

基于实测数据的原始数据压缩对 InSAR 系统测高影响研究

祁海明^{①②} 禹卫东^①

^①(中国科学院电子学研究所 北京 100080)

^②(中国科学院研究生院 北京 100039)

摘要: 为了研究原始数据压缩对InSAR系统干涉测高的影响。该文应用4组典型高程起伏的ERS-1/2实测数据,对Block Adaptive Quantization (BAQ) 与 Amplitude-Phase (AP) 两种有重要工程应用潜力的算法对干涉测高影响进行研究。试验结果表明: BAQ算法在设计时,要合理选择数据分块尺寸;同时,对于测高精度要求较高且垂直基线较短的系统,量化位数应至少为4 bit。AP算法在设计时,要合理分配幅度和相位量化位数。对于平均比特率为3 bit/sample, AP算法最优方案为幅度2 bit、相位4 bit,其性能介于3 bit BAQ和4 bit BAQ之间。

关键词: 合成孔径雷达; 原始数据; 压缩; 干涉测量

中图分类号: TN957.52

文献标识码: A

文章编号: 1009-5896(2008)11-2693-05

Study of Effect of Raw Data Compression on Space-borne InSAR Interferometry Based on Real Data

Qi Hai-ming^{①②} Yu Wei-dong^①

^①(Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China)

^②(Graduate School, Chinese Academy of Sciences, Beijing 100039, China)

Abstract: In order to study the effect of raw data compression on space-borne InSAR interferometry. This paper utilizes 4 typical ERS-1/2 raw data of different altitude to evaluate Block Adaptive Quantization (BAQ) and Amplitude-Phase (AP) algorithm. The numerical experiments results show that when BAQ is used, the block size should be chosen carefully. At the same time, 4 bit/sample is the least bit rate for precise SAR interferometry when the vertical baseline is short. When AP algorithm is used, the quantization bits should be distributed suitably. When the average bit rate is 3 bit/sample, the best choice of AP is 2 bits for amplitude quantization and 4 bits for phase quantization which performance is between 3 bit BAQ and 4 bit BAQ.

Key words: Synthetic Aperture Radar (SAR); Raw data; Compression; Interferometry

1 引言

星载SAR原始数据量大,星载存储设备容量和数据传输带宽有限。同时,由于SAR系统正向高分辨率、多极化、三维测绘、宽测绘带、多频段、多工作模式等方向发展,SAR原始数据量将变得更加庞大,因此必须对回波数据作压缩处理。压缩算法一般可以分为3类:(1)标量压缩算法;(2)矢量压缩算法;(3)变换域压缩算法。其中标量压缩算法包括:BAQ^[1], AP^[2,3], BFPQ^[4] (Block Floating Point Quantization), FBAQ^[5,6] (Fuzzy BAQ), ECBAQ^[5,6] (Entropy-Constrained BAQ), FBAQ^[7] (Flexible Block Adaptive Quantization)。矢量压缩算法包括:VQ^[5-9] (Vector Quantization), BAVQ^[5,6] (Block Adaptive Vector Quantization), TCVCQ^[5] (Trellis Coded Vector Quantization)。变换域压缩算法包括:FFT-BAQ^[5,6,10,11], WT^[5] (Wavelet Transform based compression)。虽然矢量量

化的性能高于标量量化,但是较长的码字搜索时间,无法满足传输速率的要求。FFT-BAQ算法需要对大数据块做二维FFT运算,同时需要进行多普勒中心估计,难以实时实现。WT算法是一种较有前途的压缩算法,但实现复杂,目前距离应用还有较远的距离。BAQ算法由于实现的简便性,曾被分别应用于美国Magellan飞行器的SAR系统^[1]和ENVISAT^[12]的A-SAR上。AP算法和BAQ算法相比,运算量有所增加,但性能有所改善^[2,3]。

