An Experiment-Based Quantitative and Comparative Analysis of Target Detection and Image Classification Algorithms for Hyperspectral Imagery

Chein-I Chang, *Senior Member, IEEE,* and Hsuan Ren, *Student Member, IEEE*

***Abstract*—Over the past years, many algorithms have been developed for multispectral and hyperspectral image classification. A general approach to mixed pixel classification is linear spectral unmixing, which uses a linear mixture model to estimate the abundance fractions of signatures within a mixed pixel. As a result, the images generated for classification are usually gray scale images, where the gray level value of a pixel represents a combined amount of the abundance of spectral signatures residing in this pixel. Due to a lack of standardized data, these mixed pixel algorithms have not been rigorously compared using a unified framework. In this paper, we present a comparative study of some popular classification algorithms through a standardized HYDICE data set with a custom-designed detection and classification criterion. The algorithms to be considered for this study are those developed for spectral unmixing, the orthogonal subspace projection (OSP), maximum likelihood, minimum distance, and Fisher's linear discriminant analysis (LDA). In order to compare mixed pixel classification algorithms against pure pixel classification algorithms, the mixed pixels are converted to pure ones by a designed mixed-to-pure pixel converter. The standardized HYDICE data are then used to evaluate the performance of various pure and mixed pixel classification algorithms. Since all targets in the HYDICE image scenes can be spatially located to pixel level, the experimental results can be presented by tallies of the number of targets detected and classified for quantitative analysis.**

***Index Terms*—Linear discriminant analysis (LDA), linear unmixing, maximum likelihood estimator (MLE), minimum distance, mixed-to-pure pixel (M/P) converter (M/P converter), oblique subspace projection (OBSP), orthogonal subspace projection (OSP), signature space projection (SSP), winner-take-all M/P converter (WTAMPC).**

I. Introduction

IMAGE classification is a segmentation method that aggregates image pixels into a finite number of classes by certain rules so that each class represents a distinct entity with specific properties [1]. In general, it can be viewed as a label assignment by which image pixels sharing similar properties will be assigned to the same class. Since multispectral images are acquired at different spectral wavelengths, a multispectral image pixel can be represented by a pixel vector, in which each component corresponds to a specific wavelength. As a result, criteria used for multispectral image classification are usually designed to explore spectral characteristics rather than spatial properties, as used in digital image processing [2]–[5]. A unique feature of multispectral image classification that does not exist in standard image processing is the occurrence of spectral mixtures within pixels. Spectral unmixing is particularly important with high spectral resolution imaging spectrometers. These sensors use as many as 200 contiguous bands and can uncover narrow-band diagnostic spectral features of materials that cannot be resolved by multispectral imagers. Two such important imagers currently in use are the NASA Jet Propulsion Laboratory's 224-band Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) and the Naval Research Laboratory's 210-band HYperspectral Digital Imagery Collection Experiment (HYDICE) sensor. One of major challenges in hyperspectral image processing is how to process the enormous amount of information provided by hyperspectral images without spending effort on undesired/unwanted information [6]. Additionally, the data dimensionality of hyperspectral imagery is generally tens of times more than that of multispectral imagery. As a consequence, methods developed for multispectral image processing such as principal components analysis/canonical analysis [7], minimum distance [1], maximum likelihood (ML) classification [8]–[13], and decision boundary-based feature extraction [14] can be further improved for hyperspectral imagery.

Harsanyi and Chang [15], [16] introduced an orthogonal subspace projection (OSP)-based classifier for hyperspectral image classification. It implemented an orthogonal subspace projector in conjunction with a matched filter to derive a classifier for mixed pixel classification. It has been successfully applied for HYDICE data exploitation [17]–[19]. A variety of OSP-based classifiers were also developed, such as the *a posteriori* OSP (LSOSP) classifier [20], the oblique subspace projection classifier (OBC) [21], the desired target detection and classification algorithm (DTDCA) and the automatic target detection and classification algorithm (ATDCA) [22]. In particular, the OSP-based methods were also shown in [21], [23], [24] to be equivalent to the maximum likelihood classifier, given that the noise is additive and Gaussian. So all of these classifiers turned out to perform the same spectral unmixing.

There is a lack of standardized data that can be used to evaluate individual algorithms. In addition, no unified criterion has been accepted for rigorous and impartial comparisons. The importance of this issue cannot be understated. Without standardized data and effective evaluation criteria,

Manuscript received June 12, 1998; revised December 9, 1999.

The authors are with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250 USA (e-mail: cchang@umbc.edu).

Publisher Item Identifier S 0196-2892(00)02835-7.

the performance of any new algorithm cannot be substantiated. In this paper, we take a first step by conducting a comparative study of performance analysis among several classification algorithms. We confine our study to linear spectral mixing problems only. Additionally, we consider two types of classification: mixed pixel classification and pure pixel classification. A general approach to mixed pixel classification (such as spectral unmixing) is to estimate the abundance fraction of a material of interest present in an image pixel, and then the estimated abundance fraction is used to classify the pixel. However, this generally requires visual interpretation. Such human intervention is rather subjective and may not be reliable or repeatable. With no availability of standardized data or objective criteria, a quantitative analysis for mixed pixel classification is almost impossible. By contrast, pure pixel classification does not have such a problem. Unlike mixed pixel classification, it does not require abundance fractions of spectral signatures to be used for class assignment. Its performance is completely determined by the criteria used for classification. So, two major contributions of this paper are 1) to establish a link between pure and mixed pixel classification by designing a mixed-to-pure pixel (M/P) converter and 2) to conduct experimental comparisons among a set of selected pure and mixed classification algorithms, including quantitative performance analysis. In order to validate such a study, a standardized HYDICE data set is used where all man-made targets present in image scenes have been precisely located to the pixel level and designated as either target center pixels or target masking pixels. The reason for using target masking pixels is to include partial target pixels, target background pixels, and target shadow pixels to account for all possible pixels that may have impacts on targets of interest. In addition, a custom-designed criterion for target detection and classification is also introduced for the purpose of tallying target pixels detected and classified. By making use of this data set, along with the designed criterion, a comparative analysis for classification accuracy becomes possible. The significance of these experimental results is to offer a performance evaluation of the classification algorithms in a rigorous fashion so that each algorithm is fairly compared on the same common ground.

A standardized HYDICE data set is used for evaluation. The experiments show that the OSP-based classification algorithms resulting from an M/P conversion perform better than the minimum distance-based classification algorithms, but not as well as LDA. On the other hand, the same experiments also show that the abundance-based images generated by mixed pixel classification algorithms significantly improve classification results. These facts substantiate the need for mixed pixel classification for multispectral/hyperspectral imagery.

This paper is organized as follows. Section II formulates the mixed pixel classification problem as a linear mixture model. Section III describes various approaches to abundance estimation for mixed pixel classification (e.g., OSP-based and ML classifiers). Section IV introduces the concept of mixed-to-pure pixel conversion to reduce a mixed pixel classification problem to a conventional pure pixel classification problem. Section V derives an objective criterion for target detection and classification to used for experiments. Section VI presents a comparative performance analysis for classifiers described in Sections III and IV, and Section VII concludes with some remarks.

II. Linear Mixing Problems and OSP Approach

Linear spectral unmixing is a widely used approach in remotely sensed imagery to determine and quantify individual components [25], [26]. Since every pixel is acquired by multiple spectral bands, it can be represented by a column vector where each component represents a particular band. Suppose that L is the number of spectral bands. Let $\mathbf{r}$ be an $L \times 1$ column vector in a multispectral or hyperspectral image where vectors are all boldfaced. In this case, each pixel is considered to be a pixel vector of dimension L. Assume that M is an $L \times p$ signature matrix denoted by $(\mathbf{m}_1\ \mathbf{m}_2\ \cdots\ \mathbf{m}_p)$, where $\mathbf{m}_j$ is an $L \times 1$ column vector representing the j-th spectral signature resident in the pixel $\mathbf{r}$, and p is the number of signatures of interest. Let $\alpha = (\alpha_1\ \alpha_2\ \cdots\ \alpha_p)^T$ be a $p \times 1$ abundance column vector associated with $\mathbf{r}$, where α_j denotes the fraction of the j-th signature in the pixel $\mathbf{r}$.

A. Linear Spectral Mixture Model

A classical approach to solving the mixed pixel classification problem is linear unmixing, which assumes that the materials (endmembers) present in a pixel vector are linearly mixed. A pixel vector can be described by a linear regression model as follows:

$$\mathbf{r} = M\alpha + \mathbf{n} \tag{1}$$

where $\mathbf{n}$ is an $L \times 1$ column vector that can be viewed as either noise or an error correction term resulting from data fitting.

The algorithms to be used for our comparative study only include those derived from OSP, minimum distance approaches, and Fisher's linear discriminant analysis (LDA). This selection is made for three major reasons.

1) As mentioned earlier, if the noise in a linear mixing problem is white Gaussian, ML estimation and the OSP approach for mixed pixel classification are equivalent and both can be viewed as a spectral unmixing method.
2) The white Gaussian noise assumption also simplifies and reduces the Gaussian ML classifier to a minimum distance classifier.
3) Fisher's LDA has been widely used for classification since its criterion is based on the maximization of class separability.

These facts allow us to restrict the mixed pixel classification algorithms to three classes of classification algorithms listed above (the OSP-based classifiers, minimum distance-based classifiers, and LDA). The difference between the OSP and the other approaches (i.e., minimum distance, LDA) is that the OSP was designed for mixed pixel classification, whereas the latter is for pure pixel classification. Nevertheless, we will show that by imposing appropriate constraints on the abundance fractions, the mixed pixel classification can be reinterpreted and reduced to pure pixel classification. By means of a mixed-to-pure pixel (M/P) conversion, mixed pixel classification algorithms can then be directly compared with minimum distance-based classifiers and LDA.

B. Orthogonal Subspace Projection (OSP)

Without loss of generality, we assume that there is a signature of interest in model (1), $\mathbf{d} = \mathbf{m}_p$. So the signature matrix M can be partitioned into the desired signature vector $\mathbf{d}$ and an undesired signature matrix denoted by U. By separating $\mathbf{d}$ from U, model (1) can be expressed as follows:

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

where the subscript i is suppressed throughout this paper and $U = (\mathbf{m}_1\ \mathbf{m}_2\ \cdots\ \mathbf{m}_{p-1})$. Let $\langle \mathbf{d} \rangle$, $\langle M \rangle$ and $\langle U \rangle$ be the spaces linearly spanned by $\mathbf{d}$, U, and M respectively. The reason for separating U from M in model (2) is to allow us to design an orthogonal subspace projector to annihilate U from an observed pixel $\mathbf{r}$ prior to classification. One such desired orthogonal subspace projector was derived in [15] given by $P_U^{\perp} = I - UU^{\#}$, where $U^{\#} = (U^T U)^{-1} U^T$ is the pseudo-inverse of U and the notation $\frac{\perp}{U}$ indicates that the projector $P_U^{\perp}$ maps the observed pixel $\mathbf{r}$ into the range space $\langle U \rangle^{\perp}$, the orthogonal complement of $\langle U \rangle$.

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

$$P_U^{\perp}\mathbf{r} = P_U^{\perp}\mathbf{d}\alpha_p + P_U^{\perp}\mathbf{n} \tag{3}$$

where the undesired signatures in U vanish due to orthogonal projection elimination, and the original noise $\mathbf{n}$ has been suppressed to $P_U^{\perp}\mathbf{n}$.

Equation (3) represents a standard signal detection problem and can be solved by a matched filter $M_{\mathbf{d}}$ given by $M_{\mathbf{d}}(\mathbf{x}) = \mathbf{d}^T\mathbf{x}$. So, an orthogonal subspace projection (OSP) classifier q_{OSP}^T derived in [15] can be implemented by an undesired signature annihilator $P_U^{\perp}$, followed by a desired signature matched filter $M_{\mathbf{d}}$

$$q_{\mathrm{OSP}}^T = M_{\mathbf{d}} P_U^{\perp} = \mathbf{d}^T P_U^{\perp}. \tag{4}$$

III. Hyperspectral Abundance Estimation Algorithms for Mixed Pixel Classification

Equation (1) represents a general linear model for mixed pixel classification where the signature matrix M and the abundance vector α are assumed to be known *a priori*. In reality, α is generally not known and must be estimated. In order to estimate α, a common approach is spectral unmixing via an inverse of the linear mixture model given by (1) (e.g., [27]). In this paper, we will describe two general approaches in Sections III and IV, the estimation of abundance and the classification of abundance, with the former closely related to the spectral unmixing and the latter reduced to distance-based classification.

A. A Posteriori Orthogonal Subspace Projection

In order to estimate $\alpha = (\alpha_1\ \alpha_2\ \cdots\ \alpha_p)^T$, several techniques have been developed in [20]–[24] based on *a posteriori* information obtained from the data cube. As a result, model (1) or (2) can be cast in terms of an *a posteriori* formulation and can be given by

$$\begin{aligned}\mathbf{r} &= M\hat{\alpha}(\mathbf{r}) + \hat{\mathbf{n}}(\mathbf{r})\\ &= \mathbf{d}\hat{\alpha}_p(\mathbf{r}) + U\hat{\gamma}(\mathbf{r}) + \hat{\mathbf{n}}(\mathbf{r})\end{aligned} \tag{5}$$

where $\hat{\alpha}(\mathbf{r})$, $\hat{\alpha}_p(\mathbf{r})$, and $\hat{\gamma}(\mathbf{r})$ are estimates of α, α_p, and γ, respectively, based on the observed pixel itself $\mathbf{r}$. Because of this, model (5) is called an *a posteriori* model as opposed to model (1), which can be viewed as a Bayes or *a priori* model. For simplicity, the dependency on $\mathbf{r}$ will be dropped from all the notations of estimates throughout the rest of this paper.

1) Signature Subspace Projection (SSP) [20], [21]: Using the least squares error as an optimal criterion for model (5) yields the optimal least squares estimate of α, $\hat{\alpha}_{\mathrm{LS}}(\mathbf{r})$ given by

$$\hat{\alpha}_{\mathrm{LS}}(\mathbf{r}) = (M^T M)^{-1} M^T \mathbf{r}. \tag{6}$$

Substituting (6) for the estimate of α in model (5) results in

$$\mathbf{r} = M\hat{\alpha}_{\mathrm{LS}} + \hat{\mathbf{n}}_{\mathrm{LS}} \tag{7}$$

where

$$\hat{\mathbf{n}}_{\mathrm{LS}} = \mathbf{r} - M\hat{\alpha}_{\mathrm{LS}} = M(\alpha - \hat{\alpha}_{\mathrm{LS}}) + \mathbf{n}. \tag{8}$$

From (6), we define $P_M = M(M^T M)^{-1} M^T$ to be the signature space orthogonal projector that projects $\mathbf{r}$ into the signature space $\langle M \rangle$ and apply P_M to model (5), which yields

$$P_M\mathbf{r} = P_M M\hat{\alpha}_{\mathrm{LS}} + P_M\hat{\mathbf{n}}_{\mathrm{LS}} \tag{9}$$

$$= M\hat{\alpha}_{\mathrm{LS}} \tag{10}$$

where $P_M M = M$ and the term $P_M\hat{\mathbf{n}}_{\mathrm{LS}}$ vanishes in (9) since P_M annihilates $\hat{\mathbf{n}}_{\mathrm{LS}}$.

