Atef A. Elassal**
521 National Center
U.S. Geological Survey
Reston, VA 22092

Managing Mapping Data Using Commercial Data Base Management Software

The utility of commercial data base management software in a mapping application is explored.

Abstract: *Electronic computers are involved in almost every aspect of the map-making process. This involvement has become so thorough that it is practically impossible to find a recently developed process or device in the mapping field which does not employ digital processing in some form or another. This trend, which has been evolving over two decades, is accelerated by the significant improvements in capability, reliability, and cost-effectiveness of electronic devices.*

Computerized mapping processes and devices share a common need for machine readable data. Integrating groups of these components into automated mapping systems requires careful planning for data flow amongst them. Exploring the utility of commercial data base management software to assist in this task is the subject of this paper.

Background

Two years ago, the United States Geological Survey (USGS) established a research project to explore the feasibility of a software system which could provide least-squares computational support to all adjustment programs in the allied fields of photogrammetry, geodesy, and surveying. This effort was culminated in the production of a prototype software system named Generalized Adjustment by Least Squares (GALS). The implementation of GALS achieved two primary objectives: (a) standardization of least-square computations for all network adjustment problems, and (b) creation of an environment in which highly complex network adjustment software can be brought into operation in a very timely fashion and with a great deal of confidence in the accuracy and integrity of the software.

Successful testing and evaluations of the prototype of GALS resulted in a decision to proceed with a production-oriented version. This version is scheduled for completion during the third quarter of 1983.

Having achieved a prototype for a least-squares module, planning work was initiated on the definition of an experimental data processing environment within which this module can function in relationship to the components of a typical network adjustment process. This environment was also necessary in order to test and verify GALS itself. Because the role of data management in such an environment is crucial, it was decided to delegate this role to a generalized data base management system (DBMS). The selected DBMS was to be representative of the state-of-the-art in data base management technology, providing a realistic first-hand experience with the utility of this technology in a typical mapping subsystem.

Data Base Management System

The DBMS selected for the experiment is the Model 204 software product of Computer Corporation of America.* Model 204 is a generalized data

** Presently employed by NOAA, Department of Commerce.

* Any use of trade names and trademarks in this publication is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.

Photogrammetric Engineering and Remote Sensing,
Vol. 51, No. 8, August 1985, pp. 1133-1136.

0099-1112/85/5108-1133$02.25/0

base management system designed to support users of operating systems on IBM computers, or compatible systems. The Geological Survey executes Model 204 on the Amdahl model V7 computer (using an IBM operating system) at the National Center in Reston, Virginia.

Model 204 provides rapid access to on-line files and features a user language with considerable selective power. The language is easy to learn and use, in addition to being capable of composing detailed or complex inquiries or reports. The system also supports a host language interface, thus allowing communication with application programs written in COBOL, PL/1, and FORTRAN.

Model 204 works with collections of files. Files are collections of related records and records are collections of related fields. Fields can be variable-length and variable-format. Model 204 operates by manipulating a set of inverted file indices. Each record has a unique internal number that the user is not normally aware of. Model 204 uses the inverted file to obtain record numbers in accordance with requests. In this manner, data themselves are not handled unless the user specifically requests information contained in one or more records. This mode of operation makes Model 204 very efficient in responding to unusually complicated queries.

One unusual characteristic of Model 204 is the fact that it supports access to multiple physically independent files. This facility allows cross-referencing of data among several physically distinct, separately maintained, yet logically related files.

Operating Environment

The environment chosen to experiment with Model 204 is schematically depicted in Figure 1. Specifically, photogrammetric analytical aerotriangulation was selected as the subject of the experiment. Within the data base are the related parameters and statistics, for example, camera systems, image orientations, object point coordinates, and image point coordinates. The data base also contains information which defines the data structure of these parameters and their associated statistics.

Figure 1 shows four major subsystems organized around the Model 204 DBMS. Data requirements for each subsystem were assumed by design to be different. In order to accommodate this assumption, special two-way interfaces were to be built in order to transform data between individual subsystems and Model 204 data definitions. Under this design philosophy, Model 204 would perform retrievals of selected portions of the data base for use by the attached subsystems. Results of execution of any subsystem are then used, through Model 204, to update the data base contents. Because the subsystems are totally independent, subsystems can be arbitrarily added to the environment. Furthermore, new functions can be created within each subsystem. While this design scheme results in some overhead due to the large number of interfaces, it is justified by the need to keep the experimental system structurally as flexible as possible.

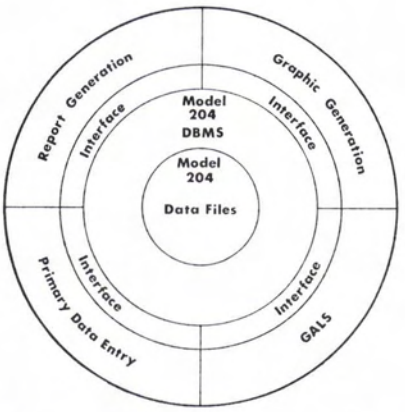

Fig. 1. Data processing environment for experimentation with Model 204.