设计SAR系统时,需要针对不同的应用,合理地选择压缩算法和压缩比。当星载SAR系统作高程测量时,原始数据压缩将影响系统测高精度^[12]。本文应用European Space Agency (ESA)提供的4对ERS-1/2系统数据,详细分析BAQ和AP两种算法对重复轨道干涉测高的影响。文章结构如下:第2节介绍BAQ算法和AP算法的基本原理。第3节给出重复轨道干涉测高误差分析的原理。第4节定义了干涉测量的评估指标。第5节给出了详细的实验方案、实验结果及其分析。第6节为文章的结论。

2 BAQ 和 AP 算法原理

BAQ 算法利用 SAR 回波的统计特性, 将 I, Q 两路数据分成相同大小的子块, 对每一子块的数据利用其相邻块的标准差归一化后, 再应用 Lloyd-MAX 量化器进行量化。显然, BAQ 算法本质上是针对直角坐标格式数据的压缩算法, 即 I, Q 两路信号必须采用统一的压缩比压缩。由于 SAR 原始数据幅度和相位统计独立^[1], 对回波数据的幅度和相位采用不同的压缩比压缩, 可以在总压缩比不变的前提下, 提高压缩性能^[2, 3]。同时, 由于相位信息在 SAR 信号处理中占有重要地位, 对相位信息采用较低的压缩比有利于后续 SAR 信号处理。

3 干涉测高误差分析原理

InSAR 测高方式分为双天线单航过测高和单天线双航过测高两种。前者精度较高, 适用于机载系统。但对于星载系统, 需要较长的基线, 实现十分困难。目前除了 2000 年 2 月 11 日~22 日, 装载 SRTM(Shuttle Radar Topography Mission)的美国“奋进”号航天飞机上应用伸出 60m 长的天线外, 星载系统均应用单天线双航过测高方式。重复轨道干涉测量作为单天线双航过测高方式, 利用了星载 SAR 受外界扰动小, 轨道参数精确的特点。在国外星载 InSAR 系统中获得了重要的应用^[13, 14]。

重复轨道干涉测量的高程误差 dh 与相位噪声 $d\phi$ 之间的关系由式(1)^[14]给出。

$$dh = \frac{\lambda r \sin \theta}{4\pi B_n} d\phi \quad (1)$$

其中 B_n 为垂直基线长度。 r 为 SAR 距目标斜距。 θ 为 SAR 的视角。 λ 为信号波长。

4 评估指标定义

(1) 相干系数与相位标准差, 对于 SLC(Single Look Complex)图像对, 其相干系数定义为^[15]

$$\gamma = \frac{\sum_{i=1}^M \sum_{j=1}^N S_1(i, j) S_2^*(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N |S_1(i, j)|^2 \sum_{i=1}^M \sum_{j=1}^N |S_2(i, j)|^2}} \quad (2)$$

其中 M, N 分别为距离向和方位向窗口尺寸。 S_1, S_2 分别为 SLC 图像对的复像素点。

由式(1)求解 InSAR 测高误差, 需要求解相位标准差。InSAR 信号处理中, 相位标准差通常是通过相干系数映射得到的。就单视而言, 映射关系如式(3)所示, 多视结果可参考文献[15]。

$$\sigma_\phi = \sqrt{\frac{\pi^2}{3} - \pi \arcsin(|\gamma|) + \arcsin^2(|\gamma|) - \frac{\text{Li}_2(|\gamma|^2)}{2}} \quad (3)$$

其中 $\text{Li}_2(\bullet)$ 是欧拉以 2 为底的对数。

(2) 测高误差, 工程上常用两种方法测量 DEM 的高程精度: (a) 放置角反射器, 用 GPS 进行实地测量, 然后与干涉计

算得到的结果比较, 计算高程值的 RMS 误差。(b) 与现有的高精度 DEM 进行对比, 计算高程值的 RMS。上述方法实质是采用已知高程的控制点来求解干涉系统的测高精度。本文应用未压缩时, 在不同区间上, 求解相干系数的窗口中心位置作为所谓的“控制点”, 来衡量由于数据压缩而产生的高程误差。处理流程如图 1 所示:

