

基于预测的双线偏振天气雷达数据无损压缩算法^{*}

曾强宇^{1,2} V Chandrasekar³ 何建新^{1,2} 史朝^{1,2} 王皓^{1,2} 李学华^{1,2}
ZENG Qiangyu^{1,2} V Chandrasekar³ HE Jianxin^{1,2} SHI Zhao^{1,2} WANG Hao^{1,2} LI Xuehua^{1,2}

1. 成都信息工程大学电子工程学院, 成都, 610225

2. 中国气象局大气探测重点开放实验室, 成都, 610225

3. 科罗拉多州立大学电气工程学院, 柯林斯堡, 80526

1. College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China

2. CMA Key Laboratory of Atmospheric Sounding, Chengdu 610225, China

3. Department of Electrical Engineering, Colorado State University, Fort Collins 80526, USA

2018-04-13 收稿, 2018-07-20 改回.

曾强宇, V Chandrasekar, 何建新, 史朝, 王皓, 李学华. 2019. 基于预测的双线偏振天气雷达数据无损压缩算法. 气象学报, 77(2):346-357

Zeng Qiangyu, V Chandrasekar, He Jianxin, Shi Zhao, Wang Hao, Li Xuehua. 2019. A lossless compression algorithm for dual-polarization weather radar data based on prediction. *Acta Meteorologica Sinica*, 77(2):346-357

Abstract With the development of dual-polarization weather radar technology, radar provides more and more products, and data accuracy continues to increase. The updating of detection performance also leads to a sharp increase in the amount of weather radar data. Data storage and transmission become an important issue to be solved in radar network applications. Data compression algorithms are used to reduce the amount of data transmitted and stored, but the characteristics of dual-polarization weather radar data are not fully considered in generic off-the-shelf data compression algorithms. This paper proposes a dual-polarization weather radar data compression algorithm (DPRC), which uses radial prediction to reduce the correlation of weather radar data and achieve efficient lossless compression of weather radar base data. The performance of the DPRC algorithm is evaluated using CSU-CHILL radar data. The experimental results show that the DPRC has a higher compression ratio than the generic off-the-shelf compression algorithm and is suitable for lossless compression of high resolution dual-polarization radar data.

Key words Weather radar, Data compression, Prediction, Data correlation, Dual-polarization

摘 要 随着双线偏振天气雷达技术的发展, 雷达提供的探测参量越来越多, 数据精度不断提高。探测性能提升使得天气雷达数据量急剧增长, 数据存储和传输是雷达网络化应用中需要解决的重要问题。数据压缩算法用于减小传输和存储的数据量, 但通用的数据压缩算法并未充分考虑双线偏振天气雷达数据的特点。文中提出适用于双线偏振天气雷达数据压缩算法 (DPRC), 使用径向预测减小天气雷达数据相关性, 实现了天气雷达基数据的高效无损压缩。使用 CSU-CHILL 雷达数据和双线偏振改造后的 CINRAD SA 雷达对 DPRC 的算法性能进行评估, 试验结果表明, DPRC 较通用的压缩算法压缩率更高, 适用于高分辨率双线偏振雷达数据无损压缩。

关键词 天气雷达, 数据压缩, 预测算法, 数据相关性, 双线偏振

中图法分类号 P413

^{*} 资助课题: 自然科学基金项目 (41405030、41375043 和 41505031)、国家留学基金委 (201608515109)、四川省教育厅项目 (13Z213)、成都信息工程大学项目 (J201503)。

作者简介: 曾强宇, 主要从事天气雷达信号处理研究。E-mail: zqy@cuit.edu.cn

1 引言

天气雷达探测数据是研究天气现象成因,预测和分析天气过程演变的重要信息源。随着高分辨率天气雷达和天气雷达网络应用于大气探测,气象灾害的监测预警能力和探测范围得到了极大的提升(杜牧云等,2013; Li, et al, 2017)。双线偏振天气雷达能够实现降水粒子相态识别,较常规多普勒天气雷达能够提供更多探测参量。双线偏振天气雷达能够获取天气目标的水平反射率因子(Z_h)、平均径向速度(V_r)、速度谱宽(W)等基数据产品,还能获取差分反射率因子(Z_{dr})、差分相移(φ_{dp})、差分相位率(K_{dp})、相关系数(ρ_{hv})、线性退偏振比(L_{dr})、归一化相干功率(NCP)、反射功率(DM)和交叉极化功率(DX)等偏振参数产品。偏振参数产品有利于天气目标微物理场分析,可以更准确地探测降水粒子的滴谱特征,提高降水粒子类型识别准确度和天气雷达定量测量降水能力(李晓敏等,2017; 梁海河等,2006; 郭凤霞等,2014)。双线偏振天气雷达较常规多普勒雷达探测数据精度也有极大的提高。CSU-CHILL 雷达是目前最先进的双线偏振雷达之一(Junyent, et al, 2015)。

高分辨率双线偏振雷达产生了大量需要实时处理、存储和传输的数据。统计表明,CHILL 雷达的体扫数据量是 CINRAD 雷达体扫的 15—30 倍,巨大的数据量给网络传输和存储带来了极大的挑战。气象业务需要在线实时共享天气雷达探测数据,各组网站点的天气雷达回波数据实时传输是天气雷达组网实现需解决的关键问题(Lakshmanan, et al, 2006)。根据双线偏振天气雷达数据特点,设计实时高效的数据压缩算法能有效提高雷达的业务应用能力。

天气雷达基数据是应用最广泛的数据。天气雷达基数据压缩算法分为有损压缩和无损压缩两种。有损压缩算法较无损算法压缩率更高,但会损失回波数据中的部分信息。使用有损或是无损压缩算法需要考虑不同的应用环境。雷达数据的定性分析对数据的精度要求相对较低,而定量计算对数据的精度要求相对较高,因此,有损压缩算法适用于定性分析的天气雷达数据,而不适用于数据保真度较高的应用需求,数据失真会影响天气雷达数据的同化和再分析。在天气过程分析中需要用到多种天气雷达

产品,数据测量精度决定了估计结果的准确性(李喆等,2016)。气象研究中需要用到 Level-Ⅲ 气象雷达产品,天气雷达基数据是 Level-Ⅱ 数据,Level-Ⅲ 气象雷达产品由 Level-Ⅱ 数据生成。上述应用环境需要对天气雷达数据无损传输。

天气雷达探测数据有损压缩研究中,使用图像压缩算法对平面位置显示(PPI)图像数据进行压缩是一种有效的数据压缩方法。黄云仙等(2012)提出了基于阶梯量化和 MED(Multiple-Differential Evolution)预测的有损天气雷达数据压缩算法,并比较了几种压缩算法的性能。Ma 等(2016)根据用户的需求剪辑天气雷达数据,实现天气雷达数据的有损压缩。Makkapati 等(2007)使用量化编码技术实现天气雷达数据的有损压缩。Mishra 等(2014)、Kawami 等(2016, 2017)提出基于三维压缩感知的方法,对相控阵天气雷达(PAWR)数据进行压缩采样,压缩比优于二维方法。但压缩感知算法性能受数据稀疏性影响较大,高分辨率雷达数据的稀疏性较低,压缩感知重建信号可能会产生较大的失真。

