

归一化自适应预测矢量量化算法压缩SAR原始数据

关振红 朱岱寅 朱兆达

(南京航空航天大学信息学院 南京 210016)

摘要 该文提出归一化自适应预测矢量量化(NAPVQ)算法压缩 SAR 原始数据。NAPVQ 算法先采用矢量线性预测器对输入矢量进行预测, 再对原矢量与预测矢量之间的残差矢量进行矢量量化。该算法可视为差分脉冲调制在矢量量化中的拓展, 其性能优于块自适应量化(BAVQ)算法以及归一化预测自适应量化(NPAQ)算法。对算法复杂度的进一步分析表明, NAPVQ 算法能获得复杂度和性能之间比较合理的折衷, 具有实用价值。

关键词 合成孔径雷达, 数据压缩, 差分脉冲调制, 矢量量化, 预测矢量量化

中图分类号: TN958, TN957.51 文献标识码: A 文章编号: 1009-5896(2006)03-0507-05

Compression of SAR Raw Data with Normalized Adaptive Predictive Vector Quantization

Guan Zhen-hong Zhu Dai-yin Zhu Zhao-da

(College of Information Science and Technology, NUA, Nanjing 210016, China)

Abstract This paper presents a new SAR raw data compression algorithm named Normalized Adaptive Predictive Vector Quantization (NAPVQ). The normalized SAR raw data are firstly processed with vector linear predictor, and then the error vectors are compressed with vector quantizer. Regarded as an extension of the differential pulse code modulation in the vector quantization, the NAPVQ achieves a better performance gain than the Block Adaptive Vector Quantization (BAVQ) and the Normalized Predictive Adaptive Quantization (NPAQ). With the analyses of algorithm complexity, the proposed algorithm shows a good performance/complexity trade-off and accommodates the requirement in engineering applications.

Key words Synthetic Aperture Radar (SAR), Data compression, Differential pulse code modulation, Vector quantization, Predictive vector quantization

1 引言

SAR 原始数据压缩最常用的算法是块自适应量化(BAQ)^[1], 因为其简单和易于实施而被应用于美国火星探测的 Magellan 任务, 在美国太空总署的 SIR-C/X-SAR 任务中也采用了此算法进行数据传输。目前新提出的算法几乎都是以 BAQ 算法为基础, 而性能的改善也是与比 BAQ 算法性能提高程度作为参考, 因此 BAQ 算法已经成为 SAR 原始数据压缩领域的“事实上”标准^[2]。目前国内外已经提出的算法主要包括基于矢量量化技术的块自适应矢量量化(BAVQ)^[3,4]、网格型矢量量化(TCVQ)^[5]、基于标量预测技术的归一化预测量化(NPAQ)^[2]及基于快速傅里叶变换和小波变换的变换编码算法^[6,7]等。BAVQ 算法和 TCVQ 算法性能较好, 但运算量大; NPAQ 算法性能优于 BAQ 算法, 但采用标量预测技术, 因此性能提高有限; 变换编码算法性能也比较好, 但复杂度高, 并且性能受图像的分辨率等因素影响

较大。算法的性能和复杂度是一对矛盾, 在 SAR 原始数据压缩领域中, 要寻求算法性能和效率之间的比较合理的折衷。

本文提出归一化自适应预测矢量量化(NAPVQ)算法, 该算法可以看作NPAQ算法在矢量中的拓展, 并将记忆机制或矢量间的相关特性融入到矢量量化中, 其性能优于NPAQ算法和BAVQ算法。详细讨论了NAPVQ算法的原理、性能及复杂度等问题, 并且结合两组实测SAR原始回波数据进行了验证。