By coupling P_M with the OSP classifier q_{OSP}^T given by (4), a classifier q_{SSC}^T, called signature space projection classifier (SSC) derived in [21] is given by

$$q_{\mathrm{SSC}}^T = q_{\mathrm{OSP}}^T P_M = \mathbf{d}^T P_U^{\perp} P_M. \tag{11}$$

Now we apply q_{SSC}^T to both *a priori* model (1) and *a posteriori* model (5), we obtain

$$\begin{aligned}q_{\mathrm{SSC}}^T\mathbf{r} &= \mathbf{d}^T P_U^{\perp} P_M M\alpha + q_{\mathrm{SSC}}^T\mathbf{n}\\ &= \mathbf{d}^T P_U^{\perp} P_M \mathbf{d}\alpha_p + q_{\mathrm{SSC}}^T\mathbf{n}\end{aligned} \tag{12}$$

and

$$\begin{aligned}q_{\mathrm{SSC}}^T\mathbf{r} &= q_{\mathrm{SSC}}^T(M\hat{\alpha}_{\mathrm{LS}} + \mathbf{n}_{\mathrm{LS}})\\ &= \mathbf{d}^T P_U^{\perp} P_M \mathbf{d}\hat{\alpha}_p.\end{aligned} \tag{13}$$

Equating (12) and (13) yields

$$\mathbf{d}^T P_U^{\perp} P_M \mathbf{d}\hat{\alpha}_p = \mathbf{d}^T P_U^{\perp} P_M \mathbf{d}\alpha_p + q_{\mathrm{SSC}}^T\mathbf{n}. \tag{14}$$

Dividing (14) by $\mathbf{d}^T P_U^{\perp} P_M \mathbf{d}$, we obtain the estimate of α_p, denoted by $\hat{\alpha}_{\mathrm{SSC},p}$.

$$\hat{\alpha}_{\mathrm{SSC},p} = \alpha_p + \frac{q_{\mathrm{SSC}}^T\mathbf{n}}{\mathbf{d}^T P_U^{\perp} P_M \mathbf{d}} = \alpha_p + \frac{q_{\mathrm{SSC}}^T\mathbf{n}}{\mathbf{d}^T P_U^{\perp} \mathbf{d}} \tag{15}$$

where the last equality holds because $\mathbf{d}^T P_U^{\perp} P_M \mathbf{d} = \mathbf{d}^T P_U^{\perp} \mathbf{d}$. The estimation error resulting from (15) is given by

$$\varepsilon_{\mathrm{SSC},p} = \hat{\alpha}_{\mathrm{SSC},p} - \alpha_p = \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} P_M \mathbf{n}. \quad (16)$$

2) Oblique Subspace Projection (OBSP) [21]: In SSP, the noise is suppressed by making use of P_M, and the undesired signatures in U are subsequently nulled by the projector $P_U^{\perp}$. It would be convenient if we could have these two operations done in one step. One such operator, called an oblique subspace projection, was developed in [21] and designates $\langle \mathbf{d} \rangle$ as its range space and $\langle U \rangle$ as its null space. In this case, the oblique subspace projection is no longer orthogonal. Furthermore, it was shown in [28] that the orthogonal subspace projector P_M can be decomposed as a sum of two oblique projectors, one of which is the oblique subspace projection.

Let E_{XY} be a projector with its range space X and null space Y. The P_M can be decomposed and expressed by

$$P_M = E_{\mathbf{d}U} + E_{U\mathbf{d}} \quad (17)$$

with

$$E_{\mathbf{d}U} = \mathbf{d} \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} \quad (18)$$

$$E_{U\mathbf{d}} = U \left(U^T P_{\mathbf{d}}^{\perp} U\right)^{-1} U^T P_{\mathbf{d}}^{\perp} \quad (19)$$

particularly, $E_{\mathbf{d}U}\mathbf{d} = \mathbf{d}$ and $E_{\mathbf{d}U}U = 0$.

In analogy with (11), an oblique subspace projection classifier (OBC) denoted by q_{OBC}^T can be constructed via (18) by

$$q_{\mathrm{OBC}}^T = \mathbf{d}^T E_{\mathbf{d}U} = E_{\mathbf{d}U} = (\mathbf{d}^T \mathbf{d}) \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} \quad (20)$$

$$q_{\mathrm{OBC}}^T = \mathbf{d}^T E_{\mathbf{d}U} \mathbf{r} = \mathbf{d}^T \mathbf{d} \alpha_p + \mathbf{d}^T E_{\mathbf{d}U}. \quad (21)$$

Applying (20) to model (1) and model (5) results in

$$\begin{aligned} q_{\mathrm{OBC}}^T \mathbf{r} &= \mathbf{d}^T E_{\mathbf{d}U} \mathbf{r} = \mathbf{d}^T \mathbf{d} \hat{\alpha}_{\mathrm{OBC},p} + \mathbf{d}^T E_{\mathbf{d}U} \hat{\mathbf{n}} \\ &= \mathbf{d}^T \mathbf{d} \hat{\alpha}_{\mathrm{OBC},p} \end{aligned} \quad (22)$$

where $\mathbf{d}^T E_{\mathbf{d}U} \hat{\mathbf{n}} = 0$.

Equating (21) and (22) yields

$$\mathbf{d}^T \mathbf{d} \hat{\alpha}_{\mathrm{OBC},p} = \mathbf{d}^T \mathbf{d} \alpha_p + \mathbf{d}^T E_{\mathbf{d}U} \mathbf{n} \quad (23)$$

and

$$\hat{\alpha}_{\mathrm{OBC},p} = \alpha_p + (\mathbf{d}^T \mathbf{d})^{-1} \mathbf{d}^T E_{\mathbf{d}U} \mathbf{n}. \quad (24)$$

So, the estimation error $\varepsilon_{\mathrm{OBC},p}$ can be obtained from (24) as

$$\varepsilon_{\mathrm{OBC},p} = \hat{\alpha}_{\mathrm{OBC},p} - \alpha_p = \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} \mathbf{n}. \quad (25)$$

3) Maximum Likelihood Estimation (MLE) [23]: In the subspace projection approaches described in Subsections 1 and 2, we only assumed that the variance of the noise $\mathbf{n}$ is given by $\sigma^2 I$ and is independent of the signatures. We further assume that $\mathbf{n}$ is an additive white Gaussian noise. Then $p(\mathbf{r})$ in model (1) can be expressed as a Gaussian distribution with mean $M\alpha$ and variance $\sigma^2 I_{L \times L}$ (i.e., $p(\mathbf{r}) \approx N(M\alpha, \sigma^2 I_{L \times L})$. The MLE of α for model (5) can be obtained in [23], [24] and [29] by

$$\hat{\alpha}_{\mathrm{MLE}} = \arg\left\{\max_{\alpha} p(\mathbf{r})\right\} = (M^T M)^{-1} M^T \mathbf{r}. \quad (26)$$

In particular, the estimate of the p-th abundance α_p is given by

$$\begin{aligned} \hat{\alpha}_{\mathrm{MLE},p} &= \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \left(\mathbf{d}^T P_U^{\perp}\right) \mathbf{r} \\ &= \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \left(q_{\mathrm{OSP}}^T\right) \mathbf{r} \\ &= \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} (\mathbf{d}\alpha_p + \mathbf{n}) \\ &= \alpha_p + \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} \mathbf{n}. \end{aligned} \quad (27)$$

and the associated estimation error is

$$\hat{\alpha}_{\mathrm{MLE},p} - \alpha_p = \left(\mathbf{d}^T P_U^{\perp} \mathbf{d}\right)^{-1} \mathbf{d}^T P_U^{\perp} \mathbf{n}. \quad (28)$$

From (6) and (26), SSC and MLE both generate an identical abundance estimate $\hat{\alpha}_{\mathrm{LS}} = (M^T M)^{-1} M^T \mathbf{r} = \hat{\alpha}_{\mathrm{MLE}}$, but different noise estimates are produced, $(\mathbf{d}^T P_U^{\perp} \mathbf{d})^{-1} \mathbf{d}^T P_U^{\perp} P_M \mathbf{n}$ for SSC in (16), and $(\mathbf{d}^T P_U^{\perp} \mathbf{d})^{-1} \mathbf{d}^T P_U^{\perp} \mathbf{n}$ for MLE in (28). However, if we further compare (24) to (27) and (25) to (28), we discover that both sets of equations are identical. This implies that MLE is indeed OBC, given the condition that the noise is white Gaussian. In this case, MLE can be replaced by OBC in mixed pixel classification.

B. *Unsupervised OSP [22]*

Until now, we have made an important assumption that the signature matrix was given *a priori*. Due to significantly improved spectral resolution, hyperspectral sensors generally extract much more information than what we expect, particularly more spectral signatures than desired. These include natural background signatures, unwanted interferers, or clutter. Under such circumstances, identifying these signatures is almost impossible and prohibitive in practice. In order to cope with this problem, an unsupervised OSP was recently developed in [22], where the undesired and unwanted signatures can be found automatically via an unsupervised process. One such algorithm, referred to as Automatic Target Detection and Classification Algorithm (ATDCA), is a two-stage process consisting of a target generation process and target classification process and can be summarized as follows.

ATDCA

Stage 1) Target Generation Process (TGP)

Step 1) Initial condition:
Select a pixel vector with the maximum length as an initial target denoted by $\mathbf{T}_0$, i.e.,

$$\mathbf{T}_0 = \arg\left\{\max_{\mathbf{r}} [\mathbf{r}^T \mathbf{r}]\right\}.$$

Set $i = 1$ and $U_0 = \phi$.

Step 2) Find the orthogonal projections of all image pixels with respect to $\mathbf{T}_0$ by applying $P_{\mathbf{T}_0}^{\perp} = (I - \mathbf{T}_0 \mathbf{T}_0^{\#})$ to all image pixel vectors $\mathbf{r}$, where $\mathbf{T}_0^{\#}$ is the pseudo-inverse of $\mathbf{T}_0$.

Step 3) Find the first target, denoted by $\mathbf{T}_1$, by finding

$$\mathbf{T}_1 = \arg\left\{\max_{\mathbf{r}} \left[\left(P_{\mathbf{T}_0}^{\perp} \mathbf{r}\right)^T \left(P_{\mathbf{T}_0}^{\perp} \mathbf{r}\right)\right]\right\}.$$

Step 4) If $\eta_1 = \mathbf{T}_0^T P_{U_1}^{\perp} \mathbf{T}_0 < \varepsilon$ with $U_1 = \mathbf{T}_1$, go to step 7. Otherwise, let $i = i + 1$ and continue.

Step 5) Find the ith target $\mathbf{T}_i$ generated by the i-th stage, i.e.,

$$\mathbf{T}_i = \arg\left\{\max_{\mathbf{r}}\left[\left(P_{(\mathbf{T}_0, U_{i-1})}^{\perp}\mathbf{r}\right)^T \left(P_{(\mathbf{T}_0, U_{i-1})}^{\perp}\mathbf{r}\right)\right]\right\}.$$

Let $U_i = (\mathbf{T}_1 \ \mathbf{T}_2 \ \cdots \ \mathbf{T}_i)$ be the target matrix generated in the ith stage.

Step 6) Stopping rule.
Calculate

$$\eta_i = \mathbf{T}_0^T P_{U_i}^{\perp} \mathbf{T}_0 \tag{29}$$

and compare it to the prescribed threshold ε. If $\eta_i > \varepsilon$, go to step 5. Otherwise, continue. (Note that each iteration from step 5 to step 6 in the ATDCA generates and detects one target at a time.)

Step 7) At this point, the target generation process will be terminated. In this case, the process is called to be convergent. The set $\{\mathbf{T}_0, U_i\} = \{\mathbf{T}_0, \mathbf{T}_1, \ldots, \mathbf{T}_i\}$ will be the desired target set used for the next stage of target classification.

Stage 2) Target Classification Process (TCP)
In this stage, the target set $\{\mathbf{T}_0, \mathbf{T}_2, \ldots, \mathbf{T}_i\}$ generated by TGP is ready for classification. Let $\mathbf{T}_k$ be the kth target for $k \leq i+1$. Apply the OSP classifier $q_{\mathrm{OSP}}^T = \mathbf{T}_k^T P_{\bar{U}_k}^{\perp}$ given by (4) to classify $\mathbf{T}_k$, where $\bar{U}_k = (\mathbf{T}_0, \ldots \mathbf{T}_{k-1}, \mathbf{T}_{k+1}, \ldots, \mathbf{T}_i)$ is the undesired signature matrix made up of all signatures in $\{\mathbf{T}_0, \mathbf{T}_2, \ldots, \mathbf{T}_i\}$ except for the desired signature $\mathbf{T}_k$.

It is worth noting that the OPCI stopping criterion $\eta_i = \mathbf{T}_0^T P_{U_i}^{\perp} \mathbf{T}_0$ given by (29), actually arises from the constant $\mathbf{d}^T P_U^{\perp} \mathbf{d}$ appearing in the estimation errors derived in (16), (25) and (28). One comment on OPCI is useful regarding implementation of ATDCA. The OPCI only provides a guide to terminate ATDCA. Unfortunately, no optimal number of targets can be set for TGP to generate. The number of targets needed to be generated by TGP is determined by the prescribed error threshold ε set for OPCI in step 6, which is determined empirically. Another way to terminate ATDCA is to preset the number of targets. In this case, there is no need to use OPCI as a stopping criterion described in step 6. Which one is a better approach depends upon different applications and varies with scene-by-scene.

IV. Conversion of Hyperspectral Abundance Estimation Algorithms to Pure Pixel Classification

The objective of mixed pixel classification algorithms is to estimate $\alpha = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_p)^T$ in a pixel vector $\mathbf{r}$ using the linear mixture model described by (1) or (5). Since the abundance vector α in the *a priori* model (1) is assumed to be known, there is no need to estimate α for OSP. On the other hand, (5) is an *a posteriori* model and requires an estimate of α. This results in *a posteriori* OSP approach where the abundance estimation is solved as an unconstrained least squares problem. In the latter case, $\hat{\alpha}_p$ is an estimate of the abundance fraction α_p of a desired signature specified by $\mathbf{d}$ in model (1). The images generated by these algorithms are presented as gray scale, with the gray level value used to represent the estimated abundance fraction of a desired signature $\mathbf{d}$ present in a mixed pixel vector. The classification of any given pixel vector $\mathbf{r}$ is then based on the estimated abundance fraction $\hat{\alpha}_p$. In the past, this has been done by visual interpretation and later supported by ground truth. So, technically speaking, OSP and *a posteriori* OSP are signature abundance estimation algorithms, not classification algorithms. In order to use these algorithms as classifiers, we need a process, called a mixed-to-pure pixel converter that can convert mixed pixel abundance estimation to mixed pixel classification. A similar process, referred to an analog-to-digital converter (A/D converter) has been widely used in communications and signal processing. Such an A/D converter is generally implemented by vector quantization. As a matter of fact, the concept of using vector quantization (VQ) to generate desired targets has been explored in [30], where each codeword in the VQ-generated codebook corresponded to one potential target in an image scene. Furthermore, to make classification fully automated, a computer-aided classification criterion must be also provided.

