Fast OMP reconstruction for compressive hyperspectral imaging using joint spatial-spectral sparsity model

Liu Haiying[a], Chen Rongli[b], Wang Yajun[b], Lv Pei*[b]
[a]School of Information Engineering, Chang'an University, Xi'an, China 710064; [b]Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China 710119

ABSTRACT

Hyperspectral imaging typically produces huge data volume that demands enormous computational resources in terms of storage, computation and transmission, particularly when real-time processing is desired. In this paper, we study a low-complexity scheme for hyperspectral imaging completely bypassing high-complexity compression task. In this scheme, compressive hyperspectral data are acquired directly by a device similar to the single-pixel camera based on the principle of compressive sensing (CS). To decode the compressive data, we propose a flexible recovery strategy by taking advantage of the joint spatial-spectral correlation model of hyperspectral images. Moreover, a thorough investigation is analytically conducted on compressive hyperspectral data and we find that the compressive data still have strong spectral correlation. To make the recovery more accurate, an adaptive spectral band reordering algorithm is directly added to the compressive data before the reconstruction by making best use of spectral correlation. The real hyperspectral images are tested to demonstrate the feasibility and efficiency of the proposed algorithm. Experimental results indicate that the proposed recover algorithm can speed up the reconstruction process with reliable recovery quality.

Keywords: Hyperspectral images, compressive sensing, orthogonal matching pursuit, joint sparsity

1. INTRODUCTION

Hyperspectral imaging refers to imaging the electromagnetic (EM) properties reflected or emitted by a scene or an object, over possible hundreds of contiguous spectral bands. It is a crucial tool to identify and quantify distinct components consisting of the observed scene. Hyperspectral imaging has a wide range of applications such as mineral exploration, forest monitoring, military surveillance and tissue spectroscopy in medicine. The price to pay for such high spatial and spectral resolution is to handle huge data size. It is particularly difficult to process directly and to transmit hyperspectral data cubes in real time or near real time. Prior to storage or transmission, an additional compression step is conventionally taken place to compact the data volume as small as possible without much degrading image quality. Obviously, this oversampling followed massive dumpling acquisition process is wasteful for sensor resource or power consumption, especially for particular applications, where large detector arrays are too expensive or computational resource are limited. Compressive sensing or compressed sampling (CS) recently proposed by Candes[1] and Donoho[2] is a novel sampling theorem for data acquisition. The CS theorem states that a sparse or compressible signal can be recovered from highly incomplete sets of linear measurements (i.e. far fewer than the number dictated by Shannon/Nyquist theorem) by a specially designed nonlinear recovery algorithm. In CS, sampling is performed by directly computing inner products between a signal and a set of random sequences. That is to say, the CS encoder is very simple and energy efficient. The CS decoder is more complicated because a large scale convex optimization problem must be solved to recover the original signal from fewer samples. Shifting complexity from onboard encoder to offline decoder is necessary for severely resource deprived applications, such as aerospace remote sensing. Another potential advantage of CS imaging is that it can work easily in low light environment or at wavelength outside the visible light due to the use of only one or few photon detectors. Thus, more sensitive detectors can be used to enhance imaging quality. Plenty of potential applications of the CS theorem have been made for compressive imaging[3], remote sensing[4] and spectral imaging[5]. The prototype of single-pixel digital camera based on CS has been presented by Duarte, *et al.*[3], which is different from the conventional cameras based on CCD (or CMOS). It uses a single photon detector to sequentially acquire random linear measurements and then use an optimization algorithm to recover the original signal from a small set of measurements. In [6], active illumination was added to the CS camera to extend its applications, such as pixel-

* Corresponding author: Lv Pei, E-mail: luupi@163.com

-level programmable gain imaging. Beside the visible spectrum, the single-pixel CS camera architecture has been used for imaging in the terahertz[7], and the short-wave infrared[8] spectrum. Ma[9] demonstrated two possible CS systems for different applications in remote sensing to reduce the cost of data acquisition, which are single-pixel but multitime (SPMT) imaging and multipixel but single-time (MPST) imaging. The former carries out sequential imaging as the single-pixel camera does, and the latter captures random projections within a single exposure by using an optical phase mask that is placed on a lens. In both systems, an iterative curvelet thresholding algorithm was used for better recovery quality. To achieve compressive spectral imaging, two classes of imagers dubbed the coded aperture snapshot spectral imager (CASSI) based on the principle of CS have been developed, i.e. single disperser CASSI (SD-CASSI) systems[10] and dual disperser CASSI (DD-CASSI) systems[11]. August *et al.*[12] proposed an efficient system for compressive hyperspectral imaging to reduce computational complexity by randomly separable encoding in both the spatial and the spectral domains. Li *et al.*[13] proposed an efficient CS unmixing scheme for hyperspectral images. Other CS applications include MRI imaging[14], ISAR imaging[15], compressive motion tracking[16] and wireless communications[17], just to mention a few.

