DOI:10.3969/j.issn.1672-2337.2024.01.011

基于多维矩特征联合的K分布杂波参数估计

施赛楠[1]，高季娟[1]，李东宸[2]

（1. 南京信息工程大学电子与信息工程学院，江苏南京 210044；2. 中国船舶工业系统工程研究院，北京 100094）

摘　要： K分布杂波参数估计在海洋雷达自适应目标检测中起着关键作用。传统的矩估计器通过联立两个矩求解参数，其估计性能受限于有限矩的信息。因此，本文提出一种基于多维矩特征联合的参数估计方法，旨在拓展矩信息的维度。首先，从观测数据中，提取多个精心设计的线性矩和对数矩，构建一个特征向量。其次，将传统基于统计分布的参数估计问题转换为非线性优化问题。然后，通过引入梯度提升决策树（Gradient Boosting Decision Tree, GBDT）算法，建立特征向量和形状参数之间的函数关系，实现形状参数的估计。此外，推导证明特征向量与尺度参数的独立性以及二阶矩只依赖于尺度参数，从而解决两个参数估计的相关性问题。最后，仿真和实测数据结果表明，所提估计器能利用多个矩的丰富信息，进一步提高参数估计性能。特别是在小形状参数时，其估计性能显著优于现有矩估计法和$z\log(z)$期望法。

关键词： 海杂波；K分布；参数估计；多维矩特征

中图分类号： TN957.51　　**文献标志码：** A　　**文章编号：** 1672-2337（2024）01-0077-10

引用格式： 施赛楠，高季娟，李东宸. 基于多维矩特征联合的K分布杂波参数估计[J]. 雷达科学与技术，2024，22（1）：77-86.
SHI Sainan，GAO Jijuan，LI Dongchen. Parameter Estimation of K-Distributed Clutter via Multidimensional Moment Feature Combination[J]. Radar Science and Technology, 2024, 22(1):77-86.

Parameter Estimation of K-Distributed Clutter via Multidimensional Moment Feature Combination

SHI Sainan[1]，GAO Jijuan[1]，LI Dongchen[2]

（1. *Shool of Electronic and Information Engineering, Nanjing University of Information Science & Technology, Nanjing* 210044, *China*;
2. *System Engineering Research Institute, China State Shipbuilding Corporation, Beijing* 100094, *China*）

Abstract: Parameter estimation of K-distributed clutter plays an important role in adaptive target detection for marine radars. The traditional moment estimator solves the parameters by combining two moments and its performance is limited by the information of finite moments. Therefore, a parameter estimation method via multidimensional moment feature combination is proposed to expand the dimension of moment information. Firstly, multiple well-designed linear moments and logarithmic moments are extracted from the observed data to construct a feature vector. Secondly, the traditional parameter estimation problem based on statistical distribution is transformed into a nonlinear optimization problem. Thirdly, the gradient boosting decision tree (GBDT) algorithm is introduced to establish the functional relationship between the feature vector and the shape parameter to achieve the estimation of the shape parameter. Besides, it is proved that the feature vector is independent of the scale parameter and the second moment only depends on the scale parameter, so as to solve the relationship of the two parameters in estimation. Finally, the simulation and measured data results show that the proposed estimator can take advantage of the rich information of multiple moments to further improve the performance of parameter estimation. Especially for small shape parameters, its estimation performance is significantly better than the existing moment estimation methods and $z\log(z)$ expectation method.

Key words: sea clutter; K-distribution; parameter estimation; multidimensional moment features

0　引　言

复合高斯模型（Compound-Gaussian Model, CGM）[1]将海杂波建模为一个慢变纹理调制一个快变复高斯散斑的过程，已广泛使用于海杂波建模中。在CGM中，不同的纹理分布产生不同的海杂波幅

收稿日期：2023-05-13；修回日期：2023-06-19
基金项目：国家自然科学基金（No.62201184）；江苏省高等学校基础科学（自然科学）研究项目（No.21KJB510036）

度分布。典型的有，具有伽马纹理的K分布[2-3]，具有逆伽马纹理的广义Pareto分布[4]、具有逆高斯纹理的复合高斯分布[5]和具有对数正态纹理的复合高斯分布[6]。这些杂波分布在研究杂波仿真、杂波抑制、最优或近最优检测器设计等方面起着至关重要的作用。

目前，K分布已广泛用于描述中低分辨率海杂波的幅度分布[2-3]。该分布的特性参数包括形状参数和尺度参数，前者决定杂波的非高斯特性，后者反映杂波功率水平。通常，采用以下三大类方法获得杂波参数的估计值。第一类，基于概率密度函数（Probability Density Function, PDF）的估计方法，比如最大似然（Maximum Likelihood，ML）估计法[7]和矩估计法（Method of Moment，MoM）[8]。ML估计精度高，但没有显式解析表达式且计算复杂度高。考虑到实际雷达系统要求实时估计的要求，发展了结构简单且快速估计的MoM估计，比如2-4阶矩估计、1-2阶矩估计、分数阶矩估计[9-10]等。此外，为了进一步挖掘矩信息，Yang等人将原点矩偏导和原点矩之间的关系用于参数估计中[11]。针对小形状参数下矩估计误差大的问题，Blacknell等人提出$z\log(z)$期望的参数估计方法[12]。同时，在该方法上进行拓展，Hu等人得到了$z^r\log(z)$期望法[13]。第二类，基于累积分布函数（Cumulative Distribution Function, CDF）的估计方法，比如分位点估计法（Method of Percentile, MoP）[14-15]。Yu提出基于三分位点估计器（Tri-percentile Estimator, TPE）[15]，两个分位点比值用于估计形状参数，第三个分位点用于估计尺度参数，适用于存在较多异常值的杂波环境中。由于K分布存在贝塞尔函数，因而很难获得分位点比值和形状参数之间的函数关系，必须借助表格法获得。第三类，基于非线性网络的参数估计方法，引入人工智能网络建立观测值和参数之间的函数关系。Fernández等人[16]提出先对数据功率进行归一化处理，然后将直方图作为神经网络的输入来估计形状参数，只需单个网络。但是，归一化处理需要提前估计尺度参数，其误差势必会影响后续形状参数的估计。因此，在设计估计器时，必须解决形状参数和尺度参数独立估计的问题。

