefore we have developed a new approach of using foveated STOne patterns in reconstruction with full resolution STOne patterns in acquisition to achieve variable foveation anywhere in the image while still preserving the ability of high quality compressive video reconstruction.

4.0 **RESULTS AND DISCUSSIONS**

4.1 HS Static Image Acquisition and Reconstruction

HS static images have been acquired and reconstructed under the measurement conditions of 24,200 Lux illumination level, 300Hz DMD frame rate, 100% sampling rate, 64x64 pixel format. The data acquisition time for each frame is t = 64x64/300 = 13.6 sec. Figure 3(a) shows a bundle of roses as the object, which consists of one (1) plastic rose (in the middle) and two (2) real roses (at the two sides). Figure 4 shows the reconstructed HS images of the rose bundle at a series of wavelengths. Figure 3(b) shows the restored color picture from the hyperspectral image frames based on CIE 1931 color space.





Figure 3. (a) Photo picture of the rose bundle with plastic and real roses and (b) restored color picture from the hyperspectral image frames based on CIE 1931 color space



440 nm 460 nm 480 nm 500 nm 520 nm 540 nm 560 nm 580 nm 600 nm 620 nm 640 nm



660 nm 680 nm 700 nm 720 nm 740 nm 760 nm 780 nm 800 nm 820 nm 840 nm

Figure 4. Reconstructed HS images of the rose bundle (in Figure 3) at a series of wavelengths.

Figure 5 shows the comparison of the spectra of the real and plastic leaves measured directly by a spectrometer versus that from the reconstructed HS images in Figure 4. The plastic and real leaf is located at the coordinates of (13, 19) and (12, 25), respectively, as shown in Figure 5. It is seen that the reconstructed spectra agree in general with the ground truth spectra measured directly on the spectrometer, although there is some degree of discrepancy.



Figure 5. Comparison of the spectra of real and plastic leave measured directly by a spectrometer and plotted from the reconstructed HS images in Figure 4.

Similarly, Figure 6 shows the comparison of the spectra of the real and plastic rose measured directly by a spectrometer versus that from the reconstructed HS images in Figure 4. The plastic and real rose is located at the coordinates of (18, 13) and (30, 22), respectively, as shown in Figure 6.



Figure 6. Comparison of the spectra of real and plastic roses measured directly by a



spectrometer and plotted from the reconstructed HS images in Figure 4.

Noise in the reconstructed HS images can be reduced to increase the image contrast and quality. Figure 7 shows the reconstructed HS image with different spectral resolution, where the upper row has $\Delta\lambda = 10$ nm spectral resolution without noise reduction; the middle row has $\Delta\lambda = 1$ nm resolution with noise reduction; and lower row has the same $\Delta\lambda = 1$ nm spectral resolution but without noise reduction, respectively. The raw data were taken from the setup shown in Figure 2 under 24,200 Lux illumination, 500 Hz DMD frame rate and 50 µm slit size. Before noise reduction, the HS image quality with $\Delta\lambda = 1$ nm spectral resolution (lower row) is rather poor with much noticeable noise leading to significantly reduced contrast. We then applied the following noise reduction approach. At a selected λ_i , average its intensity $I(\lambda_i)$ over a spectral interval of +/- $\Delta\lambda$ around λ_i to yield an averaged intensity $\langle I \rangle (\lambda_i)$ which was then used to replace $I(\lambda_i)$. The processed HS images are shown in the middle row. Obviously, the noise suppressed HS image with $\Delta\lambda = 1$ nm spectral resolution have a much improved contrast and image quality which are comparable with the HS images with $\Delta\lambda = 10$ nm lower spectral resolution.



Figure 7. Reconstructed HS image with different spectral resolution, where the upper row has $\Delta\lambda$ = 10 nm spectral resolution without noise reduction; the middle row has $\Delta\lambda$ = 1 nm resolution with noise reduction; and lower row has the same $\Delta\lambda$ = 1 nm spectral resolution but without noise reduction, respectively.

4.2 Spatially Foveated Grayscale Video Image Acquisition and Reconstruction

Figure 8 shows the reconstructed HS image frames with spatial fovea, where the ROI is at the center. The images are reconstructed from the data acquired under the following condition, 81,000 Lux illumination, 500 Hz DMD frame rate, 100 μ m spectrometer slit opening. The image format is 128×128. The spatial fovea is obtained with the log(z) pattern.² The spectral resolution is 4.1 nm over the entire image frame under the 100 μ m spectrometer slit.





Figure 8. Reconstructed HS images with spatial tovea in the center.

