

## TEMPLATE MATCHING VIA $l_1$ MINIMIZATION AND ITS APPLICATION TO HYPERSPECTRAL DATA

ZHAOHUI GUO AND STANLEY OSHER

Department of Mathematics, University of California, Los Angeles  
Los Angeles, CA 90095, USA

(Communicated by Margaret Cheney)

**ABSTRACT.** Detecting and identifying targets or objects that are present in hyperspectral ground images are of great interest. Applications include land and environmental monitoring, mining, military, civil search-and-rescue operations, and so on. We propose and analyze an extremely simple and efficient idea for template matching based on  $l_1$  minimization. The designed algorithm can be applied in hyperspectral classification and target detection. Synthetic image data and real hyperspectral image (HSI) data are used to assess the performance, with comparisons to other approaches, e.g. spectral angle map (SAM), adaptive coherence estimator (ACE), generalized-likelihood ratio test (GLRT) and matched filter. We demonstrate that this algorithm achieves excellent results with both high speed and accuracy by using Bregman iteration.

**1. Introduction.** In recent decades, hyperspectral remote sensing technology has become highly developed. Remote sensing can be used to collect the reflective spectrum of the objects in a scene at specific wavelengths simultaneously, resulting in hundreds of digital images. The data collected from a hyperspectral sensor contains not only the visible spectrum, but also ultraviolet and infrared ranges as well. It is common to list the hyperspectral data in a three-dimensional array or “cube”, with the first two dimensions corresponding to spatial dimensions and the third one corresponding to the spectrum.

Since the distinct materials leave different spectral signals, each of which is known as the spectral signature, there has been an increasing demand in using spectral imagery to detect specific objects or targets of interest. Template matching is one of the most important hyperspectral image (HSI) applications based on materials’ spectral signatures.

In hyperspectral classification and especially target detection, the main task is to find the spatial pixels in three-dimensional hyperspectral cube data for some given spectral signals of interest. However, this becomes difficult because of the uncertainty and variability of each material’s spectral signature. The difficulties include the noise from atmospheric conditions, sensor influence, location, illumination and so on, all of which depend on when and where the image was taken. Another disadvantage is that the same materials will not generate identical spectral curves for physical reasons, especially between the laboratory data and actual image data. Occasionally, two different materials might unexpectedly reflect similar spectral signals, like two different types of vegetation. Additionally, the low resolution of the

---

2000 *Mathematics Subject Classification.* 65N21, 86A21.

*Key words and phrases.* Hyperspectral, template matching, Bregman,  $l_1$  minimization.

Research supported by the Department of Defense.

hyperspectral image make it reasonable to accept the fact that a single pixel may be composed of several different materials. These are usually called mixed pixels. The spectral signal of this mixed pixel is totally different from the spectrum of the pure materials. All of these facts increase the difficulties in processing HSI data, making classification and target detection highly challenging problems.

Previous work dealing with hyperspectral target detection includes both statistical and geometrical approaches. Statistical approaches are mostly based on the assumption that HSI data can be represented by a multivariate normal distribution. The statistical models employing two hypotheses  $H_0$  and  $H_1$  corresponding to given material absent and present are well known as the likelihood ratio (LR) test (optimum) and generalized-likelihood ratio test (GLRT) [18, 27] (adaptive) detectors. Some other related algorithms with details can be found in [22, 23, 24, 31, 34]. Also, adaptive coherence estimator (ACE) [11] as an extension of generalized-likelihood ratio test (GLRT), employs a covariance matrix to identify the background, and achieves better result. However, in these models, prior knowledge including background information has to be obtained from training data, like the covariance matrix of the background. Where to set the threshold towards a reasonable compromise between probability of detection and probability of false alarms is also a question. In practice, there still exist serious challenges.

The spectral angle map (SAM) [21], and the adaptive subspace detector (ASD) [19, 20] are representative geometrical algorithms. In addition, similar methods based on hypotheses and GLRT include the matched subspace detector [29], and matched filter algorithms, which in turn include orthogonal subspace detector (OSD) [17, 29] and spectral matched filter (SMF) [12, 28].

Many classical hyperspectral classifiers have been used in remote sensing, including minimum distance to mean (MD), spectral angle mapper (SAM), maximum likelihood (ML), Fisher linear likelihood and so on. In order to reduce the error, classification is usually performed in the Principal Component Analysis (PCA) and Generalized Eigenvalue (GEV) dimensionality reduced domains.

Recent methods have largely improved the efficiency of mixed pixel computation. The method presented in [7, 16, 32] using the  $l_1$  unmixing model arose from the linear mixing model (LMM), which assumed that the observed pixel spectrum can be decomposed into a linear combination of the spectrum of the target and the background material. The distribution of each material decides the detection result.

Our method resembles SAM, but has an  $l_1$  sparsity term which seems to improve the performance. We will discuss this further. The paper is organized as follows. A brief introduction of Bregman methods for  $l_1$  minimization and the algorithm for template matching are presented in Section 2. In Section 3, we apply the proposed algorithm to a simulated data set and the real HSI image data sets. Comparisons with classical algorithms, SAM, ACE, GLRT and matched filter, are presented based on receiver operating characteristic (ROC) curves and detection results to show the obtained improvements. Finally, conclusions are given in Section 4.

**2. Template matching approach and algorithms.** In this section, we first briefly introduce the concept of Bregman iteration [25], and split Bregman [14, 15], then present our algorithm for template matching. This is done to give simple and effective methods for potential users to do the  $l_1$  minimization.

**2.1. Bregman Iteration and Algorithms.** Bregman methods [8] applied to image processing [25] and compressive sensing (CS) [13, 14, 26, 33, 35] are an emerging

field for solving sparse reconstruction and improving signal and image enhancement. Bregman iteration is based on the definition of Bregman distance. The Bregman distance for a convex function  $J(u)$  defined in  $R^n \rightarrow R \cup \{+\infty\}$  is given as:

$$D_p^J(u, v) = J(u) - J(v) - \langle p, u - v \rangle,$$

where  $p \in \partial J$  is a subgradient of  $J$  at the point  $v$ , and  $\langle \cdot, \cdot \rangle$  is used as the inner product. In general Bregman distance shows the closeness between two points  $u$  and  $v$ , even though it doesn't satisfy the symmetry and triangle inequality properties of usual distance.

