

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing [scholarworks-group@umbc.edu](mailto:scholarworks-group@umbc.edu) and telling us what having access to this work means to you and why it's important to you. Thank you.

# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://SPIDigitalLibrary.org/conference-proceedings-of-spie)

## Unsupervised automatic target generation process via compressive sensing

Adam Bekit, Charles Della Porta, Bernard Lampe, Bai Xue, Chen-I Chang

Adam Bekit, Charles Della Porta, Bernard Lampe, Bai Xue, Chen-I Chang, "Unsupervised automatic target generation process via compressive sensing," Proc. SPIE 10989, Big Data: Learning, Analytics, and Applications, 109890G (13 May 2019); doi: 10.1117/12.2518359

**SPIE.**

Event: SPIE Defense + Commercial Sensing, 2019, Baltimore, Maryland, United States

# Unsupervised Automatic Target Generation Process via Compressive Sensing

Adam Bekit<sup>a</sup>, Charles Della Porta<sup>a</sup>, Bernard Lampe<sup>a</sup>, Bai Xue<sup>a</sup>, Chen-I Chang<sup>a</sup>

<sup>a</sup>Remote Sensing Signal and Image Processing Laboratory  
Department of Computer Science and Electrical Engineering  
University of Maryland, Baltimore County, Baltimore, MD 21250

## ABSTRACT

Unsupervised target generation for hyperspectral imagery (HSI) have generated great interest in the hyperspectral community. However, most of the current unsupervised target generation algorithms have to process large HSI data, which is acquired using the traditional Nyquist-Shannon sampling theorem, resulting in data with high band-to-band correlation. As a consequence, these algorithms end up processing redundant information, raising the demand for large memory storage, processing time, and transmission bandwidth. In the past, some efforts have been dedicated to dealing with the redundant information via data reduction (DR) or data compression post-acquisition. However, to the best of our knowledge, this challenge has been addressed outside the context of Compressive Sensing (CS). This paper applies CS data acquisition process at the sensor level so that the redundant information is removed at the early stage of the data processing chain. The main advantage of our approach is that it employs a random sensing process, and the concept of universality, to randomly sense the HSI bands and produce data containing the bare minimum information. We take advantage of CS Restricted Isometric Properties (RIP), Restricted Conformal Properties (RCP), and newly derived orthogonal sub-space projection (OSP) properties to perform automatic target generation process (ATGP) in the compressively sensed band domain (CSBD), instead of in the original data space (ODS), where the HSI data contains full spectral bands. Our experimental results show that, by working in the CSBD, we avoid processing redundant data and still maintain performance results that are comparable with the performance results obtained in the ODS.

**Keywords:** Compressive sensing (CS), automatic target generation process (ATGP), compressively sensed band domain (CSBD), original data space (ODS), orthogonal sub-space projection (OSP), Restricted Orthogonal Sub-space Projection Property (ROSPP)

## 1. INTRODUCTION

Many applications in the Military and Civilian require sensing, detection, and tracking of objects or activities. Hyperspectral sensors provide passively detected spatial and spectral information allowing remote sensing applications such as mineral mapping, vegetation classification and environment analysis, etc. [1], which have received tremendous attention in the hyperspectral community. In a hyperspectral remote sensing system, there are four basic elements: radiation source, atmospheric, image surface, and sensor [2]. The sun's energy provides the illumination source, which travels through the earth's atmosphere and interacts with elements (material) on the surface of the earth before it reflects to the sensor. The sensor has the ability to measure the amount of intensity reflected back from the material to the sensor. The amount of intensity the sensor measures is a function of the composition of the imaged material on the earth's surface. Therefore, depending on the material type, the amount of energy absorbed by the material is different. In addition, the sensor measures the amount of energy reflected from the material as a function of different wavelengths or spectral bands.

Because hyperspectral sensors have the technological ability to measure many contiguous intensity of spectral bands (wavelengths), they provide tremendous amount of information, which can be used for data exploitation. One example of data exploitation is endmember finding. An endmember is defined as an idealized "pure" signature that can be used to specify a particular spectral class. As a consequence, finding endmembers in hyperspectral imagery has become a major and unique application, which cannot be found in other imagery such as multispectral or single spectral imagery. There are many supervised and unsupervised endmember finding techniques, which have been very successful in extracting endmembers (targets) from hyperspectral imagery. Another very important data exploitation application, which depends on target extraction is spectral unmixing [3] (i.e., abundance fractions estimation). The material composition of a

hyperspectral image can be inferred by performing data processing and spectral unmixing [4]. This is possible because hyperspectral sensors have very fine resolutions. As a result, they capture very detailed information about the material present on the surface.

Despite providing very rich spatial and spectral information, hyperspectral sensors provide redundant information, which many algorithms must deal within the processing chain down the stream. The direct impact is that the high band-to-band correlation present in the imagery cost high algorithm computational time, memory storage, and transmission bandwidth. Lately, much efforts and attention have been given to developing hyperspectral image processing algorithms in order to deal with this challenge. One approach to dealing with this challenge is to develop new algorithms (i.e., compression post acquisition) that remove the high band-to-band correlation present in the imagery before applying the endmember (target) finding algorithms. However, these approaches have shown to have performance issues and complicate the problem even more because they are still required to process redundant data. A new but very promising approach is to take advantage of the concept of Compressive Sensing (CS), which compresses (senses) the imagery as the signal is being acquired at the sensor level. As a result, the redundant information is removed immediately (at the sensor layer) and never enters the data processing chain.

To demonstrate the utility of using CS for unsupervised ATGP, a series of experiments are designed and their results are presented. In particular, we show the results of applying CS to ATGP in the CSBD using two synthetic images and one real image. We then compare the results with the results obtained in the ODS, where the traditional approach is used (i.e., non-CS).

## 2. COMPRESSIVE SENSING OVERVIEW

Recently, CS has been and continued to be a major research area in many applications. It is a new signal or image acquisition technique, which goes against the conventional and well-known approach, Nyquist-Shannon sampling theorem. For a perfect signal reconstruction, the conventional approach requires the signal sampling rate to be at least twice faster than the signal bandwidth. On the other hand, CS does not have such a sampling constraint. The CS theory asserts that a perfect signal reconstruction is possible with much fewer samples than samples required by the conventional method [5]–[9]. The key of CS theory relies on three major ideas: sparsity, incoherence, signal sampling and recovery.

### 2.1 Sparsity

In many real-life applications signals may not appear sparse in one domain, but may be highly sparse under some “sparsifying” domain represented by a basis matrix  $\Psi \in \mathcal{R}^{L \times L}$ . For example, a signal  $\mathbf{r} \in \mathcal{R}^L$  may not look sparse in time-domain but may be very sparse in other domains such as fast fourier transform (FFT), direct cosine transform (DCT), or wavelet [10], [11]. What this means is that the signal  $\mathbf{r}$  is well approximated by a linear combination of  $k$  vectors such that  $\mathbf{r} \approx \sum_{i=1}^k \rho_i \Psi_i$ , where  $\rho_i$  are the coefficients and  $\Psi_i$  are the basis vectors in the sparse domain. Mathematically speaking, a signal  $\mathbf{r}$  is said to be sparse if most of its coefficients are zeros. In addition,  $\mathbf{r}$  is  $k$ -sparse if

$$\|\Psi \mathbf{r}\|_0 \leq k, \quad (1)$$

where  $\|\cdot\|_0$  denotes the number of non-zero coefficients in the vector  $\Psi \mathbf{r}$ . If the information bearing signal,  $\mathbf{r}$ , is very sparse (i.e., can be represented using few coefficients) in some sparse domain, then sampling the signal at twice its bandwidth is not necessary. This is one of the central themes of the CS theory. In fact, the CS theory asserts that a perfect signal recovery is possible, with high probability, despite that sampling (sensing) the signal at a rate that is much lower than the Nyquist rate, provided that certain CS conditions are satisfied and the signal is sparse [5]–[9].