A. Winner-Take-All Mixed-to-Pure Pixel Converter (WTAMPC)

In order to compare pure pixel classification to mixed pixel classification, we need to interpret a mixed pixel classification problem in the context of pure pixel classification. One way is to convert the abundance estimation for mixed pixels to the classification of pure pixels by considering model (1) as a constrained problem with some specific restrictions imposed on the estimated abundance vector $\hat{\alpha}$.

Assume that the abundance vector α in model (1) satisfies constraints $\alpha_j \geq 0$ for all $1 \leq j \leq p$ and $\sum_{j=1}^{p} \alpha_j = 1$. Additionally, the estimate $\hat{\alpha}$ is constrained to a set of p-dimensional vectors with one in only one component and zeros in the remaining $p-1$ components. Such vectors will be denoted by p-dimensional unit vectors. If $\mathbf{u}_j$ is a p-dimensional vector with 1 in the j-th component and 0's in all other remaining components (i.e., $\mathbf{u}_j = (0, \ldots, 0, \underbrace{1}_{j\text{-th component}}, 0, \ldots, 0)^T$), then $\mathbf{u}_j$ is called the j-th p-dimensional unit vector. In this case, the estimated abundance vector $\hat{\alpha}$ is forced to be a pure signature. Thus, there are only p choices for $\hat{\alpha}$. In other words, $\hat{\alpha}$ can be assigned to only one of p classes, which reduces a mixed pixel classification to a p-class classification problem. It then can be solved by pure pixel classification techniques. With these constraints model (5) becomes

$$\chi_{\mathrm{MPC}}(\mathbf{r}) = M\mathbf{u}_j = \mathbf{m}_j \quad \text{for some } 1 \leq j \leq p \tag{30}$$

where χ_{MPC} is called a mixed-to-pure pixel (M/P) converter operating on a pixel vector $\mathbf{r}$ that assigns $\mathbf{r}$ to signature $\mathbf{m}_j$ for some j. It should be noted that the estimated noise $\hat{\mathbf{n}}$ in model (5) has been absorbed into $\mathbf{u}_j$ for classification accuracy. So if we interpret model (1) by model (30), each signature vector in M represents a distinct class, and any sample pixel vector $\mathbf{r}$ will be assigned to one of the signatures in M via an M/P converter $\chi_{\mathrm{MPC}}(\mathbf{r}) = \mathbf{u}_j$ in the sense of a certain criterion. Using (30), we can assign 1 to a target pixel and 0 otherwise. The resulting image will be a binary image which shows only target pixels. An important but difficult task is to design an effective M/P converter for (30), which will preserve as much information as possible from mixed pixels during the mixed-to-pure pixel conversion.

A simple M/P converter is to use the abundance percentage as a cut-off threshold value. If the estimated abundance fraction $\hat{\alpha}_p$ of a signature $\mathbf{d}$ accounts for more than a certain percentage within $\mathbf{r}$, we may classify $\mathbf{r}$ to the material specified by the signature $\mathbf{d}$. However, in order for such an M/P converter to be effective, a percentage value needs to be appropriately selected to threshold an abundance-based image to a binary image with target pixels assigned by 1 and others by 0. Unfortunately, this was shown not effective in [31].

An alternative way is the one proposed in [31], called the WTA thresholding criterion as described later, and is very similar to the winner-take-all learning algorithm used in neural networks [32]. This WTA thresholding criterion can be used as an M/P converter and serve as a mechanism for (30) to convert a mixed pixel to a pure pixel. Instead of focusing on the abundance estimation of the desired signature α_p, as done in all OSP-based classifiers, we look at the complete spectrum of abundance estimates for all signatures present in $\hat{\alpha}$. Assume that there are p signatures $\{\mathbf{m}_j\}_{j=1}^{p}$ where $\mathbf{m}_j$ is the j-th signature. Let $\mathbf{r}$ be a mixed pixel vector to be classified and $\alpha(\mathbf{r}) = (\alpha_1(\mathbf{r})\ \alpha_2(\mathbf{r})\ \cdots\ \alpha_p(\mathbf{r}))^T$ be the associated p-dimensional abundance vector. Let $\hat{\alpha}_j(\mathbf{r})$ be the unconstrained estimated abundance fraction of $\mathbf{m}_j$ contained in $\mathbf{r}$ produced by mixed pixel classifiers. We then compare all estimated abundance fractions $\{\hat{\alpha}_1(\mathbf{r}), \hat{\alpha}_2(\mathbf{r}), \ldots \hat{\alpha}_p(\mathbf{r})\}$ and find the one with the maximum fraction, say $\hat{\alpha}_{j^*}(\mathbf{r})$ (i.e., $j^* = \arg\{\max_{1\le j\le p}\{\hat{\alpha}_j(\mathbf{r})\}\}$). It will be used to classify the $\mathbf{r}$ by assigning $\mathbf{r}$ to the j^*-th signature $\mathbf{m}_{j^*}$. In other words, using the WTA thresholding criterion and (30), we can define a WTA-based M/P converter $\chi_{\mathrm{WTAMPC}}(\mathbf{r}) = M\mathbf{u}_{j^*} = \mathbf{m}_{j^*}$ (referred to as WTAMPC) by setting $\hat{\alpha}_{j^*}(\mathbf{r}) = 1$ and $\hat{\alpha}_j(\mathbf{r}) = 0$ for $j \neq j^*$. As a result of such assignment, the mixed abundance vector $\hat{\alpha}(\mathbf{r})$ is then converted to a pure abundance vector, the j^*-th p-dimensional unit vector $\mathbf{u}_{j^*} = (0,\ldots,0,\underbrace{1}_{j^*\text{-th component}},0,\ldots,0)^T$.

B. Minimum Distance-Based Classification Algorithms

In Section IV.A, we described a WTAMPC that directly converted the abundance estimation of a mixed pixel to the classification of a pure pixel. In the following two sections, we use (30) as a vehicle to reinterpret two commonly used pure pixel classification methods, minimum distance-based classification and Fisher's linear discriminant analysis, in the context of constrained mixed pixel classification.

As noted in (30), there is no noise term present in the equation. This is because the noise can be interpreted and described

W	W	W	W
W	B	B	W
W	B	B	W
W	W	W	W

Fig. 1. Typical mask target.

as misclassification error. So, if the noise in model (1) is reinterpreted as the error resulting from classification and is also modeled as a white Gaussian, then the mixed pixel classifiers, OSP and *a posteriori* OSP described above, become Gaussian maximum likelihood classifiers

$$\hat{\alpha}_{\mathrm{MLE}} = \arg\left\{\max_{\hat{\alpha}\in\Delta} p(\mathbf{r})\right\} \tag{31}$$

where $\Delta = \{\hat{\alpha} = (\hat{\alpha}_1,\ldots,\hat{\alpha}_p)^T \,|\, \hat{\alpha}_j = 1$ for some j, and $\hat{\alpha}_i = 0$ for all $1 \le i \le p$ and $i \neq j\}$ (i.e., $\Delta = \{\mathbf{u}_j\}_{j=1}^{p} = \{(1,0,\ldots,0)^T, (0,1,0,\ldots,0)^T, \ldots, (0,0,\ldots,0,1)^T\}$). In other words, the estimated abundance vector $\hat{\alpha}$ in (31) must be a p-dimensional unit vector. Since there are p components, there are only p options in Δ. Due to the Gaussian structure assumed in $p(\mathbf{r})$, the classification using (31) can be simplified to a classifier based on the distance between class means $\{\mathbf{m}_j\}_{j=1}^{p}$ and a pixel vector $\mathbf{r}$ as shown later.

Assume that $\mathbf{x} = (x_1,\ldots,x_L)^T$ is a general sample pixel vector to be classified in a hyperspectral image. Let $\{\omega_1,\omega_2,\ldots,\omega_p\}$ be the set of classes of interest and ω_j be the class representing the j-th signature $\mathbf{m}_j = (m_{j1},\ldots,m_{jL})^T$. Assume that $\mathbf{x}_{jk}$ is the k-th sample vector in class j, and $\Xi = \{\mathbf{x}_{jk}\}_{j=1,k=1}^{p,N_j}$ is the set of sample vectors to be used for classification where N_j is the number of sample vectors in the j-th class, and $N = N_1 + \cdots + N_p$ is the total number of sample vectors. Two types of distance-based classifiers can be considered depending upon sample statistics.

1) The first-order statistics classifier.
   Minimum distance classifier:
   a) Euclidean distance

$$\mathrm{ED}(\mathbf{x},\mathbf{m}_j) = (\mathbf{x}-\mathbf{m}_j)^T(\mathbf{x}-\mathbf{m}_j) = \sum_{l=1}^{L}(x_l - m_{jl})^2. \tag{32}$$

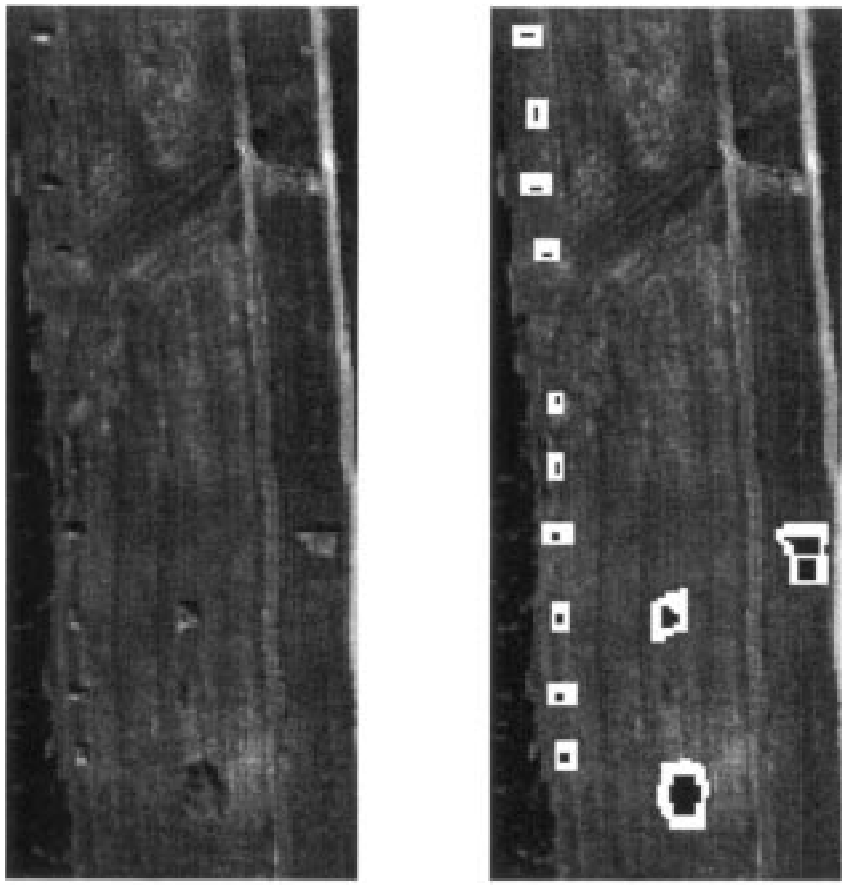
   Since the quadratic term in $\mathbf{x}$ of (32) is independent of class j, the Euclidean distance-based minimum distance classifier is a linear classifier.

   b) City block distance

$$\mathrm{CBD}(\mathbf{x},\mathbf{m}_j) = \sum_{l=1}^{L}|x_l - m_{jl}|. \tag{33}$$

   c) Tchebyshev (maxmimum) distance (TD)

$$\mathrm{TD}(\mathbf{x},\mathbf{m}_j) = \max_{1\le l\le L}|x_l - m_{jl}|. \tag{34}$$



(a) (b)

Fig. 2. (a) HYDICE image Scene (b) Same scene as Fig. 2(a) but with vehicles masked by BLACK and WHITE.

2) Second-order statistics classifiers.
   a) Mahalanobis classifier [33]

$$M(\mathbf{x}, \mathbf{m}_j) = (\mathbf{x} - \mathbf{m}_j)^T (\Sigma_j)^{-1} (\mathbf{x} - \mathbf{m}_j). \qquad (35)$$

In general, the Mahalanobis classifier is a quadratic classifier. When $\Sigma_j = \Sigma_0$ for any class j, then the Mahalanobis classifier is reduced to the minimum-distance classifier with Euclidean distance.

   b) Bhattacharyya classifier [33]

$$B_{ij} = \frac{1}{8}(\mathbf{m}_i - \mathbf{m}_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (\mathbf{m}_i - \mathbf{m}_j) + \frac{1}{2}\ln\left(\frac{\left|\frac{\Sigma_i+\Sigma_j}{2}\right|}{\sqrt{|\Sigma_i||\Sigma_j|}}\right). \qquad (36)$$

When $\Sigma_i = \Sigma_j$ for classes i and j, then the Bhattacharyya classifier is reduced to the Mahalanobis classifier.

If the covariance matrices Σ in (35) and (36) are not of full rank, their inverses will be replaced by their pseudo-inverses $\Sigma^{\#} = (\Sigma^T\Sigma)^{-1}\Sigma^T$.

