# The Matched Subspace Detector with Interaction Effects

Ziyu Wang[a,b], Jing-Hao Xue[b,*]

[a] *Department of Security and Crime Science, University College London, London WC1E 6BT, UK*

[b] *Department of Statistical Science, University College London, London WC1E 6BT, UK*

**Abstract**

This paper aims to propose a new hyperspectral target-detection method termed the matched subspace detector with interaction effects (MSDinter). The MSDinter introduces "interaction effects" terms into the popular matched subspace detector (MSD), from regression analysis in multivariate statistics and the bilinear mixing model in hyperspectral unmixing. In this way, the interaction between the target and the surrounding background, which should have but not yet been considered by the MSD, is modelled and estimated, such that superior performance of target detection can be achieved. Besides deriving the MSDinter methodologically, we also demonstrate its superiority empirically using two hyperspectral imaging datasets.

*Keywords:* Matched subspace detector (MSD), linear mixing model (LMM), bilinear mixing model (BMM), interaction effects, target detection, hyperspectral image (HSI)

## 1. Introduction

Hyperpsectral target detection aims to detect small objects from the background of a hyperspectral image (HSI) by the use of known target spectra. The number of target pixels is relatively very small compared with the total number

*Corresponding author. Tel.: +44-20-7679-1863; Fax: +44-20-3108-3105

*Email addresses:* ziyu.wang.12@ucl.ac.uk (Ziyu Wang), jinghao.xue@ucl.ac.uk (Jing-Hao Xue)

*Preprint submitted to Pattern Recognition* *October 19, 2016*

of pixels in an HSI, e.g. only a few target pixels in millions of pixels. Typical applications of the HSI target detection include the detection of specific terrain features, minerals and crops for resource management, the detection of military vehicles and aeroplanes for defence, etc. Comprehensive overviews and gentle tutorials of the HSI target detection can be found in [1, 2, 3, 4].

Target detection algorithms are typically derived from the binary hypothesis model, which consists of two competing hypotheses: the $H_0$ (absence of target) hypothesis and the $H_1$ (presence of target) hypothesis. The likelihood ratio or the generalised likelihood ratio (GLR) of functions of target and background can be used to construct a detector.

Some well-known detectors have been successfully applied to the HSI target detection, including the matched subspace detector (MSD) [5], the orthogonal subspace projection detector (OSP) [6], the spectral matched filter (SMF) [7, 8], the adaptive coherence/cosine detectors (ACEs) [9, 10] and the constrained energy minimization (CEM) [11]. Kwon et al. [12] also extend the MSD, OSP, SMF and ACEs to their corresponding kernel versions based on the kernel-based learning theory. Several methods have been developed based on the CEM specifically [13, 14, 15]. Yang et al. [13] utilise an inequality constraint on the output detector to solve the spectral variability problems, instead of the equal constraint on the CEM. An hierarchical structure of CEM [14] is proposed, which suppresses the backgrounds while preserving the target spectra to boost the performance of CEM. In a very recent work, Yang et al. [15] use total variation to constrain the spatial smoothness and show a promising detection performance when only one single target spectrum is available for training.

Sparse representation (SR)-based algorithms have also been applied to the HSI target detection [16, 17, 18, 19, 20, 21]. Chen et al. [16] propose a sparsity-based target detection (STD), linearly modelling a test pixel by the training background samples and the training target samples. Zhang et al. [17] propose an SR-based binary hypothesis model (SRBBH), which is in the similar fashion of the binary hypothesis model of the MSD. The kernel versions of the STD and SRBBH can be found in [18] and [19], respectively. Detailed reviews of SR algorithms for the HSI classification and detection can be found in [20, 21].

The assumption of these well-known detectors [5, 6, 7, 8, 9, 10, 16, 17] is the linear mixing model (LMM) [22]. The LMM assumes that the spectrum of a mixed pixel can be represented as a linear combination of component spectra (endmembers). The weight (abundance) of each endmember spectrum is proportional to the fraction of the pixel area covered by the endmember. If there are $p$ spectral bands, the $p$-variate spectrum $\mathbf{x} = [x_1, \ldots, x_p]^T$ of a mixed pixel can be expressed as a mixture of $K$ endmembers $\mathbf{m}_k$ with additive noise:

$$\mathbf{x} = \Sigma_{k=1}^{K} a_k \mathbf{m}_k + \mathbf{n} = \mathbf{M}\mathbf{a} + \mathbf{n}, \tag{1}$$

where $\mathbf{M}$ is a $p \times K$ matrix whose columns are the $K$ endmember spectra $\mathbf{m}_k = [m_{k,1}, \ldots, m_{k,p}]^T$ for $k = 1, \ldots, K$, respectively; $\mathbf{a} = [a_1, \ldots, a_K]$ is the fraction abundance vector; and $\mathbf{n} = [n_1, \ldots, n_p]^T$ represents the additive Gaussian white noise. Physical considerations dictate that the abundances have to satisfy 1) the non-negative constraint, i.e. $a_k \geq 0$, and 2) the sum-to-one constraint, i.e. $\Sigma_{k=1}^{K} a_k = 1$. Although the non-negative constraint and the sum-to-one constraint are quite meaningful, they are not always enforced because it significantly complicates the solving of detection problems. As explained in [22] and as usually the case, we can relax both constraints in target detection.

For the HSI target detection, the underlying physical assumption of the LMM is that each incident photon interacts with one earth surface component only and the reflected spectra do not mix before entering the sensor. Therefore, adopting the LMM in [5, 6, 7, 8, 9, 10, 16, 17] assumes that the target spectral signature in the scene remains linearly mixed with the surrounding background spectra after entering the sensor. However this is not true in practice, since the target spectral signatures captured by the hyperspectral sensor can appear significantly different from the true underlying spectrum. The exhibited target spectrum may be contaminated by the *interaction effect* of its true underlying spectrum and its surrounding environments. The reasons can be, but not limited to, that the sensor picks up the signal from multiple scattering of photons and as a result, the abundance vector of targets will be dependent on the characteristics of their surrounding background.

To cope with multiple scattering problems and to model interaction effects, the bilinear mixing model (BMM) has been proposed in the hyperspectral analysis, particularly for the unmixing applications [23, 24, 25, 26, 27, 28]. Nascimento et al. [23] and Fan et al. [24] address the HSI unmixing problem by taking into account of the second-order scattering interaction between endmembers, referred to as "Nascimento model" and "Fan model" hereafter, respectively. The two models are distinguished by different sum-to-one constraints imposed on the abundances. Halimi et al. [25] propose a generalised bilinear model (GBM) to unmix an HSI pixel and solve the problem by a hierarchical Bayesian algorithm. Practical analysis [26, 27, 28] also demonstrate impacts of different orders of interactions in real HSI mixing problems, such as tree cover estimates in orchards. It shows that the second-order interaction has the most significant effect of nonlinear mixing and the higher order interactions can be neglected. On top of the BMM, Heylen et al. [29] derive a multilinear mixing model (MLM) which extends the BMM to an infinite orders of interactions. Experimental studies in [23, 24, 25, 26, 27, 28, 29] have been carried out and shown superior performance of the above-mentioned nonlinear mixing models to conventional linear mixing models.

In this paper, to account for the effect of interaction between the target and their surrounding background on the target spectral signature captured by the sensor, we propose to introduce interaction effects into the models for the HSI target detection. Specifically, we propose a new model, termed the matched subspace detector with interaction effects (MSDinter), by introducing the terms that describe the interaction effects between the target and its surrounding background. To our knowledge, such model is the first one proposed for the HSI target detection. The proposed MSDinter model is able to capture better the target-background mixing effects within pixel spectrum and therefore can improve the performance of target detection.

## 2. The Matched Subspace Detector

The matched subspace detector (MSD) [5] is a popular algorithm which explores the idea of the LMM binary hypothesis model (4). The task is to determine if a test pixel $\mathbf{x}$ contains materials characterised by exemplar target spectral signatures, i.e. whether the test pixel can be represented by a linear combination of target spectral signatures and background spectral signatures. In the MSD, the target spectral signatures and background spectral signatures are represented by the bases of a target subspace and the bases of a background subspace, respectively. The underlying assumption of the MSD in the HSI target detection is that each basis vector of these subspaces represents an endmember, which follows the assumption in the LMM (1).

When a target pixel presents, the spectrum of an observed pixel can be decomposed into two components under the LMM assumption, as

$$\mathbf{x} = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \qquad (2)$$

where $\mathbf{T} = [\mathbf{t}_1, \ldots, \mathbf{t}_{r_t}]$ is a $p \times r_t$ matrix representing the target subspace, and $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_{r_b}]$ is a $p \times r_b$ matrix representing the background subspace; $\mathbf{T}$ is derived from a training target matrix $\mathbf{M}_T \in \mathbb{R}^{p \times N_t}$ whose columns are the $N_t$ target spectra $\mathbf{M}_T(\cdot, n_t)$ for $n_t = 1, \ldots, N_t$, respectively; $\mathbf{B}$ is derived from a training background matrix $\mathbf{M}_B \in \mathbb{R}^{p \times N_b}$ whose columns are the $N_b$ background spectra $\mathbf{M}_B(\cdot, n_b)$ for $n_b = 1, \ldots, N_b$, respectively; $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are the corresponding abundance vectors of the subspace $\mathbf{T}$ and the subspace $\mathbf{B}$, respectively; and $\mathbf{n}$ is the additive Gaussian white noise.

