Compressive Spectral Anomaly Detection

Vishwanath Saragadam[†], Jian Wang[†], Xin Li[‡] and Aswin C. Sankaranarayanan[†]
[†] ECE Department, Carnegie Mellon University, Pittsburgh PA
[‡] ECE Department, Duke University, Durham NC

Abstract

We propose a novel compressive sensing imager for detecting anomalous spectral profiles in a scene. The sensing strategy models the background spectrum to lie on a low-dimensional subspace while assuming the anomalies to form a spatially-sparse set of spectral profiles different from the background. Our core contributions are in the form of a two stage sensing mechanism: in the first stage, we estimate the subspace for the spectrum of background by acquiring spectral measurements at a few randomly-selected pixels. In the second stage, we acquire spatially-multiplexed spectral measurements of the scene. Projecting the spectral measurements on the complementary subspace creates a sparse matrix which is then recovered using a Multiple Measurement Vector problem. Theoretical analysis and simulations show significant speed up in acquisition time over other anomaly detection techniques. A lab prototype based on a DMD and a visible spectrometer validates our proposed imager.

1. Introduction

Hyperspectral images measure light intensity in a scene as a function of space and spectrum. The availability of spectral measurements at each spatial location helps in identifying material composition. This has found applications in facial recognition [18, 14], segmentation [23] , and numerous geoscience and remote sensing applications [9].

An important application of hyperspectral imaging is to identify the presence of anomalous materials in a scene. When the materials of interest are present in trace quantities, it is common to model them as anomalous spectral signatures [16]. The most common approach for hyperspectral anomaly detection is to represent the background using a low-dimensional subspace. Subsequently, a spatial location or a "pixel" which cannot be represented accurately by the background model is tagged as an anomaly. Estimation of the location of these anomalies typically requires acquisition of the complete hyperspectral image. Since typical hyperspectral images are extremely high-dimensional, the imaging effort is not only time consuming and also requires highly sensitive and expensive sensor elements.

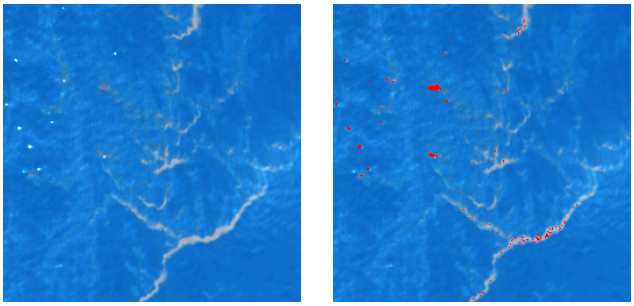


Figure 1: Anomaly detection is often used for geological surveillance. (left) False-color visualization of a $256 \times 256 \times 360$ hyperspectral image of the Deepwater horizon spill [22] with oil spills denoted in gray and background in dark blue. (right) Pixels with anomalous spectra are detected using the proposed sensing framework and marked in red. The results were obtained using 3500×360 measurements which is $1/18$ the dimensionality of the image.

In this paper, we propose a novel sensing framework that directly estimates the location of anomalous spectral signatures without sensing the entire hyperspectral image associated with a scene. We model the hyperspectral image as a sum of the background, assumed to be low-rank, and the anomalies, assumed to be sparse. Under these modeling assumptions, we estimate the locations of the anomalies using a computational camera that implements the following two-stage approach. In the first stage, we build a low-dimensional subspace for the background spectrum by acquiring a spectral measurements at a few randomly-selected spatial locations in the scene. In the second stage, we obtain spatially-multiplexed spectral measurements of the scene and estimate the sparse anomalies by projecting out the subspace corresponding to the background spectrum. Figure 1 shows estimation of 200 anomalous pixels of oil spill in $256 \times 256 \times 360$ Deepwater oil spill hyperspectral image [22] with 3500 spectral measurements. Our algorithm recovers anomalies with very high accuracy with fewer measurements than the dimension of the signal.

Our sensing framework enables a significant reduction in the number of measurements over competing schemes that require the entire hyperspectral image, thereby providing significant speedups and efficiencies at acquisition time.

We make the following contributions in this paper.

- **Novel sensing framework.** We propose a sensing framework that detects anomalies directly with fewer measurements than the dimension of the hyperspectral image.
- **Theoretical guarantees.** We provide expressions for the minimum number of measurements and recovery accuracy for our proposed sensing method.
- **Validation with a lab prototype.** We demonstrate the efficacy of our approach using a lab prototype that uses a digital micromirror device (DMD) for spatial modulation and a spectrometer for spectral measurements.

We also evaluate our technique, on synthetic data, against the classical Reed Xiaoli (RX) [21] detector as well as more recent techniques that rely on compressive imaging under the low rank and sparse signal model.

2. Prior work

Notation. We denote vectors in bold font, scalars in small letters and matrices in capital letters. The ℓ_0-norm of a vector is the number of non-zero entries. Given a matrix $X \in \mathbb{R}^{M\times N}$ and an index set $\Omega \subset \{1, 2, \ldots, N\}$, X_Ω represents a matrix of size $\mathbb{R}^{M\times|\Omega|}$ formed by selecting columns of X corresponding to the elements of Ω. The p,q-norm of a matrix is defined as

$$\|X\|_{p,q} = \left(\sum_{n=1}^{N} \|\mathbf{x}_i\|_p^q\right)^{1/q}.$$

Note that this is an entry-wise norm and is not to be confused with the induced matrix norms. Hence, $\|X\|_{2,2}$ is the Frobenius norm of a matrix, $\|X\|_{2,1}$ is the sum of ℓ_2-norms of columns, $\|X\|_{p,0}$ is the number of non-zero columns of X and $\|X^\top\|_{p,0}$ is the number of non-zero rows of X. Consider a signal $\mathbf{x} \in \mathbb{R}^{JN}$, which can be reshaped into a $J \times N$ matrix, X. The vector $\mathbf{x}$ is called K-block sparse if the number of non-zero columns in X is no more than K.