为了进一步提高参数估计精度，本文提出了一种基于多维矩特征联合的K分布杂波参数估计方法。主要创新有以下三个方面。第一，联合提取多个线性矩和对数矩，构建一个特征向量用于估计形状参数，拓展矩信息的维度，进而提高数据信息的利用率。第二，将传统基于统计分布的参数估计问题转换为非线性优化问题，引入梯度提升树（Gradient Boosting Decision Tree，GBDT）算法[17]，建立特征向量和形状参数之间的非线性关系，提高估计精度。第三，精心设计矩特征，并推导证明特征向量与尺度参数相互独立，并且二阶矩只依赖于尺度参数，实现形状参数和尺度参数的独立估计。

1 K分布杂波的矩估计

1.1 K分布杂波

根据海杂波的物理散射机理，海杂波可表述为一个慢变分量调制一个快变分量，即

$$c = \sqrt{\tau}\,u, \tau > 0 \tag{1}$$

式中：τ称为纹理分量，是一个正的随机变量，由大尺度的涌浪产生；u称为散斑分量，服从零均值、单位方差的复高斯分布，由小尺度的毛细波产生。在CGM中，不同的纹理分布会产生不同的杂波幅度分布。当纹理服从伽马分布时，海杂波幅度$x = \sqrt{\tau}\,|u|$服从K分布。K分布的PDF表达式为

$$f(x;v,b) = \frac{4}{x\Gamma(v)}\left(x\sqrt{\frac{v}{b}}\right)^{v+1}\mathrm{K}_{v-1}\left(2x\sqrt{\frac{v}{b}}\right), v,b > 0 \tag{2}$$

式中，$\Gamma(\cdot)$为伽马函数，$\mathrm{K}_{v-1}(\cdot)$为$v-1$阶的第二类修正Bessel函数，v为形状参数，b为尺度参数。形状参数决定分布的非高斯性，尺度参数与杂波功率相关。

根据式（2），计算出r阶矩为

$$m_r = E\left[x^r\right] = \frac{\Gamma(1+r/2)\Gamma(v+r/2)}{\Gamma(v)}\left(\frac{b}{v}\right)^{r/2} \tag{3}$$

式中，E为期望运算符。通常，理论矩是未知的，需要通过获取实际环境中的N个样本获得。假设N个样本$x_1, x_2, \cdots, x_N$服从K分布且相互独立，则样本矩为

$$\hat{m}_r = \frac{1}{N}\sum_{i=1}^{N} x_i^{\,r} \tag{4}$$

当样本数目N趋向于无穷时，样本矩等于理论矩。

1.2　估计器的互补性

在实际雷达探测场景中，通常需要快速获得K分布的参数估计值。因而，矩估计方法得到了广泛的应用。理论上，MoM估计可使用任意两个矩来估计形状参数和尺度参数。在式(3)中，令$r=2$，则

$$m_2 = E\left[x^2\right] = \frac{\Gamma(2)\Gamma(v+1)}{\Gamma(v)}\left(\frac{b}{v}\right) = b \tag{5}$$

可见，二阶矩只决定于尺度参数b。根据式(4)，得到尺度参数的估计为

$$\hat{b} = \hat{m}_2 \tag{6}$$

因此，矩估计的重点在于如何获得形状参数的估计。典型地，2-4阶矩估计法(MoM 2-4)、1-2阶矩估计法(MoM 1-2)的形状参数估计分别为[8]

$$\begin{aligned} &\hat{v} = \frac{2\left(\hat{m}_2\right)^2}{\hat{m}_4 - 2\left(\hat{m}_2\right)^2} \\ &\frac{\Gamma^2(\hat{v})\hat{v}}{\left(\Gamma(1.5)\Gamma(\hat{v}+0.5)\right)^2} = \frac{\hat{m}_2}{\left(\hat{m}_1\right)^2} \end{aligned} \tag{7}$$

式中，$\hat{m}_1$, $\hat{m}_2$, $\hat{m}_4$分别为一阶、二阶和四阶的样本矩。MoM 2-4具有解析表达式，计算简单。但由于存在伽马函数，MoM 1-2没有显式解析表达式，只能采用数值计算法。为了进一步应用多个矩的信息，发展了1-2-3阶矩估计法(MoM 1-2-3)，形状参数估计为

$$\hat{v} = \frac{3\hat{m}_1\hat{m}_2}{4\hat{m}_3 - 6\hat{m}_1\hat{m}_2} \tag{8}$$

此外，为了更好地拟合重杂波拖尾的杂波，文献[13]采用对数矩，称为$z^r\log(z)$期望法，其形状参数的表达式为

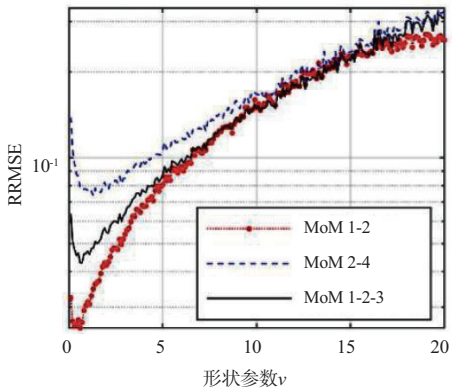
$$\frac{E\left[z^r\log(z)\right]}{E\left[z^r\right]} - E\left[\log(z)\right] = \psi(r+1) - \psi(1) + \psi(v+r) - \psi(v) \tag{9}$$

式中，$z = x^2$为杂波功率，$\psi(\cdot)$为Digamma函数。同样，$z^r\log(z)$期望法无法获得形状参数的显式解析表达式，只能借助数值法求解。

