A MICROCOMPUTER-BASED INFORMATION RETRIEVAL SYSTEM FOR PROCESSING VERTICAL PROFILES OF AEROSOL CHARACTERISTICS

L.A. Gerasimova, M.V. Panchenko, S.A. Terpugova, and V.D. Teushchekov

Institute of Atmospheric Optics,
Siberian Branch of the Academy of Sciences of the USSR, Tomsk
Received March 14, 1990

The purpose and the principles of operation of a dialog information-retrieval system (IRS) and experience in using the system for retrieving and processing of required data from several hundreds of files, created as a result of an experiment performed on an airborne laboratory, are described.

The information-retrieval system is characterized by efficient use of the working memory, by the possibility of operating under conditions when the user possesses the objects to be retrieved before the retrieval-system is created and used, and by the fact that there are no restrictions on the structure and composition of the experimental data.

The described system can be easily adapted to any object region, where the collection of objects to be retrieved is represented in the form of separate files stored in the external memory of a computer.

The problem of studying the optical characteristics of aerosol is an important problem in atmospheric optics.

The strong variability of aerosol in time and space is caused by a complicated collection of meteorological, synoptic, seasonal, and geographical factors. To determine which of the factors is most important for the modeling of the optical properties of the atmosphere at different altitudes it is important to have statistically significant experimental data.

Often, however, the data are difficult to analyze because it is impossible to have on-line access to the sets data meeting the various requirements. In this connection it is important to use an information-retrieval system (IRS). Such a system makes it possible to find quickly and to put together statistical ensembles of data according to preselected criteria.

In this paper we describe a microcomputer-based version of such a system,[1] since microcomputers are used in airborne and mobile computing systems and they make it possible to perform real-time preliminary analysis of the results obtained.

A number of dialog IRS have been developed for mini- and microcomputers,[2–4] but often they are difficult to use because the carriers and software are not compatible, the working and long time memory are not used efficiently, a long time is required to process a large volume of information, and the user often processes the objects to be retrieved (specific files of data) before the IRS is constructed and used. The system described is free of most of these drawbacks: the required volume of working memory does not exceed 30 Kb, and no restrictions are imposed on the structure and composition of the experimental data.

The information-retrieval system was developed for processing data obtained by airborne sounding of the atmosphere, performed at the Institute of Atmospheric Optics of the Siberian Branch of the Academy of Sciences of the USSR in the period from 1986 to 1988. After preliminary processing the experimental data are stored in the external memories (EMs) of a microcomputer. A separate file includes the results of a single flight: the vertical profile of the aerosol scattering component for the wavelength $\lambda = 0.52$ μm, the relative and specific moisture content of the air, the temperature, and the number density of particles with size of $d \geq 0.4$ μm with a step of 100 or 50 m in altitude depending on the character of the vertical distribution of the aerosol. The height of the profile can be different, and varies from 1 to 5 km. Each such array is accompanied by additional information about the location and time of the experiment, the character of the pressure field and air mass, the size, type, and altitude of the cloud cover, and the presence of precipitation. Separate parameters may be missing in the files.

Operation sequence. Every separate realization is characterized by a set of parameters, whose values are either given explicitly in the data set or they can be obtained by means of very simple processing. They are divided into three classes — integer, real-valued, and textual — and they are organized in the form of tables (Tables I—III).

When the IRS is started up each table is displayed on a monitor in order to inform the user about the existing parameters and to make it easier to place restrictions on their values.

0235-6880/90/07 709-04 $02.00

TABLE I.

List of integer numerical parameters of the profile

Name of parametes	Value	
	min	max
Type of flight (1 — takeoof, 2 — landing)	1	2
Number	1	31
Month	1	12
Year	86	88
Moisture content at the ground	10	100
Cloud amount	0	10

TABLE II.

List of real-valued numerical parameters of the profile

Name of parameters	Value	
	min	max
Time of day	0.00	24.00
Height of profile	0.1	5.5
Scattering coefficient at the ground	1.00E-04	1.0
Temperature at the ground	-60.0	50.0
Average temperature gradient	-3.0	3.0
Lower boundary of the cloud cover	0.0	5.5

TABLE III.

List of textual parameters of the profile

Name of Parameter	Name of file	Length of the field
Region of flight	RAJON.TXT	15
Type of air mass	MASSA.TXT	12
Weather conditions	POGODA.TXT	8
Cloud type	OBLAKA.TXT	3
Pressure field	DAWL.TXT	11
Precipitation	OSADKI.TXT	5

For a numerical parameter the required minimum and maximum values are introduced by referring to its order number.

In setting the conditions on the textual parameters the complete list of textual parameters is displayed on the monitor, and the user indicates which parameter must be restricted. Then, based on the list of possible values of the selected parameter, the user inputs the numbers of the parameters which are admissible for the files sought. After the required information has been input an automatic search is made for the required experimental realizations and a catalog of these restrictions is formed. This catalogue is then used for further processing.

The operating mode of the IRS. The search for the required experimental data can be performed in two modes: 1) direct and 2) record.

In the direct search files whose names are listed in a separate catalog are analyzed successively and they are checked in order to determine whether or not they satisfy the restrictions that were input. This search mode is simple to implement on a microcomputer and is convenient to use when the number and size of the files studied are small. As the number of searched files increases, i.e., the number of operations of reading of data from the external memory increases, the average operating time of the IRS increases substantially and can become unacceptable. A second operating mode of the IRS is provided for working under these conditions.

