

基于拖尾分布的高分辨率合成孔径雷达图像建模^{*}

孙增国^{1)†} 韩崇昭²⁾

1) (西安交通大学机械制造系统工程国家重点实验室, 西安 710049)

2) (西安交通大学电子与信息工程学院, 西安 710049)

(2009 年 5 月 4 日收到; 2009 年 6 月 9 日收到修改稿)

基于中心极限定理的合成孔径雷达(SAR)图像统计分布不能反映高分辨率 SAR 图像尖峰和厚尾的统计特征. 文中使用广义中心极限定理, 由雷达回波的实部和虚部的对称稳定分布, 得到 SAR 图像的拖尾分布(幅值图像的拖尾 Rayleigh 分布以及强度图像的拖尾指数分布), 并以拖尾 Rayleigh 分布为例, 讨论了拖尾分布的代数拖尾特征以及尖峰厚尾的统计特性. 为了实现拖尾分布对高分辨率 SAR 图像的精确建模, 基于第二类统计量, 提出了对数累积量的参数估计方法, 从而高效估计出拖尾分布的参数. 真实 SAR 图像的建模实例表明, 基于广义中心极限定理的拖尾分布可以精确描述高分辨率 SAR 图像的尖峰和厚尾的统计特征.

关键词: 高分辨率合成孔径雷达图像, 广义中心极限定理, 拖尾分布, 对数累积量估计

PACC: 4230V, 4240F, 0250

1. 引言

合成孔径雷达(SAR)图像的统计分布是 SAR 图像研究的重要内容, 只有精确描述 SAR 图像的统计特征, 才有可能获得高质量的边缘检测、分类和降噪结果^[1-7]. 基于中心极限定理, SAR 图像分辨单元内回波的实部和虚部服从高斯分布, 由此可得幅值图像的 Rayleigh 分布以及强度图像的指数分布. 这些传统的统计分布具有解析的概率密度函数, 可以精确描述较低分辨率 SAR 图像的统计特征^[2]. 但是, 当用传统的统计分布对高分辨率 SAR 图像建模时, 却发现高分辨率 SAR 图像(特别是城区图像)所具有的尖峰和厚尾的统计特征不能被精确描述^[8]. 从理论上讲, 传统的统计分布均基于中心极限定理, 而中心极限定理要求 SAR 图像分辨单元内的所有散射体都必须具有有限的方差^[9,10]. 如果这一条件不满足, 则不能使用中心极限定理来描述分辨单元内回波的统计特性. 为了精确描述高分辨率 SAR 图像的尖峰和厚尾的统计特征, 有必要对中心极限定理进行扩展, 在更一般的条件下获得 SAR 图像的统计分布.

本文使用广义中心极限定理描述 SAR 图像分辨单元内回波的统计特性. 根据广义中心极限定理, 独立同分布随机变量的和的极限服从稳定分布, 传统的高斯分布仅是稳定分布的特例^[11-13]. 稳定分布具有代数拖尾, 因此它具有尖峰和厚尾的冲激型的统计特征. 如果使用特殊的稳定分布——对称稳定分布描述 SAR 图像分辨单元内回波的实部和虚部, 就可以得到幅值图像的拖尾 Rayleigh 分布以及强度图像的拖尾指数分布, 传统的 Rayleigh 分布和指数分布分别是它们的特例^[12,14,15]. 和稳定分布类似, 拖尾分布(幅值图像的拖尾 Rayleigh 分布以及强度图像的拖尾指数分布)也具有代数拖尾, 因此它们同样具有尖峰和厚尾的统计特性, 可以精确描述高分辨率 SAR 图像的真实统计特征.

使用拖尾分布对高分辨率 SAR 图像建模, 关键问题是如何精确估计出拖尾分布的参数. 文献[14]推导出拖尾 Rayleigh 分布具有有限的负数阶矩, 由此使用比值估计获得拖尾 Rayleigh 分布的参数. 本文通过 Monte Carlo 仿真以及真实 SAR 图像的估计实验, 证明了比值估计的估计性能受到矩阶次的影响, 不合适的矩阶次可能导致错误的估计结果. 由此, 基于第二类统计量, 本文提出了对数累积量的

^{*} 国家重点基础研究发展计划(批准号:2007CB311006)、国家高技术研究发展计划(批准号:2006AA01Z126)、国家自然科学基金(批准号:60602026)和高等学校博士学科点专项科研基金(批准号:20070698002)资助的课题.

[†] E-mail: duffer2000@163.com

估计方法,可以高效估计出拖尾分布的未知参数.最后,本文给出了基于拖尾分布的高分辨率 SAR 图像的建模实例.和传统的基于中心极限定理的统计分布(幅值图像的 Rayleigh 分布以及强度图像的指数分布)相比,基于广义中心极限定理的拖尾分布可以精确描述高分辨率 SAR 图像的尖峰和厚尾的统计特征.

2. 基于中心极限定理的 SAR 图像统计分布

当地物表面的粗糙程度远大于雷达波长时,可以认为地物表面分辨单元内存在若干散射体,分辨单元内的总回波是所有散射体回波的矢量和^[2].如果 P 和 Q 分别表示分辨单元内总回波的实部和虚部,则总回波可以表示如下:

$$P + jQ = \sum_{i=1}^N A_i e^{j\phi_i}, \tag{1}$$

其中, j 是虚数单位, N 是分辨单元内散射体的个数, A_i 和 ϕ_i 分别表示第 i 个散射体回波的幅值和相角.如果分辨单元内的所有散射体统计独立,并且散射体个数非常大,同时又没有占主导地位的散射体,根据中心极限定理,分辨单元内总回波的实部 P 和虚部 Q 就服从高斯分布.如果雷达使用线性检波器,则获得的回波幅值 A ($A = \sqrt{P^2 + Q^2}$) 就服从 Rayleigh 分布;如果雷达使用平方率检波器,则获得的回波强度 I ($I = P^2 + Q^2$) 就服从指数分布.由此,幅值图像就服从 Rayleigh 分布,强度图像就服从指数分布,它们共同构成了基于中心极限定理的 SAR 图像统计分布.