Compressive sensing (CS). CS aims at recovering a signal from a set of linear measurements fewer than its dimensionality [3], i.e., we seek to acquire a signal $\mathbf{x} \in \mathbb{R}^N$ from $M \ll N$ measurements, $\mathbf{y} = \Phi\mathbf{x} + \mathbf{e}$, where $\Phi \in \mathbb{R}^{M\times N}$ is the sensing matrix, and $\mathbf{e}$ is the measurement noise. CS relies on the sensed signal being sparse either canonically or under a linear transformation. In such a scenario, we can recover the signal by solving

$$\min_{\mathbf{x}} \|\mathbf{y} - \Phi\mathbf{x}\|_2, \quad \|\mathbf{x}\|_0 \leq K.$$

There are many approaches to solving this optimization problem, including convex relaxation techniques like Basis Pursuit [24] and greedy techniques like Orthogonal Matching Pursuit [19] and CoSaMP [17]. This model is also called the single measurement vector (SMV) problem.

The multiple measurement vector (MMV) problem extends the idea of sensing of sparse signals to a scenario where we sense multiple signals that share a common sparsity pattern [10]. The MMV problem can be formulated as

$$Y = \Phi X + E, \quad \|X^\top\|_{2,0} \leq K,$$

with $Y \in \mathbb{R}^{M\times Q}$,$X \in \mathbb{R}^{N\times Q}$,$\Phi \in \mathbb{R}^{M\times N}$ and $E \in \mathbb{R}^{M\times Q}$. Here, $Y = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_Q]$ are the multiple measurements, $X = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_Q]$ are the signals that share the same sparsity pattern, and $E = [\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_Q]$ is the measurement noise. Solutions to the MMV problem can be obtained through algorithms similar in spirit to the SMV problem [10]. The MMV model and the results in Baraniuk et al. [4] are of particular interest to the results in this paper since CS of spatially sparse hyperspectral image under spatial multiplexing reduces to an MMV problem, with each spectral image representing a signal of interest.

Hyperspectral image models. Hyperspectral images can be seen as a composition of a background composed of a few spectral profiles and anomalies, with spectral profiles different from that of the background. Hence, the background spectrum can often be represented by a low-dimensional model and anomalies can be represented as spatially sparse outliers. A sum of low-rank and sparse matrix models is therefore very appealing for hyperspectral images and is of particular interest to us, as it is simple and represents the background spectrum and anomalies accurately. Consider a hyperspectral image with spatial resolution of $N_1 \times N_2$ pixels and a spectrum resolution of N_3 spectral bands. We represent such a hyperspectral image as a $N_3 \times N_1N_2$ matrix such that each column of this matrix is the spectral profile at a particular pixel and each row representing the image at a particular spectral band. Under this notation, the low rank and sparse matrix model is given as

$$X = L + S, \qquad \text{rank}(L) \leq k, \|S\|_{2,0} \leq T \quad (1)$$

Chakrabarti et al. [7] showed that hyperspectral images of natural scenes are very well represented by this signal model. The low rank and sparse model has been used for various hyperspectral imaging tasks [27, 12, 29]. While such models have found use in CS of hyperspectral images with anomaly detection in post-processing, there are few results in existing literature that directly sense the spectral anomalies directly from compressive measurements.

Hyperspectral imaging. Hyperspectral images are captured by various techniques like spatial scanning, spectral

scanning, snapshot imaging and spatio-spectral scanning [25]. Since we need to measure the spectral profile at each pixel, the acquisition process is extremely time consuming and wasteful if we are only interested in detecting anomalies. Hence, CS approaches have been pursued to reduce the number of acquired measurements [1, 26].

Identifying anomalous spectral signatures in hyperspectral images is done in the post-process stage, which involves techniques like hyperspectral unmixing [5], support vector data descriptor [2] and subspace-based methods [20]. The idea underlying these methods is to fit a model to the background spectrum in the hyperspectral image; any pixel which does not conform to this model is classified as an anomalous spectral signature. These techniques require either measuring the whole hyperspectral cube, or recovery of the whole cube from its compressive measurements.

We propose a sensing technique that directly estimates the location of spatially sparse anomalies from a few compressive measurements. Such a model, which involves sensing of spatially sparse signals by removing contribution for background has been explore in video surveillance [6]. Our paper extends this idea to hyperspectral images by knocking off the contribution from background spectra, which results in a spatially sparse hyperspectral image. Locations of the anomalous spectral signatures can then be estimated by solving it as an MMV problem.

3. Proposed methodology

We now detail the imaging architecture for obtaining measurements for the proposed method and provide details of algorithm to solve for the anomalous pixel locations.

3.1. Sensing architecture

Our optical setup (see Figure 2) consists of a DMD and a spectrometer. The scene is focused onto the DMD using an objective lens. Subsequently, light directly from the DMD along one of the mirrored optical axis is fed onto a spectrometer. If we represent the hyperspectral image X as a matrix of dimensions $N_3 \times N_1 N_2$ such that each column of the matrix is the spectral profile at a pixel, then a measurement $\mathbf{y}_i \in \mathbb{R}^{N_3}$ at the spectrometer is of the form $\mathbf{y}_i = X\boldsymbol{\phi}_i + \mathbf{e}_i$, where $\boldsymbol{\phi}_i \in \mathbb{R}^{N_1 N_2}$ is the binary pattern displayed on the DMD and $\mathbf{e}_i$ is the measurement noise. It is worth noting that this optical system is the same as a single pixel camera [11], except that the photodetector (i.e, the single pixel) has been replaced by a spectrometer. Hyperspectral imaging using such an architecture has been explored in [15].

