Article

Joint Sparse and Low-Rank Multi-Task Learning with Extended Multi-Attribute Profile for Hyperspectral Target Detection

Xing Wu ^{1,2} , Xia Zhang ^{1,*}, Nan Wang ¹ and Yi Cen ¹

¹ State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; wuxing@radi.ac.cn (X.W.); wangnan@radi.ac.cn (N.W.); cenyi@radi.ac.cn (Y.C.)

² University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: zhangxia@radi.ac.cn

Received: 24 November 2018; Accepted: 10 January 2019; Published: 15 January 2019

Abstract: Target detection is an active area in hyperspectral imagery (HSI) processing. Many algorithms have been proposed for the past decades. However, the conventional detectors mainly benefit from the spectral information without fully exploiting the spatial structures of HSI. Besides, they primarily use all bands information and ignore the inter-band redundancy. Moreover, they do not make full use of the difference between the background and target samples. To alleviate these problems, we proposed a novel joint sparse and low-rank multi-task learning (MTL) with extended multi-attribute profile (EMAP) algorithm (MTJSLR-EMAP). Briefly, the spatial features of HSI were first extracted by morphological attribute filters. Then the MTL was exploited to reduce band redundancy and retain the discriminative information simultaneously. Considering the distribution difference between the background and target samples, the target and background pixels were separately modeled with different regularization terms. In each task, a background pixel can be low-rank represented by the background samples while a target pixel can be sparsely represented by the target samples. Finally, the proposed algorithm was compared with six detectors including constrained energy minimization (CEM), adaptive coherence estimator (ACE), hierarchical CEM (hCEM), sparsity-based detector (STD), joint sparse representation and MTL detector (JSR-MTL), independent encoding JSR-MTL (IEJSR-MTL) on three datasets. Corresponding to each competitor, it has the average detection performance improvement of about 19.94%, 22.53%, 16.92%, 14.87%, 14.73%, 4.21% respectively. Extensive experimental results demonstrated that MTJSLR-EMAP outperforms several state-of-the-art algorithms.

Keywords: hyperspectral imagery (HSI); target detection; extended multi-attribute profiles; multi-task learning; sparse representation; low-rank representation

1. Introduction

Hyperspectral imagery (HSI) conveys rich spectral information over a wide range of the electromagnetic spectrum [1–3]. Each pixel can serve as a contiguous spectral fingerprint. Furthermore, the improved spatial resolution of the sensors promotes the analysis of spatial structures in the image. Target detection is an active area in the hyperspectral community, which focuses on distinguishing specific target pixels from various background pixels with *a priori* knowledge of target [2,4]. Due to its both civil and military use [5,6], target detection has been extensively applied in many HSI applications.

A large number of target detection algorithms have been proposed in the past decades. The statistical model is one of the widely used models for target detection, including constrained

energy minimization (CEM) [7], adaptive cosine/coherence estimation (ACE) [8] and hierarchical CEM (hCEM) [5]. Although the statistically based algorithms can obtain closed-form solutions and considerably low computational cost, they are affected by the estimation of background characteristics since the multivariate normal distribution assumption is too strong to satisfy in reality [9]. In recent years, the sparse representation (SR) model has drawn much attention [10–12] since there is no explicit assumption on the statistical distribution of the observed data. The key idea of SR-based methods is that a test pixel in HSI lies in a low-dimensional subspace and can be represented as a sparse linear combination of the training samples [10]. The label of the test pixel is directly determined by reconstruction error. Chen, Nasrabadi and Tran [11] proposed a joint sparse representation detector by assuming neighboring pixels are likely to have common sparse support for the training samples. A spatially adaptive sparsity model was proposed in Reference [13], which exploited different contribution of each neighboring pixel.

Although a few methods in conventional target detection models take the spatial correlation into consideration [11,13,14], most of them merely assume pixels are likely to have similar materials in a predefined neighborhood. The major drawbacks of these methods are: (1) A smaller neighborhood system may not contain adequate samples to characterize the target of interest, while a larger neighborhood system leads to the intractable computational problem [15]. (2) These detectors still treat each test patch as an ensemble of spectral measurements. It turns out that they mainly benefit from the spectral information without fully exploiting spatial structures of HSI. In addition to the inadequate use of spatial information, most of the target detection methods are affected by high redundancy among adjacent bands since they directly utilize discriminant information within all single-band images. To eliminate redundancies of hyper-dimensional data while keeping as much information as possible [16], dimensionality reduction (DR) has been applied in target detection [17,18]. However, it is still critical to preserving all informative subspace in the DR process [19]. There exists a dilemma to reduce the redundancy without loss of information.

In this study, we adopted the extended multi-attribute profile (EMAP) [20] to extract spatial features to enhance the detection performance. Although several types of spatial features were applied for analyzing hyperspectral data [15,21], EMAP has been largely proved to be a powerful tool for modeling spatial information and successfully used in many remote sensing applications [20,22]. A set of multi-scale spatial features can be derived by conducting a sequence of morphological filter operations with a family of structuring elements (SEs) [23]. To fully explore spatial information, we considered different types of attributes with a range of thresholds in filter operations. Since filters with slightly different parameters may produce similar profiles [24], there exists redundancy in EMAP. Fortunately, multi-task learning (MTL) technique has been applied in hyperspectral target detection to tackle the high-dimensional redundancy [19,25]. MTL is an inductive transfer method that improves generalization by using shared information across all related tasks [26]. By splitting the original EMAP into multiple sub-EMAP with a different set of profiles, we implemented target detection on each sub-EMAP using the MTL framework. In this way, the redundancy can be reduced while the information is retained simultaneously.