For a specific CS imaging system, how to design an efficient reconstruction algorithm is an important problem. Generally, the faster we recover the signal, the lower reconstruction performance we get. There is a tradeoff between reconstruction speed and recovery performance for one algorithm. Recently, distributed compressive sensing (DCS) has been put forward by Baran, *et al.*[18] to jointly recover statistically correlated signals, while CS encoding is individually applied to each signal. The DCS theory rests on a concept termed joint sparsity – the sparsity of the entire signal ensemble. There are three joint sparsity models (JSMs) presented in [18]: sparse common component with innovations (JSM1), common sparse supports (JSM2) and nonsparse common component with sparse innovations (JSM3). The DCS theory has been successfully applied to recover color images[19] and video sequence[20]. Hyperspectral images are captured of the same scene over contiguous spectral bands. They have similar structural information and can be regarded as a joint sparse signal ensemble. The DCS theory can be easily applied to hyperspectral images. In [21], JSM1 was applied to hyperspectral image compression.

In this paper, we study the reconstruction problem of the compressive hyperspectral imaging scheme with low imaging cost. In this scheme, compressive hyperspectral data are captured directly by a typical SPMT system. To efficiently decode the compressive data, we propose a flexible recovery strategy by taking advantage of prior knowledge that hyperspectral images have common sparse supports (JSM2) to reduce complexity. Secondly, a thorough investigation is originally carried out on the compressive hyperspectral data. To make the recovery more precise, an adaptive grouping algorithm with simple spectral band reordering algorithm is added to the compressive data before the reconstruction by making use of strong spectral correlation of the compressive hyperspectral data.

This paper is organized as follows. Section 2 contains the background of this work, notations and definitions. Section 3 briefly describes the compressive hyperspectral imaging scheme and gives a thorough correlation analysis of compressive hyperspectral data. In Section 4, we describe the recovery algorithm we propose. Section 5 presents some experimental results based on the real hyperspectral images. Finally, Section 6 gives concluding remarks.

2. BACKGROUND

2.1 Notations and definitions

We denote (column-) vectors and matrices by lowercase and uppercase boldface characters, respectively. The n-th column of a matrix $\mathbf{A}$ is $\mathbf{A}^n$.The (m,n)-th element of matrix $\mathbf{A}$ is $\mathbf{A}(m,n)$. The n-th element of a vector $\mathbf{v}$ is $\mathbf{v}(n)$. The transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^{\mathrm{T}}$. The notation $\|\mathbf{v}\|_{\ell_0}$ denotes the number of nonzero elements of a vector $\mathbf{v}$. The notation $\|\mathbf{v}\|_{\ell_1}$ denotes the ℓ_1-norm of a vector $\mathbf{v}$ and is defined as $\|\mathbf{v}\|_{\ell_1} \triangleq \sum_i |\mathbf{v}(i)|$. The notation $\|\mathbf{v}\|_{\ell_2}$ denotes the Euclidean norm of a vector $\mathbf{v}$ and is defined as $\|\mathbf{v}\|_{\ell_2} \triangleq \sqrt{\sum_i |\mathbf{v}(i)|^2}$. The notation $x \sim \mathrm{N}(\mu, \sigma^2)$ denotes a Gaussian random variable x with mean μ and variance σ^2.

The notation $C_{\mathbf{A}}$ denote the covariance matrix of $\mathbf{A}$. The (m,n)-th element of matrix $C_{\mathbf{A}}$ is $C_{\mathbf{A}}(m,n)$, which represents the covariance of columns m and n in $\mathbf{A}$. The notation $R_{\mathbf{A}}$ denote the correlation coefficient matrix of $\mathbf{A}$. The (m,n)-th element of matrix $R_{\mathbf{A}}$ is $R_{\mathbf{A}}(m,n)$, which represents the correlation coefficient of columns m and n of $\mathbf{A}$.

The notation Ev denotes the mathematical expectation of a variable v. Assume that $\mathbf{v}$ is an N-length observation vector of variable v. For simplicity, we use an unbiased estimate $\mu_{\mathbf{v}}$ to replace Ev, i.e. $\mu_{\mathbf{v}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{v}(n)$.

2.2 Compressive sensing

The CS theorem is a novel sampling theorem introduced in [1]-[2]. It states that sparse or compressible signals can be reconstructed from far fewer linear measurements than Shannon/Nyquist theorem suggests. A signal $\mathbf{x} \in \mathbb{R}^{N\times 1}$ is said sparse if it has a sparse representation in some domain $\Psi \in \mathbb{R}^{N\times N}$, i.e. $\mathbf{x}=\Psi\theta$, $\|\theta\|_{\ell_0} = K$, $K \ll N$, and for compressible signals, the transform coefficient vector θ has only K significant elements and the rest $(N-K)$ elements are small. One can use a measurement matrix $\Phi \in \mathbb{R}^{M\times N}$ to acquire a smaller vector $\mathbf{y} \in \mathbb{R}^{M\times 1}$, i.e. $\mathbf{y} = \Phi\mathbf{x}$, $K < M < N$. If the measurement matrix Φ is incoherent with Ψ, i.e. $\Phi\Psi$ satisfies a sufficient condition named restricted isometry property (RIP)[1], the sparse coefficient vector θ can be accurately recovered by solving a sparse-promoting convex optimization problem using at last $M = K+1$ measurements.

$$\min \|\theta\|_{\ell_0} \quad s.t. \quad \Phi\Psi\theta = \mathbf{y} \tag{1}$$

Since the ℓ_0- norm minimization is an NP-hard problem, one can resort to an easier ℓ_1-norm minimization by linear programming

$$\min \|\theta\|_{\ell_1} \quad s.t. \quad \Phi\Psi\theta = \mathbf{y} \tag{2}$$

To accurately recover θ, the ℓ_1-norm minimization needs more measurements, i.e. $M = O(K\log(N/K))$. Many algorithms have been proposed to recover signal $\mathbf{x}$ from its measurements $\mathbf{y}$, e.g. linear programming (LP)[22] and orthogonal matching pursuit (OMP) algorithm[23]. For a detailed overview on recovery algorithms, please refer to[24].

