Manifold-Based Classification of Underwater Unexploded Ordnance in Low-Frequency Sonar

Nick H. Klausner, *Member, IEEE*, and Mahmood R. Azimi-Sadjadi, *Life Member, IEEE*

***Abstract*—This paper addresses the problem of discriminating underwater unexploded ordnance (UXO) from non-UXO objects using manifold learning principles when applied to data collected from low-frequency sonar. Our classification hypothesis is that the sequence of measurements collected from an object lie in some low-dimensional subspace which is locally linear but globally nonlinear. These low-dimensional features and their behavior on the manifold can then be used to discriminate among various UXO and non-UXO objects that may be encountered in shallow water environments. With this goal in mind, techniques are developed to not only learn the manifold but also to provide an out-of-sample embedding for newly acquired data. The manifold features from the training set are then used to construct local linear subspaces for representing each newly embedded testing feature vector. A statistical-based technique is then used to select the most likely class label by finding the class which best represents the data. The ability of the classifier to discriminate among multiple object types is then demonstrated using a sonar data set collected from underwater objects in a controlled setting. Classification results are presented and compared with an alternative method that also relies on a set of features extracted using manifold learning principles.**

***Index Terms*—Manifold learning, nonlinear dimensionality reduction, subspace classification, underwater munition classification.**

I. Introduction

DETECTION, classification, and remediation of military munitions and unexploded ordnance (UXO) in shallow water are of utmost importance to many Department of Defense (DoD) agencies owing to the severity of threats they pose to humans and the environment. The problem is challenging due to variability in environmental conditions as well as obscuration of the munitions. Thus, new technologies are needed to rapidly and reliably assess large areas that are potentially contaminated with munitions and detect, localize, and identify each individual threat with a high degree of accuracy.

Manuscript received June 28, 2018; revised December 3, 2018 and April 23, 2019; accepted May 9, 2019. Date of publication June 19, 2019; date of current version July 14, 2020. This work was supported by the Strategic Environmental Research and Development Program under SEED Contract W912HQ-14-P-0033. *(Corresponding author: Mahmood R. Azimi-Sadjadi.)*

Associate Editor: L. Culver.

N. H. Klausner was with the Information System Technologies, Inc., Fort Collins, CO 80521 USA. He is now with the Numerica Corporation, Fort Collins, CO 80528 USA (e-mail: nklausne@gmail.com).

M. R. Azimi-Sadjadi is with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA (e-mail: azimi@engr.colostate.edu).

Digital Object Identifier 10.1109/JOE.2019.2916942

Present systems employed to search for underwater munitions in shallow water are typically based upon low frequency broadband sonar that offer longer detection range and wider surveying area while at the same time provide good penetration into the sediment for detecting partially or fully buried munitions. Moreover, the use of transmitted signals with a wider bandwidth provides high range resolution for detailed surveys with greater localization capability as well as better ability to excite those discriminatory structural modes of the proud and/or buried targets for unambiguous target classification. The theme of this paper is the development of dedicated methods for feature extraction and classification of military munitions that remain robust to operating and environmental changes using data collected from low-frequency sonar systems.

Extracting low-dimensional features that are representative of the properties of the original high-dimensional data is an important topic in pattern recognition, information retrieval, and data mining. These *dimensionality reduction* algorithms can be typically categorized into linear versus nonlinear. Among the linear methods are principal component analysis (PCA) and Fisher discriminant analysis (FDA) [1], which attempt to find a low-dimensional linear subspace that best characterizes the structure of the data. Among the popular nonlinear dimensionality reduction methods are manifold learning algorithms including isometric feature mapping (ISOMAP) [2], [3], locally linear embedding (LLE) [4], [5], maximum variance unfolding (MVU), or semidefinite embedding [6], and Laplacian eigenmaps (LE) [7]. Unlike ISOMAP [2], which is global in nature, i.e., preserves geometry at all scales, LLE and LE are local approaches in which they attempt to only preserve the local geometry of the data by mapping nearby points in the ambient space to nearby points on the low-dimensional manifold. Similar to LLE, MVU [6] also belongs to the class of *spectral embedding*, though it is based on estimating and preserving local distances and angles. These methods exhibit several problems including inability to deal with highly curved manifolds and out-of-sample extension for nonisometric manifolds. The latter implies that they fail to provide a feature mapping (explicit or implicit) to map new data points that are not included in the original training set.

Although most work has involved exploiting manifold structure for dimensionality reduction, not much attention has been paid to its use in feature extraction for classification purposes. In [8], the authors discussed exploiting manifold structure for the purposes of enhancing classification performance. Given a set of explicit generative models, one for each signal manifold, the authors devise a maximum likelihood (ML) approach to signal

classification. Under a signal-plus-noise representation with a spherically symmetric noise model, the classification problem is cast as one of finding the signal manifold that is closest to the test sample, an approach similar to the one adopted here. In [9], the authors exploited a manifold-based classification method by assuming that the data from all classes lie on the same submanifold. Low-dimensional features are subsequently extracted using the LE algorithm. Using these low-dimensional features, a linear classification strategy was employed to distinguish one class from another.

The goal of this paper is to develop and test a new feature extraction and classification framework for discriminating among various UXO and non-UXO objects using low-frequency sonar. The underlying hypothesis motivating the methods developed in this paper is that the set of sonar returns collected from an object as the platform is translated along a certain path produces a sequence of aspect-dependent observations which inherently lie on some low-dimensional manifold [10]. These low-dimensional features and their dynamic behavior on the manifold can then be used to discriminate one type of object from another. With this objective in mind, a manifold is trained using the LE algorithm [7], which is extended to include previously unseen testing data to yield an out-of-sample embedding. Once these manifold features are extracted, the training features from each object type are used to construct linear subspaces for locally representing each extracted feature. A statistically inspired classification strategy is then proposed based on ML principles which selects the class best representing the data. The performance of the proposed feature extraction and classification methods are then demonstrated by applying them to a data set collected in a freshwater pond consisting of a rail system collecting sonar backscatter from multiple object types at different range and orientation. The algorithm's classification performance is benchmarked against that in [9], which also relies on the LE algorithm for feature extraction, using metrics such as probability of correct classification and receiver operating characteristic (ROC) curve.

This paper is organized as follows. Section II gives a review of the methods used to train and embed newly acquired data points on the manifold. Using these extracted features, Section III then describes the technique introduced to perform classification on the trained manifold. Section IV-C gives the test results of the proposed classifier when trained on simulated data and then applied to two low-frequency sonar data sets containing various UXO and non-UXO objects. Finally, Section V gives concluding remarks on this paper.

II. Feature Extraction Using Manifold Learning

The first step in the proposed classification algorithm [11] is to perform feature extraction and dimensionality reduction using manifold learning. Here, it is assumed that we are given a set of N labeled training patterns $\{\mathbf{x}_i\}_{i=1}^{N}$ with each $\mathbf{x}_i \in \mathbb{R}^D$ belonging to one of L different models or classes. Given this set of data, we wish to define a mapping $\mathbf{f} : \mathbb{R}^D \rightarrow \mathbb{R}^d$ ($d \ll D$) such that the feature vector $\mathbf{y}_i = \mathbf{f}(\mathbf{x}_i) \in \mathbb{R}^d$ captures the general behavior of the data over a low-dimensional manifold.