### 2.2 Incoherence

Mutual incoherence, or coherence, plays a significant role in CS. It is used to measure the correlation between the sensing

matrix  $\Phi \in \mathbb{R}^{m \times L}$  and the sparse representation basis matrix  $\Psi \in \mathbb{R}^{L \times L}$ . According to the CS theory, the lowest coherence between the sensing matrix and the sparse representation matrix is desired for better signal sensing and reconstruction. Mathematically, mutual coherence is defined as

$$\mu(\Phi, \Psi) = \sqrt{L} \max_{1 \leq j, i \leq L} |\langle \phi_j, \psi_i \rangle|, \quad (2)$$

where  $\phi_j$  and  $\psi_i$  represent the columns of the sensing matrix  $\Phi$  and the sparse representation matrix  $\Psi$ , respectively. Low coherence between the sensing matrix and any sparse representation matrix is guaranteed by choosing a sensing matrix  $\Phi$  with random entries [5]–[9]. In our experiments, incoherence is directly achieved by using a Gaussian sensing matrix. Therefore, there is no need to check for incoherence explicitly.

### 2.3 Signal Sampling and Recovery

The CS framework treats signal sampling as a linear model given by

$$\mathbf{r}_\Phi = \Phi \Psi \mathbf{r} + \mathbf{n}, \quad (3)$$

where  $\Psi$  is sparse matrix,  $\Phi$  is a user-specified measurement (sensing) matrix,  $\mathbf{r}_\Phi \in \mathbb{R}^m$  is a vector containing the sensed signal, and  $\mathbf{n} \in \mathbb{R}^m$  is a random noise vector. It is clearly seen in (3) that the matrix  $\Phi$  represents dimensionality reduction. It maps the vector  $\mathbf{r} \in \mathbb{R}^L$  into another vector  $\mathbf{r}_\Phi \in \mathbb{R}^m$  with dimensionality of  $m \ll L$  being the number of compressively sensed bands (CSBs) size. It is important to note that the framework given by (3) is not the universality model used in CS, where the transformation,  $\Psi$  of  $\mathbf{r}$  to a sparse domain is not required during sampling (sensing). More precisely, in the CS universality model, the sparse transformation  $\Psi$  is only needed during reconstruction. Therefore, in our research, we take advantage of the so-called CS universality model to avoid any time-domain to sparse-domain (FFT, DCT, etc.) transformation of the signal  $\mathbf{r}$  by directly applying CS to  $\mathbf{r}$ . In other words, we can avoid the  $\Psi$  transformation applied to  $\mathbf{r}$  and use the universality model by reducing (3) to

$$\mathbf{r}_\Phi = \Phi \mathbf{r} + \mathbf{n}. \quad (4)$$

The CS sampling framework is shown below in

Fig. 1.

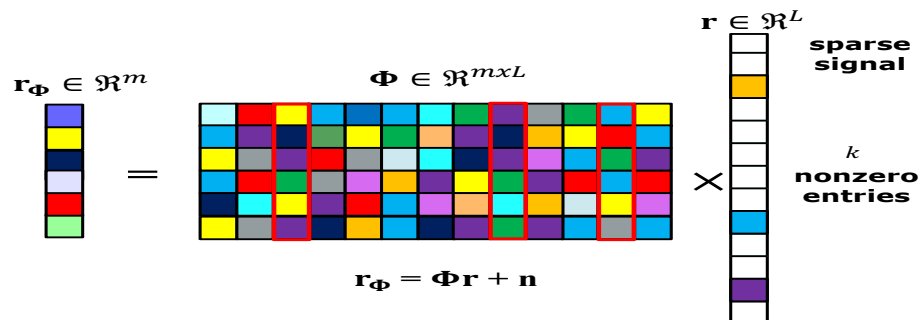


Fig. 1. Compressive Sampling

The use of the universality model given by (4) is significant since it allows us to avoid “sparsifying” the HSI. In addition, considering the HSI is very sparse to begin with, transforming the HSI to another sparse domain brings no added benefits. In fact, it introduces more complexity and unnecessary computational time. Our early experimental results did demonstrate that transforming the HSI from time-domain to the DCT-domain brought no additional performance gains.

One fundamental question in CS is how to design a sensing matrix  $\Phi$  so that  $\mathbf{r}_\Phi$  preserves as much of the bare minimum formation content from the original signal  $\mathbf{r}$ ? In other words, the sensing matrix  $\Phi$  must ensure that the signal  $\mathbf{r}$  is recovered from  $\mathbf{r}_\Phi$  as accurately as possible. One way to address this question is to impose a requirement on the sensing matrix  $\Phi$ . The CS theory establishes the Restricted Isometric Property (RIP), which is a sufficient condition for signal recovery [12], [13]. A sensing matrix  $\Phi$  satisfies the RIP condition of order  $k$  if there exists a Restricted Isometric Constant (RIC)  $\delta_k$ , which is the smallest constant [12] such that the RIP condition specified by

$$(1 - \delta_k) \|\mathbf{r}\|_2^2 \leq \|\Phi\mathbf{r}\|_2^2 \leq \|\mathbf{r}\|_2^2(1 + \delta_k) \quad (5)$$

is satisfied for a  $k$ -sparse vector  $\mathbf{r}$ . According to CS theory, if a sensing matrix  $\Phi$  satisfies (5), then the original signal  $\mathbf{r}$  can be recovered from signal  $\mathbf{r}_\Phi$  with high probability [5]–[9]. Therefore, the design requirement on the sensing matrix  $\Phi$  is equivalent to satisfying the RIP condition. From seeing (5), it is apparent that the RIP condition is a distance preserving criterion on  $\mathbf{r}$ . Alternatively speaking, the length (norm) of  $\mathbf{r}$  in the ODS is close to the length of  $\mathbf{r}_\Phi$  in the CSBD. A direct consequence of RIP is that the difference in length between  $\mathbf{r}_\Phi$  and  $\mathbf{r}$  is bounded by  $\delta_k$  as shown below.

$$\frac{\|\Phi\mathbf{r}\|_2^2 - \|\mathbf{r}\|_2^2}{\|\mathbf{r}\|_2^2} \leq \delta_k. \quad (6)$$

A related concept to RIP is the RCP [14], which relates the angle between two vectors in the ODS to their corresponding angle between two vectors in the CSBD. For example, the RCP condition asserts that the angle  $\theta$  between two vectors  $\mathbf{r}_1$  and  $\mathbf{r}_2$  in the ODS and the angle  $\theta_\Phi$  between their corresponding vectors  $\mathbf{r}_{\Phi 1}$  and  $\mathbf{r}_{\Phi 2}$  in the CSBD is preserved. More specifically, RCP can be described by

$$\frac{(1 - \delta_k)}{(1 + \delta_k)} \cos \theta \leq \cos \theta_\Phi \leq \cos \theta \frac{(1 - \delta_k)}{(1 + \delta_k)}. \quad (7)$$

In practice, checking if the RIP condition is satisfied is difficult if not impossible [5]. However, one can easily construct a sensing matrix  $\Phi$  by sampling independently identical distributed (i.i.d.) entries from a Gaussian distribution with 0 mean and variance  $\frac{1}{m}$  or construct  $\Phi$  by sampling i.i.d. entries  $\left(\pm \frac{1}{\sqrt{m}}\right)$  from a symmetric Bernoulli distribution. If the sensing matrix  $\Phi$  is constructed using random entries from a Gaussian or a Bernoulli distribution, then (5) is satisfied with high probability provided that

$$m \geq C \times k \times \log(L/k) \quad (8)$$

is true, where  $C$  depends on the desired application. It is interesting but not surprising to note that the lower bound on  $m$  to guarantee the RIP condition depends on the sparsity of  $\mathbf{r}$ . The fewer coefficients are used to represent the signal  $\mathbf{r}$ , the fewer samples, (small  $m$ ), are needed to recover the signal and to satisfy the RIP condition. Directly determining the minimum  $m$  bounded by (8) to satisfy (5) is not easy. One major issue is that accurately estimating the sparsity of a signal is difficult [15], [16]. As a result, in our experiments, we avoid estimating the sparsity and determine the minimum  $m$  empirically. Also, as mentioned earlier, we use the universality model of CS, which avoids estimating sparsity.