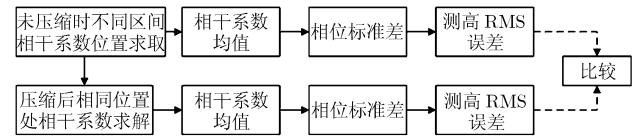


图1 本文计算高程RMS误差的流程图

具体步骤如下: 首先, 将相干系数分成若干区间, 如[0.49, 0.51], [0.59, 0.61]等。对于低相干系数区, 去相干引起的相位噪声远远大于原始数据压缩引起的相位噪声, 因此本文仅统计相干系数大于 0.35 的情况。其次, 在不同相干系数区间上, 求出未压缩情况下, 该区间相干系数对应的窗口中心位置。再次, 以窗口中心位置为“控制点”, 在压缩后的精配 SLC 图像对中, 求解上述相同窗口中心位置处的相干系数。最后, 将相干系数映射为相位的标准差, 代入式(1), 即可求得在此相干系数区间下的系统测高误差。

(3) 相对残差点数, 残差点数量反映了相位场无旋分量的水平, 残差点越多相位解缠越困难。本文应用的 4 组 ERS-1/2 数据是不同高程起伏地形的观测结果, 应用绝对残差点数无法类比数据压缩对于不同高程起伏类型数据干涉测高的影响。这里定义相对残差点数, 作为统一的标准。定义式如式(4)所示:

$$R_N = \frac{\text{Num2}}{\text{Num1}} \times 100\% \quad (4)$$

其中 Num1 和 Num2 分别为干涉图像像素数和经过数据压缩的干涉图残差点数。

5 数值实验

5.1 数据说明

本实验应用 ESA 提供的 4 组 ERS-1/2 原始数据。这 4 组数据是经过精心挑选的具有代表性数据。表现为 (1) 观测时间间隔为一天, 时间去相干影响较弱。(2) 高程起伏程度不同, 分别为较剧烈区域(Scene1&Scene2)、平坦区域(Scene3)、剧烈区域(Scene4)。(3) 具有多垂直基线长度, 便于分析不同垂直基线长度情况下的测高误差。具体数据参数汇总于表 1。

原始数据二维采样点数为 4096×4096 (距离 \times 方位)。图 2 给出相应场景的灰度图及干涉处理后的相位图(均为拉伸处理并截取的结果)。鉴于篇幅, 灰度图只列出 SLC 图像对中的一幅。其中, 图 2(a)~图 2(h) 分别为 Scene1~Scene4 的灰度图与干涉条纹图。干涉图均是去平地效应后的结果。

表 1 实验数据参数列表

场景编号	观测地点	Orbit	Frame	观测时间	中心视角	垂直基线长度
Scene1	Greenland	ERS1: 24118	2195	1996.2.24	23.721°	-103m
		ERS2: 4445	2187	1996.2.25		
Scene2	Greenland	ERS1: 23846	2259	1996.2.05	23.817°	-190m
		ERS2: 4173	2259	1996.2.06		
Scene3	Greenland	ERS1: 23760	2277	1996.1.30	23.725°	-154m
		ERS2: 4087	2277	1996.1.31		
Scene4	Italy Roma	ERS1: 39646	2763	1999.2.13	23.794°	211m
		ERS2: 19937	2763	1999.2.14		

reenland 常年积雪, 雪景在 SAR 图像上图 2(a)、图 2 (c)、图 2 (f) 表现为白色絮状景物。

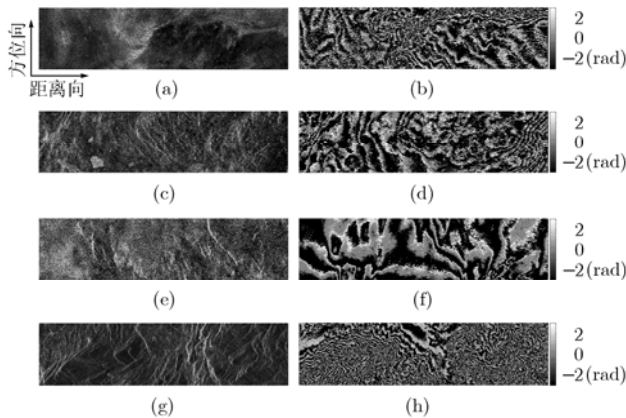


图 2 不同场景数据的灰度图与干涉条纹图(去平地后)