基于中心极限定理的 SAR 图像统计分布可以较好地描述低分辨率 SAR 图像的统计特性,但是,

它们却不能描述高分辨率 SAR 图像尖峰和厚尾的统计特征^[2,8].图 1 给出了一幅高分辨率 SAR 幅值图像,使用 Rayleigh 分布的建模结果如图 2 所示(直方图用来描述 SAR 图像的真实统计特征).为了清晰反映高分辨率 SAR 图像的拖尾特征,同时给出了对数尺度下的建模结果.显然,高分辨率 SAR 图像的统计分布具有较尖的峰以及较厚的尾,而 Rayleigh 分布不能描述这种特殊的尖峰厚尾的统计特征.因此,基于中心极限定理的统计分布不能用于高分辨率 SAR 图像的精确建模.

我们使用中心极限定理的定义来分析基于中心极限定理的统计分布的局限性^[9,10].中心极限定理用来描述独立的随机变量的和的极限分布,我们仅考虑随机变量具有相同分布的情况(对于 SAR 图像,如果分辨单元内的散射体数量巨大,又没有占主导地位的散射体,这种情况通常是合理的),即对于独立同分布的随机变量 X_i ($i = 1, 2, \cdots, N$),期望 $E(X_i) = \mu$, 方差 $D(X_i) = \sigma^2$, 则对任意的 x , 下式成立,

$$\lim_{N \rightarrow \infty} P \left(\frac{\sum_{i=1}^N X_i - N\mu}{\sqrt{N}\sigma} \leq x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \tag{2}$$

其中, $P(\cdot)$ 表示随机事件的概率.中心极限定理表明,当 $N \rightarrow \infty$ 时, $\sum_{i=1}^N X_i$ 服从高斯分布,期望为 $N\mu$, 方差为 $N\sigma^2$, 这就要求任意随机变量 X_i 的方差 σ^2 必须为有限.如果这一条件不再成立,则中心极限定理失效,基于中心极限定理的统计分布也就失去了理论依据.对于高分辨率 SAR 图像,特别是城区图像,由于分辨单元内通常存在类似于角反射器的散射体,其对应的回波往往具有冲激型的统计特征.



图 1 X 波段 SAR 幅值图像 美国新墨西哥州 Albuquerque 城局部,1 m 分辨率

此时,中心极限定理不能精确描述分辨单元内总回波的统计特性,因此,基于中心极限定理的统计分

布也就不能精确描述高分辨率 SAR 图像尖峰和厚尾的统计特征.

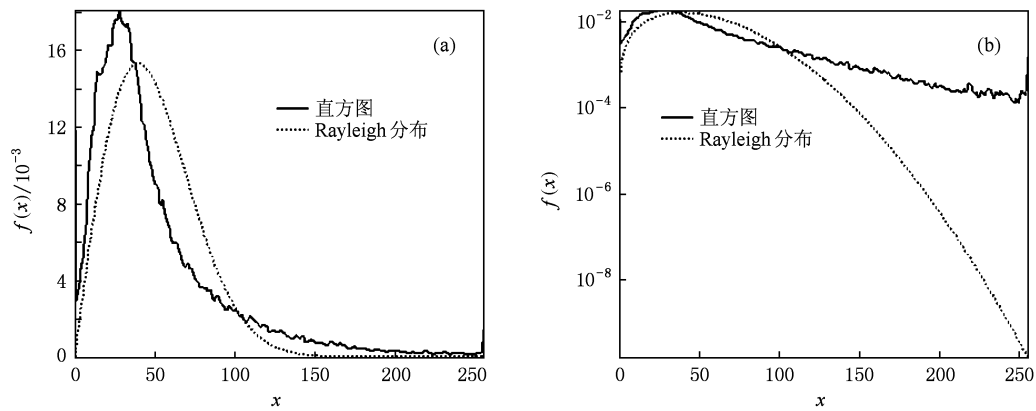


图2 使用 Rayleigh 分布对图 1 建模 (a) 线性尺度,(b) 对数尺度

3. 广义中心极限定理及其 SAR 图像的拖尾统计分布

中心极限定理要求所有的随机变量必须具有有限的方差,我们使用广义中心极限定理来描述一般情况^[12,13]. 广义中心极限定理表明, X 稳定分布的充分必要条件是 X 为如下归一化和的极限分布 (即当 $N \rightarrow \infty$ 时):

$$S_N = \frac{1}{a_N} \sum_{i=1}^N X_i - b_N, \tag{3}$$

其中, $X_i(i=1,2,\cdots,N)$ 是独立同分布的随机变量, a_N 和 b_N 是归一化系数. 由广义中心极限定理可知,独立同分布随机变量的和的极限服从稳定分布,而与随机变量是否具有有限方差没有关系,即广义中心极限定理并不要求随机变量都具有有限方差,因此,它在更一般的意义上描述了独立同分布随机变

量的和的极限分布形态. 当然,如果随机变量都具有有限方差,广义中心极限定理就简化为中心极限定理,即中心极限定理是广义中心极限定理在随机变量方差有限条件下的特例.

根据广义中心极限定理,SAR 图像分辨单元内的总回波服从稳定分布^[11,12]. 稳定分布没有解析的概率密度函数,可以由特征函数描述如下:

$$\varphi(t) = e^{j\alpha t - \gamma |t|^\alpha [1 + j\beta \text{sign}(\cdot) \omega \in (t, \alpha)]}, \tag{4}$$

其中, $\alpha(0 < \alpha \leq 2)$ 是特征指数,它描述稳定分布尖峰和厚尾的统计特征,当 $\alpha = 2$ 时,稳定分布简化为高斯分布; $\gamma(\gamma > 0)$ 是尺度参数,它类似于高斯分布中的方差; $\beta(-1 \leq \beta \leq 1)$ 是对称参数,当 $\beta < 0$ 时,稳定分布向左倾斜,当 $\beta > 0$ 时,稳定分布向右倾斜,当 $\beta = 0$ 时,稳定分布关于位置参数对称,此时,稳定分布又称为对称稳定分布,高斯分布也是对称稳定分布的特例; $a(-\infty < a < \infty)$ 是位置参数,它表示稳定分布在横坐标上的偏移; $\text{sign}(\cdot)$ 是符号函数,

