

Article

Matrix Information Geometry for Spectral-Based SPD Matrix Signal Detection with Dimensionality Reduction

Sheng Feng ¹ , Xiaoqiang Hua ^{2,*} and Xiaoqian Zhu ²

¹ College of Computer Science, National University of Defense Technology, Changsha 410073, China; fengsh14@lzu.edu.cn

² College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China; zhu_xiaoqian@nudt.edu.cn

* Correspondence: hxq712@yeah.net; Tel.: +86-138-7496-9340

Received: 12 July 2020; Accepted: 19 August 2020; Published: 20 August 2020

Abstract: In this paper, a novel signal detector based on matrix information geometric dimensionality reduction (DR) is proposed, which is inspired from spectrogram processing. By short time Fourier transform (STFT), the received data are represented as a 2-D high-precision spectrogram, from which we can well judge whether the signal exists. Previous similar studies extracted insufficient information from these spectrograms, resulting in unsatisfactory detection performance especially for complex signal detection task at low signal-noise-ratio (SNR). To this end, we use a global descriptor to extract abundant features, then exploit the advantages of matrix information geometry technique by constructing the high-dimensional features as symmetric positive definite (SPD) matrices. In this case, our task for signal detection becomes a binary classification problem lying on an SPD manifold. Promoting the discrimination of heterogeneous samples through information geometric DR technique that is dedicated to SPD manifold, our proposed detector achieves satisfactory signal detection performance in low SNR cases using the K distribution simulation and the real-life sea clutter data, which can be widely used in the field of signal detection.

Keywords: dimensionality reduction; signal detection; SPD manifold; spectrogram processing

1. Introduction

Signal detection is a challenging task in signal processing [1]. As the basic subject in object detection, it is acknowledged as a very valuable research which arouses lots of researchers [2]. However, since complex clutter noise and interference are ubiquitous in the context of signal detection, this crucial task becomes extremely difficult. The main technique in dealing with this issue rely on the Neyman–Pearson criterion [3]. Specifically, by establishing a binary signal statistical detection model, that is, the binary hypothesis H_0 (no signal exists) or H_1 (signal exists), the task for signal detection is accomplished with constant false alarm technique, which is referred to as “CFAR”. Although this classical method has made many achievements in some application fields with high efficiency [4], its detection performance still appears a serious bottleneck that limits its practicability, especially for the case of heterogeneous clutter under low signal-noise-ratio (SNR) [5]. Hence, how to achieve satisfactory detection performance under such complex detection background is an issue worthy of attention [6].

From the perspective of time-frequency analysis, known as a powerful technique of research in signal processing [6,7], the sample data processed after short time Fourier transform (STFT) [8] can be exploited for signal detection due to the significant difference between noise and signal. Recent studies have extracted relevant information after STFT which shows that it is feasible for signal detection from

the perspective of time-frequency technique [9,10]. More specifically, a signal exists in an area of the spectrogram where the energy change exceeds a certain threshold. In fact, the detection performance of these time-frequency analysis based methods greatly depends on the quality of extracted features from spectrogram. However, former researches only extracted few features maintaining local information, and thus gives inadequate descriptions on the sample data. Meanwhile, their researches are mainly under the prior condition that signal exists, in other words, the problem of false alarm has not been taken into consideration. Hence, such spectrogram processing based methods have many constructive issues to be solved [11].

On the other hand, in the area of matrix information geometry, data representation by symmetric positive definite (SPD) matrix has been widely applied in many scientific researches, e.g., pattern recognition [12], image processing [13], signal processing [14,15] and machine learning [16]. More specifically, by constructing SPD matrices, the original information extracted from the sample data is embedded on a specific SPD manifold, which is shown to outperform the Euclidean space operation. However, the drawback of using these matrix information geometry based methods is the rapidly increasing computational complexity, especially for the high-dimensional SPD manifold [17].

In this case, dimensionality reduction (DR) technique in machine learning is always imperative to reduce the redundancy and improve the discrimination in solving these high computational problems. In fact, simply using conventional Euclidean dimensionality reduction (DR) methods, e.g., Principal Component Analysis (PCA) [18], Linear Discriminant Analysis (LDA) [19], may destroy the implicit structure of these manifold-valued data such that they are unreasonable for our task. Recently, considering the special high-dimensional structure, DR work has been extended to Riemannian manifold. A generalization of PCA to Riemannian manifold named as PGA [20], which tends to flatten the Riemannian manifold via tangent space mapping, however, does not fully capture the structure of Riemannian manifold, which makes it might be suboptimal for classification. Another popular trend of considering the Reproducing Kernel Hilbert Space (RKHS) mapping by using kernel tricks [21], however, has a huge computational complexity for high-dimensional data, which limits its efficiency. To this end, the information geometric based DR technique dedicated to SPD manifold is proposed [22], by solving a manifold optimization to search a projection matrix, it maps the high-dimension SPD matrices into a lower-dimensional and more discriminative SPD manifold, which has proven its strong power in SPD manifold learning.

The contributions in this paper are following:

According to [23], the high-dimensional feature can provide richer and more discriminative information. In this case, a high-dimensional feature descriptor that contains both local and global information called the dense short time Fourier transform (SIFT) descriptor, is employed to obtain feature vectors of the 2-D spectrograms, which makes better use of information than inadequate edge features previously studied [9,11]. Furthermore, combining with the emerging theory of information geometry, our proposed detector has outperformed the existing spectrogram processing technique based signal detection methods by using information geometric DR method on SPD manifold illustrated in Figure 1. This study presents important information for future development of matrix information geometry-based signal detection methods, which can be widely used for signal detection tasks.

