

控制与决策

Control and Decision

K近邻优化估计的SAR图像建模与目标检测算法

彭书娟, 曲长文, 李健伟

引用本文:

彭书娟, 曲长文, 李健伟. K近邻优化估计的SAR图像建模与目标检测算法[J]. 控制与决策, 2020, 35(9): 2199–2206.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0051>

您可能感兴趣的其他文章

Articles you may be interested in

一种基于改进的暗通道先验的运动模糊核估计方法

A method of motion blurry kernel estimation based on improved dark channel prior
控制与决策. 2020, 35(7): 1667–1673 <https://doi.org/10.13195/j.kzyjc.2018.1687>

杂波协方差矩阵结构的融合估计方法

A fusion estimation method for covariance matrix structure of clutter
控制与决策. 2019, 34(9): 2010–2014 <https://doi.org/10.13195/j.kzyjc.2018.0052>

基于级联CNN的SAR图像舰船目标检测算法

A ship detection method based on cascade CNN in SAR images
控制与决策. 2019, 34(10): 2191–2197 <https://doi.org/10.13195/j.kzyjc.2018.0168>

复合高斯背景下基于最优控制参数的自适应检测器

Adaptive detector based on optimal control parameter in compound-Gaussian background
控制与决策. 2018, 33(8): 1532–1536 <https://doi.org/10.13195/j.kzyjc.2017.0505>

高分辨TerraSAR-X图像舰船目标几何参数提取方法

Geometric parameters extraction method of ship target in high resolution TerraSAR-X image
控制与决策. 2015(6): 1135–1138 <https://doi.org/10.13195/j.kzyjc.2014.0398>

K 近邻优化估计的 SAR 图像建模与目标检测算法

彭书娟^{1†}, 曲长文², 李健伟¹

(1. 海军航空大学 研究生三队, 山东 烟台 264001; 2. 海军航空大学, 山东 烟台 264001)

摘要: 在非均匀杂波环境下的合成孔径雷达 (synthetic aperture radar, SAR) 图像背景建模问题中, 针对非参量建模算法 Parzen 窗估计严重依赖于窗宽设置及最优核函数选择的问题, 提出一种基于 K 近邻优化的概率密度函数估计算法, 解决因固定近邻数而导致估计不准确甚至不能估计的问题. 该算法不需要图像的任何先验知识, 且无需考虑窗宽的设置及最优核函数的选择问题. 与 Parzen 窗估计、 K 分布和 G^0 分布的对比实验表明, 所提出的 K 近邻优化估计算法可以实现对单峰、多峰甚至不规则图像数据的准确建模, 优于 K 分布和 G^0 分布; 同时, 对图像首尾数据的处理优于 Parzen 窗估计. 实验结果验证了所提出方法对 SAR 图像杂波建模的精确性、鲁棒性和简便性, 以及全局恒虚警率目标检测的有效性.

关键词: SAR 图像统计建模; K 近邻优化估计; 平均区域体积; 核密度估计; 恒虚警率; 目标检测

中图分类号: TN957.52

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0051

引用格式: 彭书娟, 曲长文, 李健伟. K 近邻优化估计的 SAR 图像建模与目标检测算法 [J]. 控制与决策, 2020, 35(9): 2199-2206.

K nearest neighbors optimized estimation algorithm for SAR image statistical modeling and target detection

PENG Shu-juan^{1†}, QU Chang-wen², LI Jian-wei¹

(1. The 3rd Graduate Student Team, Naval Aviation University, Yantai 264001, China; 2. Naval Aviation University, Yantai 264001, China)

Abstract: Clutter statistical modeling for synthetic aperture radar (SAR) measurements in nonhomogeneous clutter environments is a complex and challenging task. In order to overcome the dependence of the Parzen window estimation in the non-parametric modeling algorithm for setting the window width and picking the optimal kernel function, a new K nearest neighbors optimized estimation algorithm is proposed. The proposed algorithm solves the problem of that the estimation is inaccurate or cannot be estimated due to the fixed neighbor number. Meanwhile, the algorithm neither requires any prior knowledge nor needs to set the window width and select the optimal kernel function. Compared with Parzen window estimation, K distribution and G^0 distribution, the proposed K -nearest neighbor optimization estimation algorithm can accurately model single-peak, multi-peak and even irregular image data, which is better than K distribution and G^0 distribution. And the processing of the first and last data of the image is better than the Parzen window estimation. The experimental results verify the accuracy, robustness and simplicity of the proposed method for background clutter modeling of SAR images, and the effectiveness of global constant false alarm rate target detection.

Keywords: SAR image; K nearest neighbors optimized estimation; average regional volume; kernel density estimation; CFAR; target detection

0 引言

合成孔径雷达 (synthetic aperture radar, SAR) 图像目标检测作为监测与导航方向的重要组成部分, 一直受到民用与军事领域的广泛关注, SAR 图像杂波统计建模的优劣直接影响后续检测与识别等应用. 近年来, 学者们针对 SAR 图像相干斑噪声模型和

乘积噪声模型提出了一系列参量统计分布模型^[1-2]及其改进型^[3-5], 以实现 SAR 图像杂波统计建模. 参量模型越复杂则参数越多^[4-6], 使得参数估计成为一个不容忽视的难点问题^[4,6]; 而实测 SAR 图像杂波概率密度分布大多呈现不规则甚至多峰特性^[7], 导致只具有单峰拟合性能的参量模型对于真实杂波的

收稿日期: 2019-01-09; 修回日期: 2019-05-02.

基金项目: 国家自然科学基金项目 (615714541).

责任编辑: 程龙.

[†]通讯作者. E-mail: peng_shu_juan@163.com.

