

文章编号: 1000-8349(2008)01-0062-11



我国探月工程 VLBI 相关处理机简介

朱人杰, 张秀忠, 项 英, 徐志骏

(中国科学院 上海天文台, 上海 200030)

摘要: 甚长基线干涉测量 (VLBI) 是重要的射电天文技术, 具有极高的空间分辨率, 是国际上广泛采用的深空探测器高精度角位置测量手段^[1]。相关处理机是 VLBI 数据预处理的核心设备。由于 VLBI 观测数据的相关处理具有数据密集和计算密集的双重特点, 普通计算机的性能难以达到对数据处理速度的要求, 需要有专用的高速硬件数据处理单元来完成相关处理。我国探月工程 VLBI 地面测控系统中所使用的硬件相关处理机是基于大容量、高性能的在线可编程门阵列 (FPGA), 自主研制的高速数据处理单元, 可以实时处理每台站最高速率达 1 Gb/s 的数据, 其数据处理结果, 经过软件验证并与欧空局 (ESA) 提供的数据进行了比较, 结果符合良好。

关键词: 天文观测设备与技术; 甚长基线测量; 相关处理机; 数据回放; 相关处理; 模型计算; 长期累加

中图分类号: P228.6

文献标识码: A

1 引 言

甚长基线干涉测量 (Very Long Baseline Interferometry, VLBI) 是 20 世纪 60 年代后期发展起来的射电干涉新技术^[2], 是一种重要的射电天文技术。它采用无线电干涉法, 将多个位于距离数百乃至数千公里长度基线两端的口径较小的射电望远镜, 合成为巨大的综合孔径望远镜, 其最大等效直径为望远镜之间的最大基线长度。VLBI 通过延长基线和提高观测频率可获得极高空间分辨率和基线测量精度。VLBI 是目前角分辨率最高的天文观测技术, 已获得优于亚毫角秒量级的极高分辨率, 相当于在地球上能分辨出月球上篮球大小的面积, 也能够以毫米的精度测量上千公里的基线长度。因此它在天体物理学、天体测量学和天文地球动力学等领域得到广泛的应用^[3]。

收稿日期: 2007-05-17; 修回日期: 2007-07-18

基金项目: 探月工程 -1-0-1 项目 (0384001002)

VLBI 以高稳定的氢原子钟信号作为独立频率标准, 各观测站的射电望远镜接收无线电信号, 将其变换为基频信号后, 进行数字化采样, 并记录或传输。在传统的 VLBI 技术中, 观测站获得的海量数字化数据先记录在磁带上, 再传送到数据处理中心的相关处理机进行事后相关处理 (见图 1)。

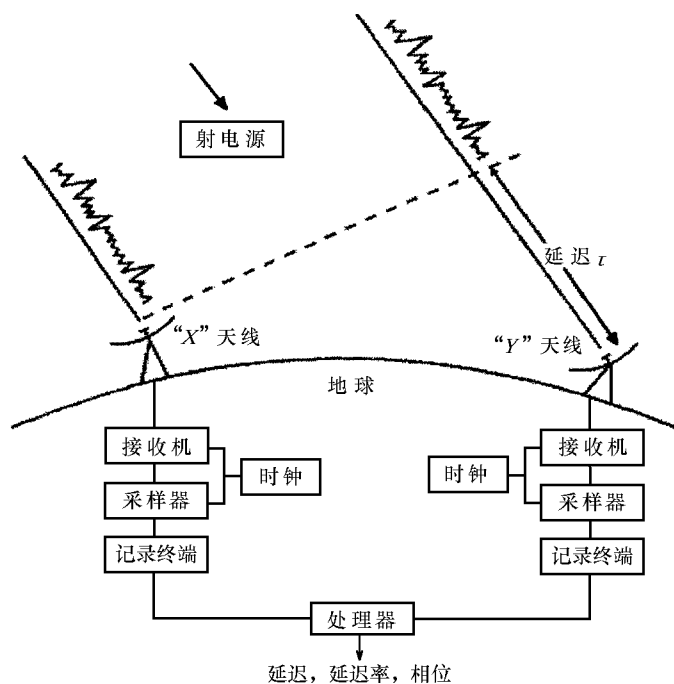


图 1 VLBI 系统原理^[4]

相关处理机是 VLBI 的数据预处理核心设备, 是复杂的高速信号处理系统, 观测信号的相关函数或复可见度函数 (互谱密度函数) 就在相关处理机中形成。目前的相关处理机是用于 VLBI 卫星观测数据处理的五台站 FX 型频域相关处理机, 完成数据的格式转换、时间同步、延时补偿、FFT 运算、复数交差相乘累加、长期累加, 最后输出计算结果, 供后处理使用。

