Neural Network Classification of Large-sized Multi-segmented Polygon with Formation of Features by the Hilbert-Huang Transform of Hyperspectral Data

Evgeny S. Nezhevenko
Institute of Automation and Electrometry SB RAS, Novosibirsk

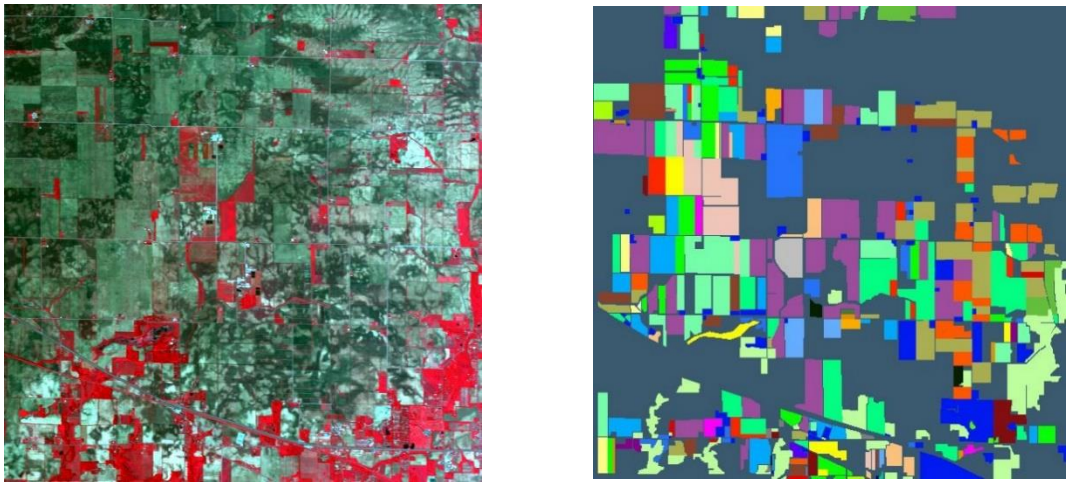
A method of classification of hyperspectral images of terrain is described. It includes preprocessing of the spectral data in the form of conversion to the principal components, spatial processing consisting of finding the empirical mode of the principal components (Hilbert-Huang transform), and classification itself using RBF neural networks. Experiments are conducted on a large-format (580x580) hyperspectral image, and difficult-to-distinguish classes are not united.

.

Keywords: hyperspectral image, principal component, Hilbert-Huang transform, neural networks.

Introduction As observation tools are developed and improved, classification of terrain regions on the basis of hyperspectral data has recently become one of the most important aspect of remote sensing science. It allows solving various problems of land tenure in agriculture, forest management, environmental science, minimization of detrimental effects induced by natural disasters, etc. This problem was discussed in many publications, and a good review can be found in [1]. This review describes a large number of classification methods based on hyperspectral data finalized by particular results obtained by the authors personally on a large-format hyperspectral image (HSI). Unfortunately, they only briefly mentioned one of the image processing methods, whereas it provided good results on previous medium-format multispectral and hyperspectral images [2, 3]. In the present work, we want to demonstrate that our approach applied to a large-format HSI yields results that can be ranked among the best results obtained by other methods. However, a few words have to be said about the method itself. It was described in much detail in [3]; therefore, only a brief presentation of the method is given below.

Object of classification As it is obvious from Introduction, we want to compare our method with other options described in publications. For this purpose, the methods have to be checked on a similar object (the best option is to use identical objects). In our study, we used a large-format HSI obtained within the framework of the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) program at the Indian Pines test field (Indiana, USA) [4]. The image size was 614 × 2677 pixels, and the resolution was 20 m/pixel. The image was divided into 58 classes. In addition to a highway, railway, and residential buildings, the image displays cultivated crops (including 15 classes of corn and 18 classes of soya produced by different method of soil treatment). The number of spectral channels is 220 in the range of 0.4-2.5 μm, and 20 most noisy channels were eliminated from consideration. An informative fragment of 580x580 pixels was chosen (it is shown in pseudo-colors in Fig. 1a). All classes captured within the chosen fragment are shown in pseudo-colors in Fig. 1b. The experiment was performed with 33 classes having a sufficient number of pixels. Among the chosen classes, there are 9 classes of corn and 12 classes of soya.



