REGULARIZED SPECTRAL MATCHED FILTER FOR TARGET DETECTION IN HYPERSPECTRAL IMAGERY

NASSER M. NASRABADI

US Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD 20783

ABSTRACT

This paper describes a new adaptive spectral matched filter that incorporates the idea of regularization (shrinkage) to penalize and shrink the filter coefficients to a range of values. The regularization has the effect of restricting the possible matched filters (models) to a subset which are more stable and have better performance than the non-regularized adaptive spectral matched filters. The effect of regularization depends on the form of the regularization term and the amount of regularization is controlled by so called the regularization coefficient. Experimental results for detecting targets in hyperspectral imagery are presented for regularized and non-regularized spectral matched filters.

Index Terms—spectral matched filter, regularization, shrinkage, hyperspectral imagery

1. INTRODUCTION

Target detection using a Spectral Matched Filter (SMF) is a well-known approach in detecting objects of interest in hyperspectral imagery [1]-[3]. SMF is based on the assumption of a linear model where the spectral signature of the target and the background clutter covariance matrix are assumed to be known. Typically, the target spectral signature is obtained from a spectral library or from a set of training data, which is used in conjunction with the estimated covariance matrix of the background clutter data. In the adaptive SMF the background clutter covariance matrix is estimated from a small number of samples surrounding the test pixel in order to adapt the matched filter to the local statistics. The expression for the adaptive spectral matched filter involves the spectral signature of the desired target and the inverse of the local clutter covariance matrix. However, this covariance matrix is usually singular and rank deficient due to the high dimensionality of the hyperspectral data and the use of a small number of samples surrounding the test pixel. This is the main reason for the poor performance of the adaptive spectral matched filter. Representing the inverse covariance matrix in terms of its eigenvalue decomposition, it becomes clear that the behavior of the inverse covariance matrix depends on the small eigenvalues which could make the inverse unstable. In order to reduce its sensitivity to statistical errors we could stabilize the inverse by discarding a number of eigenvectors with small eigenvlaues [3] or by adding a scaled identity matrix to the background covariance matrix before inverting.

In this paper, the SMF design is first formulated as maximizing the signal-to-background clutter ratio with a constraint on minimizing the L_2 norm of the filter coefficients which is referred to as the regularization term. We show that this is equivalent to adding a scaled identity matrix to the background clutter covariance in the design of the SMF. Scaling the identity matrix as shown in Section III controls the amount of regularization, the scale is referred to as the regularization coefficient. Adding a regularization term to a cost function in order to control over-fitting or reducing the model complexity is known as shrinkage methods [4]. For example, the well known ridge regression [5] (regularized least squares) includes a L_2 norm constraint on the regression coefficients in its regression model. Another shrinkage method similar to the ridge regression problem is the *lasso* estimator [6] where the L_1 norm of the regression coefficients is used as the regularization term. Similar regularization ideas are used in neural networks and linear classifiers. In neural networks, it is known as *weight decay* [5] which is used for pruning and reducing the number of weights while in the classification literature it is known as maximal margin classifiers [7].

This paper is organized as follows. Section 2 introduces the linear matched filter as maximizing the signal-to-clutter ratio. In Section 3 a regularized spectral matched filter is described by including a quadratic regularization term. Performance of the SMF and regularized SMF on hyperspectral imagery is provided in Section 4 and conclusions are given in Section 5.

2. ADAPTIVE SPECTRAL MATCHED FILTER

To define an adaptive spectral matched filter, let the input data be modeled as a linear superposition of the desired target spectral signature and background clutter noise given by

$$\mathbf{x} = a\mathbf{s} + \mathbf{n} \ , \tag{1}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_J]^T$ is the observation spectral sample with J spectral bands, $\mathbf{s} = [s_1, s_2, \ldots s_J]^T$ is the spectral signature of the desired targets, a is the attenuation constant (target abundance measure), and $\mathbf{n}$ is the background clutter noise. When $a = 0$ no target is present and when $a > 0$ a target is present.

