HYPERSPECTRAL TARGET DETECTION USING KERNEL MATCHED SUBSPACE DETECTOR

Heesung Kwon and Nasser M. Nasrabadi

U.S. Army Research Laboratory, 2800 Powder Mill Rd., Adelphi, MD 20783-1197

ABSTRACT

In this paper we present a nonlinear realization of a subspace signal detection approach based on the generalized likelihood ratio test (GLRT) – so called matched subspace detectors (MSD). The linear model for MSD is first extended to a high, possibly infinite, dimensional feature space and then the corresponding nonlinear GLRT expression is obtained. In order to address the intractability of the GLRT in the nonlinear feature space we kernelize the nonlinear GLRT using kernel eigenvector representations as well as the kernel trick where dot products in the nonlinear feature space are implicitly computed by kernels. The proposed kernel-based nonlinear detector, so called kernel matched subspace detector (KMSD), is applied to a given hyperspectral imagery – HYDICE (HYperspectral Digital Imagery Collection Experiment) images – to detect targets of interest. KMSD showed superior detection performance over MSD for the HYDICE images tested in this paper.

1. INTRODUCTION

In target detection based on airborne hyperspectral sensors, such as the HYDICE sensors, two main issues need to be addressed – spectral variability and spectral mixing [1, 2]. The spectral variability is a phenomenon that stems mainly from variations existing on the material surface. The material variations entail direct changes in reflectance that no two pixel spectra are identical even within the same types of materials. Irregular illumination and atmospheric effects often deepen the spectral variability to the extent that some pixel spectra from a local area with the same type of materials take significantly different spectral shapes.

The HYDICE sensors view normally large areas, resulting in coarse spatial resolution for sensor images. Consequently, in HYDICE images, a pixel spectrum represents a mixture of spectral signatures of different constituent materials that are present in the corresponding pixel area – an image region represented by one pixel. Therefore, the notion of subpixel targets leads to the use of binary hypothesis tests based on a linear spectral mixing model, in which the likelihood ratio test is performed to predict the presence of target within the pixel spectrum [3]. The pixel spectrum is labeled target if it includes the spectral signatures of target materials.

The linear subspace mixture model can easily be extended to a nonlinear domain by mapping the input data into a potentially infinite feature space which can be efficiently implemented by kernel methods [4, 5]. The kernel methods have emerged as new nonlinear-based learning techniques that implicitly exploit the dot product of feature vectors generated by the nonlinear mapping of the input vectors using kernel representations. Recently, kernel versions of data discrimination techniques based on spectral matched filtering, Fisher's linear discriminant, Mahalanobis distance measure, and principal component analysis (PCA) have been introduced [6, 7, 8, 9].

In this paper we combine the notion of kernel-based learning and matched subspace detection developed by Scharf et. al [3] to develop a new subpixel target detection method in hyperspectral images, so called kernel matched subspace detectors. To this end, based on [3], we first relate a hyperspectral detection problem in the original input space to the one in the nonlinear feature space and derive the corresponding general likelihood ratio test (GLRT) expression. The nonlinear GLRT needs to be kernelized before it can be implemented and the kernelization procedure is presented in this paper.

The paper is organized as follows. Brief introduction of kernel methods is presented in Section 2. In Section 3 we formulate the hyperspectral target detection problem using the hypothesis test based on linear subspace mixture model and derive the corresponding GLRT. In Section 4 we derive the nonlinear expression of the GLRT in the nonlinear feature domain and kernelize it using kernel representations. The proposed kernel-based detector is applied to the given HYDICE image in Section 5. Conclusions are summarized in Section 6.

2. KERNEL METHODS

Let $\mathcal{Y}$ be a P-dimensional input space ($\mathcal{Y} \subseteq \mathcal{R}^P$) and $\mathcal{F}$ be a nonlinear feature space associated with $\mathcal{Y}$ by a nonlinear mapping function Φ

$$\begin{aligned} \Phi : \mathcal{Y} &\rightarrow \mathcal{F}, \\ \mathbf{y} &\mapsto \Phi(\mathbf{y}), \end{aligned} \tag{1}$$

where $\mathbf{y}$ is an input vector in $\mathcal{Y}$. Because of the nonlinearity associated with the map Φ, $\mathcal{F}$ is normally in high, possibly infinite, dimensionality.