а б

Fig.1 HSI fragment (a) and its division into classes (b)

Classification method As is known from the pattern recognition theory, object preprocessing is an extremely important stage of the classification procedure. Here we consider a spectral-spatial object; hence, preprocessing is performed in the spectral-spatial domain. As it was already mentioned, each pixel is characterized by 200 spectral features. The analysis shows that they are mutually correlated to a large extent; therefore, it is logical to pass to non-correlated features. This can be done by several methods, but the most effective option is to perform transformation to the principal components (PCs). An important issue is the number of components to be left in order to lose the minimum possible amount of information. Kaiser and scree tests are applied for this purpose. After the transformation to the principal components is performed, a spatial transformation is applied. It was shown in many publications [5, 6, 7] that the use of these tests ensures significant enhancement of the overall accuracy, which is the percentage of the ratio of the number of correctly classified pixels to the total number of pixels in the sample. In these publications, similar to many others, however, the spatial transformation is reduced to changing the pixel value depending on its neighborhood. The structure of the classified fragment regions is absolutely ignored, and large errors of classification are observed in the case of difficult-to-distinguish regions. Because of that, in particular, in [6, 7, 8] all soya regions with different methods of soil treatment were united into one class. The same refers to corn. Our method is based on another procedure. Each principal component is expanded into the so-called empirical functions (or intrinsic mode functions, IMFs). In contrast to the Fourier transform or wavelet transform, IMFs are not defined analytically and are determined exclusively by the analyzed sequence itself. The basis functions of the transform are formed adaptively, directly from the input data.

The algorithm of expansion into IMFs is based on constructing smooth envelopes on the basis of the extreme (maximum and minimum) points of the sequence and subsequent subtraction of the mean value of the envelopes from the initial sequence. For this purpose, the maximum and minimum points are found and approximated by splines. These splines are the upper and lower envelopes. The process of envelope construction is illustrated in Fig. 2.

The analyzed sequence is presented in Fig. 2 by the thin blue curve. The maximum and minimum points of this sequence are marked by the red and blue colors, respectively. The

envelopes are shown by the green curve. The mean value is calculated on the basis of two envelopes (shown by the dashed curve in Fig. 2). The thus-found mean value is further subtracted from the initial sequence.

