Analysis and object markup of hyperspectral images for machine learning methods

V P Gromov[1], L I Lebedev[1] and V E Turlapov[1]

[1]National Research State University named after NI Lobachevsky, Gagarin Ave., 23, Nizhny Novgorod, Russia, 603950

e-mail: vadim.turlapov@itmm.unn.ru

Abstract. The development of the nominal sequence of steps for analyzing the HSI proposed by Landgrebe, which is necessary in the context of the appearance of reference signature libraries for environmental monitoring, is discussed. The approach is based on considering the HSI pixel as a signature that stores all spectral features of an object and its states, and the HSI as a whole - as a two-dimensional signature field. As a first step of the analysis, a procedure is proposed for detecting a linear dependence of signatures by the magnitude of the Pearson correlation coefficient. The main apparatus of analysis, as in Landgrebe sequence, is the method of principal component analysis, but it is no longer used to build classes and is applied to investigate the presence in the class of subclasses essential for the applied area. The experimental material includes such objects as water, swamps, soil, vegetation, concrete, pollution. Selection of object samples on the image is made by the user. From the studied images of HSI objects, a base of reference signatures for classes (subclasses) of objects is formed, which in turn can be used to automate HSI markup with the aim of applying machine learning methods to recognize HSI objects and their states.

1. Introduction

To successfully develop the application of hyperspectral images (HSI), a certain level of automation for their processing is required, which is demonstrated today by some examples of the successful application of deep learning methods. However, the success of deep learning methods is to a significant extent determined by the completeness of the research and formalization of HSI objects to be recognized and quantified in each applied task. All this makes it relevant to study classes of HSI objects in order to form an effective set of the HSI object descriptors and the knowledge bases (or libraries of the HSI objects and their states as the sets of signature instances) providing the stages of machine learning and using the artificial intelligence methods.

A necessary function of such a class library of HSI objects should be the function of automatic marking of objects on each training HSI for a new application task. In a number of applied tasks, the HSI objects have the states and properties that depend on season and time, which greatly complicates the task, and requires the creation of a tool capable of distinguishing small changes of the object being classified. One of the possible candidates for such a tool may be the principal component analysis (PCA) method, which is widely used today in HSI processing to minimize the dimension of the basis for the representation of the class being studied. In this paper, we discuss the HSI analysis apparatus,

which is capable of providing an adequate description of the HSI class hierarchy, as well as developing, implementing and demonstrating the performance of algorithms and software for these purposes.

2. Review of publications on the research topic

The content of this work is determined by a number of publications in recent years. However, we would like to start with the publication [1] of 20 years ago, in which the author, D.Landgrebe, summarizes the more than the 20-year period of his work in the study of Earth remote sensing (ERS). Landgrebe began research with Earth's multispectral imaging (MSI) and is one of the two authors of the well-known open-source MultiSpec software for processing ERS data, which was then extended to HSI processing and remains popular in the world to the present. This software complex is used in Russia, for example, in the GIS Lab at the Moscow State University. The latest version of the complex is available on the website of its authors' university (https://engineering.purdue.edu/~biehl/MultiSpec/).

In [1], Landgrebe establishes, as a result of the MSI and HSI study, the priority of the spectral features of an object over features associated with the neighborhood pixels, as they do not require the study of such a neighborhood. Here he formulated the requirements for the optimal definiteness of the class, and therefore for the tool (software package) with which this certainty is achieved. The Landgrebe-optimal definiteness of the class requires that the class definitions (attributes) be:

- *Exhaustive*. There must be a logical class to assign each pixel in the scene to.
- *Separable*. The classes must be separable to an adequate degree in terms of the spectral
- features available.
- *Of informational value*. The classes must be ones that meet the users needs.

In the same paper, Landgrebe proposed a nominal sequence of steps for analyzing hyperspectral images from 4 steps, which, if necessary, can be repeated iteratively:

1) Classes Delineation and Training Samples Designation.
2) Feature Extraction and Preliminary Classification.
3) Final Class Description Determination.
4) Classification.