However, use of kernel methods allows us to implicitly compute the dot products in $\mathcal{F}$ without mapping the input vectors into $\mathcal{F}$; therefore, in the kernel methods, the map Φ does not need to be identified. The kernel representation for the dot products in $\mathcal{F}$ is expressed as

$$\begin{aligned} k(\mathbf{y}_i, \mathbf{y}_j) &= < \Phi(\mathbf{y}_i), \Phi(\mathbf{y}_j) > \\ &= \Phi(\mathbf{y}_i) \cdot \Phi(\mathbf{y}_j). \end{aligned} \tag{2}$$

Equation 2 shows that the dot products in $\mathcal{F}$ can be avoided by replacing it with a kernel function, a nonlinear function which can be easily calculated without identifying the nonlinear map Φ. Two commonly used kernels are the Gaussian RBF kernel: $k(\mathbf{x}, \mathbf{y}) = \exp(\frac{-\|\mathbf{x}-\mathbf{y}\|^2}{c})$ and Polynomial kernel: $((\mathbf{x} \cdot \mathbf{y}) + \theta)^d$. See [4] for

Correspondence:Email:nnasraba@arl.army.mil

0-7803-8554-3/04/$20.00 ©2004 IEEE.

detailed information about the properties of kernels and kernel-based learning.

Given a training data in the feature space we can also determine the eigenvectors of its covariance matrix and as shown in [4] any of the eigenvectors $\mathbf{e}$ with $\lambda \neq 0$ are spanned by the training samples $\Phi(\mathbf{y}_1), \ldots, \Phi(\mathbf{y}_m)$ – i.e.

$$\mathbf{e} = \sum_{i=1}^{m} \gamma_i \Phi(\mathbf{y}_i) = \Phi_{\mathbf{Y}} \boldsymbol{\gamma}, \tag{3}$$

where $\Phi_{\mathbf{Y}} = \left[\begin{array}{cccc} \Phi(\mathbf{y}_1) & \Phi(\mathbf{y}_2) & \ldots & \Phi(\mathbf{y}_m) \end{array} \right]$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_m)^T$. As pointed out in [9]$\boldsymbol{\gamma}$ turn out to be the eigenvectors with nonzero eigenvalues of a centered kernel matrix $\mathbf{K}$ ($(\mathbf{K})_{ij} = k(\mathbf{y}_i, \mathbf{y}_j)$). Note that $\boldsymbol{\gamma}$ need to be normalized by the square root of their corresponding eigenvalues [9].

3. LINEAR SUBSPACE SPECTRAL MIXING MODEL

The detection is based on a hypothesis test in which two competing hypotheses ($\mathbf{H_0}$ and $\mathbf{H_1}$) are tested for an input spectrum to decide which hypothesis is best related to the input spectrum. The hyperspectral target detection problem in a P-dimensional input space is expressed as

$$\begin{aligned} &\mathbf{H}_0 : \mathbf{y} = \mathbf{B}\zeta + \mathbf{n}, \\ &\mathbf{H}_1 : \mathbf{y} = \mathbf{T}\theta + \mathbf{B}\zeta + \mathbf{n} = \left[\begin{array}{cc} \mathbf{T} & \mathbf{B} \end{array} \right] \left[\begin{array}{c} \theta \\ \zeta \end{array} \right] + \mathbf{n}, \end{aligned} \tag{4}$$

where $\mathbf{T}$ and $\mathbf{B}$ represent orthogonal matrices whose P- dimensional column vectors span the target and background subspaces, respectively; θ and ζ are unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of $\mathbf{T}$ and $\mathbf{B}$, respectively; $\mathbf{n}$ represents Gaussian random noise ($\mathbf{n} \in \mathcal{R}^P$) distributed as $\mathcal{N}(0, \sigma^2\mathbf{I})$; and $\left[\begin{array}{cc} \mathbf{T} & \mathbf{B} \end{array} \right]$ is a concatenated matrix of $\mathbf{T}$ and $\mathbf{B}$. The numbers of the column vectors of $\mathbf{T}$ and $\mathbf{B}$, N_t and N_b, respectively, are usually smaller than P ($N_t, N_b < P$).