Consider a general constrained minimization problem:

$$(1) \quad \min_u J(u) \quad \text{s.t. } H(u) = 0, \text{ or } H(u) < \sigma,$$

where  $H(u)$  is convex and differentiable energy function with zero as its minimum value.  $\sigma$  is the noise measure coming from the original problem.

We transform it to an unconstrained minimization problem using:

$$(2) \quad \min_u \{ \mu J(u) + H(u) \}.$$

$\mu > 0$  works as penalty weight. Bregman iteration solves a related problem iteratively by replacing  $J(u)$  by  $D_{p^k}^J(u, u^k)$ :

$$(3) \quad \begin{cases} u^{k+1} = \operatorname{argmin}_u (\mu D_{p^k}^J(u, u^k) + H(u)), \\ p^{k+1} = p^k - \partial H(u^{k+1}), \end{cases}$$

where  $\partial H(u^{k+1})$  is a subgradient of  $H$  at  $u^{k+1}$ . The problems under consideration here use  $H(u) = \frac{1}{2} \|Au - f\|_2^2$  for some linear operator  $A$  and vector  $f$ . Convergence analysis of the iterative scheme is also provided in [25]. It is also shown that under fairly weak conditions,  $H(u^k) \rightarrow 0$  as  $k \rightarrow \infty$ . These can be easily solved by split Bregman.

Split Bregman was first introduced by Goldstein and Osher [14] for solving  $l_1$ , TV, and related regularized problems and applied to various imaging problems [9, 15, 35]. Consider of the problem

$$(4) \quad \min_u \|u\|_1 \quad \text{s.t. } Au = f,$$

where  $A \in R^{m \times n}$ ,  $f \in R^m$  be given, and  $m < n$ . The related unconstrained minimization problem is

$$(5) \quad \min_u \{ \mu \|u\|_1 + \frac{1}{2} \|Au - f\|_2^2 \}.$$

To solve this sparse reconstruction model, the iteration is generated by using an auxiliary variable  $d = u$ , and given by

$$(6) \quad \min_{u,d} \{ \mu \|d\|_1 + \frac{1}{2} \|Au - f\|_2^2 \} \quad \text{s.t. } d = u,$$

or

$$(7) \quad (u, d) = \operatorname{argmin} \{ \mu \|d\|_1 + \frac{\lambda}{2} \|Au - f\|_2^2 + \frac{1}{2} \|d - u\|_2^2 \}.$$

where  $\mu, \lambda > 0$  work as penalties balancing the energy functions. Apply the Bregman formulation, we obtain the following algorithm

$$(8) \quad \begin{cases} (u^{k+1}, d^{k+1}) = \operatorname{argmin} \{ \mu \|d\|_1 + \frac{\lambda}{2} \|Au - f\|_2^2 + \frac{1}{2} \|d - u - b^k\|_2^2 \}, \\ b^{k+1} = b^k + u^{k+1} - d^{k+1}. \end{cases}$$

We “split” the  $l_1$  and  $l_2$  components of the first subproblem in (8) and perform optimization in an alternating fashion:

$$(9) \quad \begin{cases} u^{k+1} = \operatorname{argmin} \left\{ \frac{\lambda}{2} \|Au - f\|_2^2 + \frac{1}{2} \|d^k - u - b^k\|_2^2 \right\}, \\ d^{k+1} = \operatorname{argmin} \left\{ \mu \|d\|_1 + \frac{1}{2} \|d - u^{k+1} - b^k\|_2^2 \right\}, \\ b^{k+1} = b^k + u^{k+1} - d^{k+1}. \end{cases}$$

Set  $u^0 = b^0 = d^0 = 0$ , and define

$$(10) \quad \operatorname{shrink}(v, \mu) = \begin{cases} v - \mu & : v > \mu \\ 0 & : -\mu \leq v \leq \mu, \\ v + \mu & : v < -\mu \end{cases}$$

and

$$(11) \quad \operatorname{shrink}^+(v, \mu) = \begin{cases} v - \mu & : v > \mu \\ 0 & : v \leq \mu \end{cases}$$

where the  $\operatorname{shrink}^+$  is designed for computing nonnegative solutions.

We have the iterations:

$$(12) \quad \begin{cases} u^{k+1} = (\lambda A^T A + I)^{-1} (\lambda A^T f - b^k + d^k), \\ d^{k+1} = \operatorname{shrink}^+(u^{k+1} + b^k, \mu), \\ b^{k+1} = b^k + u^{k+1} - d^{k+1}, \end{cases}$$

Additionally, we can also enforce another Bregman iteration process,

$$(13) \quad \begin{cases} u^{n+1} = \operatorname{argmin} \left\{ \mu \|u\|_1 + \frac{1}{2} \|Au - f^n\|_2^2 \right\}, \\ f^{n+1} = f^n + f - Au^{n+1}, \end{cases}$$

and obtain a two-layer inner-outer iteration style.

The inner iteration is given as (9), using  $k$  as the inner iteration steps. And the outer iteration is given as  $f^{n+1} = f^n + f - Au^{n+1}$ , where  $u^{n+1}$  is generated by the inner iteration.  $n$  denotes the steps of outer iteration.

We know that  $\|u^k - u^{k+1}\|$  converges monotonically as  $k \rightarrow \infty$  in the inner iteration, so does  $\|u^k - d^k\|$ . It is also proved in [35] that the iteration converges even with only one inner step if  $0 < \lambda < \frac{1}{\|A^T A\|_2}$ . This is called Bregmanized operator splitting (BOS). In practice, we use a fixed number of inner iterations, usually 5-10 steps.  $\mu$  is a parameter that balances the  $l_1$  and  $l_2$  components. It is determined by the actual data. This turns out to be a very efficient method for  $l_1$  minimization, and we will use it below. We also found experimentally that  $\lambda = \frac{100}{\|A^T A\|_2}$  is a good choice for the problem.

