

合成孔径雷达原始数据幅相压缩算法¹

姚世超 王岩飞 张冰尘 廖蜀燕

(中国科学院电子学研究所 北京 100080)

摘 要 针对现代合成孔径雷达 (SAR) 对相位信息越来越高的要求, 该文根据 SAR 原始数据幅值和相位统计独立的分布特点, 提出了对 SAR 原始数据的幅度和相位分别进行量化的幅度相位压缩算法, 即 AP(Amplitude & Phase) 算法; 介绍了算法原理和实现流程图; 在实验基础上分析了算法的性能; 并将该算法与 BAQ 算法相比较, 验证了 AP 算法能够很好地保留原始数据的相位信息的特点。

关键词 合成孔径雷达, 数据压缩, AP 算法

中图分类号 TN951

1 引 言

合成孔径雷达 SAR(Synthetic Aperture Radar) 是一种新型的对地观测微波遥感设备, 由于它能够全天候、全天时地工作, 并具有多频段、多极化、视角可变、穿透性等优势, 已成为人们获取地表信息的重要手段。伴随着 SAR 在多种领域的广泛应用, 合成孔径雷达朝着更高分辨率、多频段、多极化、多模式等方向发展。这些发展反过来更加剧了合成孔径雷达海量的数据和数据传输、数据存储之间的矛盾。对原始数据进行压缩是解决这一矛盾的一种有效途径。在合成孔径雷达原始数据压缩领域, 人们已经发展了多种算法, 如 BAQ 算法^[1]、BFPQ 算法^[2,3]、矢量量化^[4], 以及 BAVQ^[5] 和 FFT-BAQ^[6] 算法、二维非线性压缩算法^[7] 等。其中 BAQ 算法和 BFPQ 算法由于算法简单, 容易实现的优点分别应用于美国 Magellan 飞行器上的 SAR 和航天飞机的 SAR SIR-C 上。

现代雷达的多频、多极化技术以及干涉成像技术的应用, 对于 SAR 原始数据压缩处理的相位信息保留, 提出了更高的要求。无论是 BAQ, BFPQ, 还是直接量化、矢量量化等算法, 都只是着眼于如何提高算法的信噪比, 却无法在直观上显示出对相位信息的特殊保留, 而只能通过降低压缩比, 提高信噪比来间接减小相位误差。为此, 本文提出一种新的压缩方法——AP 算法, 该算法根据 SAR 原始数据幅度和相位统计独立的分布特点, 对幅度和相位分别进行量化, 在直观上显示出对相位信息的特殊保留, 很好地解决了合成孔径雷达原始数据压缩过程中相位信息的保留问题。

2 幅相压缩算法原理及实现方案

2.1 幅值相位压缩算法的基本原理

SAR 回波信号是合成孔径雷达发射出的电磁波经照射区域内的地面物体反射之后, 通过天线接收而得到。由于成像区域内的散射体很多, 所以回波信号可以看作是大量统计独立的随机变量之和, 其数学模型可以表示为

$$A = \sum_{i=1}^N a_i \cdot e^{j\phi_i} \quad (1)$$

其中 a_i 是单个散射体的回波幅度 ϕ_i 是相位延迟, 当 N 足够大时, 由中心极限定理, 可以认为, 合成孔径雷达原始数据的幅度和相位统计独立。 ϕ 在 $[-\pi, +\pi]$ 之间呈均匀分布, 幅度 A

¹ 2001-05-08 收到, 2001-08-07 定稿

呈瑞利分布^[1,2]。幅度 A 的概率密度函数为

$$f(A) = (A/\mu^2) \exp[-(A^2/(2\mu^2))], \quad A \geq 0 \quad (2)$$

相位 ϕ 为与回波幅度 A 均值有关的一个标量。瑞利分布由该参数唯一地确定。 ϕ 的概率密度函数为

$$f(\phi) = 1/(2\pi), \quad -\pi \leq \phi \leq \pi \quad (3)$$

对于现有的合成孔径雷达系统,通常采用正交解调技术,得到 I 和 Q 两路正交信号。如果将 I 和 Q 两路正交信号转化为幅度和相位表示,根据上面的分析可知,信号的幅度和相位统计上相互独立,幅度符合瑞利分布,相位呈均匀分布特性。利用这一特点,可以分别对幅度和相位进行编码。考虑到相位 ϕ 在 $[-\pi, +\pi]$ 之间呈均匀分布的特性,我们对相位采用 m bit 均匀量化;而幅度为瑞利分布,相应地我们就对幅度进行最佳量化。