When the target is absent, the spectrum of the observed pixel is adequately described by

$$\mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \qquad (3)$$

which is a reduced order model. Therefore, to decide whether a given target is present or not, we can fit the full model and the reduced model to the test pixel spectrum and check which model provides a better fitting according to certain criterion. Formulated as a binary hypothesis test, the detection problem becomes a decision between the two competing hypotheses $H_0$ and $H_1$,

$$\begin{aligned} H_0 &: \mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \text{ target absent,} \\ H_1 &: \mathbf{x} = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \text{ target present.} \end{aligned} \tag{4}$$

Model (4) is defined as the MSD model. Using the generalised likelihood ratio test (GLRT) [3], the output detector of the MSD model is given by

$$D_{\text{MSD}}(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{P}_B^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{P}_V^{\perp} \mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \nu, \tag{5}$$

where $\mathbf{P}_B^{\perp} = \mathbf{I} - \mathbf{P}_B$ with $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ being the projection matrix onto the column space of $\mathbf{B}$; and $\mathbf{P}_V^{\perp} = \mathbf{I} - \mathbf{P}_V$ with $\mathbf{P}_V = \mathbf{V}(\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T$ being the projection matrix onto the column space of $\mathbf{V}$, where $\mathbf{V}$ is a $p \times (r_t + r_b)$ concatenated matrix of $\mathbf{T}$ and $\mathbf{B}$, i.e. $\mathbf{V} = [\mathbf{T}, \mathbf{B}]$.

The value of $D_{\text{MSD}}(\mathbf{x})$ is compared to a threshold $\nu$ to make a final decision of which hypothesis should be rejected for test pixel $\mathbf{x}$. In general, any set of orthogonal basis vectors that spans the corresponding subspace can be used as the column vectors of $\mathbf{B}$ and $\mathbf{T}$. In this paper, the significant eigenvectors (normalised by the square roots of their corresponding eigenvalues) of the background and target covariance matrices $\mathbf{C}_b$ and $\mathbf{C}_t$ are used to create the column vectors of $\mathbf{B}$ and $\mathbf{T}$, respectively.

## 3. The Matched Subspace Detector with interaction effects (MSDinter)

The linear model (2) in the MSD assumes that the abundance vector $\boldsymbol{\gamma}$ of the target subspace $\mathbf{T}$ in composing a target pixel $\mathbf{x}$ will not change if the characteristics of the background change. Specifically, the effect of one-unit change of $\mathbf{T}$ on $\mathbf{x}$ is the marginal effect of targets $\mathbf{T}$ on $\mathbf{x}$. The marginal effect is obtained by differentiating the conditional expected value of $\mathbf{x}$ with respect to $\mathbf{T}$, i.e.

$$\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \\ \vdots \\ \mathbf{\Gamma}_{r_t} \end{bmatrix}_{(pr_t)\times p}, \qquad (6)$$

where

$$\mathbf{\Gamma}_i = \begin{bmatrix} \gamma_i & 0 & \dots & 0 \\ 0 & \gamma_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma_i \end{bmatrix}_{p\times p} = \gamma_i \mathbf{I}_p,\ i = 1, \dots, r_t, \qquad (7)$$

and $\mathbf{I}_p$ denotes the $p \times p$ identity matrix. The details of the derivation are shown in section 6 of Appendix.

That is, $[\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_{r_t}]^T \in \mathbb{R}^{(pr_t)\times p}$ in (6) is the change of expected value of $\mathbf{x}$ induced by one-unit change of $\mathbf{T}$, which includes only the effect of $\mathbf{T}$ on $\mathbf{x}$, ignoring the effect of $\mathbf{B}$ on $\mathbf{x}$. In other words, no matter whether or not background spectra present in the subpixel $\mathbf{x}$ (i.e. $\boldsymbol{\beta} = \mathbf{0}$ or $\boldsymbol{\beta} \neq \mathbf{0}$), the marginal effect of $\mathbf{T}$ on the test pixel $\mathbf{x}$ does not depend on the values of $\mathbf{B}$.

However, in real applications of the HSI target detection, an observed HSI pixel will also receive multiple scattering of photons between its material and its neighbourhood materials, which the LMM cannot capture. The BMM has been introduced in the hyperspectral unmixing problems to accounts for the presence of multiple photon interactions [23, 24, 25, 26, 27, 28]. However, the interaction effects have not been studied in the hyperspectral target detection. To this end, we hypothesise that there are interaction effects of background spectra and the target spectrum on the composition of the spectrum of an observed target pixel. Therefore we introduce interaction terms into the LMM-based subspace model (2) and propose a new method called matched subspace detector with interaction effects (shortened as MSDinter).

### 3.1. The bilinear mixing model

As aforementioned, LMM (2) cannot deal with multiple scattering that often occurs in the real applications. To this end, the bilinear model (BMM) [23, 24, 25, 26, 27, 28] is proposed to model interaction effects of each pair of endmembers, so as to take account of the multiple scattering phenomena. A typical BMM called "Fan model" [24] is given by

$$\mathbf{x} = \mathbf{M}\mathbf{a} + \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \alpha_{i,j} \mathbf{m}_i \odot \mathbf{m}_j + \mathbf{n}, \tag{8}$$

where $\odot$ denotes the element-wise product operation between two vectors. It is defined as that for two vectors, $\mathbf{m}_i = [m_{i,1}, m_{i,2}, \ldots, m_{i,p}]^T$ and $\mathbf{m}_j = [m_{j,1}, m_{j,2}, \ldots, m_{j,p}]^T$ of the same length, in this case $p \times 1$, the element-wise product is still a vector of the same dimension as the operands with elements given by

$$(\mathbf{m}_i \odot \mathbf{m}_j)_l = m_{i,l} \cdot m_{j,l}, \text{ where } l = 1, \ldots, p. \tag{9}$$

So the element-wise product of two endmembers $\mathbf{m}_i$ and $\mathbf{m}_j$ is

$$\mathbf{m}_i \odot \mathbf{m}_j = \begin{pmatrix} m_{i,1} \\ \vdots \\ m_{i,p} \end{pmatrix} \odot \begin{pmatrix} m_{j,1} \\ \vdots \\ m_{j,p} \end{pmatrix} = \begin{pmatrix} m_{i,1}m_{j,1} \\ \vdots \\ m_{i,p}m_{j,p} \end{pmatrix}. \tag{10}$$

There are various BMMs with different definitions on the sum-to-one constraint to account for the hyperspectral unmixing problems. In the "Fan model" [24], it is assumed that $\sum_{k=1}^{K} a_k = 1$ and $\alpha_{i,j} = a_i a_j$, whereas in the "Nascimento model" [23], the sum-to-one constraint is based on $\sum_{k=1}^{K} a_k + \sum_{i=1}^{K-1} \Sigma_{j=i+1}^{K} \alpha_{i,j} = 1$. In the following proposed method, since we only care about the presence of the interactions terms, it does not matter whether the summation of abundance fractions is 1. Again with the explanations in the HSI target detection [22], we will relax the sum-to-one constraint as well as the non-negative constraint in the following proposed method to simplify the solution to target detection problems.

*3.2. Formulations of MSDinter*

As with the BMM (8), we introduce terms of the interaction between basis vectors of the background subspace $\mathbf{B}$ and the target subspace $\mathbf{T}$ into the MSD model (2), and then revise the alternative hypothesis $H_1$ of the MSD model (4). The proposed model with interaction effects is defined as follows:

$$\mathbf{x} = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\eta} + \mathbf{n}, \tag{11}$$

where $\mathbf{H}$ is a matrix representing the interaction terms between $\mathbf{T}$ and $\mathbf{B}$. We call the matrix $\mathbf{H}$ the interaction matrix, and $\boldsymbol{\eta}$ is the abundance vector for $\mathbf{H}$.

The interaction matrix $\mathbf{H}$ is obtained by the element-wise product of each basis $\mathbf{t}_i$ and $\mathbf{b}_j$, where $i = 1, \ldots, r_t$ and $j = 1, \ldots, r_b$, of the subspace $\mathbf{T}$ and the subspace $\mathbf{B}$, respectively. Similar to the element-wise production $\odot$ defined in (8), the element-wise product of two basis vectors $\mathbf{t}_i = [t_{i,1}, \ldots, t_{i,p}]^T$ and $\mathbf{b}_j = [b_{j,1}, \ldots, b_{j,p}]^T$ is defined as

$$\mathbf{t}_i \odot \mathbf{b}_j = \begin{pmatrix} t_{i,1} \\ \vdots \\ t_{i,p} \end{pmatrix} \odot \begin{pmatrix} b_{j,1} \\ \vdots \\ b_{j,p} \end{pmatrix} = \begin{pmatrix} t_{i,1}b_{j,1} \\ \vdots \\ t_{i,p}b_{j,p} \end{pmatrix}. \tag{12}$$

Hence, the interaction matrix $\mathbf{H}$ is formulated as

$$\mathbf{H} = [\mathbf{t}_1 \odot \mathbf{b}_1, \ldots, \mathbf{t}_1 \odot \mathbf{b}_{r_b}, \mathbf{t}_2 \odot \mathbf{b}_1, \ldots, \mathbf{t}_2 \odot \mathbf{b}_{r_b}, \ldots, \mathbf{t}_{r_t} \odot \mathbf{b}_1, \ldots, \mathbf{t}_{r_t} \odot \mathbf{b}_{r_b}], \tag{13}$$

which is a $p \times (r_t r_b)$ matrix. As a result, the abundance vector corresponding to $\mathbf{H}$ in (13) becomes

$$\boldsymbol{\eta} = [\eta_{1,1}, \ldots, \eta_{1,r_b}, \eta_{2,1}, \ldots, \eta_{2,r_b}, \ldots, \eta_{r_t,1}, \ldots, \eta_{r_t,r_b}]^T, \tag{14}$$

which is a $(r_t r_b) \times 1$ vector.

In model (11), each basis vector in $\mathbf{T}$ and $\mathbf{B}$ is still assumed to represent an endmember. The column vectors in $\mathbf{H}$, on the other hand, are assumed to represent the interactions between the corresponding basis vectors in $\mathbf{T}$ and $\mathbf{B}$, respectively. The interaction matrix $\mathbf{H}$ in fact can be regarded as a generalisation of interaction terms $\mathbf{m}_i \odot \mathbf{m}_j$ defined in model (8).