C. Fisher's Linear Discriminant Analysis (LDA)

From Fisher's discriminant analysis [1], we can form total, between-class and within-class scatter matrices as follows. Let $\mu = \frac{1}{N}\sum_{j=1}^{p}\sum_{k=1}^{N_j}\mathbf{x}_{jk}$ be the global mean.

$$S_T = \frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{N_j}(\mathbf{x}_{ij} - \mu)(\mathbf{x}_{ij} - \mu)^T \qquad (37)$$

$$S_W = \frac{1}{N}\sum_{i=1}^{p}\sum_{\mathbf{x}_{ij}\in\omega_j}\frac{1}{N}(\mathbf{x}_{ij} - \mathbf{m}_i)(\mathbf{x}_{ij} - \mathbf{m}_i)^T \qquad (38)$$

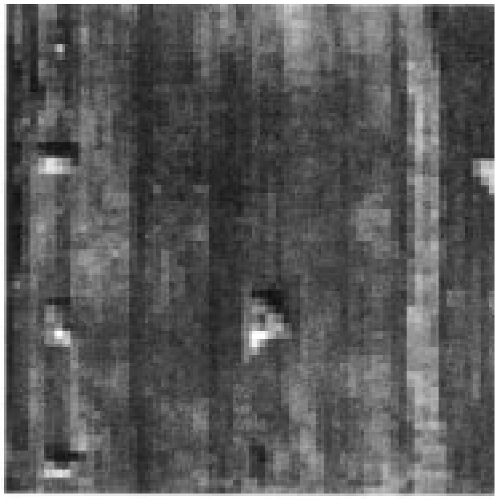
$$S_B = \sum_{i=1}^{p}\frac{N_i}{N}(\mathbf{m}_i - \mu)(\mathbf{m}_i - \mu)^T. \qquad (39)$$

From (37)–(39)

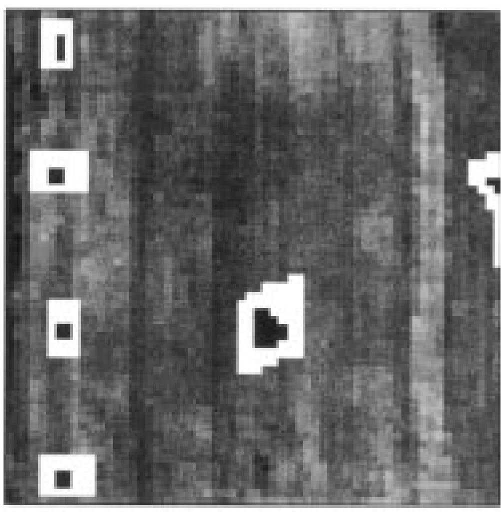
$$S_T = S_W + S_B. \qquad (40)$$

In order to minimize the misclassification error, we maximize the Raleigh quotient

$$J(Z) = \frac{Z^T S_B Z}{Z^T S_W Z} \quad \text{over } Z. \qquad (41)$$



(a)



(b)

Fig. 3. Subscene from Fig. 2(a).

Finding the solution to (41) is equivalent to solving the following generalized eigenvalue problem

$$S_B \mathbf{v}_i = \lambda_i S_W \mathbf{v}_i \tag{42}$$

or equivalently

$$S_W^{-1} S_B \mathbf{v}_i = \lambda_i \mathbf{v}_i. \tag{43}$$

where the eigenvector $\mathbf{v}_i$ is called the i-th Fisher's linear discriminant.

Since only p signatures need to be classified, there are only $p-1$ nonzero eigenvalues. Assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{p-1} > 0$ are such $p-1$ values arranged in decreasing order of magnitude. Then their corresponding eigenvectors $\{\mathbf{v}_i\}_{i=1}^{p-1}$ resulting from (42) are called Fisher's discriminants. For instance, $\mathbf{v}_1$ corresponding to λ_1 is the first Fisher's discriminant, $\mathbf{v}_2$ corresponding to λ_2 is the second Fisher's discriminant, etc. Using

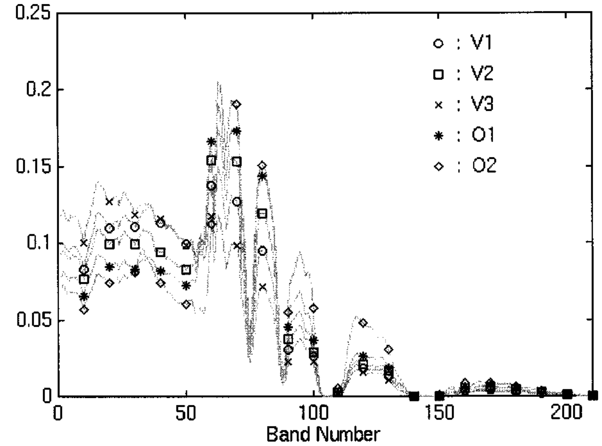


Fig. 4. Average radiances for target signatures, vehicles of Type 1, Type 2, and Type 3 and two types of man-made objects.

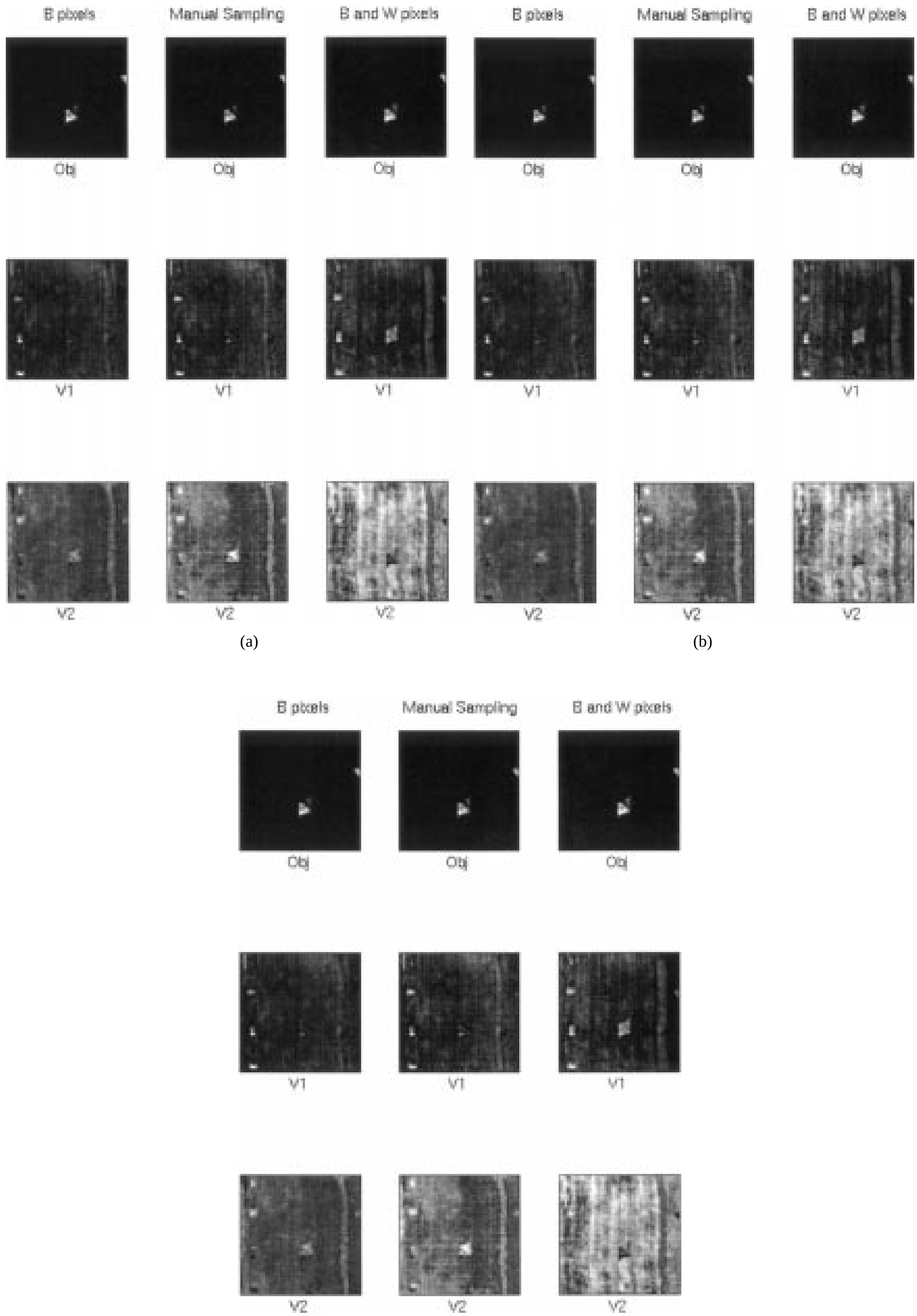
these $p-1$ Fisher's discriminants $\{\mathbf{v}_i\}_{i=1}^{p-1}$, we construct an eigenmatrix Ψ given by $\Psi = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_{p-1}]$ to map the pixel vector $\mathbf{x}$ into a new vector $\rho = \Psi\mathbf{x}$ in a new space Z linearly spanned by $\{\mathbf{v}_i\}_{i=1}^{p-1}$. Then the LDA classification is carried out in the space Z using the minimum distance measures given by (32)–(36).

D. Unsupervised Classification

Although the distance-based classifiers described above are supervised based on a set of training samples, they can be extended to unsupervised classifiers by including a clustering process such as the nearest neighboring rule [1] or a neural network-based, self-organization algorithm [32]. For example, the minimum distance classifier can be implemented by its unsupervised version, ISODATA [1].

V. Criterion for Target Detection and Classification

The standardized HYDICE data set used for the following experiments contains ten vehicles and four man-made objects. The precise spatial locations of all these targets are provided by ground truth where two types of target pixels are designated, BLACK and WHITE. The BLACK-masked (B) pixels are assumed to be target center pixels, while WHITE-masked (W) pixels may be target boundary pixels or target pixels mixed with background pixels [see Fig. 2(b)]. The positions of these two types of pixels were located in the image by (x, y) coordinates, where x and y represent row and column, respectively. The size of a mask used for a target varies and depends upon the size of the target. A typical masked target of size 4×4 is shown in Fig. 1 where black (B) pixels are centered in the mask that are considered to be the target center pixels and white (W) pixels surrounding B pixels are target pixels that may be either target boundary pixels or target pixels mixed with background pixels. Here we make a subtle distinction between a target *detected* and a target *hit*. When a target is detected, it means that at least one B target pixel is detected. When a target is hit, it means that at least either one B or one W pixel is detected. As long as one of these B or W pixels is detected, we declare the target is hit. So, by way of this definition, a target detected always implies a target hit, but not vice versa. Using these B and W pixels, we



(a) (b) (c)

Fig. 5. (a) Images produced by OSP, (b) images produced by OBSP, and (c) Images produced by SSP.

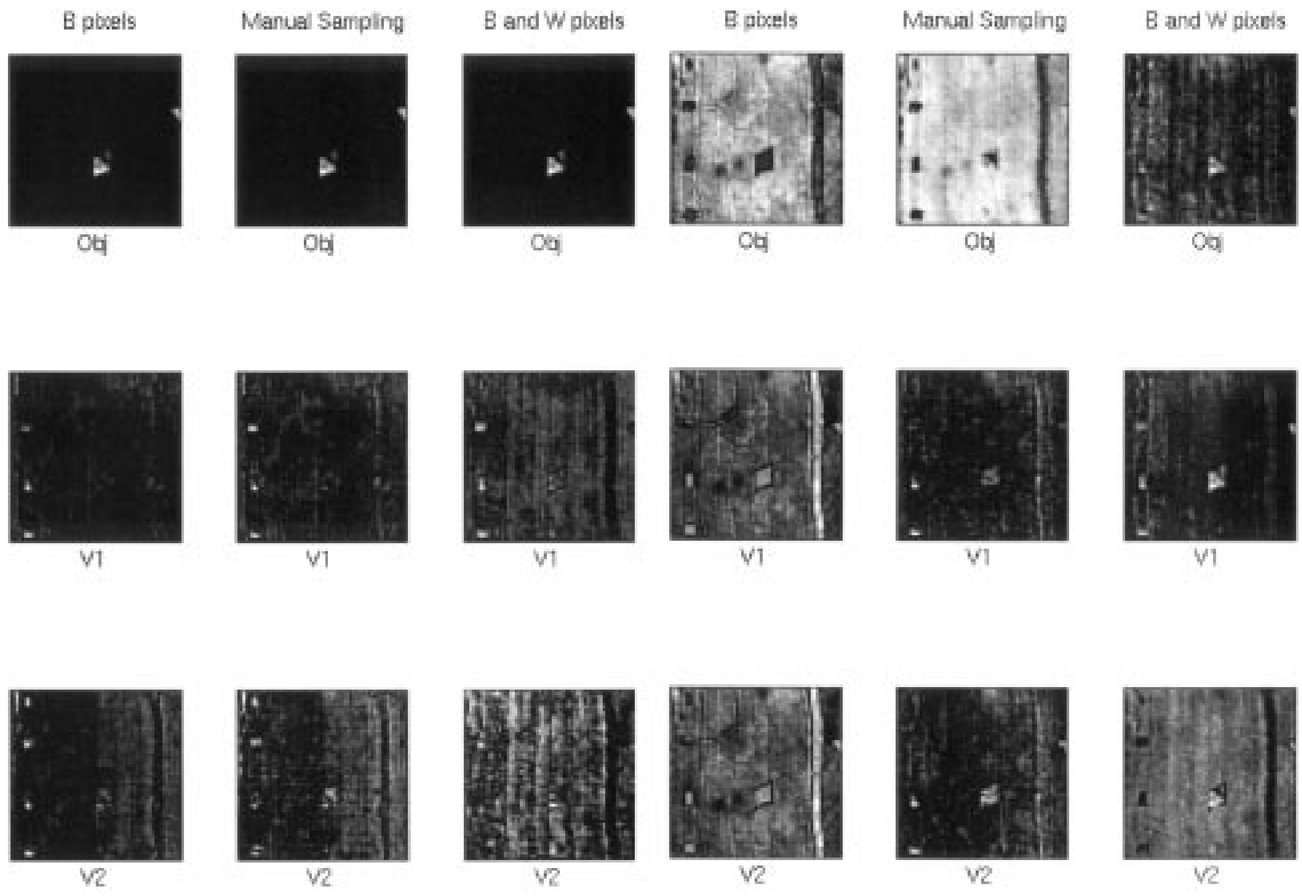


Fig. 6. (a) Error images produced by taking absolute difference between OSP-generated and OBSP-generated images. (b) Error images produced by taking absolute difference between OBSP-generated and SSP-generated images.

can actually tally the number of target pixels detected or hit by a particular algorithm.

The criteria that we use in this paper are

1) How many target B pixels are detected;
2) How many target W pixels are detected;
3) How many pixels are detected as false alarms for a target in which case neither a BLACK-masked pixel or a WHITE-masked pixel is detected;
4) How many target B pixels are missed.

For example, suppose that the shaded pixels in Fig. 1 are those detected by a detection algorithm. We declare the target to be detected with one B pixel as well as hit with one B and two W pixels. There are no false alarm pixels, but have three B pixels missed. In order to quantitatively study target detection performance, the following definitions are introduced.

N	total number of sample pixel vectors;
$\mathbf{T}$	specific target to be detected;
$N_{\mathrm{B+W}}(\mathbf{T})$	total number of BLACK-masked plus WHITE-masked $\mathbf{T}$ pixels;
$N_{\mathrm{B}}(\mathbf{T})$	total number of BLACK-masked $\mathbf{T}$ pixels;
$N_{\mathrm{W}}(\mathbf{T})$	total number of WHITE-masked $\mathbf{T}$ pixels;
$N_{\mathrm{(B+W)D}}(\mathbf{T})$	total number of either BLACK-masked or WHITE-masked pixels detected;
$N_{\mathrm{BD}}(\mathbf{T})$	total number of BLACK-masked $\mathbf{T}$ pixels detected;
$N_{\mathrm{WD}}(\mathbf{T})$	total number of WHITE-masked $\mathbf{T}$ pixels detected;
$N_{\mathrm{TPF}}(\mathbf{T})$	total number of false alarms pixels, i.e., total number of pixels which are neither BLACK-masked nor WHITE-masked $\mathbf{T}$ pixels detected;
$N_{\mathrm{TPM}}(\mathbf{T}) =$	$N_{\mathrm{B+W}}(\mathbf{T}) - N_{\mathrm{(B+W)D}}(\mathbf{T})$ total number of BLACK-masked or WHITE-masked $\mathbf{T}$ pixels missed.