At the first step, on the basis of visual perception and selection of samples of homogeneous objects belonging to a class (subclass), the features of each class are formed, maximally utilizing spectral features.

The second step practice assumes that the training sets defined for the class may be small. From which it is concluded that it is useful to exclude features that are ineffective for a particular set of classes, and at the same time reduce the dimension of features without losing information. A preliminary classification is a criterion for the success of feature extraction and can guide the expansion of both the composition of the features and the training set. As the main procedure for the formation of a set of features, the PCA method was used as one of the main ways to reduce the data dimension by losing the least amount of information. The nominal sequence of steps proposed by Landgrebe for analyzing HSI basically remains relevant to the present.

However, the most important result of many years of Landgrebe's research in our opinion is the conclusion that, due to the influence of the atmosphere, the best HSI accuracy that we can count on is from 2% to 5% (not the accuracy of the HSI fixation equipment, but the accuracy compliance of the recorded reflection of the earth's surface with the real reflection), which cannot be compared even with the measurement level of 0.1%, implied by 10-bit data. Unfortunately, Landgrebe does not specify the rules for estimating these 2-5%, but it gives hope that in many cases it would be possible to confine one-byte representation of HSI instead of two-byte, and speaks about the usefulness of having a tool for evaluating the noise component along with other HSI analysis tools. Such a tool can be based, for example, on the empirical mode decomposition method [2].

At the same time, recent publications are turning to the construction of libraries of reference spectral signatures HSI, which provide automation of the classification of HSI objects in applied problems. Thus, the publication [3] of 2014 emphasizes the need to create a database of reference spectral signatures of HSI for objects with different types of pollution. This is planned to be used to

create classification maps of polluted sites and environmental monitoring according to HSI data. The classification "with training" uses a database of pollutant signatures, which greatly speeds up and simplifies further work. A test was conducted to identify areas with the presence of oil pollution. In the publication [4] of 2017, the authors supplemented the method with a two-stage partitioning into clusters within each class: at the first stage, the algorithm clusterizes the regions of interest, maximizing the likelihood function for a mixture of Gaussian distributions; in the second, it performs clustering using Mahalanobis distances to cluster centers as a feature space.

Detection of anomalies in hyperspectral images is very similar to the task of detecting various states of an object. In the publication [5] of 2014, a number of new algorithms for detecting anomalies on HSI were proposed: spectral mismatch algorithms (spatial-spectral and adaptive spectral with modification) and an algorithm for probabilistic anomaly detection. Algorithms of spectral mismatch are used to determine the fact of "abnormality" mismatch of the statistics of the inner window data with the outer window data. The probabilistic anomaly detection algorithm uses a re-quantization of the spectral data of the signatures to a significantly smaller number of levels (for example, to 4 levels). A comparison of the proposed algorithms with widely used variants of the RXD algorithm (global and local), which showed a convincing superiority of the proposed algorithms.

The article [6] of 2016 is devoted to the development of information technology for the early recognition of crop types based on a set of MODIS satellite images in the region's territories. Early recognition is carried out in the first half of the planting season and is characterized by a lack of ground and space data to adjust the recognition algorithms. A two-step solution is proposed. At the first stage, models and databases of time series of agricultural crops and agro-seasons are formed according to data for the past years, at the second stage, data of the current year are added to detect the initial state (agro-season) for crop monitoring objects and a training sample is generated for the entire current season based on the models and DB.

In the article [7] of 2016, a comparative analysis of algorithms for recognition of forest cover objects and some other natural and man-made objects on hyperspectral aerospace images is given. For example, in one of the tasks, there are seven main classes: water surface; road surface; soils; meadow vegetation; Pine; Birch tree; aspen. For forest cover, the species composition and age were studied. The spectral dimension of the problem is limited by the PCA method to four principal components. Each class is further characterized by a reference pattern and parameters of the Gaussian distribution. The best of the participating methods turned out to be the Bayes classifier (BC) based on Gaussian mixtures, while the worst one, the metric classifier. The method of K-weighted neighbors showed the accuracy commensurate with the accuracy of nonlinear optimal classifiers and can be used along with them for the recognition of forest cover of different species composition and age. To build a class standard in the considered problems, 4 principal components were sufficient.