Typically, the signal recovery process is accomplished by using either the Orthogonal Matching Pursuit (OMP) or Basis Pursuit (BP) algorithms [17]–[20]. Generally speaking, the recovery process is formulated as

$$\arg \min_{\Psi \mathbf{r}} \{ \|\Psi \mathbf{r}\|_q \text{ s.t. } \Phi \mathbf{r} = \mathbf{r}_\Phi \}, \quad (9)$$

where  $\Psi \mathbf{r}$  is the sparse recovery of the original signal  $\mathbf{r}$ . A more traditional sparse-based recovery approach is

$$\arg \min_{\mathbf{r}} \{ \|\Phi \mathbf{r} - \mathbf{r}_\Phi\|_2 \text{ s.t. } \|\mathbf{r}\|_1 \leq k \}, \quad (10)$$

which requires estimating the sparsity of  $k$ . Since we completely avoid the HSI reconstruction process, we can also avoid directly estimating  $k$  by taking advantage of universality specified by (4), RIP, and RCP.

One way to think of CS is as a mapping process from the ODS to the CSBD, where a vector in the ODS is mapped to another vector in the CSBD via the sensing matrix  $\Phi$ . The recovery process is the inverse operation, where the sensed vector  $\mathbf{r}_\Phi$  is mapped back uniquely to its original vector  $\mathbf{r}$  in the ODS. Fig. 2. below shows how the mapping from the ODS to the CSBD is performed via the sensing matrix  $\Phi$  and how the lengths of the vectors are preserved via the RIP condition. The figure shows the length difference between  $\mathbf{r}_1$  and  $\mathbf{r}_2$  in the ODS is preserved in the CSBD, which also implies that the actual length (norm) of the individual vectors is also preserved.

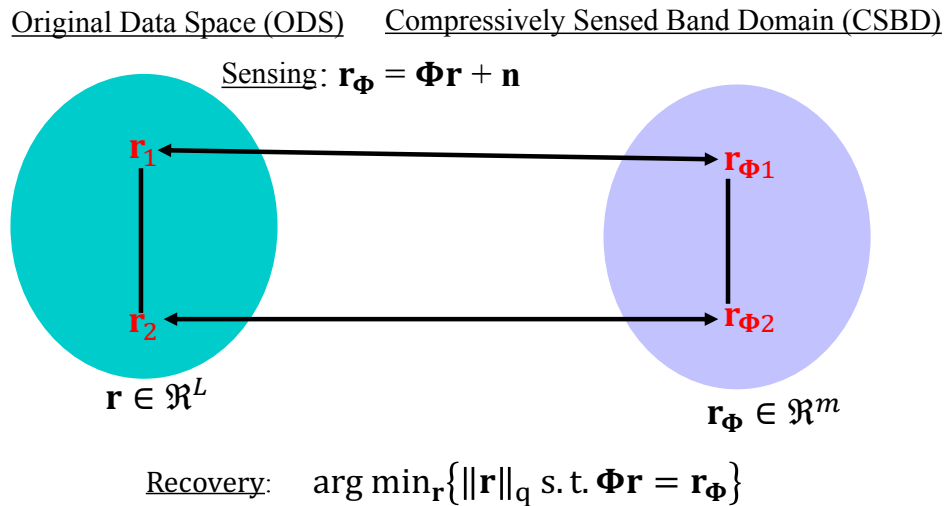


Fig. 2. Restricted Isometric Property

### 3. ORTHOGONAL SUB-SPACE PROJECTIOB PRESERVATION

The orthogonal sub-space projection (OSP) is a well-studied technique [21], which has been applied in wide range of hyperspectral applications. One such application that utilizes OSP is the Automatic Target Generation Process (ATGP). As we will show later, ATGP uses OSP to find the desired targets (endmembers) in the HSI without any prior knowledge. In ATGP, the desired target is the pixel with the maximum projected height for a given data matrix and a projection matrix.

Therefore, the ATGP techniques relies on the projection matrix and the projected heights of pixel vectors in the HSI to find the desired targets.

Our goal in this paper is to show how the sensing matrix  $\Phi$  preserves the projected heights of vectors for a given projection matrix. To show this, we will first define the data matrix and the corresponding projection matrix in the ODS. The data matrix in the ODS is defined as

$$\mathbf{U}_{(i-1)} = [\mathbf{r}_1 \ \mathbf{r}_2, \dots, \mathbf{r}_{(i-1)}], \quad (11)$$

and the corresponding projection matrix is defined as

$$\mathbf{P}_{\mathbf{U}_i}^\perp = [\mathbf{I} - \mathbf{U}_{(i-1)}(\mathbf{U}_{(i-1)}^T \mathbf{U}_{(i-1)})^{-1} \mathbf{U}_{(i-1)}^T]. \quad (12)$$

In the CSBD, the data matrix is defined as

$$\mathbf{U}_{(i-1)\Phi} = [\mathbf{r}_{1\Phi} \ \mathbf{r}_{2\Phi}, \dots, \mathbf{r}_{(i-1)\Phi}], \quad (13)$$

while the projection matrix is defined as

$$\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp = [\mathbf{I} - \mathbf{U}_{(i-1)\Phi}(\mathbf{U}_{(i-1)\Phi}^T \mathbf{U}_{(i-1)\Phi})^{-1} \mathbf{U}_{(i-1)\Phi}^T]. \quad (14)$$

However, looking at (14) closely, the data matrix  $\mathbf{U}_{\Phi(i-1)}$  can be rewritten as

$$\mathbf{U}_{(i-1)\Phi} = \Phi[\mathbf{r}_1 \ \mathbf{r}_2, \dots, \mathbf{r}_{(i-1)}] = \Phi \mathbf{U}_{(i-1)}, \quad (15)$$

and substituting this for the projection matrix  $\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp$ , we get

$$\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp = [\mathbf{I} - \Phi \mathbf{U}_{(i-1)}(\mathbf{U}_{(i-1)}^T \Phi^T \Phi \mathbf{U}_{(i-1)})^{-1} \mathbf{U}_{(i-1)}^T \Phi^T]. \quad (16)$$

With this in mind, the projected component of vector  $\mathbf{r}_\Phi$  by the projection matrix  $\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp$  is simply given by multiplying the projection matrix  $\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp$  by  $\mathbf{r}_\Phi$  as in

$$\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp \mathbf{r}_\Phi = [\mathbf{r}_\Phi - \Phi \mathbf{U}_{(i-1)}(\mathbf{U}_{(i-1)}^T \Phi^T \Phi \mathbf{U}_{(i-1)})^{-1} \mathbf{U}_{(i-1)}^T \Phi^T \mathbf{r}_\Phi]. \quad (17)$$

Recalling the fact that  $\mathbf{r}_\Phi = \Phi \mathbf{r}$ , we can write (17) as

$$\mathbf{P}_{\mathbf{U}_{i\Phi}}^\perp \Phi \mathbf{r} = [\Phi \mathbf{r} - \Phi \mathbf{U}_{(i-1)}(\mathbf{U}_{(i-1)}^T \Phi^T \Phi \mathbf{U}_{(i-1)})^{-1} \mathbf{U}_{(i-1)}^T \Phi^T \Phi \mathbf{r}]. \quad (18)$$

Letting  $\Phi^T \Phi \rightarrow \mathbf{I}$  as  $m \rightarrow L$ , we reduce (18) to



$$\mathbf{P}_{\mathbf{U}_i^\perp}^\perp \Phi \mathbf{r} \cong \left[ \Phi \mathbf{r} - \Phi \mathbf{U}_{(i-1)} (\mathbf{U}_{(i-1)}^T \mathbf{U}_{(i-1)})^{-1} \mathbf{U}_{(i-1)}^T \mathbf{r} \right]. \quad (19)$$

After factoring out the common terms  $\Phi$  and  $\mathbf{r}$ , (19) is equivalent to

$$\mathbf{P}_{\mathbf{U}_i^\perp}^\perp \Phi \mathbf{r} \cong \Phi \left[ \mathbf{I} - \mathbf{U}_{(i-1)} (\mathbf{U}_{(i-1)}^T \mathbf{U}_{(i-1)})^{-1} \mathbf{U}_{(i-1)}^T \right] \mathbf{r}. \quad (20)$$

By using the definition of  $\mathbf{P}_{\mathbf{U}_i}^\perp$  in (12), we can simplify (20) to

$$\mathbf{P}_{\mathbf{U}_i^\perp}^\perp \Phi \mathbf{r} \cong \Phi \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r}, \quad (21)$$

Squaring the norm of (21), it yields

$$\left\| \mathbf{P}_{\mathbf{U}_i^\perp}^\perp \Phi \mathbf{r} \right\|_2^2 \cong \left\| \Phi \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r} \right\|_2^2. \quad (22)$$