天气雷达探测数据无损压缩研究中,McCarroll 等(2011)分析了 WSR-88D 雷达的 Level-Ⅱ 超分辨率数据,包括拖尾 0、原始数据和差分数据的分布。基于分析结果,提出了基于径向差分的无损压缩方法,压缩效果较通用压缩算法性能有明显提升。Lakshmanan(2007)、Kruger 等(1997)比较了几种用于天气雷达数据的典型无损压缩算法,结果表明天气雷达数据中存在较大的数据冗余,利用天气雷达数据的特性能进一步提高压缩率。Ai 等(2009)提出了对天气雷达 PPI 数据的压缩算法,利用线性预测实现对 PPI 图像的压缩。Zeng 等(2018)提出了基于时间和空间预测的 CINRAD 天气雷达数据的无损压缩算法。现有的天气雷达数据压缩算法主要针对传统多普勒雷达,对于双线偏振雷达的数据压缩算法的研究很少,为提高双线偏振雷达的应用范围,研究双线偏振雷达的数据压缩算法有重要的理论意义和实用价值。

2 双线偏振天气雷达基数据的特点

天气雷达根据体扫模式不同,通常 5—10 min 完成一次体扫,并将体扫数据保存成文件。体扫数据文件包含头文件和体扫数据。头文件记录雷达参数、工作参数、扫描模式和位置信息等。Level-Ⅱ 体

扫数据即基数据,保存雷达的观测结果。基数据以径向为基本单位,径向数据是雷达在方位角的观测数据。基数据包含基本参量(Z_h 、 V_r 、 W),双线偏振雷达还包含偏振参量(Z_{dr} 、 φ_{dp} 、 K_{dp} 、 ρ_{hv} 、 L_{dr} 、NCP、DM、DX等)。

2.1 天气雷达基数据中的无效数据

体扫数据中有大量无效数据。在天气雷达数据处理中,当数据值超过了编码阈值或对应区域位于大气中不存在天气现象的区域时,该区域的距离库表示为无效数据。由于大量无效数据的存在,天气

雷达数据通常具有较高的稀疏性。无效数据可以分为两类:一类是位于径向尾部,从径向尾部逆向到非0值的部分,称为拖尾无效数据;另一类是随机分布在数据中的非拖尾无效数据。图1比较了CHILL和CINRAD-SA的体扫数据样本的无效数据分布,统计了 0.5° 仰角扫描的无效数据数目。试验CHILL数据使用Greely站点2007年6月12日16时03分33秒(北京时,下同)降雨过程探测的数据,该雷达一个仰角包含464个径向,每个径向有1000个距离库,距离分辨率150 m,方位分辨率 0.802° 。

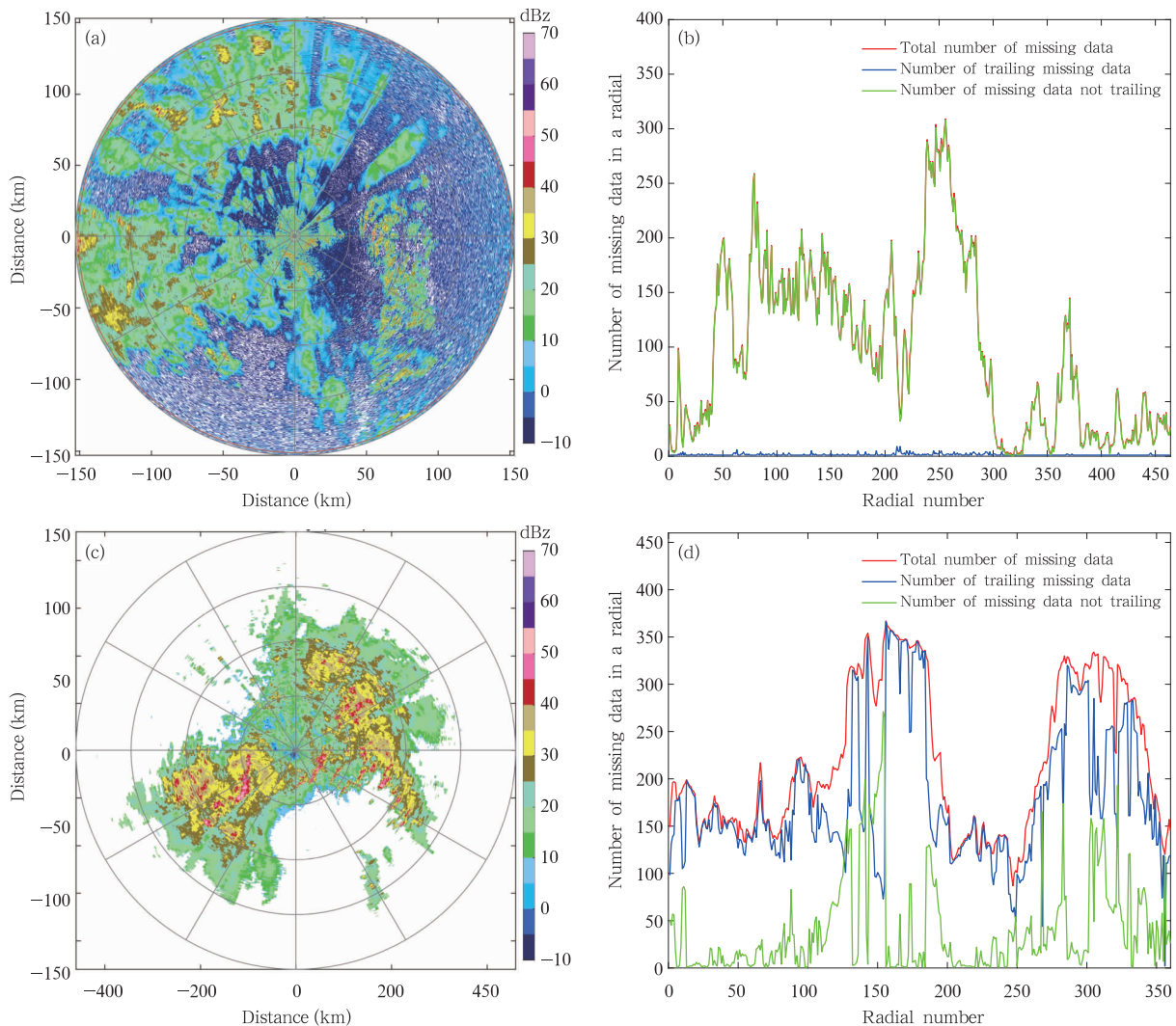


图1 雷达数据一个仰角的无效数据分布统计

(a. CHILL 反射率 PPI, b. CHILL 数据无效数据分布, c. CINRAD SA 反射率 PPI, d. CINRAD SA 数据无效数据分布)

Fig. 1 Missing data statistics of an example elevation cut

(a. CHILL radar reflectivity PPI, b. Missing data statistics of CHILL data, c. CINRAD SA radar reflectivity PPI, d. Missing data statistics of CINRAD SA data)

试验 CINRAD SA 数据使用广州 9200 站点 2016 年 5 月 20 日降雨过程探测的数据,该雷达一个仰角包含 360 个径向,每个径向有 460 个距离库,距离分辨率 1 km。天气雷达数据中的无效数据占很大比例,这种数据结构有利于数据冗余的压缩。CINRAD SA 数据中存在大量的拖尾无效数据,现有天气雷达压缩算法中通常将无效数据置 0,直接移除拖尾 0,这种方法不会造成数据损失,可以提高数据压缩率约 20%(McCarroll, et al, 2011)。但在高分辨率的双线偏振 CSU-CHILL 雷达数据中,由于探测能力的提高,拖尾无效数据数目很少,使用这种方法对压缩性能提升有限。而且无效数据在气象研究中也是有意义的,置 0 无效数据会造成数据的失真。