3. COMPRESSIVE HYPERSPECTRAL IMAGING

3.1 CS encoding for hyperspectral images

Let $\mathbf{X} = [\mathbf{x}^1,...,\mathbf{x}^{N_b}] \in \mathbb{R}^{N_p\times N_b}$ denote the matrix representing of a hyperspectral image cube, where $\mathbf{x}^i \in \mathbb{R}^{N_p\times 1}$ is the vectorized version of an image at the i-th band, N_p denotes the number of pixels and N_b denotes the number of spectral bands. We may use specially designed measurement matrix for each spectral band. For convenience, we use the same measurement matrix $\Phi \in \mathbb{R}^{M\times N_p}$ to randomly sample all the bands, where $M < N_p$ is the number of samples for each band. The CS encoding for each hyperspectral image can be described as:

$$\mathbf{y}^i = \Phi\mathbf{x}^i, \quad i \in \{1,2,...,N_b\} \tag{3}$$

Where $\mathbf{y}^i$ denotes the compressive data vector of the i-th band.

Let $\mathbf{Y} = [\mathbf{y}^1,...,\mathbf{y}^{N_b}] \in \mathbb{R}^{M\times N_b}$ denote the compressive hyperspectral data matrix. The CS encoding for hyperspectral images can be rewritten as:

$$\mathbf{Y} = \Phi\mathbf{X} \tag{4}$$

3.2 Correlation analysis

For matrices $\mathbf{X}$ and $\mathbf{Y}$, define the correlation coefficient of columns (bands) i and j of the compressive hyperspectral data matrix $\mathbf{Y}$:

$$R_{\mathbf{Y}}(i,j) = \frac{C_{\mathbf{Y}}(i,j)}{\sqrt{C_{\mathbf{Y}}(i,i)C_{\mathbf{Y}}(j,j)}} = \frac{E(\mathbf{y}^i\mathbf{y}^j) - E\mathbf{y}^i E\mathbf{y}^j}{\sqrt{[E(\mathbf{y}^i)^2 - (E\mathbf{y}^i)^2][E(\mathbf{y}^j)^2 - (E\mathbf{y}^j)^2]}} \tag{5}$$

And define the correlation coefficient of columns i and j of the hyperspectral image matrix $\mathbf{X}$:

$$R_{\mathbf{X}}(i,j) = \frac{C_{\mathbf{X}}(i,j)}{\sqrt{C_{\mathbf{X}}(i,i)C_{\mathbf{X}}(j,j)}} = \frac{E(\mathbf{x}^i\mathbf{x}^j) - E\mathbf{x}^i E\mathbf{x}^j}{\sqrt{[E(\mathbf{x}^i)^2 - (E\mathbf{x}^i)^2][E(\mathbf{x}^j)^2 - (E\mathbf{x}^j)^2]}} \tag{6}$$

Without loss of generality, assume that the measurement matrix Φ is defined as a Gaussian random sampling matrix, whose entries are independent and identically distributed (i.i.d) Gaussian random variables, i.e. $\Phi(m,n) \sim \mathrm{N}(0, \frac{1}{M})$, and set the squared column norm of Φ to 1, i.e. $\sum_{m=1}^{M} |\Phi(m,n)|^2 = 1$. According to the law of large number[25], if M is large enough, we know that $\frac{1}{M}\sum_{m=1}^{M}\Phi(m,n) = E[\Phi(m,n)] = 0$ with overwhelming probability. Since

$$E\mathbf{y}^i = \frac{1}{M}\sum_{m=1}^{M}\mathbf{y}^i(m) = \frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N_p}\Phi(m,n)\mathbf{x}^i(n) = \sum_{n=1}^{N_p}\mathbf{x}^i(n)\left(\frac{1}{M}\sum_{m=1}^{M}\Phi(m,n)\right) = 0$$

and

$$E(\mathbf{y}^i\mathbf{y}^j) = \frac{1}{M}\sum_{m=1}^{M}\mathbf{y}^i(m)\mathbf{y}^j(m) = \frac{1}{M}\sum_{m=1}^{M}\left(\sum_{n=1}^{N_p}\Phi(m,n)\mathbf{x}^i(n)\right)\left(\sum_{n=1}^{N_p}\Phi(m,n)\mathbf{x}^j(n)\right)$$

$$= \frac{1}{M}\sum_{n=1}^{N_p}x^i(n)x^j(n)\left(\sum_{m=1}^{M}|\Phi(m,n)|^2\right) = \frac{N_p}{M}E(\mathbf{x}^i\mathbf{x}^j)$$

Hence, Eq. (5) can be rewritten by

$$R_{\mathbf{Y}}(i,j) = \frac{E(\mathbf{x}^i\mathbf{x}^j)}{\sqrt{E(\mathbf{x}^i)^2 E(\mathbf{x}^j)^2}} \tag{7}$$