在我国探月工程中, 要求 VLBI 系统具有数据的实时传输与实时处理能力, 因此与传统的处理机相比, 新的相关处理机需要满足相应的技术要求。

2 相关处理原理

地球上位于不同位置的两架射电望远镜, 可以接收到来自宇宙深处射电源发出的电磁辐射。不同的望远镜接收到同一波前的信号时, 由于位置的差异, 到达时间会有差异, 即图 1 中的延迟 τ 。

以图 1 中的“X 天线”为参考, 假设 X 天线收到的信号为 $V_X = \cos \omega t$ (为简化计算, 不妨假设幅度为 1, 初始相位为 0), Y 天线收到的信号为 $V_Y = \cos \omega(t - \tau)$, 两地本振信号均为 $V_L = \cos \omega_0 t$ 。经混频后, 取各自的差频信号 $V'_X = \cos(\omega - \omega_0)t$, $V'_Y = \cos[(\omega - \omega_0)t - \omega\tau]$ 。

将两者做相关, 可以得到:

$$\begin{aligned}
 r_{X,Y}(\tau_{g0}) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} V'_X(t) V'_Y(t + \tau_{g0}) dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \cos(\omega - \omega_0)t \cos[(\omega - \omega_0)(t + \tau_{g0}) - \omega\tau] dt \\
 &= \cos[(\omega - \omega_0)\tau_{g0} - \omega\tau] \\
 &= \cos[(\omega - \omega_0)(\tau - \tau_{g0}) + \omega_0\tau], \tag{1}
 \end{aligned}$$

其中 τ_{g0} 为延时补偿, 定义几何延迟 τ 与实际补偿延时 τ_{g0} 之差为残余延迟 $\Delta\tau_{g0}$, 即:

$$\Delta\tau_g = \tau - \tau_{g0}.$$

当 τ_{g0} 无限趋近于 τ 时, $\Delta\tau_g$ 趋近于零, 式 (1) 的结果为趋近于 $\cos\omega_0\tau$, 此时, 输出仅与几何延迟 τ 有关. 式 (1) 的结果随 τ 而周期变化, 这种周期性变化称为“干涉条纹”. 相关处理机的任务就是找到干涉条纹, 计算出几何延迟 τ .

3 相关处理机组成

相关处理机主要由数据回放 (Playback)、模型计算 (Module Parameter Calculation)、相关处理 (Correlation Processing)、长期累加器 (Long Term Accumulator, LTA)、中心控制计算机 (Center Control Computer, CCC) 五大模块组成, 如图 2 所示.

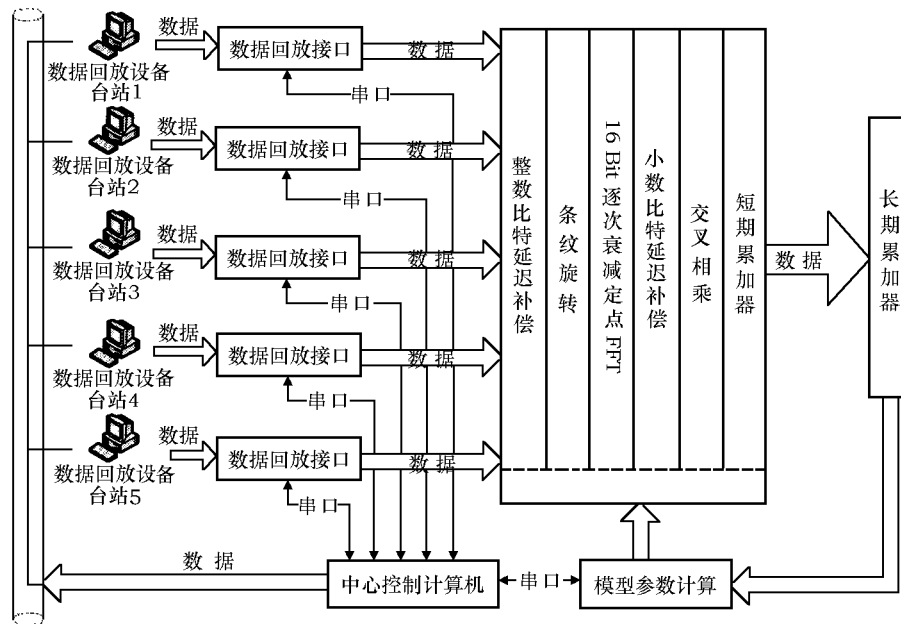


图 2 相关处理机结构示意图

数据由网络开始, 各台站数据经过各自的数据回放设备 (Playback Device, PBD) 缓存后, 由 PCI7300A 数据采集卡以 32 位并行形式送至相应的数据回放接口 (Playback Interface, PBI), 经 PBI 对数据格式进行处理后, 按照要求的回放模式送给相关处理模块。相关处理模块同时接收 PBI 的数据和模型计算模块送出的模型参数, 进行相关运算, 结果以每通道 1 位的形式送给 LTA, LTA 将其拼接成 16 位宽度, 送至 CCC, 最后仍由网络送出。

