A Fast Method for Hyperspectral Target Detection with Matrix Inverse Updating and Downdating

Xin Wu\textsuperscript{1,}\textsuperscript{*}, Mizhen Wang\textsuperscript{1} and Chen Yang\textsuperscript{2}

\textsuperscript{1}School of Physics and Optoelectronic Engineering, Xidian University, Shaanxi, Xi’an 710071, China
\textsuperscript{2}Faculty of Automation and Information Engineering, Xi’an University of Technology, Shaanxi, Xi’an 710048, China

\textsuperscript{*}Corresponding author

Abstract—The unsupervised nonnegative constrained least squares (UNCLS) method is a well-known method for hyperspectral subpixel target detection. Depending on the complexity and dimensionality of the hyperspectral data, it may be computationally expensive for large matrix inversion with the UNCLS method. This paper proposed a fast UNCLS algorithm to improve its computational performance. Matrix inverse updating and downdating theory is implemented to speed up the iterative least square estimation in the UNCLS method. Experiments on both simulated and real hyperspectral images show that the proposed fast UNCLS can efficiently detect subpixel targets with precision guaranteed. We have demonstrated that our proposed method improves performance of the UNCLS method, which achieves a >10x speedup for both the simulated and real scenes when large number of targets are detected.

Keywords—remote sensing; matrix inverse; fast least squares; subpixel target detection

I. INTRODUCTION

Mixing the spectrum of a subpixel target with the spectra of the background results in a pixel with a combined spectral signature. Subpixel target detection in hyperspectral images involves with a separation of the constituent elements in the combined spectral signatures [1–4]. Several classical subpixel target detection algorithms have proposed, which are summarized in the book by Chang [5]. The unsupervised nonnegative constrained least squares (UNCLS) algorithm is one of the popularly-used methods. The UNCLS method iteratively applies the nonnegative constrained least squares (NCLS) algorithm with the signature matrix to estimate the abundance fractions [5]. It can automatically extract the target signatures with a stable result no matter how many times of execution without a priori [6]. Nevertheless, depending on the complexity and dimensionality of hyperspectral data, the UNCLS algorithm may be time-consuming involving with big matrix inversion for the covariance of the signature matrix.

Acceleration of hyperspectral image processing methods has been an active topic in recent years. With the rapid development of remote sensing techniques, more efficient computation methods are necessary for the unprecedented size of hyperspectral images. Researchers have proposed fast algorithms with different approaches [7–9], including adoption of hardware acceleration [10–12], in order to reduce the computational costs. Typically, Chang et al. [7] proposed a fast iterative method for the pixel purity index (PPI). Guerra et al. [8] applied the orthogonal subspace projection (OSP) method to accelerate the N-FINDR algorithm. López et al. [9] selected a subset of pixels to extract endmembers in a reduced time. In addition, implementations of parallel algorithms with high performance hardware are recently popular in hyperspectral image processing. PPI algorithm has been accelerated by a GPU-based massively parallel version for real-time applications [13]. The automatic target detection and classification algorithm (ATDCA) and the hyperspectral signal subspace identification by minimum error (HySime) method have also been implemented in parallel with GPUs [14,15]. Besides, field programmable gate arrays (FPGAs) and coprocessors (Intel Xeon Phi) are promising to adopted in this domain [11,12,16]. However, there is few acceleration of the UNCLS algorithm for hyperspectral subpixel detection.

In this paper, we propose a fast UNCLS method to accelerate the original one. Firstly, we incorporated the matrix inverse updating theory [17] to accelerate the inverse of the signature covariance matrix in estimating the abundance fractions. Secondly, we derived the matrix downdating scheme to invert the steering matrix in the NCLS algorithm. Given that we have the inverse of a matrix \( R^{-1} \), it is possible that the inverse of the matrix \( R^{-1} \) can be obtained from \( R \) by appending a row or column (“update” or “upscale”) or deleting a row or column (“downscale”). In this situation, it is much more efficient to update or downdate inverse of \( R \) other than to calculate \( R^{-1} \) from scratch. Theoretical time complexity of the proposed method has been made. Experiments have been conducted on a simulated hyperspectral image (HSI) based on the Cuprite AVIRIS data and a real HSI collected by NASA AVIRIS sensor over the World Trade Center area in New York City. Experimental results indicate that the proposed fast UNCLS method can achieve a better performance compared to the original algorithm with correct results guaranteed.

