Abstract—The principal component analysis (PCA) is one of the most effective unsupervised techniques for feature extraction. To extract higher order properties of data, researchers extended PCA to kernel PCA (KPCA) by means of kernel machines. In this paper, KPCA is applied as a feature extraction procedure to dimension reduction for target detection as a preprocessing on hyperspectral images. Then the detection was done with a support vector data description (SVDD) algorithm which is another type of one-class support vector machines (SVMs). The SVDD constructs a minimum hypersphere enclosing the target objects as much as possible. For the supervised learning, SVDD has been trained with a training set which has been chosen from target class. Balanced classification rate (BCR) and F-measure have been used to evaluate the performance of proposed technique against full-band target detection. The experimental results on hyperspectral data from the HYDICE sensors show that the KPCA based dimension reduction offers high performance for target detection applications by SVDD.

Index Terms—Feature extraction, hyperspectral imagery, kernel principal component analysis, support vector data description, target detection.

I. INTRODUCTION

In hyperspectral imaging, target detection is based on the material of the light absorption and reflection characteristics. Theoretically, two different materials may have the same spectral signature, but man-made materials which constitute the scope of the target detection are separated from their natural background with different spectra. Thanks to high and detailed spectral information, the hyperspectral imaging (HSI) offers practical solutions to difficult problems of remote sensing such as target detection and recognition, anomaly detection, end member extraction, segmentation, and classification [1]. However, extraction of the most useful spectral information from full band has been still an important problem for the researchers. One of the most common approaches to feature extraction is the dimensionality reduction method which tries to eliminate the redundancy and irrelevancy caused by the high-dimensional data [2]. In general, hyper spectral data dimensionality reduction is divided into two groups which are choosing of meaningful subsets of data or compression of data with mathematical transformations [3]. Related to the dimensionality reduction of remote sensing data, many methods have been proposed such as principal component analysis (PCA) [4], linear discriminant analysis (LDA) [5], and independent component analysis (ICA) [6]. However, those methods which focus on removing linear relationship are become insufficient to remove the characteristic of hyperspectral data like other natural distribution. The kernel methods have been utilized recently, in which learning can be carried out with preserving the nonlinear properties. One of the kernel methods is kernel PCA (KPCA) which is the nonlinear version of PCA. It can provide more information from the original data set with the capability of capturing part of high order statistics [7]. In this study, KPCA has been implemented in high dimensional hyperspectral data and the performance improvement of this implementation has been investigated on nonlinear target detection.

Because of the good generalization ability and constructing optimal decision boundary, support vector machines (SVMs) have been recently obtained great interest in remote sensing area. Since SVM is inherently suitable for binary classification problem, it cannot be applied directly to the actual target detection application. But, the another type of one-class SVMs, support vector data description (SVDD) [8] has advantages of SVM and constructs a hypersphere that encloses the target objects as much as possible. The SVDD has been implemented successfully in many different areas such as classification [9] and anomaly detection [10], [11] in remote sensing. In this study, we present the performance of detection of natural and man-made targets in HSI via SVDD with KPCA. This paper is organized as follows. In section II, brief introduction to the KPCA algorithm is provided. Section III describes the SVDD and the procedure of how to implement it for target detection. The experimental results present a performance analysis for introduced framework in Section IV. Finally, conclusions are summarized in last section.

II. KPCA FOR DIMENSIONALITY REDUCTION

In this section, we review the theoretical backgrounds of the KPCA for feature extraction. The detail information can be found in [7]. A spectral vector set are given as $x^k \in \mathbb{R}^n$, $k \in [1, ..., d]$. 

