

海杂波中基于可控虚警K近邻的海面小目标检测

郭子薰^{*①} 水鹏朗^{*①} 白晓惠^① 许述文^① 李东宸^②

^①(西安电子科技大学雷达信号处理国家重点实验室 西安 710071)

^②(中国船舶工业系统工程研究院 北京 100094)

摘要: 由于高分辨海杂波具有复杂的特性以及海面小目标具有多样性, 没有精确的简单统计模型可以较好地描述海杂波和目标回波时间序列, 这导致目标检测遇到了很多阻碍。为了区分海杂波和目标回波, 分别提取它们的特征将检测问题转化为特征空间中的分类问题是一种有效的方法。基于特征的检测可以归结为在特征空间中的一种2元假设检验问题, 但是其有两个问题需要解决: 一是目标回波数据远少于杂波数据; 二是虚警率不可控。为了解决第1个问题, 一种典型小目标的仿真回波产生器被用于产生充足的典型目标回波数据, 以辅佐后续检测器的设计。K近邻(K-NN)是一种简单有效的分类方法, 但是因为无法精确地控制虚警率而不能直接在目标检测中使用。该文提出一种基于改进K-NN的海面小目标检测方法, 可以很好地实现可控虚警。经IPIX雷达数据集验证, 所提出的方法在观测时间分别为0.512 s和1.024 s时获得了85.1%和89.2%的检测概率, 相比现有的检测器获得了7%和5%的提升, 具有良好的检测效果和稳定性。

关键词: 高分辨海杂波; 海面小目标; 特征检测; K近邻; 可控虚警

中图分类号: TN957.51

文献标识码: A

文章编号: 2095-283X(2020)04-0654-10

DOI: 10.12000/JR20055

引用格式: 郭子薰, 水鹏朗, 白晓惠, 等. 海杂波中基于可控虚警K近邻的海面小目标检测[J]. 雷达学报, 2020, 9(4): 654-663. doi: 10.12000/JR20055.

Reference format: GUO Zixun, SHUI Penglang, BAI Xiaohui, *et al.* Sea-surface small target detection based on K-NN with controlled false alarm rate in sea clutter[J]. *Journal of Radars*, 2020, 9(4): 654-663. doi: 10.12000/JR20055.

Sea-surface Small Target Detection Based on K-NN with Controlled False Alarm Rate in Sea Clutter

GUO Zixun^{*①} SHUI Penglang^{*①} BAI Xiaohui^① XU Shuwen^① LI Dongchen^②

^①(National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

^②(Systems Engineering Research Institute, Beijing 100094, China)

Abstract: Owing to the complicated characteristics of high-resolution sea clutter and the diversity of sea-surface small targets, there is no precise parameter model to describe sea clutter and returns with targets. As a result, target detection faces many obstacles. To distinguish sea clutter and target returns, it is effective to extract their features to transform the detection problem into a classification problem in feature space. Feature-based detection is a binary hypothesis test in the feature space that encounters two intrinsic difficulties: one difficulty is insufficient target returns versus sufficient sea clutter; the other difficulty is an uncontrolled false alarm rate in detection. To solve the first difficulty, a generator of typical targets returns that can generate sufficient simulated targets returns is used to balance the number of samples between two classes and assist to design the detector. K Nearest Neighbors (K-NN) is the type of classification method that is simple and effective; however, it cannot be used to detect small targets directly because of the uncontrolled false alarm rate. This paper

收稿日期: 2020-05-08; 改回日期: 2020-07-06; 网络出版: 2020-07-29

*通信作者: 郭子薰 zxguo_724@stu.xidian.edu.cn; 水鹏朗 plshui@xidian.edu.cn

*Corresponding Author: GUO Zixun, zxguo_724@stu.xidian.edu.cn; SHUI Penglang, plshui@xidian.edu.cn

基金项目: 国家自然科学基金(61871303)

Foundation Item: The National Natural Science Foundation of China (61871303)

责任主编: 刘宁波 Corresponding Editor: LIU Ningbo

proposes a modified K-NN method with a controlled false alarm rate for detecting small targets. Experimental results on the IPIX radar database indicate that the proposed detector attains 85.1% and 89.2% rates of target detection for the observation time of 0.512 s and 1.024 s, respectively, compared with other existing feature-based detectors, the proposed detector exhibits 7% and 5% improvement, respectively. Thus, the proposed detector exhibits more stable and effective detection performance than other existing feature-based detectors.

Key words: High-resolution sea clutter; Sea-surface small targets; Feature-based detection; K Nearest Neighbors (K-NN); Controlled false alarm rate

1 引言

对于海面警戒雷达来说, 在小擦地角下提高海面小目标的检测能力是一项长期且困难的工作, 这类小目标主要指浮冰、小船、蛙人和飞机残骸等^[1]。大致原因如下: 第1, 小目标具有很小的雷达截面积(Radar Cross Section, RCS)和较弱的雷达回波; 第2, 海杂波具有复杂且变化的特性, 例如较宽的多普勒带宽和较强的非高斯特性等^[2]。为了检测小目标, 通常采取两种方式: 高分辨率和长观测时间, 其中前者可以降低杂波功率水平, 后者可以增加目标回波的累积增益。但是在传统雷达中, 会面临一个波位上长驻留时间和扫描效率上的矛盾。随着雷达系统的不断发展, 2003年麻省理工(Massachusetts Institute of Technology, MIT)的林肯实验室^[3,4]提出了泛探雷达(ubiquitous radars), 其同时利用多个接收波束覆盖整个观测空域实现了在所有方向下的全时观测。所以, 在长观测时间下提出一种有效的检测小目标的方法是有必要的。但是, 由于海杂波的长时非平稳特性和目标回波复杂的幅度和多普勒调制, 海杂波和目标回波都很难被建模为简单有效的参数模型^[5]。随着对海杂波和小目标回波的深入研究, 研究者提出了时频分析、分形特性和机器学习等方法^[6-8]来解决这一类问题。总的来说, 都是通过研究海杂波和目标回波在各个域上的不同特性来对二者进行区分。

