

## 一种距离聚焦后的 SAR 原始数据压缩算法

曾尚春<sup>①②</sup> 朱兆达<sup>②</sup>

<sup>①</sup>(西南民族大学电气信息工程学院 成都 610041)

<sup>②</sup>(南京航空航天大学信息科学与技术学院 南京 210016)

**摘要:** 由于合成孔径雷达(SAR)原始数据的相关性很低,直接压缩原始数据是比较困难的。该文提出一种新算法,先对 SAR 原始数据作距离聚焦处理,使其在方位向具有较强的相关性,再沿方位向作矢量线性预测,并对预测残差序列作分块自适应量化。结合一组实测 SAR 原始数据,用3种算法分别进行了压缩和解压缩,并计算了数据域及图像域信噪比,给出了3种压缩算法所成的图像。实验表明,在相同比特率条件下,该文算法得到的数据域信噪比和图像域信噪比均比 BAQ 算法高。该文算法的计算量远小于有关文献给出的距离聚焦后的压缩方法。

**关键词:** 合成孔径雷达; 距离聚焦; 矢量线性预测; 信噪比; 块自适应量化; 比特率

中图分类号: TN957.52

文献标识码: A

文章编号: 1009-5896(2008)04-0921-04

## An Algorithm for SAR Raw Data Compression after Range Focusing

Zeng Shang-chun<sup>①②</sup> Zhu Zhao-da<sup>②</sup>

<sup>①</sup>(College of Electrical & Information Engineering, Southwest University for Nationalities, Chengdu 610041, China)

<sup>②</sup>(Nanjing University of Aeronautics and Astronautics, College of Information Science and Technology, Nanjing 210016, China)

**Abstract:** It is difficult to directly compress Synthetic Aperture Radar (SAR) raw data for its low relativity. In this paper, a new algorithm is put forward. Range focusing is imposed to SAR raw data, which makes it have comparative high relativity, then a vector linear prediction is performed along the azimuth direction, and block adaptive quantization is used to the prediction error series. By using a real SAR raw data, compression and decompression are made respectively. The SQNR and SDNR are achieved. The images correspond to the three algorithms are gained. The experiments manifest that with same bit rate, SQNR and SDNR after using the algorithm proposed in this paper surpass that of BAQ. The calculation in this paper is far less than that of compression method after range focusing advised in corresponding reference.

**Key words:** Synthetic Aperture Radar (SAR); Range focusing; Vector linear prediction; Signal to Noise Ratio (SNR); Block Adaptive Quantization (BAQ); Bit rate

### 1 引言

合成孔径雷达(SAR)是一种新型雷达系统,具有常规雷达无法比拟的距离向和方位向二维高分辨特性,以及成像范围广阔和全天候等特点。现代空载合成孔径雷达系统包括发射、接收单元和模数转换等硬件处理系统以及实时数据传送或存储设备。SAR系统面临的一个重要问题是数据率高,难以实时下传,所以必须降低SAR原始数据率以满足有限的下传链路带宽。由于SAR原始数据的高熵值特性,无损压缩算法难以满足大的压缩比要求,所以目前普遍采用有损压缩技术。

在SAR原始数据压缩领域,目前已研究了多种压缩算法,如BAQ<sup>[1]</sup>,UPQ<sup>[2]</sup>,矢量量化<sup>[3]</sup>,FFT-BAQ<sup>[4]</sup>以及TCVQ<sup>[5]</sup>等。由于SAR原始数据的相关性很低,所以直接压缩SAR原始数据是比较困难的,通过并不复杂的距离聚焦,可以使

SAR数据在方位向具有较强的相关性,这会使SAR原始数据变得易于压缩,这是SAR原始数据压缩领域极具吸引力的一个研究方向。文献[6]研究了SAR原始数据距离聚焦后采用可变速率矢量量化,在性能上超过了传统的非聚焦SAR原始数据压缩算法。由于采用矢量量化,该方法的计算复杂度是比较大的。文献[7]采用TCVQ编码来压缩距离聚焦后的SAR数据,为了降低计算复杂度,使用了比文献[6]更短的码字,码字长度的减小必将导致性能的降低,但是TCVQ算法码本空间的划分更合理,这将减小量化误差。因此文献[7]采用的方法在性能上与文献[6]相当,但计算复杂度略小。由于矢量量化的计算量随码字长度作指数增加,这极大限制了上述两种方法在SAR原始数据实时传输中的运用。所以,对距离聚焦后的SAR数据,有必要研究复杂度更小的算法。在本文中,先对SAR原始数据作距离聚焦处理,再沿方位作矢量线性预测<sup>[8]</sup>,最后对预测残差作分块自适应量化。实验证明,本文算法的计算量远小于文献[6]的方法,而性能比该方法略低。