Generally, the targets of interest are usually rare in the scene, which has relatively small pixels or is distributed with a low probability. The image is dominant by several primary backgrounds. Thus the background pixel has a low-rank structure. Considering the significant difference between the background and target samples, we extended the separate modeling strategy but enforced different regularization terms. Briefly, a background pixel can be low-rank represented by the background samples while a target pixel can be sparsely represented by the target samples.

To sum up, we proposed a novel multi-task joint sparse and low-rank representation with EMAP (MTJSLR-EMAP) algorithm for hyperspectral target detection. The contributions of this paper are in threefold.

1. EMAP is adopted for spatial feature extraction in hyperspectral target detection. Compared to the conventional detectors which are susceptible to the spectral variability caused by imaging

- conditions, we take advantage of the multi-level spatial information to identify targets of interest. By introducing the spatial information, the detection performance can be significantly improved.
2. There exists high dimensional redundancy among the multiple attribute profiles. Thus, the manner that directly utilizes multiple profiles tends to degrade the accuracy [27]. To alleviate this problem, we resort to an MTL framework which can reduce the redundancy and fully exploit the information simultaneously.
 3. Based on the substantial difference between the background and target samples, we not only model the target and background pixels separately but also add a more reasonable regularization term. Compared to the existing MTL based methods, the proposed algorithm can capture the intrinsic relatedness of the background modeling tasks by enforcing the low-rank constraint.

The rest of this paper is organized as follows. Section 2 briefly introduces preliminary knowledge of EMAP and MTL. The proposed MTJSLR-EMAP method is presented in Section 3. The experimental results and analysis are given in Section 4. Finally, the discussion and conclusions are drawn in Sections 5 and 6, respectively.

2. Related Work

In this section, we first recall the basic concepts of the morphological profile (MP), attribute profile (AP) and extended morphological attribute profile (EMAP). For a complete overview of AP along with its modifications and applications in HSI, we refer the reader to [24]. Subsequently, the multi-task learning theory is briefly reviewed.

2.1. Extended Morphological Attribute Profile

Morphological profile (MP) decomposes image at multi-scale based on morphological operators of opening and closing [20,28]. The Attribute profile (AP) is an evolution of morphological profile (MP). For extracting spatial information, structure elements (SEs) with specific shapes are utilized in MP, while in AP the filtering operation is performed by connected components (CCs) that do not have particular shapes. Furthermore, AP is more computationally efficient than MP because AP is on the basis of the Max-tree representation [20].

The opening operation of AP is based on the concept of granulometry, while the closing operation of AP is based on antigranulometry [28]. Given a sequence of increasing criteria $T = \{T_\lambda : \lambda = 0, \dots, l\}$, with $T_0 = \text{true} \forall X \subseteq E$, where E is a subset of image domain $\mathbb{R}^n$ or $\mathbb{Z}^n$ (usually $n = 2$, i.e., 2-D image), X is a connected region in the image and λ is a set of scalar values used as a reference in the filtering operation. Given a gray-scale image f (with a single tone value), the attribute closing profile can be defined as

$$\prod_{\phi^{T_\lambda}}(f) = \phi^{T_\lambda}(f) \forall \lambda \in [0, \dots, l] \quad (1)$$

where $\phi^{T_\lambda}(f)$ represents the morphological attributes closing for an increasing criterion T . On the contrary, the attribute opening profile can be defined as

$$\prod_{\gamma^{T_\lambda}}(f) = \gamma^{T_\lambda}(f) \forall \lambda \in [0, \dots, l] \quad (2)$$

where $\gamma^{T_\lambda}(f)$ denotes the morphological attributes opening. AP is the concatenation of closing and opening profiles and is defined in (3)

$$AP(f) = \left\{ \underbrace{\phi^{T_{\lambda_l}}(f), \phi^{T_{\lambda_{l-1}}}(f), \dots, \phi^{T_{\lambda_1}}(f)}_{\text{thickening profile}}, f, \underbrace{\gamma^{T_{\lambda_1}}(f), \dots, \gamma^{T_{\lambda_{l-1}}}(f), \gamma^{T_{\lambda_l}}(f)}_{\text{thinning profile}} \right\} \quad (3)$$

The original image f also presents in AP since it can be viewed as the level zero of both the thickening and thinning profiles (i.e., $\phi^{T_{\lambda_0}}(f) = \gamma^{T_{\lambda_0}}(f) = f$).

Nevertheless, the extension of AP to multi-value data (e.g., HSI) is not straightforward since an ordering relation between the elements of this dataset is not natively defined [20]. To solve this problem, Dalla Mura, Benediktsson, Waske and Bruzzone [20] reduced the dimensionality of HSI through principal component analysis (PCA) and computed AP on each of the first principal components (PCs). In this way, the Extended AP (EAP) can be formalized as

$$EAP(f) = \{AP(f_1), AP(f_2) \dots, AP(f_q)\} \quad (4)$$

where q is the number of retained PCs. To further explore the spatial characteristics in the scene, it is natural to acquire multiple EAPs by considering different attributes simultaneously. The EMAP merges different EAPs and can be defined as

$$EMAP(f) = \{EAP_{a_1}(f), EAP'_{a_2}(f) \dots, EAP'_{a_k}(f)\} \quad (5)$$

where a_i is a generic attribute and $EAP' = EAP \setminus \{PC_1, \dots, PC_q\}$. It is noteworthy that the PCs in EAP' should be removed since the PCs present in each EAP .

