Convex relaxation based sparse algorithm for hyperspectral target detection

Zhongwei Huang [a], Zhenwei Shi [a,*] and Zhen Qin [a]

[a] *Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, P.R. China*

Abstract

Target detection in hyperspectral images is an important task. In this paper, we propose a sparsity based algorithm for target detection in hyperspectral images. In sparsity model, each hyperspectral pixel is represented by a linear combination of a few samples from an overcomplete dictionary, and the weighted vector for such reconstruction is sparse. This model has been applied in hyperspectral target detection and solved with several greedy algorithms. As conventional greedy algorithms may be trapped into a local optimum, we consider an alternative way to regularize the model and find a more accurate solution to the model. The proposed method is based on convex relaxation technique. The original sparse representation problem is regularized with a properly designed weighted ℓ_1 minimization and effectively solved with existing solver. The experiments on synthetic and real hyperspectral data suggest that the proposed algorithm outperforms the classical sparsity-based detection algorithms, such as simultaneous orthogonal matching pursuit (SOMP) and Simultaneous Subspace Pursuit (SSP) and conventional ℓ_1 minimization.

Key words: Hyperspectral target detection; Sparsity-based algorithm; Convex relaxation; Weighted ℓ_1 minimization.

1 Introduction

Hyperspectral imagery (HSI) detection involves important applications such as space security, attack-warning or debris detection. Hyperspectral imaging sensors provide image data containing both spatial and spectral information, referred to as a data cube [14,18]. The idea for hyperspectral imaging

* Corresponding author. Tel.: +86-10-823-39-520; Fax: +86-10-823-38-798.
Email address: shizhenwei@buaa.edu.cn (Zhenwei Shi).

is based on the fact that the amount of radiation varies with wavelength for any given material. The radiance of the materials can be measured by hyperspectral sensors within each pixel area at a very large number of contiguous spectral wavelength bands. Recent study shows that great majority of natural signals are constitutionally sparse in a certain basis or with respect to a given dictionary [2]. They can be approximately represented by a few atoms in an overcomplete dictionary. Sparsity has played a key role in many signal and image processing applications such as compressing and denoising [10] ,[13]. Moreover, the recent development in the sparse coding of signals and images has provided an extremely powerful tool for computer vision and pattern recognition [2], [11], [17]. At the moment, there are several sparsity based detection algorithms for hyperspectral target detection, especially Simultaneous Orthogonal Matching Pursuit (SOMP) and Simultaneous Subspace Pursuit (SSP) [8]. These approaches notice and utilize the sparsity of coefficients in constructing natural signals. However, without constructing the coefficients in a global fashion, greedy algorithms could trap into undesired local optimums. Thus, it is significant to seek a new algorithm for sparsity reconstruction and better solving these problems.

In this paper, we propose a novel sparsity based algorithm for HSI target detection, based on a regularized sparsity model and convex relaxation technique [3]. In sparsity model, a test pixel lies in the union of the target and background subspaces and can be approximately represented by combination of training samples considered as atoms within an overcomplete dictionary. The atoms will be assigned different weights for this representation and it is believed that very few atoms will be actually used thus the weights are usually sparse, denoted as sparse coefficients. This property of sparse representation can be used in several applications of signal, image processing and computer vision [17]. Recent studies have also suggested the effectiveness of sparsity model in HSI target detection [8]. Once the dictionary is constructed with atoms representing background signatures and atoms representing target signatures, it can be separated as background and target sub-dictionaries. The sparse coefficients with respect to these two parts respectively can recognize a test pixel discriminatively as target or background pixel. However, the sparsity model faces with a important challenge about its solution. So the novel idea here is that we provide an effective approach to analysis of weights of sparse representation in a global fashion. A weighted ℓ_1 minimization is applied to approximately solve the original problem and find a more accurate global solution.

This paper is structured as follows. The sparsity model for hyperspectral target detection is briefly reviewed in section 2. The proposed target detection algorithm is introduced in Section 3. The effectiveness of the proposed method is demonstrated by experiments in Section 4. Conclusions are drawn in Section 5.

