

# Analysis of Signals Under Compositional Noise With Applications to SONAR Data

J. Derek Tucker, *Student Member, IEEE*, Wei Wu, and Anuj Srivastava *Senior Member, IEEE*

## Abstract

We consider the problem of denoising and classification of SONAR signals observed under compositional noise, i.e., they have been warped randomly along the  $x$ -axis. The traditional techniques do not account for such noise and, consequently, cannot provide a robust classification of signals. We apply a recent framework that: (1) uses a distance-based objective function for data alignment and noise reduction, and (2) leads to warping-invariant distances between signals for robust clustering and classification. We use this framework to introduce two distances that can be used for signal classification: (a) a  $y$ -distance, which is the distance between the aligned signals, and (b) an  $x$ -distance that measures the amount of warping needed to align the signals. We focus on the task of clustering and classifying objects, using acoustic spectrum (acoustic-color), which is complicated by the uncertainties in aspect angles at data collections. Small changes in the aspect angles corrupt signals in a way that amounts to compositional noise. We demonstrate the use of the developed metrics in classification of acoustic color data and highlight improvements in signal classification over current methods.

## Index Terms

compositional noise, functional data analysis, random warping, spectral signal classification, signal registration, SONAR.

J. D. Tucker, W. Wu, and A. Srivastava are with the Department of Statistics, Florida State University, Tallahassee, FL 32306, USA (e-mail: {dtucker,wwu,anuj}@stat.fsu.edu).

J. D. Tucker is with the Naval Surface Warfare Center Panama City Division, Panama City, FL 32407-7001 USA (email: james.d.tucker@navy.mil).

This work was supported by the Naval Surface Warfare Center Panama City Division In-house Laboratory Independent Research program funded by the Office of Naval Research (ONR) and ONR grant # N00014-09-1-0064

Approved for public release; distribution is unlimited.

## I. INTRODUCTION

The problem of underwater object detection and classification, using SONAR, has attracted a substantial amount of attention [1]–[5]. This problem is complicated due to various factors, such as variations in operating and environmental conditions, presence of spatially varying clutter, and variations in target shapes, compositions, and orientation. Moreover, bottom features such as coral reefs, sand formations, and vegetation, may totally obscure a target or confuse the classification process. Consequently, a robust classification system should be able to quantify changes between the returns from the bottom and any target activity in SONAR data. Thus, a robust system designed to mitigate false alarms, in various clutter density scenarios is desirable.

Considerable research has been devoted to the development of methodologies to detect and classify underwater objects utilizing SONAR imagery. Dobeck [1], [6] utilized a nonlinear matched filter to identify mine-size regions that match the target template in a side-scan SONAR image. For each detected region, several features were extracted based on the size, shape, and strength of the target template. A stepwise feature selection process was then used to determine the subset of features that maximizes the probability of detection and classification. A  $K$ -nearest neighbor and an optimal discrimination filter classifier were used to classify each feature vector. The decisions of the two classifiers were then fused to generate the final decision. In [2], the adaptive clutter filter detector in [7] was individually applied to three different SONAR images varying in frequency and bandwidth. Final classification is done by using a nonlinear log-likelihood ratio test, on an optimal set of features, where the decisions of the individual detector and classifier are fused. Recently in [4], a new coherence-based detection framework was developed for the dual-sensor problem using Canonical Correlation Analysis (CCA) that can be applied to the data collected by two disparate SONAR systems. Using this method allows for the simultaneous detection and feature extraction of coherent target information among two SONAR images.

These methods all use traditional synthetic aperture SONAR (SAS) images, which often provides high quality images of proud targets that are useful for image based detection, localization, and identification algorithms. For targets in highly cluttered environments, or where the target is partially or fully buried in the sediment, the images are usually blurred with less structure definition; hence, target identification from these images is more difficult. Moreover, one may

have access to only a small range of angles, lacking the ability to use a full frequency versus aspect structure. Generating acoustic color data products is one way to overcome these shortcomings. Acoustic color [8], [9] is a simple, spectral-based method that generates a normalized plot showing the strengths of the return signatures off an object, in individual frequency bands at various aspects, that may provide features useful for identification. Acoustic color data provides structural acoustic and target information for a wider range of aspect angles, which can increase algorithm performance over image-based methods. The problem then shifts to statistical analysis on these acoustic color images, which are essentially a two-dimensional function of frequency and aspect representing relative power of the target return.

We can treat these spectral signatures as functions (of frequency) and analyze them using tools from statistical function analysis, such as functional PCA under Euclidean metrics [10]. However, this spectral data is sensitive to sensor placement and aspect alignment between sensor and targets. Even small changes in aspect angles can result in nonlinear shifts of peaks and valleys in spectral signatures, a phenomenon that is naturally encoded as random warping along the  $x$  axis or *compositional noise*. Shown in Fig. 1(a) are two observed signatures (upper and middle panels) of the same target from two similar angles, one as the solid line and the other as the dashed line, where it is easy to see shifts in peaks and valleys between the two signals. The bottom panel shows an alignment of the dashed line to the solid line, denoting removal of estimated warping noise, and the optimal warping function used here is shown in the right panel. The alignment is performed using the method that will be described in Section III. Fig. 1(b) is the same as (a) except that the two signatures are taken from two different target classes, where a more drastic warping is needed to align the signals. This motivates the use of a measure of warping as a separate metric by itself for target classification. Fig. 8 shows more examples of raw data from nine target classes and one can see the shifting of peaks in signals of the same class. In practical situations where the data is not collected in a controlled environment, one will not know exactly the aspect angles for data collection. Therefore, automated data warping and alignment are crucial to handling compositional noise for an improvement in classification. Fig. 2(a) presents the difference between a pair of signals for one target type with 2 degrees of separation in aspect angle before and after alignment. There is a smaller difference after alignment which confirms that small changes in aspect angle causes warping of the received signal. Fig. 2(b) presents the distance between a pair of signals for multiple pairs before and

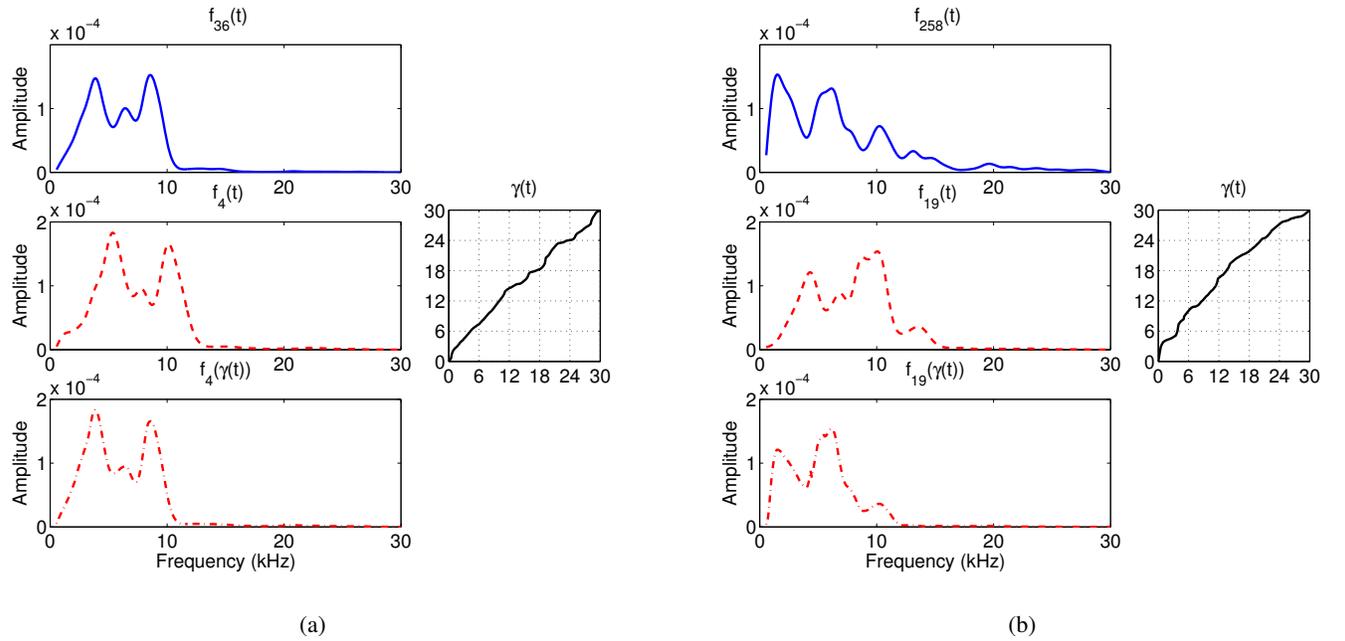
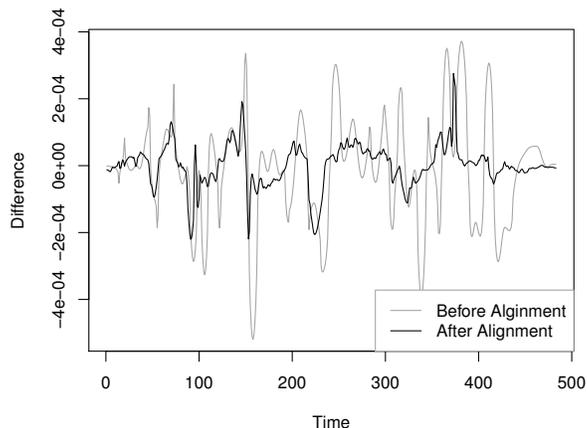


Fig. 1. (a) An example of frequency warping between two functions in the same class:  $f_{36}$  (upper panel),  $f_4$  (middle panel, left) and  $f_4 \circ \gamma$  (lower panel), where  $\gamma$  is the optimal warping (middle panel, right) between  $f_4$  and  $f_{36}$ . (b) Same as (a) except that the functions  $f_{19}$  and  $f_{258}$  are in different classes.

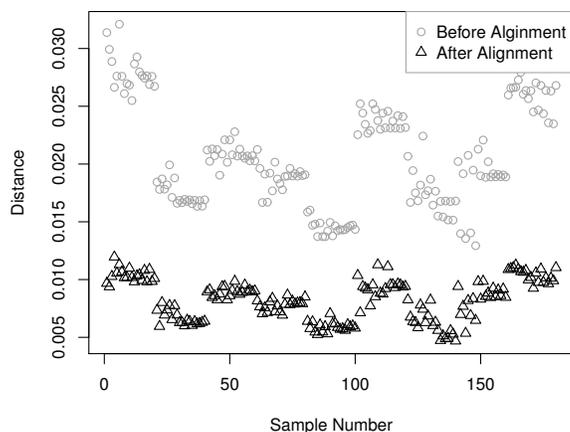
after alignment which, again, confirms our hypothesis that small degree of aspect angle separation is representative of compositional noise. We can make the conclusion that our original signals (circles) contain both compositional and additive noise and, that after alignment (triangles), contain only additive noise, as the compositional noise is removed during the alignment process.

