ISSN 8756-6990, Optoelectronics, Instrumentation and Data Processing, 2016, Vol. 52, No. 1, pp. 1–10. © Allerton Press, Inc., 2016.
Original Russian Text © S.M. Borzov, A.O. Potaturkin, O.I. Potaturkin, A.M. Fedotov, 2016, published in Avtometriya, 2016, Vol. 52, No. 1, pp. 3–14.

ANALYSIS AND SYNTHESIS OF SIGNALS AND IMAGES

Analysis of the Efficiency of Classification of Hyperspectral Satellite Images of Natural and Man-Made Areas

S. M. Borzov[a], A. O. Potaturkin[a], O. I. Potaturkin[a,b], and A. M. Fedotov[b,c]

[a] *Institute of Automation and Electrometry, Siberian Branch, Russian Academy of Sciences pr. Akademika Koptyuga 1, Novosibirsk, 630090 Russia*
[b] *Novosibirsk State University, ul. Pirogova 2, Novosibirsk, 630090 Russia*
[c] *Institute of Computational Technologies, Siberian Branch, Russian Academy of Sciences, pr. Akademika Lavrent'eva 6, Novosibirsk, 630090 Russia*
E-mail: borzov@iae.nsk.su, potaturkin@iae.nsk.su

Received February 2, 2015

Abstract—The efficiency of a number of the classical methods of supervised classification of hyperspectral data is estimated by an example of discriminating the types of the underlying surface in natural and man-made areas. The minimum distance, support vector machine, Mahalanobis, and maximum likelihood methods are considered. Particular attention is paid to studying the dependence of the data classification accuracy on the number of spectral features and the way of choosing them in the above-mentioned methods. Experimental results obtained by processing real hyperspectral images of landscapes of various types are reported.

Keywords: remote sensing, hyperspectral images, classification of surface types, reflection spectrum.

DOI: 10.3103/S8756699016010015

INTRODUCTION

The modern status of remote sensing (RS) tools is characterized by intense application of technologies of hyperspectral (HS) imaging in the visible and infrared spectral ranges. Optoelectronic devices that allow obtaining high-quality HS images of the Earth surface were created. A specific feature of data recorded by these devices is a narrow width and a large number of spectral channels. Based on this fact, numerous approaches were developed, which involve the analysis of the fine structure of image pixel spectra and their classification through comparisons with spectra of reference surface areas [1]. However, applications of HS data in practice (owing to the tremendous amount of incoming information) require multiple enhancement of the power of computational tools, parallelization of data processing, and significant expansion of data transfer channels, which are extremely expensive, thus, restraining the development of technologies of HS monitoring of the Earth surface and near-Earth space.

Two alternative approaches can be used to solve this problem. The first approach is based on reducing the amount of data for solving particular applied problems by means of selecting the most informative spectral channels and subsystems of features [2, 3]. In the second approach, the amount of data is reduced by means of preliminary segmentation of images and their zone-by-zone processing in terms of spectral and spatial features with allowance for a priori information about the position and nature of observed objects [4]. The efficiency of these approaches primarily depends on methods chosen for data clusterization in the spectral and spatial domains.

The efficiency of various methods of classification of HS images can be reliably estimated only through comparisons of results of processing of real RS data and maps of ground-based (ground-truth) observations.

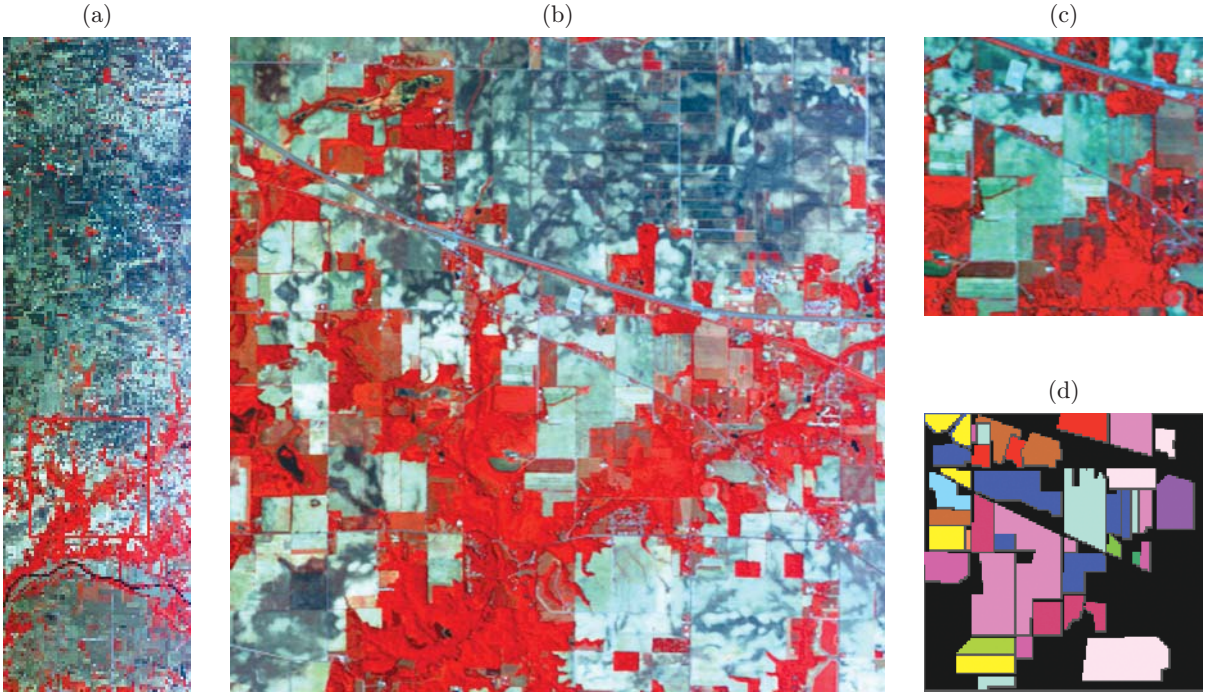

Fig. 1. Raw data: (a) *RGB* composite of the natural area image; (b) domain bounded by the frame; (c) fragment used for learning; (d) map of classes for the fragment.

