USE OF KNOWLEDGE–ENGINEERING METHODS FOR SOLVING THE PROBLEMS OF CORRECTION OF REMOTE MEASUREMENTS FOR THE EFFECT OF THE ATMOSPHERE

E.I. Pospelova

Institute of Atmospheric Optics,
Siberian Branch of the Academy of Sciences of the USSR, Tomsk
Received April 30, 1991

The works devoted to the use of the knowledge–engineering methods for processing of the remote sensing data are reviewed. An expert system which takes into account the effect of the atmosphere on the propagation of thermal radiation in the atmosphere–underlying surface system is described. It is shown how the salient features in the subject area influence the formation of a fundamental system principle which consists in the fact that the user should participate in choosing the relevant data and research methods. The range of applicability of the described system prototypes are pointed out.

INTRODUCTION

The works devoted to the investigation of the effect of the atmosphere on the radiation transfer are undoubtedly of practical importance. These investigations are based on analyzing the radiation characteristics as functions of the optical–meteorological state of the atmosphere and exhibit the following salient features:

– a large amount of diversified information (spectroscopic, meteorological, etc.), which must be used virtually at every investigation stage,

– a complicated description of the physical mechanisms associated with the radiation transformations, and

– a lack of a common viewpoint on the nature of some atmospheric processes among the investigators.

The last facts make it impossible for the specialists to formalize their knowledge about these processes, their intuitive guesses, and assumptions based on their own experience. For this reason, the conventional automation methods, e.g., development of computational programs, program packages, etc., have a limited efficiency in practice. But it is well known that in many cases one can succeed in solving analogous problems by employing advances in the artificial intelligence (AI). Thus the procedures for systematization, structurization, and reduction of relevant data (i.e., the data which refer to the subject area) which are necessary for solving any information–abundant problem, can be implemented in a comparatively easy and smooth way with the help of heuristic algorithms. If data processing requires taking into account the great number of the autonomous parameters, then the role of one or another factor may be most efficiently evaluated by the investigator himself as a user of a certain intellectual system. Concepts of the means for processing the knowledges have recently started to appear, expert systems (ES), decision support systems, and so on, have been created, which persuasively prove that a lot of problems of remote sensing (RS) can be solved with the help of the systems with the AI.

This paper describes an ES intended for examining the effect of the atmosphere on the thermal radiative transfer in the atmosphere–underlying surface system. But prior to discussing this system let us briefly review the investigations of the knowledge engineering for the RS problems.

EXPERT SYSTEMS FOR REMOTE–SENSING PROBLEMS

At present great efforts are undertaken to create different systems with the AI, which could be used for solving the RS problems. Among these works we can identify the investigations aimed at studying the role of the intellectual methods in systems of processing the satellite information.

Considering the conditionality of the AI in the satellite automation systems, the NASA investigators conclude that the methods and principles of the AI have many applications to the subject at hand because, in their opinion, the latter exhibits the following features[1]:

– a quite definite subject of investigations whose aim is clear and unambiguous;

– the existence of skilled specialists, which are ready to give their knowledge to an electronic expert;

– a frequent recurrence and labor–intensive nature of the investigation process;

– an inapplicability or large expense of the conventional software.

For this reason, we can identify three main fields of applicability of the ES's, where such systems will ensure the greatest advantages, namely, checking of the onboard devices, programming and control of the onboard resources, and reduction of the investigator's efforts in the interactive regime of operation.

The ES with different hierarchy levels have been proposed, e.g., information control system, functional satellite subsystems, and operating control system. The model of taking decisions, which has been developed in the NASA for carrying out commercial satellite operations, chooses and estimates an optimal employment of the AL for automatization of a specific spacecraft.[2] A concept for constructing the systems of the AI, which is being created in the All–Union Scientific–Research Institute of Agricultural Meteorology, pays a considerable attention to find a possible employment of the ES for processing the RS data. Their TOPIKA system involves a subsystem for supporting a software development and methods

0235-6880/91/08 625-05 $02.00

of the AI used for interpreting the RS data in agricultural meteorology.[3]

At present possibilities for integrating the conventional technologies of processing the RS data and the AI methods are intensively examined. The state–of–the–art level of the development of the recording devices predetermines an appreciable complication of the structure of the processed information, while an increase in the knowledge about the subject and an extension of the problems being solved leads to an increasing number of principles and facts which are involved in the process of taking a decision. These circumstances are spurring the attempts to combine the technologies of taking a decision, which are typical of the ES, and the methods of efficient processing the large amounts of data, belonging to the databases (DB).

