PERGAMON

Pattern Recognition 32 (1999) 1161–1174

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

An unsupervised vector quantization-based target subspace projection approach to mixed pixel detection and classification in unknown background for remotely sensed imagery

Clark Brumbley, Chein-I Chang*

Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, U.S.A.

Received 26 January 1998; accepted 19 October 1998

Abstract

A recently developed orthogonal subspace projection (OSP) approach has been successfully applied to AVIRIS as well as HYDICE data for image classification. However, it has found that OSP performs poorly in multispectral image classification such as 3-band SPOT data. This is primarily due to the fact that the number of signatures to be classified is greater than that of spectral bands used for data acquisition in which case the effects of uninterested signatures cannot be properly annihilated via orthogonal projection. This constraint, referred to as band number constraint (BNC) is generally not applied to hyperspectral images because the number of signatures resident within the images is usually far less than the total number of spectral bands. In this paper, a new approach, called unsupervised vector quantization-based target subspace projection (UVQTSP) is presented which can be implemented in an unknown environment with all required information obtained from the data to be processed. The proposed UVQTSP has practical advantages over OSP, specifically, it relaxes the band number constraint (BNC) so that it can be applied to multispectral imagery. The UVQTSP uses vector quantization to find a set of clusters representing the unknown signatures and interferers which will be eliminated prior to target detection and classification. The number of clusters can be determined by constraints such as the intrinsic dimensionality or the number of spectral bands. This process is carried out in an unsupervised manner without training data. The superiority of UVQTSP is demonstrated through real data including SPOT and HYDICE images.

Keywords: Band number constraint (BNC); Orthogonal subspace projection (OSP); Unsupervised vector quantization-based target subspace projection (UVQTSP)

Acronyms

AVIRIS	Airborne Visible/InfraRed Imaging Spectrometer
BNC	Band Number Constraint
HYDICE	HYperspectral Digital Imagery Collection Experiment
LBG algorithm	Linde–Buzo–Gray algorithm
MSS	MultiSpectral Scanner
OSP	Orthogonal Subspace Projection
SID	Spectral Information Divergence

* Corresponding author. Tel.: 001 410455 3502; Fax: 001 410 455 3969.

0031-3203/99/$20.00
PII: S0031-3203(98)00150-2

SNR	Signal-to-Noise Ratio
SPOT	Satellite Pour l'Observation de la Terra
TM	Thematic Mapper
TSC	Target Signature Classifier
UVQTSP	Unsupervised Vector Quantization-based Target Subspace Projection
VQ	Vector Quantization

1. Introduction

Over the past years, remote sensing instruments have advanced a new era where new sensors have been developed for better spatial resolution and spectral resolution. Of particular interest are image spectrometers which can implement hundreds of bands to improve sensors' capability of discriminating, identifying and quantifying materials with very similar spectral signatures which cannot be resolved by multispectral sensors such as the 4-band Multispectral Scanner (MSS) and 7-band Thematic Mapper (TM) [1]. Two types of such imaging spectrometers in current use are the 224-band Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) developed by NASA Jet Propulsion Laboratory and the 210-band HYperspectral Digital Imagery Collection Experiment (HYDICE) sensor developed by Naval Research Laboratory. While both enjoy very fine spectral resolution 10 nm meters with spectral coverage from 0.4 to 2.5 μm, they do have different spatial resolutions 30 m for AVIRIS and 1–4 m for HYDICE which result in different applications.

Since the area covered by a multispectral/hyperspectral image pixel generally contains more than one material with different spectral signatures, this results in a mixture of these signatures. Therefore, one of major challenges in multispectral/hyperspectral image classification is to detect, discriminate and identify various materials (or endmembers) embedded in a mixed pixel. However, their performance is highly dependent upon the abundance fractions of endmembers as well as their spectral characteristics. In particular, when the size of a target is smaller than the ground sampling distance (i.e., spatial resolution), detecting or classifying such a target at subpixel scale presents great difficulty because the target is completely embedded in only one pixel and cannot be seen from the image. Consequently, traditional pure-pixel-based image processing techniques cannot be directly applied to remotely sensed imagery.

Recently, an orthogonal subspace projection (OSP) approach has shown success in airborne visible/infrared imaging spectrometer (AVIRIS) [2] and hyperspectral digital imagery collection experiment (HYDICE) data in image classification and subpixel detection [2–6]. However, the applicability of the OSP to multispectral imagery has not been yet justified, but thought of as a straightforward application. A recent new finding [7–8] provided evidence that the OSP did not perform well in multispectral image classification such as SPOT (Satellite Pour l'Observation de la Terra) data as it did for hyperspectral images. This failure arises from an intrinsic constraint of the OSP, that is, the projection space into which the undesired signature space, denoted by U are projected does not have sufficient dimensionality to accommodate all signatures in U. When the OSP is applied, it first makes use of an undesired signature projector $P_U^{\perp} = I - UU^{\#}$ to annihilate effects resulting from signatures in U by projecting a hyperspectral image pixel vector into the orthogonal complement subspace of $\langle U \rangle$, denoted by $\langle U \rangle^{\perp}$. This annihilation works well if the dimensionality of $\langle U \rangle$, denoted by p, is less than the number of spectral bands, denoted by l. In this case, it may take a p-dimensional space to accommodate all p signatures in U. However, there are only l dimensions available to be used for annihilation of U. As a result, some signatures in U will inevitably blended into the projection space via $P_U^{\perp}$. These unwanted signatures are then mixed into the desired signature, $\mathbf{d}$. Such mixing further deteriorates the classification performance. As also noticed [9], when an l-dimensional pixel vector is projected into a signature space with dimensionality p, the resulting signal-to-noise ratio (SNR) is l/p. The SNR can be improved by either increasing l or decreasing p. In other words, SNR will be enhanced if $p < l$ and degraded if $p > l$. Since hyperspectral images always have relatively large data dimensionality compared to the number of signatures to be classified, SNR is usually increased by a magnitude $l/p \gg 1$. This $p > l$ problem, referred to as band number constraint (BNC) never becomes an issue and generally overlooked in hyperspectral image classification.

Interestingly, a similar finding was also observed in Refs. [10,11] where a new approach, called target signature classifier (TSC), was proposed for hyperspectral image classification. Instead of mapping a pixel vector into $\langle U \rangle^{\perp}$, TSC directly projects a hyperspectral image pixel vector into a target space $\langle \mathbf{d} \rangle$, a space linearly spanned by the desired target signature $\mathbf{d}$. This gives rise to a target signature projector, $P_{\mathbf{d}}$ as opposed to the undesired signature rejecter $P_U^{\perp}$ used in the OSP. Unfortunately, it comes at a price with a bias b which is a direct result of the effects caused by the signatures in U improperly mixed into $\langle \mathbf{d} \rangle$ through $P_{\mathbf{d}}$. Coincidentally, the creation of such bias is very closely related to the constraint, BNC.

In order to relax the BNC imposed on the OSP and the bias associated with the TSC, a new approach based on TSC, called unsupervised vector quantization-based target subspace projection (UVQTSP) is proposed in this paper. The UVQTSP employs vector quantization (VQ) to generate a set of clusters orthogonal to the target space $\langle \mathbf{d} \rangle$. The number of clusters is determined by practical

constraints such as the intrinsic dimensionality or the number of spectral bands used for data collection. The VQ-based clustering process is carried out in an unsupervised fashion with no need of a priori knowledge about the background and signatures except the desired signature **d** served as a target signature. Comparing to OSP which assumes the prior knowledge of all signatures in U, this is an important advantage because the knowledge of the background, interferences and undesired signatures is generally not known a priori in many applications and very difficult to obtain. In addition, the proposed UVQTSP can further generate unwanted interference signatures and eliminate them subsequently before classification takes place, a situation to which OSP does not apply. The VQ used in UVQTSP groups all possible undesired signatures as well as potential interferers into a number of clusters for U. Then an orthogonalization process is proposed to orthogonalize **d** with respect to U so that the bias created by the TSC vanishes [12]. In other words, this orthogonalization prevents the U from mixing into the target space $\langle \mathbf{d} \rangle$. At least three improvements yielded by UVQTSP are significant. Since UVQTSP makes use of an unsupervised vector quantization-based clustering process to find a set of clusters corresponding to **d** and U respectively to meet the constraint, BNC, UVQTSP can be used for multispectral image classification. Another is that only the prior knowledge about the signature of interest is required. This advantage removes a need of identifying undesired signatures in U used in $P_U^{\perp}$. A third advantage is that no information about the number of signatures in U is required. This can be determined by a rank curve analysis. Therefore, UVQTSP can be viewed as a generalization and an unsupervised version of OSP and TSC.