We start by formulating our adaptive spectral matched filter $\mathbf{w}$ as a projection that maximizes the signal-to-clutter ratio given by

$$SCR(\mathbf{x}) = \frac{\mathbf{w}^T \mathbf{s}}{\mathbf{w}^T \hat{\mathbf{C}} \mathbf{w}} ,$$ (2)

where $\mathbf{w} = [w_1, w_2, \ldots, w_J]^T$ defines the filter coefficients with the constraint that $\mathbf{w}^T \mathbf{s} = 1$ and $\mathbf{w}^T \hat{\mathbf{C}} \mathbf{w}$ is the background clutter variance. $\hat{\mathbf{C}} = \frac{1}{N} \mathbf{X} \mathbf{X}^T$ is the estimated covariance matrix of the background clutter and $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ is a $J \times N$ matrix of N centered (mean-removed) reference pixels obtained from the test input image. Maximizing (2) is equivalent to minimizing the cost function $E(\mathbf{w}, \lambda)$ with respect to $\mathbf{w}$

$$E(\mathbf{w}, \lambda) = \mathbf{w}^T \hat{\mathbf{C}} \mathbf{w} + \lambda (\mathbf{w}^T \mathbf{s} - 1) ,$$ (3)

where λ is the Lagrangian multiplier. Differentiating (3) with respect to $\mathbf{w}$ gives

$$\frac{\partial E}{\partial \mathbf{w}} = 2\hat{\mathbf{C}} \mathbf{w} + \lambda \mathbf{s} .$$ (4)

Setting (4) equal to zero the matched filter is given by

$$\mathbf{w} = -\frac{\lambda}{2} \hat{\mathbf{C}}^{-1} \mathbf{s} .$$ (5)

The Lagrangian multiplier, λ, is found using the constraint equation as

$$\mathbf{w}^T \mathbf{s} = -\frac{\lambda}{2} \mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{s} = 1 ,$$ (6)

and therefore,

$$-\frac{\lambda}{2} = \frac{1}{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{s}} .$$ (7)

Substituting (7) into (5) the matched filter is given by

$$\mathbf{w} = \frac{\hat{\mathbf{C}}^{-1} \mathbf{s}}{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{s}} ,$$ (8)

the output of the filter for a given input $\mathbf{x}$ is given by

$$y(\mathbf{x}) = \mathbf{w}^T \mathbf{x} = \frac{\mathbf{x}^T \hat{\mathbf{C}}^{-1} \mathbf{s}}{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{s}} .$$ (9)

Minimizing $E(\mathbf{w}, \lambda)$ is also known as the best linear unbiased estimator (BLUE) [1].

The performance of a SMF that uses global background clutter statistics has been found to be much worse than an adaptive SMF that uses the local clutter statistics around the test pixel. In our adaptive SMF, the background covariance matrix was calculated locally by using a dual window centered at each pixel. The dual window consists of two regions, an inner region approximately the same size as the target and a larger outer region representing the local background clutter. The background clutter statistics is only obtained from the pixels within the outer region, this way the spectral matched filter is made adaptive to the local clutter statistics. Since a relatively small set of samples are used to represent the background clutter statistics around each test pixel, the background clutter covariance matrix needs to be regularized in order to obtain a numerically stable pseudo-inverse as discussed in the next Section 3.

3. REGULARIZED SPECTRAL MATCHED FILTER

Due to the high dimensionality of the input data a large number of samples are required to estimate the true value of the covariance matrix of the background. However, in the adaptive matched filter the covariance matrix is estimated by using only a small number of spectral pixels surrounding the test region. Therefore, the estimated $\hat{\mathbf{C}}$ is singular and its inverse does not exist and its pseudo-inverse could be unstable due to very small eigenvalues. One approach to alleviate this problem is to eliminate some of the eigenvectors with small eigenvalues by using the singular value decomposition of the inverse matrix [3]. Another approach is to include a regularization term in the cost function $E(\mathbf{w}, \lambda)$. The regularization usually imposes a numerical constraint on the size of the matched filter coefficients. The regularized cost function $E(\mathbf{w}, \lambda, \beta)$ is now given by

$$E(\mathbf{w}, \lambda, \beta) = \mathbf{w}^T \hat{\mathbf{C}} \mathbf{w} + \lambda (\mathbf{w}^T \mathbf{s} - 1) + \beta \Omega(\mathbf{w}) ,$$ (10)