2 归一化自适应预测矢量器

首先简要介绍预测矢量量化器的基本原理, 然后在此基础上结合SAR原始数据的统计特性讨论归一化自适应预测矢量量化器的设计问题。

2.1 预测矢量量化器的基本原理

预测矢量量化器^[8]的基本结构如图1。将输入标量序列进行等间隔分段, 每段包含 k 个连续采样并组成一个输入矢量

x_n , x_n 的预测矢量为 \tilde{x}_n , 差值矢量 e_n 为

$$e_n = x_n - \tilde{x}_n \quad (1)$$

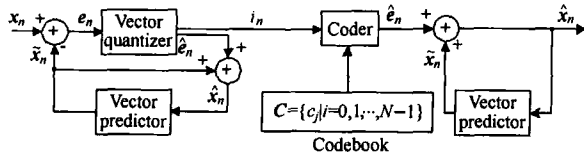


图1 预测矢量量化器的结构图

Fig.1 Block scheme of the predictive vector quantizer

通过量化 e_n 可以获得 $\hat{e}_n = Q(e_n)$ 以及码书索引 i_n , 相应的量化误差为

$$q_n = e_n - \hat{e}_n \quad (2)$$

解码端根据索引 i_n 得到差值矢量 \hat{e}_n , 此时输入矢量 x_n 的重构矢量 \hat{x}_n 为

$$\hat{x}_n = \tilde{x}_n + \hat{e}_n \quad (3)$$

如果矢量预测器采用形如 $\hat{y} = Ax$ 的一阶线性预测器, 则根据正交法则, 可得最优系数矩阵

$$A = E[x_n x_{n-1}^T] \left(E[x_{n-1} x_{n-1}^T] \right)^{-1} = R_{01} R_{11}^{-1} \quad (4)$$

式(4)中, R_{01} 和 R_{11} 是输入矢量 x_n 的相关矩阵, 由下式确定

$$R_{ij} = E[x_{n-i} x_{n-j}^T] \quad (5)$$

工程上实施NAPVQ算法时, R_{01} 和 R_{11} 可以通过SAR历史回波数据采取离线(不占用系统工作时间)的方式进行估计。

2.2 归一化自适应预测矢量器的设计

设计归一化自适应预测矢量器既需要设计矢量线性预测器, 还需要设计矢量量化器。一般来说, 设计预测矢量量化器可以采用预测器和量化器单独优化设计法或预测器和量化器联合优化设计方法^[8]。与联合优化设计方法相比, 单独优化设计方法运算量低, 并且性能下降的很小。考虑到SAR原始数据压缩的实时性要求高的特点, 本文采用单独优化的设计方法。

SAR原始数据的实部和虚部都是零均值的渐近高斯信号, 而高斯信号的最优矢量预测器是线性的^[9], 因此NAPVQ算法采用(一阶)矢量线性预测器。设计矢量线性预测器时, 将SAR原始回波信号看成自回归模型的输入信号, 通过估计历史数据获得相关系数矩阵 R_{01} 和 R_{11} 。因为 R_{11} 是Toeplitz矩阵, 通过Levinson-Durbin^[9]算法求出矢量线性预测器的最优系数矩阵 A 。需要注意的一点是, 估计相关系数矩阵的训练数据是经过归一化处理的数据, 而非SAR原始回波数据。所谓归一化处理是指利用SAR原始数据在小段时间间隔内信号的动态范围远小于整个回波数据集的动态范围, 把整个回波数据集沿方位向和距离向分成若干个小块, 每个小块可以认为是具有稳态特性的零均值高斯分布信号, 并且其分布特性可由数据块的方差唯一确定。一般地, 分块的大小为: 32×32 (方位 \times 距离)^[6]。

设计矢量量化器时, 可以采用开环、半闭环和闭环3种

方式^[8]。与其它两种方式相比, 开环方式最简单, 运算量也最低。开环方式设计矢量量化器时, 先利用设计好的矢量线性预测器以开环方式生成残差矢量训练序列, 然后根据给定训练序列以开环方式设计码书。具体步骤如下:

(1)对SAR原始数据进行分块归一化处理, 确定训练序列 $\{x_n; n = -M, \dots, 1, 2, \dots, L\}$ 及阈值 T 。

(2)对训练矢量序列 $\{x_n\}$ 进行开环预测, 获得预测矢量序列 $\{\tilde{x}_n\}$, 其中 $\tilde{x}_n = P(x_{n-1})$, 进而得到开环残差矢量序列 $\{e_n\}$, 其中 $e_n = x_n - \tilde{x}_n$ 。

