

ANALYSIS OF INFORMATIVE FEATURE CHANGES ON COLOR IMAGES USING MASS-PARALLEL PROCESSING

Summary

Remote sensing methods allow effective detecting field areas that are infected by plant diseases. The infection detected on early stages of its development reduces costs of plants protective measures. In the paper the problems of disease feature extraction as well as disease identification are considered. Three groups of potato plants with 25 images in each group were under experimental observation in laboratory conditions. The proposed algorithm of automatic definition of appearance of changes has shown good result of objects identification at use of an attribute of change of color characteristics of object. The greatest influence on job of a method renders: presence in the staff of extraneous subjects and the shadows, having color of object; non-uniformity of illumination that creates additional handicaps.

1. Introduction

Remote sensing methods allow effective detecting field areas that are infected by plant diseases. The infection detected on early stages of its development reduces costs of plants protective measures. For example, change of reflective characteristics of potato plants in infra-red area allows identifying phytophthora even before appearance of visual features [1, p. 409]. In spite of that fact development of optical method for infection detection takes place both for an independent system and for spectrometric one that increases quality of the identification.

There are different kinds of features that allow infection detection [2-8]: geometric, morphological and color or their combination in order to reduce a feature space which leads to simple accuracy schemes of identification. The one of important steps of processing is to solve the segmentation problem. Different approaches are used caused by both specific of input data and representation of results.

In the paper the problems of disease feature extraction as well as disease identification are considered. Three groups of potato plants with 25 images in each group were under experimental observation in laboratory conditions [10]:

B – Plants infected by alternaria disease. The affected plants have dark- brown stains on leaves and stems. The leaves fade, grow yellow and become black, then they wither and decay and fall off at humid weather, and the stems fracture.

Z – Plants infected by erwinia disease. Leaves of the affected plants grow yellow and roll up. Bottom parts of stems and roots decay and become black.

K – Healthy plants (check group).

Photographies, which had been done at 8, 10, 12, 14 and 16 o'clock during 7 days, were used for study.

2. Statement of a problem and technique of the analysis

Let the object on the color image of the layer submitted by set of the staff, is set by color. Thus color of object can differ a little on the different staff.

The purpose of identification is division of the image into sites of two types: the areas corresponding to objects, and all other, being a background.

The problem of identification is solved in two stages. In the beginning the expert analysis for revealing information attributes of objects is carried out.

The set of the presented images is analyzed by the expert by means of software Photoshop (the following basic operations are used: segmentation; the analysis of histograms of a hue and RGB a component of color of segments). Result of the expert analysis is value of a hue and saturation, corresponding to normal and required segments.

Then the automatic analysis for revealing changes and monitoring of dynamics of their changes is carried out. For the analysis have been developed: clustering algorithm, multi criterion_threshold algorithm of segmentation of images and the algorithm of identification which is taking into account dynamics of change of informative attributes.

The hue and saturation are used as informative features.

The hue and saturation values are calculated according to the following formulas:

$$Hue = \arctan \left(\frac{\sin(\frac{2}{3})\pi \cdot g - \sin(\frac{2}{3})\pi \cdot b}{r + \cos(\frac{2}{3})\pi \cdot g + \cos(\frac{2}{3})\pi \cdot b} \right),$$

$$Sat = \max(r, g, b) - \min(r, g, b),$$

where r, g, b are red, green and blue color components.

