

Received November 26, 2019, accepted December 17, 2019, date of publication December 30, 2019, date of current version January 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2962793

# Feature-Compression-Based Detection of Sea-Surface Small Targets

PENGLANG SHUI<sup>1</sup>, (Senior Member, IEEE), ZIXUN GUO<sup>1</sup>, AND SAINAN SHI<sup>2</sup>

<sup>1</sup>National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China

<sup>2</sup>Nanjing University of Information Science and Technology, Nanjing 210044, China

Corresponding authors: Penglang Shui (plshui@xidian.edu.cn) and Zixun Guo (zxguo\_724@stu.xidian.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61871303.

**ABSTRACT** This paper aims to develop a feature-based detector using seven existing salient features of radar returns to improve the detection ability of high-resolution maritime ubiquitous radars to sea-surface small targets. Maritime ubiquitous radars form simultaneously dwelling beams at multiple azimuths by digital array receiver and allow long observation time for detection. Due to absence or incompleteness of training samples of radar returns with various types of sea-surface small targets, the detection boils down to designing a one-class classifier in the seven-dimensional (7D) feature space mainly by using training samples of sea clutter. A feature compression method, though maximizing interclass Bhattacharyya distance, is proposed to compress the 7D feature vector into one 3D feature vector with the help of simulated radar returns of typical targets. In the compressed 3D feature space, a modified convexhull learning algorithm is given to determine one convex polyhedron decision region of sea clutter at a given false alarm rate. In this way, a feature-compression-based detector is constructed, which can exploit more features of radar returns to improve detection performance. It is verified by the recognized and open IPIX and CSIR radar databases for sea-surface small target detection. The results show that it attains obvious performance improvement.

**INDEX TERMS** High-resolution maritime ubiquitous radars, sea-surface small targets, feature compression, convexhull learning, one-class classifier, feature-compression-based detector.

## LIST OF ACRONYMS

CGM	Compound-Gaussian Model
CUT	Cell Under Test
CV	Coefficients of Variation
DAS	Doppler Amplitude Spectrum
DBF	Digital Beam Forming
LDA	Linear Discriminant Analysis
MS	Maximal Size of connected region
NHE	Normalized Hurst Exponent
NR	Number of connected Region
NTFD	Normalized TF Distribution
PRI	Pulse Repetition Interval
RAA	Relative Average Amplitude
RDPH	Relative Doppler Peak Height
RI	Ridge Integration
RVE	Relative Vector Entropy
TCL	Texture Coherent Length

The associate editor coordinating the review of this manuscript and approving it for publication was Chengpeng Hao.

## I. INTRODUCTION

As one of important tasks of sea battlefield perception, it is always a difficult problem for maritime surveillance radars to find sea-surface small targets, such as small boats, icebergs, frogmen, debris, and periscopes of submarines. Owing to low power level of sea clutter, high-resolution radars are used to improve signal-to-clutter ratios (SCR) of small target returns. High-resolution maritime radars often adopt dwelling mode or fast scan mode to obtain integration gain of target returns in radar slow time [1]–[4]. At fast scan mode, inter-scan noncoherent or binary integration is used for detection, for instance, anti-submarine radars [1], [2]. Experimental radars for sea-surface small target detection often work at dwelling mode, i.e., radar beam stares at one azimuth angle to collect data [3], [4]. In conventional surveillance radar systems, long observation time for small target detection and search efficiency at azimuth are conflictive, which is one of reasons for long time integration methods are limited in practical applications. Recently, the conflict is completely removed in ubiquitous radars [5], [6]. Ubiquitous radars using

MIMO digital array can implement all-time observation at all directions by digital beam forming (DBF) of multiple channels and thus allow long observation time for sea-surface small target detection. Therefore, it becomes imperative to develop an effective method to detect sea-surface small targets in the case of long observation time.

In high-resolution ubiquitous radars, sea-surface small target detection suffers from difficulties from three aspects. First, both high spatial and Doppler resolution of radars can lift SCR over a critical value of small targets to be detectable, but some small targets must be detected in low SCR cases. Second, high resolution complicates sea clutter characteristics and aggravates the effect of sea spikes [7], [8]. Third, interactions between sea-surface small targets and waves make that radar returns of sea-surface small targets have complex amplitude and Doppler modulations, which are difficult to be parameterized by some simple models. A series of attempts have been made, unfortunately no one attains satisfactory results in all cases, even on a special database such as the IPIX database [3] or CSIR database [4].

Traditional moving target identification/detection (MTI/MTD) techniques [7]–[9] fail to detect sea-surface small targets, because target weak returns fall inside the main clutter region of sea clutter in the Doppler domain, and have complex amplitude and Doppler modulations. As a way once expected, nonlinear time series analysis is the common foundation of the fractal-based methods [10]–[14] and the neural network (NN) learning methods [15]–[17]. Fractal features are not as a replacement of other statistics but only as additional indexes of sea clutter characteristics. The fractal-based methods require observation time long up to several seconds and have low detectability in some cases, a common defect of most of detection methods using one single feature. Neural network (NN) learning methods originate from the understanding that sea clutter time series is the output of an unknown nonlinear chaotic system [10]. Several structural NNs [15]–[17] are trained by sea clutter data to short-term predict the behavior of sea clutter, and the predictive error is naturally used as a statistic to find anomaly in data. The NN learning methods have two flaws. Radar returns with targets do not participate in the learning process from first to last, and a detection problem relevant to two-class classifiers is mandatorily reduced to one-class classifiers [18]. The NN-based detectors can be interpreted as one-class classifiers using full implicit features of sea clutter. To minimize the predictive error, the NN must capture as many implicit features as possible. Most of them serve for sea clutter description rather than discriminating target returns from sea clutter, so they degrade the generalization ability of the NN learning methods.

Due to diversity of sea-surface small targets and their complicated interactions with waves, it is impossible to collect radar returns of all sea-surface targets of interest in all cases. We refer to the phenomenon as the incompleteness of training data of returns with targets. However, it is rather important for radar returns with targets to participate in the learning

process. For instance, focusing on a special kind of small targets such as growlers [19], [20], the two-class classifiers using radar returns with targets in training attain satisfactory detection results. Besides, radar returns of simple targets can be simulated to train a detector and an experimental radar system using a NN as a tool is established [21]. Designing a detector, training samples of radar returns with targets are as important as that of sea clutter. One-class classifiers under the anomaly detection framework are unavoidable but not preferential choice in the sea-surface small target detection.

By qualitative analyses of sea clutter and radar returns with targets and introduction of three salient features of radar returns, the sea-surface small target detection boils down to designing a one-class classifier in a 3D feature space only using training samples of sea clutter [22], which is realized by means of convexhull learning algorithm [23], [24]. This tri-feature-based detector behaves well on the IPIX database [3]. It is noted that characteristics of returns with targets take effect at feature selection instead of learning stage. In comparison with the  $\nu$ -SVM algorithm for design of one-class classifiers [18], the convexhull learning algorithm can precisely control false alarm rate and gives visualized decision regions in the 3D feature space. Later, three time-frequency (TF) features are introduced to build a TF-tri-feature-based detector [25], which attains better overall performance on the IPIX database than the tri-feature-based detector [22]. Also, the tri-feature-based detector behaves better than the TF-tri-feature-based detector in some datasets. It shows that the three TF features [25] are not replacement of the amplitude and Doppler features in [22]. Exploitation of more complementary features is an approach to further improve performance. However, the dimension limitation of the convexhull computation [23], [24] impedes the cooperation of more than three features.

In this paper, seven existing salient features are available to construct a feature-compression-based detector, so as to realize effective and robust detection of sea-surface small targets. Due to the dimension limitation of the convexhull learning algorithm, the 7D feature vector must be mapped on one feature space whose dimension is no more than three. Feature compression requires quantitative statistics of the feature vectors of sea clutter and radar returns with targets [26], [27]. The incompleteness in training data on radar returns with targets is a major obstacle of feature compression. The detection problem is a semisupervised two-class classification problem with unbalanced demand of mistake probabilities and unbalanced training data for the two classes. A large amount of sea clutter data and only a small quantity of radar returns with special test targets are available in learning. Referring to characteristics of sea-surface small targets in measured data, a generator is constructed to yield radar returns of typical targets. Simulated target returns plus sea clutter yield radar returns with targets and training samples for feature compression. Using the training samples of the two classes, a feature compression method is proposed to compress the 7D feature vector to a 3D feature vector by

maximizing the interclass Bhattacharyya distance [28], [29]. In the compressed 3D feature space, a modified convexhull learning algorithm piloted by the training samples of radar returns with targets is presented to determine the decision region of sea clutter at a given false alarm rate. A feature-compression-based detector is constructed. The proposed detector is verified by data from the IPIX and CSIR databases and is compared with the early fractal-based detector, our previous detectors, and two more recent detectors [30], [31].

This paper is organized as follows. Section II reviews detection problem of sea-surface small targets. Seven existing salient features are introduced and their complementarity is analyzed. Section III presents a generator of target returns, proposes a feature compression method and constructs a feature-compression-based detector. Section IV includes experimental results and performance comparison. Finally, we conclude our paper in Section V.

