Multi-spectral compressive snapshot imaging using RGB image sensors

Hoover Rueda,1 Daniel Lau,2 and Gonzalo R. Arce1,*

1 Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, 19716, USA
2 Department of Electrical Engineering, University of Kentucky, Lexington, KY, 40507, USA

*arce@udel.edu

Abstract: Compressive sensing is a powerful sensing and reconstruction framework for recovering high dimensional signals with only a handful of observations and for spectral imaging, compressive sensing offers a novel method of multispectral imaging. Specifically, the coded aperture snapshot spectral imager (CASSI) system has been demonstrated to produce multi-spectral data cubes color images from a single snapshot taken by a monochrome image sensor. In this paper, we expand the theoretical framework of CASSI to include the spectral sensitivity of the image sensor pixels to account for color and then investigate the impact on image quality using either a traditional color image sensor that spatially multiplexes red, green, and blue light filters or a novel Foveon image sensor which stacks red, green, and blue pixels on top of one another.

© 2015 Optical Society of America

OCIS codes: (110.4234) Multispectral and hyperspectral imaging; (110.0110) Imaging systems; (170.1630) Coded aperture imaging; (230.7408) Wavelength filtering devices.

References and links

1. Introduction

Multispectral imaging in the visible light spectrum refers to color cameras that record more than three primaries [1]. Composed of a monochrome image sensor taking a series of pictures through an array of narrowband interference filters, these cameras are especially common in the fine arts since full spectral recordings avoid issues associated with metamerism – the apparent change in color of an object caused by changes in ambient light. A more recent approach to multispectral imaging, the coded aperture snapshot spectral imager (CASSI) system [2], is modeled after a traditional spectrophotometer and records all wavelengths simultaneously. Here a ray of light emanating from the scene is passed through a series of optical elements culminating in a prism that spreads that single ray of light over a lateral sequence of sensor pixels, in a wavelength dependent manner, just like a spectrophotometer would spread light across a linear CCD array. Allowing for multiple rays of light simultaneously incident upon the sensor means that the camera can resolve a two dimensional image; however, because the spreading of light rays across lateral sequences of pixels means a single pixel will collect light from multiple sources, compressive sensing techniques must be employed to decouple the spectral profiles of neighboring pixels [3, 4].

Perhaps simplifying the problem somewhat, color sensors may be employed such that each pixel will only record those wavelengths of light falling within the range of its corresponding color filter, either red, green, or blue, but such sensors lose spatial resolution since only one...
quarter of the pixels are red, one quarter of the pixels blue, and the rest green [5]. Noting that photons of light penetrate the silicon in a wavelength dependent manner before being absorbed by the crystal lattice, Foveon manufactures a color image sensor that effectively layers red, green, and blue pixels on top of one another such that a 1 megapixel sensor records 1 million unique red, 1 million unique blue, and 1 million unique green values [6, 7]. Such a stacked color image sensor, like Foveon’s, could provide a significant performance improvement for spectral recordings using CASSI, and in this paper, we investigate that improvement in both theoretical and experimental venues. Specifically, we will introduce a new sensing model that incorporates the spectral response of the image sensor, we will present and analyze its forward sensing operator, and we will show simulated and experimental data to validate the findings. In particular, the Foveon-based system attains around a 4 dB improvement over the traditional monochrome system, and the color CCDs that employs a Bayer filter array attains up to 3 dB improvement, regardless of the number of snapshots.

2. Camera model

In the CASSI system, the front end lens system effectively creates a light field in front of the image sensor where all light rays travel in parallel to the Z-axis as depicted in Fig. 1(a) such that \( f(x, y, \lambda) \) represents the light incident upon the sensor at 2-D coordinate \((x, y)\) and having wavelength \( \lambda \). Accumulating all of the light incident upon a discrete pixel of the sensor with row and column coordinate \([m, n]\) results in the discrete-space, continuous-frequency signal \( f_{m,n}(\lambda) \). If we now define \( \chi_{m,n}(\lambda) \) as the spectral sensitivity of the sensor for that same discrete pixel, then we can define the gray or monochrome pixel, \( g[m, n] \), of the resulting digital image according to

\[
g[m, n] = \int_{\lambda} \chi_{m,n}(\lambda) f_{m,n}(\lambda) d\lambda.
\]  

(1)

One would expect \( \chi_{m,n}(\lambda) \) to be either constant for all pixels and absorb light in the visible light range from 450 to 670 nm or selectively absorb red, green, or blue light in a traditional Bayer pattern as illustrated in Fig. 1(b).

