

极不均匀合成孔径雷达杂波建模及恒虚警检测

刘向阳^{1,2}, 许 稼^{2,3}, 彭应宁²

(1. 西安通信学院通信装备管理系, 陕西西安 710106; 2. 清华大学电子工程系, 北京 100084;

3. 空军雷达学院四系, 湖北武汉 430019)

摘 要: 基于合成孔径雷达(SAR)城区极不均匀杂波实际数据和 Cramer Von Mises 距离, 本文运用 Rayleigh 分布、Weibull 分布、K 分布和 G 分布对杂波进行了分布拟合检验. 结果表明 Rayleigh 分布和 Weibull 分布难以拟合极不均匀场景回波, K 分布只能部分拟合, 而 G 分布的拟合程度最好. 进而, 针对 G 分布下的目标恒虚警检测问题, 本文给出了不同参考单元条件下达到给定检测性能时信噪比随形状参数的变化曲线, 分析了恒虚警(CFAR)性能损失随分布参数变化的原因, 得到了确定恒虚警门限的有效方法.

关键词: G 分布; 合成孔径雷达; 杂波模型验证; 恒虚警检测; 极不均匀杂波

中图分类号: TN957.54; TN957.51 **文献标识码:** A **文章编号:** 0372-2112(2007)09-1617-05

Model Validation of the Extremely Heterogeneous SAR Clutter and Its CFAR Detection

LIU Xiang-yang^{1,2}, XU Jia^{2,3}, PENG Ying-ning²

(1. Department of Communication Equipment Management, Xi'an Communication Institute, Xi'an, Shaanxi 710106, China;

2. Department of Electronic Engineering, Tsinghua University, Beijing 100084, China;

3. Radar Academy of Air Force, Wuhan, Hubei 430010, China)

Abstract: Based on real extremely heterogeneous synthetic aperture radar (SAR) urban clutter data and Cramer Von Mises distance, the goodness of fit tests were done via Rayleigh distribution, Weibull distribution, K distribution and G distribution, respectively. It is shown that the Rayleigh distribution and Weibull distribution may be mismatched and the K distribution may be only partly matched, but the G distribution may model the real data very well. Furthermore, for the CFAR (Constant False Alarm Rate) detection under the G distributed clutter, the curves of the signal to noise ratio versus shape parameter and numbers of reference cells were presented to meet the given detection performance, and the reason was also analyzed that the CFAR losses decrease with the increasing of the distribution parameter. Also, an effective CFAR threshold calculating method was proposed at last.

Key words: G distribution; SAR (Synthetic Aperture Radar); validation of clutter model; CFAR (Constant False Alarm Rate) detection; extremely heterogeneous clutter

1 引言

雷达杂波统计特性是目标最优检测的基础. 同样, 杂波数据的幅度分布也构成了合成孔径雷达(SAR)目标自动检测的前提. 由于 SAR 高分辨单元中独立同分布散射点的个数有限, 传统的 Rayleigh 分布假设不再成立, 而高分辨率 SAR 杂波表现出明显的“长拖尾”特性. 针对 SAR 图像数据的幅度分布, 文[1]指出, Weibull 分布对于陆地的高分辨率 SAR 图像幅度能够较好地拟合. 但是 Weibull 分布不能像 K 分布那样描述草地和树

杂波等地杂波. 并且, 无论是 Weibull 分布还是 K 分布对城区等极不均匀 SAR 回波的拟合效果均不佳. 为此, Ferry 等^[2]提出了 G_0 分布, 该分布具有 K 分布类似的参数, 它不但对城区等极不均匀场景具有很好的建模能力, 对森林、草场等场景也具有很好的描述能力. 但是, 目前对 G_0 分布的有效性和普适性的研究工作还鲜有报道, 基于 G_0 分布的目标检测和参数估计等研究则更少. 为此, 基于 Cramer-Von Mises (CVM) 距离, 本文分别用 G_0 分布、K 分布、Weibull 分布和 Rayleigh 分布对实际星载 SAR 图像城区数据进行了分布拟合检验, 进一步验证了

收稿日期: 2006-01-12; 修回日期: 2007-05-09

基金项目: 国家自然科学基金(No. 60502012); 部委预研基金(No. 9140A07020106JW0103); 航天支撑基金(No. J042005047); 教育部科技创新工程重大项目培育基金(No. 706004)

\mathcal{G}_a^0 分布对城区等极不均匀场景的描述能力. 在此基础上, 本文对 \mathcal{G}_a^0 分布下的恒虚警检测性能进行了深入研究, 得到了信噪比随分布形状参数的变化曲线, 分析了 CFAR 损失随形状参数变化的原因, 得到了确定恒虚警门限的有效方法, 为在极不均匀 SAR 杂波环境中的目标检测打下了基础.

