Vol. 42 No. 1 *Journal of the Communications Research Laboratory* *March 1995*
Printed in Tokyo, Japan *pp. 31 – 42*

II. OVERVIEW OF THE EXPERIMENT SYSTEM

II.4 DATA PROCESSING SOFTWARE AND DATA ANALYSIS SOFTWARE

By
Yukio TAKAHASHI, Mitihito IMAE, Noriyuki KURIHARA, Yuji SUGIMOTO, Taizoh YOSHINO, Fujinobu TAKAHASHI, Shin'ichi HAMA, Hiroshi TAKABA, Takahiro IWATA, Yashuhiro KOYAMA, Yuko HANADO, Mamoru SEKIDO, Jun'ichi NAKAJIMA, Tetsuro KONDO, Jun AMAGAI, and Akihiro KANEKO

(Received on November 21, 1994)

ABSTRACT

We developed the data processing and analysis software systems for the Western Pacific VLBI network project, and we introduce the software systems in this paper. A new computer system was introduced for the Western Pacific VLBI network project. As the processing system uses the new K-4 recorder, and we accommodate the processing software to the K-4 recorder using new computer HP9000-330. This is the first data processing system available for the K-4 recorder. In the data analysis, we use Mark-III software package developed by NASA (USA), such as the software for the Mark-III data base, the software for the calculated delay and the parameters estimation software. We present an overview of the software system. We also developed the software to set up the data into the Mark-III data base for the new computer system.

Keywords: VLBI, software, analysis, data processing, K-4

1. Introduction

The "Mark-III" VLBI system including the data processing software and data analysis software has been used worldwide for the geodetic VLBI experiments since 1975[1][2]. The Communications Research Laboratory (CRL) has participated in the international and domestic VLBI experiments since 1984. As the basis of the Mark-III system, we had developed our own software system (K-3 software system) which is compatible with the Mark-III system. For the data processing system, we developed the two types of software; the correlation processing software (KROSS)[3][4] compatible with the software "COREL" in the Mark-III system[1], and the bandwidth synthesis software "KOMB"[4][5] compatible with the Mark-III software "FRNGE." We also developed the four types of software for the data analysis; the software for the data base handler (KASTL) using the HP data base handler (IMAGE1000)[4][6], the software to set up the data into data base (KASET)[4][7], the software to get the calculated delay and delay rate (KAPRI)[4][8]-[10] which corresponds to the software "CALC"[2] in the Mark-III system, and the software to estimate the parameters (KLEAR)[4][11] which corresponds to the software "SOLVE"[2] in the Mark-III system. The K-3 software system was developed using an HP1000-45F computer system.