统计建模失效. 鉴于自动阈值异常检测的恒虚警率 (constant false alarm rate, CFAR)^[1,4,5,7] 算法依赖于杂波的精确统计, 力求获得先验知识未知下的杂波精确建模, 基于数据驱动的非参量模型^[7-14] 引起了越来越多研究者的关注.

非参量模型采用样本推断总体条件概率密度估计分布, 代表方法有核密度估计^[7-14] 和 K 近邻估计^[9,15-17] 等. Parzen 窗法是一种典型的基于高斯核函数的核密度估计方法, 近几年来, 研究者们针对 Parzen 窗法, 在窗函数加权、引入预处理特征信息进行算法融合^[11-12] 等方面加以改进, 而窗宽的估算与核函数的选取仍凭借经验^[13], 在考虑全局优劣的渐近积分均方误差准则 (asymptotic mean integrated squared error, AMISE) 下, 高斯核并非最优核^[8-9], 而被估算的最优窗宽直接影响着概率密度函数的平滑度. 理论上 K 近邻估计法^[16] 在消除 Parzen 窗法估计不稳定性的同时, 不会因窗宽的选择导致估计过于平坦而严重失真; 与 Parzen 窗估计采用的固定窗宽相反, 为保证落入带宽内部的样本数一样, K 近邻估计法的带宽是动态可变化的. K 近邻估计在聚类分类^[15,17] 中应用广泛, 在纯粹概率密度函数的估计中却鲜有应用.

本文首先分析传统 K 近邻估计法统计建模的基本思想, 指出其不足并提出优化算法; 然后应用优化算法对 SAR 图像杂波进行统计建模; 最后利用确定的全局阈值实现目标的检测.

1 K 近邻优化的 SAR 图像杂波统计建模

1.1 K 近邻估计法基本思想

设 x_1, x_2, \dots, x_N 是从某个具有分布 P 和未知概率密度 $p(x)$ 的样本总体中抽出的 N 个样本观测值, 根据总样本选定一个与 N 有关的参数 $K(N)$. 求 x 处的概率密度估计时, 通过调整包含 x 的区域体积内的样本数恰好为 $K(N)$, 得到最小区域体积 $V_N(x)$, 将

$$\hat{p}_N(x) = \frac{K(N)}{NV_N(x)} \quad (1)$$

作为 $p(x)$ 的估计. $V_N(x)$ 会随着样本密度的高低而自动减小或增大, 能够较好地兼顾在高密度区域估计的分辨率和在低密度区域估计的连续性. 只需要设置 $K(N) = k_1\sqrt{N} \geq 1 (k_1 \in \mathbb{Z})$, 使 $V_N(x)$ 适应于 $K(N)$ 的变化, 而不是取决于 N , 避免估计出现严重失真. 由于 N 一定, k_1 也是常系数, $K(N)$ 是固定的, $p(x)$ 只根据 $V_N(x)$ 的变化而改变.

1.2 基于 K 近邻优化估计的模型实现

在含有非均匀杂波的实测 SAR 图像实验中发现, 各个灰度值出现的频数相差较大, 达到甚至超过

10^3 的数量级, 依据传统 K 近邻估计法将所有待观测点的近邻数设为与 N 相关的足够大的定值不再可行, 会导致计算严重耗时甚至不能估计. 同时待观测点的近邻数 $K(N)$ 也并非越多越好, 在这种情况下, $K(N)$ 的设置不能只由 k_1 和 N 决定.

首先将近邻数设成随着各观测点出现的频数而自适应改变的参变量, 利用下式确定某观测值 x_i 的近邻数 K_N^i :

$$K_N^i = k_c \cdot \lfloor \sqrt{\text{Freq}(x_i)} \rfloor. \quad (2)$$

其中: k_c 为设置的固定常系数, $\text{Freq}(x_i)$ 为观测值 x_i 出现的频数, $\lfloor \cdot \rfloor$ 为向下取整. 这样 K_N^i 即随着 $\text{Freq}(x_i)$ 的变化而自适应地改变, 定义平均区域体积如定义 1 所述.

定义 1 以观测点 x_i 为中心, 计算所有样本点到该观测点的欧氏距离 $D_{ij} = \{d_{i1}, d_{i2}, \dots, d_{iN}\}_{j=1}^N$, 并将 D_{ij} 按照升序排列得到 $D_{ij}^* = \{d_{i1}^*, d_{i2}^*, \dots, d_{iK_N^i}^*, \dots, d_{iN}^*\}_{j=1}^N$. 取距离 x_i 最近的前 K_N^i 个欧氏距离, 则包含观测点 x_i 的 K_N^i 近邻的平均区域体积为

$$V_N(x_i) = c \cdot \sum_{j=1}^{K_N^i} d_{ij}^* / K_N^i. \quad (3)$$

其中: c 为调整常系数, $\sum_{j=1}^{K_N^i} d_{ij}^*$ 为到观测点 x_i 的 K_N^i 个近邻点的欧氏距离的和. 优化后的 K 近邻估计的概率密度函数如定义 2 所述.

定义 2 当观测样本的数量 N 足够大时, 选择观测点 x_i 的近邻数 $K_N^i = k_c \cdot \lfloor \sqrt{\text{Freq}(x_i)} \rfloor$, 并同时满足条件 $\lim_{N \rightarrow \infty} V_N(x_i) = 0$, $\lim_{N \rightarrow \infty} K_N^i = \infty$, $\lim_{N \rightarrow \infty} K_N^i / N = 0$, x_i 处的概率密度 $p(x_i)$ 的估计 $\hat{p}(x_i)$ 定义为

$$\hat{p}(x_i) = \frac{K_N^i}{NV_N(x_i)} = (K_N^i)^2 / \left(c \cdot N \sum_{j=1}^{K_N^i} d_{ij}^* \right) = \frac{k_c^2}{c \cdot N} \cdot \text{Freq}(x_i) / \sum_{j=1}^{K_N^i} d_{ij}^*, \quad (4)$$

且有 $\lim_{N \rightarrow \infty} \hat{p}(x_i) = p(x_i)$.