Using the above notations, we can further define the detection rate $R_{\mathrm{BTD}}(\mathbf{T})$ for B pixels of target $\mathbf{T}$ by

$$R_{\mathrm{BTD}}(\mathbf{T}) = \frac{N_{\mathrm{BD}}(\mathbf{T})}{N_{\mathrm{B}}(\mathbf{T})} \tag{44}$$

and the detection rate $R_{\mathrm{WTD}}(\mathbf{T})$ for W pixels of target $\mathbf{T}$ by

$$R_{\mathrm{WTD}}(\mathbf{T}) = \frac{N_{\mathrm{WD}}(\mathbf{T})}{N_{\mathrm{W}}(\mathbf{T})}. \tag{45}$$

Since B pixels represent target center pixels and W pixels are target boundary pixels mixed with background pixels, a good detection algorithm must have a higher rate of target B pixels detected $R_{\mathrm{BTD}}(\mathbf{T})$. On the other hand, detecting a W pixel does not necessarily mean a target detected. Nevertheless, we can declare the target to be hit. For this purpose, we define the target hit rate $R_{\mathrm{TH}}(\mathbf{T})$ for target $\mathbf{T}$ by

$$R_{\mathrm{TH}}(\mathbf{T}) = \frac{N_{(\mathrm{B+W})\mathrm{D}}(\mathbf{T})}{N_{\mathrm{B+W}}(\mathbf{T})}. \tag{46}$$

So from (46) a higher target hit rate $R_{\mathrm{TH}}(\mathbf{T})$ does not imply a higher target detection rate $R_{\mathrm{BTD}}(\mathbf{T})$ or vice versa. This is because the number of W pixels are generally much greater than the number of B pixels. Thus, the W pixels may actually dominate the performance of $R_{\mathrm{TH}}(\mathbf{T})$. As will be shown in the experiments, a detection algorithm may detect all B pixels but no W pixels. In this case, this algorithm achieves 100% target pixel detection rate $R_{\mathrm{BTD}}(\mathbf{T}) = 1$, but $R_{\mathrm{WTD}}(\mathbf{T}) = 0$. As a result, its target hit rate $R_{\mathrm{TH}}(\mathbf{T})$ is very small because $R_{\mathrm{WTD}}(\mathbf{T}) = 0$. On the other hand, if the target hit rate $R_{\mathrm{TH}}(\mathbf{T}) = 1$, it implies that all B and W pixels are detected. In this case, even though the target is hit, we may still not be able to precisely locate where the target is. So the B target pixel detection rate $R_{\mathrm{BTD}}(\mathbf{T})$ is more important than $R_{\mathrm{TH}}(\mathbf{T})$ since it provides the information about the exact location of the target.

In addition to (44)–(46), we are also interested in target false alarm rate $R_{\mathrm{TPF}}(\mathbf{T})$ and target miss rate $R_{\mathrm{TPM}}(\mathbf{T})$ defined later

$$R_{\mathrm{TPF}}(\mathbf{T}) = \frac{N_{\mathrm{TPF}}(\mathbf{T})}{N - N_{\mathrm{B+W}}(\mathbf{T})} \tag{47}$$

$$\begin{aligned} R_{\mathrm{TPM}}(\mathbf{T}) &= 1 - R_{\mathrm{TH}}(\mathbf{T}) = \frac{N_{\mathrm{TPM}}(\mathbf{T})}{N_{\mathrm{B+W}}(\mathbf{T})} \\ &= \frac{N_{\mathrm{B+W}}(\mathbf{T}) - N_{(\mathrm{B+W})\mathrm{D}}(\mathbf{T})}{N_{\mathrm{B+W}}(\mathbf{T})}. \end{aligned} \tag{48}$$

If there are p targets $\Gamma = \{\mathbf{T}_i\}_{i=1}^{p}$ needing to be classified, the overall detection rate $R_{\mathrm{OD}}(\Gamma)$ for a class of targets Γ can be defined as

$$R_{\mathrm{OD}}(\Gamma) = \sum_{i=1}^{p} p(\mathbf{T}_i) R_{\mathrm{BTD}}(\mathbf{T}_i) \tag{49}$$

where $p(\mathbf{T}_i) = (N(\mathbf{T}_i)/\sum_{k=1}^{p} N(\mathbf{T}_k))$ for $1 \leq i \leq p$. As will be seen in the following experiments, a higher $R_{\mathrm{OD}}(\Gamma)$ does not imply higher classification accuracy, because it may happen that several targets are detected in one single image due to their similar signature spectra and it is difficult to discriminate one from another. This results in poor classification. In order to account for this phenomenon we define the classification rate for a specific target $\mathbf{T}_i$, $R_C(\mathbf{T}_i)$ as

$$R_C(\mathbf{T}_i) = \frac{N_{\mathrm{BD}}(\mathbf{T}_i)}{N_{\mathrm{B}}(\mathbf{T}_i) + N_{\mathrm{TPF}}(\mathbf{T}_i)} \tag{50}$$

and the overall classification rate as

$$R_{\mathrm{OC}}(\Gamma) = \sum_{i=1}^{p} p(\mathbf{T}_i) R_c(\mathbf{T}_i) \tag{51}$$

where $p(\mathbf{T}_i)$ and $R_c(\mathbf{T}_i)$ are defined by (49) and (50) respectively. Now using (44)–(51) as criteria, we can evaluate the detection and classification performance of various algorithms through the HYDICE experiments.

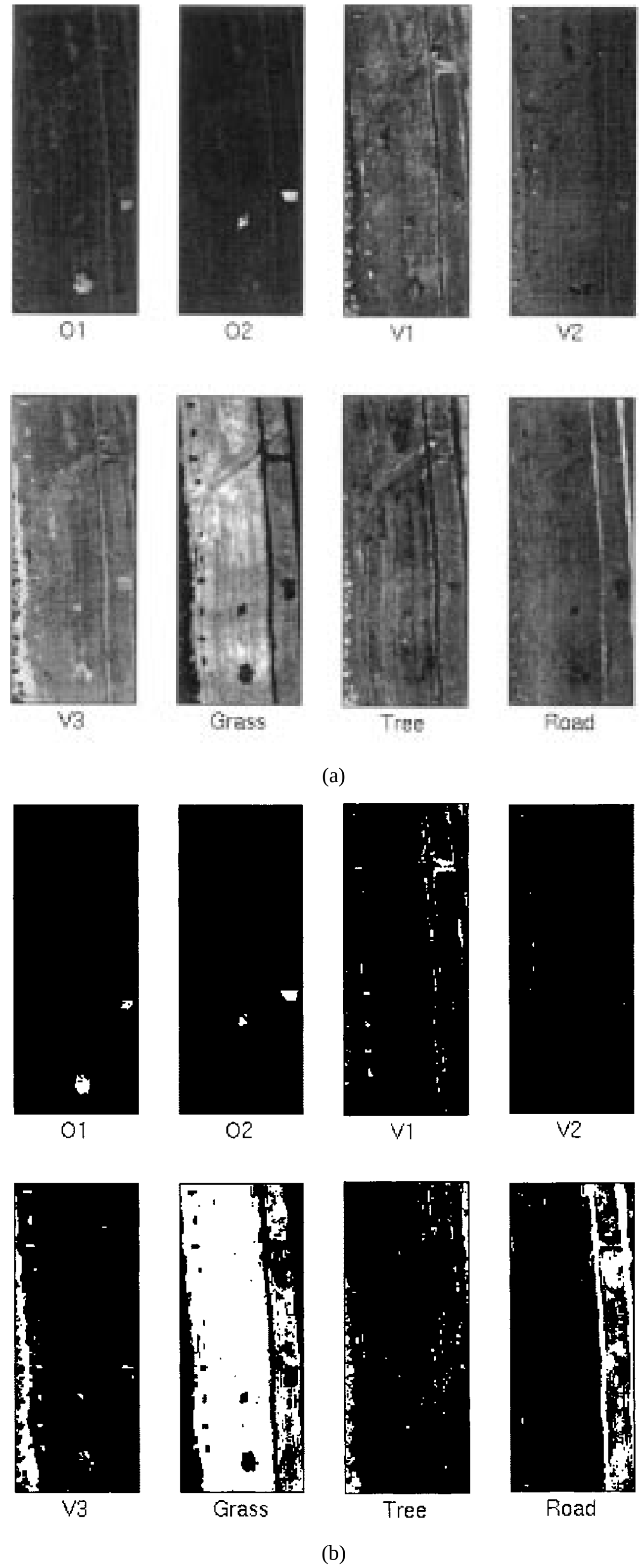


Fig. 7. (a) Abundance-based gray scale images generated by OSP using B pixels. (b) Binary images resulting from WTAMPC applied to images in Fig. 7(a).

Since the target detection and classification algorithms described in Section III are based on the abundance fractions of targets estimated from mixed pixels, the images produced by mixed pixel classification are gray-scale with the gray level

TABLE I
TALLIES OF TARGET PIXELS FOR OSP-DETECTION USING B PIXELS AFTER WTAMPC, WITH DETECTION RATES

vehicle	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	16	117	133	10	17	27	328	106	0.625	0.145	0.203	0.021	0.797
V2	7	67	74	6	1	7	4	67	0.857	0.015	0.095	0.0003	0.905
V3	9	96	105	6	22	28	1124	77	0.667	0.229	0.267	0.070	0.733
O1	73	146	219	66	54	120	0	99	0.904	0.370	0.548	0.000	0.452
O2	43	118	161	43	31	74	0	87	1.000	0.263	0.460	0.000	0.540

TABLE II
TALLIES OF TARGET PIXELS FOR OSP-DETECTION USING B AND W PIXELS AFTER WTAMPC WITH DETECTION RATES

vehicle	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	16	117	133	15	37	52	290	81	0.938	0.316	0.391	0.018	0.609
V2	7	67	74	2	35	37	2491	37	0.286	0.522	0.500	0.156	0.500
V3	9	96	105	0	42	42	4486	63	0.000	0.438	0.400	0.280	0.600
O1	73	146	219	47	37	84	2	135	0.644	0.253	0.384	0.0001	0.616
O2	43	118	161	43	30	73	0	88	1.000	0.254	0.453	0.000	0.547

TABLE III
TALLIES OF TARGET PIXELS FOR OSP-DETECTION USING MANUAL SAMPLING AFTER WTAMPC WITH DETECTION RATES

vehicle	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	16	117	133	11	18	29	44	104	0.688	0.154	0.218	0.003	0.782
V2	7	67	74	7	6	13	493	61	1.000	0.090	0.176	0.031	0.825
V3	9	96	105	3	17	20	998	85	0.333	0.177	0.191	0.063	0.810
O1	73	146	219	71	55	126	1	93	0.973	0.377	0.575	0.0001	0.425
O2	43	118	161	43	35	78	0	83	1.000	0.297	0.485	0.000	0.516

values representing the abundance fractions of targets present in mixed pixels. With the availability of standardized data and the help of the MPC algorithms developed in Section IV, we can evaluate these algorithms objectively via (44)–(51) by actually tallying the number of target pixels detected for performance analysis.

VI. COMPARATIVE PERFORMANCE ANALYSIS USING HYDICE DATA

This section contains a series of experiments which use a HYDICE standardized data set to conduct a comprehensive comparison among the OSP-based mixed pixel classification and distance-based pure pixel classification algorithms. Three comparative studies are designed. First of all, we describe the HYDICE image scene.

A. HYDICE Image Scene

The data used for the experiments are an image scene in Maryland taken by a HYDICE sensor in August 1995 using 210 bands of spectral coverage 0.4–2.5 μm with resolution 10 nm. The scene is of size 200×80, shown in Fig. 2(a), taken from a flight altitude of 10 000 ft within a GSD of approximately 1.5 m. Each pixel vector has a dimensionality of 210. This figure shows a tree line along the left edge and a large grass field on the right. This grass field contains a road along the right edge of the image. There are ten vehicles, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, and V_{10} parked along the tree line and aligned vertically. They belong to three different types, denoted by V1 for Type 1, V2 for Type 2 and V3 for Type 3. The bottom four, denoted by V_1, V_2, V_3, and V_4 belong to V1 with size approximately 4 m$\times$8 m. The middle three, denoted by V_5, V_6, and V_7, belong to V2 with size approximately 3 m$\times$ 6 m. The top three, denoted by V_8, V_9, and V_{10}, belong to V3 but have the same size as V2. In addition to vehicles, four man-made objects of two types are shown in the image. Two are located in the near center of the scene, the bottom one denoted by O_1 and the top one by O_2, and another two are on the right edge, the bottom one denoted by O_3, and the top one by O_4. O_1 and O_3 belong to the same type, indicated by O1, O_2, and O_4 belong to another type indicated by O2. In terms of class separation, there are five distinct classes of targets in the image scene, three for vehicles and two for man-made objects. It is worth noting that the HYDICE scene in Fig. 2(a) was geometrically corrected to precisely locate the spatial coordinates of all vehicles by either BLACK or WHITE masks, where the BLACK-masked pixels are center pixels of targets and WHITE-masked pixels may be part of the target pixels or target background pixels or target shadow pixels. So, BLACK-masked target pixels are always in WHITE mask frames. However, in this paper, the BLACK-masked pixels will be considered separately from WHITE-masked pixels since they

