Parametric Adaptive Modeling and Detection for Hyperspectral Imaging

Hongbin Li
ECE Department, Stevens Institute of Technology
Hoboken, NJ 07030, USA
E-mail: hli@stevens-tech.edu

James H. Michels
AFRL/SNRT, 26 Electronics Parkway
Rome, NY 13441, USA
E-mail: James.Michels@rl.af.mil

***Abstract*— Hyperspectral imaging (HSI) sensors can provide very fine spectral resolution that allows remote identification of ground objects smaller than a full pixel. Traditional approaches to the so-called subpixel target signal detection problem involve the estimation of the sample covariance matrix of the background from target-free training pixels. This entails a large training requirement and high complexity. In this paper, we investigate parametric adaptive modeling and detection for HSI applications. To deal with non-stationarity in the spectral dimension that is characteristic of HSI data, we introduce a sliding-window based time-varying (TV) autoregressive (AR) modeling and detection technique, by which the spectral data is sliced into overlapping subvectors for parameter estimation and signal whitening. Experimental results using real HSI data show that the proposed parametric technique outperforms conventional detection schemes, especially when the training size is small.**

I. Introduction

Adaptive signal detection from hyperspectral imaging (HSI) data has received significant interest recently [1]. A challenging problem for HSI applications is the so-called *subpixel target signal detection*, whereby the target object occupies only a portion of a full pixel. Traditional approaches to this problem rely on an estimate of the background covariance matrix obtained from target-free training pixels for interference or background suppression [1]. Such techniques suffer two drawbacks, namely a *large training requirement* and *high complexity*. In particular, reliable estimation of the background covariance matrix requires a large number of target-free training pixels, which may not be available in a non-homogeneous environment. Meanwhile, the covariance matrix has to be estimated and inverted for each test pixel, which involves significant computation.

To address the above issues, we consider in this paper parametric models that capture the characteristics of HSI data in the spectral dimension, and exploit such models for efficient target signal detection. Adaptive target signal detection based on multichannel parametric modeling has been successfully utilized in airborne radar systems equipped with an antenna array [2], [3]. Extending the idea for HSI applications, however, is challenging since, unlike the former case where the interference and radar clutter can be approximated reasonably well to be *stationary* under mild conditions, HSI data is highly *non-stationary* in the spectral dimension. To deal with non-stationarity, we consider a sliding-window based time-varying (TV) autoregressive (AR) modeling and detection technique, by which we slice the spectral data into overlapping subvectors for parameter estimation and signal whitening. Experimental results using real HSI data show that the proposed parametric technique outperforms conventional schemes, especially when the training size is small.

This work was supported by the Air Force Research Laboratory (AFRL) Visiting Faculty Research Program (VFRP).

II. Problem Formulation

HSI data is usually described as a datacube, whose face is a function of the spatial coordinates and depth is a function of spectral bands or wavelengths. Each pixel can be represented as an $L \times 1$ *real-valued* vector: $\boldsymbol{x} = [x(0), x(1), \ldots, x(L-1)]^T$, where L denotes the total number of spectral bands, $x(l)$ denotes the spectral response at the lth spectral band, and $(\cdot)^T$ denotes transpose. Since HSI data has non-zero mean [1], [4], a preprocessing stage is usually invoked to remove the sample mean estimated using neighboring pixels.