2 Sparsity Model for Hyperspectral Detection

Let $\mathbf{x}$ be a hyperspectral pixel observation, which is a B-dimensional vector with B being the number of spectral bands and the entries corresponding to the spectral bands [8]. An unknown test sample $\mathbf{x}$ lies in the union of the background and target subspaces and therefore can be written as a sparse linear combination of all training pixels by combining the two dictionaries $\mathbf{A}^b \in B \times N_b$ and $\mathbf{A}^t \in B \times N_t$:

$$\mathbf{x} = \mathbf{A}^b\mathbf{s}^b + \mathbf{A}^t\mathbf{s}^t = [\mathbf{A}^b \quad \mathbf{A}^t][\mathbf{s}^b \quad \mathbf{s}^t]^T = \mathbf{A}\mathbf{s}, \tag{1}$$

where $\mathbf{A} = [\mathbf{A}^b \quad \mathbf{A}^t]$ is a $B \times (N_b + N_t)$ matrix consisting of both background and target training samples, and $\mathbf{s}$ is an $(N_b + N_t)$ dimensional sparse vector.

Given the dictionary of training samples $\mathbf{A}$, the representation $\mathbf{s}$ satisfying $\mathbf{A}\mathbf{s} = \mathbf{x}$ can be obtained by solving the following sparse optimal problem:

$$(\mathbf{P}_0) : \widehat{\mathbf{s}} = \operatorname{argmin} \|\mathbf{s}\|_0 \quad \textit{subject to } \mathbf{A}\mathbf{s} = \mathbf{x}. \tag{2}$$

When we find the sparse solution $\mathbf{s}$ to the proposed problem, the class of $\mathbf{x}$ can be determined by comparing the residuals $r_{b(\mathbf{x})} = \|\mathbf{x} - \mathbf{A}^b\mathbf{s}^b\|_2$ and $r_{t(\mathbf{x})} = \|\mathbf{x} - \mathbf{A}^t\mathbf{s}^t\|_2$, where $\mathbf{s}^b$ and $\mathbf{s}^t$ represent the recovered sparse coefficients corresponding to the background and target sub-dictionaries, respectively. The output of detector is calculated by [8]:

$$D(\mathbf{x}) = r_{b(\mathbf{x})} - r_{t(\mathbf{x})} = \|\mathbf{x} - \mathbf{A}^b\mathbf{s}^b\|_2 - \|\mathbf{x} - \mathbf{A}^t\mathbf{s}^t\|_2. \tag{3}$$

If $D(\mathbf{x}) > \epsilon$, where ϵ is a given threshold, $\mathbf{x}$ is determined as a target pixel; otherwise, it is labeled as a background pixel [11], [12].

3 Proposed Algorithm

To solve the optimal problem in (2), most approaches are based on the greedy algorithms [2]. In this paper we focus on an alternative way. By replacing (2) with a continuous or even smooth approximation, which solution is more accessible. This kind of approach is called Convex Relaxation Techniques. To handle the $(\mathbf{P}_0)$ problem easier, there were several attempts focusing on its regularizing with tractable functions such as ℓ_1 norm and some smooth functions like standard log-barrier approximation. A common alternative of $(\mathbf{P}_0)$ problem was to consider the convex problem:

$$(\mathbf{P}_1) : \widehat{\mathbf{s}} = \operatorname{argmin} \|\mathbf{s}\|_1 \quad \textit{subject to } \mathbf{A}\mathbf{s} = \mathbf{x}. \tag{4}$$

Unlike ($\mathbf{P}_0$) and smooth approach mentioned above, this problem is convex and could be solved efficiently.

3.1 Weighted ℓ_1 Approach

In proposed algorithm, we consider another approach to regularize ℓ_0 norm. As recent studies suggested, ℓ_1 norm approach fails to take into consideration that the ℓ_0 norm was indifferent to the magnitude of the nonzero entries of the reconstruction coefficients [3]. To be explicit, coefficients with larger magnitudes will be penalized more heavily in the ℓ_1 norm, unlike the more democratic penalization that ℓ_0 does. To address this imbalance, a weighted formulation of ℓ_1 minimization was designed to more democratically penalize nonzero coefficients [2]:

$$(\mathbf{WP}_1): \ \hat{\mathbf{s}} = \text{argmin}\|\mathbf{Ws}\|_1 \quad \textit{subject to } \mathbf{As} = \mathbf{x}. \tag{5}$$

The matrix $\mathbf{W}$ is a diagonal positive-definite matrix that introduces the above-mentioned precompensating weights. For convenience the (i, i)th entry in diagonal line of the matrix is marked by w_i. It is also suggested that [3] the corresponding ℓ_1 relaxations ($\mathbf{P}_1$) and ($\mathbf{WP}_1$) will have different solutions in general [2], [3]. Hence, it is believed that if the weights w_i is set wisely, ($\mathbf{WP}_1$) could make more democratic penalization than ($\mathbf{P}_1$) does and therefore enhance the sparse reconstruction.

