SEGMENTATION OF AEROSPACE IMAGES BY A NON-STANDARD APPROACH USING INFORMATIVE TEXTURAL FEATURES

The article presents an analysis of a non-standard approach to the segmentation of textural areas in aerospace images. The question of the applicability of sets of textural features for the analysis of experimental data is being investigated to identify characteristic areas on aerospace images that in the future it will be possible to identify types of crops, weeds, diseases, and pests. The selection of suitable algorithms was carried out and appropriate software tools were created on Matlab 2021a and in the software package for statistical analysis Statistica 12.

The main way to extract information is to decrypt images, which are the main carrier of information about the underlying surface. The main tasks of texture area analysis include selection and formation of features describing textural differences; selection and segmentation of textural areas; classification of textural areas; identification of an object by texture.

To solve the tasks, spectral brightness coefficient (SBC), Normalized Difference Vegetation Index (NDVI), textural features of various crops and weeds. Much attention will be paid to the development of software tools that allow the selection of features describing textural differences for the segmentation of textural areas into subdomains. That is the question of the applicability of sets of textural features and other parameters for the analysis of experimental data to identify types of soils and soils, vegetation types, humidity, crop damage in aerospace images will be resolved.

This approach is universal and has great potential for identifying objects using image clustering. To identify the boundaries of areas with different properties of the image under study, images of the same surface area taken at different times are considered

Keywords: image processing, satellite images, textural features, SBC, NDVI, clustering, agricultural crops, weeds, pests

Akbota Yerzhanova*
Doctoral Student, Teacher*
S. Seifullin Kazakh Agrotechnical University
Pobeda ave., 62, Nur-Sultan, Republic of Kazakhstan, 010000
Gulzira Abdikerimova
Corresponding author
PhD*
E-mail:gulzira1981@mail.ru
Zhanar Alimova
Toraighyrov University
Lomova str., 64, Pavlodar, Republic of Kazakhstan, 140008
Assylzat Slanbekova
Master of Technical Sciences**
Aigul Tungatarova
Candidate of Pedagogical Sciences***
Raikhan Muratkhan
PhD **
Gaukhar Borankulova
Candidate of Technical Sciences, Associate Professor,
Head of Department***
Gulzat Zhunussova
Master's Degree of Pedagogical Sciences
Department of Informatics And Biostatistics
Karaganda Medical University
Gogol str., 40, Karaganda, Republic of Kazakhstan
*Department of Information Systems
L. N. Gumilyov Eurasian National University
Satpayev str., 2, Nur-Sultan, Republic of Kazakhstan, 010008
**Department of Applied Mathematics and Computer Science
Karaganda Buketov University
Universitetskaya str., 28, Karaganda, Republic of Kazakhstan, 100024
***Department of Information Systems
M. Kh. Dulaty Taraz Regional University
Tole bi str., 40, Taraz, Republic of Kazakhstan, 080000

1. Introduction

Operations for recognizing real objects in images are usually used to create conditions that increase the efficiency and quality of distinguishing and recognizing objects that are searched for or studied by processing images. Currently, the problem of building algorithms that divide images from different sources into clusters (for solving classification problems) often arises in such applied areas as, for example, the analysis of photos of the Earth's surface, the detection of defects, scratches on images, the study of images of starry sky regions. In addition, the method of image processing varies depending on the research reports, for example, the most informative is to separate fragments and increase them, increase the intensity contrast, improve the image quality, and so on. When solving such problems, image clustering algorithms based on certain criteria must have such properties as high classification accuracy and high speed. Accurate and fast video classification algorithms can simplify the task of a scientist or specialist in a particular field. Therefore, the relevance of the topic is determined by the study of textural features during a certain growing season, the extraction of

Information features from them, the analysis of the Ward Method, and the development of automated algorithms.

There are built-in methods for clustering images, such as the well-known k-means method. In the study, let's found that the k-means method showed different results for a single image because the data array was large. Therefore, before using the k-means method, let's use mathematical processing of big data (creating a vector of textual labels) and obtained accurate data when clustering.

2. Literature review and problem statement

The texture is understood as "the spatial organization of elements within a certain area of the surface". It also explains that this organization is due to a certain statistical distribution of the intensity of gray tones or tones of different colors. A plot can be considered textural if the number of intensity differences or color changes noted on it is large enough.

In texture is called "a certain way organized area of the surface".

A texture is a certain part of an image, one that has homogeneous statistical properties. This means that each texture of this class can be outlined using the corresponding characteristic common to all textures of this class. Such properties are called textural features.

The article [1] provides an analysis of the nonlinear segmentation of textured regions on microphotography in the processing of systems. Here are some questions about the use of experimental data from the experimental data collection on microphones, which can be supposed to be related to porosity, chemical reaction, and so on.

In [2], the indicators of tests on the collective use of spectral and textural features for the systematization of vegetation cover on aviation hyperspectral images are considered. Images were obtained by the experimental sensor Texture features can improve the classification quality only using a small number of bands. But to boost the informativeness it should be dealt with an individual for each scene set of texture features.

In [3], a method of computer diagnostics of osteoporotic diseases of the musculoskeletal system is proposed and investigated using a textural examination of the X-ray photo of the leg neck. The problem of selecting textural features that provide the applicable definition quality is solved. The operation of the method is investigated on a set of real X-ray images with popular diagnoses. The disadvantages of this approach are the dependence of many features on changes in the overall brightness of the image, the dependence on the choice of a polygonal area of interest by a specialist, the inability to take into account the age, gender and other features of the patient, which densitometry takes into account.

In [4], uncorrelated textural features of Haralick were used for the segmentation of single-channel images. It should be emphasized that the spatial resolution of the images should be quite high for the possibility of obtaining reliable textural information.

