

INFORMATION TECHNOLOGY

10.18372/2306-1472.68.10910

Viktor Kurochkin

CLUSTER ANALYSIS IN AGRICULTURAL DECISION SUPPORT SYSTEMS DEVELOPMENT

National aviation university, Kyiv, Ukraine
E-mail: glukozavr.91@gmail.com

Abstract

Purpose: With the development of unmanned aircraft technology raises the problem of creation systems of aerial data analysis to increase efficiency and reduce the cost of monitoring areas based on operational data and building automated systems of farm management combined with precision farming techniques. **Methods:** The review of the current state of agriculture and the development of unmanned aircraft, the use of information technologies in Ukraine compared to more developed countries made. The modern approach to imaging and conducted own research of the possibilities of image processing system based on cluster analysis. **Results:** Shown that the level of integration of information technology in Ukraine is quite low, in contrast to other developed countries, where there is active development of geographic information systems, including for agricultural purposes. Unmanned aerial vehicles developed in Ukraine, as evidenced by the development of the legal sphere and the presentation of definition of the category. Image processing based on cluster analysis provides data suitable for further analysis and give an idea of the structure of the area, the presence of abnormalities and the general condition of farmland. **Discussion:** Cluster analysis of digital images of agricultural areas is a promising area of further research that could be the basis for building complex decision support systems for the development of precision farming including in Ukraine and increase the efficiency of farming, as well as an additional tool for determining the state of the areas for evaluation and prognostication.

Keywords: aerial photography; agribusiness; cluster analysis; image processing; spatial data.

1. Introduction

With development and decreasing cost of unmanned aircraft, effective aerial photography usage problem for security, surveillance, control took place.

Cluster analysis is one of the most studied examples of image processing tasks used in data processing, as well as in "Data Mining" [1]. With clustering general classes revealed in arrays of data. The trait of the majority of cluster analysis methods is a large number of variables setting method flow. This is the reason of the complexity of adapting approaches for specific areas of development and that is why this research took place. This is about creating systems capable of operating in certain conditions, to solve specific problems, and give best results without burdening interface and complexity.

The relevance for decision support systems development for agriculture caused primarily by the

active development of unmanned aircraft and the low level of integration of information technology in Ukraine, especially with the background of "precision farming" development based on information technology integration in other developed countries such as the US Japan, China, Germany, etc. [2].

An example of the cluster analysis implementation in agriculture is modules for digital images analysis for farmland images for the geographic information systems, as described in [3] to perform analysis of the territories general state, to identify deviations, crop damage, inconsistencies, characteristics of soil and forecasting, decision-making. The results of this analysis may have a key role in creating an automated support system for agriculture that will increase its efficiency through better control of the territories.

Today Ukraine have a low level of information technologies integration to agriculture such as GIS. Possible causes of this situation considered in more detail in [3]. It can be expensive solutions, user demands etc. A partial solution might be to develop own systems, with an intuitive interface and powerful features that will be available for agronomists to understand the expediency of new technologies integration.

Problem with cultivated and agricultural areas digital images is the need to consider the specifics and peculiarities of cultures and its external manifestations typical violations to form features set cluster analysis require solving the tasks. This may be the detection of violations of Culture; dehydration; lack of nutrients in the soil; the density of vegetation; the presence of excess crops (weed etc.); crop forecasting, spread of the disease, etc.

2. The modern trends and approaches in precise farming and digital image analysis

Law development is indicator of the widespread use of unmanned aerial vehicles (UAVs). Legal status of the UAV recently defined in the legislation of Ukraine, thus unmanned aircraft - aircraft intended for the flight without a pilot on board, which flight management and control carried out through a special control station located outside the aircraft [4].

One example of the UAVs use is a complex monitoring system facility of OAO "Gazprom" [5], which rely on the system the following tasks: industrial and environmental monitoring, geophysical surveys, mapping, security, search and rescue, ice reconnaissance, and more. In monitoring tasks for civilian use not always real time information processing needed, thus data delivery to the station, where processing will take place is sufficient. Results can be reported to appropriate services if needed.