LE [7] is an algorithm that attempts to achieve this mapping by defining features such that if the data points $\mathbf{x}_i$ and $\mathbf{x}_j$ are near one another in the original high-dimensional ambient space then their corresponding feature representations $\mathbf{y}_i$ and $\mathbf{y}_j$ will be near one another in the low-dimensional space as well. The algorithm begins by defining a weighted graph whose edges are set by selecting the K nearest neighbors to each point. Each edge relating the data point $\mathbf{x}_i$ to $\mathbf{x}_j$ is then weighted using

$$w_{ij} = \begin{cases} k(\mathbf{x}_i, \mathbf{x}_j), & \text{if nodes } i \text{ and } j \text{ are connected} \\ 0, & \text{otherwise} \end{cases}$$

for some symmetric continuous function $k(\cdot,\cdot)$, typically chosen to be the Gaussian $k(\mathbf{x}_i, \mathbf{x}_j) = \exp\{-(||\mathbf{x}_i - \mathbf{x}_j||^2)/\sigma^2\}$ with smoothing parameter σ^2. Given this graph connecting each point to its nearest neighbors, we then wish to find the set of coordinates $\mathbf{Y} = [\mathbf{y}_1 \ \cdots \ \mathbf{y}_N] \in \mathbb{R}^{d \times N}$ that minimizes the objective function

$$J(\mathbf{Y}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}||\mathbf{y}_i - \mathbf{y}_j||^2. \tag{1}$$

That is, to produce a set of coordinates that are near one another if their corresponding data points are near one another in the original ambient space. Enforcing additional constraints that remove several trivial solutions associated with minimizing (1) yields the optimization problem [7]

$$\begin{aligned} \min_{\mathbf{Y}} \quad & \text{tr}\left(\mathbf{YLY}^T\right) \\ \text{s.t.} \quad & \mathbf{YDY}^T = \mathbf{I} \\ & \mathbf{YD1} = \mathbf{0} \end{aligned} \tag{2}$$

where $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is the graph Laplacian, $\mathbf{W}$ is a symmetric matrix with elements $[\mathbf{W}]_{ij} = w_{ij}$, $\mathbf{D} = \text{diag}\{\sum_j w_{j1}, \ldots, \sum_j w_{jN}\}$, and $\mathbf{1} = [1 \ \cdots \ 1]^T \in \mathbb{R}^N$. The solution to this optimization problem can be found by solving the generalized eigenvalue problem $\mathbf{Ly} = \lambda\mathbf{Dy}$ and extracting eigenvectors associated with the smallest d nonzero eigenvalues [7]. This process then embeds each data point in $\mathbb{R}^d$.

Although the LE method gives one the ability to learn the structure of an underlying manifold in the training data, the theory described above does not directly allow one to extract low-dimensional features for novel testing data. That is, the method does not allow one to embed unforeseen data on the manifold. One could simply attempt to achieve this by adding new rows and columns to the weight matrix $\mathbf{W}$ corresponding to the new data points. However, this would change the solution to (2) attained during training, i.e., adding new data would modify the structure of the manifold. Moreover, the algorithm is not directly capable of synthesizing data in the original ambient space based on its corresponding location on the manifold. Therefore, we seek a method that is able to embed test data on the manifold without modifying its structure for the purposes of extracting features from newly acquired data.

To embed novel testing data onto the manifold, we employ the use of a latent variable approach [12] which combines the advantages of the original LE method with those of latent variable models. For this method we assume that there exist a pair of data matrices $\mathbf{X}_s \in \mathbb{R}^{D \times N}$ and $\mathbf{X}_u \in \mathbb{R}^{D \times M}$ representing data previously seen during the training process and the unseen novel testing data, respectively. With these two data matrices,

we then wish to find an embedding $\mathbf{Y}_u \in \mathbb{R}^{d\times M}$ for the unseen data while leaving the embedding $\mathbf{Y}_s \in \mathbb{R}^{d\times N}$ for the training data unchanged. The most natural way to accomplish this is to extend the problem in (2) by modifying the objective function as

$$J(\mathbf{Y}_u) = \text{tr}\left([\mathbf{Y}_s\,\mathbf{Y}_u]\begin{bmatrix}\mathbf{L}_{ss} & \mathbf{L}_{su}\\ \mathbf{L}_{us} & \mathbf{L}_{uu}\end{bmatrix}\begin{bmatrix}\mathbf{Y}_s^T\\ \mathbf{Y}_u^T\end{bmatrix}\right)$$
$$= \text{tr}\left(\mathbf{Y}_s\mathbf{L}_{ss}\mathbf{Y}_s^T\right) + 2\text{tr}\left(\mathbf{Y}_s\mathbf{L}_{su}\mathbf{Y}_u^T\right) + \text{tr}\left(\mathbf{Y}_u\mathbf{L}_{uu}\mathbf{Y}_u^T\right) \tag{3}$$

where $\mathbf{L}_{ss}$ and $\mathbf{L}_{uu}$ are the graph Laplacians for $\mathbf{X}_s$ and $\mathbf{X}_u$, respectively, and $\mathbf{L}_{su} = \mathbf{L}_{us}^T$ is the graph Laplacian shared between them. Note that the constraints used in (2) are not required in this case since they are already imposed on the solution for $\mathbf{X}_s$ [12]. Taking the derivative of (3) with respect to the unknown feature matrix $\mathbf{Y}_u$ while imposing the constraint that $\mathbf{Y}_s$ remains fixed yields the solution

$$\mathbf{Y}_u = -\mathbf{Y}_s\mathbf{L}_{su}\mathbf{L}_{uu}^{-1}. \tag{4}$$

If we now consider a single novel testing sample (i.e., $M = 1$) so that $\mathbf{x}_u = \mathbf{X}_u \in \mathbb{R}^D$ and $\mathbf{y}_u = \mathbf{Y}_u \in \mathbb{R}^d$, then the two graph Laplacian used in embedding this data point on the manifold are simply given by

$$\mathbf{L}_{su} = -\mathbf{w}_{su} = -[k(\mathbf{x}_u, \mathbf{x}_1)\ \cdots\ k(\mathbf{x}_u, \mathbf{x}_N)]^T \in \mathbb{R}^N$$

and

$$\mathbf{L}_{uu} = \ell_{uu} = \mathbf{1}^T\mathbf{w}_{su} = \sum_{i=1}^{N} k(\mathbf{x}_u, \mathbf{x}_i).$$

Substituting these expressions into the solution given in (4) yields the feature vector

$$\mathbf{y}_u = \mathbf{f}(\mathbf{x}_u) = -\frac{1}{\ell_{uu}}\mathbf{Y}_s\mathbf{L}_{su} = \frac{\mathbf{Y}_s\mathbf{w}_{su}}{\mathbf{1}^T\mathbf{w}_{su}}$$
$$= \sum_{i=1}^{N}\frac{k(\mathbf{x}_u, \mathbf{x}_i)}{\sum_{j=1}^{N} k(\mathbf{x}_u, \mathbf{x}_j)}\mathbf{y}_i. \tag{5}$$

Using these results for the purposes of both embedding (analysis) and reconstructing (synthesis) data points on the manifold finally suggests the pair of relationships [12]

$$\mathbf{f}(\mathbf{x}_u) = \sum_{i=1}^{N}\frac{k(\mathbf{x}_u, \mathbf{x}_i)}{\sum_{j=1}^{N} k(\mathbf{x}_u, \mathbf{x}_j)}\mathbf{y}_i \tag{6}$$

$$\mathbf{g}(\mathbf{y}_u) = \sum_{i=1}^{N}\frac{k(\mathbf{y}_u, \mathbf{y}_i)}{\sum_{j=1}^{N} k(\mathbf{y}_u, \mathbf{y}_j)}\mathbf{x}_i. \tag{7}$$

Thus, new data points are embedded onto the low-dimensional manifold and reconstructed using a convex combination of the samples used in the training set.

III. Manifold-Based Classification Using Local Linear Representations

As discussed before, the guiding principle behind many manifold learning algorithms is that the data lie on a low-dimensional manifold which parameterizes the inherently nonlinear properties of the data. However, it is often assumed that the manifold is locally linear so that, at least on small enough scales, one may measure distances between neighboring points using Euclidean means. In keeping with this same philosophy, we develop a classification method that relies on local measures when deciding the class label of an object [11]. This is achieved by constructing local linear subspaces using the training data for each object type and selecting the one that is best capable of representing the data. Thus, unlike traditional template matching which performs replica correlation or the matched subspace classification [13], [14] which uses energy of the signal in a subspace to make decisions, the proposed method exploits locally adaptive aspect-dependent subspace using manifold features.

Fig. 1 gives a block diagram of the proposed classification system. The first step in the process involves extracting a set of low-dimensional features by embedding them onto the manifold. Let the matrix $\mathbf{Y}_s = [\mathbf{Y}_1\ \cdots\ \mathbf{Y}_L]$ denote the set of training features found by solving (2) using labeled training data where $\mathbf{Y}_j \in \mathbb{R}^{d\times N_j}$ denotes those feature vectors associated with class (or object type) $j \in [1, L]$. Also, let

$$\mathbf{X}_u = \left[\mathbf{x}_1^{(u)}\ \cdots\ \mathbf{x}_M^{(u)}\right] \in \mathbb{R}^{D\times M}$$

denote a sequence of M newly observed testing data vectors with $M \le N_j$. It is assumed that the columns of matrix $\mathbf{X}_u$ form a naturally ordered sequence (according to the vehicle path) and that every column within this matrix corresponds to the same class. In this specific application, this sequence of vectors is formed by measuring an object's acoustic response in a linear survey path leading to a sequence of M aspects from the same object. Feature extraction is subsequently performed by individually applying each vector $\mathbf{x}_i^{(u)}$ in $\mathbf{X}_u$ to the analysis equation given in (6) using the training features in $\mathbf{Y}_s$ to produce the set of feature vectors $\mathbf{Y}_u \in \mathbb{R}^{d\times M}$ in an unsupervised fashion.

