ENGINEER

international scientific journal

ISSUE 2, 2025 Vol. 3
E-ISSN
3030-3893
ISSN
3060-5172

A bridge between science and innovation

TOSHKENT DAVLAT TRANSPORT UNIVERSITETI
Tashkent state transport university

ENGINEER

A bridge between science and innovation

E-ISSN: 3030-3893

ISSN: 3060-5172

VOLUME 3, ISSUE 2

JUNE, 2025

engineer.tstu.uz

TASHKENT STATE TRANSPORT UNIVERSITY

ENGINEER

INTERNATIONAL SCIENTIFIC JOURNAL

VOLUME 3, ISSUE 2 JUNE, 2025

Founder of the international scientific journal "Engineer" – Tashkent State Transport University, 100167, Republic of Uzbekistan, Tashkent, Temiryo'lchilar str., 1, office: 465, e-mail: publication@tstu.uz.

The "Engineer" publishes the most significant results of scientific and applied research carried out in universities of transport profile, as well as other higher educational institutions, research institutes, and centers of the Republic of Uzbekistan and foreign countries.

The journal is published 4 times a year and contains publications in the following main areas:

- Engineering;
- General Engineering;
- Aerospace Engineering;
- Automotive Engineering;
- Civil and Structural Engineering;
- Computational Mechanics;
- Control and Systems Engineering;
- Electrical and Electronic Engineering;
- Industrial and Manufacturing Engineering;
- Mechanical Engineering;
- Mechanics of Materials;
- Safety, Risk, Reliability and Quality;
- Media Technology;
- Building and Construction;
- Architecture.

Tashkent State Transport University had the opportunity to publish the international scientific journal "Engineer" based on the **Certificate No. 1183** of the Information and Mass Communications Agency under the Administration of the President of the Republic of Uzbekistan. **E-ISSN: 3030-3893, ISSN: 3060-5172.** Articles in the journal are published in English language.

The problem of classification and justification of the methodology for selecting classes of objects subject to recognition in a multispectral space image

R.H. Ramazanov[1][a]
[1]National Aerospace Agency / Research Institute of Aerospace Informatics, Baku, Azerbaijan

Abstract: In the article, the specific task of Remote Sensing is being investigated to classify recognizable classes from objects consisting of 12 different types of vegetation soil features. The stages of solving this problem were selected and justified and the territorial areas where the test areas were formed for the development and testing of the classifier were determined. Based on information obtained from satellite and IKONOS, tables are provided that provide quantitative assessments for various class combinations of vegetation objects in the images formed from the satellite image of the studied area. The tables reflect statistical characteristics allow for preliminary assessments of the representativeness of the clusters. The article presents combinations of histograms for various variants of combinations of classified vegetation classes and provides tables of unsuitability using decision rules based on statistical characteristics.
Keywords: Classification, Justification, Methodology, Objects

1. Introduction

To solve the main task of classifying recognizable objects, 12 types of vegetation and soils were identified in the selected study region, the dynamics of changes in the distribution areas of which are indicators reflecting the anthropogenic impact on the study area. Below, in Table 1, some types of vegetation are listed and their names are given.

Table 1

Initial set of classes – vegetation and soil types

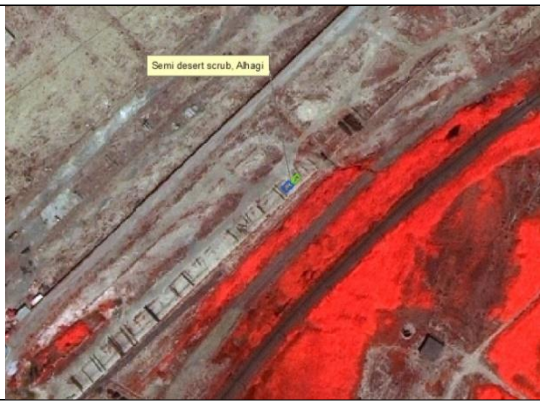
Class number	Full name of the plant
Class 1	Marshy reed vegetation
Class 2	Swamp Shrub Tamarisk (Tamarix)
Class3	Coastal zone: semi-desert vegetation
Class 4	Australian false reed (Phragmaties australis)
Class 5	Salsola ericoides
Class 6	Mountain saltwort (Salsola nodulosa)
Class 7	Mountain saltwort (Salsola nodulosa) / Lerch's wormwood (Artemesia lerchiana)
Class 8	Mountain Salsola (Salsola Nodulosa) / Herb
Class 9	Semi-desert vegetation - Caspian potash (Kalidium capsicum)
Class 10	Semi-desert vegetation dominated by Camel Thorn (Alhagi pseudoalhagi)
Class 11	Bare soil
Class 12	Mountain saltwort (Salsola nodulosa) / bare soil

At the first stage, the geographic coordinates measured at the characteristic distribution points of these plants in the studied area are determined, which were geocoded on the space image, and then the areas of the territories are determined, from which the test areas for training and testing the classifier are formed. The image below shows the contours of these areas for some types of selected sets of vegetation classes. The areas with training examples are shown in green, and the areas with testing examples are shown in blue.