The STOne transform has previously proved its utility in both static and compressive video imaging. In addition to producing an L1 reconstruction, it is also capable of generating a fast L2 inverse transform at varying resolutions with the same patterns. As seen in section 4.3, these STOne previews can also be used in conjunction with optical flow for greatly improved video reconstruction. Here we demonstrate that the multiple length scale properties of STOne patterns derived through nested embedding can also be used to generate on the fly foveation in any region of the image without having to change the original acquisition patterns. Simply by replacing the reconstruction matrix in the solver from the original full high resolution matrix with a down sampled low resolution everywhere but the region of interest which is still at full resolution as illustrated in Figure 9. A foveated reconstructed final image is achieved as illustrated in Figure 10. By performing foveated reconstruction we can make the tail numbers on the test image much more legible while at the same time reducing the reconstruction time in comparison to the non-foveated image.



Figure 9. A schematic illustrating on the left the up-sampling/down-sampling kernel property inherent in the STOne transform through nested embedding. On the right is a schematic illustrating foveated reconstruction by combining the low and high resolution patterns input into the solver while still utilizing the full high resolution patterns in acquisition.

In addition to performing foveated L1 reconstruction with the STOne patterns, we also developed a novel method of generating L2 previews from these measurements that tends to lead to a higher SSIM than the previous method proposed by Goldstein et al.¹ when compared to a ground truth image. In Goldstein et al.'s method¹, the measurement sequence is first unpermuted. Then the unpermuted measurement sequence is



down sampled in the following way: for each group of (full resolution) / (preview resolution) measurements, all values in the group are replaced by the average of the possessed measurements within that group. Down sampling the measurements in this way has the effect of approximating measurements taken by a lower-resolution STOne matrix. One can then simply perform a fast STOne transform on the down sampled measurements to acquire a lower-resolution preview. However, with fewer than 100% of the measurements, the averages will stray slightly from the correct value, leading to noisy images.





Therefore we propose a second approach to generating previews to improve the resulting image quality. Instead of replacing all values in the group with the average of the possessed measurements, one only replaces the zero-values with the average. In this way, there is no loss of the high-resolution information contained within the possessed values, and instead simply fill the measurement sequence with interpolated low-resolution information. We can then perform a fast STOne transform on this interpolated measurement sequence to obtain a high-resolution preview as shown in Figure 11.





Figure 11. Test images of the new L2 STOne method as a function of compression in comparison to the old method.

The effect of this methods in terms of quantitative image quality and time can be seen in Figure 12 where our proposed L2 preview method increases linearly with the number of measurements until it reaches an SSIM of 1 at 100%, when there are no more zeros in the sequence that need to be filled. Conversely, if one has the minimum number of measurements required to generate a preview, that is, each group contains only one measured value, then the proposed method and Goldstein et al.'s method produce the same results.





The plot in Figure 12 is brief summary of a few of the options available in generating reconstructed images with this unbiased, post-acquisition completely adaptable approach that allows foveated reconstruction within the entire field-of-view of the DMD and simultaneously allowing for optimal compressive video capture. With one set of measurements using the full resolution STOne patterns on the DMD, the user can tradeoff between reconstruction speed and fidelity depending on their requirements. While the foveated reconstruction offers a modest decrease in reconstruction time, when considering a hyperspectral data cube composed of hundreds of spectral channels this is a substantial gain. We also believe that this is likely not the most optimal way to exploit this foveation and offer as a proof of principle while other methods are being researched.



4.3 HS Video Acquisition and Reconstruction

HS Video has been acquired and reconstructed. Figure 13 shows the photo picture of the moving target which is a toy car on a linear translation stage. During the measurement, the car moving speed is ~0.5 inches/second. Other experiment parameters are: 81,000 Lux light illumination, 500Hz DMD frame rate, and 12.5% sampling rate. The video image pixel format is 128*128. Thus, the data acquisition time t = 128*128/500 = 33 sec. Figure 13 (b)-(e) show the reconstructed video images at the wavelength 450 nm, 530 nm, 650 nm, and 710 nm, respectively.



Figure 13. (a) Photo picture of the target, and (b)-(e) the reconstructed video images at wavelength 450 nm, 530 nm, 650 nm, and 710 nm, respectively.

4.4 Foveated Video Reconstruction

Figure 14, illustrate the application of this foveated STOne to compressive video reconstruction of a moving scene. A video test sequence of sliding Air Force test pattern was generated after which a data set was obtained by applying the STOne patterns temporally over the sequence. Full resolution in the foveated region was kept at 128×128 spatial equivalent while the remaining image was down sampled to either 16×16 or 32×32 resolution blocks which were then restored to the sample spatial equivalent by nearest neighbor resizing. Depending on the chosen down sampling and percent compression one can adjust the image quality versus the reconstruction speed.