2.2 天气雷达基数据的相关性

天气雷达回波数据在时间和空间上存在很强的相关,这为天气雷达数据压缩提供了理论基础。天气雷达径向内相邻的两个距离库以及空间位置上相邻的两个径向的数据存在很强的相关。连续的相邻距离库大概率出现连续相同或是近似的值。游程编码能有效压缩这种空间冗余。WSR-88D 雷达的基数据传输时,使用了 4-bit RLE (Run-Length Encoding) 编码压缩这种空间冗余(National Weather Service, 2008)。由于天气雷达相邻径向的角度差很小,在空间上存在较强的相关。图 2a—d 分别显示了图 1 中 CHILL 雷达数据当前径向的 Z_h 、 V_r 、 Z_{dr} 、 φ_{dp} 和它的前一个径向对应参数的相关系数,图 2e、f 分别显示了图 1 中 CINRAD SA 雷达数据当前径向的 Z_h 、 V_r 和它的前一个径向对应参数的相关系数。对比晴天、降水和降雪等天气的 CINRAD SA 和 CHILL 天气雷达数据的径向之间的相关,可以发现 CINRAD SA 雷达数据中相邻径向的相关比 CHILL 雷达数据明显更强,这是因为 CHILL 探测数据精度更高,数据中有更多离散分布的值,降低了数据的相关。CINRAD SA 雷达数据中大部分的相邻径向的相关系数接近 1,表明了相邻径向距离库的幅度值分布高度近似。基于这种相关,McCarroll 等(2011)提出了相邻径向差分方法,使用差分数据替代原始数据进行熵编码

$$D_i = R_i - R_{i-1} \quad (1)$$

式中, D_i 为差分数据, R_i 为当前径向距离库值, R_{i-1}

为 R_i 的前一个径向距离库值。该算法适用于 WSR-88D 雷达的反射率数据, D_i 和 R_i 的单位为 dBz。相关强的天气雷达数据差分后,能有效降低数据的动态范围,差分数据集中在 0 值附近的区域分布。但这种方法的缺点是在变化较剧烈的区域(如天气现象区域的边界处),差分会导致数据的动态范围增大,反而降低了数据的压缩效率。CHILL 数据的径向相关在 Z_h 、 L_{dr} 和 ρ_{hv} 等参数中较高,而 V_r 、 Z_{dr} 和 φ_{dp} 等参数数据的径向相关较弱,这导致了相邻径向差分方法对双线偏振雷达数据处理时,差分数据熵值大于原始数据,不适合双线偏振雷达数据压缩。由于天气雷达探测时相邻两束电磁波存在很强的相关,对相邻径向数据分块处理,使用径向预测可以有效减小数据间的相关。使用预测编码时,传输当前径向和参考径向的残差数据值和运动矢量,可以无损重建原始数据。预测算法通过降低数据相关,使得残差数据的零阶熵小于原始数据。根据 Shannon 编码理论,压缩数据的平均码长由无记忆信源的零阶熵决定,预测算法可以有效压缩天气雷达数据。

双线偏振雷达基数据也有很强的时间相关,时间相关与体扫模式和天气现象变化的剧烈程度有关。基数据的时间相关主要分为两类:第一类是在相邻的仰角同方位角的径向相关;另一类是在时间相邻的体扫数据同位置的径向相关。第一类时间相关由于仰角差导致了数据间存在几何畸变,相关受仰角差影响较大。大多数情况下天气现象是慢变的局部运动,相邻的体扫数据有很强的相关。图 3 显示了图 1 数据与其之前的两相邻体扫数据中相同位置径向间的相关。CHILL 雷达数据在时间上具有很强的相关,随着时间间隔增大,相关会逐渐降低。

3 双线偏振天气雷达基数据的压缩算法

3.1 双线偏振雷达基数据压缩算法流程

文中提出了双线偏振雷达基数据无损压缩算法(图 4)。

预处理根据头文件信息,读取各个仰角的扫描数据,将基数据按各种参数分类存储。预测算法用于去除径向数据的相关,减少数据冗余。预测得到的运动矢量和残差数据替代原始值进行存储和传