Marshy reed vegetation	Swamp Shrub Tamarisk (Tamarix)
Coastal zone: semi-desert vegetation	Australian false reed (Phragmaties australis)
	
Mountain saltwort (Salsola nodulosa)	Salsola ericoides
Mountain Salsola (Salsola Nodulosa) / Herb	Mountain saltwort (Salsola nodulosa) / Lerch's wormwood (Artemesia lerchiana)

[a]https://orcid.org/0009-0006-3244-2091

Semi-desert vegetation - Caspian potashnik, (Kalidium capsicum)



Semi-desert vegetation dominated by Camel Thorn (Alhagi pseudoalhagi)

Bare soil

Mountain saltwort (Salsola nodulosa) / bare soil

Fig. 1. Vegetation types

2. Research methodology

The capabilities of the designed GIS for recognition and classification of objects allow not only to automate the processes of extracting test examples from a space image within the boundaries of the selected areas, but also to use various algorithms for this[1,2,3] .

In the case under consideration, an algorithm is used that takes into account all pixels that have a slight overlap with the selected areas. In the second case, pixels intersecting with the boundary of the area are added to the set only if the overlap area is more than half the area of one pixel(1 sq.m., in this case).

Below, in Table 2, quantitative data are given on different classes of vegetation cover formed from a space image of the region under study from the IKONOS satellite. The table includes calculated statistical characteristics that allow obtaining preliminary estimates of the representativeness of clusters.

Table 2

Quantitative data are given on different classes of vegetation cover formed from a space image

Class number	Number of examples Procedure 1	Number of examples Procedure 2	Number of examples Procedure 1	Number of examples Procedure 2
Class 1	1215	1113	1487	1368
Class 2	3181	2955	2087	1893
Class 3	97	67	63	33
Class 4	1891	1748	2055	1859
Class 5	234	186	279	176
Class 6	2393	1992	2867	2658
Class 7	690	564	478	388
Class 8	200	182	147	124
Class 9	52	35	49	30
Class 10	40	33	28	21
Class 11	474	242	553	287
Class 12	680	604	579	472
General	11147	9721	10672	9309
Mathematical expectation	928.9167	810.0833	889.3333	775.75
Standard deviation	1036.354	950.3847	976.7583	917.4441

The table shows that the distribution of examples by classes is uneven. This is due to both the total number of examples extracted from different classes of vegetation and soil types presented on the scene, and the nature of the fragmentation of the distribution areas of different plant types [4]. Let's consider some of these estimates and, first of all, let's consider the estimates by their total number:

1. $N_T = 5 * N_W$
2. $N_T = 30 * p * N_O$
3. $N_T = 10 * N_W$
4. $N_T = 30 * N_I * (N_I + 1)$
5. $N_T = 60 * N_I * (N_I + 1)$, where,

N_I - the number of initial neurons, in our case, this is the number of spectral channels of the space image;

N_O - the number of output neurons or recognizable classes.

N_H - the number of neurons in the hidden layer.

Using the given formulas, we determine the minimum and maximum values of the required number of neurons. As a result of the calculations, we obtain;

$$N_H \min = 57, \text{ a } N_H \max = 120.$$

N_W -the number of weighting factors, which is determined by the formula

$$N_W = N_H * (N_I + N_O).$$

Thus

$N_W min = 8*(4+12) = 684$, a $N_W max = 120*(8+12) = 2400$, where

• p is the number of objects/neurons per class and ranges from 3 to 8.

Analyzing the obtained data, we have a wide range of values, which makes it difficult to determine the optimal threshold when choosing the required number of examples. However, it is clear that those estimates that take into account the number of weighting coefficients are reliable. Let us consider the estimates of the distribution of examples by object classes using the formula for determining the interval obtained from the maximum and minimum number of examples:

$$И_{HT} = [N_{av} - \alpha * N_{msd}, N_{av} + \alpha * N_{msd}], \quad (1)$$

where $N_{av} = (N_T / N_O)$ is the mathematical expectation, N_T is the total number of examples for training or testing; N_O is the number of classes, α is a real number in the range of numbers from one to three, and N_{msd} is the standard deviation of the number of examples in different classes [5,6].

Literature review. The main literature sources used are the monograph "Theory and Methods of Digital Processing of Space Images for Remote Sensing Problems". The book is in print and will be published in 2024. In addition, materials from articles published in various international publications and SUDEF 24 conferences in 2023-2024 were used.