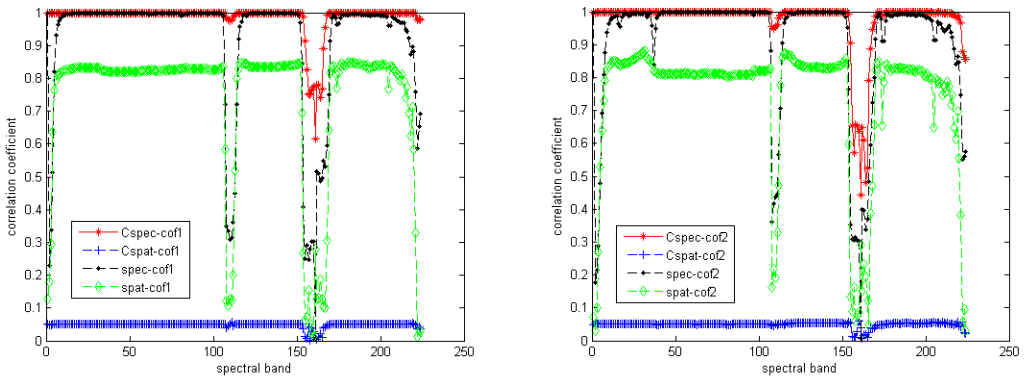
If all the bands $\mathbf{x}^i$ from the hyperspectral image matrix $\mathbf{X}$ have zero means, i.e. $E\mathbf{x}^i = 0$, $i = 1,...,N_p$, compared with Eq. (5) and Eq. (7), we have $R_{\mathrm{Y}}(i,j) = R_{\mathrm{X}}(i,j)$. It means compressive hyperspectral data have the same spectral correlation as hyperspectral images have. The CS encoding does not destroy the spectral correlation structure at all. Otherwise, when $E\mathbf{x}^i \neq 0$, we only consider the bands that have strong spectral correlations with each other. These bands can be regarded as a stationary process[26]. Assume that $\mathbf{x}^i$ and $\mathbf{x}^j$ are two bands from these bands, one has $E\mathbf{x}^i = E\mathbf{x}^j = \mu$, $E(\mathbf{x}^i)^2 = E(\mathbf{x}^j)^2 = R$, where μ and R are constant numbers. By subtracting Eq. (6) from Eq. (7), we have

$$R_{\mathbf{Y}}(i,j) - R_{\mathbf{X}}(i,j) = [1 - R_{\mathbf{X}}(i,j)]\frac{\mu^2}{R} \tag{8}$$

Since $0 \le R_{\mathbf{X}}(i,j) \le 1$, and $R > 0$, thus

$$R_{\mathbf{Y}}(i,j) - R_{\mathbf{X}}(i,j) \ge 0 \tag{9}$$

From Eq. (9), we come to a conclusion that compressive hyperspectral data have stronger or at least the same spectral correlation as hyperspectral images have. Some experiments are carried out to prove the validity of the above theoretical derivation about spectral correlation. The hyperspectral images for test are the sub-image sets of Cuprite scene 1 and Jasper Ridge scene 1 from AVIRIS (http://aviris.jpl.nasa.gov), with size 64×64 cut from the upper-left corner. Fig. 1 shows the correlation experimental results of compressive hyperspectral data and hyperspectral images. In our experiments, the identical Gaussian random measurement matrix is exploited to compressively sense all bands.



(a) (b)

Fig. 1 Correlations of compressive hyperspectral data and hyperspectral images. (a) 'Cspec-cof1' and 'Cspat-cof1' indicate the spectral and spatial correlation coefficients of compressive Cuprite scene 1, respectively; 'spec-cof1' and 'spat-cof1' denote the spectral and spatial correlation coefficients of Cuprite scene 1, respectively, and (b) 'Cspec-cof2' and 'Cspat-cof2' indicate the spectral and spatial correlation coefficients of compressive Jasper scene 1, respectively; 'spec-cof2' and 'spat-cof12' denote the spectral and spatial correlation coefficients of Jasper Ridge scene 1, respectively.

From Fig. 1, one can see that compressive hyperspectral data have stronger spectral correlation but much lower spatial correlation. It is experimentally proved that the CS encoding does not affect the spectral correlation structure at all. The low spatial correlation of compressive data is decided by the inherent property of CS. The CS samples are noise-like, as shown in Fig. 2. However, the noise-like characteristic has a useful advantage of error resilience. This is because each CS sample makes an identical contribution to reconstruction. That is to say the reconstruction performance is determined by the number of samples, not certain samples. Therefore, the CS encoding has a much strong robustness, very suitable for long distance transmission.

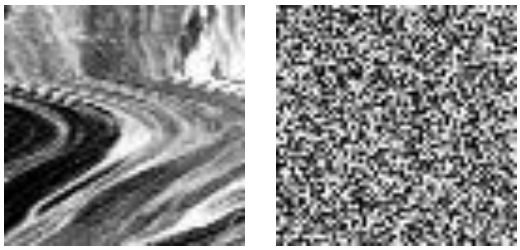


Fig. 2 Comparison of compressive hyperspectral image and hyperspectral image. The left is a sub-image set from the 40-th band of Cuprite scene 1, with size 64×64 cut from the upper-left corner, and the right is the corresponding compressive image.

