https://doi.org/10.15407/knit2020.04.031
UDC 553.98:528.854

A. V. KHYZHNIAK, Researcher, PhD in Tech.
E-mail: avsokolovska@gmail.com
O. D. FEDOROVSKYI, Head of Department, Corresponding Member of the National Academy of Sciences of Ukraine, Dr. Sci. in Phys. & Math., Professor, the Winner of State Awards in Science and Technology of Ukraine (2005)
E-mail: adfedorovsky@ukr.net

State Institution "Scientific Centre for Aerospace Research of the Earth of the Institute of Geological Science of the National Academy of Sciences"
55-b O. Honchara Str., Kyiv 54, 01601, Ukraine

INTEGRATION OF REMOTE SENSING DATA AND GROUND-BASED INFORMATION TO SOLVE NATURAL RESOURCES AND ENVIRONMENTAL PROBLEMS

We consider the rationale for the integration of remote sensing data and ground information using a statistical criterion to solve natural resources and environmental problems. The method is proposed to be implemented as the module of the "computer assistant". The algorithm that describes a set of the sequence of operations for automation of the decision-making procedure is presented. This algorithm should free the operator from a significant amount of subjective and labor-intensive work, which is performed using visual methods.

Based on the proposed methodology, we perform the object recognition in aerospace images of the territory with different geological and landscape conditions and, respectively, with standard objects of different classes with different sets of values of informative features (of different nature and dimension). For the recognition and classification of images of studied objects, the probabilities of the ratios of informative features of the studied areas to ones of each standard object present in the aerospace image were determined.

The results of testing the proposed methodology are presented on the examples of assessing oil and gas prospects in the areas of the Dnieper-Donetsk cavity and the problem of classifying crops of different varieties in different periods of vegetation in the agricultural fields in the Kyiv region.

Key words: *remote sensing, integration, natural resource, system analysis, statistical criterion.*

There are several ways to solve environmental problems based on the decoding of aerospace images: visual, interactive, and automatic [2, 8, 11]. However, the growth of the scope of the thematic tasks and the creation of new systems of remote sensing of the Earth requires the improvement and development of new methods for decoding aerospace images, including the recognition of the space images of objects' classes and their detection. For example, analyzing and classification of the aerospace images is performed by measuring and comparing the informative features of the objects with the corresponding features of the object taken as the standard [7]. The correspondence of the object and the standard class is determined by the maximum value of the calculated membership function [5].

There are possible situations when the assessed object may belong to different standard classes. In that case, the classification of this object is performed by calculating the probability of its relation to each

Цитування: Khyzhniak A. V., Fedorovskyi O. D. Integration of remote sensing data and ground-based information to solve natural resources and environmental problems. *Космічна наука і технологія*. 2020. **26**, № 4 (125). С. 31—37. https://doi.org/10.15407/knit2020.04.031