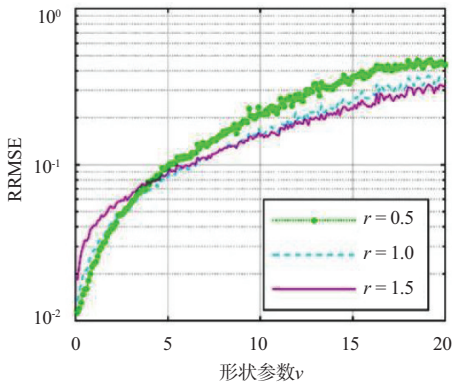
下面，讨论不同矩估计器的形状参数估计性能。由于尺度参数估计方法一致，采用形状参数的相对均方根误差(Relative Root Mean Square Error，RRMSE)作为衡量指标：

$$\mathrm{RRMSE}(v,\hat{v}) = \frac{\sqrt{E\left[\left(\hat{v}-v\right)^2\right]}}{v} \tag{10}$$

式中，v和$\hat{v}$分别为形状参数的真实值和估计值。图1给出了不同形状参数下的6种矩估计器的RRMSE曲线。在图1(a)中，低阶矩MoM 1-2的估计误差明显小于高阶矩MoM 2-4的估计误差。因而，实际矩估计器的矩阶数一般不会高于4。相对于MoM 1-2估计器，MoM 1-2-3估计器增加一个高阶矩后，在中等形状参数下有一定的性能提升。这意味着，合理地增加矩的个数是一种提升估计性能的潜在途径。在图1(b)中，当形状参数$v<2$时，对数矩的估计误差明显小于矩估计。但是，当形状参数$v>10$时，情况正好相反。此外，不同阶数r明显影响估计性能。当$r<1$时，适合估计具有小形状参数的杂波；当$r>1$时，适



(a) MoM估计器



(b) $z^r\log(z)$估计器

图1　不同估计器的性能互补性分析

合估计具有大形状参数的杂波。从整体上来看，MoM估计器在形状参数较大时的估计性能较好，而$z^t\log(z)$估计器在形状参数较小时的估计性能较好，两种估计器具有一定的互补性。因此，可以考虑将两者的优势结合起来，以降低整体的参数估计误差。

2 基于多维矩特征联合的参数估计方法

2.1 非线性参数估计框架

事实上，海杂波幅度分布的参数受到多种因素的影响[1-2]，主要包括雷达工作参数和实时海洋环境参数。雷达工作参数一般有距离分辨率、擦地角、极化方式等。在高分辨率小擦地角下，海杂波的幅度分布明显偏离高斯分布，其分布的形状参数变小。海洋环境参数包含风速、风向、海况、温度等多种气象因素。通常，气象环境是随着时间不断变化的。不同环境下采集的杂波参数特性是不同的。因此，一种可靠有效的手段是收集当前环境下的雷达回波后实时估计杂波的参数。此外，在时间上，海杂波时间序列呈现出短时平稳而长时非平稳特性。在空间上，海杂波具有局部均匀而整体非均匀的特性。然而，通常要求用于杂波参数估计的样本是独立同分布的，这意味着只有局部区域中短时的杂波序列可用于参数估计。幸运的是，高分辨雷达一旦开机后，可收集大量的局部短时杂波序列。因此，在大场景杂波环境下，需要对不同空间的杂波进行划分，实时估计杂波特性。

对于K分布杂波，需要估计形状参数和尺度参数。尺度参数反映了杂波的平均功率水平，可根据式(6)中的幅度二阶矩，获得其估计值。形状参数反映了杂波的非高斯特性，更能表征杂波的本质，并且其与最优相干检测器相关。因而，亟需解决的是K分布杂波的形状参数估计。

本质上，矩估计法是通过幅度分布的PDF建立了两个线性矩和形状参数之间的函数关系进行参数估计的。这种方法的核心在于找到函数关系，最终估计器是否存在显式解析表达式与采用的两个矩戚戚相关，比如式(7)中的MoM 2-4和MoM 1-2。启发于这种思路，本文提出了一种基于多维矩特征联合(Multidimensional Moment Feature Combination, MMFC)的估计器框架，通过丰富估计器可用信息以提升形状参数的估计性能，如图2所示。首先，提取了海杂波时间序列$c(n)$, $n=1,2,\cdots,N$中的矩特征，包含线性矩特征和对数矩特征两大类。其次，联合所有的特征，构成一个特征向量$\boldsymbol{\xi}$。在高维特征空间中，该向量凝聚了海杂波时间序列的显著特征。然后，K分布参数估计问题转换为高维空间中的一个非线性优化问题。该优化问题最终找到一个最优的非线性函数F，建立特征向量$\boldsymbol{\xi}$和形状参数v之间的函数映射关系，实现形状参数估计。最后，根据式(6)中的幅度二阶矩，获得尺度参数的估计。

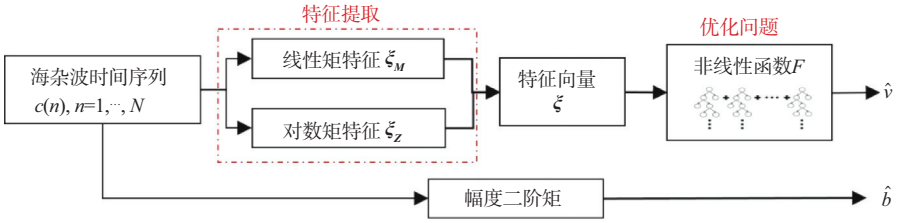


图2 MMFC估计的流程图

MMFC估计法的核心在于特征提取和非线性函数获取。由于K分布是双参数模型，式(3)中的原点矩往往由两个参数决定。但是，在实际参数估计中，为了降低估计误差，往往要求两个参数的估计是独立的。这意味着提取的特征必须只包含单个参数的信息。因而，特征设计的首要原则是必须只依赖于形状参数，完全独立于尺度参数。并且，如何从海杂波序列中提取有效的特征，这也是特征提取的难点和重点。在传统矩估计中，虽然用了两个矩，但实际上只有单个矩比值，属于一维优化问题。根据杂波分布的PDF，一般都可以寻找到非线性函数，且在特殊情况下具有显式解析