The study [6] analyze existing approaches to solving problems using digital image processing in the context of agriculture. Remote sensing (RS) is widely used in agriculture for different purposes. Satellite data used to determine characteristics, evaluate geo-biophysical properties of the earth's surface. Spatial, spectral and time-dependent data used with analytical methods using digital image processing, data mixing from different sources. These RS provide information on groundwater, which helps in irrigation, flood management, and make it possible to monitor the environment and

natural disasters, assess effects, and more. With pattern recognition, acreage potential evaluated and region inventory studied. Optical and microwave data used for the classification of sowing. Cultures discriminated using visual or digital interpretation of data, and so on. Using digital image processing the next problems is solving, including the identification of weeds approaches with borders identification, color analysis, classification based on wavelet and fuzzy sets, sorting and grading of fruits and food, and so on. The authors conclude about the prospects of this area, especially in relation to approaches by analyzing the color component images, including methods of classification and cluster analysis.

Also in [7] analyzed the capabilities of digital processing and stated that the spectral characteristics of vegetation depends on the health of the vegetation, the internal structure of leaf chlorophyll content, water and so on. These spectral properties are expressed in the form of vegetation index VNIR (visible and near infrared (Fig. 1)). In this case satellite data were examined, and therefore, a more general analysis using multiple spectral bands and their combination with determining optimal for imaging vegetation take place.

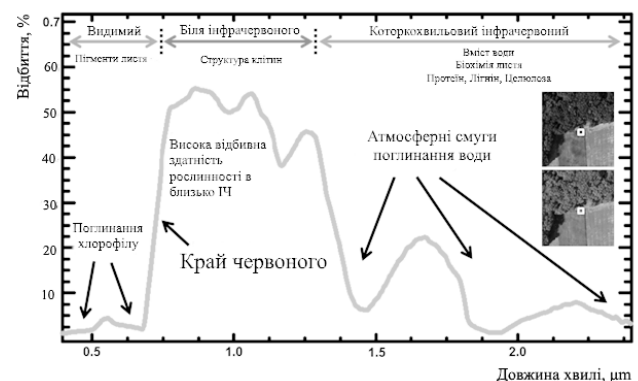


Fig. 1 Detailed description vegetative spectrum

In earlier work [8], the problem of image analysis studied with the three steps script: image enhancement, segmentation, measurements. The main purpose of the study was to determine the number and area of clusters present in the image data using smoothing and cluster analysis methods on the points set.

In Kansas Unmanned Systems Conference in 2013 raised the issue of the growing demand for post processing system of the UAVs data of territory in the context of agriculture. In [9] the researchers concluded the following about the traditional sampling and remote sensing using UAVs: the traditional collection of samples characterized by high costs and long delays, uncertain local risks

while RS UAV - a complete virtual analysis of the entire surface with high efficiency and high spatial and temporal resolution.

Special attention demands the development of precision farming \ agriculture, which based on the idea of the existence of irregularities within the same field. Based on the collected data optimum sowing density evaluated, fertilizers and plant protection products rules calculated, yield production forecasted and financial planning settled. There is a need to take into account the peculiarities of soil and climatic conditions. According to [10] Precision Agriculture rapid growth predicted in early 2000, but there is a problem of developing decision support systems for the implementation of this technology. There is a lack of quality assessment methods and efficiency of recognition. Not taken into account the change in time or only once a year score is used or only part of the assessment areas. The reason is the high cost of sampling and analysis. Therefore, a cheaper area monitoring using UAVs can solve some of these problems.

Cluster analysis is a research tool in the study of patterns, clustering, decision-making, machine learning, documents search, image segmentation and pattern recognition [11]. Solving the problem of determination of the types of structures available in the image often leads into the conditions of small amounts of priori data, but decision-making system should produce a minimum number of assumptions. Under these restrictions cluster methodology is suited to the study of the relationships between data elements to assess their structure.