Given the linear subspace detection model and the two hypotheses about how the input vector is generated (4), the generalized likelihood ratio test (GLRT) is given by [3]:

$$\mathbf{L}_2(\mathbf{y}) = \frac{\mathbf{y}^T(\mathbf{I} - \mathbf{P_B})\,\mathbf{y}}{\mathbf{y}^T(\mathbf{I} - \mathbf{P_{TB}})\,\mathbf{y}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{5}$$

where $\mathbf{P_B} = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \mathbf{B}\mathbf{B}^T$ is a projection matrix associated with the N_b-dimensional background subspace $< \mathbf{B} >$; $\mathbf{P_{TB}}$ is a projection matrix associated with the $(N_b + N_t)$- dimensional target-and-background subspace $< \mathbf{TB} >$.

4. KERNEL MATCHED SUBSPACE DETECTION

The nonlinear hyperspectral detection problem based on the target and background subspaces can be described in the feature space $\mathcal{F}$ as

$$\begin{aligned} &\mathbf{H}_{0_\Phi} : \Phi(\mathbf{y}) = \mathbf{B}_\Phi \zeta_\Phi + \mathbf{n}_\Phi, \\ &\mathbf{H}_{1_\Phi} : \Phi(\mathbf{y}) = \mathbf{T}_\Phi \theta_\Phi + \mathbf{B}_\Phi \zeta_\Phi + \mathbf{n}_\Phi \\ &\qquad\qquad = \left[\begin{array}{cc} \mathbf{T}_\Phi & \mathbf{B}_\Phi \end{array} \right] \left[\begin{array}{c} \theta_\Phi \\ \zeta_\Phi \end{array} \right] + \mathbf{n}_\Phi, \end{aligned} \tag{6}$$

where $\mathbf{T}_\Phi$ and $\mathbf{B}_\Phi$ represent full-rank matrices whose column vectors span target and background subspaces $< \mathbf{B}_\Phi >$ and $< \mathbf{T}_\Phi >$ in $\mathcal{F}$, respectively.

We now derive the GLRT of the nonlinear hyperspectral detection problem described by the model in (6):

$$\mathbf{L}_2(\Phi(\mathbf{y})) = \frac{\Phi(\mathbf{y})^T(\mathbf{P}_{\mathbf{I}_\Phi} - \mathbf{P}_{\mathbf{B}_\Phi})\Phi(\mathbf{y})}{\Phi(\mathbf{y})^T(\mathbf{P}_{\mathbf{I}_\Phi} - \mathbf{P}_{\mathbf{T}_\Phi\mathbf{B}_\Phi})\Phi(\mathbf{y})}, \tag{7}$$

where $\mathbf{P}_{\mathbf{I}_\Phi}$ represents an identity projection operator in $\mathcal{F}$; $\mathbf{P}_{\mathbf{B}_\Phi} = \mathbf{B}_\Phi(\mathbf{B}_\Phi^T\mathbf{B}_\Phi)^{-1}\mathbf{B}_\Phi^T = \mathbf{B}_\Phi\mathbf{B}_\Phi^T$ is a background projection matrix; and $\mathbf{P}_{\mathbf{T}_\Phi\mathbf{B}_\Phi}$ is a target-and-background projection matrix in $\mathcal{F}$. The basic operations associated with every term in (7) – $\Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{I}_\Phi}\Phi(\mathbf{y})$, $\Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{B}_\Phi}\Phi(\mathbf{y})$, and $\Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{T}_\Phi\mathbf{B}_\Phi}\Phi(\mathbf{y})$ – can be regarded as principal components extraction which is the projection of $\Phi(\mathbf{y})$ onto the significant eigenvectors in $\mathcal{F}$. Note (7) can not be implemented explicitly due to potential infinite dimensionality of Φ, therefore, (7) has to be kernelized in order to obtain an expression in terms of the kernel function k.