在1993年, Haykin等人通过对海杂波时间序列的分形(fractal)特性进行研究, 进一步发现分形维数可以被用作特征以帮助实现目标检测。为了使用更多的特征来提高检测性能, Shui等人^[9]进一步提取了1个幅度特征和2个多普勒特征, 并将3个特征联合起来建立特征空间中的三维凸包以实现最终检测。通过IPIX雷达数据集^[10]的验证, 基于3特征的检测器与基于单特征的检测器相比, 实现了更优的检测性能。这可以归功于多个来自不同域互补的有效特征的联合使用。但是, 目标回波有时会落入多普勒域的主杂波带内, 这大大影响了检测结果。为了解决这一问题, 另一种基于时频3特征的检测器被Shui等人^[11]提出, 大大改善了检测器的性能。

但是, 通过对检测结果的研究发现, 两个基于3特征的检测器分别在不同的数据上表现优异。考虑到不同特征对不同数据的敏感性, 联合使用更多的互补特征来设计高维空间中的检测器就成为了进一步提高检测性能的有效途径。

在机器学习领域中, 有很多分类能力优异的算法, 比如支撑向量机(Support Vector Machine, SVM)、神经网络(Neural Network, NN)等。我们尝试将雷达目标检测问题与机器学习算法结合起来, 设计在高维特征空间中的检测器。在本文中, 为了获得更好和更具有鲁棒性的检测结果, 已有的来自不同域的7个有效特征被联合使用, 包括归一化Hurst指数(Normalized Hurst Exponent, NHE)^[12,13]、相对平均幅度(Relative Average Amplitude, RAA)、相对多普勒峰高(Relative Doppler Peak Height, RDPH)、相对向量熵(Relative Vector Entropy, RVE)^[9]、脊累积(Ridge Integration, RI)、连通区域数目(Number of connected Regions, NR)和最大连通区域尺寸(Maximal Size of connected regions, MS)^[11]。由于该多个特征使得杂波与目标都处于七维特征空间, 从而一些已有的方法在高维($n > 3$)空间中计算复杂度极高或者不再适用。例如, 由于已有的基于3特征的检测器^[9,11]使用凸包来决定最终的判决区域, 相继通过收缩凸包实现虚警控制。但是, 凸包只能在小于等于三维的空间中使用, 一旦到四维或更高维, 凸包的计算代价将会非常大, 甚至无法计算, 所以这就是其不能推广到高维的原因。后来, 本团队为了突破维数限制的问题, 提出了基于特征压缩的检测器^[14,15], 可以将原本在高维空间中的特征压缩至三维空间。与以往基于单特征或3特征的检测器相比, 基于特征压缩的检测器获得了更加优异的性能。但是在特征压缩的过程中, 压缩损失是不可避免的。本文不仅为了利用更多的已知特征使得对目标有更好的检测, 同时也为了避免压缩带来的性能损失, 使用了机器学习中的K近邻(K Nearest Neighbours, K-NN)算法。但是由于K-NN无法实现对虚警率的有效控制, 所以本文提出了一种基于可控虚警的改进K-NN方法, 有效地解决了问题。

本文的工作安排大致如下：第2部分先回顾所使用的特征，并介绍一种典型的仿真目标回波产生器；第3部分提出了基于可控虚警的改进K-NN检测方法；第4部分利用实测数据对所提出的检测器进行性能评估，并与其它基于特征的检测器进行对比和分析；最后，第5部分会对本文的工作进行总结评价。

2 仿真目标回波产生器及7特征概述

雷达目标检测问题可以归结为以下2元假设检验问题^[5,9,11,13-15]

$$\left. \begin{aligned} H_0 : & \begin{cases} z(n) = c(n), n = 1, 2, \dots, N \\ z_p(n) = c_p(n), p = 1, 2, \dots, P \end{cases} \\ H_1 : & \begin{cases} z(n) = s(n) + c(n), n = 1, 2, \dots, N \\ z_p(n) = c_p(n), p = 1, 2, \dots, P \end{cases} \end{aligned} \right\} \quad (1)$$

其中， $z(n)$ 和 $z_p(n)$ 分别表示待检测单元和周围参考单元所接收的复时间序列， $c(n)$ 和 $c_p(n)$ 分别表示待检测单元和周围参考单元的纯杂波时间序列， $s(n)$ 表示目标回波时间序列， P 为参考单元数目， N 为脉冲长度。在零假设 H_0 下，待检测单元所接收的时间序列为纯杂波，其具有与周围参考单元的杂波序列相同的统计特性和多普勒特性。在备择假设 H_1 下，待检测单元所接收的时间序列为带有目标回波时间序列。

2.1 7种有效特征概述

在检测器的设计中，7种有效特征被使用，包括两个幅度特征：NHE, RAA^[9,12,13]，两个多普勒特征：RDPH, RVE^[9]，3个时频特征：RI, NR, MS^[11]。对于两个幅度特征，若雷达回波数据包含目标，则其NHE和RAA的取值均大于纯杂波。此外，我们对回波数据的多普勒幅度谱进行分析并提取有效特征RDPH和RVE，当发现待检测单元包含目标时，RDPH较大，RVE较小；当待检测单元为纯杂波时，RDPH较小，RVE较大。但是，目标回波容易落在多普勒域的主杂波带内，使得前4种特征不再适用，那么3个时频特征就可以在这种情况下发挥作用帮助完成目标检测。当雷达回波包含目标时，RI和MS取值较大、NR取值较小；反之，当回波是纯杂波时，RI和MS取值较小、NR取值较大。

通过提取7个有效的海杂波和目标的特征，目标检测问题就可以近似地被转化为一个在七维特征空间的二分类问题。由于不同的特征对不同的数据或者海态都有不同的敏感性，那么基于单个特征检测器就不会对任何情况都适用和有效。对目标检测问题来说，联合使用7个有效特征来构建高维特征空间中的检测器无疑是一种综合有效的方法。