Unfortunately, the amount of available non-classified HS images is often insufficient for such investigations; moreover, there are practically no corresponding ground-truth data for verification of the classification results, which is another essential restraining factor for further development of the above-mentioned technologies.

By an example of the analysis of real HS images, it was demonstrated [5, 6] that it is not always reasonable to use all channels simultaneously. Thus, the efficiency of classification of vegetation types remains practically at the same level with the number of features being reduced by a factor of 20 if these features are chosen correctly. It is obvious that choosing the method of formation of the system of informative features for each problem of the Earth surface monitoring is another difficult problem.

The challenge of the present activities was an experimental study of the efficiency of the classical methods of supervised classification of hyperspectral images of natural and man-made areas, depending on the number of spectral features taken into account with different methods used to choose these features.

CLASSIFICATION OF VEGETATION TYPES ON THE BASIS OF HS IMAGES OF NATURAL AREAS

The efficiency of the method of RS data classification for natural areas was studied by an example of processing a hyperspectral image (Figs. 1a and 1b) obtained within the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) program on the Indian Pines test site (Indiana, USA). The image size was 614×2677 pixels, the resolution was 20 m/pixel, and the number of channels was 220 in the range from 0.4 to 2.5 μm. Twenty channels with a high noise level were eliminated from consideration. The images show a highway, railroad, agricultural areas, forest, and residential facilities.

Based on the results of ground-based observations, the image was divided into 58 classes. Moreover, one of the fragments was divided into 16 classes (Figs. 1c and 1d) with 14 of them being different types of vegetation including, in particular, 3 classes of corn and 3 classes of soy with different ways of soil cultivation: no tilling (no-till), tilling with the use of low-intrusive technologies (min-till), and usual tilling (clear). The no-till and min-till technologies are aimed at providing the minimum effect on the soil structure. Some part of the area is not marked and is not considered further in the paper.

A comparison of the map of ground-based observations with the image shows that the map is significantly generalized. Multiple minor features visually seen in the image are not marked on the map. Thus, it cannot be expected that each pixel of a certain area is assigned to the class associated with this area. Moreover,