3. The main objectives

The main objective of this study is to search for applications of cluster analysis to solve the problems of improving the efficiency of agriculture through the development of information technologies and automation processes to save time and money.

So, let us have digital images, aerial photography data, which shows plots of the crop area, with some signs of deviation from normal culture development.

It is necessary to use methods of cluster analysis to identify the damaged areas and those that do not meet certain quality criteria.

4. Cluster analysis application in agriculture

Problem for clustering are as follows. There are sample research objects $s = \{o^{(1)}, \dots, o^{(N)}\}$ that formed in result of the selection of some members of the population Ω . It is necessary to form $K \geq 2$

classes (groups of objects). Each object of the population described by a set of variables X_1, \dots, X_n .

The set $X = \{X_1, \dots, X_n\}$ may include different types of variables (quantitative and qualitative, nominal, Boolean, order).

A typical clustering algorithm includes the following steps:

- presenting samples (can include the definition, selection and feature extraction);
- Determination of distance, proximity samples corresponding to the type of data;
- clustering, or grouping;
- abstracting data, if necessary;
- Assessment results, if necessary.

Based on software «Elfintest» [12] the use of cluster analysis to digital photos of the planted areas is studied.

Consider the typical structure that can occur on aerial photography data fields, cultivated areas, forest areas etc. (Fig. 2).

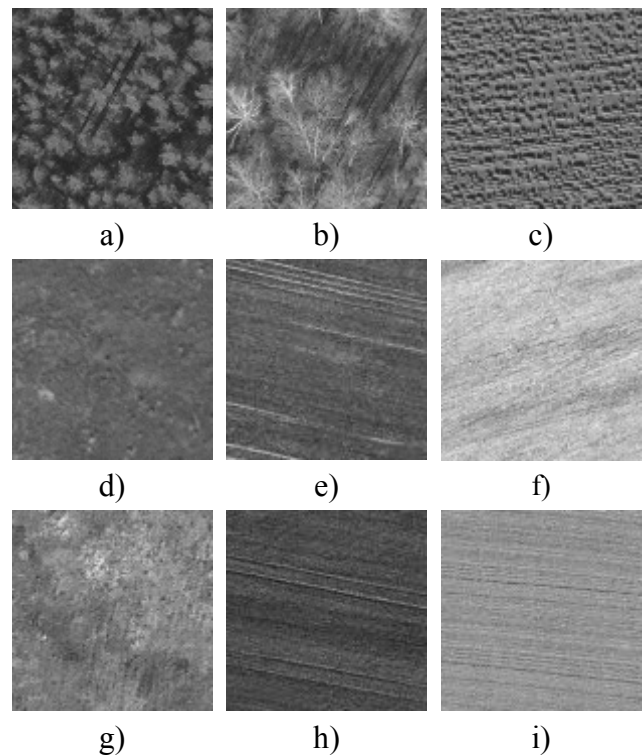
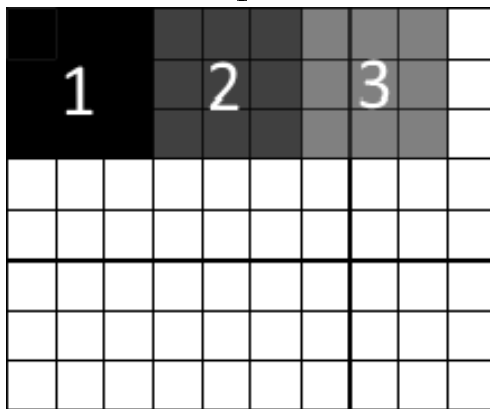
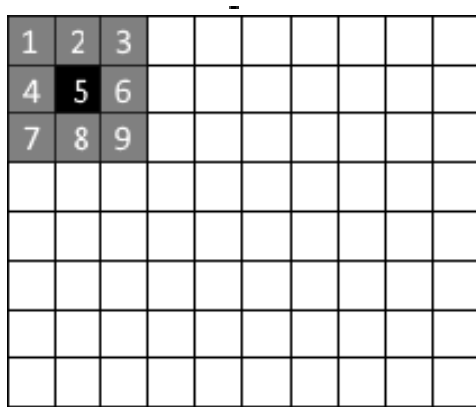