4. PROPOSED RECOVERY ALGORITHM DESCRIPTION

The reconstruction problem (Eq. (4)) of hyperspectral images can be termed a multiple measurement vectors (MMV) problem. How to efficiently solve Eq. (4) is the aim we seek. It is known that hyperspectral images are acquired from the same scene over continuous spectral bands. They have similar structure information and can be regarded as a jointly sparse ensemble. Recently, the DCS theory[18] has been shown its efficacy on reconstructing jointly sparse signals[19, 20]. We know that the OMP algorithm[23] is a representative sparse signal recovery algorithm and can achieve good performance with low complexity. It is iteratively to find out the significant coefficients and to record the corresponding locations of the signal. In this section, we extend the OMP algorithm to solve the MMV problem (Eq. (4)) of hyperspectral images by exploiting the joint sparsity model of JSM2 suggested in [18]. In JSM2, signals share the same sparse supports but have different nonzero coefficients. Moreover, from the correlation analysis in Section 3, we know that the compressive hyperspectral images have strong spectral correlations. In order to reduce complexity, an adaptive grouping based OMP algorithm is proposed to recover hyperspectral images by making best use of spectral correlation. In our algorithm, an adaptive spectral band reordering algorithm is first used to classify the spectral bands into several small groups before the reconstruction and secondly the nearly best reference band for each group is found out via a least square method. Lastly, in each group, a flexible recovery strategy is designed to reconstruct reference and non-reference bands, respectively. Generally, the high recovery performance requires strong spectral correlation and small group size. In this paper, the size of group is set to ten.

4.1 Adaptive spectral band grouping algorithm

The adaptive spectral band grouping algorithm used in this paper consists of three steps:

Step 1: Calculate the spectral band correlation coefficient $R_{\mathbf{Y}}(i,i+1)$ for all any two adjacent bands of compressive hyperspectral data matrix $\mathbf{Y}$. If $R_{\mathbf{Y}}(i,i+1)>0.995$, the two adjacent bands i and $(i+1)$ belong to the same subset n; otherwise, they belong to different subsets n and $n+1$.

Step 2: Calculate the number of bands in each subset. When there is only one band in a subset, the subset is taken as a group, for this band, the OMP algorithm is used to reconstruct the original hyperspectral image. Otherwise, if the number is smaller than ten, the whole subset is regarded as a group; else the subset is divided into groups consisting of ten adjacent bands. If the number of bands in the last group is smaller than ten, the last group is incorporated into the closest former group as a new last group.

Step 3: Find the best reference band for each group. Denote the group n by $Z \in \mathbb{R}^{M \times n_b}$, where M is the number of samples for each band, n_b is the number of bands, $10 \le n_b < 20$. The reference band index λ for group n can be found by solving the following optimization problem:

$$\lambda = \arg\max_{i \in \Omega} \sum_{j=1}^{n_b} | R_{\mathbf{Z}}(i, j) | \tag{10}$$

Where $\Omega = \{1, 2, ..., n_b\}$, $R_{\mathbf{Z}}(i, j)$ represents the correlation coefficient of columns i and j of $\mathbf{Z}$. The idea behind maximizing the sum of absolute correlation coefficients is that we wish to find that band that relate the most to all other bands of the same group.

4.2 Adaptive grouping based OMP Algorithm

Denote the reference band by $\mathbf{y}^r$ and $\mathbf{x}^r$, and the non-reference band by $\mathbf{y}^p$ and $\mathbf{x}^p$ for compressive hyperspectral images and hyperspectral images, respectively. The recovery strategy of the adaptive grouping based OMP algorithm is described as follows. First, employ the OMP algorithm (Algorithm 1) to reconstruct each reference band $\mathbf{x}^r$ from the corresponding compressive data $\mathbf{y}^r$. In this step, we find out the significant coefficients and record the sparse support of the reference band $\mathbf{x}^r$. Secondly, construct a least square algorithm (Algorithm 2) by using the recorded sparse support of $\mathbf{x}^r$ to recover the non-reference band $\mathbf{x}^p$ from $\mathbf{y}^p$. Obviously, the proposed recovery strategy can speed up the reconstruction process but with little degradation in reconstruction performance.

Algorithm 1: reconstruction of the reference band $\mathbf{x}^r$ via the OMP algorithm

Inputs: samples of the reference band $\mathbf{y}^r \in \mathbb{R}^{M \times 1}$, the sampling matrix $\Phi \in \mathbb{R}^{M \times N_p}$ and the sparse basis $\Psi \in \mathbb{R}^{N_p \times N_p}$. Let the dictionary $\Omega = \Phi\Psi = \{\boldsymbol{\omega}_1, ..., \boldsymbol{\omega}_{N_p}\} \in \mathbb{R}^{M \times N_p}$, each column vector $\boldsymbol{\omega}_i$ denote an atom. Sparsity level k.

Outputs: reconstruction of the ideal sparse signal θ^r, hence the reconstructed reference band $\mathbf{x}^r = \Psi\theta^r$. Sparse support Γ of θ^r : $\Gamma = \mathrm{supp}(\theta^r)$, i.e. set of the indices of locations that are nonzero coefficients.