由式 (4) 可知, 观测点 x_i 到其 K_N^i 个近邻点的距离越小, 则 $\hat{p}(x_i)$ 越大. 对比式 (1), 本文 $\hat{p}(x_i)$ 的定义是将概率密度的计算范围从整个观测样本集缩减为观测点 x_i 的 K_N^i 个近邻, 不同于固定值 $K(N)$, K_N^i 随着观测点出现频数的变化而改变.

1.3 K 近邻优化估计算法理论证明

求证:

$$1) \lim_{N \rightarrow \infty} K_N^i = \infty, \lim_{N \rightarrow \infty} K_N^i / N = 0, \lim_{N \rightarrow \infty} V_N(x_i) = 0.$$

$$2) \lim_{N \rightarrow \infty} \hat{p}(x_i) = p(x_i).$$

证明 1) 当观测样本的数量 N 足够大, 即 $N \rightarrow \infty$ 时, \exists 任意某观测值 x_i , 其出现的频数 $\text{Freq}(x_i) \rightarrow \infty$. 因为 $K_N^i = k_c \cdot \lfloor \sqrt{\text{Freq}(x_i)} \rfloor$, 所以有 $\lim_{N \rightarrow \infty} K_N^i = \infty$, 且必有 $K_N^i \ll N$, 进而 $\lim_{N \rightarrow \infty} K_N^i/N = 0$.

由式(3)

$$V_N(x_i) = c \cdot \sum_{j=1}^{K_N^i} d_{ij}^*/K_N^i,$$

分子为到 x_i 的 K_N^i 个近邻点欧氏距离的和, 当 $\lim_{N \rightarrow \infty} K_N^i = \infty$ 时, 有 $c \cdot \sum_{j=1}^{K_N^i} d_{ij}^* \ll K_N^i$, 故 $\lim_{N \rightarrow \infty} V_N(x_i) = 0$.

2) 求 x 处的概率密度估计时, 平均区域体积会随着样本密度的高低而自动减小或增大, 当 $N \rightarrow \infty$ 时, 通过调整包含 x 的区域体积内的样本数恰好为 K_N^i , 有

$$\lim_{N \rightarrow \infty} K_N^i = \lim_{N \rightarrow \infty} k_c \cdot \lfloor \sqrt{\text{Freq}(x_i)} \rfloor = \infty.$$

得到最小区域体积 $V_N(x_i)$, 有 $\lim_{N \rightarrow \infty} V_N(x_i) = 0$, 此时可以将式(4) $\hat{p}(x_i)$ 作为 $p(x_i)$ 的估计, 则 $\lim_{N \rightarrow \infty} \hat{p}(x_i) = p(x_i)$ 得证. \square

1.4 K 近邻优化估计算法步骤

利用 K 近邻优化估计概率密度函数, 首先统计图像中各个观测点出现的频数, 设置 k_c , 确定 K_N^i ; 然后求到该观测点的前 K_N^i 个欧氏距离的和, 得到平均区域体积; 最后由式(4)计算 $\hat{p}(x)$. K 近邻优化估计概率密度函数算法的伪代码如下所示.

算法1 K 近邻优化估计概率密度函数算法.

输入: 观测点集 X (其第 i 个元素对应第 i 个观测点), 近邻数 K_N^i , 最小观测值 x_{\min} , 最大观测值 x_{\max} , 采样步长 x_{step} ;

中间值: 欧氏距离 D_{ij} 和包含 K_N^i 个近邻的最小区域体积 $V_N(x_i)$;

输出: $[p_x(k)] = K(X, K_N^i, x_{\min}, x_{\max}, x_{\text{step}})$.

while $x < x_{\min} + x_{\text{step}}/2$

for $i = 1 : N$

$d(i) = \text{sqrt}(\text{sum}((x - X(i)).^2));$ %sum 求和, sqrt

求开方

end

$d = \text{sort}(d, 'ascend');$ % 将欧氏距离 d 按升序排列

for $j = 1 : K_N^i$

$r = r + d(j);$

end

$V_N(x_i) = (cr)/K_N^i;$

$p_x(k) = K_N^i/(NV_N(x_i));$

$k = k + 1; x = x + x_{\text{step}};$

end

2 基于 K 近邻优化估计的全局CFAR检测

用 K 近邻优化估计对 SAR 图像进行杂波统计建模后, 利用获得的全局检测阈值对图像中的每一个像素进行比较判别, 实现对强度较亮的“异常像素”的检测.

2.1 检测阈值的求取

基于 K 近邻优化估计的 SAR 图像概率密度估计为 $\hat{p}_N(x)$, 设定的虚警率 P_{fa} 与阈值门限 T 确定如下:

$$1 - P_{fa} = \sum_{t=0}^T \hat{p}_N(t). \quad (5)$$

结合式(4), 有

$$P_{fa} = 1 - \sum_{t=0}^T \hat{p}_N(t) = 1 - \frac{k_c^2}{c \cdot N} \sum_{t=0}^T \text{Freq}(t) / \sum_{j=1}^{K_N(x)} d_j^*. \quad (6)$$

其中: d_j^* 为所有样本点到观测点 x 按升序排列的欧氏距离, $K_N(x)$ 为观测点 x 的近邻数. SAR 图像 CFAR 检测的全局阈值 T 可由式(6)确定. 对于每一个像素 x , 若其强度值 I 大于等于阈值 T , 则为目标, 否则判定为背景杂波, 遍历整幅图像即可实现对疑似目标的检测.