The work [5] is devoted to the study of textural images and segmentation by textural features. The source is micrographs of plant raw materials ground in special mills. The paper deals with the calculation of various kinds of textural features and other characteristics of images. In this work, this problem has not been fully solved, it requires more in-depth research, which will be carried out in the future. In this paper, the main atten-tion is paid to the development of software tools for segmentation of texture areas of images based on fractal dimension.

In [6], the authors proposed a probabilistic way of classifying colors based on a neural network with the introduction of textural features. Suitable texture personalities, such as CTMs, GLCM, and Gabor reactions, are investigated for clustering colors. The larger the number of classes, the better the combinations of all textural features work. In the proposed method, the combination of CTM, GLAM and Gabor responses works better than the individual functions themselves.

This article [7] suggests a way to find changes in several textures by integrating macro- and micro-textural features. Macro textures are connected to information that is determined by the entire scene of the image, whereas microtextures outline the distribution and relationship of grayish levels in the local window. The recommended method of merging several structures to show changes, which was used to accurately determine seasonal and artificial changes in municipal areas, will have an actual property for predicting fast-growing municipal areas. However, one of the limitations of the proposed method is that it cannot guarantee the exact types of changes that will be at the center of our future work.

This article [8] describes the design and development of a hierarchical network by including texture features. In addition, the impact of the inclusion of textural features on classification is being studied. Traditional statistical approaches to the classification of vegetation cover use only gray values. However, they lead to incorrect classification due to strictly convex boundaries. Textural features can be included for better classification, but they are inconvenient for conventional methods. Artificial neural networks can handle non-convex solutions. The use of textural features helps to eliminate incorrect classification.

In the above-mentioned works, textural features were applied to medical images, flowers classification, selection of elements of microscopic images and classification of vegetation cover. The application of texture features to multispectral images is not shown. Despite the ubiquitous presence of textures in images, there is currently no uniform and formal approach to the description of texture and its strict definition. Texture analysis methods are usually developed for each case.

3. The aim and objectives of the study

The aim of the study is to determine the types of crops, weeds, diseases, and pests from aerospace images.

To achieve this aim, the following objectives are accomplished:

– analyze features based on statistical and structural characteristics (energy, entropy, variation, etc.) and taking into account the relative location;

– create a conceptual model to highlight the most appropriate set of features;

– implement clustering using a non-standard approach that allows to identify types of agricultural crops and weeds.

4. Materials and methods

The study was conducted in the village of Shortandy (Shortandy district, Akmola region, Republic of Kazakhstan). In the field where flax was grown (Fig. 1). The source is aerospace images obtained from [9].

The program is implemented in the MatLab environment. Images were taken as input (images in grayscale). Many approaches and algorithms have been proposed to solve this problem. The main attention was paid to the development of effective algorithms for calculating textural features. The selected regions were obtained at random (Fig. 1). Textural features are calculated for each region.

Namely, texture windows are calculated from large windows (including non-square ones), which the user can set by selecting an area of interest to it. That is, it is about calculating numerical characteristics related to vast areas, including various kinds of artifacts. For each selected fragment (Fig. 1), 14 important texture features are calculated. All of them are implemented in the program (MatLab) and put in Table 1.

Each node of the diagram (Fig. 2) represents the union of two or more clusters, the position of the nodes on the horizontal axis determines the distance at which the corresponding clusters were combined [10].

The Graph of means graph can represent the average value of each cluster in a linear graph.