## II. SEVEN EXISTING SALIENT FEATURES AND THEIR COMPLEMENTARITY ANALYSIS

### A. DETECTION PROBLEM OF SEA-SURFACE SMALL TARGETS IN HIGH-RESOLUTION UBIQUITOUS RADAR

A radar attains high range resolution by transmitting wide-band pulses and high Doppler resolution by dwelling long time at a beam position. When dwelling time is long up to a few tenths of second or several seconds, radar returns of a sea-surface small target have to be modeled as a nonlinear frequency modulated (FM) signal with complex amplitude fluctuation, because its radial velocity and RCS are severely affected by wind waves and swells [25], [32]. Due to long observation time and high range resolution, temporally nonstationary sea clutter time series is characterized by the compound-Gaussian model (CGM) with a time-varying texture or piecewise spherical invariant random vector model [7], [32], [33]. Sea clutter time series at adjacent spatial resolution cells can be regarded to have same or similar characteristics. Like in adaptive detection [9], radar returns received at spatial resolution cells around the cell under test (CUT) are used to predict characteristics of sea clutter at the CUT. Detection of sea-surface small targets boils down to the following binary hypothesis test [9]–[14], [22], [32],

$$\begin{cases} H_0 : \begin{cases} z(n) = \sqrt{\tau(n)}u(n), & n = 1, 2, \dots, N \\ z_p(n) = \sqrt{\tau_p(n)}u_p(n), & n = 1, 2, \dots, N; \\ & p = 1, 2, \dots, P \end{cases} \\ H_1 : \begin{cases} z(n) = a(n) \exp(2\pi j(\phi(n)n + \varphi_0)) + \sqrt{\tau(n)}u(n), \\ & n = 1, 2, \dots, N \\ z_p(n) = \sqrt{\tau_p(n)}u_p(n), & n = 1, 2, \dots, N; \\ & p = 1, 2, \dots, P \end{cases} \end{cases} \quad (1)$$

where  $z(n)$  is the received complex time series at the CUT,  $z_p(n)$  is sea clutter time series at the reference cells around the CUT,  $H_0$  is the null hypothesis, and  $H_1$  is the alternative hypothesis, target returns are specified by amplitude

series  $a(n)$ , Doppler frequency series  $\phi(n)$ , and initial phase  $\varphi_0$ . For simplification, the pulse repetition interval (PRI) or sampling interval of time series is neglected. Detection strategies depend on models of the target returns and sea clutter.

In the CGM, high-resolution sea clutter time series is expressed as the product of two independent stochastic sequences: slowly-varying non-negative texture sequence  $\tau(n)$  and fast-varying speckle sequence  $u(n)$ .  $\tau(n)$  degenerates to a random constant when the observation time is shorter than the texture coherent length (TCL) of sea clutter. TCL is about several hundreds of milliseconds in X-band high-resolution radars and relates to the time scale of tilt modulation from long waves [33]–[35].  $u(n)$  follows complex Gaussian distribution of unit variance and has a decorrelated time of tens of milliseconds [7]. Hence, when the observation time is long up to the order of seconds,  $\tau(n)$  must be modeled as a time-varying function [32]. Radar returns of sea-surface small targets have more complex characteristics. Target returns of sea-surface small targets have distinct amplitude fluctuation and Doppler modulation relevant to target type, movement state, and sea state and have to be modeled as nonparametric forms, such as smooth functions of time. Under nonparametric models, detection often relies on ad-hoc test statistics, referred to as features, instead of the optimum or suboptimal test statistics depending on rigorous parametric models of the target returns and clutter [36], [37].

### B. SEVEN SALIENT FEATURES IN SEA-SURFACE SMALL TARGET DETECTION

This paper focuses on exploitation of existing salient features rather than feature extraction. This subsection gives a brief review of seven existing salient features in sea-surface small target detection and more details refer to [12], [14], [22], [25]. Two amplitude features are the normalized Hurst exponent (NHE) [12], [14] and relative average amplitude (RAA) [22]. Two Doppler features are the relative Doppler peak height (RDPH) and relative vector entropy (RVE) [22]. Three TF features are extracted from the normalized TF distribution (NTFD) and thresholded NTFD [25], including the ridge integration (RI) of the NTFD and the number of connected regions (NR) and maximal size of connected regions (MS) in the thresholded NTFD. The prefixes ‘relative’ and ‘normalized’ in the names of these features imply the idea of constant false alarm rate (CFAR). The features are computed from time series at the CUT and reference cells so as to adapt spatial-temporally varying characteristics of sea clutter.

In what follows, we briefly introduce the mechanism behind each feature. It is found that amplitude time series of sea clutter exhibit multifractal characteristics in a time scale of 0.01 to several seconds [12], [13]. Hurst exponent relevant to the fractal dimension of sea surface is a salient feature to detect radar returns with targets from sea clutter. Hurst exponent related with sea state and viewing geometry of the radar is spatial-temporally varying. The NHE feature is the Hurst exponent at the CUT normalized by the average value

and standard deviation of the Hurst exponents at the reference cells, so as to adapt spatial-temporally varying characteristics of the Hurst exponent [14]. The RAA is a commonly-used test statistic in non-coherent CFAR detection. It is the average amplitude at the CUT divided by the average value of the average amplitudes at the reference cells [22]. The NHE and RAA features are dimensionless and take large positive values for returns with targets while small positive values for sea clutter.

Sea clutter and radar returns with targets exhibit differences in Doppler amplitude spectrum (DAS) [22]. The DAS of sea clutter has an obtuse peak because the power of sea clutter distributes on the wide main clutter region in the Doppler domain while that of radar returns with targets has a sharp peak, because the power of target returns concentrates on several Doppler bins. The difference is measured by the Doppler peak height (DPH): the ratio of the peak to the average amplitude on its two sides. In view of the fact that the DPH of sea clutter alters with sea state and viewing geometry of radar, the DPH at the CUT divided by the average value of the DPHs at reference cells generates the relative DPH (RDPH) [22]. It takes large positive values for radar returns with targets while positive values around one for sea clutter. The vector entropy of the DAS measures its complexity, and the relative vector entropy (RVE) [22] is the VE at the CUT divided by the average value of the VEs at the reference cells. The RDPH reflects a local difference of radar returns with targets and sea clutter in the DAS, and the RVE reflects a global difference in the DAS.

As nonstationary time series, sea clutter and target returns exhibit salient differences in the TFD. It transfers 1D complex time series into 2D image. The TFD of sea clutter is modeled as a stochastic process on the 2D TF plane with a mean function and standard deviation function, which are estimated from the TFDs at the reference cells [25]. The normalized TFD (NTFD) at the CUT is the TFD at the CUT normalized by the estimated mean and standard deviation functions. The NTFD reflects differences of the radar returns at the CUT from sea clutter in the TF characteristics. From the NTFD, the three TF features are extracted. The ridge integration (RI) is the sum of the grayscale values of the pixels along the ridge of the NTFD. Significant pixels in each time slice of the NTFD form multiple connected regions in the TF grid. The number of the connected regions (NR) and their maximal size (MS) are the other two important TF features. For radar returns with targets, the RI and MS take large values while the NR takes small values, because the power of target returns concentrates on a small number TF pixels along the instantaneous Doppler curve of the target. The situation is just converse for sea clutter. The first four features can be computed from the received time series at the CUT and reference cells in a low computational cost by the Fast Fourier Transformation (FFT). The computation of the TF features is involved in the TFD of the received time series, and the computational complexity of each TFD is  $O(NK\log N)$  where  $N$  is length of the time series and  $K$  is the frequency samples.

### C. DEFICIENCIES AND COMPLEMENTARITY

As in [25], twenty datasets from the IPIX database [3] are used to testify detection ability of each feature and to analyze the main factor to affect its ability. As shown in Fig.6 in [25], the average signal-to-clutter ratios (ASCRs), Doppler offsets and bandwidths of sea clutter alter in wide ranges. Besides ASCR, overlapping extent of sea clutter and target returns in the Doppler domain or TF domain is another factor. Test targets, an anchored spherical block of Styrofoam wrapped with wire mesh in the first ten datasets and small boats at anchor in the last ten datasets, have Doppler offsets that vary pseudo-periodically around zero owing to joint actions of anchor and waves. When sea clutter has a large Doppler offset from zero and a small Doppler bandwidth, target returns and sea clutter little overlap in the Doppler domain and on the TF plane [25]. In this way, the overlapping extent can be measured by the ratio of the Doppler offset to the Doppler bandwidth of sea clutter and it is referred to as the overlapping index.

Each among the seven salient features can be used as a test statistic to yield a single-feature-based detector like the fractal-based detector. At the same observation time and false alarm rate, the detection ability of one feature on a dataset is evaluated by the detection probability for the test target. In experiments, the observation time is 0.512s, false alarm rate is  $10^{-3}$ , and the number of reference cells is  $P = 24$ . The detection probabilities of individual features at the twenty datasets are illustrated in Fig.1 (b-d). To examine the dependence of the detection ability on ASCR, the eighty datasets at the four polarizations are arranged in the ascending order of ASCR in Fig.1 (a). From Fig.1 (b-c), the detection abilities of the two amplitude features and the RVE have obvious ascending trend with ASCR, and their abilities highly depend on ASCR. From Fig.1 (c-d), the detection abilities of the RDPH and the three TF features sharply fluctuate with ASCR, and large detection probabilities are still achieved at some datasets of low ASCRs, showing that ASCR is not

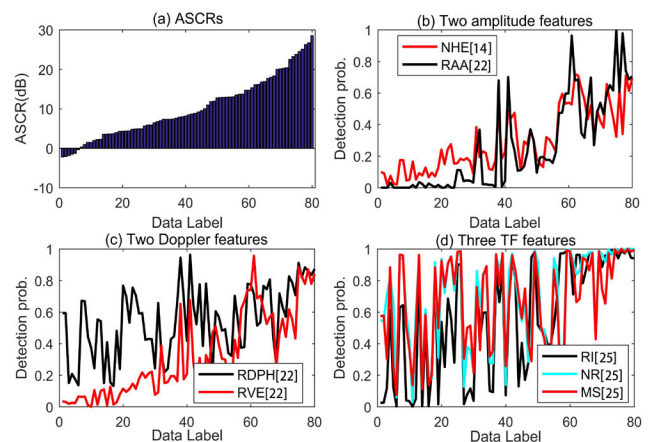


FIGURE 1. Detection probability comparison of the seven salient features on the twenty datasets.

major factor to affect their detection abilities. In fact, SCR is important for all the seven features. The ability of the RDPH mainly depends on the local SCR on the Doppler interval of target returns to occupy in the Doppler domain. The reason of the sharp fluctuation in Fig.1(c) is that large ASCRs do not always correspond to the large local SCR. Similarly, the abilities of the three TF features mainly depend on the local SCR on the TF region of target returns to occupy on the TF plane.