5.2 实验流程

数值实验的流程图如图 3 所示:

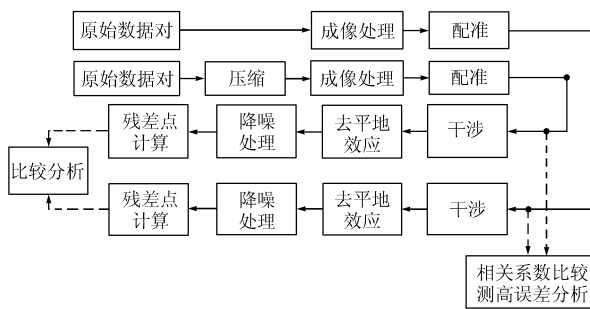


图 3 数据处理流程图

具体信号处理方法归纳如下: 压缩算法采用 BAQ 和 AP 算法。原始数据成像采用 RD 算法。SLC 图像对配准采用相干系数最大法进行, 即首先对其粗配准, 然后进行多项式拟合, 通过插值处理将配准精度提高到“亚像素级”。求解相干系数时的窗口尺寸为 5×5 。去平地效应采用距离向频谱偏移法。相位滤波采用窗口尺寸为 5×5 的矢量中值滤波器。干涉后没有进行多视处理。需要说明的是, 对不同场景、不同压缩比和不同分块下的 BAQ 与 AP 算法, 所有信号处理

参数完全相同, 这样保证了高程 RMS 误差仅由压缩算法本身引起。同时, 本文没有继续对相位降噪后的干涉图进行二维相位解缠, 这是因为二维相位解缠用于计算绝对 DEM 高程结果, 而估计数据压缩相对于没有数据压缩的高程误差, 应用式(1)即可^[12], 而且对于星载 InSAR 数据, 二维相位解缠往往需要人工干预^[12], 这样无法保证各数据处理之间的一致性, 即无法有效独立分析由于数据压缩引起的高程误差。

5.3 数值实验结果

5.3.1 测高误差分析 图 4 给出 Scene1~Scene4 在 64×64 分块下, BAQ 压缩相对未经原始数据压缩的高程 RMS 误差。

由图 4 可见: 总体上, 相干系数大的区间, 数据压缩所引起的测高误差较大。这是因为单视情况下, 相干系数较大区间的相位标准差对相干系数一阶导数的绝对值较大。因此, 在其他条件不变时, 相似的相干系数变化水平会在不同相干系数区间上引入不同的测高误差。具体表现为: 低相干系数区间上的测高误差小于高相干系数区间上的测高误差。需说明的是, 以上分析仅从理论上考虑, 对于实际数据, 这种趋势并不在每一条实验曲线的各实验点上严格满足, 但从总体上, 与理论分析相一致。