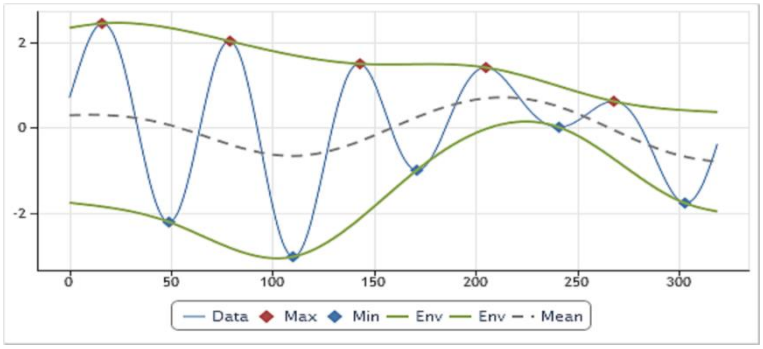


Fig.2 Construction of the intrinsic mode function of a one-dimensional function

These steps are performed to find the first approximation of the sought IMF. For definite identification of the IMF, it is necessary to find the maximum and minimum points of this IMF estimate and repeat these steps again. This process (called sieving) is continued until a threshold condition is satisfied. If the sieving process is successfully finalized, we obtain the first IMF. To find the next IMF, it is necessary to subtract the found IMF from the initial signal and repeat the procedure again. The procedure is repeated until all IMFs are found. The algorithm of expansion of the two-dimensional signal into IMFs formally does not differ from expansion of the one-dimensional signal, thought there are certainly some specific features associated with the search for extreme points and interpolation [9,10]. After *m* intrinsic mode functions of *n* principal components are found, each HIS pixel is described by an $m_x n$-dimensional vector of features. All HSI vectors together with information about the classes are fed to the neural network. Pixels of each class are randomly divided into three samples: learning sample (LS), control sample (CS), and test sample (TS). The process of learning is terminated on the basis of results of the CS. The classification accuracy is checked for all samples, but special care is applied for the TS because it is the TS that determines the neural network capability to generalization.

Experimental results The HSI shown in Fig. 1a was expanded into the principal components from which four components were chosen on the basis of the scree test (which include 99.42% of data dispersion). Then each of the PCs was expanded into five IMFs. The first PC and its five IMFs are shown in Fig.3.

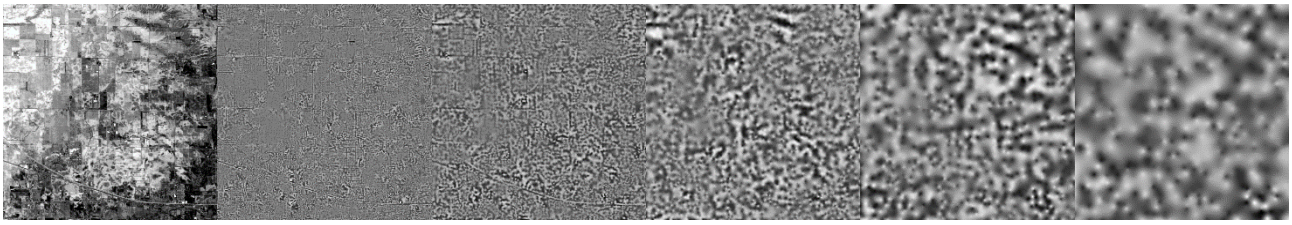


Fig.3 First PC and its five IMFs

Thus, after all these transformations, the HSI is described by a 580x580x20 array divided into 33 classes. The number of pixels in each class is divided in the ratio LS:CS:TS= 50%:25%:25%. Learning is performed in the RBF neural network. The learning time is several hours. Let us consider the results of classification after the learning process. Let us first give the list of all classes with numeration that will be used in the tables.

Table 1 List of classes with numeration

1	Background	31	Oats
2	Bare soil	33	Orchard
3	Buildings	35	Pond
4	Concrete/asphalt	36	Soybeans
7	Corn-EW	38	Soybeans-NS
8	Corn-NS	41	Soybeans Clean Till-WE
9	Corn	42	Soybeans Clean Till-NS
12	Corn Clean Till-NS irrigated	43	Soybeans Clean Till Drifted
14	Corn Min Till	44	Soybeans Clean Till Weedy
15	Corn Min Till- WE	45	Soybeans Driller
16	Corn Min Till-NS	46	Soybeans Min Till
17	Corn No Till	47	Soybeans Min Till- WE
18	Corn No Till-WE	51	Soybeans No Till EW
25	Grass Runway	52	Soybeans No Till NS
27	Hay?	53	Soybeans No Till Drilled
29	Lake	57	Wheat
		58	Woods

The classification results are summarized in Table 2. The network architecture in the left box of the table means 20 input features (number of neurons in the first layer), 851 neurons (RBF functions) in the hidden layer, and 33 neurons in the output layer (number of classes). The classification accuracy (CA) is not that high as in experiments performed with smaller fragments and a smaller number of classes [2, 3], but it should be again recalled that the present study involves difficult-to-distinguish classes of corn and soya, which are not united into two classes.

Table 2 Total classification accuracy of the fragment

	File “All regions of	class 33”	
Network architecture	Classification accuracy Learning sample	Classification accuracy Control sample	Classification accuracy Test sample
RBF 20-851-33	89.35	87.36	87.31

Let us consider the results of TS classification in more detail with indication of CA percentage (Table 3).