Table 1 lists the average, minimal, and maximal detection probabilities of the seven salient features on the twenty datasets for quantitative analysis. The average one reflects their overall ability and the order of the seven features is NR, MS, RI, RDPH, RVE, NHE, and RAA from strong to weak. The two amplitude features have the worst behavior. It is easy to explain. Sharp amplitude fluctuations of both target returns and sea clutter lower the ability of amplitude characteristics to discriminate them. In other words, non-coherent detection methods are not very effective for sea-surface small targets due to spiky high-resolution sea clutter and target returns of sharp amplitude fluctuation. The fact that the average detection probabilities of all the seven features are under 70% means that no single feature can realize effective detection of sea-surface small targets. For other features except the NHE, the maximal detection probabilities are over 95% and the minimal ones are under 15%. This fact shows that a single feature fails to provide stable detection results. For the seven salient features, the best results and worst results often occur at different datasets. This fact shows their complementary in detection ability.

**TABLE 1. Average, maximal and minimal detection probabilities of the seven salient features on the twenty datasets and correlation coefficients of their detection probability and ASCR and overlapping index.**

Salient Features	Detection probability			ASCR	Overlapping index
	average	maximal	minimal		
NHE	0.2931	0.7189	0.0172	0.7971	0.0275
RAA	0.2544	0.9957	0	0.8298	-0.0857
RDPH	0.5435	0.9634	0.1312	0.5198	0.2513
RVE	0.3227	0.9570	0	0.8480	-0.1039
RI	0.5655	1	0.0001	0.6666	0.3578
NR	0.6977	1	0.0294	0.4178	0.4668
MS	0.6707	1	0.0744	0.3951	0.4885

At last, we analyze the correlation of the detection probability and ASCR and overlapping index for each feature. The correlation coefficients of the seven features are listed in the last two columns of the Table 1. All the seven features positively correlate with ASCR and its increase lifts up the detection ability of every feature. This is just the reason for spatial high-resolution radars to be used in sea-surface small target detection for increasing ASCR. From the correlation coefficients, we draw the following conclusions. The ASCR is the major factor to affect the detection abilities of NHE, RAA, and RVE but the overlapping index hardly affects their abilities. The ASCR is still the main factor to affect

the detection abilities of RDPH and RI but the overlapping index also takes considerable effect. From the correlation coefficients of the NR and MS, it can be known that the correlation of the detection probability and overlapping index is higher than that of the detection probability and ASCR. So the overlapping index is the major factor to affect the detection abilities of NR and MS, and their abilities are also affected by ASCR. In conclusion, the seven existing features are complementary and their detection abilities depend on several factors. Therefore, in order to realize stable and effective detection of sea-surface small targets, all the seven features require to be jointly exploited.

### III. TRAINING SAMPLES GENERATOR AND FEATURE-COMPRESSION-BASED DETECTOR

Detection of sea-surface small targets encounters two unbalances. One is sufficient and complete sea clutter data versus insufficient and incomplete radar returns with targets. The other is required false alarm rate less than  $10^{-3}$  versus allowable missed probability of a few tenths. Due to the first unbalance, designing detectors mainly relies on the semisupervised learning based on characteristics of sea clutter under the anomaly detection framework. The problem boils down to design of one-class classifiers without anomaly instance or with insufficient anomaly instances as [15]–[18], [22], [25]. Due to the second unbalance, existing methods to design two-class classifiers in pattern recognition, which often require balanced mistake probabilities of two classes, fail to be directly applied in sea-surface small target detection problem. The convexhull learning algorithm in 3D feature space is an effective way to design feature-based detectors with controllable false alarm rate [22], [25] but the dimension limitation impedes exploitation of more features. Thus, it is necessary to compress the 7D feature vector to one 3D feature vector. The compression requires training samples or probability models of the two classes rather than only that of sea clutter. In this section, a generator of typical radar returns of sea-surface small targets is given. By the generator and measured sea clutter, training samples of the  $H_1$  hypothesis in the 7D feature space are generated. Using the training samples of the two hypotheses, a feature compression method to maximize interclass Bhattacharyya distance [28], [29] is proposed to transfer the 7D feature vector to a 3D feature vector. A modified convexhull learning algorithm in the compressed 3D feature space is presented to design a detector.

#### A. GENERATION OF RADAR RETURNS OF SEA-SURFACE SMALL TARGETS

Because of complex interactions between sea-surface small targets and waves, radar returns of various types of sea-surface small targets are difficult to be characterized by a universal model. Sea-surface small target detection in long observation time focuses on amplitude fluctuation and Doppler modulation of radar returns. Amplitude fluctuation originates from RCS change during observation, and Doppler modulation originates from the change of its radial velocity.

It is a long-term and time-consuming task to collect returns data of as many types of sea-surface small targets as possible by field tests. It is an approach to simulate typical returns of sea-surface small targets to assist design of a detector.

The RCS of a sea-surface small target has fluctuation, because its posture is severely affected by waves. The posture change rate relates to the change rate of swells on the sea surface, which is much slowly relative to the time scale of the radar's PRI of several milliseconds. The RCS or amplitude of the target slowly varies with pulses in a wide dynamic range. Based on this point, amplitude series of typical target returns are simulated by a highly correlated positive stochastic sequence. In the observation time of the order of seconds, the radial velocity of the sea-surface small target alters in a complex form for complex interactions between it and waves. The radial velocity is simulated by a simple linear model. In other words, the target is assumed to be of constant radial acceleration. Typical target returns are simulated by

$$s(n) = \bar{A}\sqrt{P_c}a(n) \times \exp\left(j\left[\frac{4\pi}{\lambda}(\vartheta_0(1 - \frac{n}{N}) + \vartheta_1\frac{n}{N})n\Delta t + \varphi_0\right]\right), \quad n = 1, 2, \dots, N \quad (2)$$

where  $P_c$  is the power of sea clutter,  $\bar{A}$  is a positive factor to adjust SCR,  $a(n)$  is a highly correlated amplitude sequence,  $\lambda$  is the radar wavelength,  $\vartheta_0, \vartheta_1$  are the initial and last radial velocities,  $\Delta t$  is the PRI of the radar, and  $\varphi_0$  is a random initial phase of the uniform distribution.

Below we discuss the parameter selection of the model (2). The amplitude sequence  $a(n)$  is a unit-mean positive stochastic sequence. When simulated target returns are added to sea clutter of power  $P_c$ , the SCR of radar returns is  $20 \log_{10}(\bar{A})$ . The parameter  $\bar{A}$  is generated by the uniform distribution of the interval  $[10^{-1}, 10^{1/2}]$ , corresponding to the SCR range from  $-20\text{dB}$  to  $10\text{dB}$ . In fact, targets can be always detected as the SCR is too high while fail to be detected as the SCR is too low. It is significant to simulate target movement and amplitude fluctuation. Sea-surface small targets generally have small velocity and acceleration. The simple model of constant acceleration is enough for sea-surface small targets when the observation time for a test is within several seconds. Target velocity is assumed to follow the uniform distribution of  $[-\eta, \eta]$ , target acceleration is restricted in  $[-\zeta, \zeta]$  in the observation time interval of length  $N\Delta t$ , and the angle between moving direction and the sight line of the radar follows the uniform distribution of  $[-\pi, \pi]$ . Under these assumptions, the initial and last radial velocities are generated by

$$\begin{aligned} x, y, z &\sim \text{Uniform}[-1, 1] \\ v_0 &= \eta x, \quad v_1 = \eta y, \quad |v_1 - v_0| \leq \zeta N\Delta t, \quad \theta = \pi z \\ \vartheta_0 &= v_0 \cos \theta, \quad \vartheta_1 = v_1 \cos \theta \end{aligned} \quad (3)$$

where the random numbers  $x, y$ , and  $z$  are mutually independent. Note that the random numbers  $x$  and  $y$  are generated again if the constraint on acceleration is not satisfied.

In simulation,  $\eta$  is taken as  $5 \text{ m/s}$  and  $\zeta$  is taken as  $2\text{m/s}^2$ , corresponding to the fact that the velocity of small targets is within 10 knots and their maximal acceleration is  $2\text{m/s}^2$  [7].

$a(n)$  is modeled as a nonnegative, highly correlated, unit-power stochastic sequence with adjustable dynamic range and decorrelated time, and is generated in four steps. First, an independent and identical distributed (IID) sequence  $u(n)$  is generated that follows the uniform distribution of  $[-1, 1]$  and has zero mean and a variance of  $1/3$ . Second, it is input into a first-order auto-regressive system to generate a highly correlated sequence  $v(n)$ ,

$$v(1) = u(1), \quad v(n+1) = \rho v(n) + u(n+1), \quad n = 1, 2, \dots; \rho \in (0, 1) \quad (4)$$

It can be proved that  $v(n) \in [-1/(1-\rho), 1/(1-\rho)]$ . Third, the sequence in (4) is shifted to generate a highly correlated non-negative stochastic sequence,

$$v_+(n) = \frac{1}{1-\rho} + v(n) \in \left[0, \frac{2}{1-\rho}\right], \quad n \in \mathbb{N} \quad (5)$$

Sea-surface small targets are occasionally fully shadowed by swells as the radar works at low grazing angle and the amplitudes reduce to zero. Shadowing easily occurs at high sea states. For examples, in the dataset #17 of the third sea state in the IPIX database [3], the test target is fully shadowed in about 36 seconds out of the observation time of 131 seconds [22].