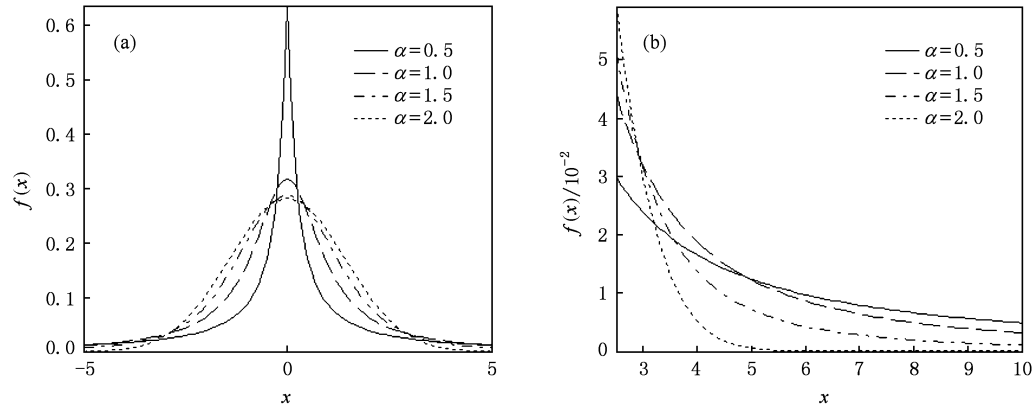


图3 对称稳定分布的概率密度函数 $\gamma=1, a=0$. (a) 概率密度函数,(b) (a) 的尾部

当 $\alpha = 1$ 时, $\omega(t, \alpha) = \frac{2}{\pi} \ln |t|$; 除此之外, $\omega(t, \alpha) = \tan \frac{\alpha \pi}{2}$. 稳定分布同时由 4 个参数决定, 它可以灵活描述分辨单元内回波的各种形态.

稳定分布和高斯分布的最大差别, 在于它具有代数拖尾, 即对于非高斯的稳定分布, 下式成立^[12]:

$$\lim_{t \rightarrow \infty} t^\alpha P(|X| > t) = \gamma C(\alpha), \tag{5}$$

其中, X 是服从稳定分布的随机变量 ($\alpha \neq 2, a = 0$), $C(\alpha)$ 是依赖于 α 的常数. 可见非高斯的稳定分布具有代数拖尾, 而高斯分布只具有指数拖尾. 图 3 给出了对称稳定分布在 α 取不同值时的概率密度函数. 可见对称稳定分布具有明显的尖峰和厚尾的统计特征, 并且 α 越小, 尖峰和厚尾越明显. 当 $\alpha = 2$ 时, 对称稳定分布简化为高斯分布, 因此它的尖峰和厚尾最不明显.

文献[12,14,15]使用特殊的对称稳定分布(位置参数 $a = 0$)来描述 SAR 图像分辨单元内总回波的实部和虚部, 其特征函数如下:

$$\varphi(t) = e^{-\gamma |t|^\alpha}. \tag{6}$$

由此推导出幅值图像服从拖尾 Rayleigh 分布, 其概率密度函数如下:

$$f_A(x) = x \int_0^\infty \rho e^{-\gamma \rho^\alpha} J_0(\rho x) d\rho, \tag{7}$$

其中, $J_0(\cdot)$ 是第一类零阶 Bessel 函数. 由于强度图像的概率密度函数 $f_I(x)$ 可以写为^[2]

$$f_I(x) = \frac{1}{2/x} f_A(\sqrt{x}), \tag{8}$$

其中, $f_A(x)$ 是相应的幅值图像的概率密度函数 ($I =$

A^2), 因此, 强度图像的拖尾指数分布可以直接由 (7) 式得到

$$f_I(x) = \frac{1}{2} \int_0^\infty \rho e^{-\gamma \rho^\alpha} J_0(\rho/\sqrt{x}) d\rho. \tag{9}$$

幅值图像的拖尾 Rayleigh 分布以及强度图像的拖尾指数分布, 共同构成了 SAR 图像基于广义中心极限定理的拖尾统计分布. 当 $\alpha = 2$ 时, 拖尾 Rayleigh 分布和拖尾指数分布分别简化为 Rayleigh 分布和指数分布. 因此, 基于中心极限定理的统计分布是拖尾分布的特例. 由于对称稳定分布具有代数拖尾, 因此, 拖尾 Rayleigh 分布和拖尾指数分布也具有代数拖尾, 即它们同样具有尖峰和厚尾的统计特性(这也是拖尾 Rayleigh 分布和拖尾指数分布名称的由来). 由于幅值图像应用最为广泛, 我们就以拖尾 Rayleigh 分布为例, 具体讨论拖尾分布的代数拖尾特征以及尖峰厚尾的统计特性.

对于拖尾 Rayleigh 分布, 下式成立^[12]:

$$\lim_{t \rightarrow \infty} t^\alpha P(A > t) = \left(\frac{2\pi}{\alpha}\right) B(\alpha, \gamma), \tag{10}$$

其中, A 是服从拖尾 Rayleigh 分布的随机变量 ($\alpha \neq 2$), $B(\alpha, \gamma)$ 是依赖于 α 和 γ 的常数. 可见拖尾 Rayleigh 分布确实具有代数拖尾, 这是拖尾 Rayleigh 分布和 Rayleigh 分布的最大差异. 图 4 给出了拖尾 Rayleigh 分布在 α 取不同值时的概率密度函数, 由图 4 可见, 拖尾 Rayleigh 分布具有明显的尖峰和厚尾的统计特征, 并且 α 越小, 尖峰和厚尾越明显. 当 $\alpha = 2$ 时, 拖尾 Rayleigh 分布简化为 Rayleigh 分布, 因此它的尖峰和厚尾最不明显. 相应地, 拖尾指数分布也具有类似的特点.

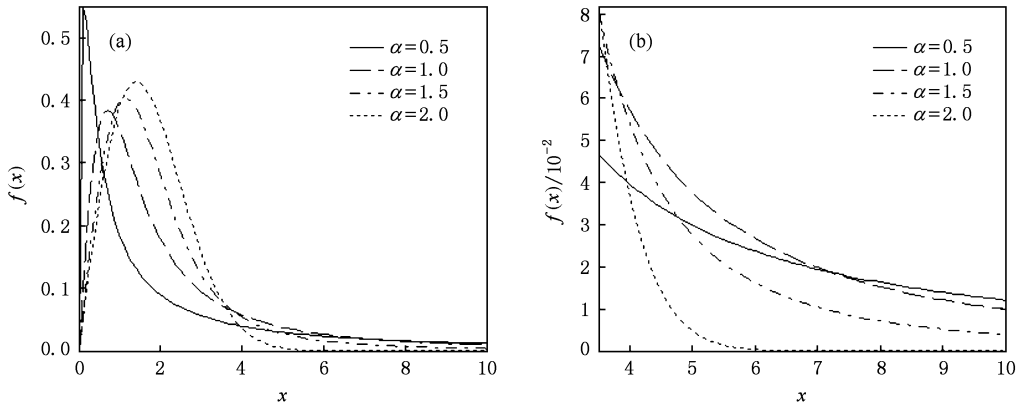


图 4 拖尾 Rayleigh 分布的概率密度函数 $\gamma = 1$. (a) 概率密度函数, (b) (a) 的尾部

拖尾 Rayleigh 分布具有代数拖尾,因此它具有不同于一般 Rayleigh 分布的特殊的统计特性. 图 5 给出了拖尾 Rayleigh 分布的仿真数据实现,服从拖尾 Rayleigh 分布的样本由文献[16]中所提出的方法产生. 显然,拖尾 Rayleigh 分布具有冲激型的统计特性. 图 6 给出了使用拖尾 Rayleigh 分布对图 1