Fig. 1. Analyzed areas in the original image

Fig. 2. Dendrogram for 4 textural features: Ward's method, Euclidean distance

Table 1

Values of textural features of selected areas

No.	T11	T7	T8	T2	T10	T6	T4	T5	T1	T14	T3	T9	T13	T12	T11
1	31317.97	353.5469	5.587903	7.890208	3.23E-05	0.011703	2.678242	3.562355	0.00728	0.005656	0.38573	-0.27672	0.055401	2.46E-07	0.00728
2	30026.95	333.0048	7.66911	10.22352	5.47E-05	0.00966	3.273813	12.58858	0.002924	0.006005	0.315083	-0.44878	0.019311	9.80E-08	0.002924
3	35957.44	378.2851	6.626275	9.241705	2.84E-05	0.014686	3.079664	6.844632	0.002794	0.005287	0.317655	-0.33325	0.026437	9.32E-08	0.002794
4	23546.05	303.4873	7.104676	9.589619	4.61E-05	0.013985	2.976648	5.635921	0.00197	0.006587	0.326013	-0.40578	0.018685	6.55E-08	0.00197
5	45278.31	425.0712	5.403173	7.736198	2.25E-05	0.00828	2.817762	5.579671	0.008815	0.004705	0.376377	-0.24699	0.065701	2.98E-07	0.008815
6	17656.33	250.7452	7.767993	10.34587	9.48E-05	0.011541	3.165404	8.457648	0.001299	0.007979	0.309941	-0.45273	0.012093	4.32E-08	0.001299
7	31897.82	350.7296	6.870924	9.605984	3.87E-05	0.011909	3.194179	8.810401	0.003015	0.005707	0.301694	-0.36315	0.028527	1.01E-07	0.003015
8	44461.71	420.3436	6.028294	8.526484	2.38E-05	0.009624	2.948206	6.275669	0.005172	0.00476	0.34847	-0.30718	0.042321	1.73E-07	0.005172
9	39183.03	384.0501	6.705	9.208319	3.84E-05	0.009517	3.12462	10.56912	0.005823	0.005214	0.334221	-0.39502	0.041732	1.97E-07	0.005823
10	23624.41	301.5574	7.65277	10.21512	5.02E-05	0.009431	3.172406	8.958266	0.00131	0.006629	0.311852	-0.43578	0.012873	4.35E-08	0.00131
11	27677.41	328.6756	7.413613	9.966054	4.01E-05	0.012802	3.090141	7.076281	0.001526	0.006083	0.318747	-0.4204	0.014268	5.07E-08	0.001526
12	8638.033	173.6772	7.362724	9.485392	0.00018	0.012816	2.91472	6.201213	0.002581	0.011486	0.371073	-0.4696	0.018156	8.60E-08	0.002581
13	38639.96	392.6781	5.971867	8.52789	2.61E-05	0.017786	2.894925	4.97377	0.004726	0.005093	0.340457	-0.28295	0.042095	1.58E-07	0.004726
14	26391.24	320.9444	7.29016	9.803909	4.18E-05	0.013717	2.989473	5.746063	0.001653	0.006229	0.324942	-0.42003	0.015526	5.50E-08	0.001653
15	36893.57	383.2336	5.889077	8.422391	2.79E-05	0.011725	2.971965	5.768131	0.006223	0.005221	0.332949	-0.27801	0.053613	2.09E-07	0.006223
16	18133.08	267.2438	7.008422	9.453067	5.87E-05	0.013208	3.154898	7.292224	0.002232	0.007482	0.292307	-0.3666	0.022449	7.42E-08	0.002232
17	34205.44	367.4255	6.688209	9.274523	3.22E-05	0.008011	3.113452	10.40366	0.003084	0.00544	0.332544	-0.36654	0.028098	1.03E-07	0.003084
18	34781.58	359.7624	7.205591	9.710888	4.47E-05	0.009821	3.228997	11.53604	0.00536	0.005565	0.325905	-0.42999	0.031636	1.81E-07	0.00536
19	16446.28	254.7322	6.833754	9.227018	6.41E-05	0.014861	3.120622	6.875937	0.002492	0.007848	0.297904	-0.35302	0.024907	8.29E-08	0.002492
20	28767.95	333.3336	7.10537	9.909922	4.16E-05	0.011947	3.237029	9.02488	0.001986	0.005996	0.288186	-0.36769	0.021843	6.62E-08	0.001986
21	34353.5	369.7798	6.619902	9.249682	2.97E-05	0.014547	3.104878	7.095407	0.002647	0.005408	0.309407	-0.32687	0.02663	8.83E-08	0.002647
22	12400.55	221.8543	6.182726	8.077419	8.34E-05	0.02044	2.697629	3.567247	0.006308	0.009015	0.393212	-0.36736	0.042003	2.11E-07	0.006308
23	14047.78	223.0676	7.884799	10.29808	0.000137	0.011868	3.134123	7.425656	0.00128	0.008983	0.318199	-0.46931	0.010953	4.25E-08	0.00128
24	7768.148	168.2065	7.333617	9.698023	0.000177	0.00948	2.903595	5.396266	0.00217	0.011879	0.350705	-0.44897	0.017496	7.23E-08	0.00217
25	36172.78	371.5041	6.797504	9.284456	3.84E-05	0.01074	3.016968	6.931508	0.00479	0.005385	0.341192	-0.39662	0.035366	1.61E-07	0.00479
26	39656.8	390.9352	6.531197	8.933665	3.38E-05	0.011189	2.842022	4.94864	0.004728	0.005118	0.360226	-0.38629	0.037319	1.59E-07	0.004728
27	43211.32	415.5309	5.724139	8.201036	2.33E-05	0.026305	2.880914	4.949922	0.006257	0.004814	0.350861	-0.26092	0.048681	2.10E-07	0.006257
28	35895.78	376.197	6.40724	9.039983	3.08E-05	0.010051	3.028831	6.559822	0.004101	0.005315	0.327035	-0.33027	0.037207	1.38E-07	0.004101
29	36532.07	379.2953	6.309376	9.029008	3.00E-05	0.010865	3.099946	7.077423	0.004259	0.00527	0.306484	-0.30379	0.042398	1.43E-07	0.004259
30	29575.16	330.4289	7.742631	10.48673	5.63E-05	0.009414	3.423205	16.66033	0.001821	0.006052	0.296291	-0.43259	0.014869	6.07E-08	0.001821
31	12354.85	221.2068	6.314986	8.488715	8.46E-05	0.019633	2.700513	3.604566	0.004438	0.009041	0.382142	-0.3688	0.034069	1.48E-07	0.004438
32	30206.03	346.3201	6.317738	9.005859	3.44E-05	0.012032	3.151413	7.911882	0.003893	0.005776	0.30389	-0.29394	0.038275	1.30E-07	0.003893
33	43438.2	416.5961	5.710789	8.156419	2.31E-05	0.028529	2.85193	4.741304	0.006588	0.004801	0.356945	-0.26908	0.049406	2.22E-07	0.006588
34	19660.77	278.356	7.035444	9.675277	5.40E-05	0.013429	3.156057	7.383788	0.00177	0.007183	0.29044	-0.36159	0.01949	5.88E-08	0.00177

The division of the selected fragment into 2 classes calculated using Ward's method in the Statistica soft software environment is shown in Table 2

Result of two clustering

Selected fragment numbers	Cluster Membership Linkage distance=70112.2 Ward's method. Euclidean distances
	Cluster Membership
1	2
2	2
3	2
4	1
5	2
6	1
7	2
8	2
9	2
10	1
11	2
12	1
13	2
14	2
15	2
16	1
17	2
18	2
19	1
20	2
21	2
22	1
23	1
24	1
25	2
26	2
27	2
28	2
29	2
30	2
31	1
32	2
33	2
34	1

The most important result obtained as a result of four clusterings is a hierarchical tree. Ward's method is different from all other methods because it uses variance analysis methods to estimate distances between clusters. The method minimizes a certain sum of squared distances. More precisely, the increase in the sum of the squares of the distances of objects to the cluster centers obtained as a result of their union is taken as the distance between clusters. In general, the method seems to be very effective, but it tends to create small clusters.