## 2 算法分析

### 2.1 距离聚焦

SAR 回波数据是方位和距离向的二维复矩阵,实部和虚部相互独立,均为零均值高斯分布,具有相同的方差。因此,

可以将实部和虚部分开来处理,需要处理的数据块大小为  $2048 \times 2048$ (方位 $\times$ 距离)。SAR 聚焦处理通常分两步进行:(1)距离聚焦;(2)依赖于距离的方位聚焦。方位聚焦是比较复杂的,它的参考函数与斜距  $r$  有关,而距离聚焦对应于每个回波的参考函数(与发射信号有关)是相同的,相当于回波信号通过一个相同的线性滤波器。所以本文只进行距离聚焦处理。

### 2.2 数据相关性分析

随机变量  $X$ ,  $Y$  之间的相关系数定义如下

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)D(Y)}} \quad (1)$$

式(1)中,  $\text{cov}(X, Y)$  表示  $X$ ,  $Y$  之间的协方差,  $D(X)$ ,  $D(Y)$  表示方差,分别定义如下

$$\begin{aligned} \text{cov}(X, Y) &= E\{[X - E(X)][Y - E(Y)]\} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [x - E(X)][y - E(Y)] p_{XY}(x, y) dx dy \quad (2) \end{aligned}$$

$$D(X) = E\{[X - E(X)]^2\} = \int_{-\infty}^{+\infty} [x - E(X)]^2 p_X(x) dx \quad (3)$$

$$D(Y) = E\{[Y - E(Y)]^2\} = \int_{-\infty}^{+\infty} [y - E(Y)]^2 p_Y(y) dy \quad (4)$$

式(2)中  $p_{XY}(x, y)$  为  $X$ ,  $Y$  之间的联合概率密度,式(3)中  $p_X(x)$  和式(4)中  $p_Y(y)$  为边缘概率密度。

表1、表2分别给出了一块 SAR 原始数据实部及距离聚焦后的 SAR 数据实部的行间(方位向)相关系数和列间(距离向)相关系数(因为同一块 SAR 数据实部和虚部的相关性差别很小,所以表中只给出了实部数据的相关性值),可以看出,经过距离聚焦后的 SAR 数据行间相关系数变化较大,而列间相关系数变化很小,也就是说距离聚焦后的 SAR 数据在方位向呈现出较强的相关性,而距离向相关性仍然很弱。

表1 行间(方位向)相关系数

阶数	1	2	3	4	5
SAR 原始数据	0.276529	0.026456	0.005475	0.004203	0.002426
距离聚焦后数据	0.574028	0.118157	0.031627	0.005648	0.004357

表2 列间(距离向)相关系数

阶数	1	2	3	4	5
SAR 原始数据	0.210556	0.135121	0.118994	0.048223	0.036329
距离聚焦后数据	0.223461	0.14399	0.120682	0.051468	0.038206

### 2.3 距离聚焦后数据统计特性

距离聚焦后的 SAR 数据具有比原始数据更大的动态范围,所以直接对它进行压缩是比较困难的,借鉴 BAQ 算法的思想,本文仍然对  $2048 \times 2048$  的大数据块进行分块归一化,分块大小取  $32 \times 32$ 。图1示出了  $2048 \times 2048$ (方位 $\times$ 距离)数据块实部的统计特性(由于所取数据为一方阵,因此它包含了距离向和方位向的统计特性),容易看出它具有高斯分布特性。由此可以得出,由于距离聚焦相当于一个线性变换,因此并没有改变 SAR 原始数据的高斯分布特性。对  $2048 \times 2048$ (方位 $\times$ 距离)数据块进行分块( $32 \times 32$ )归一化,每个小块数据都为标准高斯分布,因此,由这些小块( $32 \times 32$ )组合成的大块数据( $2048 \times 2048$ )也为标准高斯分布。