的建模结果. 和 Rayleigh 分布相比,拖尾 Rayleigh 分布可以精确描述高分辨率 SAR 图像的尖峰和厚尾的统计特征. 总之,基于广义中心极限定理的拖尾分布具有尖峰和厚尾的统计特征,因此,它们可以用于高分辨率 SAR 图像的精确建模.

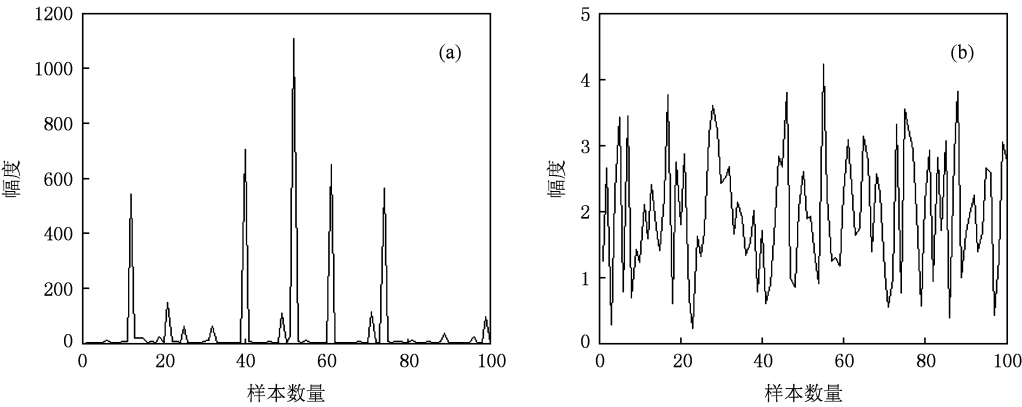


图 5 拖尾 Rayleigh 分布的仿真数据实现 样本数量为 100. (a) $\alpha=0.5$, (b) $\alpha=2$ (Rayleigh 分布)

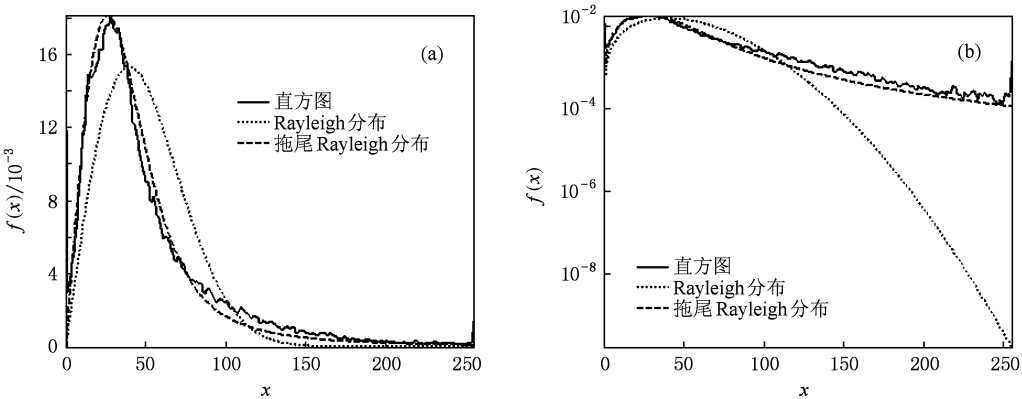


图 6 使用拖尾 Rayleigh 分布对图 1 建模 (a) 线性尺度, (b) 对数尺度

4. 拖尾分布的对数累积量参数估计

要实现拖尾分布对高分辨率 SAR 图像的精确建模,参数估计是一个关键问题. 本节首先对 Kuruoglu 等^[14]提出的比值估计进行分析,着重说明矩阶次对比值估计的影响. 其次,基于第二类统计量,本文提出了对数累积量的估计方法,从而实现对拖尾分布未知参数的高效估计.

4.1. 比值估计

Kuruoglu 等^[14]证明了拖尾 Rayleigh 分布具有

有限的负数阶矩,由此提出了用于估计拖尾 Rayleigh 分布参数的比值估计如下:

$$\frac{E(A^{2p})}{E^2(A^p)} = \frac{\Gamma(p+1)\Gamma^2\left(-\frac{p}{2}\right)}{2\Gamma(-p)\Gamma^2\left(\frac{p}{2}+1\right)} \frac{\alpha\Gamma\left(-\frac{2p}{\alpha}\right)}{\Gamma^2\left(-\frac{p}{\alpha}\right)}$$
$$\left(-1 < p < -\frac{1}{2}\right), \tag{11}$$

$$\gamma = \left[\frac{\alpha\Gamma\left(-\frac{p}{2}\right)E(A^p)}{2^{p+1}\Gamma\left(\frac{p}{2}+1\right)\Gamma\left(-\frac{p}{\alpha}\right)} \right]^{\frac{\alpha}{p}}, \tag{12}$$

其中, A 是服从拖尾 Rayleigh 分布的随机变量, p 是

矩阶次, $\Gamma(\cdot)$ 是 Gamma 函数. 可见如果使用样本矩代替 $E(A^{2p})$ 和 $E(A^p)$, 特征指数 α 就可由 (11) 式估计得出, 而尺度参数 γ 就可由 (12) 式估计得出. 在估计 α 时, 由于没有解析的估计表达式, 因此必须使用数值算法 (文献 [14] 推荐使用二分法). 表 1 给出了在不同矩阶次和参数真值条件下比值估计的 Monte Carlo 仿真结果, 包括多次实验结果的平均值和标准差 (括号中的数值), 服从拖尾 Rayleigh 分布的样本由文献 [16] 中所提出的方法产生, 使用二分法估计特征指数 α . 可见, 比值估计性能严重依赖矩阶次的选择, 如果矩阶次选择不当 (如 $p =$

