

## Scenario Approach for Image Processing in Smart City

*Filipp Galiano, Nataly Zhukova, Maksim Pelevin*

(Dr. Filipp Galiano, SPIIRAS-HTR&DO Ltd, Russia, 199178, St. Petersburg, 14 line, 39, galiano@oogis.ru)

(Dr. Nataly Zhukova, SPIIRAS, Russia, 199178, St. Petersburg, 14 line, 39, gna@oogis.ru)

(Mag. Maksim Pelevin, SPIIRAS, Russia, 199178, St. Petersburg, 14 line, 39, maks.pelevin@oogis.ru)

### 1 ABSTRACT

Modern city is described by fast change of a situations and large volume of available data about it. One of principal data types are images (satellite pictures, video from surveillance cameras, etc.). Fast manual analysis of such data is often complicated or impossible. Widespread practice is development of analysis system for each separate problem (for example identification of car identification numbers or QR-codes) that leads to a number of difficulties; in particular, it is difficult to integrate and modify such systems. In article the structure of raster image analysis system (both single images and a video stream) is proposed. The system is developed taking into account peculiarities of input data and requirements imposed by the smart city. As a basis of system the set of algorithms, that implements all image analysis stages (preliminary processing, construction of feature space and classification) is proposed. That algorithms are elements of data processing scenarios. Usage of scenario approach, based on subject domain ontology, allows to significantly simplify adaptation of image processing system to the peculiarities of subject domain without limitation of allowed algorithms set. A example of application of the developed system for real-world problem is given.

### 2 BASIC CONCEPTS

“In the future, cities will account for nearly 90% of global population growth, 80% of wealth creation, and 60% of total energy consumption“ (MIT). In addition, widespread use of surveillance systems recovers the problem of processing heterogeneous data. This data almost always has geospatial part. The idea of IGIS (Popovich, Voronin, 2005) forms a theoretical base for analysis this big heterogeneous data. Significant part of the data are videos and still images. Smart cities make demands on image analysis systems, because they are characterized by:

- Different problems and demands of analysis speed and accuracy.
- Considerable volume of heterogeneous data.
- Distributed analysis.
- Inequality of demands.

This data can be later processed by more complicated and deeper analysis for different purposes, but this problem isn't described in this paper.

The problem of image analysis is reviewed below. We consider that analysis problems have 3 objected included to it:

- Recognized image — set of images that need to be detected.
- Training set — set of objects (instances) with its classes, that's should be selected on images.
- Recognition — set of operations, running on recognized image and training set. These operations set a class for every pixel or group of image's pixels.

Today there's no solid theory of digital image analysis, therefore analysis system should include wide set of d processing methods, based on different properties of objects of interest (Gonzales, Woods, 2007).

Every analysis system that works with digital images includes different set of operations of image processing. Most part of these operations can be divided in next classes:

- Image preprocessing.
- Selecting informative features — segments, contours, points and etc.
- Creation of description space (usually— feature space) (Potapov, 2007).
- Classification of the description space elements.

Every step of analysis can use different set of algorithms and entire process can be presented as these algorithms sequence. We discuss more detailed description of basic concepts of the scenario approach in section 3, classification of the analysis algorithms in section 4, and set of used algorithms in section 5.

### 3 BASICS OF SCENARIO APPROACH

The scenario can be defined formally as sequence of stages and decisions (Popovich and others, 2013). The stage is a set of the elementary actions which are carried out in sequence or in parallel. The decision represents a point in which the process can change the direction depending of some conditions which take place in this moment.

The elements, corresponding to actions, can represent both separate algorithms and the nested scenarios. Elements of the decision include one or several logical rules which are used for a choice of process's branch in the presence of several possible options. Usage of both strict and fuzzy logic is allowed.

Usage of scenario approach allows:

- to describe difficult, hardly formalizable or even not formalizable spatial processes, which assumes interaction of a large number of technical and natural objects;
- to specify descriptions of processes just before the beginning of their execution and during their execution without restarting;
- to have vivid graphical representation of described processes;
- to carry out parallel data processing. Some possible alternatives of parallel processing can be considered: parallel performance of various stages of processing for various data sets, application of the parallel computing algorithms, parallel processing of separate fragments of one data set.

As a rule, work with scenarios assumes:

- development of scenarios using knowledge of experts from different subjects domains;
- formalized description of the created scenarios;
- definition of conditions under which scenarios can be applied;
- implementation of scenarios;
- qualitative and quantitative evaluation of the data, obtained as a result of whole scenarios execution or execution of their separate fragments.