As a result, the calculation method of coefficients was proposed. The coefficients specify disease features. The method includes the following two steps:

- extraction of required and normal color segments;
- calculation of numerical values of the appearance changes features:

- threshold coefficient is

$$Tresh = 100 \cdot \frac{N1}{(N1 + N2)},$$

where N1 and N2 are number of segments pixels of normal and required color respectively;

- percents of hue average shifts toward required are

$$Shift1 = 100 \cdot \left(\frac{1}{N} \sum_{i=1}^N hue_i \right) / (Hue_1^{norm} - Hue_1^{abnorm}),$$

$$Shift2 = \frac{100 \cdot \left((Hue_1^{norm} + Hue_2^{norm}) / 2 - \frac{1}{N} \sum_{i=1}^N hue_i \right)}{\left((Hue_1^{norm} + Hue_2^{norm}) / 2 - (Hue_1^{abnorm} + Hue_2^{abnorm}) / 2 \right)}$$

- hue dispersion of required and normal color segments is

$$Disp = \sqrt{\frac{1}{N} \sum_{i=1}^N (hue_i)^2 - \left(\frac{1}{N} \sum_{i=1}^N hue_i\right)^2},$$

where $N = N1 + N2$, Hue^{norm} – normal hue value, Hue^{abnorm} – required hue value.

The effectiveness of proposed coefficients is analyzed below.

3. Segmentation algorithms

The following approaches have been tested.

1. Allocation of segments on the basis of criterion of hit in the set interval of value of a shade of color of the points, based on the following ratio:

$$x = R \cdot \cos(30) - G \cdot \cos(30); \quad y = B - R \cdot \cos(60) - G \cdot \cos(60);$$

$$h = \begin{cases} \arctan(x/y), x \geq 0, y \geq 0; \\ \arctan(x/y) + 180, x \geq 0 \text{ and } y \leq 0 \text{ or } x \leq 0 \text{ and } y \geq 0; \\ \arctan(x/y) + 360, x \leq 0, y \leq 0. \end{cases}$$

Thus it is necessary to leave such points of the image, for which $h \in [a; b]$, where a and b – the values received at a stage of selection of informative attributes.

2. Selection of segments on the basis of criterion of implementation of the following empirical ratio:

$$abs(R - G) < c; \quad \frac{R + G}{4} > B; \quad G > d,$$

where c and d - the values received at a stage of allocation of informative attributes.

However the best result gives selection of segments on a basis clustering, described below.

Clustering based algorithm of segmentation. The idea of algorithm consists in clustering of sets of image points in coordinate space of color components. As clustering algorithm the modified algorithm "Trout" which feature is dynamic addition new clusters during receipt of new points that has allowed to reduce essentially quantity of iterations

on set of points was used. Then reduction of clusters quantity up to the set size is made by merge most relatives clusters. Cluster-"winner" who is most to relatives on color parameters to the set color at last gets out; segmentation of the image is made according to criterion of the least remoteness of a considered point from the center chosen cluster.

That algorithm has the following disadvantage: if there are no segments of desired color in the image, the algorithm finds segments similar to desired color, that results in an inaccuracy of result of segmentation. The algorithm resulted below is relieved of this disadvantage.

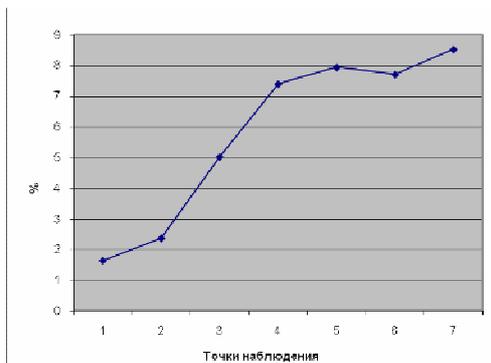
Multi criterion threshold algorithm of segmentation.

This algorithm of segmentation is proposed that finds segments which color characteristics exactly coincide with desired color. Moreover, the developed algorithm has higher performance because of its simplicity.