Using the RIP, we can write the right side of (22) as

$$\left\| \Phi \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r} \right\|_2^2 \leq \left\| \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r} \right\|_2^2 (1 + \delta_k). \quad (23)$$

Therefore, combining (21) and (23), we can assert that

$$\left\| \mathbf{P}_{\mathbf{U}_i^\perp}^\perp \Phi \mathbf{r} \right\|_2^2 \leq \left\| \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r} \right\|_2^2 (1 + \delta_k). \quad (24)$$

By using the RIP symmetry, we can claim

$$(1 - \delta_k) \left\| \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r} \right\|_2^2 \leq \left\| \mathbf{P}_{\mathbf{U}_i^\perp}^\perp \Phi \mathbf{r} \right\|_2^2 \leq \left\| \mathbf{P}_{\mathbf{U}_i}^\perp \mathbf{r} \right\|_2^2 (1 + \delta_k). \quad (25)$$

The result obtained in (25) shows that the projected length in the CSBD is bounded exactly the same way any vector length is bounded by RIP. Again, this finding was expected considering that the projection of a vector using a projection matrix results in an orthogonal vector, which has length. It is worth mention that it is this OSP preservation that allows the ATGP and any other target (endmember) finding algorithms to extract the desired targets from the HSI in the CSBD. We call this OSP preservation Restricted Orthogonal Sub-space Projection Property (ROSPP).

#### 4. COMPRESSIVE SENSING AUTOMATIC TARGET GENERATION PROCESS

In this section, we show how the concept of CS is directly applied to the Automatic Target Generation Process (ATGP) [22]–[25] algorithm for unsupervised target detection (i.e., endmember finding) by using the important CS property of RIP. The ATGP algorithm relies on the well-developed OSP concept, which is described in detail in [26]. The algorithm implements a sequence of OSP to detects targets (endmembers) imbedded in the HSI. At each iteration, it finds a pixel vectors (endmembers), which provides the maximum orthogonal projects. The research in this paper compares the CS-based ATGP with the traditional ATGP, which does not take CS into account. For convenience, the two different approaches are shown in Fig. 3. below.

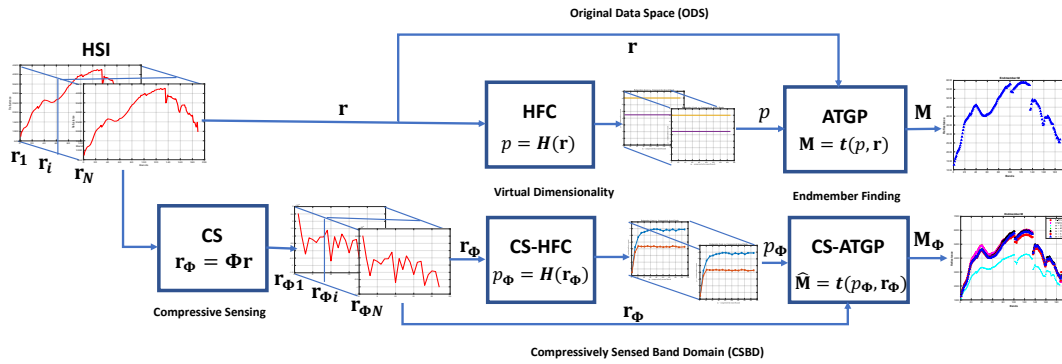


Fig. 3. Traditional AGTP setup vs. CS-ATGP setup

In CS-ATGP algorithm, the first step is determining the initial target in the HSI. This is typically done by finding the pixel with the maximum energy as seen below

$$\mathbf{t}_{0\Phi} = \max_{\mathbf{r}_\Phi}(\mathbf{r}_\Phi^T \mathbf{r}_\Phi) = \max_{\mathbf{r}_\Phi}(\mathbf{r}^T \Phi^T \Phi \mathbf{r}) \quad (26)$$

Applying the property that  $\Phi^T \Phi \rightarrow \mathbf{I}$  as  $m \rightarrow L$ , we can see that (26) approaches the regular (non-CS) ATGP given by

$$\mathbf{t}_{0\Phi} = \max_{\mathbf{r}_\Phi}(\mathbf{r}_\Phi^T \mathbf{r}_\Phi) \cong \max_{\mathbf{r}}(\mathbf{r}^T \mathbf{r}). \quad (27)$$

As a result, the length of all pixel vectors in the HSI are preserved due to RIP, which also implies that energy of all pixels is preserved. Therefore, applying CS to ATGP, does not change the first target picked, despite using very fewer CSBs,  $m$ , to determine the initial target pixel.

After determining the initial target, the next step is determining the initial data matrix given by

$$\mathbf{U}_{0\Phi} = \Phi \mathbf{U}_0 = \Phi \mathbf{t}_{0\Phi}, \quad (28)$$

whose length is also preserved due to RIP. Therefore, the initial data matrix will not be affected. In other words, the initial data matrix  $\mathbf{U}_0$  selected in the ODS is equivalent to the initial data matrix selected in the CSBD, except it will be sensed using the sensing matrix  $\Phi$ . It is also worth mentioning that if the selected initial data matrix  $\mathbf{U}_{0\Phi}$  in the CSBD is the same as the initial data matrix in the ODS, then the subsequent data matrices will also be the same in the CSBD and in the ODS. In other words,  $\mathbf{U}_{i\Phi} = \Phi \mathbf{U}_i$ . After determining the initial data matrix, the next step is figuring out the projection matrix  $\mathbf{P}_{\mathbf{U}_{0\Phi}}^\perp$ , which was defined earlier, but it is re-defined again below for convience

$$\mathbf{P}_{\mathbf{U}_{0\Phi}}^\perp = \mathbf{I} - \mathbf{U}_\Phi (\mathbf{U}_\Phi^T \mathbf{U}_\Phi)^{-1} \mathbf{U}_\Phi^T = [\mathbf{I} - \Phi \mathbf{U} (\mathbf{U}^T \Phi^T \Phi \mathbf{U})^{-1} \mathbf{U}^T \Phi^T]. \quad (29)$$

From the OSP length preservation property and proof given in (21), we claim that

$$\mathbf{P}_{\mathbf{U}_\Phi}^\perp \mathbf{r}_\Phi \cong \Phi \mathbf{P}_\mathbf{U}^\perp \mathbf{r}, \quad (30)$$

and

$$\|\mathbf{P}_{\mathbf{U}_\Phi}^\perp \mathbf{r}_\Phi\|_2 \cong \|\mathbf{P}_\mathbf{U}^\perp \mathbf{r}\|_2 \quad (31)$$

The most important step in ATGP is determining the current target pixel  $\mathbf{t}_{i_\Phi}$  after determining the current projection matrix and data matrix. In the CSBD, the current target is determined using

$$\mathbf{t}_{i_\Phi} = \max_{\mathbf{r}_\Phi} \left( (\mathbf{P}_{\mathbf{U}_{(i-1)\Phi}}^\perp \mathbf{r}_\Phi)^T (\mathbf{P}_{\mathbf{U}_{(i-1)\Phi}}^\perp \mathbf{r}_\Phi) \right) = \max_{\mathbf{r}_\Phi} \left( (\Phi \mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r})^T (\Phi \mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r}) \right), \quad (32)$$

where  $\mathbf{U}_{(i-1)\Phi} = [\mathbf{t}_{1_\Phi} \ \mathbf{t}_{2_\Phi} \ \dots \ \mathbf{t}_{(i-1)\Phi}]$ . As  $m \rightarrow L$  and using (15), it can farther be reduced to

$$\mathbf{t}_{i_\Phi} = \max_{\mathbf{r}_\Phi} \left( \mathbf{r}^T (\mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp)^T \Phi^T \Phi \mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r} \right) \cong \max_{\mathbf{r}} \left( (\mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r})^T \mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r} \right). \quad (33)$$

It is important to observe that (33) approaches the case when ATGP is executed in the ODS. In other words,  $\mathbf{t}_{i_\Phi} \cong \mathbf{t}_i$  as  $m \rightarrow L$ . Therefore, when CS is applied to ATGP using compressively sensed band vectors, it will pick the same target as when ATGP is used in ODS, as long as the sampling size  $m$  is large enough. This is a significant finding since, as it will be shown later, the CS-based ATGP only requires a very few CSBs,  $m$ , to find the same targets as the targets found in the ODS. The CS-ATGP algorithm is summarized below in Table 1.