-0.81), 估计性能就会严重降低. 进一步, 我们使用拖尾 Rayleigh 分布对图 1 建模, 选择不同的矩阶次, 对应的比值估计的参数估计结果如图 7 所示. 显然比值估计的估计结果仍然随着矩阶次的变化而变化. 值得注意的是, 对于图 1 所示的 X 波段 SAR 幅值图像, 在使用 (11) 式估计特征指数时, 当矩阶次取值较小时, α 的估计值可能大于 2, 此时比值估计失效, 因为 α 只能在 $(0, 2]$ 上取值. 比如, 当 $p = -0.87$ 时, $\hat{\alpha} = 3.3044$. 为了克服矩阶次对比值估计性能的影响, 基于第二类统计量, 我们提出不依赖矩阶次的高效参数估计方法.

表 1 比值估计的 Monte Carlo 仿真

参数真值	$p = -0.81$		$p = -0.75$		$p = -0.51$	
	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$
$\alpha = 0.5$	0.5537	0.9392	0.5268	0.9724	0.5021	0.9990
$\gamma = 1$	(0.0688)	(0.1217)	(0.0538)	(0.0824)	(0.0156)	(0.0158)
$\alpha = 1.0$	1.2499	0.9656	1.1125	0.9869	1.0031	1.0009
$\gamma = 1$	(0.2835)	(0.0739)	(0.1793)	(0.0500)	(0.0527)	(0.0167)
$\alpha = 1.5$	1.8141	1.0132	1.7415	1.0127	1.5117	1.0024
$\gamma = 1$	(0.3482)	(0.0602)	(0.3184)	(0.0309)	(0.1027)	(0.0192)

注: 括号中的数值是多次实验结果的标准差. 独立运行 1000 次, 每次 10000 个样本.

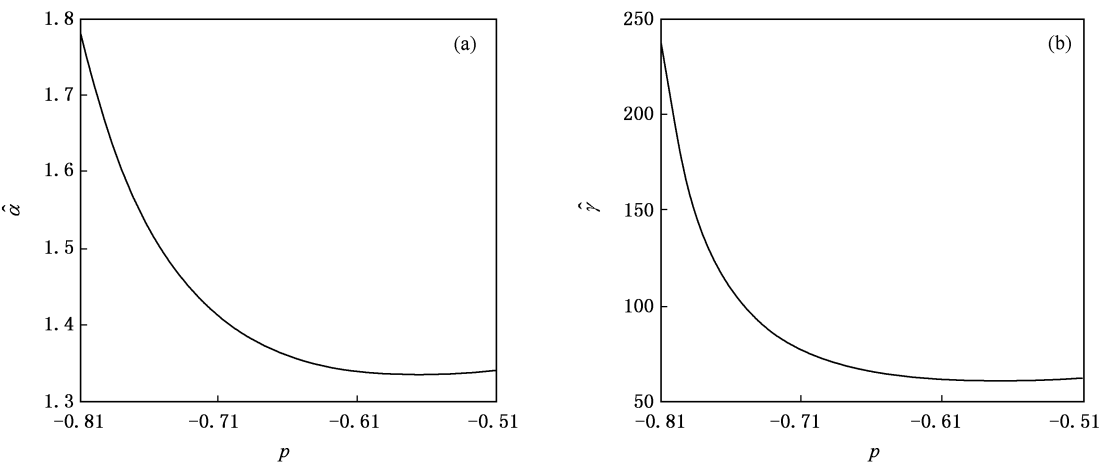


图 7 图 1 的比值参数估计结果 (a) $\hat{\alpha}$ 随 p 的变化关系, (b) $\hat{\gamma}$ 随 p 的变化关系

4.2. 对数累积量参数估计

我们知道, 普通统计量是由傅里叶变换定义的, 即第一特征函数 (一般的特征函数) 是相应概率密度函数的傅里叶变换, 第二特征函数是第一特征函数的对数, 矩是第一特征函数在原点处的导数, 累积量是第二特征函数在原点处的导数^[10]. 类似地, 如果概率密度函数定义在正半轴上, 我们就可

以把普通统计量中的傅里叶变换替换为 Mellin 变换, 从而定义相应的第二类统计量^[15,17,18].

对于函数 $g(x) (0 \leq x < \infty)$, 其 Mellin 变换定义如下:

$$M[g(x)](s) = \int_0^\infty x^{s-1} g(x) dx,$$

(13)

其中, s 是复数变量. 可见 Mellin 变换要求 $g(x)$ 定义在正半轴上, 这是它和傅里叶变换的显著差异.

由于幅值图像和强度图像的概率密度函数只在正半轴上有定义,因此可以使用 Mellin 变换定义相应的第二类统计量. 对于概率密度函数 $f(x)$ ($0 \leq x < \infty$), 其第二类第一特征函数定义如下:

$$\Phi(s) = \int_0^\infty x^{s-1} f(x) dx. \tag{14}$$

可见, $f(x)$ 的第二类第一特征函数正是它的 Mellin 变换. 由此, $f(x)$ 的第二类第二特征函数定义如下:

$$\Psi(s) = \ln \Phi(s). \tag{15}$$

最终, 我们定义 r 阶对数矩 (第二类矩) $\tilde{m}_r$ 和 r 阶对数累积量 (第二类累积量) $\tilde{k}_r$ 分别如下:

$$\tilde{m}_r = \left. \frac{d^r \Phi(s)}{ds^r} \right|_{s=1}, \tag{16}$$

$$\tilde{k}_r = \left. \frac{d^r \Psi(s)}{ds^r} \right|_{s=1}. \tag{17}$$

在第二类统计量中, 对数矩 $\tilde{m}_r$ 和对数累积量 $\tilde{k}_r$ 的关系和普通统计量中矩和累积量的关系完全相同. 假定 $\tilde{m}_{X(r)}$ 表示随机变量 X 的 r 阶对数矩, 根据 Mellin 变换的性质, $\tilde{m}_{X(r)} = E[(\ln X)^r]$, 这就是对数矩和对数累积量名称的由来.