where β is the regularization coefficient and $\Omega(\mathbf{w})$ is the regularization term given as

$$\Omega(\mathbf{w}) = \sum_{j=1}^{J} |w_j|^q ,$$ (11)

for $q \geq 0$. The effect of the regularization term is two folds first it makes sure that the estimated $\hat{\mathbf{C}}$ is non-singular and secondly it shrinks the filter coefficients to a range of values with mean zero. The behavior of regularization depends on the value of q. When $q = 1$ it corresponds to the lasso shrinkage method which forces some of the filter coefficients to zero resulting in a sparse matched filter. Regularization with $q = 2$ corresponds to ridge regression or weight decay in the neural network literature.

The regularization constraint used in our penalized matched filter is $q = 2$ which corresponds to the quadratic regularization $\Omega(\mathbf{w}) = \beta \mathbf{w}^T \mathbf{w}$.

The cost function is now given by

$$E(\mathbf{w}, \lambda, \beta) = \mathbf{w}^T \hat{\mathbf{C}} \mathbf{w} + \lambda(\mathbf{w}^T \mathbf{s} - 1) + \beta \mathbf{w}^T \mathbf{w} . \quad (12)$$

Differentiating (12) with respect to $\mathbf{w}$ gives

$$\frac{\partial E}{\partial \mathbf{w}} = 2\hat{\mathbf{C}}\mathbf{w} + \lambda \mathbf{s} + 2\beta \mathbf{w} , \quad (13)$$

and setting it to zero we can solve for $\mathbf{w}$

$$\mathbf{w} = -\frac{\lambda}{2}(\hat{\mathbf{C}} + \beta \mathbf{I})^{-1}\mathbf{s} . \quad (14)$$

Using the constraint equation $\mathbf{w}^T \mathbf{s} = 1$, the Lagrangian multiplier λ is found as

$$-\frac{\lambda}{2} = \frac{1}{\mathbf{s}^T(\hat{\mathbf{C}} + \beta \mathbf{I})^{-1}\mathbf{s}} . \quad (15)$$

Substituting (15) into (14) the regularized matched filter is now given by

$$\mathbf{w} = \frac{(\hat{\mathbf{C}} + \beta \mathbf{I})^{-1}\mathbf{s}}{\mathbf{s}^T(\hat{\mathbf{C}} + \beta \mathbf{I})^{-1}\mathbf{s}}, \quad (16)$$

the output of the filter for a given input $\mathbf{x}$ is given by

$$y(\mathbf{r}) = \mathbf{W}^T \mathbf{x} = \frac{\mathbf{s}^T(\hat{\mathbf{C}} + \beta \mathbf{I})^{-1}\mathbf{x}}{\mathbf{s}^T(\hat{\mathbf{C}} + \beta \mathbf{I})^{-1}\mathbf{s}} . \quad (17)$$

This is the expression for the regularized spectral matched filter used in our implementation. Compared with SMF (9) it can be seen that the regularized matched filter suppresses the influence of the low eigenvalues by forcing the background clutter covariance matrix to become more isotropic (whitened). At high regularization ($\beta \gg 1$), (17) is equivalent to a SMF with a uniform background clutter noise, the resulting filter will be given by $\mathbf{w} = \mathbf{s}/\mathbf{s}^T\mathbf{s}$. At no regularization ($\beta = 0$) (17) is equivalent to (8) an unstable adaptive spectral matched filter.