2.2 仿真目标回波产生器

由于海面小目标的多样性和目标与海表面之间复杂的相互运动，获取所有种类的目标的有效信息是不可能的。考虑到雷达的工作模式，大量的杂波数据可以被快速地收集，相比之下，只有少量的目标数据被获取，这就导致了两类样本数量的不均衡。在我们的工作中，使用已有的杂波信息对典型目标回波进行仿真以辅助完成后续检测器的设计。具有匀速和匀加速运动状态的典型小目标回波可以被表示为^[15]

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

通过对公开IPIX数据集的研究发现，当信杂比低于-10 dB时，目标无法被检测到；当信杂比高于20 dB时，目标一定可以被检测到，所以将经验参数 \bar{A} 设置为[-10 dB, 20 dB]的对应区间 $[10^{-1}, 10^{1/2}]$ 上的均匀分布，用于调节信杂比， P_c 是杂波的平均功率， $a(n)$ 是一个高度相关的取值为非负的随机序列，被用于模拟小目标的幅度波动， λ 是雷达的工作波长， ϑ_0 和 ϑ_1 分别是目标的初始和终止径向速度， Δt 是雷达的脉冲重复周期， φ_0 是随机分布于区间 $[0, 2\pi]$ 上的初始相位， N 是仿真目标回波信号的长度。

考虑到小目标幅度的物理特性，幅度序列 $a(n)$ 被建模为一个非负的、高度相关的、单位功率的随机序列，并具有可调动态范围和去相关时间。序列 $a(n)$ 的生成步骤分为如下4步：第1，生成一个独立、同服从在区间 $[0, 1]$ 上均匀分布的序列 $u(n)$ ；第2，由于幅度序列具有高度的空时相关性，所以将随机序列 $u(n)$ 作为一个1阶自回归模型的输入^[15]

$$v(1) = u(1), v(n+1) = \rho v(n) + u(n+1), \quad n = 1, 2, \dots, \rho \in (0, 1) \quad (3)$$

即可产生高度相关的序列 $v(n)$ ，其中 $\rho \in [0.95, 0.99]$ 表示1阶相关系数，经计算，此处序列 $v(n)$ 的取值范围在区间 $[-1/(1-\rho), 1/(1-\rho)]$ 内；第3，由于目标的幅度为非负值，所以将序列 $v(n)$ 转化为一个非负的序列 $v_+(n)$

$$v_+(n) = \frac{1}{1-\rho} + v(n) \in \left[0, \frac{2}{1-\rho} \right], n \in \mathbb{N} \quad (4)$$

第4，由于刚开始产生的幅度序列不够稳定，所以为了避免这种情况我们只选取后续产生的幅度序列 $v_+(n+M)$ ，其中 M 是一个足够大的正数，再对

序列进行功率归一化, 即可得到最终的幅度序列 $a(n)$

$$a(n) = \frac{v_+(n+M)}{\sqrt{E\{v_+^2\}}} = \frac{\sqrt{3(1+\rho)}(1-\rho)}{\sqrt{2(2+\rho)}} v_+(n+M),$$

$$n = 1, 2, \dots, N \quad (5)$$

其中, $E\{v_+^2\}$ 是序列 $v_+(n)$ 的功率。

此外, 初始径向速度 v_0 和终止径向速度 v_1 可以被表示为^[15]

$$\left. \begin{aligned} x, y, z &\sim \text{Uniform}[-1, 1] \\ v_0 &= \eta x, v_1 = \eta y, |v_0 - v_1| \leq \zeta N \Delta t, \theta = \pi z \\ \vartheta_0 &= v_0 \cos \theta, \vartheta_1 = v_1 \cos \theta \end{aligned} \right\} \quad (6)$$

其中, θ 为雷达视线与目标运动方向的夹角, 服从 $[-\pi, \pi]$ 上的均匀分布, v_0 和 v_1 分别表示目标的初始速度和终止速度, 随机数 x, y, z 是相互独立的。值得注意的是, 当加速度的限制条件不满足时, 随机数 x 和 y 需要重新产生。在目标回波仿真过程中, η 的经验取值为 5 m/s, ζ 的经验取值为 2 m/s², 其对应于小目标的速度在 $[-5 \text{ m/s}, 5 \text{ m/s}]$ 之间, 不超过 10 节, 最大加速度为 2 m/s²。通过上述步骤, 一个典型的仿真目标回波产生器被构建。需要注意的是, 该仿真目标回波产生器旨在生成足量的不完备的目标回波样本, 为后续检测器的设计提供更多有效信息, 而不是覆盖各种各样的目标回波。

图1是在HV极化下第1组数据的真实目标回波和仿真目标回波的7个特征的对比如, 其中真实目标特征用蓝色表示, 仿真目标回波用红色表示。由此可知仿真目标回波产生器的有效性, 进一步仿真

目标回波的特征与真实目标回波的特征也是十分接近的。

3 基于可控虚警的改进K-NN方法

很多机器学习算法可以有效地解决分类问题, 比如SVM, K-NN, NN等。但是普通的二分类问题和目标检测问题存在着两点不同: 第一, 在普通的二分类问题中, 两类样本的数量是均衡的, 但是在目标检测中, 杂波数据远远多于目标数据; 第二, 普通的二分类问题中, 两类错分概率是等价的, 但是在目标检测中, 虚警概率比漏检概率更重要(通常, 虚警概率要低于 10^{-3} 而漏检概率可能达十分之几)。这两点不同导致机器学习中用于分类问题的算法不能直接使用在目标检测问题中。其中, 两类样本数量不均衡的问题已经被所提出的仿真目标回波产生器所解决, 但是实现虚警可控仍是一个难题。

3.1 传统K-NN算法概述

许多学者都在研究机器学习算法, 用于分类和回归问题。K-NN算法是一种简单有效的非参数分类算法, 其经常用于许多的模式识别问题。K-NN算法使用训练数据本身直接对测试样本进行分类^[6,17]。首先寻找到与测试样本最相似的 k 个训练样本(称之为近邻), 之后将 k 个近邻中最主要的类别定为待测样本的类别, 其中样本之间相似性度量以距离计算为准。为了获得 k 个近邻, 测试样本需与每个训练样本进行距离计算。

具体算法步骤可描述如下:

(1) 给出一个测试样本 $x_{\text{test}} \in \mathbb{R}^d$ 和训练样本集

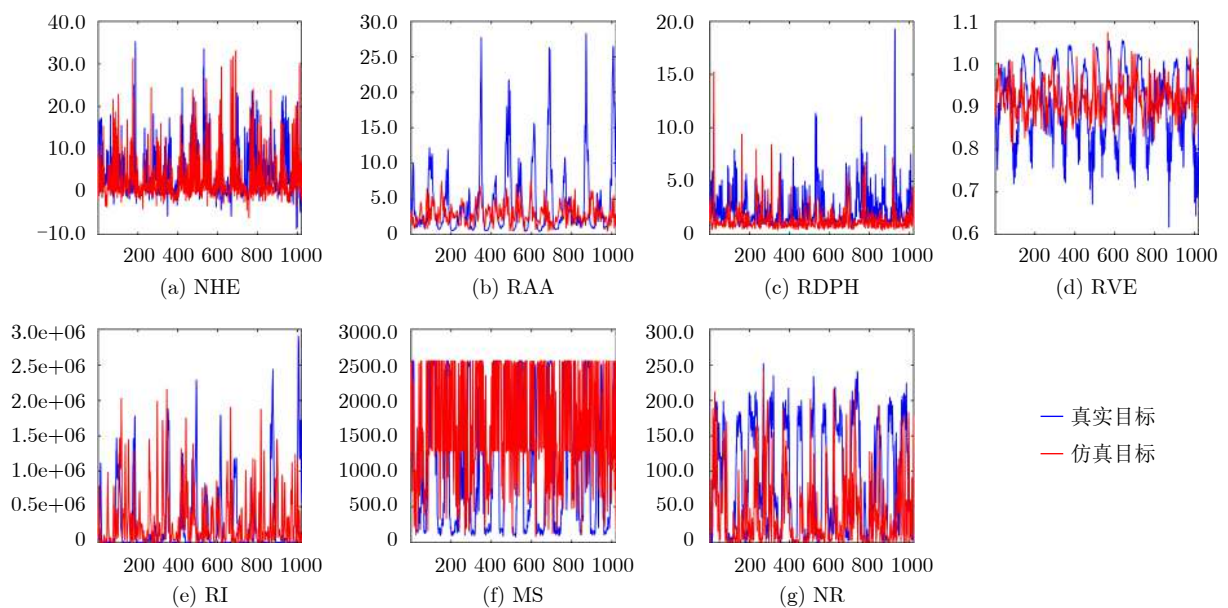


图1 仿真目标7特征与真实目标7特征对比图

Fig. 1 The comparisons of seven features of simulated targets returns and real targets returns

合 \mathbf{X}_{set} , 其中, x_i 是训练集中第 i 个训练样本, 计算测试样本和每个训练样本之间的欧式距离

$$d(x_{\text{test}}, x_i) = \|x_{\text{test}} - x_i\|^2, i = 1, 2, \dots, m \quad (7)$$

其中, m 是训练样本的数目, $d(x, y)$ 表示计算 x 与 y 之间欧氏距离。

(2) 将所有距离 $\{d(x_{\text{test}}, x_i), i=1, 2, \dots, m\}$ 从小到大进行排序, 选取 k 个最小的距离值 $\{d(x_{\text{test}}, x_{i_1}), d(x_{\text{test}}, x_{i_2}), \dots, d(x_{\text{test}}, x_{i_k})\}$ 所对应的 k 个训练样本 $\{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}$, 此处称之为 k 个近邻。

(3) 令 l_i 为训练样本 x_i 的类标 (也可称之为类别), 并认为 k 个近邻中最主要的类标即为测试样本的类标, 即

$$l_{\text{test}} = \text{mode} \{l_{i_1}, l_{i_2}, \dots, l_{i_k}\} \quad (8)$$

其中, l_{test} 表示测试样本的类标, 操作 $\text{mode} \{\cdot\}$ 表示计算集合中所有元素的众数。

由于简单的判决准则, 即测试样本的类标取决于 k 个近邻中最主要的类标, K-NN 算法被广泛使用。但是, 在目标检测中, 虚警概率和漏检概率是不均等的, 往往我们对虚警概率的控制要求更为严格。此外, 机器学习算法通常要求两类样本的数目近似相等。所以, 机器学习算法中可用于解决二分类问题的算法不可以直接使用在雷达目标检测中。

3.2 基于可控虚警的改进K-NN海面小目标检测

之前的很多关于K-NN算法的研究普遍集中于寻找最优的 k 值以提高检测器的性能。如果要将K-NN算法使用在雷达目标检测问题中, 准确地控制虚警概率则是重要而且必不可少的。虽然凸包学习算法^[9,11]可以精准地控制虚警概率, 但是它只能被用于单分类问题, 以及只能在低维空间 ($n \leq 3$) 中使用。因此, 凸包学习算法便不再适用于高维空间两分类问题。一种改进的基于可控虚警的K-NN方法被提出, 具体过程如下:

(1) 通过雷达接收机获取海杂波时间序列 $c(n)$;

(2) 仿真目标的产生: 使用仿真目标回波产生器产生与海杂波数据等量的仿真目标回波数据 $s(n)$;

(3) 提取有效特征: 对杂波序列 $c(n)$ 和目标回波序列 $s(n)$ 分别提取7个有效的特征, 构成七维特征向量, 并将其组成特征矩阵 \mathbf{S}_0 和 \mathbf{S}_1 , 则训练样本集为 $\mathbf{S} = \mathbf{S}_0 \cup \mathbf{S}_1$, 其中, \mathbf{S} 为 $m \times n$ 的矩阵, $m=7$ 是特征数量, n 是海杂波和仿真目标样本数目之和;

(4) 对于当前测试样本集中的每一个测试样本也提取同样的7特征并构成特征向量, 计算其与所有训练样本之间的欧式距离, 得到距离向量 \mathbf{D} ;

(5) 将距离向量 \mathbf{D} 中的元素按照从小到大的顺序进行排列, 并取 k 个最小距离值所对应的 k 个训练样本, 构成 k 近邻;