A. Sequence Matching

Once this set of low-dimensional feature vectors has been extracted, the next step in the classification procedure shown in Fig. 1 involves finding the subset of features in each training data matrix $\mathbf{Y}_j,\ j \in [1, L]$ that best matches the sequence of extracted features in $\mathbf{Y}_u$. Since in this application the columns of data matrix $\mathbf{X}_u$ form a naturally ordered sequence of observations from an object over a range of aspect angles, it makes sense in this case to find the corresponding range of aspect angles in the training data that best matches the extracted features. In this way, one may define each point's nearest neighbors by finding where the data matrix $\mathbf{Y}_u$ *as a whole* best matches the training data for each class as opposed to relying on point-wise estimates of proximity through Euclidean distance. Fig. 2(a) gives a depiction of the idea behind this process for two sequences of features lying on a low-dimensional manifold.

Finding the set of training features that provides the best match is accomplished by taking a length-M sliding window of the feature vectors (i.e., for M aspects) in $\mathbf{Y}_j$ and computing the normalized inner product (or matching index) between the features in that window and the extracted sequence of features in $\mathbf{Y}_u$. That is, if we let $\mathbf{Y}_j^{(m)} \in \mathbb{R}^{d\times M}$ denote the subset of features corresponding to the mth window, the match between these two

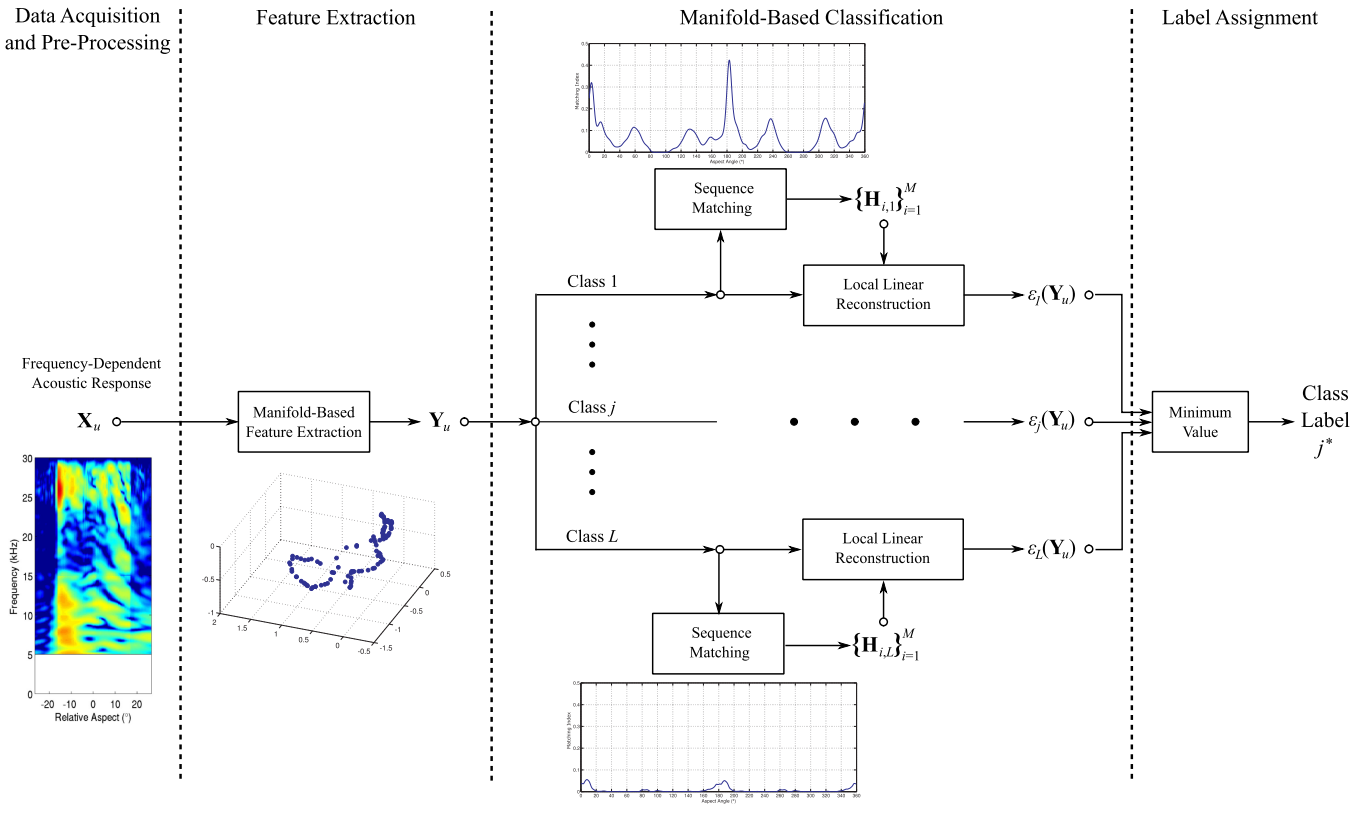

Fig. 1. Block diagram of the entire classification system. Low-dimensional manifold features are extracted from the frequency-dependent acoustic response of an unknown object. Sequence matching determines candidate aspect sequences with high similarity indices. Aspect-class dependent subspace classifier decides the class membership of the unknown object.

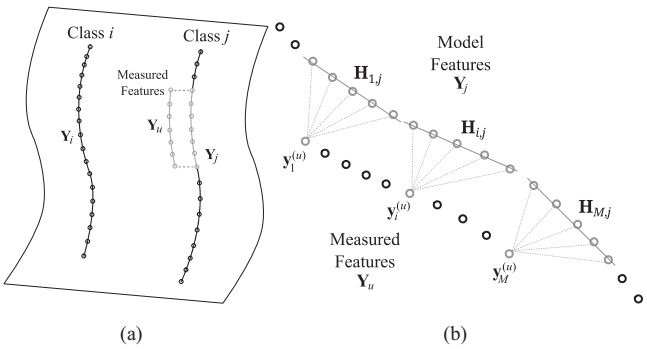

Fig. 2. Performing classification on the manifold using the extracted sequence of features $\mathbf{Y}_u$. (a) Finding the sequence of features in $\mathbf{Y}_j$ that best matches $\mathbf{Y}_u$. (b) Constructing local linear subspaces $\{\mathbf{H}_{i,j}\}$ from the training data nearest each sample $\mathbf{y}_i^{(u)}$.

sets of features is measured by finding the index m^* that gives

$$m^* = \arg\max_m C_j(m) = \frac{\left|\langle \mathbf{Y}_j^{(m)}, \mathbf{Y}_u \rangle\right|^2}{||\mathbf{Y}_j^{(m)}||_F^2 ||\mathbf{Y}_u||_F^2}. \quad (8)$$

In this equation, the expression $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{B}^T\mathbf{A})$ represents the Frobenius inner product between matrices $\mathbf{A}$ and $\mathbf{B}$ and $||\mathbf{A}||_F = \sqrt{\text{tr}(\mathbf{A}^T\mathbf{A})}$ is the Frobenius norm of matrix $\mathbf{A}$. Note that this matching index takes the value 1 when $\mathbf{Y}_j^{(m)} = \mathbf{Y}_u$ and zero when columns of $\mathbf{Y}_j^{(m)}$ are orthogonal to the corresponding columns of $\mathbf{Y}_u$.

Fig. 3 illustrates the plots of the matching index $C_j(m)$ produced by the manifold features of the measured response from an aluminum UXO object when applied to those of the model-generated [13], [15] training data from an aluminum UXO and aluminum cylinder. The top image in Fig. 3 displays the portion of the frequency-aspect acoustic color response around 20° aspect angle for the actual UXO object. The top image in each subplot shows the acoustic color templates of the model-generated data at all aspects for each of these two objects while the bottom graphs give the plot of feature space matching index given in (8) as m is varied. The vertical dashed line in each plot denotes the point where the matching index reaches its maximum value and the corresponding window is shown above it in the acoustic color template using a red box. From this figure, one can see that aluminum UXO class produces a much higher level of match. More importantly, one can see that the matching index in (8) for the aluminum UXO indeed correctly estimates the object's aspect angle as the graph achieves its maximum value near 20°. Note that the fact that there are two distinct maxima points in this plot is due to the rotational symmetry exhibited by a cylindrical object.