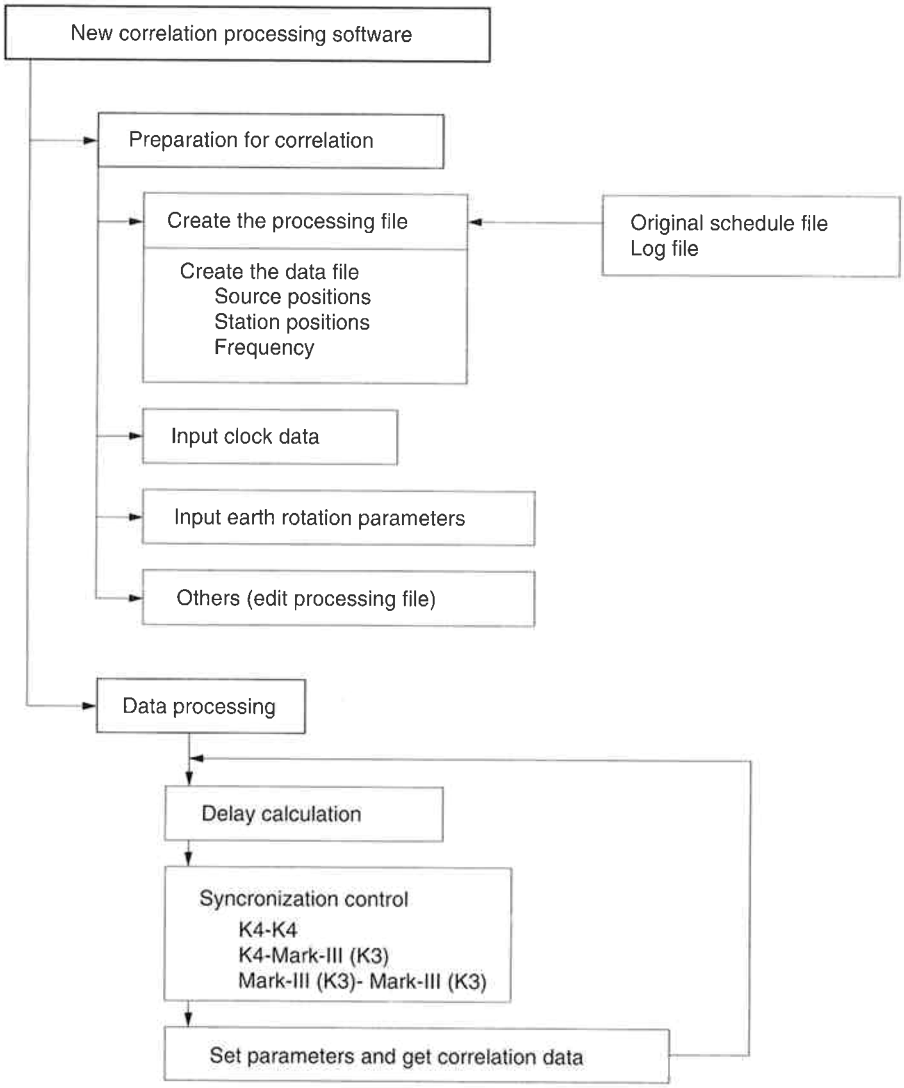

Fig. 1 The outline of the new correlation software

The Marcus and Daito stations in the Western Pacific VLBI network project are located on the island in the Pacific Ocean, and therefore the VLBI acquisition system had to be transportable. We developed the new VLBI system which uses the compact K-4 recorder(12). We call it "K-4" system. In the Western Pacific VLBI network project, the new computer systems (HP1000 A900 and HP9000 330) which are three times faster than the previous systems were introduced for the data processing and data analysis. We developed the new data processing system for the K-4 recording system and the analysis system for the new computer system(4).

The Mark-III data analysis software used in the HP1000 A900 computer are also developed by NASA. For the compatibility, we adopted the Mark-III data analysis software, such as the software for the data base handler, the software (CALC) for the calculated delay and delay rate, and the parameter estimation software (SOLVE). Thus, they are easy to install in the computer system and they are compatible with the worldwide data base used in geodetic VLBI.

However, we had developed our own software to set up the data into the Mark-III data base. We conducted some of the experiments in the Western Pacific VLBI network project together with the stations in the Crustal Dynamics Project (CDP), such as Kauai (Hawaii), Gilmore-Creek (Alaska). In this case we had to send a Mark-III (K-3) tape which we copied from a K-4 tape recorded at Marcus. We developed the data converting software, which will be described later in this paper.

2. Data Processing Software

CRL developed a new recording system and called it the K-4 recording system[12]. New correlation processing software was also developed for this K-4 recording system.

The correlation processing software must control both the recorders and correlation processor. We developed the processing software using BASIC language using HP9000-330 personal computer since it can control the commands of GP-IB sequentially and it is especially easy to check and modify.

Bandwidth synthesis software runs in the HP1000-A900 minicomputer. Correlation data is transferred from the HP9000-330 computer for correlation processing, to the HP1000-A900 computer.

Figure 1 shows an outline of the new correlation processing software. Some files in the software are created in order to manage the correlation processing. These files are automatically created using only the original schedule file for individual experiments.

In the correlation processing, the data of a remote station is adjusted to the data of the reference station. This time adjustment is made using the calculated delay and delay rate for every parameter period (PP) of a few seconds, and the correlation data is integrated. Their calculated delay and delay rate for every PP are obtained using the 0th deviation (τ), the 1st deviation ($\dot{\tau}$), the 2nd deviation ($\ddot{\tau}$) and the 3 order derivation ($\dddot{\tau}$) with respect to time. In the old software, these values are the coefficients of Taylor expansion at the center of the observation. The difference between the calculated delay and the true delay is larger at the beginning and end of each observation in according to be far from the center of the observation. In this software, the 0th, 1st, 2nd, 3rd deviations are obtained using the approximation of the four order polynomial formulation. The entire duration of each observation is divided into 4 parts ($T_0 - 2\Delta T$~$T_0 - \Delta T$, $T_0 - \Delta T$~T_0, T_0~$T_0 + \Delta T$, $T_0 + \Delta T$~$T_0 + 2\Delta T$ while $4\Delta T$ is the complete duration, and T_0 is the center of the observation and it is reference time of the observation). The values of 0th, 1st, 2nd, 3rd order deviations are calculated for the least square fittings as follows;