输。模式选择用于确定最优的分块模式和预测模式,使得输出的运动矢量和残差数据熵值最小。最

后熵编码进一步压缩数据间的结构冗余。对于其余的数据(如头文件等)使用算术编码压缩。

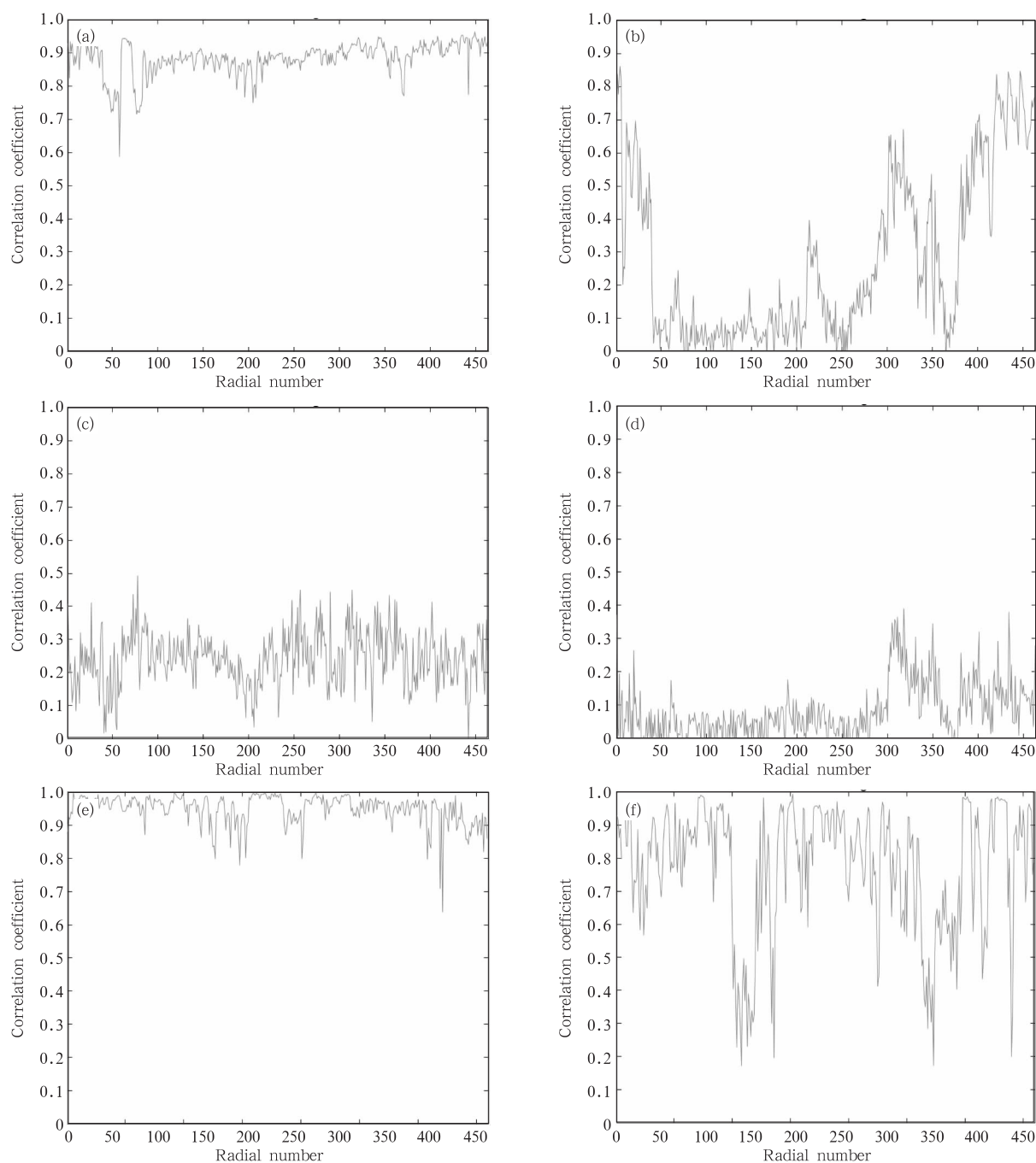


图2 CHILL 和 CINRAD SA 雷达相邻径向的相关系数

(a—d, CHILL Z_h , V_r , Z_{dr} , φ_{dp} ; e—f, CINRAD Z_h , V_r)

Fig.2 Radial-to-radial correlation coefficients between different types of CHILL radar and CINRAD radar products

(a—d, CHILL Z_h , V_r , Z_{dr} , φ_{dp} ; e, f, CINRAD Z_h , V_r)

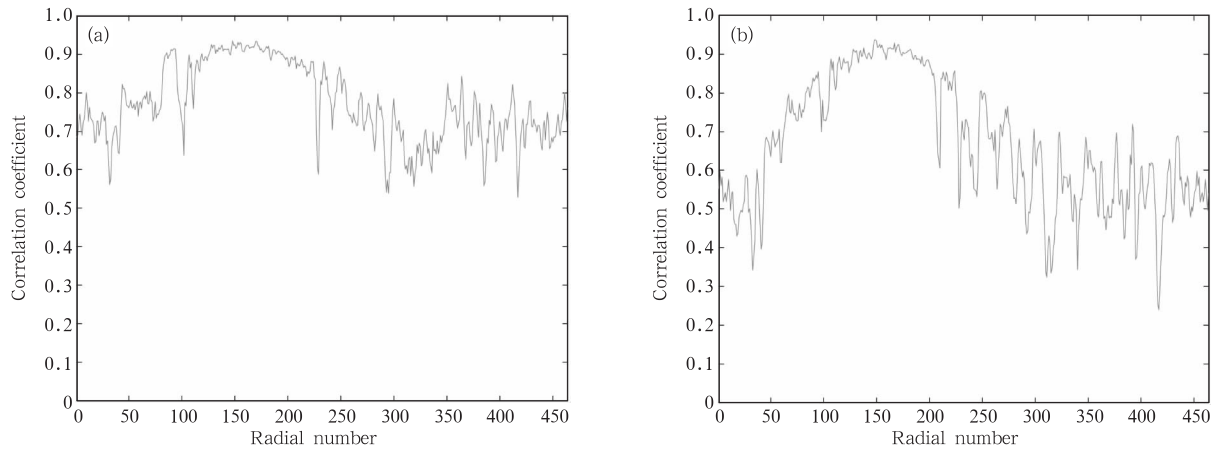


图3 CHILL 雷达相邻体扫数据相同位置径向的相关系数

(a. 16 时 03 分 33 秒与 15 时 59 分 16 秒, b. 16 时 03 分 33 秒与 15 时 55 分 19 秒)

Fig. 3 Volume-to-volume reflectivity radial correlation coefficients of CHILL radar products

(a. 16:03:33 BT and 15:59:16 BT; b. 16:03:33 BT and 15:55:19 BT)

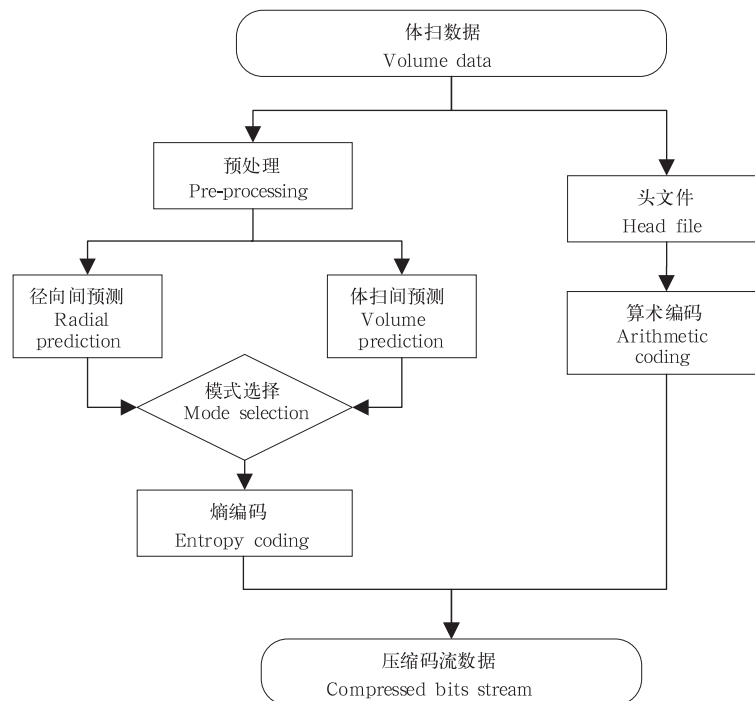


图4 双线偏振雷达基数据无损压缩算法流程

Fig. 4 Lossless compression algorithm for dual-polarization weather radar data flow

3.2 预测算法

预测算法将径向作为处理的基本单元。径向以块为处理单位进行预测。块的大小会影响预测精度。较小的块,能够有更高的预测精度,残差数据动

态范围更小,但缺点是需要用更多的运动矢量来表示最优块的位置。选择块大小需要平衡残差数据和运动矢量的编码消耗。本算法将相邻的 32 个距离库分成一个块,在边界部分没有足够距离库时,使用

补“0”的方法构成完整的块。每个块在预测时根据数据的分布特性分别使用图 5 所示的 3 种分块模式进行预测处理。分块模式 1 预测时,整个块只有 1 个运动矢量,适用于天气变化平缓的区域;分块模式 2 预测时,整个块分成两个各有 16 个距离库的子块分别预测得到各自的运动矢量;分块模式 3 预测时,

整个块分成 4 个各有 8 个距离库的子块分别预测得到各自的运动矢量。分块的目的是在数据分布较平缓的区域使用较大的块进行预测以减少运动矢量的开销,对于数据存在剧烈变化的区域(如存在天气现象的边界区域),使用较小的块进行精细预测以减小残差数据的动态范围。