Using an iterative algorithm to construct the weight w_i tends to allow for successively better estimation of the nonzero coefficient locations. Since in this kind of methods, the larger signal coefficients are more likely to be identified as nonzero [4]. In the proposed iterative algorithm, once locations for nonzero entries are identified, their influences are down weighted. Therefore more sensitivity is provided for identifying the remaining small coefficients. Based on this kind of idea, we hope to find an effective way for setting w_i, which makes minimization in each iteration is a process of finding solution to ($\mathbf{P}_0$) approximately. In this paper, we propose a novel method for regularizing ℓ_0 norm via ($\mathbf{WP}_1$) problem and solved with reweighted ℓ_1 minimization.

3.2 Construct the Weights

To regularize the ℓ_0 norm with weighted ℓ_1, the selection of the weights is considered most important. In the proposed approach, we found a continuous function $F(a, s_i)$ (where a is a parameter) applied to each elements of $\mathbf{s}$, and the sum of those values will approach the ℓ_0 norm of $\mathbf{s}$. Other than classic methods, like $(\mathbf{P}_0)$, this attempt aims to make democratic penalization as the ℓ_0 norm does. Moreover, in order to put forward our $(\mathbf{WP}_1)$ construction, we apply the first-order approximation of function $\sum_i F(a, s_i)$ and this approximation maintain the property of converging to ℓ_0. Finally, we propose our iterative algorithm for solving our $(\mathbf{WP}_1)$ problem, which solution proved to be effective in the detection task.

Consider that ℓ_0 norm is indifferent to the magnitude of the nonzero entries in $\mathbf{s}$, we are looking for a function $F(a, s_i)$, so that $\sum_i F(a, s_i)$ obtain the same property. We hope that the convergence order of $F(a, s_i)$ is considerably high. We set $F(a, s_i)$ as:

$$F(a, s_i) = e^{-\frac{1}{a}e^{-as_i}}. \tag{6}$$

Compared with some recently applied function in approaching ℓ_0 norm we get faster in convergence, which suggests at the right neighborhood of point 0, $F(s_i)$ approaches 1 as quick as possible.

Actually, it is not hard to observe the convergence order of $F(a, s_i)$ is super-exponential, (compared with $g(s_i) = e^{-\frac{1}{a}s_i}$) as a $\to \infty$ when s_i is fixed, i.e.

$$\lim_{a\to\infty} \frac{-\frac{1}{a}e^{-as_i}}{-\frac{1}{a}s_i} = 0.(s_i = s_0 \neq 0). \tag{7}$$

In order to regularize our optimal problem to be convex, i.e., the $(\mathbf{WP}_1)$ problem discussed above, we apply the first-order approximation of $\sum_i F(a, s_i)$. First, by applying the Lagrange mean value theorem, we get the first-order approximation of $F(s_i)$:

$$F(s_i) \approx F'(s_0)(s_i - s_0) + F(s_0) = F'(s_0)s_i + c, \tag{8}$$

where c is a constant related to the point s_0. To be explicit, we find a tangent line of $F(s_i)$ at the point s_0. Thus, we put forward the first-order approximation of $\sum_i F(a, s_i)$ as following:

$$\sum_i F(a, s_i) \approx \sum_i F'(s_0)s_i + c_i, (s_0 \neq 0). \tag{9}$$

By utilizing this first-order approximation, we replace the $(\mathbf{P}_0)$ problem with

$$(\mathbf{WP}_1):\ \widehat{\mathbf{s}} = \operatorname{argmin} \|\mathbf{Ws}\|_1 \quad \textit{subjet to } \mathbf{As} = \mathbf{x}, \tag{10}$$