Our proposed MSDinter is then modelled as follows:

$$\begin{aligned} &H_0 : \mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \text{ target absent}, \\ &H_1 : \mathbf{x} = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\eta} + \mathbf{n}, \text{ target present}. \end{aligned} \tag{15}$$

For a simple representation, let $\mathbf{U}$ be the concatenated matrix of $\mathbf{T}$, $\mathbf{B}$ and $\mathbf{H}$ (13), i.e.

$$\begin{aligned} \mathbf{U} &= [\mathbf{T}, \mathbf{B}, \mathbf{H}] \\ &= [\mathbf{t}_1, \ldots, \mathbf{t}_{r_t}, \mathbf{b}_1, \ldots, \mathbf{b}_{r_b}, \mathbf{t}_1 \odot \mathbf{b}_1, \ldots, \mathbf{t}_{r_t} \odot \mathbf{b}_{r_b}], \end{aligned} \tag{16}$$

which is a $p \times (r_t + r_b + r_t r_b)$ matrix. Then the abundance vectors $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ of model $H_1$ in (15) can be concatenated into a single vector, denoted as $\boldsymbol{\upsilon}$, i.e.

$$\boldsymbol{\upsilon} = [\boldsymbol{\gamma}^T, \boldsymbol{\beta}^T, \boldsymbol{\eta}^T]^T, \tag{17}$$

which is a $(r_t + r_b + r_t r_b)$-dimensional vector. Hence model $H_1$ in the proposed MSDinter (15) can be rewritten as

$$H_1 : \mathbf{x} = \mathbf{U}\boldsymbol{\upsilon} + \mathbf{n}, \text{ target present}, \tag{18}$$

and thus the MSDinter model (15) becomes

$$\begin{aligned} &H_0 : \mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \text{ target absent}, \\ &H_1 : \mathbf{x} = \mathbf{U}\boldsymbol{\upsilon} + \mathbf{n}, \text{ target present}. \end{aligned} \tag{19}$$

To align with the MSD [5], we also adopt the least squares estimate (LSE) to solve the abundance vector $\boldsymbol{\beta}$ in $H_0$ and the abundance vector $\boldsymbol{\upsilon}$ in $H_1$, respectively. Hence it is easily to see that the LSE of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{x} \tag{20}$$

and the LSE of $\boldsymbol{\upsilon}$ is

$$\hat{\boldsymbol{\upsilon}} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{x}, \tag{21}$$

respectively.

Based on (20) and (21), the residual sums of squares (RSS) $e_0$ and $e_1$ given $H_0$ and $H_1$ of MSDinter (19) are computed as

$$H_0 : e_0 = \left\| \mathbf{x} - \mathbf{B}\hat{\boldsymbol{\beta}} \right\|_2^2 = \mathbf{x}^T(\mathbf{I} - \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T)\mathbf{x}, \tag{22}$$

and

$$H_1 : e_1 = \|\mathbf{x} - \mathbf{U}\hat{\boldsymbol{\upsilon}}\|_2^2 = \mathbf{x}^T(\mathbf{I} - \mathbf{U}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T)\mathbf{x}, \tag{23}$$

respectively, where $\mathbf{I}$ is a $p \times p$ identity matrix.

Therefore the generalised test ratio of the MSDinter model is then given by

$$D_{\text{MSDinter}}(\mathbf{x}) = \frac{e_0}{e_1} = \frac{\mathbf{x}^T(\mathbf{I} - \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T)\mathbf{x}}{\mathbf{x}^T(\mathbf{I} - \mathbf{U}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T)\mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \nu. \tag{24}$$

Referring to the final results of MSD (5), we reformulate the output detector of the MSDinter model (24) by utilising the projection matrices. The numerator of (24) is the same as that of the MSD (5), where $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ is the projection matrix onto the subspace $\mathbf{B}$ spanned by the basis vectors $\mathbf{b_1}, \ldots, \mathbf{b}_{r_b}$ and $\mathbf{P}_B^{\perp} = \mathbf{I} - \mathbf{P}_B$ is the orthogonal complement of $\mathbf{P}_B$. The denominator of (24) can be derived in the same way, where

$$\mathbf{P}_U = \mathbf{U}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T \tag{25}$$

is the projection matrix onto the subspace $\mathbf{U}$ spanned by the column vectors in (16) and

$$\mathbf{P}_U^{\perp} = \mathbf{I} - \mathbf{P}_U, \tag{26}$$

is the orthogonal complement of $\mathbf{P_U}$. Hence the final output detector of the MSDinter is formulated as

$$D_{\text{MSDinter}}(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{P}_B^{\perp}\mathbf{x}}{\mathbf{x}^T\mathbf{P}_U^{\perp}\mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \nu. \tag{27}$$

The value of $D_{\text{MSDinter}}(\mathbf{x})$ is compared with the threshold $\nu$ to make a final decision of which hypothesis should be rejected for the test pixel $\mathbf{x}$.

### *3.3. Underlying assumption of adding interaction terms in target detection*

In the proposed MSDinter model (15), we assume that the marginal effect of targets $\mathbf{T}$ on $\mathbf{x}$ varies in different surrounding backgrounds. Specifically, the abundance of target is not only $\boldsymbol{\gamma}$ when an interaction with the background presents. The abundance of the target can be decomposed into the main effect of $\boldsymbol{\gamma}$ plus a contribution from the interactions.

Differentiating the conditional expected value of $\mathbf{x}$ given model (11) with respect to $\mathbf{T}$, we can obtain the following result:

$$\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \\ \vdots \\ \mathbf{\Gamma}_{r_t} \end{bmatrix}_{(pr_t)\times p} + \begin{bmatrix} \mathbf{\Pi}_1 \\ \mathbf{\Pi}_2 \\ \vdots \\ \mathbf{\Pi}_{r_t} \end{bmatrix}_{(pr_t)\times p}, \tag{28}$$

where

$$\mathbf{\Pi}_i = \begin{bmatrix} \mathbf{B}_{1,\cdot}^T \boldsymbol{\eta}_i & 0 & \dots & 0 \\ 0 & \mathbf{B}_{2,\cdot}^T \boldsymbol{\eta}_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{B}_{p,\cdot}^T \boldsymbol{\eta}_i \end{bmatrix}_{p\times p}, i = 1, \dots, r_t, \tag{29}$$

which is a diagonal $p \times p$ matrix; $\boldsymbol{\eta}_i$ is an $r_b \times 1$ vector which is a segment of $\boldsymbol{\eta}$ (14) with

$$\boldsymbol{\eta} = [\boldsymbol{\eta}_1^T, \dots, \boldsymbol{\eta}_i^T, \dots, \boldsymbol{\eta}_{r_t}^T]^T \tag{30}$$

where

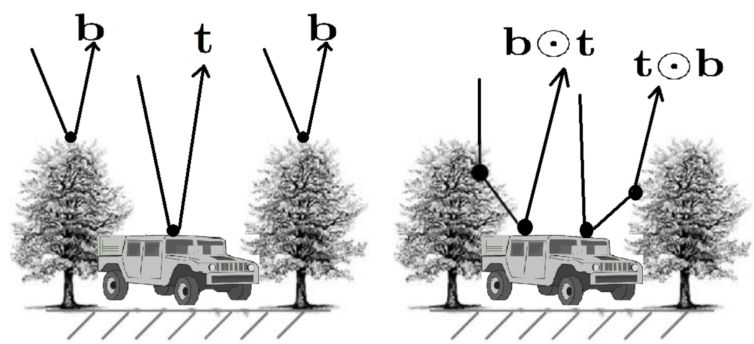
$$\boldsymbol{\eta}_i = [\eta_{i,1}, \dots, \eta_{i,r_b}]^T; \tag{31}$$

and $\mathbf{B}_{l,\cdot}$ denotes a column vector representing the $l$th row of matrix $\mathbf{B}$. The details of the derivation are also presented in section 6 of Appendix.

In (28), when $\boldsymbol{\eta} = \mathbf{0}$, the marginal effect of targets $\mathbf{T}$ on an observed test pixel $\mathbf{x}$ is $[\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_{r_t}]^T \in \mathbb{R}^{(pr_t)\times p}$ only; when $\boldsymbol{\eta} \neq \mathbf{0}$, the marginal effect is $[\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_{r_t}]^T + [\mathbf{\Pi}_1, \dots, \mathbf{\Pi}_{r_t}]^T \in \mathbb{R}^{(pr_t)\times p}$. In other words, the abundance of targets can be variable and dependent on the values of $\mathbf{B}$, when there are interactions between target spectra and background spectra.

The underlying physical assumption of model (11) is that given an observed target pixel, the hyperspectral sensor will not only receive the reflectance of the target and the background independently (modelled by a linear combination of $\mathbf{T}\boldsymbol{\gamma}$ and $\mathbf{B}\boldsymbol{\beta}$), it will also receive the multiple scattering of the target and the background (modelled by additional interaction effects $\mathbf{H}\boldsymbol{\eta}$ between the target and the background).

Similarly to the explanation of the model used for unmixing of HSIs [25], for example, we assume that there are only two components “trees” and “vehicle” presented in an observed target pixel, where the ‘vehicle” is the target to be detected and “trees” are backgrounds. Illustrations of complex photons paths possible to occur are shown in Fig. 1.




(a) (b)

Figure 1: Examples of complex photon paths possible to occur: (a) LMM; (b) interaction effects.

In the assumption of LMM, the hyperspectral sensor will receive signals backscattered by the trees and the vehicle independently, which are represented by the terms $\beta\mathbf{b}$ and $\gamma\mathbf{t}$, respectively as illustrated in Fig. 1(a). However, if a signal is first backscattered by the vehicle to trees (or vice versa), and then backscattered to the sensor, this will result in multiple scattering and the hyperspectral sensor will receive interaction effects between endmembers “trees” and “vehicle”, which we assume to be represented by the interaction term $\eta(\mathbf{t} \odot \mathbf{b})$. This multiple scattering process is illustrated in Fig. 1(b). It is possible that higher order interactions are also received by the hyperspectral sensor. However, as with the analysis of unmixing of HSI [25, 26, 27, 28], these higher order terms can be neglected.