Unsupervised Nonlinear Feature Extraction Method And Its Effects On Target Detection In High-Dimensional Data
Classical PCA solves the eigenvalue problem:

\[
\lambda v = \mathbf{K} v, \quad \text{s.t. } \|v\| = 1, \quad (1)
\]

where \(\lambda\) is the eigenvalue and \(v\) is the eigenvector. In this work, one of the most popular kernels, Gaussian radial basis function (RBF) has been used. It is calculated as:

\[
\phi(x) = \exp\left(-\frac{|x - \mu|^2}{2\sigma^2}\right), \quad \sigma \in \mathbb{R}^+.
\]

Following kernel matrix can be produced with any kernel which is formed the basis of the Mercer’s Theorem:

\[
K = \begin{pmatrix}
    k(x_1, x_1) & \cdots & k(x_1, x_d) \\
    \vdots & \ddots & \vdots \\
    k(x_d, x_1) & \cdots & k(x_d, x_d)
\end{pmatrix}
\]

In which \(K\) matrix should be centered:

\[
K_{cc} = K - \mathbf{1}_d \mathbf{1}_d^T - \mu \mathbf{1}_d + \mathbf{1}_d \mu^T
\]

where identity matrix \(\mathbf{1}_d\) is a \(d \times d\) matrix which has the same elements of 1/d. The projection is accomplished as in PCA case by using solutions of (3).

Projection on first \(t\) eigenvectors:

\[
\Phi_{T}^k(x) = \sum_{i=1}^{t} v_i k(x_i, x)
\]

### III. TARGET DETECTION BASED ON SVDD

Although the SVM is a powerful technique for multi-pattern classification, since it has binary characteristics it cannot be applied naturally for one-class target detection tasks. To overcome this drawback, Tax and Duin [13] developed a support vector method called support vector data description to solve one-class problems. The SVDD tries to fit a boundary with center \(a\) and radius \(R\) around the target class. Again, A spectral vector set are given as \(x^k \in \mathbb{R}^a, k \in [1, \ldots, d]\). The following optimization procedure is constructed to minimize the \(R^2\):

\[
\min_{a,R,b}\left\{R^2 + \sum_k \xi_k \right\}
\]

\[
s.t. \quad \|\phi(x^k) - a\|^2 \leq R^2 + \xi_k,
\]

where \(\phi\), in principle unknown, is a mapping function, and \(C\) is the user-defined parameter that controls the tradeoff between errors and the volume of the hypersphere. After solving (7), a test vector \(x\) is predicted to be the outlier if

\[
\|\phi(x) - a\|^2 > R^2
\]

The problem (7) is usually considered as a Lagrange dual problem [14] given by the following:

\[
\max_{\alpha_k} \sum_k a_k K_{b,k} - \sum_k \alpha_k \alpha_l K_{k,l}
\]

\[
s.t. \quad 0 \leq \alpha_k \leq C, k = 1, \ldots, d.
\]

In (9), \(K\) is a kernel function and defined such that

\[
K_{k,l} = \langle \phi(x_k), \phi(x_l) \rangle.
\]

Thanks to the kernel trick, the SVDD can be performed employing only the kernel function instead of knowing the mapping function. More details about the SVDD may be found in [15].

### IV. EXPERIMENTAL RESULTS

#### A. Data Sets

In this study, a hyperspectral data cube is used to evaluate the performance of presented detection procedure. The dataset is an airborne remote sensing data [16] captured by the HYperspectral Digital Imagery Collection Experiment (HYDICE) sensor. The HYDICE data, gathered over the Washington, DC Mall, has seven labeled classes and 1208 lines with 307 pixels each. The HYDICE sensor originally contains 210 bands covering the 0.4-2.4 \(\mu\)m portion of the spectrum. After removing noisy bands due to water absorption, 191 bands were used in the experiments. A false color composite image of the scene is shown in Fig. 1.

Fig. 1. Hyperspectral data. HYDICE DC Mall, rotated and false colored image of (R:90, G:120, B:40).
For detection purposes, 4 classes of data set were selected (see Table I).