所谓量化是对数据样本的取值范围进行分层,然后用单个值代表一层内的所有可能值。最优量化则是指在某种误差意义下的最优量化方案。本文采用的最优量化原则为均方误差最小的量化原则。

假设样本的取值在 $[X_0, X_K]$ 内,概率密度函数分布为 f , 样本落入 $[X_i, X_{i+1}] (i = 0, 1, \dots, K-1)$ 子区间的输出电平为 y_i , 量化误差为 $x - y_i$, 其中 $x \in [X_i, X_{i+1}]$, 这样, 在 K 个子区间内总的均方误差为

$$\overline{\varepsilon^2} = \sum_{i=0}^{K-1} \int_{X_i}^{X_{i+1}} (x - Y_i)^2 f(x) dx \quad (4)$$

为使均方误差最小, Y_i 和 X_i 的取值应满足

$$\partial \overline{\varepsilon^2} / \partial Y_i = 0 \quad (5)$$

$$\partial \overline{\varepsilon^2} / \partial X_i = 0 \quad (6)$$

对于相位的均匀分布, 将 (3) 式代入 (4)-(6) 式可解得判决电平为

$$X_i = (\phi_{i-1} + \phi_i) / 2 \quad (7)$$

输出电平为

$$\phi_i = (X_{i+1} + X_i) / 2 \quad (8)$$

对于幅度的瑞利分布, 将 (2) 式代入 (4)-(6) 式可解得判决电平为

$$X_i = (Q_{i-1} + Q_i) / 2 \quad (9)$$

输出电平为

$$Q_i = \int_{X_i}^{X_{i+1}} x f(x) dx / \left[\int_{X_i}^{X_{i+1}} f(x) dx \right] \quad (10)$$

由 (9), (10) 两式, 我们就可以采用逐次逼近的计算方法, 在设定的误差下, 在计算机上求得幅度瑞利分布的输出电平和判决电平^[8,9]。图 1 即为分布系数 $\mu = 1$ 的瑞利分布 1bit 最佳量化示意图。

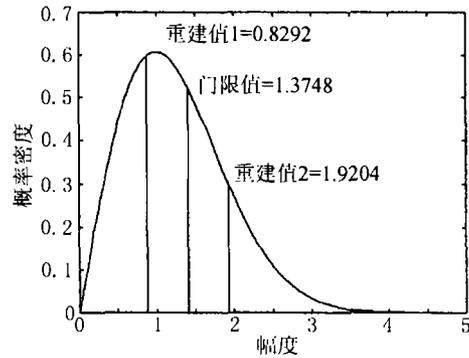


图 1 瑞利分布 1bit 量化示意图

2.2 幅相压缩算法的实现

在数据压缩的实际应用中, 为了降低数据的动态范围, 取得较好的压缩效果, 一般对原始数据进行分块, 然后对分块内的数据进行压缩处理。分块应该足够大, 使分块内的数据统计特性符合压缩处理的要求, 即分块内数据的幅度和相位在统计上相互独立, 且幅度符合瑞利分布, 相位呈均匀分布特性; 同时由于分块越大, 压缩的计算量就越大, 不利于算法的电路实现。所以分块既不能太大, 也不能太小。根据实际的计算结果, 一般说来, 当分块内的样本数多于 100 个时, 就已经基本上满足要求了。

对于 AP 算法的实现, 如图 2 所示的 AP 算法压缩流程图, 各为 $N \times N$ 分块的 I, Q 两路数据经计算后得到 $N \times M$ 分块的幅值数组和相位数组。I, Q 两路数据经 $N \times M$ 后用数组表示为

$$I = \begin{pmatrix} i_{11} & i_{12} & \cdots & i_{1M} \\ i_{21} & i_{22} & \cdots & i_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ i_{N1} & i_{N2} & \cdots & i_{NM} \end{pmatrix} \quad (11)$$

$$Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1M} \\ q_{21} & q_{22} & \cdots & q_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1} & q_{N2} & \cdots & q_{NM} \end{pmatrix} \quad (12)$$

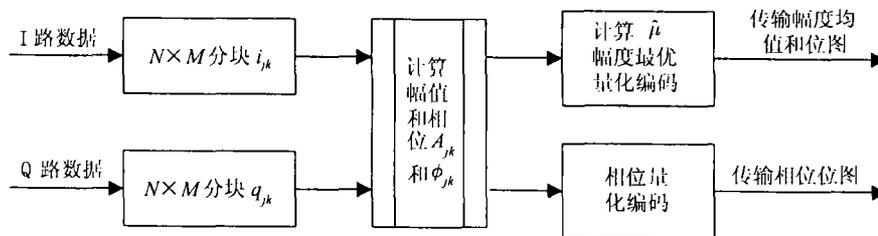


图 2 AP 算法压缩流程图

转化为极坐标, 即可求得其幅度和相位分别为

$$A_{jk} = \sqrt{i_{jk}^2 + q_{jk}^2}, \quad k = 0, 1, \dots, N, \quad j = 0, 1, \dots, M \quad (13)$$

$$\theta_{jk} = \text{atan}(q_{jk}/i_{jk}), \quad k = 0, 1, \dots, N, \quad j = 0, 1, \dots, M \quad (14)$$