or

$$(\mathbf{WP}_1^\delta):\ \widehat{\mathbf{s}} = \operatorname{argmin} \|\mathbf{Ws}\|_1 \quad \textit{subjet to } \|\mathbf{As} - \mathbf{x}\|_2 < \delta, \tag{11}$$

where $w_i = f'(s_0)$. In practice, we use an unconstrained version and regularize the ℓ_2 term in order to apply a quadratic programming:

$$(\mathbf{WP}_1^\lambda):\ \widehat{\mathbf{s}} = \operatorname{argmin} \|\mathbf{Ws}\|_1 + \frac{\lambda}{2}\|\mathbf{As} - \mathbf{x}\|_2^2, \tag{12}$$

where λ is a Lagrange multiplier. Note that we focus on the weights of this $(\mathbf{WP}_1)$, the significant part of this approximation lies in the fact that it differentiates the weights based on the gradient at point near zero or much larger than zero. In proposed algorithm, a nature choice of s_0 for iith entry in the weights matrix is set to be the ith entry of the last solution, i.e., w_i in kth iteration is obtained as:

$$w_i^k = F'(s_i^{k-1}). \tag{13}$$

4 Iterative algorithm for sparse reconstruction

As the reweighted ℓ_1 minimization does, we would like to construct our iterative algorithm which renews the weights $\mathbf{W}$ in each iteration. As $w_i = F'(s_0)s_i$, it is nature to consider using s_{i-1} in the iteration to replace s_0, since when s_{i-1} is large, its corresponding weights $F'(s_{i-1})s_i$ becomes very small (near zero), which helps to keep the large signal coefficients in later iterations. We are more likely to find the dominant coefficients in the sparse reconstruction. The proposed convex relaxation based target detector (CRBTD) hyperspectral image detection is presented as follows:

Task: Approximate the solution of $((\mathbf{P}_0) : \hat{\mathbf{s}} = \text{argmin}\|\mathbf{s}\|_0 \quad \textit{subject to}\ \mathbf{As} = \mathbf{x})$.

Parameters: We are given the matrix $\mathbf{A}$, the vector $\mathbf{x}$, and the error threshold ε_0.

Initialization: Initialize $k = 0$, and set the initial solution $\mathbf{s}^0 = min_s\|\mathbf{As} - \mathbf{x}\|_2$ and the initial residual $\mathbf{R}_0 = \mathbf{0}$.

Main Iteration: Increment k by 1 and perform the following steps:

a. Weight Update: $w_i^k = F'(s_i^{k-1})$

b. Update Solution: Compute $\mathbf{s}^{k+1} = \text{argmin}\|\mathbf{Ws}\|_1 + \frac{\lambda}{2}\|\mathbf{As} - \mathbf{x}\|_2^2$

c. Update Residual: Compute $\mathbf{R} = \mathbf{s}^{k+1} - \mathbf{s}^k$

d. Stopping Rule: If $\|\mathbf{R}\|_2$ is smaller than the predetermined threshold ε_0, stop. Otherwise, apply another iteration.

Output: The desired result is $\mathbf{s}^k$. Apply (3) for labeling the test pixel.

The optimal problem (11) in each iteration of the proposed algorithm is solved with the quadratic programming by expanding the ℓ_2 term. We use the existing function provided by Matlab called *quadprog* to apply the quadratic

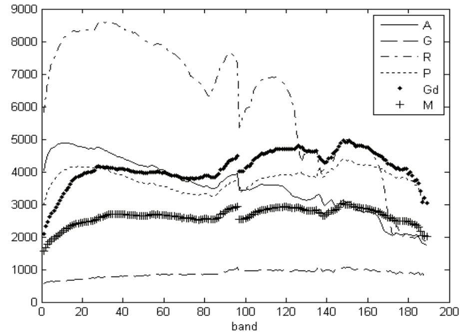


Fig. 1. Spectra of A, G, R, P, Gd and M.

programming. The parameter a in the reweigthed algorithm we proposed has a range to set with respect to its sensitiveness. As we have discussed, the larger coefficients in current iteration tends to make smaller weights in next one, a is the parameter responsible for how large a coefficient should be considered as nonzero [9], [17]. As $a \to \infty$, more smaller coefficients are taken as nonzero and get less penalization, so these coefficients are likely to be remained. In our experiments, the parameter a is set to be around 100, which allowed the algorithm provide great performance.