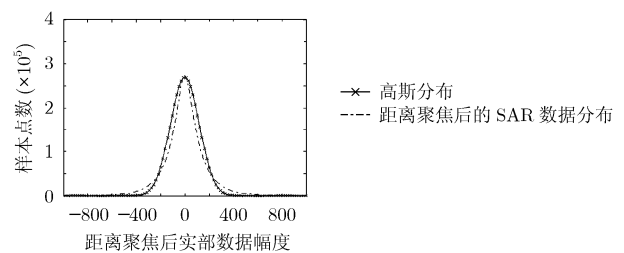


图1 距离聚焦后数据实部概率分布

### 2.4 预测残差序列统计特性及 BAQ 量化

从理论上说,随着预测阶数的增加,计算量也会增大。本文在方位向采用3阶矢量线性预测,矢量长度为3。如果对不同距离门数据都设计一个预测器会使计算量过大,不利于实时处理,本文采用一种简化的方法,每间隔512个距离门设计一个预测器,实验表明这样处理带来的误差是很小的。另外,由于距离聚焦后的 SAR 数据动态范围很大,因此应先分块归一化(分块大小  $32 \times 32$ )后再作线性预测。由于 SAR 成像区域很广,不同目标的后向散射系数也不一样,所以,不排除某些数据块会偏离高斯分布。文献[9]指出,如果原序列呈渐进高斯分布,则预测得到的残差序列比原序列更具有高斯分布特性。另外残差序列接近0值的数目显著增多,动态范围小,这会更有利于量化。对矢量线性预测得到的残差序列采用分块自适应量化,分块大小应遵循以下原则:分块必须足够小以保证每一分块中估计出的方差值恒定,分块必须足够大以保证能有效地估计出每一分块的方差值。一般分块大小取  $32 \times 32$  就能满足上述原则。由于 SAR 原始数据是距离和方位向的二维随机信号,每一个小块数据( $32 \times 32$ )相当于大块数据( $2048 \times 2048$ )的一个抽样,因此它们具有相同的分布特性,均为高斯分布,对小块数据进行归一化处理,就得到标准高斯分布数据。另外,本文采用最小均方误差准则,因此 BAQ 量化器是最优量化器。各数据块( $32 \times 32$ )采用相同的比特量化。

## 3 实验结果及分析

本文采用一组实测正侧视机载 SAR 原始数据,系统主

要参数:脉冲宽度  $9.2\mu\text{s}$ , 脉冲重复频率  $735\text{Hz}$ , 带宽  $60\text{MHz}$ , 采样频率  $73.33\text{MHz}$ 。由于 BAQ 算法是目前实用的比较成熟的 SAR 原始数据压缩算法, 因此将本文算法的结果与 BAQ 算法作比较。另外, 文献[6]与文献[7]给出的两种算法性能相当, 为进一步比较, 本文计算了文献[6]所示算法的性能参数。表 3 是 3 种算法在不同比特率下的计算结果。

本文采用数据域信噪比(SQNR)和图像域信噪比(SDNR)为压缩效果评价参数。同时计算了图像域的空间分辨率( $\rho$ )、峰值旁瓣比(PSLR)和积分旁瓣比(ISLR)<sup>[10]</sup>。可以看出, 在相同比特率条件下, 本文算法的数据域信噪比和图像域信噪比均比 BAQ 算法高, 但略低于文献[6]所示算法(我们在实验中采用了 3 组数据, 经过多次实验都证实了这一结论, 限于篇幅, 本文只给了一组实验数据的结果)。在比特率分别为 1, 2 和 3 比特/采样的情况下, 本文算法的 SQNR 分别比 BAQ 算法高  $1.75\text{--}2.77\text{dB}$ , 本文算法的 SDNR 比 BAQ 算法高  $1.53\text{--}2.29\text{dB}$ ; 本文算法的 SQNR 比文献[6]所示算法低  $0.24\text{--}0.65\text{dB}$ , 本文算法的 SDNR 比文献[6]所示算法低  $0.18\text{--}0.64\text{dB}$ 。同时可以看到, 在不同比特率时, 分辨率的变化很小, PSLR 和 ISLR 变化也不大。这说明数据压缩几乎不影响 SAR 图像的分辨率, 这同时也说明 SAR 图像对原始数据的量化比特数不敏感。另外, 在同一比特率下, 同一种算法的 SDNR 总是比 SQNR 高, 这是由于 SAR 成像处理时还要进行方位聚焦, 这一过程对信号进行相干叠加而对噪声进行非相干叠加, 这样就部分抵消了噪声对图像的影响, 增加了图像域信噪比。

结合上述 3 种压缩算法的性能优劣, 它们的应用也有一定的针对性, BAQ 算法主要用于简单快速、对性能要求不高的 SAR 实时数据压缩中; 本文算法主要用于性能较好的 SAR 实时数据压缩中; 文献[6]算法主要用于对性能要求很高、对实时性要求不高的 SAR 数据压缩中。

图 2 为 BAQ 算法所成的图像, 图 3 为文献[6]算法所成的图像, 图 4 为本文算法所成的图像(截图), 均采用 R-D 成



图 2 BAQ 算法所成的图像(比特率: 2 比特/采样)