The rest of this paper is organized as follows. Section 2 briefly describes the UNCLS method and then describes the proposed fast UNCLS method and then our implementation in detail. Section 3 shows the experimental results. Section 4 concludes the paper with some remarks.
II. PROPOSED APPROACH

A. The UNCLS Method Review

The UNCLS is a popular method for hyperspectral subpixel target detection. We briefly describe the UNCLS method here. More details can be found in the paper by Chang et al. [6]. The UNCLS method uses a linear mixture model, which assumes that \( \mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_p \) are \( P \) target signatures resident in a multispectral/hyperspectral image pixel vector \( \mathbf{r} \).

\[
\mathbf{r} = \mathbf{M}\alpha + \mathbf{n}, \quad (1)
\]

where \( \mathbf{n} \) denotes the noise from instruments or measurement. Denoting \( \mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_p] \) an \( L \times P \) target spectral signature matrix, the abundance column vector \( \alpha = (\alpha_0, \alpha_1, \ldots, \alpha_P)^T \) can be estimated by a linear regression model as \( \hat{\alpha}_{LS} \):

\[
\hat{\alpha}_{LS} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{r} \quad (2)
\]

To fulfill the non-negative constraint, \( \forall j \in 1, p; \alpha_j > 0 \) the method of Lagrange Multipliers is introduced to solve the optimization problem:

Minimize LSE = \( (\mathbf{M}\alpha - \mathbf{r})^T(\mathbf{M}\alpha - \mathbf{r}) \), subject to \( \alpha \geq 0 \) \( (3) \)

Equation (4) is adopted to estimate the non-negative abundance:

\[
\hat{\alpha}_{NCLS} = (\mathbf{M}^T\mathbf{M})^{-1}\lambda\mathbf{r} - (\mathbf{M}^T\mathbf{M})^{-1}\lambda, \quad (4)
\]

where \( \lambda \) is the multiplier vector \([18]\). \( \hat{\alpha}_{NCLS} \) is the abundance vector estimated by the NCLS algorithm. The NCLS algorithm is one of the key steps of the UNCLS method involving with matrix inversion \( (\mathbf{M}^T\mathbf{M})^{-1} \) recursively. In the NCLS algorithm, two sets of indicators are introduced to decide which row/column to be deleted according to the Kuhn-Tucker conditions \([18,19]\). The passive set is initialized as an empty vector, and the active set is initialized with the indices from 1 to \( P \). Indices in the passive and active sets will be changed according to the Lagrange multiplier vector \( \lambda \), the complete scheme of the NCLS algorithm can be referred to reference \([6]\).

Since the NCLS algorithm requires a complete knowledge of target signatures, the UNCLS method iterates the NCLS algorithm to achieve an unsupervised version. The UNCLS algorithm have two major components, namely two major loops. The first one is the implementation of the NCLS with an inner loop to form a steering matrix \( \Phi_k^{(k)} \). Each steering matrix has to be inversed for next calculation until the estimated abundance \( \hat{\alpha}_{NCLS} \) fulfills the non-negative constrain.

As the NCLS algorithm describes, the steering matrix is updated by deleting some rows and columns from the matrix \( (\mathbf{M}^T\mathbf{M})^{-1} \) and then inverse the newly-formed one. More detailed steps can be found in reference \([6]\).

B. Proposed Fast UNCLS Method

We introduce matrix inverse updating and downdating theory to improve the UNCLS algorithm for a better computational performance. The comprehensive derivation of our fast UNCLS method is as follows. In the original UNCLS algorithm, each newly-detected target signature will be added to \( \mathbf{M} \), the target signature matrix. In order to express the derivation clearly, we denote \( \mathbf{R}_k = \mathbf{M}_k^T\mathbf{M}_k \) and its inverse as \( \mathbf{R}_k^{-1} \), where the subscript \( k \) is the number of target
signatures and \( \mathbf{e}_k \) is the k-th detected signature. Then \( \mathbf{R}_k \) is expressed in Equation 5 and 6.

\[
\mathbf{R}_k = \begin{bmatrix} \mathbf{M}_k^\top \mathbf{M}_{k-1} & \mathbf{M}_k^\top \mathbf{e}_k \\ \mathbf{e}_k^\top \mathbf{M}_{k-1} & 1 \end{bmatrix}, \quad (5)
\]

\[
\mathbf{R}_k = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_k^\top \mathbf{M}_{k-1} & \mathbf{M}_k^\top \mathbf{e}_k \\ \mathbf{e}_k^\top \mathbf{M}_{k-1} & 1 \end{bmatrix}, \quad (6)
\]