In [8], 2018, the use of imaging hyperspectrometers for precision farming in a field experiment to determine the humidity index of soil covered with vegetation is considered. It is shown that the humidity index at each point of the hyperspectral image, where greens are present, is determined by the ratio of $I_{0.9}$ to $I_{0.97}$, where $I_{0.9}$ is the spectral brightness at a wavelength of 0.9 μm, and $I_{0.97}$ is at a wavelength of 0, 97 microns.

In the publication [9] of 2018, focused on the use of high-resolution multispectral images (1-5m), an information technology for processing ERS data for estimating plant habitats, based on the calculation of local signs, was proposed and investigated. The basic stages of the proposed technology are a super-pixel segmentation (the selection of connected areas of local homogeneity in the multi-channel image), the calculation of the segments' parameters, clustering of the segments, the estimates of plant habitats. The stages of calculating parameters and superpixel segmentation are combined into one process. This process allows obtaining both splitting into locally homogeneous sections and their parameters in one pass through the image, reducing the computational complexity of the algorithm and the complexity of producing the training set by the operator. The parameters include minimum and maximum brightness values of superpixels, the mathematical expectation of their brightness for each channel, and superpixel's areas. The data from ground studies of 4 elementary vegetation classes are used: 1) sparse grass vegetation; 2) low-grass vegetation; 3) tall grass; 4) trees and shrubs. These data

are used to form the parameters of classes in the synthesis of model images, as well as for training sets construction and classifying ERS images of plant habitats.

Publication [10], also in 2018, demonstrates the potential of hyperspectral and multispectral data for monitoring and assessing land cover across the Russian Federation region, and highlights the development of a spectral library for land cover classes with the construction of a standardized classification system.

The study used 1 hyperspectrometer (Hyperon) and 2 multi-spectrometers (ALI and OLI). For all three images, the following were applied: 1) the k-nearest-neighbor algorithm (KNN) as supervized learning method for a classification; 2) the method of migrating means clustering (MMC) as unsupervized learning method. Training set data are collected on the basis of field works and supported by the reference to topographic maps. Initially, 27 subclasses were allocated along with training sites for them, which were then combined into 8 classes: 1. Water; 2. Vegetation; 3. Settlement; 4. Wetlands; 5. Naked land; 6. Transportation; 7. Bare rocks; 8. Sand dunes. Field measurements were performed on a FieldSpec 3 ASD portable spectroradiometer (range: VNIR, visible, near-infrared, and SWIR, shortwave infrared, between 350 and 2500 nm; accuracy: 1 nm) and GPS (Trimble PRO XRS) with a reference accuracy of 1 m. 7-8 measurements were made for each class, in random points of 60 zones of each image. The classification accuracy assessment experiment confirmed the expected that supervised classification is better than unsupervised classification for all three images (Hyperion, ALI, and OLI), and hyperspectral images are more informative than multispectral ones. The main thing, however, is that the automatic classification via the collected reference samples was quite effective for both HSI and MSI.

All considered publications of recent years ultimately confirm the relevance of work on creating databases of standards of HSI objects and standards of states of these objects for different tasks of territorial management, as well as the relevance of developing the HSI analysis methods that ensure the information content of the standards. The same applies to multispectral images, which are easier, due to greater availability, to be used in the practice of managing territories, but for them more often you have to use, along with spectral, to textural attributes of the classification objects' states.

3. Methods of analysis and markup of HSI objects

The survey noted that the nominal sequence of steps for analyzing HSI proposed by Landgrebe [1] largely remains relevant to the present. This is true, however, there are features in the described scheme that lead to the loss of essential HSI information. Actually, the use of PCA to significantly reduce the data dimension initially implies a loss of information. The main question is what information we agree to lose. It is also important that the result of PCA depends on the content of a particular image. In this regard, we would like to clarify the place of the PCA method in the nominal sequence of steps for analyzing HSI in cases requiring really "fine" analysis. Consider the first two steps of the Landgrebe nominal HSI analysis sequence.