In vector notation, the *subpixel target signal detection* problem is described by the following composite hypothesis test [1]:

$$\begin{aligned} H_0: &\quad \boldsymbol{x} = \boldsymbol{b}, && \text{target absent} \\ H_1: &\quad \boldsymbol{x} = a\boldsymbol{s} + \boldsymbol{b}, && \text{target present} \end{aligned} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^{L\times 1}$ is the *demeaned* test pixel, $\boldsymbol{s} \in \mathbb{R}^{L\times 1}$ is the signature vector of the target object with amplitude a, and $\boldsymbol{b} \in \mathbb{R}^{L\times 1}$ denotes the background plus system noise. We adopt the standard assumption that the signature vector $\boldsymbol{s}$ is deterministic and known to the detector;[1] the amplitude a, however, is assumed unknown. For the background, we follow a statistical approach that models the background interference $\boldsymbol{b}$ as a multivariate Gaussian random vector with zero mean and covariance matrix $\boldsymbol{R}_b \triangleq E\{\boldsymbol{b}\boldsymbol{b}^T\}$. The Gaussian assumption has been experimentally justified for multispectral data (e.g., [4]), and the extension to HSI data is widely adopted [1]. It leads to mathematical tractability and, most importantly, provides good performance in many practical situations.

The problem is to find the decision rule for hypothesis testing (1), given knowledge of the test pixel $\boldsymbol{x}$, target signal signature $\boldsymbol{s}$, and a number of training pixels $\{\boldsymbol{x}_n\}_{n=1}^N$ that are target-free.

III. Conventional Techniques

If the covariance matrix $\boldsymbol{R}_b$ is known exactly, the optimum detector for (1) is the *matched filter (MF)* [5]:

$$\frac{|\boldsymbol{s}^T\boldsymbol{R}_b^{-1}\boldsymbol{x}|^2}{\boldsymbol{s}^T\boldsymbol{R}_b^{-1}\boldsymbol{s}} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\text{MF}}, \tag{2}$$

[1]The spectral signature may vary due to variations in atmospheric conditions and other factors. Such uncertainty can be captured by a linear mixing model [1].

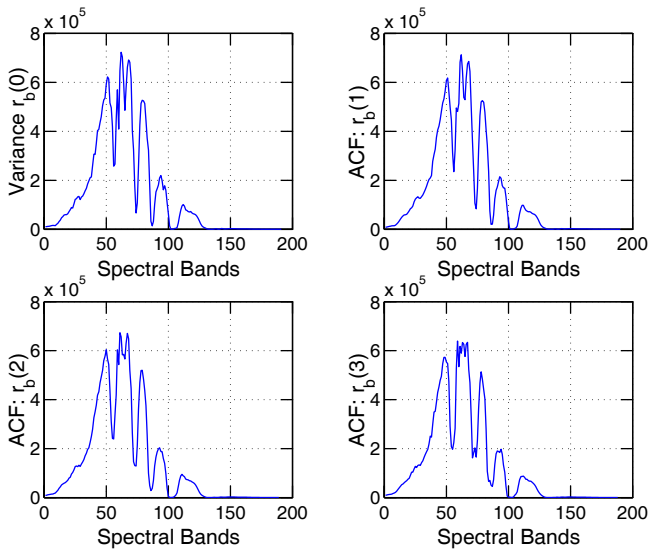

Fig. 1. Sample estimates of the autocorrelation function (ACF) at frequency lag 0 (variance), lag 1, lag 2 and lag 3 across the spectral bands.

where t_{MF} denotes the MF threshold. In practice, the unknown $\boldsymbol{R}_b$ is usually replaced by some estimate, such as the sample covariance matrix: $\hat{\boldsymbol{R}}_b = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{x}_n \boldsymbol{x}_n^T$, where $\{\boldsymbol{x}_n\}_{n=1}^{N}$ denote a set of training pixels, assumed target-free. Using $\hat{\boldsymbol{R}}_b$ in (2) leads to the *adaptive matched filter (AMF)* [5]:

$$\frac{|\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}|^2}{\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{s}} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\text{AMF}}, \tag{3}$$

where t_{AMF} denotes the AMF threshold. Another popular detector is the *adaptive coherence estimator (ACE)* test [6]:

$$\frac{|\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x}|^2}{(\boldsymbol{s}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{s})(\boldsymbol{x}^T \hat{\boldsymbol{R}}_b^{-1} \boldsymbol{x})} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\text{ACE}}, \tag{4}$$

which can be thought of as a normalized version of the AMF test (3). The normalization incurred by the ACE test is important since it accounts for power variations between the training and test data. By the Schwartz inequality, the ACE test statistic is bounded between zero and one.