## 4. Experimental studies

We conduct comparative experiments on two publicly available hyperspectral datasets. One is for synthetic target detection analysis and the other is for real target detection analysis:

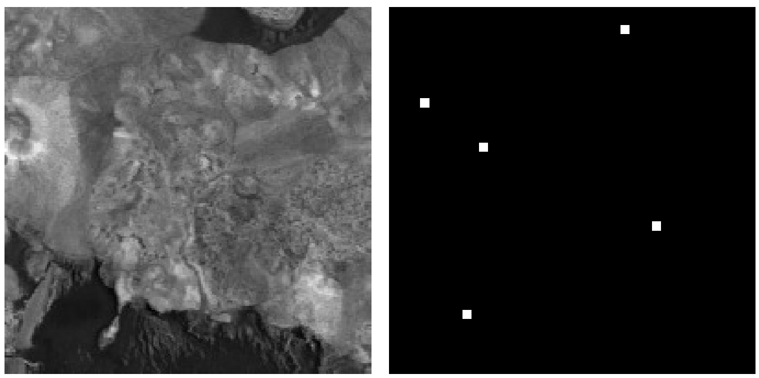
1) *Synthetic targets*: the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) dataset was captured at the Lunar Crater Volcanic Field (LCVF) in northern Nye County, Nevada, USA (http://aviris.jpl.nasa.gov/data/). It has a total of 224 spectral bands covering the spectral range of 400nm-2500nm. The dataset has been widely used for simulated HSI target detection such as in [30, 31]. We use a $200 \times 200$ sub-image in our experiment. There is no defined target in the scene. We manually implant target pixels into the image and simulate the target detection process, to explore the capability of the proposed method.

2) *Real targets*: the Hymap dataset contains ground-truth spectra of targets and has them readily deployed in the scene. It was captured at the location of a small town of Cook City, USA. This image is published by Rochester Institute of Technology (RIT), Rochester, NY, USA [32]. The dataset comes with the locations and pure spectra for all the desired targets. It has a total of 126 spectral bands and is of size $280 \times 800$, covering the spectral range of 453nm-2496nm. Thy Hymap dataset serves as standard target detection dataset and is widely used, such as in [21, 30, 31, 33, 34].

### *4.1. Synthetic targets: the AVIRIS dataset*

In the AVIRIS image, five target pixels are manually implanted using two mixing models that simulate the possible linear/multi-scattering behaviour of hyperspectral sensors. This experiment focuses on exploring the capability of the proposed method in capturing the interaction effects between the target spectrum and the background spectra.

The AVIRIS image is shown in Fig. 2(a). The locations of the five implanted pixels are depicted in Fig. 2(b). The implanted target is a species of mineral called almandine, which is not from the AVIRIS dataset. As with [31], the spectrum of the target almandine is rescaled and resampled to match the AVIRIS image wavelength. The target spectrum and five background spectra originally at implanted locations are show in Fig. 3. In this simulation, we only conduct comparative experiments on MSD and MSDinter, to explore the potential of MSDinter.



(a) (b)

Figure 2: (a) The AVIRIS sub-image (200 × 200) of the third spectral band. (b) Locations of the implanted targets.

#### *4.1.1. Experimental settings*

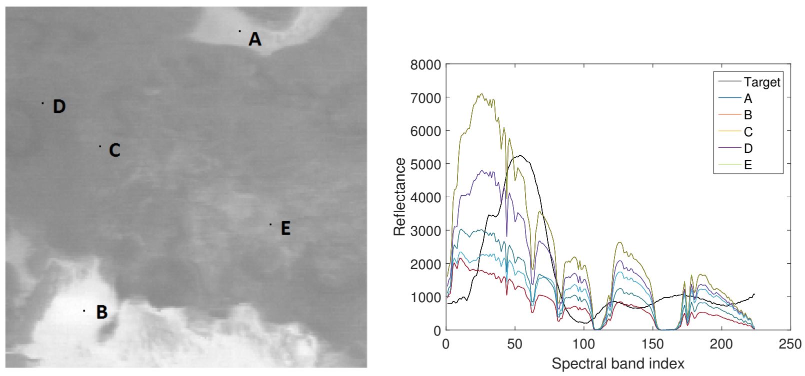
The implanted target pixel $\mathbf{x}$ is mixed with the prior target spectrum $\mathbf{t}$ and the original background spectrum $\mathbf{b}$ at each implanted location shown in Fig. 2(b). Two mixing models are used:

- Linear mixing model (LMM):

$$\mathbf{x} = f_t\mathbf{t} + f_b\mathbf{b}, \tag{32}$$

- Bilinear mixing model (BMM):

$$\mathbf{x} = f_t\mathbf{t} + f_b\mathbf{b} + (1 - f_t - f_b)\mathbf{t} \odot \mathbf{b}, \tag{33}$$




(a) (b)

Figure 3: (a) The locations of the representative background spectral samples. (b) The pure target spectrum and the representative background spectra located in (a).

where $f_t$ and $f_b$ are implanted fractions of the target spectrum and of the background spectrum, respectively. The fractions of all terms are sum to 1 in LMM (32) and BMM (33), respectively. We simulate four datasets for LMM and BMM, respectively, and details of the implanted fractions are shown in Table 1.

Table 1: Details of the implanted fractions for the AVIRIS dataset.

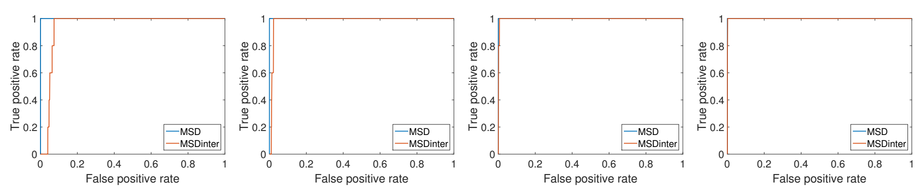
| | LMM | | BMM | | |
|---|---|---|---|---|---|
| Fraction | $f_t$ | $f_b$ | $f_t$ | $f_b$ | $1 - f_t - f_b$ |
| Simulation 1 | 5% | 95% | 1% | 5% | 94% |
| Simulation 2 | 7% | 93% | 1% | 7% | 92% |
| Simulation 3 | 9% | 91% | 1% | 9% | 90% |
| Simulation 4 | 10% | 90% | 1% | 10% | 89% |

As the spectra of the mixed target pixels may appear very different from the spectra in the original image, the detection may become trivial and the performances of both detectors are not distinguishable. Therefore we randomly add white noise with mean $\mathbf{0}$ to the whole image after implanting the target pixels, which mimics the distortion caused by the sensors in real applications. In this experiment, the added white noise is measured in terms of the Signal-to-Noise Ratio (SNR). The SNR in decibels is defined as

$$\mathrm{SNR}_{dB} = 10\log_{10}\left(\frac{\sigma_i^2}{\sigma_{noise}^2}\right), \tag{34}$$

where $\sigma_i$ is the standard deviation of the $i$th band image for $i = 1, \ldots, 224$ and $\sigma_{noise}$ is the standard deviation of the noise added to each band image. We set $\mathrm{SNR}_{dB} = 20\mathrm{dB}$ and therefore add white noise with $\sigma_{noise}^2 = \sigma_i^2/100$ in each band image in the following simulations.

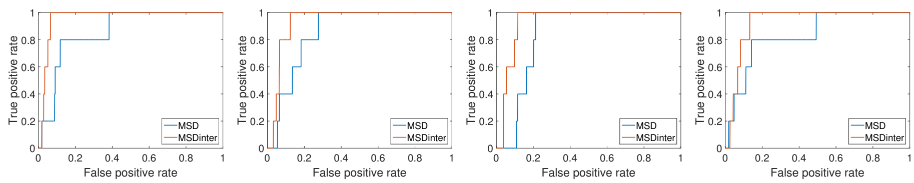
We use the single target spectrum and five background spectra shown in Fig. 3 as the target subspace $\mathbf{T}$ and the background subspace $\mathbf{B}$, respectively. The receiver operating characteristic (ROC) curve is adopted to measure the detection performances. The ROC is a threshold-free measurement. For each detector result, the threshold varies in a range to obtain a set of pairs of the true positive rate and the false positive rate, which is then used to plot the ROC curve. We also employ the area under curve (AUC) statistics to measure the detection performance quantitatively, in pair with the ROC curve.




(a) (b) (c) (d)

Figure 4: ROC curves of detecting implanted target pixels mixed by LMM: (a) $f_t = 5\%$, $f_b = 95\%$; (b) $f_t = 7\%$, $f_b = 93\%$; (c) $f_t = 9\%$, $f_b = 91\%$; (d) $f_t = 10\%$, $f_b = 90\%$.

The ROC curves of detecting the LMM-based implanted targets pixels and the BMM-based implanted targets pixels by MSD and MSDinter are shown in Fig. 4 and Fig. 5, respectively. The AUC performances corresponding to Fig. 4 and Fig. 5 are listed in Table 2.




(a) (b) (c) (d)

Figure 5: ROC curves of detecting implanted target pixels mixed by BMM: (a) $f_t = 1\%$, $f_b = 5\%$, $1 - f_t - f_b = 94\%$ ; (b) $f_t = 1\%$, $f_b = 7\%$, $1 - f_t - f_b = 92\%$ ; (c) $f_t = 1\%$, $f_b = 9\%$, $1 - f_t - f_b = 90\%$; (d) $f_t = 1\%$, $f_b = 10\%$, $1 - f_t - f_b = 89\%$.

Table 2: AUC statistics of MSD and MSDinter for the AVIRIS dataset.

| | LMM | | BMM | |
|---|---|---|---|---|
| AUC | MSD | MSDinter | MSD | MSDinter |
| Simulation 1 | 1 | 0.945 | 0.860 | 0.961 |
| Simulation 2 | 1 | 0.984 | 0.857 | 0.933 |
| Simulation 3 | 1 | 0.998 | 0.839 | 0.931 |
| Simulation 4 | 1 | 1 | 0.837 | 0.930 |

#### 4.1.2. Results on LMM-mixed targets

From the results listed in Table 2 and shown in Fig. 4, where implanted target pixels are synthesised by LMM, we can observe at least two patterns. Firstly, MSD achieves perfect performance for LMM-mixed targets, i.e. AUC = 1 on detecting all implanted targets with enumerated fractions. That is, it implies that if target pixels captured by the HSI sensor are mixed by the linear combination of the target spectrum and the background spectrum, MSD can perform perfectly. Secondly, as the implanted target fraction $f_t$ increases, e.g. slightly increasing from 5% to 10%, the detection performance of MSDinter improves from 0.945 to 1. It implies that MSDinter can also achieve nearly perfect to perfect performance even when targets are linearly mixed without any interaction effect.