ICSEGT 2024; 74th International Astronautical Congress 2023 y.

3. Results and Discussion

Statistical tests for representativeness and separability of samples.

Let's consider the general statistical characteristics:

Table 3

Statistical characteristics of training clusters from the "Main set" for four channels

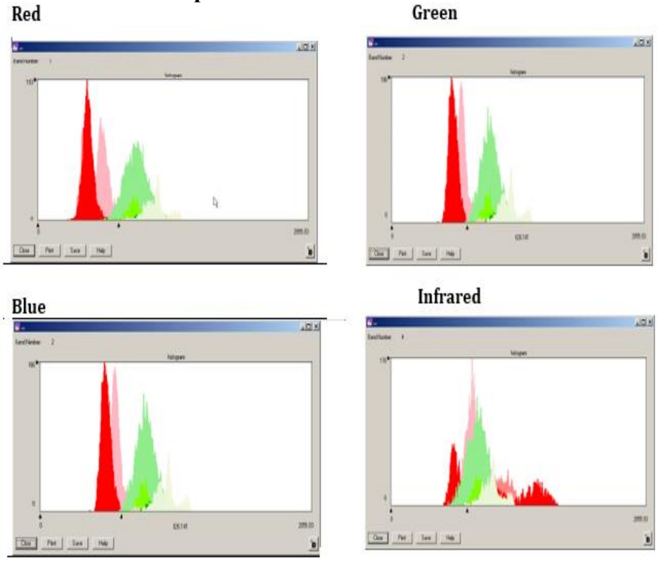
Class name	Red Channel		Green Channel		Blue Channel		Infrared channel	
	Minimum-Maximum	Average ± Stan.Off	Minimum-Maximum	Average ± Stan.Off	Minimum - Maximum	Average ± Stan.Off	Minimum-Maximum	Average ± Stan.Off
Class 1	241 – 655	363.5 ± 40.5	398.0 - 725.0	497.9 ± 31.4	356 - 553	418.0 ± 20.0	466 - 1132	820.5 ± 115.8
Class 2	264 - 738	497.8 ± 65.9	391 - 779	574.4 ± 53.2	354 - 605	468.6 ± 33.9	396 - 821	631.8 ± 52.6
Class 3	555 - 826	704.7 ± 60.4	648 - 878	772.5 ± 56.5	511 - 655	586.1 ± 36.7	481 - 759	649.4 ± 57.4
Class 4	216 - 697	380.7 ± 44.5	362 - 737	498.7± 38.0	329 - 553	424.7 ± 22.4	416 - 1394	806 ± 289.4
Class 5	544 – 981	769.6 ± 95.3	653 - 1008	839.7 ± 77.2	493 - 724	619 ± 50.1	481 - 941	713.7 ± 100.8
Class 6	395 – 1058	743.2 ± 89.7	523 - 1051	799.2 ± 75.9	409 - 760	609.3 ± 42.8	362 - 1000	698.1 ± 87.8
Class 7	554 – 930	729 ± 55.5	631 - 973	780.1 ± 50.7	493 - 691	583.7 – 30.1	534 - 896	700.4 – 50.2
Class 8	597 – 972	784.1 ± 63.5	694 - 994	835.8 ± 51.2	517 - 715	618.6 ± 33.2	541 - 928	738.3 ± 65.7
Class 9	639 – 761	689.0 ± 30.6	684 - 781	722.4 ± 24.4	528 - 598	553 ± 16.1	642 - 783	698 ± 30.2
Class 10	389 – 870	657.1 ± 139.6	520 - 914	742.8 ± 115.7	424 - 672	561.4 ± 70.2	432 - 909	696.8 ± 137.9
Class 11	760 – 1104	941.4 ± 64.6	841 - 1141	1001.8 ± 60.2	607 - 800	709.8 ± 38.0	673 - 1001	851.5 ± 61.1
Class 12	613 – 1047	854.0 ± 75.9	715 - 1078	912 ± 63.3	530 - 768	660.6 ± 40.7	523 - 975	780.0 ± 76.4

It follows from the table that some clusters either intersect or partially coincide. The article considered various combinations of histograms for various combinations of classified vegetation classes. For ease of analysis, the recognizable classes were presented in different colors.

Table 4

Joint histograms of clusters for individual spectral channels

Class 1	
Class 2	
Class 3	
Class 4	
Class 5	
Class 6	
Class 7	
Class 8	
Class 9	
Class 10	
Class 11	
Class 12	



The article analyzes various combinations of histograms, unfortunately the limited volume of the article does not allow us to present the results obtained in full.

Thus, the preliminary tests characterize one of the possible options for modifying the original classification scheme.