$$\begin{aligned}
\tau &= \tau_0 + \frac{\left(12(\tau_2 + \tau_3 - 2\tau_0) - 3(\tau_1 + \tau_4 - 2\tau_0)\right)}{35} \\
\dot{\tau} &= \frac{\left(\tau_1 - \tau_4 - 8(\tau_2 - \tau_3)\right)}{12\Delta T} \\
\ddot{\tau} &= \frac{\left(2(\tau_1 - \tau_0) - (\tau_2 - \tau_4) - (\tau_3 - \tau_4)\right)}{7(\Delta T)^2} \\
\dddot{\tau} &= \frac{\left(2(\tau_2 - \tau_3) + (\tau_4 - \tau_1)\right)}{2(\Delta T)^3}
\end{aligned} \qquad \cdots\cdots (1)$$

where τ_1, τ_2, τ_0, τ_3, τ_4 are the delays at $T_0 - 2\Delta T$, $T_0 - \Delta T$, T_0, $T_0 + \Delta T$, $T_0 + 2\Delta T$ for each observation.

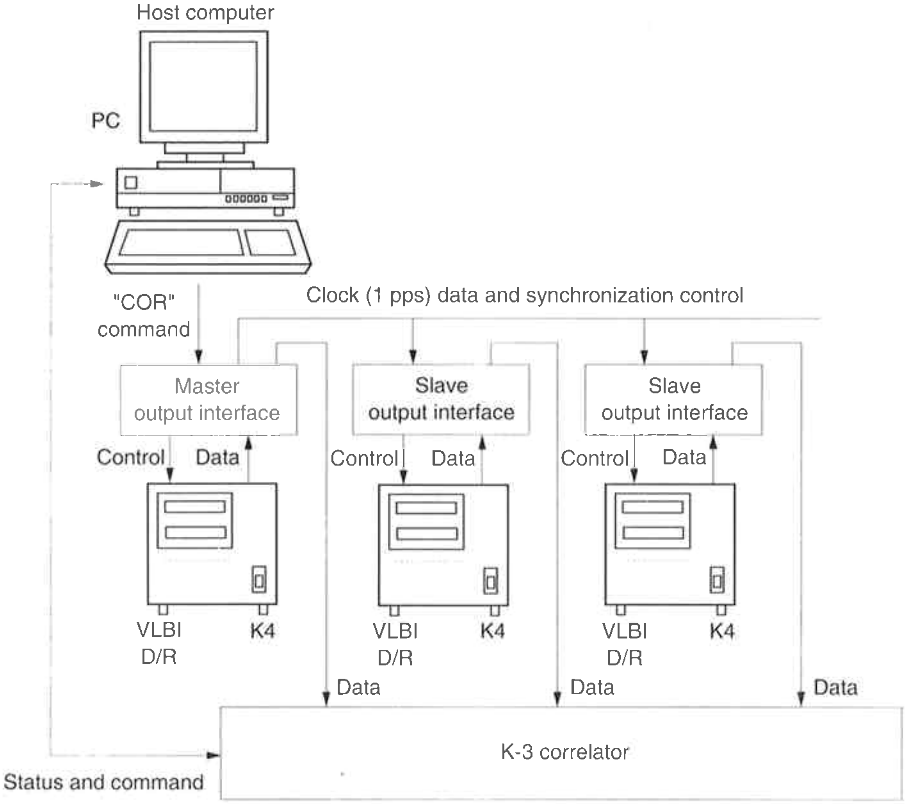

Fig. 2 The correlation processing among K4 recorders

The parts of each physical model used in the correlation processing software are separated into subroutines similar to KAPRI. The software structure is arranged to use the rotation matrices of precession, nutation, diurnal rotation and wobble. Atmospheric delay is calculated by using the fixed zenith path delay and the elevation mapping function. Some stations are located in elevated regions which are over 1000 m above sea level, and the zenith path delay varies by about 100 mb (10%). Then, we corrected for height in the zenith path delay.

This software can facilitate data processing both among K-3 (Mark-III) recorders, and among the new K-4 recorders, and also the correlation between K-3 recorders and K-4 recorders. In correlations between the K-3 and K-4 recorders, the K-4 recorder serves as the master for synchronization and it is not controlled. This control of the synchronization is the same as the control among the conventional K-3 systems. One command "COR" in the recording system is used to automatically synchronize K-4 system with other K-4 systems. Before fine automatic adjustment, rough tape synchronization to within 1 sec is achieved by using the position search command "PRL." Figures 2 and 3 show the correlation processing for each system combination.

The bandwidth synthesis software developed in the HP1000-45F computer system[4][5], was modified for the new computer system (HP1000 A900).