Table 1 CS-ATGP Algorithm Summary

<b>Algorithm 1:</b> CS-ATGP Algorithm Summary
1. Initial conditions: Choose an initial target pixel vector $\mathbf{t}_{\Phi_0}$ using the method below: $\mathbf{t}_{0_\Phi} = \max_{\mathbf{r}_\Phi} (\mathbf{r}_\Phi^T \mathbf{r}_\Phi) = \max_{\mathbf{r}_\Phi} (\mathbf{r}^T \Phi^T \Phi \mathbf{r}),$ using all pixel vectors in the HSI Set $\mathbf{U}_{0_\Phi} = \mathbf{t}_{0_\Phi} = \Phi \mathbf{U}_0 = \Phi \mathbf{t}_0$ Set $i = 1$
2. At $i^{\text{th}}$ iteration apply the projection matrix to all pixel vectors in the HSI: $\mathbf{t}_{i_\Phi} = \max_{\mathbf{r}_\Phi} \left( (\mathbf{P}_{\mathbf{U}_{(i-1)\Phi}}^\perp \mathbf{r}_\Phi)^T \mathbf{P}_{\mathbf{U}_{(i-1)\Phi}}^\perp \mathbf{r}_\Phi \right) \cong \max_{\mathbf{r}} \left( (\mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r})^T \mathbf{P}_{\mathbf{U}_{(i-1)}}^\perp \mathbf{r} \right),$ where $\mathbf{U}_{(i-1)\Phi} = [\mathbf{t}_{1_\Phi} \ \mathbf{t}_{2_\Phi} \ \dots \ \mathbf{t}_{(i-1)\Phi}]$ , $\mathbf{U}_{(i-1)} = [\mathbf{t}_1 \ \mathbf{t}_2 \ \dots \ \mathbf{t}_{(i-1)}]$ .
3. Stopping rule: If $i < (n_{\text{NATGP}} - 1)$ , then $i \leftarrow i + 1$ and go step 2. Otherwise, $[\mathbf{t}_{1_\Phi} \ \mathbf{t}_{2_\Phi} \ \dots \ \mathbf{t}_{n_{\text{NATGP}\Phi}}]$ are the desired $n_{\text{NATGP}}$ targets in the HSI.

## 5. EXPERIMENTAL IMAGES

In order to show how the concept of CS is applied to unsupervised target detection using ATGP, experiments were designed. The designed experiments used both real and synthetic hyperspectral images. For the proposed experiments, one real image and two synthetic images were used. Each hyperspectral image is described below.

### 5.1 AVIRS Cuprite Real Image

The Airborne Infrared Imaging Spectrometer (AVIRIS) Cuprite hyperspectral image is one of the most widely used image in the hyperspectral community. The image contains a spatial resolution of 20 m collected using 224 bands (only  $L = 189$  used for experiment) and 10 nm spectral resolution in the range of 0.4-2.5  $\mu\text{m}$ . The image contains five pure pixels that can be identified and located. The pure pixels are alunite (A), buddingtonite (B), calcite (C), kaolinite (K), and muscovite (M). The Cuprite image has  $350 \times 350$  pixels (i.e.,  $N = 122,500$ ). The details of the Cuprite image are also covered in [27]. Fig. 4. below shows the Cuprite image.

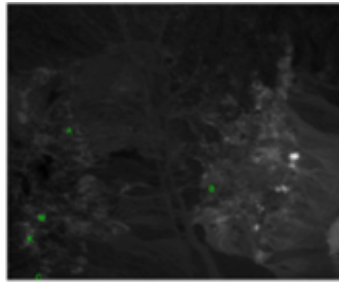


Fig. 4. AVIRIS Cuprite Real Image

### 5.2 Target Implantation 2 (TI2) Synthetic Image

The Target Implantation 2 (TI2) image is a  $200 \times 200$  (i.e.,  $N = 4,000$ ) synthetic image, which is simulated using the real Cuprite image. The TI2 image is simulated by implanting (inserting) target pixels into noise background, which were corrupted by adaptive white Gaussian noise and made to maintain a signal-to-noise (SNR) ratio of 20:1 after removing the background pixels. Therefore, the image simulates the cases where clean targets are present in noisy image background. The image contains  $5 \times 5 = 25$  pixel panel as seen in Fig. 5. Each row in the image panel contains the same mineral signature and each column in the image panel contains the same size. The pixels in the first column are  $4 \times 4$  pure pixels, while the pixels in the second column are pure pixels with size of  $2 \times 2$ . The third column contains  $2 \times 2$  mixed pixels, while the fourth and the fifth columns contain sub-pixels. The details of the TI2 image are covered in [27].

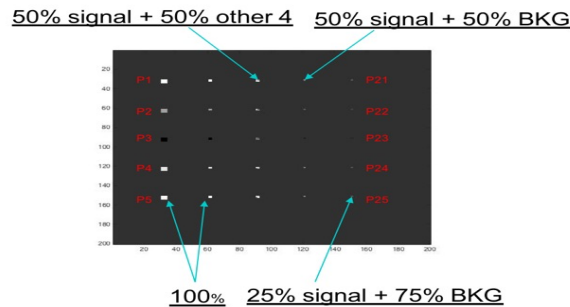


Fig. 5. TI2 Synthetic Image

### 5.3 Target Embeddedness 2 (TE2) Synthetic Image

The Target Embeddedness 2 (TE2) is the same as the TE1 image, except the way the image was simulated is little different. For the TE1 image, the background pixels were removed before inserting the pure pixels. For the TE2 image, the background pixels are not removed. Instead, the targets are embedded or superimposed into the background pixels without removing the background pixels. Therefore, the final simulated TE2 image has clean targets imbedded into a noisy background as seen in Fig. 6. The details of the TE2 image are covered in [27].

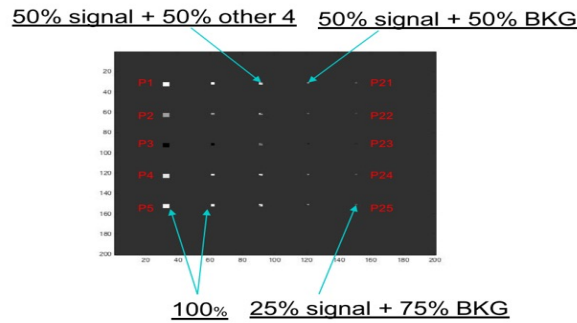


Fig. 6. TE2 Synthetic Image

## 6. EXPERIMENTAL RESULTS

### 6.1 Experimental Results for TE2 Image

Fig. 7. below shows the result of applying CS to ATGP. As it can be seen from Fig. 7(b). through Fig. 7(f)., applying CS to ATGP was successful in finding the targets (endmembers). However, in Fig. 7(b)., when the number of CSBs,  $m$ , was only 6, the spatial locations of the endmembers were off compared to their true locations. After increasing  $m$ , however, the locations of the endmembers were very close to their truth locations. It is interesting to note that only  $m = 46$  or less (see Fig. 7(c).) CSBs were needed to exactly determine the endmember locations. The third endmember's location, however, was not the same as its true locations, despite using all CSBs (i.e.,  $m = L$ ). This is the case because the third endmember's spectral values are very close to the image background. In other words, the third target is background.