The ES operating onboard a satellite may be directly linked with the systems of data acquisition. In this case, the acquired data are archived in a complex DB, which already contains conventional geographical data. Such an ES can preliminarily process the data right onboard (cloudiness classification, atmospheric correction, image classification, and so on), while employment of the additional geographical information makes it possible to improve the accuracy and reliability of the results.[4] In this connection, a methodology, which allows us to solve the problem of processing the large arrays of expert knowledge in addition to the classical data, combining the technologies of constructing the ES and the DB, is of interest. The problems and methods for creating an interface between the ES and the DB have been discussed in Ref. 5.

When carrying out experiments on integrating the geographical information about the forest areas in the British Columbia province (Canada) and the images obtained by the systems of analyzing the digital images of LANDSAT (LDIAS) and SPOT satellites, some scientific and technical problems have arisen, and in order to solve them it was necessary to introduce symbolic descriptions and expert estimates. Thus the hierarchical ES was constructed called the Analytical Advisor, which controls all devices and computers of the LDIAS system as well as all problems being solved in the system. The ES incorporates the software program packages, each of them model the work of an individual expert. These programs form a hierarchy with the help of a switching system with a common memory. This switching system is used to link the different levels. Now the administration of this province has a timely and exact information about the regeneration of the forest resources in the entire region.

The MOPA has been developed for job scheduling of the UARS artificial earth satellites. A laborious process of making up the programs for daily satellite operations consists mainly in determining the operational regimes of the ten main devices and in their coordination. The device operation has a cyclic character, only some of its characteristics are changed every day; and this makes it convenient to formalize the knowledge about the device performance in the form of frames. The device actions are arranged in the form of a hierarchical scheme, whose upper levels perform more general operations and the lower levels – elementary minute operations. At first, the needed level is chosen. This is followed by specifying the direct operations. In the process, a consecutive conversion of the symbolic description of the device performance into specific command sets is performed.[8]

All intellectual systems, which are being developed now for assisting in solving the RS problems, belong to the class of the so–called special–purpose systems. These systems are characterized by a fixed collection of problems being solved, which is predetermined in the course of designing such a system. Now, in order to be employed they require only that the knowledge base be filled with the facts and rules, which are appropriate for the chosen application. Alongside with the special–purpose systems, there are general–purpose systems, deliberately designed for processing data from a quite wide spectrum of the subjects and, by virtue of their versatility, exhibiting a redundancy (so–called shells of the ES). In contrast to the general–purpose systems, the special–purpose systems provide a maximum rational employment of all computational resources.

EXPERT SYSTEM OF CORRECTION OF THE RS DATA FOR THE EFFECT OF THE ATMOSPHERE

When developing complex intellectual systems, a technology has been worked out, which is based on the so–called evolution method. In short this method can be characterized as follows. Quite elementary prototype system is created, in which the basic properties of the future ES have been implemented. The prototype performance is comprehensively tested and it is elucidated at this stage, how efficient is the chosen method of knowledge representation, the requirements to the system are clarified in order to extract an additional knowledge and to improve the mechanism of interaction with the user, etc.

Such a method allows the designer to improve constantly his creation obtaining a working program product at each stage.

Let us now consider the ES, which is currently developed based on the principles of the evolution approach and is intended for studying the effect of the atmosphere on the radiation transfer in the atmosphere–underlying surface system. The first version of this system was described in Ref. 10. It was called an information–interactive system (IIS) and was created for analyzing the principles of formation and employment of the software as well as of the means for interactive mode of operation between the system and a user. The functional loading of the IIS was limited by calculations of the optical parameters of the atmosphere within the IR transparency windows. In order to obtain a more detailed information about the algorithms and techniques, which have been incorporated in this IIS, the reader is referred to Refs. 10 and 11. In this paper I present the second version of the system, which has been called an expert system of correction of the data (ESCD) of remote sensing for the effect of the atmosphere. I briefly describe the algorithms and the functions of the system, which are of interest for an investigator–physicist and focus my attenuation mostly on the presentation of the principles and examples of intellect of the ESCD.