For development of scenarios specialized graphic tools, which allow to describe the general structure of scenarios and their separate elements are used. Such tools also provide means for debugging the scenarios, allowing revealing and eliminating possible errors.

The formalized description of scenarios assumes that:

- All concepts, used for description of processes (types of objects, types of actions, etc.) are defined in the dictionary of subject domain;
- The format of the scenarios description conforms to standards (for example, BPEL(Web Services Business Process Execution Language));
- The logical rules, that are set for elements of decisions elements, aren't inconsistent;

Descriptions of the restrictions, imposed on use of scenarios, and also conditions of scenarios application are intended for end users in subjects domains and are used for scenarios choice. The developed scenarios are stored arranged in the library of scenarios which is as an element of domain's model description. Execution of scenarios requires three principal components:

- The engine of processes execution - provides execution of actions according to the given sequence;
- The inference engine - allows to make decisions in the presence of alternative options of processes execution;
- The graphic interface - allows to display processes, described scenarios and to control the flow of their execution;

The assessment of results of scenarios execution is based on of an assessment of results of operation of separate algorithms. If the received results not meet requirements, the cycle of scenarios execution can be repeated partially or completely.

In case of scenarios creation for image processing it is necessary to define:

- list of algorithms for each processing stage;
- sequence of algorithms execution for each stage;
- list of the computing procedures applied in algorithms;
- parameters of execution for each algorithm;
- the criteria, applied for an assessment of algorithms results.

The heterogeneous structure of input data and tasks at hand, and also presence of various factors which can have impact on output data, result in need to develop a large number of scenarios. Besides, it is often needed to delete earlier created scenarios, and also develop of some new.

This problem is solved due to use of the generalized scenarios which describe general processing stages. Detailing of scenarios takes place during the execution of tasks, after necessary additional information is available. As a rule, context-dependent information is part of additional information. For example, on the basis of a signal / noise ratio assessment can be defined the filter, used at preliminary preprocessing stage or temporal and frequency scales at which image processing is carried out.

For setup of scenarios to the solution of tasks in different conditions it is necessary to have possibility to specify the boundary values applied in logical rules. For example, in case of processing of seabed images, obtained from unmanned underwater vehicle (Zhukova, Pankin, 2013) boundary values of such properties as average brightness depends from a relief of a seabed and can significantly differ. In some cases, for calculation of boundary values additional computation can be demanded, and for specification of boundary conditions the results of statistical processing of historical data can be used.

During a execution of scenarios the set of the used software modules, that implements algorithms, is defined by a set of available services, and also computing resources. The composition of services and resources, as a rule, changes dynamically. As such, the choice of algorithms implementation is carried out before execution of scenarios. The description of scenarios contain names of algorithms according to the nomenclator and the description of their parameters; explicit links to algorithms libraries in scenarios should not be specified.

Thus, for the solution of image processing tasks in the applications developed for Smart Cities it is desirable to use the generalized scenarios formalized in the form of templates which are detailed and set up before the solution of application-oriented tasks.

It is necessary to mention that use of scenario approach for data processing also allows:

- to create new algorithms on the basis of a combination the existing algorithms;
- to execute different variants of data processing and - on their basis - create final decisions.

For a long time scenario approach is successfully used in intellectual geoinformation systems - IGIS (Popovich and others, 2013). Using IGIS for working with scenarios gives a set of important advantages - first of all, IGIS collect actual data about the environment, secondly, provides GIS interface, convenient for the end user, in the third, artificial intelligence techniques are integrated into IGIS, in particular, the inference engine and the expert system, used for implementation of scenario approach.

#### **4 CLASSIFICATION OF DATA ANALYSUS ALGORITHMS**

It is needed to develop onotology of image analysis algorithms for using algorithms of remote sensing in term of scenario approach. The basic to this ontology is classification of the subject domain by its main features. There are list of objects, that are subjects of this classification:

- Data processing and analysis algorithms.
- Variants of formal description of raw, intermediate and output data.
- Data and algorithms' results estimation.
- Tarfget objects.

Analysis algorithms classification is the most important part, because these algorithms are used to analysis sequence of data. Data analysis algorithms can be classified by next features:

Feature 1. General analysis process step. According to this algorithms can be divided as shown in table 1.

Feature 2. For algorithms of 1 and 2 classes - image type, which can be used by these algorithms (multispectral, hiperspectral, binary). It is important to expand the scope of application of the algorithms, that requires their generalization to process hyperspectral images, which is a non-trivial task.