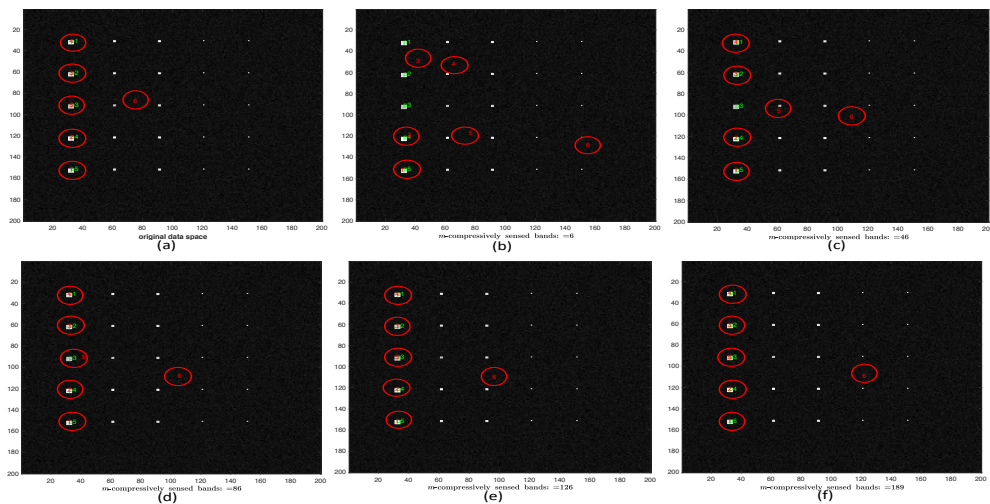


Fig. 7. TE2: (a) Endmembers in ODS (b)-(f) Found endmembers in CSBD for different value of  $m$

Fig. 8. below depicts the spectral angle map (SAM) [28] results between the found targets and the truth targets. The same figure also shows reflectance of the found targets (endmembers). The first figure, Fig. 8(a)., shows the SAM values calculated by using the endmembers extracted by applying CS to ATGP in the CSBD for different values of  $m$  and using the ground truth endmembers. Looking at the figure, we can see that the spectral signatures of the extracted endmembers using CS-ATGP are very close to the spectral signatures of the true endmembers. The rest of the figures, Fig. 8(b). through Fig. 8(f)., show the radiance of the five endmembers in the first left column (top to bottom) of the TE2 HSI. By looking at the figures, we can see that the radiance of the extracted targets (endmembers) approach that of the ground truth radiance for  $m \leq 46$ . This is a significant result considering that it only took  $m \leq 46$  to find the desired targets in the CSBD, while processing less data compared to the ODS.

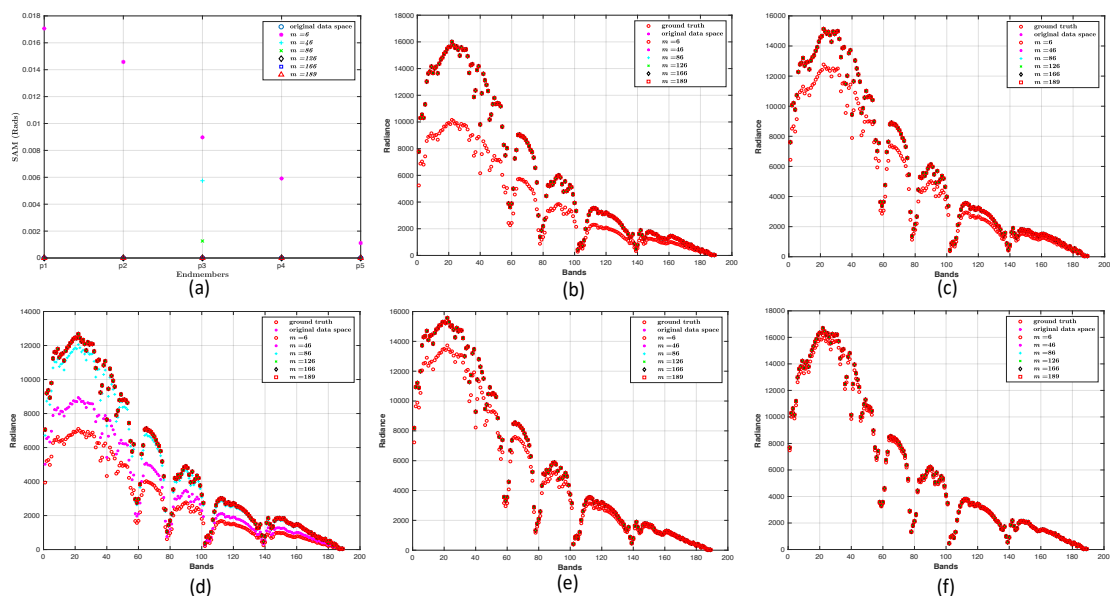


Fig. 8. TE2: (a) Found endmembers' SAM for different value of  $m$  (b) Found endmembers' radiance for different value of  $m$

The algorithm computational times obtained by CS-ATGP in the CSBD as well as the case when CS is not applied in the ODS are shown in Fig. 9. below. Comparing the two cases, it easily seen that when the CSBs size,  $m$ , is small, the CS-ATGP case takes less computational time to run the algorithm, despite producing the correct endmember locations. As expected, as  $m$  increases, it approaches the case seen in the ODS.

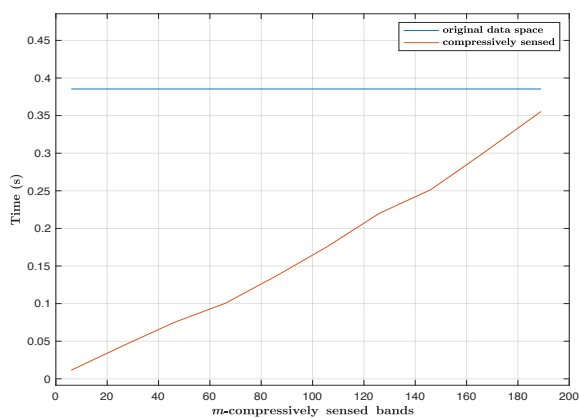


Fig. 9. TE2: CS-ATGP computational time for different values of  $m$

## 6.2 Experimental Results for TI2 Image

In order to increase the test data diversity, the synthetic TI2 image was also used for the experiment. The results of using the TI2 image are shown in Fig. 10. through Fig. 12. Paying a close attention to the figures, we can see that the same results are obtained as using TE2. In other words, it takes a small number of CSBs ( $m \leq 46$ ) to find the desired targets. In addition, it takes less computational time to find the desired target in the CSBD compared with the ODS.

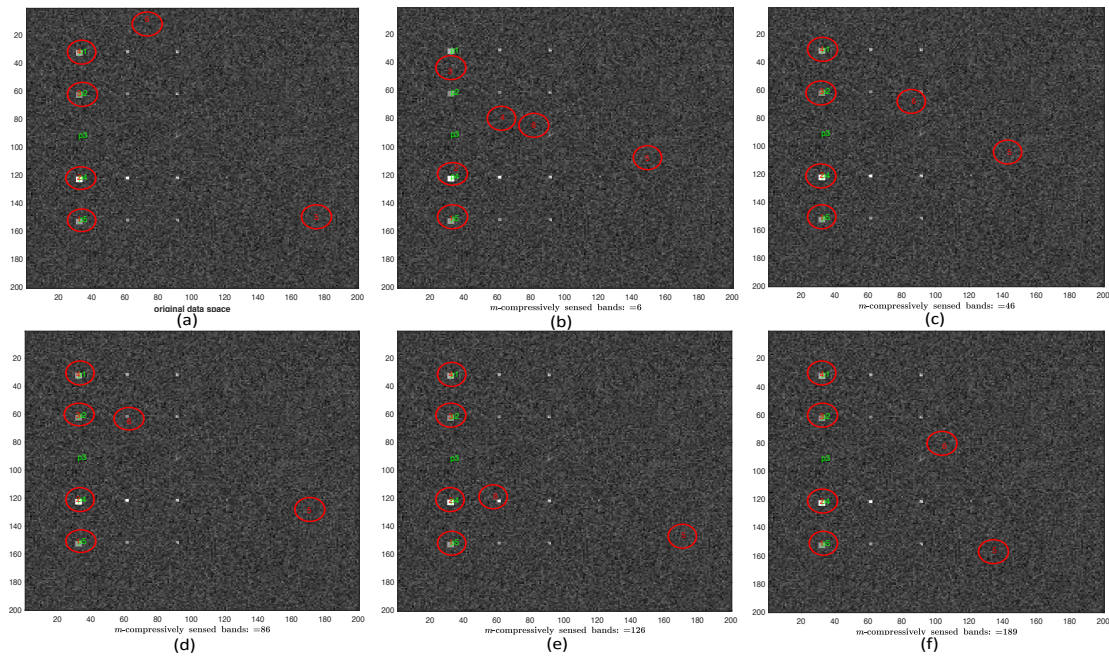


Fig. 10. TI2: (a) Endmembers in ODS (b)-(f) Found endmembers in CSBD for different value of  $m$