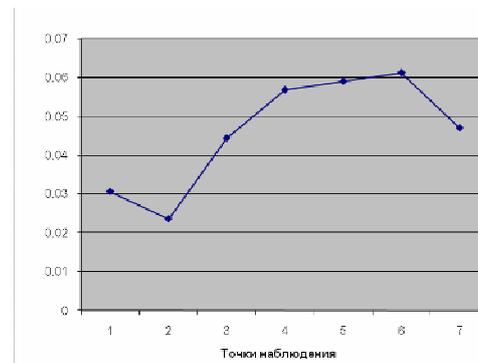
- Set the following segmentation conditions:
 - For normal segments: $1 Hue^{norm}_1 < hue < Hue^{norm}_2$ и $Sat^{norm}_1 < sat < Sat^{norm}_2$.
 - For required segments: $Hue^{abnorm}_1 < hue < Hue^{abnorm}_2$ и $Sat^{abnorm}_1 < sat < Sat^{abnorm}_2$.
- Detect pixels according to a) or b) conditions.
- Calculate the values of features *Treash*, *Shift1*, *Shift2*, *Disp* for selected pixels.

4. Analysis of changes dynamics

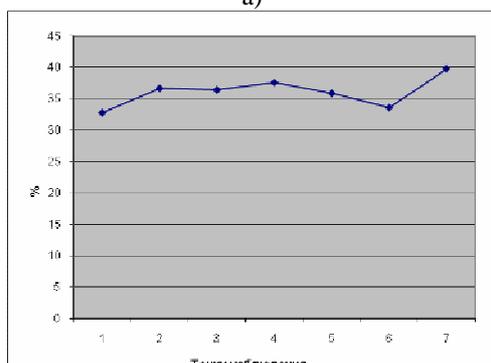
Analysis of the graphs of values *Treash*, *Shift1*, *Shift2* and *Disp*, depending on vantage points, showed that the most effective coefficients is *Treash* that correlates with *Disp*. Therefore *Treash* was used during further analysis. Examples of *Treash* values graphs are shown in Figure 3.



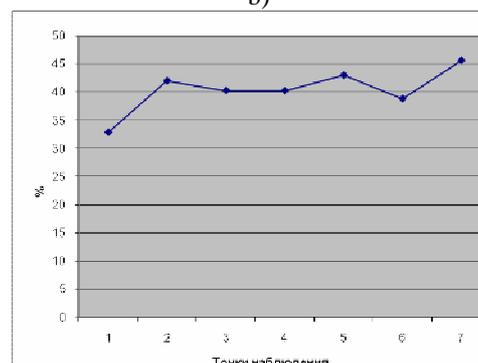
a)



b)



c)



d)

Fig. 1. Feature graphs example: a) Treash; b) Disp; c) Shift1; d) Shift2

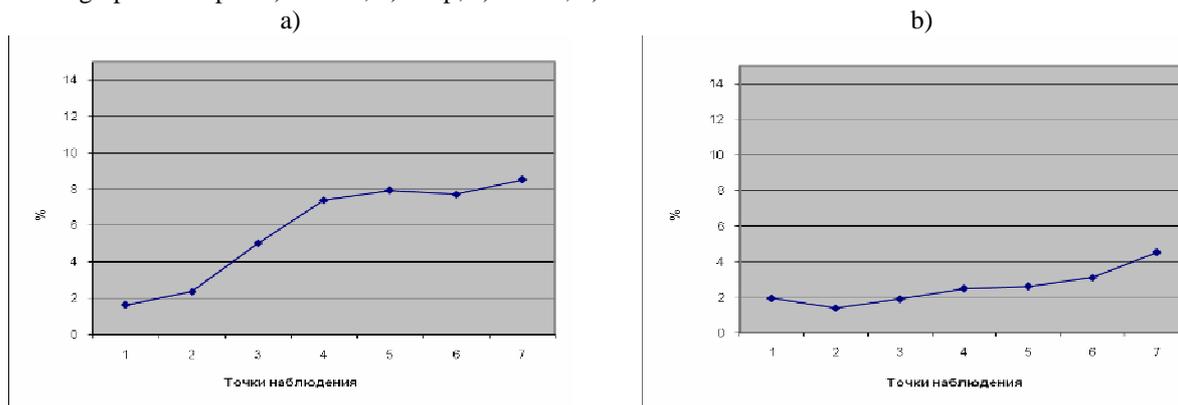


Fig. 2. Example of *Treash* feature graphs for varying (a) and not varying (b) objects