5. Results of clustering satellite images

5. 1. Analysis of objects based on statistical and structural characteristics

Textural features play an important role when dividing an image into separate areas. Let's take a closer look at the system of signs. In the problem under consideration, gray (NDVI) images are investigated. Thus, the image is set

as a matrix, the values of which are pixel brightness values in the range from 0 to 255. The standard approach for calculating textural features is as follows. It is necessary to select the so-called running window with an odd side: 3, 5, 7 pixels. The attribute is calculated inside the running window. The size of the local fragment is the carrier of texture properties. The attribute value is written to a new matrix of the same size as the original one. In the new matrix, the value is written to a point with coordinates equal to the coordinates of the center of the running window. The elements of the new matrix are obtained in a certain interval (A, B). Then, usually, this interval is linearly mapped into a segment $(0,255)$. After that, it is possible to visualize the result of calculating the textural feature.

Experiments have shown that the standard approach in our case is not informative enough. Therefore, it was decided to use a non-standard approach. The steps of the non-standard approach are given below.

Namely, texture windows are calculated from large windows (including non-square ones), which the user can set by selecting an area of interest to it. That is, it is about calculating numerical characteristics related to vast areas, including various kinds of artifacts. The 14 most important textural features are considered, and all of them are implemented in the program. The calculation of textural features for the selected segments was implemented in the MatLab environment and the analysis of these textural features in the Statistica software package for statistical analysis. A large number of approaches and algorithms have been proposed to solve this problem. The main efforts were focused on creating efficient algorithms for calculating textural features. The highlighted areas were randomly selected. Textural features are calculated for each area:

1. Energy – determines the measure of uniformity of the image:

$$T_1 = \sum_{i=1}^{N} \sum_{j=1}^{N} \left[p(i,j) \right]^2. \tag{1}$$

2. Entropy – uneven distribution of brightness properties of image elements:

$$T_2 = -\sum_{i=1}^{N} \sum_{i=1}^{N} p(i,j) \cdot \log_2 \left(p(i,j) \right). \tag{2}$$

3. Uniformity (homogeneity) – means a measure of uniformity:

$$T_3 = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{p(i,j)}{1+|i-j|}. \tag{3}$$

4. Differential entropy – the same as the total entropy of the Matrix and the entropy of the adjacent Matrix, but calculated for a histogram of differences in brightness values:

$$T_4 = -\sum_{i=2}^{N} p_{x+y}(i) \log \left(p_{x+y} + (i) \right), \tag{4}$$

here

$$p_{x+y}(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{i+j,k} p(i,j), \quad k = 2,3,...,2N,$$

$$\delta_{m,n} = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n. \end{cases} \tag{5}$$

5. Variance sum – serves as a variation of the brightness change relative to the sum of the average value:

$$T_5 = \sum_{i=2}^{2N} \left(i - T_4\right)^2 p_{x+y}(i). \tag{6}$$

6. Correlation – a measure of the dependence of video brightness on linear regression:

$$T_6 = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\left(i - \mu_i\right)\left(j - \mu_j\right) \cdot p(i,j)}{\sigma_i \sigma_j}; \tag{7}$$

here μ_i, μ_j, σ_i, σ_j – $p(i, j)$ for mean values and mean square deviations.

7. The total average value is determined by the sum histogram of brightness values for pairs of elements in the image $p_{x+y}(i)$ that are closely related to the adjacent Matrix:

$$T_7 = \sum_{i=2}^{2N} i \cdot p_{x+y}(i); \tag{8}$$

here

$$p_{x+y}(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} i + j = k^{p(i,j)} \tag{9}$$

$k=2, 3, ..., 2N$ – distribution on an indirect diagonal.

8. Total entropy – determined by the classical value of statistical information theory and represents an uneven distribution of the brightness properties of video elements:

$$T_8 = -\sum_{i=0}^{N-1} p_{x-y}(i) \log p_{x-y}(i); \tag{10}$$

here

$$p_{x-y}(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} |i - j| = k^{p(i,j)}, \tag{11}$$

$k=0, 1, 2, 3, ..., N-1$ – head diagonal distribution.

9. 1. information criterion of correlation:

$$T_9 = \frac{T_2 - HXY_1}{\max(HX, HY)}; \tag{12}$$

here

$$HX = -\sum_{i} p_x(i) \log\left(p_x(i)\right),$$

$$HY = -\sum_{j} p_y(j) \log\left(p_y(j)\right), \tag{13}$$

$$HXY_1 = -\sum_{i} \sum_{j} p(i,j) \log\left(p_x(i) p_y(j)\right),$$

$$p_x(i) = \sum_{j=1}^{N} p(i,j); \quad p_y(j) = \sum_{j=1}^{N} p(i,j). \tag{14}$$

Information criteria are determined by the ratio of statistical information theory for elements of an adjacent Matrix, a histogram of the sum of brightness values, and a histogram of the difference of brightness values.

10. Reverse torque – a measure of local similarity that determines the sharpness of a short series:

$$T_{10} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i,j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}. \tag{15}$$

11. Moments – the weight of a line that has a constant optical density. This sign is characterized by the value of increasing the weight of the length of each line for any gray level:

$$T_{11} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}. \tag{16}$$

12. Brightness heterogeneity – distribution of the gray level. This sign takes the least value in cases where the number of optical density lines is evenly distributed over the gray level:

$$T_{12} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left(p(i,j)\right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}. \tag{17}$$

13. Series length heterogeneity – distribution of the length of a line with a constant optical density. In a uniform distribution, the minimum value is:

$$T_{13} = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} \left(p(i,j)\right)\right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}. \tag{18}$$

14. Share of images in the series:

$$T_{14} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} jp(i,j)}. \tag{19}$$

After checking each feature for informativeness, 4 textural features were selected from 14 features: T11, T7, T10, T14. The analyzed areas in the original image are shown in Fig. 1. The values of the textural features of the selected areas are shown in Table 1.