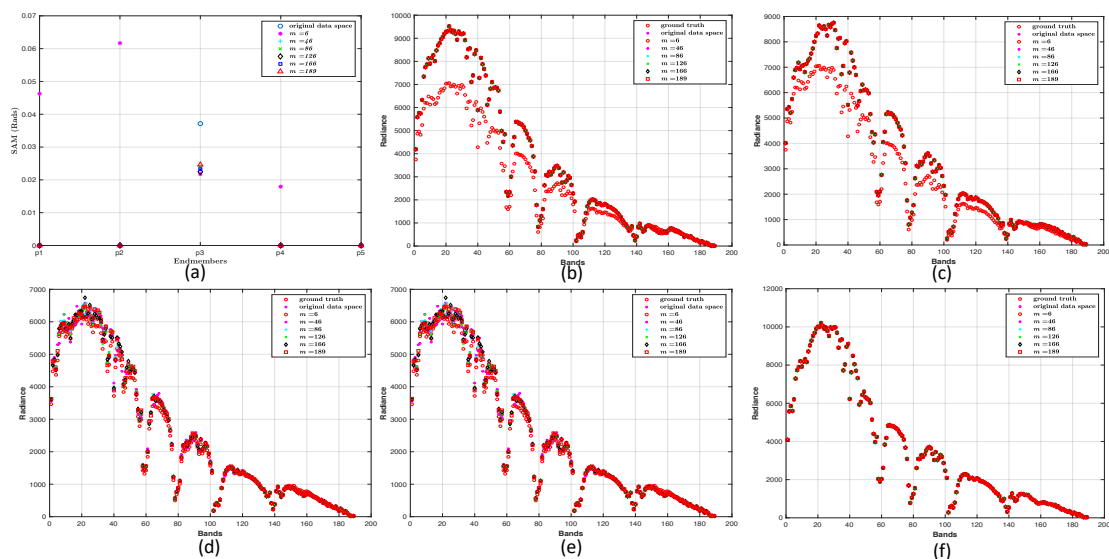


Fig. 11. TI2: (a) Found endmembers' SAM for different value of  $m$  (b)-(f) Found endmembers' radiance for different value of  $m$

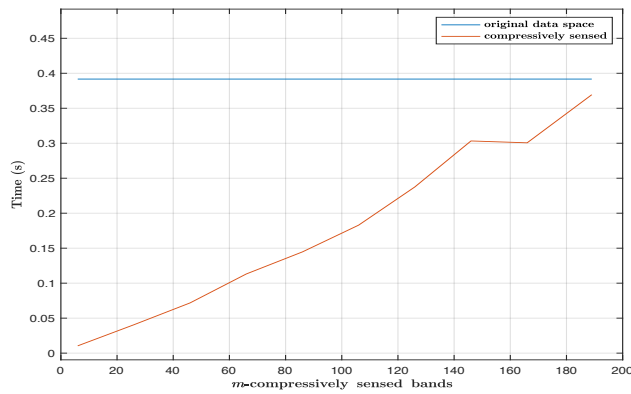


Fig. 12. T12: (a) CS-ATGP computational time for different values of  $m$

### 6.3 Experimental Results for Cuprite Image

The third HSI used for the ATGP experiment was the real Cuprite image. Figures Fig. 13. through Fig. 15. below show the experimental results. The first figure of Fig. 13., Fig. 13(a), shows the case when CS is not applied (i.e., ODS). The rest of the figures show the case when CS is used with different CSBs sizes. Looking at the rest of the figures, we can see that when  $m = 22$ , the three endmember locations (M), (K), and (C) are very close to their ground truth locations. However, the other two endmember locations (A) and (B) are not close to their ground truth locations. This is the case because there are other pixels in the image that have similar spectral signatures (i.e., reflectance) as the two endmembers, (A) and (B). In this case, the CS-based ATGP is finding other pixels that have spectral signatures that are very close to the true endmembers. At any rate, it is very evident that the CS-ATGP in CSBD requires a very few CSBs to find the desired targets. In doing so, it utilizes less data, and as a result, less storage memory compared to the ODS.

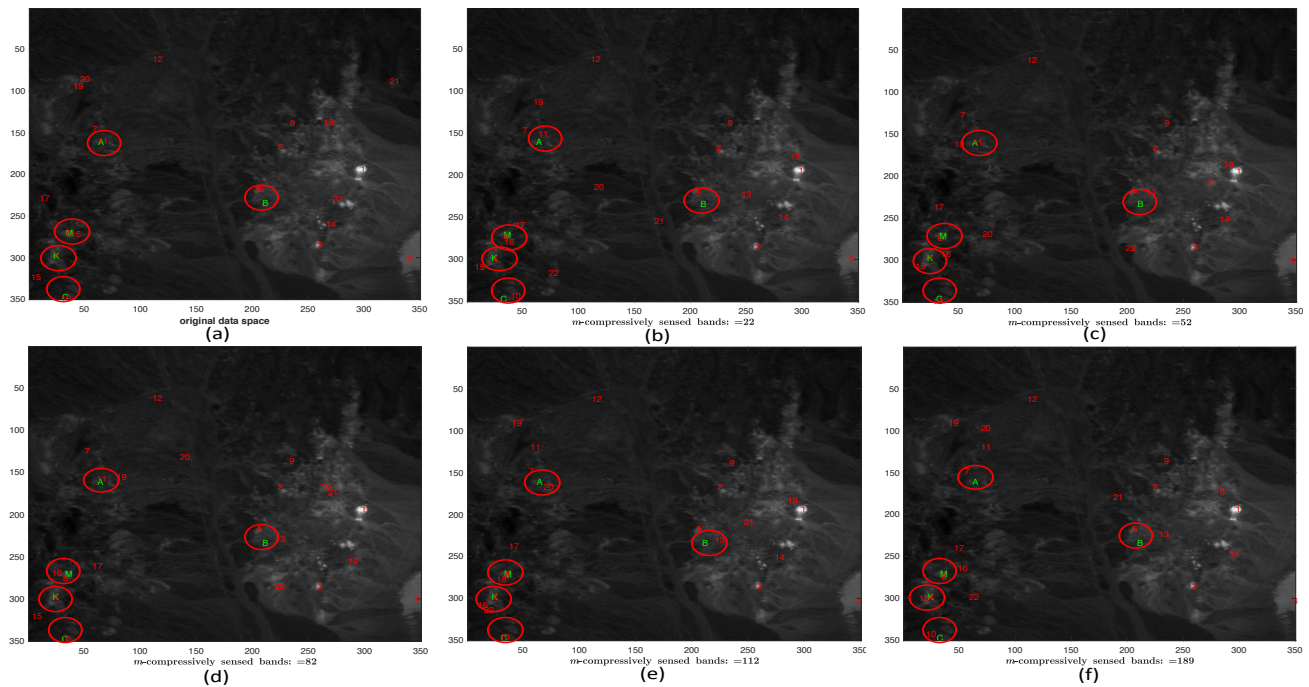


Fig. 13. Cuprite: (a) Endmembers in ODS (b)-(f) Found endmembers in CSBD for different value of  $m$



The calculated SAM values and the reflectance for the found targets (endmembers) for the Cuprite image are shown in Fig. 14. The SAM results are depicted in Fig. 14(a), where the SAMs are calculated for different values of CSBs,  $m$ , in the CSBD. In addition, the SAM results for the ODS is also shown in the same figure. Again, the same type of performance is observed again, where the CS-ATGP is successful in finding the targets (endmembers) in the Cuprite image, despite using a very few numbers of CSBs. This observation was made by comparing the SAM results obtained in CSBD with the SAM values seen from the ODS. The reflectance for the found targets (A), (B), (C), (K), and (M) in the CSBD are shown in Fig. 14(b) through Fig. 14(f), respectively. In addition, the same figures show the reflectance for the ground truth targets and the reflectance for the targets obtained in the ODS. Again, the reflectance produces by the CS-ATGP are very similar to the reflectance obtained in the ODS. However, the CS-ATGP requires less data to produce results that are similar to results produced in the ODS.