图 3 文献[6]算法所成的图像(比特率: 2 比特/采样)



图 4 本文算法所成的图像(比特率: 2 比特/采样)

像算法<sup>[11]</sup>。限于篇幅, 仅给出比特率为 2 比特/采样时的图像。

#### 4 算法复杂度分析

设  $N_r$  为距离向数据长度,  $N_a$  为方位向数据长度,  $N$  为对原始数据归一化操作中数据块的长度,  $k$  为矢量线性预测器的矢量长度,  $m$  为矢量线性预测器的阶数。在算法的运算复杂度方面, 由于 DSP 对加法、比较及乘法运算都可在同一时钟周期内完成, 因此将算法中每一个浮点数加法、比较和乘法各看成一次操作。BAQ 算法量化一个采样值需要 2 次操作估计均值、2 次操作进行数据归一化及  $2^b$  ( $b$  表示量化比特数)次操作进行 Lloyd-Max 编码。本文算法中距离聚焦相当于卷积运算, 对一个回波数据 ( $N_r$  个复数)共需要  $N_r^2$  次复数乘法和  $N_r^2$  次复数加法, 每一次复数乘法需要 6 次操作, 所以每个浮点数对应的操作数为

$$\frac{N_r^2 \times 6 + N_r^2 \times 2}{2N_r} = 4N_r \quad (5)$$

如果用快速傅里叶变换将时域卷积转化为频域乘积处理, 则每个浮点数对应的操作数降为

$$\frac{(2N_r \log_2 N_r \times 2 + N_r) \times 6}{2N_r} = (4 \log_2 N_r + 1) \times 3 \quad (6)$$

表 3 3 种压缩算法在不同比特率下的性能参数

算法	压缩比	比特率 (比特/采样)	SQNR (dB)	SDNR (dB)	分辨率(m)		PSLR(dB)		ISLR(dB)	
					距离	方位	距离	方位	距离	方位
BAQ 算法	8	1	6.52	8.12	3.49	2.63	-8.7	-10.87	-5.64	-6.29
BAQ 算法	4	2	11.38	14.16	3.35	2.59	-9.13	-11.12	-6.18	-7.06
BAQ 算法	2.67	3	16.64	20.09	3.32	2.54	-9.56	-10.96	-6.64	-6.84
文献[6]算法	8	1	8.51	9.83	3.71	2.57	-8.68	-11.58	-6.46	-6.13
文献[6]算法	4	2	14.21	16.63	3.36	2.64	-9.37	-10.9	-6.47	-6.89
文献[6]算法	2.67	3	20.06	23.02	3.35	2.62	-9.54	-10.62	-6.39	-6.92
本文算法	8	1	8.27	9.65	3.51	2.58	-8.69	-11.41	-6.37	-6.08
本文算法	4	2	13.79	16.17	3.36	2.67	-9.31	-10.92	-6.58	-6.76
本文算法	2.67	3	19.41	22.38	3.34	2.61	-9.52	-10.76	-6.81	-6.27
SAR 原始数据	-	-	-	-	3.31	2.62	-9.58	-10.72	-6.79	-6.96

对  $m$  阶矢量线性预测, 每个浮点数需要  $(2km + m/k)$  次操作, 对预测差值序列作块自适应量化, 每个浮点数需要  $(2^b + 4)$  次操作。因此本文算法每个浮点数需要的操作次数为

$$(4\log_2 N_r + 1) \times 3 + 2km + m/k + 4 + 2^b \\ = 12\log_2 N_r + 2km + m/k + 2^b + 7 \quad (7)$$

文献[6]算法对距离聚焦的 SAR 数据采样矢量量化, 设比特率为  $R$ , 码字长度为  $L$ , 则码字搜索过程中每一分量(浮点数)对应的操作次数为  $2L \cdot 2^{RL}$ , 故该算法每个浮点数需要的操作次数为

$$(4\log_2 N_r + 1) \times 3 + 2L \cdot 2^{RL} = 12\log_2 N_r + 2L \cdot 2^{RL} + 3 \quad (8)$$

就算法的空间复杂度来看, BAQ 算法的 RAM 存储区主要用来存放原始数据、估计数据块的均值和方差以及对数据进行归一化操作, 其中估计方差和均值有简化的方法而占用更少内存。文献[6]方法由于要对距离聚焦后数据作分块归一化处理, 故 RAM 存储区大小与 BAQ 算法相同。本文算法的 RAM 存储区主要用来存放进行距离聚焦的 SAR 数据和沿方位向预测得到的差值序列。