As light approaches the sensor, CASSI employs a coded aperture or diaphragm to block all but a single ray of light from reaching the sensor as indicated in Fig. 1(c). We can account for this aperture in our model by means of the binary modulation function \( T_{m,n} \in [0, 1] \). In order to separate the individual wavelengths of the incoming signal, the CASSI system places a dispersive element in front of the image sensor such that \( f_{m,n}(\lambda) \) is divided into discrete bands.
spread across \( K \) consecutive pixels along the lateral or \( n \) dimension in a wavelength dependent manner, indicated by \( S(\lambda) \). We depict this discrete-space behavior in Fig. 1(d) such that,

\[
g[m, n] = \sum_k \chi_{m, n, k} T_{m, n+k} f_{m, n+k},
\]

where \( k = -\frac{K}{2}, \ldots, 0, \ldots, \frac{K}{2} \) indexes the pixels affected by the spreading of \( f_{m, n}(\lambda) \) as a result of \( S(\lambda) \). At this point, we note, as indicated in Fig. 1(e), that we could use the coded aperture to allow light rays sufficiently far apart so that there is no overlap of the corresponding light spectrums; however, doing so means only a fraction of the incoming light field can be recorded. We could, therefore, take a series of snapshots where the open pixels in the coded aperture sweep across the focal plane such that, in this case, we would need to take exactly \( K \) spectrums; however, doing so means only a fraction of the incoming light field can be recorded.

In order to establish a matrix representation of CASSI consistent with CS, \( f \) is defined as a column vector containing \( f_{m, n, k} \) for all spectral components \( k \) and all pixels \( m \) and \( n \) ordered first by row, then column, and finally by wavelength. Assuming we have a coded aperture with \( M \times N \) pixels, \( f \) will be \( MNK \) in length. We now define \( g^\ell \) as a column vector holding all of the recorded pixel values, \( g[m, n] \), from the monochrome image sensor for the \( \ell \)th image in our sequence of \( L \) images. Noting the laterally dispersion of light, the length of \( g^\ell \) is \( M(N + K - 1) \) such that all light is accounted for on the sensor. So while the coded aperture has an \( M \times N \) pixel array, the sensor has an extra set of \( K \) pixel columns. Assuming that \( f \) stays constant over the \( L \) snapshots, we can relate \( g^\ell \) to \( f \) according to,

\[
g^\ell = XPT^\ell f = H^\ell f,
\]

where \( H^\ell \) represents the combined effects of the sensor’s pixel-wise spectral sensitivity in matrix form \( (X) \), the dispersion effect in matrix form \( (P) \), and the \( \ell \)th coded aperture \( (T^\ell) \).

As an illustration of the system transfer function, Fig. 2 depicts the non-zero entries for single \( H^\ell \) for an \( M = N = 4 \) camera sensor used to record a \( K = 6 \) multi-spectral image where the matrix width \( Q = MNK \) is the total number of scalar measurements \( (K \) values per pixel) while \( V = M(N + K - 1) \) is the total number of observed pixels. In the case of Fig. 2 (top), the system transfer function is for a typical camera where none of the incoming pixels are blocked by the coded aperture nor is there any dispersion. Hence, there is no need to decompose the incoming light into spectral bins. For Fig. 2 (center), the system transfer function divides the incoming light into \( K \) spectral bins, but without any dispersion, the resulting observation \( g^\ell \) is the same as Fig. 2 (top). In Fig. 2 (bottom), the depicted transfer function models dispersion by off-setting the diagonal structure of the previous matrices vertically.