5 Experimental Results

In this section, we use one synthetic hyperspectral image and three real hyperspectral images for target detection experiments. The proposed algorithm CRBTD is compared with conventional sparse algorithms including Simultaneous Orthogonal Matching Pursuit(SOMP), Simultaneous Subspace Pursuit (SSP) and unweighted ℓ_1 minimization. SOMP and SSP falls in the general class of greedy algorithms and unweighted ℓ_1 minimization is compared as a classical convex minimization [4], [16].

5.1 Synthetic Image Experiment

we use the synthetic image designed method introduced by Chang et al. [7]. Five pure pixels were chosen: Airplane (A), Grass (G), Roof (R), Parking apron (P), and Ground (Gd) from a real hyperspectral image collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor. Six kinds of spectra were utilized to design a 189-band synthetic image with size of 200×200 called synthetic image, where target spectra are simulated by A, G, R, P and Gd, and the background spectra are filled by the spectrum of M. The mean spectrum of the synthetic image is removed. In Fig. 1, spectra of A, G,

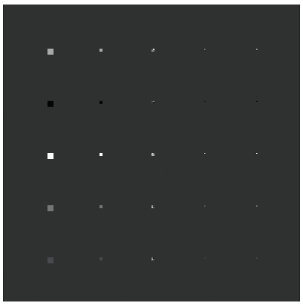


Fig. 2. The first band image of the synthetic image.

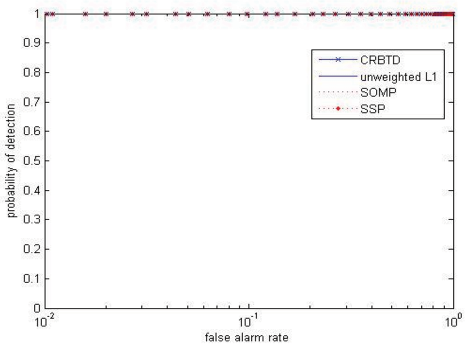


Fig. 3. ROC curves of different algorithms for the synthetic image.

R, P, Gd and M are shown. Fig. 2 shows the first band of the synthetic image. The pure-pixel panels are in the column 1 and 2, and mixed-pixel panel in in column 3, 4 and 5. White Gaussian noise with 50dB Signalto-Noise Ratio (SNR)is added to the synthetic image.

The synthetic image is used to detect the five kinds of target spectra: A, G, R, P, and Gd [15]. The ROC curves show the varying relationship of detection probability and false alarm rate. We pick thousands of thresholds in the range of the minimum and maximum of the detector output to calculate the ROC curve. The class labels for all pixels, targets or non-targets, in the test region are determined at each threshold. The PFA is calculated via the number of false alarms (background pixels determined as target) over the number of pixels in the test region, and the PD is the ratio of the number of hits (target pixels determined as target) and the total number of true target pixels. It could evaluate the performance of detectors and compare detectors independently of discriminative threshold selections [14]. A better detector performance contributes to a higher ROC curve [5], [6]. Fig. 3 shows the ROC curves of different algorithms, namely, CRBTD, SOMP, SSP and unweighted ℓ_1 method. In Fig. 3 ROC curves of different algorithms are overlapped, and all locate at 1. It means no matter what the false alarm rates are, the probabilities of detection are always 1. From Fig. 3, we can see all algorithms can detect targets well.

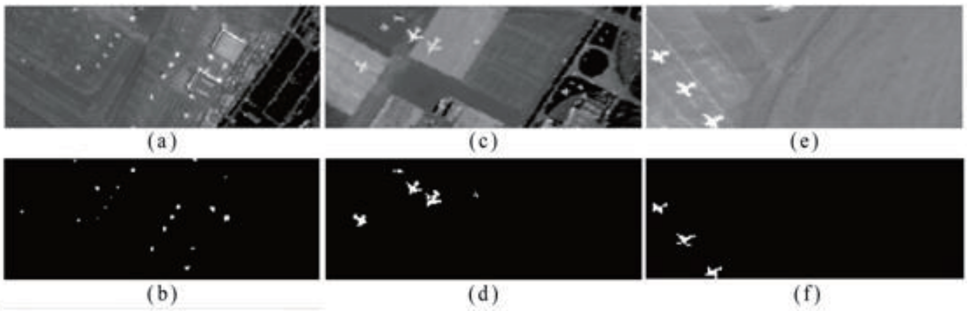
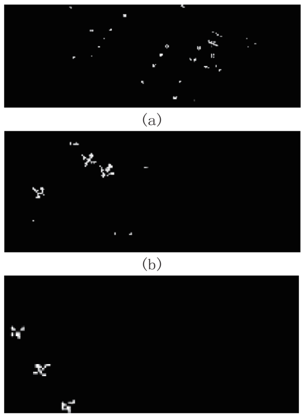