各部分在中心控制计算机的协调下工作, 每台站峰值数据速率可达 1 Gb/s, 经处理后的数据量大大减小, 每台站约为原数据量的几十分之一至几百分之一。

3.1 数据回放

数据回放模块分为 PBD 和 PBI。PBD 由工控机担任, 主要任务是从网络上接收数据并做缓存, 再经 PCI7300A 数据采集卡送给 PBI。

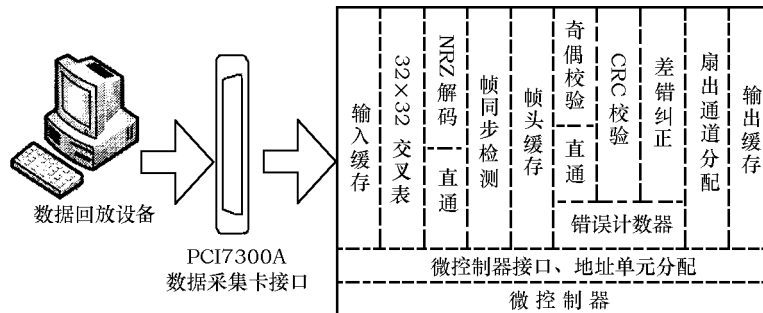


图 3 数据回放模块内部结构

PBI 部分功能由 Xilinx 公司的 FPGA — XC2V3000 外加一片 Winbond 的微控制器 — W77LE58 实现的。微控制器负责与中心计算机通过串口进行通信, 处理各种指令, 将指令转换成各种参数, 对 FPGA 内部各模块进行配置。FPGA 中各子模块将对数据进行解码、帧同步检测、校验、差错纠正、提取帧头信息等操作, 并按所要求的时间进行同步, 同时将时间信息通过串口返回给中心计算机。

相关处理机接收的数据有两种格式, 一种是 MK IV 原始数据格式^[5], 符合 MK IV 标准。帧长 22 500 位, 有 NRZ 编码, 帧头 180 位, 帧头有 CRC 校检, 全帧除帧头同步字带偶校验外, 其余部分均为奇校验。另一种是经 MK V 网络回放后的数据, 与 MK IV 原始数据格式不同的是, 其帧数据无 NRZ 编码, 且已剔除奇偶校验位, 故帧长为 20 000 位, 帧头 160 位, 帧头有 CRC 校验。

FPGA 中实现的逻辑功能是 PBI 的核心功能, 它包括了 12 个子模块, 工作在 32 MHz 时钟下, 对 32 位并行数据作并行的逻辑运算, 总的带宽可达到 1 Gb/s。

3.1.1 输入缓存

输入缓存是一个 $8\ 191 \times 32$ 比特数据宽度, 先进先出的异步双端口 RAM, 称为异步 FIFO。其一端与 PCI7300A 数据采集卡时序相匹配, 使用数据采集卡的时钟, 另一端与 FPGA 内部逻辑匹配, 使用系统时钟, 实现了输入部分两个时钟域的过渡。

3.1.2 交叉表

交叉表是一个 32×32 的数据选择器, 可以将 32 位的输入数据通过参数设置的方式以任

意的位组合方式分配到 32 位输出端。

3.1.3 NRZ 解码

对于 MK IV 格式的原始数据, 为避免传输过程中出现相位模糊, 采用了信源编码, 用相邻时刻电平的变化来表示真正的数据“0”和“1”。若当前时刻电平和前一时刻电平相同(同为“0”或同为“1”), 表示当前时刻真实数据为“0”; 若当前时刻接收电平和前一时刻接收电平不同(前一时刻为“0”, 当前时刻为“1”; 或者前一时刻为“1”, 当前时刻为“0”), 表示当前时刻真实数据为“1”。因此, 要得到真实数据, 必须进行解码, 即将当前时刻和前一时刻的接收数据进行异或, 得到当前时刻的真实数据。

然而从 MK V 中通过网络送出的数据, 是已经经过解码的真实数据。因此, 该子模块分为两个通道—解码通道和直通通道。根据数据格式的不同, 通过不同的指令对通道进行选择。