Both the AMF and ACE tests have constant false alarm rate (CFAR). They also suffer two drawbacks. The first is a *large training requirement*. The covariance matrix $\boldsymbol{R}_b$ has a dimension of $L \times L$. Typical values for L in real HSI systems are in the range of hundreds. An accurate estimate of the covariance matrix would require a large number of target-free training pixels, which may not be available, especially in non-homogeneous environments. The second is *high complexity*, since $\boldsymbol{R}_b$ has to be estimated and inverted for each test pixel.

IV. Parametric Adaptive Modeling and Detection

It is well-known that the interference suppression ability of the MF detector (2) comes from a whitening procedure. Specifically, the whitening operation takes as inputs the signature vector $\boldsymbol{s}$ and test pixel $\boldsymbol{x}$, and outputs whitened versions $\tilde{\boldsymbol{s}} \triangleq \boldsymbol{R}_b^{-1/2}\boldsymbol{s}$ and $\tilde{\boldsymbol{x}} \triangleq \boldsymbol{R}_b^{-1/2}\boldsymbol{x}$. Following the whitening, the MF reduces to simple correlation of the whitened outputs:

$$\frac{|\tilde{\boldsymbol{s}}^T \tilde{\boldsymbol{x}}|^2}{\tilde{\boldsymbol{s}}^T \tilde{\boldsymbol{s}}} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\text{MF}}. \tag{5}$$

If the whitening operation can be designed or approximated via a parametric model without explicitly estimating $\boldsymbol{R}_b$, then it is conceivable that fewer training pixels are needed, provided that the parametric model is described by a few parameters. This is the essence of parametric model based adaptive detection.

Autoregressive (AR) models have been popular choices for parametric modeling in various applications [7]. Parametric adaptive detection based on multichannel AR models has been considered in [2], [3], [8] for airborne radar systems equipped with multiple antennas. For the problem under study, the $L \times 1$ background vector $\boldsymbol{b}$, or equivalently the observed signal $\boldsymbol{x}$ under H_0, may be assumed to be a scalar AR process which produces the L samples of $\boldsymbol{b}$. If an AR model is appropriate for HSI data, then the detection problem amounts to first estimating the AR coefficients from training data, whitening the signals by a whitening filter constructed from the AR coefficient estimates, and computing the decision statistic from the whitened signals followed by thresholding. For brevity, the above approach is referred to as the *parametric adaptive matched filter (PAMF)*, or *normalized PAMF (NPAMF)* [8] if the decision variable is normalized, similar to the normalization imposed by the ACE detector of (4).

We have tested an AR-based version of PAMF/NPAMF with real HSI data and found they work poorly since AR models are in general not suitable for HSI data. In particular, *HSI data are non-stationary in the spectral dimension while fixed parameter AR models characterize stationary processes*. To see this, we have computed the sample covariance matrix $\hat{\boldsymbol{R}}_b$, from a total of $K = 24 \times 46 = 1104$ training pixels drawn from a homogeneous grass region. Figure 1 depicts the main and 3 subdiagonals of $\hat{\boldsymbol{R}}_b$, which correspond to the autocorrelation function (ACF) at frequency lag 0 (i.e., variance), lag 1, lag 2 and lag 3, respectively, versus the spectral bands. Clearly, the signal is not stationary since the variance and ACF at other lags vary significantly across the spectral bands.

