基于小波包变换的 SAR 原始数据压缩

胡晓新　　王岩飞
(中国科学院电子学研究所　北京　100080)

摘　要　该文介绍了小波包变换和网格编码量化的优点，结合合成孔径雷达(SAR)原始数据的特点，提出了基于小波包变换和网格编码量化的 SAR 原始数据压缩算法。利用模拟和实际数据分析了该压缩算法的性能，通过与其他原始数据压缩算法的比较说明了这一方法的有效性。

关键词　合成孔径雷达，数据压缩，小波包变换，网格编码量化
中图分类号：TN958　　**文献标识码**：A　　**文章编号**：1009-5896(2006)08-1476-04

SAR Raw Data Compression Using Wavelet Packet Transform

Hu Xiao-xin　　Wang Yan-fei
(*Institute of Electronics, Chinese Academy of Sciences, Beijing* 100080, *China*)

Abstract　The advantages of wavelet packet transform and Trellis-Coded Quantization (TCQ) are introduced and compression algorithm for SAR raw data based on wavelet packet transform and TCQ is given. Compression performance of the algorithm on simulated and real raw data analysis is presented and compared with other techniques by different quality parameters, which shows that the algorithm is suit for compression of SAR raw data.

Key words　Synthetic Aperture Radar(SAR), Data compression, Wavelet packet transform, Trellis-Coded Quantization (TCQ)

1　引言

合成孔径雷达(SAR)系统具有较高的分辨率、较大的测绘带宽和较大系统动态范围，所接收到的回波数据量相当大，同时 SAR 系统逐步向多极化、多波段、多模式发展会导致数据量越来越大，这与下传信道的有限带宽相矛盾。因而如何降低数据量，同时保持系统动态范围成为需要解决的热点问题。

根据SAR系统原理，SAR回波数据可以视为地面大量独立小反射体回波的叠加，因此SAR原始数据的I，Q两路信号为互相独立的近似零均值高斯分布信号，沿距离向和方位向方差缓变，具有类白噪的频谱，数据动态范围很大，熵值较高[1]。

针对SAR数据的特点和受小波变换编码用于SAR图像压缩的启发，Pascazio[2]提出将小波变换编码技术直接用于SAR原始数据，根据每一频带小波系数的统计特征分配量化比特，优化相应的量化器和熵编码器。Magli[3]将BAQ算法对I，Q系数的分块归一化的思想用于离散小波变换(DWT)去相关数据的过程，使得基于DWT的压缩算法受输入数据特性的统计波动影响减小，从而改善数据压缩的效果。

小波变换的良好“时频局部性”使得它常用于数据压缩中，但其具有“高频部分时间分辨率低”的问题，对于富含中高频分量的信号而言，值较大的变换系数所占比例往往高于预期值，会影响压缩比的提高。小波包变换考虑对高频分量的进一步分解，并根据分解信号的特征，按照一定的准则选择更恰当的频率分解，所以利用小波包变换可获得更好的压缩性能。Pascazio[4]首先对SAR原始数据进行小波包变换，然后根据不同频率成分的重要性分配比特数，对各频带进行相应的标量量化，从而实现对SAR原始数据的压缩，取得了较好的压缩效果。

为了进一步提高编码性能，本文引入网格编码量化(TCQ)小波包变换系数。TCQ 作为一种简化矢量量化方法，利用样点间的相关性，性能优于标量量化。因此本文结合小波包变换和 SAR 原始数据的特点，提出了基于小波包和网格编码量化的 SAR 原始数据压缩算法。通过计算解压缩后数据和成像后图像的相关质量参数，以及与其他原始数据压缩算法结果的比较说明了这一方法的有效性。

2　基于小波包变换和网格编码量化的压缩算法

本文提出的 SAR 原始数据压缩算法结合小波包变换和 TCQ 的优点，获得更大的编码增益。压缩算法主要包括预处理，小波包变换，比特率分配和 TCQ 4 个部分。算法框图如图 1 所示。