2 分布模型以及分布拟合检验

\mathcal{G}_a^0 分布^[2]的概率分布密度函数如下,

$$f_{\mathcal{G}_a^0}(x/n, \alpha, \gamma) = \frac{2n^n \Gamma(n-\alpha) \gamma^{-\alpha} x^{2n-1}}{\Gamma(n) \Gamma(-\alpha) (\gamma + nx^2)^{n-\alpha}}, \quad -\alpha, \gamma, x > 0 \quad (1)$$

其中 n 为 SAR 图像有效视数, 单视处理时 $n=1$; 多视处理时 n 的具体值需要根据具体的成像条件进行估计. 在 n 给定后, \mathcal{G}_a^0 分布的形状由 α 决定, α 可称为形状参数. γ 为尺度参数. 不同 γ 的随机变量可通过尺度变换互相转化. 令

$$y = \frac{x}{\sqrt{\gamma}} \quad (2)$$

则式可转化为

$$f(y) = \frac{2\Gamma(n-\alpha)n^n y^{2n-1}}{\Gamma(n)\Gamma(-\alpha)(1+ny^2)^{n-\alpha}} \quad (3)$$

因此, 当 n 和 α 已知时可以通过尺度变换的方法来实现恒虚警处理.

目前, Kolmogorov-Smirnov (K-S) 分布拟合检验是验证数据集是否满足给定分布的有效方法^[3]. 但是, K-S 检验要求计算经验分布函数的样本是独立同分布的, 为了进行 K-S 检验, 须对 SAR 图像进行空间重采样, 如每 4×4 个图像块选择一个幅度数据. 然而, 正如文[4]所指出的, 即使存在相关性, 经验分布和理论分布之间的 Cramer-Von Mises (CVM) 距离仍旧是它们的相似性的正确度量. 为了使结果更加可信, 除了采用传统的 K-S 检验外, 本文还采用 CVM 距离来衡量上述分布对实际数据的拟合程度.

给定 N 个同分布的随机变量 X_1, \dots, X_N , 设待检验的分布为 $F(x)$, 经验分布函数为 $F_N(x)$, 则 $F(x)$ 和 $F_N(x)$ 间的 CVM 距离为

$$W^2 = N \int_{-\infty}^{+\infty} |F_N(x) - F(x)|^2 dF(x) \quad (4)$$

该积分可以通过下式计算^[3]

$$W^2 = \frac{1}{12N} + \sum_{i=1}^N \left| F(x_{(i)}) - \frac{2i-1}{2N} \right|^2 \quad (5)$$

其中, $x_{(i)}$ 是随机变量集合中的第 i 个有序统计量.

在对 SAR 图像幅度的统计分析中, 样本集合所服从的分布的参数是未知的, 它们可以通过样本的矩来获得. 下面具体给出本文采用的参数估计方法. K 分布和 \mathcal{G}_a^0 分布的参数中视数 n 可以通过特定的方法估计出

来^[5], 这里假定 n 已知. K 分布和 \mathcal{G}_a^0 分布的其余两个参数可以分别通过式(6)和式(7)估计出^[2].

$$\begin{cases} \frac{\Gamma(n)\Gamma(n+1/2)\Gamma(\alpha)\Gamma(\alpha+1/2)}{(\Gamma(n+1/4)\Gamma(\alpha+1/4))^2} = \frac{m_1}{m_1^2} \\ \lambda = \left(\frac{\Gamma(n+1/2)\Gamma(\alpha+1/2)}{\Gamma(n)\Gamma(\alpha)m_1} \right)^2 \frac{1}{n} \end{cases} \quad (6)$$

$$\begin{cases} \frac{\Gamma^2(n+1/4)\Gamma^2(-\alpha-1/4)}{\Gamma(-\alpha-1/2)\Gamma(n+1/2)\Gamma(-\alpha)\Gamma(\alpha)} = \frac{m_1^2}{m_1} \\ \gamma = n \left(\frac{m_1\Gamma(-\alpha)\Gamma(n)}{\Gamma(-\alpha-1/2)\Gamma(n+1/2)} \right)^2 \end{cases} \quad (7)$$

其中, m_k 表示样本的第 k 阶矩, 设 x_1, \dots, x_N 为一组同分布样本, 则其 k 阶矩可通过下式估计

$$m_k = \frac{1}{N} \sum_{i=1}^N x_i^k \quad (8)$$

由于实际获得的样本可能是相关的, 此时矩估计的性能比样本独立时差, 具体表现在样本相关时矩估计的方差比独立时要大. 为了达到同样的估计性能, 样本相关时比样本独立时需要更多的样本数^[6]. 实际用于参数估计的样本数为 10^4 , 这样可以保证参数估计的结果的可信度.