where \( \mathbf{A} \), \( \mathbf{b} \), \( \mathbf{c} \) and \( \mathbf{d} \) denote the block matrix elements. To solve the least squared estimation problem in Equation 3, matrix inverse updating is applied to the block-wise matrix inverse. The analytic block matrix inverse is given in Equation 7 and 8:

\[
\mathbf{R}_1^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{r}_{12} \\ \mathbf{r}_{21}^\top & \mathbf{r}_{22} \end{bmatrix}, \quad (7)
\]

\[
\mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{b} (\mathbf{d} - \mathbf{c} \mathbf{A}^{-1} \mathbf{b})^{-1} \mathbf{c} \mathbf{A}^{-1} + \mathbf{r}_{12} = -\mathbf{A}^{-1} \mathbf{b} (\mathbf{d} - \mathbf{c} \mathbf{A}^{-1} \mathbf{b})^{-1} \mathbf{c} \mathbf{A}^{-1}, \quad (8)
\]

\[
\mathbf{r}_{21} = (\mathbf{d} - \mathbf{c} \mathbf{A}^{-1} \mathbf{b})^{-1} \mathbf{c} \mathbf{A}^{-1}, \quad (9)
\]

\[
\mathbf{r}_{22} = (\mathbf{d} - \mathbf{c} \mathbf{A}^{-1} \mathbf{b})^{-1}. \quad (10)
\]

\[
\mathbf{R}_{11} = \mathbf{R}_{k-1}^{-1} + \mathbf{r}_{22} \cdot \mathbf{R}_{k-1}^{-1} \mathbf{M}_{k-1}^\top \mathbf{e}_k \mathbf{M}_{k-1} \mathbf{R}_{k-1}^{-1}, \quad (11)
\]

\[
\mathbf{r}_{12} = -\mathbf{r}_{22} \cdot \mathbf{R}_{k-1}^{-1} \mathbf{M}_{k-1}^\top \mathbf{e}_k, \quad (12)
\]

For block-wise matrix inverse, \( \mathbf{A} \) and \( \mathbf{d} \) must be square matrices and the sub-blocks of the \( \mathbf{b} \) and \( \mathbf{c} \) can be of arbitrary size \([20, 21]\), and yet for our fast UNCLS method \( \mathbf{b} \) and \( \mathbf{c} \) are vectors and \( \mathbf{d} \) is a scalar. It is clear that \( r_{22} \) can be directly inverted without a matrix inverse. Given the transpose relation between \( \mathbf{R}_{11} \) and \( \mathbf{R}_{12} \) as shown in Equations 11 and 12, we can also gain some improvement on the computational performance. Since the updating and downdating scheme has been derived above in Equations (8), (10) and (14), we modified the UNCLS algorithm and applied matrix inverse updating and downdating method to the fast UNCLS algorithm. The comprehensive algorithm of the proposed fast UNCLS method is shown in Figure 2. The pseudo code contains some symbols, which are used in the origin UNCLS algorithm by Chang \([6]\). According to the pseudo code, we calculated the time complexity of the proposed method \([22]\). Firstly, let \( n \) be the number of pixels in one band of the HSI to be processed, which can be regarded as the scale of the UNCLS algorithm. As we applied the matrix inverse updating and downdating technique, the matrix inverse can be done outside the inner loop, thus the time complexity can be reduced from\( T(n) = O(n^2) \) to \( T(n) = O(n^2 + n \log_2 n) \). Theoretically, when the computing scale increases, the proposed algorithm can overcome the original one. However, the time complexity analysis is based on an abstract flow of the UNCLS algorithm, the computing performance evaluation with synthetic and real HSI are present in the next section.

Firstly, the HSI data is loaded to the memory and the number of subpixel targets is declared as initial condition. The first target signature candidate is selected by finding the maximum power through the whole HSI, then enter the inner loop with a minimum error threshold is set as 1.0e-24. When the number of the detected subpixel targets is greater than two, matrix inverse updating and downdating was applied to the inner loop of the original NCLS algorithm (In Figure 2, from Step 2 to Step 9). In the modified NCLS algorithm, downdating scheme is applied to build a steering matrix \( \mathbf{F}_a^{(k)} \), all rows and columns specified in the negative indices are deleted from the matrix \( (\mathbf{M} \mathbf{M}^\top)^{-1} \). To calculate inverse of the newly formed matrix \( (\mathbf{F}_a^{(k)})^{-1} \), parts of \( \mathbf{M} \mathbf{M}^\top \) can be reused for simple combination according to the block-wise matrix inverse theory. Suppose that the i-th row and column of \( (\mathbf{M} \mathbf{M}^\top)^{-1} \) is to be deleted, the new matrix can be formed by the other four block-wise matrices. By moving the deleted row and column to the bottom and the left respectively as shown in Figure 3, inverse of the new matrix can be calculated by Equation 14 from the block element \( \mathbf{A} \), \( \mathbf{b} \), \( \mathbf{c} \), \( \mathbf{d} \) in \( \mathbf{M} \mathbf{M}^\top \).
Besides, negative indices may include more than one index for the rows and columns to be deleted in the NCLS algorithm, so B, C, and D can be block-wise matrices instead of vectors or a scalar. Hence, we can express matrix inverse downdating in a block-wise format:

\[
R_{11}^{-1} = A - BD^{-1}C. \tag{15}
\]

Note that Equation 15 can be more efficient only when \(D^{-1}\) is a small size matrix. When the sizes of \(R_{11}\) and D are close in the fast UNCLS algorithm, the downdating scheme cannot make an obvious improvement on the performance. If matrices \(R_{11}\) and D are close in size, the downdating operation will consume almost the same computing time as inverting matrix D directly. Therefore, a predetermined threshold \(\theta\) is defined to make a limitation on the size of downdating matrix scale in Step 6. The threshold is necessary since computing performance may be reduced if too many columns and rows are deleted from the steering matrix, which is also observed in our experiments. It means that the threshold constrains the maximum size of matrix D in Equation 15. The value of the threshold \(\theta\) is important for the overall efficiency of the proposed fast UNCLS method. It is related to the size of the target signature matrix with respect to the number of signatures, once the number of signatures is determined the value can be predefined accordingly. Details about how we set the value of \(\theta\) in the experiments are present in the next section.

Moreover, the performance of matrix inverse updating and downdating is compared with the built-in matrix inverse function of Matlab in Section 3.1. The proposed fast UNCLS method is evaluated with both simulated and real HSI data in the following section in Section 3.2.

III. EXPERIMENTS

A. Simulated Hyperspectral Image Evaluation

We conducted simulated hyperspectral image experiments for two reasons: one is to verify the correctness of the proposed fast UNCLS algorithm; the other one is to evaluate the performance of the matrix inverse with updating/downdating versus Matlab built-in inv function. As illustrated in reference [5], TE (Target Embeddedness) is a simulated HSI primarily designed for target detection. Five pure spectra were extracted from the public Cuprite AVIRS hyperspectral image, which are alunite, buddingtonite, kaolinite, calcite, and muscovite as shown in Figure 4(a) and (c). There are 25 target panels in the simulation as shown in Figure 4(b). Each row is simulated by the same mineral signature and each column of five targets has the same pixel size. The first and second columns are pure spectral targets with 4x4 and 2x2 pixels, respectively. The third column is also in a 2x2 pixel size but the spectra is a mixture of the five pure signatures. The last two columns contain ten 1x1 pixel targets with a mixture of the five mineral signatures. The mixture rate is different between the fourth (1:1) and fifth column (1:3), which is of same with the TE2 scenario in reference [5]. Thus, we can simulate ten different subpixel targets in the last two columns. Besides, the targets pixels are embed into the background with a SNR = 20:1 Gaussian noise, which is simulated by the sample mean of the real Cuprite image. Figure 5(a) shows the embed targets as false color image with background noise, bands 40, 80, and 189 are used. For clarity, we have removed the noisy background for illustration of target detection in the following.
The hyperspectral target signature detection results by the original and our fast UNCLS method are identical as shown in Figure 5. The six signatures (the background was detected as one target signature) were marked with yellow circles and the numbers besides indicating the detection order.

We also verified the abundance of the 5 detected target signatures, and illustrated the detection result in Figure 6. As seen from the figures, all the hyperspectral subpixel targets located in the last two columns were detected.

Comparing with the “ground truth” data [5,7], we calculated the RMSE (root mean square errors) of the detection results obtained by our fast UNCLS and the original one, as shown in Table I.

Identical RMSE of the original and our fast UNCLS verify the correctness of the proposed method tested on the simulated HSI. Taking advantage of the simulation of a hyperspectral image, we can create synthetic HSI with different spatial size, spectral signatures, and subpixel targets for experiments. Thus, we also evaluated the performance of the matrix inverse updating comparing with the traditional inv Matlab build-in function. Since we set the number of subpixel targets from 1 to 100, which means a set of 1×1 to 100×100 square matrices were inverted by the two methods, we simulated a set of different spatial size HSI with 224 bands. To guarantee an adequate data, we selected a 200 (samples) x 200 (lines) x 224 (bands) with 100 different signatures as research object. And all the tests were iterated by 100 times to retrieve the average computation time. The computing performance result is shown in Figure 7.