4. EXPERIEMNTAL RESULTS

The real hyperspectral images are from a HYDICE (HYperspectral Digital Imagery Collection Experiment) sensor. The HYDICE imaging sensor generates 210 bands across the whole spectral range (0.4 -- 2.5 µm). But we only use 150 bands by discarding water absorption and low SNR bands; the spectral bands used are the 23rd--101st, 109th--136th, and 152nd--194th. Two HYDICE images from the Desert Radiance II (DR-II) data collection and the Forest Radiance I (FR-I) data collection were used to test both the conventional and regularized spectral matched filters. The DR-II image contains 6 military targets located in the dirt road; FR-I image includes total of 14 military targets along the tree line. Fig. 1 (a) and (b) show FR-I and DR-II hyperspectral images of 150 spectral bands, respectively. All the pixel vectors in a test image are first normalized by a constant, which is a maximum value obtained from all the spectral components of the spectral vectors in the corresponding test image, so that the entries of the normalized pixel vectors fit into the interval of spectral values between zero and one.
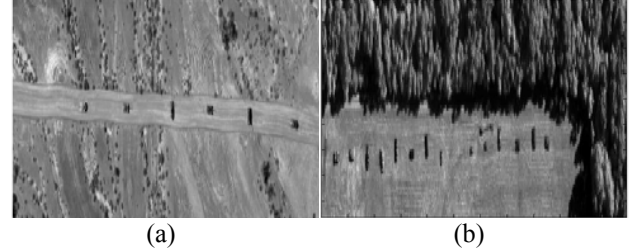

(a)   (b)

Fig. 1. (a) Original Forest Radiance, (b) original Desert Radiance

The receiver operating characteristics (ROC) curves representing detection probability versus false alarm rates were generated to provide quantitative performance comparison as well as qualitative performance comparison. To generate the ground truth information we obtained the coordinates of all the target pixels within a rectangular region containing each target. All the pixels on the target were considered as desired target points to be detected. Our ROC curves, based on the ground truth information from the HYDICE images, represent the number of target pixels detected verses false alarm rate. In all the experimental results the ground truth target spectral signatures for DR-II and FR-I were obtained by averaging the target samples collected from the left most target in the corresponding test image.
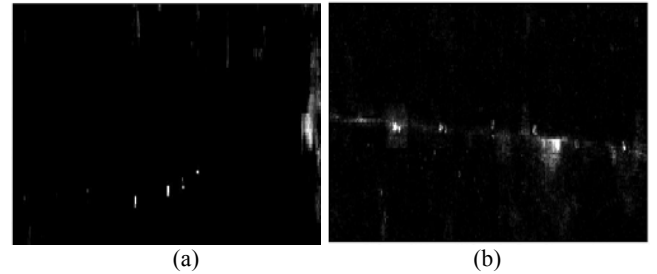

(a)   (b)

Fig. 2. (a) SMF with $\beta = 0$ no regularization for FR-I image b) SMF with $\beta = 0$ no regularization for DR-II image

Simulation was performed on several hyperspectral images and ROC curves were plotted for the regularized and the conventional matched filters. Experimental results for the regularized matched filter for different values of the regularization coefficient β is provided through ROC curves and compared with the non-regularized $\beta = 0$ result. As can be seen from the ROC plots in Fig. 4, the best values for the regularization coefficient was found to be $\beta = 0.1$. Fig. 2 (a) and (b) show the outputs of the conventional

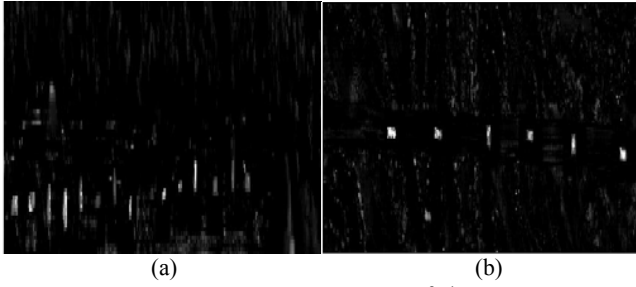adaptive SMF for FR-I and DR-II with no regularization, respectively.