Initialization: iteration counter $t = 0$, the residual $\mathbf{r}^{(0)} = \mathbf{y}^r$, the sparse solution $\theta^{r(0)} = 0$, $\Gamma^{(0)} = \phi$, set of the indices of atoms that are allowed to be selected in the next iteration $\Lambda^{(0)} = \{1, 2, ..., N_p\}$.

For t from 1 to k execute the following procedures:

(1) Pick from the dictionary Ω the atom that best matches the residual:

$$i_t = \arg\max_{i' \in \Lambda^{(t-1)}} \frac{|< \mathbf{r}^{(t-1)}, \boldsymbol{\omega}_{i'} >|}{\| \boldsymbol{\omega}_{i'} \|_2}$$

(2) Renew the set $\Lambda^{(t)} = \Lambda^{(t-1)} \setminus i_t$ and $\Gamma^{(t)} = \Gamma^{(t-1)} \bigcup i_t$.

(3) Renew the weights for all the already selected atoms via the least square algorithm:

$$\hat{\mathbf{u}}^{(t)} = \arg\min_{\mathbf{u}} \| \mathbf{y}^r - \sum_{t'=1}^{t} \mathbf{u}(t') \boldsymbol{\omega}_{i_{t'}} \|_2$$

(4) Renew the sparse solution: $\theta^{r(t)}(\Gamma^{(t)}) = \hat{\mathbf{u}}^{(t)}$, which means the significant coefficients of the indices in $\Gamma^{(t)}$ equal to $\hat{\mathbf{u}}^{(t)}$.

(5) Renew the residual: $\mathbf{r}^{(t)} = \mathbf{y}^r - \sum_{t'=1}^{t} \mathbf{u}(t') \boldsymbol{\omega}_{i_{t'}}$.

Algorithm 2: reconstruction of the non-reference band $\mathbf{x}^p$ via the least square algorithm

Inputs: samples of the non-reference band $\mathbf{y}^p$, the dictionary $\Omega = \Phi\Psi = \{\boldsymbol{\omega}_1, ..., \boldsymbol{\omega}_{N_p}\}$, the recorded sparse support $\Gamma^{(k)}$ for the reference band $\mathbf{x}^r$.

Outputs: reconstruction of the non-reference band $\mathbf{x}^p$.

Execute the following procedures:

(1) Calculate the weights for all the already selected atoms via the least square method. Let

$\Omega_k = \Omega(\Gamma^{(k)}) = (\boldsymbol{\omega}_{i_1}, ..., \boldsymbol{\omega}_{i_k})$. Then $\sum_{t'=1}^{k} \mathbf{u}(t')\boldsymbol{\omega}_{i_{t'}} = \Omega_k \mathbf{u}$. According to linear algebra,

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \| \mathbf{y}^p - \Omega_k \mathbf{u} \|_2 = (\Omega_k^{\mathrm{T}} \Omega_k)^{-1} \Omega_k^{\mathrm{T}} \mathbf{y}^p$$

(2) Define the sparse solution $\theta^p(\Gamma^{(k)}) = \hat{\mathbf{u}}$, which means the significant coefficients of the indices in $\Gamma^{(k)}$ equal to $\hat{\mathbf{u}}$.

(3) Calculate the reconstructed non-reference band $\mathbf{x}^p = \Psi\theta^p$.

5. EXPERIMENTAL RESULTS

In order to test the performance of the proposed algorithm, some experiments are carried out. The hyperspectral images for test are the sub-image set of Cuprite scene 1 from AVIRIS (http://aviris.jpl.nasa.gov), with size 64×64 cut from the upper-left corner, for the purpose of reducing simulation time. The Gaussian random matrix is exploited as the measurement matrix for all bands. The popular (9,7) wavelet transform is used as the sparse basis. Currently, most CS performance is evaluated by adopting peak signal to noise ratio (PSNR) vs. measurement rate (MR), where $\mathrm{MR} = M / N \times 100\%$, M is the length of the measurement vector and N is the length of the original signal. For uniformity, we also use the measurement rate here.

First, an experiment is employed to demonstrate that a CS encoder has a stronger robustness with lower encoding complexity. For the sake of fairness, the measurement rate is set to 100%, i.e. for a sub-image which is shown in Fig. 3(a), we acquire the number of CS measurements equal to the original image size. As we know, a classical conventional encoder used to compress the acquired images is the wavelet transform-based encoder, such as that employed in JPEG2000[27], which consist of two steps (e.g. wavelet transform and zero-tree coding). In the wavelet transform-based encoder, the low frequency wavelet coefficients contain most of the image information, and if one of these coefficients is destroyed, the recovered image would be intolerable since some of information is lost completely. Fig. 3(b) illustrates a case that one of the low frequency coefficients is lost, and the reconstructed PSNR is only 40.25dB. Contrarily, the CS encoder has a much stronger robustness since it delivers the information to all the CS samples (e.g. every CS measurement makes the same contribution to the reconstruction). The recovered images are not affected much when one or some CS samples are destroyed. Fig. 3(c) shows the case that one measurement is destroyed, and the reconstructed PSNR is 107.86dB, much higher than the former. Furthermore, Fig. 3(d) shows the case that a ten percent of measurements are destroyed, and the reconstructed PSNR is 104.71dB.