Because we have arranged vector \( f \) in row-major order, a lateral shift of light one pixel to the right on the sensor corresponds to a vertical shift of \( M \) pixels. As we progress through all \( K \) pixels on the sensor, we end up with a transform matrix that has \( M(K - 1) \) additional rows compared to Fig. 2 (center). Now having modeled the collection of incoming light rays on a per frame basis, we can now assemble the complete set of all \( L \) snapshots into a single transform by concatenating each \( g^\ell \) vector end to end to create the \( LM(N + K - 1) \times 1 \) vector \( g = (g^0)^T, \ldots, (g^{L-1})^T)^T \), which can be modeled in Eq. (3) by similarly concatenating the
matrices $H^c$ on top of one another to form the single matrix $H = [(H^0)^T, \ldots, (H^{L-1})^T]^T$ and the transformation $g = Hf$.

Having now modeled the spectral imaging system as a single vector to single vector transform, CS requires us to find a linear transform of $f$ from its current spatial-spectral $(m,n,k)$ space into some other $MNK$-dimensional space with basis $\Psi$ such that $\theta = \Psi^Tf$ is the vector coordinate of $f$ in the new space. For this process to be worth-while, $f$ needs to be $S$-sparse where only a small subset, $S \ll (M \cdot N \cdot K)$, of the basis vectors $\Psi$ can largely reconstruct $f$ with little or no distortion [13]. Formally, $f$ is $S$-sparse or has sparsity $S$ in a basis $\Psi$ if $\|\theta\|_0 = S$, where $\|\theta\|_0$ denotes the $\ell_0$-norm, which simply counts the number of nonzero entries in the vector.

Traditional basis functions encompass the wavelet ($\Psi_{W2D}$) and discrete cosine transforms ($\Psi_{DCT}$) as well as pre-trained dictionaries [12, 14, 15], but it is also possible to simultaneously exploit the sparsity properties of a multidimensional signal along each of its dimensions [16–18]. Particularly for spectral images, we will use the basis expressed as the Kronecker product between $\Psi_{W2D}$ and $\Psi_{DCT}$ where $\Psi_{DCT}$ is used to make sparse the $k$-axis while $\Psi_{W2D}$ makes sparse the $(m,n)$ coordinates.

Having our $\Psi$, CS allows $f$ to be recovered from a single $g^c$ vector since the length of $g^c$, $V$, is greater than $S\log(M \cdot N \cdot K)$ which is much smaller than $M \cdot N \cdot K$ [12]. Having multiple
g' will, therefore, only improve the result since it corresponds to increasing the dimensionality, S, of the sub-space, in which we are projecting f. To finally estimate f, we need to solve the minimization problem,

\[ \tilde{f} = \Psi^T \left( \arg \min_{\theta'} \| g - H \Psi \theta' \|_2 + \tau \| \theta' \|_1 \right), \]  

(4)

where \( \theta' \) is an S-sparse representation of f on the basis \( \Psi \), and \( \tau \) is a regularization constant. To solve Eq. (4), different methodologies and frameworks have been proposed in the literature [19–21]. In particular, the gradient projection for sparse reconstruction (GPSR) algorithm has been proposed in [21]. GPSR finds a sparse solution to the non-linear, unconstrained minimization problem, where the first term minimizes the euclidean or \( \ell_2 \) distance between the detector measurements, g, and the contribution from the estimate \( \theta' \); while the second term is a penalty term that encourages sparsity of the reconstruction in the basis domain and controls the extent to which piecewise smooth estimates are favored. In this formulation, \( \tau \) is a tuning parameter for the penalty term and higher values yield sparser estimates. An estimate for the data cube, \( \tilde{f} \), for a chosen value of \( \tau \), is found by an iterative optimization procedure. This method searches for a data cube estimate with a sparse representation in the chosen basis, i.e. the coefficients in \( \theta' \) are mostly zeros numerically.

4. Color sensing model

As a means of extending the CS reconstruction to allow for color sensitive sensors, we note that all prior works assumed a monochrome sensor model [2] where \( \chi_{m,n}(\lambda) = \gamma(\lambda) \) was constant for all \((m, n)^{th}\) pixels, as depicted in Fig. 3(left). In this paper, we intend to employ two color sensitive sensors: a traditional filter-based RGB Bayer sensor [5] shown in Fig. 3(center), and an absorption-based stacked color sensor developed by Foveon [6], presented in Fig. 3(right). Color sensors are employed such that each sensor pixel records those wavelengths falling within small spectral ranges corresponding to either the blue, green or red spectra. Figure 3 summarizes the 3 different sensors along with their structure and spectral sensitivity curve in the visible window of the electromagnetic spectrum.