In order to demonstrate the superiority of UVQTSP, several experiments are conducted based on 3-band SPOT data and 210-band HYDICE data. UVQTSP performs better than the OSP in classifying hyperspectral images because UVQTSP can find potential interferers and then eliminate them prior to target classification. In general, these interferers are a direct result of unwanted signatures which are also extracted at the expense of improved spectral resolution. They presents more serious problem than the noise. Under this circumstance, UVQTSP can take care of the effects caused by such interferers. As expected, UVQTSP performs better than OSP.

This paper is organized as follows. In Section 2, we briefly review linear mixing problems and subspace projection approach. In Section 3, we describe the UVQTSP approach which includes a VQ-based clustering process. In Section 4 we present a series of data experiments based on SPOT and HYDICE data to demonstrate the advantages of UVQTSP. Finally, we include a brief conclusion in Section 5.

2. Subspace projection approach

2.1. Linear mixing problem

Linear spectral unmixing is a widely used approach in remotely sensed imagery to discriminate, identify and quantify individual spectral signatures (or endmembers) in a mixed pixel. Let $\mathbf{r}_i$ be an $l \times 1$ column vector and denote the ith pixel vector in a multispectral or hyperspectral image where l is the number of spectral bands and the bold face is used for vectors. Assume that M is an $l \times p$ signature matrix denoted by $(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_p)$ where $\mathbf{m}_j$ is an $l \times 1$ column vector represented by the jth spectral signature resident in the pixel $\mathbf{r}_i$ and p is the number of signatures. We also let α_i be a $p \times 1$ abundance column vector associated with $\mathbf{r}_i$ given by $(\alpha_{i1}\ \alpha_{i2}\ \ldots\ \alpha_{ip})^{\mathrm{T}}$ where α_{ij} denotes the fractional abundance of the jth signature present in pixel $\mathbf{r}_i$.

A linear spectral mixture model for $\mathbf{r}_i$ is described by

$$\mathbf{r}_i = M\alpha_i + \mathbf{n}_i, \tag{1}$$

where $\mathbf{n}_i$ is an $l \times 1$ column vector representing additive noise with zero mean and covariance matrix $\sigma^2 I$ and I is the $l \times l$ identity matrix.

2.2. Orthogonal subspace projection

In the following, we briefly review the Orthogonal Subspace Projection (OSP) approach [2]. First of all, we rewrite model (1) as

$$\mathbf{r} = \mathbf{d}\alpha_p + U\gamma + \mathbf{n}, \tag{2}$$

where the subscript i is suppressed, $\mathbf{d} = \mathbf{m}_p$ is a desired signature and $U = (\mathbf{m}_1 \mathbf{m}_2 \ldots \mathbf{m}_{p-1})$ is the undesired spectral signature matrix made up of a set of $p-1$ undesired signatures. Here, we assume without loss of generality that the last signature is the desired signature **d**. It should be noted that Eq. (2) can be extended straightforwardly to more than one desired signature. The reason of separating U from M is to allow us to design an orthogonal subspace projector to eliminate U from the pixel vector prior to classification. One such a projector was developed in Ref. [2] and called an undesired signature annihilator given by

$$P_U^{\perp} = I - UU^{\#}, \tag{3}$$

where $U^{\#} = (U^{\mathrm{T}}U)^{-1}U^{\mathrm{T}}$ is the pseudo-inverse of U and the notation $^{\perp}_{U}$ in $P_U^{\perp}$ indicates that the projector $P_U^{\perp}$ maps the observed pixel **r** into the range $\langle U \rangle^{\perp}$, the orthogonal complement of $\langle U \rangle$.

Now, applying $P_U^{\perp}$ to model (2) results in a new spectral signature model

$$P_U^{\perp}\mathbf{r} = P_U^{\perp}\mathbf{d}\alpha_p + \boldsymbol{P}_{\mathrm{U}}^{\perp}\mathbf{n}, \tag{4}$$

where the undesired signatures in U have been eliminated and the original noise has been suppressed to $P_U^{\perp}\mathbf{n}$.

Eq. (4) represents a standard signal detection problem. If the optimal criterion for the signal detection model specified by Eq. (4) is to maximize the Signal-to-Noise Ratio (SNR)

$$SNR = \frac{(\mathbf{x}^{\mathrm{T}}P_U^{\perp}\mathbf{d})\,\alpha_p^2(\mathbf{d}^{\mathrm{T}}P_U^{\perp}\mathbf{x})}{\mathbf{x}^{\mathrm{T}}P_U^{\perp}E[\mathbf{n}\mathbf{n}^{\mathrm{T}}]P_U^{\perp}\mathbf{x}} \quad \text{over} \quad \mathbf{x}, \tag{5}$$

then the maximum SNR of Eq. (5) can be obtained by a matched filter, denoted by $M_{\mathbf{d}}$ with letting $\mathbf{x} = \kappa\mathbf{d}$ where κ is a constant and the designed matched signal is given by $\mathbf{d}$.

It is easily shown that maximizing Eq. (5) is equivalent to finding the maximum eigenvalue of the following generalized eigenvalue problem [3]:

$$[P_U^{\perp}\mathbf{d}\alpha_p^2\mathbf{d}^{\mathrm{T}}P_U^{\perp}]\,\mathbf{x} = \lambda P_U^{\perp}\mathbf{x}. \tag{6}$$

Since Eqs. (5) and (6) present a two-class classification problem, the rank of the matrix on the left of Eq. (6) is one. This implies that the only nonzero eigenvalue is the maximum eigenvalue which also solves Eq. (6) with letting $\mathbf{x} = \kappa\mathbf{d}$. Accordingly, this eigenvalue can be obtained via Eq. (5) as follows:

$$\lambda_{\max} = SNR_{\max} = \frac{\alpha_p^2[\mathbf{d}^{\mathrm{T}}P_U^{\perp}\mathbf{d}][\mathbf{d}^{\mathrm{T}}P_U^{\perp}\mathbf{d}]}{\mathbf{d}^{\mathrm{T}}P_U^{\perp}\mathbf{d}} = \alpha_p^2\mathbf{d}^{\mathrm{T}}P_U^{\perp}\mathbf{d}. \tag{7}$$