(a) (b) (c) (d)

Fig. 3 Reconstructed images of the 40-th band of Cuprite scene 1. (a) is the original sub-image, (b) is the reconstructed image from wavelet coefficients with one low frequency coefficient destroyed, (c) is the reconstructed image from CS measurements with one measurement destroyed, and (d) is the reconstructed image from CS measurements with a ten percent of measurements destroyed.

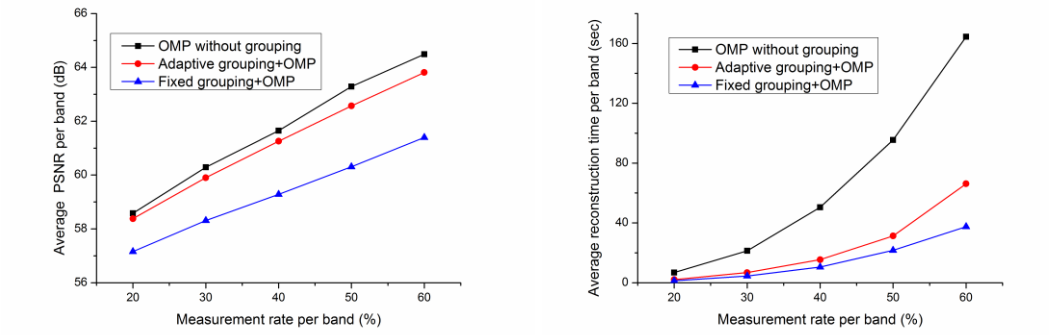
Next, we test the performance of the proposed adaptive spectral band reordering algorithm added to our recovery strategy, compared with that when there is no reordering, by using the OMP algorithm proposed in [31]. No reordering means that the 224 bands are divided into groups with ten adjacent bands, and then use the proposed recovery strategy to reconstruct images. Denote our algorithm by the adaptive grouping based OMP algorithm. In the proposed algorithm, the adaptive spectral band reordering algorithm described in Section 4 is employed to adaptively divide the 224 bands into groups. Denote the algorithm with no reordering by the fixed grouping based OMP algorithm. The reconstruction performances are shown in Table 1 in terms of PSNR and the reconstruction time on average. Algorithms are implemented in MATLAB version 7.11.0 on a workstation with Intel(R) Core(TM) i5 CPU and 4 GB memory. The reconstruction time is estimated using the etime function available in MATLAB.

Table 1 Reconstruction performance of Cuprite scene 1 with the measurement rate MR=50%

Algorithm	The average PSNR per band (dB)	The average reconstruction time per band (seconds)
Fixed grouping + OMP	60.31	21.63
Proposed adaptive grouping + OMP	62.37	31.28

From Table 1, it is clear that our proposed adaptive grouping based OMP algorithm shows better PSNR performance than that of the fixed grouping based OMP algorithm when there is no reordering. A significant improvement of 2.06 dB in PSNR on average can be seen when MR=50%. The fixed grouping based OMP algorithm exploits a fixed grouping algorithm to simply divide the 224 bands into groups with adjacent bands. This fixed grouping algorithm has some shortcomings. It might make bands with high correlation belong to different groups, while bands with low correlation belong to the same group. These shortcomings decrease the whole reconstruction performance. In our algorithm, the adaptive spectral band grouping with the simple spectral band reordering algorithm can overcome this shortcoming and make sure the bands in the same group have high correlation. It is also seen from Table 1 that the proposed algorithm consumes little more time. This is because more groups may be produced by the proposed adaptive grouping algorithm than that of the fixed grouping algorithm, and there are some groups with only one band in our experiment. For instance, when the measurement rate MR=50%, our algorithm needs 31.28 seconds on average, while the fixed grouping based OMP algorithm needs 21.63 seconds.

Lastly, we compare the reconstruction performance of the proposed adaptive grouping based OMP algorithm with the fixed grouping based OMP algorithm and the OMP algorithm without grouping under different measurement rates. The OMP algorithm without grouping means that the 224 bands are individually reconstructed using the OMP algorithm. In order to reduce complexity, both the fixed grouping based OMP algorithm and the adaptive grouping based OMP algorithm use a grouping algorithm. The experimental results of Cuprite scene 1 are shown in Fig. 4. From Fig. 4, it is clear that both the grouping based OMP algorithms perform much faster than the OMP algorithm without grouping. The price to pay for the faster speed is the decreasing performance. We can see that if one just uses the simpler fixed grouping algorithm, its performance is much worse than the OMP algorithm without grouping. However, when an adaptive grouping algorithm is added, like in our proposed algorithm, the reconstruction performance has great improvement and is only slight worse than the OMP algorithm without grouping. For example, when MR=40%, the performance of our algorithm is only 0.39 dB worse than that of the OMP algorithm without grouping, while there is a loss of 2.37 dB on the fixed grouping based OMP algorithm. It shows again that the fixed grouping algorithm is simple but not a good grouping algorithm for hyperspectral images. Moreover, when the measurement rate increases, the loss becomes larger. Thus, we can say that the adaptive grouping algorithm has a good effect on the proposed recovery strategy and our algorithm can greatly reduce computational complexity with reliable recovery quality.