2.2 检测阈值的数值计算方法

由于每个观测点的近邻数都是变化的, 由式(6)求阈值 T 得不到解析表达式, 必须按点分段, 寻求数值计算. 按照文献[5, 7]的处理思路, 不妨设 $\hat{p}_N(x)$ 的累积分布函数为 $P_N(x)$, 其相互关系为

$$P_N(x) = \sum_{t=0}^x \hat{p}_N(t). \quad (7)$$

式(6)可进一步表示为

$$P_{fa} = 1 - P_N(T). \quad (8)$$

设给定的 SAR 强度图像为 I_{img} , 像素强度集合可表示为 $I_{\text{img}} = \{I_1, I_2, \dots, I_N\}$, 其中 N 为图像像素个数. 令 $I_{\max} = \max\{I_1, I_2, \dots, I_N\}$ 表示像素强度集合的最大值, 将区间 $[0, I_{\max}]$ 按步长 γ 划分为 I_{\max}/γ 个等间隔区间, 并令 $I_{\text{imgnew}}\{I'_1, I'_2, \dots, I'_n\}$, 其中 I'_κ ($1 \leq \kappa \leq n$) 为区间的节点, 则总的节点数 $n = I_{\max}/\gamma + 1$. 显然, 检测阈值门限 $T \in [0, I_{\max}]$. 由式(7)分别计算累积分布函数 $P_N(I'_\kappa)$, $1 \leq \kappa \leq n$, 得到

$$P_N(I'_\kappa) = \sum_{t=I'_1}^{I'_\kappa} \hat{p}_N(t) = \frac{k_c^2}{c \cdot N} \sum_{t=I'_1}^{I'_\kappa} \frac{\text{Freq}(t)}{\sum_{j=1}^{K_N(t)} d_j^*}. \quad (9)$$

进一步, $P_N(I'_\kappa)$ 可以用 $I'_\kappa/\gamma = \kappa - 1$ 个近似等高梯形的面积求和得到

$$P_N(I'_\kappa) \doteq \begin{cases} \sum_{k=1}^{\kappa-1} \frac{(\hat{p}_N(I'_k) + \hat{p}_N(I'_{k+1})) \cdot \gamma}{2}, & 2 \leq \kappa \leq n; \\ 0, & \kappa = 1. \end{cases} \quad (10)$$

由式(9)和(10)可以得到 $P_N(I'_\kappa)$ 的迭代表式为

$$P_N(I'_\kappa) \doteq P_N(I'_{\kappa-1}) + \frac{(\hat{p}_N(I'_{\kappa-1}) + \hat{p}_N(I'_\kappa)) \cdot \gamma}{2}, \quad 2 \leq \kappa \leq n. \quad (11)$$

由文献[7], 令

$$P' = 1 - P_N = 1 - \{P_N(I'_1), P_N(I'_2), \dots, P_N(I'_n)\} = \{P'(1), P'(2), \dots, P'(n)\}, \quad (12)$$

由于 $P_N(\cdot)$ 的单调递增性, 可知 $P'(\cdot)$ 是单调递减的. 对于给定的虚警概率 P_{fa} , 当满足 $P'(k) < P_{fa} < P'(k - 1)$ 时, 不难得到阈值 $T = I'_k$, 至此检测门限值确定.

2.3 复杂度分析

K 近邻优化估计的 CFAR 检测算法的时间复杂度包含两部分, 一是估计概率密度函数, 二是进行 CFAR 目标检测. 假设图像有 N 个像素值, 设任意一个观测值 x_i 的近邻数为 K_N^i , 则由算法1得到 K 近邻优化估计的时间复杂度为 $O(N) + O(K_N^i)$. 因 $K_N^i \ll N$, 故其时间复杂度仅与 N 有关, 即 $O(N)$. 求取全局阈值后, 遍历图像中的每一个像素点进行比较判决, 其时间复杂度仍为 $O(N)$. 所以算法总的时间复杂度仍与图像尺寸有关, 为 $O(N)$.

杂波统计建模与目标检测流程如图1所示.

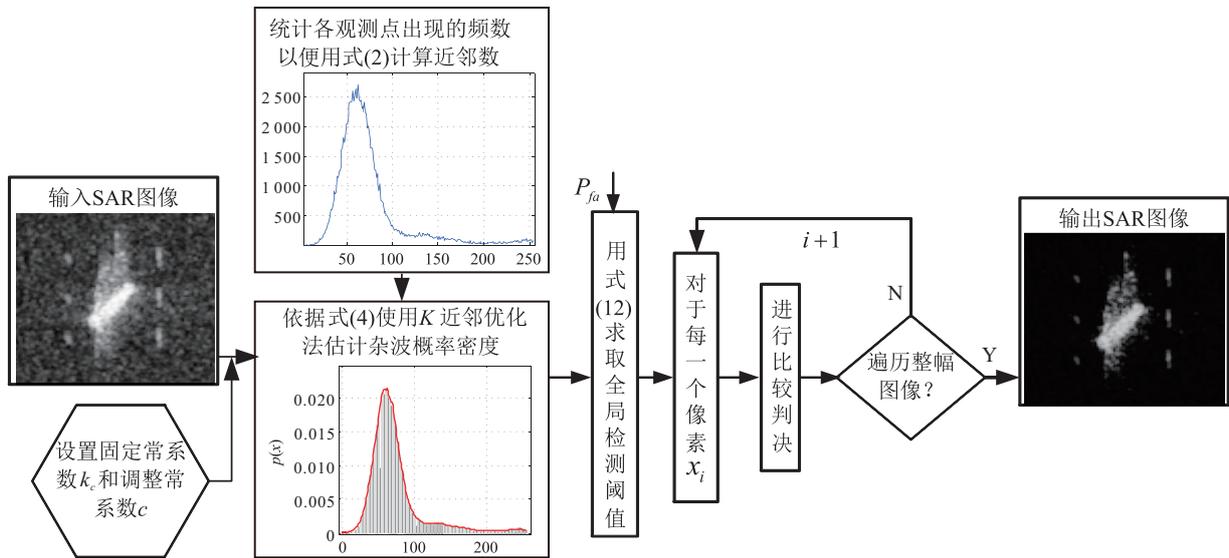


图1 基于 K 近邻优化估计的 CFAR 检测流程

3 实验结果与分析

分别使用 RADARSAT-2 数据、MSTAR 数据库的地面静止车辆数据、TerraSAR-X 图像的子图像数据, 验证本文方法对 SAR 图像杂波建模以及 CFAR 目标检测的有效性, 图像数据信息如表1所示.