It must be pointed out that although in this example this aspect-dependent sequence matching resulted in higher matching for the correct object class, it lacks robustness to object orientation,

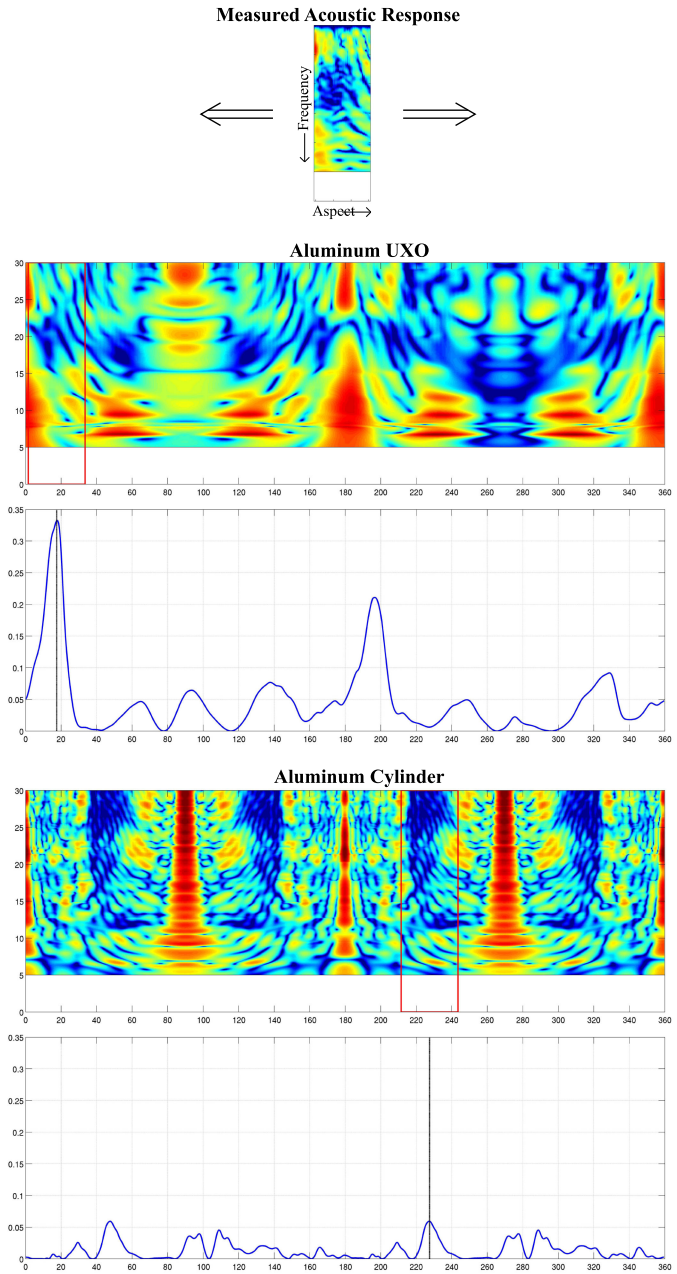

Fig. 3. Sequence matching results of an actual aluminum UXO around a 20° aspect angle (top) with those extracted from simulated aluminum UXO and cylinder. The red box shows the corresponding window of aspect angles where the matching index reached its maximum value.

range, and variation in seafloor properties. That is, it is possible that for a certain range of aspects two objects of different classes exhibit similar matching results at some object orientations.

B. Class Determination Using Local Linear Representations

Once it is determined where the sequence of feature vectors $\mathbf{Y}_u$ best fits (with respect to aspect sequence) with the training sequence from each class in $\mathbf{Y}_j$, the final stage of our classification process [11] in Fig. 1 involves building locally linear representations and selecting the class that best represents the data. This is accomplished by expressing each feature vector in $\mathbf{Y}_u$ as a linear combination of its P nearest neighbors in $\mathbf{Y}_j^{(m^*)}$ found by applying (8), measuring the error in each representation, and selecting the class that yields the least error. Fig. 2(b) gives a depiction of this process where each point is connected to its $P = 5$ nearest neighbors using a dashed line.

For each $\mathbf{y}_i^{(u)}$ in $\mathbf{Y}_u$ for $i = 1, \ldots, M$, let $\mathbf{H}_{i,j} \in \mathbb{R}^{d \times P}$ be the local linear subspace formed from the P vectors in $\mathbf{Y}_j^{(m^*)}$ nearest in aspect to $\mathbf{y}_i^{(u)}$. That is, for each object type $j \in [1, L]$ and for every sample $i \in [1, M]$, a linear subspace is constructed using the training data that provides the best match with the extracted features. Note that there can and will be some overlap in the feature vectors used to construct each subspace for nearby samples, i.e., many of the feature vectors used to build $\mathbf{H}_{i,j}$ will also be used to build $\mathbf{H}_{i+1,j}$, and so on. Then, each feature vector $\mathbf{y}_i^{(u)}$ is represented using the linear model

$$\mathbf{y}_i^{(u)} = \mathbf{H}_{i,j}\boldsymbol{\theta}_i + \mathbf{n}_i \tag{9}$$

where $\boldsymbol{\theta}_i \in \mathbb{R}^P$ is a vector of deterministic but unknown parameters and $\mathbf{n}_i \in \mathbb{R}^d$ is a vector containing elements that are *independent identically distributed* realizations of a zero-mean normal random variable with unknown variance σ^2. Thus, in this model the unknown vector $\boldsymbol{\theta}_i$ describes the coordinates in the local linear subspace $\mathbf{H}_{i,j}$ for $\mathbf{y}_i^{(u)}$ and the unknown variance σ^2 in some sense measures the inaccuracy in representation between measurement and model. Under these assumptions, the data matrix $\mathbf{Y}_u$ has the likelihood function

$$\ell_j(\mathbf{Y}_u) = \frac{1}{(2\pi\sigma^2)^{dM/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{M}\left|\left|\mathbf{y}_i^{(u)} - \mathbf{H}_{i,j}\boldsymbol{\theta}_i\right|\right|_2^2\right\} \tag{10}$$

where $||\mathbf{x}||_2$ denotes the ℓ_2 norm of vector $\mathbf{x}$. Replacing the unknown parameters $\boldsymbol{\theta}_i$ and σ^2 in (10) with their ML estimates

$$\hat{\boldsymbol{\theta}}_i = \left(\mathbf{H}_{i,j}^T\mathbf{H}_{i,j}\right)^{-1}\mathbf{H}_{i,j}^T\mathbf{y}_i^{(u)} \tag{11}$$

$$\hat{\sigma}^2 = \frac{1}{dM}\sum_{i=1}^{M}\left|\left|\mathbf{y}_i^{(u)} - \mathbf{P}_{\mathbf{H}_{i,j}}\mathbf{y}_i^{(u)}\right|\right|_2^2 \tag{12}$$

results in the likelihood function

$$\begin{aligned}\ell_j(\mathbf{Y}_u) &= \frac{1}{(2\pi\hat{\sigma}^2)^{dM/2}} \exp\left\{-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{M}\left|\left|\mathbf{y}_i^{(u)} - \mathbf{H}_{i,j}\hat{\boldsymbol{\theta}}_i\right|\right|_2^2\right\} \\ &= e^{-dM/2}\left(\frac{2\pi}{dM}\sum_{i=1}^{M}\left|\left|\mathbf{y}_i^{(u)} - \mathbf{P}_{\mathbf{H}_{i,j}}\mathbf{y}_i^{(u)}\right|\right|_2^2\right)^{-\frac{dM}{2}}.\end{aligned} \tag{13}$$