<center>(a)   (b)</center>

Fig. 3. (a) Regularized SMF with $\beta = 0.1$ for FR-I image, b) Regularized SMF with $\beta = 0.1$ for DR-II image

Fig. 3 (a) and (b) show the outputs of the regularized SMF for FR-I and DR-II with the regularization coefficient equal to 0.1, respectively. Fig. 4 (a) and (b) show the ROC plots for the conventional adaptive SMF and for the regularized adaptive SMF at several different regularization coefficients for FR-I and DR-II, respectively. Overall, the performance of the regularized spectral matched filter is much better than the conventional spectral matched filter which uses no regularization.
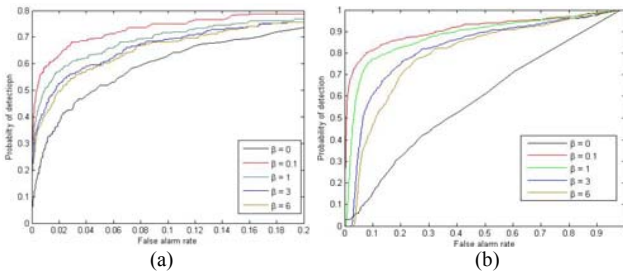

<center>(a)   (b)</center>

Fig. 4 ROC plots for (a) FR-I image and (b) DR-II image at different values of $\beta = 0, 0.1, 1, 3,$ and 6.
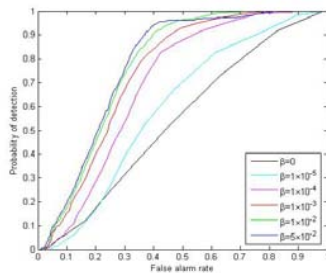

Fig. 5 ROC plots for a global matched filter applied on DR-II image at different values of β.

To obtain the effect of regularization on the filter coefficients we designed a single global matched filter for the test image DR-II by using a background covariance matrix obtained from the whole image. Fig. 5 show the ROC curves for several different β values. Fig. 6 (a) shows

the matched filter coefficients for $\beta = 0$ which has extreme positive and negative values. Fig. 6 (b) shows the filter coefficients for several different β values. At large β values the filter coefficients tend to become less oscillatory and more smooth.

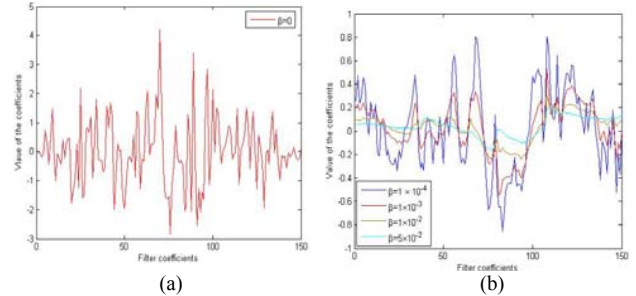<center>(a)   (b)</center>

Fig. 6 (a) Filter coefficients values for $\beta = 0$ and (b) filter coefficients values for several values of β.

5. CONCLUSIONS

We have extended the conventional SMF detector to a regularized version by including a regularization term that forces the filter coefficients to shrink and become smooth. The conventional SMF and the regularized SMF were implemented with several different values for the regularization coefficient. In general, regularized SMF with appropriate regularization coefficient showed a superior detection performance when compared to the conventional SMF for the HYDICE images tested in this paper. Different forms of regularization will impose different prior distributions on the filter models which need to be studied in the future research. Kernel spectral matched filter [8] can also be regularized for improved performance.

6. REFERENCES

[1] Kay, S.M., *Fundamentals of Statistical Signal Processing Estimation Theory*, Prentice Hall, 1993.

[2] D. Manolakis, G. Shaw, and N. Keshava, "Comparative Analysis of Hyperspectral Adaptive Matched Filter Detector," SPIE, vol. 4049, pp. 2-17, April 2000.

[3] P.V. Villeneuve, H.A. Fry, J. Theiler, and W.B. Clodius, "Improved Matched-Filter Detection Techniques," SPIE, vol. 3753, pp. 278-285, 1999.

[4] Hastie, T., and R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Springer, 2001.

[5] Bishop, C.M., *Pattern Recognition and Machine Learning*, Springer, 2006.

[6] Tibshirani, R., "Regression Shrinkage and Selection via the Lasso," *J. Royal. Statist. Soc. B.*, vol. 58, pp. 671-686, 1999.

[7] Scholkopf, B., and A.J. Smola, *Learning with Kernels*, The MIT Press, 2002.

[8] H. Kwon, and N.M. Nasrabadi, "Kernal Spectral Matched filter," *International Journal of Computer Vision*, vol. 71, no. 2, pp. 127-141, 2007.