3.1.4 帧同步检测

帧同步检模块中, 包含了帧头检测和时间同步两部分。帧头检测是指找到每一帧的帧同步字, 并输出每帧的帧同步信号。通过帧同步字的定位可以从帧头中提取信息, 包括台站信息、记录模式信息、时间信息, 并做一些校验等。利用帧同步信号, 可以进行帧完整性检测。时间同步是找到具有与中心控制计算机送来的同步时间相同时间的帧, 并从这一帧开始, 将以后的数据送入后级缓存。

3.1.5 帧头缓存

在一帧数据中, 帧同步字位于帧头的中间。当检测到同步字时, 为了不使同步字之前的数据丢失, 需要一个缓存, 其大小至少需容纳每帧同步字之前的数据。在本文中, 该缓存是由一个 160×32 位宽度的同步 FIFO 完成。

3.1.6 奇偶校验

MK IV 格式的原始数据中, 同步字部分采用偶校验, 其他部分均采用奇校验。这样的好处是不会将数据误当成同步字。每发现一个错误, 错误计数器加 1。

在 MK V 网络输出的数据中, 不存在奇偶校验。因此, 该子模块和解码模块类似, 分为两个通道, 根据数据格式的不同, 通过不同的指令进行选择。

3.1.7 CRC 校验

帧头部分在剔除奇偶校验位后, 为 160 位。前 148 位为信息数据, 最后 12 位为 CRC 校验值。生成多项式为 $x^{12} + x^{11} + x^3 + x^2 + x + 1$ 。在该模块中对帧头部分进行 CRC 校验, 每发现一个错误, 错误计数器加 1。

3.1.8 差错纠正

在数据记录或数据回放过程中, 可能会受到来自外界的干扰影响, 使得数据出错, 如帧长度改变, 同步头受损等。若不加以纠正, 会给后端的相关处理造成很大的困难, 甚至导致处理无法进行。因此在这部分子模块中, 采用了两套措施来尽可能保证帧的完整性和前后帧时间的连续性。一是计算相邻帧同步信号之间的数据位数是否等于帧长, 若大于帧长, 则舍弃多余部分, 直到下一个帧同步信号出现。若小于帧长, 则插入数据位, 直至达到帧长。二是根据记录模式, 计算出两帧之间时间步长, 当相邻两帧时间大于一个步长时, 在这两帧之间插入足够的的数据位(帧长的整数倍), 直至两帧之间满足时间连续性。

3.1.9 错误计数器

在奇偶校验、CRC 校验、帧完整性检测过程中, 一旦发现错误, 该计数器会加 1, 用于统计出错次数。

3.1.10 扇出通道分配

根据数据记录的模式, 按 1:1、1:2 或 1:4 扇出。通过前端的交叉表对通道的分配, 扇出部分能实现对输入 32 个通道中任意 1 个、2 个或 4 个通道的组合。

3.1.11 输出缓存

缓存输出数据, 使得与后级接口的逻辑相对独立, 便于接口逻辑的更改与实现。

3.1.12 微控制器接口、地址分配

配合微控制器的总线时序, 将微控制器从中心计算机接收并转换好的参数送到相应的模块。

3.2 模型计算

模型计算模块负责硬件相关处理机线性模型的实时计算, 它需要由中心控制计算机实时提供当前模型时段所有观测站的时延多项式模型^[6], 以及相关的辅助信息; 根据这些信息, 实时计算每个模型周期所需的参数(模型参数), 送相关处理模块, 相关处理模块根据这些参数进行时延修正、条纹旋转等操作。另外, 还可以根据中心控制计算机的指令置某个台站无效, 即在处理中忽略该台站的数据。

3.2.1 模型参数定义

硬件处理机的线性模型周期 Δ 为 64 个 FFT 周期, 即 $\Delta = \frac{64 \times N}{f_s}$ (其中 N 为单个采样周期内的点数, f_s 为采样频率), 计算线性模型的 5 个参数: a 为相位延迟(周); b 为相位延迟率(周/位); c 为整数位延迟(位); d 为小数位延迟(位); e 为延迟率(位/64)。

3.2.2 模型参数的计算

时延五次多项式 $x(t) = a_5 t^5 + a_4 t^4 + a_3 t^3 + a_2 t^2 + a_1 t + a_0$, 设 Δ 为如前所述的线性模型周期, 则 $t = 0, \Delta, 2\Delta, \dots, (N-1)\Delta$, 由此可计算出 1 min 内每个模型周期的时延值 τ ; 另对时延多项式进行求导得四次多项式为时延率多项式 $x'(t) = 5a_5 t^4 + 4a_4 t^3 + 3a_3 t^2 + 2a_2 t + a_1$, 时延率多项式表示延迟的变化率, 由此可计算出一分钟内每个模型周期的时延率 $\dot{\tau}$ 。