The sequence (5) is not of unit power. When  $n$  is large enough or after the transition effect from the initial value vanishes,  $v(n)$  is zero-mean and its power equals to  $1/(3(1-\rho^2))$ . The power of the sequence in (5) is

$$E\{v_+^2(n)\} = \left(\frac{1}{1-\rho}\right)^2 + E\{v^2(n)\} = \frac{2(2+\rho)}{3(1-\rho)^2(1+\rho)}. \quad (6)$$

Fourth,  $a(n)$  is given by

$$\begin{aligned} a(n) &= \frac{v_+(n+M)}{\sqrt{E\{v_+^2\}}} \\ &= \frac{\sqrt{3(1+\rho)}(1-\rho)}{\sqrt{2(2+\rho)}} v_+(n+M), \quad n=1, 2, \dots, N \\ E\{a(n)\} &= \sqrt{\frac{3(1+\rho)}{2(2+\rho)}}, \quad \text{CV}(a(n)) = \sqrt{\frac{1-\rho}{3(1+\rho)}}. \end{aligned} \quad (7)$$

An enough large integer  $M$  is taken to avoid the transition effect from the initial value in the auto-regressive system. The sequence  $a(n)$  is correlated and unit-power but its mean is close to one only when the one-lag correlation coefficient  $\rho$  approximates to one. Its coefficients of variation (CV), the ratio of standard deviation to the mean, approximates to zero and little amplitude fluctuation occurs when the  $\rho$  is close to one. Besides the fluctuation of the  $a(n)$ , the  $\rho$  also controls the decorrelated time of the  $a(n)$ . Its  $k$ -lag correlation coefficient is

$$c_a(k) = \frac{E\{(a(n+k) - E\{a\})(a(n) - E\{a\})\}}{E\{(a(n) - E\{a\})^2\}} = \rho^k \quad (8)$$

For a time series, what are interested in applications are the coherent time and decorrelated time [34]. Change of the time series within the coherent time is ignored to be processed as a random constant. Two samples outside the decorrelated time are regarded to be uncorrelated. The coherent time and decorrelated time are all relevant to the  $k$ -lag correlation coefficient. It attenuates with the increase of the lag  $k$ . The largest lag of the correlation coefficients over 0.5 is referred to as the coherent length and the smallest lag of correlation coefficient under 0.1 is referred to as the decorrelated length. When the  $\rho = 0.95$ , the coherent length is 14 and the decorrelated length is 45. When the  $\rho = 0.99$ , the coherent length is 70 and the decorrelated length is 229. When the PRI of radar is one millisecond and the  $\rho \in [0.95, 0.99]$ , the coherent time of the  $a(n)$  is in [14ms, 70ms] and the decorrelated time is in [45ms, 229ms]. The analysis of the radar returns with targets in the IPIX database shows that the coherent time and the decorrelated time of the returns with targets in most of datasets fall into the two intervals mentioned above. In the simulation, the  $\rho$  randomly takes values by the uniform distribution of [0.95, 0.99].

The simulation of the target returns aims at generation of the training samples of the  $H_1$  hypothesis in the 7D feature space. A mass of sea clutter data can be collected even in several minutes once the radar starts to work. Thus, sufficient training samples of the  $H_0$  hypothesis are easy to obtain. By means of adding the simulated target returns to sea clutter, sufficient training samples of the  $H_1$  hypothesis are generated by the flowchart in Fig.2. Training sample collection is finished in a short time after the radar starts to work and thus hardly affects the normal operation of the radar. In the generation, more random parameters are used so that the generated training samples cover as many characteristics of sea-surface small targets as possible.

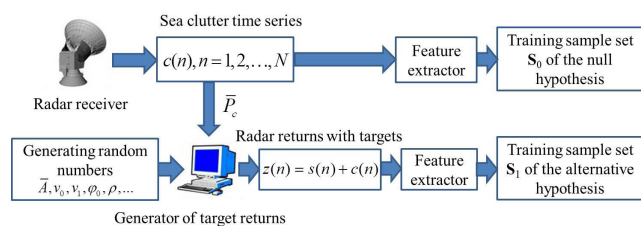


FIGURE 2. Flowchart diagram to obtain the training samples of the two hypotheses in the 7D feature space.

### B. FEATURE COMPRESSION METHOD BASED ON BHATTACHARYYA DISTANCE

When training samples of the two classes are available, the detection problem corresponds to design of a two-class classifier with quite unbalanced mistake probabilities of the two classes. Most of existing methods to train two-class classifiers fail to train two-class classifiers with this demand. More importantly, target returns generation can generate sufficient training samples of the  $H_1$  hypothesis. However,

these training samples remain incomplete. In other words, they fail to cover the characteristics of all sea-surface small targets of interest. Therefore, the detector is still designed in anomaly detection framework to assure the generalizability of the detector to various small targets. In order to utilize the convexhull learning algorithm [22], [25], the 7D feature vector must be compressed into one 3D feature vector with as little performance loss as possible from compression.

In what follows, we deal with feature compression problem. From the training sample sets of the two classes, the means and covariance matrices are computed by

$$\begin{aligned} \mu_0 &= \frac{1}{\#\mathbf{S}_0} \sum_{\mathbf{x} \in \mathbf{S}_0} \mathbf{x}, & \Sigma_0 &= \frac{1}{\#\mathbf{S}_0} \sum_{\mathbf{x} \in \mathbf{S}_0} (\mathbf{x} - \mu_0)(\mathbf{x} - \mu_0)^T, \\ \mu_1 &= \frac{1}{\#\mathbf{S}_1} \sum_{\mathbf{x} \in \mathbf{S}_1} \mathbf{x}, & \Sigma_1 &= \frac{1}{\#\mathbf{S}_1} \sum_{\mathbf{x} \in \mathbf{S}_1} (\mathbf{x} - \mu_1)(\mathbf{x} - \mu_1)^T, \end{aligned} \quad (9)$$

where  $\#\mathbf{S}$  denotes the cardinality of the finite set  $\mathbf{S}$ . Since  $\#\mathbf{S}_0$  and  $\#\mathbf{S}_1$  are very large, the means and covariance matrices can be regarded to be precise. It is impossible to theoretically build the probability models of the two classes in the 7D feature space because of complex characteristics of target returns and sea clutter and nonlinear operations in feature extraction.

Herein, the feature compression is based on the first and second-order statistics from the training samples. We consider a linear map. The feature compression is to find a  $3 \times 7$  projective matrix  $\mathbf{A}$  to map the 7D feature vector to a 3D vector so that the compressed feature vector keeps a classification ability as good as possible. Let the linear map be  $\mathbf{y} = \mathbf{A}\mathbf{x}$ ,  $\mathbb{R}^7 \rightarrow \mathbb{R}^3$ . Then, the compressed feature vector  $\mathbf{y}$  has the mean  $\mathbf{A}\mu_0$  and covariance matrix  $\mathbf{A}\Sigma_0\mathbf{A}^T$  under the  $H_0$  hypothesis and the mean  $\mathbf{A}\mu_1$  and covariance matrix  $\mathbf{A}\Sigma_1\mathbf{A}^T$  under the  $H_1$  hypothesis. A quantitative measure needs to be introduced to evaluate the separability of the two classes in the compressed 3D feature space so as to select a ‘good’ projective matrix  $\mathbf{A}$ . The interclass Bhattacharyya distance (B-distance) is the most effective measure in texture discrimination and other applications [38], [39]. It is applied to our problem. The interclass B-distance is a function of the projective matrix  $\mathbf{A}$  and is given by

$$\begin{aligned} d_B(\mathbf{A}) &= (\mu_0 - \mu_1)^T \mathbf{A}^T (\mathbf{A}\Sigma\mathbf{A}^T)^{-1} \mathbf{A} (\mu_0 - \mu_1) \\ &\quad + \ln \left( \frac{|\mathbf{A}\Sigma\mathbf{A}^T|}{\sqrt{|\mathbf{A}\Sigma_0\mathbf{A}^T| |\mathbf{A}\Sigma_1\mathbf{A}^T|}} \right), \\ \Sigma &= \Sigma_0 + \Sigma_1 \end{aligned} \quad (10)$$

where  $\Sigma_0$ ,  $\Sigma_1$ ,  $\Sigma$  are all positive-definite matrices and  $|\mathbf{A}|$  denotes the determinant of the matrix  $\mathbf{A}$ . The B-distance exists only when the matrix  $\mathbf{A}$  is full-rank in row, i.e.,  $rank(\mathbf{A}) = 3$ . A full-rank  $\mathbf{A}$  assures that the compressed features are at least linearly uncorrelated if so are the seven salient features. However, the  $rank(\mathbf{A}) = 3$  specifies an open manifold in the matrix space  $\mathbb{R}^3 \times \mathbb{R}^7$ . In some cases, the values of the determinants in (10) are very close to zero and

as a result the matrix  $\mathbf{A}\Sigma\mathbf{A}^T$  has a large condition number and the interclass B-distance is unstable in numerical computation. The full-rank in row is strengthened to the row-orthonormality, i.e.,  $\mathbf{A}\mathbf{A}^T = \mathbf{I}_3$ . It assures that the computation of the interclass B-distance is stable if the feature extraction makes the matrices  $\Sigma_0$  and  $\Sigma_1$  have good condition numbers. In this way, the feature compression boils down to an optimization

$$\begin{aligned} \max_{\mathbf{A} \in \mathbb{R}^{3 \times \mathbb{R}^7}} & \left\{ (\boldsymbol{\mu}_0 - \boldsymbol{\mu}_1)^T \mathbf{A}^T (\mathbf{A}\Sigma\mathbf{A}^T)^{-1} \mathbf{A} (\boldsymbol{\mu}_0 - \boldsymbol{\mu}_1) \right. \\ & \left. + \ln \left( \frac{|\mathbf{A}\Sigma\mathbf{A}^T|}{\sqrt{|\mathbf{A}\Sigma_0\mathbf{A}^T| |\mathbf{A}\Sigma_1\mathbf{A}^T|}} \right) \right\}, \\ \text{subject to } & \mathbf{A}\mathbf{A}^T = \mathbf{I}_3. \end{aligned} \quad (11)$$

An optimal compression matrix is searched to maximize the interclass B-distance.

The interclass B-distance degenerates to Mahalanobis distance (M-distance) without the second term in (10) [38], [39] if the covariance matrices in (10) are equal. The M-distance is used in Fisher linear discriminant analysis (LDA) to find a linear fusion of multiple features to characterize or separate the two classes. Equivalently, the feature vector is compressed into a scalar test statistic by a linear combination. In some sense, the feature compression in (10) is a generalization of the LDA. Owing to the fact that a linear combination is replaced by three linear combinations, the feature compression suffers from less loss from dimension reduction in comparison with feature fusion. For our problem, the interclass B-distance fails to be replaced by the M-distance in two reasons. On the one hand, from sea clutter data and radar returns with the test targets on the IPIX database [3], the covariance matrices of the two classes are quite different for each dataset. The condition of the interclass B-distance to degenerate to the M-distance does not hold. On the other hand, we try the feature compression using the M-distance instead of the B-distance and found that it brings significant loss in performance.