Based on the approach outlined by Eqs. (4)–(7), a mixed pixel classification can be carried out by a two-stage process, an undesired signature annihilator $P_U^{\perp}$ followed by a matched filter, $M_{\mathbf{d}}$. More precisely, if we want to classify a desired signature, say $\mathbf{d}$ in a mixed pixel at the subpixel scale based on model (1), we first apply $P_U^{\perp}$ to model (2) to eliminate U, then use the matched filter $M_{\mathbf{d}}$ to extract $\mathbf{d}$ from the signal detection model (4). The operator coupling $P_U^{\perp}$ with $M_{\mathbf{d}}$ is called an Orthogonal Subspace Classifier, P_{OSP}, the one derived in Ref. [2] and denoted by

$$P_{OSP} = M_{\mathbf{d}}P_U^{\perp}. \tag{8}$$

2.3. Target signature classifier

Since model (1) assumes prior knowledge about signature abundance vector α, OSP was extended to a posteriori version OSP referred to as SSC [10] where model (1) was converted to a posteriori model by including a posteriori information obtained from the data. As a result, no a priori knowledge is required for α. The idea of SSC is to project pixel vectors into the signature space $\langle M\rangle$ linearly spanned by the signature matrix M to achieve abundance estimate while enhancing signatures and suppressing noise. The projector induced by M is called signature space projector, P_M and was derived by

$$P_M = M(M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}} = MM^{\#}, \tag{9}$$

where $M^{\#}$ is the pseudo-inverse of M [9,10]. Applying P_M in conjunction with P_{OSP} to model (1) yields a signature subspace classifier, P_{SSC} given by

$$P_{SSC} = M_{\mathbf{d}}P_U^{\perp}P_M = P_{OSP}P_M. \tag{10}$$

A similar classifier to P_{SSC} called Target Signature Classifier (TSC) was also derived based on Eq. (10) [10,11]. Instead of projecting pixels into $\langle M\rangle$, TSC designates a desired signature as a target signature and replaces the signature space projector P_M by the target signature projector $P_{\mathbf{d}}$ with using $\langle\mathbf{d}\rangle$, as the projection space. The resulting classifier is P_{TSC} given by

$$P_{TSC} = M_{\mathbf{d}}P_U^{\perp}P_{\mathbf{d}} = P_{OSP}P_{\mathbf{d}}. \tag{11}$$

If the target signature $\mathbf{d}$ is orthogonal to the undesired signature space $\langle U\rangle$, P_{TSC} will outperform P_{OSP} and P_{SSC}. Unfortunately, it is often not case in practice. A bias is generally produced as a direct result of non-orthogonality. It is a mixing of undesired signatures and $\langle\mathbf{d}\rangle$ and degrades the performance of P_{TSC}. A detailed comparison among three classifiers can be found in Refs. [10,11].

3. Unsupervised vector quantization-based subspace projection (UVQTSP)

Although the unknown signature abundance can be estimated by P_{SSC} and P_{TSC}, the knowledge of the signatures in M is still needed to be given a priori. In practice, this may be difficult to obtain since in most cases the information about how many signatures present in the pixel vectors is generally not available. In order to mitigate this dilemma the idea of TSC may be worth pursuing because it designates a signature of interest as a target signature $\mathbf{d}$ for projection space rather than using undesired signatures for projection space as do P_{OSP} and P_{SSC}. As mentioned, if $\langle\mathbf{d}\rangle$ is orthogonal to $\langle U\rangle$, there is no need of using $P_U^{\perp}$ given in Eq. (11). This orthogonality also results in no bias. So, the only required information is the knowledge of the target signature. In this section, we present an approach, called Unsupervised Vector Quantization-Based Target Subspace Projection (UVQTSP) that solves this problem and requires nothing more than the target signature.

The proposed UVQTSP consists of a sequence of operations which include an unsupervised VQ-based clustering preprocessing (a Gram–Schmidt-like orthogonalization projection procedure) followed by a matched filter. Since no prior knowledge is given except the desired signature, we appeal for VQ to find undesired

signatures in an unsupervised manner. The VQ will generate a set of clusters, each of whose centers or centroids represents one or more unknown undesired signatures. The number of clusters can be determined beforehand by the intrinsic dimensionality of data or the number of spectral bands. Nevertheless, it should be noted that the number of clusters can be determined by different applications. For example, in SPOT data, there are only three bands to be used for data collection. If there are more than three signatures, any of P_{OSP}, P_{SSC} and P_{TSC} breaks down and performs poorly. However, the classifier designed based on UVQTSP approach overcomes this problem by making use of VQ to group all other signatures excluding the desired signature into two clusters to represent undesired signatures. In this case, the undesired signatures also include interferers. Combining these two clusters' centers with the desired signature constitutes a three-dimensional signature matrix M. As a result, the constraint BNC imposed on OSP is relaxed and UVQTSP can be directly applied to multispectral image classification.

After all clusters are generated by VQ, an orthogonalization process is necessary to eliminate the correlation between $\mathbf{d}$ and U to prevent the bias. In UVQTSP, we employ an algorithm proposed by Chen et al. in [12] to orthogonalize the desired signature with respect to all cluster centers. Since the target signature is orthogonal to all clusters, there is no need of using a target signature projector $P_{\mathbf{d}}$ as described in TSC to remove or suppress noise and interferences. The only required subsequent operation is the matched filter $M_{\mathbf{d}}$ to extract the target signature $\mathbf{d}$. In terms of mathematical operators, UVQTSP can be expressed as follows:

$$P_{UVQTSP} = M_{\mathbf{d}} O_{LS}(VQ), \tag{12}$$

where VQ is a vector quantizer obtained by the LBG algorithm [13] and O_{LS} is an operator carried out by the Chen et al.'s algorithm. It is worth noting that no undesired signature annihilator $P_U^{\perp}$ or $P_{\mathbf{d}}$ is required in Eq. (12) as they appear in Eqs. (8) and (11) for OSP and TSC. This is because the target signature is already orthogonal to U.

In the following, we briefly describe the LBG algorithm and a least-squares-based orthogonalization algorithm O_{LS}.

3.1. Unsupervised vector quantization

Suppose that the number of clusters is p. The LBG algorithm is used to find an optimal set of p clusters based on the k-means nearest-neighbor criterion. The LBG algorithm is an iterative algorithm, also known as the Generalized Lloyd algorithm.

The LBG algorithm starts with an initial code book, $\text{Code}^1 = \{\mathbf{x}_j^1\}_{j=1}^p$ which will be generated using the Katsavounides et al.'s algorithm [14]. The data are then partitioned into clusters using the k-means nearest-neighbor rule. If a tie occurs, the data point is assigned to the cluster corresponding to the vector with the lowest index. After the image data have been partitioned, the center or centroid of each cluster set is calculated using the following Eq. (13).

$$\mathbf{x}_j = E[X \mid X \in R_j], \tag{13}$$

where X is a random vector used to describe the entire image and R_j is the jth cluster determined by the previous codebook to update and generate the next codebook. If an empty region is encountered, the center cannot be calculated by Eq. (13). In this case a code vector is assigned to the centroid of the cluster containing the most data points. These steps are repeated until a stopping criterion is satisfied. Here, the stopping criterion is chosen to be the squared Euclidean distance between the data vector and its nearest cluster center. A summarized LBG algorithm is given below.

Step 1. Initialization: $\text{Code}^1 = \{\mathbf{x}_j^1\}_{j=1}^p$ where $\{\mathbf{x}_j^1\}_{j=1}^p$ is a set of initial clusters generated by the Katsavounides et al.'s algorithm [14].

Step 2. Iterative procedure for reclustering at step $i > 1$ to generate the ith code book $\text{Code}^i = \{\mathbf{x}_j^i\}_{j=1}^p$:

$$\mathbf{x}_j^i = E[X \mid X \in R_j^{i-1}], \tag{14}$$

where R_j^{i-1} is the jth cluster produced by the code book at step $i-1$, $\text{Code}^{i-1} = \{\mathbf{x}_j^{i-1}\}_{j=1}^p$.

Step 3. Stopping rule: The reclustering will be terminated when no more data vectors are shuffled from one cluster to another. That is, as the algorithm iterates the squared Euclidean distance between data vectors and their nearest cluster centers will be reduced until there is no change in the codebook. As a result, either the squared Euclidean distances in two consecutive iterations remain unchanged or their difference is below a prescribed threshold. In either case, no data vector will be shuffled.