For example, the algorithm of the classification scheme is considered, in which an additional set consisting of 7 classes of vegetation types is considered [7].

Using this algorithm, we create new clusters by extracting examples from one or more clusters formed on the basis of the original classification scheme.

The main problem of the initial formation of training examples is the uneven distribution of them across clusters. It is theoretically possible to redistribute the examples so that they are more evenly distributed between the classes [8].

4. Conclusion

We have proposed a technique where the problem statement related to the definition of the structure of the classification scheme was changed. As a rule, the proposed approach is more effective for solving problems of natural resource management, as well as control and aerospace monitoring. When applying this approach, the priority is the selection of recognizable indicators for assessing the situation. After performing the initial tests, an option was proposed for combining classes with similar characteristics and new classification schemes were created.

This allows us to improve the results of training the classifier for classifying objects of recognizable classes. The approach we proposed, combining the methods of modern information technologies and data models on which the proposed methodology is based, always allows us to find the optimal set of indicators - classes and solve the problem [9,10].

Calculation of confusion tables using decision rules based on statistical characteristics.

After conducting statistical tests and evaluations, we created two additional sets from the original 12-cluster set, consisting of 7 and 5 clusters, respectively. Additional sets were compiled before the training procedure of the neural classifier in order to determine the boundaries of possible changes in the classification scheme.

In order to minimize the risks associated with the uncertainty of the training procedure, it is necessary to form representative sets of training examples. Before the training process of neural classifiers, sets of sets of training examples are tested based on the calculation of confusion tables using decision rules based on statistical characteristics.

In this paper, confusion tables are compiled using the maximum likelihood rule on clusters from the maximum 12-class set of examples. Each set has two types: sets of training examples and examples for testing, called training and test clusters. The full version of the article provides tables for the two types of sets, as well as a table of inaccuracies in the classification of examples.

References

[1] Sulaymanov S., Kamilov X. Developing a method [1] Acton, S. T.: On unsupervised segmentation of remotely sensed imagery using nonlinear regression. International Journal of Remote Sensing. 17, 1407–1415 (1996)

[2] Blaschke, T.: Object based image analysis for remote sensing. ISPRS International Journal of Photogrammetry and Remote Sensing . 65 , 2-16 (2010). doi:10.1016/j.isprsjprs.2009.06.004

[3] Burrough, P. A., Mcdonnel, R. A.: Principle of Geographical Information Systems, Oxford University Press .(1998). doi:10.1080/01431160210144499

[4] Bruin, D.S., Molenaar, M.: Remote sensing and geographical information systems. In Spatial Statistics for Remote Sensing. Stein, A., Meer, F.V., Gorte, B.G.H.:(eds) Dordrecht Kluwer Academic . 41–54 (1999)

[5] Frohn, R. C., Autrey, B. C., Lane, C. R., Reif, M.: Segmentation and object-oriented classification of wetlands in a karst Florida landscape using multi-season Landsat-7 ETM+ imagery. International Journal of Remote Sensing . 32, 1471-1489 (2011). doi:10.1080/01431160903559762

[6] Geneletti, D., Gorte, B. G. H.: A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. International Journal of Remote Sensing. 24, 1273-1286 (2003). https://doi.org/10.1080/01431160210144499

[7] Navulur, K.: Multispectral Image Analysis Using the Object-Oriented Paradigm. CRC Press, Taylor & Francis Group. (2007)

[8] Stein, A., Meer, F.V., Gorte, B.G.H. : (eds) Spatial Statistics for Remote Sensing. Dordrecht: Kluwer Academic. (1999)

[9] GEOINFORMATION TECHNOLOGIES IN REMOTE SENSING TACKS BASED ON HIGH RESOTION SATTELITE IMAGES E.RAMAZANOVA, R.RAHIMOV, D.QAHRAMANOVA, F.SAMEDOV INTERNATIONAL CONFERENCE DEDICATED TO THE 90-TH ANNIVERSARY OF ACADEMICIAN AZAD MIRZAJANZADE , BAKU, AZERBAIJAN ,13-14 DESEMBER (2018)

[10] Rauf Ragimov, Elchin Isgenderzade, Rovshan Ramazanov, Farid Samadov, Shana Jahidzada ”An innovative technology for aerospace monitoring of geotechnical systems based on the use of unmanned aerial vehicles (UAVs). Azerbaijan National Aerospace Agency 74th International Astronautical Congress Baku, Azerbaijan 2-6 October (2023).

Information about the author

Rovshan Ramazanov

National Aerospace Agency / Research Institute of Aerospace Informatics, Baku, Azerbaijan
PhD, Deputy Director for Science,
E-mail: r.r_90@mail.ru
Tel.: +994504887805
https://orcid.org/0009-0006-3244-2091

CONTEXT / MUNDARIJA