### B. Nonlinear Feature Extraction

In RBF kernel, σ is a main and important parameter that is determined to control boundary tightness. Before solving the eigenvalue problem, the parameter σ in the RBF kernel should be chosen appropriately. In this study, the kernel parameter is set to:

\[
\sigma = \frac{\sum_{i=1}^{d}||x_i - \bar{x}||^2}{d},
\]

(11)

where \(\bar{x}\) is the centroid of the total \(d\) training data. The eigenvalues and cumulative variance for KPCA by the solving the eigenvalue problem (3) for each data are given in Table II. The kernel matrix in each case has been produced using only labeled samples, i.e., 8079 samples. As seen in cumulative values, five kernel principal components (KPCs) contain 95% of the total variance for the KPCA transformation. Therefore, the new dimension of the DC Mall data set is set to 5. Using these five KPCA features, we have performed target detection experiment which is explained in next section to evaluate if the KPCA features are useful or not.

| Component | DC Mall
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
</tr>
<tr>
<td>1</td>
<td>49.98</td>
</tr>
<tr>
<td>2</td>
<td>20.36</td>
</tr>
<tr>
<td>3</td>
<td>16.21</td>
</tr>
<tr>
<td>4</td>
<td>5.53</td>
</tr>
<tr>
<td>5</td>
<td>3.37</td>
</tr>
</tbody>
</table>

### C. Target Detection

The SVDD has been performed for target detection using the LIBSVM toolbox [17]. In this study, we have utilized RBF kernel in SVDD. RBF kernel has some parameters, i.e., σ and C, which should be tuned. Because C is not a critical parameter, it is fixed to default value. But, σ should be carefully selected for achieving good detection performance. That's why; several values of kernel width are tested between zero and ten in increments of \(5 \times 10^{-3}\). The SVDD has been trained and implemented on an independent and randomly selected subset which includes 10% of the target and the clutter signatures by using each candidate kernel parameter value. The σ value that produces the highest BCR, defined in (12), is chosen as the proper value. The final σ has been performed to train the SVDD for target detection on raw and feature extracted versions of hyperspectral data.

Target detection is a different type of imbalanced data classification. For this reason, the detection accuracy was assessed with the balanced classification rate (BCR). In a binary decision problem, i.e, target and clutter case, the BCR is defined as the average between the sensitivity and the specificity:

\[
BCR = \frac{0.5(\frac{TP}{TP + FN} + \frac{TN}{TN + FP})}{2}
\]

(12)

where TP is the count of true detections, FN of false negatives, TN of true negatives and FP of false alarms which are elements of a confusion matrix. In addition to BCR, another metric consisting of recall and precision, namely F-measure, was used to evaluate the detection performance.

\[
Recall = \frac{TP}{(TP + FN)}
\]

\[
Precision = \frac{TP}{(TP + FP)}
\]

(13)

\[
F\text{-measure} = 2 \times \frac{\frac{Recall \times Precision}{Recall + Precision}}
\]

In ideal case, the BCR and the F-measure will yield the value of one. The final SVDD was trained with the ten percent of the randomly selected target pixels and experiments were repeated ten times. The results are reported in Table III for the DC Mall data set. Looking at the overall results, we can say that the KPCA features perform significantly better than the raw features in terms of BCR and F-measure. In other words, feature extraction helps for the target detection. Regarding the class performance with the BCR, the highest improvements were obtained for Street and Grass. In F-measure case, all targets except Trees have shown great enhancement. The improvement of the Trees is limited.

<table>
<thead>
<tr>
<th>Feature</th>
<th>No. of features</th>
<th>BCR</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Roof</td>
<td>Street</td>
</tr>
<tr>
<td>Raw</td>
<td>191</td>
<td>71.80 ± 15.31</td>
<td>72.36 ± 15.26</td>
</tr>
<tr>
<td>KPCA</td>
<td>5</td>
<td>74.76 ± 0.68</td>
<td>79.87 ± 11.20</td>
</tr>
</tbody>
</table>

Unsupervised Nonlinear Feature Extraction Method And Its Effects On Target Detection In High-Dimensional Data
CONCLUSION

In this paper, we have proposed a framework for hyperspectral target detection consists of the kernel PCA and the SVDD. Instead of using full band, we have utilized KPCA method which is exploited to extract nonlinear features of the hyperspectral data and trained SVDD with these features. The performance of the proposed algorithm is evaluated on HYDICE HSI images. BCR and F-measure have been calculated and used for comparison. The experimental results in terms of BCR and F-measure of the real hyperspectral images verify that the SVDD has better detection performance for relatively low dimension of the feature space generated by the KPCA.

REFERENCES