Fig. 3. FPA detector architectures along with their traditional spectral responses. (left) Monochromatic (Source: Stingray F-033B), (center) RGB-Bayer (Source: TMC) and (right) Foveon (Source: Foveon Quattro sensor).
Formally, each \((m,n)\) pixel from an RGB-Bayer, or Foveon sensor exhibits a filter function given by:

\[
\chi_{n,m}(\lambda) = \begin{cases} 
  b_B(\lambda), & \forall n,m \in \Omega_b \\
  g_B(\lambda), & \forall n,m \in \Omega_g \\
  r_B(\lambda), & \forall n,m \in \Omega_r 
\end{cases} 
\]

\[
\chi_{n,m}(\lambda) = \begin{cases} 
  b_F(\lambda), & \forall n,m \\
  g_F(\lambda), & \forall n,m \\
  r_F(\lambda), & \forall n,m, 
\end{cases} 
\]

where \(\Omega_b, \Omega_g\) and \(\Omega_r\) represent the sets of blue, green and red pixels of the Bayer color pattern. Notice that the Foveon sensor entails the absorption of the total light incident in every pixel, whereas the Bayer sensor filters out wavelengths lying outside the spectral-sensitivity of the pixel. Remark that the relative response of the three layers of the Foveon sensor \(b_F(\lambda), g_F(\lambda), r_F(\lambda)\) exhibit a broader sensitivity compared to the corresponding RGB filters \(b_B(\lambda), g_B(\lambda), r_B(\lambda)\) from the Bayer sensor. Due to the blue on-top-of green on-top-of red layer stack of the Foveon sensor, the blue layer absorbs part of the green and red spectrum, and the green layer will absorb part of the red spectrum as well.

In order to modify our transform functions \(H^I\), to include variations in \(X\), we will essentially stack three monochrome transforms, as depicted in Fig. 4, on top of one another with the first transform weighted for blue, the second weighted for green, and the third weighted for red, according to \(b_F(\lambda), g_F(\lambda), r_F(\lambda)\) for the Foveon, and \(b_B(\lambda), g_B(\lambda), r_B(\lambda)\) for the Bayer sensor. Due to the stacking process, and in order to be fair in the comparison between the color sensors and the monochrome sensor, 3-monochrome snapshots will be compared against 1-color snapshot. Therefore the compression ratio of the system will be \(K : 3L\). Fig. 4 shows our proposed color transform models for both Fig 4 (left) a spatially multiplexed Bayer sensor, versus Fig 4 (right) the stacked-color Foveon sensor. It can be observed that both matrices have the same size, but the Bayer sensor entails completely blanked rows due to the black holes on each Bayer layer compared to the Foveon sensor, which absorbs all the blue, green, and red spectrum in every \((m,n)\) sensor pixel. Further, the Bayer pixel sensitivities can be seen as, \(b_B(\lambda)\) filters out the last portion (4 out of 6 bands) of the spectrum, \(g_B(\lambda)\) filters out the first (1 out of 6 bands) and the last portion (1 out of 6 bands) of the spectrum, and \(r_B(\lambda)\) filters out the first portion (4 out of 6 bands) of the spectrum. Similarly, the Foveon pixel absorption is modeled...

Fig. 4. Non-zero elements of the system transfer functions, \(H^I\), for a \(4 \times 4 \times 6\) data cube, using (left) RGB-Bayer and (right) Foveon sensors.
as, \( b_F(\lambda) \) absorbs the first 3/4 of the spectrum in average, due to the broader spectral sensitivity, \( g_F(\lambda) \) absorbs the central portion of the spectrum and part of the last portion, and \( r_F(\lambda) \) just the last portion of the spectrum. Aside from this change in \( \mathbf{H} \), deriving \( \mathbf{f} \) from \( \mathbf{g} \) is largely unchanged from the CS process used for monochrome sensors.