The remainder of this paper is organized as follows: Section 2 introduces some useful backgrounds regarding the SPD Riemannian manifold and the construction of desired covariance matrix for SPD manifold modeling. Section 3 shows how the novel information geometric DR technique be exploited to our signal detection task. Our experimental results with simulation and real-life sea clutter data are presented in Section 4 and our ongoing works are discussed in Section 5.

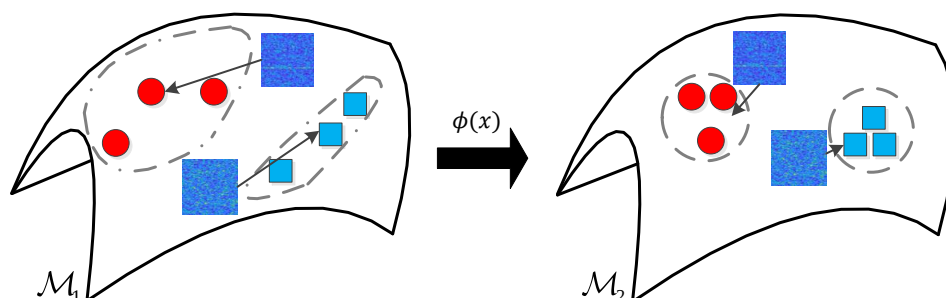


Figure 1. Geometrical interpretation of signal detection method based on information geometric dimensionality reduction. Function $\phi(x)$ means a non-linear mapping from high-dimensional symmetric positive definite (SPD) manifold $\mathcal{M}_1$ to lower-dimensional SPD manifold $\mathcal{M}_2$.

2. Constructing Region Covariance Matrix on SPD Manifold

In this section, we first provide related knowledge regarding the SPD Riemannian manifold. Then we show the construction of desired SPD matrix transformed by high-dimensional vector using the 128-dimensional dense SIFT descriptor for feature extraction. So far, the high-dimensional SPD manifold consists of dense SIFT features based SPD matrices is modeled. In other words, the labeled two heterogeneous samples are embed into the constructed SPD manifold.

2.1. The Symmetric Positive Definite (SPD) Manifold

In fact, a Riemannian manifold $\mathcal{M}_H$ with intrinsic geometry structure shown in Figure 2 is a differentiable topological space in which the tangent space T_p of each point p on the manifold is defined by a smoothly varying inner product. As is well acknowledged, the SPD manifold $\mathcal{S}_{++}^n$ viewed as the space of $n \times n$ SPD matrices, forms the interior of a convex cone in the $n(n+1)/2$ dimensional Euclidean space. Unlike the linear Euclidean space, most properties and vector operations are not suitable for SPD manifold due to its nonlinearity. In this case, to encode a valid SPD manifold $\mathcal{M}_H$, appropriate Riemannian metric should be applied so that the similarity between two SPD matrices can be analyzed.

Among these Riemannian distances designed for SPD manifold, two most popular metrics are Affine-invariant Riemannian metric (AIRM) [24] and Log-Euclidean metric [25]. More specifically, with the property of affine-invariance, the AIRM metric, which induces the true geodesic distance has been most widely used. However, one of the drawbacks is that the AIRM metric accumulates a high computational burden due to the curvature of such high-dimensional manifold space. Thus, AIRM-based algorithms perform less efficiently than other Riemannian metrics in some engineering cases [26].

On the other hand, induced from Lie group, the Log-Euclidean metric embeds the SPD manifold via matrix logarithm mapping into its tangent space T_p so that linear operations can be directly performed. In addition, this Riemannian metric between $\mathbf{X}$ and $\mathbf{Y}$ enjoys a variety of useful properties [27] that have a simple form of distance measure and low computational complexity, which can be defined as:

$$d_{LE}(\mathbf{X}, \mathbf{Y}) = \|\log \mathbf{X} - \log \mathbf{Y}\|_F \quad (1)$$

where $\log(\cdot)$ denotes the matrix logarithm operation, and $\|\cdot\|_F$ means the Frobenius norm.

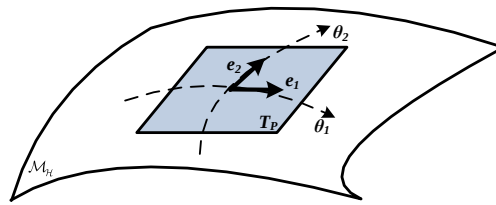


Figure 2. The Riemannian manifold and its tangent space T_p

2.2. The Region Covariance Matrix Based on Dense SIFT Descriptor

The region covariance matrix (RCM) is known as a special kind of SPD matrix with strong robustness and discrimination, which can implicitly capture the second-order statistical characteristics of the sample data. Furthermore, the size of the RCM only depends on the characteristic dimension, hence, regions with different sizes can be directly compared without any transformations. As a consequence of these useful properties, the RCM has achieved remarkable improvement in texture classification [28], pedestrian detection [29], 3D human motion sequence [30], target tracking [31] and many other aspects.

Given a $W \times H$ image I with n image pixels. The RCM can be given by:

$$C_R = \frac{1}{n-1} \sum_{k=1}^n (z_k - \mu)(z_k - \mu)^T \quad (2)$$

where z_k is the characteristic value corresponding to the k -th pixel, and $\mu = \frac{1}{n} \sum_{i=1}^n z_k$ is the average value of the corresponding feature.