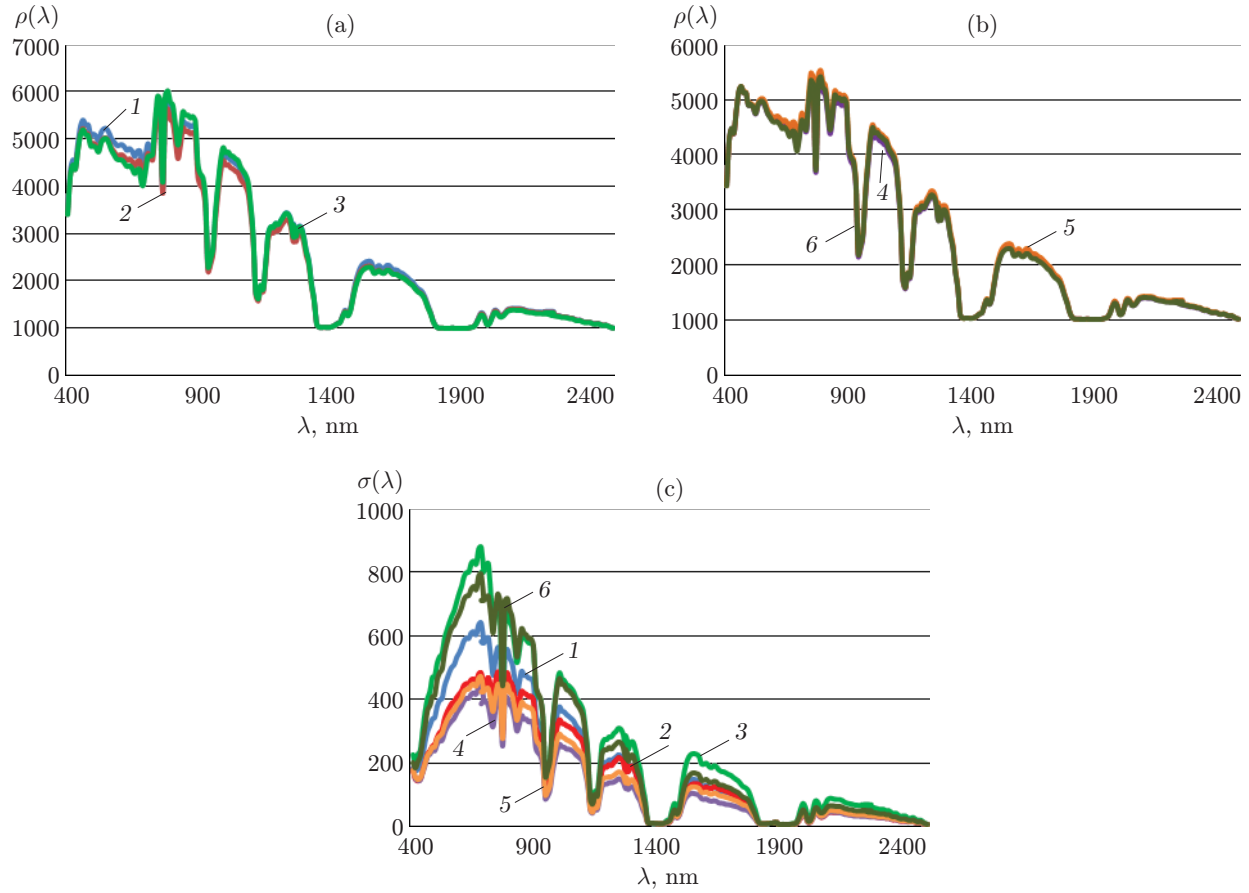

Fig. 2. Spectral characteristics of various classes: (a) corn; (b) soy; (c) scatter of spectral coefficients of reflection within the classes; the curves are given for corn-no (*1*), corn-min (*2*), corn (*3*), soy-no (*4*), soy-min (*5*), and soy (*6*).

the map provides a better illustration of the classes of land utilization rather than the classes of vegetation in these areas. For example, the area marked as corn-no does contain bare soil regions and remainders of last-year vegetation, while the young crops of corn cover only a small percentage of the area because of the early date of image recording (June 12).

The calculation of the corresponding pixel-averaged regions of the image spectra confirms the above-discussed facts. Thus, different classes of corn and soy have similar spectral characteristics (Figs. 2a and 2b), but there is a significant scatter of values within each region. The root-mean-square (RMS) deviations of the spectral coefficients of reflection $\sigma(\lambda)$ are shown in Fig. 2c.

The fragment shown in Fig. 1c was used for learning in studying various methods of classification of HS data. Using this fragment (for each method), we constructed the boundaries between the regions corresponding to the chosen classes in a multidimensional space of features formed from the values of brightness of different spectral channels. After that, image classification was performed (see Fig. 1a). The efficiency of the methods was estimated both in terms of discriminability of classes in the learning sample and in terms of the accuracy of classification of the entire image. In the latter case, the available division of the area into 58 classes was used as the data for objective monitoring.

The classification was performed by several most popular methods. Let us recall the essence of these methods.

The minimum distance (MinDist) method determines the cluster whose center is located at the minimum Euclidean distance from the classified pixel in the multidimensional space of features. The spectral angle mapper (SAM) method [7] determines the cluster with the minimum angle between the vectors directed from the origin of the coordinate system to the cluster center and to the classified pixel. The spectral information divergence (SID) method [8] is based on comparisons of the spectral curves of the reference and classified pixels based on calculating the Kullback–Leibler divergence.

Table 1. Efficiency of classification of the natural area image fragment

Classification methods	Accuracy, %	Processing time, s
MinDist	37.9	3
SAM	43.0	3
SID	51.0	3
SVM	85.7	120
MahDist	79.2	15
ML	99.9	15

Table 2. Efficiency of classification of the full 200-channel image of the natural area

Classification methods	Accuracy, %	Processing time, min
MinDist	33.9	6
SAM	40.5	7
SID	43.1	7
SVM	50.5	120
MahDist	56.3	22
ML	52.0	22

In contrast to these methods, the support vector machine (SVM) method [9] takes into account only the nearest pixels to the dividing boundary and forms it in such a way that the distance between the boundary and the classified object is maximized. If it is impossible to find the surface separating pixels of different classes, then the number of intersections of the classes is taken into account (on the basis of calculating the sum of penalties).

The methods of classification on the basis of the Mahalanobis distance (MahDist) and the maximum likelihood (ML) [10] involve determination of the distribution density of pixels of the learning sample in a chosen system of features. The Euclidean distance from the classified pixel to the cluster center normalized to the RMS deviation of the class in this direction is used as a proximity measure. It follows from here that the classification result in the case with identical distances from the classified pixel to two clusters is determined by data scatter. The main difference between these two methods is the fact that the MahDist method employs a collective (averaged over all data) covariation matrix, while the ML method employs individual covariation matrices for each class.

The efficiency of discriminating the learning data (image fragment) by different classification methods is illustrated in Table 1. Here the accuracy is understood as the fraction of correctly classified pixels. The last column shows the time necessary for the classification on a computer with an IntelCore2 Duo processor in order to illustrate the relative performance of the algorithms.

For a correct analysis of classification methods that take into account data covariation in the learning process and require large-volume samples, four classes containing less than 200 points (54, 26, 20, and 95 pixels) or $\sim 2\%$ of the marked domain of the fragment were eliminated from consideration. Thus, the efficiency was estimated on the basis of 12 classes. Moreover, three subclasses of corn and three subclasses of soy were additionally united in the full image, i.e., eight combined classes were used. The estimates of the efficiency of classification of the full image are presented in Table 2.

The analysis of the results obtained shows that the accuracy of discriminating learning data for methods that ignore the distribution of classes is rather low (MinDist, SAM, and SID); the SVM and MahDist methods are more effective; finally, the ML method ensures almost complete separation of the learning samples. If the full image is classified, however, the results are much worse, which testifies to significant spectral variability of different cultivated crops over the image. In this case, the SVM and ML methods yield worse results than the MahDist method. The reason is a small number of points in the learning samples. According to [10], for correct operation of statistical methods, the sample size for each class should exceed the number of spectral channels at least by an order of magnitude. In this aspect, the MahDist method is less vulnerable (which is actually observed) because it is based on a covariation matrix averaged over all classes.