(6) 计算 k 近邻中, 原本属于海杂波样本 \mathbf{S}_0 的数量 n_0 和原本属于目标样本 \mathbf{S}_1 的数量 n_1 , 其中 $n_0 + n_1 = k$;

(7) 设置一个阈值 w , 若 $n_0 \geq w$, 则测试样本被分类为海杂波, 否则认为其是目标;

(8) 通过多次实验, 经验地选取一个阈值 w^* , 固定其保持不变, 改变近邻数目 k 的取值, 使得实现的虚警概率 P_{fa} 和预设的虚警概率 P_f 之间差的绝对值小于一个极小值, 将最终符合虚警要求的近邻数记为 k^* ;

(9) 对于每个测试样本, 在 k^* 个近邻中, 若 $n_0 \geq w^*$, 测试样本被分为海杂波, 否则被分为目标。

图2是所提检测器的流程图, 主要分为两个部分: 用蓝色表示的离线操作部分和用黄色表示的在线操作部分, 红色为共用部分。离线部分包括海杂波的获取、仿真目标回波的产生、零假设和各择假设下两类特征向量的提取和根据虚警要求对应的改进K-NN算法中参数的计算; 在线部分包括待检测回波的获取、待检测的特征向量的构成以及判决部分。

4 实验与分析

4.1 实验数据

实验所用数据为20组(4种极化)驻留模式下的实测IPIX^[7]海杂波数据, 其雷达工作在X波段, $f_r=1000$ Hz, 除了第18组和19组数据的距离分辨率

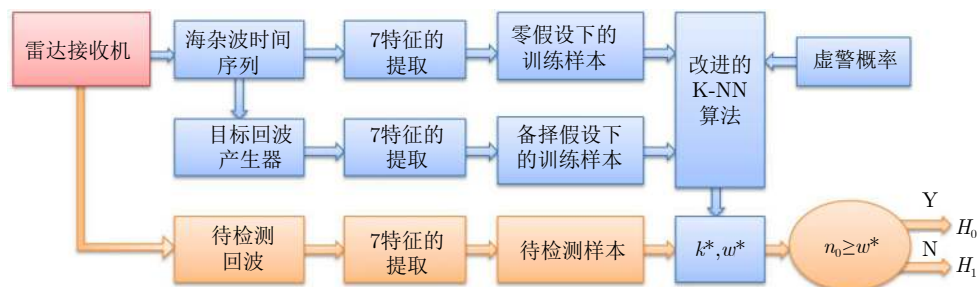


图 2 所提检测器的流程图

Fig. 2 The flowchart of the proposed detector

为15 m和9 m以外, 其余均为30 m。前10组数据采集于1993年加拿大东海岸, 测试目标是一个用锚固定的漂浮的直径为1 m的塑料小球, 随着海浪上下运动, 每组数据的时间序列长度为 2^{17} , 距离单元数目为14; 后10组数据采集于1998年在加拿大安大略湖, 测试目标是一艘低速运动的小船, 每组数据由28个连续距离单元构成, 每个距离单元包含60000个脉冲序列。具体的数据信息, 如风速(Wind Speed, WS)、有效浪高(Significant Wave Height, SWH)、雷达视线与风速夹角以及目标所在单元和周围影响单元如表1所示。

4.2 检测器性能与分析

在图3中, 分别画出了在IPIX雷达数据集上, 4种极化下、观测时间为0.512 s的多种检测器平均检测概率对比图。通过比较发现, 所提检测器在76组数据上呈现出最优的检测结果, 剩下4组也是接近于最优检测结果。可以得出结论, 所提检测器具有良好的检测性能和稳定性。在表2中, 通过观察各种基于特征的检测器^[9,11,12]的平均检测概率可以得知, 使用单一特征或少量特征的检测器的性能远差于联合使用多个有效特征的检测器的性能。在图4中, 分别画出了在IPIX雷达数据集上, 四种极

化下、观测时间为1.024 s的多种检测器平均检测概率对比图。通过对比发现, 当观测时间从0.512 s提升至1.024 s时, 所有检测器的性能均有提升。其中, 基于分形的检测器的性能从0.329提升至0.435, 它是所有检测器中提升效果最明显的。但是所提出的检测器的性能只从0.851提升至0.892, 这是由于天花板效应。因为在观测时间为0.512 s时, 所提出的检测器的平均检测概率已经接近于1, 所以它只有很小的空间以供性能的提升。

更进一步, 为了验证所提方法在多种虚警概率下的更优性能, 本文对比了所提检测器和其余检测器在虚警概率为0.01时的检测结果。如表3所示, 通过实验表明, 所提检测器在更高虚警率下, 表现仍旧优异。

除此之外, 为了进一步的分析7个特征在检测过程中的贡献, 本文设计了7个仅使用6个特征的基于KNN的检测器, 即每个检测器分别去掉1个特征。通过这7个基于6特征的KNN检测器与所提出检测器的性能差值来评估这些特征的重要性。具体结果如表4。

其中, 性能损失为在相同条件下, 所提出的使用7个特征的检测器的检概率减去使用6个特征的检

表 1 IPIX数据集描述^[10]

Tab. 1 Description of IPIX radar database^[10]

序号	数据	WS(km/h)	SWHs(m)	Angle (degree)	目标单元	影响单元
1	19931107_135603_starea17	9	2.2	9	9	8,10,11
2	19931108_220902_starea26	9	1.1	97	7	6,8
3	19931109_191449_starea30	19	0.9	98	7	6,8
4	19931109_202217_starea31	19	0.9	98	7	6,8,9
5	19931110_001635_starea40	9	1.0	88	7	5,6,8
6	19931111_163625_starea54	20	0.7	8	8	7,9,10
7	19931118_023604_stareC0000280	10	1.6	130	8	7,9,10
8	19931118_162155_stareC0000310	33	0.9	30	7	6,8,9
9	19931118_162658_stareC0000311	33	0.9	40	7	6,8,9
10	19931118_174259_stareC0000320	28	0.9	30	7	6,8,9
11	19980204_202225_ANTSTEP	-	-	165	24	23,25,26
12	19980204_202525_ANTSTEP	-	-	180	7	6,8,9
13	19980204_163113_ANTSTEP	-	-	165	24	23,25,26
14	19980205_171437_ANTSTEP	-	-	180	7	6,8,9
15	19980205_180558_ANTSTEP	-	-	180	7	6,8,9
16	19980212_195704_ANTSTEP	-	-	180	7	6,8,9
17	19980223_164055_ANTSTEP	-	-	165	31	30,32,33
18	19980223_173317_ANTSTEP	-	-	165	32	31,33,34
19	19980223_173950_ANTSTEP	-	-	165	29	28,30-34
20	19980304_184537_ANTSTEP	-	-	-	21	20,22