We can obtain multiple set of spectral measurements by displaying different patterns on the DMD. The resulting measurement model can be expressed as $Y = X\Phi + E$, where $Y = [\mathbf{y}_1, \mathbf{y}_2, ...\mathbf{y}_M]$ is the collection of spectral measurements, $\Phi \in \mathbb{R}^{N_1 N_2 \times M}$ is the sensing matrix and $E \in \mathbb{R}^{N_3 \times M}$ is the measurement noise.

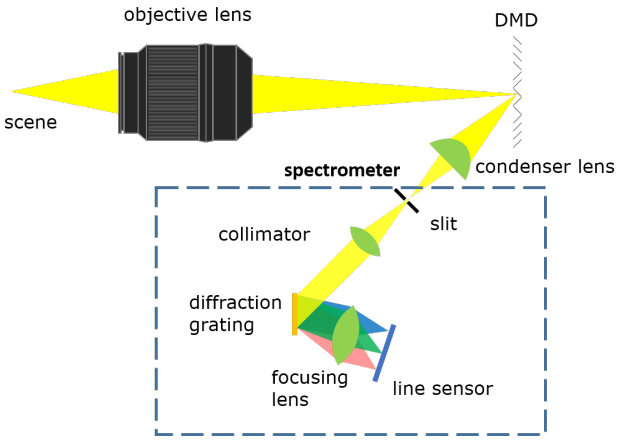


Figure 2: Our built imager consists of a DMD which displays desired patterns and a spectrometer which obtains spectrum of the scene's dot product with DMD pattern.

3.2. Anomaly detection algorithm

We model a hyperspectral image X as a sum of a low rank matrix L and a sparse matrix S. Our aim is to estimate the non-zero columns of S with as few measurements as possible, since each column represents the spectrum at a particular scene pixel. We achieve this using a two-stage sensing mechanism. In the first stage, we obtain a low-rank subspace U corresponding to the column-span of the matrix L; this subspace provides a model for the background spectra. In the second stage, we obtain a few spatially-multiplexed measurements. Complementary projection onto the previously estimated subspace leaves only the sparse outliers, which are recovered using sparse approximation techniques.

Stage 1 – Subspace estimation. The goal is to estimate a subspace for the background pixels by acquiring n_1 spectometer measurements. Let Ω_1 represent the indices of n_1 randomly-chosen spatial locations. We acquire measurements of the form $Y_1 = X_{\Omega_1} + E_1$, where $E_1 \in \mathbb{R}^{N_3 \times n_1}$ is the measurement noise. We estimate a basis $U \in \mathbb{R}^{N_3 \times k}$ for the k-dimensional subspace, such that

$$Y_1 = L_{\Omega_1} + S_{\Omega_1} + E_1 = U\Lambda_{\Omega_1} + S_{\Omega_1} + E_1, \quad (2)$$

We estimate the basis U by solving (2) under a Robust PCA [28] formulation.

Stage 2 – Finding outliers. In the second stage, we obtain M spatially-multiplexed measurements of the form

$$Y_2 = X\Phi + E_2 = L\Phi + S\Phi + E_2. \quad (3)$$

Since we have estimated the basis for the low rank representation in the first stage, we can write $L = U\Lambda$, $\Lambda \in \mathbb{R}^{k \times N_1 N_2}$. Thus $Y_2 = U\Lambda\Phi + S\Phi + E_2$. Let $P_U^{\perp}$ be the complementary projection operator for U.

$$\begin{aligned}\tilde{Y} = P_U^{\perp} Y_2 &= P_U^{\perp} U\Lambda + P_U^{\perp} S\Phi + P_U^{\perp} E_2 \\ &= 0 + P_U^{\perp} S\Phi + P_U^{\perp} E_2 = \tilde{S}\Phi + \tilde{E}_2. \end{aligned} \quad (4)$$

Note that $\tilde{S}$, which is the projection of the sparse matrix onto the complementary subspace of U, retains its sparsity structure, as the projection operation is only for the spectra, spanned by the columns. Estimating $\tilde{S}$ given $\tilde{Y}$ and Φ is then an MMV problem [10].

4. Recovery guarantees

We provide guarantees on recovery in terms of minimum number of measurements and estimation error.

4.1. Minimum number of measurements

We now provide an estimate of the minimum number of measurements required to estimate the anomalies with our sensing strategy. We rely heavily on results from model-based compressive sensing [4] and Chen et al. [8].

Proposed method. To estimate the column span of the low-rank component, we need measurements proportional to its rank and hence, we need $\mathcal{O}(kN_3)$ samples for stage-one. Stage-two requires compressive measurements of a block sparse signal. Baraniuk et al. [4] have shown that for a block sparse signal of size $J \times N$, with K non-zero blocks, measurements as few as $\mathcal{O}(JK + K\log(N/K))$ are sufficient for robust recovery, as long as the measurement matrix satisfies sub-Gaussianity. In our setup, we have $J = N_3$ and $K = T$. Hence, the overall measurement rate of the proposed technique is

$$M_{\text{proposed}} \gtrsim (kN_3 + N_3 T + T\log(N_1 N_2/T)),$$

where $\gtrsim$ implies greater-than upto a multiplicative constant.