Our contributions in this manuscript are: 1) We particularize the framework presented in [11], [12] for the alignment and denoising of spectral signatures for targets in presence of compositional noise; and evaluate it empirically in presence of both compositional and additive noise. 2) We introduce two metrics for classification of targets from SONAR data: one of them is invariant to compositional noise and the other measures the warping itself.

The rest of this paper is organized as follows: Section II reviews the main ideas in function registration in the current literature and discusses their major limitations. Section III presents a differential geometric approach for signal alignment and suggests two metrics for target classification. It also presents a computational procedure for alignment of signals observed under



(a) Pair of Signals



(b) Multiple Pairs

Fig. 2. Compositional noise amongst pairs of signals with  $2^\circ$  of separation in aspect angle.

random warping. Section IV describes experimental results on acoustic color signals, including the classification problem, and compares results with current conventional methods. Finally, conclusions and observations are offered in Section V.

## II. CURRENT IDEAS, LIMITATIONS, AND CHALLENGES

First, we introduce some notation which is shown in Table I. Let  $f_1, f_2, \dots, f_n : [0, \Omega] \rightarrow \mathbb{R}$  be real-valued signals that are observed and let  $\mathcal{F}$  be the set of all such signals. Let  $\Gamma$  be the set

Symbol	Meaning
$[0, \Omega]$	Frequency Domain
$\omega$	element of $[0, \Omega]$
$f_i$	real-valued signal on $[0, \Omega] \rightarrow \mathbb{R}$
$\mathcal{F}$	set of all $f_i$
$g$	mean function
$\Gamma$	set of boundary-preserving diffeomorphisms on $[0, \Omega]$
$\gamma_i$	element of $\Gamma$
$\gamma_{id}$	identity element of $\Gamma$

TABLE I  
SUMMARY OF NOTATION.

of boundary-preserving diffeomorphisms of the interval  $[0, \Omega]$ :  $\Gamma = \{\gamma : [0, \Omega] \rightarrow [0, \Omega] \mid \gamma(0) = 0, \gamma(\Omega) = \Omega, \gamma \text{ is a diffeomorphism}\}$ . Elements of  $\Gamma$  play the role of warping functions or compositional noise. That is, for any  $\gamma \in \Gamma$ ,  $f_i \circ \gamma$  denotes a composition or a warping of the signal  $f_i$  by  $\gamma$ . With the composition operation, the set  $\Gamma$  is a group with the identity element  $\gamma_{id}(\omega) = \omega$ .

For a variety of reasons, the observed signals get corrupted, including by random warping, during observations. What we mean by warping is that a real-valued signal  $g$  on the frequency domain  $[0, \Omega]$  is composed with a random warping function  $\gamma_i : [0, \Omega] \rightarrow [0, \Omega]$ , resulting in a nonlinear frequency-shift of the *locations* of peaks and valleys, but not the heights of those peaks and valleys. More generally, a signal model incorporating compositional as well as a additive noise model is given by:

$$f_i(\omega) = c_i g(\gamma_i(\omega)) + n_i(\omega), \quad (1)$$

where  $c_i$ 's are random scalars,  $\gamma_i$ 's are random warping functions, and  $n_i$ 's are independent Gaussian noise. This model contains three types of noise sources: multiplicative, compositional, and additive. The goal is to use observations  $\{f_i, i = 1, 2, \dots, n\}$  to denoise the signal, i.e., estimate  $\gamma_i$ , and to classify  $f_i$  into pre-determined classes.

**Signal Alignment and Denoising:** Although the classical cross-sectional mean (i.e.,  $\bar{f} = \frac{1}{n} \sum_{i=1}^n f_i$ ) is optimal for denoising a signal in presence of (zero-mean) additive Gaussian noise, it is not optimal for the case of compositional noise. Consider the reduced model of  $f_i(\omega) = g(\gamma_i(\omega))$

for all  $i$ . (We will consider this reduced model for the rest of the paper). There are three sets of variables here,  $\{f_i\}$ ,  $g$ , and  $\{\gamma_i\}$ , and, given any two, one can easily estimate the third. As an example, given  $f_i$  and  $g$ , the solution for  $\gamma_i$  is based on the dynamic programming algorithm (DPA) [13] or other estimation theoretic methods [14], [15]. This estimation is popular in speech processing where the observation signal is matched to signals of known speech using DPA [16], often called dynamic time warping in this context. The other problem, of estimating  $g$  given  $\gamma_i$  and  $f_i$ , is also simple, since  $g = f_i \circ \gamma_i^{-1}$ , as one has to just compute the inverse of the  $\gamma_i$ 's. Thus, the problem of alignment of signals is identical to the problem of denoising signals when only the compositional noise is present.

The situation gets more complicated if both  $g$  and  $\{\gamma_i\}$  are unknown. If we assume that  $g$  takes a parametric form; e.g., a superposition of Gaussians or exponentials with different parameters, then perhaps we can set up an estimation problem for these parameters from the observed data [17]. Our interest is in the nonparametric case, i.e., when  $g$  is a full function on  $[0, \Omega]$ . Notice that the estimation problem for this case, as stated, is ill-defined. One solution comes from Kurttek et al. [11], [12], where they derive a fundamental theoretical framework for estimation of  $g$  given  $\{f_i\}$ , under the assumption that the mean of  $\{\gamma_i^{-1}\}$  is  $\gamma_{id}$ . The estimator for  $g$  is obtained by **aligning** the given signals  $\{f_i\}$  by warping their horizontal axes (which is mathematically equivalent to estimating  $\gamma_i$ 's), and then computing the mean of the aligned signals. Since this estimation is based on multiple signal alignment, a well-studied problem in the literature, we need to discuss the strengths of this approach relative to the current literature. The most common approach for *pairwise* alignment of functions is based on the following optimization:

$$\min_{\gamma \in \Gamma} (\|(f_1 \circ \gamma) - f_2\|^2 + \lambda \mathcal{R}(\gamma)) \quad , \quad (2)$$

where  $\mathcal{R}$  penalizes the roughness of  $\gamma$  and  $\lambda > 0$  is a constant. The minimization over  $\gamma$  is typically done using DPA. One obvious problem with this alignment procedure is that it is not symmetric, i.e., the solution for  $(f_1, f_2)$  is not the same as for  $(f_2, f_1)$ . Another issue comes in the choice of  $\lambda$  in Eqn. 2. While it allows a user to adapt the regularization term to an application, it may not be straightforward to select this automatically in a general situation. It will be good to have an objective function that already incorporates both the terms in a single functional, as is the case in the proposed approach.

In alignment of multiple signals  $\{f_i, i = 1, 2, \dots, n\}$ , a common iterative approach is to define

a *template* and use this template to align the individual signals. In each iteration one: (1) updates the template by taking the cross-sectional mean of the currently aligned signals, and (2) aligns each  $f_i$  to the template in a pairwise manner, described above. Unfortunately, this approach often does not provide desirable results, both theoretically and empirically, unless a proper distance function is used to align and average signals. The framework introduced in [11], [12] avoids these problems by defining a distance-based objective function that changes only with the change in the registration. It is symmetric by definition and does not require a choice of  $\lambda$  since we do not have two separate terms in the objective function. Additionally, it has been shown in [11], [12] that as the sample size  $n \rightarrow \infty$ , the mean of the aligned functions in our framework converges to  $g$ .

**Metrics for Signal Classification:** The second goal is deriving metrics for comparing signals, especially metrics that are either invariant or robust to compositional noise. Given any  $f_1$  and  $f_2$ , the problem is how to compare  $f_1$  and  $f_2$  in a manner that is invariant to random warpings:  $\gamma_1, \gamma_2$ . In other words, we would like a distance  $d(\cdot, \cdot)$  such that

$$d(f_1 \circ \gamma_1, f_2 \circ \gamma_2) = d(f_1, f_2) \quad \forall \gamma_1, \gamma_2 \in \Gamma . \quad (3)$$

In presence of only the zero-mean additive noise, a commonly used distance is the  $\mathbb{L}^2$  norm of their difference  $\|f_1 - f_2\|$ . One may be tempted to extend this distance to the case of compositional noise using a quantity of the type:  $\min_{\gamma \in \Gamma} \|f_1 - f_2 \circ \gamma\|$ , or  $\min_{\gamma_1, \gamma_2 \in \Gamma} \|f_1 \circ \gamma_1 - f_2 \circ \gamma_2\|$ , or  $\min_{\gamma \in \Gamma} (\|(f_1 \circ \gamma) - f_2\|^2 + \lambda \mathcal{R}(\gamma))$ , but they all suffer from the same issue. They are not proper distances and/or they do not satisfy Eqn. 3! We present a comprehensive framework for alignment and estimation of acoustic signals and use it for the classification of spectral acoustic signatures into pre-determined classes. It is based on introducing a mathematical expression for representing a function where the standard metric is isometric. An important reason for selecting this representation is that it leads to a distance between signals that is invariant to their random warpings. Not only is it robust to compositional noise present in the signals, but it also leads to another metric that measures the amount of warping. Different levels of warpings in different classes also lead to classification using such metrics.

### III. MATHEMATICAL FRAMEWORK

In this section, we adapt the theoretical framework presented in a recent report [12] and conference papers [11], [18] for use in SONAR signal processing. This resulting framework achieves three important goals: (1) completely automated alignment of signals using nonlinear warpings, (2) estimation of underlying signals observed under random warpings, and (3) derivation of individual phase and amplitude metrics for comparing and classifying signals. For a more comprehensive introduction on this theory, including asymptotic theorems and estimator convergences, we refer the reader to [12].