Feature 3. For algorithms of class 2 – type of objects, which can be extracted by algorithm, such as points, contours and segments (see table 1).

No.	Algorithm class	Subclass	Input	Output	Settings	Examples
1	Preprocessing	Spatial processing methods	Raster image	Raster image	Level of image approximation, etc.	Median filtration, brights average, etc.
		Frequency domain methods	Raster image	Raster image	Level of image approximation, etc.	
2	Extracting infomative features	Segmentation	Raster image	Image model (result of image partitioning)	Level of image approximation, etc.	Statistical Region merging (Nock, Nielsen, 2004), flood fill, Full-Lambda schedule (Robinson, Redding, Crisp, 2002), etc.
		Points selection	Raster image	Image model (set of points and its properties).	Algorithm-specific.	SIFT / SURF descriptors
		Edge detection	Raster image	Image model as a set of contours.	Algorithm-specific.	Canny edge detector
3	Features space construction	-	Image model	Description space (Potapov, 2007)	Algorithm-specific, frecuently absent.	Calculation of segment properties—geometrical, bright, contextual, etc.
4	Classification	-	Description space, training set	Classes tags for model elements.	Algorithm-specific.	Nearest neighborho od.

Table 1. Algorithm classification by analysis process step.

Origin data (images) can be classified by next features:

Feature 1. For remote sensing — spatial resolution (meter per pixel or pixel per meter).

Feature 2. Image size in pixels.

Feature 3. Spectral band.

Feature 4. Color depth (radiometric resolution), defined by bits per pixel for one spectral band of image.

Feature 5. Capturing time and geospatial data.

Feature 6. Images can be divided to static and dynamic (videos).

Feature 7. Data format. They can be divided to lossy (jpeg, jpeg2000, etc.) and lossless (png, bmp, etc.) compression.

Feature 8. File size.

Also there are some specific videostream features (bit rate and others).

Formalized representation of data is different depending on data type. Raw (input) image can be represented by sequences of matrices (two-dimension array) with non-negative values. Value of array element is visualized as pixel brightness.

Types of image's model storage can differ essentially (see table 2). In case of segmented image, when it's pixels are arranged to group (segments), dynamic trees can be used as data structure for storage (Sleator, Tarjan, 1985; Tarjan, 1975). In this case if levels of segmentation in different levels of approximation compose strict hierarchy, then this hierarchy can be stored in indexed tree (Kharinov, 2006). It store different levels of hierarchy with low memory usage. In cases when object of image cannot be presented as solid segment "marked segments" can be used for selection instead. Marked segments are specific for each class objects; extract such segments is equal to extracting whole object. Contours of image can be presented as pixels set, i. e. segments. It is needed to keep in mind, that image segment shouldn't be coherenced. Some points can be stored in one-dimension array, where index is their linear coordinates. The formulas for mapping two-dimensional pixels coordinates to one-dimensional array are evident.

Objects class	Subclass	Type of formalized representation	Advantages	Disadvantages
Single-band image	Raw image	Array of integer numbers	No accumulation of errors for many operations	The need to check errors accumulation for many operations
		Array of real numbers	Can be used as a temporary representation for many operations	Cannot be visualized directly
	Segmented image (one level of segmentation)	Assosiative array (key – coordinae of pixel, value – pixel bright)	Memory saving due to store only meaningful pixels.	Slow segments union operations.
		Disjoint-set data structure	High speed of segments union and enumeration operations.	Complex software implementation.
	Segmented image (multi-level hierachic segmentation)	Multi-dimensional assosiative array	Simplicity of software implementation.	High memory consumption
		Ierarchic tree	High speed of segments union and enumeration operations.	Complex software implementation.
Feature space	Explicit representation	Two-dimensional array of properties	Can be used for analisys of relations between properties	Computationally constly.
	Implicit representation	One-dimensional array (indexed tree)	Economy of computaional resources.	Low classification accuracy.

Table 2. Image representation at different stages of analisys.

Types of formalized representation of feature space are classification algoirhm-specific. Volume of training set is often less than the number of image pixels, that lowers demands to data structure. In most cases using matrices of features vectors is smart choice. These matrices are stored as two-dimensional array where real numbers are values. Using indexed trees for storage segments features (contours) on some level of hierarchy

is another option. Numbers of elements in features vectors in training set is equal to number of image pixels (in the worst case), and dimension of each feature vector is production of spectral band number by number of features, calculated for each band. If classification algorithm, during training process, works with each dimension of the feature space one-by-one, then we can reduce demands to CPU, because we do not need to storage the whole training set. But algorithms of these type often demonstrate worse classification results than algorithms that analyse features correlation.