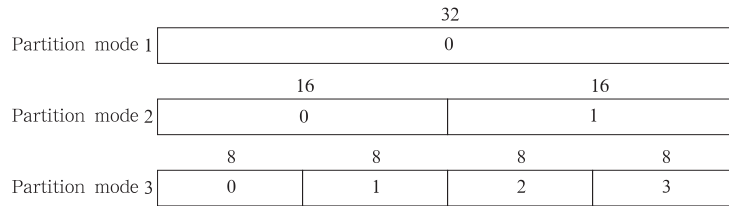


图 5 径向数据分块模式
Fig. 5 Radial data block mode

预测算法基于块匹配,通过比较每个块的绝对误差和(Sum of Absolute Difference, SAD),找到参考径向中与当前块最相似的块,并保存这个预测块的运动矢量(位置信息)和残差数据(当前块和预测块的差值)。

$$x_{i,j} = vx_{k,l} + r_{i,j} \quad (2)$$

式中, $x_{i,j}$ 为第 i 个径向中的第 j 个块, $x_{k,l}$ 为第 k 个径向(参考径向)中的运动匹配的最优块, x 的单位由对应的天气雷达产品类型决定, v 为运动矢量,指示最优块的位置。 v 包含两个参数,最优块所在径向的索引和最优块与当前块在径向方向上的位置偏

移量。 $r_{i,j}$ 是预测后当前块与最优块对应位置距离库的差值,即残差数据。最优块的选择准则是在运动搜索范围内搜索与当前块 SAD 值最小的块。SAD 计算方法为

$$SAD = \sum_{i=1}^{blocksize} |C(i) - P(i)| \quad (3)$$

式中, $C(i)$ 表示当前块中第 i 个距离库, $P(i)$ 表示最优预测块中第 i 个距离库。当前块 C 和最优预测块 P 位置如图 6 所示,其单位取决于对应的天气雷达产品类型,如天气雷达产品为反射率时,单位为 dBz,天气雷达产品为径向速度时,单位为 m/s。

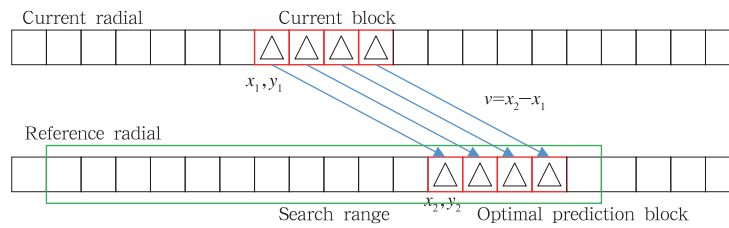


图 6 径向数据预测模式
Fig. 6 Radial data prediction mode

解码端通过运动补偿可以无损重建原始数据。预测编码在相邻径向间变化较小时,压缩效率较高,和径向差分相比,预测编码对存在天气现象变化较大的数据和离散分布的数据也有很高的压缩性能。

径向间的预测算法如图 6 所示。使用滑动窗,计算每个参考径向中搜索范围内的每个块和当前块的 SAD,寻找 SAD 最小值对应的最优块。搜索范围会影响预测精度和残差数据的动态范围。使用大的搜

索范围可以提高预测精度,但也会导致编码时的运算量提升,同时也会影响运动矢量的大小。例如,搜索范围设置为 $[-15, 15]$ 时,表示径向方向上的位置偏移量需要 5 bit,而搜索范围设置为 $[-31, 31]$ 时,需要 6 bit。参考径向的数目也会影响预测精度和运动矢量。根据参考径向的位置可以将预测算法分成径向间预测和体扫间预测。径向间预测模式中参考径向和当前径向在同一个仰角切面中;体扫间预测模式中参考径向在当前径向之前的体扫数据相同仰角的仰角切面中。两种模式的参考径向位置如图 7 所示。为保证编码的实时性,参考径向只能使用当前径向之前已经编码的数据。

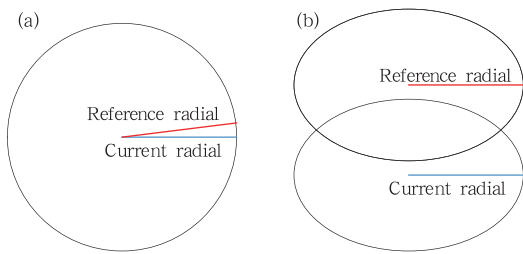


图 7 预测模式

(a. 径向间预测模式, b. 体扫间预测模式)

Fig. 7 Prediction type

(a. Radial-to-radial prediction,
b. volume-to-volume prediction)

模式选择用于确定最优的预测模式和分块模式,分别计算 3 种分块模式下一个 1×32 块的数据最优预测的残差数据的 SAD 和运动矢量绝对值的和,根据残差数据的 SAD 和运动矢量绝对值的和的最小值确定分块模式和预测模式。每一个子块都对应各自的运动矢量,分块模式 1,1 个 1×32 块有 1 个运动矢量;分块模式 2,有 2 个运动矢量;分块模式 3,有 4 个运动矢量。模式选择用于平衡残差数据和运动矢量的编码消耗。

3.3 熵编码

根据天气雷达数据预测后得到的残差数据和运动矢量的数据分布特点,熵编码使用 LZW(Lempel Ziv Welch)。LZW 编码根据符号出现频率冗余度进行压缩,对数据流中连续重复出现的符号有很高的压缩比。通常残差数据值的分布范围远小于原始值,集中在 0 附近,即 0 附近的值出现频率很高,使用 LZW 编码能有效压缩。LZW 编码对不同的块编码后长度不同。在原始数据中每个距离库的值大小是 1—2 byte,LZW 编码后,每个距离库的值根据出现概率的大小,编码成不同码长的符号,并在传输和储存时以码流的形式保存。LZW 码对出现概率高的符号用较短的码进行编码,低概率的符号用较长的码字进行编码,以实现减少数据的平均码长。LZW 编码后的数据仍以字节的形式存放,但每个字节不再含有独立意义,解码器需要接收完一个径向的数据后才能解压。

4 双线偏振天气雷达基数据的压缩算法 (DPRC)

为验证 DPRC 的压缩性能,首先使用 S 波段的 CHILL 雷达 Greely 站点 2007 年 6 月 12 日 16 时 03 分 33 秒降雨过程探测的数据对预测算法的性能进行验证。图 8—10 分别显示了 Z_h 、 Z_{dr} 和 φ_{dp} 预测前、后的数据分布。数据精度如表 1 所示,搜索范围设置为 $[-31, 31]$ 。参考径向设置为 5 个,包括径向间预测模式参考径向 2 个(当前径向的前两个径向)和体扫间预测模式参考径向 3 个(前一个体扫中和当前径向相同位置的径向及其前后各一个径向)。图 8 是 CHILL 雷达 Z_h 数据预测前、后的数据分布差异。 Z_h 数据预测后,数据的分布范围从 $[-32, 80]$ 降至 $[-5, 5]$,并且残差值的出现频率也明显提高,这种数据结构有利于后续使用变长编码进一步

表 1 不同天气原始数据和残差数据的平均熵值比较

Table 1 Average entropy of raw data and residual in different weather conditions

	Z_h	V_r	W	Z_{dr}	L_{dr}	φ_{dp}	ρ_{hv}
降雨数据	6.37	7.37	5.39	4.28	3.79	6.60	2.80
降雨残差	3.83	5.45	4.39	3.35	2.81	5.24	1.44
降雪数据	5.92	7.55	5.82	4.67	3.88	7.60	2.17
降雪残差	3.34	5.51	4.77	3.88	2.92	6.14	1.11
晴天数据	5.29	7.80	5.63	4.09	3.67	7.43	1.23
晴天残差	3.55	6.08	4.86	3.60	2.77	6.28	0.73