表达式，如MoM 2-4估计器。不同于传统矩估计，特征向量包含了多个矩的信息，需要在三维甚至更高维的特征空间中寻找非线性函数F。因此，以寻找与真实形状参数误差最小的F为目标，可构建如下的最优化问题：

$$\min_F \left\{ \left(F\left(\boldsymbol{\xi} | v\right) - v \right)^2 \right\},\ \text{s.t.}\ F\left(\boldsymbol{\xi} | v\right) > 0 \tag{11}$$

式中，$F(.|v)$表示只依赖于v的高维非线性函数且$F\left(\boldsymbol{\xi} | v\right) = \hat{v}$，高维空间维度取决于特征向量的维度。通常，为了提升估计性能，需要联合更多的矩特征，势必会增加非线性函数的维度。因此，式(11)的优化问题很难从理论上获得具体解析函数表达式。此外，基于多维特征的估计方法是一个开放的框架，可以自由地设计特征以及寻找非线性函数。同时，该框架不局限于K分布杂波，不依赖于杂波的概率分布，后续可用其他不同的杂波分布。

2.2　多维矩特征提取

在海杂波分布参数估计中，可用的信息全部来自于当前局部区域获得的海杂波时间序列。通常，从海杂波时间序列中获得统计矩，用于估计杂波分布的形状参数和尺度参数。本文将反映杂波特性的统计特征，统称为特征。相对于海杂波时间序列，特征凝聚了杂波特性，大大减少了数据的存储空间。在这种情况下，如何提取有效特征是估计的核心。目前，K分布下发展了大量的矩估计器，这表明统计矩是一种有效的特征。

那么，下一步是如何设计矩特征。需要指出的是，在式(3)中，除了二阶矩，其他任意r阶矩由尺度参数和形状参数同时决定。若直接使用矩特征，需要同时估计尺度参数和形状参数，因为这两者是相互关联的。这无疑大大增加了参数估计的困难。目前，有两种途径可以解决上述问题。第一种途径，先采用二阶矩估计尺度参数，然后将海杂波时间序列的功率除以尺度参数，获得归一化的海杂波时间序列，去除尺度参数影响[17]。这种方式易于寻找特征，但尺度参数的估计误差势必会进入形状参数的估计中，从而降低估计性能。第二种途径，设计不依赖于尺度参数的特征，比如MoM 2-4中四阶矩比二阶矩的平方，完全去除尺度参数和形状参数的相关性，使两个参数的估计误差不再相互影响，有利于提升估计性能。但这就需要设计者具有一定的先验知识去寻找具体的特征。

为了减少估计性能的损失，本文采用第二种途径。假设存在a, c, d阶的3个矩，矩比值定义为

$$\frac{m_a}{m_c m_d} = \frac{\Gamma(1+a/2)\Gamma(v+a/2)\Gamma(v)}{\Gamma(1+c/2)\Gamma(1+d/2)\Gamma(v+c/2)\Gamma(v+d/2)} \left(\sqrt{\frac{b}{v}} \right)^{a-c-d} \tag{12}$$

当$a = c + d$时，矩比值仅取决于形状参数v，独立于尺度参数b。

综合考虑现有的MoM估计器和式(12)特性，本文精心提取了7个矩比值，构成一个线性矩向量

$$\boldsymbol{\xi}_M = \left[\frac{m_1}{\left(m_{0.5}\right)^2}, \frac{m_2}{m_{0.5} m_{1.5}}, \frac{m_2}{\left(m_1\right)^2}, \frac{m_3}{m_{0.5} m_{2.5}}, \frac{m_3}{m_1 m_2}, \frac{m_4}{m_{0.5} m_{3.5}}, \frac{m_4}{\left(m_2\right)^2} \right]^{\mathrm{T}} \tag{13}$$

该线性矩向量凝聚了0.5阶、1.5阶、2.5阶、3.5阶分数矩和1阶、2阶、3阶、4阶矩的信息，丰富了信息的多样性，同时拓展了信息的维度。

同样地，根据式(9)，可证明对数矩比值只与形状参数有关。因此，本文提取了3个对数矩比值特征，构成一个对数矩向量

$$\boldsymbol{\xi}_Z = \left[\frac{E\left[z^{0.5} \log(z)\right]}{E\left[z^{0.5}\right]} - E\left[\log(z)\right], \frac{E\left[z \log(z)\right]}{E\left[z\right]} - E\left[\log(z)\right], \frac{E\left[z^{1.5} \log(z)\right]}{E\left[z^{1.5}\right]} - E\left[\log(z)\right] \right]^{\mathrm{T}} \tag{14}$$

考虑到对数矩和线性矩估计器两者的互补性，将两个矩向量联合，构建一个高维特征空间。相应地，得到一个10维的特征向量

$$\boldsymbol{\xi} = \left[\boldsymbol{\xi}_M ; \boldsymbol{\xi}_Z \right] \tag{15}$$

显然，该特征向量$\boldsymbol{\xi}$只与形状参数v有关。由于式(12)和式(9)中存在伽马函数和Digamma函数，理论上很难推出$\boldsymbol{\xi}$和v之间的显式解析表达式。

2.3　非线性函数获取

在高维特征空间中，式(11)的优化问题需要找到一个高维函数关系，实现从特征向量中估计出形状参数，且该函数必须保证估计值与真实值之间的误差最小。理论上，存在这样的高维函数