In order to fully exploit the information implied in the 2-D spectrograms, the 128-dimensional dense SIFT descriptor with scale invariance is employed for each grid point, which has been shown as a powerful tool in image recognition [32]. The difference between the dense SIFT descriptor and classical SIFT descriptor is that the former is used to extract global image features, while the latter only obtains SIFT features of several key points. Through Gaussian smoothing, the dense SIFT descriptor is obtained by sliding a specific window to record the gradient in 8 directions of each grid point, forming a $4 \times 4 \times 8$ -dimensional feature vector shown in Figure 3.

Clearly, using such a global high-dimensional descriptor has its pros and cons: the dense SIFT descriptor can extract abundant information from the sample data but brings unnecessary redundancy, which motivates us to consider effective dimensionality reduction methods.

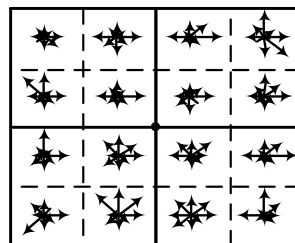


Figure 3. The 128-dimensional dense short time Fourier transform (SIFT) descriptor of one grid point.

3. Matrix Information Geometric Dimensionality Reduction Technique

In order to reduce the redundancy and enhance the discrimination, a manifold optimization based information geometric DR technique is employed. The key idea is to find a projection matrix that maps the original SPD manifold into a low-dimensional one by performing a manifold optimization without changing the intrinsic structure of “the manifold-valued data”. Formally, for a set of SPD matrices $X \in S_{++}^n$, our goal is to seek for a mapping $f : S_{++}^n \times \mathbb{R}^{n \times m} \rightarrow S_{++}^m$ with the learned projection matrix $W \in \mathbb{R}^{n \times m}$, $m < n$. Hence, with the elimination of redundant information by using this technique, “the manifold-valued data” becomes more discriminative, which is shown to overcome the limitations of Euclidean dimensionality reduction methods.

3.1. Affinity Graph Embedding

In feature extraction, the training data are generally described by in terms of measurable features that tend to be high-dimensional. In order to reduce the cost of feature measurement, and improve the learning performance, dimensionality reduction that drives new lower-dimensional features from the original features is always employed. However, new feature descriptions invariably cause the discard of the original information, and thus, may give poor overall accuracy. To this end, an affinity function $A(i, j)$ is added in order to force the samples sharing the same label more concentrated, that is, pulling the homogeneous samples closer and pushing the heterogeneous samples away. More specifically, let $\{(X_i, Y_i)\}_{i=1}^p$ denotes p labeled spectrograms, here $X_i \in S_{++}^n, Y_i \in \{0, 1\}$ represents the label of the two classes. In this case, the affinity function $A(i, j)$ can be established by two nearest neighbor graphs, namely the within-class similarity graph $G_w(i, j)$ and the between-class similarity graph $G_b(i, j)$, which can be defined by:

$$G_w(i, j) = \begin{cases} 1, & \text{if } Y_i = K_w(Y_j) \text{ or } Y_j = K_w(Y_i) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$G_b(i, j) = \begin{cases} 1, & \text{if } Y_i \neq K_b(Y_j) \text{ or } Y_j \neq K_b(Y_i) \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Here, parameter $K_w(X_i)$ and $K_b(X_i)$ represent the number of the neighbor of X_i . Hence, the affinity function can be written as:

$$A = G_w - G_b \quad (5)$$

which is a binary matrix consisting of 0 and 1. For the balance of $G_w(i, j)$ and $G_b(i, j)$, we usually take $k_b < k_w$.

3.2. Cost Function

In particular, to ensure that the new constructed SPD manifold is valid, that is, $W^T X W \succ 0$, $\forall X \in S_{++}^n$, an orthogonal constraint $W^T W = \mathbf{I}_m$ is included because the projection matrix $W \in \mathbb{R}^{n \times m}$ ($m < n$) is required to be full rank. After learning the embedding affinity function $A(i, j)$, we confront with the following optimization problem [22]:

$$\begin{aligned} \min \sum_{i,j} A(i, j) \delta_{TE}^2(W^T X_i W, W^T X_j W) \\ \text{s.t. } W^T W = \mathbf{I}_m \end{aligned} \quad (6)$$

By substituting Formula (1):

$$\begin{aligned} \min_{W \in \mathbb{R}^{n \times m}} \sum_{i,j=1}^p a(X_i, X_j) \left\| \log(W^T X_i W) - \log(W^T X_j W) \right\|_F^2 \\ \text{s.t. } W^T W = \mathbf{I}_m. \end{aligned} \quad (7)$$

With Talyor expansion that $\log(W^T X W)$ can be approximated as $W^T \log(X) W$, then the optimization problem (7) can be rewritten as:

$$\begin{aligned} \min_{W \in \mathbb{R}^{n \times m}} \sum_{i,j=1}^p a(X_i, X_j) & \left\| W^T \log(X_i) W - W^T \log(X_j) W \right\|_F^2 \\ \text{s.t. } W^T W &= \mathbf{I}_m. \end{aligned} \quad (8)$$

For simplicity, let

$$\begin{aligned} F(W) &= \sum_{i,j=1}^p a(X_i, X_j) (\log(X_i) - \log(X_j)) W W^T \\ &\quad \times (\log(X_i) - \log(X_j)) \end{aligned} \quad (9)$$