2.2. Multi-task Learning Framework

It is common to split the large problem into multiple small and independent subproblems. Each task is considered to be independent and learned separately in single-task learning (STL). However, this method sometimes ignores the potential information available in many real-world applications. Multi-task learning (MTL) is an inductive transfer method that improves generalization by using the domain information contained in the training signals of related tasks [26]. A common assumption in MTL is that all tasks are intrinsically related to each other [29]. Under this assumption, MTL enhances the overall learning efficiency and prediction accuracy by incorporating shared information across multiple tasks. Therefore, MTL has been successfully employed in many applications, such as spam filtering, face recognition and so forth.

There are two crucial steps in MTL. One is the designation of multiple tasks with relatedness. The multiple tasks can be generated differently according to the various application. For example, some MTL-based classification methods construct pertinent classification tasks according to different features extracted from an image scene [23]. The other vital step is the relevance analysis of multiple tasks. More specifically, two commonly used approaches are involved: (1) all tasks are close to each other in some norm and (2) all tasks share a common underlying representation. The paradigm for MTL problem is to minimize the penalized loss, specified as

$$\min_W \sum_{i=1}^K \ell(W^i) + \Omega(W) \quad (6)$$

where W is the parameter to be estimated from the training samples, $\ell(W^i)$ is the i -th loss function on the training set, K is the number of total tasks and $\Omega(W)$ is the regularization term that encodes task relatedness. Different assumptions on task relatedness lead to different regularization terms. Therefore, it is vital to enforce a reasonable regularization on task relatedness. In the field of hyperspectral target detection, the target samples are selected globally using the *a priori* knowledge of target training samples, while the background samples are normally generated through a dual window. Recently, the ℓ_1 , $\ell_{2,1}$ -norm have been used in MTL-based hyperspectral target detectors by assuming different tasks share the same sparsity pattern [19,25]. Moreover, the low-rank regularization on background pixels can be performed to encode the correlation among background samples.

3. Proposed Algorithm

3.1. MTJSLR with EMAP Model

Instead of using the original hyperspectral data, we exploited the spatial structure of HSI through EMAP. Four conventional attributes were utilized in this study, they are the area of the regions, diagonal of the box bounding the region, moment of inertia and standard deviation [20]. The EMAP was generated via a series of morphological attribute filters. From an alternative view, each pixel of EMAP dataset records a spectrum of spatial features. This property provides an excellent opportunity to handle EMAP similar to handling HSI.

The proposed detector relies on the ideas of binary hypothesis model [12]. In brief, a background pixel can be low-rank represented by the background dictionary under the null hypothesis (target absent), while a target pixel can be sparse represented by the target dictionary under the alternative hypothesis (target present). Considering the EMAP dataset with N pixels and L features, $D_s = [D_s^b, D_s^t]$ are the training samples, where D_s^b and D_s^t are the background and target dictionary consisting of N_b and N_t atoms, respectively. Let x_s be a test pixel in the original EMAP, $\{x_s^k\}_{k=1}^K \in R^{L^k}$ represents the partial pixel in each sub-EMAP. For a background pixel x_s , it can be modeled as

$$\begin{aligned} x_s^1 &= D_s^{1b} w^{1b} + \zeta_s^1 \\ &\vdots \\ x_s^K &= D_s^{Kb} w^{Kb} + \zeta_s^K \end{aligned} \quad (7)$$

Correspondingly, if x_s is a target pixel, it can be represented as

$$\begin{aligned} x_s^1 &= D_s^{1t} w^{1t} + \zeta_s^1 \\ &\vdots \\ x_s^K &= D_s^{Kt} w^{Kt} + \zeta_s^K \end{aligned} \quad (8)$$

where $D_s^{kb} \in R^{L^k \times N_b}$, $D_s^{kt} \in R^{L^k \times N_t}$ and $\zeta_s^k \in R^{L^k}$ represent the background, target dictionary and random noise in k -th detection task. $w^{kb} \in R^{N_b \times 1}$ and $w^{kt} \in R^{N_t \times 1}$ indicate the coefficient vectors corresponding to D^{kb} and D^{kt} .

The targets are typically small in size or distributed with a low probability (i.e., spatially sparse). Thus, the ℓ_1 -norm is applied to targets. Due to the high correlations among background samples, it is assumed that the background pixel lies in a low dimensional subspace, we enforce the backgrounds with a low-rank regularization. The models in (7) and (8) can be rewritten as

$$\min_{W^b} \sum_{k=1}^K \|x_s^k - D_s^{kb} w^{kb}\|_2^2 + \rho_1 \|W^b\|_* \quad (9)$$

$$\min_{W^t} \sum_{k=1}^K \|x_s^k - D_s^{kt} w^{kt}\|_2^2 + \rho_2 \|W^t\|_1 \quad (10)$$

where $W^b \in R^{N_b \times K}$ and $W^t \in R^{N_t \times K}$ are the coefficient matrix formed by stacking the vectors w^{kb} and w^{kt} , respectively. $\|W^b\|_*$ is matrix nuclear norm ($\|W^b\|_* = \sum_i \delta_i(W^b)$, $\delta_i(W^b)$ denotes the i -th singular value of W^b), which is a good surrogate for matrix rank. $\|W^t\|_1$ is the maximum absolute column sum of the matrix ($\|W^t\|_1 = \max_{1 \leq j \leq K} \sum_{i=1}^{N_t} |w_{ij}|$). Parameter ρ_1 and ρ_2 balance the data fidelity term and the regularization term.