关系。但是，由于高维函数的非线性和复杂性，实际上很难给出这样的数学解析表达式。因此，本文引入模式识别里的GBDT算法[17]，获得非线性函数F的最优表达式。

GBDT算法是一种以分类回归树（Classification And Regression Tree, CART）为基学习器的集成学习算法。它能够学习输入向量中蕴含的多维信息，建立输入向量和真实值之间的复杂函数关系，在小样本数据上具有较好的估计性能。假设存在M棵CART树，每棵树的深度为D，级联构成GBDT模型。对于任意一棵树，按照损失函数的最小值进行迭代更新参数。损失函数定义为平方误差，即

$$L\left[f\left(\boldsymbol{\xi}\right),v\right]=\left(f\left(\boldsymbol{\xi}\right)-v\right)^2 \tag{16}$$

式中，$f(\cdot)$表示某棵CART树的函数。与式（11）中的优化问题一致，所有的CART树都寻找与真实形状参数值误差最小的函数。最终，将M棵CART树的结果累加获得最终的估计结果。当GBDT的输入为式（15）中的特征向量$\boldsymbol{\xi}$时，最终输出为

$$\hat{v}=f_M\left(\boldsymbol{\xi}\right) \tag{17}$$

式中，$f_M(\cdot)$表示第M棵CART树的参数函数。因此，借助GBDT算法，$f_M(\cdot)$可以作为式（11）中最终优化函数F的近似解。当然，后续也可以采用其他的算法，获得优化问题的最优解或近似最优解。

下面，问题的关键是如何获得GBDT模型的参数。这就需要从大量的海杂波时间序列中自主学习模型的最优参数。假设训练集Ω包含Q个特征向量$\boldsymbol{\xi}_i, i=1,2,\cdots,Q$和相对应的真实形状参数值$v_i, i=1,2,\cdots,Q$，简记为

$$\Omega=\left\{\left(\boldsymbol{\xi}_1,v_1\right),\left(\boldsymbol{\xi}_2,v_2\right),\ldots,\left(\boldsymbol{\xi}_Q,v_Q\right)\right\} \tag{18}$$

需要指出的是，训练集样本需要尽可能地覆盖不同形状参数，保证建立较为稳健的特征向量和形状参数之间的关系。此外，需要保证不同形状参数下的样本数量是足够的，能深入挖掘到特征向量所包含的形状参数信息。

图3给出了GBDT模型的训练过程，用于获取模型的最优权重参数，具体步骤如下。

步骤1：初始化

对于输入的训练集Ω，初始化参数为

$$f_0\left(\boldsymbol{\xi}\right)=\arg\min_{\gamma}\sum_{i=1}^{Q}L\left(v_i,\gamma\right) \tag{19}$$

式中，γ表示使损失函数达到最小的常数。

步骤2：生成第m棵CART树

对于第i个样本，计算损失函数的负梯度

$$r_{mi}=-\left[\frac{\partial L\left[v_i,f_{m-1}\left(\boldsymbol{\xi}_i\right)\right]}{\partial f_{m-1}\left(\boldsymbol{\xi}_i\right)}\right]=-2\left(v_i-f_{m-1}\left(\boldsymbol{\xi}_i\right)\right), i=1,2,\cdots,Q \tag{20}$$

由此，得到J个叶节点以及其对应的叶节点区域$R_{mj}, j=1,2,\cdots,J$。然后，计算第j个节点的最佳残差拟合值

$$\gamma_{mj}=\arg\min_{\gamma_m}\sum_{\boldsymbol{\xi}_i\in R_{mj}}\left[L\left(v_i,f_{m-1}\left(\boldsymbol{\xi}_i\right)\right)+\gamma_m\right], j=1,2,\cdots,J \tag{21}$$

最后，生成第m棵CART树的参数函数

$$f_m\left(\boldsymbol{\xi}\right)=f_{m-1}\left(\boldsymbol{\xi}\right)+\rho\sum_{j=1}^{J}\gamma_{mj}I\left(\boldsymbol{\xi}_i\in R_{mj}\right) \tag{22}$$

式中，$\rho\in[0,1]$为学习率，避免过拟合现象。当$\boldsymbol{\xi}_i\in R_{mj}$成立时，$I(\boldsymbol{\xi}_i\in R_{mj})$取值为1，反之为0。

步骤3：生成第M棵CART树

重复步骤2，直到生成第M棵CART树。最终，GBDT的输出为

$$f_M\left(\boldsymbol{\xi}\right)=f_0\left(\boldsymbol{\xi}\right)+\rho\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}I\left(\boldsymbol{\xi}\in R_{mj}\right) \tag{23}$$

至此，获得给定训练样本下的GBDT模型的最优权重参数，完成式（11）中的最优化问题求解。

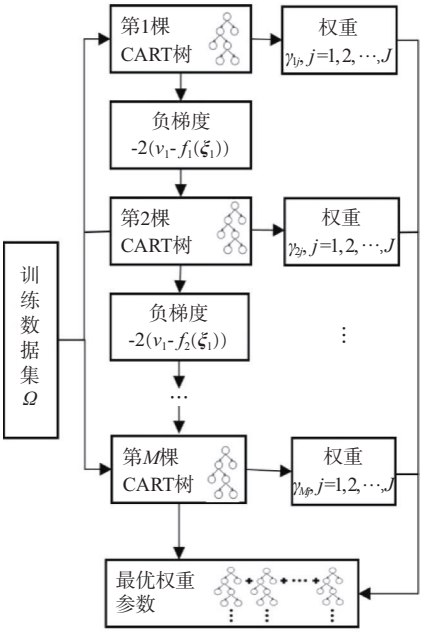


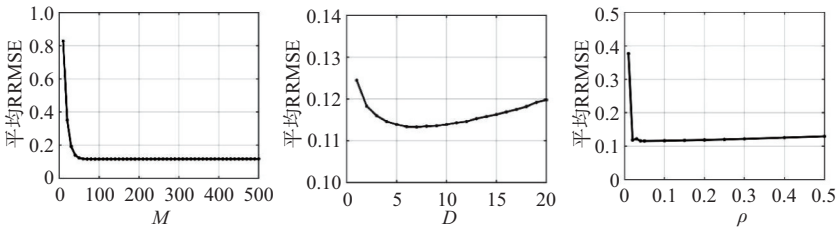
图3 GBDT模型训练过程

3　实验结果与性能分析

3.1　GBDT参数设置

事实上，形状参数的估计精度在很大程度上依赖于GBDT模型的参数。通常，GBDT模型主要分为结构参数和学习参数两大类。结构参数决定了GBDT模型的结构，比如CART树的数目M，每棵树的最大深度D，学习率ρ等。学习参数需要从训练集中学习获得，比如叶子节点数J，叶子节点对应的最佳残差拟合值等。当训练集更改时，学习参数必须重新从训练集中学习获得，即高度依赖于数据。然而，结构参数对数据的依赖度较小，可以提前设置。理论上，结构参数设置得越大，即联合更多和更深的基分类器，可以提高估计性能。但同时也带来了过拟合和计算量高的问题。因此，有必要研究如何设置合理的结构参数。