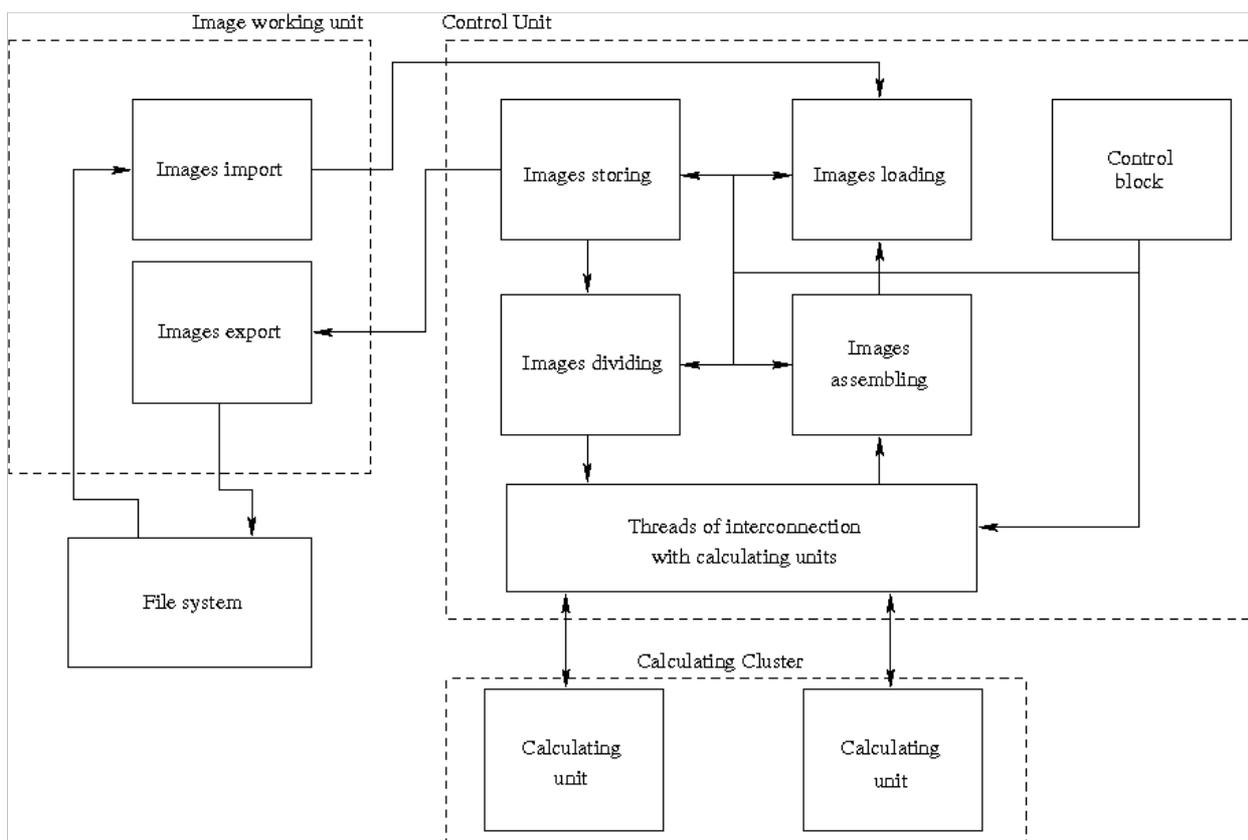


Fig. 3. Generalized structural schema of program system

The automated algorithm of analysis of changes dynamics includes the following steps:

1. Image segmentation and segments sizes calculation.
2. *Treash* feature of the disease calculation.
3. Average of the feature values, which were calculated in different vantage points, in order to remove random noises.
4. Building graph of *Treash* feature.
5. If the feature value has increased during last three vantage points, the segment is considered as object and the first such point is set as the reference point.