Due to highly nonlinear objective function and quadratic constraints, a good initial projective matrix and gradient ascend algorithm are cooperated to solve the optimization (11). The initial projective matrix is obtained by global search in all the 3D combinations of the seven features and all the 3D combinations of the seven eigenvectors of the positive-definite matrix  $\Sigma$ . Let  $\mathbf{e}_k, \mathbf{v}_k, k = 1, 2, \dots, 7$  be the coordinate vectors of the 7D feature space and the seven row eigenvectors of the matrix  $\Sigma$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_7$ , respectively. The initial projective matrix is selected from the finite set of projective matrices

$$\Lambda = \left\{ \mathbf{C}(k_1, k_2, k_3) = \begin{bmatrix} \mathbf{e}_{k_1} \\ \mathbf{e}_{k_2} \\ \mathbf{e}_{k_3} \end{bmatrix}, \mathbf{B}(k_1, k_2, k_3) = \begin{bmatrix} \mathbf{v}_{k_1} \\ \mathbf{v}_{k_2} \\ \mathbf{v}_{k_3} \end{bmatrix} \right\}, \quad 1 \leq k_1 < k_2 < k_3 \leq 7 \quad (12)$$

The matrix set  $\Lambda$  consists of seventy row-orthonormal projective matrices. The initial projective matrix is given by

$$\mathbf{A}_0 = \arg \max_{\mathbf{A} \in \Lambda} \{d_B(\mathbf{A})\}. \quad (13)$$

Starting from  $\mathbf{A}_0$ , the penalty function method [40] is used to find a solution of (11). The row-orthonormal constraint is not a rigid demand. Once the matrix  $\mathbf{A}$  is full-rank in row and the condition number of  $\mathbf{A}\mathbf{A}^T$  is not too large, it is a feasible solution of the problem. Thus, a penalty function is introduced as follows,

$$\Phi_\rho(\mathbf{A}) = d_B(\mathbf{A}) - \omega \|\mathbf{A}\mathbf{A}^T - \mathbf{I}_3\|_F^2 \quad (14)$$

where  $\omega > 0$  is the penalty factor. Using  $\mathbf{A}_0$  as the initial point, the gradient ascend algorithm [40] is used to maximize (14) to attain a solution  $\mathbf{A}_*$ . Generally speaking, a large penalty factor  $\omega$  corresponds to a small value of  $\|\mathbf{A}_*\mathbf{A}_*^T - \mathbf{I}_3\|_F^2$ . Therefore, the penalty factor can be tuned to make  $\|\mathbf{A}_*\mathbf{A}_*^T - \mathbf{I}_3\|_F^2 \approx \varepsilon (=0.1)$ , so as to assure approximate row-orthonormality of the projective matrix. By the matrix  $\mathbf{A}_*$ , the two sets of the training samples are transferred to the 3D feature space,

$$\begin{aligned} \tilde{\mathbf{S}}_0 &= \{ \mathbf{y}_k = \mathbf{A}_* \mathbf{x}_k \in \mathbb{R}^3 : \mathbf{x}_k \in \mathbf{S}_0 \}, \\ \tilde{\mathbf{S}}_1 &= \{ \mathbf{y}_k = \mathbf{A}_* \mathbf{x}_k \in \mathbb{R}^3 : \mathbf{x}_k \in \mathbf{S}_1 \}. \end{aligned} \quad (15)$$

Based on them, a detector is designed in the compressed 3D feature space.

### C. MODIFIED CONVEXHULL LEARNING ALGORITHM

In [22], [25], the two tri-feature-based detectors are designed in the anomaly detection framework by only using training samples of sea clutter. From the set  $\mathbf{S}_0$  of the training samples of the  $H_0$  hypothesis, designing a one-class classifier with a given false alarm rate  $P_F$  boils down to the optimization

$$\begin{aligned} \min_{\Omega \subset \mathbb{R}^3 \text{ is bounded convex set}} & \{ \text{volume}(\Omega) \} \\ \text{subject to } & \# \{ \mathbf{x} \in \Omega \cap \mathbf{S}_0 \} = (1 - P_F) \# \mathbf{S}_0. \end{aligned} \quad (16)$$

where the  $\text{volume}(\Omega)$  denotes the volume of a 3D convex set  $\Omega$  and  $\#\mathbf{S}$  denotes the cardinality of a finite point set  $\mathbf{S}$ . The solution of (16), the convexhull of minimal volume that contains  $(1-P_F)\#\mathbf{S}_0$  out of  $\#\mathbf{S}_0$  training samples of sea clutter, is the decision region of the  $H_0$  hypothesis in the tri-feature-based detector with a desired false alarm rate  $P_F$ . In (16), the convexity on the decision region is a regularity constraint to assure the generalizability of the one-class classifier and computability to search for a decision region. Solving (16) is a combination explosion problem, and a greedy convexhull learning algorithm is given to attain a solution [22], [25].

Differently, the sufficient but incomplete training samples of the  $H_1$  hypothesis are available. Thus, the volume of the decision region  $\Omega$  in (16) is replaced by the probability of the training set  $\tilde{\mathbf{S}}_1$  to fall outside the region  $\Omega$ . The convexity as regularity constraint is kept to assure the generalizability of



the detector. In the compressed 3D feature space, designing a detector boils down to the optimization,

$$\begin{aligned} & \min_{\Omega \subset \mathbb{R}^3 \text{ is bounded convex set}} \left\{ \# \left\{ \Omega \cap \tilde{\mathbf{S}}_1 \right\} \right\} \\ & \text{subject to } \# \left\{ \Omega \cap \tilde{\mathbf{S}}_0 \right\} = (1 - P_F) \# \tilde{\mathbf{S}}_0. \quad (17) \end{aligned}$$

It minimizes the missed probability on the training sample set of the  $H_1$  hypothesis subject to the false alarm rate on the training sample set of the  $H_0$  hypothesis, which accords with the Neyman-Pearson criterion [7] except the convex constraint on the decision region for generalizability and computability. For the combination explosion problem (17), a greedy convexhull learning algorithm, which is a modified version of the algorithm in [22], [25], is given to attain its sub-optimal solution. The steps of the algorithm are listed in Table 2.

**TABLE 2. Modified convexhull learning algorithm.**

**Input:** The sets of the training samples of the  $H_0$  and  $H_1$  hypotheses in the compressed 3D feature space,  $\tilde{\mathbf{S}}_0$  and  $\tilde{\mathbf{S}}_1$ , and the desired false alarm rate  $P_F$ .

**Output:** Decision region of the  $H_0$  hypothesis at the false alarm rate  $P_F$ , the convexhull  $\Omega_*$ .

1. **Initialization:** Compute the number of iterative shrinkages  $L = \lceil \# \tilde{\mathbf{S}}_0 \times P_F \rceil$ , where  $\lceil x \rceil$  is the integer part of a real number  $x$ . Let  $l = 0$ .
2. **WHILE**  $l <$  the iteration number  $L$
3. **Convexhull computation:** Compute the convexhull of the set of the training samples of the null hypothesis  $\mathcal{CH}(\tilde{\mathbf{S}}_0)$ . Set its vertices be  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$  and these vertices are all samples of the training set  $\tilde{\mathbf{S}}_0$ .
4. **Find the best shrinkage:** For  $q$  from 1 to  $r$ , compute the convexhull  $\Omega_q = \mathcal{CH}(\tilde{\mathbf{S}}_0 - \{\mathbf{v}_q\})$  where a vertex is removed from the training set of the  $H_0$  hypothesis and the convexhull is shrunk. Compute the number of the samples of  $\tilde{\mathbf{S}}_1$  to fall into the set the convexhull  $\Omega_q$ ,
 
$$N(q) = \# \{ \tilde{\mathbf{S}}_1 \cap \Omega_q \}, q = 1, 2, \dots, r. \quad (18)$$
 The best shrinkage is given by
 
$$\tilde{q} = \arg \min_{q=1,2,\dots,r} \{ N(q) \}. \quad (19)$$
5. **Shrink the training set of the  $H_0$  hypothesis:** Replace  $\tilde{\mathbf{S}}_0$  by  $\tilde{\mathbf{S}}_0 - \{\mathbf{v}_{\tilde{q}}\}$  and set  $l = l + 1$ .
6. **END WHILE**
7. **Output the decision region:**

$$\Omega_* = \Omega_{\tilde{q}}. \quad (20)$$

Once the decision region of the  $H_0$  hypothesis is obtained, a decision is made by the extraction of the 7D feature vector, feature compression, and binary decision. When the compressed feature vector falls into the decision region, the  $H_0$  hypothesis holds at the CUT. Otherwise, the  $H_1$  hypothesis holds and a target is declared to be in the CUT. The decision is made by computing multiple  $3 \times 3$  determinants [22], [25] and thus can be fast implemented. Comparing with other learning algorithms, the convexhull learning algorithm can accurately control false alarm rate. It is worthy of noting that the learning process is designed under the anomaly detection framework by mainly using the training samples of sea clutter. This lowers the effect of the incompleteness of the training samples

of the  $H_1$  hypothesis and assures that the proposed feature-compression-based detector possesses strong generalizability to other types of sea-surface small targets.

As a detector based on learning the characteristics of sea clutter from data, it needs to adapt the change of characteristics of sea clutter. Below, we discuss its use in practical radar systems. For an island-based or shipborne high-resolution radar with fixed operating parameters, the characteristics of sea clutter are mainly affected by sea states and viewing geometry of the radar. Generally, sea state varies at the time scale of tens of minutes and at the spatial scale of hundreds of square kilometers. The viewing geometry of the radar includes the grazing angle and angle between wave direction and the sight line of the radar. The grazing angle varies with radial distance, and the latter varies with azimuth. Detecting sea-surface small targets, the surveillance scene of the radar is partitioned into multiple sections along the radial distance and azimuth. On each section, one feature-compression-based detector is trained for detection. Moreover, the decision region is aperiodically learned again as sea state in the section obviously alters. In this way, the detector based on online training can adapt spatial-temporally varying oceanic environment. Each training process spends short time and little influences the normal work of the radar. In addition, when the radar is at work, the feature vectors at the resolution cells that are declared to have a target can be collected to enrich the set of training samples of the  $H_1$  hypothesis to improve the detection performance.