图4给出了不同结构参数下估计器的平均RRMSE曲线，其中平均RRMSE为形状参数从0到20之间的平均误差值。在图4(a)中，随着CART树的数目增加，平均RRMSE的值逐步减少。这表明多个弱分类器联合后，确实可以提高估计性能。但是，这种性能增加不是无止尽的。当$M>100$后，估计性能趋于稳定。在图4(b)中，随着树的深度D的增加，平均RRMSE的值呈现先下降后上升的趋势。这说明较大的深度可以提升估计性能，但过大的D值，导致GBDT算法存在过拟合的问题。类似地，在图4(c)中，随着学习率ρ的增加，导致GBDT算法过度依赖于先前的数据特性而性能降低。因此，综合考虑性能和计算量，在后续的实验中，设置$M=200$，$D=15$，$\rho=0.1$。



(a) CART树的数目　　(b) 树的最大深度　　(c) 学习率

图4　结构参数对估计器性能的影响

3.2　仿真数据结果

除了结构参数，GBDT的估计性能还受到学习参数的影响。GBDT模型需要根据当前的数据不断更新学习数据本身包含的形状参数和矩特征之间的关系。因此，需要合理地设置训练集，以保证GBDT模型估计性能的广泛性和优越性。考虑到K分布的特性，设置形状参数的范围从0.1到20，间隔为0.1均匀取值。由于矩特征不依赖于尺度参数，因而尺度参数可任意取，比如设定为1。对于每个给定的形状参数v，生成长度为10^4服从K分布的海杂波时间序列，按照式(13)和式(14)，提取对应的线性矩特征和对数矩特征，最终构成式(15)中的特征向量。该特征向量和真实的形状参数，构成一个样本。为了增加样本数目，该过程重复1 000次。因此，对于式(18)的训练集，总样本数为Q=200 000，用于训练GBDT模型的最优学习参数。

图5给出了5种估计器的RRMSE误差曲线。随着形状参数的增加，所有误差曲线呈现出增大的趋势。这是因为在大形状参数下K分布海杂波趋向于高斯分布，导致PDF曲线之间差异性变小。MoM 1-2估计器[8]性能整体优于MoM 2-4估计器[8]，再次验证了低阶矩在估计性能上的优势。然而，$z\log z$估计器[12]性能与MoM 1-2估计器[8]接近，并且在形状参数$v<2$时，其估计误差最小。为了提升估计器性能，TPE估计器[15]的一个分位点设置为0.9。该估计器在形状参数$v<2$时具有较好的性能，但在形状参数较大时，性能损失严重。这是由其估计器本身特性决定的。不同于矩估计，分位点估计只采用了3个分位点的信息，因而信息量非常有限。在大形状参数情况下，存在一定的性能损失。但这些有限的信息，在异常环境下，具有较好的抗噪声性能。从整体上来说，本文所提的MMFC估计器的平均RRMSE值最小，具有最佳的

估计性能。相较于MoM 1-2[8]和$z\log(z)$估计器[12]，MMFC估计器不仅在小形状参数($v<2$)时获得更精准的估计结果，并且大大降低了在大形状参数($v>10$)下的估计误差。这种性能优势主要得益于多个线性矩特征和对数矩特征的联合使用，集成了线性矩估计和对数矩估计的优势。

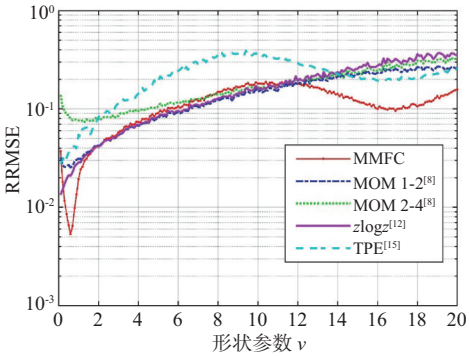


图5 不同形状参数下估计器性能对比

为了更加量化性能对比结果，表1列出了不同形状参数区间下，5种估计器的平均RRMSE值。从整体来看，即$\nu\in(0,20]$，所提估计器的平均RRMSE最小。相对于MoM 1-2估计器[8]、MoM 2-4估计器[8]、$z\log z$估计器[12]和TPE估计器[15]，所提MMFC估计器分别有23%，34%，33%，48%的估计性能提升。当形状参数非常小时，即$\nu\in(0,0.5]$，K分布海杂波完全偏离高斯分布，其概率密度函数表现出严重的重拖尾现象。通常，这种杂波出现在高海况、低擦地角条件下。此时，MMFC估计器仅次于最优的$z\log z$估计器，具有部分性能损失。当形状参数较小时，即$\nu\in(0.5,2]$，此时K分布海杂波具有重拖尾现象，出现在中高海况下，这也是大多数雷达工作的环境。相对次最优$z\log z$估计器[12]，MMFC估计器的性能最优，具有19%的性能提升空间。当形状参数中等大小时，即$\nu\in(2,10]$，此时K分布海杂波较为平稳，出现在低中海况下。除了TPE估计器[15]存在较大的性能损失，剩下的4种估计器性能接近。当形状参数较大时，即$\nu\in(10,20]$，K分布海杂波接近高斯分布，往往出现在低海况下或远海区。所提估计器能够充分利用多个矩的信息，相对于次优MoM 1-2估计器具有36%的性能提升。综上，所提MMFC估计器具有最佳的综合性能，有望应用于不同的杂波环境下。