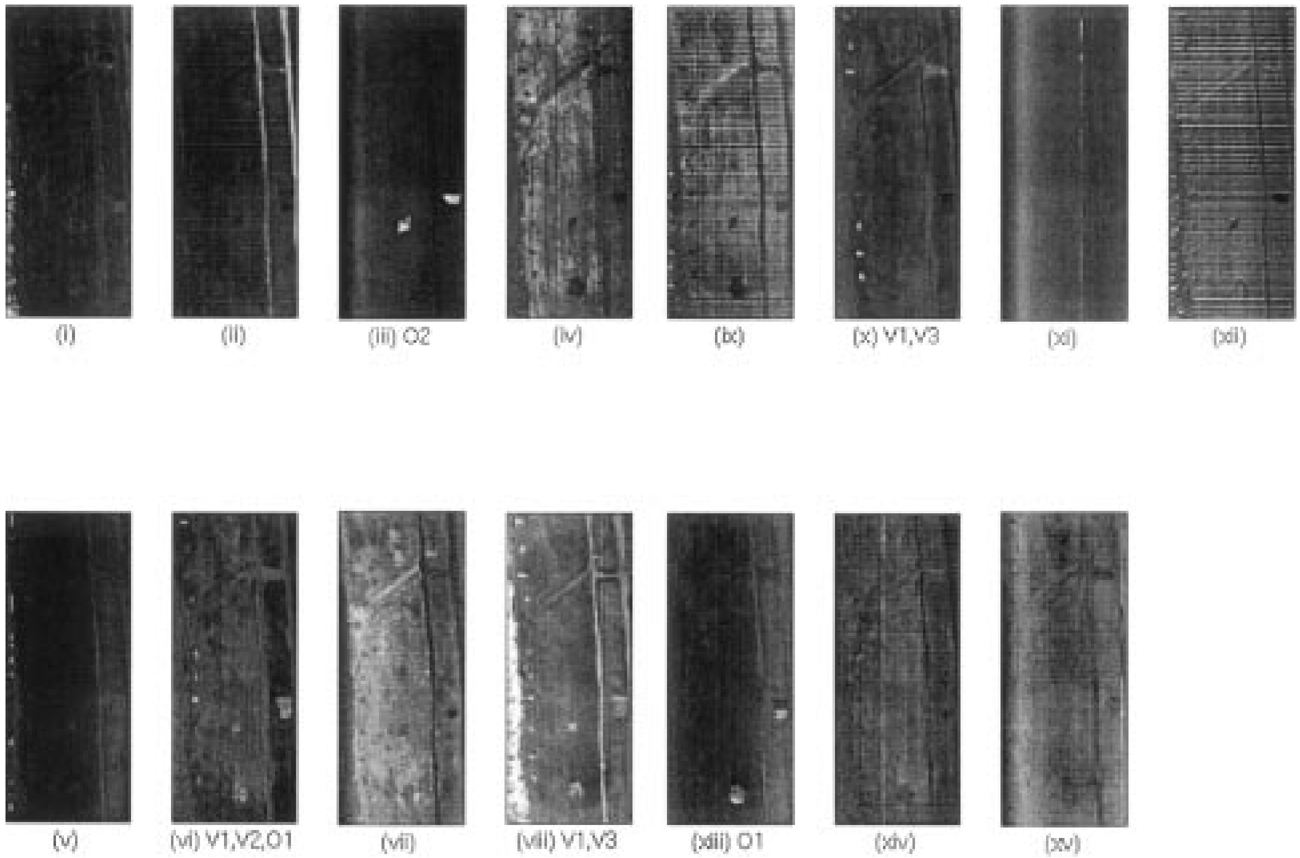


Fig. 8. (top) Abundance-based gray scale images generated by ATDCA. (bottom) Binary images resulting from WTAMPC applied to images in Fig. 8(a).

TABLE IV
TALLIES OF TARGET PIXELS FOR ATDCA AFTER WTAMPC WITH DETECTION RATES

targets	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
O2 → Fig. 10(c)	43	118	161	43	52	95	0	66	1.000	0.441	0.590	0.000	0.410
V1 →	16	117	133	4	6	10	48	123	0.250	0.051	0.075	0.003	0.925
V2 →	7	67	74	7	3	10	48	64	1.000	0.045	0.135	0.003	0.865
O1 → Fig. 10(f)	73	146	219	21	13	34	24	185	0.289	0.089	0.155	0.015	0.848
V1 →	16	117	133	1	12	17	870	120	0.063	0.103	0.098	0.054	0.902
V3 → Fig. 10(h)	9	96	105	6	11	13	866	88	0.067	0.115	0.162	0.054	0.838
V1 →	16	117	133	11	16	27	8	106	0.688	0.137	0.203	0.0005	0.797
V3 → Fig. 10(j)	9	96	105	3	5	8	27	97	0.333	0.052	0.076	0.002	0.924
O1 → Fig. 10(m)	73	146	219	49	32	81	0	138	0.671	0.219	0.370	0.000	0.630

will be used as target signatures for classification. This information allows us to perform a quantitative analysis and comparative study of various classification algorithms. A smaller scene shown in Fig. 3, cropped from the lower part of Fig. 2 will be also used for more detailed studies. It is the exact same image scene studied in [6], [7], [19], [31] and has a different GSD 0.78 meters with the image turned upside down. It contains only four vehicles V_2, V_3, V_4, and V_5 and one man-made object O_2. The top vehicle (V_5) belongs to V2 and the bottom three (V_2, V_3, V_4) belong to V1.

B. HYDICE Experiments

Since the exact locations of all the vehicles and man-made objects in Fig. 2 are available, we can extract target center pixels masked by BLACK and mixed pixels masked by WHITE directly from the image scene for each vehicle. The

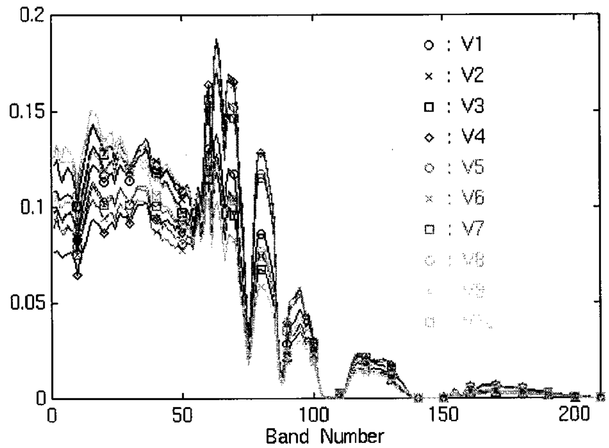


Fig. 9. Spectral signatures of the ten targets in Fig. 2.

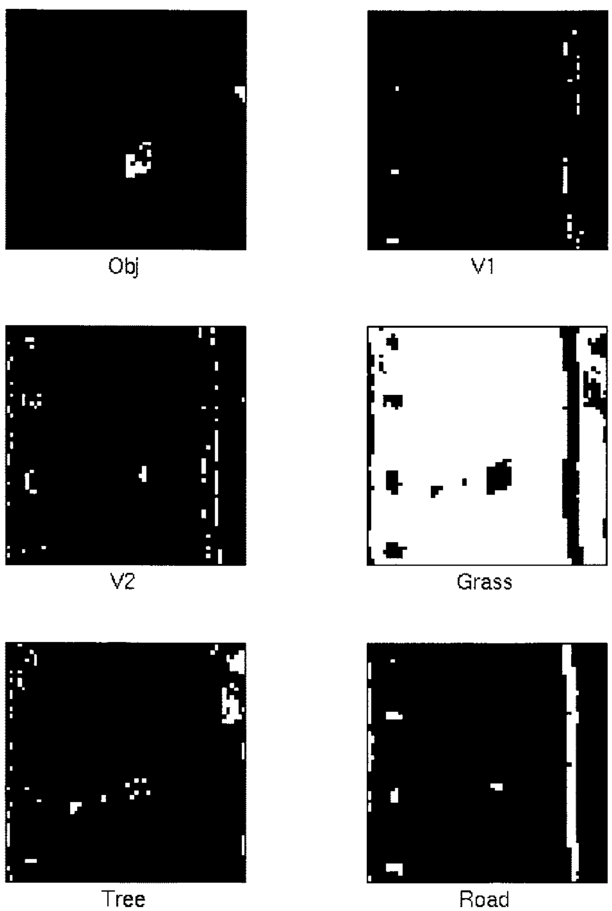


Fig. 10. Images generated by ED using B pixels.

average radiances for three types of vehicles were calculated and plotted in Fig. 4. The spectral signatures in Fig. 4 were used as the desired target information in implementation of the algorithms.

Example 1: The theoretical studies on comparative analysis among subspace projection methods were investigated previously and separately in [15], [16], [20], [21] based on AVIRIS data. In this example, we conduct an experiment-based comparison among OSP, OBSP, MLE and SSP using standardized

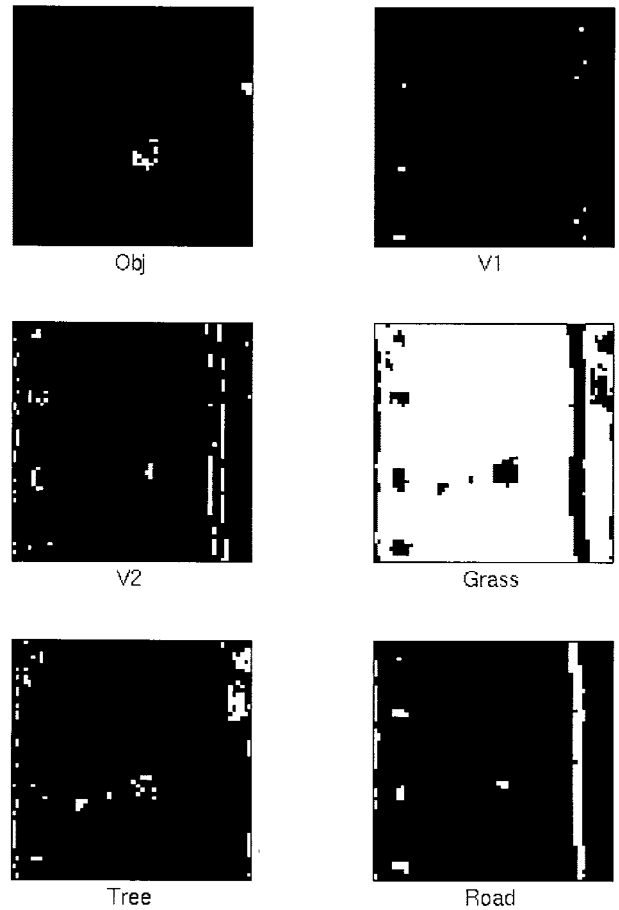


Fig. 11. Images generated by MD using B pixels.

HYDICE data. Since both OBSP and MLE generate an identical estimation error given by (25) and (28), a fact also reported in [21], [23] and [24], we will only focus our experiments on OSP, OBSP and SSP. It is interesting to note that if we apply a scaled OSP classifier, $(\mathbf{d}^T P_U^{\perp}\mathbf{d})^{-1} q_{\rm OSP}^T$ to model (2), it results in the same equations given by Eqs. (24) and (28) with both $\hat{\alpha}_{{\rm OBC},p}$ and $\hat{\alpha}_{{\rm MLE},p}$ replaced by α_p. This implies that if the knowledge about the abundance vector α is given *a priori*, then OBSP and MLE are reduced to OSP. On the other hand, if the abundance vector α is not known and needs to be estimated by $\hat{\alpha}$, then OBSP and MLE will be used to replace OSP. Consequently, OSP can be viewed as the *a priori* version of OBSP and MLE, while OBSP and MLE can be thought of as *a posteriori* version of OSP. So, the experiments done in [15] were actually based on the *a posteriori* version of OSP.

As shown in (4) and (20), OSP and OBSP produced an identical classification vector, $\mathbf{d}^T P_U^{\perp}$ with an extra scaling constant $(\mathbf{d}^T P_U^{\perp}\mathbf{d})^{-1}$ appearing in OBSP classifier. As reported in [23] and [24], this scaling constant accounts for the amount of the abundance fractions resident in classified pixels and results in two completely different gray level ranges for OSP and OBSP. However, an interesting finding was observed. The scaling constant does not have impact on images displayed on computer because the images generated by OSP and OBSP for computer display are all scaled to 256 gray levels. In this case, the scaling constant $(\mathbf{d}^T P_U^{\perp}\mathbf{d})^{-1}$ is absorbed in the scaling process for

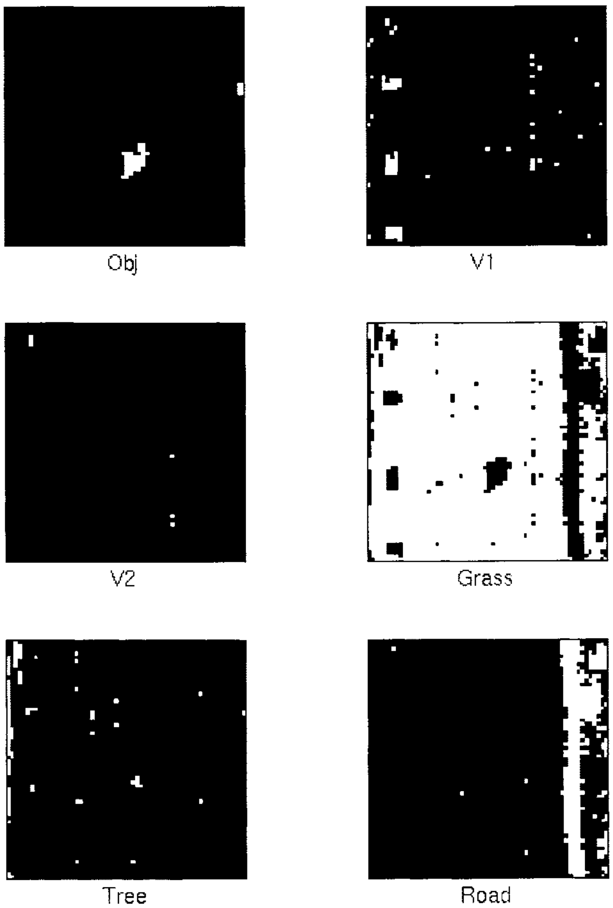


Fig. 12. Images generated by LDAED using B pixels.