Fig. 2 Typical textures for RS of agricultural crops

For the most adequate analysis, it is necessary to determine the vector of features for the submitted samples. Pixel by pixel analysis, when the sample is a pixel intensity values, usually gives insignificant results because each separated pixel cannot represent the texture. There are two approaches to image analysis. The first - the most detailed, but extremely

difficult due computational complexity - a pixel by pixel analysis when sample an array of pixels around (Fig. 3a) where the pixels 1, 2, 3, ..., 9 are used to form feature vector pixel at number 5).



b)

Fig. 3. Features sets building approaches

The second approach described in [13]. The image divided into some segments (Fig. 3 B), which used to form feature vectors, so the number of samples is less than the number of pixels. After analysis on a new set of samples, approximation used with results to get an estimated value in every pixel.

The following image properties can be used to build the feature vector: the expectation of intensity values for the three components (if RGB - red, green and blue); distribution of intensity values, ie, standard deviation and other statistical parameters of any population. RGB is not the best color scheme for cluster analysis, since three components describes the colors content ratio. More information is stored in color schemes like YUV [14], where the components represent brightness and color

component separately, thus allowing distinguishing darkened areas and different color areas.

In constructing this vector of features, there is a problem of data normalization, assessing the impact of each feature on cluster analysis. As example, the use of mathematical expectation, standard deviation and more. The impact on the difference between textures based on quantitative differences of expectation or deviation is difficult to estimate objectively. During the study concluded that more efficiently will be independent analysis of each feature and the next mixing of the results for a single decision about clusters rather than building mixed feature. After analysis of each feature, the result is a new vector of normalized features.

Different approaches and methods of cluster analysis used to solve the problems of interpreting digital images. For example, in Fig. 4 shows an image on the crop with dehydration vegetation in section 1 and section 2.

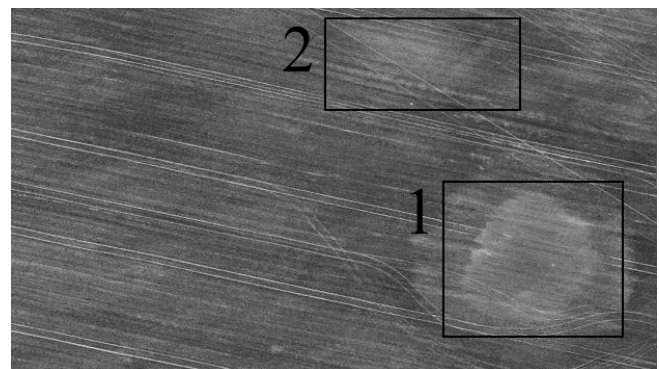


Fig. 4 Crop photo example with problems in 1 and 2 areas

In this case, samples of an array of pixels represented by intensities from 0 to 255. As a measure of the distance simple metric used - the absolute value of the difference of intensities. Using a simple per pixel clustering method [11] we can build original form [13], for easy visual analysis (Fig. 5), which shows two main areas: 1 - dry area, 2 - fertile area. In this case, the type of zones defined by visual analysis.

To gather more information, more sophisticated methods used, such as ISODATA [15]. In the same example, we can get 4 area (Fig. 6), which can be roughly characterized as follows: 1 - critical dehydrated; 2 - dehydrated; 3 - normal; 4 - saturated.

In this case the relative dehydration value for areas received. When working with a certain type of

tasks to specific areas, crops, etc., evaluation criteria generated in process, followed by a determination of the absolute values of "normal", "critical" and other estimates of the analysis.