To deal with non-stationarity, we consider a time-varying (TV) [2] AR modeling technique. Specifically, we slice $\boldsymbol{x}_n \triangleq [x_n(0), \ldots, x_n(L-1)]^T$ that corresponds to the nth training pixel into $L - L_s + 1$ overlapping subvectors: $\boldsymbol{x}_n(l) \triangleq \big[x_n(l), \quad \ldots, \quad x_n(l+L_s-1)\big]^T$, $l = 0, \ldots, L - L_s$, where $L_s \leq L$ denotes the length of the subvectors. Equivalently, these subvectors can be thought of as being obtained through a sliding window of size L_s. For small enough L_s, each subvec-

[2] Frequency-varying would be a more appropriate term for the HSI spectral data. Nevertheless, we will retain TV because the spectral data are essentially treated and analyzed using time series analysis techniques.

tor $\boldsymbol{x}_n(l)$ can be modeled as an Mth order AR process:

$$x_n(k) = -\sum_{m=1}^{M} a_m(l)x_n(k-m) + w_n(k), \quad k = l, l+1, \ldots, l+L_s-1;\ n = 1, \ldots, N. \tag{6}$$

Note that different subvectors are associated with different sets of AR coefficients. For simplicity, we consider fixed AR model order M. The choice of M and window size L_s should be made with tradeoffs among several issues. A large M might be desirable since it can provide better fitting to the HSI data. Increasing M, however, would require the window size L_s to increase accordingly since more parameters are to be estimated and, therefore, more data should be provided within each subvector to reduce the variance of parameter estimates. Finally, if L_s is too large, the assumption of stationarity within the subvector may be violated, which can cause significant degradation. Our studies show that the following choices in general provide good tradeoffs among the above issues:

$$8 \le L_s \le 15, \quad M = \lfloor L_s/2 \rfloor, \tag{7}$$

where $\lfloor \cdot \rfloor$ denotes the flooring operator.

Once we have made choices for L_s and M, the next step is to estimate the TV-AR coefficients $\{\boldsymbol{a}(l)\}$. While there are a multitude of estimators available, the least squares (LS) estimator [7] is considered here. Let

$$\boldsymbol{a}(l) \triangleq \left[a_1(l), \quad \ldots, a_M(l)\right]^T,$$
$$\boldsymbol{y}_n(l) \triangleq \left[x_n(l+M), \quad \ldots, \quad x_n(l+L_s-1)\right]^T,$$
$$\boldsymbol{Y}_n(l) \triangleq \begin{bmatrix} x_n(l+M-1) & \ldots & x_n(l) \\ \vdots & \vdots & \vdots \\ x_n(l+L_s-2) & \ldots & x_n(l+L_s-M-1) \end{bmatrix}.$$

The LS estimates of the TV-AR coefficients based on (6) using N training pixels are given by

$$\hat{\boldsymbol{a}}(l) = -\left(\sum_{n=1}^{N} \boldsymbol{Y}_n^T(l)\boldsymbol{Y}_n(l)\right)^{-1}\left(\sum_{n=1}^{N} \boldsymbol{Y}_n^T(l)\boldsymbol{y}_n(l)\right), \quad l = 0, 1, \ldots, L-L_s. \tag{8}$$

The matrix within the first pair of brackets is assumed full rank. This can be ensured with probability one if the number of training pixels N satisfies the following condition: $N \ge \frac{M}{L_s - M}$.

Following parameter estimation, a TV moving-average (MA) filter, defined by the TV taps $\{\hat{a}_l(m)\}_{m=1}^{M}$, is formed to whiten the test pixel $\boldsymbol{x}$ and target signature $\boldsymbol{s}$ (cf. (1)) as follows:

$$\epsilon(l) = x(l) + \sum_{m=1}^{M} \hat{a}_m(l-L_s)x(l-m),$$
$$u(l) = s(l) + \sum_{m=1}^{M} \hat{a}_m(l-L_s)s(l-m), \quad l = L_s, \ldots, L,$$

where $\epsilon(l)$ and $u(l)$ denote the lth output sample of the TV-whitening filter when the input is $\boldsymbol{x}$ and $\boldsymbol{s}$, respectively. Effectively, each set of AR coefficient estimates is used to compute only *one* output sample; as the sliding window shifts to the next position, we use the next set of AR coefficients for whitening.