对数累积量可以由下式估计得出 (即样本对数累积量):

$$\begin{aligned} \hat{\tilde{k}}_1 &= \frac{1}{M} \sum_{i=1}^M \ln y_i, \\ \hat{\tilde{k}}_2 &= \frac{1}{M-1} \sum_{i=1}^M (\ln y_i - \hat{\tilde{k}}_1)^2, \end{aligned} \tag{18}$$

其中, y_i 是样本, M 是样本个数.

对于拖尾 Rayleigh 分布, 将其概率密度函数 (7) 式代入到 (14) 式中, 可得相应的第二类第一特征函数如下:

$$\Phi(s) = \frac{2^s \Gamma((s+1)/2) \gamma^{(s-1)/\alpha} \Gamma((1-s)/\alpha)}{\Gamma((1-s)/2) \alpha}, \tag{19}$$

接着把 (19) 式依次代入 (15) 和 (17) 式中, 最终可得拖尾 Rayleigh 分布的前两阶对数累积量如下:

$$\tilde{k}_1 = C_E \left(\frac{1}{\alpha} - 1 \right) + \frac{\ln \gamma}{\alpha} + \ln 2, \tag{20}$$

$$\tilde{k}_2 = \frac{\pi^2}{6\alpha^2}, \tag{21}$$

其中, C_E 是欧拉常数 ($C_E \approx 0.5772$). 可见, 使用样本对数累积量 (18) 式代替对数累积量 $\tilde{k}_1$ 和 $\tilde{k}_2$, 特征指数 α 可以方便地由 (21) 式估计得出, 而尺度参数 γ 可以由 (20) 式估计得出, 上面两式就是拖尾

Rayleigh 分布基于第二类统计量的对数累积量估计式. 类似地, 对于拖尾指数分布, 其对数累积量估计式 (前两阶对数累积量) 可以推导如下:

$$\tilde{k}_1 = 2C_E \left(\frac{1}{\alpha} - 1 \right) + \frac{2 \ln \gamma}{\alpha} + 2 \ln 2, \tag{22}$$

$$\tilde{k}_2 = \frac{2\pi^2}{3\alpha^2}. \tag{23}$$

我们以拖尾 Rayleigh 分布为例, 分析对数累积量参数估计方法. 和文献 [14] 提出的比值估计 ((11) 和 (12) 式) 相比, 对数累积量估计式不含矩阶次, 因此它的估计性能完全由样本决定, 而不受矩阶次的影响. 对数累积量估计式具有解析的表达形式, 不需要像比值估计那样求解非线性方程, 易于计算. 表 2 给出了对数累积量估计的 Monte Carlo 仿真, 包括多次实验结果的平均值和标准差 (括号中的数值), 图 8 给出了在 α 取不同真值时对数累积量估计和比值估计的估计性能比较, 服从拖尾 Rayleigh 分布的样本由文献 [16] 所提出的方法产生. 可见, 无论 α 真值取为何值, 对数累积量估计均能获得较高的估计精度. 当矩阶次 p 取值合适时 (如 $p = -0.51$), 比值估计的参数估计结果和对数累积量估计接近 (α 真值较小时的情况除外). 但是, 一旦矩阶次取值较小时 (如 $p = -0.75$), 比值估计的估计性能就严重退化. 总之, 由于对数累积量估计式不受矩阶次的影响, 又具有解析的表达形式, 同时可以获得较高的估计精度, 因此, 和比值估计相比, 对数累积量估计是拖尾分布高效的参数估计方法. 当使用拖尾分布对高分辨率 SAR 图像建模时, 我们使用对数累积量估计来获得拖尾分布的未知参数.

表 2 对数累积量估计的 Monte Carlo 仿真

$\alpha = 0.5$	$\gamma = 1$	$\alpha = 1$	$\gamma = 1$	$\alpha = 1.5$	$\gamma = 1$
$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$
0.5001	0.9996	1.0006	0.9997	1.5003	1.0003
(0.0052)	(0.0107)	(0.0102)	(0.0115)	(0.0161)	(0.0133)

注: 括号中数值是多次实验结果的标准差. 独立运行 1000 次, 每次 10000 个样本.

5. 高分辨率 SAR 图像建模实例

对于图 1 的高分辨率 X 波段 SAR 幅值图像, 使用不同的分布对其建模, 结果如图 9 所示. Weibull 分布和 K 分布是描述 SAR 图像的常用统计分布, Weibull 分布中的 c 和 b 分别是形状参数和尺度参数, K 分布中的 ν 和 b 分别是形状参数和尺度参数 [2, 19]. 表 3 给出了不同分布的参数估计结果以及