To kernelize (7) first consider its numerator,

$$\begin{aligned} &\Phi(\mathbf{y})^T(\mathbf{P}_{\mathbf{I}_\Phi} - \mathbf{P}_{\mathbf{B}_\Phi})\Phi(\mathbf{y}) \\ &= \Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{I}_\Phi}\Phi(\mathbf{y}) - \Phi(\mathbf{y})^T\mathbf{B}_\Phi\mathbf{B}_\Phi^T\Phi(\mathbf{y}). \end{aligned} \tag{8}$$

Using (3) $\mathbf{B}_\Phi$ and $\mathbf{T}_\Phi$ can be written as

$$\mathbf{B}_\Phi = \left[\begin{array}{cccc} \mathbf{e_b}^1 & \mathbf{e_b}^2 & \ldots & \mathbf{e_b}^{N_b} \end{array} \right] = \Phi_{\mathbf{Z_B}}\mathcal{B}, \tag{9}$$

$$\mathbf{T}_\Phi = \left[\begin{array}{cccc} \mathbf{e_t}^1 & \mathbf{e_t}^2 & \ldots & \mathbf{e_t}^{N_t} \end{array} \right] = \Phi_{\mathbf{Z_T}}\mathcal{T}, \tag{10}$$

where $\mathbf{e_b}^i$ and $\mathbf{e_t}^j$ are the significant eigenvectors of $\mathbf{C}_{\mathbf{B}_\Phi}$ and $\mathbf{C}_{\mathbf{T}_\Phi}$, respectively; $\Phi_{\mathbf{Z_B}} = \left[\begin{array}{cccc} \Phi(\mathbf{y}_1) & \Phi(\mathbf{y}_2) & \ldots & \Phi(\mathbf{y}_M) \end{array} \right]$, $\mathbf{y}_i \in \mathbf{Z_B}$ and $\Phi_{\mathbf{Z_T}} = \left[\begin{array}{cccc} \Phi(\mathbf{y}_1) & \Phi(\mathbf{y}_2) & \ldots & \Phi(\mathbf{y}_N) \end{array} \right], \mathbf{y}_i \in \mathbf{Z_T}$; the column vectors of $\mathcal{B}$ and $\mathcal{T}$ represent only the significant eigenvectors ($\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \ldots, \boldsymbol{\beta}^{N_b}$) and ($\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2, \ldots, \boldsymbol{\alpha}^{N_t}$) of the background kernel matrix $\mathbf{K}(\mathbf{Z_B}, \mathbf{Z_B}) = (\mathbf{K})_{ij} = k(\mathbf{y}_i, \mathbf{y}_j)$, $\mathbf{y}_i, \mathbf{y}_j \in \mathbf{Z_B}$ and the target kernel matrix$\mathbf{K}(\mathbf{Z_T}, \mathbf{Z_T}) = (\mathbf{K})_{ij} = k(\mathbf{y}_i, \mathbf{y}_j)$, $\mathbf{y}_i, \mathbf{y}_j \in \mathbf{Z_T}$, respectively.

Using (9) the projection of $\Phi(\mathbf{y})$ onto $\mathbf{B}_\Phi$ becomes

$$\begin{aligned} \mathbf{B}_\Phi^T\Phi(\mathbf{y}) &= \left[\begin{array}{cccc} \mathbf{e_b}^1 & \mathbf{e_b}^2 & \ldots & \mathbf{e_b}^{N_b} \end{array} \right]^T \Phi(\mathbf{y}) \\ &= \left[\begin{array}{c} \boldsymbol{\beta}^{1T}\Phi_{\mathbf{Z_B}}^T\Phi(\mathbf{y}) \\ \boldsymbol{\beta}^{2T}\Phi_{\mathbf{Z_B}}^T\Phi(\mathbf{y}) \\ \cdot \\ \boldsymbol{\beta}^{N_bT}\Phi_{\mathbf{Z_B}}^T\Phi(\mathbf{y}) \end{array} \right] = \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{y}), \end{aligned} \tag{11}$$