3.2. Framework of the MTJSLR-EMAP Detector

Given the dictionaries D_s^b and D_s^t , the low-rank vector W^b and sparse vector W^t can be obtained by solving the problem in (9) and (10), which, in this work, is achieved with accelerated proximal

gradient algorithm [30,31]. The reconstruction error accumulated from all tasks for the target and background can be easily derived as follows.

$$r^b = \sum_{k=1}^K \left\| x_s^k - D^{kb} w^{kb} \right\|_2 \quad (11)$$

$$r^t = \sum_{k=1}^K \left\| x_s^k - D^{kt} w^{kt} \right\|_2 \quad (12)$$

For a test pixel x_s , the final detection is determined by Equation (13)

$$D(x_s) = r^b - r^t \quad (13)$$

Ultimately, the framework of the proposed MTJSLR-EMAP algorithm is illustrated in Figure 1. For a hyperspectral data, the EMAP dataset was first calculated. Then multiple tasks were constructed through a band cross-grouping strategy [19]. Specifically, each sub-EMAP is generated from the original EMAP according to the band order at equal intervals. The target dictionary D_s^t was selected globally using the *a priori* knowledge of target training samples, while the background dictionary D_s^b was generated locally through a dual window [10]. The dual window splits the local area into two regions. A small inner window region (IWR) centered within a larger outer window region (OWR). The dual window can prevent the potential target pixels from entering the background dictionary. The background pixel is modeled by low-rank representation via the background dictionary while the target pixel is modeled by sparse representation via the target dictionary. Once the coefficient matrixes are acquired, the final detection is in favor of the class that has the lowest total reconstruction error accumulated from all tasks.

Figure 1. Illustration of the multi-task joint sparse and low-rank representation with extended morphological attribute profile (MTJSLR-EMAP) algorithm.

4. Experimental Results and Analysis

In this section, the effectiveness of the proposed algorithm was validated on three HSIs. Several target detection algorithms were used as benchmarks in the experiments for comparison. Additionally, the effects of various parameters on the detection performance of MTJSLR-EMAP were further analyzed.

4.1. Dataset Description

The first synthetic image with 64×64 pixels and 224 bands was created by [5] and labradorite HS17.3B from USGS spectral library was used as the target spectrum. There are two targets in the image, which consist of 12 pixels. The dataset is available on <http://levir.buaa.edu.cn/code>. In this experiment, the image was corrupted by a Gaussian white noise with 20dB SNR. Figure 2 shows the band 150 of the synthetic image and ground truth. Due to the high contamination of noise, it is hard to

identify targets from Figure 2a. We chose two pixels marked in red as the target atoms, $N_t = 2$ [19]. And their spectral and EMAP signatures are shown in Figure 3a,d.

The second data set was collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) from the San Diego airport area, CA, USA. It often serves as a benchmark dataset in target detection algorithm evaluation. This dataset consists of 100×100 pixels and 224 bands in a wavelength ranging from 370 to 2510 nm. After filtering out the water absorption, low SNR and bad bands (1–6, 33–35, 97, 107–113, 153–166 and 221–224), 189 bands were retained. As shown in Figure 2c, three airplanes in the top right corner of the image are targets, including 58 pixels. We selected one pixel labeled in red from each airplane as the target atoms, $N_t = 3$ [19]. Their spectral and spatial signatures are shown in Figure 3b,e.

The last data set is an airborne HSI, which was captured over Xiong'an, China in October 2017. This dataset consists of 251 spectral bands with a wavelength range from 0.4 to 1.0 μm . The spatial resolution is 0.5 m/pixel. After removing the water absorption, low-SNR and bad bands (1–19, 132–143, 174–187 and 219–251), 173 bands were retained. The image scene is 120×120 pixels and the main backgrounds include vegetation, highway, soil and shadow. The ground truth of this image was manually interpreted. As shown in Figure 2e, three vehicles in the middle of the image are targets, including 217 pixels. We took one pixel from each car labeled in red as the target atoms, $N_t = 3$. Their spectral and spatial signatures are displayed in Figure 3c,f.

Figure 2. The visualization of the datasets and their ground truth maps. The first column shows the image scene of (a) synthetic dataset, (c) AVIRIS dataset, and (e) Xiong'an dataset respectively. The second column shows the ground truth maps of (b) synthetic dataset, (d) AVIRIS dataset, and (f) Xiong'an dataset respectively. The target training samples are shown in red in each ground truth maps.

Figure 3. Spectral and Spatial signatures with respect to the above labeled red pixels. The first row shows target training spectra of (a) synthetic dataset, (b) AVIRIS dataset, and (c) Xiong'an dataset. The second row shows EMAP curves of (d) synthetic dataset, (e) AVIRIS dataset, and (f) Xiong'an dataset.

4.2. Experimental Settings

For the synthetic and AVIRIS datasets, the sizes of OWR and IWR were set as 17×17 and 7×7 [19]. For the Xiong'an dataset, we set the sizes of OWR and IWR as 27×27 and 19×19 . The number of background samples is $N_b = 240$, $N_b = 240$ and $N_b = 368$, respectively.

In the stage of EMAP calculation, we used the routine provided by [32]. As suggested in Reference [20], we retained the first PCs that could account more than 99% of the total variance of the original data. For the three datasets, we used the first 6, 3 and 3 PCs for subsequently spatial feature extraction. Each EMP was computed with a disk-shaped structure element of radius increased with a step size of 2. And four typical attributes were exploited in this study, namely, the area of the regions, diagonal of the box bounding the region, moment of inertia and standard deviation [20]. It leads to a stack of 297, 99 and 99 profiles for the three datasets, respectively. Finally, each EMAP dataset was normalized to the range of 0 to 1.