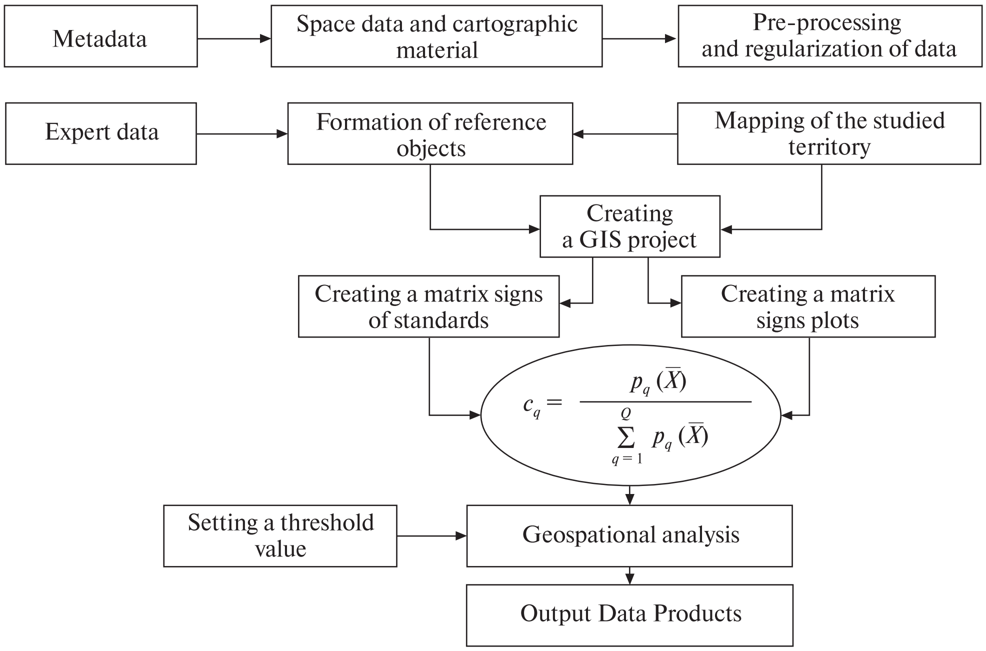

Fig. 1. Algorithm for automation of decision making

of possible standard classes, and belonging is determined by the maximum value of the probability of its relation to the particular standard class. The calculation of the above probabilities is done using the heuristic criterion (1), which involves previously entered informative features of each standard class and automatic calculation of probability ratio of the belonging of the studied object to the standard classes [1]:

$$c_q = \frac{p_q(\bar{X})}{\sum_{q=1}^{Q} p_q(\bar{X})}, \qquad (1)$$

where:

$$p_q(\bar{X}) = \left(\frac{1}{2\pi}\right)^{K/2} \prod_{k=1}^{K} \exp\left[\frac{(L_k - L_{q,k})^2}{2\sigma_{q,k}}\right] / \sigma_{q,k}$$

— multidimensional distribution density,

$$L_{q,k} = \frac{1}{N_q} \sum_{n=1}^{N_q} L_{q,k,n}$$

— average value of spectral brightness,

$$\sigma_{q,k} = \left[\frac{1}{N_q} \sum_{n=1}^{N_q} (L_{q,k,n} - L_{q,k})^2\right]^{-1/2}$$

— average dispersion value,

Q — the number of classes of objects to be identified, q — the current number (index) of a particular class of objects, K — number of informative features used, k — the current number of a specific informative feature, L — the result of a particular measurement of an informative feature (vector of dimension K with coordinates $L_1,..., L_k$), N_q — the sample size for the random variable $L_{q,k}$, n — current value.

Model decision support systems can be considered as the "computer assistant operator", and in case of weakly structured tasks, the decision maker is able to make the best decision using the computer together with personal efforts. Such kind of tasks characterized by the uniqueness and lack of information is widespread in aerospace research of natural resources management. The solution of well-structured tasks by verified algorithms practically does not require the participation of a decision maker and the usage of a computer assistant operator [6].

The main advantage of objects' classification method based on the heuristic criterion in comparison with other statistical methods is the structuring of a complex problem by calculating the level of belonging of object parameter values concerning the data of the object considered as standard. The latter is proposed to be used as a "computer assistant operator" module when decoding aerospace images. An operator develops and takes a balanced decision using a computer, having the databases, methods, models, and criteria. The basis of the “computer assistant operator” is mathematical modeling, which allows a decision maker to obtain the missing information for him to make a decision.

Probabilistic and statistical methods used to solve optimization problems, statistical theories of identification, recognition and number theory [3, 4, 9, 10] are applied in [6] for the representation of the heuristic criterion of object classes determination by the results of measuring their informational features.

Processing of multizonal aerospace images for objects’ classification according to known standards is interesting for the considered problems.

Based on heuristic criterion, the recognition of objects in aerospace images of the territory with different geological and landscape conditions and, respectively, the conditions with standard objects of different classes with different sets of values of informative features (of different nature and dimension) was performed. For the recognition and classification of investigated objects’ images, the probabilities of the ratios of the studied areas’ informative features to ones of each standard object present in the aerospace image was calculated. The class of the studied object was determined by the probability value which is higher or equal to the threshold value for the particular standard object by assessment of abovementioned parameters based on a heuristic criterion, which includes the introduction of informative features of all the standards and automatic calculation of probabilities of the ratio of each studied object to all referenced objects.

The sequence of operations to implement the proposed method of automation of the decision-making procedure should free the operator from a significant amount of subjective and labor-intensive work which is performed using visual methods. The above sequence of execution is shown in Fig. 1.

Consider two examples of approbation of the proposed method below.