Compressive sensing. Chen et al. [8] showed that for a matrix of size $m \times n$ which can be expressed as a sum of a rank-k matrix and sparse non-zero columns, robust recovery is guaranteed with high probability under the following conditions,

$$\begin{aligned} p &\gtrsim \frac{r\log^2(m+n)}{\min(m,n)} \\ \gamma &\lesssim \frac{p}{k\sqrt{r}\log^3(m+n)}. \end{aligned} \quad (5)$$

p is the fraction of observations and γ is the fraction of non-zero columns in the sparse matrix. Adapting it to our method, we have

$$\begin{aligned} M_{\text{joint}} &\gtrsim \frac{N_1 N_2 N_3 r\log^2(N_1 N_2 + N_3)}{\min(N_1 N_2, N_3)} \\ &= rN_1 N_2 \log^2(N_1 N_2 + N_3), \end{aligned}$$

where M_{joint} is the number of measurements for joint recovery of sparse and low rank matrices. It is clear that the number of measurements for our proposed method is less than the number for joint recover of sparse and low-rank matrix. As an example, consider a hyperspectral image of size $256 \times 256 \times 300$ and 100 outliers. Let us assume the background is approximated by a rank-5 matrix. The number of measurements needed for the proposed sensing technique would be approximately $10 \times 5 \times 300 + 100 \times 300 + 100 \times 16 = 33{,}100$, whereas that for compressive sensing would be approximately $5 \times 256 \times 256 \times 16 = 524{,}288$. This provides us a gain of $15\times$ in number of measurements.

4.2. Recovery guarantees for outliers

We look at three conditions for recovery of outliers: noiseless measurements with knowledge of the exact subspace estimate, noisy measurements with knowledge of the exact subspace estimate, and noisy measurements with an inexact, rank-k approximation of the subspace. The bounds and guarantees are based on the MMV variant of CoSaMP and we will repeatedly invoke the following result from [4].

Theorem 1. *(Adapted from Theorem 6 in [4]) Let S be a K block-sparse signal and let $Y = S\Phi + E$ be a set of noisy compressive sensing measurements, where $\Phi \in \mathbb{R}^{N_1 N_2 \times M}$. If Φ is a random sub-Gaussian matrix, then the estimate obtained from iteration i of block-based CoSaMP satisfies*

$$\begin{aligned}\|S - \hat{S}_i\|_{2,2} \le \frac{1}{2^i}\|S\|_{2,2} + 20\Big(\|S - S_T\|_{2,2} + \cdots \\ \frac{1}{\sqrt{T}}\|S - S_T\|_{2,1} + \|E\|_{2,2}\Big), \end{aligned} \quad (6)$$

where S_T is the best T-term approximation of S, and E is the additive measurement noise.

In essence, the theorem states that recovery is guaranteed if M is sufficiently large. Further details about the sensing matrices and number of measurements can be obtained in [4]. Next, we show the bounds on error with the three conditions of measurements.

Noiseless and exact subspace. In this case, we have the following measurement model, $\widetilde{Y} = P_U^{\perp} Y$, where the sparsity of $\widetilde{S}$ is the same as S. Invoking (6), we have

$$\|\widetilde{S} - \hat{\widetilde{S}}_i\|_{2,2} \le \frac{1}{2^i}\|\widetilde{S}\|_{2,2}, \quad (7)$$

where $\widetilde{S} = [\widetilde{\mathbf{s}}_1, \widetilde{\mathbf{s}}_2, ..., \widetilde{\mathbf{s}}_{N_1 N_2}]$ is the projection of S on complementary subspace of U and $\hat{\widetilde{S}}_i$ is the estimate of $\tilde{S}$ after i iterations of MMV CoSaMP.

Noisy and exact subspace. Assume that the added noise is bounded as, $\|E_2\|_{2,2} \leq \epsilon \|\widetilde{S}\|_{2,2}$. Here, ϵ is a measure of the signal-to-noise ratio. As before, taking complementary projection, we have $\widetilde{Y} = \widetilde{S}\Phi + \widetilde{E}_2$. Invoking (6), we have

$$\begin{aligned} \|\widetilde{S} - \hat{\widetilde{S}}_i\|_{2,2} &\leq \frac{1}{2^i}\|\widetilde{S}\|_{2,2} + 20\|\widetilde{E}_2\|_{2,2} \\ &\leq \left(\frac{1}{2^i} + 20\epsilon\right)\|\widetilde{S}\|_{2,2}. \end{aligned} \tag{8}$$

Noisy and inexact subspace. Let $\hat{U}$ be estimated rank k estimate of the columns of L and $\Lambda = [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, ..., \boldsymbol{\lambda}_{N_1 N_2}]$ represent coefficients for each column. Further, let the following bound hold

$$\|P_{\hat{U}}^{\top} U \boldsymbol{\lambda}_k\| \leq \epsilon_s \|\widetilde{\mathbf{s}}_k\|, \tag{9}$$

where ϵ_s is a small constant. Also, let $\|\Phi\|_{2,2} \leq \eta$. Now taking the complementary projection as before, we have

$$\begin{aligned} P_{\hat{U}}^{\top} Y &= P_{\hat{U}}^{\top} U \Lambda \Phi + P_{\hat{U}}^{\top} S_2 \Phi + P_{\hat{U}}^{\top} E_2 \\ &= E_s + \widetilde{S}_2 \Phi + \widetilde{E}_2, \end{aligned} \tag{10}$$

where $E_s = P_{\hat{U}}^{\top} U \Lambda \Phi$. $P_{\hat{U}}^{\top} U \Lambda$ can be treated as the error in T-term approximation of $\widetilde{S}$.The $\|\cdot\|_{2,2}$ and $\|\cdot\|_{2,1}$ error bounds on this term,

$$\|E_s\|_{2,2} = \left(\sum_{k=1}^{|\Omega_3|} \|P_{\hat{U}}^{\top} U \boldsymbol{\lambda}_k \Phi\|_2^2\right)^{1/2} \tag{11}$$

$$\leq \left(\sum_{k=1}^{|\Omega_3|} \|P_{\hat{U}}^{\top} U \boldsymbol{\lambda}_k\|_2^2 \|\Phi\|_{2,2}^2\right)^{1/2} \tag{12}$$