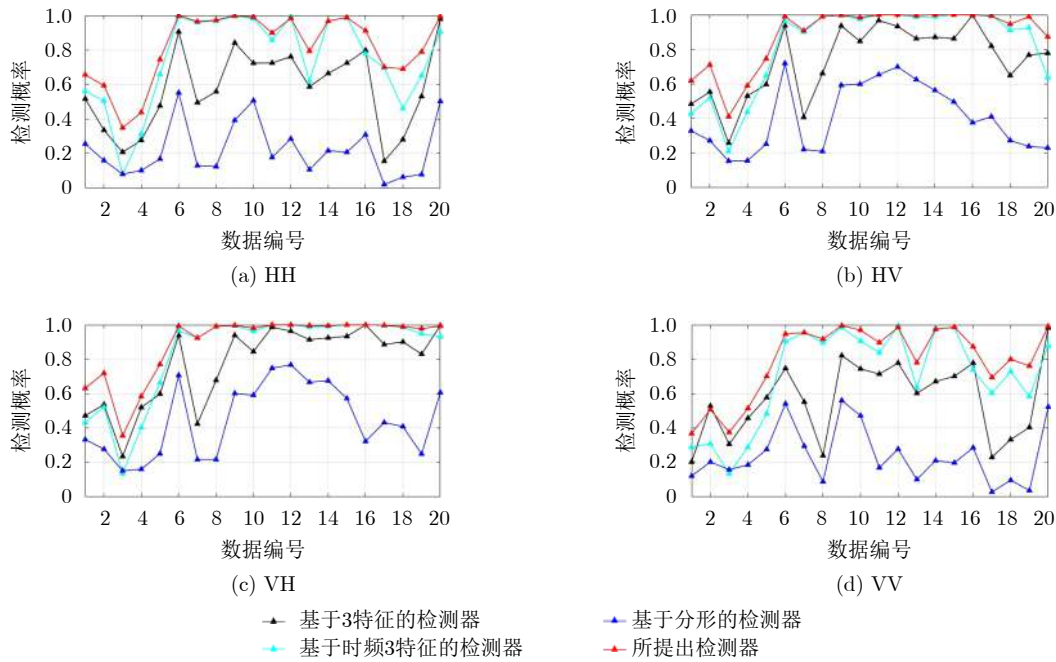


图 3 所提检测器与其余检测器的检测概率

Fig. 3 Detection probabilities of the proposed detector and other detectors

表 2 IPIX数据集上多种检测器的平均检测概率

Tab. 2 The average detection probabilities of detectors on IPIX radar database

检测器	观测时间(s)	HH	HV	VH	VV	平均
基于分形的检测器 ^[12]	0.512	0.223	0.404	0.448	0.241	0.329
	1.024	0.301	0.536	0.576	0.328	0.435
基于3特征的检测器 ^[9]	0.512	0.577	0.736	0.776	0.569	0.665
	1.024	0.622	0.797	0.813	0.598	0.708
基于时频3特征的检测器 ^[11]	0.512	0.747	0.826	0.842	0.706	0.780
	1.024	0.821	0.882	0.877	0.789	0.842
所提检测器	0.512	0.821	0.887	0.895	0.800	0.851
	1.024	0.868	0.922	0.921	0.858	0.892

测器的检测概率。由表4的结果可知，当所有特征中去掉RDPH时，对检测性能的影响最大(其中HH极化下性能损失高达5.14%)，其他依次是RI, MS, RAA, NHE, RVE和NR。缺失某个特征后，在4种极化下的性能损失均有不同，不存在某个特征在4种极化下的贡献都最低的情况。当然，也不存在某个特征在每组数据下都表现最优的情况，即各个特征在不同情况下或多或少都对检测结果做出了贡献，所以联合使用多个来自不同域的特征也是很有必要的。总的来说，该检测方法为一个开放的理论框架，允许更多有效且互补的特征加入以提高检测器的性能。

下面，为了印证本文所提方法对于虚警率的控制情况，图5展示了k值的变化对虚警率的影响。通过改变k的取值，就可以达到不同的虚警概率。

本文中，当预设虚警率为0.01时，实际的虚警与预设的虚警之差的绝对值小于0.001即满足要求；当预设虚警为0.001时，实际的虚警与预设的虚警之差的绝对值小于0.0001即满足要求。最终所提出的检测器在20组IPIX雷达数据集上实现的虚警如表5所示，其中4种极化的平均实现虚警率和所有极化的平均实现虚警率已列出，与预设虚警率相差甚微。由此可见，所提检测器很好地控制了虚警率。

5 结论

本文提出了一种高维空间中基于可控虚警K-NN的海面小目标检测方法。现有的基于特征的检测方法存在维数限制问题，维数限制问题严重地阻止了更多有效特征在目标检测过程中的使用，更进一步限制了性能的提升。所提检测器很好地解决了上述