#### 4.1.3. Results on BMM-mixed targets

In this simulation, the implanted target fraction $f_t$ is fixed to be 1%, and the implanted background fraction is ranged from 5% to 10%. The rest of fractions are occupied by the interaction terms $\mathbf{t} \odot \mathbf{b}$. The performances of MSD and MSDinter on detecting the BMM-based implanted targets are listed in Table 2 and shown in Fig. 5. We can observe that MSDinter outperforms MSD on detecting all BMM-based implanted targets with enumerated fractions. It reveals that if the interaction between the background spectrum and the target spectrum does exist, MSDinter can achieve better performance than that of MSD, as the latter fails to take the interaction effects into consideration.

#### 4.1.4. Detection statistics of MSD and MSDinter

We further compare the test statistics of all pixels in the AIVRIS image processed by MSD and MSDinter. The test statistics of 40,000 pixels in the LMM-based simulation and BMM-based simulation are shown in Fig. 6 and Fig. 7, respectively. Due to the nature of MSD and MSDinter, the test statistics are always greater than 1 and the pixels with higher statistics are considered more likely to be targets.

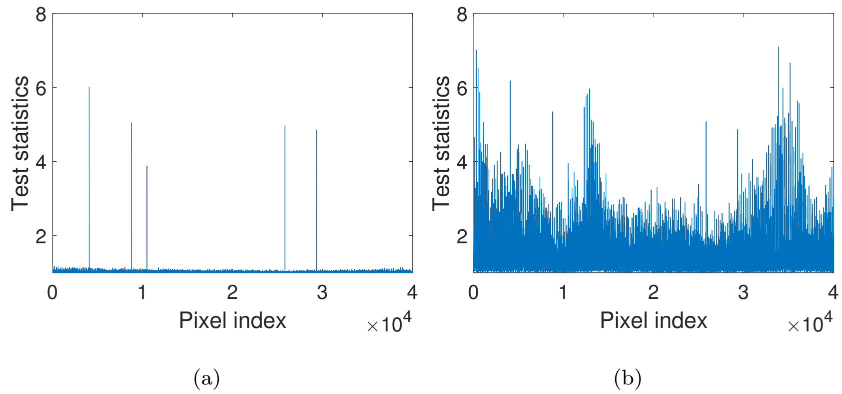



Figure 6: Test statistics of the AVIRIS image implanted by LMM with mixing fractions $f_t = 9\%$, $f_b = 91\%$: (a) MSD, AUC = 1; (b) MSDinter, AUC = 0.998.

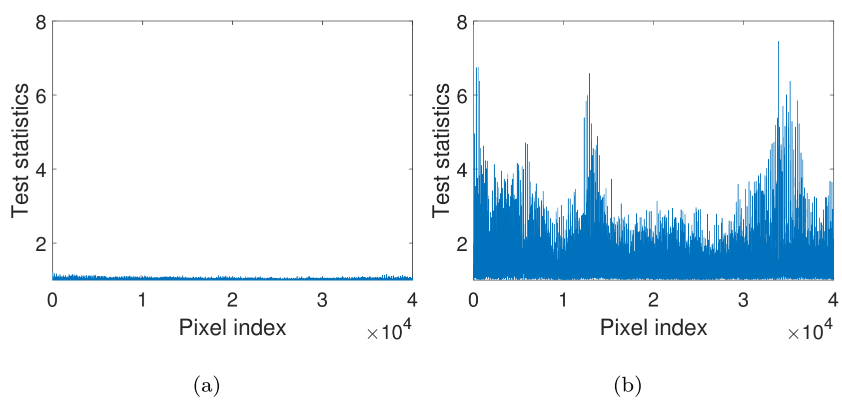



Figure 7: Test statistics of the AVIRIS image implanted by BMM with mixing fractions $f_t = 1\%$, $f_b = 9\%$ $1 - f_t - f_b = 90\%$: (a) MSD, AUC = 0.839; (b) MSDinter, AUC = 0.931.

In Fig. 6(a), we can observe that MSD has very distinguishable test statistics of the implanted targets which are linearly mixed without interaction. However in Fig. 7(a), the test statistics of MSD on targets not only largely decrease but also become undistinguishable when the implanted targets are bilinearly mixed with interaction, and the performance of MSD drops significantly, from AUC = 1 (6(a)) to AUC = 0.839 (7(a)). On the other hand, the test statistics of MSDinter are more stable than those of MSD, whether or not the implanted pixels are mixed by LMM or BMM, which are depicted in Fig. 6(b) and 7(b). It indicates that MSDinter can handle both simple and complex mixing effects, with much more stable performance than MSD.

### *4.2. Real targets: the Hymap dataset*

For the real hyperspectral dataset, i.e. the Hymap dataset where targets are deployed in the scene, the proposed MSDinter method is evaluated against not only MSD but some other well-known detectors, such as ACE [10], CEM [11] and OSP [6]. We also compare the MSDinter method with an SR-based method termed STD [16].

The Hymap image is shown in Fig. 8. As the desired targets are mainly located in the central part of the whole image and the materials lie around the margin of the image are homogeneous which are mainly composed of trees, we cropped a $100 \times 300$ sub-image from the central part of the original Hymap image for evaluating the performances of detectors. Such a sub-image setting has been widely used and well accepted by researchers, such as in [21, 35, 36]. Different experimental settings for analysing the Hymap image can also be found in [13, 15, 30, 31, 33, 34] for different illustrative purposes.

There are seven types of targets in the Hymap dataset, including four types of fabric panels (F1, F2, F3, F4) and three types of vehicles (V1, V2, V3). There are two samples with different sizes deployed in the scene for F3 and F4, termed F3a and F3b, F4a and F4b, respectively. The rest of targets, i.e. F1, F2, V1, V2 and V3, have only one sample each. When one type of target is to be detected, e.g. F3a and F3b, the other targets, i.e. F1, F2, F4a, F4b, V1, V2 and V3, are regarded as background pixels. The seven types of targets and their central coordinates of region of interests (ROIs) are shown in Table 3. Since the spatial resolution of the Hymap dataset is about 3m, we can infer that F1 (3m × 3m), F2 (3m × 3m) are nearly full pixels, whereas all the other targets are smaller than a pixel and appear as subpixels. Therefore a mixture model should be considered for all the targets, and the interaction effects between the target and the background are likely to occur. The cropped sub-image as well as ROIs of seven types of targets are shown in Figure 8 and Figure 9, respectively.

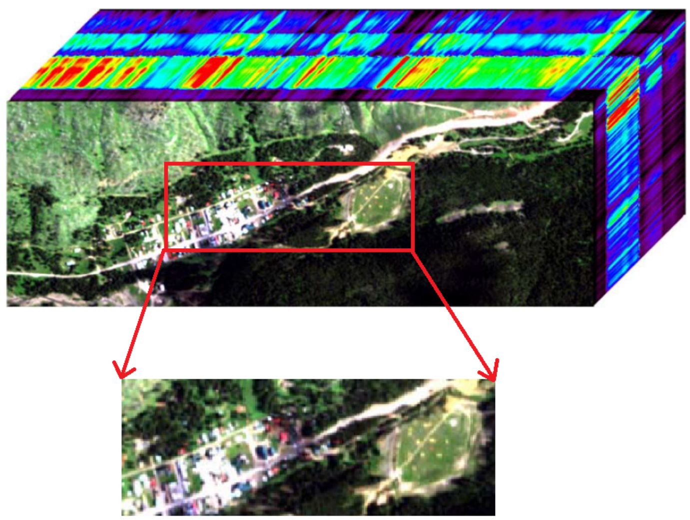
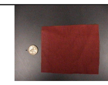
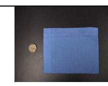
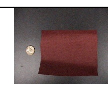
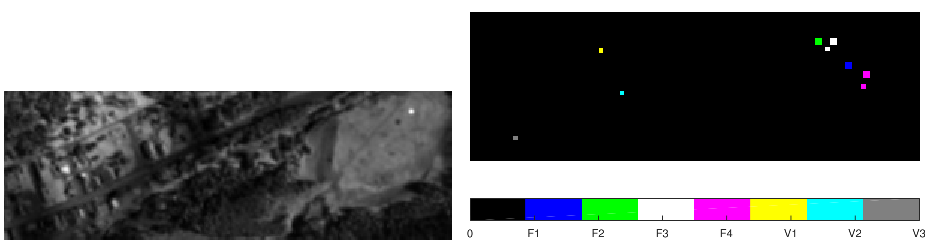


Figure 8: The Hymap image with a spatial size of 280 × 800 [32]. We cropped a spatial size of 100 × 300 sub-image for evaluation in this experiment.

The spectrum of each desired target (F1-F4 and V1-V3) is provided by projected-equipped SPL files [32]. As with [31], we rescale and resample the SPL spectra according to the Hymap HSI wavelength. Preprocessed target spectra are given in Fig. 10. We randomly select one sample spectral signature of each target in the scene, and plot them in Fig. 11. Comparing Fig. 10 with Fig. 11, we can clearly see that target spectra signatures in the scene are very different from those ground-truth spectra in Fig. 10, and the pattern of how the sampled target spectra are mixed with the background spectra is complicated.