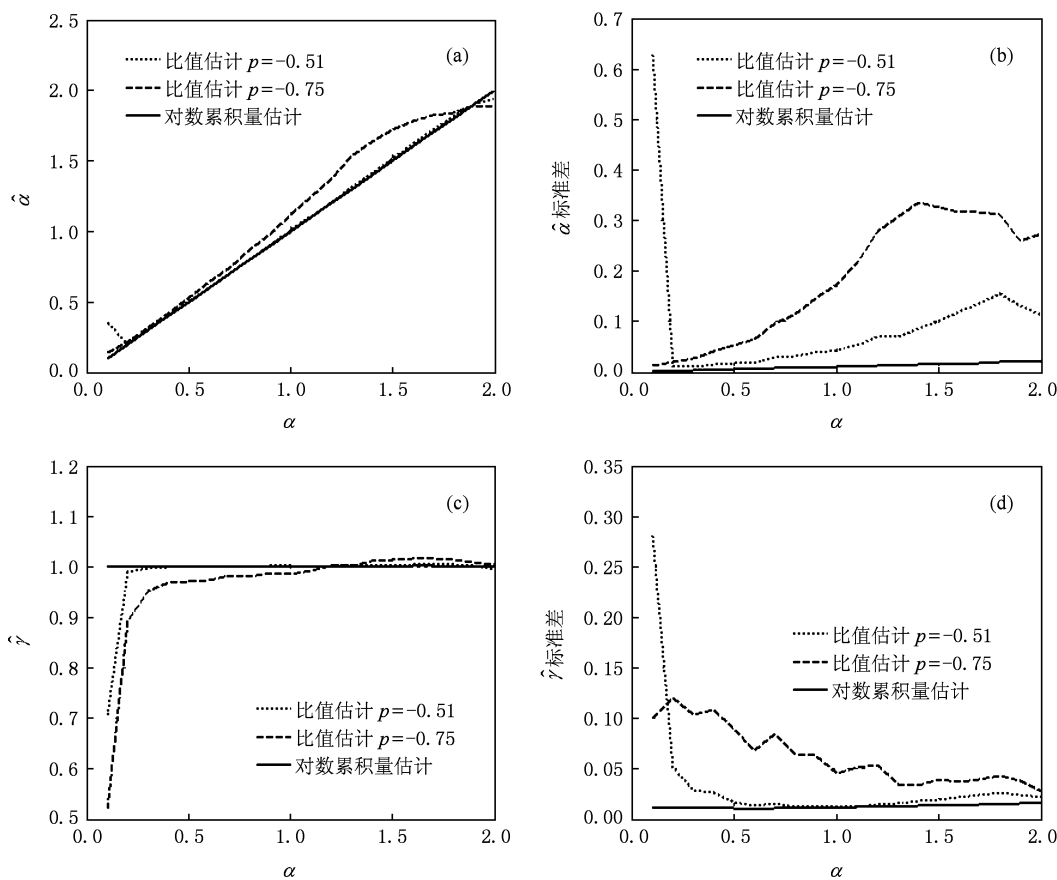


图 8 对数累积量估计和比值估计的估计性能比较 Monte Carlo 仿真, $\gamma=1$, 独立运行 1000 次, 每次 10000 个样本. (a) $\hat{\alpha}$ 随 α 真值的变化关系, (b) $\hat{\alpha}$ 标准差随 α 真值的变化关系, (c) $\hat{\gamma}$ 随 α 真值的变化关系, (d) $\hat{\gamma}$ 标准差随 α 真值的变化关系

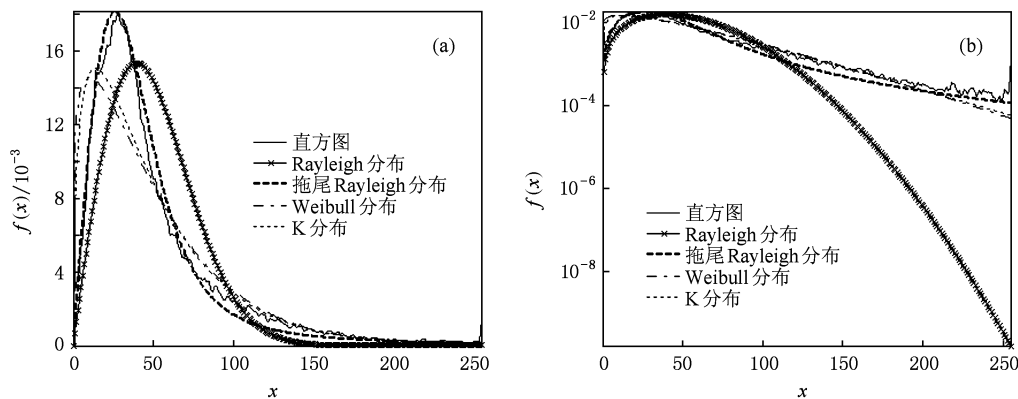


图 9 不同分布对图 1 的建模结果 (a) 线性尺度, (b) 对数尺度

K-S 概率. K-S 概率用来定量描述某一分布对图像统计分布的接近程度, K-S 概率越大, 说明该分布越接近图像的统计分布^[2,20]. 显然, Rayleigh 分布既不能反映图像的拖尾特征, 也不能反映图像的尖峰特征, 因此它不能用于高分辨率 SAR 幅值图像的建模. Weibull 分布和 K 分布可以在一定程度上描述图像的拖尾特征, 但它们不能反映图像的尖峰特

征. 只有拖尾 Rayleigh 分布能够同时描述高分辨率 SAR 幅值图像的尖峰和厚尾的统计特征, 因此它最接近图像的真实统计特性. 其次, 使用拖尾指数分布对图 10 的高分辨率 SAR 强度图像建模, 结果如图 11 所示, 相应的参数估计结果和 K-S 概率如表 4 所示. 和其他分布相比, 拖尾指数分布可以反映高分辨率 SAR 强度图像的拖尾特征, 因此对应的 K-S

概率最大,即拖尾指数分布最接近图像的真实统计特性. 总之,不论是幅值图像还是强度图像,也不论是 X 波段图像还是 Ku 波段图像,基于广义中心极限定理的拖尾分布可以描述高分辨率 SAR 图像的尖峰和厚尾的统计特征,因此,它们可以用于高分辨率 SAR 图像的精确建模.

表 3 对图 1 建模分布的参数估计结果和 K-S 概率

	参数估计结果	K-S 概率
Rayleigh 分布	$\hat{\gamma} = 784.5940$	0.8568
拖尾 Rayleigh 分布	$\hat{\alpha} = 1.5310 \quad \hat{\gamma} = 116.3923$	0.9651
Weibull 分布	$\hat{c} = 1.1765 \quad \hat{b} = 52.5036$	0.9316
K 分布	$\hat{v} = 0.7779 \quad \hat{b} = 0.00018269$	0.9431

表 4 对图 10 建模分布的参数估计结果和 K-S 概率

	参数估计结果	K-S 概率
指数分布	$\hat{\gamma} = 1819.5$	0.8828
拖尾指数分布	$\hat{\alpha} = 1.7869 \quad \hat{\gamma} = 707.9520$	0.9704
Weibull 分布	$\hat{c} = 1.6812 \quad \hat{b} = 81.5041$	0.9587
K 分布	$\hat{v} = 4.6607 \quad \hat{b} = 0.00064037$	0.9603

6. 结 论

本文使用理论分析与实验验证的方法,指出中心极限定理的局限性,由此基于广义中心极限定理,由 SAR 图像分辨单元内回波的实部和虚部的对称稳定分布,依次得到幅值图像的拖尾 Rayleigh 分布以及强度图像的拖尾指数分布. 拖尾分布具有代数拖尾,因此可以描述高分辨率 SAR 图像的尖峰和厚尾的统计特性. 为了实现拖尾分布对高分辨率 SAR 图像的精确建模,基于第二类统计量,本文提出了对数累积量的参数估计方法,可以高效估计出拖尾分布的未知参数. 高分辨率 SAR 图像的建模实例表明,基于广义中心极限定理的拖尾分布可以精确反映高分辨率 SAR 图像的真实统计特性,因此它是高分辨率 SAR 图像建模的有效工具.