Then, the optimization problem for finding the projection matrix W is given as:

$$\begin{aligned} \min_{W \in \mathbb{R}^{n \times m}} \text{Tr}(W^T F(W) W) \\ \text{s.t. } W^T W &= \mathbf{I}_m \end{aligned} \quad (10)$$

In fact, function $\text{Tr}(W^T F(W) W)$ is independent from the choice of basis spanned by projection matrix W , that is, for an arbitrary orthogonal group $R \in \text{O}(m)$, $\text{Tr}(W^T F(W) W) = \text{Tr}((WR)^T F(W) (WR))$, in this case, the solution for Formula (10) known as a Grassmann problem can be optimized by Riemannian Conjugate Gradient (RCG) [33]. However, the RCG method on the Grassmannian manifold converges much more slowly than the iterative two-stage method using eigen-decomposition proposed in [22]. Thus, we take advantage of this eigen-decomposition-based approach for solving Formula (10), which can be summarized as follows. First, $F(W)$ is fixed by assuming that $F(W)$ does not depend on the projection matrix W . With the property that matrix trace is equal to the sum of the eigenvalues, the current solution can be obtained by taking the m smallest eigenvectors of $F(W)$. Then we have a new W to update the corresponding $F(W)$. These steps are repeated until convergence for the optimal solution of W , which achieves the mapping from a high-dimensional SPD manifold to a low-dimensional one with more discrimination.

4. Experimental Results

To strengthen the algorithm validation, our experiment revolves around simulation and semi-simulation data for signal detection comparing with the state of art techniques. In the simulation data case, we use K distribution to simulate the clutter, and a target signal with Doppler frequency $f_d = 0.15$ Hz. Additionally, we use the IPIX radar data of sea clutter with 27 cells collected by the McMaster University in 1998 [34] for semi-simulation data, likewise, a simulation target with Doppler frequency $f_d = 0.15$ Hz is added.

In particular, based on the K simulation data, we have in total two labelled spectrogram categories, namely, (a) negative samples in simulation set (no signal exists), (b) positive samples in simulation set (signal exists), which can be clearly seen in Figure 4.

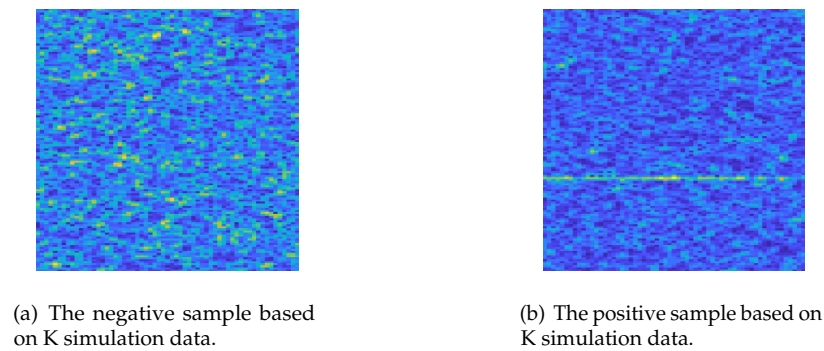


Figure 4. Generated spectrograms on K simulation data with two categories in case of SNR = −10. (a) K simulation data without signal labelled as the negative sample. (b) K simulation data containing signal labelled as the positive sample.

After the short time Fourier transform (STFT), we resized the generated 2-D spectrograms to 200×200 , using the dense SIFT descriptor to extract feature vector on a regular grid with 2 pixel spacing. We note that the 128-dimensional covariance matrices were constructed by using the method proposed in [35]. In this way, these spectrograms were mapped into an originally 128-dimensional SPD manifold. The overall experimental procedure shown in Figure 5 can be summarized as follows: processed by STFT, the sample data generated a high-precision 2-D spectrogram. Then feature vectors were obtained using SIFT descriptor in feature extraction, which was constructed into SPD matrices for SPD manifold modeling. Finally, a manifold optimization based information geometric DR technique was employed by solving Formula (10) to improve the discrimination, and thus enhance the detection performance.

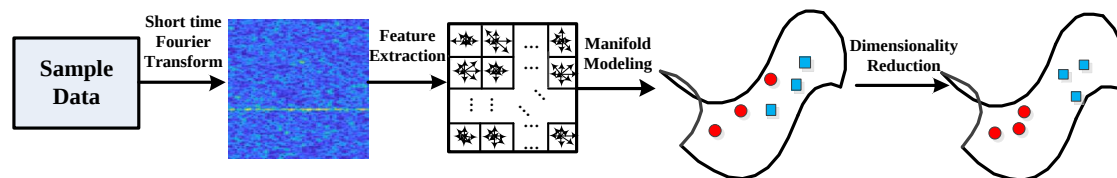


Figure 5. The flowchart of our experiment which can be mainly divided into 4 steps.

4.1. Simulation Data Experiment

According to our former work in ICSP 2020 [36], we have proved that our proposed approach leads to significant improvement over the original SPD manifold in the case of SNR from −5 to −10, which has shown that the manifold-valued data after information geometric DR become more discriminative, however, we have not explicitly given a practical detector for real-time detection. Furthermore, the signal detection ability of our proposed algorithm in the larger range of SNR is not demonstrated.