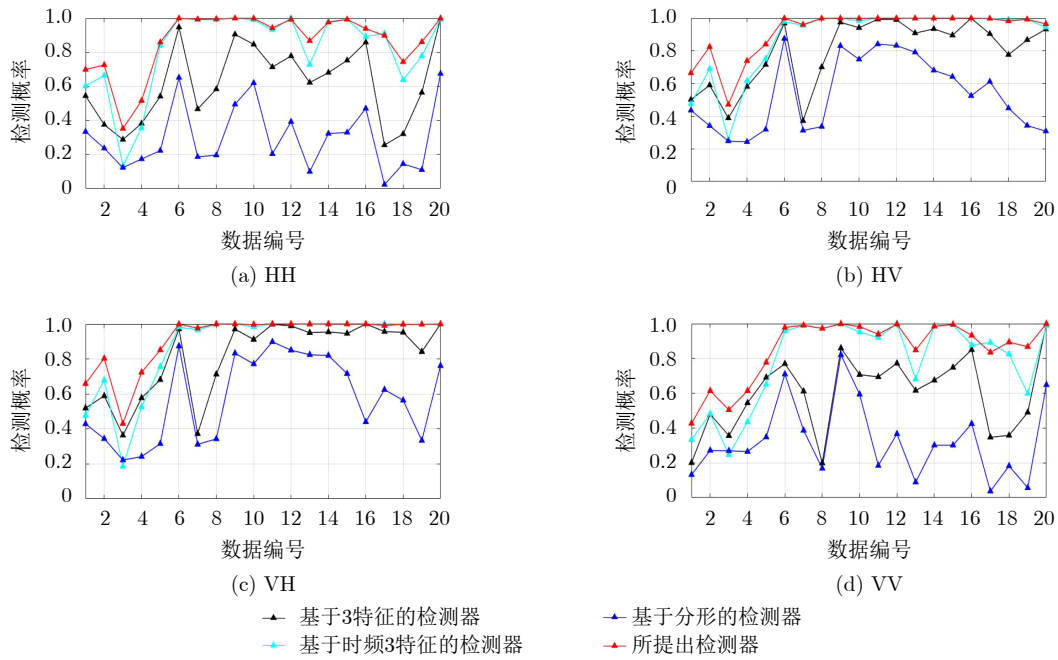


图4 所提检测器与其余检测器的检测概率

Fig. 4 Detection probabilities of the proposed detector and other detectors

表3 IPIX数据库中20组数据的平均检测结果对比

Tab. 3 The comparisons of average detection results of 20 datasets on IPIX radar database

检测器	虚警概率	HH	HV	VH	VV	平均
基于分形的检测器 ^[12]	0.010	0.322	0.514	0.554	0.335	0.431
	0.001	0.223	0.404	0.448	0.241	0.329
基于3特征的检测器 ^[9]	0.010	0.673	0.812	0.841	0.657	0.745
	0.001	0.577	0.736	0.776	0.569	0.665
基于时频3特征的检测器 ^[11]	0.010	0.806	0.878	0.882	0.781	0.837
	0.001	0.747	0.826	0.842	0.706	0.780
所提检测器	0.010	0.887	0.932	0.932	0.868	0.905
	0.001	0.821	0.887	0.895	0.800	0.851

表4 基于6特征的KNN检测器在IPIX数据库上20组数据的平均检测结果对比 (%)

Tab. 4 The average detection results comparisons of KNN-based detectors using six features at 20 datasets on IPIX radar database (%)

检测器所去掉的特征	NHE	RAA	RDPH	RVE	RI	MS	NR
性能损失	HH	0.01	0.79	5.14	0.50	3.01	0.20
	HV	0.36	0.02	2.50	0.01	4.19	0.17
	VH	0.67	0.21	2.68	0.08	5.45	0.12
	VV	0.01	0.63	4.62	0.01	0.89	0.02
	平均	0.26	0.41	3.74	0.15	3.38	0.13

问题。考虑到普通的二分类算法中两类样本数目均衡的需求，使用一种典型的仿真目标回波产生器可产生与杂波等量的典型的仿真目标回波，为后续检测器的设计起到辅佐的作用，使得目标信息被使用到训练检测器的过程。此外，通过改进机器学习中

K-NN算法，实现了目标检测问题中最重要的虚警率可控，即有效地将K-NN算法应用到了目标检测中。最后通过公认的IPIX雷达数据集的验证，本文所提检测器与其余基于单特征或者3特征的检测器相比，具有良好的检测结果和鲁棒性。

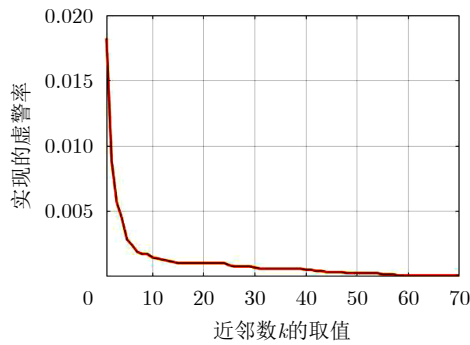


图 5 k 值不同时, 所实现的虚警率变化图, 其中 $w^*=3$

Fig. 5 Realized false alarm rate when k takes different values, where the $w^*=3$

表 5 所提检测器在 20 组 IPIX 雷达数据集上实现的虚警概率

Tab. 5 The realized false alarm rate of the proposed detector of 20 datasets on the IPIX radar database

所需的虚警率	实现的虚警率				
	HH	HV	VH	VV	平均
0.010	0.008774	0.009153	0.008776	0.008563	0.008816
0.001	0.001011	0.001028	0.001010	0.001015	0.001016