提高数据压缩率。图 9、10 的测试结果表明,对径向间相关较弱的 Z_{dr} 和 φ_{dp} 数据,预测算法也能有效改变数据的分布,便于后续压缩。分块预测方法比径向差分更适用于双线偏振雷达数据。DPRC 的预测算法适用于 CHILL 雷达的各种基数据产品,预测后残差值的动态范围明显小于原始数据,残差值的分布集中在“0”附近的区域,去除了天气雷达回波数据间的相关。这种数据结构在“0”附近高概率分布,熵编码对高概率的值以短码长进行编码,能有效压缩 CHILL 雷达数据。

为验证 DPRC 对不同天气情况的鲁棒性,选取降雨、降雪和晴天每种数据20个PPI,统计原始数

据和残差数据的平均熵值。从表 2 的结果可以看出 DPRC 能有效降低双线偏振雷达各种基数据产品的

表 2 CSU-CHILL 数据不同算法压缩率比较
Table 2 Compression ratios of different methods using CSU-CHILL data

	Bzip2	Gzip	Winzip	LP	DPRC
压缩率	3.33	3.21	2.85	2.76	4.05

表 3 双线偏振 CINRAD SA 数据不同算法压缩率比较
Table 3 Compression ratios of different methods using CINRAD SA data

	Bzip2	Gzip	Winzip	LP	DPRC
压缩率	4.62	5.01	4.33	4.27	5.62

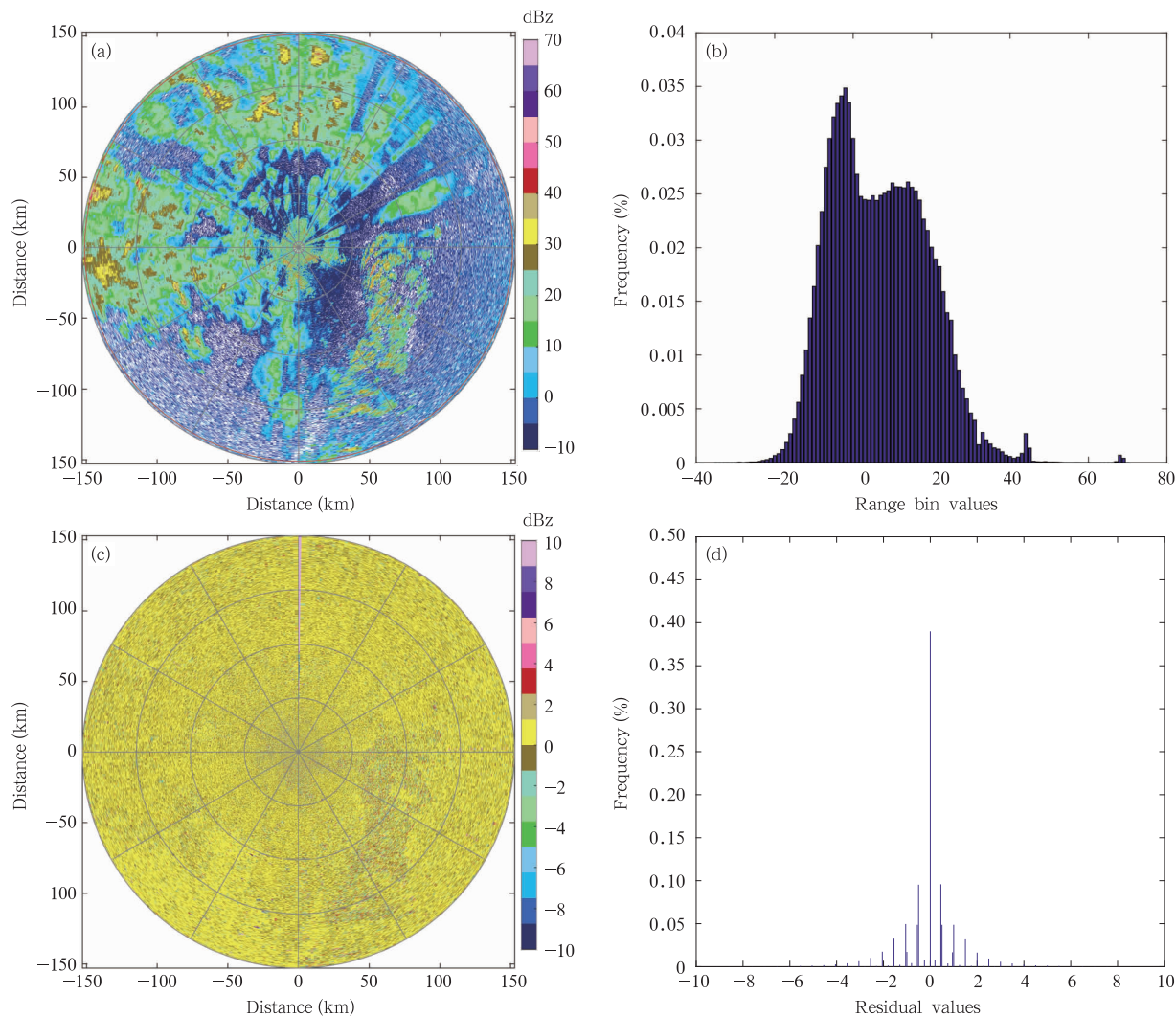


图 8 CHILL 雷达 Z_h 原始数据(a,b)和残差数据(c,d)的 PPI(a,c)和频率直方图(b,d)
Fig.8 Distributions of Z_h raw data (a, b) and residual raw (c, d) reflectivity PPI (a, c) and frequency histogram (b, d)