If  $A$  is fixed and saved as a prior knowledge,  $(\lambda A^T A + I)^{-1}$  in (12) could be precomputed, resulting in run-time savings. However, the size of square matrix  $(\lambda A^T A + I)$  is dependent on the number of columns of  $A$ . Note that if  $Au = f$  is underdetermined and the number of column is very large, it is impractical to compute the inverse. An alternative method could be used here. Instead of solving

$$(14) \quad u^{k+1} = (\lambda A^T A + I)^{-1} (\lambda A^T f^n - b^k + d^k)$$

in (12), we insert two new variables  $y$  and  $g$ ,

$$(15) \quad \begin{cases} g = \lambda A^T f^n - b^k + d^k, \\ y = (\lambda A A^T + I)^{-1} A g, \\ v^{k+1} = g - \lambda A^T y, \end{cases}$$

Using (15), the computation of inverse matrix is only done for a much smaller size matrix of  $(\lambda A A^T + I)$ . Equivalence between (14) and (15) can be proved easily through a few steps of simple linear transformation.

**2.2. Template Matching Approach.** The template matching approach basically employs the  $l_1$  sparsity idea arising in compressed sensing [10, 13]. Suppose the hyperspectral image is given as  $N$  by  $N$  pixels, each containing an  $M$  channel spectrum (a spectral band), as a three-dimensional cube. We aim to locate the pixels of interest in the image corresponding to a given spectral signal  $f$ , which is a specific object containing the same  $M$  channel spectrum.

We arrange  $A$  as an  $M$  by  $N^2$  matrix, with generally  $M < N^2$ . The signal  $a_{i, 1 \leq i \leq N^2}$  is the column of the spectrum matrix  $A$ , and corresponds to each pixel in the hyperspectral image.  $f$  is a given material's spectral signal. The aim is to find the nonzero component  $u_i$  in the solution  $u \in R^{N^2}$  of the constrained minimization problem:

$$(16) \quad \min_u |u|_1 \quad \text{s.t.} \quad \|Au - f\|_2 < \sigma \quad \text{and} \quad u \geq 0.$$

The  $u_{i, 1 \leq i \leq N^2}$  corresponds to the same pixel as  $a_i$ . By searching for the nonzero component  $u_i$ , we can locate the pixels of interest. With the introduction of  $l_1$  minimization, the algorithm becomes much more robust.

This is different from the Basis Pursuit problem (4), which is used in compressed sensing. We are not solving the exact solution of  $Au = f$ , but approaching the solution of  $\min_u \|Au - f\|_2^2, u \geq 0$  until the tolerance is reached. There are many methods proposed to do this numerical computation. In this paper, we apply split Bregman to the following unconstrained minimization problem.

$$(17) \quad \min_u \left\{ \mu |u|_1 + \frac{1}{2} \|Au - f\|_2^2 \right\} \quad \text{s.t.} \quad u \geq 0.$$

To get a nonnegative solution,  $shrink^+$  should be employed in the algorithm. In practice, the solution for hyperspectral data turns out to be nonnegative even without the nonnegative constraint employed here because of the nonnegativity of the data.

Surprisingly this simple idea works well because the  $l_1$  norm term forces the energy to find a sparse solution  $u$  for  $\mu$  large enough, which will enable the solution to pick up the matching pixels compared with the desired spectral signal  $f$ , and ignore nonmatching pixels. In detection applications, it is highly possible that the matching pixels are sparsely distributed, or the number of matching pixels is limited, which makes the problem difficult and interesting. The sparseness of the matching pixels means the sparseness of  $u$ . We say that given a proper value of  $\mu$ , the solution results in  $u$  with a limited number of positive components, which just correspond to the matching pixels in the image with a given spectral signal  $f$ .

However, even in the nonsparse (numerous) case, this also works well. We can easily illustrate this by a simple clean case of classification. We randomly choose  $n$  pixels in the image and assign them to be the material of interest. We assume the spectrum of these pixels identically equal to  $f$ , or proportional to  $f$ , which

means  $kf$ ,  $0 < k < 1$  works as a simulated shadow effect. After solving the  $l_1$  minimization problem, we get a sparse solution  $u$  with each nonzero component  $u_i$  corresponding to the same  $n$  pixels we chose. The values of these  $u_i$  are, close to  $\frac{1}{n}$  after normalization. Of course, it does make sense since the algorithm aims to find the proper solution  $u$  to make  $\|Au - f\|_2$  as small as possible, while ignoring the spectral signatures distinct from  $f$ .

In the actual computation, the only significant parameter that should be considered is the quantity  $\mu$  to control the sparsity of the solution, while  $\lambda$  could be chosen fixed. There exists a suitable range of  $\mu$  for each set of HSI data. In detection applications, the number of matching pixels is much smaller than the total number of image pixels,  $\mu$  could vary in a much large range. The result wouldn't be affected so much. But in the classification case, sometimes the number of matching pixels is numerous, the result depends largely on  $\mu$ . If  $\mu$  is chosen to be too large, the algorithm will only identify a few correct matches. To deal with this problem, we can remove the detected matches and do the algorithm again on the remaining pixels using a new matrix with fewer columns.

Simultaneously, if the background spectral signatures are quite distinct from  $f$ , we get no false alarms despite the change of  $\mu$ . Otherwise, it will depend on  $\mu$ . However, in practice the number of incorrect matches turns out to be much fewer than the number of correct matches, and the coefficients  $u_i$  of most incorrect matches are much smaller than those of the correct matches, usually by 1 or 2 orders of magnitude. Thus we can set up a simple threshold to identify the correct result.

We notice the similarity of the template matching model proposed here and the unmixing model in [7, 16, 32]. However, the connotation of parameters in the two models are totally different:  $A$  in the template matching model is listed as matrix of HSI data, while  $A$  in the unmixing model is arranged as an endmember matrix;  $f$  in the template matching model denotes the spectrum of target, but in the unmixing model is the spectrum of one pixel in HSI data. In addition, the former model is typically underdetermined, but the later is generally overdetermined. Moreover, the indications of two solutions are not the same: one measures the similarities between spectrum, and the other illustrates the fractional values of endmembers occupying the pixel. Therefore, these are two different methods and two different applications.

In contrast to the other methods, the template matching algorithm involves neither any requirement of background information nor the assumption of multivariate normal distribution of the background and designed material during the whole computation process. Only the spectrum of material of interest is needed in the algorithm. The method can deal with the case when the matches are sparse or numerous and when background models are unavailable and unreliable.