and, similarly, using (10) the projection onto $\mathbf{T}_\Phi$ is

$$\begin{aligned} \mathbf{T}_\Phi^T\Phi(\mathbf{y}) &= \left[\begin{array}{cccc} \mathbf{e_t}^1 & \mathbf{e_t}^2 & \ldots & \mathbf{e_t}^t \end{array} \right]^T \Phi(\mathbf{y}) \\ &= \left[\begin{array}{c} \boldsymbol{\alpha}^{1T}\Phi_{\mathbf{Z_T}}^T\Phi(\mathbf{y}) \\ \boldsymbol{\alpha}^{2T}\Phi_{\mathbf{Z_T}}^T\Phi(\mathbf{y}) \\ \cdot \\ \boldsymbol{\alpha}^{N_tT}\Phi_{\mathbf{Z_T}}^T\Phi(\mathbf{y}) \end{array} \right] = \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{y}), \end{aligned} \tag{12}$$

where $\mathbf{K}(\mathbf{Z_B}, \mathbf{y})$ and $\mathbf{K}(\mathbf{Z_T}, \mathbf{y})$ are column vectors whose entries are $k(\mathbf{x}_i, \mathbf{y})$ for $\mathbf{x}_i \in \mathbf{Z_B}$ and $\mathbf{x}_i \in \mathbf{Z_T}$, respectively. Therefore, $\Phi(\mathbf{y})^T\mathbf{B}_\Phi\mathbf{B}_\Phi^T\Phi(\mathbf{y})$ can be written as

$$\Phi(\mathbf{y})^T\mathbf{B}_\Phi\mathbf{B}_\Phi^T\Phi(\mathbf{y}) = \mathbf{K}(\mathbf{Z_B}, \mathbf{y})^T\mathcal{B}\mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{y}). \tag{13}$$

The projection onto the identity operator $\Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{I}_\Phi}\Phi(\mathbf{y})$ also needs to be kernelized. $\mathbf{P}_{\mathbf{I}_\Phi}$ is defined as $\mathbf{P}_{\mathbf{I}_\Phi} := \mathbf{\Omega_\Phi}\mathbf{\Omega}_\Phi^T$, where $\mathbf{\Omega_\Phi} = \left[\begin{array}{ccc} \mathbf{e_q}^1 & \mathbf{e_q}^2 & \dots \end{array}\right]$ is a matrix whose columns are all the eigenvectors with $\lambda \neq 0$ that are in the span of $\Phi(\mathbf{y}_i)$, $\mathbf{y}_i \in \mathbf{Z_T} \cup \mathbf{Z_B}$. From (3) $\mathbf{\Omega_\Phi}$ can be expressed as

$$\mathbf{\Omega_\Phi} = [\mathbf{e_q}^1, \mathbf{e_q}^2, \dots \mathbf{e_q}^{N_{bt}}] = \Phi_{\mathbf{Z_{TB}}}\Delta, \tag{14}$$

where $\Phi_{\mathbf{Z_{TB}}} = \Phi_{\mathbf{Z_T}} \cup \Phi_{\mathbf{Z_B}}$ and $\Delta = (\boldsymbol{\kappa}^1, \boldsymbol{\kappa}^2, \dots, \boldsymbol{\kappa}^{N_{bt}})$ are the eigenvectors of the kernel matrix $\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{Z_{TB}}) = (\mathbf{K})_{ij} = k(\mathbf{y}_i, \mathbf{y}_j)$, $\mathbf{y}_i, \mathbf{y}_j \in \mathbf{Z_T} \cup \mathbf{Z_B}$ with nonzero eigenvalues, normalized by their associated eigenvalues. Using $\mathbf{P}_{\mathbf{I}_\Phi} = \mathbf{\Omega_\Phi}\mathbf{\Omega}_\Phi^T$ and (14)