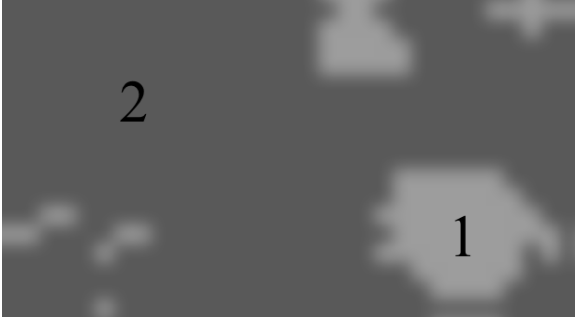


Fig. 5 Simple clustering result to reveal dehydrated area 1

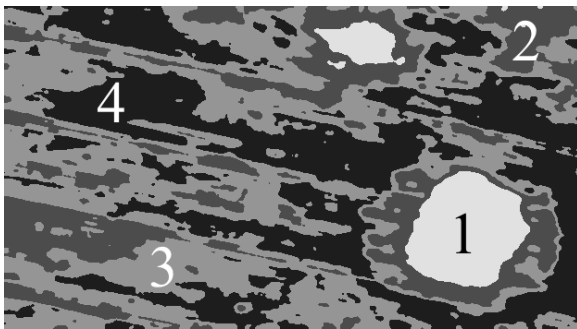


Fig. 6. ISODATA clustering result with 4 detected areas

Another task is to determine impurities in the homogeneous culture, such as weed, various diseases of trees and vegetation, crops or unplanned development (Fig. 7). In this case, to obtain better results primary image processing take place, in order to highlight differences and get rid of noise, provide color correction. Results of cluster analysis (Fig. 8) reveal two different textures. In this case, texture that is typical for this area should be set, so the system can signal the spread of other, understanding that the identified is typical.

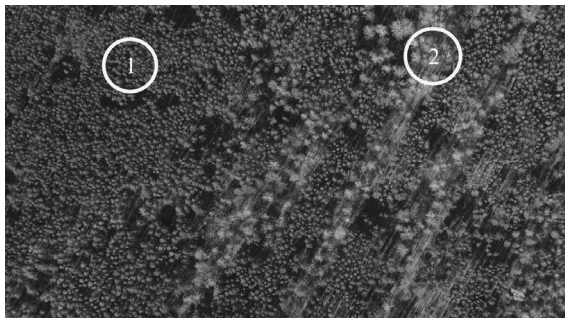


Fig. 7 Image of forest with two types of trees

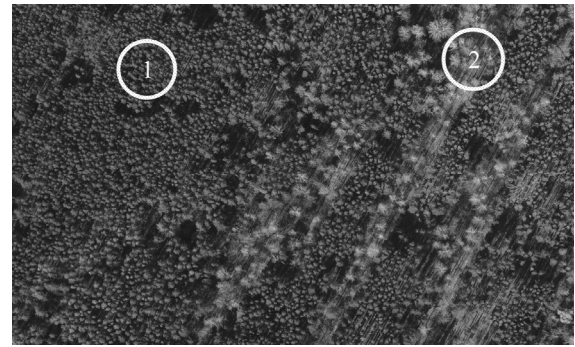


Fig. 8. Clustering results separating two trees types

5. Conclusions

Thus, there is interest in exploring the capabilities of digital image processing and analysis, including cluster analysis, especially considering the development and distribution of UAVs and use the RS.

Cluster analysis is capable to give meaningful results, but requires knowledge of the features of the phenomenon for the selection of feature vectors and can be self-sufficient. However, this is high-quality tool, a part of the information system that can carry out pre-processing, and even learning, since clusters analysis can't give the full information but only a part of it, that can be used to interpret for the end user.

Using the capabilities of cluster analysis, together with a powerful apparatus of primary image processing [12] and integration with GIS we may create a system that is capable to assess the state of the territory, detect various kinds of deviations to prevent accidents and timely response to water shortages, certain micronutrients lack based on aerial photography, which will greatly facilitate the task of farming, reducing risks related to oversight. Always be aware of the current state-controlled territories. In future work we can apply the results to the map, using the GIS layers that supported in most of these systems [3], specify definitions for such zones and be able to determine the automated data types.