The first step is "Classes Delineation and Training Samples Designation". Recall that Landgrebe-optimal classes should be: exhaustive (can be specified to each HSI pixel); separable; informationally valued. It is recommended, if possible, to build a set of classes divided by spectral features. At this step, the number of classes must be predetermined and sets of training samples for each class must be selected. The analysis process always begins with a visual review of HSI. It usually uses 3-color coloration of the visible and infrared ranges, or some other combination of channels, which are useful for preliminary visual classification to be separated by color. This operation is useful in any cases and tasks, and of course must be present as a mandatory HSI analysis tool. In the same way, we should welcome the desire to maximally ensure the spectral separability of classes, which we will be able to check by the separability in trying to mark class samples automatically via spectral features.

The second step: "Feature Extraction and Preliminary Classification". At this step, it is intended to significantly reduce the dimension of the HSI space by applying the PCA method or a method similar in capabilities and choosing the number of principal components (m) corresponding to the specified accuracy. After this, we can verify that the class samples are separable in the space of m principal components from the point of view of linear discriminant analysis (LDA). This can be monitored visually by observing clusters of color-labeled samples on the coordinate planes of the principal

component basis. If any classes are inseparable, additional components are introduced into the basis. Checking whether the list of classes is sufficiently comprehensive is carried out by checking for the presence of pixels that are not assigned to any of the classes. If there are any, unclassified pixels are marked manually with labels of existing classes, to which they should belong, or labels of additional classes. An example of markup using HSI analysis tools included in the Terra Plus GIS [11] is shown in Figure 1.
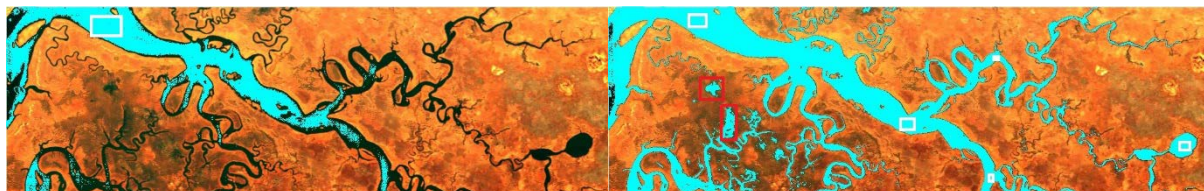

Figure 1. The choice of one sample (1440 pixels, window with white frame) for marking the river on HSI and the area of the river (24935 pixels) are automatically colored in cyan in response to this choice (on the left); selection of 5 samples (4191 pixels), which ensured complete filling of the main watercourse, all tributaries and lakes (95,935 pixels, on the right). In both cases, only one recognition basis is constructed. Cases of over-marking zones (wetlands) are marked with red outlines.

The data source for Figure 1 is HSI Suwannee_0609-1331_ref.dat: wetlands of the Suwannee River Basin in Florida, USA, Spectir spectrometer, the number of channels $n = 360$.

The efficiency of the procedure of the second step has been repeatedly tested by practice. However, we consider it necessary to introduce some development into it. First, we insist on the complete preservation of the reference HSI pixel signatures, regardless of the number of channels that are necessary for the current classification in a particular application (perhaps with the exception of completely noisy channels). This will allow us to build a classification, which is not so much dependent on the conditions of a specific task and a specific HSI, but is reusable and is supported by reference (standard) signatures, which were confirmed in field tests, as in [3], [4], [6-10].