Moreover, the template matching algorithm only matches the spectrum of image pixels with the spectrum of material interested, and doesn't involve spatial information. So the method is parallelizable and can work locally under the condition of inhomogeneous background. Thus the  $l_1$  minimization based algorithm is realistically applicable.

**3. Numerical experiments and results.** In the following section, the split Bregman algorithm indicated in Section 2 is implemented in Matlab to assess this applicability of  $l_1$  minimization for hyperspectral template matching. Simulated data and real hyperspectral images data will be tested.

The code was written in MATLAB using 3 GHz, Intel Core 2 Duo CPU. The most time-consuming part is a matrix inverse computation. But the matrix  $A$  is dependent on the HSI data, which could be used as prior knowledge. Depending on the size of data to be tested, one can choose to use the alternative method (15). Furthermore, because the problem is parallelizable, one can also split the image into smaller patches, and the original  $A$  into smaller size matrices.

Concerning the computational complexity, we note that, given  $A \in R^{m,n}, m < n$ . The complexity of the algorithm per iteration using the alternative method (15) is thus  $O(m^2n)$ .

**3.1. Simulated Data Set.** The simulated data set we tested here is based on the a HYDICE hyperspectral image called Urban [2] which is publicly available. Firstly we performed a data correction by deleting the water absorption noise bands from the original 210 bands, leaving a 163-band image for actual computation. We chose a 50 by 50 pixels patch image in Urban to perform the background, basically including dirt road, tree and grass materials, as shown in the left of Fig 1. So the size of matrix  $A$  to be computed was 163 by 2500, and the solution  $u \in R^{2500}$ .

We manually chose another material to perform the target or material of interest, and here we used a material of metal rooftop appeared in another part of Urban to be distinct from the background, as shown in Fig 1. Ten pixels in the patch were randomly selected as the locations of matching pixels, and were assigned to be metal rooftop. After that, we finished the test data simulation.

Solving the constrained minimization problem via split Bregman algorithm generates a sparse signal  $u$ . This enables us to locate the matched pixels by just checking the nonzero components  $u_i$ . In the following, we list the different kinds of examples we tested, and show the performance of the algorithm under different cases. The result is also compared with other two methods. We used fixed value of parameter  $\lambda$  and varying  $\mu$  in the experiments. In practice, we found choosing  $\lambda$  larger will speed up the algorithm, here, we use  $\lambda = \frac{100}{\|A^T A\|_2}$ , where works well even without the proof of convergence.

Besides the time used to compute the matrix inversion, it took only 1-2 seconds to compute the sparse  $u$  of the simulated data to determine the targets using split Bregman.

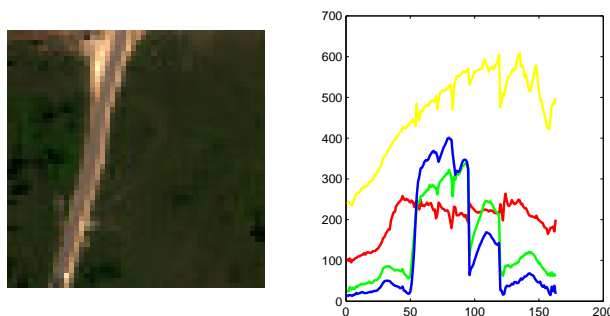


FIGURE 1. Simulated data image. Left: RGB view of patch chosen in the Urban. Right: Spectrum of materials in the patch. Red: target-metal rooftop. Yellow: dirt road. Blue: trees. Green: grass.

3.1.1. *Robustness to Noise.* To simulate the effect of the environmental variations, illumination and so on, we added noise  $n \in N(0, \sigma)$  to the spectrum of 10 simulated pixels to make this simulation more physically meaningful. In order to characterize the noise level, we use SNR (signal to noise ratio) defined as

$$(18) \quad SNR = \tilde{u}/\sigma,$$

where  $\tilde{u}$  is the mean of the signal, and  $\sigma$  is the standard deviation of the noise.

With  $\mu = 0.01$ , and SNR=20.3, we got the the solution  $u$  with the exact 10 nonzero components  $u_i$  to be [0.1029, 0.0832, 0.1107, 0.1119, 0.0906, 0.1143, 0.0759, 0.1294, 0.1065, 0.0849], the other  $u_i = 0$ , and the  $l_1$  norm of  $u = 1.0101$ . And each of the nonzero  $u_i$  corresponded to one correct simulated pixel.

Increasing noise to SNR=15, with the same  $\mu$ , we got 10 nonzero components  $u_i$  to be [0.0929, 0.0790, 0.1000, 0.0954, 0.1671, 0.1181, 0.0430, 0.1079, 0.0766, 0.1068], and the  $l_1$  norm of  $u = 0.9869$ . We also located the correct pixels.

As the noise increased, we can still locate the correct pixels, but a few incorrect positive components  $u_i$  may arise. We introduce some measure quantification to evaluate the performance of algorithm, true positive rate (TPR) and false positive rate (FPR), which are defined as

$$(19) \quad TPR = \frac{TP}{TP + FN}; FPR = \frac{FP}{FP + TN},$$

where  $TP$  is the pixels detected correctly (true positive),  $FP$  is the false alarm pixels (false positive),  $TN$  is background pixels with zero components (true negative) and  $FN$  is the matching pixels with zero components (false negative).

In the test SNR=10, we obtained the average  $TPR = 98.6\%$  and  $FPR = 0.004\%$  based on 100 computations. Increasing the noise to SNR=5, we found the algorithm gave the average  $TPR = 98.1\%$  and  $FPR = 0.87\%$ . The number of these incorrect pixels detected is limited since they just work to reduce the residual. Moreover, the magnitudes of the incorrect positive components is relatively small. A simple threshold can be employed to reduce the number of  $FP$  to zero.

In addition, we can also add a small perturbation to  $f$ , or change the magnitude of  $f$ . The algorithm still works well for moderate changes.

3.1.2. *The Choice of  $\mu$ .*  $\mu$  is a parameter that balances the sparsity of the solution and how closely  $Au$  describes the spectrum signal  $f$ . A good choice for  $\mu$  is determined by the actual data. If the number of matching pixels is very large or the background materials have similar spectrum with  $f$ , variation of  $\mu$  will affect the result. Larger values of  $\mu$  may only detect a subset of of these correct pixels. This encourages us to process the algorithm in the following iterative way:

- i Choose  $\mu$  large, get the solution  $u$  and identify the correct pixels.
- ii Remove correct pixels identified in i from  $A$ , get a matrix  $A_2$  with fewer columns. Repeat i until you get as many good pixels as possible.