(3)对残差矢量序列 $\{e_n\}$ 运用 LBG^[10]算法设计矢量量化器的码书, 直到矢量量化器的总体失真小于给定的阈值 T 。

矢量量化器采用开环方式设计时, 训练数据是开环残差矢量序列 $\{x_n - P(x_{n-1})\}$, 而系统实际量化的是闭环残差矢量序列 $\{x_n - P(\hat{x}_{n-1})\}$ 。由于闭环残差矢量和开环残差矢量统计特性并不完全相同, 因此矢量量化器的性能一般不能达到最优。

本文给出一种改进方法, 即采用半闭环方式设计矢量量化器, 其与开环方式相比, 设计复杂度略有增加, 但算法性能更好。半闭环方式在前述开环方式基础上增加如下步骤:

(4)根据(3)获得的码书, 对输入训练矢量序列 $\{x_n\}$ 进行预测, 获得闭环残差矢量序列 $\{\tilde{e}_n\}$, 其中 $\tilde{e}_n = x_n - P(\hat{x}_{n-1})$ 。

(5)根据闭环残差矢量序列 $\{\tilde{e}_n\}$ 训练码书, 直到矢量量化器的总体失真小于给定的阈值 T 。

采用半闭环方式设计的归一化自适应预测矢量器与开环方式相比, 性能有所提高, 而运算量增加有限。理论上, 矢量量化器设计完成之后, 还可以利用设计好的量化器改进预测器, 但此时算法开销增大, 而且性能增加很小。

3 实验结果

为了检验 NAPVQ 算法的性能, 使用实测数据进行压缩成像处理, 通过比较数据域和图像域的性能指标来评价算法性能, 实验数据是两部不同参数的实测机载 SAR 原始数据。考虑到 SAR 成像的特点, SAR 原始数据压缩的性能指标除了包括数据域信噪比(SNR)和图像域信号失真比(SDNR)外, 也对峰值旁瓣比(PSLR)、积分旁瓣比(ISLR)以及图像的空间分辨率^[11]等指标进行考察。表1和表2分别给出了采用2bit量化(压缩比为4)时, BAQ、NPAQ和BAVQ及NAPVQ4种算法量化两组实测数据的实验结果。

根据表1和表2, 压缩算法对SAR图像的影响主要体现在SNR和SDNR降低(增加了量化噪声), 其它性能指标与原始图像相比几乎没有变化(微小变化主要是由于测量造成的), 并且由于经过二维SAR成像处理降低了数据域的量化噪声对图像的影响, 所以SDNR一般要大于SNR。因此, 目前在SAR原始数据压缩领域, 评估压缩算法性能时选择的主要指标是SNR和SDNR。

表 1 4 种压缩算法压缩数据 1 的性能分析结果

Tab.1 Image quality analysis of data 1 using 4 compression algorithms

| 算 法 | 比特率 (比特/采样) | SNR (dB) | SDNR (dB) | 分辨率(<i>m</i>) | | PSLR (dB) | ISLR (dB) |
|-------------|----------------|----------|--------------|-----------------|------|--------------|--------------|
| | | | | 距离 | 方位 | | |
| 未压缩 | 8.00 | - | - | 3.06 | 3.44 | 17.53 | 11.52 |
| BAQ | 2.00 | 9.31 | 12.46 | 3.06 | 3.44 | 17.32 | 11.61 |
| NPAQ | 2.00 | 10.35 | 13.72 | 3.06 | 3.44 | 17.41 | 11.43 |
| BAVQ | 2.00 | 11.29 | 14.63 | 3.06 | 3.44 | 17.52 | 11.52 |
| NAPVQ (开环) | 2.00 | 12.17 | 15.32 | 3.06 | 3.44 | 17.41 | 11.48 |
| NAPAQ (半闭环) | 2.00 | 12.41 | 15.58 | 3.06 | 3.44 | 17.34 | 11.47 |