The proposed algorithm was compared with the following detectors: (1) CEM; (2) ACE; (3) hierarchical CEM [5]; (4) STD; (5) JSR-MTL; and (6) IEJSR-MTL. To demonstrate the effectiveness of spatial features in target detection, the EMAP counterparts of these detectors were also analyzed. The detection performance was evaluated by receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) value. The ROC curve illustrates the relationship between target detection rate and false alarm rate at a set of given thresholds. In this paper, a base 10 logarithmic scale for the false alarm rate was used to display the details of different detectors. For all detectors, we used the same given target signatures as input. All the experiments were carried out using MATLAB 2016a on a desktop with 3.2-GHz CPU and 16-GB Memory. As for CEM and hCEM, the mean of the target samples was used as the target signature. We set the number of detection task $K = 3$ for all MTL based detectors. The optimal parameters of the sparsity level for STD and regularization parameters for JSR-MTL and IEJSR-MTL were selected for each dataset according to the corresponding AUC values. For the proposed algorithm, the low-rank parameters ρ_1 were set as 10 and the sparsity parameters ρ_2 were set as 0.001, 1, 0.001 for the three datasets, respectively. The detailed parameter analysis of the proposed algorithm is given in Section 4.4.

4.3. Detection Performance

A good detection ROC curve should lie near to the top left. For the synthetic dataset, as shown in Figure 4a, the ROC curve of MTJSLR is above that of the other detectors, except for IEJSR-MTL. For other datasets, the ROC curve of the proposed algorithm is broadly above those of other detectors, especially when the false alarm rate ranges from 0.01 to 1. It should be noted that the ROC curve of ACE for EMAP in Figure 4d is unavailable because the inverse of the covariance matrix is ill-conditioned and unstable. Generally, it is observed that the ROC curve of each detector for the original spectral data is below that of its EMAP version.

The AUC values and computation times for the different detectors with the three datasets are shown in Table 1. The best results are labeled in bold. In spectral space, the AUC values obtained by the second best detector IEJSR-MTL are improved from 0.9892 and 0.8794 to 0.9992 and 0.9614 for AVIRIS and Xiong'an data respectively by the proposed algorithm but the AUC value of IEJSR-MTL decreases from 0.9845 to 0.9618 (MTJSLR) for the synthetic image. There are two reasons for the lower AUC value of MTJSLR for the synthetic dataset: (1) This data was generated through a linear mixing model and low-pass filter thus all spectra are highly mixed. The assumption of MTJSLR that a low-rank structure under background pixels may not be perfectly tenable. (2) With the contamination of Gaussian white noise, as shown in Figure 3a, target spectra are much different from each other, the spectral variability leads to impaired performance. By incorporating the spatial texture information, as shown in Figure 3, the similarity of EMAP curves of target pixels is higher than that of spectra. The AUC values achieved by the second best detector are improved from 0.9996, 0.9988 and 0.9620 to 0.9999, 0.9991 and 0.9805 by MTJSLR-EMAP for the three datasets, respectively. In terms of computation efficiency, in the spectral domain, the statistical-based methods (e.g., CEM, ACE, hCEM) are much faster than the sparse representation based method. The multi-task learning (MTL) based detectors are most time-consuming due to the extra burden introduced by sparse representation and MTL operations. The same phenomenon can also be found in the spatial domain.

Figure 4. Cont.

Figure 4. Detection performance of 7 detectors for three datasets. The first column is the ROC curves for original (a) synthetic dataset, (b) AVIRIS dataset, and (c) Xiong'an dataset. The second column is the ROC curves for their EMAP counterparts (d) synthetic EMAP dataset, (e) AVIRIS EMAP dataset, and (f) Xiong'an EMAP dataset.

Table 1. AUC values and running time (seconds) for the different detectors with the three datasets.

Detectors		Synthetic Dataset		AVIRIS Dataset		Xiong'an Dataset	
		Spectral	EMAP	Spectral	EMAP	Spectral	EMAP
CEM	AUC	0.7027	0.9966	0.9950	0.9963	0.6836	0.9369
	Time	0.02	0.09	0.03	0.02	0.15	0.09
ACE	AUC	0.7579	-	0.9881	0.9811	0.5576	0.9653
	Time	0.13	0.14	0.28	0.18	0.59	0.43
hCEM	AUC	0.6171	0.9998	0.9157	0.9248	0.9390	0.8631
	Time	1.15	1.62	2.01	1.65	1.76	1.80
STD	AUC	0.8070	0.9414	0.9708	0.9945	0.7557	0.8286
	Time	8.60	9.85	11.30	7.81	18.30	14.78
JSR-MTL	AUC	0.8748	0.9966	0.9133	0.9983	0.7495	0.7216
	Time	1.36E+03	1.49E+03	2.39E+03	2.26E+03	4.01E+03	4.72E+03
IEJSR-MTL	AUC	0.9845	0.9996	0.9892	0.9988	0.8794	0.9620
	Time	1.42E+03	1.49E+03	2.92E+03	2.14E+03	4.11E+03	4.91E+03
proposed	AUC	0.9618	0.9999	0.9992	0.9991	0.9614	0.9805
	Time	2.40E+03	2.48E+03	3.22E+03	3.26E+03	4.05E+03	5.16E+03

For illustrative purposes, the two-dimensional detection results of all detectors with three datasets are shown in Figures 5–7. For the synthetic dataset, as shown in Figure 5, the proposed approach gets high detection values for the target pixels as well as STD-EMAP, CEM-EMAP, hCEM-EMAP, JSR-MTL-EMAP. However, compared with hCEM-EMAP, other detectors also show high response values for some background pixels. For the AVIRIS dataset, as shown in Figure 6, it is observed that MTJSLR obtains a satisfactory detection result as well as STD, JSR-MTL, IEJSR-MTL and their EMAP counterparts. The ACE and hCEM algorithm only show high response values near the given target training pixels. For the Xiong'an dataset, in the spectral domain, as shown in Figure 7a–f, none of these detectors show a distinguishable detection map except for hCEM and MTJSLR. However, the detection performances are getting better by using spatial features. The EMAP versions of these detectors outperform their spectral versions in most cases, especially for traditional detectors (e.g., CEM, ACE). For example, from the detection maps shown in Figures 5–7, the target pixels in CEM-EMAP maps are more distinguishable than that in CEM maps. The AUC values increase from 0.7027, 0.9950 and 0.6838 to 0.9966, 0.9963 and 0.9369 for the synthetic, AVIRIS and Xiong'an dataset, respectively.