The first one is an assessment of the oil and gas availability of the sites on the territory of the Dnieper-Donetsk cavity, which was carried out using three standard classes: two of them — productive wells and

Fig. 2. Territory of the study of the Dnieper-Donetsk cavity (fragment of Landsat 8, 10.08.2017)

the other – empty one (Fig. 2). The following archive data were selected to assess the oil availability: a map of residual anomalies of the gravitational field, a geothermal grade map, a gravitational field diagram, an anomalous magnetic field map, temperature charts at the cut –3500 m and –5000 m. As a result of the pre-processing of the total a priori information, maps of various anomalies related to hydrocarbon deposits in one format are obtained. Also, multi-spectral (hyperspectral) and radar aerospace images (Landsat / ETM +, EO1 / Hyperion, Sentinel 1A, Sentinel 2A, and others) were used to calculate the Haralick’s texture parameters and to create the scheme of relative neotectonic activity of blocks from the minimum to the maximum (in points from 1 to 10) (in the process of structural decoding and structurally-geomorphological analysis, schemes of block and high-altitude

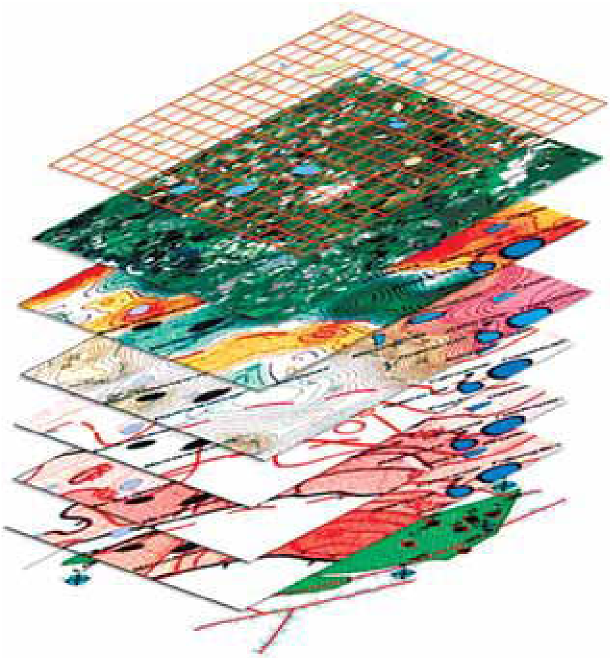

Fig. 3. A set of input layers of heterogeneous geospatial data

landscape fields are built). In addition, for obtaining the values of all types of features in the single points of the selected area, the grid (299 separate plots) with a size of 4×4 km was created. The set of input data is presented on Fig. 3.

Then, according to the algorithm (Fig. 1), all relevant iterations were carried out, and the threshold value of the probability of the ratio of the values of the informative features of each studied object to the values of the informative features of each object of a standard class of oil and gas availability present in the aerospace image was determined.

The final product of this process is the mapping of interdisciplinary, integrated assessment of the oil and gas availability in the studied areas and the determination of zones of reservoirs on the investigated territory, which is presented in Fig. 4 as the gray-scaled levels — the black color corresponds to the maximum value of oil and gas availability. Creating a map diagram characterizing the oil and gas availability in the zone and each of its sites, based on the results of remote and terrestrial studies, allows us to distinguish abnormal areas in the studied area.

The second example of approbation of the proposed method is the classification of agricultural fields of the Kyiv region. Fig. 5 depicts the part of the area being studied with fields for standard classes and fields for classification.

The satellite image, which closely corresponds to the date of the growing season and ground-based data, was used to carry out the research. The Sentinel 2A image for August 1, 2017, spectral channels 2—12 (496.6 nm — 2202.4 nm) was selected [9]. Verification of the obtained results was carried out using the ground-based data for August 1, 2017, on the studied territory. According to the preconditions of the heuristic criterion, data for the study were selected for each of 5 crops (sunflower, corn, soybean, sugar beet, buckwheat) for 15 reference values of spectral characteristics. The separate class was added for soybean with wild grass. Also, separate classes for stubble and plowing were added, since the period of growth has ended.

The validation of the obtained results was carried out on 8 separate sites in accordance with the above-mentioned points with ground-based data on the same fields to assess the accuracy of the classification. The results of the validation are shown in Table 1. The following metrics were used for comparison: overall classification accuracy, producer's accuracy (PrAc), and user's accuracy (UsAc). The producer's accuracy shows how well the classification result for this class matches the test data. The user's accuracy, in turn, shows how likely it is that this class coincides with the results of the classification.

The mismatch matrix for the test sample using the proposed method is shown in Table 2 for more detailed consideration.