So developing software to analyze data from the UAV aerial photography is a promising research area. Especially due to the presence of growing interest in precision farming and finding the most effective ways of farming, increasing the need to increase food supply and distribution using economical, compact unmanned aircraft that can provide the necessary data for similar systems.

References

- [1] *Witten I. H.* Data Mining: practical machine learning tools and techniques. 3rd ed. Ian H. Witten, Fran Eibe, Mark A. Hall The Morgan Kaufmann series in data management systems, 2011, 630 p.
- [2] *Adamova A. M.* The use of modern information technology in agriculture. Scientific conference Science and Practice: Innovation 2007. [Electronic resource]. Access: <http://www.pdaa.edu.ua/np/pdf2/27.pdf> (In Russian)
- [3] *Kurochkin V. M.* Analysis and prospects of development of modern geographic information solutions to optimize agriculture. Problems of information and management, 4(52). 2016.8 p. (In Ukrainian)
- [4] *Boyarska Z. I.* use of airspace Ukraine unmanned aircraft. Air and Space Law: Proceedings of the Second International Conference "Formation of law in the context of globalization: a theoretical and practical aspects"- K: HAY. 2012. 366-368 p.(In Ukrainian)
- [5] *Voronov V.* Complex system of monitoring objects of OJSC "Gazprom" with help of UAVs [Electronic resource]. 2009. Access: <http://uav.ru/articles/KSKM.pdf> (In Russian)
- [6] *Vibhute A.* Applications of Image Processing in Agriculture: A Survey. A. Vibhute, S. K. Bodhe [Electronic resource]. International Journal of Computer Applications (0975. 8887) Volume 52, No.2. 2012. Access: <http://research.ijcaonline.org/volume52/number2/pxc3881495.pdf>
- [7] *Slave C.* Analysis of Agricultural Areas Using Satellite Images. International Journal of Academic Research in Environment and Geography Vol. 1, No.2. 2014, 26. 32 p.
- [8] *Glasbey C.* Image analysis in agricultural research \ C.A. Glasbey, G.W. Horgan. Quantitative Approaches in Systems Analysis, 23. 2001. 43-54 p.
- [9] *Merwe D.* Using sUAS for high resolution characterization of harmful algal. Deon van der Merwe, Kevin P. Price [Electronic resource]. Kansas Unmanned Systems Conference, 2013. Access: <https://krex.kstate.edu/dspace/bitstream/handle/2097/17033/VanderMerweppt.pdf?sequence=1&isAllowed=y>
- [10] *Mcbratney A.* Future Directions of Precise Agriculture [Electronic resource]. A. B. Mcbratne, B. M. Whelan, J. Bouma. Precision Agriculture. 2005. Access: https://www.researchgate.net/publication/226724989_Future_Directions_of_Precision_Agriculture
- [11] *Jain A. K.* Data Clustering: A Review. M. N. Murty, P. J. Flynn [Electronic resource]
- [12] *Kurochkin V. M.* The system «Elfintest» Environmental Monitoring data based clustering. knowledge based technology, №2 (26). 2015. 127-133 p. (In Ukrainian)
- [13] *Kurochkin V. M.* Analysis of heterogeneous textures acreage estimates based on mixture distributions. V. M. Kurochkin, P. A. Pristavka. Knowledge based technology, N4 (28). 2015.. 305-310 p. (In Ukrainian)
- [14] *Lyhun A. O.* Computer Graphics. A. O. Lyhun, O. O. Shumeiko.. D.: Bila K. O., 2010.. 83 p. (In Ukrainian)
- [15] *Memarsadeghi. N. A.* Fast Implementation of the ISODATA Clustering Algorithm [Electronic resource]. N. Memarsadeghi, N.S. Netanyahu, J. LeMoigne. International Journal of Computational Geometry and Applications. Access: <https://www.cs.umd.edu/~mount/Projects/ISODATA/ijcga07-isodata.pdf>