#### A. Warping-Invariant Metrics

Let  $f$  be an acoustic signal viewed as a real-valued function with the domain  $[0, \Omega]$ . For concreteness, only functions that are absolutely continuous on  $[0, \Omega]$  will be considered; let  $\mathcal{F}$  denote the set of all such functions. In practice, since the observed data is discrete, this assumption is not a restriction. Our first goal is to find a distance function that will be invariant to random warpings of the input functions. To address this, Srivastava et al. [12] introduced a square-root representation for functions. This function,  $q : [0, 1] \rightarrow \mathbb{R}$ , is called the *square-root slope function* or SRSF of  $f$ , and is defined in the following form:

$$q(\omega) = \text{sign}(\dot{f}(\omega))\sqrt{|\dot{f}(\omega)|} .$$

SRSF is a particularization of the square-root velocity function used in shape analysis of Euclidean curves [19]. It can be shown that if the function  $f$  is absolutely continuous, then the resulting SRSF is square-integrable. Thus, we will define  $\mathbb{L}^2([0, \Omega], \mathbb{R})$ , or simply  $\mathbb{L}^2$ , to be the set of all SRSFs. For every  $q \in \mathbb{L}^2$  and a fixed  $t \in [0, \Omega]$ , the function  $f$  can be obtained precisely using the equation:

$$f(\omega) = f(0) + \int_0^\omega q(s)|q(s)|ds . \quad (4)$$

Thus, the representation  $f \Leftrightarrow$  pair  $(f(0)$  and  $q)$  is invertible. The most important property of this representation is, if we warp a function  $f$  by  $\gamma$ , the SRSF of  $f \circ \gamma$  is given by:  $\tilde{q}(\omega) = (q, \gamma)(\omega) = q(\gamma(\omega))\sqrt{\dot{\gamma}(\omega)}$ . With this expression, it can be shown that for any  $f_1, f_2 \in \mathcal{F}$  and  $\gamma \in \Gamma$ , we have

$$\|q_1 - q_2\| = \|(q_1, \gamma) - (q_2, \gamma)\| .$$

where  $q_1$  and  $q_2$  are the SRSFs of  $f_1$  and  $f_2$ , respectively. This is called the *isometry* property.

Why is this property important? The reason is that it leads to a distance between signals that is invariant to their random warpings. This distance is defined as follows.

*Definition 1 (Amplitude or  $y$  distance):* For any two signals,  $f_1, f_2 \in \mathcal{F}$ , and the corresponding SRSFs,  $q_1, q_2 \in \mathbb{L}^2$ , we define the amplitude or the  $y$  distance  $D_y$  to be:

$$D_y(f_1, f_2) = \inf_{\gamma \in \Gamma} \|q_1 - (q_2 \circ \gamma)\sqrt{\dot{\gamma}}\|.$$

The advantage of this is that  $D_y$  is symmetric, positive definite, and satisfies the triangle inequality. Technically, this is a proper distance, not on  $\mathcal{F}$  but on the quotient space  $\mathcal{F}/\Gamma$ . Moreover, it is invariant to the random warpings of the input signals, i.e.,  $D_y(f_1 \circ \gamma_1, f_2 \circ \gamma_2) = D_y(f_1, f_2)$  for all  $\gamma_1, \gamma_2 \in \Gamma$ .

It is quite possible that the level of warping may be different in different signal classes and one can also use that for classification. Toward this goal, we define another metric  $D_x$  that compares the relative warping needed to align any two signals. Therefore, for any two functions  $f_1, f_2 \in \mathcal{F}$  and the corresponding SRSFs,  $q_1, q_2 \in \mathbb{L}^2$ , let  $\gamma^*$  be given by:

$$\gamma^* = \operatorname{argmin}_{\gamma \in \Gamma} \|q_1 - (q_2, \gamma)\| = \operatorname{argmin}_{\gamma \in \Gamma} \|q_1 - (q_2 \circ \gamma)\sqrt{\dot{\gamma}}\|. \quad (5)$$

If  $\gamma^* = \gamma_{id}$ , then no warping is needed or the functions are perfectly aligned. Therefore, it makes sense to use the difference between  $\gamma^*$  and  $\gamma_{id}$ , in the set  $\Gamma$ , to define  $D_x$ . If we let  $\gamma_1, \gamma_2, \dots, \gamma_n \in \Gamma$  be a set of observed warping functions, our goal is to find a proper distance on this space.

However, since  $\Gamma$  is a nonlinear manifold, it is difficult to compute a proper distance. If we use a convenient transformation similar to the definition of SRSF, we can overcome the nonlinearity. We are going to represent  $\gamma \in \Gamma$  by the square-root of its derivative,  $\psi(\omega) = \sqrt{\dot{\gamma}(\omega)}$ . This is the same as the SRSF as defined earlier and takes this form since  $\dot{\gamma} > 0$  by definition. The identity function,  $\gamma_{id}$ , maps to a constant function of  $\psi_{id}(\omega) = 1$  and one can reconstruct  $\gamma$  from  $\psi$  using  $\gamma(\omega) = \int_0^\omega \psi(s)^2 ds$ . The advantage of this representation is that  $\|\psi\|^2 = \int_0^\Omega \psi(\omega)^2 d\omega = \int_0^\Omega \dot{\gamma}(\omega) d\omega = \gamma(\Omega) - \gamma(0) = \Omega$ . This means that the set of all  $\psi$ s is a Hilbert sphere in  $\mathbb{L}^2$ . In other words, the square-root representation simplifies the nonlinear structure of  $\Gamma$  to a sphere with much simpler geometry.

Therefore, we can define a distance between two warping functions as the arc-length between their corresponding SRSFs on a sphere with radius  $\sqrt{\Omega}$ . This is defined as follows.

*Definition 2 (Phase or  $x$  distance):* For any two functions,  $f_1, f_2 \in \mathcal{F}$ , let  $\gamma^*$  be the optimal frequency warping function as given in Eqn. 5. Then, the horizontal distance between them,  $D_x(f_1, f_2)$ , is defined to be:

$$\begin{aligned} D_x(f_1, f_2) &= \sqrt{\Omega} \cos^{-1} \left( \frac{\langle \psi^*(\omega), \psi_{id}(\omega) \rangle}{\Omega} \right) \\ &= \sqrt{\Omega} \cos^{-1} \left( \frac{\int_0^\Omega \psi^*(\omega) d\omega}{\Omega} \right). \end{aligned}$$

where  $\langle \cdot, \cdot \rangle$  denotes the standard inner product operation in the  $\mathbb{L}^2$  space.

Figure 3 shows a toy example which illustrates the advantages of using  $D_y$  and  $D_x$  in presence of compositional noise. This example has two well separated signal classes, each class having three functions. The functions from Class 1 are all one-period sine waves with amplitude 1, slightly shifted from each other. The functions in Class 2 are the same except smaller amplitudes (0.5) and a larger shift from the first class. The  $6 \times 6$  distance matrices for these signals under  $D_y$ ,  $D_x$ , and  $\mathbb{L}^2$  are shown in the remaining panels. One can see that both  $D_y$  and  $D_x$  will outperform  $\mathbb{L}^2$  in classifying these signals.

### B. Multiple Signal Alignment

The second goal of this paper is to align signals observed in the presence of compositional noise. In this section, we describe an algorithm for the alignment of  $\{f_i\}$ , and refer the reader to [11], [12] for the proof of consistency and underlying asymptotic behavior. This alignment is a two step process – in the first step, we find a set of mean functions and, in the second step, we align the given functions to a particular member of this set.

For a given collection of functions  $f_1, f_2, \dots, f_n \in \mathcal{F}$ , let  $q_1, q_2, \dots, q_n$  denote their SRSFs, respectively. A notion of a mean of these functions is defined as follows:

*Definition 3:* Define the Karcher mean of the given functions as the function that minimizes

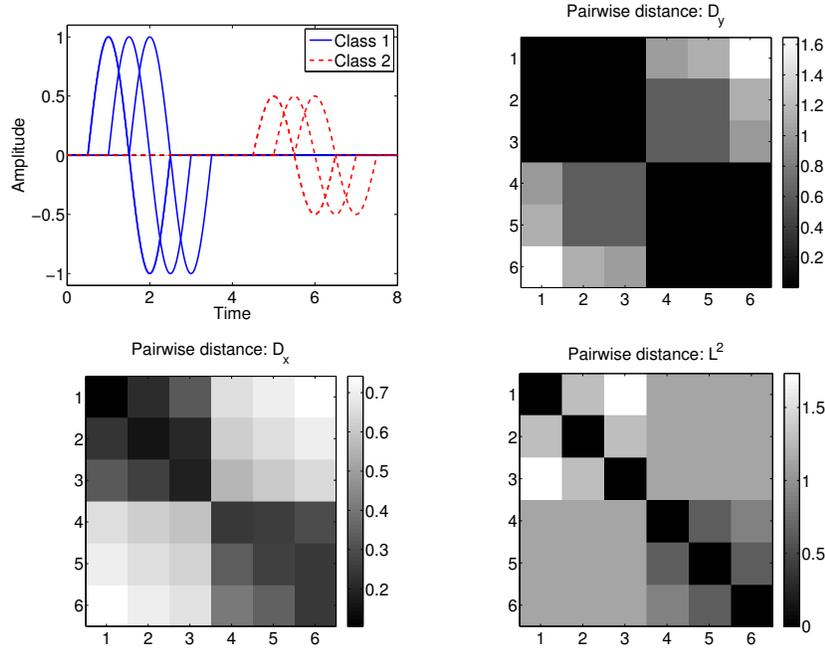


Fig. 3. An example of two-class data where  $D_y$  and  $D_x$  outperform  $\mathbb{L}^2$  distance in classification.

the sum of squares of distances according to:

$$\begin{aligned} \mu_f &= \operatorname{argmin}_{f \in \mathcal{F}} \sum_{i=1}^n D_y(f, f_i)^2 \\ \mu_q &= \operatorname{argmin}_{q \in \mathbb{L}^2} \sum_{i=1}^n \left( \inf_{\gamma_i \in \Gamma} \|q - (q_i, \gamma_i)\|^2 \right). \end{aligned} \quad (6)$$

We denote the mean function by  $\mu_f$  and its SRSF by  $\mu_q$ . It is important to note that if  $\mu_f$  is a minimizer in the above equation, then so is  $\mu_f \circ \gamma$  for any  $\gamma \in \Gamma$ , since  $D_y$  is invariant to random warping of its input variables. Therefore, this definition provides a whole set of functions, rather than an individual function, as a mean. So, we have an extra degree of freedom in choosing an arbitrary element of the set  $\{\mu_f \circ \gamma \mid \gamma \in \Gamma\}$ . To make this choice unique, we can define a special element of this set as follows. Let  $\{\gamma_i^*\}$  denote the set of optimal warping functions in Eqn. 6; then, we choose the element such that the mean of  $\{\gamma_i^*\}$  is  $\gamma_{id}$ .

The following algorithm can be used to compute the mean set for the given data:

*Algorithm 1:* (Karcher Mean of  $\{f_i\}$  under  $D_y$ )

- 1) Compute SRSFs  $q_1, q_2, \dots, q_n$  of the given functions  $f_1, f_2, \dots, f_n$ .
- 2) Initialization Step: Select  $\mu_q = q_j$ , where  $j = \operatorname{argmin}_{1 \leq i \leq n} \|q_i - \frac{1}{n} \sum_{k=1}^n q_k\|$ .