4.5 Data Acquisition Time

Data acquisition and reconstruction time is another important topic to discuss. Data acquisition time "t" is determined by $t = N \times N/f$ where N is the image format and f is the data acquisition frame rate. Currently, f is limited to 1,000 Hz in our system which uses the Ocean-FX spectrometer. For 128× 128 image format and 1,000 Hz DMD frame rate, $t = 128 \times 128/1000 = 16$ sec. If using single pixel detector whose response time is in nanosecond, we can use f=17,857 Hz for the DMD. Thus we have $t = 128 \times 128/17857 = 0.92$ sec.





4.86 seconds/frame

5.07 seconds/frame

Figure 9. Comparison of down sampling resolution and percent measurement compression in the ability to reconstruct foveated video at full resolution (highlighted by the red box).

4.10 seconds/frame

For the sensor system as shown Figure 2, different DMD frame rate yields different imagery quality. Since the 1-D detector array in the spectrometer is synchronized with the DMD, faster DMD frame rate forces shorter detector integration time. The raw data were taken under 48,000 Lux illumination. For the Ocean-FX Spectrometer with 50 µm slit for high spectral resolution, the reconstructed image can be barely recognized under 1,000 Hz DMD frame rate. At this condition, 500 Hz DMD rate yields pretty good result, as shown in Figure 15. If the slit width is 100 µm, even under 1,000 Hz DMD rate, the reconstructed image is still good. When the DMD rate increases to 1,500 Hz, the image is barely recognized. This result indicates that the image quality gets worse for higher spectral resolution (i.e., smaller slit width), given the same illumination level. The image quality is also worse with the increase of the DMD frame rate which limits the data acquisition time.



Figure 10. Reconstructed images under different DMD frame rate and spectrometer slit width.



4.6 Computation Time for Reconstruction

Given the current computer power of 4 cores, 3.2 GHz clock frequency, and 32 Gb memory, we have clocked the reconstruction time for generating grayscale video. The results are shown in Table 4, which have been obtained with the parameters of 128×128 pixel format, mu = 10 (the parameter imbedded in the STOne solver), and 256 shift. The shortest time is 2.33 sec under 75% sampling rate which involves 16 frames of image per video set. It is concluded that under the current situation, it is difficult to generate real time video at 30 fps frame rate. More powerful algorithm as well as computer are then necessary.

Sampling rate (%)	Frame	Binning time (s)	Reconstruction	Converting time (s)	Elapsed time (s)
Tute (70)	Trume	time (b)			
75.0%	16	0.0113	1.8	0.0040	2.33
62.5%	24	0.006755	3.2	0.0030	3.84
50.0%	32	0.012999	4.8	0.0070	6.13
43.1%	36	0.008258	7.4	0.0046	8.4
37.5%	40	0.008629	10.0	0.0049	11.2
12.5%	56	0.014215	30.5	0.0117	32.14
	Frames in			Input data to	Total time
Note	each video set		Per video set	image data	per video set

Table 4. Computer clocked computation time in reconstructing a set of grayscale video

5.0 CONCLUDING REMARKS

A breadboard compressive sensing (CS) sensor capable of hyperspectral video has been assembled to demonstrate the feasibility of multi-mission capability. Self-similarity of STOne patterns enables efficient reconstruction of image and video data at different spatial and temporal resolutions. Further this can be done at the time of reconstruction without the change of the DMD patterns. Also the STOne patterns enables ROI foveated image and video which when coupled with fast low resolution previews can allow adaptive sampling driven by the scene content. Significant challenges however remain to fulfil the promise of true multi-mission EO/IR sensor system providing an adaptability to trade between spectral, spatial, temporal resolution on the fly during data collection. Video rates are limited by available hardware. Reconstruction algorithms that are efficient in terms of sampling are computationally expensive. Better hardware and sensing architectures are needed to improve optical efficiency.

6.0 ACKNOWLEDGEMENT

This work is supported by the contract W909MY-17-C-0006 under US Army Small Business Innovation Research (SBIR) Program. The support from the US Army CCDC C5ISR Center and US Army SBIR Program is greatly appreciated.

REFERENCES

- [1] Goldstein, T., Xu, L., Kelly, K.F. and Baraniuk, R., 2015. The stone transform: Multi-resolution image enhancement and compressive video. *IEEE Transactions on Image Processing*, 24(12), pp.5581-5593.
- [2] D.B.Philips et al., 2017. Adaptive foveated single-pixel imaging with dynamic super-sampling, *Science*

Advances 3 e1601782.

[3] Sankaranarayanan, A.C., Studer, C. and Baraniuk, R.G., 2012, April. CS-MUVI: Video compressive sensing for spatial-multiplexing cameras. In 2012 IEEE International Conference on Computational Photography (ICCP) (pp. 1-10).