Experimentally, in a few iterations (2 or 3), we obtain all the correct pixels. The last iteration usually contains both correct and incorrect pixels.

On the other hand, smaller  $\mu$  can solve the case that matching pixels are relatively numerous in the image. Under that situation, we are not supposed to look for a sparse solution, but a quite non-sparse solution.

In the follow contrast examples, we first chose 10 pixels out of 2500 to be the assigned matching pixels, with SNR=20.3. The result of computation found all the 10 correct pixels with 0 false alarms, even when  $\mu$  changed from 1e-2 to 1e-6.



Secondly we chose 2000 pixels out of 2500 to be the assigned matching pixels, with SNR=20.3. The difference was when  $\mu$  changed from 1e-2 to 1e-4, the number of nonzero  $u_i$  changed from 100 to 2003 (also containing a few incorrect pixels), as shown in Table 1. The different results shows the power of  $\mu$  in controlling the sparsity of solution.

TABLE 1. Sparseness Control Results Using Different Value of  $\mu$

| $\mu$ | Correct pixels | Incorrect pixels |
|-------|----------------|------------------|
| 1e-2  | 100            | 0                |
| 1e-3  | 1062           | 0                |
| 1e-4  | 2000           | 3                |

3.1.3. *Mixed Pixels Test.* In many situations, the designed material only partially occupies the mixed pixels. We generated synthetic mixed pixels by adding one or two background materials proportionately to the pure pixels according to a linear mixing model (LMM). We tested the algorithm under different proportions.

We have observed that the algorithm can obtain the most correct matching pixels if the amount of pure material's contribution to the mixed pixels is higher than the others, ranging from 95% to 50%. The result is shown in Table 2 (using SNR=20.3 and  $\mu = 1e - 2$ ).

TABLE 2. Experiment Results Using Mixed Pixels

| Metal | Dirt road | Grass | Correct pixels | Incorrect pixels |
|-------|-----------|-------|----------------|------------------|
| 95%   | 0%        | 5%    | 10             | 0                |
| 95%   | 5%        | 0%    | 10             | 0                |
| 90%   | 0%        | 10%   | 10             | 1                |
| 90%   | 10%       | 0%    | 10             | 2                |
| 70%   | 0%        | 30%   | 10             | 1                |
| 70%   | 30%       | 0%    | 10             | 2                |
| 70%   | 15%       | 15%   | 9              | 1                |
| 60%   | 20%       | 20%   | 9              | 2                |
| 50%   | 30%       | 20%   | 9              | 2                |
| 40%   | 30%       | 30%   | 6              | 2                |

3.1.4. *Comparisons.* The spectral angle map (SAM) involves computing the normalized inner product

$$T(a_i) = \frac{\langle a_i, f \rangle}{\|a_i\| \cdot \|f\|},$$

the cosine of the angle between two vectors, the spectrum of target  $f$  and pixels  $a_i$ , which is always positive because spectral vectors contain only nonnegative components.

The adaptive coherence estimator (ACE) algorithm uses a covariance matrix to qualify the background. It computes a ratio at each pixel

$$T(a_i) = \frac{(f^T \Sigma^{-1} a_i)^2}{(f^T \Sigma^{-1} f)(a_i^T \Sigma^{-1} a_i)},$$

where  $f$  is the target of interest,  $a_i$  is a pixel in the test image,  $\Sigma$  is the covariance matrix of background with zero mean.

Thresholding is done after  $T(a_i)$  is generated for each pixel. Then SAM and ACE denote those pixels as positive for which the value of  $T(a_i)$  exceeds a threshold  $\tau$ , a training data, and denote the others as negative. The performances of SAM and ACE are highly dependent on the training data.

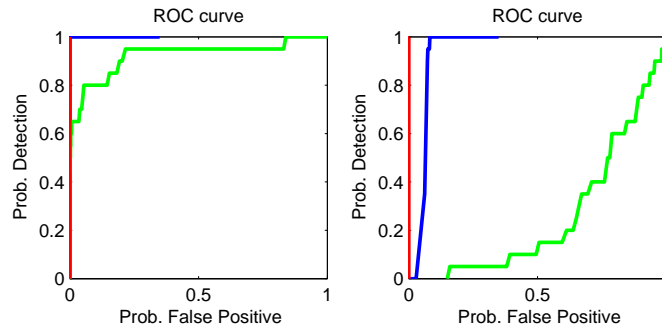


FIGURE 2. ROC Curves of SAM in blue and ACE in green vs  $l_1$  template matching in red for simulated dataset. From left to right, SNR=20, SNR=5.

ROC curves in Fig 2 were drawn to show target detection probability versus false alarms probability by varying a threshold  $\tau$ . We evaluated the performances of the three algorithms on the same simulated data sets. Comparisons between the three algorithms were shown at different noise levels. In Fig 2, red and blue curves in subimage of SNR=20 and red curve in subimage of SNR=5 are all located on the y axis, which are probably invisible. We can see with small enough noise, all three algorithms can detect the correct matching pixels well, while ACE needs a good threshold  $\tau$  to decrease the false alarms. Increasing the noise to SNR=5, we found that  $l_1$  template matching essentially shows zero probability of false positives and worked better than SAM and ACE, which intended to obtain low detection rate when the noise increased.

### 3.2. HSI Data Set.

3.2.1. *Fake Leaves.* We performed  $l_1$  template matching on the Fake Leaves indoor image provided by Surface Optics Company [1] using the SOC-700 hyperspectral imaging system as shown in Fig 3. The HSI data includes 640x640 pixels and 120 bands in the 400-900 nm range. Under RGB model, fake leaves look exactly same as true leaves, however, true leaves have a strong feature in the near infrared.

To perform the algorithm, we selected 15 patches that contain most of the objects present in the image for simplicity. This is shown in Fig 4. Using the template matching algorithm, we detected almost all the correct pixels. We used a red color to show the pixels detected to be fake leaf. As shown in the image, we detected 678

correct and missed around 16 fake leaf pixels among 6000 pixels, the true positive rate equals was 97.69%. All of the missed pixels were located between the fake leaf and true leaves.