表1 实测 SAR 数据

| No. | object/location | sensor | mode | resolution/m |
|-----|----------------------|------------|-----------|--------------|
| 1 | 烟台港 | RADARSAT-2 | Spotlight | 1 |
| 2 | HB15196 | X波段 SAR | Spotlight | 0.3 |
| 3 | Strait of Gibraltar | TerraSAR-X | StripMap | 1.25 |
| 4 | Strait of Singapore | TerraSAR-X | StripMap | 1~5 |
| 5 | Netherlands offshore | RADARSAT-2 | ScanSAR | — |

实验1 SAR 图像单峰杂波建模.

该实验采用表1中 No.1 的子图像和 No.2 图像数据, 如图2(a)和图2(b)所示. 两幅图像均有大量强相干斑噪声, 前者受海杂波干扰, 后者受地杂波污染. 图2(c)和图2(d)分别给出两幅图像的灰度-频数/累积频率图, No.1 子图像的灰度值集中在 $0 \sim 100$ 之间, 且有尖峰及长拖尾, 而 HB15196 灰度值分布较平均.

由图2(e)和图2(f)可见, 二者均呈现单波峰的分布形式. 文献[7]和本文方法均对图像灰度直方图进行了较好地拟合, 而 K 分布和 G^0 分布的拟合效果并不理想, 这也直接凸显了非参量模型建模的优势.

为了定量验证各个估计方法的优劣, 将估计的概率密度函数 (PDF) $P_{estimation}$ 与直方图 P_{data} 的比值进行比较, 结果如图2(g)和图2(h)所示. 观察 No.1 子