The above-mentioned requirement can be satisfied in two ways: by increasing the sample size and by decreasing the number of features. If the first option is used, the computational complexity of the procedures

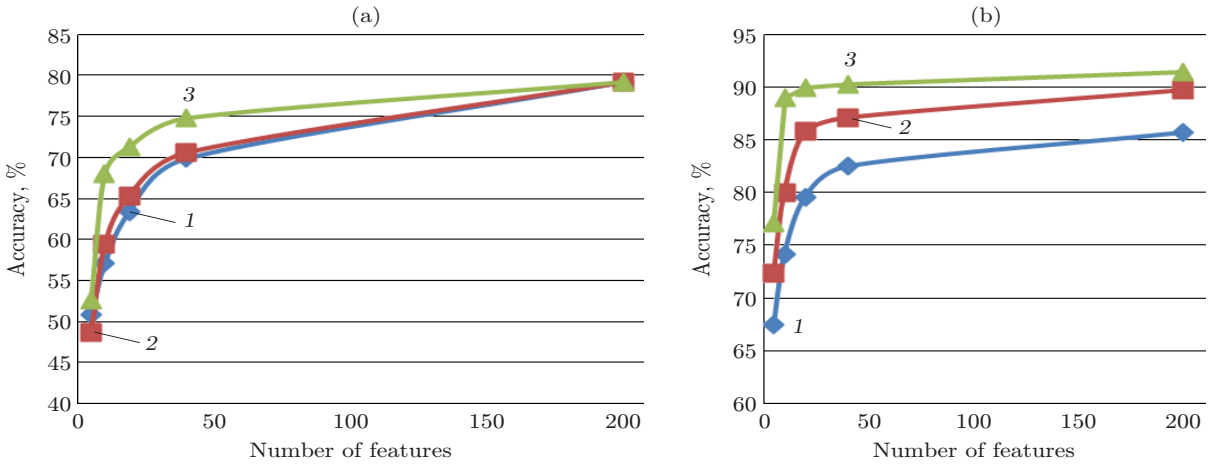

Fig. 3. Accuracy of classification of the learning samples with different numbers of features by the MahDist (a) and SVM (b) methods with regular decimation (*1*), PCA (*2*), and MNF (*3*).

increases; moreover, this option is not always possible in solving many applied problems. The second option is more practical, but an inappropriate choice of features may result in losing information determining the difference between the classes.

Let us consider the efficiency of classification of HS data as a function of the number of spectral channels used. Regular decimation, principal component analysis (PCA), and minimum noise fraction transform (MNF) [11] can be used for choosing the spectral channels.

Figure 3 shows the efficiency of separation of the learning samples as a function of the number of features for the SVM and MahDist classification methods. It is seen that the most effective systems of features are formed by using MNF. However, the use of more than 20 features for the chosen image is unreasonable.

Figure 4 shows the results of classification of the full image by the MinDist method and three most effective methods for different numbers of spectral features chosen by means of regular decimation, PCA, and MNF.

If the size of the space of features is reduced by simple regular decimation, the efficiency of the classification based on the Mahalanobis distance gradually decreases. For the SVM and MinDist methods, the efficiency weakly depends on the number of features (up to ten features). However, reduction of the number of features in the case of application of the ML method ensures not only a 20% increase in the classification efficiency, but also reduction of the computational complexity by more than two orders of magnitude. Thus, the best results were obtained by using only five PCA features. In this case, the most effective methods are the ML and SVM methods, as in the case of separation of the learning sample based on the full set of spectral features. Thus, in analyzing images with poorly discriminated types of vegetation (with close mean values for the separated classes and with a high scatter inside the class), it is reasonable to use HS data for choosing a moderate number of informative features with subsequent application of statistical classification methods. For the ML and MahDist methods, the efficiency is almost independent of the method of choosing the features compared in this study. For the SVM and MinDist methods, the most effective method is MNF.

CLASSIFICATION OF TYPES OF UNDERLYING SURFACE ON THE BASIS OF HYPERSPECTRAL IMAGES OF MAN-MADE AREAS

The efficiency of various methods of classification of RS data for man-made areas was studied by an example of processing of the HS image of the National Mall in the historical center of Washington, which was take from an aircraft. The image size was 307×1208 pixels, and there were 191 channels in the range of 0.4 to 2.4 μm (the channels in the range of water absorption (0.9–1.4 μm) were eliminated). Figure 5a shows the *RGB* composite of this image (channels 60, 27, and 17).