5. 2. Creating a conceptual model for selecting the most appropriate set of functions

Despite the absence of a strict definition, four methods of formal approach and texture can be considered at the stage of definition formation:

– statistical approach: the texture is calculated as a quantitative characteristic of the distribution of intensity values in the image area;

– structural approach: the texture is considered as a set of textured primitives arranged in simple or repetitive order;

– fractal approach: the description of a wide range of processes and phenomena, such as limited diffusion aggregation processes, the formation of sticky fingers in Porous Media, diffusion processes called leaks, and the characteristics of clouds, the Earth's surface, and other natural objects, has revealed a new direction called fractal analysis in the term fractal geometry. In such methods, the authors consider the object as a fractal, not as a texture [11];

– spectral approach: Fourier and wavelet analysis are used to work with the texture of an image at different scales. In wavelet analysis, the signal is divided into basic functions (wavelets) built based on the creator's wavelet, using Shift and zoom operations. For the initial image, a layout with a basic function (filter) is created, and then a second projection is created from the difference between the received and the initial signal, etc. Each base function is an extension and displacement of the previous function by some number.

One of the heuristic clustering methods is methods based on a sequential agglomeration procedure. The advantage of these methods is the simplicity of the computational procedure and algorithms. The work [12] shows the stages of clustering satellite images. Clustering was performed using vector textural features.

In our study, let's use combined textural features from among statistical and structural features. The 14 textural features were made up of these two textural features. Based on the results of calculations, as shown in Fig. 3, textural features were selected from statistical and structural features.
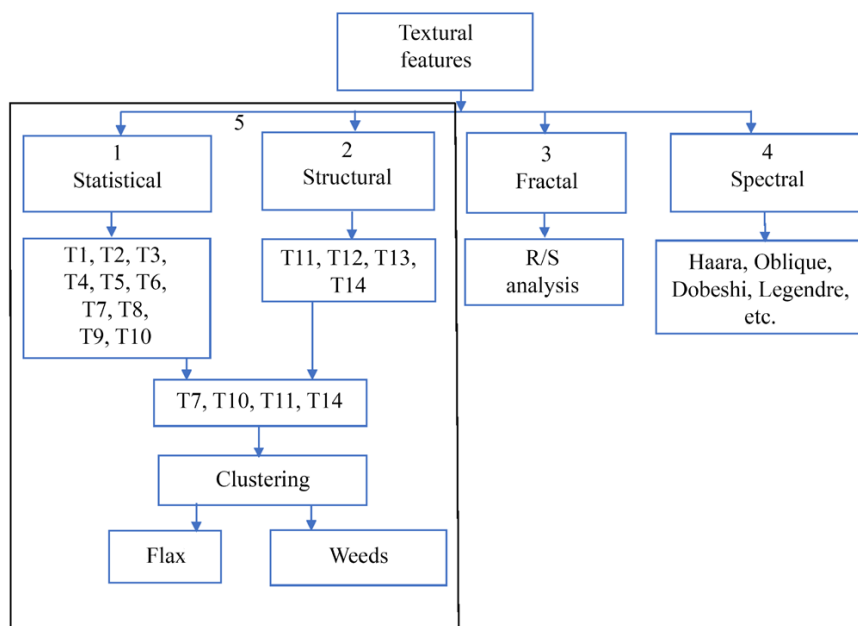

Fig. 3. Conceptual model for the isolation of weeds on linen for the growing season

Looking at the results, it is possible to cluster weeds on a flax plot.

This conceptual model can be used for clustering other crops and negatively influencing factors (weeds, pests, etc.).

5. 3. Implementation of clustering using a non-standard approach

As a result of the research, it was found that due to a large number of textural features, it was unknown which of them were informative. Therefore, it was decided to use a non-standard method to select informative textural features for clustering with a standard approach. As a result, clustering by the standard method was time-efficient.

The numerical values of the textural features calculated for the selected fragment are shown in the following table (Table 4).

The average values graph allows to view the average values for the selected cluster on a linear graph (Fig. 4, Table 4).

Each node in the diagram above combines two or more clusters, nodes on the horizontal axis determine the distance between the combined clusters.

According to the average values of textural features from Table 5, it is possible to say that they do not intersect, they are well separated from each other. This arrangement of the coordinates of the cluster centers indicates a good clustering quality.

For five clusters, the results of the analysis of variance indicate a good quality of clustering: the significance level of p everywhere is less than 5 % (Table 6).

Based on the visual representation of the results and the Euclidean distance, it is possible to determine the composition of each cluster (Table 7).

The composition of clusters isolated by the Ward and k-means method is presented in Table 8.