Table 3 Results of fragment classification into classes

Class number	1	2	3	4	7	8	9	12	14	15	16	17	18	25	27	29	31	33	35	36	38	41	42	43	44	45	46	47	51	52	53	57	58
Percentage of correct data	6.2	19.0	57.1	12.3	8.0	12.1	19.8	6.1	11.1	9.5	15.8	11.4	9.1	32.6	28.2	26.6	27.4	10.8	15.4	1.8	22.4	18.7	29.6	8.8	11.0	34.3	20.2	30.0	10.1	9.0	3.8	3.7	0.7
Percentage of erroneous data	93.8	81.0	42.9	87.7	92.0	87.9	80.2	93.9	88.9	90.5	84.2	88.6	90.9	67.4	71.8	73.4	72.6	89.2	84.6	98.2	77.6	81.3	70.4	91.2	89.0	65.7	79.8	70.0	89.9	91.0	96.2	96.3	99.3
Correctly	552	374	109	501	88	623	206	41	324	395	163	102	493	71	66	79	34	243	276	7	431	141	243	48	569	135	68	334	52	621	22	123	1
Erroneously	8363	1593	82	3582	1013	4547	837	635	2592	3752	871	796	4949	147	168	218	90	2006	1514	372	1489	614	577	495	4615	259	268	781	464	6258	561	3234	143
Total	8915	1967	191	4083	1101	5170	1043	676	2916	4147	1034	898	5442	218	234	297	124	2249	1790	379	1920	755	820	543	5184	394	336	1115	516	6879	583	3357	144

It is seen that the CA percentage varies from 42% (class 3 “Buildings”) to 99.3% (class 58 “Forest”). It should be noted that higher CA values are normally provided by classes with a large number of pixels: the mean percentage of correct classification is 83.63% based on classes and 87.31% based on pixels.

Let us consider Table 4, which is the error matrix showing the classes to which erroneously classified data are referred to. Each row of the matrix is normalized to the number of correctly classified pixels in the corresponding classes.