表1 不同形状参数区间下的平均RRMSE值

v区间	MMFC	MoM 1-2[8]	MoM 2-4[8]	$z\log z$[12]	TPE[15]
(0,0.5]	**0.016 40**	0.027 6	0.102 7	0.017 3	0.032 0
(0.5,2]	**0.026 3**	0.034 4	0.077 2	0.032 4	0.057 7
(2,10]	0.110 6	**0.095 5**	0.119 2	0.099 0	0.241 3
(10,20]	**0.139 3**	0.219 0	0.241 8	0.263 3	0.244 6
(0, 20]	**0.116 3**	0.151 0	0.177 0	0.174 1	0.224 0

注：黑色加粗表示每一行的最小值。

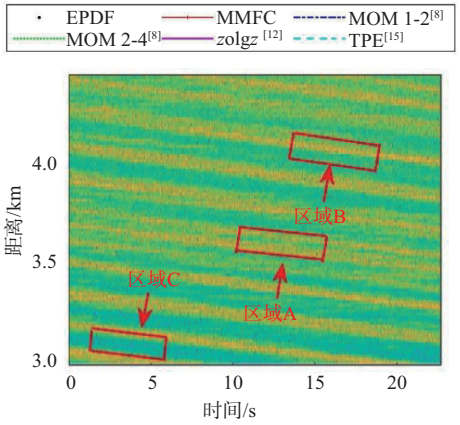
3.3 实测数据结果

在仿真实验中，RRMSE可以用来准确地衡量形状参数的估计误差。但是，该指标要求已知真实形状参数，这在实际探测环境下是不可能实现的。因此，当用实测数据评估性能时，采用KS距离(Kolmogorov-Smirnov Distance, KSD)

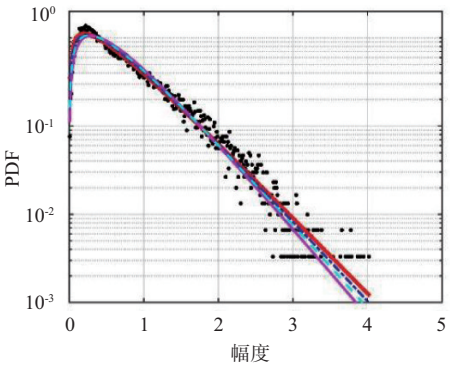
$$\mathrm{KSD}(\nu,b;\hat{\nu},\hat{b})=\max_{x\in(0,+\infty)}\left\{\left|C(x;\nu,b)-C(x;\hat{\nu},\hat{b})\right|\right\} \tag{24}$$

式中，$C(\cdot)$表示K分布的CDF。在实际杂波环境中，K分布杂波的所有真实参数都是未知的。此时，杂波序列真实的CDF通常用经验CDF代替。因此，KSD反映了由参数估计值得到的经验CDF与真实CDF之间的差值大小。KSD值越小，表明拟合误差越小，估计精度越高。

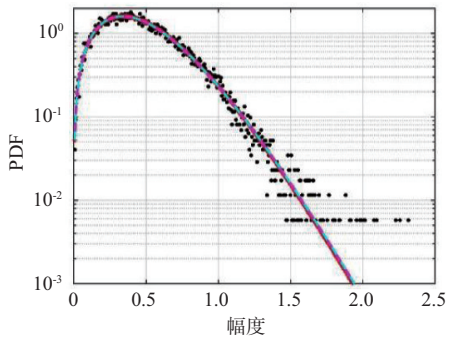
实测数据来自X波段Fynmeet雷达于2006年8月1日采集的数据[18]，文件名为TFC15_004。该雷达架设在岸边，工作在VV极化下的驻留模式，距离分辨率为15 m，脉冲重复频率为5 kHz，探测范围距离岸边3 km到4.4 km。图6(a)画出了时间-距离的杂波幅度图，呈现出明暗交替的“斑马”条



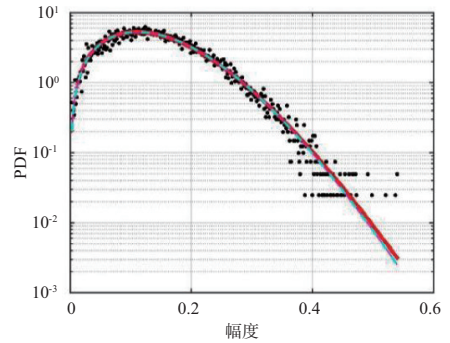
(a) 实测数据幅度图



(b) 区域A



(c) 区域B



(d) 区域C

图6　Fynmeet数据上5种估计器的拟合结果

纹现象。颜色越亮,表明杂波幅度值越高,反之杂波幅度越低。区域A包含大量的亮点,杂波幅度起伏较大。区域B同时包含亮点和暗点,杂波幅度起伏中等。区域C包含大量的暗点,杂波幅度起伏较小。这3个区域作为不同杂波特性的典型区域,用于验证不同估计器的性能,结果如图6(b)~(d)所示。

表2列出了各个估计器的形状参数估计值和KSD值。在区域A内,形状参数v的值较小,杂波具有很强的非高斯性。MMFC估计器的KSD值最小,其次是$z\log z$估计器[12],剩下3种估计器都遭受了较为严重的性能损失。在区域B内,形状参数的值中等,杂波幅度较为均匀。因而,5种估计器的KSD值较为接近,但MMFC估计器的KSD值仍是最小的。在区域C内,形状参数v的值较大,杂波趋向于高斯分布。MMFC估计器的估计精度最高,$z\log(z)$估计器[12]的估计误差最大。由此可见,5种估计器在实测数据3个区域内的参数估计结果和仿真实验的结果一致。并且,所提MMFC估计器在3个区域内都实现了最佳的拟合,这表明MMFC估计器具有较好的稳健性,能适用于不同的K分布杂波环境。