In expressions (12) and (13), the matrix

$$\mathbf{P}_{\mathbf{H}_{i,j}} = \mathbf{H}_{i,j}\left(\mathbf{H}_{i,j}^T\mathbf{H}_{i,j}\right)^{-1}\mathbf{H}_{i,j}^T$$

is the orthogonal projection matrix onto the P-dimensional subspace spanned by the columns (training vectors) of matrix $\mathbf{H}_{i,j}$. Ignoring all constants that are not dependent on the data, one can see from (13) that finding the class index j that maximizes likelihood is equivalent to finding the class which solves the optimization problem

$$j^* = \arg\min_j \epsilon_j(\mathbf{Y}_u) \tag{14}$$

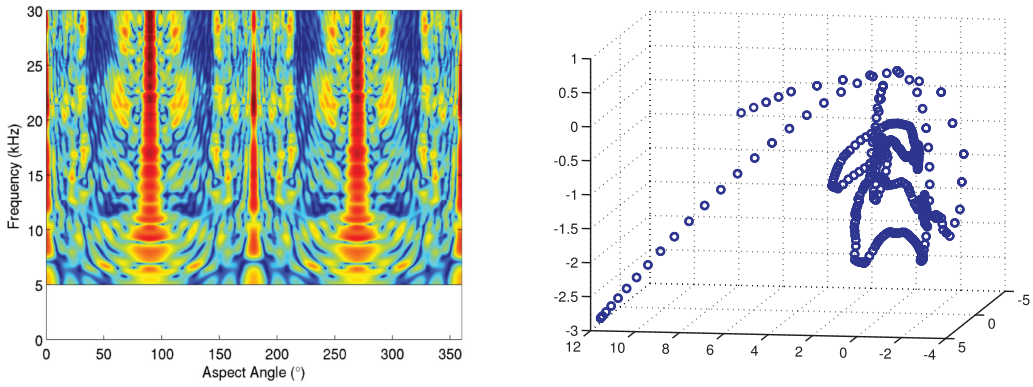

Fig. 4. Training data and corresponding manifold features for an aluminum cylinder object. (a) Acoustic color training template. (b) Manifold features in three dimensions.

with the objective function

$$\epsilon_j(\mathbf{Y}_u) = \sum_{i=1}^{M} \left|\left| \mathbf{y}_i^{(u)} - \mathbf{P}_{\mathbf{H}_{i,j}} \mathbf{y}_i^{(u)} \right|\right|_2^2. \tag{15}$$

Thus, by projecting the observation (manifold feature vectors) onto the local linear subspace $\mathbf{H}_{i,j}$ through the projection operator $\mathbf{P}_{\mathbf{H}_{i,j}}$, the minimization problem in (14) selects the class which best models the data in the sense of minimizing the error in representing the data. This representation indeed captures the object-dependent dynamical behaviors on the manifold by employing the local features that are used to build the subspace matrices $\mathbf{H}_{i,j}$.

IV. Results and Discussion

A. Training Data Set and Manifold Learning

The target-in-the-environment-response model [15], which combines an acoustic ray model for propagation and a free-field target scattering model was utilized to generate raw sonar returns for different simulated runs. Using this model and procedures described in [13] and [16], raw sonar runs for four (out of five) different objects, two "replica UXO" objects (aluminum and steel UXO) and two non-UXO objects (aluminum cylinder and pipe), used in the PondEX experiments were modeled. The real UXO object in the PondEX experiments was not included to test the generalization ability of the overall classification system. The modeled runs were designed to mimic a circular path with the object at the center 10 m from the sonar. The sonar elevation was 3.8 m above the sediment. The synthetic sonar data sets were generated for two different environments with water sound speeds matching those conditions in PondEx experiments, however, one used the sediment sound speed of sand and the other used that of a slightly denser material like a sand-clay mixture. These raw synthetic sonar returns were then processed to create acoustic color data [13], [16]. All acoustic color data from these experiments are on a calibrated absolute target strength scale.

The acoustic color data for each object are subsequently used as training data set to build the low-dimensional manifold. Thus, the original data vector $\mathbf{x}_i \in \mathbb{R}^D$ referred to throughout Section II represents the acoustic color response of a particular object at a given aspect angle. That is, each element of the vector $\mathbf{x}_i$ represents the magnitude of the frequency response over the 25 kHz bandwidth of the sonar system with a 50 Hz resolution resulting in a $D = 501$ dimensional data vector at each aspect. As mentioned before, for all of the object types in the test set except for the real UXO, model-generated acoustic color templates containing the spectral information for that object over the entire 360° range in aspect were used to train the manifold mapping process. Figs. 4(a) and 5(a) give two examples of the training acoustic color data for an aluminum cylinder and a UXO object, respectively. The training acoustic color data from all objects included in the training set are then used together to form the weight matrix $\mathbf{W}$ and graph Laplacian $\mathbf{L}$ used in (2). Here, the weighted graph was constructed by finding the $K = 64$ nearest neighbors to each training point and weighting them with the Gaussian

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{ -\frac{||\mathbf{x}_i - \mathbf{x}_j||}{\sigma^2} \right\}$$

with smoothing parameter $\sigma^2 = 2.5$. Using the smallest $d = 48$ eigenvectors of the graph Laplacian $\mathbf{L}$, the set of coordinates $\mathbf{Y}$ associated with each object type are then used as training features to represent that particular object. Figs. 4(b) and 5(b) give examples of the first three (out of 48) manifold coordinates for an aluminum cylinder and aluminum UXO objects corresponding to their acoustic color data given in Figs. 4(a) and 5(a), respectively. From these two figures one can see that, although the manifold for each object does indeed form a definitive track as one moves from one aspect to another, that track is very complicated and somewhat erratic at least in three dimensions.

B. Testing Data Set and Procedure

To test the ability of the proposed method for discriminating UXO from non-UXO objects, the classification method in Section III was applied to the experimental PondEx09 and PondEx10 data sets [15] collected at Naval Surface Warfare Center Panama (NSWC), Panama City, FL, USA. The pond facility used in these experiments was designed to collect sonar

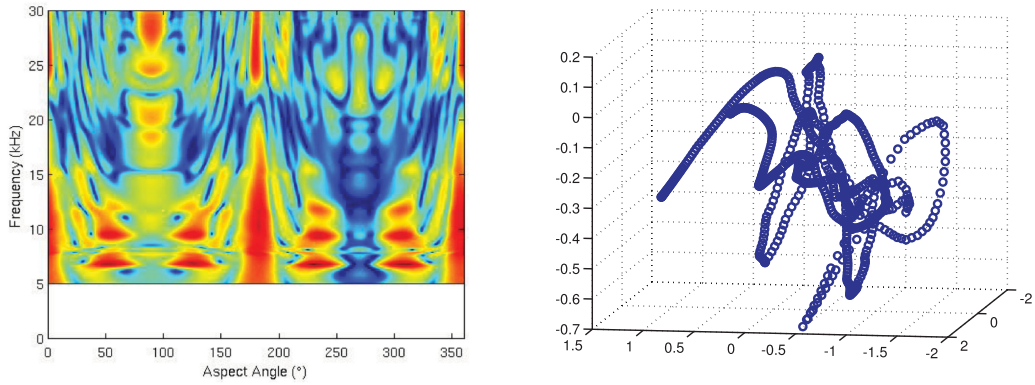

Fig. 5. Training data and corresponding manifold features for an aluminum UXO object. (a) Acoustic color training template. (b) Manifold features in three dimensions.

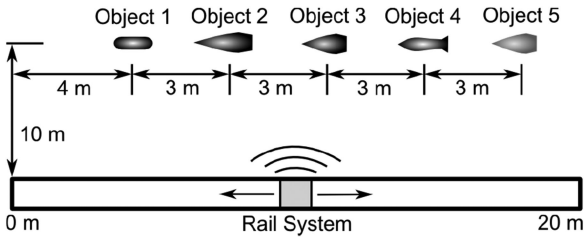

Fig. 6. Layout of the target fields for PondEx experiments.

TABLE I
OBJECTS IN THE PONDEX09-10 TESTING DATA SET

Object Type	Class
Aluminum UXO	UXO
Steel UXO	UXO
Real UXO	UXO
Aluminum Cylinder	Non-UXO
Aluminum Pipe	Non-UXO

data from underwater objects in a relatively controlled and clutter-free environment. Fig. 6 shows the layout of the test setup for PondEx10 experiments including the relative locations of the rail-mounted sonar system and the objects in the target field. Note that for the PondEx09 data set, the set up was the same but the objects were placed one at a time. As can be seen from Fig. 6, both experiments consisted of a 20-m rail system collecting sonar returns from one or more objects with varying shapes, sizes, and compositions located at a certain range from the rail. In most cases, the object was placed at a range of 10 m but there were several experiments where the object was located only 5 m from the rail. There were a total of nine object orientations ranging from -80° to $+80^\circ$ in 20° increments where a 0° object orientation designates a configuration where the object's major axis is parallel to the rail system. The sonar transmit signal was a 6-ms linear frequency modulated (LFM) pulse over 0.5–30 kHz with a 10% taper between the leading and trailing edges to minimize ringing in the transmitted signals. For the studies conducted here, five different object types were used. Table I gives a list of the object types used in this study which consists of three UXO objects of different material properties as well as two non-UXO objects, namely an aluminum cylinder and pipe.