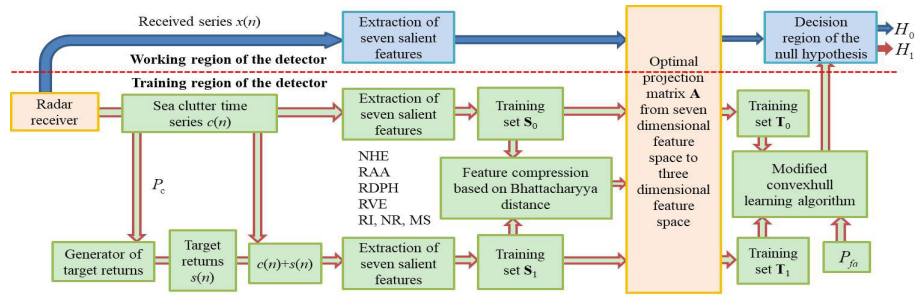
At last, a flowchart diagram of the proposed feature-compression-based detector is illustrated in Fig.3. It consists of the training branch and working branch. The pink boxes denote the common blocks, the green boxes denote the blocks in the training stage, and the baby blue boxes denote the blocks in the detection stage. The training branch includes more blocks, which may be carried out offline. The detection branch includes only three blocks: extraction of the 7D feature vector, feature compression via the linear transform, and the decision whether it falls into the convexhull or not. And thus the detection can be performed fast. It is worthy to be noted that as an open framework the feature-compression-based method allows more salient features to be utilized to improve detection performance.

#### IV. EXPERIMENTAL RESULTS AND PERFORMANCE COMPARISON

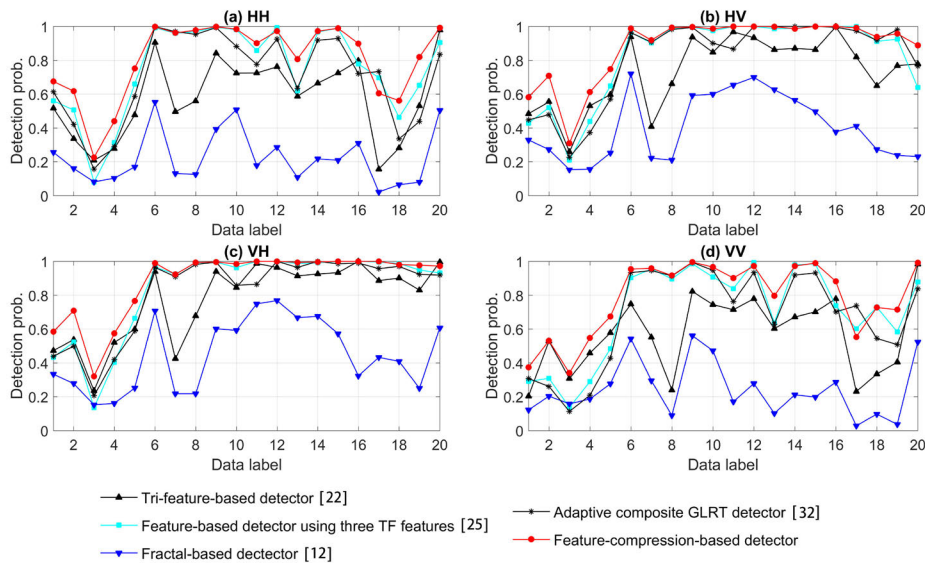
In this section, experimental results of the proposed detector on the IPIX database [3] and two datasets of the CSIR database [4] are reported. The performance comparisons with existing detectors are given. Moreover, some limitations are discussed.

##### A. EXPERIMENTAL RESULTS ON TWENTY DATASETS OF IPIX DATABASE

Twenty datasets at the dwelling mode from the IPIX database are available for performance evaluation. The first ten datasets were collected at the Dartmouth, Nova, Scotia,



**FIGURE 3.** Flowchart of the feature-compression-based detector including its working and training branches.

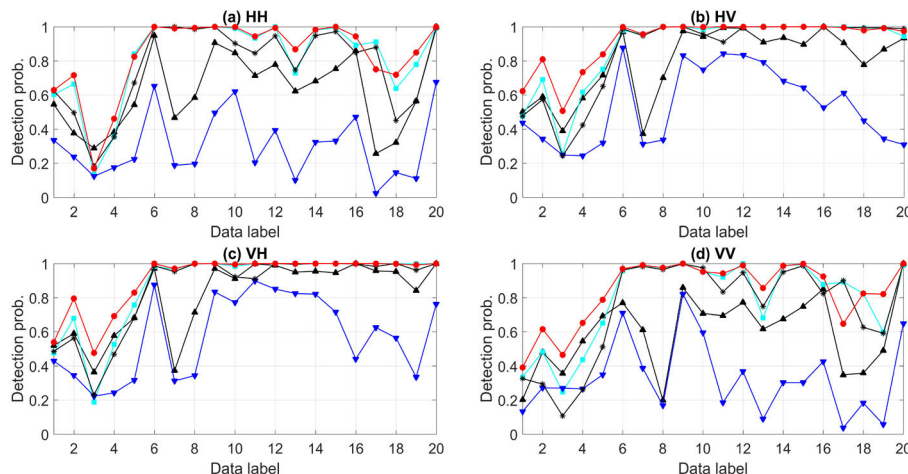


**FIGURE 4.** Detection probabilities of the feature-compression-based detector and four existing detectors at the twenty datasets where the false alarm rate is 0.001 and the observation time is 0.512s.

Canada, in 1993. The radar was mounted on a cliff 100 feet above sea level facing the Atlantic ocean at a low grazing angle (about  $0.33^\circ$ ). Each dataset contains radar returns synchronously collected at the HH, HV, VH, and VV polarizations. At each polarization, the data consist of complex returns time series of length  $2^{17}$  at 14 adjacent range cells (about 131 seconds). The test target is an anchored spherical block of Styrofoam wrapped with wire mesh and diameter of about 1m. The range cell of the target to occupy is referred to as the primary cell. Around it, two or three cells whose returns are affected by the target are called the secondary cells. The last ten datasets were collected by the IPIX radar with an improved quantizer at Grimsby, Ontario, Canada, in 1998. The test targets are small boats at anchor and each dataset consists of complex returns of length  $6 \times 10^4$  (one minute) at 28 contiguous range cells at the four polarizations. The range resolution of the first eighteen datasets is 30m, that of the 19<sup>th</sup> dataset is 15m and that of the 20<sup>th</sup> dataset is 9m. These datasets were collected at different sea states and thus can give high-confidence assessments. In experiments, for each dataset at each polarization, the sets of training samples of

the 7D feature vector under the  $H_0$  and  $H_1$  hypotheses are computed from sea clutter time series and the simulated target returns. And then, the two sets are used to train the detector. The set of the test samples of the 7D feature vectors are computed from the time series at the primary cells. This set is used to compute detection probabilities. In the experiments, take  $N = 512$  and  $1024$ , i.e., the observation time for one decision is 0.512s and 1.024s respectively, and the false alarm rate is 0.001.

On the twenty datasets, the proposed feature-compression-based detector is compared with the early fractal-based detector [12] and our previous three detectors [22], [25], [32]. Fig.4 and Fig.5 illustrate the detection probabilities of the five detectors on the twenty IPIX datasets, where the four polarizations are separately plotted in subfigures. The feature-compression-based detector attains the largest detection probabilities at 75 out of the eighty datasets, and its detection probabilities are quite close to the largest ones at the five exceptions. The early fractal-based detector has the smallest detection probabilities due to limited ability of a single feature.



**FIGURE 5.** Detection probabilities of the feature-compression-based detector and four existing detectors at the twenty datasets where the false alarm rate is 0.001 and the observation time is 1.024s.

**TABLE 3.** Average detection probabilities comparison at the twenty datasets of the IPIX database.

Detectors	Observation time	HH	HV	VH	VV	Average
Fractal-based detector [12]	0.512s	0.223	0.404	0.448	0.241	0.329
	1.024s	0.301	0.536	0.576	0.328	0.435
Tri-feature-based detector [22]	0.512s	0.577	0.736	0.776	0.569	0.665
	1.024s	0.622	0.797	0.813	0.598	0.708
TF-tri-feature-based detector [25]	0.512s	0.747	0.826	0.842	0.706	0.780
	1.024s	0.821	0.882	0.877	0.789	0.842
Composite GLRT detector [32]	0.512s	0.705	0.817	0.821	0.677	0.755
	1.024s	0.770	0.857	0.856	0.739	0.806
Feature-compression-based detector	0.512s	<b>0.807</b>	<b>0.880</b>	<b>0.888</b>	<b>0.788</b>	<b>0.841</b>
	Upper bounds	<b>0.831</b>	<b>0.888</b>	<b>0.889</b>	<b>0.810</b>	<b>0.855</b>
	1.024s	<b>0.841</b>	<b>0.920</b>	<b>0.914</b>	<b>0.839</b>	<b>0.879</b>

Table 3 lists the average detection probabilities of the five detectors at each polarization when  $N = 512$  and 1024. Doubling observation time brings some benefits for all the five detectors. It is found that the proposed detector attains the least benefit due to the ceiling effect. Or rather, when the observation time is 0.512s, it attains the detection probabilities quite close to one (so-called ceiling). So doubling observation time has little space to further improve detection performance. In terms of the overall performance, the proposed detector is the best one, followed by the TF-tri-feature-based detector [25], adaptive composite GLRT detector [32], tri-feature-based detector [22], and fractal-based detector [12]. In comparison with the second best one, the feature-compression-based detector improves the overall detection probability of 6.1% at the observation time of 0.512s and of 3.7% at the observation time of 1.024s. In fact, due to usage of more salient features and their complementarity, the proposed detector provides better and more stable detection performance.