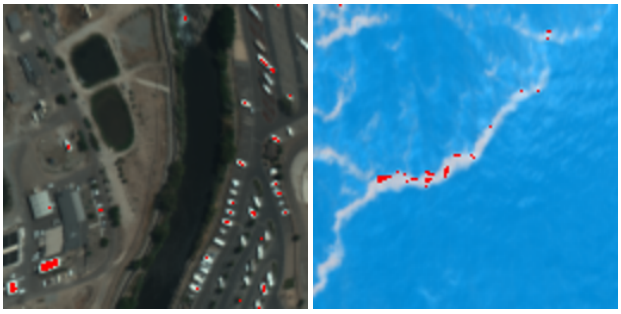
$$\leq \left(\eta^2 \sum_{k=1}^{|\Omega_3|} \|\widetilde{\mathbf{s}}_k\|_2^2 \epsilon_s^2\right)^{1/2} = \eta \epsilon_s \|\widetilde{S}\|_{2,2}.$$

Similarly, we have

$$\|E_s\|_{2,1} \leq \epsilon_s \|\widetilde{S}\|_{2,1}. \tag{13}$$

Invoking (6), we get

$$\begin{aligned} \|\widetilde{S} - \hat{\widetilde{S}}_i\|_{2,2} \leq \frac{1}{2^i}\|\widetilde{S}\|_{2,2} + \\ 20\left(\|E_s\|_{2,2} + \frac{1}{\sqrt{T}}\|E_s\|_{2,1} + \|\widetilde{E}_2\|_{2,2}\right) \\ \leq \left(\frac{1}{2^i} + 20(\epsilon + \epsilon_s)\right)\|\widetilde{S}\|_{2,2} + \\ 20\epsilon_s \frac{1}{\sqrt{T}}\|\widetilde{S}\|_{2,1}. \end{aligned} \tag{14}$$



(a) Image of urban aerial snapshot with anomalous locations highlighted in red.

(b) Image of deepwater horizon spill with anomalous locations highlighted in red.

Figure 3: Synthetic images used for experiments. The urban image has 71 anomalous pixels and the deepwater spill image has 48 anomalous pixels.

Equation (14) establishes the effect of model mismatch as well as measurement noise on our estimation algorithm. Provided that the model is well estimated and the noise energy is bounded, the estimated anomalous locations can be very accurate. The synthetic and real experiments section gives a brief insight into the effect of noise and model error.

Joint recovery. Ha et al. [13] discussed error bounds for recovery of low rank and sparse matrix from compressed and noisy data. For brevity, we skip the complete expression of error and request readers to refer to [13] for analysis on noise. Table 1 summarizes the number of measurements and the measurement noise for various methods.

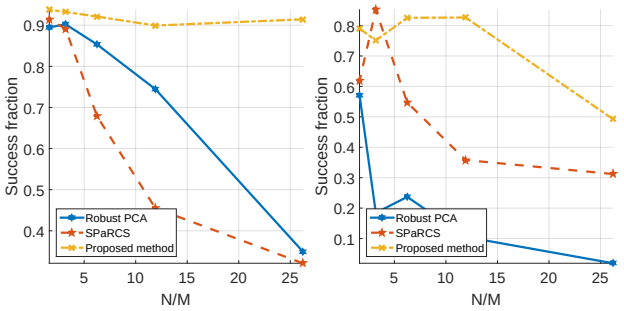
5. Synthetic experiments

We compare our proposed technique against two compressive sensing techniques for recovering data from a low rank and sparse matrix, SPaRCS [27], which is a greedy technique and ADMM implementation of Robust PCA [28]. We create synthetic data by decomposing the data as sparse and low rank component using Robust PCA and then adding the two together. Figure 3 shows the two hyperspectral images with anomalous locations highlighted as red pixels.

Recovery accuracy. We performed recovery of the anomalous locations with $n_1 = 500$ measurements to obtain subspace and by sweeping n_2 from 1,000 to 15,000 random-permuted Hadamard measurements. For comparison, we recovered a rank-6 matrix as well as a column-sparse matrix from the compressive measurements using SPaRCS and Robust PCA under corrupted columns model [8]. We define success ratio as the fraction of anomalous locations correctly identified. Figure 4 and 5 compares the success fraction for both the hyperspectral images under two different noise conditions. Our method outper-

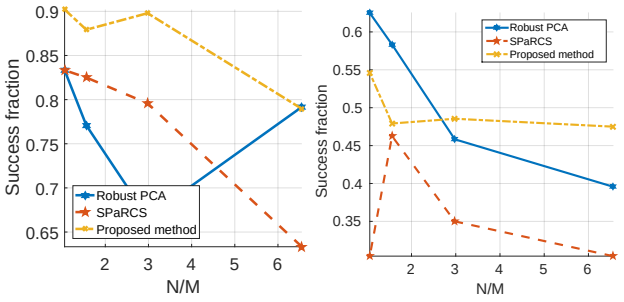
Method	Number of measurements	Bounds on recovery accuracy
Full measurements	$N_1N_2N_3$	Measurement noise
Compressive sensing	$\mathcal{O}(rN_1N_2\log^2(N_1N_2+N_3))$	Joint low rank and sparse recovery [13]
Proposed method	$\mathcal{O}(rN_3+N_3T+T\log(N_1N_2/T)))$	$a\|\|\widetilde{S}\|\|_{2,2}+\frac{b}{\sqrt{T}}\|\|\widetilde{S}\|\|_{2,1}$

Table 1: Comparing various methods based on number of measurements and the estimation error in anomalous pixels. Our proposed method requires far fewer measurements than even compressive sensing.