To this end, in this paper, we trained one general matrix W to form a detector for multiple SNRs by taking all labeled samples into the training set at once rather than training objective W for each SNR. More specifically, we had 200 signal sequences per signal-to-noise ratio (SNR −1 to −15) containing the desired signal labelled as the positive samples, and 3000 noise-only samples as the negative samples. To evaluate the detection performance, 4000 signal sequence samples per SNR (−1 to −20) were collected as the test set (2000 positive samples, and 2000 negative samples, respectively), using KNN classifier for efficiency on the low-dimensional and discriminative SPD manifold. Taking false alarm probability P_{fa} and detection probability P_d as the two most crucial factors in signal detection as criterion, our experimental results with the K simulation data were the following:

In our first experiment, parameters K_w and K_b in affinity graph embedding were 40 and 10. The parameter K in KNN classifier was fixed to 4200, whose effect on detection performance will be discussed in Section 4.1.2. We first give the detection probability curves of scheme ① the Radon Transform detector [37] that were widely used for linear signal detection from spectrogram processing, and scheme ② our proposed detector with the objective dimension dim at 40 shown in Figure 6.

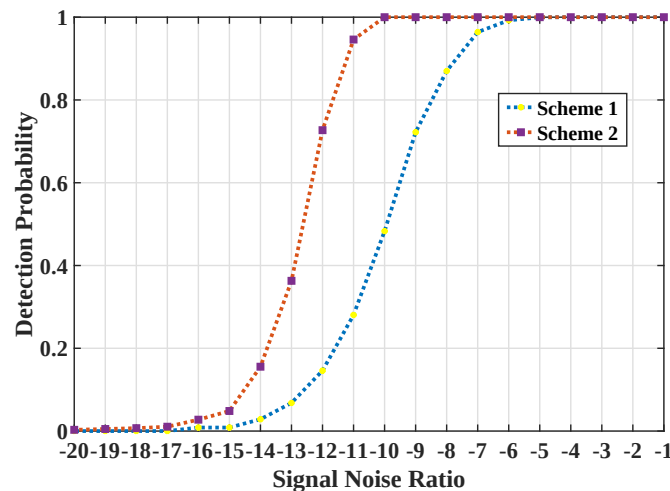


Figure 6. The detection probability curves of the two detectors (scheme ① the Radon Transform detector and scheme ② our proposed detector at $dim = 40$ with $P_{fa} = 0.001$) for varying the signal-noise-ratio based on K simulation.

It can be seen that the proposed detector outperformed the Radon Transform detector with almost 5 SNR performance improvement. Specifically, our proposed detector led to a significant signal detection capability with a low false alarm rate for the cases with SNR over -10 . Since the intra-class matrices constructed under lower SNR cases had little discrimination on the SPD manifold, the detection probability dropped significantly at SNR of interval $[-10, -15]$.

4.1.1. Detection Performance for Various Objective Dimension

In fact, it is obvious that the discrimination of the manifold-valued data was different on various dimensional SPD manifolds, thus the binary classification performance was surely affected. To this end, we compared the detection performance for various objective dimension dim (from 20 to 120 with an interval of 20) on the simulation data, which is shown in Table 1.

Table 1. Comparison of detection performance for varying the value of dim based on K simulation.

	$dim = 20$ $P_{fa} = 0.004$	$dim = 40$ $P_{fa} = 0.001$	$dim = 60$ $P_{fa} = 0.0035$	$dim = 80$ $P_{fa} = 0.0085$	$dim = 100$ $P_{fa} = 0.0205$	$dim = 120$ $P_{fa} = 0.02$
SNR = -1	1	1	1	1	1	1
SNR = -5	1	1	1	1	1	1
SNR = -10	1	1	1	0.9975	1	1
SNR = -15	0.0925	0.0485	0.1035	0.0085	0.1975	0.1895
SNR = -20	0.0105	0.003	0.011	0.0085	0.0375	0.0355

The results are shown for SNR of -1 , -5 , -10 , -15 , and -20 . We note that since the negative samples on the SPD manifold were naturally lying close, there was little difference in false alarm frequency so that the probability P_{fa} for each SNR was almost a constant.

In general, our proposed detector could achieve full signal detection over SNR = -10 in any objective dimension on the simulation dataset. It can be seen that P_{fa} had a proportional correlation with P_d , and P_d decreased rapidly in the SNR range of -10 to -15 because signals at this level became extremely insignificant on the spectrogram. Furthermore, the minimum P_{fa} occurred at $dim = 40$ while

the maximum P_d was achieved for $dim = 100$ but caused a higher false alarm in lower SNR cases.

4.1.2. Parameter Sensitivity Experiment on K

To further control false alarm probability P_{fa} , we analyze the sensitivity on parameter K in KNN classifier. We selected the case of $SNR = -12$ with the objective dimension $dim = 40$ and fixed the parameter K_w to 40, K_b to 10. Table 2 depicts the detection performance for various values of K in the interval [100, 5100].

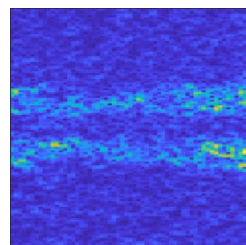
It can be seen that both the value of P_{fa} and P_d decreased with the increase of K , which is considered to be because fewer positive samples for each SNR were trained, while the negative samples were relatively large. However, simply increasing the positive samples under each SNR caused more false alarms at low SNR cases because of the small discrimination with the two heterogeneous samples, which also shows that our proposed detector could achieve a satisfactory P_d with a very low P_{fa} for SNR over -11 on the simulation dataset.