The image contains buildings with roofs of different types, highways, pathways, lawns in different conditions, individual trees and areas with dense forest, and water surface. The main classes can be easily

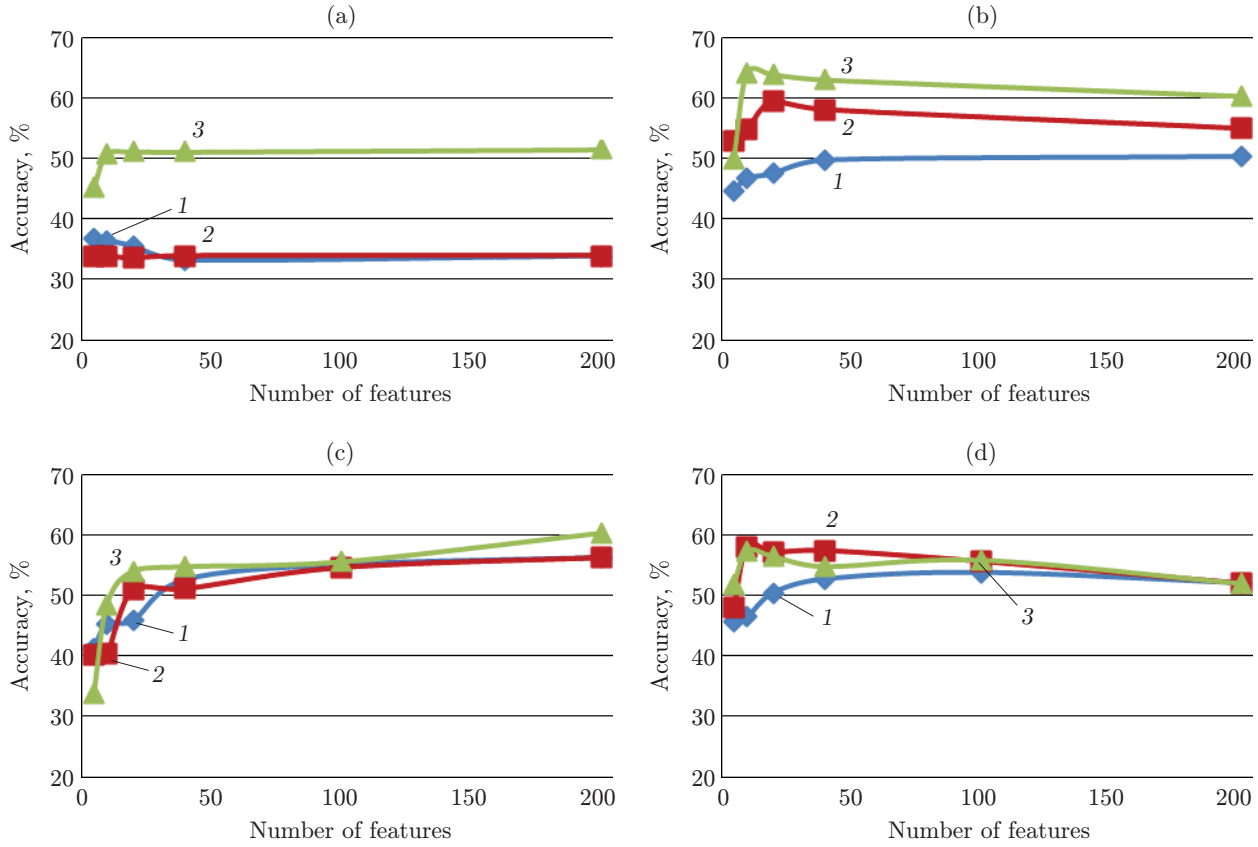

Fig. 4. Efficiency of classification of the natural area for different numbers of features by the MinDist (a), SVM (b), MahDist (c), and ML (d) methods. The notation of the curves is the same as that in Fig. 3.

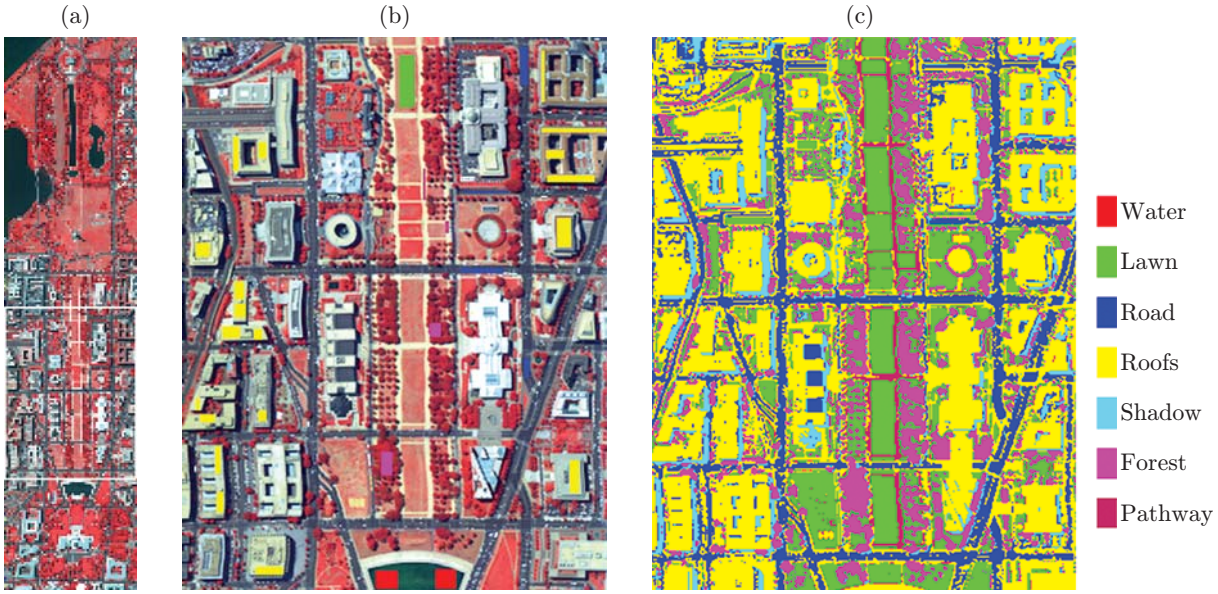

Fig. 5. Classification of the man-made areas: (a) *RGB* composite of the image; (b) *RGB* composite of the image fragment with reference sample regions; (c) result of classification.