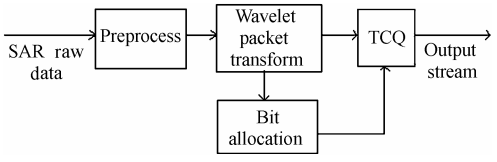

图 1　基于小波包和网格编码量化的 SAR 原始数据压缩算法
Fig.1 SAR raw data compression algorithm based on wavelet packet and TCQ

2004-12-13 收到，2005-03-30 改回

压缩算法的预处理中，原始数据块被划分成恰当大小的不重叠块，这些数据块是后续变换编码运算的基本单元，称为压缩数据块。所以小波包变换和量化编码可作用于小的压缩数据块，而不是整个数据集合。这有助于区别对待不同数据分块的不同统计特性，利用最优位分配方案改善整个数据集合的编码效益。此外这一操作有利于实现压缩算法的并行处理，从而提高算法的实时实现性。而通过对 SAR 模拟数据的分析发现数据分块处理后，大多数压缩数据块的动态范围减小。由于小波包变换的能量聚积作用容易导致硬件实现时出现数据溢出，因此信号动态范围小，有利于避免小波包变换系数溢出。数据块大小应该综合考虑实际存储计算能力和后期量化处理的要求。

SAR原始数据相关性差，熵值较高，不利于直接量化，因此借助小波包变换以获得易于量化编码的数据。基于小波变换，Wickerhauser, Coifman[5]等人提出小波包的概念，并进行了严密的数学推导。小波变换仅分解每一级的低频分量，将信号分解成一组多尺度空间中的展开系数。而小波包对每一级的各高频分量选择性地进行细分，克服了小波变换时间分辨率高时频率分辨率低的缺陷，可以获得适应于信号特征的频率划分。在小波包分析中，信号被映射到大量正交子空间 $W_{i,n}$ 上，信号空间 V 被分解成正交子空间，满足

$$V=\oplus_{i,n}W_{i,n} \tag{1}$$

其中 i 表示分解层，n 表示该层的第 n 个子空间，$\oplus$ 表示空间的直和。

小波包变换考虑进一步分解高频部分，提供了多种空间分解集合，所以可以按照一定的标准(代价函数)选取满足应用要求的空间分解及对应的小波基。本文选择的代价函数为最大熵代价函数，定义为

$$E(S)=-\sum_i \varepsilon(s_i)=-\sum_i s_i^2\log_2\left(s_i^2\right) \tag{2}$$

其中 s_i 为信号的小波包分解系数，规定 $s_i=0$ 时 $\varepsilon(s_i)=0$。具体的最佳小波基选取步骤参见文献[6]。

普通图像数据一般具有丰富的低频成分，且人眼对高频成分不敏感，所以压缩普通图像时简单地放弃高频信息也不会对人眼识别产生明显影响。与之不同，SAR 原始数据相关性差，信号分布在整个频带范围内，中、高频信息也很丰富，且对于地物识别和分类而言，高频分量可能包含小目标等重要信息。所以对于 SAR 原始数据，中高频子带的作用不能忽视。细化高频子带的小波包变换有助于获得中高频信息的准确表示，因此本文选择小波包变换，获得更恰当的频率分解，对不同子带采取不同的量化方案，达到要求的压缩比和失真限度。为了便于分析，图 2 给出了 2D 信号的两层完全小波包分解的空间结构图。

LL,LL	LL,LH	LH,LL	LH,LH
LL,HL	LL,HH	LH,HL	LH,HH
HL,LL	HL,HL	HH,LL	HH,LH
HL,HL	HL,HH	HH,HL	HH,HH

图 2　2D 信号两层完全小波包分解空间结构图
(H 表示高通，L 表示低通)
Fig.2 Space structure of 2D wavelet packet transform
(H:high-pass, L:low-pass)