Received 10 May 2016

В.М. Курочкін

Використання кластерного аналізу для побудови системи прийняття рішень для сільського господарства
 Національний авіаційний університет, просп. Космонавта Комарова, 1 Київ, Україна, 03680
 E-mail: glukozavr.91@gmail.com

Мета: З розвитком безпілотної авіаційної техніки постає задача створення систем аналізу даних аерофотозйомки для збільшення ефективності та зменшення вартості моніторингу територій на основі обробки оперативних даних та побудови автоматизованих систем управління господарством в сполученні з методиками точного фермерства. **Методи:** Зроблено огляд сучасного стану сільського господарства та розвитку безпілотної авіатехніки, використання інформаційних технологій в Україні у порівнянні з більш розвинутими країнами. Розглянуто сучасні підходи до обробки зображення та проведено власне дослідження про можливості обробки зображень на основі апарату кластерного аналізу. **Результати:** Показано, що рівень інтеграції інформаційних технологій в Україні досить низький, на противагу іншим розвинутим країнам, де йде активний розвиток геоінформаційних систем, в тому числі для агропромислових цілей. Безпілотні літальні апарати розвиваються в

Україні, про що свідчить розвиток юридичної сфери та сформульовано визначення даної категорії. Обробка зображень на основі кластерного аналізу дає дані що підходять для подальшого аналізу і дають уявлення про структуру територій, наявність відхилень та загальний стан посівних площ. **Обговорення:** Кластерний аналіз цифрових зображень агропромислових територій є перспективним напрямом досліджень, що може бути основою для побудови складних систем підтримки прийняття рішень для розвитку точного фермерства в тому числі в Україні та збільшення ефективності ведення господарства, а також як додатковий інструмент визначення стану територій для оцінки і прогнозування.

Ключові слова: аерофотографія; агропромисловість; кластерний аналіз; обробка зображень; просторові дані.

В.М. Курочкин

Использование кластерного анализа для построения системы принятия решений для сельского хозяйства

Национальный авиационный университет, просп. Космонавта Комарова, 1, Киев, Украина, 03680

E-mail: glukozavr.91@gmail.com

Цель: С развитием беспилотной авиатехники появляется задача создания систем анализа данных аэрофотосъемки для увеличения эффективности та уменьшения стоимости мониторинга территорий на основе обработки оперативных данных и построение автоматизированных систем управления хозяйством в сочетании с методикой точного фермерства. **Методы:** Сделано обзор современного состояния сельского хозяйства и развития беспилотной авиатехники, использования информационных технологий в Украине в сравнении с более развитыми странами. Рассмотрено современные подходы к обработке изображений и проведено собственно исследование возможностей обработки изображений на основе аппарата кластерного анализа.

Результаты: Показано, что уровень интеграции информационных технологий в Украине достаточно низкий, на противовес более развитым странам, где идет активное развитие геоинформационных систем, в том числе ждя агропромышленных целей. Беспилотные летательные аппараты развиваются в Украине, про что свидетельствует развитие юридической сферы и тот факт что сформулировано определение данной категории. Обработка изображений на основе кластерного анализа даёт данные, что подходят для дальнейшего анализа и дают представление о структуре территорий, наличии отклонений и общее состояние посевных площадей.

Обсуждение: Кластерный анализ цифровых изображений агропромышленных территорий является перспективным направлением развития, что может быть основой для построения более сложных систем поддержки принятия решений для развития точного фермерства, в том числе в Украине и увеличения эффективности ведения хозяйства, а также как дополнительный инструмент определения состояния территорий для оценок и прогнозирования.

Ключові слова: аерофотографія; агропромисленність; кластерний аналіз; обробка зображень; просторові дані.

Kurochkin Viktor (1991). Graduate student.

National Aviation University, Kyiv, Ukraine.

Education: National Aviation University, Kyiv, Ukraine (2013).

Research area: image processing, cluster analysis, pattern recognition.

E-mail: glukozavr.91@gmail.com