- 3) For each  $q_i$ , find  $\gamma_i^*$  by solving:  $\gamma_i^* = \operatorname{argmin}_{\gamma \in \Gamma} \|\mu_q - (q_i \circ \gamma)\sqrt{\gamma}\|$ . The solution to this optimization comes from the dynamic programming algorithm.
- 4) Compute the aligned SRSFs using  $\tilde{q}_i = (q_i \circ \gamma_i^*)\sqrt{\gamma_i^*}$ .
- 5) If the increment  $\|\frac{1}{n} \sum_{i=1}^n \tilde{q}_i - \mu_q\|$  is small, then stop; the solution is given by  $\mu_q$ . Convert  $\mu_q$  into  $\mu_f$  using Eqn. 4.

Else, update the mean using  $\mu_q = \frac{1}{n} \sum_{i=1}^n \tilde{q}_i$  and return to step 3.

As mentioned previously, this algorithm results in a whole set of functions, but we need a specific element of that set for alignment. To introduce that element, we first need to define the mean of warping functions.

*Definition 4:* For a set of warping functions  $\gamma_1, \gamma_2, \dots, \gamma_n \in \Gamma$ , we define their cross-sectional mean as:

$$\bar{\gamma}_n(\omega) = \frac{1}{n} \sum_{i=1}^n \gamma_i(\omega) \quad \omega \in [0, \Omega]. \quad (7)$$

It is easy to see that  $\bar{\gamma}_n$  is an element of  $\Gamma$ .

*Definition 5:* For a given set of functions,  $f_1, f_2, \dots, f_n$  and  $f$ , define an element  $\tilde{f}$  as the center of the set  $\{f \circ \gamma | \gamma \in \Gamma\}$  where the set of warping functions  $\{\gamma_i\}$ ,  $\gamma_i = \operatorname{argmin}_{\gamma \in \Gamma} \|\tilde{q} - (q_i, \gamma)\|$ , has a cross-sectional mean  $\bar{\gamma}_{id}$ . Here  $q_i$ 's and  $q$  are the SRSFs of  $f_i$ 's and  $f$ , respectively.

Such an element exists by construction and is found using the following algorithm:

*Algorithm 2:* (Finding Center of the set  $\{f \circ \gamma | \gamma \in \Gamma\}$ ):

- 1) Let  $q_1, \dots, q_n, q$  be the SRSFs of given functions  $f_1, \dots, f_n, f$ , respectively.
- 2) For each  $q_i$  find  $\gamma_i$  by solving:  $\gamma_i = \operatorname{argmin}_{\gamma \in \Gamma} (\|q - (q_i \circ \gamma)\sqrt{\gamma}\|)$ .
- 3) Compute the mean  $\bar{\gamma}_n$  of all  $\{\gamma_i\}$  using Eqn. 7 and compute  $\tilde{q} = (q, \bar{\gamma}_n^{-1})$ . The center of the set  $\tilde{f}$  is obtained by mapping  $\tilde{q}$  back to the function space using Eqn. 4.

Thus, we choose the element of  $\{f \circ \gamma | \gamma \in \Gamma\}$  which ensures that the cross-sectional mean of  $\{\gamma_i^*\}$ , denoted by  $\gamma_\mu$ , is  $\gamma_{id}(\omega) = \omega$ , the identity element of the group  $\Gamma$ .

Now we can utilize Algorithms 1 and 2 to present the full procedure for alignment of  $\{f_i\}$ .

**Signal Alignment Algorithm:** Given a set of functions  $f_1, f_2, \dots, f_n$  on  $[0, \Omega]$ , let  $q_1, q_2, \dots, q_n$  denote their SRSFs, respectively.

- 1) Compute the Karcher mean of  $f_1, f_2, \dots, f_n$  in SRSF space using Algorithm 1. Denote it by  $\mu_f$ .
- 2) Find the center  $\tilde{f}$  of the set  $\{\mu_f \circ \gamma | \gamma \in \Gamma\}$  with respect to  $\{f_i\}$  using Algorithm 2; name the SRSF of that center  $\mu_q$ . (Note that this algorithm requires a step for computing the mean of warping functions using Eqn. 7).
- 3) For  $i = 1, 2, \dots, n$ , find  $\gamma_i^*$  by solving:  $\gamma_i^* = \operatorname{argmin}_{\gamma \in \Gamma} \|\mu_q - (q_i, \gamma)\|$ .
- 4) Compute the aligned SRSFs  $\tilde{q}_i = (q_i, \gamma_i^*)$ , their average  $\hat{\mu}_q = \frac{1}{n} \sum_{i=1}^n \tilde{q}_i$ , and the aligned functions  $\tilde{f}_i = f_i \circ \gamma_i^*$ .

#### IV. EXPERIMENTAL RESULTS

In this section, we describe some experimental results to demonstrate the estimation of SONAR signals and their classification using distances  $D_y$  and  $D_x$ . We choose the acoustic color data (spectral response), over spatial impulse response data, to exploit resonances that occur in the frequency domain for different materials.

##### A. Data Description

The data set used in these experiments was collected at the Naval Surface Warfare Center, Panama City Division (NSWC PCD), test pond. For a description of the pond and experimental setup, the reader is referred to [20]. The raw SONAR data was collected using a 1 - 30kHz LFM chirp and data was collected for nine proud targets that included a solid aluminum cylinder, an aluminum pipe, an inert 81mm mortar (filled with cement), a solid steel artillery shell, two machined aluminum un-exploded ordinances (UXOs), a machined steel UXO, a de-militarized 152mm TP-T round, a de-militarized 155mm empty projectile (without fuse or lifting eye), and a small aluminum cylinder with a notch. The aluminum cylinder is 2ft long with a 1ft diameter; while the pipe is 2ft long with an inner diameter of 1ft and 3/8 inch wall thickness.

The acoustic signals were generated from the raw SONAR data to construct relative power as a function of frequency and aspect angle. Due to the relatively small separation distances between the targets in the experimental setup, the scattered fields from the targets overlap. To generate the acoustic templates, SAS images were formed and then an inverse imaging technique was used to isolate the response of an individual target and to suppress reverberation noise. A brief summary of this process is as follows: The raw SONAR data is matched filtered and the

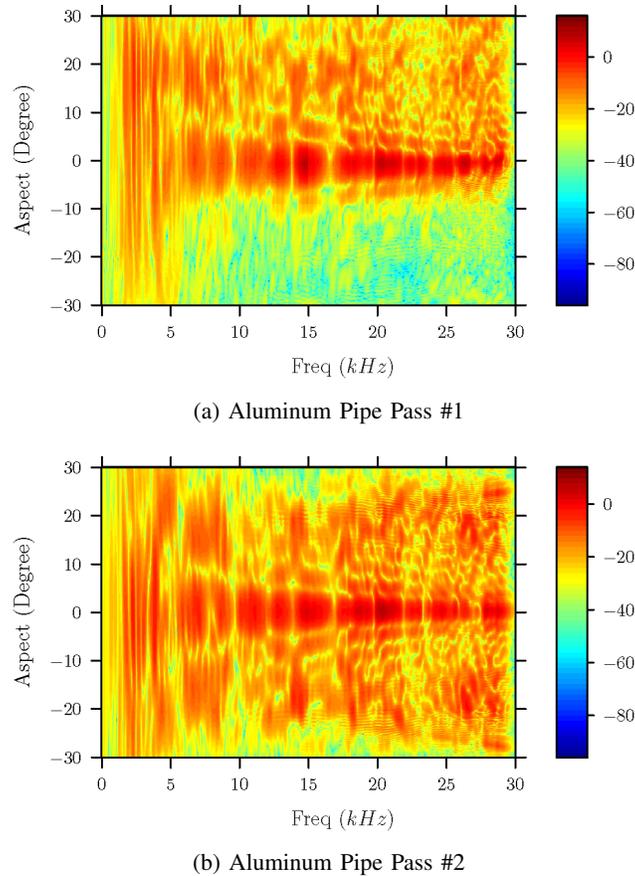


Fig. 4. Example acoustic color map for the aluminum pipe for two different target passes.

SAS image is formed using the  $\omega - k$  beamformer [21]. The target is then located in the SAS image and is then windowed around the selected location. This windowed image contains the information to reconstruct the frequency signals associated with a given target via inverting the  $\omega - k$  beamformer [22] and the responses were then aligned in range using the known acquisition geometry. For the nine targets, 2000 different data collections runs were done and 1102 acoustic color templates were generated using the method described above from the data set. Figure 4 presents an example acoustic-color map for the aluminum pipe for two different data collection passes. From the acoustic color maps, one-dimensional functional data was generated by taking slices at aspect value of  $0^\circ$  and therefore generating 1102 data samples.

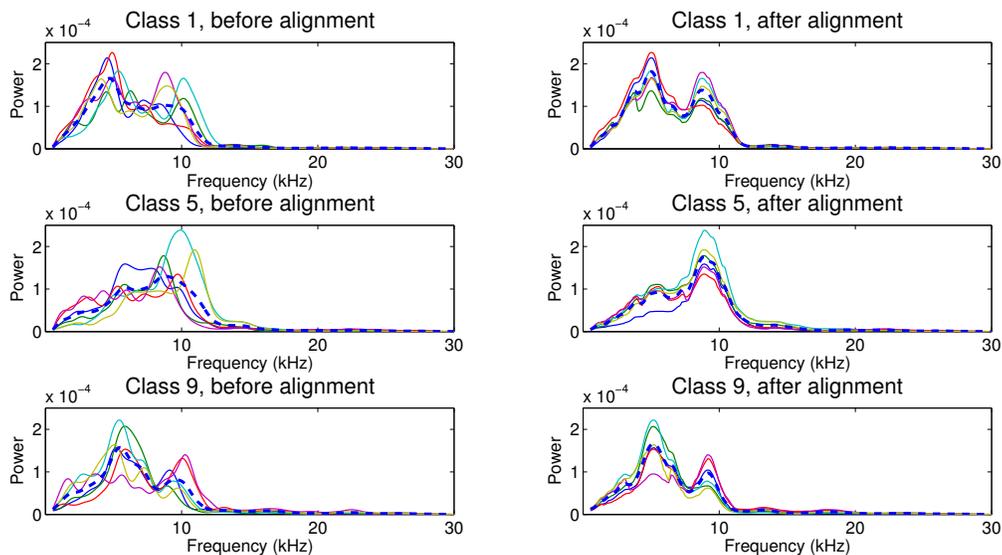


Fig. 5. Compositional noise removal on SONAR functions. First row: The left panel shows 6 typical functions  $\{f_i\}$  (solid lines) from Class 1 before alignment and their cross-sectional mean (dashed line). The right panel shows aligned functions  $\{\tilde{f}_i\}$  and their cross-sectional mean by the complete alignment algorithm. Second and third rows: Same as the first row except for Classes 5 and 9, respectively.