Table 4

Values of informative textural characteristics of selected areas

Number of the selected fragment	T11	T7	T10	T14
1	2	3	4	5
1	31317.97	353.5469	3.23E-05	0.005656
2	30026.95	333.0048	5.47E-05	0.006005
3	35957.44	378.2851	2.84E-05	0.005287
4	23546.05	303.4873	4.61E-05	0.006587
5	45278.31	425.0712	2.25E-05	0.004705
6	17656.33	250.7452	9.48E-05	0.007979
7	31897.82	350.7296	3.87E-05	0.005707
8	44461.71	420.3436	2.38E-05	0.00476
9	39183.03	384.0501	3.84E-05	0.005214
10	23624.41	301.5574	5.02E-05	0.006629
11	27677.41	328.6756	4.01E-05	0.006083
12	8638.033	173.6772	0.00018	0.011486
13	38639.96	392.6781	2.61E-05	0.005093
14	26391.24	320.9444	4.18E-05	0.006229
15	36893.57	383.2336	2.79E-05	0.005221

Continuation of the Table 4

| 1	2	3	4	5
16	18133.08	267.2438	5.87E-05	0.007482
17	34205.44	367.4255	3.22E-05	0.00544
18	34781.58	359.7624	4.47E-05	0.005565
19	16446.28	254.7322	6.41E-05	0.007848
20	28767.95	333.3336	4.16E-05	0.005996
21	34353.5	369.7798	2.97E-05	0.005408
22	12400.55	221.8543	8.34E-05	0.009015
23	14047.78	223.0676	0.000137	0.008983
24	7768.148	168.2065	0.000177	0.011879
25	36172.78	371.5041	3.84E-05	0.005385
26	39656.8	390.9352	3.38E-05	0.005118
27	43211.32	415.5309	2.33E-05	0.004814
28	35895.78	376.197	3.08E-05	0.005315
29	36532.07	379.2953	3.00E-05	0.00527
30	29575.16	330.4289	5.63E-05	0.006052
31	12354.85	221.2068	8.46E-05	0.009041
32	30206.03	346.3201	3.44E-05	0.005776
33	43438.2	416.5961	2.31E-05	0.004801
34	19660.77	278.356	5.40E-05	0.007183
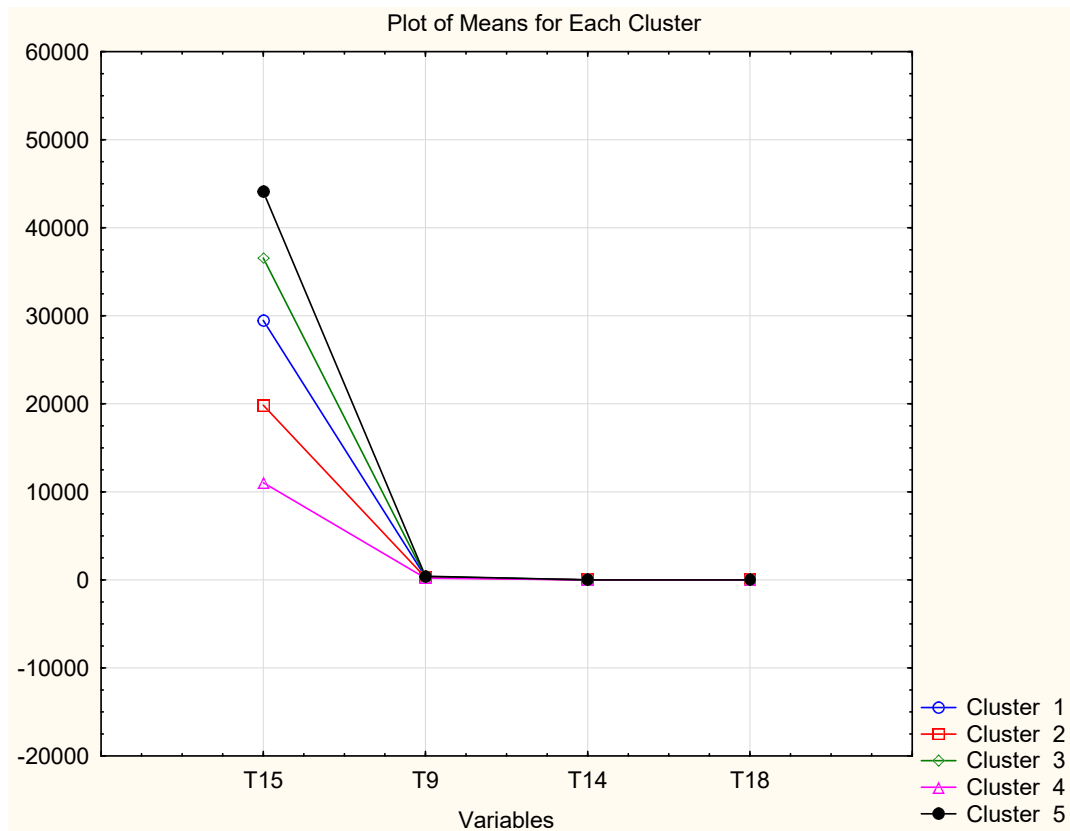
Fig. 4. Average value of textural features for the selected cluster

Table 5

Mean value, standard deviation, and variance of textural features

Variable	Descriptive Statistics for Cluster 2		
	mean	standards (deviation)	variance
T11	19844.49	3075.314	9457560
T7	276.02	22.720	516
T10	0.00	0.000	0
T14	0.01	0.001	0

Table 6

The result of the analysis of variance

Vari-able	Analysis of Variance					
	between (SS)	df	within (SS)	df	F	signif. (p)
T11	3.659152E+09	4	137929200	29	192.3367	0.000000
T7	1.571342E+05	4	7740	29	147.1897	0.000000
T10	4.070985E-08	4	0	29	26.1335	0.000000
T14	9.798875E-05	4	0	29	65.2409	0.000000

Table 7

Clusters to which the observation belongs, Euclidean distance and observation number