Table 3. Accuracy of classification of the reference sample of the man-made area image for different numbers of channels

n	MinDist			SVM			MahDist			ML		
	Decimation	PCA	MNF	Decimation	PCA	MNF	Decimation	PCA	MNF	Decimation	PCA	MNF
2	93.3	89.7	69.1	94.8	92.8	67.2	93.5	91.1	68.6	95.6	94.8	73.7
5	95.0	94.5	85.9	95.4	95.5	89.3	84.9	92.8	78.1	96.5	93.8	87.2
10	95.2	94.5	90.6	93.7	95.6	96.2	83.0	85.9	87.2	94.7	89.3	90.1
40	94.5	94.6	87.7	93.9	96.4	93.0	85.6	85.6	84.5	88.5	84.9	86.4
100	94.6	94.6	82.3	93.8	94.4	88.0	89.8	83.7	84.9	79.1	77.1	80.3

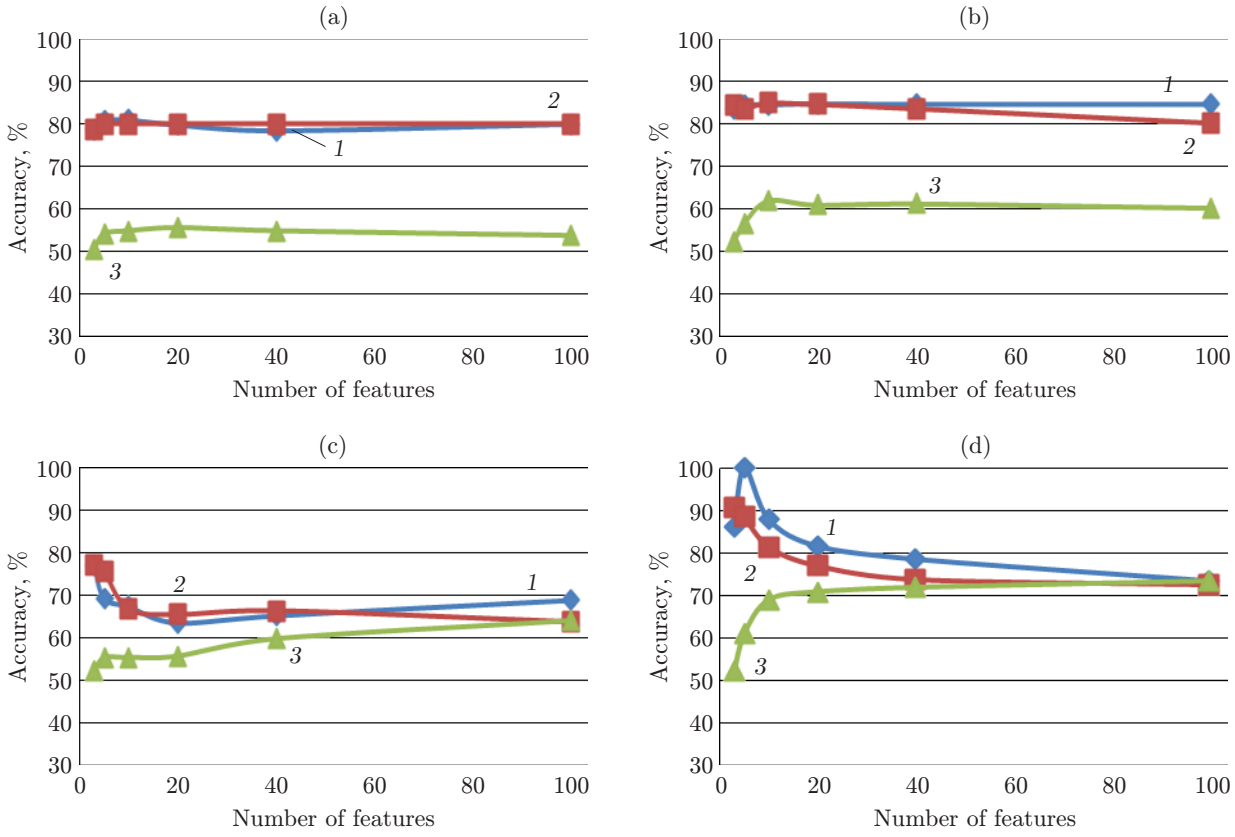

Fig. 6. Efficiency of classification of the man-made area for different numbers of features by the MinDist (a), SVM (b), MahDist (c), and ML (d) methods. The notation of the curves is the same as than in Fig. 3.

discriminated visually. Unfortunately, there is no map of ground-based observations for this area; nevertheless, a reference sample with separation into seven classes (water, lawn, road, roofs, shadow, forest, and pathway) is available. The locations of the reference sample areas for the central fragment are shown in Fig. 5b.

Image areas of 13 classes were used for classifier learning (five roof types, four lawn types, and forest, water surface, pathways, and shadow — one class for each area).

The estimates of the efficiency of different classification methods (with different methods used for reduction of the space of features) on the basis of the reference sample are listed in Table 3 (n is the number of channels used). It is seen that each of the examined methods ensures the classification accuracy close to 93–97% if the features are chosen correctly.