For a full comparison, the feature-compression-based detector is compared with two recent detectors using the IPIX database: the decision tree based detector [30] and detector based on graph description of amplitude series [31]. The former uses three features, Hurst exponents in the time domain and frequency domain, and RDPH, and the detector is designed by learning two-class decision tree in the 3D feature space from radar data. The latter is based on subtle graph modelling of amplitude series of sea clutter, and is a substantially single-feature-based detector. Table 4 lists detection probabilities of the three detectors on the ten datasets in 1993 of the IPIX database, where the false alarm rate is 0.001. In terms of the analysis method in [22], in the third dataset (#30), the test target is fully shadowed by swells during about 60 seconds out of 131 seconds. This dataset is ignored in computation of the average detection probabilities. The proposed detector has detection probability loss of 0.02 in comparison with the decision tree based detector for the ten datasets of the IPIX database at the HH

**TABLE 4.** Comparison of detection performance of the IPIX database.

Detectors	Datasets for test	Observation time	Average detection prob.
Decision Tree-based detector [30]	Ten datasets(HH) IPIX 1993	0.512s	0.850
		1.024s	0.860
		2.048s	0.870
		4.096s	0.890
Feature-compression-based detector	Ten datasets(HH) IPIX 1993	0.512s	0.823
		1.024s	0.846
		2.048s	0.908
		4.096s	0.960
Graph-based detector [31]	#17(HH, HV) #54(HH, HV, VH, VV)	0.512s	0.643
Feature-compression-based detector	#17(HH, HV) #54(HH, HV, VH, VV)	0.512s	0.864

**TABLE 5.** Average B-distances at each polarization for the three tri-feature-based detectors.

Detector	Polarizations	HH	HV	VH	VV
Tri-feature-based detector [22]	Training sample set versus test sample set	1.697	2.310	2.683	1.528
TF-tri-feature-based detector [25]		4.809	6.027	6.658	4.211
Feature-compression-based detector	Initial projective matrices $A_0$	4.831	6.080	6.617	4.335
	Final projective matrices $A_*$	4.945	6.170	6.740	4.370

polarization when the observation time is 0.512s and 1.024s. As the observation time becomes longer, the proposed detector attains some improvement. Besides, the proposed detector provides the average detection probability improvement of 0.22 in comparison with the graph-based detector for six datasets of the IPIX database. Due to a single feature and absence of phase information of radar returns in detection, the detector using graph description has poor performance though it provides obvious improvement relative to single-feature-based detectors.

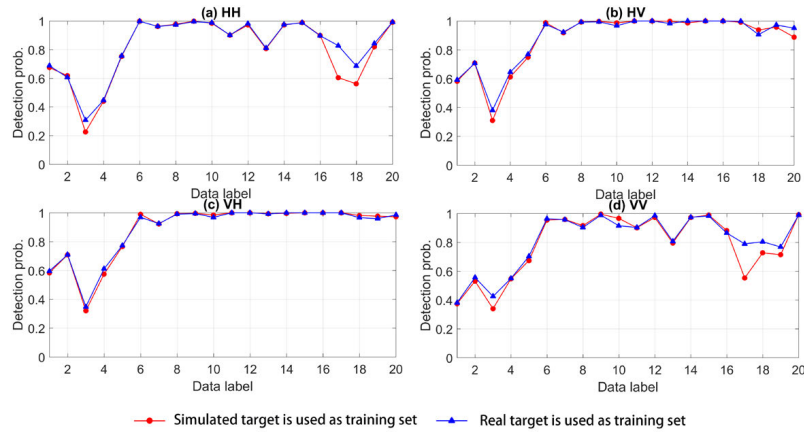
In what follows, we discuss the effectiveness of the interclass B-distance in the feature compression. The two tri-feature-based detectors [22], [25] involve in two sample sets in the 3D feature spaces: the training sample set of the  $H_0$  hypothesis from sea clutter, the test sample set from radar returns with test targets. Besides the two sets, the feature-compression-based detector still involves in the training sample set of the  $H_1$  hypothesis from the simulated target returns plus sea clutter. For the three detectors, the average B-distances between the first two sets in the 3D feature spaces at each polarization are listed in Table 5. Contrasting the average detection probabilities in Table 3 and the average interclass B-distance in Table 5, it is found that larger average B-distances always correspond to larger average detection probabilities. This fact shows that the interclass B-distance is a good measure for feature selection and compression. The average B-distances under the initial projective matrices  $A_0$  and the optimized projective matrices  $A_*$  are also listed in Table 5. It is seen that optimization increases the average interclass B-distances of the training sample set of the  $H_0$  hypothesis and the test sample set of the  $H_1$  hypothesis,

though the optimization is based on the statistics of the training sample sets of the  $H_0$  and  $H_1$  hypotheses. This fact shows that the generation method of target returns is effective at least for the test targets in the twenty datasets of the IPIX database.

In order to further testify the effectiveness of the generation method of target returns, we use the training sample set of the  $H_0$  hypothesis and the test sample set of the  $H_1$  hypothesis for the feature compression and convexhull learning. For each dataset, the obtained feature-compression-based detector should give an upper bound of the detection performance because the test samples for assessment are also used for learning. The extent of the detection probability of the feature-compression-based detector to be close to the upper bound can assess the effectiveness of the generation method of target returns. Fig.6 is the detection probabilities of the feature-compression-based detector and the upper bounds when the observation time is 0.512s and the false alarm rate is 0.001. The detection probabilities are almost equal to the upper bounds at 71 out of the 80 datasets. This fact shows that the generation method of target returns is effective.

## B. EXPERIMENTAL RESULTS ON TWENTY DATASETS OF CSIR DATABASE

In order to testify the adaptivity of the proposed detector for other types of sea-surface small targets, two datasets (TFA17-005 and TFC15-007) from the CSIR database [4] are used, where the test targets are a moving wooden fishing boat of length of about 5.5m and a moving wave-rider rigid inflatable boat of length 5.7m, respectively. The details of the two datasets are listed in Table 6. Each dataset consists of



**FIGURE 6.** Comparison of the detection probabilities of the feature-compression-based detector and the upper bounds using the test sample set for feature compression and training as the observation time is 0.512s and the false alarm rate is 0.001.

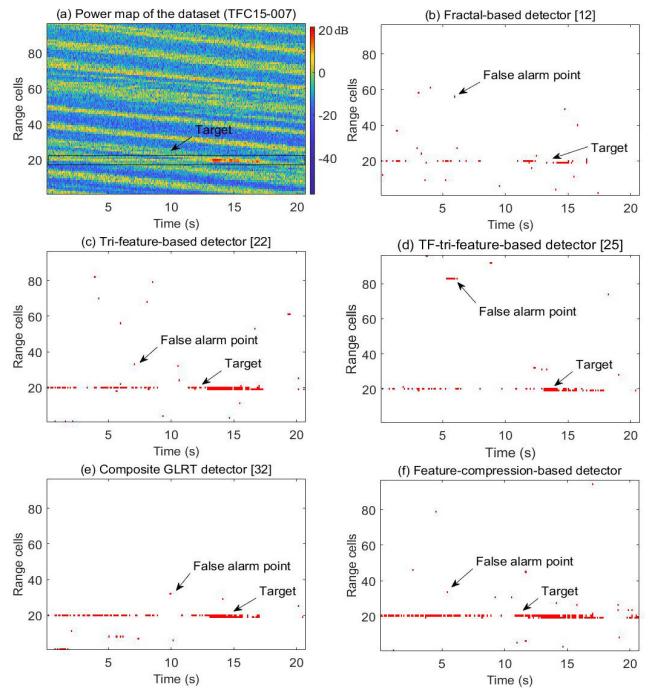
**TABLE 6.** Parameters of the two datasets in the CSIR database [4].

File Name	TFA17-005	TFC15-007
Carrier frequency	6.9GHz (C-band)	9GHz (X-band)
Polarization	VV	VV
Range resolution	15m	15m
PRF	5000 Hz	5000 Hz
Significant wave height	2.26m	3.22m
SCR	11.6 dB	10.2 dB
Target range cells	29 <sup>th</sup> -31 <sup>th</sup>	19 <sup>th</sup> -21 <sup>th</sup>

**TABLE 7.** Detection probabilities of the five feature-based detectors.

Dataset	TFA17-005	TFC15-007
Fractal-based detector [12]	0.14	0.22
Tri-feature-based detector [22]	0.63	0.51
TF-tri-feature-based detector [25]	0.78	0.35
Composite GLRT detector [32]	0.76	0.58
Feature-compression-based detector	<b>0.80</b>	<b>0.64</b>

radar returns of 25 seconds at 96 contiguous range cells at tracking mode. In experiment,  $N = 512$  corresponds to the observation time of 0.124 and the false alarm rate is 0.001. The detection probabilities of the five detectors at the two datasets are listed in Table 7. The proposed detector attains the largest detection probabilities for the two test targets. Differently, the composite GLRT detector [32], that focuses on nonstationary movement of the targets, obtains very good performance. Fig.7 illustrates the power map of the second dataset and detection results, where each subplot has eighteen false alarm points. The fractal-based detector hardly attains the track of the fishing boat, because phase information of the radar returns is not used. The proposed detector obtains a more continuous track of the boat. The proposed detector possesses a strong generalizability to other types of sea-surface small targets, owing the fact that it is designed in the



**FIGURE 7.** Detection performance comparison at the second dataset of the CSIR database [4]. (a) Power map of the data, (b) fractal-based detector [12], (c) tri-feature-based detector [22], (d) TF-tri-feature-based detector [25], (e) composite GLRT detector [32], and (f) feature-compression-based detector.

anomaly detection framework and the radar returns of special test targets do not participate in design of the detector.

## V. CONCLUSION

In this paper, a feature-compression-based detector was developed for effective and robust detection of sea-surface small targets. By means of the generation of target returns, the training samples of radar returns with targets are generated and cooperate with the training samples of sea clutter for feature compression. The feature compression based on interclass B-distance was proposed to compress the

7D feature vector into one 3D feature vector. In the compressed feature space, the modified convexhull learning algorithm is given to determine the decision region of sea clutter at a given false alarm rate. Experimental results on the IPIX radar database and two datasets of CSIR database show that the feature-compression-based detector attains better and more robust detection performance. Besides better detection performance, the other merit is to allow more salient features to be exploited.