Table 2. Comparison of detection performance at the case of $SNR = -12$ with the objective dimension $dim = 40$ for varying the value of K based on K simulation.

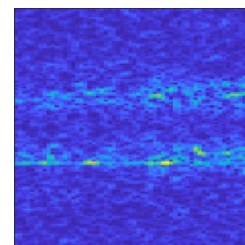
K	100	1100	2100	3100	4100	5100
P_{fa}	0.016	0.012	0.0085	0.0055	0.0015	0
P_d	0.9885	0.8875	0.8755	0.8475	0.752	0.4465

4.2. Semi-Simulation Data Experiment

Simulation data results have shown that our proposed approach was effective, to further evaluate our algorithm, experiments based on semi-simulation data were also conducted. Specifically, in training phase for optimizing the projection matrix W , by using IPIX radar data of real-life sea clutter, we had 600 signal sequences per signal-to-noise ratio (SNR 5 to 1) containing the desired signal labelled as the positive samples, and 3000 noise-only samples as the negative samples shown in Figure 7.



(a) The negative sample based on real sea clutter.



(b) The positive sample based on real sea clutter.

Figure 7. Generated spectrograms on real sea clutter data with two categories in case of $SNR = -10$. (a) Real sea clutter labelled as the negative sample. (b) Semi-simulation data containing signal labelled as the positive sample.

Likewise, to evaluate the detection performance, 4000 signal sequence samples per SNR (5 to -10) were collected as the test set (2000 positive samples, and 2000 negative samples, respectively), using KNN classifier on the low-dimensional and discriminative SPD manifold. For the semi-simulation experiment, parameters K_w and K_b in affinity graph embedding are set to 20 and 4. The parameter K in KNN classifier was fixed to 1, known as the nearest neighbor classifier.

In particular, the Radon transform detector misjudged the edge of the clutter as a hidden signal due to the complexity and heterogeneity in our semi-simulation experiment, which showed its drawback in heterogeneous clutter signal detection tasks, thus it was not acceptable for practical application. In this case, we referred to another algorithm as comparison evaluated in our experiments, namely,

scheme (a) a mathematical morphology signal detector [38] based on Sobel edge detection, and scheme (b) our proposed information geometric DR based detector.

4.2.1. Detection Performance for Various Objective Dimension

Similarly, as for the real sea clutter data, the impact of the objective dimension dim (from 20 to 120 with an interval of 20) on the detection performance was analyzed shown in Table 3, from which we can clearly see that the detection performance of the proposed detector varied with different dimensions. Generally, the geometric DR detector with lower dimension yielded better detection performance, that is, lower P_{fa} and high P_d within a certain SNR range, which indicated that our proposed detector could effectively reduce the redundancy brought by high dimensionality, and thus, improve the detection performance because of its stronger discrimination.

To verify this conclusion, we give the detection probability curves of scheme (a) and scheme (b) with the objective dimension dim at 10 shown in Figure 8. It can be seen that the proposed detector outperformed the Sobel edge detector with Sobel threshold 280. Specifically, scheme (b) gained very limited detection probability P_d when considering false alarm $P_a = 0.1$. In particular, our proposed detector led to a significant signal detection probability P_d with a low false alarm rate P_{fa} for most SNR cases. Since the intra-class SPD matrices under lower SNR cases yielded less discrimination on the manifold, the detection probability P_d dropped significantly at SNR of interval $[0, -10]$.

Table 3. Comparison of detection performance for varying the value of dim based on real sea clutter.

	dim = 20 $P_{fa} = 0.0015$	dim = 40 $P_{fa} = 0.002$	dim = 60 $P_{fa} = 0.0045$	dim = 80 $P_{fa} = 0.0035$	dim = 100 $P_{fa} = 0.003$	dim = 120 $P_{fa} = 0.0045$
SNR = 5	0.989	0.9785	0.9715	0.9645	0.9635	0.9575
SNR = 0	0.9795	0.97	0.9475	0.9415	0.9345	0.921
SNR = −5	0.561	0.575	0.5775	0.526	0.528	0.51
SNR = −10	0.073	0.0765	0.099	0.096	0.102	0.1075

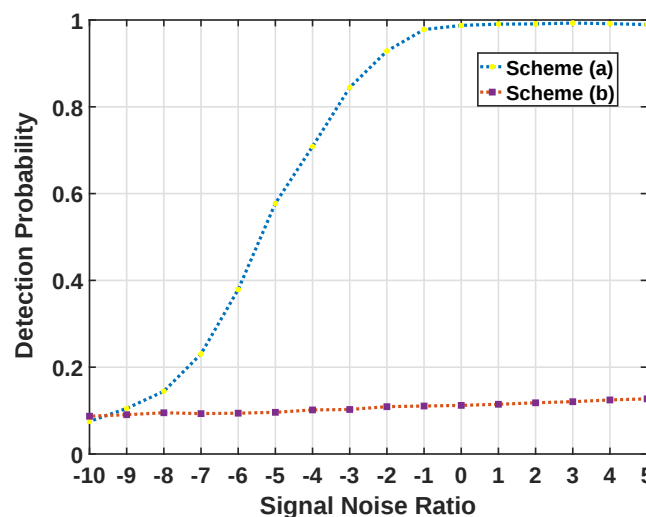


Figure 8. The detection probability curves of the two detectors (scheme (a) the Sobel edge detector and scheme (b) our proposed detector at $dim = 10$ with $P_{fa} = 5 \times 10^{-4}$) for varying the signal-noise-ratio based on the real sea clutter.