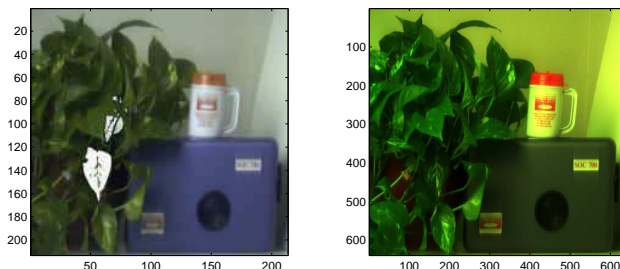


FIGURE 3. Fake Leaves. Left: Image downloaded from SOC website [1], indicating two fake leaves. Right: RGB view of Fake Leaves indoor image.

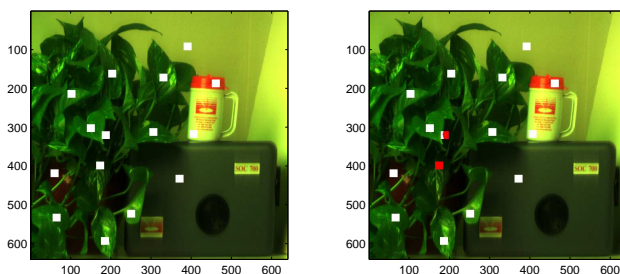


FIGURE 4. Fake leaf detection result. Left: Patches chosen as test pixels. Right: Red pixels are detected to be fake leaves.

3.2.2. *NGA HSI Data Set.* We worked with the hyperspectral image data provided by NGA, as an experimental data for classification given the materials' signatures. It contains 3349 pixels with 106 bands. A total of 9 endmembers are extracted from the data, including Coniferous, Deciduous, Grass, Lake1, Lake2, Crop, Road (Asphalt), Concrete, Gravel. To make matters worse, most of them look quite similar to each other as seen in Fig 5. We performed  $l_1$  minimization and split Bregman to all 9 endmember targets and achieved satisfactory results, which were shown in Table 3.

To get a better result, we used  $\mu = 1e - 4$ ,  $\lambda = \frac{100}{\|A^T A\|_2}$ . Based on the evidence that the coefficients of incorrect pixels were relatively small, a simple threshold can be set up to decrease the number of incorrect pixels as in Table 4.

3.2.3. *Smith Island Data Set.* The Smith Island data set [3, 4, 5, 6] has 946x679 pixels and 113 clean spectral bands after data correction. Fig 6 shows the RGB picture and the 2743 pixels ground truth available based on 22 materials. Fig. 6 shows the spectrum of total 22 materials. Because of the huge number of materials,

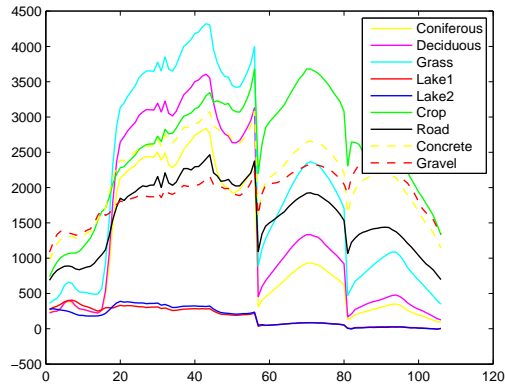


FIGURE 5. Spectrum of endmembers in NGA HSI data.

TABLE 3. Classification Results for HSI Data without Thresholding

| Endmembers     | Total pixels | Correct pixels | Incorrect pixels |
|----------------|--------------|----------------|------------------|
| Coniferous     | 150          | 135            | 53               |
| Deciduous      | 163          | 131            | 83               |
| Grass          | 1338         | 1331           | 98               |
| Lake1          | 202          | 199            | 0                |
| Lake2          | 112          | 112            | 1                |
| Crop           | 1026         | 1026           | 61               |
| Road (Asphalt) | 197          | 161            | 80               |
| Concrete       | 74           | 71             | 8                |
| Gravel         | 87           | 87             | 29               |

TABLE 4. Classification Results for HSI Data with Thresholding

| Endmembers     | Total pixels | Threshold | Correct pixels | Incorrect pixels |
|----------------|--------------|-----------|----------------|------------------|
| Coniferous     | 150          | 1e-3      | 125            | 26               |
| Deciduous      | 163          | 1e-3      | 124            | 41               |
| Grass          | 1338         | 1e-4      | 1323           | 13               |
| Crop           | 1026         | 1e-5      | 1024           | 9                |
| Road (Asphalt) | 197          | 1e-3      | 149            | 35               |
| Gravel         | 87           | 1e-3      | 85             | 17               |

we split them into two groups and show the spectrum separately as in Fig 7. We employed this 2743 pixels ground truth to test our algorithm. The true positive rate (TPR) and false positive rate (FPR) were computed for each material.

In order to see the effect of the algorithm, we compared the result of the spectral angle map (SAM) and adaptive coherence estimator (ACE) using ROC curves while changing the threshold  $\tau$ . Fig 8 shows the comparison between the three experiment's result for all 22 materials. Some curves are probably invisible because they are located on the y axis. As shown in Fig 8, almost all the red lines stay to the upper left position of green lines and blue lines, which means the experiments get better result than SAM and ACE. The reason for this is because our algorithm is more robust when the noise increases.

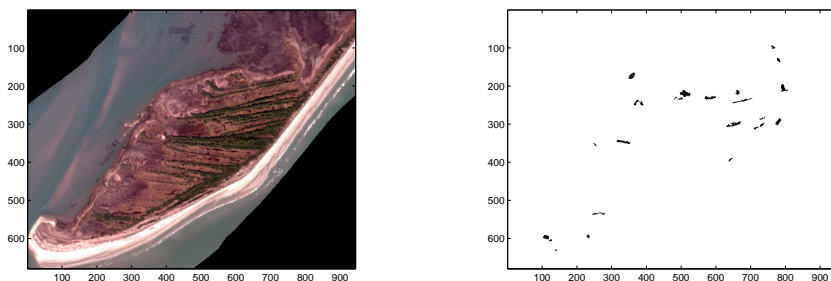


FIGURE 6. Smith Island: Up: RGB view of Smith Island data. Low: Ground truth of 22 materials.