3 基于实际数据的统计分析

本文采用加拿大的星载合成孔径雷达 (RADARSAT) 对某城市的单视成像结果进行分析, 以验证实际数据的统计特性. RADARSAT 的分辨率为 $10\text{m} \times 10\text{m}$, 系统参数如表 1 所示.

图 1 给出了某城区

λ (m)	PRF (Hz)	f_r (Hz/s ²)	R_{cell} (m)
0.0566	1257	1800	4.64

的 SAR 幅度图像.

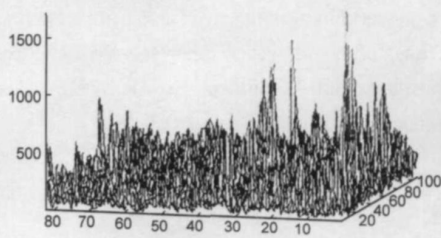


图 1 某城镇 SAR 幅度图像

该测试数据集包含 8100×8100 个像素点, 幅度数据经过了 16 位量化和归一化处理. 为保证检验的完备性, 本文将整幅图像分割成 6561 个 100×100 的子图像, 并将每一幅图像都进行同样的统计分析.

对图 1 所示的数据集的统计分析如表 2 所示. 因为是单视成像, 所以 K 分布和 \mathcal{G}_a^0 分布中的视数 $n=1$, P_{KS} 表示进行 K-S 检验时假设的分布与数据的真实分布契合的概率. 设定检验的显著性水平为 0.05, 即若 $P_{\text{KS}} < 0.05$, 认为给定数据不满足假设的分布. 在此标准下, 在这几种分布中 \mathcal{G}_a^0 分布和 K 分布的 K-S 检验结果都为 0 (表示假设的分布和实际分布吻合), 但 \mathcal{G}_a^0 的 P_{KS} 值要大

于 K 分布的相应值. 另外, 从 CVM 距离上分析, 它们的差异更明显, 相差约两个数量级. 显然, \mathcal{G}_a^0 分布是四种分布中对该数据集拟合得最好的分布模型.

表 2 图 1 数据的统计分析结果

	Rayleigh	Weibull	K	\mathcal{G}_a^0
W^2	662.6	28.4	2.4	3.6×10^{-2}
P_{KS}	5.4×10^{-84}	3.9×10^{-8}	0.14	0.56
KS	1	1	0	0

虽然有文献指出, Weibull 分布也能拟合 SAR 图像, 但是根据以上数据的分析, Weibull 分布对城区这类极不均匀 SAR 图像的拟合能力是较差的. 对全部图像块的分析结果也证明了上述结论. 图 2 给出了 Rayleigh、Weibull、K 和 \mathcal{G}_a^0 这四种分布的 Cramer-Von Mises 距离在各个子图像块上从小到大的变化情况. 从中可见 \mathcal{G}_a^0 分布的 CVM 距离的变化是最小的. 在 \mathcal{G}_a^0 分布假设下, 75.5% 的图像块的 CVM 距离小于 1, 而 K 分布只有 21%, Weibull 分布和 Rayleigh 分布的 CVM 距离小于 1 的比例更小, 分别只有 8.35% 和 1.78%. 因此, 从 CVM 距离指标可以看出, \mathcal{G}_a^0 分布具有最好的拟合性能, Rayleigh 分布对城区等非均匀场景的描述能力最差, 虽然 Weibull 分布对城区等极不均匀场景的建模能力比 Rayleigh 分布强, 但是从统计分析的角度讲也是很差的. K 分布的建模能力较好, 但是比 \mathcal{G}_a^0 分布也有很大的差距.