为进一步说明小波包变换对于 SAR 原始数据压缩的作用，取模拟 SAR 原始数据进行 2 层完全 D4 小波包变换，分析变换系数各项统计特性，如表 1 所示。由表可见，信号的能量主要集中在 LL 和 LH 方向上，而 LH 和 HH 方向的能量较少。(LL,LL)，(LL,HL)，(LH,LH)和(LH,HH)这 4 个子带的数据量为总数据量的四分之一，而它们的能量超过总能量的 50%。因此 LL 和 LH 子带的系数对于信号重构影响很大。除了(LL,HH)，(LH,LL)和(HH,HH)子带以外，其他子带的均值都较小。LL 和 LH 方向上的方差高于 HL 和 HH 方向子带的方差，且(LL,LL)和(LH,LH)子带的方差和能量高于其他子带。在 HL 方向上，小波包系数的水平相关性基本大于垂直相关性，而 LH 方向上的垂直相关性大于水平相关性。

基于对小波包变换系数统计特性的分析，进一步研究如何选择相应的量化器和如何恰当地分配量化比特数。为了最大可能的合理利用编码资源，提高信噪比，根据不同子带的特点，分配量化比特数。本文考虑的位分配准则[4]是：

$$R_i=\bar{R}+\frac{1}{2}\log_2\sigma_i^2-\frac{1}{2}\sum_{i=1}^{K}\alpha_i\sigma_i^2 \tag{3}$$

其中 R_i 为第 i 个子带的比特率，σ_i^2 为第 i 个子带的方差，K 为子带数目，α_i 为子带相对面积且满足 $\sum_{i=1}^{K}\alpha_i=1$。

普通图像数据经过小波包变换后，各子带系数聚集在零点处附近，而 SAR 原始数据本身的统计特性导致其小波包系数不具有这样的特性。同时 SAR 数据小波包分解中，各子带数据内存在残余的水平和垂直相关性。所以本文利用网格编码量化(TCQ)获取更好的数据压缩性能。

TCQ源自Ungerboeck提出的网格编码调制(TCM)，由Marcelin等推广应用于压缩编码[7]。作为一种简化的矢量量化，它利用输入样点之间的相关性，使得编码器可以模拟大的矢量量化器的性能，而复杂性不像矢量量化器那样惊人。同时相对于标量量化，网格编码使用非矩形的量化单元，可以获得更大的量化有效性。所以即使用于编码不相关的数据，由于其独特的编码结构和编码大数据块的特点，编码性能也能得到补偿，从而获得更好的压缩效果。

表 1 小波包变换系数的统计特性 (D4 小波，两层完全分解)

Tab.1 Statistical properties of wavelet packet coefficient

子带	最大值	最小值	均值	均方差	水平 相关系数	垂直 相关系数	能量 百分比(%)
LL,LL	151.89	− 142.98	0.47	48.09	0.4113	0.0928	22.44
LL,HL	89.20	− 83.47	0.95	27.85	−0.2617	0.1472	7.52
LL,LH	121.71	− 143.26	0.83	36.44	− 0.2143	− 0.0143	12.88
LL,HH	70.89	− 84.21	− 1.28	25.91	0.1388	0.0580	6.51
HL,LL	27.33	− 31.25	− 0.05	8.53	−0.3142	0.0202	0.71
HL,HL	47.01	− 50.99	− 1.12	17.33	0.2695	0.0253	2.91
HL,LH	45.89	− 37.13	− 0.26	7.81	−0.2865	− 0.0113	0.59
HL,HH	50.43	− 51.32	0.26	16.78	0.2934	0.0063	2.72
LH,LL	83.02	− 88.83	− 2.26	24.75	0.0866	−0.5020	5.94
LH,HL	69.53	− 64.65	− 0.01	21.16	−0.1365	−0.3810	4.34
LH,LH	116.17	− 192.63	0.13	42.05	0.1057	− 0.1202	17.15
LH,HH	90.58	− 91.27	− 0.03	31.39	−0.1112	− 0.1040	9.55
HH,LL	20.98	− 24.76	− 0.10	6.59	−0.1264	− 0.2927	0.42
HH,HL	42.57	− 41.42	0.10	13.43	0.3176	− 0.3323	1.75
HH,LH	31.13	− 35.28	0.10	8.79	−0.1784	− 0.0913	0.75
HH,HH	65.93	− 78.46	0.79	19.75	−0.3531	− 0.0986	3.78