Figure 5. Detection results for the synthetic dataset. (a) CEM, (b) ACE, (c) hCEM, (d) STD, (e) JSR-MTL, (f) IEJSR-MTL, (g) MTJSLR, (h) CEM-EMAP, (i) ACE-EMAP, (j) hCEM-EMAP, (k) STD-EMAP, (l) JSR-MTL-EMAP, (m) IEJSR-MTL-EMAP, (n) MTJSLR-EMAP, (o) Ground truth.

Figure 6. Detection results for the AVIRIS dataset. (a) CEM, (b) ACE, (c) hCEM, (d) STD, (e) JSR-MTL, (f) IEJSR-MTL, (g) MTJSLR, (h) CEM-EMAP, (i) ACE-EMAP, (j) hCEM-EMAP, (k) STD-EMAP, (l) JSR-MTL-EMAP, (m) IEJSR-MTL-EMAP, (n) MTJSLR-EMAP, (o) Ground truth.

Figure 7. Detection results for the Xiong'an dataset. (a) CEM, (b) ACE, (c) hCEM, (d) STD, (e) JSR-MTL, (f) IEJSR-MTL, (g) MTJSLR, (h) CEM-EMAP, (i) ACE-EMAP, (j) hCEM-EMAP, (k) STD-EMAP, (l) JSR-MTL-EMAP, (m) IEJSR-MTL-EMAP, (n) MTJSLR-EMAP, (o) Ground truth.

4.4. Parameter Analysis

We exploited the EMAP datasets to investigate the effects of various parameters on detection performances. There are four key parameters in MTJSLR-EMAP: the low-rank parameter ρ_1 , the sparsity parameter ρ_2 , the size of the dual window and the number of detection tasks K . We kept the other parameters unchanged (as mentioned in Section 4.2) and focused on one specific parameter at a time. The range of ρ_1 and ρ_2 were set as $[1e-4, 1e-3, 1e-2, 0.1, 1, 10, 100, 1e3, 1e4]$ and the range of K was set from 1 to 9. In regard to the dual window, the size of the inner window region (IWR) is related to the size of targets. When the size of IWR is set too large, the background samples in the outer window region (OWR) cannot generally represent the local background characteristic. Thus, the sizes of the IWR were fixed as above mentioned 7×7 for the synthetic and AVIRIS datasets and 19×19 for Xiong'an dataset. The range of the size of OWR was set as $[17, 19, 21, 23, 25, 27]$ and $[27, 29, 31, 33, 35, 37]$ for the first two datasets and Xiong'an dataset, respectively. The detection performance was evaluated by the AUC value.

Figure 8 illustrates the impacts of varying regularization parameter ρ_1 and ρ_2 on MTJSLR-EMAP. For the synthetic dataset in Figure 8a, the AUC values exceed 0.99 over a wide range of ρ_1 and ρ_2 . A sudden decrease can be noticed when ρ_2 exceeds 100, with the AUC value dropping to 0.3575. For the AVIRIS dataset, as shown in Figure 8b, the AUC value improves slightly as ρ_1 increase from $1e-4$ to 1 and then maintain a high value (about 0.98). When ρ_2 is greater than 100, the AUC values decline immediately. For the Xiong'an dataset, as shown in Figure 8c, a similar trend could be found. The algorithm remains a high accuracy (about 0.97) and reaches the peak of 0.9805 at $\rho_1 = 100$, $\rho_2 = 0.001$. Since then AUC values begin to decrease as ρ_2 exceeds 10. Based on the above analysis, the MTJSLR-EMAP detector shows its robustness to regularization parameters ($\rho_1 \in [1e-4, 1e4]$, $\rho_2 \in [1e-4, 100]$). This property provides excellent convenience for parameter settings.

(a) Synthetic dataset

(b) AVIRIS dataset

(c) Xiong'an dataset

Figure 8. Detection performance of MTJSLR-EMAP versus the low-rank (ρ_1) and sparsity (ρ_2) regularization parameters. (a) Synthetic dataset. (b) AVIRIS dataset. (c) Xiong'an dataset.

We further analyzed the performance of the MTJSLR-EMAP under the varying size of OWR. For the synthetic dataset in Figure 9a, the AUC value decreases at first and then increases at OWR=21. Finally, it keeps declining as the size of OWR increases. For the AVIRIS dataset, as shown in Figure 9b, the AUC value continues to climb as the size of OWR increases from 17 to 27. For the Xiong'an dataset in Figure 9c, the AUC value increases with the growing size of OWR and then would slightly decrease after reaching the maximum 0.9846 at OWR=33.

Figure 9. Detection performance of MTJSLR-EMAP versus the size of the OWR. (a) Synthetic dataset. (b) AVIRIS dataset. (c) Xiong'an dataset.