表 2 4 种压缩算法压缩数据 2 的性能分析结果

Tab.2 Image quality analysis of data 2 using 4 compression algorithms

| 算 法 | 比特率 (比特/采样) | SNR (dB) | SDNR (dB) | 分辨率(<i>m</i>) | | PSLR (dB) | ISLR (dB) |
|-------------|----------------|----------|-----------|-----------------|------|--------------|--------------|
| | | | | 距离 | 方位 | | |
| 未压缩 | 8.00 | - | - | 1.00 | 1.00 | 13.41 | 8.90 |
| BAQ | 2.00 | 9.60 | 12.92 | 1.00 | 1.00 | 13.51 | 8.83 |
| NPAQ | 2.00 | 10.78 | 13.99 | 1.00 | 1.00 | 13.24 | 8.92 |
| BAVQ | 2.00 | 11.52 | 14.70 | 1.00 | 1.00 | 13.22 | 8.58 |
| NAPVQ (开环) | 2.00 | 12.21 | 15.23 | 1.00 | 1.00 | 13.44 | 8.48 |
| NAPAQ (半闭环) | 2.00 | 12.25 | 15.37 | 1.00 | 1.00 | 13.44 | 8.37 |

图 2 和图 3 给出了采用本文讨论的算法压缩数据获得的图像与原始数据直接获得的图像(截取), 从图中可以看出随着算法性能的提高, 图像质量越来越好。

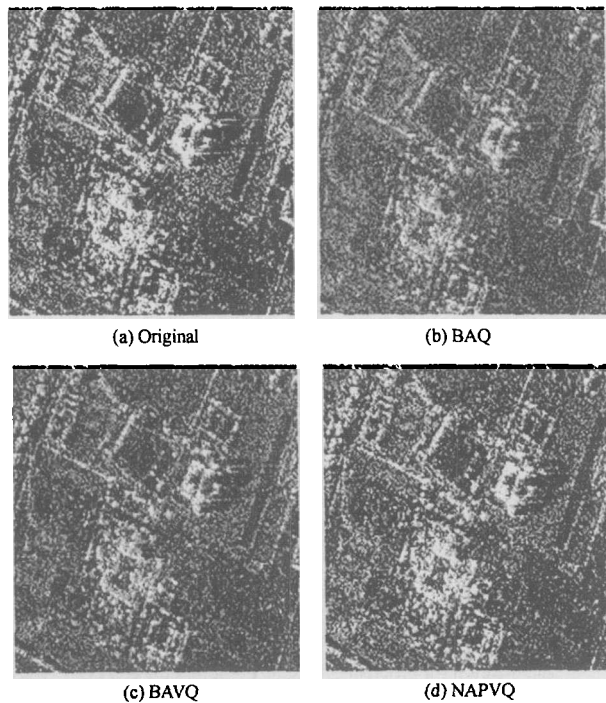


图 2 数据 1 采用不同压缩算法的成像结果

Fig.2 Images of data 1 obtained from different compression algorithms

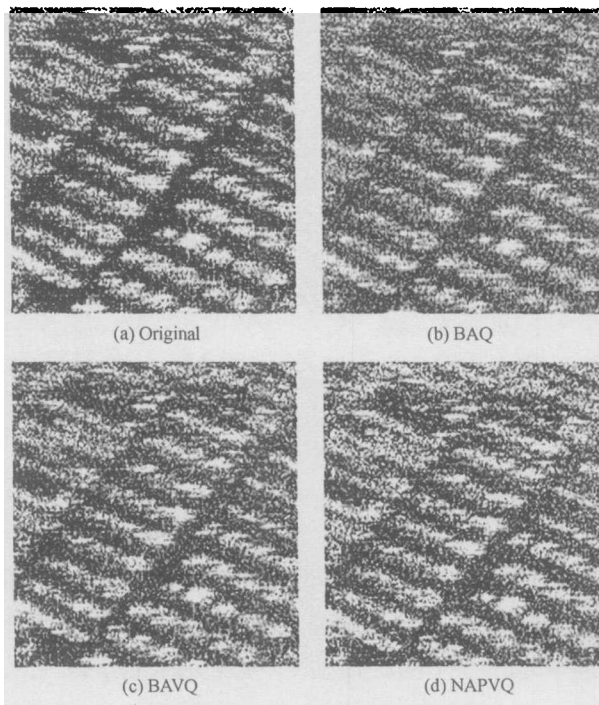


图 3 数据 2 采用不同压缩算法的成像结果

Fig.3 Images of data 2 obtained from different compression algorithms

4 算法性能及复杂度分析

一般来说, 如果增加矢量的维数, 总可以提高矢量量化