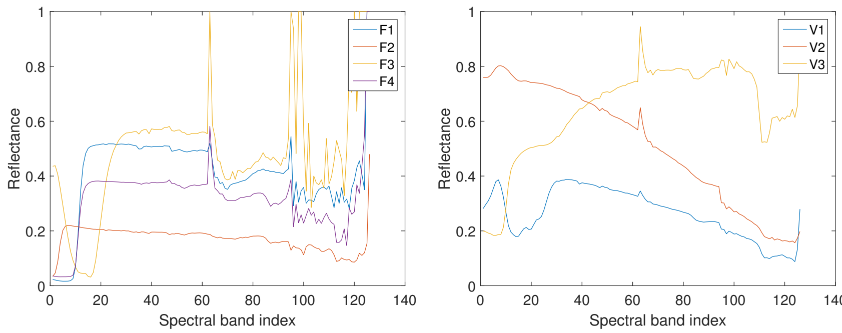
Table 3: List of the targets in the Hymap dataset

| Target | Description and pixel size of ROI | Central coordinates of ROI | Photo |
|---|---|---|---|
| F1 | Red cotton (3m × 3m)<br>( 5 × 5 pixels) | (138, 504) |  |
| F2 | Yellow nylon (3m × 3m)<br>( 5 × 5 pixels) | (122, 484) |  |
| F3 a&b | Blue cotton<br>(2m × 2m & 1m × 1m)<br>( 5 × 5 pixels &3 × 3 pixels ) | (122, 494)<br>& (127, 490) |  |
| F4 a&b | Red Nylon<br>(2m × 2m & 1m × 1m)<br>( 5 × 5 pixels &3 × 3 pixels) | (144, 516)<br>& (152, 514) |  |
| V1 | Green Chevy Blazer<br>( 3 × 3 pixels) | (128, 339) |  |
| V2 | White Toyota T100<br>( 3 × 3 pixels) | (156, 353) |  |
| V3 | Red Subaru GL<br>( 3 × 3 pixels) | (186, 282) |  |



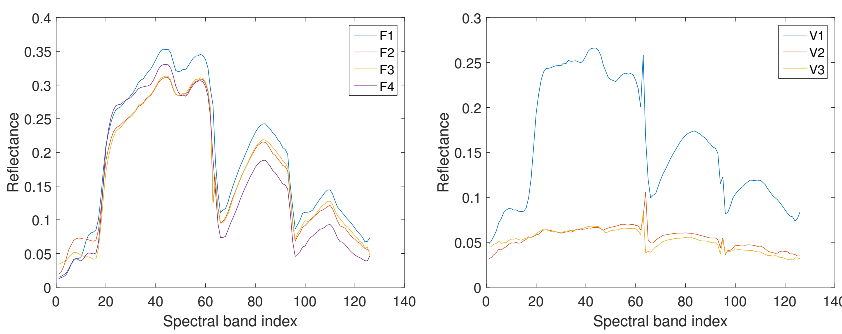

(a) (b)

Figure 9: (a) The Hymap sub-image (100 × 300) of the 33th spectral band; (b) ROIs of seven types of targets (F1, F2, F3, F4, V1, V2 and V3) in the Hymap sub-image. There are two samples of targets F3 and F4 each, termed F3a and F3b, and F4a and F4b, respectively. The pixel sizes of the ROI of targets F1, F2, F3a, F3b, F4a, F4b, V1, V2 and V3 are 25, 25, 25, 9, 25, 9, 9, 9 and 9, respectively. Different types of targets are shown in different colours.




(a) (b)

Figure 10: Rescaled prior spectra of all the targets in the SPL files: (a) fabric panels; (b) vehicles.




(a) (b)

Figure 11: Rescaled sample spectra of all targets in the Hymap scene: (a) fabric panels; (b) vehicles. The selected sample spactra are located in the central coordinates of the ROIs of F1, F2, F3a, F4a, V1, V2 and V3, respectively, which are shown in Table 3.

*4.2.1. Experimental settings*

In realistic target detection problems, the background statistics are usually unknown. As explained in [37], the statistics of background can be estimated by all pixels within the area of interest when detectors are applied in a sparse target environment. In our experiment, there are 30,000 pixels in the cropped Hymap sub-image and among which there is only 1 target pixel to be detected for each desired target. The number of target/image ratio is 1/30000, which means our detection environment is sufficiently sparse. Therefore we can use all pixels of the cropped Hymap image to estimate the mean $\boldsymbol{\mu}_b$ and the covariance $\mathbf{C}_b$ of the background. In this way, the detector of each test pixel has global and identical background statistics (mean $\boldsymbol{\mu}_b$ and covariance $\mathbf{C}_b$). In addition, detectors used in this paper, including MSD, MSDinter, ACE, CEM, OSP, all adopt the same aforementioned background samples for fair comparison. For the SR-based method STD, the background dictionary for each test pixel is constructed by 29,999 pixels of the cropped image excluding the test pixel itself.

Among the compared detectors, MSD, MSDinter and OSP involve the construction of background subspace $\mathbf{B}$. We use the mean-centred HSI (removing the estimated mean $\boldsymbol{\mu}_b$ from the HSI) to compute the covariance matrix $\mathbf{C}_b$ and then preserve significant eigenvectors of $\mathbf{C}_b$ to create columns of $\mathbf{B}$. For MSD and MSDinter, we should also construct target subspace $\mathbf{T}$. Since there is only one prior spectrum of each desired target $\mathbf{m}_t$, we actually do not need to do eigen-decomposition on $\mathbf{m}_t$ to obtained the target subspace $\mathbf{T}$. Instead, we subtract the background mean $\boldsymbol{\mu}_b$ from the prior target spectrum $\mathbf{m}_t$, i.e. $\mathbf{m}_t - \boldsymbol{\mu}_b$, and then normalise $\mathbf{m}_t - \boldsymbol{\mu}_b$ to have a unit $L_2$-norm as the target subspace $\mathbf{T}$. As a result, the estimated background endmembers $\mathbf{b}$ and the target endmember $\mathbf{t}$ all have unit $L_2$-norm and are independent of each other. For STD, the union dictionary is constructed by the concatenation of 29,999 pixels and the single prior spectrum of each desired target for each test pixel. Again, each column of the dictionary is normalised to have unit $L_2$-norm. In this paper, the STD method is solved by a greedy algorithm called orthogonal matching pursuit (OMP) [38].

We should note that each target deployed in the scene has an ROI [32], which means that the target may appear in any coordinates within the ROI. For example, F1 has a $5 \times 5$ pixels size of ROI and the central coordinates of ROI are (138, 504). It implies that if we detect at least one pixel as a target in the ROI, then this detection is regarded as a 100% correct detection. As with [31] and [36], we use the false alarm rate (FAR) to measure the detection performances of the compared methods. The FAR in this experiment is defined as the number of pixels not in the target ROI but have test statistic values equal to or greater than that of the pixel with the highest statistic value within the target ROI, normalised by the total number of pixels in the Hymap HSI (i.e. 30,000 pixels).

Among the methods to be compared, MSD, MSDinter, OSP and STD have parameters to tune. For MSD, MSDinter and OSP, the parameter is $r_b$, which is the number of eigenvectors to be preserved for the background subspace $\mathbf{B}$. For STD, the parameter is the sparse level, termed $L$, which is the number of HSI pixels to be selected for the sparse representation. As ACE and CEM only use the target endmembers and the whole HSI to construct detectors, no tuning parameters are involved. Due to the limited number of target samples in the dataset, it is infeasible to tune parameters via cross validation. Hence as with most published works of HSI target detections conducted on the Hymap dataset such as [31, 33, 34], the parameter of each detector is manually tuned to show the optimal performance of the algorithms for illustrative purposes. The number of preserved eigenvectors $r_b$ of the background subspace $\mathbf{B}$ for MSD, MSDinter and OSP and the sparse level $L$ of representation for STD are listed in Table 4, respectively.

#### *4.2.2. Experimental results*

The detection performances of all detectors are list in Table 5. We can observe that the proposed MSDinter outperforms MSD, ACE, CEM, OSP and STD in detecting all seven types of targets. Specifically, MSDinter can achieve

Table 4: The parameter $r_b$ of OSP, MSD and MSDinter and the parameter $L$ of STD.

| Target | $r_b$ | | | $L$ |
|---|---|---|---|---|
| | OMP | MSD | MSDinter | STD |
| F1 | 9 | 110 | 5 | 10 |
| F2 | 118 | 111 | 5 | 12 |
| F3 | 58 | 11 | 5 | 12 |
| F4 | 118 | 88 | 6 | 10 |
| V1 | 91 | 91 | 6 | 10 |
| V2 | 43 | 43 | 2 | 4 |
| V3 | 105 | 106 | 10 | 12 |

the best detection performance on detecting F1, F2, F3 with FAR equal to 0. Compared with MSD, MSDinter significantly improves FARs for all targets. It implies that these observed target pixels captured by the HSI sensor are more likely to contain the interaction of background spectra and target spectra. In this sense, as MSDinter models the interaction effects, it achieves better performance than MSD, which fails to model the interaction effects.

Table 5: FAR under 100% detection of ACE, CEM, OSP, MSD, STD and MSDinter for the Hymap dataset. Boldface indicates the best performance.

| FAR | ACE | CEM | OSP | MSD | STD | MSDinter |
|---|---|---|---|---|---|---|
| F1 | 1.02e-02 | 1.19e-02 | 0.01e-02 | 0.76e-02 | 0.06e-02 | **0.00e-02** |
| F2 | 8.55e-02 | 1.11e-02 | 0.01e-02 | 0.14e-02 | 0.53e-02 | **0.00e-02** |
| F3 | 0.57e-02 | 1.35e-02 | 0.27e-02 | 0.0057e-02 | 0.08e-02 | **0.00e-02** |
| F4 | 0.21e-02 | 0.51e-02 | 0.08e-02 | 0.0037e-02 | 0.31e-02 | **0.0027e-02** |
| V1 | 1.37e-02 | 1.41e-02 | 0.86e-02 | 0.62e-02 | 24.76e-02 | **0.0013e-02** |
| V2 | 1.34e-02 | 2.22e-02 | 0.85e-02 | 0.40e-02 | 0.52e-02 | **0.31e-02** |
| V3 | 19.94e-02 | 24.87e-02 | 1.82e-02 | 1.48e-02 | 11.36e-02 | **0.54e-02** |

For illustration purposes, we select one of the seven types of targets, i.e. F1, and plot prediction maps resulted from all compared methods. The prediction maps are shown in Fig. 12, in which the test statistic of each HSI pixel is colour coded. We can observe that the proposed MSDinter produces the most distinguishable detection results, as shown in Fig. 12(c). In the MSDinter prediction map (Fig. 12(c)), the test statistics of pixels within the ROI of F1 have the highest values compared with the statistics of all the other pixels, which result in the best detection performance with FAR equal to 0. On the other hand, the prediction maps of MSD, ACE, CEM, OSP and STD are not easy to distinguish F1 and the background, and their detection performances are not as good as that of MSDinter. In addition, comparing the prediction maps of MSD and MSDInter shown in Fig. 12(b) and Fig. 12(c), we can see that MSDinter eliminates the high statistics of background pixels and thus reduces FAR, which indicates that taking the target-background interaction effects into account can significantly improve the performance of the HSI target detection.