时延值 τ 换算成位数, 其整数部分即为整数位延迟补偿, 小数部分为小数位延迟补偿, 换算计算式:

$$c = \left[\frac{\tau}{T_s} \right] = [\tau \times f_s], \quad [] \text{ 表示取整}$$

小数位延迟补偿 $= \tau \times f_s - c < 1$, 整数位延迟补偿通过移动数据流的位数来实现, 小数位延迟则需要 FFT 以后(频域)通过角度的补偿实现。根据 FFT 的性质:

$FFT\{x(n + \tau_f)\} = FFT\{x(n)\} \times e^{j \times 2\pi \times i \times \tau_f / N}$, $i = 0, 1, \dots, N/2 - 1$ (根据 FFT 的对称性, 只要计算一半的点就可以了), 在频域每点都需要进行补偿, 故: $d =$ 小数位, $c = \dot{\tau} \times N$ 条纹旋转是为了补偿多普勒频移, 计算为

$$a = f \bmod (\text{delay} \times \text{skyFreq}, 1),$$

$f \bmod$ 表示求余数, a 为小于 1 的数。

$$b = \tau \times \text{skyFreq} / f_s.$$

在实际使用中, 由于硬件处理机的运算在 FPGA 芯片中用定点运算实现, 考虑到精度的要求, 除参数 c 采用 32 位表示外, 其余的 a 、 b 、 d 、 e 都采用 40 位表示, 共 192 位数据。

3.3 相关处理

相关处理模块是相关处理机的核心模块, 该模块将利用从模型计算模块得到的模型参数对数据进行一系列的数学运算, 包括整数比特延迟补偿、条纹旋转、快速傅里叶变换、小数比特延迟补偿、交叉相乘, 并进行短期累加, 最后输出给长期累加器。

3.3.1 模型参数接收

该子模块负责从模型计算模块中串行地获得所需要的 5 个模型参数, 即前文所述的 a 、 b 、 c 、 d 、 e 。在 5 台站情况下, 一个模型周期内总共需要 192×5 位的数据。模型周期设定为 64×1024 位。

3.3.2 整数位延迟补偿 (DELAY)

DELAY 子模块在每个模型周期开始时, 根据得到的 c 对每个台站的数据进行整数比特延迟补偿。

当 $c > 0$ 时, 表示该台站需要去掉 c 位数据, 此时, 让该站的 PBI_EN 信号先有效 c 个时钟周期, 则 PBI 会输出 c 位数据, 若本地不接收, 则此 c 位数据丢弃, 从而实现了 c 位数据的延迟补偿。

当 $c < 0$ 时, 表示该台站需要补充 c 位的数据。此时, PBI_EN 信号会在系统读取数据后等待 c 时钟周期再为有效, 则等效于该台站补充了 c 位数据。

3.3.3 条纹旋转 (FRINGE STOPPING)

首先由相位延迟参数 a 和相位延迟率参数 b 得到相位延迟:

$$\phi_i = a + b \cdot (i - 1) \quad i = 1, 2, 3 \cdots \text{模型周期总位数} (64 \times 1024 \text{ 位})。$$