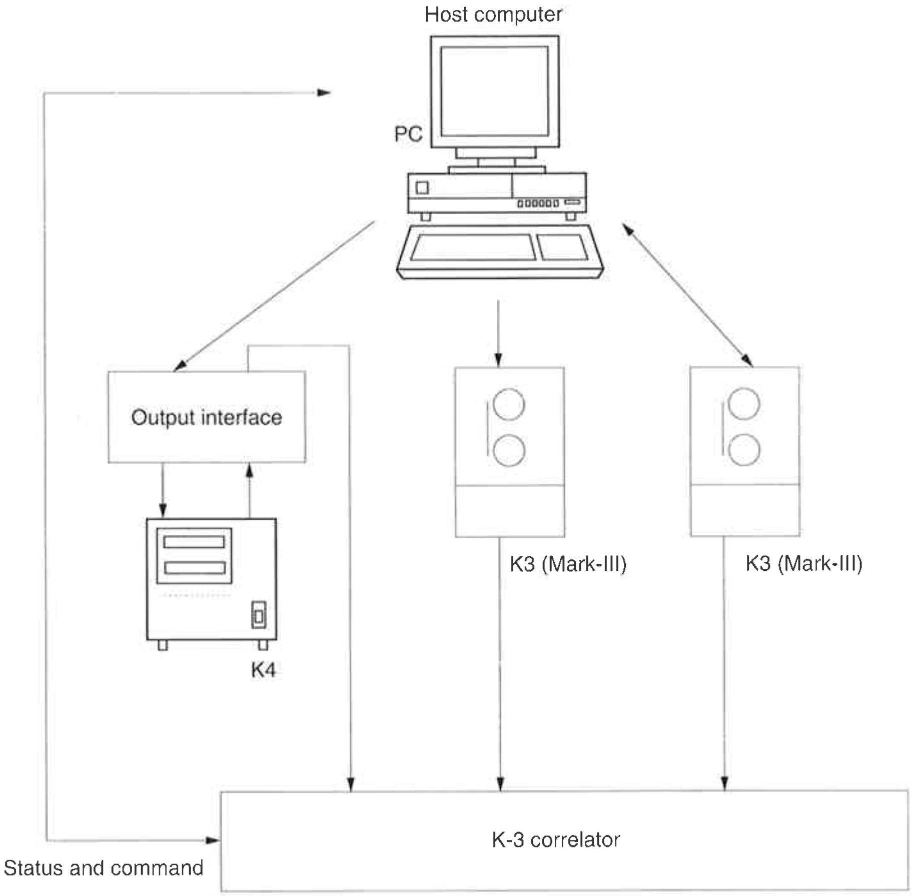

Fig. 3 The correlation processing between K4 recorder and Mark-III recorder

3. Analysis System

The analysis system consists of the data base handler, the software for the calculated delay, the parameter estimation software, and the setup software into the data base. Figure 4 shows the outline of the analysis system together with the data processing software.

3.1 Mark-III Data Base

The Mark-III data base[(2)(6)] was invented in the geodetic VLBI group of CDP. This data base is used for the exchange of VLBI data in general, and our new system uses this system, too.

The Mark-III data base consists of a single file. The data is sequentially read from the data base in the analysis, and a lot of data base must be transferred. The single file Mark-III data base is convenient. The access speed is also fast.

The data is classified as two type. One is the common data in an experiment, such as the earth rotation parameters, station positions and source positions. The other is the observation data and