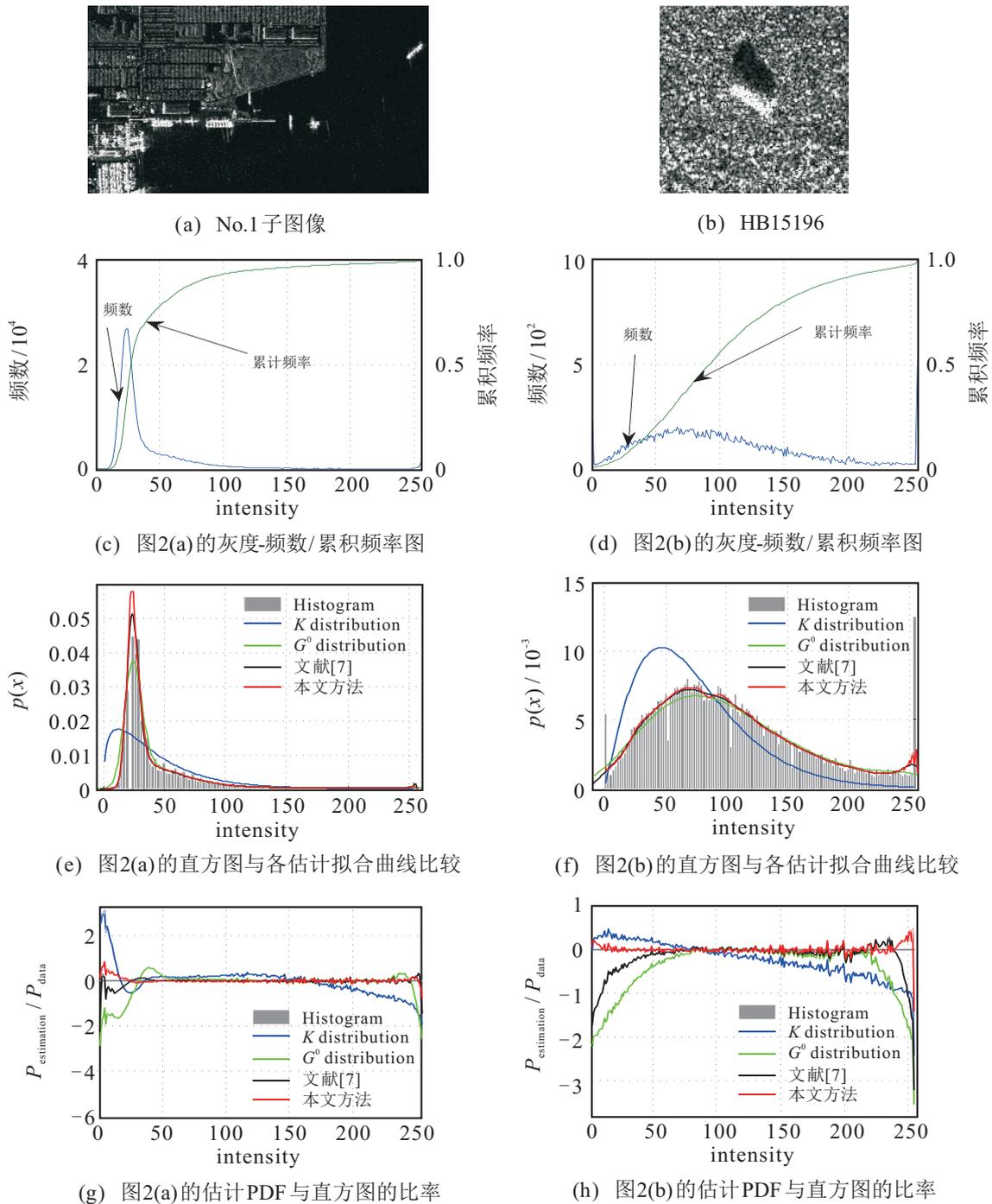


图 2 对单峰SAR图像的统计建模比较

图像的定量描述, K 分布对于长拖尾的拟合最差, G^0 分布次之. 显然文献[7]的估计性能优于 K 分布和 G^0 分布, 比率值保持在 ± 1 dB 之内, 但文献[7]在小于零和大于 255 的两段灰度值区间内会出现无意义的野值. 由图2(e)定性的目视效果和图2(g)定量的表征可以得出, 本文方法比率值在 ± 0.5 dB 之内, 优于核密度估计, 同时本文方法对首尾数据的处理更到位, 对概率密度函数的估计也更精确.

对于目标与背景对比度较差的HB15196, 本文方法的处理结果仍然是最优的, 如图2(h)所示, 文献[7]方法次之, K 分布和 G^0 分布因为建模失配和参数估

计所产生的误差等原因, 使得统计建模结果出现了较大的偏差.

实验2 SAR图像多峰杂波建模.

图3(a)和图3(b)为截取的表1中图像No.3和No.4的子图像, 其背景杂波的直方图呈现双峰或多峰的形式, 任何一种单一的经典分布均无法拟合这类图像的概率密度函数, 此时非参量背景杂波建模的优势凸现出来. 实验2采用高分辨率的TerraSAR-X数据验证本文方法的直观有效性. 两幅子图像中均包含了一些疑似舰船目标的“异常像素”, 各自的灰度-频数图和灰度直方图均出现了多峰的现象, 分别

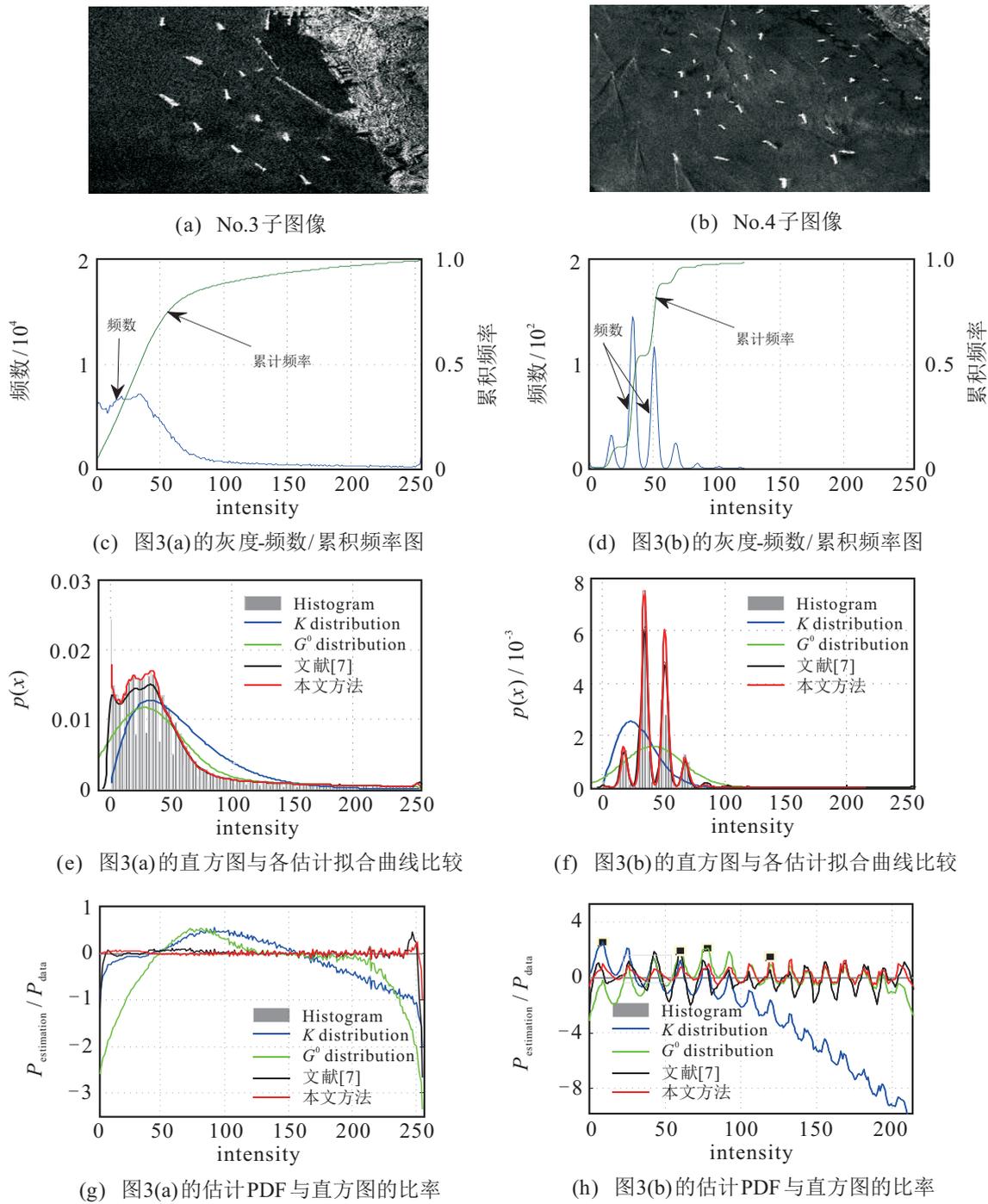


图3 对多峰SAR图像统计建模的比较

如图3(c)和图3(d)所示, K 分布和 G^0 分布已经完全没有能力估计图3(a)和图3(b)的直方图,如图3(e)和图3(f)所示.