The four subsystems which interact with Model 204 are

- *Primary Data Entry Subsystem.* The initial loading of the data base is expected to come from a multitude of sources. The nature of these sources will vary, as will the techniques for data capture, verification, and reporting. The types of equipment and organizational standards will dictate the functional requirements for the components of this subsystem. The all-important function of data verification prior to entry into the data base will reside within this subsystem in any production-oriented system. In the current experimental system, the Primary Data Entry subsystem is simulated by employing a text editor. Here the text editor is used to create all primary input files to Model 204 in the required formats.
- GALS *Subsystem.* The least-squares adjustment of the photogrammetric block or sub-block supplied through Model 204 is performed by the GALS Subsystem. GALS requires three data sets for its operations, and these three data sets can be automatically produced by Model 204. The first is a network definition data set which informs GALS of the structure of the network that will be subjected to least-squares adjustment (photogrammetric block, geodetic triangulation, leveling network, etc.). The second data set contains all basic observations on the network. In the case of photogrammetric aerotriangulation, the basic observations are photo-image coordinates and associated statistics. The third data set contains initial estimates for unknown parameters and for associated statistical data. For this experiment, photo position and attitude, camera interior orientations, and spatial coordinates of object points are the unknown parameters. The results of an adjustment are reflected in the form of numerous transformations of the content of these three data sets. Therefore, they can be sub-

sequently used to update the contents of the Model 204 data base.

- *Graphic Generation Subsystem.* Graphic portrayals of adjustment results are produced by this Subsystem. In a mature and fully developed system, graphics are expected to assume the primary role in the command and control of the system functions. Various plots showing related elements of the data base are planned. Figure 2 shows a CRT plot produced by the system for a selected image from a test photo block. The graph contains photo and camera system identifications, adjusted values and corrections for all elements of exterior orientation, and image points with plots of their coordinate residuals. Graphics are produced in a two-step process. In the first step, a portion of the data base is retrieved based on some selection criterion. During the second step, the retrieved data are used to implement a function which is capable of producing the desired graphics.
- *Report Generation Subsystem.* Reports of triangulation adjustment results and statistics are created by this Subsystem. Report generators can operate on all the contents of the data base or on a selected portion of it. Figure 3 shows an example of a statistical report of triangulation residuals for a three-by-four photo block retrieved from the data base. This type of report can be readily generated for any portion of the data base using the selective retrieval capabilities of Model 204. Several other types of reports are planned to report triangulation results in various formats.

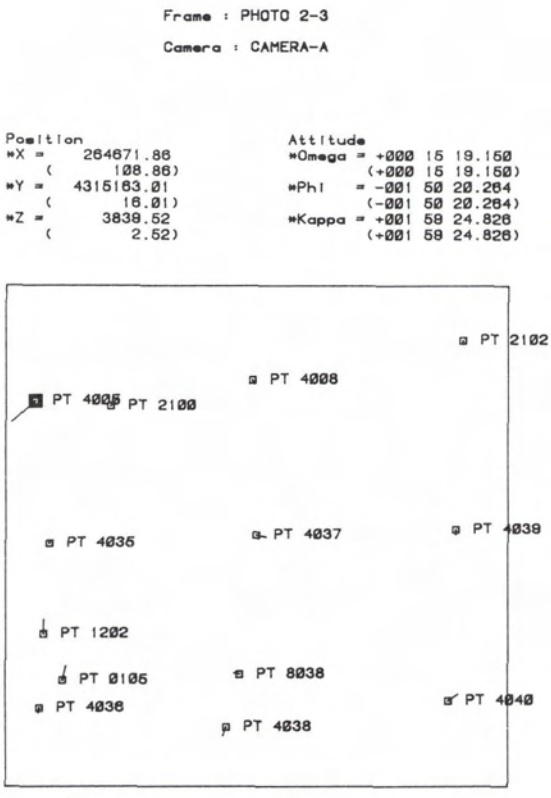

FIG. 2. Plot produced by the Graphics Generation Subsystem.

DATA BASE COMMUNICATIONS

Communications with Model 204 are achieved either through a self-contained user language in an interactive mode or through a host-language interface for a batch mode. A complicated query can be greatly facilitated by using the stored procedures facility of Model 204. Stored procedures are programmed, tested, and subsequently stored within the data base itself for later reference. A user can then invoke any of these stored procedures by their name only. For example, the information which is related to a specific photograph consists of data and statistics about the corresponding camera system, exterior orientation, image, and object points covered by the photograph. This information can be extracted from the data base by the procedure shown in Figure 4. Notice the English-like nature of the query language. A user invoking this procedure will simply type the name of the procedure (*Photo*).Model 204 will respond by displaying the sentence: ENTER.PHOTO.IDENTIFICATION., to which he would respond by typing the identification name of the desired photograph. The extracted data can then be relayed to any subsystem for further processing.