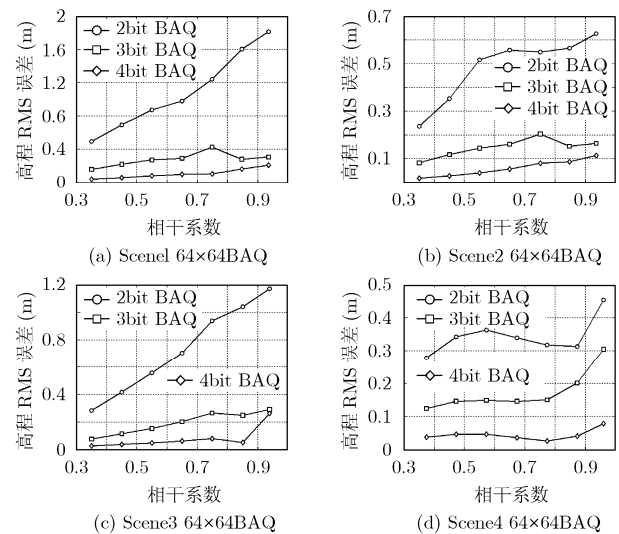


图 4 64×64 分块 BAQ 压缩相对于未压缩的各景数据干涉测高误差

图4中,垂直基线较长的场景(如 Scene2 和 Scene4),其高程 RMS 误差小于垂直基线较短的场景(如 Scene1 和 Scene3)。这是由于较长的垂直基线对应较高的高程测量精度,由式(1)可见,相似的相位噪声水平,较长垂直基线具有较小的高程 RMS 误差。因此,InSAR 系统压缩比的选择要综合考虑垂直基线的选择。较长垂直基线的 InSAR 系统即使压缩比较大,其高程 RMS 误差亦可小于较短垂直基线且压缩比较小的 InSAR 系统。同时,不同高程起伏数据的干涉测高误差变化趋势基本相同,说明分析结果对不同高程起伏类型数据具有普适性。

图5列出同一压缩比不同数据分块下的 BAQ 压缩相对于未经原始数据压缩的高程 RMS 误差。

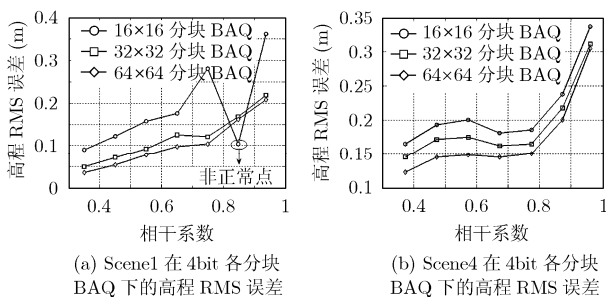


图5 同一压缩比不同分块下 BAQ 压缩相对于未压缩的高程 RMS 误差

由图5可见,随着分块尺寸增加,高程 RMS 误差减小。这种现象与分块内统计特性有关,因为 BAQ 的理论依据是无记忆信源满足标准高斯分布。块内数据较少时,数据统计特性偏离高斯分布程度较大,因此压缩引起的额外高程 RMS 误差增加。(图5(a)中的非正常点是由于该相关系数区间上样本点数少而引起的)实际选择分块尺寸时不仅要考虑满足数据统计特性,还要考虑硬件实现的复杂性,一般分块尺寸方位向要小于合成孔径点数,距离向要小于发射脉冲所对应的采样点数。

图6以 Scene4 为例给出 64×64 分块下,平均比特率为 3bit/sample 的 AP 算法与 3bit/sample 和 4bit/sample BAQ

压缩后高程 RMS 误差对比。其中 AP 算法标注 1-5 表示幅度 1bit/sample, 相位 5bit/sample, 其余照此类推。

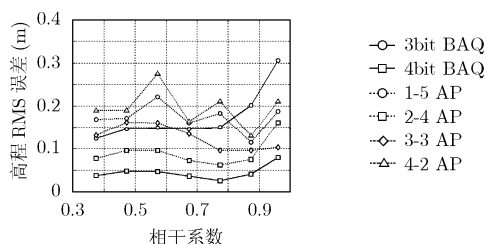


图6 AP 算法与 BAQ 算法高程 RMS 误差对比

由图6可见,高程 RMS 误差从小到大的 AP 算法方案依次为:(1)幅度 2bit, 相位 4bit; (2)幅度 3bit, 相位 3bit; (3)幅度 1bit, 相位 5bit; (4)幅度 4bit, 相位 2bit。说明仅仅通过减少幅度量化位数,来增加相位量化位数并不能总是取得最佳效果。同时方案(1)介于 3bit BAQ 和 4bit BAQ 之间,方案(2),方案(3),方案(4)不如 3bit BAQ。因此,如果选用 AP 算法时,对于平均比特率为 3bit/sample 应选择方案(1)。