In order to perform the correction of the remote measurements with the help of the ESCD, one can retrieve the vertical profiles of the attenuation coefficients, transmission function, optical thickness, and intensity of the outgoing thermal radiation. All these parameters can be retrieved with an arbitrary spectral resolution in any frequency range, under any meteorological conditions, and for any paths of the radiation propagation. The radiation intensity in the atmosphere–underlying surface system is given by the sum

$$I = I_s + I_a + I_r ,$$

here the terms represent the intensity of radiation of the underlying surface and the atmosphere and of the reflected radiation, respectively. Their values in the frequency range $\Delta\nu = \nu_2 - \nu_1$ are given as follows:

$$I_s = \frac{1}{\Delta\nu}\int\limits_{\nu_1}^{\nu_2} d_\nu B(\nu, T_0) \exp\left(-\int\limits_0^L K(l, \nu)\, dl\right) d\nu ;$$

$$I_a = \frac{1}{\Delta\nu}\int\limits_{\nu_1}^{\nu_2}\int\limits_0^L B(\nu, T(l))\, K(l, \nu) \exp\left(-\int\limits_0^L K_t(l', \nu) dl'\right) dl d\nu ;$$

$$I_r = \frac{2}{\Delta\nu}\int\limits_{\nu_1}^{\nu_2}\left\{(1 - \delta_\nu) \exp\left(-\int\limits_0^L K_t(l, \nu)\, dl\right) \times\right.$$

$$\times \int\limits_0^L B(\nu, T(l))\, K(l, \nu) \times$$

$$\left.\times\left[\int\limits_0^1 \exp\left(-\frac{1}{\mu}\int\limits_0^L K(h', \nu)\, dh'\right) d\mu\right] dl\right\} d\nu,$$

where δ_ν is the emittance of the underlying surface and $B(\nu, T)$ is the Planck function given by the formula

$$B(\nu, T) = \frac{c_1 \nu^3}{\exp\left(\frac{c_2 \nu}{T} - 1\right)} ,$$

in which $c_1 = 0.1191 \cdot 10^{-4}$ and $c_2 = 1.43879$ are the first and second radiation constants, L is the length of a nonuniform path, T is the temperature, $\mu = \cos\theta$, θ is the viewing angle, $K(l, \nu)$ is the attenuation coefficient, and $K_t(l, \nu)$ is the total extinction coefficient. The transmission function is given by the relation

$$P_\nu = \frac{1}{\Delta\nu}\int\limits_{\nu_1}^{\nu_2} \exp\left(-\int\limits_0^L K(l, \nu)\, dl\right) d\nu .$$

The optical thickness $K_{Dm} = -\ln(P_m)$. The total attenuation coefficient is given by the sum

$$K_t(l, \nu) = K(l, \nu) + K_s ,$$

where K_s is the coefficient of molecular and aerosol scattering. In order to determine its molecular component the techniques described in Ref. 12 are used.

$$K(l, \nu) = K_a + K_m ,$$

the terms determine the absorption by aerosol and molecular components of the atmosphere. The aerosol absorption and scattering are taken into account in the form of the vertical profiles of the corresponding coefficients.[13] The absorption due to molecular components is given by the sum of the selective and continual absorptions[14]

$$K_m = K_{sel} + K_{con} .$$

In order to calculate the coefficients of selective and continual absorptions, the data on the fine structure of absorption spectra of the atmospheric gases are used. These are the quantitative characteristics of the spectral lines, namely, the position of the center, intensity, half–width, energy of the ground state, and information about the isotopes. They are arranged in the catalog of the initial information. The continual absorption is taken into account based on the formulas for the generalized spectral line shape.[15] In order to avoid lengthy computations from these formulas during the ESCD operation, a catalog of the continual absorption is employed. This catalog incorporates preliminary calculated coefficients of the continual absorption due to the water vapour in the cases of line self–broadening and broadening by nitrogen as well as the coefficients of interpolating cubic splines. The data are structurized using a uniform frequency and temperature grid. The physical principles and expedience for creating such an information base are given in Ref. 11. It is also noted there that the catalog is quite universal and can be used for solving any problem, in which it is necessary to take into account the continual absorption due to water vapour.