器的性能,但对于预测矢量量化器来说并非一直正确。分析矢量维数对 NAPVQ 算法性能的影响要综合矢量预测器和矢量量化器两方面因素。NAPVQ 算法的 SNR 由下式确定

$$SNR = 10 \lg G_{clp} + SNR_{vq} \quad (6)$$

式(6)中, G_{clp} 为闭环预测增益比, SNR_{vq} 为矢量量化器的信噪比。增加矢量维数可以增加矢量量化器的 SNR_{vq} , 但增加矢量维数会对闭环预测增益比 G_{clp} 的产生两方面的影响: 一方面增加矢量维数相当于增加了过去的观测值, 因此增加了闭环预测增益比; 另一方面相当于要预测更多的未来信息, 因此增加了预测器的负担从而降低了闭环预测增益比^[9]。

NAPVQ 算法的 SNR 是预测器的闭环预测增益和矢量量化器的信噪比综合作用的结果, 矢量维数的变化对 SDNR 的影响与 SNR 类似。图4给出了不同矢量维数下的 BAVQ 和 NAPVQ 算法的 SDNR 变化曲线。根据图4, 当矢量维数较低时, NAPVQ 算法比 BAVQ 算法的 SDNR 可以提高 1dB 以上; 但当矢量维数增加时, 预测矢量与原矢量误差增大, 因此性能的提高程度越来越小。

图5给出本文讨论的几种算法在不同输出编码比特率条件下性能变化曲线。从图5可以观察到, 与 NPAQ 算法和 BAQ 算法相比, NAPVQ 算法的性能提高显著, 而且也优于 BAVQ 算法。

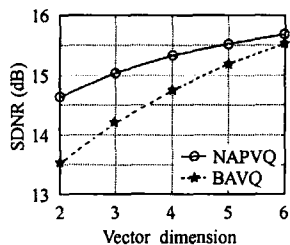


图4 不同矢量维数下的BAVQ和NAPVQ算法性能曲线
Fig.4 Performance curves of BAVQ and NAPVQ versus vector dimension

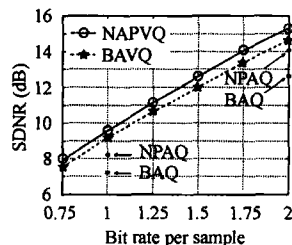


图5 BAQ、NPAQ、BAVQ和NAPVQ算法性能比较
Fig.5 Comparison of the performance of BAQ, NPAQ, BAVQ and NAPVQ

SAR 原始数据压缩算法不仅要考虑算法的性能, 还要求算法的运算量低并且占用存储空间少。表3给出了输出编码比特率为 R 时, 4种压缩算法硬件实现复杂度的结果。在表3中, N_r 为距离向数据长度、 N 是对原始数据归一化操作中数据块的长度、 p 为线性预测器的阶数、 L 为矢量量化算法中矢量的维数。设待压缩的 SAR 原始数据为 2048×2048 复采样数据, 则可取 $N_r = 2048$, $N = 32$ 。