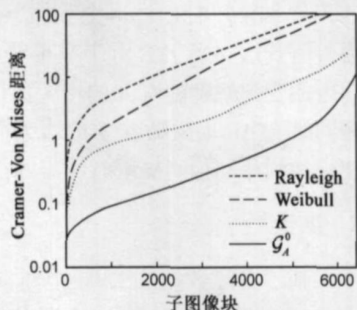


图 2 Cramer-Von Mises 距离按从小到大的排列的变化曲线

下面从 K-S 检验的角度对上述数据进行分析. 检验的结果显示, 有 90% 的图像块的幅度服从 \mathcal{G}_a^0 分布, 62% 服从 K 分布, 30.7% 服从 Weibull 分布, 只有 21.8% 的图像块服从 Rayleigh 分布. 本文所采用的 SAR 图像不仅包括城区, 也包括郊区, 山地、农田和湖泊等. 地形的复杂性导致了极不均匀场景具有极强建模能力的 \mathcal{G}_a^0 分布也不能拟合所有的图像块. 本文的结果和文[2]对 K 分布与 \mathcal{G}_a^0 分布的拟合结果是一致的.

总之, 基于一种更精确的 CVM 测度以及 K-S 检验, 本节进一步验证了 \mathcal{G}_a^0 分布对非均匀杂波的建模能力.

4 \mathcal{G}_a^0 分布下的 CFAR 检测器设计

在独立同分布的 \mathcal{G}_a^0 杂波中检测斯威林 II 目标等价于如下的二元假设检验问题:

$$\begin{aligned} H_0: X &= c \\ H_1: X &= \sqrt{v^2 + c^2 + 2vc \cos(\theta - \phi)}, i = 1, \dots, M \end{aligned} \quad (9)$$

其中, v 是第 i 个脉冲的目标回波信号幅度, v 服从参数 λ 的 Rayleigh 分布. 由于 \mathcal{G}_a^0 分布具有三个参数, 其中有效视数 n 可以通过特定的方法估计出来^[5], 因此, 在后面的分析中假设 n 已知. 因此, 为实现 CFAR 只需考虑另外两个参数 α 和 γ .

4.1 形状参数 α 已知条件下的 CA-CFAR 检测

自适应门限 T 可以通过式(10)来设置

$$T = \left(\sum_{i=1}^N x_i \right) T_0 \quad (10)$$

其中, x_1, \dots, x_N 为服从 \mathcal{G}_a^0 分布 $f_{\mathcal{G}_a^0}(x/n, \alpha, \gamma)$ 的独立同分布样本. 令 $y_i = x_i/\sqrt{\gamma}$, 则

$$\begin{aligned} P_{fa} &= \int_0^\infty \cdots \int_0^\infty \int_0^\infty \left(\sum_{i=1}^N x_i \right) T_0 f_{\mathcal{G}_a^0}(x/n, \alpha, \gamma) dx \\ &\quad \cdot \prod_{i=1}^N f_{\mathcal{G}_a^0}(x_i/n, \alpha, \gamma) dx_1 \cdots dx_N \\ &= \int_0^\infty \cdots \int_0^\infty \int_0^\infty \left(\sum_{i=1}^N y_i \right) T_0 f_{\mathcal{G}_a^0}(y/n, \alpha, 1) dy \\ &\quad \cdot \prod_{i=1}^N f_{\mathcal{G}_a^0}(y_i/n, \alpha, \gamma) dy_1 \cdots dy_N \end{aligned} \quad (11)$$

显然, P_{fa} 与尺度参数 γ 无关, 因而该方法对 γ 是恒虚警的, 因此通过单元平均可实现恒虚警. 遗憾的是难以通过式由虚警概率得到门限因子 T_0 , 因而 T_0 只能通过 Monte Carlo 仿真的方法来获得.

4.2 形状参数 α 和尺度参数 γ 均未知时

当形状参数 α 未知时, 可以先通过式估计 α . 实际中可以先计算出 α 已知时的门限, 当得到估计值后通过查表可以得到检测门限. 由于 α 的估计存在误差, 这会导致虚警率上升. 检测器结构如图 3 所示. 因为当参考单元个数和虚警概率确定后, 门限因子和 \mathcal{G}_a^0 分布的形状参数 α 存在一一对应的关系. 可以编制一张它们之间的对应表以备实时查询.