(a) Success ratio with 60dB added noise. **(b)** Success ratio with 40dB added noise.

Figure 4: Comparison of success ratio with various methods for the urban scene with 79 anomalies. Success fraction is defined as the number of outliers correctly identified. N/M is the ratio of signal dimension to number of measurements.

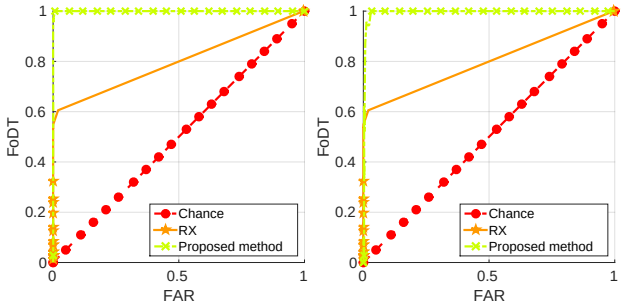


(a) Success ratio with 60dB added noise. **(b)** Success ratio with 40dB added noise.

Figure 5: Comparison of success ratio with various methods for deepwater spill with 483 anomalies. Success fraction is defined as the number of outliers correctly identified.

forms other compressive sensing based methods, particularly when the measurements are far fewer than the dimension of the signal.

RoC curve. We compared our proposed method against the Reed Xiaoli (RX) detector for anomalies [21]. We artificially introduced 71 targets in the deepwater image shown in Figure 3. The comparison is made with the receiver operating characteristic (RoC) curve, which is a plot of the fraction of detected targets (FoDT) against the false alarm rate (FAR). RoC curve for the RX detector was obtained by varying the threshold of detection. RoC curve for our proposed method was obtained by varying the number of anomalous targets to detect. Number of measurements for our method were increased linearly with the number of anomalies to estimate. Figure 6 shows RoC for the two methods. Our method outperforms the RX detector by a large margin and also needs fewer measurements, which is strongly in favor of our model.



(a) Detection with 60dB noise. **(b)** Detection with 40dB noise.

Figure 6: RoC curve comparing our method and the Reed Xiaoli (RX) detector. Our method outperforms the RX detector and needs fewer measurements.

Effect of noise and number of measurements. Table 2 shows success fraction as a function of noise and the number of initial measurements for a $128 \times 128 \times 366$ dimensional hyperspectral image with 71 anomalous pixels. It is evident that noise plays a significant roles in estimation of outliers. For estimating the subspace, measurements greater than 400 are sufficient for good accuracy. Effect of high noise can also be seen in Figure 5b, with Robust PCA and SPaRCS outperforming our method for large number of measurements. This is consistent with our theoretical guarantees, where we showed in (14) that there is a significant effect of measurement noise on estimation accuracy.

SNR (dB)	Percentage of correctly detected targets $n_1 = 100$	$n_1 = 400$	$n_1 = 800$	$n_1 = 1000$
40	60.06	68.99	68.20	70.79
60	88.59	91.92	92.37	91.66
80	96.59	98.53	98.28	98.31

Table 2: Percentage of correctly detected targets as a function of SNR and number of initial measurements for $128 \times 128 \times 356$ hyperspectral image of an urban scene. n_1 represents the number of spatial locations at which spectral signature was recorded to estimate the subspace.

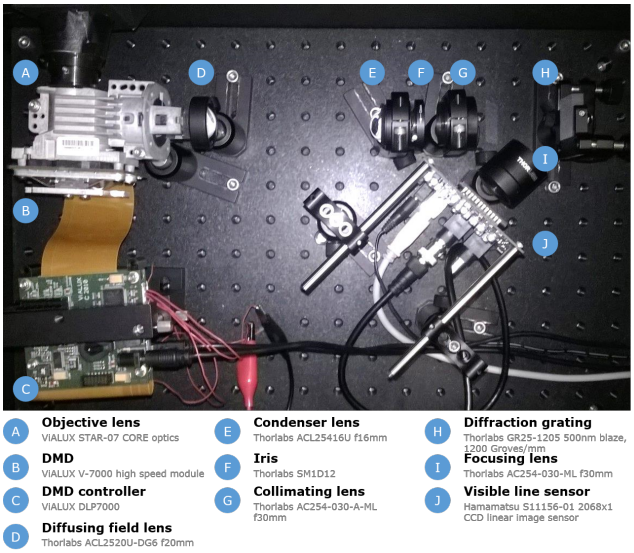


Figure 7: Setup for DMD-based hyperspectral imaging. The objective lens (A) focuses light onto the DMD (B). The light is then measured by the spectrometer consisting of a condenser lens (E) that focuses light onto an iris (F), which is then collimated by the lens (G). The collimated light is diffracted by a reflective diffraction grating (H), which is then focused by lens (I) onto the line sensor (J).

6. Real data experiments

To validate our proposed sensing method, we built a visible (400-700nm) hyperspectral camera based on architecture proposed by Li et al. [15] (see Figure 7). We show three examples of anomaly detection to demonstrate success of our proposed method.