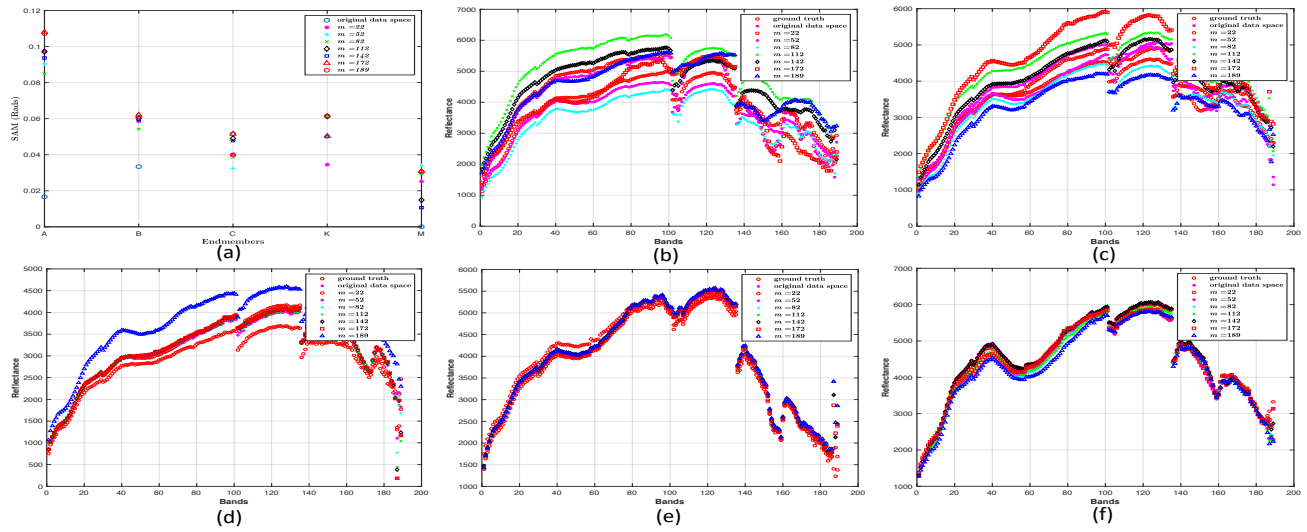


Fig. 14. Cuprite: (a) Found endmembers' SAM different value of  $m$ , (b)-(f) Found endmembers' reflectance for different value of  $m$

Fig. 15. below shows the algorithm computational times for CS-ATGP in the CSBD and the non-CS ATGP in the ODS. As expected, the CS-ATGP takes less processing time to produce results that are comparable with results obtained in the ODS, the traditional approach, where CS is not considered. However, as the value of the CSBs is increased, the computational time of the CS-ATGP increases linearly.

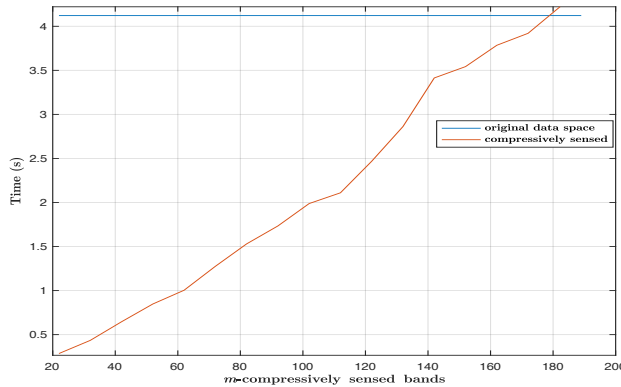


Fig. 15. Cuprite: (a) CS-ATGP computational time for different values of  $m$

## 7. CONCLUSIONS

In this paper, we connected the concept of CS with one of the most popular targets finding (generation) algorithms, ATGP, in the hyperspectral community. Mathematically, we demonstrated the role the sensing matrix plays in ATGP, and the impact it has on the projection matrix as well as the projected vector's length. We also showed that when the projected vector's length is preserved in the CSBD, the energy of the projected vector is also preserved. In addition, we showed that, when the number of CSBs is high relative to the number of full bands used in the ODS, the objective function used by the CS-ATGP algorithm in the CSBD approaches the same objective function used in the ODS (non-CS). However, experimentally and using both synthetic and real images, we showed that it actually takes a very few numbers of CSBs to find the same targets as the targets found in the ODS. This implied that the CS-ATGP will use less memory since it processes less data compared to the traditional non-CS ATGP. Our experimental results also showed that the CS-ATGP requires less computational time to find the same targets found in the ODS using the non-CS ATGP.

## REFERENCES

- [1] B. Zhang and L. Zheng, "HYPERSPETRAL REMOTE SENSING TECHNOLOGY AND APPLICATIONS IN CHINA," p. 10.
- [2] T. Adão *et al.*, "Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry," *Remote Sensing*, vol. 9, no. 11, p. 1110, Oct. 2017.
- [3] N. Keshava, J. Kerekes, D. Manolakis, and G. Shaw, "An Algorithm Taxonomy for Hyperspectral Unmixing," p. 22.
- [4] N. Keshava, "A Survey of Spectral Unmixing Algorithms," vol. 14, no. 1, p. 24, 2003.
- [5] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [6] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. New York, NY: Springer New York, 2013.
- [7] R. M. Willett, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Sparsity and structure in hyperspectral imaging: Sensing, reconstruction, and target detection," *IEEE signal processing magazine*, vol. 31, no. 1, pp. 116–126, 2014.
- [8] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," *arXiv preprint arXiv:0901.3403*, 2009.
- [9] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, "Introduction to compressed sensing," *preprint*, vol. 93, no. 1, p. 2, 2011.
- [10] N. Amhed, T. Natarajan, and K. R. Rao, "Discrete Cosine Transform," *IEEE transactions on Computers*, vol. 100, no. 1, pp. 90–93, 1974.
- [11] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on image processing*, vol. 1, no. 2, pp. 205–220, 1992.
- [12] T. GALLY and M. E. PFETSCH, "Computing restricted isometry constants via mixed-integer semidefinite programming," *preprint, submitted*, 2016.
- [13] R. Baraniuk, M. Davenport, R. De Vore, and M. B. Wakin, "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Constructive Approximation*, vol. 3, pp. 253–263, 2008.
- [14] L.-H. Chang and J.-Y. Wu, "Achievable angles between two compressed sparse vectors under RIP-induced norm/distance constraints," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*, 2014, pp. 523–528.
- [15] M. E. Lopes, "Compressed sensing without sparsity assumptions," *arXiv preprint arXiv:1507.07094*, 2015.
- [16] M. Lopes, "Estimating Unknown Sparsity in Compressed Sensing," *CoRR*, vol. abs/1204.4227, 2012.
- [17] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Puschel, "Distributed Basis Pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1942–1956, Apr. 2012.
- [18] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [19] E. Candes and J. Romberg, "l1-magic: Recovery of sparse signals via convex programming," *URL: www.acm.caltech.edu/l1magic/downloads/l1magic.pdf*, vol. 4, p. 14, 2005.
- [20] C. Li, L. Ma, Q. Wang, Y. Zhou, and N. Wang, "Construction of Sparse Basis by Dictionary Training for Compressive Sensing Hyperspectral Imaging," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, pp. 1442–1445, 2013.
- [21] C. I Chang, "Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis," *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, vol. 43, no. 3, 2005.
- [22] X. Li, B. Huang, and K. Zhao, "Massively Parallel GPU Design of Automatic Target Generation Process in Hyperspectral Imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2862–2869, Jun. 2015.
- [23] S. Bernabe, S. Lopez, A. Plaza, R. Sarmiento, and P. G. Rodriguez, "FPGA Design of an Automatic Target Generation Process for Hyperspectral Image Analysis," in *2011 IEEE 17th International Conference on Parallel and Distributed Systems*, Tainan, Taiwan, 2011, pp. 1010–1015.
- [24] Y. Li, C. Gao, and C.-I Chang, "Progressive band processing of automatic target generation process," in *2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Tokyo, Japan, 2015, pp. 1–4.
- [25] Cheng Gao and Chein-I Chang, "Recursive automatic target generation process for unsupervised hyperspectral target detection," in *2014 IEEE Geoscience and Remote Sensing Symposium*, Quebec City, QC, 2014, pp. 3598–3601.
- [26] C.-I Chang, *Real-Time Progressive Hyperspectral Image Processing*. New York, NY: Springer New York, 2016.
- [27] Chein-I Chang, *Hyperspectral Data Exploitation: Theory and Applications*. Hoboken, NJ: John Wiley and Sons, 2007.
- [28] C.-I Chang, "New hyperspectral discrimination measure for spectral characterization," *Optical Engineering*, vol. 43, no. 8, p. 1777, Aug. 2004.