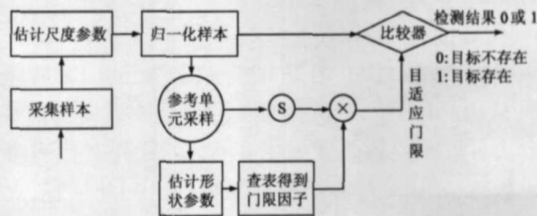


图 3 形状参数和尺度参数均未知时检测器的结构

4.3 形状参数 α 已知时门限因子的确定

在所有参数已知时, 检测门限 T 可以通过下式获得:

$$P_{fa} = \frac{n^a \gamma^{-a} T^{2a} \Gamma(n-a)}{\Gamma(n) \Gamma(1-a)} {}_2F_1(-a, n-a; 1-a; \frac{-\gamma}{nT^2}) \quad (12)$$

其中, ${}_2F_1$ 为超几何函数^[7]. 需要指出的是, 在采用重要采样仿真 CA-CFAR 的检测门限时, 超几何函数的计算是很麻烦的, 并且由于不同 α 时分布的差异很大, 基于

重要采样的门限确定方法在某些参数上性能很差,并且需要通过迭代才能求出需要的门限,而且估计的精度很难事先确定.为此本文采用基于有序统计量的稳健方法来仿真给定虚警概率时的检测门限,该方法的估计精度是事先可以确定的并且易于实现.

令

$$S = x \left(\sum_{i=1}^R y_i \right)^{-1} \quad (13)$$

其中, R 是参考单元个数, y_i 是参考采样, x 是检测单元采样. 在没有目标回波时它们是独立同分布的, S 是与尺度参数无关的. 仿真产生 K 个与 S 独立同分布的独立样本, 将它们从小到大排序, 得

$$S_{(1)} \leq S_{(2)} \leq \cdots \leq S_{(k)} \leq \cdots \leq S_{(K)} \quad (14)$$

若令门限因子 $t_0 = S_{(k)}$ ($k > K - \kappa$), 则此时检测器的虚警概率 \hat{P}_{fa} 如式(15)所示. 由于 $S_{(k)}$ 是随机变量, 所以 \hat{P}_{fa} 也是随机的. 通过这种方法得到的虚警概率的均值和方差分别如式(16)和(17)所示.

$$\hat{P}_{fa} = \Pr \left(x \geq S_{(k)} \sum_{i=1}^R y_i \right) = \Pr(S \geq S_{(k)}) \quad (15)$$

$$E(\hat{P}_{fa}) = (K+1-k)(K+1)^{-1} \quad (16)$$

$$\text{var}(\hat{P}_{fa}) = \frac{k(K+1-k)}{(K+1)^2(K+2)} \quad (17)$$

则估计的相对精度 δ 为

$$\delta = \frac{\sqrt{\text{var}(\hat{P}_{fa})}}{E(\hat{P}_{fa})} = \left(\frac{k}{(K+1-k)(K+2)} \right)^{1/2} \quad (18)$$

令式(16)等于 P_{fa} , 再根据精度 δ , 就可以确定出 K 和 k . 本文中的门限都是通过这种方法确定的, 其中 $P_{fa} = 10^{-5}$ 时 $K = 4 \times 10^7$, $k = K - 399$; $P_{fa} = 10^{-4}$ 时 $K = 4 \times 10^6$, $k = K - 399$, 估计的精度 $\delta = 0.05$. 由于 $\alpha \geq -1$ 时 \mathcal{G}_α^0 分布的二阶矩为无穷大, 所以本文只研究 $\alpha < -1$ 时的情况.

4.4 仿真试验

下面, 考察 CFAR 损失. 在形状参数已知, 虚警概率分别为 10^{-4} 和 10^{-5} 时达到 50% 检测概率所需的信噪比与形状参数的关系曲线如图 4 所示. 图中 R 代表参考单元数目, $R = \infty$ 对应于所有参数均已知的情况, 此时

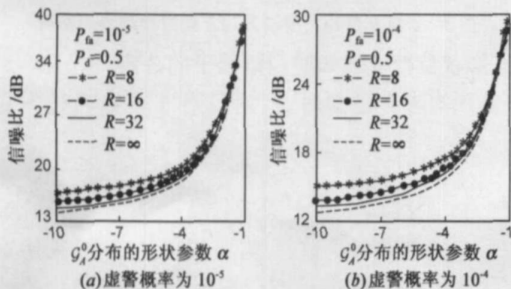


图 4 信噪比随形状参数变化曲线

具有最佳检测性能.

从图 4 可以得到以下结论:

- (1) 随着参考单元数 R 的增加 CFAR 损失越来越小;
- (2) CFAR 损失随着 α 的增大而减小;
- (3) 虚警概率固定时, α 越大, 达到相同检测概率时所需的信噪比就越高;
- (4) 当虚警概率减小时, 为达到相同检测概率, 所需增加的信噪比随着 α 的增大而增大.