5.3.2 相对残差点数计算结果 Scene1~Scene4 采用不同分块及压缩比 BAQ 压缩与未经原始数据压缩的相对残差点数计算结果列于表2。由表2可见:Scene2 和 Scene3 的相对残差点数较 Scene1 和 Scene4 少,这是因为 Scene2 和 Scene3 的地势平坦,而 Scene1 和 Scene4 对应高程起伏剧烈区域。高程起伏剧烈区域由于地形起伏引起的顶底位移、雷达阴影等因素都会引入不同程度的相干噪声,导致相位场的有旋性,因此干涉结果中相对残差点数较高。随着量化比特数的提高,Scene1~Scene4 的干涉图中,相对残差点数均有所降低。说明数据压缩对不同高程起伏特性具有普适性。

Scene1~Scene4 采用 64×64 分块 AP 算法压缩原始数据的残差点数计算结果如表3所示。

由表3可见,相对残差点数由少到多的 AP 压缩方案依次为(1)幅度 2bit, 相位 4bit; (2)幅度 3bit, 相位 3bit; (3)幅度 1bit, 相位 5bit; (4)幅度 4bit, 相位 2bit。这与高程 RMS 误差分析结果相一致。

表2 相对残差点数计算结果

场景编号	分块尺寸	2bit/sample (%)	3bit/sample (%)	4bit/sample (%)	未压缩 (%)
Scene1	16×16	3.194	2.948	2.887	2.869
	32×32	3.146	2.926	2.878	
	64×64	3.150	2.924	2.819	
Scene2	16×16	2.187	1.927	1.852	1.735
	32×32	2.085	1.834	1.812	
	64×64	2.094	1.829	1.777	
Scene3	16×16	1.879	1.421	1.358	1.237
	32×32	1.670	1.294	1.267	
	64×64	1.654	1.304	1.247	
Scene4	16×16	2.932	2.631	2.576	2.428
	32×32	2.883	2.523	2.444	
	64×64	2.875	2.521	2.437	

表3 AP 算法相对残差点数计算结果

场景编号	幅度 1bit 相位 5bit (%)	幅度 2bit 相位 4bit (%)	幅度 3bit 相位 3bit (%)	幅度 4bit 相位 2bit (%)	未压缩 (%)
Scene1	3.346	2.921	2.999	3.435	2.869
Scene2	2.194	1.813	1.931	2.371	1.735
Scene3	1.725	1.297	1.388	1.846	1.237
Scene4	2.623	2.502	2.604	3.206	2.428

6 结束语

BAQ 算法一方面要合理选择数据分块尺寸,另一方面要

结合系统对测高精度的指标要求以及系统选择的垂直基线, 设计相应的压缩比。数据压缩对 InSAR 系统测高影响对不同高程起伏数据具有普适性。需说明的是, 2bit/sample~4bit/sample 压缩后均可以干涉处理, 在数据试验中发现程序运行的流畅性没有破坏, 但对于高程 RMS 误差而言, 不同的量化比特数影响不同。对于测高精度要求较高且垂直基线较短时, 应至少选择 4bit/sample, 反之可选其它的方案。AP 算法仅仅通过减少幅度量化位数, 来增加相位量化位数并不能总是取得最佳效果。因此在实际应用时, 应该合理分配幅度和相位量化位数, 对于平均比特率为 3bit/sample 而言, 最优的分配方案为幅度 2bit, 相位 4bit, 其性能介于 3bit BAQ 和 4bit BAQ 之间。

致谢 感谢 ESA 为本研究工作提供的 ESR-1/2 实测原始数据。感谢 University of British Columbia 的 Ian G. Cumming 教授多次在 E-mail 中对作者的指导。感谢中科院电子所向茂生研究员在高程测量精度及高程误差计算上给与的指导。感谢中科院电子所的赵志伟、柳祥乐、陈兵、潘志刚 4 位博士及金廷满硕士与作者的有益讨论。