Table 1. **The results of classification accuracy assessment**

№ of class	Crop	PrAc, %	UsAc, %
1	Stubble	100	100
2	Sunflower	100	98
3	Corn	95.31	61
4	Soybean	557.56	99
5	Plowing	94.34	100
6	Soybean with weed	95.45	63
7	Sugar beet	100	94
8	Buckwheat	100	100
Overall accuracy OvAc = 89.37 %			

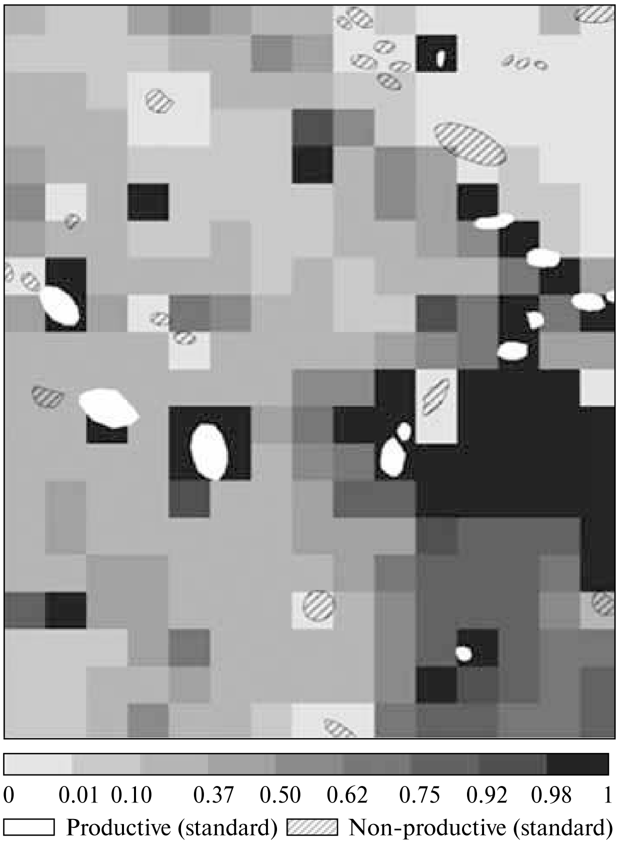

***Fig. 4*.** Mapping of oil prospecting of areas of the territory of Dnieper-Donetsk cavity. Relative perceptivity is presented in brightness gradations - the black color corresponds to the maximum value of oil prospect

Fig. 5. The part of the test area on the Sentinel 2A for the 01.08.2017 (*1* — stubble, *2* — sunflower, *3* — corn, *4* — soybean, *5* — plowing, *6* — sugar beet, *7* — fields with unknown crops)

Table. 2. **Mismatch matrix for test sampling**
(PrAc — producer's accuracy, UsAc — user's accuracy, Com — commission, Om — omission)

№ of class	1	2	3	4	5	6	7	8	UsAc, %	Om
1	1	0	0	0	0	0	0	0	100	0.0
2	0	0.98	0.02	0	0	0	0	0	98	2.0
3	0	0	0.31	0.47	0	0.2	0	0	61	39
4	0	0	0	0.99	0	0.01	0	0	99	1.0
5	0	0	0	0	1	0	0	0	100	0.0
6	0	0	0.01	0.46	0	0.53	0	0	63	37.0
7	0	0	0	0	0,06	0	0.94	0	94	6.0
8	0	0	0	0	0	0	0	1	100	0.0
PrAc, %	100	100	95.31	57.56	94.34	95.45	100	100		
Com	0.0	0.0	4.69	42.44	5.66	4.55	0	0		

As it is seen from the table above, the results of the classification of certain crops have a rather low correlation coefficient with ground-based data. This is due to the difference in vegetative periods. Therefore, the next step of research should be the study of agricultural crops in different periods of vegetation and the classification of crops for different varieties.

As a result of the study, it was found that the proposed method makes it possible to obtain the results not only of the classification of cultures but also of their belonging to each other in percentage. The latter expands the possibility of further evaluation of the particular agricultural crop condition. Also, a certain assessment of the crop condition has been performed. For example, it was managed to divide soybean crops into conditionally clean and inbred ones.

RESULTS

The use of the heuristic criterion for solving the problems of the nature resources usage with the processing of aerospace images and in-situ measurements as the module of the "computer assistant" is substantiated. The proposed method was tested on the examples of assessing the oil and gas availability in the regions of the Dnieper-Donetsk cavity and the problem of classifying crops of different varieties in different periods of vegetation in the agricultural fields of the Kyiv region.