3.1. The role of signatures, their linear dependence and the values of standard deviations in the classification procedure
In [10], an example of classification of 27 subclasses of 8 dissimilar classes is shown, showing the similarity of the signatures of all 27 subclasses in the range of 82-96 channels of the Hyperion spectrometer. This, on practical material, indicates the importance of detecting the fact of a linear relationship between signatures and using the scale factor of this relationship as a feature of a class and a subclass. In our practice, the presence of a linear dependence of class signatures is regular. Figure 2 shows an example of such pixel signatures for points randomly selected on the river area and on the land area in HSI Suwannee_0609-1331_ref.dat (see also figure 1).
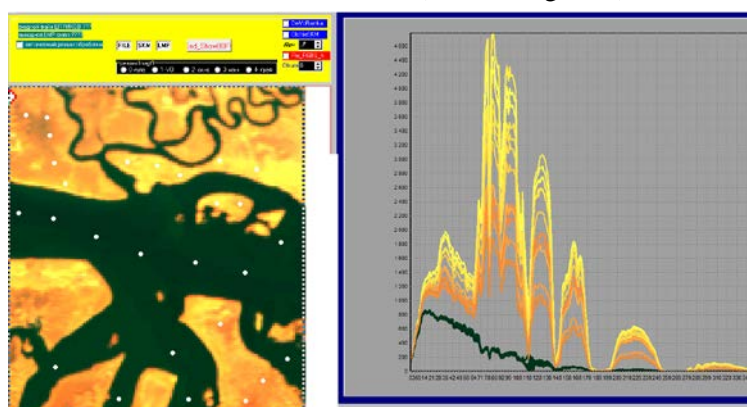

Figure 2. 28 pixels of HSI Suwannee for different parts of the river and land (on the left, white dots) and their signatures, demonstrating a linear relationship within the class (the color of the graph curve corresponds to the color of the marked pixel in the original image).

Covariance (1)

$$\text{cov}(\mathbf{x},\mathbf{y}) = E[(\mathbf{x} - E\mathbf{x})(\mathbf{y} - E\mathbf{y})], \tag{1}$$

and the Pearson correlation coefficient (2)

$$r(\mathbf{x},\mathbf{y}) = \text{cov}(\mathbf{x},\mathbf{y})/(\sigma\mathbf{x} \cdot \sigma\mathbf{y}) = \text{cov}(\mathbf{x},\mathbf{y})/(\sqrt{D\mathbf{x}} \cdot \sqrt{D\mathbf{y}}) , \tag{2}$$

where E is the moment of the first order, D is the variance, σ is the standard deviation, can be used as a detector of the linear relationship between the $\mathbf{x}$ and $\mathbf{y}$ signature vectors of two HSI pixels, with the number of channels n.

To detect a linear relationship with reference signatures, we need the unbiased values of the signatures of the current HSI, so we assume in (1) $E\mathbf{x} = E\mathbf{y} = 0$, after which we have:

$$\text{cov}(\mathbf{x},\mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})/n ; \quad r(\mathbf{x},\mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})/(|\mathbf{x}| \cdot |\mathbf{y}|), \tag{3}$$

where $|\mathbf{x}|$ and $|\mathbf{y}|$ are modules of vectors $\mathbf{x}$ and $\mathbf{y}$.

The Pearson correlation coefficient turns into a normalized scalar product of signature vectors and seems to us to be a more interesting detector, since arccos (r) can be used as a measure of the angular deviation of the signature $\mathbf{x}$ from the signature $\mathbf{y}$. The values of standard deviations $\sigma\mathbf{x}, \sigma\mathbf{y}$ (4)

$$\sigma\mathbf{x} = \sqrt{D\mathbf{x}} = \sqrt{E[(\mathbf{x} - E\mathbf{x})^2]}, \tag{4}$$

their ratios

$$s = \sigma\mathbf{x}/\sigma\mathbf{y} , \tag{5}$$

or quantized values will be used as the feature of a class (subclass).

With this approach, the reference (standard) class signature (or time series of standards) becomes the real basis for the classification, and changes in this role the PCA method of the Landgrebe nominal sequence, which depends on the content of a particular hyperspectral image.