## 5. Conclusion

In this paper we have proposed a new method called MSDinter for the hyperspectral target detection. The MSDinter method introduces interaction terms into the popular MSD to model and capture the interaction between target and background spectra. Compared with MSD, the proposed MSDinter method produces superior detection performance on the synthetic dataset of AVIRIS and the real dataset of Hymap, demonstrating the benefit of taking target-background interaction into modelling for target detection.

It is worthwhile to mention that, besides the platform of MSD, the proposed concept of *interaction effects* can also be applied to other target detection methods which have not yet considered target-background interaction. It is of our research interests to further work in this direction to investigate its potential of improving other established algorithms of target detection from hyperspectral images.

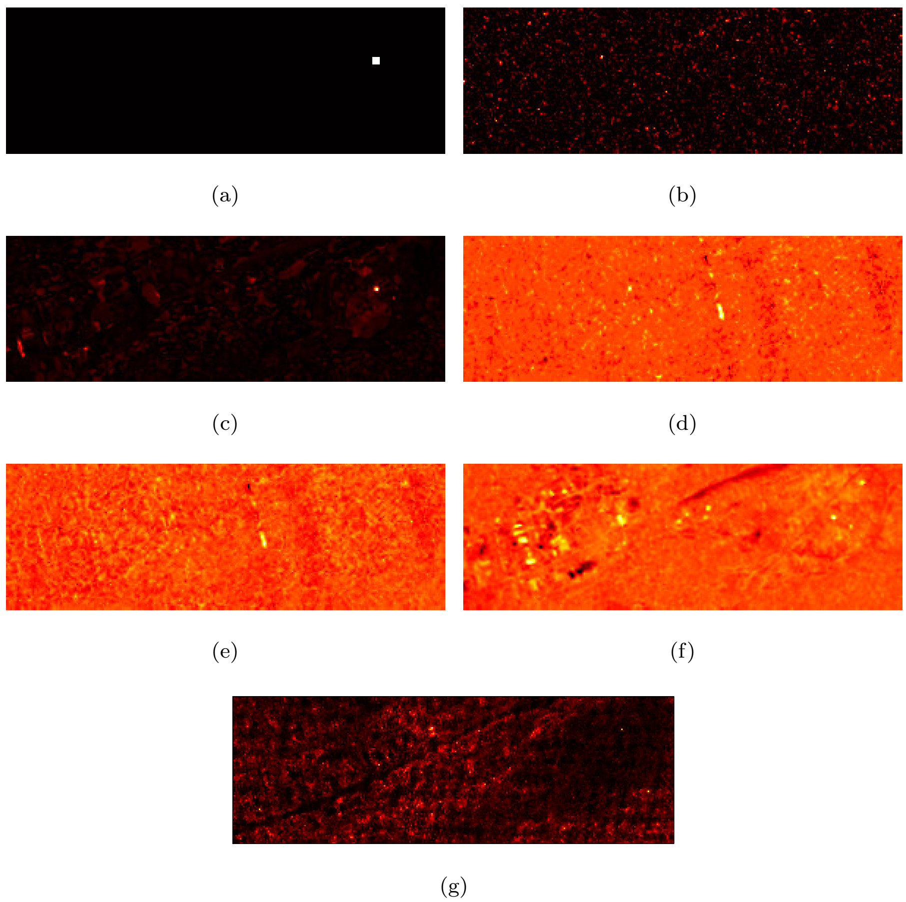


Figure 12: Test statistics for detecting F1 in the Hymap image. Brighter pixels have higher test statistics and therefore are more likely to be targets. (a) Ground-truth labels of F1; (b) MSD, FAR = 0.76e-02; (c) MSDinter, FAR = 0.00e-02; (d) ACE, FAR = 1.02e-02; (e) CEM, FAR = 1.19e-02; (f) OSP, FAR = 0.01e-02; (g) STD, FAR = 0.06e-02.

## 6. Appendix

This section describes in detail how to differentiate the conditional expected value of $\mathbf{x}$ with respect to $\mathbf{T}$, i.e. $\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}}$, for model (2) and model (11), respectively.

To start with, assume that matrix $\mathbf{T}$ contains only one vector $\mathbf{t}$. Then the model (2) of $\mathbf{x}$ is simplified as

$$\mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{t}\gamma + \mathbf{n}, \tag{35}$$

where $\gamma$ is a scalar. It follows that the derivative $\frac{\partial E[\mathbf{x}|\mathbf{t},\mathbf{B}]}{\partial \mathbf{t}}$ effectively measures the impact on the expected value of $\mathbf{x}$ from one-unit change of each element in $\mathbf{t}$. According to the definition of the Jacobian matrix, the resultant derivative of $\frac{\partial E[\mathbf{x}|\mathbf{t},\mathbf{B}]}{\partial \mathbf{t}}$ will be a $p \times p$ matrix, given a $p \times 1$ vector $\mathbf{x}$ and a $p \times 1$ vector $\mathbf{t}$. That is:

$$\frac{\partial E[\mathbf{x}|\mathbf{t},\mathbf{B}]}{\partial \mathbf{t}} = \begin{bmatrix} \gamma & 0 & \dots & 0 \\ 0 & \gamma & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma \end{bmatrix}_{p \times p} = \gamma \mathbf{I}_p, \tag{36}$$

which turns out to be a diagonal $p \times p$ matrix $\gamma \mathbf{I}_p$, where $\mathbf{I}_p$ denotes the $p \times p$ identity matrix.

When matrix $\mathbf{T}$ contains multiple vectors $\mathbf{t}_i$ for $i = 1, \dots, r_t$, which is the case of model (2), the derivative of $\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}}$ measures the impact on the expected value of $\mathbf{x}$ from one-unit change of each element in $\mathbf{T}$. Let us rewrite model (2) as

$$\mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{T}\boldsymbol{\gamma} + \mathbf{n} = \mathbf{B}\boldsymbol{\beta} + [\mathbf{t}_1, \dots, \mathbf{t}_{r_t}]\boldsymbol{\gamma} + \mathbf{n}, \tag{37}$$

where $\boldsymbol{\gamma}$ is an $r_t$-variate vector. Then the resultant derivative $\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}}$ will be a $(pr_t) \times p$ matrix, with $\mathbf{x}$ being a $p \times 1$ vector and $\mathbf{T}$ being a $p \times r_t$ matrix.

Based on the results in (36) and letting $\boldsymbol{\Gamma}_i$ denote the $p \times p$ diagonal matrix with $\gamma_i$ on the diagonal, i.e.

$$\boldsymbol{\Gamma}_i = \begin{bmatrix} \gamma_i & 0 & \dots & 0 \\ 0 & \gamma_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma_i \end{bmatrix}_{p \times p} = \gamma_i \mathbf{I}_p, \tag{38}$$

it follows that the derivative in the case of model (2) is

$$\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}} = \begin{bmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}_2 \\ \vdots \\ \boldsymbol{\Gamma}_{r_t} \end{bmatrix}_{(pr_t) \times p}, \tag{39}$$

which is a concatenated matrix.

For model (11), the addition of interaction term $\mathbf{H}\boldsymbol{\eta}$ introduces complexity to the computation, but due to the nature of linear algebra, the derivative can be found in a similar fashion. With the added interaction term, the model (11) of $\mathbf{x}$,

$$\mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{T}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\eta} + \mathbf{n}, \tag{40}$$

has the derivative as

$$\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}} = \begin{bmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}_2 \\ \vdots \\ \boldsymbol{\Gamma}_{r_t} \end{bmatrix}_{(pr_t) \times p} + \frac{\partial \mathbf{H}\boldsymbol{\eta}}{\partial \mathbf{T}}. \tag{41}$$

For the derivation $\frac{\partial \mathbf{H}\boldsymbol{\eta}}{\partial \mathbf{T}}$, we can follow the same steps by which we get results (36) and (39). Firstly, recall that the interaction matrix $\mathbf{H}$ has been expanded in (13):

$$\mathbf{H} = [\mathbf{t}_1 \odot \mathbf{b}_1, \dots, \mathbf{t}_1 \odot \mathbf{b}_{r_b}, \mathbf{t}_2 \odot \mathbf{b}_1, \dots, \mathbf{t}_2 \odot \mathbf{b}_{r_b}, \dots, \mathbf{t}_{r_t} \odot \mathbf{b}_1, \dots, \mathbf{t}_{r_t} \odot \mathbf{b}_{r_b}].$$

Thus $\frac{\partial \mathbf{H}\boldsymbol{\eta}}{\partial \mathbf{t}_i}$, where $i = 1, \dots, r_t$, can be written as

$$\frac{\partial \mathbf{H}\boldsymbol{\eta}}{\partial \mathbf{t}_i} = \begin{bmatrix} \sum_{j=1}^{r_b} b_{j,1}\eta_{i,j} & 0 & \dots & 0 \\ 0 & \sum_{j=1}^{r_b} b_{j,2}\eta_{i,j} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{j=1}^{r_b} b_{j,p}\eta_{i,j} \end{bmatrix}_{p\times p}$$

$$= \begin{bmatrix} \sum_{j=1}^{r_b} \mathbf{B}_{1,j}\eta_{i,j} & 0 & \dots & 0 \\ 0 & \sum_{j=1}^{r_b} \mathbf{B}_{2,j}\eta_{i,j} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{j=1}^{r_b} \mathbf{B}_{p,j}\eta_{i,j} \end{bmatrix}_{p\times p} \tag{42}$$

$$= \begin{bmatrix} \mathbf{B}_{1,\cdot}^T\boldsymbol{\eta}_i & 0 & \dots & 0 \\ 0 & \mathbf{B}_{2,\cdot}^T\boldsymbol{\eta}_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{B}_{p,\cdot}^T\boldsymbol{\eta}_i \end{bmatrix}_{p\times p},$$

which is a diagonal $p \times p$ matrix, where $\boldsymbol{\eta}_i$ is a segment of $\boldsymbol{\eta}$ with

$$\boldsymbol{\eta} = [\eta_{1,1}, \dots, \eta_{i,j}, \dots, \eta_{r_t,r_b}]^T = [\boldsymbol{\eta}_1^T, \dots, \boldsymbol{\eta}_i^T, \dots, \boldsymbol{\eta}_{r_t}^T]^T, \tag{43}$$

and $\mathbf{B}_{l,\cdot}$ denotes a column vector representing the $l$th row of matrix $\mathbf{B}$.