Selected fragment numbers	CASE_NO	CLUSTER	DISTANCE
1	1	1	917.74
2	2	1	272.20
3	3	3	306.37
4	4	2	1850.83
5	5	5	590.47
6	6	2	1094.15
7	7	1	1207.65
8	8	5	182.16
9	9	3	1306.43
10	10	2	1890.00
11	11	1	902.59
12	12	4	1202.00
13	13	3	1034.92
14	14	1	1545.69
15	15	3	161.72
16	16	2	855.71
17	17	3	1182.38
18	18	3	894.34
19	19	2	1699.13
20	20	1	357.31
21	21	3	1108.35
22	22	4	679.42
23	23	4	1502.99
24	24	4	1636.95
25	25	3	198.72
26	26	3	1543.33
27	27	5	443.04
28	28	3	337.20
29	29	3	19.07
30	30	1	46.42
31	31	4	656.56
32	32	1	361.76
33	33	5	329.60
34	34	2	91.87

Table 8

Composition of classes allocated by Ward and k-means methods

Class number	Number of objects in the class	Class composition
1	8	Selected regions 1, 2, 8, 11, 14, 20, 30 and 32
2	6	Selected areas 4, 6,10, 16, 19 and 34
3	11	Selected areas 3, 9, 13, 15, 17, 18, 21, 25, 26, 28 and 29
4	5	Selected areas 12, 22, 23, 24 and 31
5	4	Selected areas 5, 8. 27 and 33

Table 8 shows the cluster number and the distance to the selected regions.

The identified sites in the 2nd cluster revealed the presence of weeds and pests on the flax site according to a set of 4 textural features.

The selected area of 16, 19, and 34 belongs to Class 2 (Table 8) of the studied area of flax (Fig. 1). Using the flax section (Fig. 5), where this selected area is located, let's construct a graph of the spectral brightness coefficient of flax and weed (Fig. 6).

Let's define our research according to the graph of the spectral brightness coefficient (SBC) shown in Fig. 6.

The graph shows that NDVI is higher in this growing season. This means that a lot of weeds grow on the flax plot.
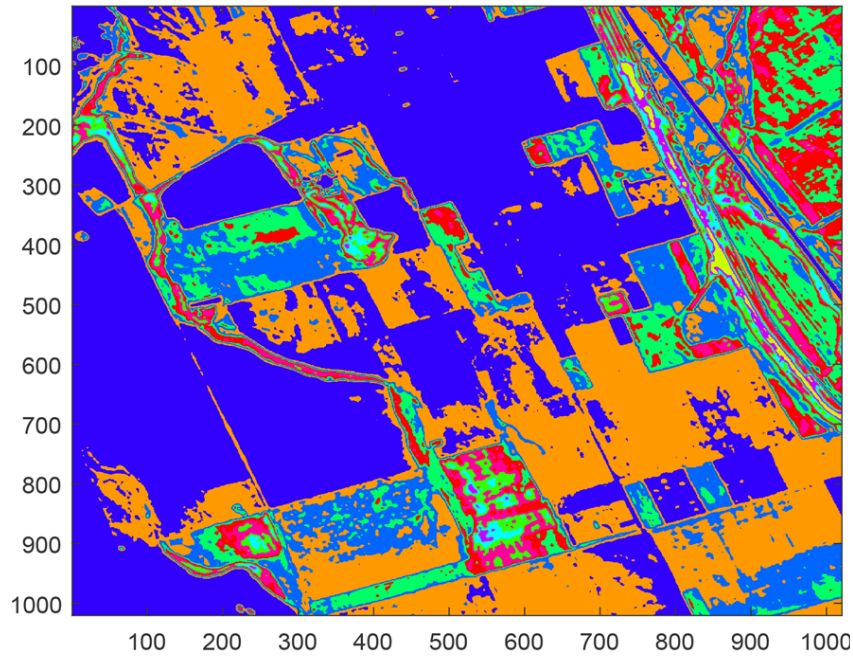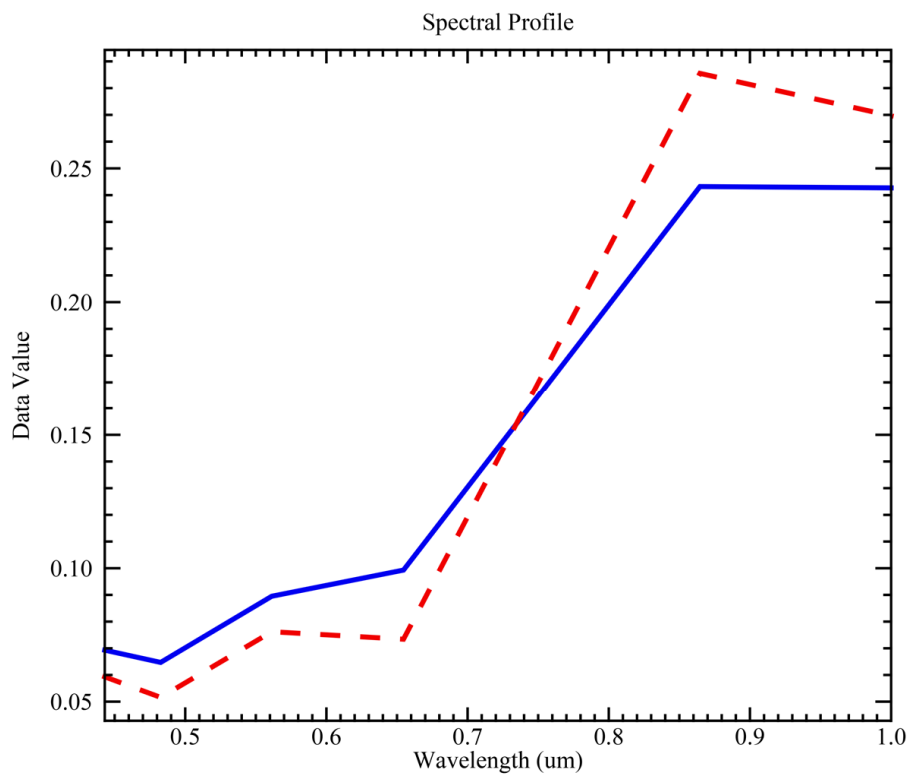
Fig. 5. Image obtained after clustering

Fig. 6. Spectral brightness coefficient of flax and weeds: – – – – – – weed; ———— – flax

6. Discussion of experimental results of the non-standard approach with the help of informative textural features

Cluster analysis was considered in a standard and non-standard hierarchical manner and experiments were conducted. Practical studies have shown that classification by selected textural features gives real results in many cases.