Figure 10 exhibits the sensitiveness of the MTJSLR-EMAP to the number of tasks K . For the synthetic data set in Figure 10a, the AUC value increases at first and then gradually decreases with the growing K . For the AVIRIS data set in Figure 10b, the AUC value keeps a downward trend as K increases from 1 to 9. As shown in Figure 10c, the AUC value increases at first and then would gradually decrease after reaching the maximum at $K = 4$. MTJSLR-EMAP performs the best for the first data at $K = 6$, while offers the best detection for AVIRIS and Xiong'an dataset when $K = 2$ and $K = 4$, respectively. This experiments demonstrated that MTL detection is superior to the single task ($K = 1$) detection.

Figure 10. Detection performance of MTJSLR-EMAP versus the number of detection tasks. (a) Synthetic dataset. (b) AVIRIS dataset. (c) Xiong'an dataset.

5. Discussion

Experimental results with three datasets show the superiority of the proposed algorithm. The exceeding performance is expected because MTJSLR-EMAP applies the spatial texture information and MTL technique. Compared with CEM, ACE, hCEM, STD, JSR-MTL and IEJSR-MTL, the proposed method has the obvious improvement of about 19.94%, 22.53%, 16.92%, 14.87%, 14.73%, 4.21% respectively in terms of the average AUC results on all spectral datasets, while increases about 1.66%, 1.67%, 6.39%, 7.17%, 8.77%, 0.64% to the average AUC results on all EMAP datasets.

Overall, some interesting findings can be drawn: (1) The proposed algorithm generally achieves the best detection performance in both spectral and spatial spaces. (2) In the spectral domain of the synthetic image and AVIRIS dataset, hCEM tends to get worse results than CEM because hCEM relies on a layer-by-layer filtering procedure. Thus, some potential weak targets could be suppressed. (3) The EMAP versions of these detectors are superior to their spectral versions in most cases. Our experimental results show that EMAP has great potential for discriminating targets from backgrounds. (4) The detection performances are boosted mostly by using spatial features, especially for some traditional detectors (e.g., CEM, ACE, STD). This reveals that spatial features also play an essential role in target detection. It is feasible to join spectral and spatial features together, which is likely to promote the detection performance.

There are four parameters of the MTJSLR-EMAP algorithm, which have been thoroughly analyzed. MTJSLR-EMAP is relatively robust to ρ_1 and ρ_2 but the empirical settings of other parameters still have a minor influence on the detection performance. We fixed $\rho_1 = 100$, $\rho_2 = 0.01$ and randomly chose the size of OWR and the number of tasks in their respective parameter range as mentioned in Section 4.4. For each dataset, the detection of the proposed algorithm was implemented 20 times with the randomly chosen parameters. The ranges of the AUC values are illustrated in Figure 11. It is evident that our algorithm achieves comparative performance in terms of AUC values more than 0.9970, 0.9939 and 0.9756 for the three datasets, respectively. However, the AUC value varies with the change of window size and the number of tasks, especially for the Xiong'an dataset. Therefore, how to adaptively decide the optimal parameters for the proposed algorithm needs to be further investigated.

Figure 11. AUC ranges for the empirical parameters (OWR and number of tasks) settings of the proposed algorithm on three datasets.

Additionally, it is worth noting that the construction of multiple related tasks is a critical step in multi-task learning. In this study, the cross-grouping strategy [19] was employed to create pertinent detection tasks. Although this strategy exploits the inter-band similarity and obtains solid performance, the band selection at equal interval for the task construction is unnecessarily the optimal approach. Exploiting another effective method for task designation needs to be investigated in the future.

6. Conclusions

In this paper, a novel joint sparse and low-rank MTL with EMAP (MTJSLR-EMAP) for target detection algorithm was proposed. We took advantage of the multiple attributes spatial information to discriminate target of interest. In each task, a background pixel can be low-rank represented by the background dictionary while a target pixel can be sparsely represented by the target dictionary. The MTL technique was applied to integrate the multiple detection tasks together. Finally, the label of the test pixel was determined by comparing which class yielded the minimum reconstruction residual. Extensive experimental results with three datasets demonstrated that the proposed MTJSLR-EMAP algorithm outperforms several state-of-the-art detectors. It should be noted that the performance of the proposed algorithm will slightly degrade when processing highly mixed and noisy dataset, since the assumption that backgrounds have a low-rank structure may not be perfectly tenable. Thanks to the increased availability of data and computational resources, the use of deep learning is taking off in remote sensing community [33–35]. Exploring deep learning techniques to extract high-level abstract features from HSI to promote target detection performance will be the focus of our future research.

Author Contributions: All the authors made contributions to the work. X.W. and X.Z. conceived, designed and performed the experiments. N.W. and Y.C. provided advice for the preparation and revision of the paper.

Funding: This research was funded by National Natural Science Foundation of China (41671360).