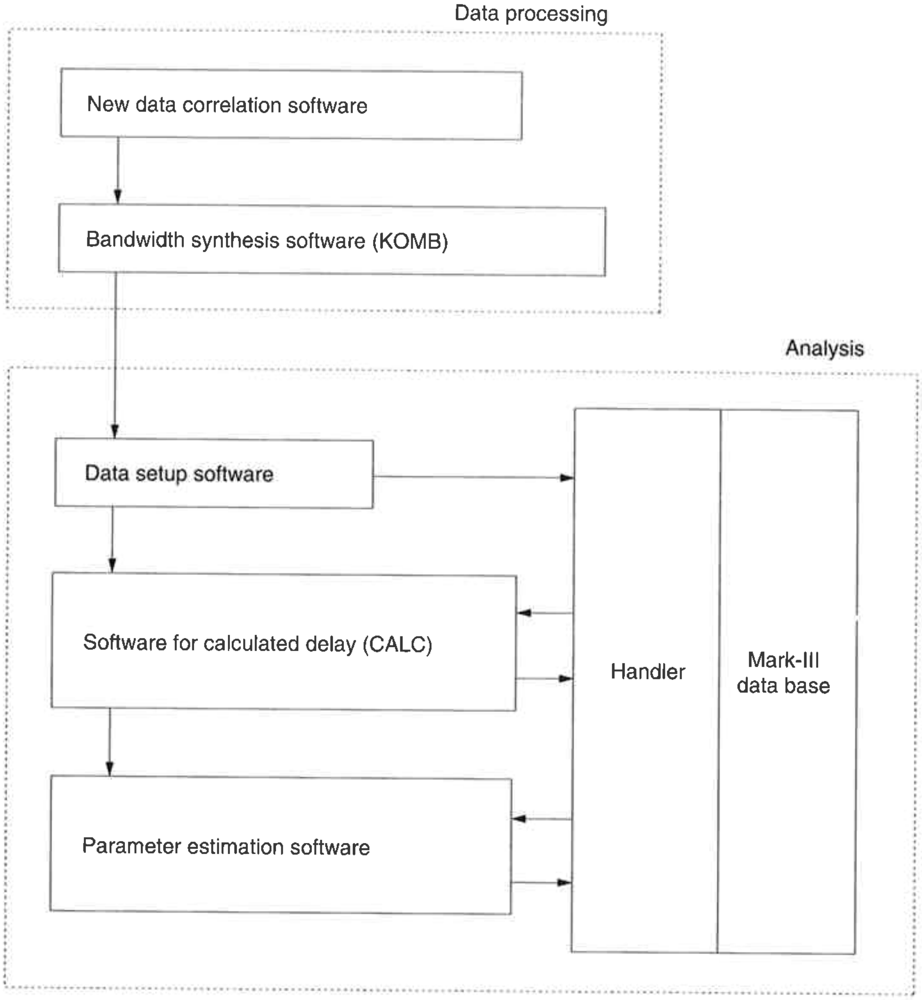

Fig. 4 The outline of the data processing and analysis software

calculated values of each occurrence. The occurrence is defined as each baseline in each observation. The data in the data base is divided into TOC (Table of Contents) 1, TOC2 and TOC3. The data of TOC1 is the common data in the experiment. The data of TOC2 and TOC3 are produced for every occurrence. The TOC2 data is mainly used for the baseline analysis, and the TOC3 data consists of the other correlation data.

The file of the data base consists of the header, the history, the definition of each data such as the array information and items, the data of TOC1, and the repeating data of TOC2 and TOC3. The file structure of the data base is shown in Table 1.

The data can be read from the data base and written in the data base using the data base handler. Figure 5 shows the method used by the program to read the data. First, the data base is opened by "KAI." The record pointer is moved by "MVREC," and the data is obtained by "GET." Finally the data base is closed by "FINIS."

Table 1 The file structure of Mark-III data base

Header of the data base		
File name, start date, experiment name, version		
History	HS	
	....	
	ZZ	
TOC1 item	TC, 1	
	TE, 1, n1, n2, n3, N1, N2, N3, n4, n5, N4, N5	
	8 character items, version, array × data number	
	32 character explanation × data number	
	n1–n5:	last number of 6 byte real, integer, ascii, 8 byte real, 2 byte integer
	N1–N5:	number of 6 byte real, integer, ascii, 8 byte real, 2 byte integer
TOC2 item	TC, 2	
	TE, 1, n1, n2, n3, N1, N2, N3, n4, n5, N4, N5	
	8 character items, version, array × data number	
	32 character explanation × data number	
TOC3 item	TC, 3	
	TE, 1, n1, n2, n3, N1, N2, N3, n4, n5, N4, N5	
	8 character items, version, array × data number	
TOC1 data	Dr, 1	
	DE, 1, ...	
	R 8 D (6 byte real data) × n1	
	+G (Integer data) × n2	
	11 (Ascii data) × n3	
	Q B. 8 D (8 byte real data) × n4	
	B: 5G (2 byte real data) × n5	
	ZZ	
repeating		
TOC2 data	Dr, 2	
	DE, 1, ...	
	R 8 D (6 byte real data) × n1	
	+G (Integer data) × n2	
	11 (Ascii data) × n3	
	Q B. 8 D (8 byte real data) × n4	
	B: 5G (2 byte real data) × n5	
	ZZ	
TOC3 data	Dr, 3	
	DE, 1, ...	
	R 8 D (6 byte real data) × n1	
	+G (Integer data) × n2	
	11 (Ascii data) × n3	
	Q B. 8 D (8 byte real data) × n4	
	B: 5G (2 byte real data) × n5	
	ZZ	

3.2 Software for the Calculated Delay (CALC)

The data analysis is made to estimate the parameters using the observation delays and delay rates. We use the least square method to estimate the parameters, such as the baseline vector, clock parameters, atmospheric delay parameters, earth rotation parameters and source positions. In this