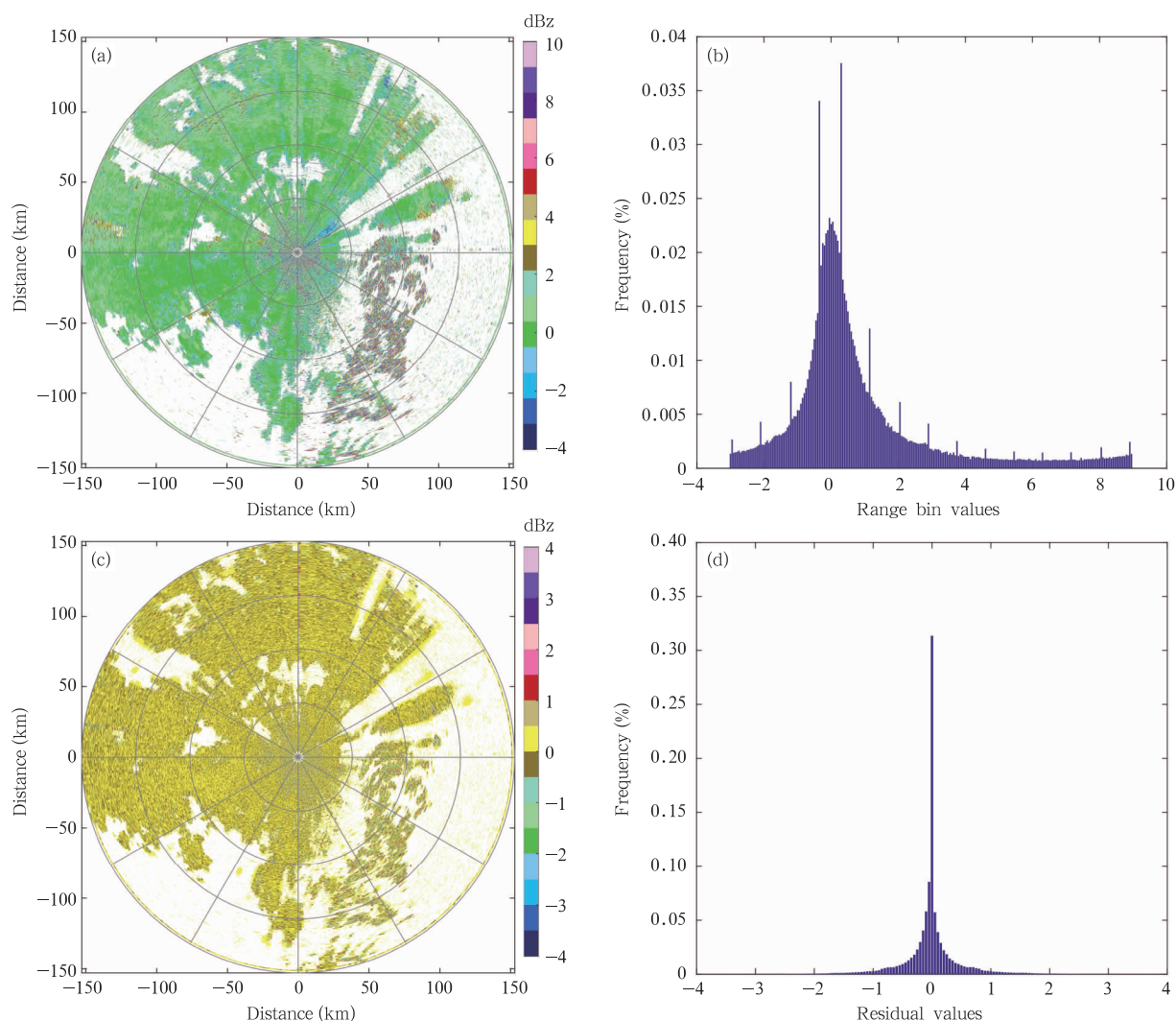


图9 CHILL 雷达 Z_{dr} 原始数据(a,b)和残差数据(c,d)的 PPI(a,c)和频率直方图(b,d)
 Fig. 9 Distributions of Z_{dr} raw data (a, b) and residual data (c, d), reflectivity PPI (a, c) and frequency histogram (b, d)

相关性,适用于各种天气现象,预测后的残差数据较原始数据熵值明显降低,为提高压缩率提供了可能。需要说明的是,表2的平均熵值中运动矢量的熵值并未统计。

随机选取30个体扫数据用于测试DPRC的压缩率。使用CHILL雷达2011—2016年的数据集,每个体扫数据根据扫描模式不同包含9—18个仰角,仰角分布 0.5° — 19.5° 压缩率的计算方法为原始数据大小与压缩后的数据大小的比值。表3比较通用压缩算法Bzip2、Gzip、WinZip与天气雷达数据的专用压缩算法最优线性预测LP(Linear Prediction)

(Lakshmanan, 2001)、DPRC的压缩率。Bzip2是基于块的压缩算法。Gzip使用算术编码,包括LZW和Huffman编码。试验结果表明,DPRC压缩效率明显高于现有的无损压缩算法,较次优算法压缩性能大约提升19%。

针对中国双线偏振天气雷达的业务应用,使用广州气象局双线偏振雷达数据对DPRC算法性能进行测试。随机选取了30个体扫数据用于测试DPRC的压缩率。数据使用双线偏振CINRAD SA雷达2016年5月18—28日的华南强降水观测试验数据集,每个体扫数据扫描模式包含9个仰角,仰角

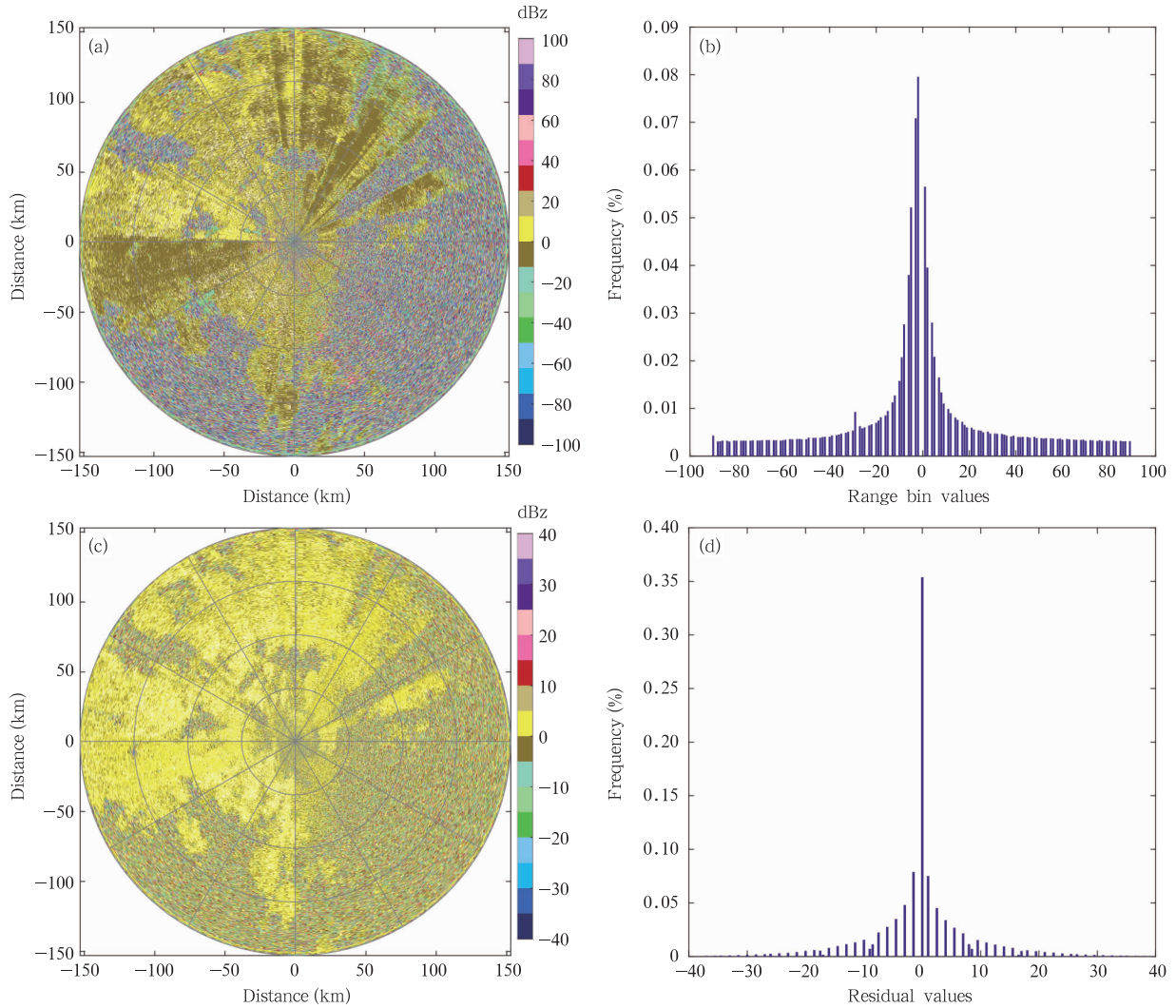


图 10 CHILL 雷达 φ_{dp} 原始数据(a,b)和残差数据(c,d)的 PPI(a,c)和频率直方图(b,d)
Fig. 10 Distributions of φ_{dp} raw data (a, b) and residual data (c, d), reflectivity PPI (a, c) and frequency histogram (b, d)