(a) (b)

Fig. 4 Comparison of reconstruction performance for Cuprite scene 1. (a) is the average PSNR per band, and (b) is the average reconstruction time per band.

6. CONCLUSION

The CS theorem is a novel sampling approach with much lower data acquisition cost, compared with the conventional oversampling followed by massive dumpling acquisition method. The CS imaging has a high potential in

resource deprived applications, especially for aerospace remote sensing, such as hyperspectral imaging. In this paper, we propose a flexible reconstruction algorithm for a compressive hyperspectral imaging scheme by exploiting the spatial-spectral correlation model of hyperspectral images. In addition, an adaptive grouping algorithm with simple spectral band reordering algorithm is added to make the recovery more accurate by making best use of strong spectral correlation of compressive hyperspectral data. Experimental results show that the proposed algorithm can provide reliable recovery quality with lower computational complexity.

ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Shaanxi Province (No. 2017JQ6064), the Natural Science Foundation of China (Nos. 61405244, 61572083) and Fundamental Research Funds for the Central Universities (No.310824161006).

REFERENCES

[1] E. Candes and M. Wakin, "An introduction to compressive sampling," IEEE Signal Process. Mag., 25(2), 21-30 (2008).
[2] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, 52(4), 1289-1306 (2006).
[3] M. Duarte, M. Davenport, D. Takhar, et al. "Single-pixel imaging via compressive sampling", IEEE Signal Process. Mag., 83, 83-91 (2008).
[4] J. Ma, "Single-pixel remote sensing," IEEE Geosci. Remote Sens. Lett. 2, 199-203 (2009).
[5] 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. Opt. 52(10), D46-D54 (2013).
[6] F. Magalhaes, F. M. Araujo, M. V. Correia, M. Abolbashari, and F. Farahi, "Active illumination single-pixel camera based on compressive sensing," Appl. Opt. 50(4), 405-414 (2011).
[7] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R.G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging systems based on compressed sensing," Appl. Phys. Lett. 93, 121105 (2008).
[8] L. McMackin, M. A Herman, B. Chatterjee, and M. Weldon, "A high-resolution SWIR camera via compressed sensing," Proc. SPIE 8353, 835303 (2012).
[9] J. Ma, "A single-pixel imaging system for remote sensing by two-step iterative curvelet thresholding," IEEE Geosci. Remote Sens. Lett. 6, 676-680 (2009).
[10] Y.Wu, I. O.Mirza, G. R. Arce, and D.W. Prather, "Development of a digital-micromirror-device-based multishot snapshot spectral imaging system," Opt. Lett. 36, 2692–2694 (2011).
[11] M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," Opt. Express 15, 14013-14027 (2007).
[12] 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. Opt. 52(10), D46-D54 (2013).
[13] C. Li, T. Sun, K. F. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," IEEE Trans. Image Process. 21, 1200-1210 (2012).
[14] A. Majumdar and R. K. Ward, "Joint reconstruction of multiecho MR images using correlated sparsity," Magnetic Resonance Imaging 29 (7), 899-906 (2011).
[15] M. S. Kang, J. H Bae, S. H. Lee, and K. T. Kim, "Bistatic ISAR imaging and scaling of highly maneuvering target with complex motion via compressive sensing", IEEE Trans. On Aerospace and Electronic Systems, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8350032 (2018).
[16] D. J. Townsend, P. K. Poon, S. Wehrwein, T. Osman, A. V. Mariano, E. M. Vera, M. D. Stenner, and M. E. Gehm, "Static compressive tracking," Opt. Express 20, 21160-21172 (2012).
[17] Z. Qin, J. Fan, Y. Gao, and G. Y. Li, "Sparse representation for wireless communication: a compressive sensing approach", IEEE Signal Processing Magazine, 35(3), 40-58 (2018).
[18] D. Baron, M. F. Durate, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," arXiv: 0910.3401v1 [cs. IT] (2009).
[19] P. Nagesh and B. Li. "Compressive imaging of color images," in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Taipei, 1261-1264 (2009).

[20] L. Knag and C. Lu, “Distributed compressive video sensing,” in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Taibei, 1169-1172, (2009).
[21] L. Deng, Y. Zheng, P. Jia, S. Lu, and J. Yang, “Adaptively group based on the first joint sparsity models distributed compressive sensing of hyperspectral image”, in Proceedings of the International Conference on Computer Science and Network Technology, Dalian, China, 429-434 (2017).
[22] E. Candes, T. Tao, “Decoding by linear programming”. IEEE Trans. Inf. Theory 51, 4203–4215 (2005).
[23] C. Liu, Y. Fang, and J. Liu, “Some new results about sufficient conditions for exact support recovery of sparse signals via orthogonal matching pursuit”, IEEE Trans. on Signal Processing, 65(17), 4511-4524 (2017).
[24] Pope G. Compressive sensing: a summary of reconstruction algorithms. Master thesis. Zurich: Eidgenossische Technische Hochschule (2009).
[25] J. L. Doob, “Stochastic Processes,” Wiley, New York, 1953.
[26] G. Motta, F. Rizzo and J. Storer. “Hyperspectral data compression,” New York: Springer, 35-55 (2006).
[27] D. Taubman and M. Marcellin, “JPEG2000: image compression fundamentals, standards and practice,” New York: Springer, 2002 edition (2001).