5. Calculations parallelization

Generalized structural schema of program system is shown at Figure 3. It contains three main units:

1. Image working unit. At this unit tools allowed storing, processing and keeping up integrity of geographical data, represented as raster images, are contained. It contains followed elements, which work with file system directly:

1.1. Unit of images import. This unit is destined for translation of raster images stored at the file system to system internal format.

1.2. Unit of images export. This unit is destined for translation of system internal format to raster images.

2. Control unit. This unit is destined for realizing of coordination calculating process, dividing of source image to parts, assembling result of image processing from parts. Unit contains three elements.

2.1. Calculations control. This unit is destined for realizing of control at stages of image processing.

2.2. Loading of data/Storing of result.

2.3. Dividing/Assembling of image

3. Calculating cluster. At this system unit parallel processing of source image parts is realized.

3.1. Processing unit. This unit realizes directly data processing.

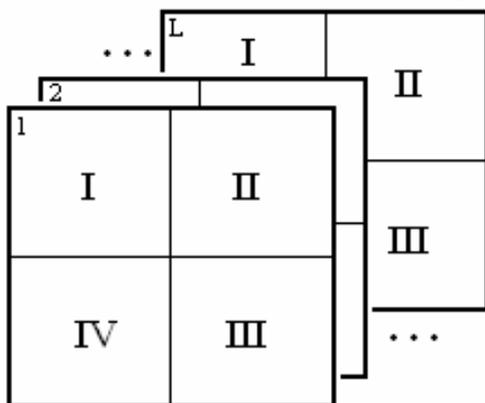


Fig. 4. Example of simple image dividing

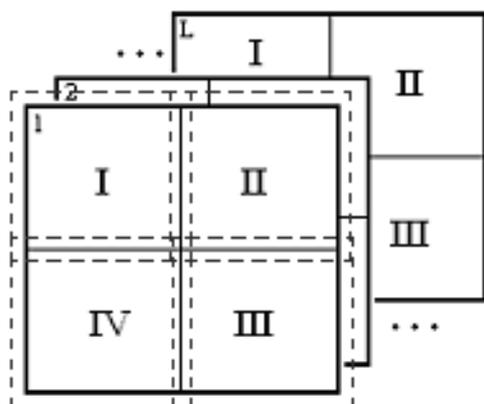


Fig. 5. Example of dividing with taking into consideration environment of image element

Image processing is realized in the following way:

Source image is divided to parts, each of them contains some section of original multispectral image. Example of simple dividing is shown at the Figure 4.

Some image processing algorithms requires taking into consideration environment of image element. In this case image element environment should be added to this element, as it is shown at the Figure 5.

After dividing image to parts processing stage follows. At this stage realizes parallel multispectral image processing by parts.

Processing technology MPI (Message Passing Interface), which is explained in detail in [9], was selected as basis of parallel. This technology was selected because it allows simple organizing interaction of calculating units and them synchronization.

TCP connection, which allows transmit required data without losing, is used for data exchange between calculating and control units. At the same time interaction must be equal for each connection with corresponding calculating unit. Also this interaction must happen parallel for decreasing of performance losses when realizing interaction. It can be realized by allocation of individual thread for service each data exchange connection between corresponding calculating and control units. Multithreading using in GNU/Linux environment is explained in detail in [11].

Schema of system working is shown at the Figure 6.

Tests of system of processing of images were made on a supercomputer «SKIF K-1000». At carrying out of experiments it was involved from 1 up to 64 computing nodes for color images in the size 2000x2000 and 1000x1000 pixels. Dependence of an operating time of system on a number of the involved computing nodes is resulted on the graphs shown on Figure 7.