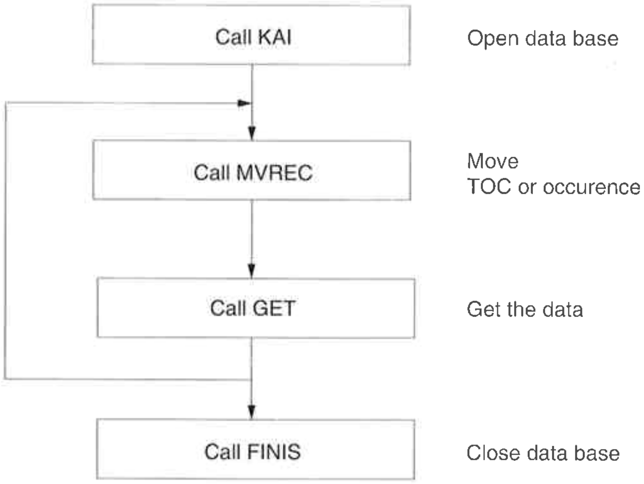

Fig. 5 The method in the program to read the data

case, precise calculated delay and delay rate are necessary. The calculated delay is required within an accuracy of less than 0.1 ns for the estimation within the precision of 0.1 ns. Many physical models and many calculations are needed to establish the calculated values. The partial differential of each parameter is also used to estimate the parameters. We repeatedly make a baseline analysis using the calculated delays and their differentials with regard to some parameters. It is effective for the calculated delays and their differentials to be sored in the data base.

Geometric delay is calculated using the baseline vector, the unit vector of source direction, and the rotation matrix which means the coordinate transformation due to precession, nutation, the diurnal earth rotation and the direction of the earth rotation and the rotation of the earth. The geometric delay is the delay within the solar system, and it should be transferred to the delay on the earth. This transformation represents the relativistic effects corresponding to the aberration and the movement of the stations during the delay. After the correction, the atmospheric delay is added to the calculated delay. We adopt CALC developed by GSFC (Goddard Space Flight Center)/NASA for the calculated delay. The physical models adopted will be presented in another paper in detail[2][8]-[10].

3.3 Parameter Estimation Software (SOLVE)

The "SOLVE" software makes an analysis by using observation data (O) and the calculated delays (C) using the least square method[4][11]. The estimated parameters are obtained so that $S = \sum_{i=1}^{N} [(O_i - C_i - \Delta C_i)^2/\sigma_i^2]/\sum_{i=1}^{N}(1/\sigma_i^2)$ becomes the minimum, where S is a weighted root mean square of the residuals, ΔC_i represents the correction based on the estimated parameters and σ_i is the error of each data. The estimated parameters can easily be selected from the computer display. The available estimated parameters are mainly clock offset and rate, atmospheric zenith path delay, station positions (X, Y, Z), earth rotation parameters, source positions and nutation. The clock parameters and atmospheric delay are estimated between epochs that are selected at the discretion of the analyst.

Initially, the ambiguity caused by the bandwidth synthesis should be eliminated from the observation delay. Next, data of poor or dubious quality are marked, and then the parameters and errors are estimated to minimize the value of S. The residual of $(O_i - C_i - \Delta C_i)$ is plotted, and an