表2　3个典型区域杂波拟合值

估计方法	区域A	区域B	区域C
MMFC估计	**v=0.91** **KSD=0.021 8**	**v=6.48** **KSD=0.006 2**	**v=19.14** **KSD=0.008 1**
MoM 1-2估计[8]	v=0.80 KSD=0.031 3	v=6.20 KSD=0.006 3	v=20.00 KSD=0.012 7
MoM 2-4估计[8]	v=1.13 KSD=0.054 7	v=6.12 KSD=0.006 6	v=20.00 KSD=0.012 7
$z\log(z)$估计[12]	v=0.77 KSD=0.029 0	v=6.13 KSD=0.006 6	v=15.67 KSD=0.015 7
TPE估计[15]	v=1.01 KSD=0.038 1	v=5.80 KSD=0.007 2	v=18.61 KSD=0.008 8

注:黑色加粗表示每一列的最小值。

4　结束语

本文主要研究了K分布杂波的参数估计方法。不同于基于概率密度函数的矩估计法,将参数估计问题转换为高维特征空间中的非线性最优化问题。因此,提出了一种基于多维矩特征联合的估计方法,丰富了参数估计的信息量。所提出的估计器具有稳健的估计性能,能运用于实际复杂杂波环境。此外,本文给出的是一种新的参数估计框架,后续可根据杂波特性深入研究特征提取和非线性函数获取的问题。

参考文献:

[1] HUANG Penghui, ZOU Zihao, XIA Xianggen, et al. A Statistical Model Based on Modified Generalized-K Distribution for Sea Clutter[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19:1-5.

[2] WARD K, TOUGH R J A, WATTS S. Sea Clutter: Scatter-

ing, the K Distribution and Radar Performance[M]. UK: Institute of Engineering Technology, 2013:101-134.

[3] 张坤,水鹏朗,王光辉.相参雷达K分布海杂波背景下非相干积累恒虚警检测方法[J].电子与信息学报,2020,42(7):1627-1635.

[4] WANG Zhihang, HE Zishu, HE Qin, et al. Persymmetric Adaptive Target Detection with Dual-Polarization in Compound Gaussian Sea Clutter with Inverse Gamma Texture [J]. IEEE Trans on Geoscience and Remote Sensing, 2022, 60:1-17.

[5] WANG Jiaheng, WANG Zhihang, HE Zishu, et al. GLRT-Based Polarimetric Detection in Compound-Gaussian Sea Clutter with Inverse-Gaussian Texture[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19:1-5.

[6] XUE Jian, LIU Jun, XU Shuwen, et al. Adaptive Detection of Radar Targets in Heavy-Tailed Sea Clutter with Lognormal Texture[J]. IEEE Trans on Geoscience and Remote Sensing, 2022, 60:1-11.

[7] 胡文琳,王永良,王首勇.基于矩方法的K分布杂波参数估计研究[J].雷达科学与技术,2007,5(3):194-198.

[8] 石志广,周剑雄,付强.K分布海杂波参数估计方法研究[J].信号处理,2007,23(3):420-424.

[9] ISKANDER D R, ZOUIR A M. Estimation of the Parameters of the K-Distribution Using Higher Order and Fractional Moments[J]. IEEE Trans on Aerospace and Electronic Systems, 1999, 35(4):1453-1457.

[10] LI Zhiyuan, SHUI Penglang, ZOU Pengjia. Order-Adaptive Fractional-Order Moment-Based Estimation of Shape Parameter of K-Distribution[C]//2019 IEEE 4th International Conference on Signal and Image Processing,Wuxi, China:IEEE, 2019:597-600.

[11] 杨立儒,刘永祥,杨威,等.基于原点矩偏导的K分布杂波参数估计[J].信号处理,2022,38(11):2424-2431.

[12] BLACKNELL D, TOUGH R J A. Parameter Estimation for the K-Distribution Based on [zlog(z)] [J]. IEE Proceedings of Radar, Sonar and Navigation, 2001, 148(6): 309-312.

[13] 胡文琳,王永良,王首勇.基于z'log(z)期望的K分布参数估计[J].电子与信息学报,2008,30(1):203-205.

[14] YU Han, SHUI Penglang, ZENG Weiliang, et al. Bipercentile Parameter Estimators of Bias Reduction Generalised Pareto Clutter Model[J]. IET Radar Sonar and Navigation, 2020, 14(7):1105-1112.

[15] YU Han, SHUI Penglang, LU Kai. Outlier-Robust Tri-Percentile Parameter Estimation of K-Distributions[J]. Signal Processing, 2021, 181:1-13.

[16] FERNANDEZ J R M, VIDAL J. Improved Shape Parameter Estimation in K Clutter with Neural Networks and Deep Learning[J]. International Journal of Interactive Multimedia and Artificial Intelligence, 2016,3(7):96-103.

[17] ZHANG Zhendong, JUNG Cheolkon. GBDT-MO: Gradient-Boosted Decision Trees for Multiple Outputs[J]. IEEE Trans on Neural Networks and Learning Systems, 2021, 32(7):3156-3167.

[18] CONTE E, DE MAIO A, GALDI C. Statistical Analysis of Real Clutter at Different Range Resolutions[J].IEEE Trans on Aerospace and Electronic Systems, 2004, 40(3): 903-918.

作者简介:

施赛楠 女,博士,副教授,主要研究方向为信号处理和微弱目标检测。

高季娟 女,硕士研究生,主要研究方向为杂波建模和目标检测。

李东宸 男,高级工程师,主要研究方向为图像处理和无人机任务规划。

(上接第76页)

mitters[D]. USA: Naval Postgraduate School, 2015.

作者简介:

司梦祥 男,硕士研究生,主要研究方向为非合作多基地雷达目标探测。

杨风暴 男,博士,教授,主要研究方向为多源图像融合。

宋 杰 男,博士,教授,主要研究方向为无源双基地雷达目标检测。