由 (13), (14) 式, 可以得到幅值和相位的计算结果, 按照上面所述的最优化原则, 就可以对其分别进行压缩. 计算表明: 对于幅度的瑞利分布, 采用 1bit, 2bit 量化, 就可以获取比较理想的压缩效果. 而对于相位信息, 考虑到其均匀分布、熵值较高的特性, 则必须采用较多的量化比特数, 根据系统的压缩比和相位误差的要求, 相应地可采用 2bit, 3bit, 4bit 和 5bit 量化.

幅度的均值为

$$\bar{x} = \frac{1}{N \times M} \sum_{i,j} A_{ij} \quad (15)$$

我们知道, 对于瑞利分布 (2) 式, 其均值 \bar{x} 和 μ 之间存在以下关系^[10]:

$$\bar{x} = \sqrt{\pi/2}\mu \quad (16)$$

利用 (16) 式, 得到瑞利分布系数 μ , 根据幅度量化的判决门限, 就可以按照系统的要求, 将幅度值量化为 1bit 或 2bit, 形成幅度位图, 然后传输 8bit 表示的均值 \bar{x} 和幅度位图.

同样, 根据 (14) 式, 就可以将相位在 $[-\pi, +\pi]$ 上分为相等的若干份 ($2^m, m = 2, 3, 4, 5$, 在 2bit 量化的情况下, 即为 4 等份, 3bit 量化, 则为 8 等份, 以此类推), 量化计算形成位图, 并将这个位图传输到接收端.

$$\hat{i}_{jk} = Q_{jk} \cos \phi_{jk} \quad (17)$$

$$\hat{q}_{jk} = Q_{jk} \sin \phi_{jk} \quad (18)$$

从图 3 的解压缩流程图可以看出: 在解压缩端, 对于幅值编码结果, 根据解码得到的均值 \bar{x} , 求得瑞利分布系数估计值 $\hat{\mu}$, 再由输出电平和幅度位图, 重建相应的幅度值; 对于相位编码结果, 根据 (8) 式所得的重建电平值, 可以得到相位的重建值. 有了幅度和相位的重建值, 利用 (17), (18) 式, 就可以得到 I, Q 两路数据的重建值.

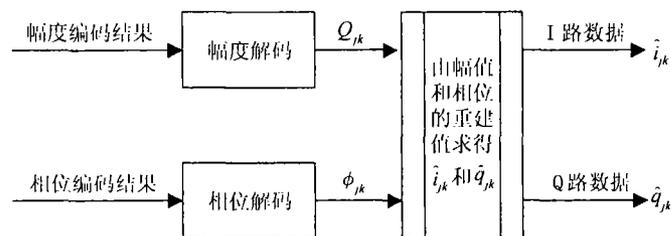


图 3 AP 算法解压缩流程图

3 实验分析

根据上述的分析, 我们对该算法在各种可能的幅值和相位的量化组合方式下的性能、分块大小对压缩性能的影响, 以及该算法保留相位信息的特点方面, 进行了计算分析. 考虑到迄今为止, 在 SAR 原始数据压缩领域, 虽然已经提出了多种算法, 但只有 BAQ 和 BFPQ 算法曾投入过实际的 SAR 原始数据压缩的应用, 因此, 我们还将 AP 算法同 BAQ 算法做一个性能计算分析比较.

实验主要采用 32×32 和 64×64 分块, 相位则分别采用 2bit, 3bit, 4bit, 5bit 量化, 幅度分别采用 1bit, 2bit, 3bit 量化. 所用的 SAR 原始数据为实际数据, 每个数据经 A/D 转换后表示为 8bit.