Let $\boldsymbol{\Pi}_i$ denote the resultant derivative with respect to $\mathbf{t}_i$ in (42):

$$\boldsymbol{\Pi}_i = \begin{bmatrix} \mathbf{B}_{1,\cdot}^T\boldsymbol{\eta}_i & 0 & \dots & 0 \\ 0 & \mathbf{B}_{2,\cdot}^T\boldsymbol{\eta}_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{B}_{p,\cdot}^T\boldsymbol{\eta}_i \end{bmatrix}_{p\times p}. \tag{44}$$

The derivative of $\frac{\partial \mathbf{H}\boldsymbol{\eta}}{\partial \mathbf{T}}$ is then the concatenation of $\boldsymbol{\Pi}_i$:

$$\frac{\partial \mathbf{H}\boldsymbol{\eta}}{\partial \mathbf{T}} = \begin{bmatrix} \boldsymbol{\Pi}_1 \\ \boldsymbol{\Pi}_2 \\ \vdots \\ \boldsymbol{\Pi}_{r_t} \end{bmatrix}_{(pr_t)\times p}. \tag{45}$$

By substituting (45) back to (41), the derivative of the expected value of $\mathbf{x}$ given the interaction model (11) is then

$$\frac{\partial E[\mathbf{x}|\mathbf{T},\mathbf{B}]}{\partial \mathbf{T}} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \\ \vdots \\ \mathbf{\Gamma}_{r_t} \end{bmatrix}_{(pr_t)\times p} + \begin{bmatrix} \mathbf{\Pi}_1 \\ \mathbf{\Pi}_2 \\ \vdots \\ \mathbf{\Pi}_{r_t} \end{bmatrix}_{(pr_t)\times p} . \quad (46)$$

## Acknowledgment

The authors would like to thank the four anonymous reviewers for their constructive comments, and thank Dr. Lefei Zhang and the RIT researchers for their help with the datasets. This work was partly supported by University College London's Security Science Doctoral Training Centre under Engineering and Physical Sciences Research Council (EPSRC) grant EP/G037264/1.

## References

[1] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, M. Brueggeman, Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms, Signal Processing Magazine, IEEE 31 (1) (2014) 24–33.

[2] N. M. Nasrabadi, Hyperspectral target detection: An overview of current and future challenges, Signal Processing Magazine, IEEE 31 (1) (2014) 34–44.

[3] S. Matteoli, M. Diani, G. Corsini, A tutorial overview of anomaly detection in hyperspectral images, Aerospace and Electronic Systems Magazine, IEEE 25 (7) (2010) 5–28.

[4] S. Matteoli, M. Diani, J. Theiler, An overview of background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 7 (6) (2014) 2317–2336.

[5] L. L. Scharf, B. Friedlander, Matched subspace detectors, Signal Processing, IEEE Transactions on 42 (8) (1994) 2146–2157.

[6] J. C. Harsanyi, C.-I. Chang, Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach, Geoscience and Remote Sensing, IEEE Transactions on 32 (4) (1994) 779–785.

[7] D. G. Manolakis, G. A. Shaw, N. Keshava, Comparative analysis of hyperspectral adaptive matched filter detectors, in: AeroSense 2000, International Society for Optics and Photonics, 2000, pp. 2–17.

[8] D. R. Fuhrmann, E. J. Kelly, R. Nitzberg, A CFAR adaptive matched filter detector, Aerospace and Electronic Systems, IEEE Transactions on 28 (1) (1992) 208–216.

[9] S. Kraut, L. L. Scharf, The CFAR adaptive subspace detector is a scale-invariant GLRT, Signal Processing, IEEE Transactions on 47 (9) (1999) 2538–2541.

[10] S. Kraut, L. L. Scharf, L. T. McWhorter, Adaptive subspace detectors, Signal Processing, IEEE Transactions on 49 (1) (2001) 1–16.

[11] Q. Du, H. Ren, C.-I. Chang, A comparative study for orthogonal subspace projection and constrained energy minimization, IEEE Transactions on Geoscience and Remote Sensing 41 (6) (2003) 1525–1529.

[12] H. Kwon, N. M. Nasrabadi, A comparative analysis of kernel subspace target detectors for hyperspectral imagery, EURASIP Journal on Applied Signal Processing 2007 (1) (2007) 193–193.

[13] S. Yang, Z. Shi, W. Tang, Robust hyperspectral image target detection using an inequality constraint, IEEE Transactions on Geoscience and Remote Sensing 53 (6) (2015) 3389–3404.

[14] Z. Zou, Z. Shi, Hierarchical suppression method for hyperspectral target detection, IEEE Transactions on Geoscience and Remote Sensing 54 (1) (2016) 330–342.

[15] S. Yang, Z. Shi, Hyperspectral image target detection improvement based on total variation, IEEE Transactions on Image Processing 25 (5) (2016) 2249–2258.

[16] Y. Chen, N. M. Nasrabadi, T. D. Tran, Sparse representation for target detection in hyperspectral imagery, Selected Topics in Signal Processing, IEEE Journal of 5 (3) (2011) 629–640.

[17] Y. Zhang, B. Du, L. Zhang, A sparse representation-based binary hypothesis model for target detection in hyperspectral images, Geoscience and Remote Sensing, IEEE Transactions on 53 (3) (2015) 1346–1354.

[18] Y. Chen, N. M. Nasrabadi, T. D. Tran, Kernel sparse representation for hyperspectral target detection, in: SPIE Defense, Security, and Sensing, International Society for Optics and Photonics, 2012, pp. 839005–839005–9.

[19] Y. Zhang, L. Zhang, B. Du, S. Wang, A nonlinear sparse representation-based binary hypothesis model for hyperspectral target detection, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 8 (6) (2015) 2513–2522.

[20] W. Li, Q. Du, A survey on representation-based classification and detection in hyperspectral remote sensing imagery, Pattern Recognition Letters (2015) –doi:10.1016/j.patrec.2015.09.010.

[21] W. Li, Q. Du, B. Zhang, Combined sparse and collaborative representation for hyperspectral target detection, Pattern Recognition 48 (12) (2015) 3904–3916.

[22] D. Manolakis, C. Siracusa, G. Shaw, Hyperspectral subpixel target detection using the linear mixing model, Geoscience and Remote Sensing, IEEE Transactions on 39 (7) (2001) 1392–1409.

[23] J. M. Nascimento, J. M. Bioucas-Dias, Nonlinear mixture model for hyperspectral unmixing, in: SPIE Europe Remote Sensing, International Society for Optics and Photonics, 2009, pp. 74770I–74770I.

[24] W. Fan, B. Hu, J. Miller, M. Li, Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data, International Journal of Remote Sensing 30 (11) (2009) 2951–2962.

[25] A. Halimi, Y. Altmann, N. Dobigeon, J.-Y. Tourneret, Nonlinear unmixing of hyperspectral images using a generalized bilinear model, Geoscience and Remote Sensing, IEEE Transactions on 49 (11) (2011) 4153–4162.

[26] T. W. Ray, B. C. Murray, Nonlinear spectral mixing in desert vegetation, Remote sensing of environment 55 (1) (1996) 59–64.

[27] X. Chen, L. Vierling, Spectral mixture analyses of hyperspectral data acquired using a tethered balloon, Remote Sensing of Environment 103 (3) (2006) 338–350.

[28] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. Van der Zande, W. W. Verstraeten, P. Coppin, Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards, Remote Sensing of Environment 113 (6) (2009) 1183–1193.

[29] R. Heylen, P. Scheunders, A multilinear mixing model for nonlinear spectral unmixing, IEEE Transactions on Geoscience and Remote Sensing 54 (1) (2016) 240–251.

[30] B. Du, L. Zhang, Target detection based on a dynamic subspace, Pattern Recognition 47 (1) (2014) 344–358.

[31] L. Zhang, L. Zhang, D. Tao, X. Huang, B. Du, Hyperspectral remote sensing image subpixel target detection based on supervised metric learning, IEEE Transactions on Geoscience and Remote Sensing 52 (8) (2014) 4955–4965.

[32] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, S. Hager, Development of a web-based application to evaluate target finding algorithms, in: Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, Vol. 2, IEEE, 2008, pp. II–915.

[33] L. Zhang, L. Zhang, D. Tao, X. Huang, Sparse transfer manifold embedding for hyperspectral target detection, IEEE Transactions on Geoscience and Remote Sensing 52 (2) (2014) 1030–1043.

[34] L. Gao, B. Yang, Q. Du, B. Zhang, Adjusted spectral matched filter for target detection in hyperspectral imagery, Remote Sensing 7 (6) (2015) 6611–6634.

[35] W. Li, Q. Du, Collaborative representation for hyperspectral anomaly detection, IEEE Transactions on Geoscience and Remote Sensing 53 (3) (2015) 1463–1474.

[36] Y. Zhang, B. Du, L. Zhang, S. Wang, A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection, IEEE Transactions on Geoscience and Remote Sensing 54 (3) (2016) 1376–1389.

[37] D. Manolakis, D. Marden, G. A. Shaw, Hyperspectral image processing for automatic target detection applications, Lincoln Laboratory Journal 14 (1) (2003) 79–116.

[38] J. A. Tropp, A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Transactions on information theory 53 (12) (2007) 4655–4666.