5. Simulations and experiments

In this section, we will first evaluate the proposed method through the simulation of the process using the two color-checker patterns shown in Fig. 5(a) and Fig. 6(a), using discrete versions of the FPA spectral responses described in Fig. 3. Afterward, we will collect real data with the DMD-based CASSI imaging system built in our lab [22] using the input target depicted in Fig. 10(a), and commercially available RGB color filters.

5.1. Simulations

In order to study the effect of the different filtering functions entailed by the color sensors under study, a set of compressive measurements is first simulated using the forward model in Eq. (3), varying the \( \mathbf{X} \) function for a monochrome, an RGB Bayer, and a Foveon sensor according to Fig. 3. These measurements are constructed employing two test spectral data cubes acquired.
by illuminating the color-checker target shown in Fig. 5(a) using a broadband Xenon lamp as the light source, and a visible monochromator spanning the spectral range between 450nm and 670nm. The test data cubes, $\mathbf{F}$, have $256 \times 256$ pixels of spatial resolution and $K = 24$ spectral bands. Although the coded apertures are a key optical element in the compressive spectral imaging systems, and their design has a great impact on the attained reconstructions quality [2,12,18,23], in this paper we will use simple random realizations of a Bernoulli random variable with parameter $p = 0.5$; that is, 50% of the coded aperture entries will be 0 and the rest will be 1. The latter is justified because this paper wants to focus on the impact of color sensors on the quality of the reconstructions, but independently of the effect of the coded apertures. The proper design of the coded apertures will improve even more the results presented in this paper. Although the color sensors are simulated through the use of a monochrome sensor, they
take into account the aliasing effects resulting of demosaicing in the Bayer sensor, and the real absorption/transmission curves of the Foveon sensor. Moreover, the three compressive spectral imaging systems compared in the manuscript are assumed to be ideal in terms of noise, due to the different sources producing it: exposure time, kind of sensor, photon count, quantization, heat, etc. Analysis of the noise characteristics of the sensors lies out of the scope of this paper.

Figure 5(b) shows the reconstructions PSNR of the first test data cube, averaged over all wavebands, as function of the compression ratio ($R = L/K$). Notice that $R = 3/24 = 1/8$ represents a single snapshot for the color sensors, and 3 for the monochrome sensor. In Fig. 5(b) it can be seen that the PSNR attained in the reconstructions with the RGB-Bayer sensor improves up to 3dB over the monochrome sensor overall. Similarly, the Foveon sensor improves an extra dB over the RGB-Bayer. For a visual comparison, Fig. 5(c) depicts the $K = 24$ reconstructed data cube by the three sensors mapped to a pseudo-color map when $R = 1/8$ and $R = 1/2$. The improvement in the spatial quality can easily be observed as $L$ increases (from top to bottom) or as the sensor changes from monochrome to Bayer to Foveon (from left to right).

The second test data cube shown in Fig. 6 consisted of a modified version of the first data cube, where white diagonal lines were added to the top region of the color-checker and equally spaced white bars were added to the bottom region. Notice that the left portion of the bottom bars are two pixels wide, whereas the right portion correspond to one pixel width bars. This test wanted to show the reliability of the reconstructions for high-frequency changes, as well as the impact of the demosaicing task performed by the Bayer reconstruction. Figure 6(b) presents the averaged PSNR of the reconstructions, where the Bayer sensor overcomes the monochrome sensor by about 1.5 dB, and the Foveon maintains an advantage of about 0.5 dB over the Bayer sensor reconstructions overall. Remark that the PSNR is dramatically affected (about 10 dB less) for the three sensors compared with Fig. 5(b), when the scene presents high-frequencies changes. This issue arises in most of the algorithms that solve Eq. (4). It can be noticed in Fig. 6(c) that the diagonals in the top of the scene along with the bars in the bottom left are recovered somehow, but the bars in the bottom right are barely reconstructed. The zoomed versions show large color aliasing artifacts in the reconstruction attained with Bayer sensor unlike Foveon sensor, which presents artifacts, but not color aliasing.