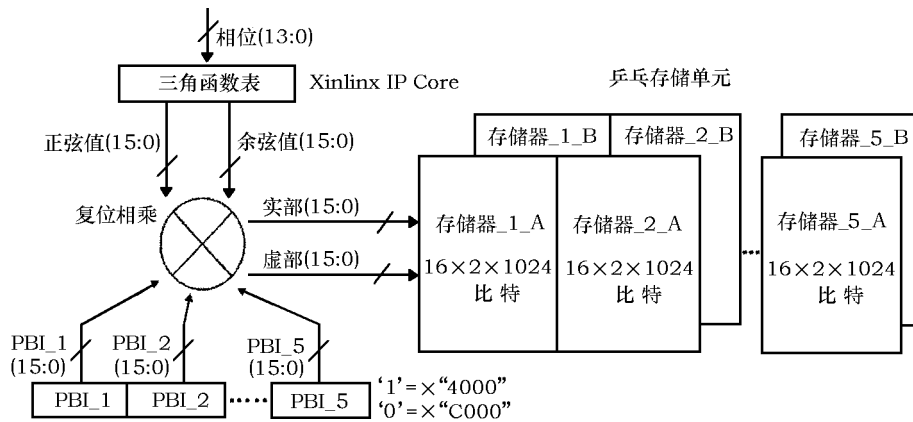


图 4 条纹旋转

相位延迟为 14 位无符号数, 送入正弦-余弦查找表, 这个使用的是 Xilinx 公司的 IP Core, 版本为 5.0。输出为 2 个 16 位的正弦和余弦参数, 这 2 个参数与输入数据进行复数相乘。输入数据为 PBI 的数据经过转化得到, PBI 数据为 '1' 时输入数据为 \times "4000", 为 '0' 时则为 \times "C000"。条纹旋转公式为 $X_i \times e^{j \times 2\pi \times \phi_i}$ 。

复数相乘后的数据存入 5 组乒乓 RAM 中。每组容量为 $16 \times 2 \times 1024$ 比特。每一组对应一个台站数据的处理结果。

3.3.4 快速傅里叶变换 (FFT)

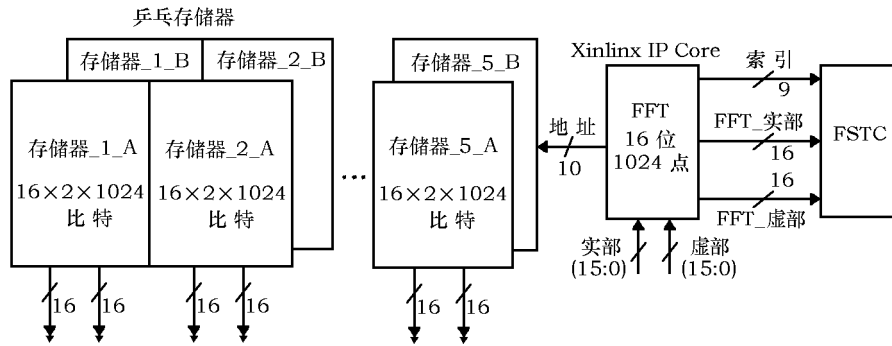


图 5 快速傅里叶变换

FFT 模块将 5 组乒乓 RAM 的数据依次取出进行 FFT, 顺序为存储器_1_A \Rightarrow 存储器_2_A \Rightarrow 存储器_3_A \Rightarrow 存储器_4_A \Rightarrow 存储器_5_B \Rightarrow 存储器_1_B \Rightarrow 存储器_2_B \Rightarrow 存储器_3_B \Rightarrow 存储器_4_B \Rightarrow 存储器_5_B \Rightarrow 存储器_1_A \dots 如此循环。

FFT 的 IP CORE 采用的是 Xilinx 的 Fast Fourier Transform v3.2。参数为 1024 点 FFT, 逐次递减。输入为 16 位实部和 16 位虚部, 输出同样为 16 位实部和 16 位虚部。输出的数据将送入下个模块 FSTC。

3.3.5 小数位延迟 (FSTC)

小数位延迟补偿先根据 d 和 e 得到当前模型周期第 f 个 FFT 周期的小数位延迟:

$$\tau_f = d + e \cdot (f - 1).$$

$f = 1, 2, 3, \dots$ 模型周期总 FFT 个数 (1024)。

设 FFT 周期包含 N 个复数点, 数据经过 FFT, 复数点为 $Y_i, i = 0, 1, 2, \dots, N/2 - 1$ 。则第 i 个复数点的 FSTC 为

$$Y_i \times e^{j \times 2\pi \times i \times \tau_f / N}.$$

复数相乘后的结果送入 5 组乒乓 RAM 中, 每组容量为 $16 \times 2 \times 512$ 比特, 每一组对应一个台站数据的处理结果。

3.3.6 交叉相乘 (MAC)

MAC 是对补偿后的数据交叉相乘, 得每台站的自相关和每 2 台站的互相关谱 $Z_1 \times \text{conj}(Z_2)$ 。

FSTC 输入的 16 位实部和虚部经过复数相乘后, 得到 32 位的相关结果, 送入短期累加器进行累加。

3.3.7 短期累加器 (STA)

累加周期为 4096 次, 为了防止溢出, 需要增加 12 位。累加时, 当前 32 位的相关结果首先增加 12 个符号位成为 44 位的数据, 该数据与累加存储器内对应地址的数据进行累加, 并将结果覆盖入该地址。当最后一次累加完成后 (4096 次), 数据将存入输出存储器准备输出到 LTA 板, 并清空累加存储器。