There are many books about estimatione criterias for input data, for example the classic book (Pratt, 1978). Result of algorithms estimation is calculated twice:

- directly after algrorithm stops
- on next steps of processing and analysis.

Method of algorithm result estimation is iterative. When analysis can not be done, the algorithm returns to previous step of processing. We use following methods in case of direct analysis:

- comparison of given results with actual information,
- comparison with results, calculated by different methods,
- comparison with result of historical data processing,
- comparison with results of combinated analysis methods,
- expert estimate (can be given automaticly with rules, that expert forumated); calculation of estimate by total calculated criterias.

The data sources can be classified by multiple parameters, for example:

- an area of basing (earth, air, space),
- a volume of transmitted data,
- noise-immunity (Verba and others, 2010), etc.

Desired objects can be classified by different features. The simplest classification is dividing object to natural and artificial. Extracting of artificial objects are more complicate, however for smart cities purposes they are more interesting. While analysis of remote sensing data we can divide objects to (Verba and others, 2010):

- Dotted objects (less than 1 pixel),
- Group objects (is around 1 pixel),
- Distributive object (more than 1 pixel).

This classification is base of ontology, that provides capability of developing image analysis scenarios without using algorithms restrictions.

## 5 BASIC SET OF ALGORITHMS

For the user of developed system the most important parameter is set of using analysis algorithma. We selecting algothms for system prototype based on the following criterias:

- High speed of analysis is nessesary condition for videostream analysis and huge remote sensing data volumes.
- For preprocessing algorithms - no “bluring” of the image, adjustable level of noise reduction and possibility of light inequality compensation.
- For algorithms of feature extracting (primarily segmentation):
  - Hierarchy segmentation (lets us to use the dynamic trees).
  - Monochrome and color image segmenation.
  - High quality of the segmenation, estimated directly or by system work results.
- For algorithms of feature calculation:
  - Use of different segment features: geometry, brightness, spectral, context, etc.

- Using dynamic trees as form of storage feature values.
- Possibility of image analysis with arbitrary number of spectral intervals.
- For classification algorithms:
  - Possibility of analysis feature vector with high dimension.
  - Possibility one-by-one feature analysis or including interrelations.
  - Adjustable accuracy of analysis.

According to this set of criteria for developed prototype following algorithms were chosen:

- Median filtration, set of histogram changing algorithms, including invariant image representation (Kharinov, 2006).
- Segmentation algorithms: Statistical region merging (Nock, Nielsen, 2004), «Flood fill», Zero Lambda Schedule – modification of Full-Lambda Schedule with  $\lambda=0$  (Robinson, Redding, Crisp, 2002).
- Algorithms of properties calculation - (dozens of algorithms), including, but not limited to: average bright, perimeter, square, compactness (Gonzales, Woods, 2007), variance of brights, axis moment (Shapiro, Stockman, 2001), etc.
- Classification algorithms: one-dimensional linear classifier (threshold processing), nearest neighborhood, SVD-classifier (Galiano, 2010), «naive» Bayes classifier (Duda, Hart, Stork, 2001), etc.

Based on this approach a prototype of image processing system based on an intelligent GIS was developed.

## 6 CASE STUDY

### 6.1 Ships detection

One of tasks which can be solved by means of the analysis of Remote Sensing Data with high resolution is selection of the objects on water surface (vessels). Data at hand are characterized by the big size (in pixels), however requirements to analysis speed here are rather insignificant. The task of vessel selection, rather simple in a case when the vessel is in open water, becomes much more difficult if the vessel is at the mooring.