3 种算法需要的缓存空间和运算量大小如表 4。在本文实验中, 表 4 的一些变量取值如下,  $N_r = N_a = 2048$ ,  $N = 32$ ,  $b = R = 2$ ,  $L = 4$ ,  $k = m = 3$ 。表 5 定量给出了存储空间、运算量及微机运算时间值。容易看出, 本文算法所需的存储空间比文献[6]算法大, 但运算量及运算时间比文献[6]算法大大减小。

表 4 算法复杂度分析

算法	RAM 存储空间 (浮点数)	运算量(操作/采样)
BAQ	$2 \times N_r \times N$	$2^b + 4$
文献[6]算法	$2 \times N_r \times N$	$12\log_2 N_r + 2L \cdot 2^{RL} + 3$
本文算法	$2 \times N_r \times N_a$	$12\log_2 N_r + 2km + m/k + 2^b + 7$

表 5 3 种算法存储空间、运算量及运算时间

算法	RAM 存储空间 (浮点数)	运算量 (操作/采样)	运算时间(s)
BAQ	131072	8	0.23
文献[6]算法	131072	2183	5.36
本文算法	8388608	162	1.21

## 5 结束语

SAR 原始数据作距离聚焦后在方位向呈现出较强的相关性, 这有利于数据压缩。本文对 SAR 原始数据作距离聚焦后, 在方位向采用矢量线性预测, 并对得到的残差序列采用分块自适应量化。实验表明, 在相同比特率条件下, 本文

算法的数据域信噪比和图像域信噪比均比 BAQ 算法高, 但略低于文献[6]所示算法。在比特率分别为 1、2 和 3 比特/采样的情况下, 本文算法的 SQNR 分别比 BAQ 算法高 1.75-2.77dB, 本文算法的 SDNR 比 BAQ 算法高 1.53-2.29dB; 同时可以得出, 本文算法的 SQNR 比文献[6]所示算法低 0.24-0.65dB, 本文算法的 SDNR 比文献[6]所示算法低 0.18-0.64dB。由于本文对预测得到的残差序列采用了标量量化, 与文献[6]相比, 本文算法的计算复杂度大大降低, 这对实时性要求很高的 SAR 原始数据压缩非常有利。本文算法具有一定实用价值。

## 参考文献

- [1] Kwok R and Johnson W T K. Block adaptive quantization of Magellan SAR data[J]. *IEEE Trans. on Geoscience and Remote Sensing*, 1989, 27(4): 375-383.
- [2] 朱永明, 朱岱寅, 朱兆达. 一种改进的 UPQ 算法在 SAR 原始数据压缩中的应用[J]. *空载雷达*, 2004, 145(1): 25-30.
- [3] 关振红, 朱岱寅, 朱兆达. 块自适应球形矢量量化算法压缩 SAR 原始数据[J]. *航空学报*, 2006, 27(1): 82-86.
- [4] Benz U, Strodl K, and Moreia A. A comparison of several algorithms for SAR raw data Compression[J]. *IEEE Trans. on Geoscience and Remote Sensing*, 1995, 33(5): 1266-1276.
- [5] Michael W, Marcellin, and Fischer T R. Trellis coded quantization of memoryless and Gauss-Markov sources[J]. *IEEE Trans. on Communications*, 1990, 38(1): 82-93.
- [6] Poggi G, Ragozini A R P, and Verdoliva L. Compression of SAR data through range focusing and variable-rate vector quantization[J]. *IEEE Trans. on Geoscience and Remote Sensing*, 2000, 38(3): 1282-1289.
- [7] D'Elia C, Poggi G, and Verdoliva L. Compression of SAR data through range focusing and variable-rate trellis-coded quantization[J]. *IEEE Trans. on Image Processing*, 2001, 10(9): 1278-1286.
- [8] Cuperman V and Gersho A. Vector predictive coding of speech at 16kb/s[J]. *IEEE Trans. on Communications*, 1983, 33(7): 685-696.
- [9] Magli E and Olmo G. Lossy predictive coding of SAR raw data[J]. *IEEE Trans. on Geoscience and Remote Sensing*, 2003, 41(5): 977-987.
- [10] 王晓军, 孙洪, 管鲍. SAR 图像相干斑抑制滤波性能评价[J]. *系统工程与电子技术*, 2004, 26(9): 1165-1171, 1183.
- [11] Smith A M. A new approach to range-Doppler SAR processing[J]. *Int. J. Remote Sensing*, 1991, 12(2): 235-251.

曾尚春: 男, 1971 年生, 博士生, 研究方向为雷达信号处理。

朱兆达: 男, 1939 年生, 教授, 博士生导师, 研究方向为雷达信号处理、信号检测。