Finally, the outputs of the TV-MA whitening filter corresponding to the test pixel $\boldsymbol{x}$ and target signature $\boldsymbol{s}$, respectively, are correlated to one another to form the decision variable. The resultant test is referred to as the *TV-PAMF*, or *TV-NPAMF* if the decision variable is normalized:

$$\frac{\left|\sum_{l=L_s}^{L} u(l)\epsilon(l)\right|^2}{\left(\sum_{l=L_s}^{L} u^2(l)\right)} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\text{TV-PAMF}}, \tag{9}$$

$$\frac{\left|\sum_{l=L_s}^{L} u(l)\epsilon(l)\right|^2}{\left(\sum_{l=L_s}^{L} u^2(l)\right)\left(\sum_{l=L_s}^{L} \epsilon^2(l)\right)} \underset{H_0}{\overset{H_1}{\gtrless}} t_{\text{TV-NPAMF}}. \tag{10}$$

There is an additional small modification to the above TV-PAMF/TV-NPAMF detectors. In particular, we found that HSI data exhibit small oscillations, which do not contribute much to detection, meanwhile making parameter estimates more noisy. Lowpass (LP) filtering to first remove those oscillations before applying the above TV-AR modeling appears helpful. For lowpass filtering, we have used a simple FIR moving-average filter with impulse response given by a Kaiser window, whose length is equal to the sliding window size L_s and the shape parameter is 3 (not critical). After the FIR lowpass filtering, we follow the procedure of TV-PAMF or TV-NPAMF. The resulting methods are called the *TV-LP-PAMF* or *TV-LP-NPAMF* detectors.

V. Experimental Results

We have compared our NPAMF, TV-NPAMF and TV-LP-NPAMF detectors with the ACE detector (4) using real HSI data. Only the normalized detectors are considered here since their test statistics are bounded between 0 and 1, which makes it convenient to compare. The results corresponding to their non-normalized counterparts, i.e, AMF, PAMF, TV-PAMF and TV-LP-PAMF, are similar to those presented below, for the case considered here, and are thus omitted.

Figure 2 shows a simulated color infrared (IR) view of an airborne hyperspectral data flightline over the Washington DC area. The data contains $L = 191$ spectral bands in the 0.4 to 2.4 μm region of the visible and IR spectrum. Other information about the data set can be found in [9]. The image shown in Figure 2 was made using bands 60, 27 and 17 for the red, green and blue colors, respectively. The results to be presented are with respect to a test region highlighted in yellow in Figure 2, which is relatively homogeneous and formed by grass. We have tested with other regions and obtained similar results. To simulate the H_1 condition, we superimpose a target signal to the test pixel. The target signal corresponds to the spectral signature of a man-made object, and is scaled according to different target fill factors [1]. The *figure of merit* considered here

Fig. 2. HSI image of the Washington DC Mall with L=191 spectral bands. The test region is highlighted in yellow.