Each run of data consists of 800 pings in which the sonar platform moved along the fixed rail in increments of 0.025 m, transmitting and receiving once for each fixed position. Sonar backscatter was received with six hydrophone elements that are arranged in a linear array that is approximately normal to the seafloor, though only data collected by the hydrophone closest to the seafloor were used in this study. The data were sampled at 1 MHz and the sonar platform was tilted at a fixed 20° angle for all runs (angle of the sonar main response axis with respect to the horizontal plane). Since the objects in the PondEx10 experiments are relatively close to one another and moreover presence of reverberation and other artifacts are inevitable, a postprocessing filter [16] was applied to the collected data to isolate the response of each individual object of interest while removing all the interference. This processing utilizes a reversible SAS imaging process, a spatial filtering process using a 2-D Tukey window, and a pseudo-inverse filtering [17]. This inverse filter maps the SAS image back to the pulse-compressed version, which is subsequently transformed to the acoustic color data vectors $\mathbf{x}_i^{(u)}$ for $i = 1, \ldots, M$ using FFT. Thus, each element of the data vector $\mathbf{x}_i^{(u)}$ gives the magnitude response for the ith aspect at a particular frequency. Given the ping rate and beamwidth of the sonar system used here, a total of $M = 40$ aspects are used to classify the object corresponding to a window spanning a range of approximately 20° in aspect angle. Fig. 7(a) and (b) gives an example of the filtered time series and measured acoustic color for an aluminum UXO at a 20° orientation angle, respectively. This acoustic color data is then applied to the synthesis equation given in (7) to extract the manifold features for this UXO object as shown in Fig. 7(c) for the first two features.

As described in Section III, the first step in the classification process depicted in Fig. 1 involves identifying the sequence of

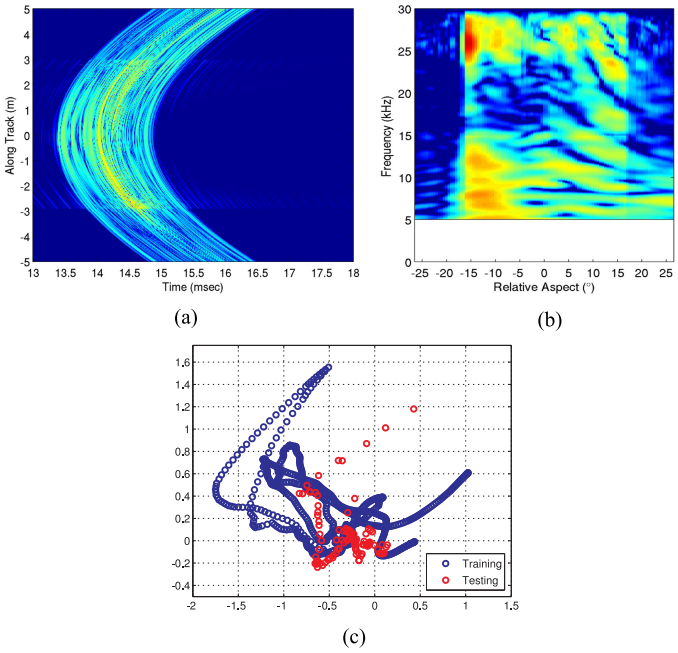

Fig. 7. Filtered time series, target strength, and manifold features for an aluminum UXO at a 20° orientation angle. (a) Filtered time series. (b) Associated target strength. (c) Extracted manifold features (first two) versus training features.

training features for each object type that best matches the extracted feature vectors using the matching index in (8). As mentioned before, this ordered sequence corresponds to the range of aspect angles observed for a given target so that finding the sequence that best matches the data corresponds to estimating the aspect angle of the target. Fig. 3 once again gives an illustration of this process when the features from an aluminum UXO are applied to the training data from an aluminum UXO and aluminum cylinder. Once the aspect is estimated for each object type, the final step in the classification algorithm involves representing each observed feature vector $\mathbf{y}_i^{(u)}$ for $i = 1, \ldots, M$ using training feature vectors local to that observation. To accomplish this, the $P = 10$ training manifold feature vectors corresponding to the jth object that are nearest in aspect to $\mathbf{y}_i^{(u)}$ are used to construct the dictionary matrix $\mathbf{H}_{i,j}$ used in the linear model (9). More specifically, if we let $\left\{\mathbf{y}_k^{(i,j)}\right\}_{k=1}^{P}$ be the P training feature vectors from object type j nearest to the extracted feature vector $\mathbf{y}_i^{(u)}$, then this dictionary matrix is constructed as

$$\mathbf{H}_{i,j} = \left[\mathbf{y}_1^{(i,j)} \ \cdots \ \mathbf{y}_P^{(i,j)}\right] \in \mathbb{R}^{d \times P}.$$

Fig. 8(a) and (b) plots the first two features of $\mathbf{y}_i^{(u)}$ for an aluminum cylinder at an 80° orientation angle with each green dot in both of these figures corresponding to single aspect angle. The blue dots in each of these plots denote the subset of the trained manifold for two different objects that best matches the extracted features using sequence matching index (8): Fig. 8(a) shows the manifold for an aluminum cylinder while Fig. 8(b) shows that for an aluminum UXO. Given the extracted features as well as the trained manifold for each of these two objects,

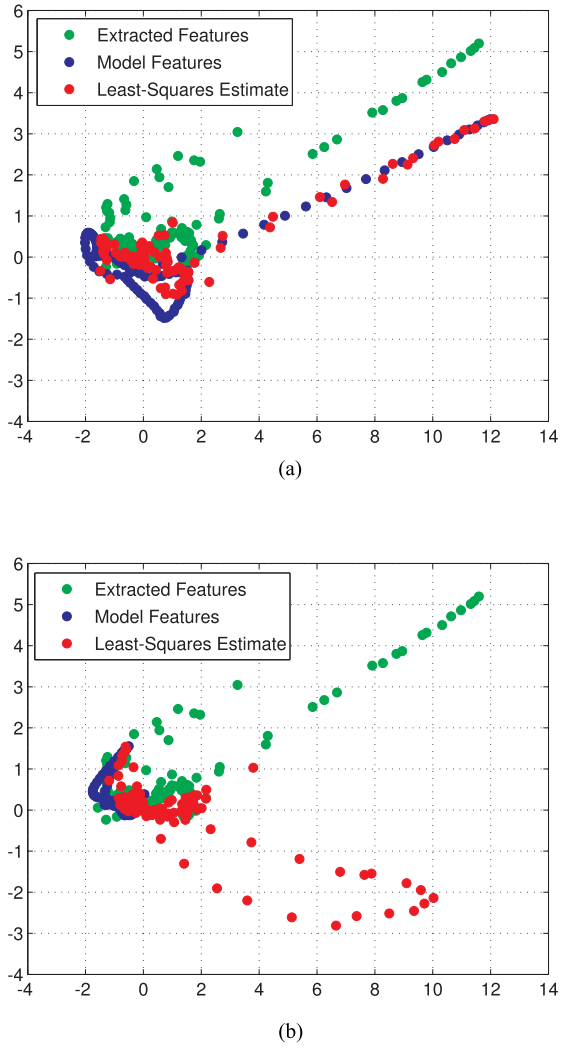

Fig. 8. Comparing the extracted features from an aluminum cylinder to its least-squares estimate $\hat{\mathbf{y}}_{i,j}^{(u)} = \mathbf{H}_{i,j}\hat{\boldsymbol{\theta}}_i$ for two objects. (a) Aluminum cylinder manifold. (b) Aluminum UXO manifold.

each red dot in these figures plots the estimated feature vector

$$\hat{\mathbf{y}}_{i,j}^{(u)} = \mathbf{H}_{i,j}\hat{\boldsymbol{\theta}}_i = \mathbf{P}_{\mathbf{H}_{i,j}}\mathbf{y}_i^{(u)}$$

where $\hat{\boldsymbol{\theta}}_i$ is the least-squares estimate of the unknown vector $\boldsymbol{\theta}_i$ in (11). The vector $\hat{\mathbf{y}}_{i,j}^{(u)}$ provides the best estimate of the observation $\mathbf{y}_i^{(u)}$ in the linear subspace $\mathbf{H}_{i,j}$ by orthogonally projecting the observation into that subspace. Note that, given this definition, the objective function in (15) can be rewritten as

$$\epsilon_j(\mathbf{Y}_u) = \sum_{i=1}^{M} \left|\left|\mathbf{y}_i^{(u)} - \hat{\mathbf{y}}_{i,j}^{(u)}\right|\right|_2^2. \tag{16}$$

That is, the classifier selects the object that produces the closest (in the sum-squared sense) estimates to the measured data. Looking at Fig. 8 for the first two manifold features, it appears that the manifold corresponding to the aluminum cylinder does a much better job of producing estimates that reflect the extracted features than those produced by the manifold corresponding the aluminum UXO.