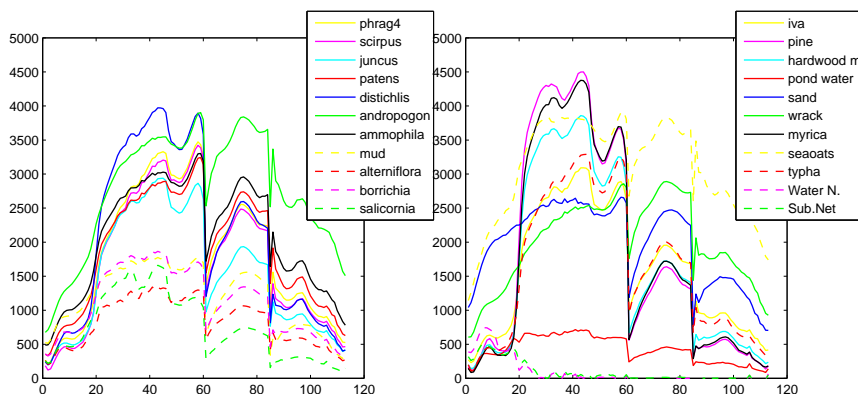


FIGURE 7. Spectrum of 22 materials in Smith Island data.

3.2.4. *RIT self test Data and Comparisons.* The data set selected for this comparative evaluation is one that is used for target detection scoring [30], which is available at <http://dirs.cis.rit.edu/blindtest/>. The self test data used in the experiment was collected in the small town of Cooke City, Montana, USA as in Fig 9. Multiple vehicles and fabric panels were provided as the targets.

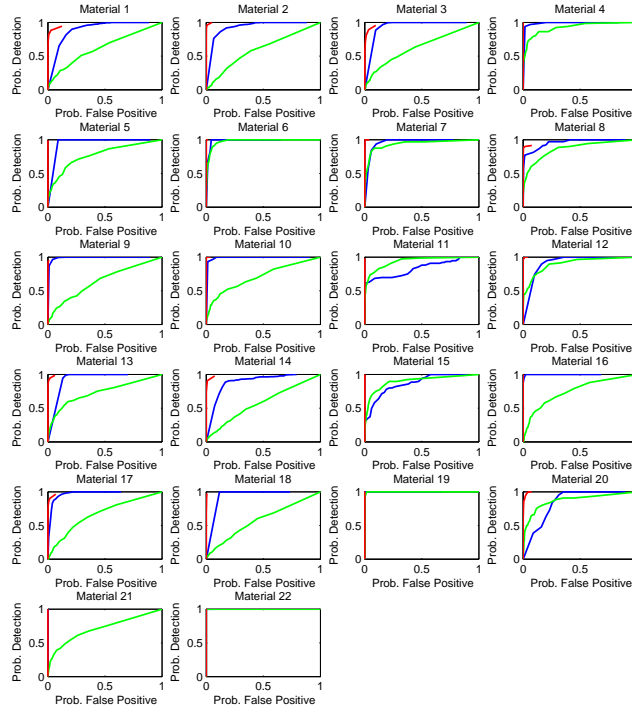


FIGURE 8. ROC Curves of SAM (blue line), ACE (green line) and  $l_1$  algorithm (red line)

We choose a 60x60 spatial ground truth with 126 spectral bands, and material “3m Red Cotton” as the target to evaluate the results between five algorithms:  $l_1$  template matching, spectral angle map (SAM), adaptive coherence estimator (ACE), generalized-likelihood ratio test (GLRT) and matched filter. The parameters of all the algorithms have been tuned to be the best. Fig 10 shows  $l_1$  template matching outperforms the other matching algorithms.

**4. Conclusion.** In this paper, we propose an  $l_1$  minimization based approach for template matching and apply it to hyperspectral classification and target detection. This algorithm is simple to understand and implement, and only involves a few lines of code. Numerical results show that the algorithm is efficient in finding the designed pixels in a hyperspectral image data for a specific material, without using any background information. There are many other applications of this algorithm worth exploring. We will extend the applicability of this algorithm in the future. The limitation of the algorithm is that it requires the materials’ spectral signatures to do the computation which is dependent on the existence of endmember extraction methods.



FIGURE 9. RGB view of self test data showing Cooke City.

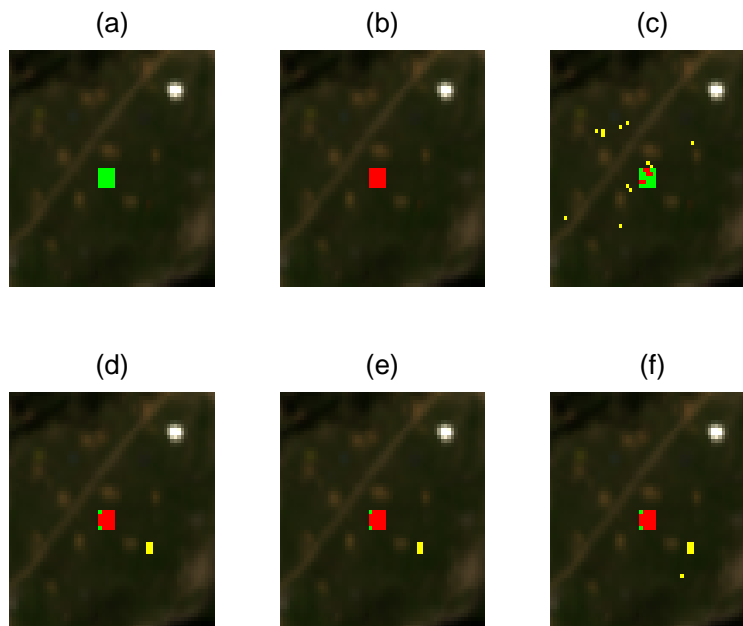


FIGURE 10. Illustrating the computation results, with ground truth in green, detected pixels in red and false alarms in yellow: (a) The ground truth of target material. (b)-(f) Results obtained by  $l_1$  template matching, SAM, ACE, GLRT and matched filter.