### B. Signal Alignment and Denoising

Here we present results for the alignment of signals from some select classes. Recall that the aligned signals represent data from which the (estimated) compositional noise has been removed. The results are similar in the remaining classes. Also, for improving the clarity, we display only a selected few observed signals  $\{f_i\}$  in each class and show the estimate  $\hat{g}$ . However, all the available signals were used in the calculation. The left panel of the first row in Fig. 5 shows 6 observed signals in Class 1 where most of the signals have two dominant peaks, but their locations vary across samples. When we take the cross-sectional mean (dashed line)  $\bar{f}$ , some peak information is averaged out and the mean does not retain the original bimodal pattern in the original sample. We then apply our alignment algorithm and the result is shown in the right panel of the first row in Fig. 5. We see that all peaks and valleys in the aligned functions are well-aligned. In this case, the average after alignment keeps the two-peak pattern and reasonably represents the variability in the frequency domain. Similar preservation of patterns is observed in the remaining examples involving Classes 5 and 9. These results generally indicate that the

overall alignment algorithm can appropriately remove the compositional noise in the functional data. Since we do not have a ground truth for these signals, we cannot evaluate these results more formally.

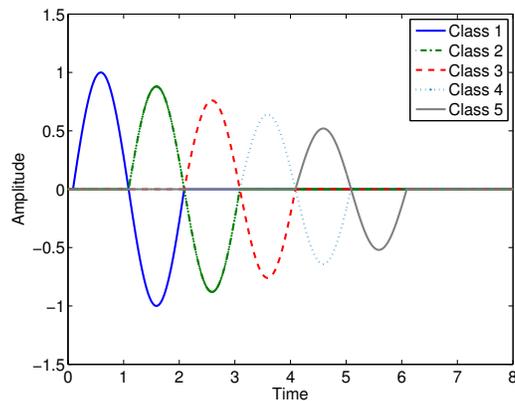
### C. Classification using Pairwise Distances

In this section, we present experimental results for classification of signals using the metrics developed in this paper. In addition to studying the real SONAR data, we will use a simulation experiment to study the effect of traditional additive noise on our metrics.

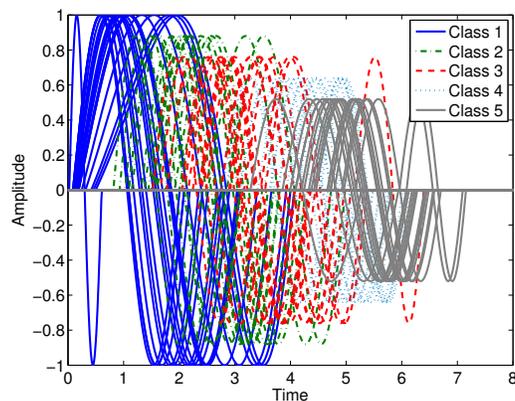
**Simulated Data:** We conducted a simulation study using a data set similar to the one presented in Fig. 3, where each class is represented by a sine waveform and each class has a different amplitude and phase shift. We studied a five-class problem with the amplitudes for the 5 classes being 1, 0.88, 0.76, 0.64, and 0.52, respectively. The phase shift for each class was  $((k-1) * \pi)$  for  $k = 1, \dots, 5$  classes and the data was generated by randomly warping the classes.

The original classes are shown in Fig. 6(a) and the randomly warped data for the 5 classes is shown in Fig. 6(b); the warping functions are generated randomly. The additive noise was generated as white Gaussian noise with mean zero and variance  $\sigma$ , where  $\sigma$  was changed for the desired signal-to-noise ratio (SNR), and then smoothed using a moving average with a window of size 3. The choice of the smoothing allows for numerical robustness in the calculation of the SRSFs. Moreover, this noise is representative of the correlated noise we see in the spectral response data. The pairwise distances were calculated for the standard  $\mathbb{L}^2$ ,  $D_{Naive}$ ,  $D_x$ , and  $D_y$  and classification was performed using the leave-one-out (LOO) cross-validated nearest-neighbor classifier for varying degrees of noise.  $D_{naive}$  corresponds to the quantity  $\min_{\gamma} \|f_1 - f_2 \circ \gamma\|$  that has often been used in the literature for signal alignment. We denote this method as a “naive” warping method and refer the associated distance matrix to as  $(D_{Naive})_{ij} = \|f_i - f_j \circ \tilde{\gamma}_{ij}\|$ . Note that the data in the original domain does not obey the isometry property. Therefore, the distance matrix  $D_{Naive}$  is not symmetric.

Fig. 7 presents the classification rates for this data set versus SNR. We can see that  $D_y$  outperforms the Naive classifier and the  $\mathbb{L}^2$  method. This implies that when the compositional noise is accounted for in the model, we get better classification performance. In this example, the distance  $D_x$  expectedly gives poor performance since the same random warping was used for all the classes.



(a) Original Classes



(b) Warped Data

Fig. 6. Simulated Data of 5 classes with 20 functions in each class.

Overall, our nonlinear warping method using a proper distance performed well on the simulated data and greatly increases classification performance. Moreover, this system shows vast improvement over the standard  $\mathbb{L}^2$  distance and current alignment techniques such as  $D_{Naive}$ .

**Real Data:** Next, we applied our metrics for classifying SONAR data containing  $n = 1102$  SONAR signals with nine target classes. The numbers of observations in the nine classes are:

$$\{n_i\}_{i=1}^9 = \{131, 144, 118, 118, 121, 119, 120, 114, 117\},$$

respectively. A selected subset of functions in each class is shown in Fig. 8. We observe that the original data is quite noisy, due to both the compositional and the additive noise, increasing variability within class and reducing separation across classes. This naturally complicates the task of target classification using SONAR signals.

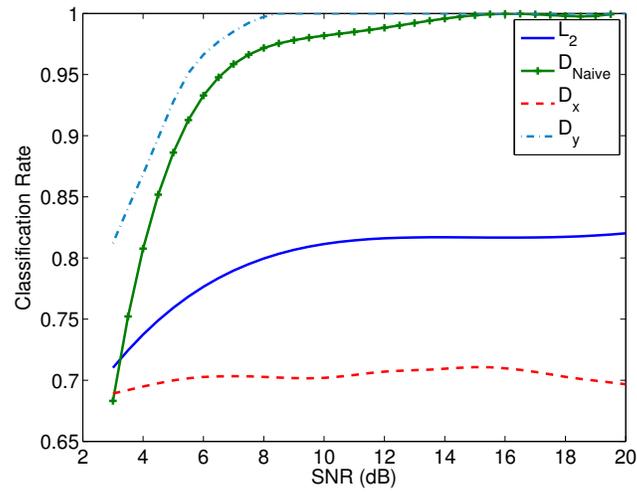


Fig. 7. Classification rates in the presence of additive noise.

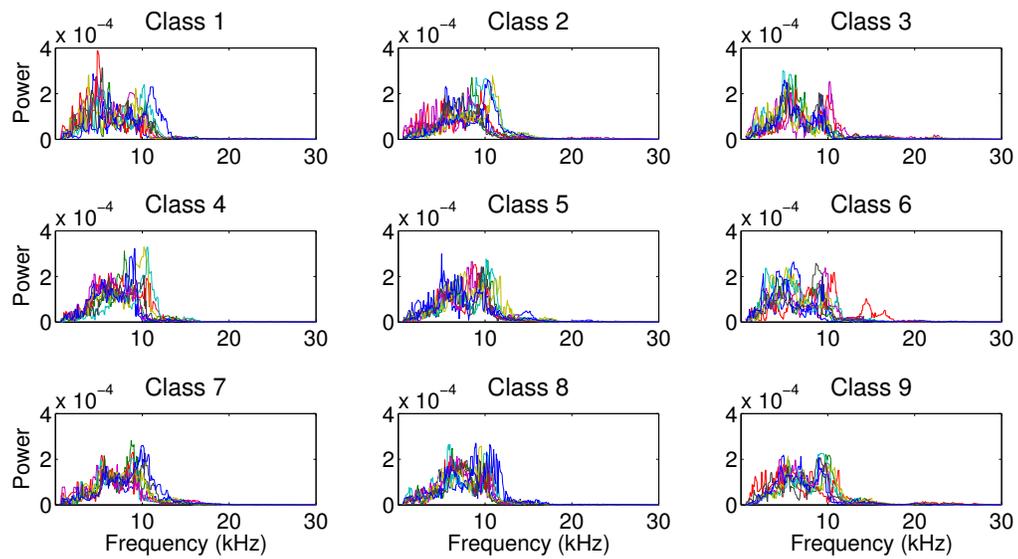


Fig. 8. Original SONAR functions in each of the 9 classes.

Amount of Smoothing	0	25	75	125	175
$D_x$	0.57	0.58	0.59	0.58	0.55
$D_y$	0.63	0.73	0.67	0.64	0.60
$D_{Naive}$	0.61	0.64	0.62	0.57	0.51
$\mathbb{L}^2$	0.43	0.44	0.45	0.45	0.44

TABLE II

CLASSIFICATION RATES VERSUS AMOUNT OF SMOOTHING APPLIED.

To have a robust estimate of the SRSF  $\{q_i\}$ , we first smooth the original signals  $\{f_i\}$  using a standard box filter  $[1/4, 1/2, 1/4]$ . That is, numerically, we update the signals at each discrete point by  $f_i(\omega_k) \rightarrow (\frac{1}{4}f_i(\omega_{k-1}) + \frac{1}{2}f_i(\omega_k) + \frac{1}{4}f_i(\omega_{k+1}))$ . To determine the effect of smoothing on the classification performance, we conducted a small study on the number of times the smoothing filter is applied. Table II presents the classification performance versus applying the smoothing filter 0, 25, 75, 125, and 175 times. It is interesting to note that the performance is quite stable with respect to smoothing and smoothing 25 times gives only slightly better performance. Hence, we use that level of smoothing for each signal for the rest of the analysis.

We first compute the standard  $\mathbb{L}^2$  distance between each pair, i.e.,  $(\mathbb{L}^2)_{ij} = \|f_i - f_j\|$ ,  $i, j = 1, \dots, n$ . The matrix of pairwise  $\mathbb{L}^2$  distances are shown as a gray scale image in Fig. 9(a). This image of the pairwise distances looks very noisy, underlying the difficulty of classification using SONAR data. Based on this distance matrix, we perform classification by using the LOO cross-validation on the standard nearest-neighbor method. It is found that the accuracy is 0.44 (489/1102). Then, we computed distances  $D_y$  and  $D_x$  between all pairs of signals and these distance matrices are shown as gray scale images in Fig. 9(b) and (c), respectively. Note that, in theory,  $D_x$  and  $D_y$  should lead to symmetric matrices, but, in practice, due to the numerical errors, these matrices are not exactly symmetric. So, we force them to be symmetric using  $D_x \rightarrow (D_x + D_x^T)/2$ ,  $D_y \rightarrow (D_y + D_y^T)/2$ , where the superscript  $\top$  indicates the transpose of a matrix.