6.1. Experiment on data from SpecTIR

We performed anomaly detection on Deepwater Horizon oil-spill data obtained from SpecTIR [22], by obtaining 500 random pure spectral measurements for estimating subspace for background spectra and then obtaining 3000 spatially multiplexed measurements for estimating a total of

(a) Scene **(b)** Detected anomalies

Figure 8: Output of our algorithm on first real scene (a). The scene is made of "ICCP 2017" printed on a paper with a hole in the letter "P". A white LED is situated behind a diffuser at the location of this hole. Detected anomalies are shown in red in image (b).

200 outliers. Figure 1 shows the $256 \times 256 \times 360$ hyperspectral image we used for estimating outliers and the result of our algorithm. We observe that except a few points, all the estimated anomalies lie on the oil slick locations, identified by gray colored streaks in the image.

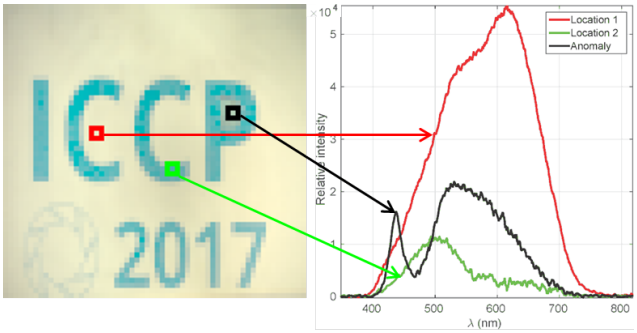
6.2. Experiment on data from our imaging setup

We tested our algorithm on two scenes to estimate 5-8 outliers with the optical setup we built. We performed anomaly detection on 64×64 dimensional images with 50 spectral bands between $350 - 820$nm.

Subspace estimation. Instead of obtaining subspace by randomly sampling the scene, we initially acquire a very low spatial resolution hyperspectral image, as this maximizes light throughput. If only a few anomalies are present, this method would still give a very good estimate of the subspace. In our experiments, we estimated subspace by taking an 8×8 hyperspectral image of the scene.

Scene 1. In our first experiment, we tested our sensing method on a scene made of "ICCP 2017" printed on a white paper with a hole in the letter "P", as shown in red circle in Figure 8. A white LED light with a spectral signature significantly different from the paper is used to illuminate the hole; hence, the region corresponding to the hole is the anomaly that we seek to detect. We estimated ten anomalous pixels from 250 permuted Hadamard measurements. Results are shown in Figure 8 and spectral signatures at some representative locations are shown in Figure 9.

Scene 2. In our second experiment, we tested our sensing method on a scene printed on paper with primarily red colors and four pixels illuminated by a laser, as show in



(a) Scene **(b)** Spectral signatures

Figure 9: Spectral signatures at some representative locations for the first scene. The noticeable peak of the anomalous pixel is due to a white LED.

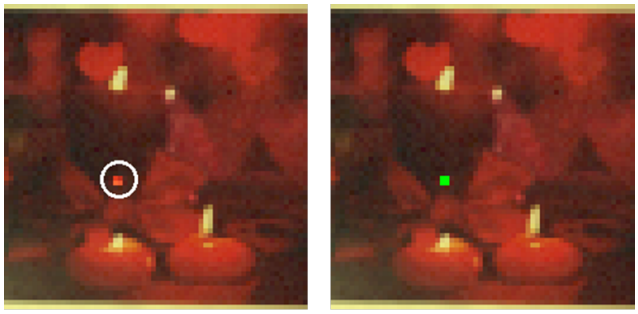
white circle in Figure 10. Since the laser is also red, it is not easy to differentiate with naked eyes. We estimated a four anomalies from 50 permuted hadamard measurements. Results are showing in Figure 10 and spectral signatures at some representative locations are shown in Figure 11.

Discussion. Success of anomaly detection with subspace modeling for the background relies on obtaining a good estimate of the subspace. A lower rank representation would not sufficiently represent the background which would give erroneous results. Since we estimated subspace from a downsampled version of the hyperspectral image, the spectrum of anomalous pixels may show up if higher rank is used. Figure 12 shows two failure cases with a lower and a higher rank subspace. Figure 13 shows residue after removing contribution from background for a subspace of rank 5, 8 and 12. Under estimation of rank gives error in regions other than the anomalies and an over estimate removes contribution due to anomalies. This problem exists with all anomaly detection algorithms. Erroneous representation of background spectra results in error in anomalies detected.

Prior knowledge of the complexity of the scene is definitely of help to estimate the rank of subspace. Further, smoother profiles are harder to differentiate. In our two examples, the stark difference between spectrum of laser and background made it easier to estimate outliers with a 4 dimensional subspace. On the other hand, the spectrum of white LED is smooth, making it harder to differentiate.

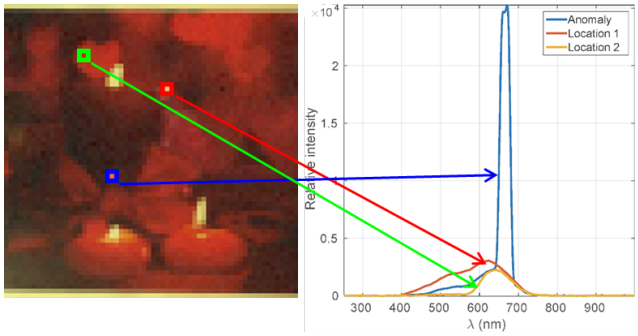
7. Conclusion

We presented a novel sensing method that directly obtains anomalies in a hyperspectral image without requiring the measurement of the complete image. Simulations and theoretical guarantees are strongly in favor of our proposed method. We also demonstrated a lab prototype that successfully detected anomalies in the scene. The method is particularly appealing as it detects anomalies with few measurements. As a consequence, we can develop imagers for geological surveillance that have higher temporal resolutions or better noise properties enabled by temporal averaging.