3.4 长期累加器

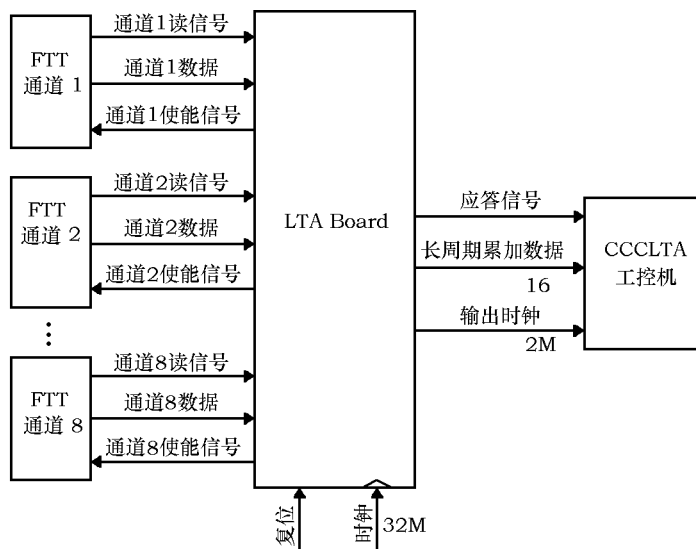


图 6 长期累加

LTA 首先判断 CCCLTA 的 ACK 信号, 当 ACK 为高时, 系统开始分别接收 8 个 FFT 通道的数据, 顺序为第一个通道到第 8 个通道, 并不断循环。接收的 32M 1 位数据串转并成 16 位数据, 并通过 CLOCK_OUT 2M 的时钟信号, 将 LTA_DATA 数据送入中心计算机的 PCI7300A 数据采集卡。

3.5 中心控制计算机

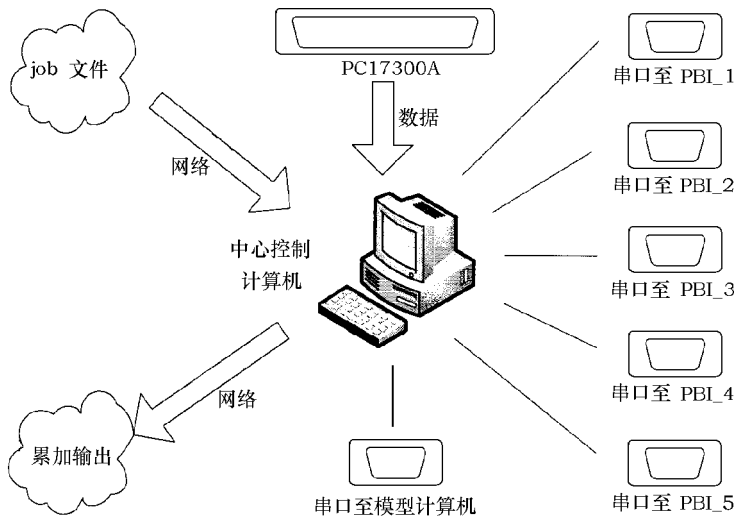


图 7 中心控制计算机

中心控制计算机中同时装有千兆以太网卡、多串口卡、以及 PCI7300A 数据采集卡, 实现从网络硬盘中读取作业文件, 按照作业要求, 通过串口对 PBI 进行参数配置, 给模型计算提供拟合

数据, 同时通过数据采集卡接收 LTA 送来的数据, 进行累加后回送至网络硬盘供后处理使用。

4 相关处理机处理结果

2006 年 5 月 27 日—6 月 2 日, 硬件相关处理机参加了对欧空局的 Smart-1 月球探测器测轨工作。根据 ESA 轨道, 轨道计算组对硬件相关处理结果和软件相关处理结果分别做了残差计算, 图 8 是 2006 年 5 月 30 日的延迟残差图。黑色代表硬件结果, 灰色代表软件结果。