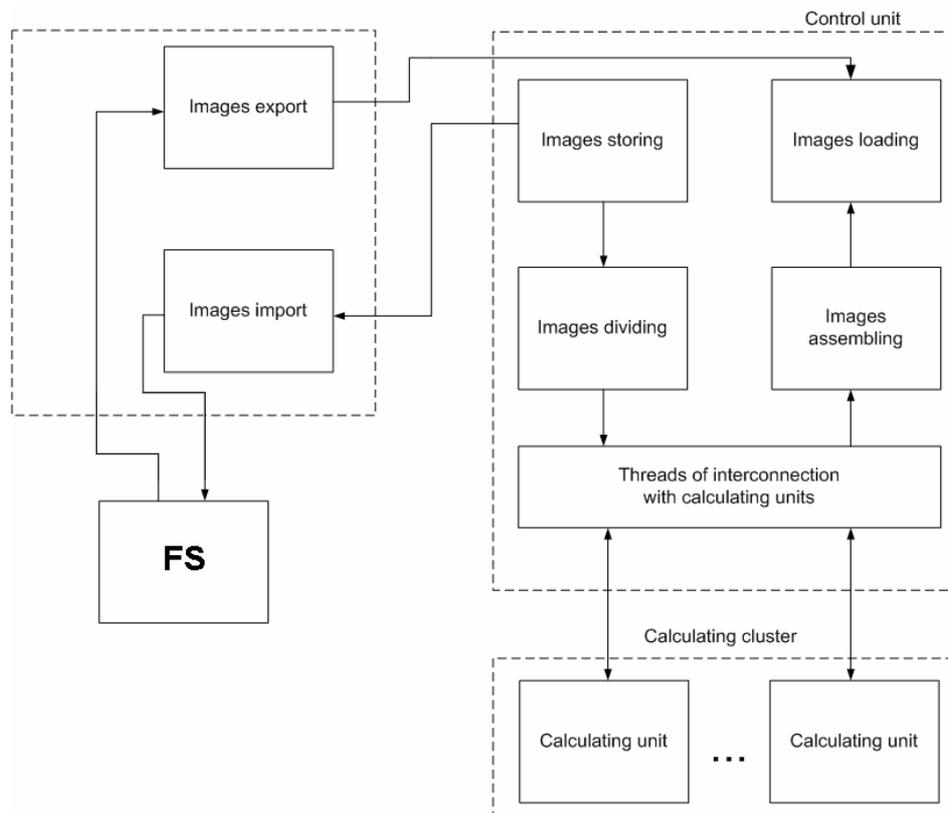


Fig. 6. Schema of system working

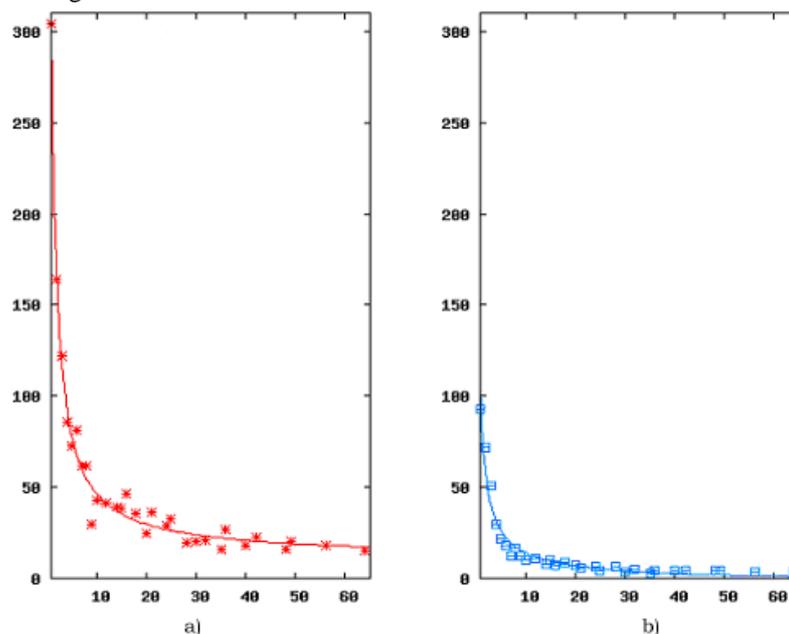


Fig. 7. Graphs of dependence of an operating time of system from a number of the involved nodes for color images in the size 2000x2000 (a) and 1000x1000 (b) pixels

6. Conclusions

The proposed algorithm of automatic definition of appearance of changes has shown good result of objects identification at use of an attribute of change of color characteristics of object [11]. The greatest influence on job of a method renders: presence in the staff of extraneous subjects and the shadows, having color of object; non-uniformity of illumination that creates additional handicaps. Experimental results have shown, that on the average on all changed objects appeared possible to reveal occurrence of changes on the fourth supervision.