C. Classification Results

This section presents classification results for the test data sets described in Section IV-B. Two experiments were conducted using data collected from objects at different ranges from the sonar. For both experiments, the training data set included model-generated data as described in Section IV-A for four object types at a 10-m range to construct the manifold features. The trained classifier was then applied to all the test objects listed in Table I from both PondEx09 and PondEx10 data sets. However, in the first experiment test data sets for only 10-m range were used while for the second experiment data collected from an aluminum UXO at a 5-m range were also included. The intent of the first study was to demonstrate the algorithm's ability to discriminate among UXO and non-UXO objects and also study whether the manifold features of the training objects can adequately represent those of the previously unseen objects i.e., the real UXO. However, the goal of the second study was to determine the effects of target range on the manifold features.

The results of the proposed classification method for both experiments were then compared to those of an alternative method in [9] which also relies on the same set of manifold features found by solving (2). Given the set of labeled training features $\mathbf{Y}_s \in \mathbb{R}^{d\times N}$ described in Section II with each feature vector $\mathbf{y}_i^s \in \mathbb{R}^d$ assigned the binary label $c_i \in \{+1, -1\}$ (in this case used to differentiate between UXO and non-UXO objects), the alternative method in [9] performs linear classification in the manifold domain by finding the vector $\mathbf{a} \in \mathbb{R}^d$ that solves

$$\min_{\mathbf{a}\in\mathbb{R}^d} \left|\left|\mathbf{c} - \mathbf{Y}_s^T\mathbf{a}\right|\right|^2 \tag{17}$$

where $\mathbf{c} = [c_1\ c_2\ \cdots\ c_N]^T \in \mathbb{R}^N$. Solving (17) yields the least-squares solution $\mathbf{a} = \left(\mathbf{Y}_s\mathbf{Y}_s^T\right)^{-1}\mathbf{Y}_s\mathbf{c}$. Upon extracting a new test sample $\mathbf{y} \in \mathbb{R}^d$ using the synthesis equation in (6), the estimated class label $\hat{c}$ is assigned using the following rule

$$\hat{c} = \begin{cases} +1, & \text{if } \mathbf{a}^T\mathbf{y} > 0 \\ -1, & \text{if } \mathbf{a}^T\mathbf{y} \leq 0. \end{cases} \tag{18}$$

Thus, one can see that the classifier in (18) relies on the same set of features but, rather than building local linear subspaces that are matched to the aspect of the object, a simple linear combination of the features is taken to perform classification. When applied to the PondEx data sets used here, the class label in (18) is computed for each of the $M = 40$ aspect angles and a majority rule is enforced to assign an overall label to that object.

Fig. 9 displays the ROC curve for both the proposed classifier in (14) (plotted with a dark shaded line) as well as that of the method in (18) (plotted with a lighter shaded line). This figure plots the probability of correct UXO classification P_{cc} versus the probability of false alarm P_{fa}. The knee-point of the ROC curve (the point at which $P_{cc} + P_{fa} = 1$) denoted by a circle in these plots indicates an approximate 10% improvement for the proposed method over the alternative in [9].

Table II gives the overall correct classification rates for each object in the PondEx data set and for both classification methods. Comparing the results in this table, one can see that with the exception of the aluminum pipe, the proposed classifier

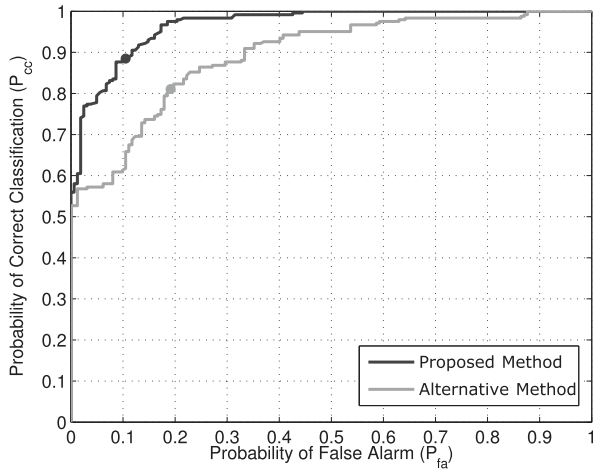

Fig. 9. ROC curve for the PondEx data set.

TABLE II
PONDEX DATA SET CORRECT CLASSIFICATION RATES

		Proposed Method	Alternative Method
UXO	Al UXO	93.8%	71.6%
	Steel UXO	88.9%	75.3%
	Real UXO	66.7%	55.6%
Non-UXO	Al Cylinder	92.6%	77.8%
	Al Pipe	91.4%	100.0%

in (14) achieved 10%–20% improvement in classification performance. This improvement in classification accuracy observed in both Fig. 9 and Table II is attributed to the fact that the alternative method in [9] relies on the use of a global linear classifier to discriminate one class from another while the proposed method relies on local subspace representations to model the measured data. From Table II, it can also be noted that both methods performed relatively well on the four objects (i.e., the aluminum UXO, steel UXO, aluminum cylinder, and aluminum pipe) learned by each classifier though the training data were synthetically generated. However, the discrimination performance for both classifiers was severely degraded for the real UXO, which was not included in the training data set. This degradation in performance is likely because the manifold features of the synthetic data for other UXO objects were not capable of capturing the properties of this particular object.

As described before, for the second experiment the trained classifiers in the first experiment were applied to the test data of a proud aluminum UXO at 5-m range to study the impact of sonar range (or grazing angle θ_g) on the performance of both classifiers. Fig. 10(a) and (b) clearly shows major differences in acoustic color responses observed for this target at 10- and 5-m range. Note that the data at 270° aspect angle was not used here since it exhibits discontinuity due to wrapping artifact. This is likely from the edge effects of the interpolation process in the

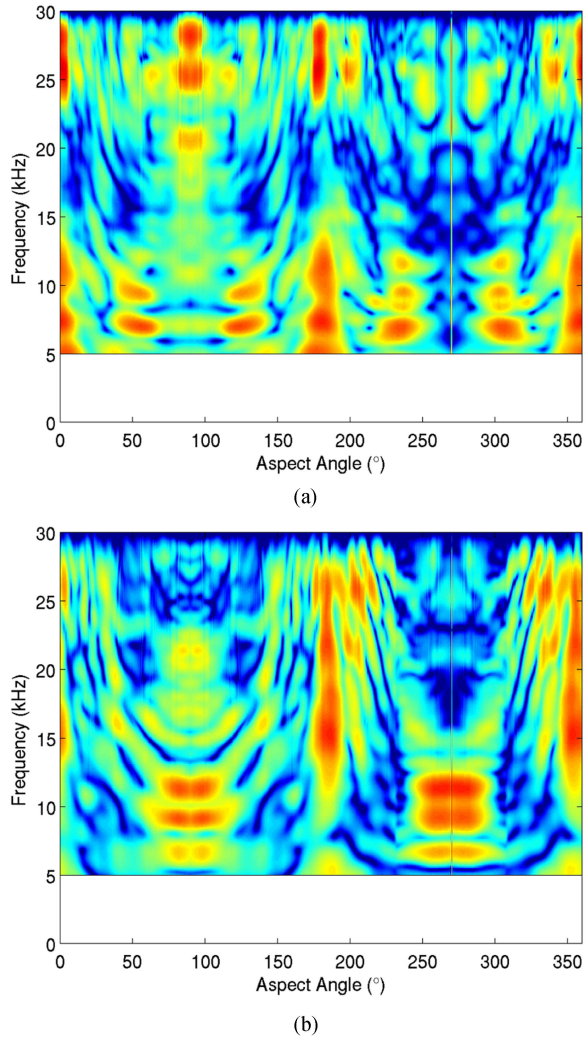

Fig. 10. Acoustic color responses for a proud aluminum UXO object at various range from the sonar: (a) 10-m range; and (b) 5-m range.