As an example of taking an expert decision let us consider, how the ESCD with direct participation of the user chooses the algorithm and the relevant data for determining the coefficient of selective absorption. In general the formula for calculating this coefficient can be written as follows:

$$K_{sel} = \sum_{i=1}^{N}\sum_{j=1}^{n_j} f_j\, q_i ,$$

where q_i is the concentration of the ith gas, $i = \overline{1, N}$; n_i is the number of spectral lines of the ith gas, and f_j is the spectral line shape.

The following methods for determining the spectral line shape are implemented in the ESCD: the Lorentz method, the complete Lorentz method, the Voigt method, the Naumov–Zhevakin method, and the generalized line shape method. There are some considerations concerning the admissibility of employing one or another shape in one or another situation in the references (see, e.g., Refs. 12 and 14). In addition, various investigators adhere to different opinions on this problem. It is pertinent to note that it is this ambiguity which has determinated the basic principle of the ESCD operation, i.e., solving a problem under control of a user. Therefore, in the interactive mode the user is always informed, how his problem is being solved and what initial data and algorithms are being used. The ESCD gives him an exhaustive information about every situation arising in the process of the operation, about the state of the object being studied, about the availability of the relevant data, and about the computational methods and the possible modifications of the investigation course. There are two levels of the information:

– operational data, which are intended for a skilled specialist trained in the interactive mode of operation with computers. Here, the situation is briefly described and it is suggested to take a decision and

– more complete data, which the user may retrieve after pressing a special key in the case, in which the available data are insufficient, and a more detailed critical analysis of the problem is required.

To complete, the ESCD gives recommendations, which of the available algorithms is more preferable for calculation, so that it is up to the user, if he should accept a recommendation of the system or choose another algorithm from those which are in the menu. At each stage of the operation the user may return to a preceding step and change a previous decision.

Let us now consider the problem of retrieving vertical profiles of temperature, pressure, and concentrations of the atmospheric gases. The ESCD allows the user to choose the source of such an information by himself, either the meteorological models, averaged over seasons and regions, or real measurements, performed at meteorological stations. The data for meteorological models were taken from different published works. The information of such a kind is quite often used and is well known to the specialists. As regards the DB of meteorological parameters, it is pertinent to consider the structure and peculiarities of its employment in ample detail. Geographical coordinates and measurement time are the descriptor keys, i.e., the search keys for this base. The user by himself makes an inquiry according to the above–indicated keys. In so doing in the menu regime, he may choose either the coordinates of one of the meteorological stations, recorded in the DB, or an arbitrary point on the globe. In the latter case, the profiles of temperature, pressure, and concentrations of the gaseous components are chosen in the database from the meteorological stations closest to this point, and the method of optimal interpolation[16] is used for spatial correction of these data. The profiles, incorporated in the DB, by virtue of errors in radiosonde operations may have missing values. In order to retrieve them, the ESCD uses the barometric formula or the methods described in Refs. 17 and 18. Each of these methods provides higher or lower accuracy depending on the state of the initial profile and the availability of the correlation information. For this reason, the user himself, just similar to the case of determinating the spectral line shape, chooses the retrieval method, being supported by the ESCD recommendations.

The ESCD constantly supervises the user, pointing out and explaining his errors, if any. In the process, it is checked, if the functional keys have been correctly pressed and if the values of the input numerical parameters are admissible.

The conceptual knowledge about the algorithms, the situations which have arisen in the course of the system operation, etc., is expressed in terms of the subject area on the natural language. It consists of verbal semantic constructions (i.e., semantic textual expressions) with free variables, whose values are filled with specific contents depending on the calculated results and analyzed situations, after which they are suggested to the user in the form of either a menu at an alternative time moment or an advice, if he is mistaken. This knowledge is arranged in situation ensembles, and this makes it possible to easily modify the functional orientation of the ESCD. One may, e.g., by cutting off the individual branches without any harm, develop a means for solving the problems for narrower applications, or by using the ESCD as a nucleus and incorporating new algorithms, data, and pieces of knowledge into its structure, build up the system capable of solving more global problems. In the nearest future, the following is planned based on the proposed base version of the ESCDA:

– creation of a system for ecological and regional exploration of natural resources of forest areas based on processing of the aerospace data;

– carrying out an ecological expert inspection of the scorched woods in the Aleksandrovskii district of the Tomsk region based on the data of aerospace observations.