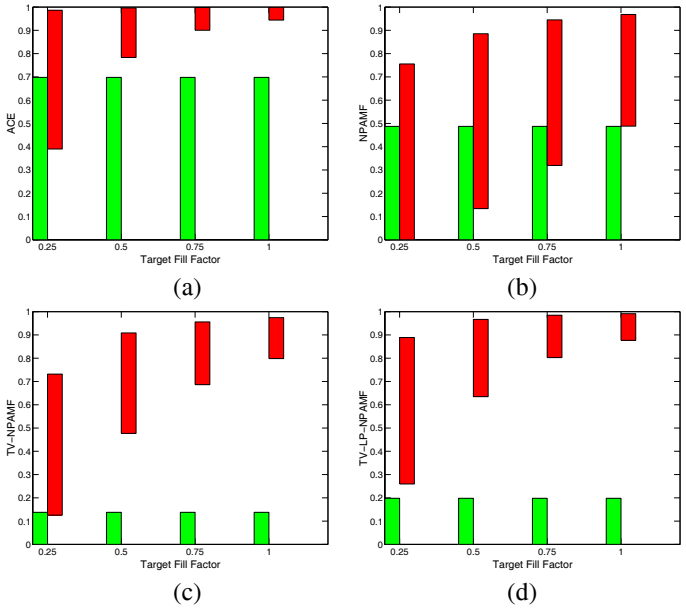

Fig. 3. Target-background separation versus target fill factor, where the red (dark) bars correspond to the range of test statistics under H_1, while the green (light) bars show the counterpart under H_0. (a) ACE. (b) NPAMF. (c) TV-NPAMF. (d) TV-LP-NPAMF.

is the *separation* of test statistics under H_0 and H_1, which is also used in [1]. For all methods, we use $N = 8$ training pixels, which correspond to a 3×3 region without counting the center pixel (i.e., test pixel), for sample covariance matrix or AR coefficient estimation. The sample covariance matrix $\hat{\boldsymbol{R}}_b$ is rank deficient in this case. As suggested in [1], we use the approximation $\hat{\boldsymbol{R}}_b^{-1} \approx \boldsymbol{I} - \boldsymbol{U}_1 \boldsymbol{U}_1^T$, where $\boldsymbol{U}_1$ is formed by the principle eigenvectors of $\hat{\boldsymbol{R}}_b$, for the ACE detector. For the parametric methods, the AR model order is set to $M = 5$, and for the TV parametric methods, the subvector length (i.e., sliding window length) is $L_s = 10$.

Figures 3(a) to 3(d) depict the test statistics separation for the four detectors, namely ACE, NPAMF, TV-NPAMF and TV-LP-NPAMF, respectively, as a function of the target fill factor. We note that NPAMF is the worst of all detectors, which suggests that stationary AR modeling is not adequate for the problem. Both TV-NPAMF and TV-LP-NPAMF outperform the ACE test. TV-LP-NPAMF is slightly better than TV-NPAMF. Specifically, we see that the former achieves full target-background separation when the fill factor is 0.25, while the latter does not.

VI. Concluding Remarks

We have explored parametric modeling and investigated its applicability in HSI systems. We have shown that HSI data are non-stationary in the spectral dimension, which makes parametric adaptive modeling and detection significantly more challenging than some earlier works that assume the data is stationary. To deal with non-stationarity, we have proposed a sliding-window based TV-AR modeling scheme, and developed the corresponding parameter estimation, whitening, and detection techniques. The proposed scheme is shown to outperform the popular ACE detector, in particular when the training size is small. In our experimental study, we have considered only homogeneous background. Spatial heterogeneity may be dealt with by training selection methods based on power sorting [10], and will be examined in the near future.

References

[1] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 29–43, January 2002.

[2] J. R. Román, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 677–692, April 2000.

[3] J. H. Michels, et al., "Multichannel parametric adaptive matched filter receiver," U.S. Patent No. 6,226,321, issued 1 May 2001.

[4] I. S. 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*, vol. 38, no. 19, pp. 1760–1770, October 1990.

[5] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, January 1992.

[6] L. T. McWhorter and L. L. Scharf, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proceedings of 30th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1996, pp. 1114–1117.

[7] S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[8] J. H. Michels, B. Himed, and M. Rangaswamy, "Performance of STAP tests in Gaussian and compound-Gaussian clutter," *Digital Signal Processing*, vol. 10, no. 4, pp. 309–324, October 2000.

[9] D. A. Landgrebe, *Signal Theory Methods in Multispectral Imaging*, Wiley, Hoboken, NJ, 2003.

[10] J. H. Michels et al., "Robust STAP detection in a dense signal airborne radar environment," to be published.