图3(e)的直方图信息较图3(f)起伏少,后者海杂波的情况更加复杂,疑似舰船目标又包含诸多尾迹,使得背景杂波的概率密度函数估计更加棘手.各方法估计的 $P_{estimation}$ 与 P_{data} 的比率如图3(g)和图3(h)所示.图3(g)中,本文方法对直方图的拟合效果是最好的,与直方图的比率一直保持在0dB左右;其次是核密度估计,但该方法对始端与末端直方图数据

的处理仍出现突兀的估计结果,中间数据估计的比率值大于本文方法.图3(h)中,本文方法和核密度估计方法均在0dB附近出现振荡形式,本文方法比率值的最高点出现在1.513dB附近,核密度估计方法最高点出现在1.947dB附近,大约相差0.4dB,本文方法更加优异.另外,尽管 G^0 分布估计性能优于 K 分布,但二者已然不能再对多峰直方图进行拟合.

实验3 K 近邻优化估计的CFAR目标检测.

在SAR图像杂波概率密度函数拟合后,使用CFAR验证本文方法对目标检测的有效性.选用表1

中图像 No.4 和图像 No.5 的子图像进行实验,如图 4 和图 5 所示.

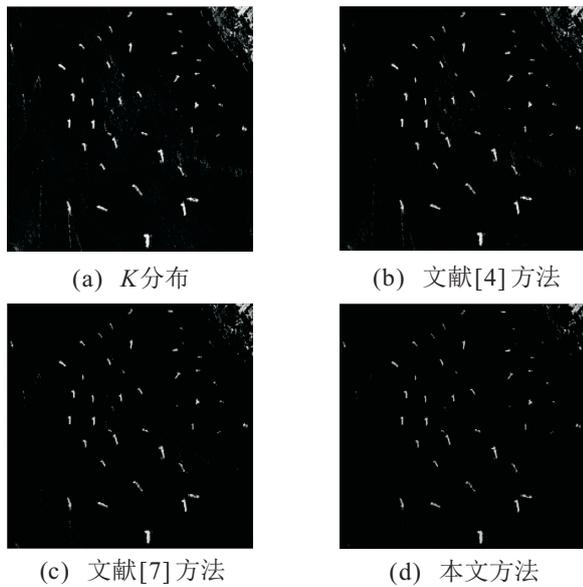


图 4 各估计方法对 No.4 子图像的 CFAR 检测比较

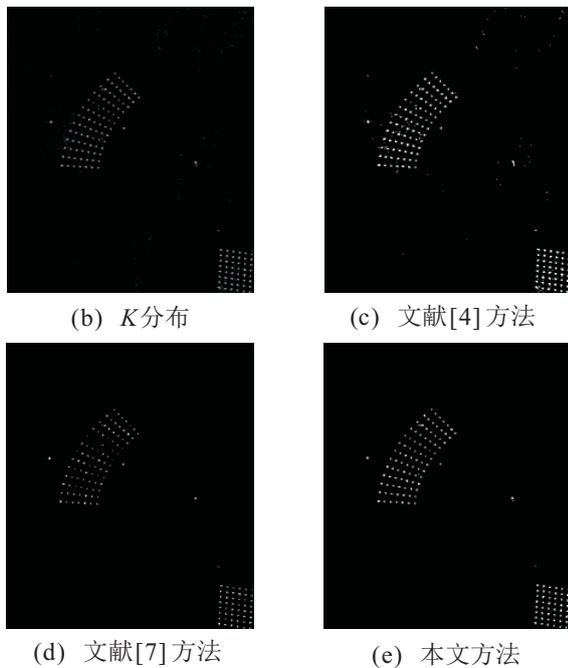


图 5 各估计方法对 No.5 子图像的 CFAR 检测比较

图像 No.4 显示的是用于目标检测的复杂海杂波情况,整个场景中包含多个目标,目标受海杂波污染严重.复杂海况之间存在明显的杂波跃迁,同时在图

像上出现许多云团状的杂波结构,强弱目标交互,目标尺寸大小不一,形态各异并伴随着大量行进目标的尾迹.作为对比,图 4 分别给出了 K 分布、文献 [4] 和文献 [7] 与本文方法的 CFAR 检测结果,在相同的虚警概率下,基于 K 分布和文献 [4] 的恒虚警算法对比文献 [7] 与本文方法,仍有目标尾迹和团雾状的杂波虚警存在,这是因为统计建模的失配带来了显著的恒虚警损失.因为本文方法可以获得比文献 [7] 更逼近于真实杂波的统计结果,其确定的检测阈值是最精确的,可以最大限度地减少虚警,所以在杂波边缘未观察到明显的虚警.

图 5(a) 中,非均匀海杂波中出现大量成集群式规则排列的目标,受到相干斑干扰的目标强弱交互,并被淹没在海杂波背景中.如图 5(b) 和图 5(c) 所示,在这种异常恶劣的杂波环境下, K 分布和 G^0 分布的恒虚警检测结果显示,目标周围有一定数量的杂波虚警出现,由于模型失配及参数估计误差的影响,造成检测结果不理想,目标区域仍存在杂波干扰.文献 [7] 因考量核函数和窗宽的选择问题,使得估计的概率密度函数曲线低于图像直方图而使全局检测门限抬高,造成弱目标漏检,如图 5(d) 所示.通过对比文献 [7],本文方法不仅将大的亮目标检测出来,也得到了许多微弱的小目标,同样在目标的边缘没有出现明显的杂波虚警.