TCQ 借鉴了 TCM 中信号集合扩展、信号集合划分和网格状态转换的思想。以常用的 4 态 1/2 卷积编码器为例说明 TCQ 编码实现，其码书子集划分图和网格图见图 3。它将码书空间扩大一倍；根据编码率的要求，将扩大后的码书集合划分成 4 个子集，利用扩大的量化空间增加码字之间的距离，从而减小量化误差；利用网格约束编码两个相邻输入样点时的状态转换，优化编码序列整体的率－失真性能，取输入样点与 4 个子集的量化误差作为网格分支的标记，采用 Viterbi 算法寻找最佳网格路径，使得输入序列与该路径决定的重构序列之间的总体误差最小(编码失真最小)。

根据前述分析，基于小波包变换的 SAR 原始数据压缩算法利用小波包变换获得最大可能的能量紧缩，使变换域内信号更适合压缩；根据各子带不同的统计特性分配量化比特数，优化选择量化器；利用网格编码量化提高压缩性能，获得更高的压缩信噪比。

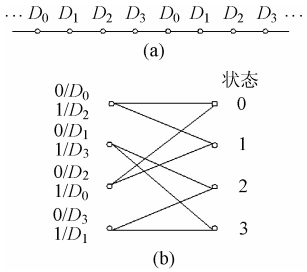

图 3 (a) TCQ 编码器的 4 子集码书划分图 (b) 4 子集 4 态网格图

Fig.3 (a) TCQ sets (b) 4-state TCQ trellis with subset labeling

3 试验结果

为验证算法的压缩性能，取实际 SAR 原始数据进行压缩实验。考虑到计算复杂性和变换性能，以 D4 小波为实验小波。图 4 为其尺度函数和小波函数的波形，其中尺度函数 ϕ 对应为低通滤波器，小波函数 ψ 对应为带通滤波器。

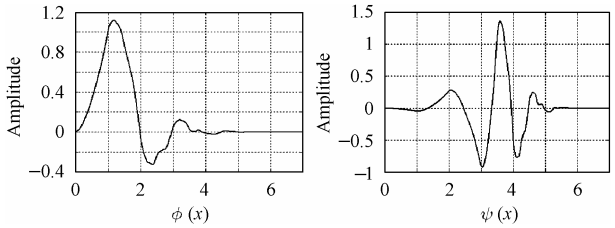

图 4 D4 小波的尺度函数和小波函数

Fig.4 Wavelet function and scale function of wavelet D4

取数据编码率 R 分别为 2.00，1.75，1.50，1.25 和 1.00 比特/样点，分析基于小波包的压缩算法(WPT−TCQ)压缩 SAR 原始数据的效果，并与矢量量化(VQ)、分块自适应矢量量化(BAVQ)的压缩结果相比较。本文选择信噪比(SNR)进行原始数据域的压缩算法评价。在均方误差意义下，SNR 定义如下。

$$\mathrm{SNR}=10\times\log10\left[\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}g(i,j)^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}[g(i,j)-f(i,j)]^2}\right] \tag{4}$$

其中 $g(i,j)$ 和 $f(i,j)$ 分别为原始数据和重建数据在 (i,j) 点的值，SNR 的单位为分贝。

表 2 为 SAR 原始数据压缩试验性能对比表，其中 R 为编码率，CR 为压缩比。与 VQ 比较，基于小波包的压缩算法获得的信噪比改善最大可达到约 2.3dB。当编码率 R 为 1bit/sample 时，信噪比改善最小也达到近 1.5dB；与 BAVQ 比较，基于小波包的压缩算法获得的信噪比改善也可以达到至少 1.0dB。压缩试验说明，在压缩比很大时，基于小波包的压缩算法仍然可获得很高的量化信噪比。因此基于小波包的压缩算法对 SAR 原始数据压缩有效。需要指出是，由于小波包的分解中涉及到选择最佳基，因此记录和下传重建结构可能需要一些额外码率，但所需数据量相对较小，这里计算压缩比时没有考虑这一花费。