The best accuracy of the reference sample classification was provided by using the maximum likelihood method with five components chosen by means of regular decimation. The result of classification of the full image by this method was further used as a reference map of classes (Fig. 5c).

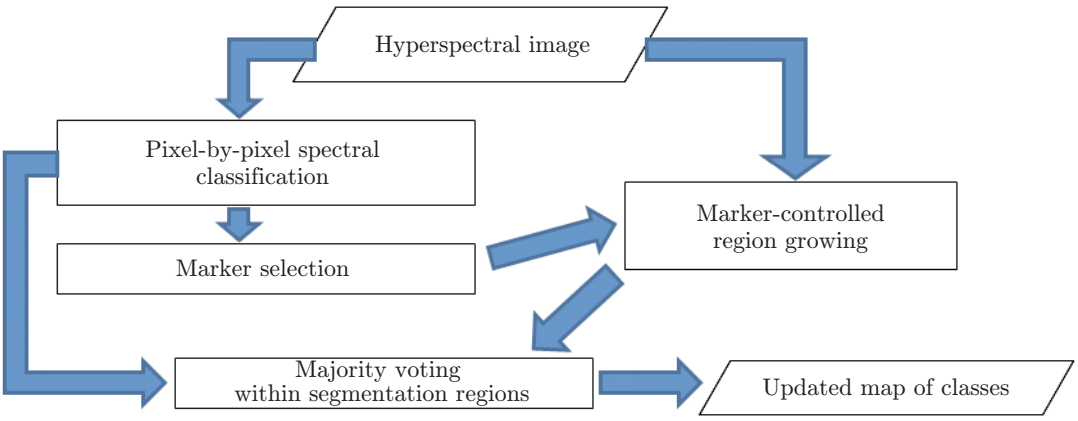

Fig. 7. Diagram of spectral-spatial classification.

Figure 6 shows the accuracy of the full image classification performed by the MinDist, SVM, ML, and MahDist methods with different numbers of spectral features chosen by three methods: regular decimation, PCA, and MNF. In contrast to results obtained by processing the natural area images, the classification obtained by the minimum distance method provides results similar to statistical methods in terms of accuracy.

The accuracy of classification of the considered image exhibits only a minor dependence on the number of features for all methods. The efficiency of choosing the features by means of regular decimation is close to the PCA efficiency as a whole. However, the choice of the features on the MNF basis is not so good because of incorrect evaluation of the noise levels in the channels: the noise is estimated on the basis of the level of the high-frequency component of the signal, while there are many significant differences in the brightness levels in the image of man-made areas. As a result, the first MNF components are low-informative.

CLASSIFICATION OF THE UNDERLYING SURFACE TYPES WITH COMBINED APPLICATION OF SPECTRAL AND SPATIAL FEATURES

In hyperspectral satellite observations of the Earth surface, the images of the majority of the objects are inhomogeneous despite the high probability that the neighboring pixels refer to the same class (a scene usually consists of homogeneous areas). First, the spectral composition of the pixel is a combination of the spectral characteristics of objects forming this pixel. Second, different surface areas are in different conditions. Some of them are directly illuminated by the Sun; other areas are shaded and are illuminated by reflected or scattered light. Moreover, their surfaces are aligned at different angles with respect to the light source. Significant distortions of the resultant spectrum are also induced by secondary reflection between the layers. The classification methods discussed above consider the processed data as a set of independent spectral measurements rather than images and ignore these conditions.

To solve the problem, it is reasonable to analyze the spectral and spatial features simultaneously [12]. In particular, methods that ensure preliminary segmentation of the image (its division into physically homogeneous areas) became very popular. After that, each segment (area) is considered as an individual object, and its classification is performed on the basis of choosing a dominating class (by means of majority voting) in the map of the pixel-by-pixel spectral classification. Such an approach has been intensely developed for the last decade, and it is the most promising method for processing data of hyperspectral observations of the Earth surface [13]. The diagram of one of the methods of spectral-spatial classification is shown in Fig. 7. The efficiency of this approach is basically determined by the correctness of preliminary segmentation of RS data, which is rather difficult in processing medium- and high-resolution images of natural areas (especially with vegetation classes difficult to discriminate).

As an alternative, it is possible to use an approach that avoids preliminary segmentation; instead, each pixel is classified on the basis of analyzing the pixel neighborhood [14]. In this case, in view of spatial coupling, the values of the neighboring pixels are considered as hypotheses for refining the value of the central pixel; moreover, there are various methods for choosing the most reliable hypothesis among all available hypotheses. In the simplest case, median smoothing can be used at the pre-processing stage. The accuracy of classification of the image fragment (see Fig. 1d) by the SVM method on the basis of linear (Lin) and radial basis functions (RBF) for ten principal components with preliminary median filtration of

Table 4. Accuracy of classification of the image fragment with different sizes of the pre-processing filter

Classification methods	No filter	3×3	5×5	7×7	9×9
SVM (Lin)	76.7	84.9	91.4	92.1	91.7
SVM (RBF)	70.8	83.6	88.0	89.7	89.8

the raw data is given in Table 4. In this case, learning was performed on the basis of a random sample of pixels of each class (15 values for classes represented by a small number of pixels in the image and 50 values for classes represented by a large number of pixels in the image).