In the image of  $D_y$  (Fig. 9(b)), we find that the pairwise distances are more structured than the  $\mathbb{L}^2$  distances. We also perform classification using the LOO cross-validated nearest-

neighbor based on the  $D_y$  distances. The accuracy turns out to be 0.73 (803/1102), a significant improvement over the result (0.44) in the standard  $\mathbb{L}^2$  distances. Interestingly, we find that the  $D_x$  distances also have strong indication of the target class in the data. In Fig. 9(c), we see that the  $D_x$  image have some clusters (dark squares) along the main diagonal. The classification accuracy by  $D_x$  turns out to be 0.58 (643/1102), which is also higher than the classification performance of the standard  $\mathbb{L}^2$  norm in the function space.

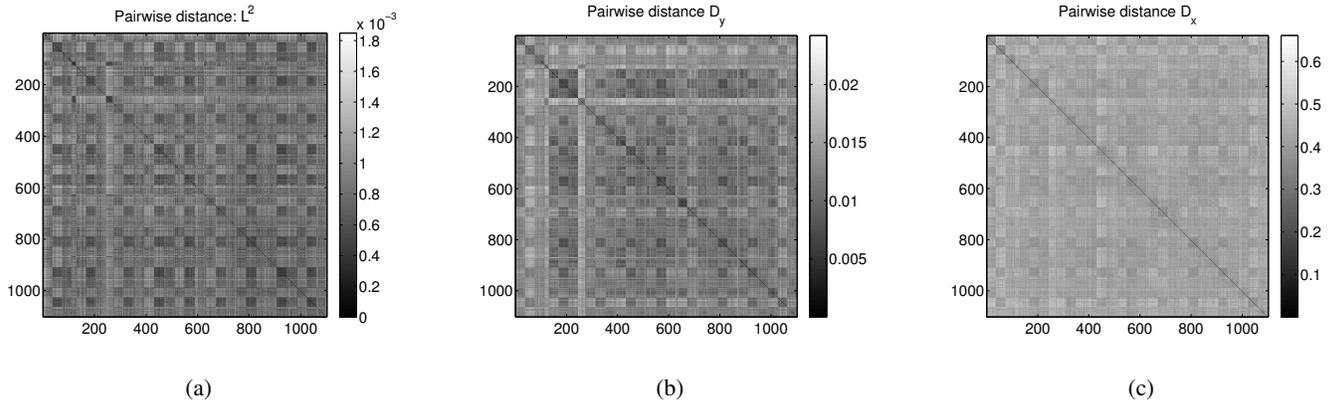


Fig. 9. The pairwise distances using the  $\mathbb{L}^2$  (a),  $D_y$  (b), and  $D_x$  (c) metrics.

Since  $D_x$  and  $D_y$  each only partially describe the variability in the data, which corresponds to the phase and amplitude differences between the functions, there is a possibility of improvement if  $D_x$  and  $D_y$  are used jointly. One simple idea is to linearly combine these two distances and use the weighted distance to perform classification on the data. Here the amplitude and phase are being treated as two different “features” of the signals. To accurately represent the contribution from each distance, we first normalize  $D_x$  and  $D_y$  by the maximum values in the matrices, respectively. That is,  $D_x \rightarrow \frac{D_x}{\max D_x}$ ,  $D_y \rightarrow \frac{D_y}{\max D_y}$ . Then, for  $\tau \in [0, 1]$ , we define

$$D_\tau = \tau D_x + (1 - \tau) D_y.$$

$D_\tau$  is a weighted average of  $D_x$  and  $D_y$  with  $D_0 = D_y$  and  $D_1 = D_x$ .

The next step is the estimation of an optimal  $\tau$ . Towards this end, we randomly select 50% of the given signals as training data and evaluate the LOO classification performance for different values of  $\tau$ . Since this selection is random, the resulting evolution is potentially random. Figure 10(a) shows the performance profile versus  $\tau$  for 100 randomly selected training data. An average

of these curves is superimposed on the same plot (thick solid line). A histogram of the optimal values of  $\tau$  for different random selections of the training data is shown in (b). Both these figures show that a broad range of  $\tau$  values, from 0.3 to 0.7, all result in a decent increase in the classification performance over the individual metrics  $D_x$  and  $D_y$ , and the general pattern of increase is similar. In fact, if we use the full data and plot the LOO classification performance versus  $\tau$ , we obtain the plot shown in (c). The overall shape (and the location of the maximizer) of this curve is very similar to the curves in (a) and underscores the independence of different observations. From this study, we select a value, say  $\tau = 0.41$  and use that to perform LOO classification on the full data.

When  $\tau = 0.41$  is used, we get an accuracy of 0.76 (839/1102), which is higher than the accuracy of the  $\mathbb{L}^2$ ,  $D_y$ , and  $D_x$  distances. This indicates that the variability in the SONAR signals are better characterized when we separate the phase and amplitude variabilities and better classification can be achieved when both variabilities are utilized.

Next, we compute the “naive” distance between any two signals presented in the previous section, according to  $(D_{Naive})_{ij} = \operatorname{argmin}_{\gamma \in \Gamma} \|f_i - f_j \circ \gamma\|$ . We also perform the cross-validated nearest-neighbor using the  $D_{Naive}$  and find that the accuracy is 0.64 (702/1102). This is slightly better than the accuracy by  $D_x$ , but worse than that of  $D_y$ . This indicates that even a simple method of warping can help remove certain warping noise in the SONAR data, but the performance is not as good as a more formal SRSF-based warping.

Next we generated a cumulative match characteristic (CMC) curve [23] for the distances  $D_x$ ,  $D_y$ ,  $D_\tau(\tau = 0.41)$ ,  $D_{Naive}$ , and  $\mathbb{L}^2$ . A CMC curve plots the probability of classification against the returned candidate list size and is presented in Fig. 11. Initially,  $D_y$  and  $D_\tau$  outperform the other distances with  $D_{Naive}$  slightly outperforming  $D_x$ . After a slight increase in the returned list size,  $D_x$  begins to outperform  $D_{Naive}$  and our method rapidly approaches over 0.90 classification rate, in contrast to the  $D_{Naive}$  and the standard  $\mathbb{L}^2$  distances.

Next, in Table III(a)-(e), we present the confusion matrices for the distances  $\mathbb{L}^2$ ,  $D_{Naive}$ ,  $D_x$ ,  $D_y$  and the weighted  $D_\tau$ , respectively. The top number in the table represents the classification rate for  $0^\circ$  aspect angle.

Overall, we see an increase in correct classification rates in  $D_y$  and  $D_\tau$  over the other methods and the incorrect classification rates are distributed evenly across the classes. Moreover, the weighted distance,  $D_\tau$ , reduced the false classifications to nearly zero in many cases.

Class	1	2	3	4	5	6	7	8	9
1	<b>0.58</b> <b>0.4</b>	0.05 0.18	0.03 0.06	0.08 0.06	0.05 0.03	0.02 0.16	0.08 0.03	0.05 0.05	0.05 0.03
2	0.03 0.08	<b>0.49</b> <b>0.52</b>	0.01 0.02	0.14 0.03	0.13 0.1	0.01 0.08	0.10 0.08	0.08 0.08	0.01 0.03
3	0.03 0.06	0.12 0.14	<b>0.34</b> <b>0.52</b>	0.02 0.01	0.12 0.06	0.04 0.13	0.17 0.03	0.13 0.03	0.04 0.03
4	0.02 0.01	0.16 0.13	0.01 0.07	<b>0.50</b> <b>0.39</b>	0.09 0.02	0.01 0.16	0.06 0.07	0.13 0.07	0.03 0.08
5	0.02 0.05	0.14 0.19	0.06 0.06	0.13 0.02	<b>0.27</b> <b>0.32</b>	0.03 0.04	0.17 0.11	0.11 0.15	0.06 0.05
6	0.07 0.08	0.04 0.05	0.04 0.03	0.03 0.09	0.05 0.01	<b>0.69</b> <b>0.54</b>	0.05 0.05	0.03 0.08	0 0.08
7	0.02 0.05	0.13 0.17	0.03 0.03	0.13 0.02	0.14 0.08	0.01 0.09	<b>0.35</b> <b>0.38</b>	0.10 0.12	0.09 0.07
8	0.01 0.02	0.08 0.21	0.04 0.04	0.16 0.04	0.11 0.07	0 0.07	0.11 0.15	<b>0.43</b> <b>0.33</b>	0.06 0.07
9	0.05 0.04	0.08 0.11	0.03 0.09	0.03 0.04	0.08 0	0.01 0.15	0.15 0.1	0.16 0.04	<b>0.32</b> <b>0.42</b>

(a)  $\mathbb{L}^2$ 

Class	1	2	3	4	5	6	7	8	9
1	<b>0.56</b> <b>0.57</b>	0.03 0.13	0.08 0.04	0.08 0.01	0.05 0.06	0.09 0.1	0.05 0.05	0.04 0.03	0.03 0.01
2	0.01 0.02	<b>0.82</b> <b>0.63</b>	0 0	0.06 0.01	0.03 0.16	0 0.03	0.06 0.16	0.03 0.09	0.01 0
3	0.02 0.1	0.05 0.03	<b>0.62</b> <b>0.63</b>	0.01 0	0.14 0.08	0.03 0.14	0.03 0.01	0.1 0.01	0.03 0.01
4	0 0.01	0.15 0.11	0.01 0.02	<b>0.66</b> <b>0.59</b>	0.05 0.06	0.01 0.03	0.07 0.08	0.04 0.09	0.01 0.01
5	0.02 0.02	0.17 0.24	0.03 0.03	0.07 0.02	<b>0.52</b> <b>0.42</b>	0.02 0.03	0.07 0.08	0.07 0.14	0.03 0.03
6	0.06 0.01	0.03 0.05	0.02 0.03	0.05 0.05	0.04 0.02	<b>0.73</b> <b>0.74</b>	0 0	0.02 0.01	0.05 0.07
7	0.01 0.02	0.11 0.13	0.03 0	0.07 0.02	0.1 0.09	0.02 0.04	<b>0.58</b> <b>0.61</b>	0.05 0.05	0.03 0.05
8	0 0.03	0.1 0.12	0.04 0.01	0.15 0.04	0.04 0.02	0.03 0.03	0.09 0.07	<b>0.54</b> <b>0.64</b>	0.04 0.02
9	0 0.02	0 0.14	0.03 0.02	0.04 0.03	0.09 0.09	0.04 0.02	0.07 0.09	0.07 0.08	<b>0.68</b> <b>0.52</b>