REFERENCES

[1] Surface Optics Corporation, <http://www.surfaceoptics.com>.  
 [2] Urban hyperspectral data set, <http://www.agc.army.mil/Hypercube/>.  
 [3] C. Bachmann, T. Donato, G. Lamela, W. Rhea, M. Bettenhausen, R. Fusina, K. Du Bois, J. Porter and B. Truitt, *Automatic classification of land cover on Smith Island, VA, using*

- HyMAP imagery*, IEEE Transactions on Geoscience and Remote Sensing, **40** (2002), 2313–2330.
- [4] C. Bachmann, *Improving the performance of classifiers in high-dimensional remote sensing applications: An adaptive resampling strategy for error-prone exemplars (ARESEPE)*, IEEE Transactions on Geoscience and Remote Sensing, **41** (2003), 2101–2112.
- [5] C. Bachmann, T. Ainsworth and R. Fusina, *Exploiting manifold geometry in hyperspectral imagery*, IEEE Transactions on Geoscience and Remote Sensing, **43** (2005), 441–454.
- [6] C. Bachmann, T. Ainsworth and R. Fusina, *Improved manifold coordinate representations of large scale hyperspectral imagery*, IEEE Transactions on Geoscience and Remote Sensing, **44** (2006), 2786–2803.
- [7] J. Bioucas-Dias and M. Figueiredo, *Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing*, 2<sup>nd</sup> Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing - WHISPERS, Reykjavik, Iceland, (2006).
- [8] L. Bregman, *The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming*, USSR Comput Math and Math. Phys., **7** (1967), 200–217.
- [9] J. Cai, S. Osher and Z. Shen, *Split Bregman methods and frame based image restoration*, Multiscale Model. Simul., **8** (2009), 337–369.
- [10] E. Candes, J. Romberg and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Transactions on Information Theory, **52** (2006), 489–509.
- [11] E. Conte, M. Lops and G. Ricci, *Asymptotically optimum radar detection in compound-Gaussian clutter*, IEEE Transactions on Aerospace Electron. Syst., **31** (1995), 617–625.
- [12] G. Dimitris, A. Gary and K. Nirmal, *Comparative analysis of hyperspectral adaptive matched filter detectors*, SPIE., **4049** (2000), 2–17.
- [13] D. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, **52** (2006), 1289–1306.
- [14] T. Goldstein and S. Osher, *The split Bregman algorithm for  $L_1$  regularized problems*, SIAM Journal on Imaging Sciences, **2** (2009), 323–343.
- [15] T. Goldstein, X. Bresson and S. Osher, “Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction,” UCLA CAM Report, **9** (2009).
- [16] Z. Guo, T. Wittman and S. Osher,  *$L_1$  unmixing and its application to hyperspectral image enhancement*, in Proc. SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery, **XV** (2009).
- [17] J. Harsanyi and C. Chang, *Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach*, IEEE Transactions on Geoscience and Remote Sensing, **32** (1994), 779–785.
- [18] S. Kay, “Fundamentals of Statistical Signal Processing,” Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [19] S. Kraut and L. Scharf, *The CFAR adaptive subspace detector is a scale-invariant GLRT*, IEEE Transactions on Signal Processing, **47** (1999), 2538–2541.
- [20] S. Kraut, L. Scharf and L. McWhorter, *Adaptive subspace detectors*, IEEE Transactions on Signal Processing, **49** (2001), 1–16.
- [21] F. Kruse, A. Lefkoff, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon and A. Goetz, *The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data*, Rem. Sens. Environ., **44** (1993), 145–164.
- [22] H. Kwon and N. Nasrabadi, *Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery*, IEEE Transactions on Geoscience and Remote Sensing, **43** (2005), 388–397.
- [23] D. Manolakis and G. Shaw, *Detection algorithms for hyperspectral imaging applications*, IEEE Signal Processing Magazine, **19** (2002), 29–43.
- [24] D. Manolakis, *Detection algorithms for hyperspectral imaging applications: A signal processing perspective*, IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, (2003), 378–384.
- [25] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin, *An iterative regularization method for total variation based image restoration*, Multiscale Model. Simul., **4** (2005), 460–489.
- [26] S. Osher, Y. Mao, B. Dong and W. Yin, *Fast linearized Bregman iteration for compressive sensing and sparse denoising*, Commun. Math. Sci., **8** (2010), 93–111.
- [27] I. Reed and X. Yu, *Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution*, IEEE Transactions on Acoustics, Speech and Signal Processing, **38** (1990), 1760–1770.



- [28] F. Robey, D. Fuhermann, E. Kelly and R. Nitzberg, *A CFAR adaptive matched filter detector*, IEEE Transactions on Aerospace and Electronic Systems, **28** (1992), 208–216.
- [29] L. Scharf and B. Friedlander, *Matched subspace detectors*, IEEE Transactions on Signal Processing, **42** (1994), 2146–2157.
- [30] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive and S. Hager, *Development of a web-based application to evaluate target finding algorithms*, Proceedings of the 2008 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), **2** (2008), 915–918.
- [31] D. Stein, S. Beaven, L. Hoff, E. Winter, A. Schaum and A. Stoker, *Anomaly detection from hyperspectral imagery*, IEEE Signal Processing Magazine, **19** (2002), 58–69.
- [32] A. Szlam, Z. Guo and S. Osher, *A split Bregman method for non-negative sparsity penalized least squares with applications to hyperspectral demixing*, UCLA CAM report, **10-06** (2010).
- [33] W. Yin, S. Osher, D. Goldfarb and J. Darbon, *Bregman iterative algorithms for  $l_1$ -minimization with applications to compressed sensing*, SIAM J. Imaging Sci., **1** (2008), 143–168.
- [34] X. Yu, I. Reed and A. Stocker, *Comparative performance analysis of adaptive multispectral detectors*, IEEE Transactions on Signal Processing, **41** (1993), 2639–2656.
- [35] X. Zhang, M. Burger, X. Bresson and S. Osher, *Bregmanized nonlocal regularization for deconvolution and sparse reconstruction*, preprint, (1993).

Received May 2010; revised August 2010.

*E-mail address:* annegzh@math.ucla.edu

*E-mail address:* sjo@math.ucla.edu