Since the target signature $\mathbf{d}$ is considered to be a fixed cluster, it should be designated as a code vector. So, only $p-1$ clusters are needed to be generated by VQ. The above LBG algorithm is modified for this application. It should be pointed out that $\mathbf{d}$ will not be updated during iterations in the VQ process because it is always a desired cluster. In other words, the initial code book had the desired signature as the first vector in the code book. This vector was retained during the course of updating code books. This allows one of the clusters to always contain the target signature. Let $C = \{C_2, C_3, \ldots, C_p\}$ be the set of the resulting $p-1$ clusters generated by VQ. Assume that $\mathbf{X} = (\mathbf{x_2}, \mathbf{x_3}, \ldots, \mathbf{x_p})$ is the undesired signature matrix

where $\mathbf{x_k}$ is the center of cluster C_k. Then $\mathbf{d}$ and $\mathbf{X}$ form a signature matrix M

$$M = (\mathbf{dX}) = (\mathbf{dx_2 x_3} \ldots \mathbf{x}_p). \tag{15}$$

Here, $\mathbf{X}$ corresponds to U in Eq. (2) except that $\mathbf{d}$ is rearranged in the first component for the purpose of the subsequent orthogonalization procedure, that is, $\mathbf{m}_1 = \mathbf{d}$ and $\mathbf{m}_k = \mathbf{x}_{k-1}$ for $k = 2, \ldots, p$

3.2. Least-squares-based orthogonalization algorithm

The operator O_{LS} in Eq. (12) is a least-squares-based orthogonalization procedure derived from the classical Gram–Schmidt method. It was originally developed for a radial basis function (RBF) neural network to select an appropriate subset of centers from a large number of training samples. In this paper, this algorithm is used for a different purpose to orthogonalize the target signature $\mathbf{d}$ with respect to $\mathbf{X}$ in Eq. (15). It decomposes M in Eq. (15) into a form $M = WA$ where A is a $p \times p$ triangular matrix with ones on the diagonal line and zeros below the diagonal line and W is an $l \times p$ matrix with l being the number of spectral bands. The W can be further expressed as $W = [\mathbf{w}_1 \mathbf{w}_2 \ldots \mathbf{w}_p]$ where $\{\mathbf{w}_k\}_{k=1}^{p}$ is a set of orthogonal column vectors such that

$$W^T W = H \tag{16}$$

and H is a $p \times p$ diagonal matrix with the kth element h_k given by $h_k = \mathbf{w}_k^T \mathbf{w}_k$. The orthogonalization procedure can be described as follows.

Step 1. $\mathbf{w}_1 = \mathbf{v}_1$.

Step 2. $a_{ik} = \dfrac{\mathbf{w}_i^T \mathbf{v}_k}{\mathbf{w}_i^T \mathbf{w}_i}$ for $1 \leqslant i < k$ and $k = 2, \ldots, p$.

Step 3. $\mathbf{w}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} a_{ik} \mathbf{w}_i$.

More details of implementing this algorithm can be found in Ref. [12].

4. Experimental results

In this section, we conduct experiments using hyperspectral and multispectral images to illustrate advantages of UVQTSP over OSP. In the first experiment, a HYDICE image is used to show that when the complete knowledge about undesired signatures is not available, UVQTSP performs better than OSP because UVQTSP also generates interferers as part of undesired signatures. In the second experiment, a SPOT image is used to show that when BNC is violated, OSP fails but UVQTSP performs reasonably well.

4.1. Hyperspectral image: HYDICE data

The HYDICE data used in the experiments are an image scene taken in Maryland in August 1995 using 210 bands with spectral coverage 0.4–2.5 μm of resolution 10 nm. The ground sampling distance (GSD) generally ranges from 1 to 4 m. However, for illustration purpose we used an image scene acquired by a low flight altitude and the GSD is approximately 0.78 m. A 160×160

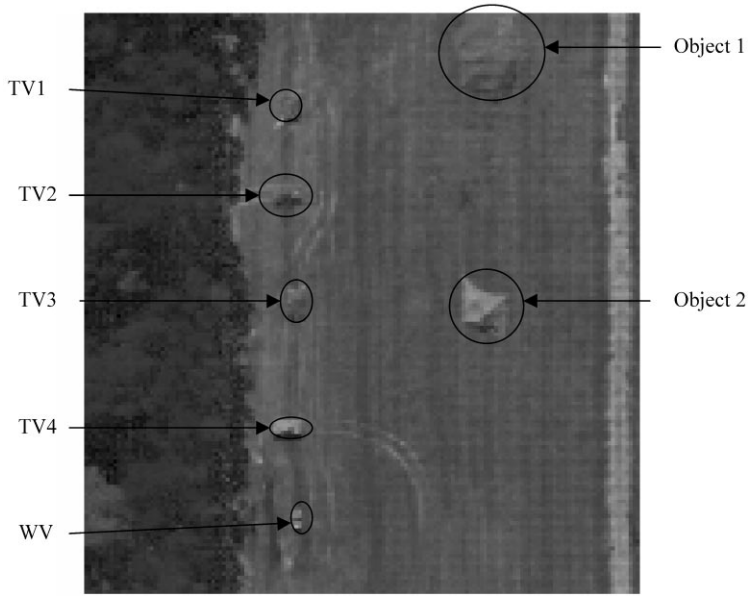

Fig. 1. A scene of HYDICE image.

subimage was cropped from this image and a single-band enlarged image is shown in Fig. 1. This figure shows a scene with trees along the left one-third and a grass field in the right two-thirds. This grass field also contains a road running along the right edge of the image, five vehicles along the tree line and two types of objects in the field. The five vehicles are four treaded vehicles in the upper half of the scene and one wheeled vehicle in the lower half of the image. The size of treaded vehicles is approximately 4×8 m and the size of the wheeled vehicle is about 3×6 m. The first object is at the top edge of the scene while the second object is located in the middle of the scene. All these seven man-made targets are circled in the image and denoted by TV1, TV2, TV3, TV4, WV, Objects 1 and 2. The following experiment is designed to demonstrate a situation that when only partial knowledge of undesired signatures is available, the performance of the OSP will be degraded by interference which is not considered to be part of undesired signatures. In this case, UVQTSP outperforms OSP by eliminating the clusters which correspond to interferers generated by VQ process.

Four signatures were obtained for the experiments, one for the treaded vehicles, one for the wheeled vehicle, one for object 1 and one for object 2. These signatures were obtained by extracting pixels directly from the scene. OSP and UVQTSP were then applied to the data using these four signatures. The results of these tests are shown in Fig. 2 for OSP, and Fig. 3 for UVQTSP, where additional clusters were generated by VQ. As shown in these figures OSP detected the objects in the scene, but UVQTSP clearly outperformed OSP. The image labeled treaded vehicles of Fig. 2 shows that OSP detected these vehicles, but significant amounts of energy from the trees and grass in the scene were also detected. This shows that OSP could not eliminate the energy from these interferers