As has already been noted, the proposed version of the ESCD is a modification of the previously described information–interactive system[10] and, in its turn, it will be modified in order to extend the functional capabilities and to formalize new knowledge.

CONCLUSION

In spite of many developments in the knowledge–engineering field, employment of the AI for processing data of remote sensing is still at the initial evolution stage. Ideas for intellectual interpretation of remote measurements are formulated, main principles, which provides the basis for the structures of databases and knowledge, are elucidated, optimal ways of interaction with the user are studied, and attempts are made to formalize the new knowledge. This situation is associated first of all with the peculiarities of the subject area, its ambiguity, and information abundance. And, in addition, there exists quite a lot of investigators, who are untrained users of computers, but at the same time they need a computer system, the communication with which is as easy as with a teacher or a colleague. There are also specialists, who are ready to apply their knowledge for solving the problems, little known to them, in the process of being acquainted with a new subject. Therefore, the research efforts aimed at an intellectualization of the RS and designing the expert systems capable of incorporating new knowledge, are urgent and promising.

REFERENCES

1.M. Boarnet, C.Culbert, and T.R.Savely, in: Proc. IEEE Int. Conf. Rob. and Autom., Raleigh (1987), Vol. 1, pp. 131–136.
2. G.L. Vernon, Proc. Soc. Photo–Opt. Instrum. Eng. **580**, 21–27 (1985).
3. V.V. Fedorov, Tr. Vses. Nauchno–Issled. Ins. Sel'skokhoz. Meteorol., No. 25, pp. 70–83 (1985).
4. J.E. Ester, M.A. Friedl, and J.L. Star, Proc. Soc. Photo–Opt. Instrum. Eng. **924**, 297–304 (1988).
5. A. Reuter, Informationstechnik **29**, No. 3, 164–175 (1988).
6. M. Goldberg, D.G. Goodenough, and M. Alvo, Proc. IEEE **73**, No. 6, 119–130 (1985).
7. D.G. Goodenough, Photogramm. Eng. and Remote Sens. **54**, No. 2, 167–176 (1988).
8. J.G. Schuetzle and D.R. Zoch, in: *Abstracts of Reports of Exp. Syst. Gov. Symp. McLesn*, Va., (1986), pp. 210–220.
9. M.S.H. Heng, Future Generat. Comput. Syst. **3**, No. 2, 103–109 (1987).
10. A.G. Gendrin, N.V. Novikova, and E.I. Pospelova, "*Information–interactive system of correction of the data of remote sensing of the parameters of the underlying surface for the effect of the atmosphere*", Preprint No. 22, Institute of Atmospheric Optics of the Siberian Branch of the Academy of Sciences of the USSR, Tomsk (1989), 73 pp.
11. S.V. Afonin and A.G. Gendrin, in: *The Software for Atmospheric Optics Problem* (Nauka, Novosibirsk, 1988) pp. 38–64.
12. V.E. Zuev, *Propagation of Visible and Infrared Radiation in the Atmosphere,* (Halsted Press, New York, 1974).
13. G.M. Krekov and R.F. Rakhimov, *Optical–Radar Model of Continental Aerosol* (Nauka, Novosibirsk, 1982), 200 pp.
14. V.E. Zuev, *Propagation of Laser Radiation in the Atmosphere* (Radio i svyaz', Moscow, 1981), 288 pp.
15. G.V. Telegin and V.V. Fomin, "*On a possibility of analytic approximation of the line shapes in the molecular vibrational–rotational spectra*", Preprint No. 26, Institute of Atmospheric Optics, Siberian Branch of the Academy of Sciences of the USSR, Tomsk (1979), 52 pp.
16. K.G. Rubinshtein and V.B. Shilyaev, Metereol. Gidrol. , No. 9, 26–33 (1987).
17. W. McCartney, *Atmospheric Optics* [Russian translation] (Mir, Moscow, 1979), 421 pp.
18. R.G. Reitenbakh and B.G. Sherstyukov, in: *Tr. 5th All–Union Conference on Application of Statistical Methods in Meteorology*, Kazan' (1985), pp. 11–113.