OBSERVATIONS

By the end of 1982, experimentation with Model 204 was still fairly new and therefore quite limited. There are, however, several remarks that are worth reporting which came about as a result of experience gained while programming various interfaces within the system: Model 204, like most commercially available DBMS, is designed and implemented with business applications in mind. This usually results in systems which are strongly biased toward optimized handling of character type data. Searches which are based on numerical values are often cumbersome and performed inefficiently. This must be considered when designing the data definition for the data base and when formulating retrieval procedures. Report-writing capabilities are available within Model 204's User Language, but facilities to perform computations in the course of report generation were found to be limited and difficult to implement. Presently, host language interfaces are being used to generate the necessary reports, and therefore overcome this problem. The efficiency with which Model 204 performs retrievals is very remarkable. Naturally, this efficiency is closely related to the design of the data base and the selection of the query fields.

PHOTO STATISTICS

I - POSITION :

Photo (PHOTO 3-1) has Max total correction = 1457.15
Max component correction : 1455.09 146.93 73.94
At photos : PHOTO 3-1 PHOTO 1-3 PHOTO 3-4

II - ATTITUDE :

Photo (PHOTO 2-4) has Max total correction = +003 21 58.418
Max component correction : +001 28 09.153 +001 50 20.205 +003 15 18.459
At photos : PHOTO 3-1 PHOTO 1-3 PHOTO 3-4

OBJECT STATISTICS

Point (PT 4053) has Max total correction = 2969.27
Max component correction : 1517.38 1449.98 2466.43
At photos : PT 4051 PT 4053 PT 4053

IMAGE STATISTICS

	Photo	Image Residuals (X,Y,R)			Image Identifications		
	PHOTO 1-1	8.8	11.0	14.0	PT 4005	PT 4003	PT 4005
	PHOTO 1-2	7.7	10.4	10.4	PT 4005	PT 8000	PT 8000
	PHOTO 1-3	4.6	3.9	5.7	PT 4005	PT 4006	PT 4006
	PHOTO 1-4	3.5	2.8	3.8	PT 4008	PT 4006	PT 4008
	PHOTO 2-1	6.1	9.8	11.6	PT 4036	PT 4036	PT 4036
	PHOTO 2-2	6.7	8.4	9.0	PT 4038	PT 4037	PT 4037
****	PHOTO 2-3	16.4	14.8	22.1	PT 4005	PT 4005	PT 4005
	PHOTO 2-4	8.8	6.8	9.6	PT 8038	PT 4037	PT 8038
	PHOTO 3-1	3.1	10.8	10.9	PT 4036	PT 4054	PT 4054
	PHOTO 3-2	5.8	9.3	9.3	PT 0106	PT 4038	PT 4038
	PHOTO 3-3	7.3	10.8	10.8	PT 8038	PT 2125	PT 2125
	PHOTO 3-4	7.6	11.7	11.7	PT 4042	PT 4056	PT 4056

Note: **** = PHOTO with Largest Vector Residuals

FIG. 3. Statistical report produced by the Report Generator Subsystem.

CONCLUSIONS

The United States Geological Survey has created a data processing environment within which state-of-the-art software for data base management, represented by the Model 204 software product, can be assessed in relation to its utility for mapping processes. Initial experience is very encouraging, especially with respect to the powerful retrieval capabilities of Model 204. While the present experimental environment is configured specifically for analytical photogrammetry problems, it is expected to yield valuable information about the utility of commercial DBMS software for mapping data.

```
PROCEDURE PHOTO
BEGIN
1.  FIND ALL RECORDS FOR WHICH
    RT = 4
    NM = ' 1??ENTER.PHOTO.IDENTIFICATION '
2.  PLACE RECORDS IN 1 ON LIST LOBS
3.  FOR EACH RECORD IN 1
     3.1 FOR EACH OCCURENCE OF NM
         3.1.1 FIND ALL RECORDS FOR WHICH
               RT = 5
               NM = VALUE IN 3.1
         3.1.2 PLACE RECORDS IN 3.1.1 ON LIST LPAR
END PROCEDURE
```

FIG. 4. Example of Model 204 procedure.

REFERENCES

Computer Corporation of America, 1979a. *User Language Manual, Model 204 Data Base Management System.*

Computer Corporation of America, 1979b. *Host Language Interface Manual, Model 204 Data Base Management System.*

Elassal, A. E., 1982. Generalized Adjustment by Least Squares (GALS), Proceedings of the Symposium on Mathematical Models, Accuracy Aspects, and Quality Control, *International Archives of Photogrammetry*, Volume 24-III.

(Received 30 April 1983; accepted 26 March 1985)