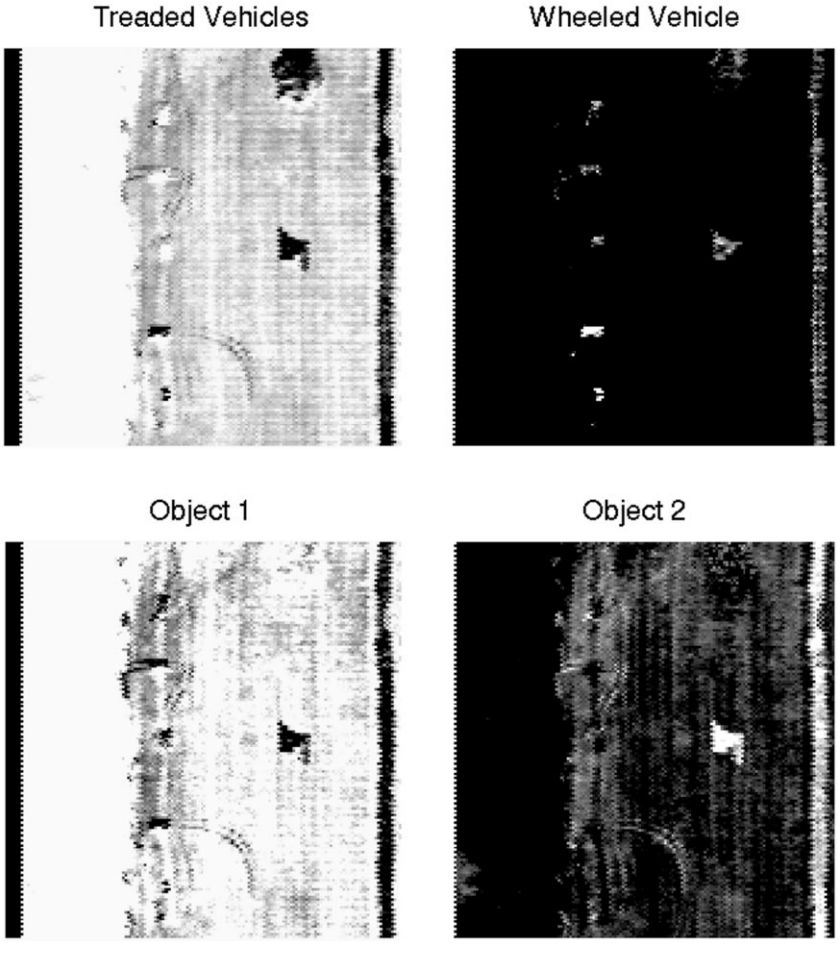

Fig. 2. Images resulting from the OSP applied to Fig. 1.

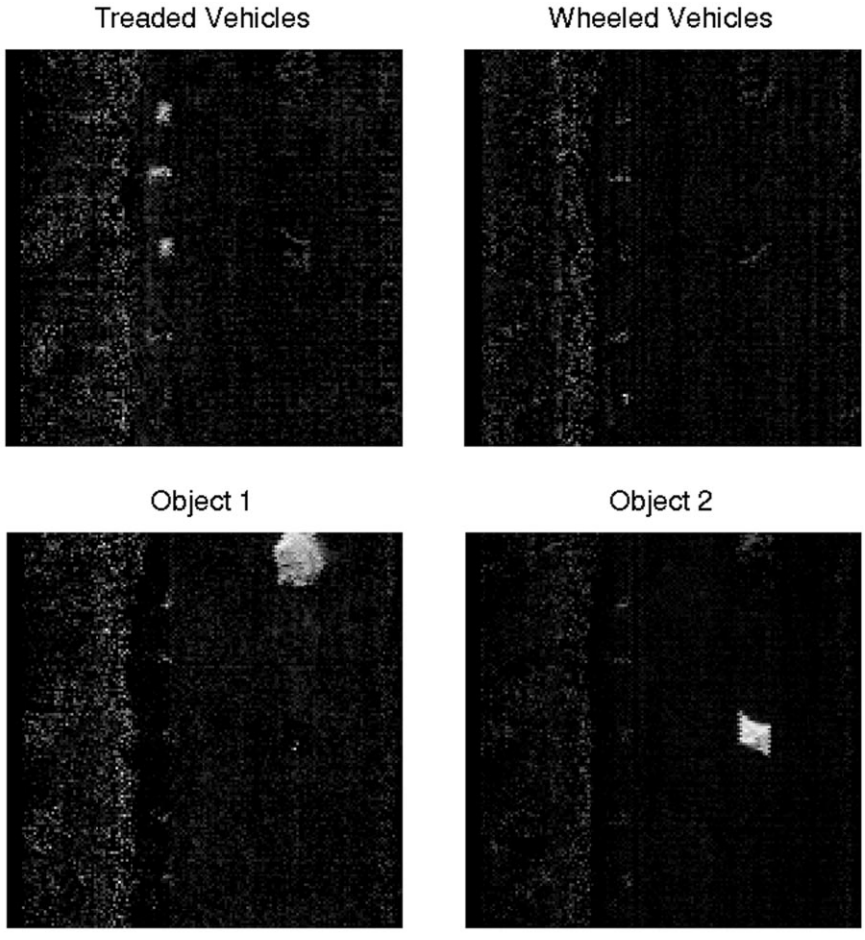

Fig. 3. Images resulting from the UVQTSP applied to Fig. 1.

because no signature was given for these objects. UVQTSP did not have this problem because the VQ produced clusters that corresponded to these signatures and these energies were eliminated subsequently. Similar results were also obtained for UVQTSP and are shown in the other images in Fig. 3. However, all of the OSP-geneated images in Fig. 2 displayed some interference. The wheeled vehicle had interference from the fourth treaded vehicle, the road and object 2. Object 1 had interference from the trees and grass. Object 2 had interference from the road. The signatures for the grass, the trees and the road could have been added to OSP to eliminate these interferers, but this would have required more knowledge of the scene than was assumed. UVQTSP did not require this additional knowledge, but obtained it through the VQ algorithm. Fig. 3 clearly shows the strength of UVQTSP method. Hyperspectral images typically have low signal to interference ratio and it is more important to eliminate the interference in the scene than the noise. UVQTSP used VQ to generate a set of clusters corresponding to interferers. These interferers are then eliminated. In each of the UVQTSP-generated images in Fig. 3, the interferers that were seen in the corresponding image in Fig. 2 were eliminated. This demonstrates an advantage of UVQTSP over OSP. If the clusters generated by the VQ were used together with the four signatures for OSP, both OSP and UVQTSP would produce similar results. The number of clusters to be generated by the VQ algorithm was determined by Eq. (17)

$$\eta = \mathbf{d}^{\mathbf{T}} P_U^{\perp} \mathbf{d}. \tag{17}$$

The variable $\mathbf{d}$ is the desired signature, and the U is the matrix of the signatures generated by the VQ algorithm. η is a measure of the energy corresponding to $\mathbf{d}$ in the

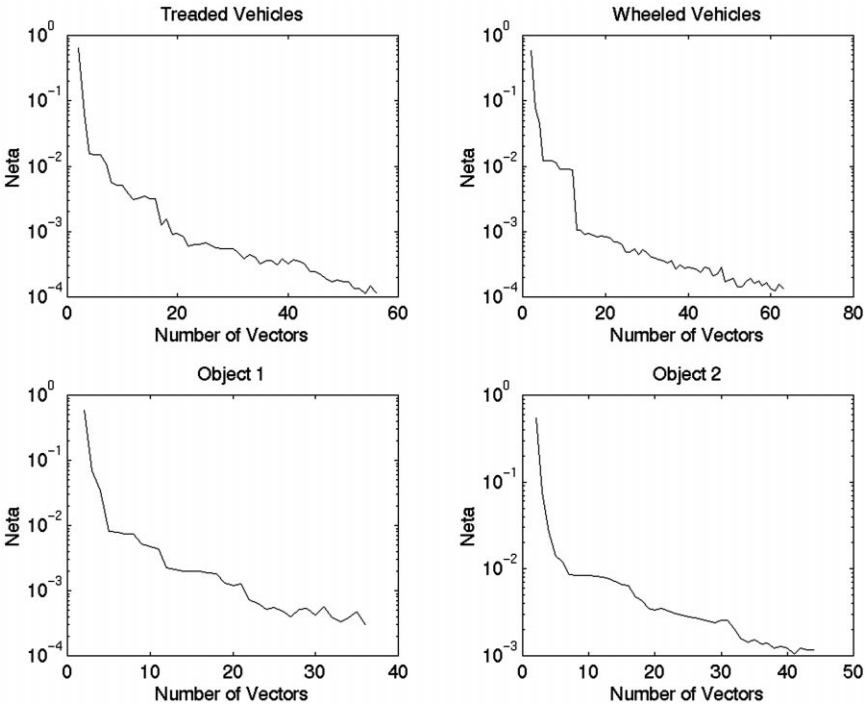

Fig. 4. Rank curves of η versus the number of clusters.

subspace that is orthogonal to the undesired signatures. The VQ algorithm was allowed to add a new cluster until the value of η was less than 10^{-4}. Fig. 4 shows the rank curve of η versus the number of clusters. The treaded vehicles had 54 clusters, the wheeled vehicles had 41 clusters, object 1 had 37 clusters and object 2 had 45 clusters.