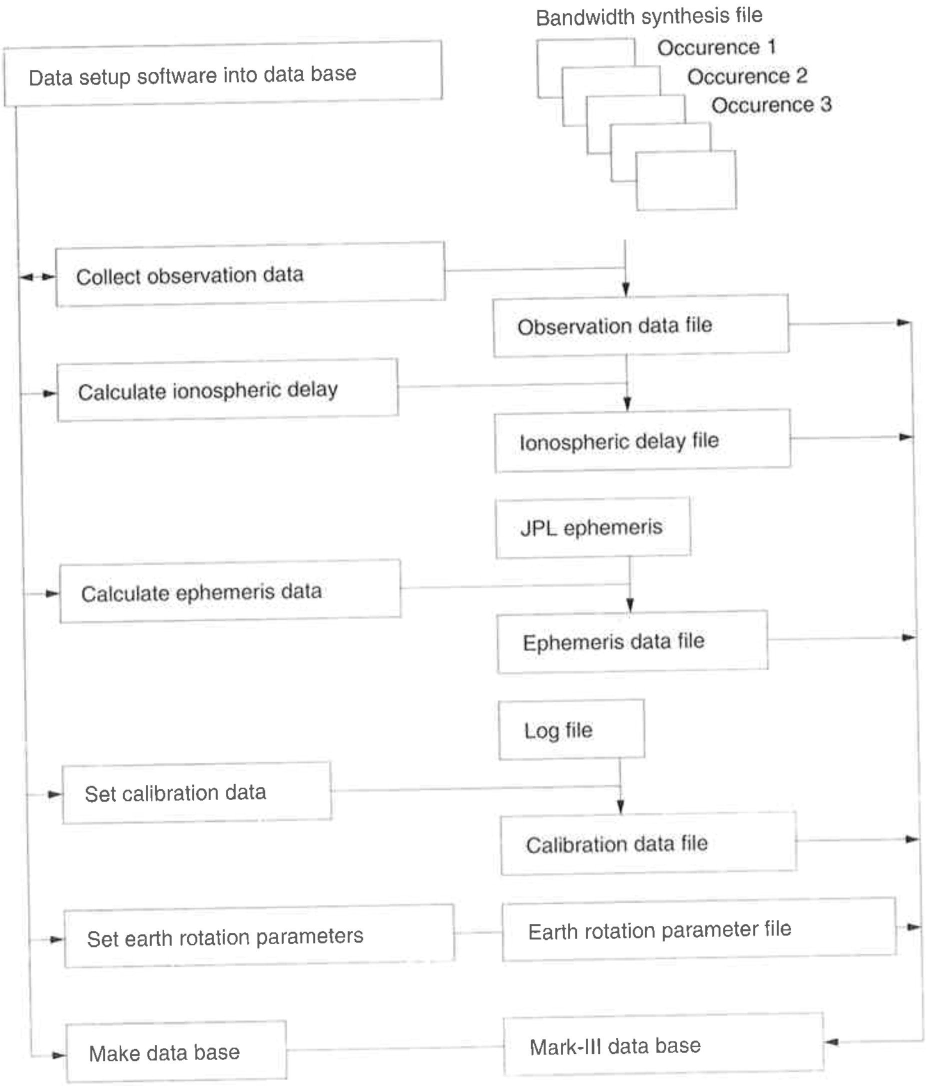

Fig. 6 The algorithm of the data setup software

assessment is made whether any systematic error exists. When the residual becomes the random distribution, the analysis is complete. Finally, re-weighting is conducted. The observation error is calculated from the SNR. The data of strong sources and the baseline between large antennas are more heavily weighted in the analysis since their observation errors are minute. However, the residuals of $(O_i - C_i - \Delta C_i)$ regarding sources are almost the same in standard analysis since analysis errors and errors caused by atmospheric scintillation should be added to observation errors in additional to the noise error by SNR[13]. It is difficult to estimate an analysis error. In re-weighting, the formal error is added to the observation error for each baseline as an analysis error. Formal error is obtained by making the chi square equal 1 for each baseline.

4. Data Base Setup Software

The data base setup software inputs various data for VLBI into the data base. We originally developed this software. Figure 6 shows the algorithm for setting up the data into the data base.

The bandwidth synthesis software produces the files including the observation delay and delay rate for every occurrence (on each baseline for each observation). First, we collect the data from these files and produce a new observation data file whose record includes the necessary observation data for each occurrence at S/X bands. The data is used as the latest results of the bandwidth synthesis.

Secondly, we calculated the ionospheric delay and delay rate corrections on both S and X bands using the observation delay and delay rate on S and X bands as follows;

$$\Delta\tau_{ion}(X) = \frac{f_s^2}{f_s^2 - f_x^2}(\tau_x - \tau_s)$$
$$\Delta\tau_{ion}(S) = \frac{f_x^2}{f_s^2 - f_x^2}(\tau_x - \tau_s) \qquad \cdots\cdots (2)$$

these ionospheric corrections are used for the input into the ionospheric correction data file.

Thirdly, we calculate the ephemeris data such as the velocity of the earth barycenter at the solar barycentric coordinate and the distances between the earth, the moon and the sun using the JPL ephemeris "DE200/LE200." The data of the original JPL ephemeris are the coefficients of Tschebyscheff polynomial expressions which are fitted to the positions of the 9 planets, the moon and sun every 32 days. We need to calculate the ephemeris data for each occurrence and we input these data into the new ephemeris data file. The calculation method used to obtain the positions, velocities and acceleration for each observation was described in detail in the review of CRL[(7)].

Fourthly, the temperature, the atmospheric pressure, the humidity and the cable delay calibration are included in the log file of each station. The data are retrieved from the log file, and the interpolated data is calculated for each observation. The interpolated data is input into the new calibration data file for each station.

Fifthly, we prepare for the precise earth rotation parameters such as UT1 and wobble, and we set up the earth rotation parameters into the earth rotation parameter file. These constant interval data such as the 5 day data of IRIS (International Rotation Interferometry Surveying) are used in the software for the calculated delay to be interpolated for each observation.

Finally, we collect the data in the observation data file, the ionospheric correction file, the ephemeris data file, the new calibration data file and the earth rotation parameter file, and we set up all the data into the data base.