参 考 文 献

- [1] Kwok R and Johnson W T K. Block adaptive quantization of Magellan SAR data. *IEEE Trans. on Geosciences Remote Sensing*, 1989, 27(4): 375-383.
- [2] 姚世超, 王岩飞等. 合成孔径雷达原始数据幅相压缩算法. 电子与信息学报, 2002, 24(11): 1627-1633.
Yao Shi-chao and Wang Yan-fei, *et al.* Amplitude and phase compression algorithm for SAR raw data. *Journal of Electronics & information Technology*, 2002, 24(11): 1627-1633.
- [3] 潘志刚. 低比特率合成孔径雷达数据压缩算法研究. [博士论文]. 北京: 中科院电子学所, 2006.
- [4] Jordan R L and Huneycutt B L. The SIR-C/X-SAR Synthetic Aperture Radar system. *IEEE Trans. on Geosciences Remote Sensing*, 1995, 33(4): 829-839.
- [5] Boustani A E, and Branham K, *et al.* A review of current raw SAR data compression. Canadian Conference on Electrical and Computer Engineering, Toronto, Ont., Canada, May 2001: 925-930.
- [6] Benz U, Strodl K, and Moreria A. A comparison of several algorithms for SAR raw data compression. *IEEE Trans. Geosciences Remote Sensing*, 1994, 33(5): 1266-1276.
- [7] 关振红, 朱兆达. 使用矢量量化技术压缩 SAR 原始数据. 南京理工大学学报, 2006, 30(2): 157-161.
Guan Zhen-hong and Zhu Zhao-da. SAR raw data compression using vector quantization techniques. *Journal of Nanjing University of Science and Technology*, 2006, 30(2): 157-161.
- [8] 关振红, 朱兆达等. 块自适应球形矢量量化算法压缩 SAR 原始数据. 航空学报, 2005, 27(1): 82-86.
Guan Zhen-hong and Zhu Zhao-da, *et al.* Compression of SAR raw data by block adaptive spherical vector quantization, *ACTA Aeronautica et Astronautica Sinica*, 2005, 27(1): 82-86.
- [9] 关振红, 朱岱寅等. 归一化自适应预测矢量量化算法压缩 SAR 原始数据. 电子与信息学报, 2006, 28(3): 507-511.
Guan Zhen-hong and Zhu Dai-yin, *et al.* Compression of SAR raw data with normalized adaptive predictive vector quantization. *Journal of Electronics and information Technology*, 2006, 28(3): 507-511.
- [10] Fischer J, and Benz U, *et al.* Efficient SAR raw data compression in frequency domain. In proceedings IGARSS'99 Hamburg Germany, 1999: 2261-2263.
- [11] 杨云志, 黄顺吉等. SAR 原始数据的频域量化编码方法研究. 系统工程与电子技术, 2005, 27(12): 2107-2111.
Yang Yun-zhi and Huang Shun-ji, *et al.* Frequency region quantization coding method researching of SAR raw data. *Systems Engineering & Electronics*, 2005, 27(12): 2107- 2111.
- [12] Mcleod I and Cumming I. ENVISAT ASAR data reduction: Impact on SAR interferometry. *IEEE Trans. on Geosciences Remote Sensing*, 1998, 36(2): 589-602.
- [13] 舒宁. 雷达影像干涉测量原理. 武汉: 武汉大学出版社, 2003, 第2章.
- [14] Abdelfattah R and Nicolas J M. Topographic SAR interferometry for high-precision DEM generation. *IEEE Trans. on Geosciences and Remote Sensing*, 2002, 40(11): 2415-2426.
- [15] 王超, 张红等. 星载合成孔径雷达干涉测量. 北京: 科学出版社, 2002, 第1-3章.

祁海明: 男, 1980年生, 博士生, 研究方向为 SAR 数据压缩技术与 SAR 系统设计.

禹卫东: 男, 1969年生, 研究员, 博士生导师, 研究方向为 SAR 系统设计与 SAR 信号处理技术.