It should be noted that when we classified the treaded vehicle, the fourth treaded did not show up. Instead, it was detected when the wheeled vehicle were classified as shown in Fig. 3. This is because the spectrum of the fourth treaded vehicle is very similar and closer to that of the wheeled vehicle than that of the other three treaded vehicles. A measure for spectral similarity defined below, called the spectral divergence provides explanation [15].

The idea of the spectral divergence is to normalize the spectra $\mathbf{m_1}$ and $\mathbf{m_2}$ to unity so that each spectrum can be viewed as a histogram described by a probability mass function (PMF) defined on the spectral range. Let $\boldsymbol{p} = \{p_i\}_{i=1}^{l}$ and $\boldsymbol{q} = \{q_i\}_{i=1}^{l}$ be the resulting PMFs of these two histograms. From information theory, a criterion called *divergence* can be used to measure the discrepancy between any two probability distributions. By means of the divergence, the spectral information divergence between p and q, denoted by $\mathrm{SID}(\boldsymbol{p},\boldsymbol{q})$ can be defined as

$$\mathrm{SID}(\boldsymbol{p},\boldsymbol{q}) = L(\boldsymbol{p};\boldsymbol{q}) + \mathrm{L}(\boldsymbol{q};\boldsymbol{p}), \tag{18}$$

where $D(\boldsymbol{p}\,\|\,\boldsymbol{q})$ and $D(\boldsymbol{q}\,\|\,\boldsymbol{p})$ are defined, respectively, by

$$D(\boldsymbol{p}\,\|\,\boldsymbol{q}) = \sum_{\mathrm{i}=1}^{1} p_i \log\frac{p_i}{q_i} \tag{19}$$

and

$$D(\boldsymbol{q}\,\|\,\boldsymbol{p}) = \sum_{i=1}^{1} q_i \log\frac{q_i}{p_i}. \tag{20}$$

Note that $D(\boldsymbol{p}\,\|\,\boldsymbol{q})$ is generally referred to as relative entropy, directed divergence, cross entropy or Kullback–Leibler distance between p and q. From Eq. (18), the spectral information divergence can be interpreted as a similarity measure between two spectra. The closer this spectral information divergence is to zero, the more similar the two signatures are to each other. The values of the spectral information divergence of the all target signatures (four treaded vehicles, one wheeled vehicle and two objects) in Fig. 1 were calculated and tabulated in Table 1. As shown in Table 1, the spectrum of the fourth vehicle is much closer to that of the wheeled vehicle than to that of the other three treaded vehicles.

Table 1
Values of spectral information divergence calculated for treaded vehicles, wheeled vehicle, objects 1 and 2 in Fig. 1

	TV1	TV2	TV3	TV4	WV	Obj1	Obj2
TV1	0	0.0067	0.0076	0.0411	0.0363	0.0430	0.0766
TV2		0	00025	0.0522	0.0464	0.0532	0.0891
TV3			0	0.0619	0.0592	0.0684	0.1086
TV4				0	0.0066	0.0197	0.0679
WV					0	0.0092	0.0393
Obj1						0	0.0289
Obj2							0

4.2. Multispectral image: SPOT data

The image data to be used in the experiments were collected by the SPOT system from three bands, two of which are from the visible region of electromagnetic spectrum referred to as band 1 (0.5–0.59 μm) and band 2 (0.61–0.68 μm), and the third band is from the near-infrared region of electromagnetic spectrum referred to as band 3 (0.79–0.89 μm). The ground sampling distance is 20 m. These three images shown in Fig. 5 were registered and combined into an image cube where each pixel is represented by a 3×1 column vector with each component corresponding to one band of the SPOT data. The experiment given below is a four-signature classification using 3 band images. In this case, the constraint BNC is violated. As expected, OSP performs poorly. On the contrary, UVQTSP works reasonably well.

The first set of pixels was extracted from the image scene corresponding to Falls Church High School. The sample average of 9 pixels was used to generate a signature of a building. A second set of 9 pixels was extracted from the scene and corresponded to the Little River Turnpike. Similarly, the sample average was used to generate a signature of a road. A third set of 9 pixels was extracted from the scene to represent the lake in the upper right of the image. Again, the sample average of 9 pixels was used to generate a signature of the water. A fourth set of 9 pixels was extracted from the scene to indicate the Mill Creek Park with the sample average used to generate a signature of vegetation. All these four signatures were normalized and shown in Fig. 6 to form the signature matrix M. OSP was then applied to classify the image with each of the four signatures used for as the desired signature. Since there are three bands in the SPOT data, only two of the remaining three signatures can be considered as undesired signatures. The fourth remaining signature, neither a desired or undesired signature is considered to be an interferer. It should be noted that if all the three remaining signatures were used as undesired signature, OSP would not have enough dimensions to project in a subspace orthogonal to all the three.

Fig. 5. A scene of SPOT image.

Three OSP classifiers, one with each of the remaining signatures as an interferer, were generated for each of the four signatures. These 12 OSP classifiers were applied to the data and the results are shown in Fig. 7. The rows in Fig. 7 contain the images generated using the labeled signatures as the desired signature. Thus, rows contain the results of the three OSP classifiers with the labeled signature as the desired signature. The columns in Fig. 7 contain the images generated by three OSP classifiers using the labeled signature as the interfering signature. For example, the first row contains the results of the three OSP classifiers with buildings as the desired signature and the first column contains the results of the three OSP classifiers with buildings as the interfering signature. The first row of Fig. 7 shows that the OSP classifier

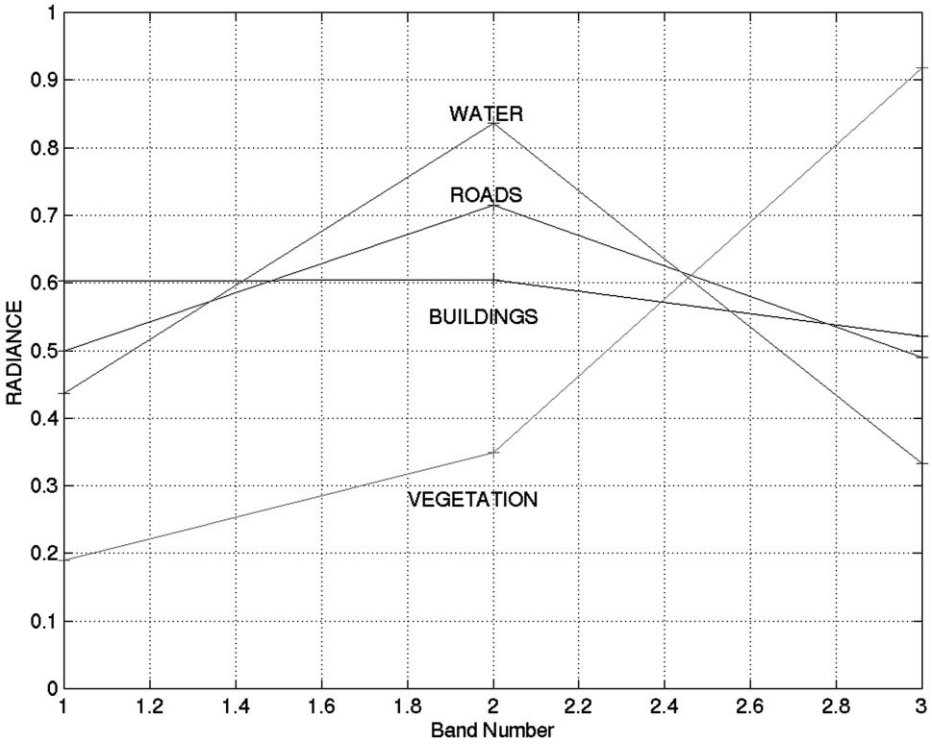

Fig. 6. Four normalized signatures for buildings, roads, water and vegetation.