Conducting research on texture marks in a non-standard method, let's use 4 texture marks on Windows T11, T7, T10, T14 with a resolution of 200×150 pixels. In this method, segmentation by region was performed using agglomerate and K-mean methods (Table 7). As a result, the signs perfectly identified homogeneous regions in the image, and these signs are considered informative for marking homogeneous regions.

Analysis of variance outputs a table of variance analysis (Table 6). The table shows values between groups (Between SS) and within groups (within SS). The more the variance of values within a group and the greater the variance of values between groups, the more the object belongs to the cluster. In addition, the quality of clusterization can be determined by: the value of the values of the F-criterion (the greater the better) and the significance of the level (the smaller the better).

The advantage of the work over other works was that it used a non-standard approach to the segmentation of textural zones for information on aerospace images (NDVI) during the growing season. During the study, several classifications were made for the selected area. Then the SBC of the selected region was built (Fig. 6). A conceptual model for determining weeds in flax during the growing season was built.

The detected textural features, as shown by the research work, do not detect weeds of all crops. Experiments have shown that the disadvantage of this approach is the identification of informative textural features for the weeds of each crop must be calculated separately.

In the future, it is planned to use neural networks to automate the identification of weeds and agricultural crops.

7. Conclusions

1. A large set of textural research methods was analyzed and the most informative ones were selected from the methods considered. In total, 14 textural features were calculated and 4 textural features were selected.

2. According to the results of the study, a conceptual model was built for the isolation of weeds on flax by textural features based on statistical and structural characteristics. According to this conceptual model, it is possible to find weeds of any agricultural crop.

3. An experiment was conducted in a non-standard version. After conducting a study on textural features, four features were selected, such as T11, T7, T10, T14 for which they were segmented by areas. Using the k-means and agglomerative methods, the features have well highlighted the areas in the images. The results show that these signs are informative.

References

1. Abdikerimova, G. B., Murzin, F. A., Bychkov, A. L., Tussupov, J., Khayrulin, S., Xinyu, W. E. I., Rybchikova, E. I. (2018). Software tools for cell walls segmentation in microphotography. Journal of Theoretical and Applied Information Technology, 96 (15), 4783–4793.

2. Chaban, L. N., Beriozina, K. V. (2018). Analysis of the informativeness of spectral and texture features while classifying the vegetation on hyperspectral airborne imagery. Geodesy and Aerophotosurveying, 62 (1), 85–95. doi: https://doi.org/10.30533/0536-101X-2018-62-1-85-95

3. Gaidel, A. V., Pervushkin, S. S. (2013). Research of the textural features for the bony tissue diseases diagnostics using the roentgenograms. Computer Optics, 37 (1), 113–119. doi: https://doi.org/10.18287/0134-2452-2013-37-1-113-119

4. Rodionova, N. V. (2012). Teksturnaya segmentatsiya odnokanal'nyh izobrazheniy: primery primeneiya. Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa, 9 (3), 65–69. Available at: http://d33.infospace.ru/d33_conf/sb2012t3/65-69.pdf

5. Tussupov, J. A., Abdikerimova, G. B., Murzin, F. A. (2018). Application of fractal dimension for the analysis of microphotographs. Vestnik KazNITU im. K.I.Satpaeva, 5, 137–142. Available at: https://official.satbayev.university/download/document/7429/%D0%92%D0%95%D0%A1%D0%A2%D0%9D%D0%98%D0%9A-2018%20%E2%84%965.pdf

6. Guru, D. S., Sharath Kumar, Y. H., Manjunath, S. (2011). Textural features in flower classification. Mathematical and Computer Modelling, 54 (3-4), 1030–1036. doi: https://doi.org/10.1016/j.mcm.2010.11.032

7. Li, Q., Huang, X., Wen, D., Liu, H. (2017). Integrating Multiple Textural Features for Remote Sensing Image Change Detection. Photogrammetric Engineering & Remote Sensing, 83 (2), 109–121. doi: https://doi.org/10.14358/pers.83.2.109

8. Rashmi, S., Mandar, S. (2011). Textural Feature Based Image Classification Using Artificial Neural Network. Advances in Computing, Communication and Control, 62–69. doi: https://doi.org/10.1007/978-3-642-18440-6_8

9. Planet. Available at: https://www.planet.com/explorer/

10. Sidorova, V. S. (2012). Hierarchical cluster algorithm for remote sensing data of earth. Pattern Recognition and Image Analysis, 22 (2), 373–379. doi: https://doi.org/10.1134/s1054661812020149

11. Fisenko, V. T. (2013). Fraktal'nye metody segmentatsii teksturnyh izobrazheniy. Priborostroenie, 56 (5), 63–70. Available at: https://pribor.ifmo.ru/file/article/6254.pdf

12. Yerzhanova, A., Kassymova, A., Abdikerimova, G., Abdimomynova, M., Tashenova, Z., Nurlybaeva, E. (2021). Analysis of the spectral properties of wheat growth in different vegetation periods. Eastern-European Journal of Enterprise Technologies, 6 (2 (114)), 96–102. doi: https://doi.org/10.15587/1729-4061.2021.249278