Thus, the procedure for detecting linear dependencies between signatures by the value of the Pearson correlation coefficient becomes a necessary procedure of the second step and applied before the PCA. To the same extent, this is also necessary in the first step of the Landgrebe nominal sequence.

3.2. The role of the PCA method in the classification procedure

However, the PCA method does not lose its significance as an effective HSI analysis tool. We consider it necessary to apply it to a local homogeneous sample, which determines the local standard of the object class on the sample area. If sample homogeneity is provided by a qualified expert, then we can set the local standard of the class equal to the average of the sample signature, which corresponds to the PCA procedure. A ready-made standard can also be used as a local standard. Further, the PCA procedure provides orthogonality and statistical independence for the components of the decomposition. This is expressed in the diagonality of the covariance matrix, with the result that the expression of the Mahalanobis distance between two signatures takes the form (6):

$$d(\mathbf{x},\mathbf{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2/\sigma_i^2} , \tag{6}$$

where σ_i is standard deviation of x_i from y_i in the sample, d is the normalized Euclidean distance.

As a result, each of the principal components gets in σ_i value its own scale to solve the problem of linear discriminant analysis (LDA).

The main task of applying the PCA method on a homogeneous sample is, in our opinion, a "fine" analysis: identifying subclasses of local classification that have an applied meaning. For clarity, when assessing the applied value of selected subclasses, it is advisable to use standard techniques: 1) visualization of different subclasses with different color and brightness; 2) visualization of the projections of the signature samples on the coordinate planes of the principal components for evaluating their clustering and LDA capabilities (Figure 3).
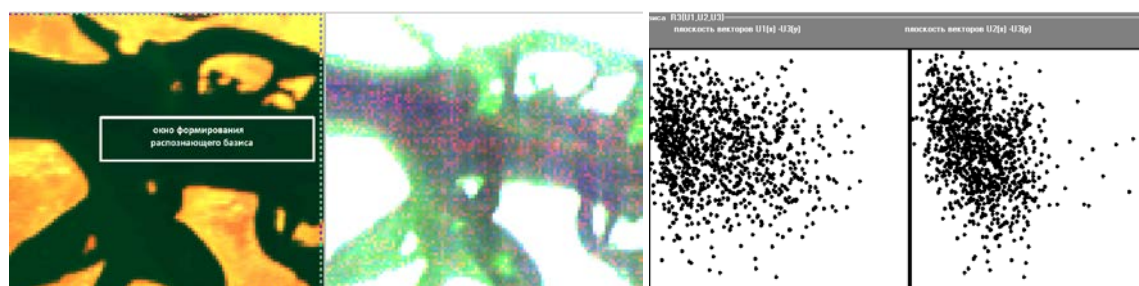
Figure 3. Color-luminance illustration of the discriminant properties of a three-component basis constructed using the PCA method (middle) for a sample of river pixels with the window shown on the left. The dark area of the color-brightness illustration corresponds to the signature of the local class and the deeper part of the river, the green - to the subclass of the shallow part of the river. Projections of 1155 sample pixels on 2 basis planes (right).

Depending on the degree of homogeneity of the sample, PCA may give different results. Figure 3 shows an example of the analysis of HSI Suwannee_0609-1331_ref.dat, which represents a fairly frequent situation when choosing an area that is completely homogeneous in appearance (the rectangle in the left image). The dark area of the color-brightness illustration (in the second image from the left) corresponds to the signature of the local class and the deeper part of the river, the green one to the subclass of the shallow part of the river, the cyan is the shallowest part; the red component is the noise. The projections of 1155 pixels of the sample on 2 base planes (on the right) are presented in two images on the right and, by visual impression, are inseparable by the LDA method before PCA is applied. We can also face the case when PCA on a uniform sample can only give noise components, which, together with an estimate of their standard deviations, is also very useful in determining the threshold values for classification error.