$$\begin{aligned} & \Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{I}_\Phi}\Phi(\mathbf{y}) \\ = \;& \Phi(\mathbf{y})^T\Phi_{\mathbf{Z_{TB}}}\Delta\Delta^T\Phi_{\mathbf{Z_{TB}}}^T\Phi(\mathbf{y}) \\ = \;& \mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y})^T\Delta\Delta^T\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y}), \end{aligned} \tag{15}$$

where $\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y})$ is a concatenated vector $\left[\begin{array}{cc} \mathbf{K}(\mathbf{Z_T}, \mathbf{y})^T & \mathbf{K}(\mathbf{Z_B}, \mathbf{y})^T \end{array}\right]^T$. The kernelized numerator of (7) is then given by

$$\begin{aligned} & \mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y})^T\Delta\Delta^T\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y}) - \\ & \quad \mathbf{K}(\mathbf{Z_B}, \mathbf{y})^T\mathcal{B}\mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{y}), \end{aligned} \tag{16}$$

We now kernelize $\Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{T}_\Phi\mathbf{B}_\Phi}\Phi(\mathbf{y})$ in the denominator of (7) to complete the kernelization process. From (9), and (10)

$$\begin{aligned} & \Phi(\mathbf{y})^T\mathbf{P}_{\mathbf{T}_\Phi\mathbf{B}_\Phi}\Phi(\mathbf{y}) = \Phi(\mathbf{y})^T\left[\begin{array}{cc} \mathbf{T}_\Phi & \mathbf{B}_\Phi \end{array}\right] \times \\ & \left[\begin{array}{cc} \mathbf{T}_\Phi^T\mathbf{T}_\Phi & \mathbf{T}_\Phi^T\mathbf{B}_\Phi \\ \mathbf{B}_\Phi^T\mathbf{T}_\Phi & \mathbf{B}_\Phi^T\mathbf{B}_\Phi \end{array}\right]^{-1}\left[\begin{array}{c} \mathbf{T}_\Phi^T \\ \mathbf{B}_\Phi^T \end{array}\right]\Phi(\mathbf{y}) \\ & = \left[\begin{array}{cc} \mathbf{K}(\mathbf{Z_T}, \mathbf{y})^T\mathcal{T} & \mathbf{K}(\mathbf{Z_B}, \mathbf{y})^T\mathcal{B} \end{array}\right] \times \\ & \left[\begin{array}{cc} \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{Z_T})\mathcal{T} & \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{Z_B})\mathcal{B} \\ \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{Z_T})\mathcal{T} & \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{Z_B})\mathcal{B} \end{array}\right]^{-1} \times \\ & \left[\begin{array}{c} \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{y}) \\ \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{y}) \end{array}\right]. \end{aligned} \tag{17}$$

Finally, substituting (13), (16), and (17) into (7) the kernelized GLR is given by

$$\begin{aligned} \mathbf{L_{2K}} = & (\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y})^T\Delta\Delta^T\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y}) - \\ & \mathbf{K}(\mathbf{Z_B}, \mathbf{y})^T\mathcal{B}\mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{y}))\,/ \\ & (\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y})^T\Delta\Delta^T\mathbf{K}(\mathbf{Z_{TB}}, \mathbf{y}) - \\ & \left[\begin{array}{cc} \mathbf{K}(\mathbf{Z_T}, \mathbf{y})^T\mathcal{T} & \mathbf{K}(\mathbf{Z_B}, \mathbf{y})^T\mathcal{B} \end{array}\right]\Lambda_1^{-1} \times \\ & \left[\begin{array}{c} \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{y}) \\ \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{y}) \end{array}\right]), \end{aligned} \tag{18}$$

where $\Lambda_1 = \left[\begin{array}{cc} \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{Z_T})\mathcal{T} & \mathcal{T}^T\mathbf{K}(\mathbf{Z_T}, \mathbf{Z_B})\mathcal{B} \\ \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{Z_T})\mathcal{T} & \mathcal{B}^T\mathbf{K}(\mathbf{Z_B}, \mathbf{Z_B})\mathcal{B} \end{array}\right]$. (18) can now be implemented with no knowledge of the mapping function Φ. The only requirement is a good choice for the kernel function k. Note that each kernel matrix $\mathbf{K}$ needs to centered properly, as shown in [4].