Acknowledgments: The authors would like to thank Prof. Wang Yueming, Shanghai Institute of Technical Physics (SITP) of the Chinese Academy of Sciences, China, for providing the Xiong'an airborne hyperspectral image.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *113*, S110–S122. [\[CrossRef\]](#)
2. Nasrabadi, N.M. Hyperspectral target detection: An overview of current and future challenges. *IEEE Signal Process. Mag.* **2014**, *31*, 34–44. [\[CrossRef\]](#)
3. Du, B.; Zhang, L. A discriminative metric learning based anomaly detection method. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6844–6857.
4. Du, B.; Zhang, L. Target detection based on a dynamic subspace. *Pattern Recognit.* **2014**, *47*, 344–358.
5. Zou, Z.; Shi, Z. Hierarchical suppression method for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 330–342. [\[CrossRef\]](#)
6. Du, B.; Zhang, Y.; Zhang, L.; Tao, D. Beyond the sparsity-based target detector: A hybrid sparsity and statistics-based detector for hyperspectral images. *IEEE Trans. Image Process.* **2016**, *25*, 5345–5357. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Du, Q.; Ren, H.; Chang, C.I. A comparative study for orthogonal subspace projection and constrained energy minimization. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1525–1529.
8. Manolakis, D.; Marden, D.; Shaw, G.A. Hyperspectral image processing for automatic target detection applications. *Linc. Lab. J.* **2003**, *14*, 79–116.
9. Lu, X.Q.; Zhang, W.X.; Li, X.L. A hybrid sparsity and distance-based discrimination detector for hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1704–1717. [\[CrossRef\]](#)
10. Chen, Y.; Nasrabadi, N.M.; Tran, T.D. Sparse representation for target detection in hyperspectral imagery. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 629–640. [\[CrossRef\]](#)
11. Chen, Y.; Nasrabadi, N.M.; Tran, T.D. Simultaneous joint sparsity model for target detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 676–680. [\[CrossRef\]](#)
12. Zhang, Y.; Du, B.; Zhang, L. A sparse representation-based binary hypothesis model for target detection in hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1346–1354. [\[CrossRef\]](#)
13. Zhang, Y.; Du, B.; Zhang, Y.; Zhang, L. Spatially adaptive sparse representation for target detection in hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1923–1927. [\[CrossRef\]](#)
14. Yang, S.; Shi, Z. Hyperspectral image target detection improvement based on total variation. *IEEE Trans. Image Process.* **2016**, *25*, 2249–2258. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in spectral-spatial classification of hyperspectral images. *Proc. IEEE* **2013**, *101*, 652–675. [\[CrossRef\]](#)
16. Jia, X.; Kuo, B.C.; Crawford, M.M. Feature mining for hyperspectral image classification. *Proc. IEEE* **2013**, *101*, 676–697. [\[CrossRef\]](#)
17. Sun, K.; Geng, X.; Ji, L. A new sparsity-based band selection method for target detection of hyperspectral image. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 329–333.
18. Farrell, M.D.; Mersereau, R.M. On the impact of pca dimension reduction for hyperspectral detection of difficult targets. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 192–195. [\[CrossRef\]](#)
19. Zhang, Y.; Du, B.; Zhang, L.; Liu, T. Joint sparse representation and multitask learning for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 894–906. [\[CrossRef\]](#)
20. Dalla Mura, M.; Benediktsson, J.A.; Waske, B.; Bruzzone, L. Extended profiles with morphological attribute filters for the analysis of hyperspectral data. *Int. J. Remote Sens.* **2010**, *31*, 5975–5991. [\[CrossRef\]](#)
21. Ghamisi, P.; Yokoya, N.; Li, J.; Liao, W.; Liu, S.; Plaza, J.; Rasti, B.; Plaza, A. Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2018**, *5*, 37–78. [\[CrossRef\]](#)
22. Song, B.Q.; Li, J.; Mura, M.D.; Li, P.J.; Plaza, A.; Bioucas-Dias, J.M.; Benediktsson, J.A.; Chanussot, J. Remotely sensed image classification using sparse representations of morphological attribute profiles. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5122–5136. [\[CrossRef\]](#)

23. Li, J.Y.; Zhang, H.Y.; Zhang, L.P.; Huang, X.; Zhang, L.F. Joint collaborative representation with multitask learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5923–5936. [[CrossRef](#)]
24. Ghamisi, P.; Dalla Mura, M.; Benediktsson, J.A. A survey on spectral-spatial classification techniques based on attribute profiles. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2335–2353. [[CrossRef](#)]
25. Zhang, Y.; Ke, W.; Du, B.; Hu, X. Independent encoding joint sparse representation and multitask learning for hyperspectral target detection. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1933–1937. [[CrossRef](#)]
26. Caruana, R. Multitask learning. *Mach. Learn.* **1997**, *28*, 41–75. [[CrossRef](#)]
27. Imani, M. Attribute profile based target detection using collaborative and sparse representation. *Neurocomputing* **2018**, *313*, 364–376. [[CrossRef](#)]
28. Pesaresi, M.; Benediktsson, J.A. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Geosci. Remote Sens. Lett.* **2001**, *39*, 309–320. [[CrossRef](#)]
29. Chen, J.H.; Liu, J.; Ye, J. Learning incoherent sparse and low-rank patterns from multiple tasks. *ACM Trans. Knowl. Discov. Data* **2012**, *5*, 22. [[CrossRef](#)]
30. Chen, X.; Pan, W.; Kwok, J.T.; Carbonell, J.G. Accelerated gradient method for multi-task sparse learning problem. In Proceedings of the 2009 9th IEEE International Conference on Data Mining, Miami, FL, USA, 6–9 December 2009; pp. 746–751.
31. Ji, S.; Ye, J. An accelerated gradient method for trace norm minimization. In Proceedings of the International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; pp. 457–464.
32. Dalla Mura, M.; Benediktsson, J.A.; Waske, B.; Bruzzone, L. Morphological attribute profiles for the analysis of very high resolution images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3747–3762. [[CrossRef](#)]
33. Zhu, X.X.; Tuia, D.; Mou, L.C.; Xia, G.S.; Zhang, L.P.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [[CrossRef](#)]
34. Yao, X.W.; Han, J.W.; Zhang, D.W.; Nie, F.P. Revisiting co-saliency detection: A novel approach based on two-stage multi-view spectral rotation co-clustering. *IEEE Trans. Image Process.* **2017**, *26*, 3196–3209. [[CrossRef](#)] [[PubMed](#)]
35. Han, J.W.; Cheng, G.; Li, Z.P.; Zhang, D.W. A unified metric learning-based framework for co-saliency detection. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 2473–2483. [[CrossRef](#)]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).