其中(1)是显然的. 在解释其余三个结论之前, 首先看一下, $\gamma = 1, n = 1$ 时, 达到

到 10^{-4} 和 10^{-5} 的虚警概率时理想的检测门限与 α 的变化关系, 结果如图 5 所示. 可以看出当 $\alpha < -4$ 时, 从曲线上看不出在这两种虚警概率时检测门限

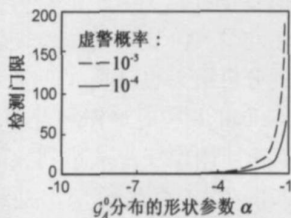


图 5 检测门限与 α 的变化关系

的明显变化, 但是当 α 大于 -4 时, 随着 α 的增加, 二者门限的差距也越来越大, $\alpha = -2.3$ 时为 4.8, $\alpha = -1.5$ 时为 25, 若 $\alpha = -1.1$, 则门限差值竟达到了 122. 由于当虚警概率发生变化时检测门限的变化幅度随着 α 的增大而增加, 从而使参考单元相同时 CFAR 损失随着 α 的增大而减小, 门限的提高也使得达到相同检测概率时所需的信噪比相应地提高.

5 结论

基于实际数据和 CVM 距离, 本文验证了 \mathcal{G}_α^0 分布是描述 SAR 图像中极不均匀场景的一种理想的幅度分布. 另外, 实际数据表明 Weibull 分布和 Rayleigh 分布对于城区这类极不均匀场景的 SAR 幅度图像难以有效地拟合. 进而, 本文提出, 在形状参数 α 已知时, \mathcal{G}_α^0 分布下的恒虚警检测可以通过 CA-CFAR 实现. 进而, 给出基于有序统计量的检测门限确定方法, 并指出达到给定的检测性能时所需的信噪比随着 α 的增大而增大, 而恒虚警损失随着 α 的增大而减小. 最后, 基于检测门限随形状参数 α 的变化, 给出了以上 CFAR 性能的进一步解释. 本文对 \mathcal{G}_α^0 分布下的恒虚警检测进行了初步的研究, 参考单元的独立同分布假设也是一种理想假设. 下一步需要进一步研究如何在参考样本相关的条件下进行恒虚警检测以及实用的 SAR 图像的自动目标检测算法.

参考文献:

- [1] Bisceglie M D, Galdi C. CFAR detection of extended objects in high resolution SAR images[J]. IEEE Trans on GRS, 2005, 43 (4): 833-843.
- [2] Frery A C, Mueller H J, Yanasse C C F, et al. Model for extremely heterogeneous clutter[J]. IEEE Trans on GRS, 1997, 35

(3): 648– 658.

- [3] Agostino R B D, Stephens M A. Goodness of Fit Techniques [M]. New York: Marcel Dekker, 1986.
- [4] Conte E, De Maio A, Galdi C. Statistical analysis of real clutter at different range resolutions[J]. IEEE Trans on AES, 2004, 40 (3): 903– 918.
- [5] Gierull C H, Sikaneta I C. Estimating the effective number of

looks in interferometric SAR data[J]. IEEE Trans on GRS, 2002, 40(8): 1733– 1742.

- [6] Redding N J. Estimating the Parameters of the K Distribution in the Intensity Domain [R]. Salisbury: DSTO Electronics and Surveillance Research Laboratory, 1999. 57– 58.
- [7] 沈永欢, 梁在中, 许履瑚, 等. 实用数学手册[M]. 北京: 科学出版社, 2000.

作者简介:



刘向阳 男, 1976 年 2 月出生于河南省禹州市, 2001 年 3 月毕业于西安电子科技大学电子工程学院, 获信息与通信工程专业工学硕士学位. 2007 年 1 月毕业于清华大学电子工程系, 获工学博士学位. 现为西安通信学院通信装备管理系讲师. 研究方向包括: 雷达杂波建模及其恒虚警检测, 分布式信号检测及其在无线传感器网络中的应用等等.

E-mail: liuxiangyang03@mails.tsinghua.edu.cn



许 稼 1974 年生, 清华大学电子工程系博士后, 副教授. 研究领域包括雷达及水声领域的目标检测和识别、参数估计、仿真模拟、合成孔径/逆合成孔径成像、混沌非线性理论等. 目前已经在《IEEE Trans. on GRS》、《IEEE Trans. on AES》、《IEE Proceeding of RSN》、《中国科学》、《声学学报》、《电子学报》等国内外刊物及各类学术会议上发表和录用论文五十余篇, 其中被 SCI、

EI、ISTP 等检索四十余篇.