Figures 4-5 show an example of analyzing a fragment of the water surface of the San Francisco Bay at HSI Moffett Field. This bay is known as the first wetland sanctuary within the United States urban area (Don Edwards San Francisco Bay National Wildlife Refuge). It is characterized by a large variety of classes and subclasses of wetland areas, visited by 280 bird species. In particular, in Figure 4 for the study a fragment of heterogeneous tidal pollution is highlighted for the study. According to this sample, water areas belonging to the class specified by the fragment are distinguished.

Figure 4. HSI Moffett Field with automatic classification by training window (shown with white frame on the left side of the image).

In Figure 5 (left), the selected fragment and the surrounding area of interest are shown with zoom to increase the accuracy of selection (Zoom tool). The right shows the color reproduction of the result of the PCA method for the 3 principal components. Blue (in the middle) and green (on edges) colors correspond to the bulk of pollution. The open water corresponds to the red color (a small part of the pixels in the selection window). The right side shows the projections of 2030 pixels of the selection

window on 3 coordinate planes of the principal components: classification clusters are noticeable, and are quite distinctly separated by color.
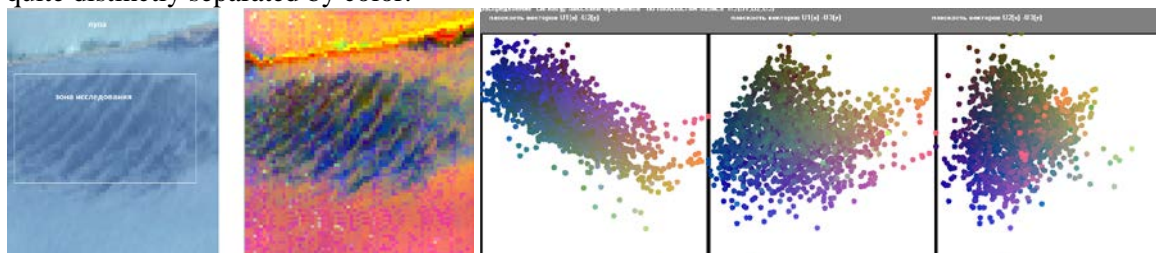

Figure 5. Fragment of the bay with a selection window (on the left, white frame); visualization of orthogonal components of the basis of the classification constructed by the PCA method with different colors (blue and green colors correspond to pollution, red to free water of the bay, the road along the dam separating the bay and the salt lake above the road is yellow). The right side shows the pixel projections of the selection window on the 3 coordinate planes of the principal components: classification clusters are noticeable, which are quite distinctly separated by color.

4. Discussion of the results

Feature extraction is one of the most important operations in the classification of hyperspectral images. This operation must be performed before the classification and best of all should be based on the analysis of the classified HSI. In the case of using standard signatures from the database, or when such standards are developed on the HSI under study, the first necessary analysis step is the procedure for detecting the linear dependence between signatures by the Pearson correlation coefficient (see Figure 2). This procedure is an adequate tool for cluster analysis of HSI by the criterion of the angular deviation of signatures already on the first step of the nominal HSI analysis sequence.

So, in the second step of the Landgrebe sequence, the procedure for detecting a linear dependence of signatures precedes the use of a fine analysis of a homogeneous area via PCA. In the experiments shown above, the PCA method was used as a method for studying the existence of subclasses in HSI object classes, which have been defined by homogeneous image fragments. The use of PCA in conjunction with the visual and quantitative assessment of the principal components allows you to see whether the constructed principal components have real application value or they reflect the noise component of the image, as in Figures 3 and 5.

Components with small standard deviations can have useful information for distinguishing classes in the classification process and for detecting different states of an object of the same class in space and time, which can make even the smallest values of standard deviations noticeably significant.

5. Conclusion

The paper discusses the development of methods for recognizing objects of hyperspectral images (HSI) in the direction of creating and using banks of reference (standard) signatures for objects of applied tasks of environmental monitoring. This direction needs modification of the Landgrebe nominal sequence of steps for analyzing HSI.

In this modification, the HSI pixel is considered as an n-dimensional vector of the signature, which stores all spectral features of both the object and its state. And HSI is generally viewed as a two-dimensional field of signatures. For these reasons, the modification is built on per-signature analysis of HSI rather then on per-channel analysis.