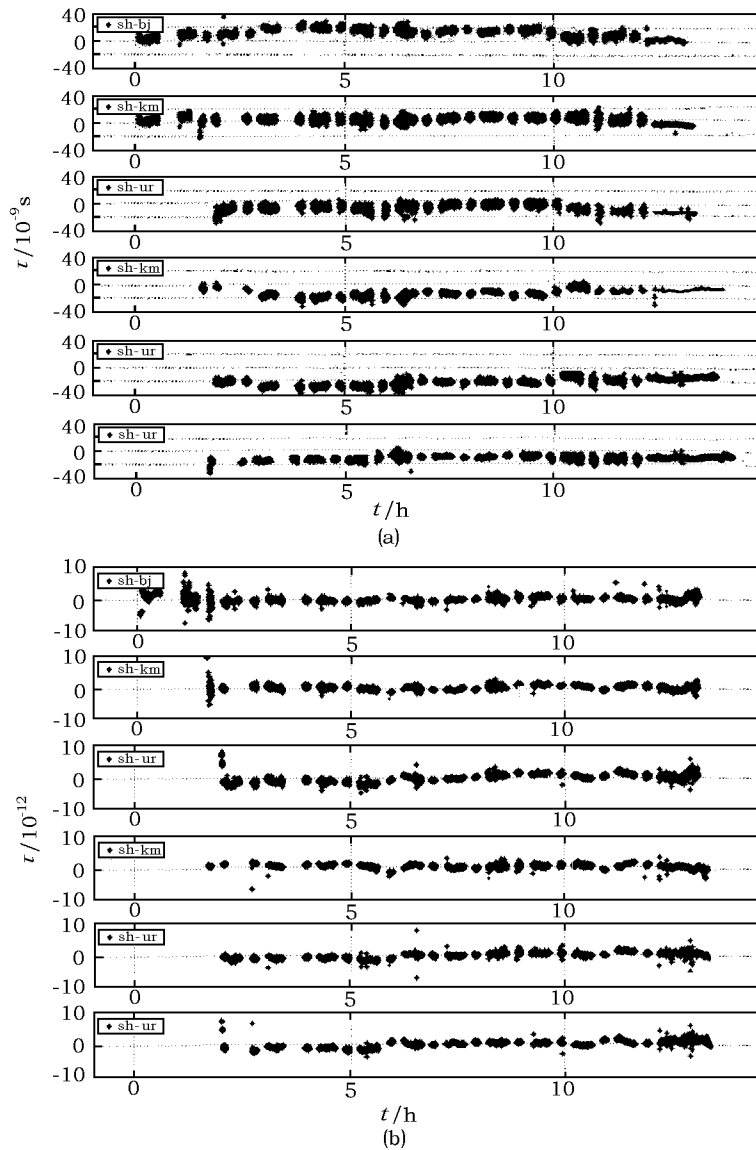


图 8 相关处理机结果比较

(a) 延迟残差; (b) 延迟率残差

可以看出, 软、硬件相关处理机的结果符合。两相关处理机的结果与根据 ESA 轨道计算的延迟残差和延迟率残差结果比较, 符合较好。

5 结束语

该处理机是目前中国第一台国内自己研制的可同时用于卫星观测和天文观测的硬件相关处理机, 其技术接近国际先进水平, 现已通过测试, 并将投入我国探月工程一期 VLBI 测控系统的正式运行。

参考文献:

- [1] 郑为民, 杨艳. 世界科技研究与发展, 2005, 27(5): 7-15
- [2] Thompson A R, Moran J M, Jr Swenson G W. Interferometry and Synthesis in Radio Astronomy. New York: John Wiley & Sons, 1986: 29-33
- [3] James U, Miller G. Sky & Telescope, 1999, 98(4): 36-41
- [4] Whitney A R. IVS 2000 General Meeting Proceedings. Vandenberg N R, Bayer K D, eds. Maryland: NASA, 2000: 187-205
- [5] Whitney A R. Mark IIIA/IV/VLBA Tape Formats, Recording Modes and Compatibility Revision 1. 21 <http://www.haystack.mit.edu/geo/mark4/memos/230.pdf>, MIT: Haystack Observatory, 2005
- [6] Fujinobu T, Tetsuro K, Yukio T, et al. Very Long Baseline Interferometer, Tokyo: Ohmsha Ltd, 1997: 73-80

An Introduction to the VLBI Correlator Used in the Chinese Lunar Exploration Program

ZHU Ren-jie, ZHANG Xiu-zhong, XIANG Ying, XU Zhi-jun

(Shanghai Astronomical Observatory, Chinese Academy of Science, Shanghai 200030, China)

Abstract: As a radio astronomical technique with very high angular resolution, Very Long Baseline Interferometry (VLBI) has been widely used in the high precision deep space probe tracking activities. The VLBI correlator is one of the key data-preprocessing device. Since the correlation process in VLBI has the characteristic of intensive data and computation, ordinary computer can hardly achieve the request for high speed data processing. Therefore, a specially desined high-speed hardware correlator is normally needed. The hardware correlator used in the Chinese Lunar Exploration Program (CLEP) is an independently developed, high-speed data processing system based upon Field Programmable Gate Array (FPGA) with large volume and high performance, which can deal with data up to 1Gbps per station at real time. The result is good through validations with soft correlator and data from ESA.

Key words: astronomical facilities and technique; Very Long Baseline Interferometry; correlator; playback; correlation; model calculation; long term accumulator