Fig. 4. (a) The first band image of AVIRIS-I. (b) The ground truth of AVIRIS-I. (c) The first band image of AVIRIS-II. (d) The ground truth of AVIRIS-II. (e) The first band image of AVIRIS-III. (f) The ground truth of AVIRIS-III.

5.2 Real Image Experiments

The three real hyperspectral images used in the experiments, AVIRIS-I, AVIRIS-II and AVIRIS-III respectively were collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor. The spectral data in 224 wavelengths spans from 0.4 to 2.5 μm. After removing the water absorption and low SNR bands, 189 available bands are left. Fig. 4 shows AVIRIS-I, AVIRIS-II and AVIRIS-III hyperspectral images. To be specifical, Figs. 4(a), (c) and (e) illustrate the first band images of AVIRIS-I, AVIRIS-II and AVIRIS-III hyperspectral images, and Figs. 4 (b), (d) and (f) show the ground truth of the three hyperspectral images. These image contains dozens of military targets, such as planes, as shown in Fig. 4. For these three AVIRIS images, every pixel on the targets is considered a target pixel when computing the ROC curves. We use a small target subdictionary containing 2 kinds of target spectra. For the background subdictionary, we selected 498 natural spectra. So, the dictionary consists of 500 spectra in all.

The convex relaxation based target detector (CRBTD) detection output of AVIRIS-I, AVIRIS-II and AVIRIS-III HSI is shown in Fig. 5 in the form of gray scale images. CRBTD method clearly performed to be effective. Under the same settings (i.e., same target and background training samples for all detectors), the classical detectors SOMP, SSP and unweighted ℓ_1 method are also applied to detect the targets of interest. The results are compared in both visual and quantitative aspects by the receiver operating characteristic (ROC) curves. ROC curves for AVIRIS-I HSI using CRBTD, SOMP, SSP and unweighted ℓ_1 method are shown in Fig. 6. Overall from this figure, one can observe that for CRBTD method yields the best performance. Under the same settings (i.e., same target and background training samples for all detectors), CRBTD could perceive the target distribution well. In terms of SOMP and SSP algorithm, they involve the matrix inverse which could results in the singular or ill-conditioned form and irritating issue for solving the detection



(a)
(b)
(c)

Fig. 5. The CRBTD detection results of AVIRIS-I (a), AVIRIS-II (b) and AVIRIS-III HSI (c).

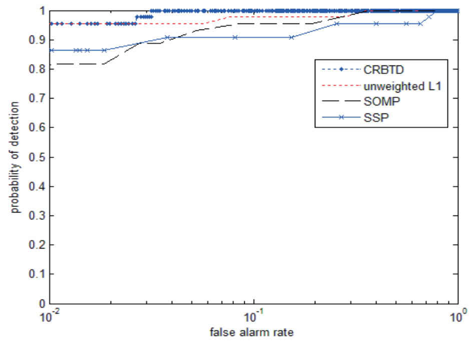


Fig. 6. The ROC curves of CRBTD, SOMP, SSP and unweighted ℓ_1 method in AVIRIS-I HSI.

problem. The unweighted ℓ_1 method is well efficient to some extent. When comparing unweighted ℓ_1, a weighted method suggests a noticeable advantage.