4.2.2. Parameter Sensitivity Experiment on K

In addition, the impact of parameter K on detection performance was analyzed by using our proposed algorithm for the semi-simulation data, which can be seen in Table 4. We selected the case of SNR = −5 with the objective dimension $dim = 20$ and fixed the parameter K_w to 20, K_b to 4. Table 2 depicts the detection performance for various K in the interval $[1, 5100]$. It can be clearly seen that

both the false alarm rate P_{fa} and the detection probability P_d increased with K , which indicated the same variation trend with the simulation experiment. In particular, to obtain a satisfactory detection performance, a reasonable value of K should be taken. For instance, in this part, the false alarm rate P_{fa} could be well controlled by setting a small K , that is, the nearest neighbor classifier.

Table 4. Comparison of detection performance at the case of SNR = −5 with the objective dimension $dim = 20$ for varying the value of K based on real sea clutter .

K	1	100	1100	2100	3100	4100	5100
P_{fa}	0.015	0.0185	0.086	0.1715	0.253	0.3075	0.4265
P_d	0.561	0.5985	0.923	0.99	1	1	1

5. Discussion

Our work focuses on the task of complex signal detection, establishing a novel signal detection framework based on matrix information geometric dimensionality reduction. The results demonstrate that our proposed detector achieve a satisfactory detection performance in complex signal detection tasks and hence can be widely used in field of signal detection. More specifically, our proposed detector has outperformed the existing signal detection techniques on both the simulation and semi-simulation data with satisfactory detection performance, although the improvement is limited to a certain extent, especially for the real-life sea clutter data. Meanwhile, the constructed SPD matrices in two cases obviously have different distribution on SPD manifold, which indicates that parameters for various signal detection tasks are required to tune for optimum function.

In conclusion, the proposed detector still has some drawbacks required to overcome: The dimensionality reduction algorithm on such high-dimensional signal data constructed SPD manifold has a very high computational complexity, resulting in a large computational cost, which limits the efficiency of our method. In this case, how to accelerate these Riemannian geometry related algorithms is a subject that merits further study. Meanwhile, the form of the constructed SPD matrix has a great influence on detection performance, to this end, appropriate feature descriptors and better SPD matrix construction are areas we intend to study. Additionally, considering the high computing and storage cost on SPD matrices, we choose the KNN algorithm in our experiment, a simple classifier without training process rather than complex predict model-based classification methods, then obtain a satisfactory detection performance. However, the KNN classifier is sensitive to outliers and closely related to the distribution of training samples, which denotes that there might be potential improvement on detection performance. Thus, for future research, we may extend our approach by using other effective classification methods on these spectral-based SPD matrices.

Author Contributions: conceptualization, S.F., X.H. and X.Z.; formal analysis, S.F.; funding acquisition, X.H.; investigation, X.Z.; software, S.F.; validation, X.Z.; visualization, S.F.; writing—original draft, S.F.; writing—review and editing, X.H. All authors have read and agreed the published version of the manuscript.

Funding: This work research was funded by the National Natural Science Foundation of China under Grant No. 61901479 and by the Basic Research Cultivation Program for Young Excellent Talents in Meteorology and Oceanography of National University of Defense Technology of China under Grant 4345161111p.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Poor, H.V. *An Introduction to Signal Detection and Estimation*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
2. Seu, R.; Phillips, R.J.; Biccari, D.; Orosei, R.; Masdea, A.; Picardi, G.; Safaeinili, A.; Campbell, B.A.; Plaut, J.J.; Marinangeli, L. SHARAD sounding radar on the Mars Reconnaissance Orbiter. *J. Geophys. Res. Planets* **2007**, *112*, E05S05.