5. SIMULATION RESULTS

In this section, we apply the proposed kernel matched subspace detector to the HYDICE test image that includes targets of interest (military vehicles). The HYDICE imaging sensor generates 210 bands across the whole spectral range (0.4 – 2.5 μm), but we use only 150 bands by discarding water absorption and low signal to noise ratio (SNR) bands; the bands used are the 23rd–101st, 109th–136th, and 152nd–194th.

The HYDICE image from the Desert Radiance II data collection was used to test the proposed detectors. The Desert Radiance II (DR-II) image contains 6 targets located in the dirt road, as shown in the sample band image in Fig. 1. The targets in the image are all military vehicles.

Fig. 1. Sample band image (48th) from the Desert Radiance II image.

We compare the detection performance between the matched subspace detector (MSD) (5) defined in the original input domain and the kernel matched subspace detector (KMSD) (18) using the HYDICE test image. The Gaussian RBF kernel, $\mathbf{k}(\mathbf{x}, \mathbf{y}) = \exp(\frac{-\|\mathbf{x}-\mathbf{y}\|^2}{c})$, was used to implement the kernel-based detector; the value of c used was 40. The numbers of the column vectors (eigenvectors) of $\mathcal{B}$ and $\mathcal{T}$, N_b and N_t, used were 6 and 3, respectively.

All the pixel vectors in the 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. The rescaling of pixel vectors was mainly performed to effectively utilize the dynamic range of Gaussian RBF kernel.

The receiver operating characteristics (ROC) curves representing detection probability P_d versus false alarm rates N_f, were also generated to provide quantitative performance comparison. For ROC curves generation, based on the ground truth information for the HYDICE image, we obtain the coordinates of all the rectangular target regions. Every target pixel inside the target regions is then considered as a candidate to be detected. P_d and N_f are defined as

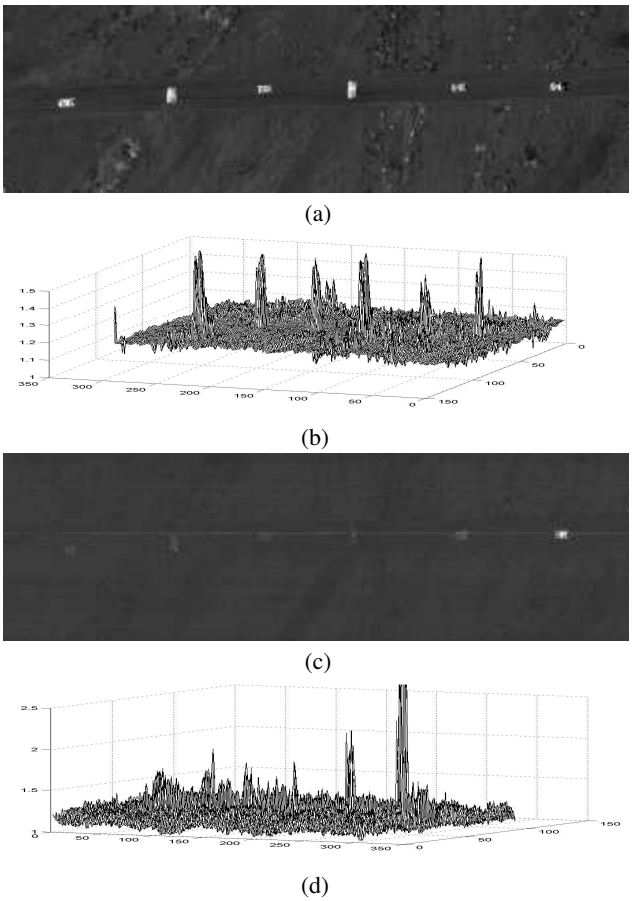
$$P_d := \tfrac{N_{hit}}{N_t} \quad \text{and} \quad N_f := \tfrac{N_{miss}}{N_{tot}}, \tag{19}$$

where N_{hit} represents the number of target pixels detected given a certain threshold; N_t represents the total number of target pixels in the image; N_{miss} represents the number of background pixels detected; and N_{tot} represents the total number of pixels in the image. P_d becomes one only when all the individual target pixels within a target are detected; perfect detection is, therefore, difficult to achieve and the values of P_d for both the KMSD and MSD reported in this paper are usually less than one.