It is a long-term work to improve the detection ability of maritime surveillance radars to sea-surface small targets. As an open framework, the future research devotes to the exploitation of more features to further improve detection performance and construction of replacements of interclass B-distance. It is known that the acquisition of training samples of the  $H_1$  hypothesis is a key to relax the limitation of one-class classifiers. Though the simulation of target returns can realize the sufficiency of the training samples of the  $H_1$  hypothesis, it fails to overcome substantially the incompleteness of training samples of the  $H_1$  hypothesis in the feature space to all types of sea-surface small targets at all sea states. It is a potential way to achieve the completeness of the training samples by long-term collecting data of the  $H_1$  hypothesis during the radar works.

## REFERENCES

- [1] J. J. Ousborne, D. Griffith, and R. W. Yuan, "A periscope detection radar," *Johns Hopkins APL Technical Dig.*, vol. 19, no. 1, pp. 125–133, 1997.
- [2] M. McDonald and S. Lycett, "Fast versus slow scan radar operation for coherent small target detection in sea clutter," *IEE Proc. Radar, Sonar Navigat.*, vol. 152, no. 6, pp. 429–435, 2005.
- [3] Cognitive Systems Laboratory, McMaster University, Canada. *The IPIX Radar Database[OL]*. Accessed: Oct. 11, 2012. [Online]. Available: <http://soma.mcmaster.ca/ipix.php>
- [4] The Defence, Peace, Safety, and Security Unit of the Council for Scientific and Industrial Research. *The Fynmeet Radar Database[OL]*. Accessed: Dec. 13, 2018. [Online]. Available: [http://www.csir.co.za/small\\_boat\\_detection](http://www.csir.co.za/small_boat_detection)
- [5] D. J. Rabideau and P. Parker, "Ubiquitous MIMO multifunction digital array radar," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, 2003, pp. 1057–1064.
- [6] P. Carta, G. Galati, E. G. Piracci, F. Madia, and R. Ronconi, "Implementation of the ubiquitous radar concept with a conformal array," in *Proc. Eur. Microw. Conf. (EuMC)*, Paris, France, 2015, pp. 1463–1466.
- [7] K. D. Ward, R. J. A. Tough, and S. Watts, *Sea Clutter: Scattering, the K Distribution and Radar Performance*, 2nd ed. London, U.K.: IET, 2013, pp. 101–134.
- [8] J. Zhu, X. Wang, X. Huang, S. Suvorova, and B. Moran, "Detection of moving targets in sea clutter using complementary waveforms," *Signal Process.*, vol. 146, pp. 15–21, May 2018.
- [9] P.-L. Shui and M. Liu, "Subband adaptive GLRT-LTD for weak moving targets in sea clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 423–437, Feb. 2016.
- [10] H. Leung and T. Lo, "Chaotic radar signal processing over the sea," *IEEE J. Ocean. Eng.*, vol. 18, no. 3, pp. 287–295, Jul. 1993.
- [11] J. Guan, N. B. Liu, J. Zhang, and J. Song, "Multifractal correlation characteristic for radar detecting low-observable target in sea clutter," *Signal Process.*, vol. 90, no. 2, pp. 523–535, 2010.
- [12] J. Hu, W.-W. Tung, and J. Gao, "Detection of low observable targets within sea clutter by structure function based multifractal analysis," *IEEE Trans. Antennas Propag.*, vol. 54, no. 1, pp. 136–143, Jan. 2006.
- [13] X.-K. Xu, "Low observable targets detection by joint fractal properties of sea clutter: An experimental study of IPIX OHGR datasets," *IEEE Trans. Antennas Propag.*, vol. 58, no. 4, pp. 1425–1429, Apr. 2010.
- [14] D. Li and P. Shui, "Floating small target detection in sea clutter via normalised Hurst exponent," *Electron. Lett.*, vol. 50, no. 17, pp. 1240–1242, Aug. 2014.
- [15] G. Hennessey, H. Leung, A. Drosopoulos, and P. Yip, "Sea-clutter modeling using a radial-basis-function neural network," *IEEE J. Ocean. Eng.*, vol. 26, no. 3, pp. 358–372, Jul. 2001.
- [16] H. Leung, T. Lo, and J. Litva, "Nonlinear adaptive signal processing based on rational function," *Signal Process.*, vol. 38, no. 2, pp. 153–168, Jul. 1994.
- [17] N. Xie, H. Leung, and H. Chan, "A multiple-model prediction approach for sea clutter modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1491–1502, Jun. 2003.
- [18] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 75–79, 2009.
- [19] T. Nohara and S. Haykin, "AR-based growler detection in sea clutter," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1259–1271, Mar. 1993.
- [20] T. Bhattacharya and S. Haykin, "Neural network-based radar detection for an ocean environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 2, pp. 408–420, Apr. 1997.
- [21] R. Vicen-Bueno, M. Rosa-Zurera, M. P. Jarabo-Amores, and D. de la Mata-Moya, "Coherent detection of swerling 0 targets in sea-ice weibull-distributed clutter using neural networks," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 12, pp. 3139–3151, Dec. 2010.
- [22] P.-L. Shui, D.-C. Li, and S.-W. Xu, "Tri-feature-based detection of floating small targets in sea clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1416–1430, Apr. 2014.
- [23] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa, "The quick hull algorithm for convex hulls," *ACM Trans. Math. Softw.*, vol. 22, no. 2, pp. 460–483, 1996.
- [24] M. D. Berg, M. V. Kreveld, M. Overmars, and O. Schwarzkopf, *Computational Geometry: Algorithms and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2000, pp. 233–248.
- [25] S.-N. Shi and P.-L. Shui, "Sea-surface floating small target detection by one-class classifier in time-frequency feature space," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6395–6411, Nov. 2018.
- [26] T. Islam, M. A. Rico-Ramirez, D. Han, and P. K. Srivastava, "Artificial intelligence techniques for clutter identification with polarimetric radar signatures," *Atmos. Res.*, vols. 109–110, pp. 95–113, Jun. 2012.
- [27] D. Gallagher, J. Burger, and A. K. Mishra, "A machine learning approach to radar sea clutter suppression," in *Proc. Radar Conf.*, Seattle, WA, USA, May 2017, pp. 1222–1227.
- [28] E. Choi and C. Lee, "Feature extraction based on the Bhattacharyya distance," *Pattern Recognit.*, vol. 36, no. 8, pp. 1703–1709, Aug. 2003.
- [29] C. Reyes-Aldasoro and A. Bhalerao, "The Bhattacharyya space for feature selection and its application to texture segmentation," *Pattern Recognit.*, vol. 39, no. 5, pp. 812–826, May 2006.
- [30] H. Zhou and T. Jiang, "Decision tree based sea-surface weak target detection with false alarm rate controllable," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 793–797, Jun. 2019.
- [31] K. Yan, Y. Bai, H.-C. Wu, and X. Zhang, "Robust target detection within sea clutter based on graphs," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7093–7103, Sep. 2019, doi: 10.1109/TGRS.2019.2911451.
- [32] S.-N. Shi and P.-L. Shui, "Detection of low-velocity and floating small targets in sea clutter via income-reference particle filters," *Signal Process.*, vol. 148, pp. 78–90, Jul. 2018.
- [33] J. Carretero-Moya, J. Gismero-Menoyo, and A. Blanco-del-Campo, "Statistical analysis of a high-resolution sea clutter database," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 2024–2037, 2010.
- [34] M. Greco, F. Bordoni, and F. Gini, "X-band sea-clutter nonstationarity: Influence of long waves," *IEEE J. Ocean. Eng.*, vol. 29, no. 2, pp. 269–283, Apr. 2004.
- [35] J. Gemmrich, J. Thomson, W. E. Rogers, A. Pleskachevsky, and S. Lehner, "Spatial characteristics of ocean surface waves," *Ocean Dyn.*, vol. 66, no. 8, pp. 1025–1035, Aug. 2016.
- [36] X. Shuai, L. Kong, and J. Yang, "Performance analysis of GLRT-based adaptive detector for distributed targets in compound-Gaussian clutter," *Signal Process.*, vol. 90, no. 1, pp. 16–23, Jan. 2010.
- [37] K. J. Sangston, F. Gini, and M. S. Greco, "Coherent radar target detection in heavy-tailed compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 64–77, Jan. 2012.
- [38] D. Ververidis and C. Kotropoulos, "Information loss of the mahalanobis distance in high dimensions: Application to feature selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2275–2281, Dec. 2009.

- [39] S. Xiang, F. Nie, and C. Zhang, "Learning a Mahalanobis distance metric for data clustering and classification," *Pattern Recognit.*, vol. 41, no. 12, pp. 3600–3612, Dec. 2008.
- [40] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 1999, ch. 9, pp. 490–526.



**ZIXUN GUO** was born in Xi'an, Shaanxi, China, in 1994. She received the B.S. degree in electrical engineering from Xidian University, Xi'an, in 2016, where she is currently pursuing the Ph.D. degree in signal and information processing with the National Laboratory of Radar Signal Processing.

Her research interests include radar signal processing, weak target detection, and their applications.



**PENGLANG SHUI** (Senior Member, IEEE) was born in Xi'an, China, in 1967. He received the B.S. degree in applied mathematics from Xidian University, Xi'an, in 1989, the M.S. degree in mathematics from Nanjing University, Nanjing, China, in 1992, and the Ph.D. degree in electrical engineering from Xidian University, in 1999.

He is currently a Professor with the National Laboratory of Radar Signal Processing, Xidian University. His research interests include multirate filter bank theory and applications, image processing, and radar target detection and tracking.



**SAINAN SHI** was born in Nantong, Jiangsu, China, in 1990. She received the B.S. degree in electrical engineering from Xidian University, Xi'an, China, in 2013, and the Ph.D. degree from the National Laboratory of Radar Signal Processing, Xidian University, in 2018.

She is currently a Lecturer with the College of Electronic and Information Engineering, Nanjing University of Information Science and Technology. Her research interests include radar signal processing and weak target detection in sea clutter.

...