Table 4 Classification error matrix for the fragment

	1	2	3	4	7	8	9	12	14	15	16	17	18	25	27	29	31	33	35	36	38	41	42	43	44	45	46	47	51	52	53	57	58
1	100	0.71	0.23	1.30	0.66	2.06	0.60	0.35	0.86	1.11	0.01	0.07	1.49	0.01	0.22	0.00	0.04	0.04	0.19	0.01	1.21	1.03	1.09	0.51	2.18	1.04	0.29	1.20	0.00	2.88	0.02	0.29	0.00
2	1.44	100	1.57	0.88	0.00	0.38	0.06	0.00	0.31	0.13	0.00	0.13	1.76	1.32	0.25	0.00	0.13	6.28	0.06	0.00	0.75	0.00	0.82	0.25	0.75	0.00	0.00	0.06	0.00	0.44	0.38	2.32	0.00
3	1.22	1.22	100	0.00	0.00	6.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.47	0.75	0.00	100	0.00	0.64	1.17	0.00	0.45	0.00	0.00	0.00	0.11	0.00	0.00	0.08	0.00	0.03	2.74	0.00	0.42	0.08	0.06	0.00	1.54	0.00	0.00	0.45	0.56	0.39	0.00	0.11	0.00
7	2.86	0.10	0.00	0.00	100	0.00	0.10	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.97	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00
8	2.24	0.53	1.10	2.27	0.00	100	0.33	0.00	0.24	0.53	0.00	0.37	0.90	0.04	0.00	0.09	0.04	0.22	0.02	0.07	1.14	0.09	1.67	0.00	2.60	0.18	0.31	0.37	0.00	1.08	0.00	0.07	0.00
9	0.96	0.00	0.00	4.18	0.48	0.00	100	0.00	0.24	0.96	0.00	0.00	0.00	0.00	0.00	1.67	0.00	0.00	0.00	0.12	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.72	0.00	0.00	0.12
12	0.31	0.00	0.00	0.00	0.00	0.00	0.47	100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.54	0.08	0.00	0.00	0.12	0.04	0.19	0.00	100	0.46	0.00	0.04	0.04	0.04	0.00	0.04	0.00	0.08	0.12	0.00	0.08	0.00	0.77	0.00	0.54	0.00	0.23	0.08	0.00	1.16	0.00	0.04	0.00
15	1.65	0.24	0.00	0.03	0.40	0.56	0.19	0.00	0.85	100	0.00	0.21	0.91	0.00	0.00	0.00	0.00	0.00	0.16	0.03	1.36	0.72	0.00	0.00	1.07	0.11	0.03	2.11	0.59	1.95	0.00	0.11	0.00
16	0.23	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.11	100	0.00	0.92	0.00	0.00	0.00	0.00	0.00	1.95	0.00	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.87	0.00	0.00	0.00
17	0.38	0.13	0.00	0.13	0.00	7.16	0.00	0.00	0.63	1.51	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00	0.88	0.00	0.13	0.00	0.00	0.00	2.64	0.00	0.00	0.50	0.00	0.63	0.00	0.00	0.00
18	0.61	0.10	0.04	0.12	0.00	0.53	0.22	0.00	0.61	1.29	0.81	0.00	100	0.12	0.00	0.14	0.00	0.14	0.18	0.00	0.30	0.12	0.73	0.00	0.89	0.14	0.06	0.44	0.00	2.28	0.00	0.06	0.00
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.00
27	2.38	3.57	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.19	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.00
29	0.00	0.00	0.00	0.00	0.46	0.00	0.00	0.00	2.75	0.00	0.00	0.00	0.00	0.00	0.00	100	0.92	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	2.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
33	0.40	6.18	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.15	0.00	0.00	0.80	0.85	1.79	0.00	0.00	100	0.55	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.60	0.00	0.10	0.25	0.70	0.00
35	0.59	0.00	0.00	4.82	0.00	2.18	0.00	0.00	0.00	0.07	3.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	0.00	0.40	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00
36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.61	0.00	0.00	0.00	0.00	0.00
38	1.68	0.87	0.07	2.75	0.40	1.21	1.48	0.00	0.54	0.94	0.00	0.47	0.07	0.00	0.00	0.00	0.00	0.00	0.07	0.00	100	0.00	0.00	0.00	1.41	0.00	0.07	0.20	0.00	0.20	0.00	0.00	0.00
41	2.93	0.00	0.00	0.00	0.00	0.81	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.14	100	0.16	0.16	0.00	0.00	0.65	0.00	0.00	0.00	0.00	0.00	0.00
42	0.00	0.52	0.00	0.00	0.00	0.69	0.00	0.00	0.35	0.00	0.00	0.00	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.39	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	100	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00
44	1.67	0.30	0.17	1.47	0.04	2.15	0.59	0.07	1.82	0.95	0.67	0.46	1.91	0.30	0.00	0.69	0.02	0.15	0.30	0.02	1.73	0.11	0.02	0.00	100	0.13	0.26	1.39	0.00	0.85	0.00	0.07	0.00
45	3.09	3.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00	0.00
46	1.87	0.00	0.00	1.12	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.12	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00
47	0.26	0.00	0.26	0.51	0.00	0.00	0.38	0.00	0.51	0.51	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.26	0.00	0.00	0.00	2.18	0.00	0.00	100	0.00	0.90	0.00	0.38	0.00
51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.90	0.00	0.00	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100	1.08	0.00	0.00	0.00
52	1.52	0.35	0.02	0.50	0.03	2.19	0.19	0.14	0.69	1.28	0.56	0.62	2.27	0.11	0.00	0.21	0.24	0.48	1.01	0.00	1.15	0.06	0.03	0.00	0.56	0.21	0.03	0.03	0.16	100	0.00	0.21	0.00
53	0.89	0.18	0.18	0.00	0.00	1.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	100	1.96	0.00
57	0.09	1.67	0.00	0.37	0.00	0.06	0.03	0.00	0.00	0.00	0.00	0.00	0.06	0.25	0.12	0.25	2.35	0.00	0.00	0.00	0.03	0.00	0.03	0.00	0.03	0.00	0.00	0.09	0.00	0.00	0.28	100	0.00
58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100