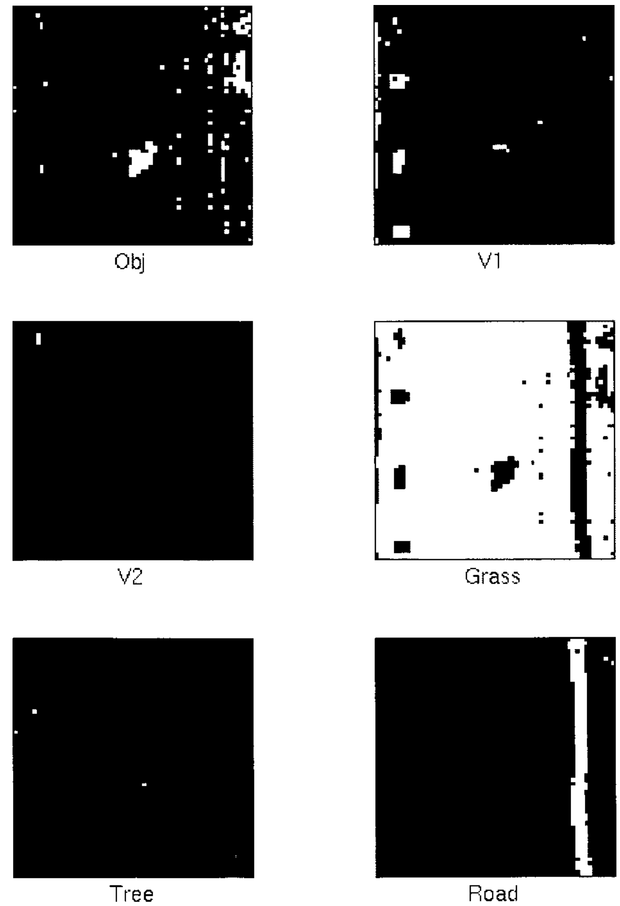
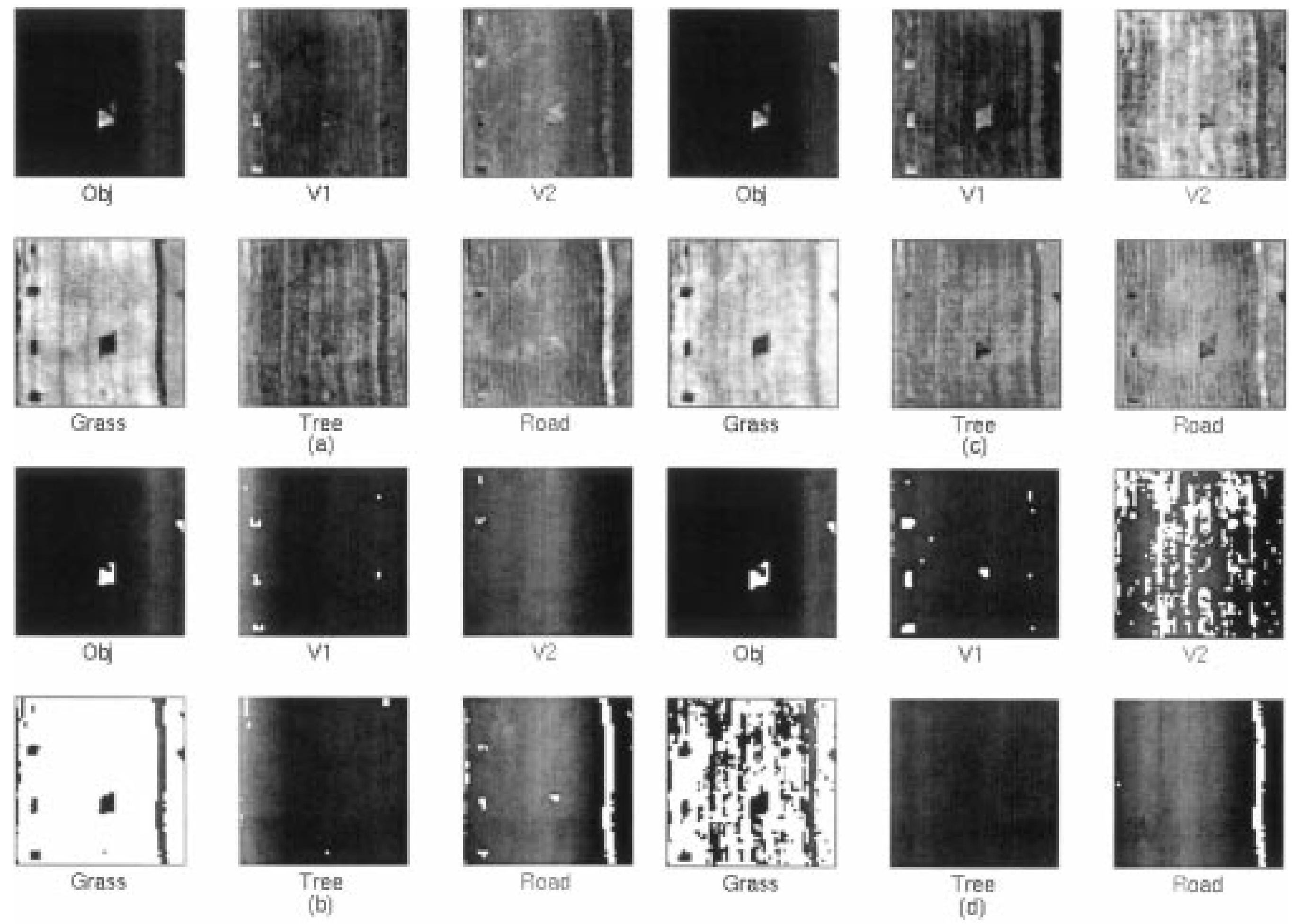


Fig. 13. Images generated by LDAMD using B pixels.

computer display. So, from a display point of view, they all produce identical results as shown in Fig. 5(a) and (b), where the man-made object O2 and a small portion of O4 in the scene in Fig. 3 were classified. In addition, this scaling process is also invariant to the abundance percentage, as mentioned in the end of Section V. This is because the abundance percentage is calculated based on relative proportions among abundance fractions. In order to overcome this problem, we took their absolute differences to substantiate the difference between the abundance fractions generated by OSP and OBSP and display their error images in 256 gray scales in Fig. 6(a). If OSP and OBSP generate identical results, their absolute difference should be 0 and their corresponding error images should be all black. Obviously, this is not true as we can see in Fig. 6(a), where only targets to be classified are shown in the images. This further justifies the subtle difference between OSP and OBSP. On the other hand, SSP is quite different from OBSP in that SSP includes an additional signature subspace projector P_M in its classifier. As a result, the SSP-generated estimation error given by (16) is different from (25). In [20], it was shown via ROC (receiver operating characteristic) analysis that SSP greatly improved OSP in terms of signal to noise ratio if the additive noise is assumed to be Gaussian. An error theory using ROC analysis for *a posteriori* OSP and OSP is further investigated in [34]. The error images resulting from the absolute difference between the OBSP-generated and SSP-generated images are shown in Fig. 6(b). Unlike Fig. 6(a), which largely shows targets of interest, the images in Fig. 6(b) contain more random noise which blurs the targets, and particularly, the classification of the object. Unfortunately, such improvements and differences cannot be visualized on a 256-gray scale computer display device because the dynamic range of the abundance is far beyond 256 scales, ranging from some negative values due to noise to numbers in thousands. So, when we display the OSP, the OBSP and SSP-generated images by scaling down to a 256-gray level range, their differences are suppressed and cannot be substantiated. As a result, the images turned out to be identical as shown in Fig. 5(a)–(c). This further simplifies our comparative analysis where the OSP can be selected as a representative for comparison in the following experiments. Nevertheless, it should be noted that the superior performance of OBSP and SSP to that of OSP in abundance estimation has been demonstrated by computer simulations in [35].

Example 2: This example is designed to demonstrate the difference between *a priori* knowledge and *a posteriori* knowledge as used in the algorithms. In the case of *a priori* knowledge, we assume that the B pixels are available. If *a posteriori* knowledge is assumed, the target pixels will be extracted directly from an image scene by manual sampling (OSP), or by computer (ATDCA) which may include either B or W pixels or both. If the signatures are not correctly extracted from the data, i.e., no B pixels, what is the effect on the detection and classification performance and how robust are OSP and ATDCA? Four signature extraction methods were compared, (1) the use of B



(a) (b)

Fig. 14. (a) Abundance-based gray scale images generated by OSP using B pixels, (b) binary images of Fig. 14(a) resulting from WTAMPC, (c) abundance-based gray scale images generated by OSP using B and W pixels, and (d) binary images of Fig. 14(c) resulting from WTAMPC.

pixels provided by the standardized data set; (2) the use of all masking pixels, i.e., both B and W pixels provided by the standardized data set; (3) manual sampling by visual inspection as done in previous research [6], [15], [16], [20], [21]; (4) unsupervised ATDCA which requires no human intervention [22]. Three types of vehicles, V1, V2, V3, and two types of objects, O1, O2, were used for classification where the desired signatures were the average values of all target sample pixels of interest. For instance, to classify V1 (i.e., the vehicles of Type 1), the desired signature was obtained by averaging target pixels of all four vehicles: V_1, V_2, V_3, V_4. Similarly, the target pixels of O_1 and O_3 were averaged to generate the desired signature for O1, etc. Fig. 7(a) is the results of using B pixels for OSP, where a total of 16 000 pixels in Fig. 2 were used for classification. In order to tally target pixels detected, we need to convert abundance-based mixed pixels to pure target pixels.

Table I is a tally of target pixels in Fig. 7(b) resulting from WTAMPC where target B pixels were used the sample pixels for OSP. Similarly, Table II is a tally of target pixels and their detection rates resulting from WTAMPC where target B and W pixels were used the sample pixels for OSP. Table III is a tally of target pixels and their detection rates resulting from WTAMPC where the sample target pixels were selected manually by visual inspection. ATDCA deserves more attention here. Unlike OSP which made use of sample pixels for target detection and classification, ATDCA does not require any such *a priori* information. It automatically searched for all targets of interest and further detected and classified the targets. So, Fig. 8(i) shows the target detection and classification results generated by ATDCA based on 15 target signatures it found in the image scene. Since ATDCA does not have prior knowledge about vehicles and objects, it detected all possible targets and then classified them subsequently. For instance, Fig. 8(iii) shows the object O_2 while Fig. 8(x) shows the vehicles V_1, V_2 and the object O_1. Similarly, both Fig. 8(xi) and (vi) show the vehicles V_1 and V_3 while Fig. 8(xiii) only shows O_1. So, Table IV is different from Tables I–III. The first column of the table specifies different types of targets in separate images as indicated and tabulates the number of detected target pixels and their corresponding detection rates using WTAMPC. In all the figures, images labeled by (a) are abundance-based images, images labeled by (b) are binary images thresholded by WTAMPC. As shown in these figures, there is no visible difference between using B pixels and manual sampling in abundance-scaled images. However, when we used full masks including B and W pixels in our experiments, the results were very poor and are not comparable to the results obtained by manual sampling and ATDCA. This is because W pixels are target-background mixed pixels and their number is much greater than that of B pixels. As a consequence, the W pixels dominate target signatures and smeared the purity of target signatures. Also shown in this example, ATDCA is comparable to OSP by visually interpreting their abundance-based images. This observation demonstrates that the unsupervised OSP can do as well as OSP and allows us to replace OSP with ATDCA in unknown or blind environment where no *a priori* knowledge is required. This advantage is substantial in many real applications because obtaining the prior information about the signatures is considered to be very difficult or sometimes impossible.

One worthy comment is the following. Although the targets shown in Fig. 2 are ten different targets, their spectral characteristics are not necessarily very distinct. As shown in Fig. 9, the spectral signatures of some targets are very similar even though the targets themselves are completely distinct. For example, the signature of V_4 is very close to those of V_5, V_6, V_7 and the signature of V_1 is also very close to those of V_2, V_3, and V_8. However, they belong to completely different vehicle types. But if we classify V_5 using its spectral signature, it was extracted along with V_6 as shown in the above experimental results, and vice versa. Similarly. it is also true for V_1, V_2, V_3, and V_8. Some studies on this phenomenon were reported in [6] and [31]. More detailed analysis on the results on Figs. 2 and 7–9 can be found in [31].

Example 3: In the previous two examples, comparisons were made among abundance estimated-based algorithms for mixed pixel classification. The example presented here will compare these algorithms against popular pure-pixel classification algorithms widely used in pattern classification as described in Section IV. In order to make the experiments simple, we again used the image scene in Fig. 3, which is of size 60×60 and has a total of 3600 pixels. In addition to vehicles and the object, we also included signatures of tree, road and grass field in the signature matrix M. So, a total of 6 classes will be considered for this example with each class represented by a distinct signature.

Since each target (including the man-made objects) contains no more than 16 B pixels whose number is far less than the number of bands. Supervised second-order minimum distance-based classification algorithms are generally not applicable because the ranks of covariance matrices used in (35) and (36) will be very small due to a very limited set of training samples. Similarly, it is also true for LDA using MD described by (42), referred to as LDAMD. Under this circumstance, we need to create more samples to augment the training pool. One way to do so is to adopt an approach proposed in [36] which uses the second-order statistics to generate additional nonlinear correlated samples from the available samples. These new generated samples can improve the classification performance. In order to further simplify experiments, ED and MD were used for comparisons because they are representatives of the first-order and second-order minimum distance-based classification algorithms. We refer for details to [31].

Figs. 10–13 are results generated by ED, MD, LDAED (LDA using ED) and LDAMD respectively. The images in Figs. 14(a)–(b) and 15(a) are abundance-based gray scale images generated by OSP and ATDCA using six signatures

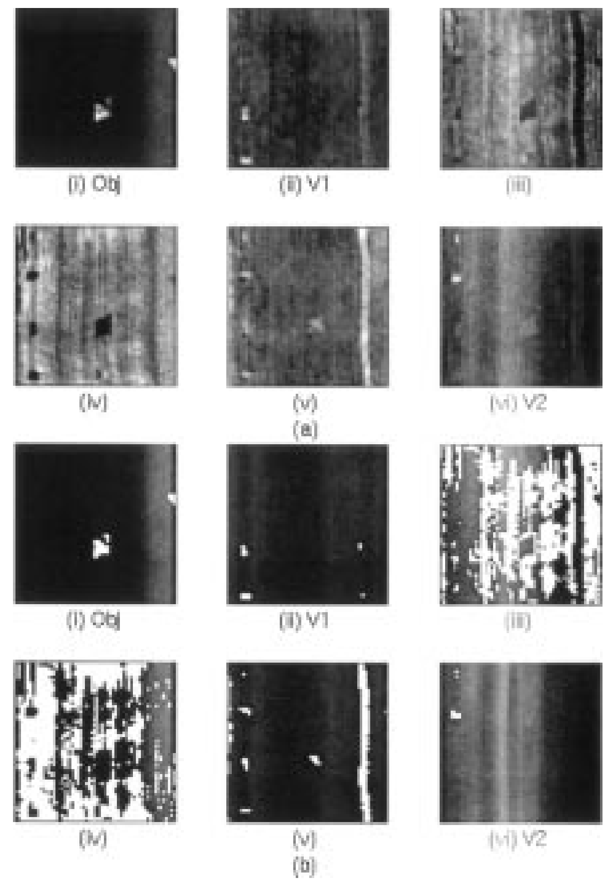


Fig. 15. (a) Abundance-based gray scale images generated by the ATDCA and (b) binary images resulting from WTAMPC.

while images in Figs. 14(c)–(d) and 15(b) are binary images thresholded by WTAMPC. Tables V–X tabulate the number of detected target pixels and their corresponding detection rates for ED, MD, LDAED, LDAMD, OSP and ATDCA respectively. It should be noted that the tallies for OSP and ATDCA were calculated after WTAMPC was applied. Their overall detection and classification rates $R_{\rm OD}$ and $R_{\rm OC}$ were also calculated by (49)–(51) and are tabulated in Table XI. The experiments demonstrate several facts.

1) The abundance-based gray scale images in Figs. 14(a)–(b) and 15(a) produced by mixed pixel classification algorithms, OSP and ATDCA are among the best since the gray levels provide significant visual information, which improves the classification results considerably.
2) If the abundance-based gray scale images in Figs. 14(a)–(b) and 15(a) are thresholded by the WTAMPC, the resulting images along with tallies shown in Figs. 14(c)–(d), 15(b), and Tables IX–X are better than those in Figs. 10 and 11 with tallies given in Tables V–VI (produced by the minimum distance-based classifiers, ED and MD), but not as good as those in Figs. 12–13 with tallies given in Tables VII–VIII (produced by LDAED and LDAMD). Among these cases, LDA produced the best results. This can be also seen in Table XI where the overall target detection rate of WTAMPC is right in between LDA and minimum distance classification.