To analyze the reconstruction results at wavelength level, Fig. 7 presents the absolute error of the $4^{th}$, $8^{th}$, $12^{th}$, $16^{th}$, $20^{th}$ and $24^{th}$ reconstructed spectral wavebands when a single snapshot
is captured ($R = 1/8$). It can be observed that the error of the reconstructions attained with the color sensors is less than that attained with the monochrome sensor along the whole reconstructed spectrum. Remark that the major errors are concentrated along the high-frequencies of the scene, thus confirming the downside of the estimation algorithms.

5.2. Experiments

The testbed proposed in [22] formed by an imaging arm and an integration arm, is used to obtain the experimental results in this paper. In this system, the imaging arm is composed of an objective lens and a DLP spatial light modulator (DMD); the integration arm is composed of a relay lens, a dispersive element, and a CCD camera, as shown in Fig. 8. To follow the mathematical model, a target scene is illuminated with its reflected light captured by the objective lens and focused onto the mirrors of the spatial light modulator image plane, which plays the role of the coded aperture. When properly aligned, the mirrors of the modulator reflect light into the integration arm, which relays light through a second lens and then through the prism, such that the dispersed field focuses in the CCD image plane, which imposes its spectral filtering and integration based on its wavelength sensitivity. The DMD coded aperture patterns are the same as described in the simulations section, random realizations of $256 \times 256$ pixels with 50% transmittance. The prism used in the testbed is a non-linear double Amici prism, which disperses the visible spectrum between 450 - 670 nm onto 24 sensor pixels. The 24 spectral channels have central wavelengths 455, 460, 465, 470, 476, 482, 488, 494, 500, 506, 512, 519, 527, 536, 545, 555, 565, 576, 587, 599, 612, 626, 643, and 660 nanometers. Notice that the bandwidth of the spectral channels is non-uniform due to the non-linearity of the prism. The spectral resolution of the testbed is 5 nm for the short wavelengths, and about 15 nm for the long wavelengths. Therefore, the attained reconstructions with the testbed will exhibit $256 \times 256$ pixels of spatial resolution and 24 spectral channels of spectral resolution.

![Fig. 8. DMD-based CASSI testbed setup used in the experiments. The illuminated target scene is imaged onto the image plane of the DMD which plays the role of the coded aperture. Subsequently, the relay lens transmits the coded light through the Amici prism which disperses it onto the image plane of the CCD array.](image-url)
The three CCD sensors under study are emulated based on their system transfer functions, depicted in Figs. 2 and 4, while using of a single, monochromatic sensor (Stingray F-033B) with pixel size of $\Delta_d = 9.9\,\mu m$ and 8 bits of dynamic range. The emulation of the monochrome sensor is straightforward, but taking 3-times more snapshots per each snapshot of the color sensors. To emulate the Bayer sensor, the monochrome sensor captures 3 snapshots imaged through either a red, green, or blue glass plate and then selecting the $(m, n)$ positions corresponding to $\Omega_b, \Omega_g, \Omega_r$ in Eq. (5), and blocking the remaining. Similarly, the Foveon sensor is emulated as in the Bayer case, but the selection procedure is not required since every pixel captures all the red, green, and blue channels. Figure 9 depicts the emulated compressive measurements attained by each sensor, and their corresponding zoomed versions in order to appreciate their differences.

![Fig. 9. Real measurements for a single snapshot of the monochrome, the Bayer and the Foveon sensor. Zoomed versions highlight the sensors differences.](image)

To evaluate the CS capabilities of each sensor, the target scene depicted in Fig. 10(a) is imaged through the system. Fig. 10(b) depicts the reconstructed data cubes mapped to a pseudo-color profile when $R = 1/8$ and $R = 1/2$. Without a perfect image to compare for calculating PSNR, visual evaluations confirm that the color sensors improve upon the monochrome, with the Foveon and Bayer sensors, exceptionally so. As a means of quantitatively evaluating the various sensors, the spectral signatures of two different spatial points, taken from the target scene (a point from the green chest, and a red point from the helmet), are plotted in Fig. 10(c) against a reference signature measured using a commercially available spectrometer assumed as the ground truth. Compared to a traditional monochrome sensor, both the Bayer and Foveon spectra offer substantial improvement, the Foveon sensor especially so at the longer wavelengths. Note that the strong background attained in the reconstructions of the monochrome sensor is the result of the broadband spectral responsivity of the sensor. The strong background reduces its intensity in the color reconstructions signatures due to the spectral filtering and separability imposed by the RGB pixel-wise spectral sensitivity.