3. Blum, R.S. Necessary conditions for optimum distributed sensor detectors under the Neyman-Pearson criterion. *IEEE Trans. Inform. Theory* **1996**, *42*, 990–994.
4. Kassam, S.A. *Signal Detection in Non-Gaussian Noise*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
5. Schinkel, S.; Dimigen, O.; Marwan, N. Selection of recurrence threshold for signal detection. *Eur. Phys. J. Spec. Top.* **2008**, *164*, 45–53.
6. Thayaparan, T.; Stankovic, L.; Amin, M.; Chen, V.; Cohen, L.; Boashash, B. Time-frequency approach to radar detection, imaging, and classification. *IET Signal Process.* **2010**, *4*, 325–328.
7. Boashash, B. *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*; Academic Press: Cambridge, MA, USA, 2015.
8. Portnoff, M. Time-frequency representation of digital signals and systems based on short-time Fourier analysis. *IEEE Trans. Acoust. Speech Signal Process.* **1980**, *28*, 55–69.
9. Luo, S.e.; Luo, L.y. Adaptive detection of an unknown FH signal based on image features. In Proceedings of the 2009 5th International Conference on Wireless Communications, Networking and Mobile Computing, Beijing, China, 24–26 September 2009; pp. 1–4.
10. Lampert, T.A.; O’Keefe, S.E. A survey of spectrogram track detection algorithms. *Appl. Acoust.* **2010**, *71*, 87–100.
11. Gillespie, D. Detection and classification of right whale calls using an ‘edge’ detector operating on a smoothed spectrogram. *Can. Acoust.* **2004**, *32*, 39–47.
12. Yin, M.; Guo, Y.; Gao, J.; He, Z.; Xie, S. Kernel sparse subspace clustering on symmetric positive definite manifolds. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 5157–5164.
13. Huang, Z.; Wang, R.; Shan, S.; Li, X.; Chen, X. Log-Euclidean metric learning on symmetric positive definite manifold with application to image set classification. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 720–729.
14. Hua, X.; Fan, H.; Cheng, Y.; Wang, H.; Qin, Y. Information Geometry for Radar Target Detection with Total Jensen–Bregman Divergence. *Entropy* **2018**, *20*, 256.
15. Hua, X.; Cheng, Y.; Wang, H.; Qin, Y.; Li, Y.; Zhang, W. Matrix CFAR detectors based on symmetrized Kullback–Leibler and total Kullback–Leibler divergences. *Dig. Signal Process.* **2017**, *69*, 106–116.
16. Huang, Z.; Van Gool, L. A riemannian network for spd matrix learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
17. Feng, S.; Hua, X.; Wang, Y.; Lan, Q.; Zhu, X. Matrix Information Geometry for Signal Detection via Hybrid MPI/OpenMP. *Entropy* **2019**, *21*, 1184.
18. Abdi, H.; Williams, L.J. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 433–459.
19. Balakrishnama, S.; Ganapathiraju, A. Linear discriminant analysis—a brief tutorial. *Inst. Signal Inf. Process.* **1998**, *18*, 1–8.
20. Fletcher, P.T.; Lu, C.; Pizer, S.M.; Joshi, S. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imaging* **2004**, *23*, 995–1005.
21. Jayasumana, S.; Hartley, R.; Salzmann, M.; Li, H.; Harandi, M. Kernel Methods on Riemannian Manifolds with Gaussian RBF Kernels. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 2464–2477.
22. Harandi, M.; Salzmann, M.; Hartley, R. Dimensionality Reduction on SPD Manifolds: The Emergence of Geometry-Aware Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 48–62. doi:10.1109/TPAMI.2017.2655048.
23. Wang, Q.; Li, P.; Zuo, W.; Zhang, L. RAID-G: Robust Estimation of Approximate Infinite Dimensional Gaussian With Application to Material Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
24. Pennec, X.; Fillard, P.; Ayache, N. A Riemannian Framework for Tensor Computing. *Int. J. Comput. Vis.* **2006**, *66*, 41–66. doi:10.1007/s11263-005-3222-z.
25. Gu, Q.; Zhou, J. A Similarity Measure under Log-Euclidean Metric for Stereo Matching. In Proceedings of the 19th International Conference on Pattern Recognition (ICPR 2008), Tampa, FL, USA, 8–11 December 2008.
26. Arsigny, V.; Fillard, P.; Pennec, X.; Ayache, N. Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices. *SIAM J. Matrix Anal. Appl.* **2006**, *29*, 328–347.

27. Arsigny, V.; Fillard, P.; Pennec, X.; Ayache, N. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **2007**, *29*, 328–347.
28. Tou, J.Y.; Tay, Y.H.; Lau, P.Y. Gabor filters as feature images for covariance matrix on texture classification problem. In *Lecture Notes in Computer Science, Proceedings of the International Conference on Neural Information Processing, Auckland, New Zealand, 25–28 November 2008*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 745–751.
29. Zhang, Y.; Li, S. Gabor-LBP based region covariance descriptor for person re-identification. In *Proceedings of the 2011 Sixth International Conference on Image and Graphics, Hefei, China, 12–15 August 2011*; pp. 368–371.
30. Hussein, M.E.; Torki, M.; Gawayyed, M.A.; El-Saban, M. Human action recognition using a temporal hierarchy of covariance descriptors on 3d joint locations. *In Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, Beijing, China, 3–9 August 2013*.
31. Porikli, F.; Tuzel, O.; Meer, P. Covariance tracking using model update based on lie algebra. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 1*, pp. 728–735.
32. Mortensen, E.N.; Deng, H.; Shapiro, L. A SIFT descriptor with global context. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 1*, pp. 184–190.
33. Huang, Z.; Wang, R.; Shan, S.; Chen, X. Projection Metric Learning on Grassmann Manifold with Application to Video based Face Recognition. *In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015*; pp. 140–149.
34. Bakker, R.; Currie, B.; Haykin, S. The McMaster IPIX radar sea clutter database, 1998. *Available online:* <http://soma.mcmaster.ca/ipix/> (access on: 15 February 2020)
35. Tuzel, O.; Porikli, F.; Meer, P. Region covariance: A fast descriptor for detection and classification. In *Lecture Notes in Computer Science, Proceedings of the European Conference on Computer Vision, Graz, Austria, 7–13 May 2006*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 589–600.
36. Feng, S.; Hua, X.; Wang, J.; Zhu, X. A signal detection method based on matrix information geometric dimensionality reduction. *J. Phy. Conf. Ser.* **2020**, *1544*, 012129. doi:10.1088/1742-6596/1544/1/012129.
37. Wu, J.X.; Kucheryavskiy, S.V.; Jensen, L.G.; Rades, T.; Müllertz, A.; Rantanen, J. Image Analytical Approach for Needle-Shaped Crystal Counting and Length Estimation. *Cryst. Growth Des.* **2015**, *15*, 4876–4885, doi:10.1021/acs.cgd.5b00720.
38. Witschi, M.; Schild, J.; Nyffenegger, B.; Stoller, C.; Dellsperger, F. Detection of Modern Communication Signals Using Frequency Domain Morphological Filtering. In *Proceedings of the 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 29 August–2 September 2016*.