分布 $0.5^{\circ}—19.5^{\circ}$, 数据中包含 3 种基本参量 (Z_h 、 V_r 、 W) 和 3 种偏振参量 (Z_{dr} 、 φ_{dp} 、 ρ_{hv})。试验结果如表 2 和 3 所示, DPRC 压缩效率明显高于现有的无损压缩算法, 较次优算法压缩性能大约提升 11%。相较于 CSU-CHILL 雷达数据, DPRC 对于双线偏振 CINRAD SA 雷达数据压缩率提升略低, 这是因为双线偏振 CINRAD SA 雷达数据精度略低, 数据稀疏性较高。预测算法的主要作用是提高残差数据的稀疏性, 对于稀疏度较高的原始数据, 稀疏性在预测后提升相对较低; 而对于稀疏度较低的高精度数据, 预测算法能有效提高残差数据的稀疏性, 提高压缩率。DPRC 对高分辨率雷达数据较传统压缩算法

有明显优势。

5 结 语

双线偏振雷达较传统多普勒雷达在数据精度和分辨率上有了极大的提升, 提高了对天气现象的探测能力。但双线偏振雷达巨大的数据量不利于数据的存储和传输。文中提出了 DPRC 压缩算法, 使用预测技术降低了雷达数据的相关性, 实现了双线偏振雷达数据的无损压缩, 克服了通用无损压缩算法压缩率低的缺点, 为高分辨率双线偏振雷达组网提供了技术支持。DPRC 适用于低码率信道的天气雷达数据传输和天气雷达探测数据的网络传输, 提高

数据压缩率。现有双线偏振雷达数据压缩率仍有进一步提升的空间,研究针对特定应用环境的有损压缩算法和适应雷达数据特点的算术编码算法,将是值得关注的研究方向。

致 谢:本文是作者在科罗拉多州立大学访学期间完成的,使用了CSU-CHILL 双线偏振雷达探测资料,并得到了 V Chandrasekar 教授团队的帮助;使用了广东省气象局的 CINRAD SA 雷达探测资料,在此表示衷心的感谢。

参考文献

- 杜牧云, 刘黎平, 胡志群等. 2013. 双线偏振多普勒雷达资料质量分析. 气象学报, 71(1): 146-158. Du M Y, Liu L P, Hu Z Q, et al. 2013. An analysis of dual-linear polarimetric Doppler radar data quality. Acta Meteor Sinica, 71(1): 146-158 (in Chinese)
- 郭凤霞, 马学谦, 王涛等. 2014. 基于 X 波段双线偏振天气雷达的雷暴云粒子识别. 气象学报, 72(6): 1231-1244. Guo F X, Ma X Q, Wang T, et al. 2014. An approach to the hydrometeors classification for thunderclouds based on the X-band dual-polarization Doppler weather radar. Acta Meteor Sinica, 72(6): 1231-1244 (in Chinese)
- 黄云仙, 马烁, 艾未华. 2012. 一种新的多普勒天气雷达回波数据无损压缩. 解放军理工大学学报(自然科学版), 13(2): 232-236. Huang Y X, Ma S, Ai W H. 2012. Novel lossless compression of Doppler weather radar data. J PLA Univ Sci Technol (Nat Sci Edit), 13(2): 232-236 (in Chinese)
- 李喆, 李柏, 赵坤等. 2016. 国产双偏振天气雷达差分反射率测量性能分析. 气象科技, 44(6): 855-859, 895. Li Z, Li B, Zhao K, et al. 2016. Measurement performance analysis of differential reflectivity for dual-polarization weather radar made in China. Meteor Sci Technol, 44(6): 855-859, 895 (in Chinese)
- 李晓敏, 周筠珩, 肖辉等. 2017. 基于 X 波段双偏振雷达对雷暴单体中水成物粒子演变特征的研究. 大气科学, 41(6): 1246-1263. Li X M, Zhou Y J, Xiao H, et al. 2017. Research on the evolution characteristics of hydrometeors in a thunderstorm cell with X-band dual-polarimetric radar. Chinese J Atmos Sci, 41(6): 1246-1263 (in Chinese)
- 梁海河, 徐宝祥, 刘黎平等. 2006. CINRAD/SA 雷达天线的辐射特性分析及其支持双线偏振升级的可行性研究. 气象学报, 64(6): 699-708. Liang H H, Xu B X, Liu L P, et al. 2006. CINRAD/SA radiation pattern and feasibility for dual polarization upgrades. Acta Meteor Sinica, 64(6): 699-708 (in Chinese)
- Ai W H, Yan W, Li X. 2009. Efficient lossless compression of weather radar data. Int J Environ Ecol Eng, 3(8): 212-215
- Junyent F, Chandrasekar V, Bringi V N, et al. 2015. Transformation of the CSU-CHILL radar facility to a dual-frequency, dual-polarization Doppler system. Bull Amer Meteor Soc, 96(6): 975-996
- Kawami R, Hirabayashi A, Tanaka N, et al. 2016. 2-Dimensional high-quality reconstruction of compressive measurements of phased array weather radar // Proceedings of the 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Jeju, South Korea: IEEE, 1-7
- Kawami R, Hirabayashi A, Ijiri T, et al. 2017. 3-Dimensional compressive sensing and high-quality recovery for phased array weather radar // Proceedings of the 2017 International Conference on Sampling Theory and Applications (SampTA). Tallin, Estonia: IEEE, 658-661
- Kruger A, Krajewski W F. 1997. Efficient storage of weather radar data. Softw Pract Exp, 27(6): 623-635
- Lakshmanan V. 2001. P2.1 Lossless coding and compression of radar reflectivity data // Proceedings of the 30th International Conference on Radar Meteorology. Munich, Germany: AMS, 50-52
- Lakshmanan V, Smith T, Hondl K, et al. 2006. A real-time, three-dimensional, rapidly updating, heterogeneous radar merger technique for reflectivity, velocity, and derived products. Wea Forecasting, 21(5): 802-823
- Lakshmanan V. 2007. Overview of radar data compression // Proceedings of Satellite Data Compression, Communications, and Archiving III. San Diego, California, United States: SPE, 6683: 668308
- Li X H, He J X, Wang C Z, et al. 2017. Evaluation of surface clutter for future geostationary spaceborne weather radar. Atmosphere, 8(1): 14
- Ma K, Li C Y, Li H Z. 2016. A new hybrid data compression algorithm for weather radar // Proceedings of the 11th International Symposium on Antennas, Propagation and EM Theory Propagation and EM Theory. Guilin, China: IEEE, 886-889
- Makkapati V V, Mahapatra P R. 2007. Extreme compression of weather radar data. IEEE Trans Geosci Remote Sens, 45(11): 3773-3783
- McCarroll S, Yearly M, Hougen D, et al. 2011. Approaches for compression of super-resolution WSR-88D data. IEEE Geosci Remote Sens Lett, 8(2): 191-195
- Mishra K V, Kruger A, Krajewski W F. 2014. Compressed sensing applied to weather radar // Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium. Quebec City, QC, Canada: IEEE, 1832-1835
- National Weather Service. 2008. Radar Operations Center, Interface Control Document for the RPG to Class 1 User, 2620001
- Zeng Q Y, He J X, Shi Z, et al. 2018. Weather radar data compression based on spatial and temporal prediction. Atmosphere, 9(3): 96