TABLE III
AL UXO PERCENT CORRECT CLASSIFICATION RATES AT DIFFERENT RANGES

	Proposed Method	Alternative Method
10-m range	93.8%	71.6%
5-m range	81.5%	72.8%

original data. Table III gives the correct classification rates of this UXO object under both scenarios and for both classification methods. Similar to the results in Table II, one can once again see a decline in the performance of both classifiers though the proposed method exhibits about 10% improvement in performance over that of the alternative method in [9]. The decline in the performance is mainly attributed to the fact that compact sand has a critical angle of around $\theta_c \approx 34^\circ$ and while for 10-m range $\theta_g \approx 21^\circ < \theta_c$, for 5-m range $\theta_g \approx 37^\circ > \theta_c$ which results in significant ($\approx$60%) loss of the acoustic energy into the sediment [15]. Nevertheless, the proposed method performed fairly well with a correct classification rate of over 80% even though the simulated training data set did not include sonar data at 5-m range.

V. CONCLUSION

This paper considered the development and testing of a manifold-based feature extraction and classification strategy for discriminating UXO versus non-UXO objects. The proposed feature extraction and classification system are designed based on the assumption that the data lies in some unknown low-dimensional subspace, which is globally nonlinear but locally linear. Based on this premise, a feature extraction technique using the LE [7] algorithm was proposed, which produces a set of low-dimensional features that respect distances among training points in the high-dimensional space. Extending the algorithm to newly observed testing data yielded an out-of-sample embedding procedure for the purposes of feature extraction.

Given this set of low-dimensional features, a multiaspect classification method was introduced to discriminate among UXO and non-UXO objects. The first step in the algorithm involves identifying the subset of training features that best matches the extracted features using a sequence matching measure. This process corresponds to estimating the aspect angle of the object. Once this sequence has been identified, the set of training features from each object type nearest the extracted features is then used to form a local linear subspace to represent those features. The most likely class label is then selected by finding the class that minimizes the error in representing the extracted features.

The performance of the classifier was then demonstrated on the PondEx data sets and compared to that of an alternative method that relies on the same set of features but constructs a linear classifier to discriminate one class from another. For this study, both methods were trained using model-generated sonar data to discriminate two non-UXO objects (cylinder and pipe) from two replica UXO (aluminum and steel UXO) objects with different shapes and material properties. The method was then applied to testing data from these four objects as well as a previously unseen real UXO when each object was located 10 m from the rail and sitting proud on the seafloor. For the four objects used to train the classifier, the proposed method was able to correctly classify over 93% of the testing data in these data sets versus an 82% correct classification rate for that in [9]. Moreover, the method was able to correctly classify nearly 70% of the testing data for the real UXO object which was not included in the simulated training data set. A second test was then conducted to observe how changes in the target's range or grazing angle affect the performance of the classifiers. It was observed that the method did indeed exhibit a deterioration in classification performance with the sonar grazing angle significantly deviates from what was used for training. Nevertheless, the system trained with the simulated data of 10-m range was still able to correctly classify over 80% of the data in these scenarios versus only a little over 70% for the alternative method in [9]. Future testing should include more realistic data sets with more object variations.

ACKNOWLEDGMENT

The authors would like to thank NSWC-Panama City and the Applied Physics Lab—University of Washington for support and providing the data used in this study.

REFERENCES

[1] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Recognition*, 2nd ed. Hoboken, NJ, USA: Wiley, 2000.

[2] J. Tenenbaum, V. de Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, Dec. 2000.

[3] V. de Silva and J. Tenenbaum, "Global versus local methods for nonlinear dimensionality reduction," *Adv. Neural Inf. Process. Syst.*, vol. 15, pp. 705–712, 2003.

[4] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, 2000.

[5] L. K. Saul and S. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *J. Mach. Learn. Res.*, vol. 4, pp. 119–155, Dec. 2003.

[6] K. Weinberger and L. Saul, "Unsupervised learning of image manifolds by semidefinite programming," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, 2004, vol. 2, pp. 988–995.

[7] M. Belkin and P. Niyogi, "Laplacian Eigenmaps and spectral techniques for embedding and clustering," *Adv. Neural Inf. Process. Syst.*, vol. 14, pp. 585–591, 2001.

[8] M. A. Davenport, C. Hegde, M. B. Wakin, and R. G. Baraniuk, "Manifold-based approaches for improved classification," in *Proc. Neural Inf. Process. Syst. Workshop Topology Learn.*, Whistler, BC, Canada, Dec. 2007.

[9] M. Belkin and P. Niyogi, "Using manifold structure for partially labelled classification," *Adv. Neural Inf. Process. Syst.*, vol. 15, pp. 929–936, 2003.

[10] J. Lee and M. Verleysen, *Nonlinear Dimensionality Reduction*. Berlin, Germany: Springer, 2007.

[11] N. H. Klausner, "Environmentally adaptive UXO detection and classification systems," Inf. Syst. Technol., Inc., Fort Collins, CO, USA, Final Rep. SERDP Project MR-2417, Apr. 2016.

[12] M. A. Carreira-Perpinan, "Continous latent variable models for dimensionality reduction and sequential data reconstruction," Ph.D. dissertation, Dept. Comput. Sci., Univ. Sheffield, Sheffield, U.K., 2001.

[13] J. Hall, M. Azimi-Sadjadi, S. Kargl, Y. Zhao, and K. Williams, "Underwater unexploded ordnance (UXO) classification using a matched subspace classifier with adaptive dictionaries," *IEEE J. Ocean. Eng.*, 2018, DOI: 10.1109/JOE.2018.2835538.

[14] A. H. A. Salberg and L. L. Scharf, "Robust multidimensional matched subspace classifiers based on weighted least-squares," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 873–880, Mar. 2007.

[15] S. Kargl, A. Espana, K. Williams, J. Kennedy, and J. Lopes, "Scattering from objects at a water-sediment interface: Experiment, high-speed and high-fidelity models, and physical insight," *IEEE J. Ocean. Eng.*, vol. 40, no. 3, pp. 632–642, Jul. 2015.

[16] S. Kargl and K. Williams, "Full scale measurement and modeling of the acoustic response of proud and buried munitions at frequencies from 1–30 kHz," Washington Univ. Seattle Applied Physics Lab, Seattle, WA, USA, Final Rep. SERDP Project MR-1665, pp. 3–5, May 2012.

[17] T. Marston, P. Marston, and K. Williams, "Scattering resonances, filtering with reversible SAS processing, and applications of quantitative ray theory," in *Proc. MTS/IEEE OCEANS Conf.*, Seattle, WA, USA, Sep. 2010, pp. 1–9.

Nick H. Klausner (S'08–M'14) received the M.S. and Ph.D. degrees in electrical and computer engineering from Colorado State University, Fort Collins, CO, USA, in 2010 and 2014, respectively.

During 2014–2017, he was a Research Scientist with Information System Technologies, Inc., Fort Collins. He is currently with Numerica Corporation, Fort Collins, CO, USA. His main research interests include statistical signal and image processing.

Mahmood R. Azimi-Sadjadi (M'81–LM'18) received the M.S. and Ph.D. degrees from the Imperial College of Science and Technology, University of London, London, U.K., in 1978 and 1982, respectively, both in electrical engineering with specialization in digital signal/image processing.

He is currently a Full Professor with the Electrical and Computer Engineering Department, Colorado State University (CSU), Fort Collins, CO, USA. He is also serving as the Director of the Digital Signal/Image Laboratory, CSU. He has coauthored the book *Digital Filtering in One and Two Dimensions* (Plenum Press, 1989). His main areas of interest include statistical signal and image processing, target detection, classification and tracking, sensor array processing, machine learning, and distributed sensor networks.

Dr. Azimi-Sadjadi served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON NEURAL NETWORKS.