6 Conclusions

In this paper, we proposed a novel sparsity based algorithm for target detection in hyperspectral images. The proposed method was based on convex relaxation technique. Since hyperspectral pixels can be represented by a linear combination of a few samples from the spectra library, the sparsity model is believed to be effective for such task. Since the conventional greedy algorithms for solving the sparsity model may get inaccurate local solution, we construct a convex relaxed model with weighted ℓ_1 minimization. A continuous convex function is applied to create the reweighted vector. The experiments with respect to syndetic and real hyperspectral data suggest the advance of the proposed algorithm on sparse reconstruction. The results show that the proposed algorithm outperforms the classical sparsity based detection algorithms, such as simultaneous orthogonal matching pursuit (SOMP) and Simultaneous Subspace Pursuit (SSP).

Acknowledgments

The work was supported by the National Natural Science Foundation of China under the Grants 61273245 and 91120301, the 973 Program under the Grant 2010CB327904, the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No. BUAA-VR-12KF- 07), and Program for New Century Excellent Talents in University of Ministry of Education of China under the Grant NCET-11-0775. The work was also supported by Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, P.R. China.

References

[1] S. P. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[2] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Rev., vol. 51, no. 1, pp. 34-81, Feb. 2009.

[3] E. J. Candes, M. B. Wakin, and S. P. Boyd, Enhancing Sparsity by Reweighted .1 Minimization, Journal of Fourier Analysis and Applications, Springer, vol. 14, no. 5, pp. 877-905, 2008.

[4] T. T. Cai, G. Xu, and J. Zhang, On Recovery of Sparse Signals Via ℓ_1 Minimization, Information Theory, IEEE Transactions on, vol. 55, no. 7, pp. 3388-3397, July 2009.

[5] C. I. Chang, Hyperspectral imaging: techniques for spectral detection and classification, NewYork, USA: Kluwer Academic/Plenum Publishers, 2003.

[6] C. I. Chang, Multiparameter receiver operating characteristic analysis for signal detection and classification, IEEE Sensors Journal, vol. 10, no. 3, pp. 423-442, 2010.

[7] Y. C. C. Chang, H. Ren, C. I. Chang, and R. S. Rand, How to design synthetic images to validate and evaluate hyperspectral imaging algorithms, in: Proceedings of SPIE - The International Society for Optical Engineering, 2008.

[8] Y. Chen, N. M. Nasrabadi, and T. D. Tran, Simultaneous Joint Sparsity Model for Target Detection in Hyperspectral Imagery, Geoscience and Remote Sensing Letters, IEEE, vol. 8, no. 4, pp. 676-680, July 2011.

[9] O. Egecioglu, H. Ferhatosmanoglu, and U. Ogras, Dimensionality reduction and similarity computation by inner-product approximations, Knowledge and Data Engineering, IEEE Transactions on, vol. 16, no. 6, pp. 714-726, June 2004.

[10] S. M. Kay. Fundamentals of statistical signal processing. Prentice Hall, New Jersey, 1998.

[11] S. Kraut and L. L. Scharf. The CFAR adaptive subspace detector is a scale-invariant GLRT. IEEE Trans. Signal Process, vol. 47, no. 9, pp. 2538-2541, 1999.

[12] S. Kraut, L. L. Scharf, and L. T. McWhorter. Adaptive subspace detectors. IEEE Trans. Signal Process, vol. 49, no.1, pp. 1-16, 2001.

[13] D. G. Manolakis, V. K. Ingle, and S. M. Kogon. Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering and array processing. McGraw-Hill, Boston, 2000.

[14] D. G. Manolakis, D. Marden and G. A. Shaw. Hyperspectral image processing for automatic target detection applications. Lincoln Laboratory Journal, vol. 14, no. 1, pp. 79-116, 2003.

[15] Z. Shi and S. Yang, Robust high-order matched filter for hyperspectral target detection, Electronics Letters, vol. 46, no. 15, pp. 1065-1066, July 2010.

[16] J. A. Tropp, Just relax: convex programming methods for identifying sparse signals in noise, Information Theory, IEEE Transactions on, vol. 52, no. 3, pp. 1030-1051, March 2006.

[17] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. Huang, and S. Yan. Sparse representation for computer vision and pattern recognition, Proc. IEEE, vol. 98, no. 6, pp. 1031-1044, Jun. 2010.

[18] Z. Zhang, Z. Shi, and Z. An. Hyperspectral and panchromatic image fusion using unmixing-based constrained nonnegative matrix factorization. Optik-International Journal for Light and Electron Optics, 2012, http://dx.doi.org/10.1016/j.ijleo.2012.04.022