The modification of the nominal sequence of HSI analysis according to Landgrebe also consists in the fact that as a first step of the analysis a procedure is proposed for detecting a linear dependence of signatures by the magnitude of the Pearson correlation coefficient. This procedure is also the basis of HSI clustering performed in the first step. The primary feature of a class and the clustering criterion is the angular proximity to the standard signature of the class. The secondary feature of a class (subclass) is the value of the standard deviation of the class standard signature **y**, and the ratio $s = \sigma\mathbf{x} / \sigma\mathbf{y}$ for the class instances.

In the second step of the Landgrebe-nominal sequence HSI analysis, the PCA method remains the main tool along with clustering. However, it is no longer used to build classes, and is applied to investigate the presence in the class of subclasses essential for the applied area. This turns PCA into a multivariate differential analysis tool.

The content is illustrated by results of analysis tools for the study of homogeneous fragments of hyperspectral images that are contained in the GIS Terra Plus [11]. Experimental material includes such objects as water, swamps, and wetlands, soil, vegetation, concrete, salt, tidal pollutions. Selection of object samples on the image is made by the user. From the studied images of objects of hyperspectral images in the GIS Terra Plus, a base of standard signatures for classes (subclasses) of objects is formed, which, in turn, can be used to automate HSI markup with the aim of using machine learning methods to recognize HSI objects and their states.

6. References

[1] Landgrebe D 1999 Information extraction principles and methods for multispectral and hyperspectral image data *Information Processing for Remote Sensing* (River Edge, NJ: World Scientific Publishing Company) **1** 3-38

[2] Pakhomov P A, Borusyak A V and Turlapov V E 2018 Investigation of hyperspectral image pixel signatures by the empirical mode decomposition method *CEUR Workshop Proceedings* **2210** 352-364

[3] Ramenskaya E V, Guryanova A O, Mandra A G and Ermakov V V 2014 Obtaining of support spectral signals during hyperspectral survey *Ecology and Industry of Russia* **10** 44-47 (in Russian)

[4] Ramenskaya E V, Kuznetsov M P, Ermakov V V, Barkova O R and Bran A A 2017 Hyperspectral image classification using cluster data structure *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* **14(7)** 9-19 (in Russian)

[5] Denisova A Yu and Myasnikov V V 2014 Anomaly detection for hyperspectral imaginary *Computer Optics* **38(2)** 287-296

[6] Vorobiova N S, Sergeyev V V and Chernov A V 2016 Information technology of early crop identification by using satellite images *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938

[7] Kozoderov V V, Kondranin T V and Dmitriev E V 2016 Comparative analysis of recognition algorithms for forest cover objects on hyperspectral aerospace images *Issledovanie Zemli iz kosmosa* **6** 45-55 (in Russian)

[8] Podlipnov V V, Shchedrin V N, Babichev A N, Vasilyev S M and Blank V A 2018 Experimental determination of soil moisture on hyperspectral images *Computer Optics* **42(5)** 877-884 DOI: 10.18287/2412-6179-2017-42-5-877-884

[9] Varlamova A A, Denisova A Y and Sergeev V V 2018 Earth remote sensing data processing for obtaining vegetation types maps *Computer Optics* **42(5)** 864-876 DOI: 10.18287/2412-6179-2018-42-5-864-876

[10] Boori M S, Paringer R, Choudhary K and Kupriyanov A 2018 Comparison of hyperspectral and multi-spectral imagery to building a spectral library and land cover classification performance *Computer Optics* **42(6)** 1035-1045 DOI: 10.18287/2412-6179-2018-42-6-1035-1045

[11] Lebedev L I, Yasakov Yu V, Utesheva T Sh, Gromov V P, Borusjak A V and Turlapov V E 2019 Complex analysis and monitoring of the state of the environment based on Earth sensing data *Computer Optics* **43** (in press)

Acknowledgments