4 结 论

杂波统计建模对于非同态杂波环境下的精确目标检测尤为重要,一方面针对 K 分布和 G^0 分布参量模型,无法拟合不规则的、存在多个波峰的背景概率密度函数的缺陷,另一个方面针对非参量模型中 Parzen 窗估计法依赖于窗宽设置和核函数选择的局限,本文通过巧妙设置待估计观测点的近邻数,重新定义了与 K 近邻信息相关的平均区域体积和概率密度函数,提出了基于 K 近邻优化的非参量估计模型,并分析得到了算法的时间复杂度只与图像的尺寸有关.对实测 SAR 图像的实验结果验证了所提出的算法可以实现对任意分布概率密度函数的估计,且简单精确而有效.

参考文献 (References)

[1] Erfanian S, Vakili V T. Introducing excision switching-CFAR in K distributed sea clutter[J]. Signal Process, 2009, 89(6): 1023-1031.
 [2] El-Darymli K, McGuire P, Power D, et al. Target detection in synthetic aperture radar imagery: A state-of-the-art survey[J]. Journal of Applied Remote Sensing, 2013,

- 7(1): 1-35.
- [3] Ai J Q, Yang X Z, Song J T, et al. An adaptively truncated clutter-statistics-based two-parameter CFAR detector in SAR imagery[J]. *IEEE Journal of Oceanic Engineering*, 2018, 43(1): 267-279.
- [4] 宋文青, 王英华, 刘宏伟. 高分辨SAR图像自动区域筛选目标检测算法[J]. *电子与信息学报*, 2016, 38(5): 1017-1025.
(Song W Q, Wang Y H, Liu H W. An automatic block-to-block censoring target detector for high resolution SAR image[J]. *Journal of Electronics and Information Technology*, 2016, 38(5): 1017-1025.)
- [5] Tao D, Doulgeris A P, Brekke C. A segmentation-based CFAR detection algorithm using truncated statistics[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(5): 2887-2898.
- [6] 王彬, 王国宇. 基于广义Gamma分布的高分辨率SAR图像海岸线检测[J]. *电子学报*, 2018, 46(4): 827-833.
(Wang B, Wang G Y. A coastline detection method in high-resolution SAR images based on the generalized gamma distribution[J]. *Acta Electronica Sinica*, 2018, 46(4): 827-833.)
- [7] Gao G. A parzen-window-kernel-based CFAR algorithm for ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(3): 557-561.
- [8] 刘远龙. 核密度估计中的窗宽选择方法[D]. 合肥: 中国科学技术大学统计与金融系, 2013.
(Liu Y L. Bandwidth selection methods in kernel density estimation[D]. Hefei: Department of Statistics and Finance, University of Science and Technology of China, 2013.)
- [9] 曾勇. 广义近邻模式分类研究[D]. 上海: 上海交通大学自动化系, 2009.
(Zeng Y. Study on generalized nearest neighbor pattern classification[D]. Shanghai: Department of Automation, Shanghai Jiao Tong University, 2009.)
- [10] Rouhani M H, Mohammadi M, Kargarian A. Parzen window density estimator-based probabilistic power flow with correlated uncertainties[J]. *IEEE Transactions on Sustainable Energy*, 2016, 7(3): 1170-1181.
- [11] Swetha N, Sastry P N, Rao Y R, et al. Fast sequency-ordered complex hadamard transform-based parzen window entropy detection for spectrum sensing in cognitive radio networks[J]. *IEEE Communications Letters*, 2016, 20(7): 1401-1404.
- [12] 张宏稷, 杨健, 李延, 等. 基于条件熵和Parzen窗的极化SAR舰船检测[J]. *清华大学学报: 自然科学版*, 2012, 52(12): 1693-1697.
(Zhang H J, Yang J, Li Y, et al. Ship detection polarimetric SAR images based on the conditional entropy and Parzen windows[J]. *Journal of Tsinghua University: Science & Technology*, 2012, 52(12): 1693-1697.)
- [13] Duda P, Rutkowski L, Jaworski M, et al. On the parzen kernel-based probability density function learning procedures over time-varying streaming data With applications to pattern classification[J]. *IEEE Transactions on Cybernetics*, 2018, DOI: 10.1109/TCYB.2018.287767.
- [14] Lang H, Zhang J, Xi Y, et al. Fast SAR sea surface distribution modeling by adaptive composite cubic bézier curve[J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(4): 505-509.
- [15] 谢娟英, 高红超, 谢维信. K 近邻优化的密度峰值快速搜索聚类算法[J]. *中国科学: 信息科学*, 2016, 46(2): 258-280.
(Xie J Y, Gao H C, Xie W X. K -nearest neighbors optimized clustering algorithm by fast search and finding the density peaks of a dataset[J]. *Scientia Sinica Informationis*, 2016, 46(2): 258-280.)
- [16] 孙即祥. 现代模式识别[M]. 第2版. 北京: 高等教育出版社, 2008: 195-197.
(Sun J X, *Modern pattern recognition*[M]. 2nd ed. Beijing: Higher Education Press, 2008: 195-197.)
- [17] Okfalisa, Gazalba I, Mustakim, et al. Comparative analysis of k -nearest neighbor and modified k -nearest neighbor algorithm for data classification[C]. *The 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering*. Yogyakarta: IEEE, 2017: 294-298.

作者简介

彭书娟(1980—), 女, 博士生, 从事SAR图像处理的研究, Email: peng_shu_juan@163.com;

曲长文(1963—), 男, 教授, 博士生导师, 从事电子对抗、SAR图像处理等研究, Email: qcwby@sohu.com;

李健伟(1989—), 男, 博士生, 从事SAR图像处理、深度学习的研究, Email: ljm_jw@163.com.

(责任编辑: 郑晓蕾)