6. Discussion

Although the color sensors are simulated and emulated through the use of a monochrome sensor, they take into account the aliasing effects resulting of demosaicing in the Bayer sensor, and the real absorption/transmission curves of commercially available Foveon/RGB sensors. Regarding the noise characteristics, they were not taken into account in the manuscript. Several papers have studied the noise characteristics of monochrome, Bayer and previous Foveon sen-
sors [24–27]. It is important to point out that, although the Foveon sensors have been labeled as noisy sensors, recent versions of Foveon, such as the Quattro sensor has reported an alleviation of this issue [28]. To the best of our knowledge a full characterization of the noise characteristics in the Foveon Quattro sensor has not been released nor published yet. Furthermore, the simulations and experimental results assumed single sensor sensing, although the Quattro sensor provides finer resolution due to the top (blue) layer partitions pixels in 4 subpixels that can be sensed independently. Analysis of the noise characteristics as well as the higher resolution capabilities of newer Foveon sensors, such as the Quattro, will be considered and analyzed in a future work.

In terms of the quality of the reconstructions, a 1 dB improvement is attained in average by the Foveon sensor over the Bayer sensor. However, when compared against the monochrome sensor, the color sensors overcome the monochrome sensor by about 2-4 dBs. In addition, and most importantly, the color sensors have the advantage of just requiring a single exposure as opposed to the three-times snapshots required by the monochrome sensor to attain the same amount of compressive measurement.

Recent research works have proposed different approaches to recover high resolution spectral images from compressive measurements, such as the use of side information in hybrid camera
architectures [29,30]. These hybrid cameras usually rely on a second image sensor (usually an RGB sensor) to capture the side information of a scene. Although improved results are attained by these recent approaches, the architecture proposed in this manuscript mainly differs in that a single sensor is used. Remark that a number of issues arises that need to be accounted for, in a thorough comparison of these 2 distinct sensing and reconstruction methods. These issues include the registration post-processing and double alignment issues (due to the beam-splitter), and the extra sensor/lens requirement for the side information method.

7. Conclusions

Compressive sensing is a powerful signal processing sensing and reconstruction framework to extract clean signals in noisy environments, and for spectral imaging, creates many new opportunities for research. But because of its novelty, few works have been performed to understand how the CS process will merge with existing methods of collecting spectral images. In this paper, we established both theoretical and experimental evaluations of how using a readily available color sensor improves upon the so far published works that rely on monochrome image sensors. And as more work is done to expand the stacked-color image sensor first pioneered by Foveon to using more colors or a wider range of wavelengths, the CS method described here will take those systems ever higher in spectral resolution.

In closing, let us present Fig. 11 as a final evaluation of color sensor abilities to reconstruct multispectral images where we show the $6^{th}$, $12^{th}$, $18^{th}$ and $24^{th}$ wavebands attained with as
few as a single snapshot. Here, it can be noticed that the Bayer and Foveon detectors attain better artifact-free reconstructions in their respective wavebands compared with the monochrome sensor, and in particular, Fig. 11 shows that using just a single snapshot with the Foveon sensor produces spectral bands with greater fidelity than the monochrome and the Bayer sensor.

Acknowledgments

This research was supported in part by the ARO under the contract W911NF-12-1-0380, and a Defense University Research Instrumentation Program (DURIP) grant.

Disclaimer

The target object presented in the experimental results was not endorsed by the trademark owners and it is used here as fair use to illustrate the quality of reconstruction of compressive spectral image measurements. LEGO is a trademark of the LEGO Group, which does not sponsor, authorize or endorse the images in this paper. The LEGO Group. All Rights Reserved. http://aboutus.lego.com/en-us/legal-notice/fair-play/.