It is seen that pre-processing ensured approximately 15% improvement of the fragment classification accuracy. For further increasing the efficiency of processing of hyperspectral and multispectral images, it seems promising to develop this approach in terms of combining the choice of the most informative features and the analysis of 3D neighborhoods of pixels (over the spectral and two spatial coordinates) for the purpose of determining the most probable hypothesis.

CONCLUSIONS

A comparative study of the efficiency of several methods of classification of hyperspectral images for discriminating different types of the underlying surface of natural and man-made areas was performed.

It was found that the best accuracy of the supervised classification of difficult-to-discriminate types of vegetation is reached by using methods that take into account the estimate of the probability density distribution function of reference classes (Mahalanobis distance and maximum likelihood methods). However, a large volume of learning samples is required in this case. If the volume of learning samples is insufficient, it is reasonable to use the support vector machine method.

It was demonstrated that it makes sense to use a moderate number of features (10–20) chosen by the MNF method for processing HS images of natural areas. Implementation of this method made it possible to increase the efficiency of classification of vegetation areas by 10% with a simultaneous decrease in the computational complexity by two orders of magnitude.

In processing of HS images of man-made areas, classification based on the minimum distance has almost the same accuracy as that obtained by statistical methods, and the best efficiency is provided by the maximum likelihood and support vector machine methods. The accuracy of classification by all methods exhibits a weak dependence on the number of features. The efficiency of features chosen by means of regular decimation and PCA is approximately identical as a whole. The MNF-based choice of features is less effective because of incorrect evaluation of the noise levels in the channels.

The advantage of HS data over multispectral data is ensured by the possibility of choosing a moderate number of informative features for each problem of monitoring of the Earth surface and near-Earth space with due allowance for the scene type. Nevertheless, the method of forming such a system is another severe problem.

The efficiency of image processing can be further improved by combined application of spatial and spectral features. Approaches based on segmentation of images and determination of the dominating class in the resultant segments and approaches combining the analysis of 3D neighborhoods of pixels with the choice of the most informative spectral features can be considered.

This work was supported by the Russian Foundation for Basic Research (Grant No. 13-07-12202) and by the Presidium of the Russian Academy of Sciences (Grant No. 15.3).

REFERENCES

1. V. N. Ostrikov, O. V. Plakhotnikov, and A. V. Kirienko, "Processing of Hyperspectral Data Obtained from Aviation and Space Carriers," Sovr. Probl. Dist. Zond. Zemli iz Kosmosa **10** (2), 243–251 (2013).
2. T. H. Chan, A. Ambikapathi, W. K. Ma, and C. Y. Chi, "Robust Affine Set Fitting and Fast Simplex Volume Max-Min for Hyperspectral Endmember Extraction," IEEE Trans. Geosci. Remote Sensing **51** (7), 3982–3997 (2013).

3. K. Cawse-Nicholson, S. B. Damelin, A. Robin, and M. Sears, "Determining the Intrinsic Dimension of a Hyperspectral Image Using Random Matrix Theory," IEEE Trans. Image Process. **22** (4), 1301–1310 (2013).
4. S. M. Borzov, A. O. Potaturkin, and O. I. Potaturkin, "Change Detection in Build-up Areas on the Basis of Structural Features of Satellite Images," Avtometriya **51** (4), 3–11 (2015) [Optoelectron., Instrum. Data Process. 51 (4), 321–328 (2015)].
5. S. M. Borzov and O. I. Potaturkin, "Classification of Vegetation Types on the Basis of Hyperspectral Data of Remote Sensing," Vestnik NGU, Ser. Inform. Tekhnol., No. 4, 13–22 (2014).
6. O. I. Potaturkin, S. M. Borzov, A. O. Potaturkin, and S. B. Uzilov, "Methods and Technologies of Processing of High-Resolution Multi- and Hyperspectral Data of Remote Sensing," Vych. Tekhnol. **18** (special issue), 53–60 (2013).
7. F. A. Kruse, A. B. Lefkoff, J. B. Boardman, et al., "The Spectral Image Processing System (SIPS) — Interactive Visualization and Analysis of Imaging Spectrometer Data," Remote Sensing of Environment **44** (2–3), 145–163 (1993).
8. H. Du, C. Chang, H. Ren, et al., "New Hyperspectral Discrimination Measure for Spectral Characterization," Opt. Eng. **43** (8), 1777–1786 (2004).
9. T. Joachims, "Making Large-Scale Support Vector Machine Learning Practical," in *Advances in Kernel Methods — Support Vector Learning*, Eds. by B. Schoelkopf, C. J. C. Burges, and A. J. Smola (MIT Press, Cambridge, USA, 1999, pp. 169–184).
10. J. A. Richards, *Remote Sensing Digital Image Analysis* (Springer-Verlag, Berlin, 2013, 494 pp.).
11. A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal," IEEE Trans. Geosci. Remote Sensing **26** (1), 65–74 (1988).
12. A. Plaza, J. A. Benediktsson, J. W. Boardman, et al., "Recent Advances in Techniques for Hyperspectral Image Processing," Remote Sensing of Environment **113** (Suppl. 1), 110–122 (2009).
13. V. G. Bondur, "Modern Approaches to Processing Large Fluxes of Hyperspectral and Multispectral Aerospace Information," Issled. Zemli iz Kosmosa, No. 1, 4–16 (2014).
14. C. Chen, W. Li, E. W. Tramel, et al., "Spectral-Spatial Preprocessing Using Multihypothesis Prediction for Noise-Robust Hyperspectral Image Classification," IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. **7** (4), 1047–1059 (2014).