拖尾统计分布的参数(特别是特征指数)具有实际的物理含义,它和地物表面的物理特征有关. 使用拖尾分布描述高分辨率 SAR 图像时,不同的地貌(如森林、农田、城区、海洋等等)应该对应不同的参数,同一种地貌在不同的成像条件下(如下雨前后的农田,即在不同的土壤湿度条件下)也应该对应不同的参数. 因此,我们可以利用估计出的拖尾分布的参数,来对不同的地物成像表面进行分类和分析,这将是下一步要研究的内容.

本文进行建模实验的高分辨率 SAR 图像均由美国 Sandia 国家实验室获得,在此对该实验室表示衷心的感谢.



图 10 Ku 波段 SAR 强度图像 美国新墨西哥州 Albuquerque 城局部,1 m 分辨率

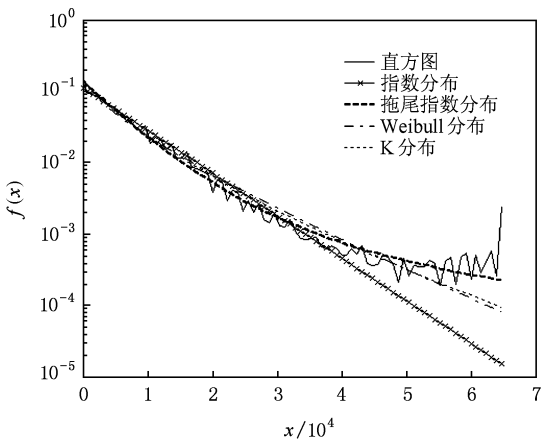


图 11 不同分布对图 10 的建模结果 对数尺度

[1] Guo H D 2000 *Theory and Application of Radar Earth Observation* (Beijing: Science Press) (in Chinese) [郭华东 2000 雷达对地观测理论与应用(北京:科学出版社)]

[2] Oliver C, Quegan S 1998 *Understanding Synthetic Aperture Radar Images* (Boston: Artech House)

[3] Zong X P, Xu Y, Dong J T 2006 *Acta Phys. Sin.* **55** 3223 (in Chinese) [宗晓萍、徐 艳、董江涛 2006 物理学报 **55** 3223]

[4] Jiang B, Wang H Q, Li X, Guo G R 2006 *Acta Phys. Sin.* **55** 3985 (in Chinese) [姜 斌、王宏强、黎 湘、郭桂蓉 2006 物理学报 **55** 3985]

[5] Jiang K Y, Cai Z M 2005 *Acta Phys. Sin.* **54** 4596 (in Chinese) [姜可宇、蔡志明 2005 物理学报 **54** 4596]

[6] Wu Y G, Tao M D 2006 *Chin. Phys.* **15** 1137

[7] Teng S Y, Cheng C F, Liu M, Gui W L, Xu Z Z 2005 *Chin. Phys.* **14** 1990

[8] Anastassopoulos V, Lampropoulos G A, Drosopoulos A, Rey M 1999 *IEEE Trans. Aerospace and Electronic Systems* **35** 43

[9] Wang R X 2003 *Mathematical Statistics* (Xi'an: Xi'an Jiaotong University Press) (in Chinese) [汪荣鑫 2003 数理统计(西安:西安交通大学出版社)]

[10] Papoulis A 1984 *Probability, Random Variables, and Stochastic Processes* (Singapore: McGraw-Hill)

[11] Shao M, Nikias C L 1993 *Proc. IEEE* **81** 986

[12] Nikias C L, Shao M 1995 *Signal Processing with Alpha-Stable Distributions and Applications* (New York: Wiley)

[13] Georgiou P G, Tsakalides P, Kyriakakis C 1999 *IEEE Trans. Multimedia* **1** 291

[14] Kuruoglu E E, Zerubia J 2004 *IEEE Trans. Image Processing* **13** 527

[15] Achim A, Kuruoglu E E, Zerubia J 2006 *IEEE Trans. Image Processing* **15** 2686

[16] Szajnowski W J, Wynne J B 2001 *IEEE Signal Processing Letters* **8** 151

[17] Nicolas J M 2002 *European Signal Processing Conference* (Toulouse: EURASIP) p197

[18] Tison C, Nicolas J M, Tupin F, Maitre H 2004 *IEEE Trans. Geoscience and Remote Sensing* **42** 2046

[19] Ward K D, Tough R J A, Watts S 2006 *Sea Clutter: Scattering, the K Distribution and Radar Performance* (London: The Institution of Engineering Technology)

[20] Press W H, Teukolsky S A, Vetterling W T, Flannery B P 1994 *Numerical Recipes in C* (Cambridge: Cambridge University Press)

Modeling high-resolution synthetic aperture radar images with heavy-tailed distributions^{*}

Sun Zeng-Guo^{1)†} Han Chong-Zhao²⁾

1) (*State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China*)

2) (*School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China*)

(Received 4 May 2009; revised manuscript received 9 June 2009)

Abstract

Statistical distributions of synthetic aperture radar (SAR) images based on central limit theorem cannot reflect the statistical characteristics of sharp peak and heavy tail of high-resolution SAR images. By using the generalized central limit theorem, the heavy-tailed distributions (heavy-tailed Rayleigh distribution for amplitude image and heavy-tailed exponential distribution for intensity image) are obtained from the symmetric stable distributions of real and imaginary parts of echoes. Taking the heavy-tailed Rayleigh distribution as an example, the algebraic tails of heavy-tailed distributions are explained as well as the statistical properties of sharp peak and heavy tail. In order to model the high-resolution SAR images with the heavy-tailed distributions, based on second-kind statistical, Characteristics the log-cumulant estimator is proposed to efficiently estimate the parameters of the heavy-tailed distributions. Modeling experiments on real SAR images demonstrate that the heavy-tailed distributions based on the generalized central limit theorem can accurately describe the sharp-peaked and heavy-tailed statistical characteristics of high-resolution SAR images.

Keywords: high-resolution synthetic aperture radar images, generalized central limit theorem, heavy-tailed distributions, log-cumulant estimator

PACC: 4230V, 4240F, 0250

^{*} Project supported by the State Key Development Program for Basic Research of China (Grant No. 2007CB311006), the National High Technology Research and Development Program of China (Grant No. 2006AA01Z126), the National Natural Science Foundation of China (Grant No. 60602026), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20070698002).

[†] E-mail: duffer2000@163.com