(a) Scene **(b)** Detected anomalies

Figure 10: Output of our algorithm on second real scene (b). The scene is made of a printed paper with primarily red color and four of the pixel locations are shone by a red laser. It is not easy to discern the location of laser with naked eye, but our algorithm identifies the anomalous locations with only 100 spectral measurements, shown as green dots in image (b).



(a) Scene **(b)** Spectral signatures

Figure 11: Spectral signatures at some representative locations for the second scene. The noticeable peak of the anomalous pixel is laser light. Although the peak for the anomalous pixel is high, the energy is similar to all the other pixels.

8. Acknowledgements

The authors were supported in part by Intel ISRA on Compressive Sensing. Sankaranarayanan was also supported in part by the NSF CAREER grant CCF-1652569.

Figure 12: Effect of model mismatch for anomaly detection. Shown here are anomaly detection results with (left) underestimated subspace rank of 5 and (right) overestimated subspace rank of 12. Over estimation captures spectral profile of anomalies as well, and hence are not identified as anomalies.

Figure 13: Residual image with subspace ranks of (left) five, (middle) eight, and (right) twelve. If the subspace rank is low, background and the anomalies have similar error. If rank is high, the residue is only noise, as the subspace captures anomalies as well. With a subspace of rank 8, the anomalous pixels within the letter "P" is clearly evident.

References

[1] Y. August, C. Vachman, Y. Rivenson, and A. Stern. Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains. *Appl. Optics*, 52(10):D46–D54, 2013. 3

[2] A. Banerjee, P. Burlina, and R. Meth. Fast hyperspectral anomaly detection via svdd. In *IEEE Intl. Conf. Image Processing*, 2007. 3

[3] R. G. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, 24(4):118–121, 2007. 2

[4] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde. Model-based compressive sensing. *IEEE Trans. Info. Theory*, 56(4):1982–2001, 2010. 2, 4

[5] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Selected Topics in Appl. Earth Observations and Remote Sensing*, 5(2):354–379, 2012. 3

[6] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa. Compressive sensing for background subtraction. In *European Conf. Computer Vision*, 2008. 3

[7] A. Chakrabarti and T. Zickler. Statistics of Real-World Hyperspectral Images. In *IEEE Conf. Computer Vision and Pattern Recognition*, 2011. 2

[8] Y. Chen, H. Xu, C. Caramanis, and S. Sanghavi. Robust matrix completion and corrupted columns. In *Intl. Conf. Machine Learning*, 2011. 4, 5

[9] E. Cloutis. Review article hyperspectral geological remote sensing: evaluation of analytical techniques. *International J. Remote Sensing*, 17(12):2215–2242, 1996. 1

[10] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Trans. Signal Processing*, 53(7):2477–2488, 2005. 2, 4

[11] M. F. Duarte, M. A. Davenport, D. Takbar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83, 2008. 3

[12] M. Golbabaee and P. Vandergheynst. Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery. In *IEEE Intl. Conf. Acoustics, Speech And Signal Processing*, 2012. 2

[13] W. Ha and R. F. Barber. Robust pca with compressed data. In *Adv. Neural Info. Processing Systems*, 2015. 5, 6

[14] J. C. Harsanyi and C.-I. Chang. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Trans. Geoscience and Remote Sensing*, 32(4):779–785, 1994. 1

[15] C. Li, T. Sun, K. F. Kelly, and Y. Zhang. A compressive sensing and unmixing scheme for hyperspectral data processing. *IEEE Trans. Image Processing*, 21(3):1200–1210, 2012. 3, 7

[16] S. Matteoli, M. Diani, and G. Corsini. A tutorial overview of anomaly detection in hyperspectral images. *IEEE Aerospace and Electronic Systems Magazine*, 25(7):5–28, 2010. 1

[17] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009. 2

[18] Z. Pan, G. Healey, M. Prasad, and B. Tromberg. Face recognition in hyperspectral images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(12):1552–1560, 2003. 1

[19] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Asilomar Conf. Signals, Systems and Computers*, 1993. 2

[20] K. I. Ranney and M. Soumekh. Hyperspectral anomaly detection within the signal subspace. *IEEE Geoscience and Remote Sensing Letters*, 3(3):312–316, 2006. 3

[21] I. S. Reed and X. Yu. Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38(10):1760–1770, 1990. 2, 6

[22] SpecTIR. Spectir, Advanced hyperspectral and geospatial solutions. http://www.spectir.com/free-data-samples/. 1, 7

[23] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognition*, 43(7):2367–2379, 2010. 1

[24] J. A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Trans. Info. Theory*, 52(3):1030–1051, 2006. 2

[25] F. Vagni. Survey of hyperspectral and multispectral imaging technologies. *DTIC Document*, 2007. 3

[26] A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Optics*, 47(10):B44–B51, 2008. 3

[27] A. E. Waters, A. C. Sankaranarayanan, and R. Baraniuk. Sparcs: Recovering low-rank and sparse matrices from compressive measurements. In *Adv. Neural Info. Processing Systems*, 2011. 2, 5

[28] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In *Adv. Neural Info. Processing Systems*, 2009. 3, 5

[29] Y.-Q. Zhao and J. Yang. Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Trans. Geoscience and Remote Sensing*, 53(1):296–308, 2015. 2