表 1 所示为 AP 算法在 32×32 分块下, 不同幅度和不同相位量化比特数的各种组合方式下, I 路数据的压缩性能。表 2 所示则为 AP 算法在 64×64 分块, 部分组合方式下, I 路数据的压缩性能。Q 路数据压缩性能和 I 路相差很小, 这里就不再列出。表 3 所示为 BAQ 算法 I 路数据压缩性能。其中所列的性能指标有: 压缩比、原始数据域信噪比 (SNR), 平均相位误差以及图像数据域信噪比。

表 1 AP 算法 32×32 分块 I 路数据压缩性能

幅度量化比特数	相位量化比特数	压缩比	原始数据域 SNR(dB)	平均相位误差 (rad)	图像域 SNR(dB)
1	2	5.33	5.43	0.45	18.41
	3	4.00	7.75	0.23	22.44
	4	3.20	10.48	0.10	25.04
	5	2.67	10.91	0.06	25.47
2	2	4.00	6.11	0.45	20.36
	3	3.20	9.23	0.23	23.80
	4	2.67	14.61	0.10	28.40
	5	2.29	15.48	0.06	28.93
3	3	2.67	9.83	0.23	24.52
	4	2.29	16.83	0.10	29.83
	5	2.00	19.39	0.06	30.53

表 2 AP 算法 64×64 分块 I 路数据压缩性能

幅度量化比特数	相位量化比特数	压缩比	原始数据域 SNR(dB)	平均相位误差 (rad)	图像域 SNR(dB)
1	2	5.33	5.43	0.45	18.38
	3	4.00	7.78	0.23	22.40
	4	3.20	10.44	0.10	24.94
2	4	2.67	14.58	0.10	28.42
3	5	2.00	19.40	0.06	30.55

表 3 BAQ 算法 I 路数据压缩性能

分块大小	量化比特数	压缩比	原始数据域 SNR	平均相位误差 (rad)	图像域 SNR
32×32	2	4.00	9.00	0.28	23.62
	3	2.67	14.23	0.17	27.56
	4	2.00	19.30	0.10	29.85

信噪比的定义为

$$\text{SNR} = 10 \log_{10}(s/\varepsilon) \quad (19)$$

其中 $\varepsilon = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [g(i,j) - f(i,j)]^2$ 为噪声功率, $s = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N g(i,j)^2$ 为信号功率; $g(i,j)$ 为原数据灰度, $f(i,j)$ 为重建数据灰度。

平均相位误差的定义

$$\bar{\phi} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |\phi_0(i,j) - \phi_1(i,j)| \quad (20)$$

其中 $\phi_0(i,j)$ 为原数据相位, $\phi_1(i,j)$ 为重建数据相位。

分析表 1 可以看出, 随着相位量化比特数的增多, 无论是原始数据域还是图像数据域的信噪比都有显著提高; 幅度量化比特数的增多, 也能使压缩性能得到改善, 但是改善幅度不如前者。在同样的压缩比下, 幅度和相位量化比特数的组合方式对压缩性能的影响十分显著。例如, 1, 4 组合 (幅度 1bit, 相位 4bit) 比之 2, 3 组合, 性能明显为高; 2, 4 组合也明显优于 3, 3 组合, 这说明相位量化比特数低于 4bit 时, 相位量化比特数对于压缩性能的影响, 比幅

度量化比特数的影响为大。同时我们也注意到：2, 4 组合优于 1, 5 组合, 3, 4 组合优于 2, 5 组合, 这说明当相位量化比特数大于 4bit 时, 低的幅度量化比特数已上升成为限制压缩性能提高的主要因素。

相位误差只与相位的量化比特数有关, 与幅度的量化比特数完全无关, 相位的量化比特数越高, 相位误差就越小, 且基本上与理论误差接近。(2bit 量化的理论误差为 $\pi/8 = 0.3926$, 3bit 量化的理论误差为 $\pi/16 = 0.1963$, 4bit 量化的理论误差为 $\pi/32 = 0.0981$, 5bit 量化的理论误差为 $\pi/64 = 0.049$)。

比较表 1 和表 2, 我们发现分块的大小对压缩性能影响比较小, 这说明只要分块内样本数足够多, 分块内数据的统计特性就基本上可以看作是其幅值和相位统计独立, 且幅值呈瑞利分布、相位在 $[-\pi, +\pi]$ 之间呈均匀分布。

比较表 1 和表 3, 我们可以得出以下结论:

(1) AP 算法不仅在直观上显示出对相位信息的特殊保留, 而且计算结果也同样验证了该方法对于相位信息的保留大大优于 BAQ 算法。

(2) AP 算法具有更多可供选择的压缩方案, 有多种不同的组合方式, 根据不同的需要, 我们可以选择适当的压缩比, 以获取最佳的系统性能。

(3) 相同的压缩比下, AP 算法总能选择出一种幅度和相位的组合方式而得到接近或超过 BAQ 算法的性能。

为比较 AP 算法和 BAQ 算法在图像域的性能, 我们选取了 4 幅图像: 图 4 是未经压缩的 SAR 原始数据数据成像结果; 图 5 是 AP 算法 32×32 分块, 幅度 1bit、相位 2bit 压缩处理后的成像结果; 图 6 是 AP 算法 32×32 分块, 幅度 2bit、相位 4bit 压缩处理后的成像结果; 图 7 为经 BAQ 算法 32×32 分块 3bit 量化压缩过的数据成像结果。压缩比在括号内标明。