REFERENCES

1. Arkhipov A. I., Glazunov N. M., Khyzhniak A. V. (2018). Heuristic criterion for class recognition by spectral brightness. *Cybernetics and Systems Analysis*, **54**(1), 105—110.
2. Kemker R., Salvaggio C., Kanan C. (2018). Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS J. Photogramm. *Remote Sens*, **145**, 60—77.
3. Knopov P. S., Kasitskaya E. J. (2002). *Empirical estimates in stochastic optimization and identification*. New York: Kluwer Acad. Publ.
4. Kolmogorov A. N. (2005). Selected Works. Vol. 2. Probability theory and mathematical statistics. Moscow: Mat. Institute named after V. A. Steklova RAS. Science, 584 p. [in Russian].
5. Lyalko V. I., Fedorovsky O. D., Popov M. O., et al. (2004). Using satellite data to study natural resource issues. *Space Research in Ukraine. 2002—2004*. Kyiv, 7—14 [in Russian].
6. Popov M. A., Stankevich S. A., Arkhipov A. I., Titarenko O. V. (2018). About possibility of hydrocarbon deposit remote detection using computer assistance. *Ukrainian J. Remote Sensing*, **16**, 34—40. URL: https://www.ujrs.org.ua/ujrs/article/view/119 (Last accessed 08.08.2019).
7. Popov M. A., Stankevich S. A., Topolnytskyj M. V., Titarenko O. V. (2017). Integration of remote and geological-geophysical data in the search for oil and gas deposits on land. *Ecological safety and nature management,* Vol. 1-2 (23), 36—43 [in Ukrainian].
8. Ren S., He K., Girshick R., Sun J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, **39**, 1137—1149.
9. Vapnik V. N. (1998). *Statistical learning theory*. New York: Wiley.
10. Webb A. R., Copsey K. D. (2011). *Statistical Pattern Recognition* (3rd ed.). New York: John Wiley.
11. Yakimchuk V. G., Sukhanov K. Yu. (2018). Estimation of hydrophysical characteristics of the aquatic environment using satellite images in the context of incomplete information, **19**, 33—38. URL: http://ujrs.org.ua/ujrs/ (Last accessed 08.08.2019).

Received 08.08.2019

А. В. Хижняк, наук. співроб., канд. техн. наук
E-mail: avsokolovska@gmail.com
О. Д. Федоровський, зав. відділу, чл.-кор. НАН України, д-р фіз.-мат. наук, проф.,
Лауреат Державної премій України в галузі науки і техніки (2005 р.)
E-mail: adfedorovsky@ukr.net

Державна установа «Науковий центр аерокосмічних досліджень Землі
Інституту геологічних наук Національної академії наук України»
вул. О. Гончара 55-б, Київ 54, Україна, 01601

ІНТЕГРУВАННЯ ДАНИХ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ ТА НАЗЕМНОЇ ІНФОРМАЦІЇ ДЛЯ РОЗВ'ЯЗУВАННЯ ЗАДАЧ ПРИРОДОКОРИСТУВАННЯ

Обґрунтовується інтегрування даних дистанційного зондування Землі та наземної інформації із застосуванням для розв'язування задач природокористування статистичного критерію як модуля комп'ютерного асистента оператора. Представлено алгоритм, який описує сукупність операцій та послідовність їхнього виконання, що реалізують запропонований метод автоматизації процедури прийняття рішень. Метод має звільнити оператора-дешифрувальника космічних знімків від значного об'єму суб'єктивної та трудомісткої роботи, яка виконується на основі візуально-інструментальних методів. На основі запропонованої методики було виконано розпізнавання зображень об'єктів на аерокосмічних знімках території з різними геологічними і ландшафтними умовами і відповідно до тих умов з об'єктами-еталонами різного класу з різним набором значень інформативних ознак (різної природи і розмірності). Для розпізнавання і класифікації зображень досліджуваних об'єктів визначались ймовірності відношення інформативних ознак досліджуваних ділянок до інформативних ознак кожного наявного на аерокосмічному знімку об'єкта-еталона. Представлено результати апробації запропонованої методики на прикладі оцінки нафтогазоперспективності ділянок Дніпровсько-Донецької западини та класифікації агрокультур сільськогосподарських полів у Київській області на різні сорти в різні періоди вегетації.

Ключові слова: дистанційне зондування Землі, інтеграція, природокористування, системний аналіз, статистичний критерій.