表 2　实际 SAR 原始数据压缩试验性能比较

Tab.2 Performance comparison of data compression algorithms for SAR raw data

R (bit/sample)		2.00	1.75	1.50	1.25	1.00
CR		4.00	4.57	5.33	6.40	8.00
SNR(dB)	VQ	9.140	7.852	6.535	5.263	4.603
	BAVQ	9.391	8.214	7.007	5.756	5.059
	WPT-TCQ	11.452	10.470	9.125	7.485	6.066

图 5 给出了原始数据直接成像以及小波包变换的压缩算法压缩后的解压缩数据成像结果。解压缩数据成像结果表明，数据压缩比逐渐增大时，图像主观质量也大多较好，绝大部分的图像纹理和细节被很好保存，由此可见本文算法的高效性。

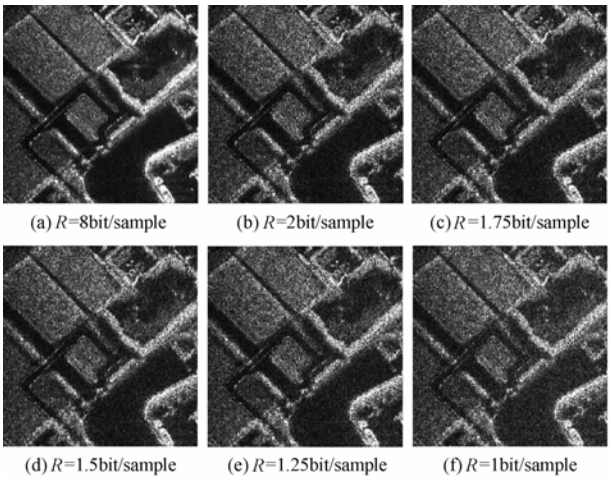

图 5　原始数据及解压缩数据成像

Fig.5　SAR image of raw data and decompressed data

4　结束语

本文在前人将小波包变换用于 SAR 原始数据压缩经验的基础上，结合小波包和网格编码量化的优势，提出了基于小波包变换和网格编码量化的 SAR 数据压缩算法。小波包变换克服了小波变换“高频部分时间分辨率低”的问题，有利于分析富含中高频信息的信号。网格编码量化利用独特的编码结构和数据间的相关性，可获得优于标量量化的性能。基于小波包的压缩算法在 SAR 原始数据压缩中取得了较好的效果，在数据域，SNR 较高，在图像域，成像图像主观质量良好。因而基于小波包的压缩算法对于 SAR 原始数据压缩算法的选择具有参考意义。此外本文仅仅探讨了在最佳基的选择时采用 Shannon 熵作为代价函数。因此为了提高压缩性能，如何选择最佳基的代价函数将是进一步研究的重点。

参 考 文 献

[1] Benz U, Strodl K, Moreira A. A comparison of several algorithms for SAR raw data compression. *IEEE Trans. on Geoscience and Remote Sensing*, 1995, 33(5): 1266−1276.

[2] Pascazio V, Schitinzi G. Wavelet transform coding for SAR raw data compression. IGARSS, Hamburg, Germany, 1999, (4): 2251−2253.

[3] Magli E, Olmo G, Penna B. Wavelet-based compression of SAR raw data. IGARSS, Toronto, Canada, 2002, (2): 24−28.

[4] Pascazio V, Schitinzi G. SAR phase history data compression by using wavelet packets. IGARSS, Honolulu, USA, 2000, (6): 2639−2641.

[5] Mallat S 著(法), 杨力华等译. 信号处理的小波导引(2nd). 北京: 机械工业出版社, 2002, 第八章.

[6] 杨福生. 小波变换的工程分析与应用. 北京: 科学出版社, 1999, 第七章.

[7] Marcellin M W, Fischer T R. Trellis coded quantization of memoryless and Gauss-Markov sources. *IEEE Trans. on Commun*, 1990, 38(1): 82−93.

胡晓新：　女，1977 年生，博士生，研究方向为合成孔径雷达数据压缩算法.

王岩飞：　男，1963 年生，博士生导师，研究方向为微波成像雷达系统及其理论、数字信号处理等.