参 考 文 献

- [1] WANG Shibin, CHEN Xufeng, WANG Yan, *et al.* Nonlinear squeezing time-frequency transform for weak signal detection[J]. *Signal Processing*, 2015, 113: 195–210. doi: [10.1016/j.sigpro.2015.01.022](https://doi.org/10.1016/j.sigpro.2015.01.022).
- [2] WARD K, TOUGH R, and WATTS S. Sea Clutter: Scattering, the K Distribution and Radar Performance[M]. 2nd ed. London: The Institution of Engineering and Technology, 2013. doi: [10.1049/PBRA025E](https://doi.org/10.1049/PBRA025E).
- [3] RABIDEAU D J and PARKER P. Ubiquitous MIMO multifunction digital array radar[C]. The 37th Asilomar Conference on Signals, Systems & Computers, Pacific Grove, USA, 2003: 1057–1064. doi: [10.1109/ACSSC.2003.1292087](https://doi.org/10.1109/ACSSC.2003.1292087).
- [4] CARTA P, GALATI G, PIRACCI E G, *et al.* Implementation of the ubiquitous radar concept with a conformal array[C]. 2015 European Radar Conference (EuRAD), Paris, France, 2015: 1463–1466. doi: [10.1109/EuRAD.2015.7346332](https://doi.org/10.1109/EuRAD.2015.7346332).
- [5] SHI Sainan and SHUI Penglang. Detection of low-velocity and floating small targets in sea clutter via income-reference particle filters[J]. *Signal Processing*, 2018, 148: 78–90. doi: [10.1016/j.sigpro.2018.02.005](https://doi.org/10.1016/j.sigpro.2018.02.005).
- [6] 陈世超, 罗丰, 胡冲, 等. 基于多普勒谱非广延熵的海面目标检测方法[J]. 雷达学报, 2019, 8(3): 344–354. doi: [10.12000/JR19012](https://doi.org/10.12000/JR19012).
CHEN Shichao, LUO Feng, HU Chong, *et al.* Small target detection in sea clutter background based on Tsallis entropy of Doppler spectrum[J]. *Journal of Radars*, 2019, 8(3): 344–354. doi: [10.12000/JR19012](https://doi.org/10.12000/JR19012).
- [7] 左磊, 产秀秀, 禄晓飞, 等. 基于空域联合时频分解的海面微弱目标检测方法[J]. 雷达学报, 2019, 8(3): 335–343. doi: [10.12000/JR19035](https://doi.org/10.12000/JR19035).
ZUO Lei, CHAN Xiuxiu, LU Xiaofei, *et al.* A weak target detection method in sea clutter based on joint space-time-frequency decomposition[J]. *Journal of Radars*, 2019, 8(3): 335–343. doi: [10.12000/JR19035](https://doi.org/10.12000/JR19035).
- [8] 苏宁远, 陈小龙, 关键, 等. 基于卷积神经网络的海上微动目标检测与分类方法[J]. 雷达学报, 2018, 7(5): 565–574. doi: [10.12000/JR18077](https://doi.org/10.12000/JR18077).
SU Ningyuan, CHEN Xiaolong, GUAN Jian, *et al.* Detection and classification of maritime target with micro-motion based on CNNs[J]. *Journal of Radars*, 2018, 7(5): 565–574. doi: [10.12000/JR18077](https://doi.org/10.12000/JR18077).
- [9] SHUI Penglang, LI Dongchen, and XU Shuwen. Tri-feature-based detection of floating small targets in sea clutter[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2014, 50(2): 1416–1430. doi: [10.1109/taes.2014.120657](https://doi.org/10.1109/taes.2014.120657).
- [10] Cognitive Systems Laboratory. IPIX radar database[EB/OL]. <http://soma.mcmaster.ca//ipix.php>, 2012.
- [11] SHI Sainan and SHUI Penglang. Sea-surface floating small target detection by one-class classifier in time-frequency feature space[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(11): 6395–6411. doi: [10.1109/TGRS.2018.2838260](https://doi.org/10.1109/TGRS.2018.2838260).
- [12] HU Jing, TUNG W W, and GAO Jianbo. Detection of low observable targets within sea clutter by structure function based multifractal analysis[J]. *IEEE Transactions on Antennas and Propagation*, 2006, 54(1): 136–143. doi: [10.1109/TAP.2005.861541](https://doi.org/10.1109/TAP.2005.861541).
- [13] LI Dongchen and SHUI Penglang. Floating small target detection in sea clutter via normalised Hurst exponent[J]. *Electronics Letters*, 2014, 50(17): 1240–1242. doi: [10.1049/el.2014.1569](https://doi.org/10.1049/el.2014.1569).
- [14] GUO Zixun and SHUI Penglang. Sea-surface floating small target detection based on feature compression[J]. *The Journal of Engineering*, 2019, 2019(21): 8160–8164. doi: [10.1049/joe.2019.0694](https://doi.org/10.1049/joe.2019.0694).
- [15] SHUI Penglang, GUO Zixun, and SHI Sainan. Feature-compression-based detection of sea-surface small targets[J]. *IEEE Access*, 2019, 8: 8371–8385. doi: [10.1109/ACCESS.2019.2962793](https://doi.org/10.1109/ACCESS.2019.2962793).
- [16] ZHANG Shichao, LI Xuelong, and ZONG Ming. Efficient kNN classification with different numbers of nearest neighbors[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, 29(5): 1774–1785. doi: [10.1109/TNNLS.2017.2673241](https://doi.org/10.1109/TNNLS.2017.2673241).
- [17] FAYED H A and ATIYA A F. A novel template reduction approach for the K-nearest neighbor method[J]. *IEEE Transactions on Neural Networks*, 2009, 20(5): 890–896. doi: [10.1109/TNN.2009.2018547](https://doi.org/10.1109/TNN.2009.2018547).

作者简介



郭子薰(1994-), 女, 陕西西安人, 西安电子科技大学博士生。主要研究方向为雷达目标检测、机器学习和海杂波信号处理。

E-mail: zxguo_724@stu.xidian.edu.cn



水鹏朗(1967-), 男, 陕西西安人, 博士, 教授。1999年在西安电子科技大学获得博士学位, 现担任西安电子科技大学电子工程学院雷达信号处理国家重点实验室教授、硕导、博导。主要研究方向为海杂波建模、雷达目标检测和图像处理。

E-mail: plshui@xidian.edu.cn



白晓惠(1998-), 女, 陕西宝鸡人, 西安电子科技大学博士生。主要研究方向为雷达目标检测、机器学习和海杂波信号处理。

E-mail: xhbai@stu.xidian.edu.cn



许述文(1985-), 男, 安徽黄山人, 博士, 副教授。2011年在西安电子科技大学获得博士学位, 现担任西安电子科技大学电子工程学院雷达信号处理国家重点实验室副教授、硕导、博导。主要研究方向为雷达目标检测、机器学习、时频分析和SAR图像处理。

E-mail: swxu@mail.xidian.edu.cn



李东宸(1988-), 男, 陕西西安人, 博士, 高级工程师。2016年在西安电子科技大学获得博士学位, 现就职于中国船舶工业系统工程研究院。主要研究方向为目标检测、图像处理、无人机任务规划。

E-mail: charlieli.xidian@outlook.com