TABLE V
TALLIES OF TARGET PIXELS FOR ED-DETECTION USING B PIXELS WITH DETECTION RATES

	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	12	86	98	2	4	6	28	92	0.167	0.047	0.061	0.008	0.939
V2	3	21	24	2	1	3	90	21	0.667	0.048	0.125	0.025	0.875
Obj	19	81	100	17	18	35	0	65	0.895	0.222	0.350	0.000	0.650

TABLE VI
TALLIES OF TARGET PIXELS FOR MD-DETECTION USING B PIXELS WITH DETECTION RATES

	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	12	86	98	2	4	6	13	92	0.167	0.047	0.061	0.004	0.939
V2	3	21	24	1	3	4	115	20	0.333	0.143	0.167	0.032	0.833
Obj	19	81	100	16	16	32	0	68	0.842	0.198	0.320	0.000	0.680

TABLE VII
TALLIES OF TARGET PIXELS FOR LDAED-DETECTION USING B PIXELS WITH DETECTION RATES

	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	12	86	98	12	30	42	32	56	1.000	0.349	0.429	0.009	0.571
V2	3	21	24	3	0	3	3	21	1.000	0.000	0.125	0.001	0.875
Obj	19	81	100	19	19	38	0	62	1.000	0.235	0.380	0.000	0.620

TABLE VIII
TALLIES OF TARGET PIXELS FOR LDAMD-DETECTION USING B PIXELS WITH DETECTION RATES

	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	12	86	98	12	31	43	55	56	1.000	0.368	0.439	0.015	0.561
V2	3	21	24	3	0	3	0	21	1.000	0.000	0.125	0.000	0.875
Obj	19	81	100	19	19	38	0	62	1.000	0.000	0.380	0.000	0.620

TABLE IX
TALLIES OF TARGET PIXELS FOR OSP-DETECTION USING B AND W PIXELS AND MANUAL SAMPLING AFTER WTAMPC WITH DETECTION RATES

vehicle	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	12	86	98	7	14	21	4	77	0.583	0.163	0.214	0.001	0.786
V2	3	21	24	3	0	3	4	21	1.000	0.000	0.125	0.001	0.875
Obj	19	81	100	19	19	38	0	62	1.000	0.235	0.380	0.000	0.620

It makes sense since LDA is based on the criterion of class separability. It further showed that the minimum distance-based pure pixel classification is among the worst. This means that without taking advantage of the visual information provided by abundance-based gray levels, the minimum distance-based classification simply cannot compete against LDA and WTAMPC. These results justify a very important conclusion. Pure pixel classification is generally not as informative as mixed pixel classification as demonstrated in Figs. 14(a), (c) and 15(a). The visual information generated by abundance-based gray scale images offers very useful and valuable knowledge that can significantly help interpret classification results.

3) There is no obvious advantage of using the second-order statistic-based classifier MD over the first order statistics-based classifier ED, as shown in Tables VII–VIII. This is probably due to the fact that there is not much spatial

TABLE X
TALLIES OF TARGET PIXELS FOR ATDCA USING 6 SIGNATURES AFTER WTAMPC WITH DETECTION RATES

vehicle	N_B	N_W	N_{B+W}	N_{BD}	N_{WD}	$N_{(B+W)D}$	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
V1	12	86	98	5	9	14	1	84	0.416	0.105	0.143	0.003	0.857
V2	3	21	24	2	0	2	10	22	0.667	0.000	0.083	0.003	0.917
Obj	19	81	100	19	23	42	0	58	1.000	0.284	0.420	0.000	0.580

TABLE XI
OVERALL DETECTION AND CLASSIFICATION RATES FOR ED, MD, LDAED, LDAMD, OSP AND ATDCA

measure / rate	ED	MD	LDAED	LDAMD	OSP			ATDCA
					B	W	Manual	
R_{OD}	0.618	0.557	1.000	1.000	0.853	0.971	0.912	0.765
R_{OC}	0.5195	0.4996	0.6992	0.7103	0.7511	0.7604	0.6598	0.7081

correlation, that a second-order statistic-based classifier can take advantage, because the pool of training target samples is relatively small.

4) For the purpose of illustration, all the images produced by pure pixel classification and WTAMPC were binary to show a specific classified target.

However, as shown in [31] this is not always the case for pure pixel classification. There are in some experiments where several targets were detected in a single binary image but could not be discriminated from one another. For instance, for an unsupervised LDAED (i.e., ISODATA(LDAED)), the three targets V1, V2, and Object were detected in a single binary image with detection rates defined by (44) as high as 100%, 100%, and 95% respectively. At the same time, the number of false alarm target pixels was also very high, e.g., 87 false alarm pixels as opposed to 12 B-pixels for V1, 125 false alarm pixels as opposed to 3 B-pixels for V2 and 95 false alarm pixels as opposed to 19 B-pixels for Object. As a result, the overall classification rate among three targets can be as low as 5% while each target detection rate is very high close to 100%. This demonstrates that higher target detection rates do not necessarily result in high classification rates. For details, we refer to [31].

VII. CONCLUSION

Many hyperspectral target detection and image classification algorithms have been proposed in the literature. Comparing one relative to another has been very challenging due to a lack of standardized data. Another difficulty arises from the fact that there are no rigorous criteria to substantiate an algorithm. This paper first considered the mixed pixel classification problem and then reinterpreted mixed pixel classification from a pure pixel classification point of view by imposing some constraints on the signature abundances. As a result, the classes of classification algorithms to be evaluated in this paper were reduced to three categories: OSP-based mixed pixel classifiers, minimum distance-based pure pixel classifiers and Fisher's LDA. In addition, a winner-take-all based mixed-to-pure pixel converter (WTAMPC) was developed to translate a mixed pixel classification problem into a pure pixel classification problem so that conventional pure pixel classification techniques could be readily applied. Although WTAMPC performed better than the minimum distance-based pure pixel classification against a standardized data set, it unfortunately did not do as well as the class separability-based LDA due to the fact that WTAMPC results in the loss of gray level information about abundance fractions. Such information, provided by the abundance-based gray scale images that are generated by mixed pixel classification algorithms, contains very useful visual features which can substantially improve image interpretation of classification results. Pure pixel classification algorithms cannot provide such information. Despite our effort to conduct comprehensive and rigorous comparative analysis of various classification algorithms for hyperspectral imagery, completion is not claimed. In particular, the WTA-based converter used in this paper for tallying target pixels was a simple thresholding technique and may not necessarily be optimal. There may exist an effective MPC which can produce better pure pixel classification performance. Many thresholding algorithms are available in the literature [37]. Most of them, however, were developed based on pure pixel image processing and may not be directly applicable to our problem. A further study on this issue may be worth pursuing. Finally, it should be noted that all the algorithms considered in this paper are unconstrained in the sense that no constraints are imposed on signature abundances, such as the abundance fractions must be summed to one or must be nonnegative. Investigation of constrained mixed pixel classification problems is a separate issue and has been recently reported in [35], [38].

ACKNOWLEDGMENT

The authors would like to thank Dr. M. L. G. Althouse and A. Ifarragaerri for proofreading this paper and the anonymous reviewers for their comments which helped to improve the paper quality and presentation.

REFERENCES

[1] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[2] J. G. Moik, *Digital Processing off Remotely Sensed Images*. Washington, DC: NASA SP-431, 1980.
[3] R. A. Schowengerdt, *Techniques for Image Processing and Classification in Remote Sensing*. New York: Academic, 1983.
[4] J. A. Richards, *Remote Sensing Digital Image Analysis*, 2nd ed. Berlin, Germany: Springer-Verlag, 1993.
[5] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[6] C.-I Chang, T.-L. E. Sun, and M. L. G. Althouse, "An unsupervised interference rejection approach to target detection and classification for hyperspectral imagery," *Opt. Eng.*, vol. 37, pp. 735–743, Mar. 1998.
[7] C.-I Chang and Q. Du, "Interference and noise adjusted principal components analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2387–2396, Sept. 1999.

[8] S. D. Zenzo, S. D. Degloria, R. Bernstein, and H. C. Kolsky, "Gaussian maximum likelihood and contextual classification algorithms for multicrop classification," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-25, pp. 805–814, Nov. 1987.
[9] S. D. Zenzo, R. Bernstein, S. D. Degloria, and H. C. Kolsky, "Gaussian maximum likelihood and contextual classification algorithms for multicrop classification experiments using thematic mapper and multispectral scanner sensor data," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-25, pp. 815–824, Nov. 1987.
[10] B. Kim and D. A. Landgrebe, "Hierarchical classifier design in high-dimensional, numerous class cases," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 792–800, July 1991.
[11] C. Lee and D. A. Landgrebe, "Analyzing high-dimensional multispectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 792–800, July 1993.
[12] B. M. Shahshahani and D. A. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hugh phenomenon," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 1087–1095, Sept. 1994.
[13] X. Jia and J. A. Richards, "Efficient maximum likelihood classification for imaging spectrometer data sets," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 274–281, Mar. 1994.
[14] C. Lee and D. A. Landgrebe, "Feature extraction based on decision boundaries," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 4, pp. 388–400, Apr. 1993.
[15] J. C. Harsanyi and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 779–785, July 1994.
[16] J. Harsanyi, "Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences," Ph.D. Dissertation, Dept. Elect. Eng., Univ. Maryland, Baltimore County, Aug. 1993.
[17] DARPA Spectral Exploitation Workshop, The Defense Adavanced Research Projects Agency, Annapolis, MD, July 1–2, 1996.
[18] C. Brumbley and C.-I Chang, "An unsupervised vector quantization-based target signature subspace projection approach to classification and detection in unknown background," *Pattern Recognit.*, vol. 32, pp. 1161–1174, July 1999.
[19] C.-I Chang, Q. Du, T. S. Sun, and M. L. G. Althouse, "A joint band prioritization and band decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2631–2641, Nov. 1999.
[20] T. M. Tu, C. H. Chen, and C.-I Chang, "A posteriori least squares orthogonal subspace projection approach to desired signature extraction and detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 127–139, Jan. 1997.
[21] C.-I Chang, X. Zhao, M. L. G. Althouse, and J.-J. Pan, "Least squares subspace projection approach to mixed pixel classification in hyperspectral images," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 898–912, May 1998.
[22] H. Ren and C.-I Chang, "A computer-aided detection and classification method for concealed targets in hyperspectral imagery," in *Int. Symp. Geoscience and Remote Sensing'98*, Seattle, WA, July 5–10, 1998, pp. 1016–1018.
[23] J. J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 1045–1046, July 1996.
[24] C.-I Chang, "Further results on relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1030–1032, May 1998.
[25] J. B. Adams and M. O. Smith, "Spectral mixture modeling: A new analysis of rock and soil types at the Viking lander 1 suite," *J. Geophys. Res.*, vol. 91, pp. 8098–8112, July 10, 1986.
[26] Y. E. Shimabukuro and J. A. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 16–20, Jan. 1991.
[27] J. W. Boardman, "Inversion of imaging spectrometry data using singular value decomposition," in *Proc. IEEE Symp. Geoscience and Remote Sensing*, 1989, pp. 2069–2072.
[28] R. T. Behrens and L. L. Scharf, "Signal processing applications of oblique projections operators," *IEEE Trans. Signal Processing*, vol. 42, pp. 1413–1423, June 1994.
[29] T. M. Tu, H. C. Shy, C.-H. Lee, and C.-I Chang, "An oblique subspace projection to mixed pixel classification in hyperspectral images," *Pattern Recognit.*, vol. 32, pp. 1399–1408, Aug. 1999.
[30] C. Brumbley and C.-I Chang, "Unsupervised linear unmixing Kalman filtering approach to signature extraction and estimation for remotely sensed images," in *Int. Symp. Geoscience and Remote Sensing'98*, Seattle, WA, July 5–10, 1998, pp. 1590–1592.
[31] H. Ren, "A Comparative Study of Mixed Pixel Classification versus Pure Pixel Classification for Multi/Hyperspectral Imagery," M.S. Thesis, Dept. Comp. Sci. Elect. Eng., Univ. Maryland, Baltimore County, May 1998.
[32] S. Haykin, *Neural Networks*: Macmillan, 1994.
[33] K. Fukunaga, *Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.
[34] C. I Chang, "Least squares error theory for linear mixing problems with mixed pixel classification for hyperspectral imagery," *Recent Res. Devel. Opt. Eng.*, vol. 2, pp. 241–268, 1999.
[35] D. Heinz and C.-I Chang, "Subpixel spectral detection for remotely sensed images," , to be published.
[36] H. Ren and C.-I Chang, "A generalized orthogonal subspace projection approach to unsupervised multispectral image classification," *SPIE Conf. Image and Signal Processing for Remote Sensing IV*, vol. 3500, pp. 42–53, Sept. 21–25, 1998.
[37] P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen, "A survey of thresholding techniques," *Comput. Vis., Graph. Image Process. (CVGIP)*, vol. 41, pp. 233–260, 1988.
[38] D. Heinz and C.-I Chang, "Fully constrained least squares-based linear unmixing," in *Int. Geoscience and Remote Sensing Symp. '99*, Hamburg, Germany, June 28–July 2, 1999, pp. 1401–1403.

Chein-I Chang (S'81–M'87–SM'92) received the B.S., M.S., and M.A. degrees from Soochow University, Taipei, Taiwan, R.O.C., in 1973, the Institute of Mathematics at National Tsing Hua University, Hsinchu, Taiwan, in 1975, and the State University of New York, Stony Brook, in 1977, respectively, all in mathematics, and the M.S. and M.S.E.E. degrees from the University of Illinois, Urbana, in 1982. He then received the 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, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore. He was a Visiting Specialist in the Institute of Information Engineering at the National Cheng Kung University, Tainan, Taiwan. from 1994 to 1995. He is an Editor for the *Journal of High Speed Network* and the Guest Editor of a special issue on Telemedicine and Applications. His research interests include automatic target recognition, multispectral/hyperspectral image processing, medical imaging, information theory and coding, signal detection and estimation, and neural networks.

Dr. Chang is a member of SPIE, INNS, Phi Kappa Phi, and Eta Kappa Nu.

Hsuan Ren (S'98) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, R.O.C., Taiwan, in 1994, and the M.S. degree in computer science and electrical engineering from the University of Maryland Baltimore County, Baltimore (UMBC), in 1998, where he is currently a Ph.D. candidate.

He is currently a research assistant in the Remote Sensing, Signal and Image Processing Laboratory, UMBC. His research interests include data compression, signal and image processing and pattern recognition.

Mr. Ren is a Member of Phi Kappa Phi.