using the building signature as the desired signature and the vegetation signature as the interfering signature gave the best results. This classifier eliminated the undesired signatures, roads and water, and the interfering signature was different enough from the desired signature that the interfering effect did not have as significant impacts on the performance of the matched filter as did the roads and water. The second row of Fig. 7 shows that neither of the OSP classifiers using the road signature as the desired signature could not eliminate the effects of the remaining interfering signature. The first OSP classifier with the road signature as the desired signature did detect the roads, but also detected the buildings (the interfering signature) in the image as roads. The second OSP classifier with the road signature as the desired signature did detect the roads, but also detected the bodies of water (the interfering signature) in the image as roads. The third OSP classifier with the road signature as the desired signature and with vegetation as the interfering signature was a complete failure. This is because the road signature is very close to both water and building signatures shown in Fig. 6. From Table 2 where the values of spectral divergence of the four signatures in Fig. 6 were calculated, the spectral information divergence between road and building is 0.0228 and the spectral information divergence between road and water 0.0498. They are very small compared to the spectral information divergence between road and vegetation, 0.5020. Thus, when the OSP projects an input pixel into an orthogonal subspace to eliminate the undesired signatures, it also eliminates much of the energy associated with the road signature. In this subspace much of the energy is associated with the interfering vegetation signature and this interfering energy is causing a bias as seen in the TSC. Unlike the TSC, this bias is from the energy not being eliminated. This can be seen by observing that there is no bias where the building and water appear in the image. This lack of a bias is because this energy has been eliminated. The third row of Fig. 7 shows that the OSP classifier using the water signature as the desired signature and the building signature as the interfering signature gave the best results. This classifier eliminated the undesired signatures, roads and vegetation, and the interfering signature was different enough from the desired signature that the matched filter could effectively extract water without much the interfering effects. Finally, the fourth row shows that all three classifiers using the vegetation signature as the desired signature produced good results. The reason for all three OSP classifiers performing well can be seen from Table 2 where the vegetation signature is

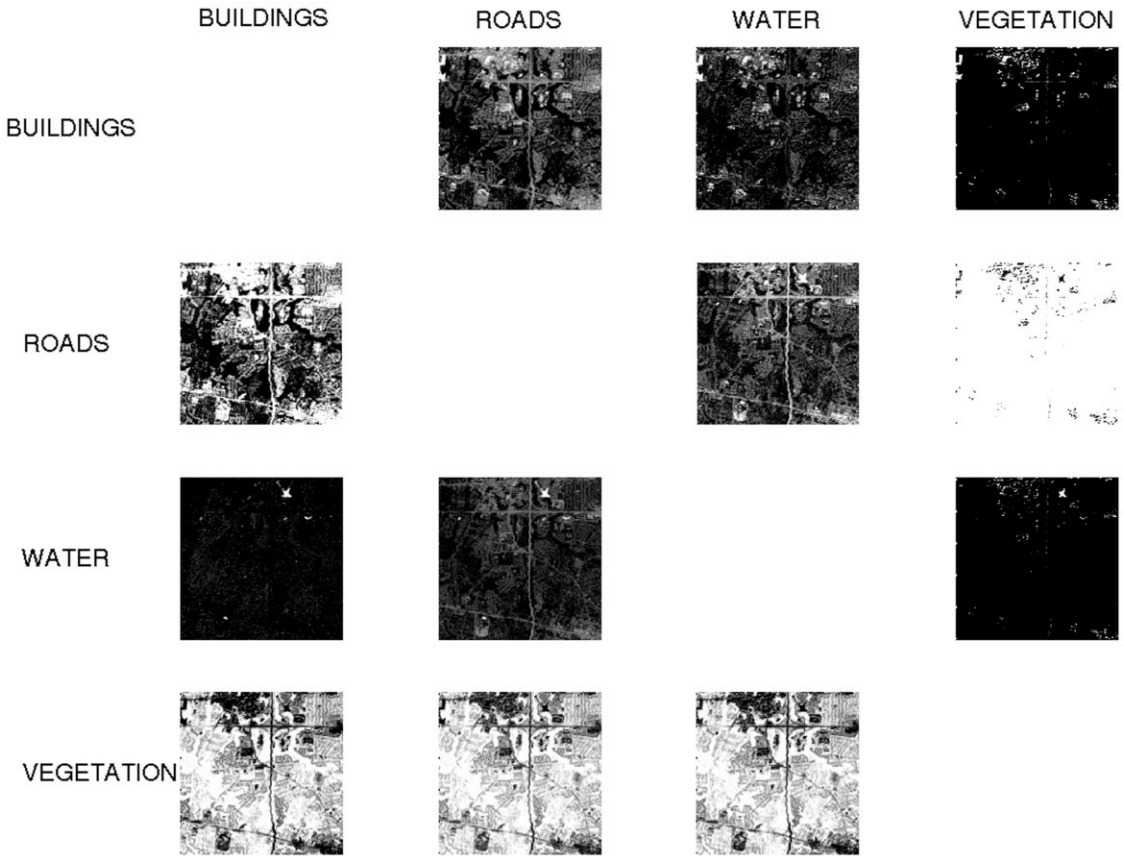

Fig. 7. Images resulting from the OSP applied to Fig. 5.

Table 2
Values of spectral information divergence calculated for any two spectra in Fig. 6

	Buildings	Roads	Water	Vegetation
Buildings	0	0.0228	0.1223	0.04993
Roads		0	0.0498	0.5020
Water			0	0.7922
Vegetation				0

distinctive enough that the matched filter can eliminate any remaining interference from the interfering signature. These images show that three bands are not sufficient for the OSP classifiers to perform in general. These images also show that in cases that the number of desired and undesired signatures is greater than the number of bands additional knowledge about the signatures is needed to optimize performance.

UVQTSP was applied to the same image with each of four signatures as a target signature. Unlike the OSP which requires three undesired signatures UVQTSP used VQ to generate two clusters corresponding to unknown undesired signatures. The results are shown in Fig. 8 where each of the four images is labeled by the target signature used in UVQTSP. Comparing Fig. 8 against Fig. 7, the results produced by UVQTSP are comparable to the best OSP-generated images except the case of water where some energy of the road was detected. However, in order for the OSP to outperform UVQTSP, a total of 12 OSP classifiers are required to be implemented to select the best results. Additionally, the best results are not necessarily generated by the same OSP classifier. This demonstrates that when no prior knowledge of undesired signatures is available, UVQTSP took advantage of unsupervised nature to generate the two most significant interferers and then effectively eliminated them to achieve the best possible classification.

As a concluding remark, since the OSP and UVQTSP classifiers specified by Eqs. (8) and (12) are basically abundance estimators, the images produced by OSP and UVQTSP are gray scaled where the gray scale of each pixel is actually the estimated target abundance.