The software for construction of the circuit of change of

the color, realizing the proposed approach to identification of required sites and selecting segments with strong changes, average changes and usual is developed. To each type of segments the factor of change, for example, strongly changed – $K1 = 100$, middle – $K2 = 50$, and usual – $K3 = 0$ is appropriated. In a window of the size 11×11 points the a number of points of segments of the first type $C1$, the second – $C2$, and the third - $C3$ is counted up. The factor of changes for the central point of a window under the formula $f = \frac{C1 \cdot K1 + C2 \cdot K2 + C3 \cdot K3}{C1 + C2 + C3}$ pays off. The given procedure repeats for all possible positions of a window. In result we

receive the circuit of change of color.

7. References

- [1] Belyaev B.I., Katkovsky L.V. Optical remote sensing. – Minsk: BSU, 2006. – 455 p.
- [2] N. Kumar, S. Pandey, A. Bhattacharya, and P. S. Ahuja, “Do leaf surface characteristics affect agrobacterium infection in tea [camellia sinensis (l.) o kuntze]?” J. Biosci., 2004. – Vol. 29, No. 3. – P. 309–317.
- [3] P. Soille. Morphological image analysis applied to crop field mapping // Image and Vision Computing. – 2000. – Vol. 18, No. 13. – P. 1025–1032.
- [4] Panagiotis Tzionas, Stelios E. Papadakis, Dimitris Manolakis. Plant leaves classification based on morphological features and a fuzzy surface selection technique // 5th Int. Conf. on Technology and Automation ICTA'05, 15-16 October 2005 Thessaloniki, Greece. – 2005. – P. 365-370.
- [5] Margarita Torre, Petia Radeva. Agricultural-Field Extraction on Aerial Images by Region Competition Algorithm // Int. Conf. on Pattern Recognition (ICPR'00), September 3-8, 2000, Barcelona, Spain. – 2000. – Vol. 01, No. 1. – P. 1313-1316.
- [6] A.V. Inyutin. The algorithm of image segmentation by grayscale pseudo-skeleton // Proc. of the III Int. Conf. on Neural Networks and Artificial Intelligence (ICNNAI 2003), November 12-14, Minsk. Belarus. – 2003. – P.263-265.
- [7] Apan, Armando and Kelly, Rob and Jensen, Troy and Butler, David and Strong, Wayne and Basnet, Badri. Spectral Discrimination And Separability Analysis Of Agricultural Crops And Soil Attributes Using Aster Imagery // In 11th Australasian Remote Sensing and Photogrammetry Conference, 2-6 September, Brisbane, Queensland. – 2002. – P. 396-411.
- [8] Burks, T.F., S.A. Shearer and F.A. Payne. Classification of weed species using color texture features and discriminant analysis // Transactions of ASAE. – 2000. – Vol. 43(2). – P. 441–448.
- [9] Antonov S.A. Parallel programming using MPI / S. A. Antonov – Moscow, MSU, 2004. – 71 p. [In russian]
- [10] A. Petrovsky, A. Doudkin, V. Ganchenko, M. Vatkin. Potato detection using color leaves characteristics. // Pattern Recognition and Information Processing (PRIP'2007), May 22-24, Minsk, Belarus, 2007. – P. 83–88.
- [11] Mitchell M., Advanced Linux Programming / Mark Mitchell, Jeffrey Oldham, Alex Samuel – Indianapolis: New Riders Publishing, 2001. – 368 p.