The gray regions are the classes of one plant type: all classes of corn (left) and all classes of soya (right). It can be concluded from the comparison of these regions with others that the total cross error in the soya region is several times smaller than that in other regions. It fact testifies that soya regions with different types of soil processing and other characteristics are indeed hard to distinguish. This phenomenon for the corn regions is observed to a smaller extent. Probably, this difference is caused by specific aspects of processing of corn and soya regions.

Conclusions Thus, the experiments aimed at classification of a large-format hyperspectral image confirmed the high efficiency of the method including preprocessing in the form of the transformation to the principal components of the spectral components, spatial transformation in the form of expansion of the principal components into intrinsic mode functions, and classification in the neural network. The method ensures a high probability of correct classification of hyperspectral

images having similar spectral compositions and effectively operates in situations with terrain areas that can be hardly distinguished by conventional methods.

This work was supported by the Presidium of the Russian Academy of Sciences within the framework of the Complex Program of Basic Research of the Siberian Branch of the Russian Academy of Sciences (Project No. II.1.37).

REFERENCES

[1] *S.M. Borzov and O.I. Potaturkin*, Spectral-spatial methods for hyperspectral image classification. Review, Optoelectron., Instrum., Data Process. 2018, Vol. 54, No.. 6, pp. 582-599.

[2] *E.S. Nezhevenko and A.S. Feoktistov*, Investigation of the efficiency of neural network classification with the use of the Hilbert-Huang transform, in: Proc. Intern. Congress "Interexpo Geo-Sibir", April 18-22, 2016, Vol. 1, pp. 60-64.

[3] *E.S. Nezhevenko, A.S. Feoktistov, and O.Yu. Dashevskii*, Neural network classification of hyperspectral images on the basis of the Hilbert-Huang transform, Optoelectron., Instrum., Data Proc. 2017, Vol. 53, No. 2, pp 165-170.

[4] *M.F. Baumgardner, L.L. Biehl, and D.A. Landgrebe*, 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3. Purdue University Research Repository. 2015. doi:10.4231/R7RX991C.

[5] *A.V. Kuznetsov and V.V. Myasnikov*, Comparison of algorithm of controlled elemental classification of hyperspectral images, Komp. Optika, 2014, Vol. 38, No. 4, pp. 494– 502.

[6] *S.M. Borzov, A.O. Potaturkin, O.I. Potaturkin, and A.M. Fedotov*, Analysis of the efficiency of classification of hyperspectral satellite images of natural and man-made areas, Optoelectron., Instrum., Data Proc., 2016, Vol. 52, No. 1, pp. 1-10.

[7] *S.M. Borzov and O.I. Potaturkin*, Efficiency of the spectral-spatial classification of hyperspectral imaging data, Optoelectron., Instrum., Data Process., 2017, Vol. 53, No. 1, pp. 26-34.

[8] *S.M. Borzov and O.I. Potaturkin*, Classification of hyperspectral images with different methods of training set formation, Optoelectron., Instrum., Data Process., 2018, Vol. 54, No. 1, pp. 76-82.

[9] *L. Vincent*, Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms, IEEE Trans. Image Process., 1993, Vol.2, No. 2, pp. 176–201.

[10] *J.C. Carr, W.R. Fright, and R.K. Beatson*, Surface interpolation with radial basis functions for medical imaging, Proc. of the Conf. ACM SIGGRAPH. Los Angeles, USA, August 12–17, 2001, pp. 67–76.