(b)  $D_{Naive}$ 

Class	1	2	3	4	5	6	7	8	9
1	<b>0.76</b> <b>0.59</b>	0.04 0.06	0.04 0.03	0.01 0.05	0.02 0.03	0.06 0.08	0.02 0.08	0.02 0.07	0.02 0.03
2	0.03 0.08	<b>0.48</b> <b>0.41</b>	0.04 0.02	0.12 0.06	0.14 0.08	0.01 0.07	0.05 0.09	0.1 0.09	0.03 0.1
3	0.02 0.08	0.03 0.03	<b>0.66</b> <b>0.61</b>	0.03 0.08	0.11 0	0.03 0.05	0.06 0.03	0.03 0.03	0.05 0.08
4	0.02 0.1	0.09 0.06	0.05 0.1	<b>0.57</b> <b>0.44</b>	0.08 0.05	0.03 0.01	0.07 0.07	0.07 0.1	0.03 0.08
5	0.02 0.07	0.1 0.02	0.06 0.06	0.04 0.07	<b>0.55</b> <b>0.49</b>	0.03 0.03	0.05 0.1	0.07 0.08	0.07 0.08
6	0.05 0.08	0.04 0.03	0.03 0.05	0 0.03	0.06 0.05	<b>0.77</b> <b>0.66</b>	0.03 0.02	0.02 0.05	0.01 0.03
7	0.03 0.1	0.1 0.07	0.08 0.06	0.07 0.14	0.14 0.16	0.06 0.04	<b>0.36</b> <b>0.28</b>	0.09 0.08	0.07 0.07
8	0.04 0.07	0.1 0.04	0.02 0.02	0.17 0.06	0.16 0.08	0 0.09	0.12 0.07	<b>0.39</b> <b>0.49</b>	0.02 0.07
9	0.02 0.07	0.04 0.07	0.03 0.07	0.04 0.07	0.03 0.03	0.06 0.04	0.03 0.05	0.03 0.03	<b>0.71</b> <b>0.57</b>

(c)  $D_x$ 

Class	1	2	3	4	5	6	7	8	9
1	<b>0.85</b> <b>0.63</b>	0.03 0.05	0.03 0.08	0 0.03	0 0.03	0.04 0.08	0.02 0.04	0.01 0.04	0.02 0.02
2	0.01 0.03	<b>0.74</b> <b>0.61</b>	0.01 0.02	0.11 0.03	0.05 0.09	0 0.02	0.06 0.1	0.01 0.08	0.01 0.03
3	0.03 0.08	0.01 0.03	<b>0.66</b> <b>0.73</b>	0.02 0.02	0.13 0.03	0.03 0.08	0.03 0.02	0.04 0.02	0.05 0
4	0 0.03	0.11 0.05	0.01 0.03	<b>0.75</b> <b>0.66</b>	0.04 0.03	0.01 0.03	0.08 0.04	0.06 0.07	0 0.03
5	0.02 0.07	0.02 0.07	0.04 0.02	0.07 0.03	<b>0.67</b> <b>0.58</b>	0.05 0.04	0.06 0.08	0 0.07	0.04 0.04
6	0.02 0.03	0.01 0.02	0.02 0.02	0.02 0.01	0.03 0.03	<b>0.87</b> <b>0.85</b>	0.01 0.01	0.01 0.03	0.02 0.01
7	0 0.05	0.12 0.13	0.02 0.03	0.06 0.02	0.1 0.08	0.02 0.04	<b>0.6</b> <b>0.5</b>	0.06 0.12	0.03 0.03
8	0 0.02	0.07 0.12	0.03 0.02	0.17 0.04	0.06 0.07	0.01 0.06	0.07 0.08	<b>0.59</b> <b>0.53</b>	0.01 0.06
9	0.03 0.03	0.01 0.04	0.04 0.03	0.01 0.07	0.07 0.06	0 0.01	0.02 0.06	0.02 0.04	<b>0.81</b> <b>0.65</b>

(d)  $D_y$ 

Class	1	2	3	4	5	6	7	8	9
1	<b>0.89</b> <b>0.76</b>	0.02 0.03	0.02 0.03	0 0.02	0 0.03	0.03 0.03	0.02 0.05	0.01 0.03	0.01 0.02
2	0.01 0.03	<b>0.76</b> <b>0.68</b>	0 0.03	0.07 0.03	0.08 0.07	0 0.02	0.03 0.05	0.04 0.07	0.01 0.03
3	0.02 0.09	0.02 0.01	<b>0.72</b> <b>0.82</b>	0.02 0	0.11 0	0.03 0.07	0.04 0	0.02 0.01	0.03 0
4	0 0.03	0.08 0.03	0.01 0.03	<b>0.79</b> <b>0.71</b>	0.03 0.02	0.01 0.04	0.01 0.03	0.08 0.06	0 0.04
5	0.02 0.05	0.02 0.03	0.02 0.02	0.07 0	<b>0.71</b> <b>0.69</b>	0.03 0.05	0.05 0.07	0.01 0.07	0.07 0.02
6	0.02 0.02	0 0.01	0.03 0.02	0 0.02	0.03 0.02	<b>0.89</b> <b>0.89</b>	0.03 0	0.01 0.03	0 0
7	0 0.03	0.07 0.1	0.02 0.03	0.08 0.03	0.11 0.06	0.03 0.04	<b>0.63</b> <b>0.58</b>	0.05 0.09	0.03 0.03
8	0 0.02	0.06 0.09	0.04 0.01	0.16 0.05	0.09 0.06	0.01 0.06	0.05 0.05	<b>0.59</b> <b>0.64</b>	0 0.02
9	0.03 0.05	0.01 0.03	0.03 0.02	0.01 0.03	0.06 0.02	0 0.02	0.03 0.04	0.01 0.03	<b>0.83</b> <b>0.76</b>

(e)  $D_\tau$  ( $\tau = 0.41$  for  $0^\circ$ ) and ( $\tau = 0.43$  for  $20^\circ$ )

TABLE III

CONFUSION MATRICES FOR THE DISTANCES  $\mathbb{L}^2$ ,  $D_{Naive}$ ,  $D_x$ ,  $D_y$  AND THE WEIGHTED  $D_\tau$  FOR THE  $0^\circ$  AND  $20^\circ$  ASPECT ANGLE DATA SETS. THE TOP NUMBER IS FOR  $0^\circ$  AND THE BOTTOM IS FOR  $20^\circ$ .

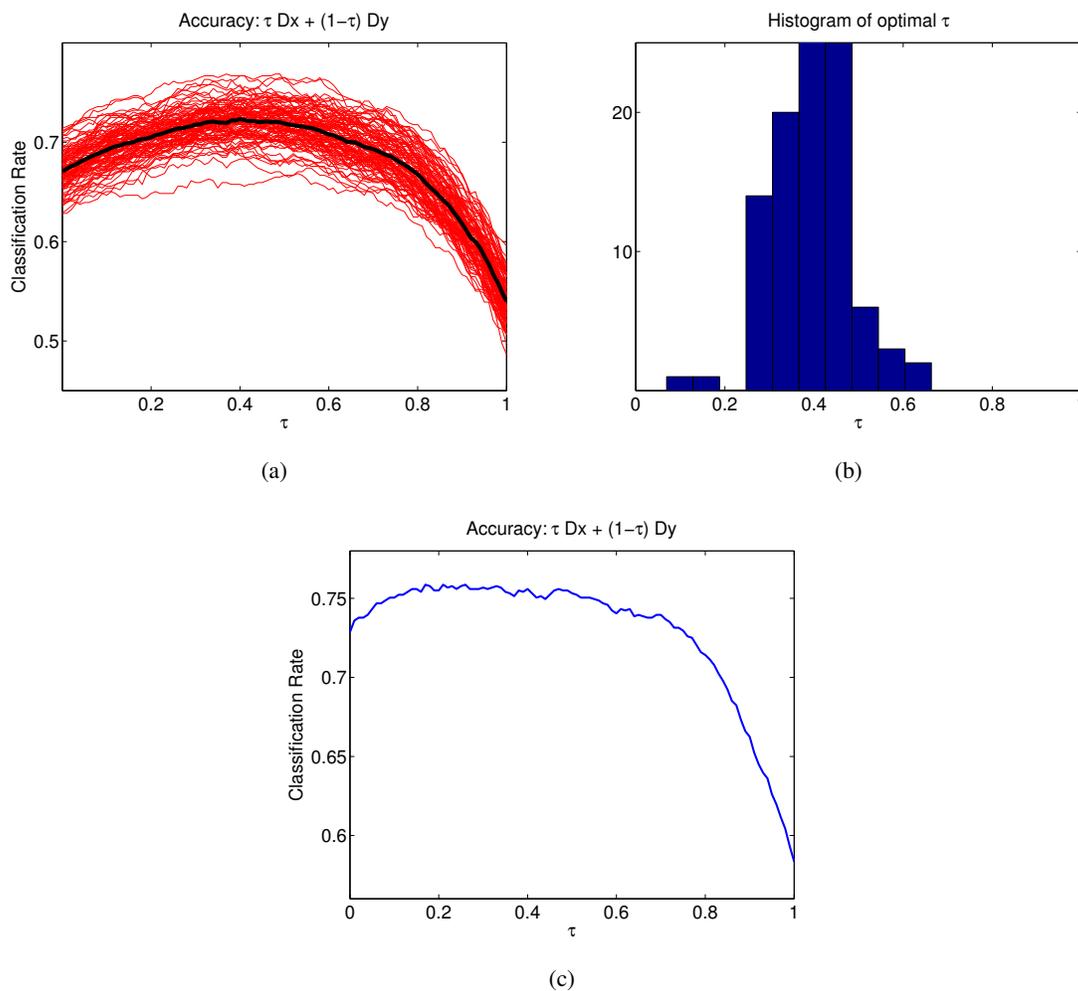


Fig. 10. (a) Evolution of classification performance versus  $\tau$  for randomly selected training data. The average of these curves is drawn on the top. (b) The histogram of optimal  $\tau$  values for different random selections of training data. (c) Overall performance versus  $\tau$  for the full data.

To compare, the classification accuracy at  $0^\circ$  aspect angle another set of data was extracted at  $20^\circ$  aspect angle and was classified using the same method described previously. The classification rates for  $\mathbb{L}^2$  and  $D_{Naive}$  was found to be 0.41 (453/1102) and 0.57 (631/1102), respectively. As with  $0^\circ$  aspect, our distances of  $D_y$  and  $D_\tau$  offered an improvement over these methods. The distance  $D_x$  had a classification rate of 0.49 (537/1102) and  $D_y$  had a classification rate of 0.61 (677/1102). We also found an optimal  $\tau = 0.43$  for the weighted distance  $D_\tau$  and had a classification rate of 0.70 (774/1102).