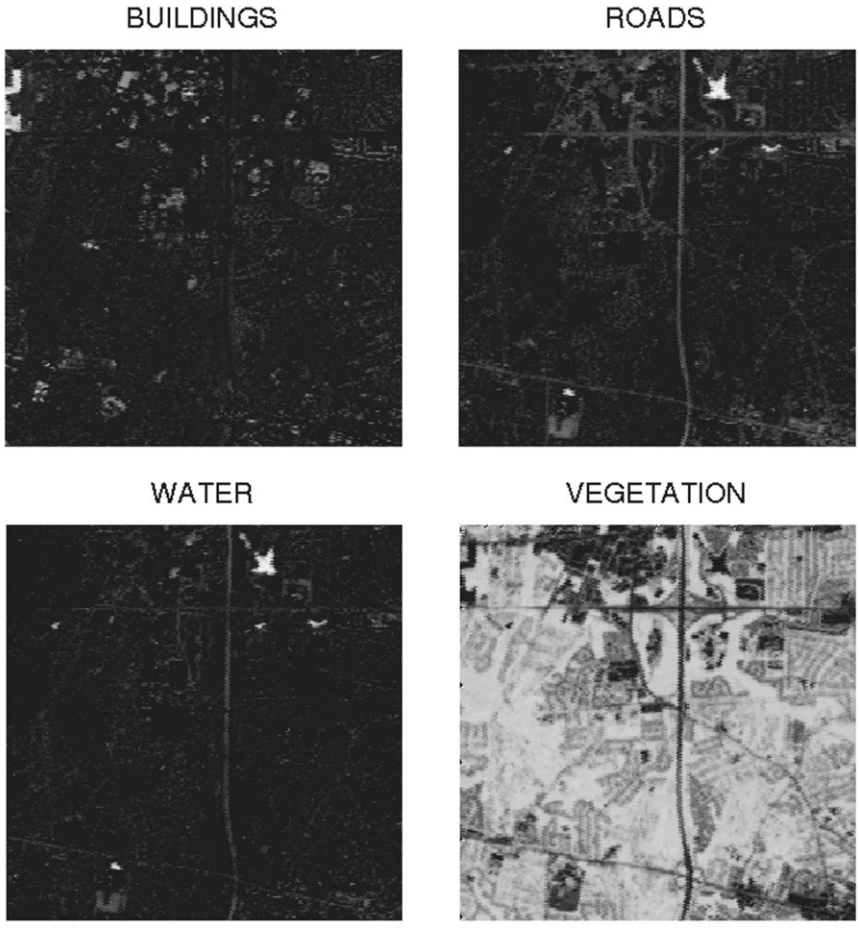

Fig. 8. Images resulting from the UVQTSP applied to Fig. 5.

Consequently, their classification performance is determined by accuracy of target abundance estimation. This can be done by visual examination. For a quantitative analysis we refer to Ref. [16].

5. Conclusion

Target signature classification and detection with limited knowledge of ground truth generally presents some difficulty for real data since ground truth may be either impossible to obtain or difficult or costly to obtain if possible. UVQTSP offers a solution to this problem. Although UVQTSP described in this paper required only information about the target signature, this requirement can be further relaxed by setting the quantity η given by Eq. (17) to a prescribed threshold [7]. In this situation, the unknown target will be one of the clusters produced by VQ and the η will determine the number of clusters needed to be generated. Several advantages resulting from UVQTSP are significant. It makes no assumption that all signatures must be known a priori. It also circumvents the constraint that the number of signatures to be classified is no greater than that of spectral bands, an inherent limitation for OSP. More importantly, UVQTSP generates potential interferers in addition to undesired signatures and eliminates them prior to detection and classification. As a result, when UVQTSP is applied to hyperspectral image data, it outperforms OSP, SSC and TSC which do not take interference into account. More experiments using 224-band AVIRIS data were also implemented and reached similar conclusions [7].

Acknowledgements

The authors would like to the Spectral Information Technology Applications Center for providing

hyperspectral data set used for experiments in this paper. They also thank Dr. Shin-Yi Hsu with Susquehanna Resources and Environment, Inc. for providing SPOT data.

References

[1] G. Vane, A.F.H. Geotz, Terrestrial imaging spectroscopy, Remote Sensing of Environ. 24 (1988) 1–29.
[2] J.C. Harsanyi, C.-I Chang, Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach, IEEE Trans. Geosci. Remote Sensing GE-32 (1994) 779–785.
[3] J.C. Harsanyi, Detection and classification of subpixel spectral signatures in hyperspectral Image sequences, Ph.D. Dissertation, Department of Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, 1993.
[4] DARPA Spectral Exploitation Workshop, Defense Advanced Research Projects Agency, Annapolis, MD, 1–2 July 1997.
[5] Spectroradiometric Working Group, VA, 2–3 February, 1997.
[6] Spectroradiometric Science Symp. 2–6 November, San Diego, CA, 1997.
[7] C. Brumbley, An unsupervised Kalman Filtering and subspace projection approaches to multispectral and hyperspectral image classification and detection. Ph.D. Dissertation, Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, 1998.
[8] C.-I Chang and C. Brumbley, A Kalman filtering approach to multispectral image classification with signature abundance estimation, IEEE Trans. on Geoscience and Remote Sensing. Vol 37, No. 1, Jan. 1999.
[9] T.M. Tu, C.H. Chen, C.-I Chang, A posteriori least squares orthogonal subspace projection approach to weak signature extraction and detection, IEEE Trans. Geosci. Remote Sensing GE-35 (1997) 127–139.
[10] C.-I Chang, X. Zhao, M.L.G. Althouse, J.J. Pan, Least squares subspace projection approach to mixed pixel classification for hyperspectral images, IEEE Trans. Geosci. Remote Sensing GE-36 (1998) 898–912.
[11] X. Zhao, Subspace Projection to multispectral/hyperspectral image classification using linear spectral mixture modeling, M.S. Thesis, Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, 1996.
[12] S. Chen, C.F.N. Cowan, and P.M. Grant, Orthogonal least squares learning algorithm for radial basis function networks, IEEE Trans. Neural Networks NN-2 (1991) 302–309.
[13] Y. Linde, A. Buzo, R.M. Gray, An algorithm for vector quantizer design, IEEE Trans. Commun. COM-28 (1980) 84–95.
[14] I. Katsavounides, C.C.J. Kuo, Z. Zhang, A new initialization technique for generalization Lloyd iteration, IEEE Trans. Signal Process. Lett. 1 (1994), 144–146.
[15] C.-I. Chang, C. Brumbley, Linear unmixing Kalman filtering approach to signature abundance detection, signature estimation and subpixel classification for remotely sensed images, IEEE Trans. Aerospace and Electronic Systems, Vol. 37, No. 1, January 1999.
[16] C.-I Chang, H. Ren, An experiment-based quantitative and comparative analysis of hyperspectral target detection and image classification algorithms, submitted to IEEE Transaction on Geoscience and Remote Sensing, 1998 (revised).

About the Author—CLARK M. BRUMBLEY received his BSEE degree from Old Dominion University, 1985, MSEE and Ph.D. from the University of Maryland Baltimore County in 1993 and 1998, respectively. He is currently a Senior Electronic Engineer for the Department of Defense at Fort Meade, MD. His research interests include signal detection and estimation, multispectral/hyperspectral image processing. Dr. Brumbley is a member of IEEE, Tau Beta Pi, Eta Kappa Nu and Sigma Xi.

About the Author—CHEIN-I CHANG received his BS, MS and MA degrees from Soochow University, Taipei, Taiwan, 1973, the Institute of Mathematics at National Tsing Hua University, Hsinchu, Taiwan, 1975 and the State University of New York at Stony Brook, 1977, respectively, all in mathematics, and MS and MSEE degrees from the University of Illinois at Urbana-Champaign in 1982 respectively and Ph.D. in electrical Engineering from the University of Maryland, College Park in 1987. He was a visiting assistant professor from January 1987 to August 1987, assistant professor from 1987 to 1993, and is currently an associate professor in the Department of Computer Science and Electrical Engineering at the University of Maryland Baltimore County. He was a visiting specialist in the Institute of Information Engineering at the National Cheng Kung University, Tainan, Taiwan from 1994–1995. His research interests include information theory and coding, signal detection and estimation, multispectral/hyperspectral image processing, neural networks, pattern recognition. Dr. Chang is a senior member of IEEE and a member of SPIE, INNS, Phi Kappa Phi and Eta Kappa Nu.