5. Data Converting System From K-3 to K-4 Tapes

A part of the experiments in the Western Pacific VLBI network project were conducted in cooperation with the CDP experiments. In that situation, NASA produced the correlation processing for MARK-III, and we had to prepare the observation tape for K-3 (MARK-III) type. In our observations, we observed by K-4 recorder at MARCUS station and CDP requested the MARCUS data. Therefore, we developed the data converting software from K-4 tape to MARK-III tape. Apart from the experiments of the Western Pacific VLBI network project, this software is used for many experiments which need the conversion from K-4 tape to MARK-III tape for both high and low density.

6. New Parameter Estimation Software

We developed the new parameter estimation software for the NEC computer system (ACOS) and HP work station (Apollo). This estimation method is almost the same as the estimation software (SOLVE). This software has the correction file. We conducted experiments with the 3 m antenna on Daito island. This antenna had only X band receiver, thus we should correct the ionospheric delay obtained by the separate ionospheric measurement system (TEC meter) using GPS. This new estimation software can correct the delay and delay rate using data from the correction file. This file also includes a bad data flag. Normally, the bad data is automatically determined by the SNR (Signal-to-Noise Ratio) and the quality code in the bandwidth synthesis, and it is also manually judged by the residuals after the baseline analysis.

7. Conclusion

We established a new data processing system and a new analysis system for the Western Pacific VLBI network project, and they are used for any VLBI experiments conducted by CRL. The computer system HP1000-A900 and personal computer HP9000-330 system were introduced. K-4 VLBI acquisition terminal was used in this project. We had develop the data processing software for the personal computer HP9000-330. This software is the first software which can make a correlation processing using K-4 recorders, and it is the only software system that can facilitate a mixing correlation processing between K-3 (MARK-III) and K-4 recorders.

The analysis software such as software to calculate the delay and delay rate, the parameter estimation software and the data base handler are basically the MARK-III type software developed by GSFC/NASA for the HP1000-A900 computer system. We developed the data base setup software by ourselves. These software systems were used not only in the Western Pacific VLBI network project but also for other experiments. Furthermore, the software used for converting from K-4 tape to MARK-III tape was also developed and it was used for many experiments.

References

(1) A.R. Whitney, "Precision geodesy and astrometry via very long baseline interferometry," Ph.D. Themis, M. I. T., 1974.

(2) C. Ma, "Very long baseline inteferometry applied to polar motion, relativity and geodesy," NASA Technical Memorandum 79582, 1978.

(3) H. Kunimor and S. Hama, "III-4 Cross correlation software (KROSS)," The Review of the Radio Research Laboratories, **30**, pp. 185–198, November 1984 (in Japanese).

(4) Y. Takahashi, S. Hama, and T. Kondo, "III-5 K-3 software system for VLBI and new correlation processing software for K-4 recording system," The Special Issue of the J. Commun. Res. Lab. Results of VLBI Experiments at the Communications Research Laboratory, **38**, 3, 1991.

(5) T. Kondo and H. Kunimori, "III-5 Bandwidth synthesis software (KOMB)," The Review of the Radio Research Laboratories, **30**, pp. 199–216, November 1984 (in Japanese).

(6) K. Koike and T. Yoshino, "IV-4 Data base handler utility," The Review of the Radio Research Laboratories, **30**, pp. 247–262, November 1984, (in Japanese).

(7) Y. Takahashi and T. Kondo, "IV-3 Data base setup software (KASET)," The Review of the Radio Research Laboratories, **30**, pp. 237–246, November 1984 (in Japanese).

(8) Y. Takahashi, K. Koike, T. Yoshino, and S. Manabe, "An analysis of the baseline determination between Japan and U.S. stations by using the VLBI data in the system-level experiments," The Journal of the Radio Research Laboratory, **32**, 136, pp. 99–121, July 1985.

(9) Y. Takahashi and S. Manabe, "IV-5 A priori calculation software (KAPRI)," The Review of the Radio Research Laboratories, **30**, pp. 263–293, November 1984 (in Japanese).

(10) S. Manabe, Y. Takahashi, H. Hanada, T. Ishikawa, M. Fujishita, T. Sato, K. Koike, and T. Yoshino, "Physical models adopted in KAPRI," The Publication of the International Latitude Observatory of Mizusawa, **18**, 2 pp. 93–104, 1984.

(11) F. Takahashi and T. Yoshino, "IV-6 Parameters estimation software (KLEAR)," The Review of the Radio Research Laboratories, **30**, pp. 295–309, November 1984 (in Japanese).

(12) S. Hama and H. Kiuchi, "III-3 K-3 and K-4 VLBI data recorders," The Special Issue of the J. Commun. Res. Lab., Results of VLBI Experiments at the Communications Research Laboratory, **38**, 3, 1991.

(13) Y. Takahashi, "The estimation of the error in VLBI and position error," J. of the Geodetic Society of Japan, **40**, 4, 1994.