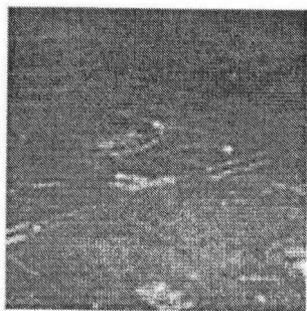


图 4 未压缩原始数据成像



图 5 AP 算法成像 (5.28:1)

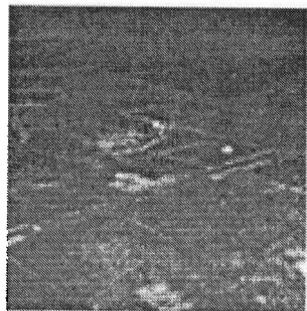


图 6 AP 算法成像 (2.67:1)



图 7 BAQ 算法成像 (2.67:1)

4 结 论

AP 算法对相位采用均匀量化, 提高相位的量化比特数, 就可以很好地保留数据的相位信息, 不仅在直观上反映出对相位信息的特殊保留, 而且实验结果也证明了它的有效性。这一优点对于相干成像和极化雷达等对于相位信息较为敏感的数据压缩应用尤为重要。考虑到其幅度和相位分别压缩的算法特点, 很容易进一步应用到极化原始数据压缩的领域。

参 考 文 献

- [1] R. Kwok, Block adaptive quantization of Magellan SAR data, IEEE Trans. on Geosci. Remote Sensing, 1989, GRS-27(4), 375-383.
- [2] T. H. Joo, D. Held, An adaptive quantization method for burst mode synthetic aperture radar, Int. Radar Conf., Arlington, 1985, 385-390.
- [3] U. Benz, A fuzzy block adaptive quantizer(FBAQ) for synthetic aperture radar, In Proc. Fuzz-IIEEE, Orlando, FL, 1994, 1583-1585.
- [4] Y. Linde, A. Buzo, R. M. Gray, An algorithm for vector quantizer design, IEEE Trans. on Comm., 1980, COM-28(1), 84-95.
- [5] A. Moreira, F. Blaeser, Fusion of block adaptive and vector quantizer for efficient SAR data compression, In Proc., IGARSS'93, Tokyo, 1993, 1583-1585.
- [6] U. Benz, *et al.*, A comparison of several algorithms for on-board SAR raw data compression, IEEE Trans. on Geosci. and Remote Sensing, 1995, GRS-33(5), 1266-1277.
- [7] 王岩飞, 贾守新, 陈宗鹭, 合成孔径雷达原始数据的二维压缩算法, 卫星遥感图象数据传输与压缩技术专题研讨会论文集, 北京, 1996, 217-222.
- [8] 李象霖, 数字图像处理, 北京, 中国科学技术大学研究生院, 1996, 83-91.
- [9] W. A. Pearlman, G. H. Senge, Optimal quantization of the Rayleigh probability distribution, IEEE Trans. on Comm., 1979, COM-27(1), 101-112.
- [10] 沈凤麟, 钱玉美, 信号统计分析基础, 北京, 中国科学技术大学出版社, 11-12.

AMPLITUDE AND PHASE COMPRESSION ALGORITHM OF SAR RAW DATA

Yao Shichao Wang Yanfei Zhang Bingchen Liao Shuyan

(*Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract The phase information becomes more and more important in the field of SAR applications. Considering the statistical independent property between phase and amplitude of SAR raw data, a new method is provided that the amplitude and phase of raw data are compressed respectively. In this paper, a flow chart shows the procession of the algorithm. The algorithm performance is analyzed based on real SAR raw data. Finally, the AP(Amplitude & Phase) algorithm is compared with the BAQ algorithm. The performance of the AP algorithm is proved to be fine, especially in terms of the reservation of phase information.

Key words Synthetic aperture radar, Data compression, AP algorithm

姚世超: 男, 1974 年生, 硕士, 主要研究方向为数据压缩。

王岩飞: 男, 1963 年生, 研究员, 主要研究领域包括信号处理和微波成像技术。

张冰尘: 男, 1973 年生, 助理研究员, 主要研究领域包括信号处理和微波成像技术。

廖蜀燕: 女, 1970 年生, 工程师, 主要研究领域包括数据压缩和微波成像技术。