In the record mode the IRS operates in two stages. At first one-time processing of all experimental data is performed in the direct-search mode and one or several lists of their records are formed. Each record contains the name of the corresponding file in the external memory and the values of the parameters of the experiment which have already been determined. At the second stage the user inputs into the IRS the necessary conditions, and the search is performed based on a scan of the indicated list of records.

In this mode it is possible to sort data according to several indicators and the further search can be performed not through the entire archive but rather only within subarchives. This significantly reduces the operating time of the IRS.

The structure and composition of the software of the IRS. The programs used in the IRS were written in the FORTRAN language. The structure of the system is shown in Fig. 1.

Most of the subprograms forming the nucleus of the system implement different numerical processes in the operation of the system that are not connected with a specific subject region of the investigations. However most of them are tied to a specific operational system RAFOS (RT-11).[5]

Independent subprograms perform the following:

— input of restrictions on the parameters of the experiment in the dialog mode;

— conversion of the values of the textual parameters into integers for efficient storage in the record of the experiment;

— initialization of a record file before it is filled; and,

— recording of the operation of the IRS.

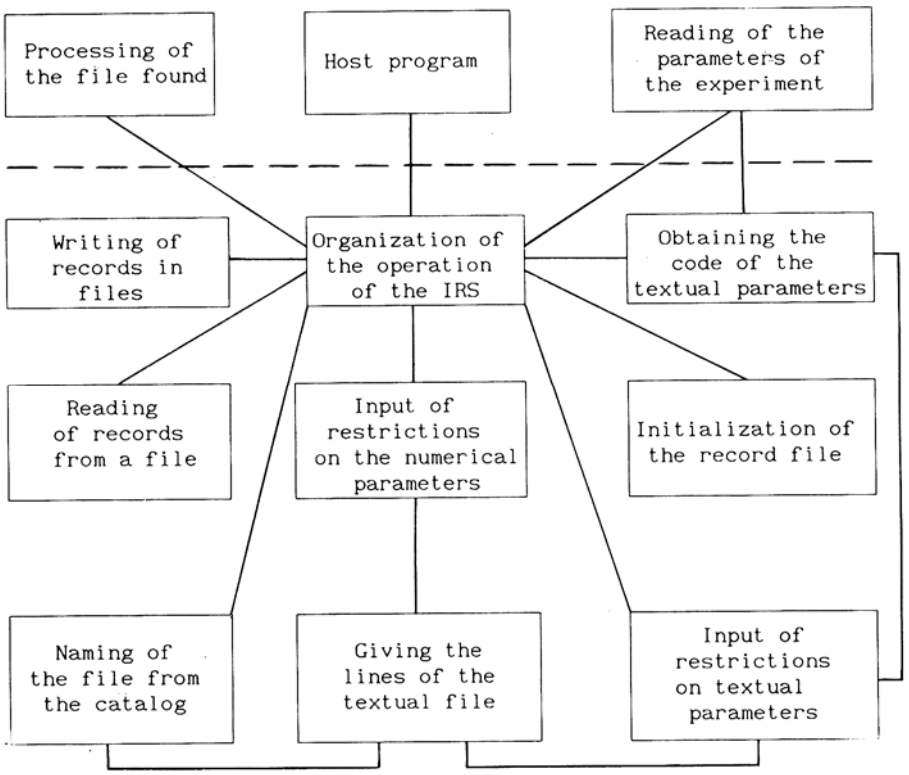

FIG. 1. Structure of the software used in the IRS.

The program units that depend on the characteristics of a specific experiment are the host program of the system and two subprograms. One subprogram obtains, for checking purposes, from a file in the external memory the actual values of the parameters of the next experiment under consideration. It is created taking into account the data present in the recording system and the sequences of parameters recorded in Tables I—III. The second subprogram performs the operations required by the user in finding an experimental data file whose parameters satisfy all restrictions. The working program of the IRS is constructed in the form of an overlay program, which makes it possible for the program to operate in the original microcomputer with on the order of 60 experimental parameters.

Application of the IRS. The software described above was first testfed for retrieving data from an archive containing more than 500 files. The IRS required on the average 40 min to search for the required files in the direct-search mode and 3 min in the record mode.

Experience in using the IRS has shown that it is simple and convenient to use and the software operates efficiently with the base model of the microcomputer.

The IRS can be easily adapted to any object region, where the collection of objects to be retrieved is represented in the form of separate files in the external memory of a computer.

REFERENCES

1. V.D. Teushchekov, in: *Abstracts of Reports at the 10th All-Union Symposium on Laser and Acoustic of Sounding of the Atmosphere,* Tomsk Scientific Center, Siberian Branch of the Academy of Sciences of the USSR, Tomsk (1989), pp. 284–288.
2. M. Meldman, D. McLeod, R. Pellicore, and M. Squire, *RISS: Relation Database Control System: for Small Computers* [in Russian), (Finansy i Statistika, Moscow, 1982).
3. V.P. Semik, G.P. Ostapenko, A.L. Friedman, and V.E. Gorskii, *Dialogue Multiterminal System for SM Computers* [in Russian], (Finansy i Statistika, Moscow, 1983).
4. Yu.I. Protasov and V.N. Sitnikov, *The Optical Properties of the Earth's Atmosphere* [in Russian], Tomsk Scientific Center, Siberian Branch of the Academy of Sciences of the USSR, Tomsk (1988), pp. 102–105.
5. *Software for the SM Computers. RAFOS Operating System with Separation of Functions. Introduction to the System. Description of Applications* [in Russian] (1980), Vol. 1, Pt. 1.