根据表3, 4种算法使用相同大小的RAM, 但BAVQ和

表3 4种压缩算法的硬件需求和计算复杂度

Tab.3 Hardware requirements and computational complexity of 4 compression algorithms

| 算法 | 存储空间 | | 加法和乘法运算量 (运算次数/采样) |
|-------|-------------------------|----------|--------------------------|
| | RAM | ROM | |
| BAQ | $2 \times N_r \times N$ | --- | $4 + 2^R$ |
| NPAQ | $2 \times N_r \times N$ | -- | $4 + 2^R + 2(p+1)$ |
| BAVQ | $2 \times N_r \times N$ | 2^{LR} | $2 \times L2^{RL}$ |
| NAPVQ | $2 \times N_r \times N$ | 2^{LR} | $2 \times L2^{RL} + L^2$ |

NAPVQ 算法需要使用 ROM 存储码书。就算法的运算复杂度来看, BAQ 算法主要的运算是对每个数据块进行归一化操作, NPAQ 算法在 BAQ 算法的基础上增加了标量预测的操作, BAVQ 算法增加了搜索码书的操作, NAPVQ 算法在 BAVQ 算法基础上增加了矢量预测的操作。总的来说, NAPVQ 算法和 BAVQ 算法运算量基本相当。

5 结束语

数据压缩的目的是去除信号中的冗余信息, 冗余概念与预测信号概念密切相关, 文中提出的 NAPVQ 算法通过矢量预测技术去除矢量间的相关性, 可以用更少的比特数编码而保持重构信号和原信号之间的总误差不变。矢量量化器的增益比标量量化器高, 它补偿了矢量预测器的预测增益。采用适当维数的矢量(如4维), 预测矢量量化器的总增益仍比标量量化器高。总的来说, NAPVQ 算法具有较好的性能, 并且算法实现复杂度又在工程允许的范围, 因此可以获得算法性能与复杂度之间较为合理的折衷。随着 DSP 技术的发展, 使用矢量量化技术代替标量量化技术已经成为必然, 而 NAPVQ 算法是很好的选择。

参考文献

- [1] Kwok R. Adaptive quantization of Magellan SAR data, *IEEE Trans. on Geoscience and Remote Sensing*, 1989, 27(4): 375 - 383.
- [2] Enrico M, Olmo G. Lossy predictive coding of SAR raw data. *IEEE Trans. on Geoscience and Remote Sensing*, 2003, 41(3): 977 - 987.
- [3] Moreira A, Blasler F. Fusion of block adaptive and vector quantization for efficient SAR data compression. IGARSS, Tokyo, 1993: 1583 - 1585.
- [4] 杨云志, 黄顺吉, 王建国. 矢量量化在 SAR 原始数据压缩中应

- 用. 系统工程与电子技术, 2002, 24(6): 42 - 44.
- [5] Owens J W, Marcellin M W, Hunt B R. Compression of synthetic aperture radar video phase history data using trellis-coded quantization techniques. *IEEE Trans. on Geoscience and Remote Sensing*, 1999, 37(2): 1080 - 1085.
- [6] Benz U, Strodl K, Moreia A. A comparison of several algorithms for SAR raw data compression. *IEEE Trans. on Geoscience and Remote Sensing*, 1995, 33(5): 1266 - 1276.
- [7] Pascazio V, Schirinzi G. SAR raw data compression by subband coding. *IEEE Trans. on Geoscience and Remote Sensing*, 2003, 41(2): 964 - 976.
- [8] 孙圣和, 陆哲明. 矢量量化技术及应用. 北京: 科学出版社, 2002: 116 - 136.
- [9] Jayant N S, Noll P. *Digital Coding of Waveforms*, NJ: Prentice-Hall, 1984, 40 - 90.
- [10] Linde Y, Buzo A, Gray R. An algorithm for vector quantization design. *IEEE Trans. on Communications*, 1980, 28(1): 84 - 95.
- [11] 魏钟铨. 合成孔径雷达卫星. 北京: 科学出版社, 2001: 202 - 210.
- 关振红: 男, 1972年生, 博士生, 研究方向为雷达信号处理.
- 朱岱寅: 男, 1974年生, 博士, 副研究员, 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.