Next, we generated a CMC curve for the  $20^\circ$  aspect angle and is presented in Fig. 12. As

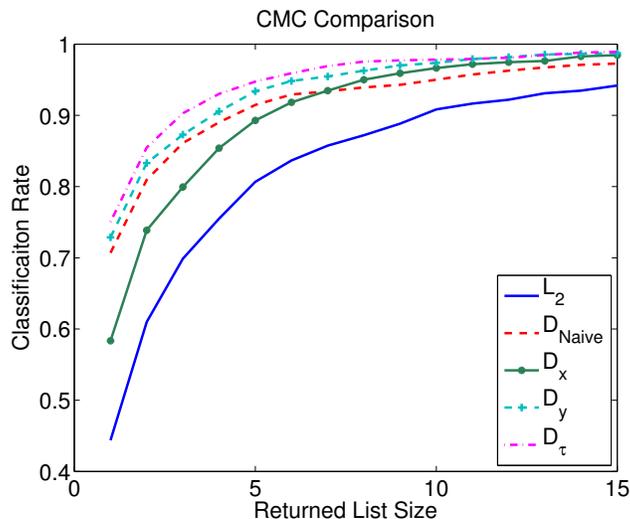


Fig. 11. CMC Comparison of  $L_2$ ,  $D_{Naive}$ ,  $D_x$ ,  $D_y$  and the weighted  $D_\tau$  ( $\tau = 0.41$ ) distances for  $0^\circ$  aspect angle.

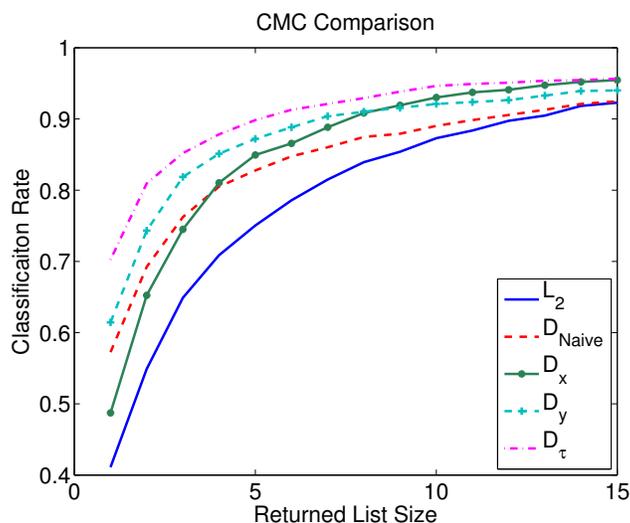


Fig. 12. CMC Comparison of  $L_2$ ,  $D_{Naive}$ ,  $D_x$ ,  $D_y$  and the weighted  $D_\tau$  ( $\tau = 0.43$ ) distances for  $20^\circ$  aspect angle.

with  $0^\circ$ , we see that both  $D_y$  and  $D_\tau$  approach 0.90 classification rate faster than  $L_2$  and  $D_{Naive}$ . However, the performance and rate of increase is lower than the  $0^\circ$  aspect angle case.

The confusion matrices for  $20^\circ$  aspect angle are presented in Table III(a)-(e) and are the bottom numbers in the table. Again, we see performance similar to  $0^\circ$  with decreasing incorrect classification rates for  $D_\tau$  and  $D_y$  over the other distances. However, the overall classification

rate is lower for most of the classes except Class 3 and Class 8. This can be attributed to more information for these classes captured at that aspect angle.

## V. CONCLUSIONS AND OBSERVATIONS

The statistical analysis and classification of targets using acoustic signatures is a challenging task. In particular, it is complicated by the presence of compositional noise in the observed signals. We have adapted a recent comprehensive approach that solves the problem of signal alignment and denoising by comparing signals in a unified framework and using a cost function that is eventually a warping-invariant distance between the two signals. This framework is applied to both real and simulated data. It provides two distances –  $D_x$  and  $D_y$  – that can be used for classifying noisy signals using any metric-based classifier. We have used the leave-one-out classifier in this paper to demonstrate the improvements over traditional methods for signal comparisons. In experiments involving real data, we demonstrate a LOO performance of almost 0.76, which easily outperforms the standard  $L^2$  distance (0.44), and current methods using a naive alignment (0.64).

The method presented in this paper only solves the one-dimensional problem. To process all available aspect angles and relationship between the angles a two-dimensional warping is needed. An extension to this work would be to explore a two-dimensional warping for distance-based classification based upon current ideas in shape analysis of surfaces [24]. Moreover, one could combine multiple results using the one-dimensional method in a machine learning framework for increased performance.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Frank Crosby and Dr. Quyen Huynh at NSWC PCD for their technical support during this work and the two anonymous referees and Associate Editor for their constructive comments and suggestions, which led to a significant improvement of the paper.

## REFERENCES

- [1] G. J. Dobeck, "Image normalization using the serpentine forward-backward filter: application to high-resolution sonar imagery and its impact on mine detection and classification," *Proc. SPIE*, vol. 5734, pp. 90–110, April 2005.

- [2] T. Aridgides and M. Fernandez, "Image-based ATR utilizing adaptive clutter filter detection, LLRT classification, and volterra fusion with application to side-looking sonar," *Proc. SPIE*, vol. 7664, pp. 1R–11R, March 2010.
- [3] N. Klausner, M. R. Azimi-Sadjadi, and J. D. Tucker, "Multi-sonar target detection using multi-channel coherence analysis," *Proc. of MTS/IEEE Oceans 2010 Conference*, pp. 1–7, Sept. 2010.
- [4] J. D. Tucker and M. R. Azimi-Sadjadi, "Coherence-based underwater target detection from multiple sonar platforms," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 1, pp. 37–51, Jan 2011.
- [5] J. D. Tucker, J. T. Cobb, and M. R. Azimi-Sadjadi, "Generalized likelihood ratio test for finite mixture model of K-distributed random variables," *Proc of IEEE DSP/SPE 2011 Workshop*, pp. 443–448, Jan 2011.
- [6] G. J. Dobeck, "Fusing sonar images for mine detection and classification," *Proc. SPIE*, vol. 3710, pp. 602–614, April 1999.
- [7] T. Aridgides, P. Libera, M. Fernandez, and G. J. Dobeck, "Adaptive filter/feature orthogonalization processing string for optimal LLRT mine classification in side-scan sonar imagery," *Proc. SPIE*, vol. 2765, pp. 110–121, April 1996.
- [8] N. Intrator, N. Neretti, and Q. Huynh, "Sonar object discrimination via spectral density analysis," *Proc. of MTS/IEEE OCEANS 2004 Conference*, vol. 2, pp. 740–742, Nov 2004.
- [9] D. Miklovic and M. Bird, "Quantitative acoustic color displays for classification with broadband sonars," *Proc. of MTS/IEEE Oceans 1998 Conference*, vol. 1, pp. 592–597, Sept. 1998.
- [10] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*. Springer, 2005.
- [11] S. Kurtek, A. Srivastava, and W. Wu, "Signal estimation under random time-warpings and nonlinear signal alignment," in *Proceedings of Advances in Neural Information Processing Systems (NIPS), Grenada, Spain*, 2011, pp. 676–683.
- [12] A. Srivastava, W. Wu, S. Kurtek, E. Klassen, and J. S. Marron, "Registration of functional data using fisher-rao metric," *arXiv:1103.3817v2 [math.ST]*, 2011.
- [13] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.
- [14] A. Willsky, "Fourier series and estimation on the circle with applications to synchronous communication–I: Analysis," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 577 – 583, sep 1974.
- [15] M. Tsang, J. H. Shapiro, and S. Lloyd, "Quantum theory of optical temporal phase and instantaneous frequency," *Phys. Rev. A*, vol. 78, no. 5, p. 053820, Nov 2008.
- [16] H. Silverman and D. Morgan, "The application of dynamic programming to connected speech recognition," *ASSP Magazine, IEEE*, vol. 7, no. 3, pp. 6 –25, jul 1990.
- [17] H. V. Trees, *Detection, Estimation, and Modulation Theory, vol. I*. N.Y.: John Wiley, 1971.
- [18] J. D. Tucker and A. Srivastava, "Statistical analysis and classification of acoustic color functions," *Proc SPIE*, vol. 8017, pp. O1–Q10, April 2011.
- [19] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn, "Shape analysis of elastic curves in euclidean spaces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 7, pp. 1415 –1428, July 2011.
- [20] S. Kargl, K. Williams, T. Marston, J. Kennedy, and J. Lopes, "Acoustic response of unexploded ordnance (UXO) and cylindrical targets," *Proc. of MTS/IEEE Oceans 2010 Conference*, pp. 1 –5, 2010.
- [21] M. Soumekh, *Synthetic Aperture Radar Signal Processing*. Wiley, 1999.
- [22] A. Khwaja, L. Ferro-Famil, and E. Pottier, "SAR raw data simulation using high precision focusing methods," *European Radar Conference*, pp. 33 –36, 2005.
- [23] R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior, "The relation between the ROC curve and the CMC," in *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*, Oct. 2005, pp. 15 – 20.

- [24] S. Kurtek, E. Klassen, Z. Ding, S. Jacobson, J. Jacobson, M. Avison, and A. Srivastava, "Parameterization-invariant shape comparisons of anatomical surfaces," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 3, pp. 849–858, 2011.



**J. Derek Tucker** (S'04) received the B.S. and M.S. degrees from Colorado State University, Fort Collins, CO in 2007 and 2009, respectively, both in electrical engineering, with an emphasis on statistical signal processing.

Currently he is pursuing a Ph.D. Degree in Statistics from Florida State University and is employed as research engineer in automatic target recognition by the Naval Surface Warfare Center Panama City Division (NSWC PCD). His research interests include statistical image understanding, target detection, classification, functional data analysis, and statistical learning methods.



**Wei Wu** Wei Wu received the Ph.D. degree in applied mathematics from Brown University, Providence, RI, in 2004. He joined the faculty at Florida State University in 2006, where he is an Associate Professor in the Department of Statistics and an Associate Faculty Member in the Program of Neuroscience. His research explores registration and alignment of functional data, and statistical modeling/analysis of spike trains and other neural signals.



**Anuj Srivastava** Anuj Srivastava received the M.S. and Ph.D. degrees in electrical engineering from Washington University, St. Louis, Missouri, in 1993 and 1996, respectively. After spending the year 1996-1997 at Brown University as a visiting researcher, he joined the Department of Statistics at Florida State University (FSU), Tallahassee, as an assistant professor in 1997. He is currently a professor in that department. He has held visiting appointments at University of Lille, France; INRIA, Sophia Antipolis, France; and, UCL, Louvain-Le-Neuve, Belgium. His research is focused on pattern theoretic approaches to problems in image analysis, computer vision, and signal processing. In particular, he has developed computational tools for performing statistical inferences on certain nonlinear manifolds, including shape manifolds of curves and surfaces. He has published more than 170 peer-reviewed journal and conference articles in these areas. He is a senior member of the IEEE.