It is straightforward that matrix inverse updating method overcomes built-in inv function of Matlab as the size of the HSI matrix getting larger. In matrix inverse updating, each iteration of the calculation reuses the last inversion result, which gains a significant time saver. From the experimental results, we can conclude that the matrix inverse updating scheme is better at calculating an iterative inverse than the built-in inv function, and the speedup will go higher when the number of subpixel targets increases.
In a similar way, we conducted a test for matrix inverse downdating method over built-in inv function of MATLAB. Figure 8 shows the computing time and speedup for the two methods. It can be seen that the speedup gets reduced as the deleted target signatures increased, which means downdating is suitable for inverting a block matrix only when deleting a small number of rows and columns. That is why a threshold is set in our fast algorithm. Besides, after testing various sizes of HSI with different number of subpixel targets (here a 200 x 200 with 60 signatures HSI is present), we found that the performance of matrix inverse downdating is better than the built-in inv function when the number of rows/columns to be deleted is less than half of the whole targets. Inspired by this observation, we can predetermine the value of threshold as half of the input number of expected target signatures in our real HSI experiments. The detailed experimental results are shown in the next section.

It should also be noted that the performance will improved by matrix inverse updating and downdating only when we need to calculate matrix inverse iteratively and the matrix size is big enough. Fortunately, usage of the fast UNCLS algorithm fulfills the above two requirements in hyperspectral subpixel target detection. There are two reasons why updating and downdating scheme gains performance improvement with the proposed fast UNCLS method: one is that a hyperspectral data is usually with massive data in spatial resulting in a large matrix $M^T M$; the other one is that the UNCLS calculated the inverse of $M^T M$ iteratively by least square optimization.

B. Real Image Evaluation

The WTC real HSI collected by NASA AVIRIS over the World Trade Center on 16 September 2001 was used in our experiment; it comprises 614 x 512 pixels with 224 spectral bands making about 140 MB of binary data. Since the spatial resolution is 1.7 meters, there exist many subpixel targets in this complex urban scene. The RGB (red, green, blue) false color image of WTC is formed using the 577.37, 836.27 and 451.72 nm channels, as shown in Figure 9(a). The green area delineates vegetation, while the purple smoke was over the burned area, and from the detailed center area (in the yellow rectangle) some thermal hot spots caused by the fire can be observed and are denoted from ‘A’ to ‘H’ in Figure 9(b). The hot spots are subpixel targets to be detected, the original data was collected by U.S. Geological Survey (USGS) and made available online (http://speclab.cr.usgs.gov/wtc/).

Both the original and our fast UNCLS algorithms were executed on the WTC real HSI. Firstly, we checked the correctness of the detection results, and we got the identical positions of signatures, as shown in Table II.

And we marked all the signatures on the false color WTC image as shown in Figure 10, where the number of target signatures varies from 1 to 30. The positions of the detected signatures have been marked with red circles in the figure. For the fire and smoke area, signatures 3, 10 and 29 can be used to extract all subpixel targets in the WTC data.
We made further performance comparison of the proposed fast UNCLS with the original one on the real WTC HSI scene. The experimental platform is a computer with an Intel® Core™ i5-3210M @2.5GHz CPU and 4 GB RAM. Matlab R2013b was used on a Windows 7 64-bits operating system. For both the original and our fast UNCLS implementations, we ran 10 times to obtain a reliable result. The running times of the two methods and the speedups are plotted in Figure 11.

From the speedup (blue-square line) shown in Figure 11, we can see that our fast UNCLS algorithm gains a better performance as compared to the original one for each target signature detected. Note that the original UNCLS algorithm would take as much as 5396.8 seconds (about 1.5 hours) to finish detecting 30 target signatures from the WTC real HSI data, while our fast version reduced that to 518.5 seconds (about 8.6 minutes), a ~10x speedup achieved.

IV. CONCLUSION

The unsupervised nonnegative constrained least squares (UNCLS) method is a popular method for hyperspectral subpixel target detection. Depending on the complexity and dimensionality of the hyperspectral data, it may be time consuming for large matrix inversion iteratively in the original UNCLS algorithm. In this paper, we developed a fast UNCLS algorithm with matrix inverse updating and downdating theory. The correctness and performance are evaluated and analyzed with both the simulated and real HSI scenes. The proposed method improves the computational performance, which achieves a >10x speedup for both the simulated Cuprite and real WTC scenes when large number of targets are detected.