Figs. 2 and 3 show the detection results including the ROC curves generated by applying KMSD and MSD to both the DR-II test image. In implementing KMSD and MSD the background samples were obtained from outside the test image to estimate the background subspace. Due to a lack of available target samples (the test image include all the targets, military vehicles, present in

the given data sets), the target samples were collected from one of the targets in each HYDICE test set: the right most target in the DR-II image in Fig. 1.

For the DR-II image, KMSD detected the majority of the target pixels with high contrast in GLRT values with respect to those of the background while MSD found a relatively small number of the target pixels with low contrast, especially, for the first four targets from the left, as shown in Fig. 2. The ROC curves in Fig. 3 also showed the superiority of KMSD over MSD for the DR-II imagery.



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

Fig. 2. Detection results for the Desert Radiance II image using the kernel matched subspace detector (KMSD) and the matched subspace detector (MSD). (a) KMSD, (b) 3-D plot of (a), (c) MSD, and (d) 3-D plot of (c).

Given the experimental results based on our limited hyperspectral data set, it can be said that if the target subspace is properly estimated to represent the target spectral variability, KMSD is expected, in general, to perform better than MSD.

6. CONCLUSIONS

We have presented a nonlinear version of the matched subspace detector by kernelizing the corresponding nonlinear GLRT expression defined in a high dimensional feature space. The kernelization procedure for the nonlinear GLRT was completely derived. The proposed nonlinear detector, referred to as the kernel matched subspace detector (KMSD), was applied to hyperspectral subpixel target detection and the detection performance was compared to the original matched subspace detector (MSD). The detection results based on the given HYDICE image confirmed that the kernel-based learning, which was implicitly performed in the nonlinear feature space, was a powerful approach to understand the underlying structures of the given data set and representing unseen targets based on the limited training samples.

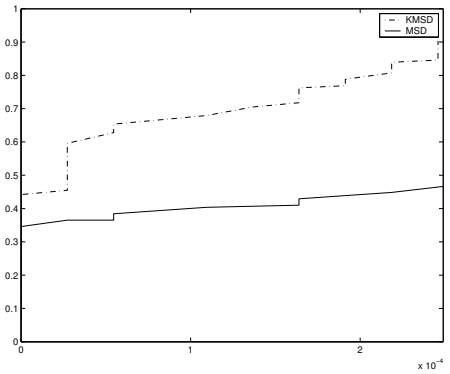


Fig. 3. ROC curves obtained by KMSD and MSD for the Desert Radiance II image.

7. REFERENCES

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

[2] H. Kwon, S. Z. Der, and N. M. Nasrabadi, "Target classification using adaptive feature extraction and subspace projection for hyperspectral imagery," in *Computer Vision Beyond the Visible Spectrum, edited by Ioannis Pavlidis and Bir Bahnu*. 2004, to appear, Springer.

[3] L. L. Scharf and B Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

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

[5] K. R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B Schökopf, "A introduction to kernel-based learning algorithms," *IEEE Trans. Neural Networks.*, , no. 2, pp. 181–202, 2001.

[6] H. Kwon and N. M. Nasrabadi, "Hyperspectral target detection using kernel spectral matched filter," in *IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, July 2004, to appear.

[7] G. Baudat and F Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, , no. 12, pp. 2385–2404, 2000.

[8] A. Ruiz and López de Teruel P., "Nonlinear kernel-based statistical pattern analysis," *IEEE Trans. Neural Networks.*, vol. 12, no. 1, pp. 16–32, 2001.

[9] B Schökopf, A. J. Smola, and K.-R. Müller, "Kernel principal component analysis," *Neural Computation*, , no. 10, pp. 1299–1319, 1999.