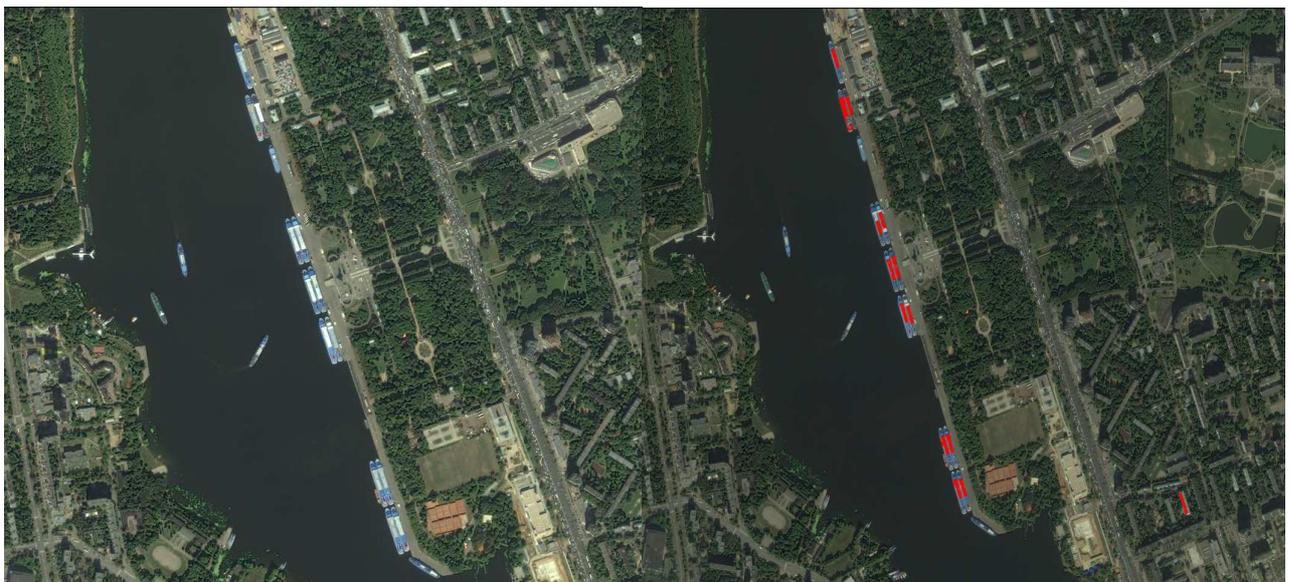


Fig. 1. The source image (at the left) and the result of its analysis (on the right) – selection of vessels of the given type on Remote sensing Data provided by the SCANEX (SCANEX Research and Development Center). In the picture the Krasnogorsk city is showed. The handle segments of the found objects are marked with red color.

However, for certain ships classes this task can be solved using a combination of standard analysis algorithms, as shown below. The purpose of the analysis was selection of vessels with big size which are at the mooring. As an algorithm of preliminary processing median filtering was used, for segmentation we use

SRM, for classification – simple linear classifier. For speed-up the recognition (but not training) we use single-level segmentation, where value of segmentation level was defined as an arithmetic average of all levels at which the object was selected during training. As "object markers" we use segments, uniq for this ships class. The list of segment properties, used for classification, includes the area and segment perimeter (in pixels), average brightness of a segment and dispersion of brightness, and also a segment inertia moment round the principal axis (Shapiro, Stockman, 2001). Total time of the analysis using normal PC is about one one minute for one color image about 10 megapixels in size (see fig. 1). In the right lower part an false ship is detected; this result can be improved using cartographical information.

## 6.2 Realtime selection of region of interest based on unmanned aerial vehicles

Other task of considerable interests is selection of region of interest on video stream, taking into account the spectral band of video. In particular, a unmanned aerial vehicles (UAV) video, see a figure 2, can be used as data sources. When processing a video stream high requirements to analysis speed are imposed, that restricts the possibilities of use computationally costly algorithms. However, use of indexed dynamic trees (Kharinov, 2006) allows to quickly analyse image at several levels of segmentation that permits to receive satisfactory analysis results.

The analysis was made for video received only in an infrared range. The analysis of video impose high requirements to the processing speed, however a separate frame it is significantly smaller, than typical Remote Sensing Data ( $720 \times 576$  pixels, in this example). As algorithm of segmentation the modified algorithm of "flood fill" was used. It was modified by creation of the minimum spanning tree, and it gives the possibility to achieve essential increase in speed of segmentation (Galiano and others, 2013). The small feature set of a segment (average brightness and the square) was used; and as algorithm of classification simple linear classifier was used. Recognition take place in the RGB color space; before the analysis median filtering was executed and the linear normalization of the histogram to all range of brightness [0; 255] was performed. In the stage of recognition one level of segmentation was used. For training stage a set of supply algorithms was developed which are creating the training set and performing selection of algorithms parameters on the basis of the image classified by the expert. For analysis of one image the normal personal computer was used, and the analysis takes from 200 milliseconds to 1 second, depending on a used feature set. For an assessment of algorithm's accuracy part of a video stream was partitioned into separate images. The assessment of accuracy is given in table 3.

Parameter	Value
Total number of images	1931
Total number of opixels	800824320
Absolute number of correctly classified pixels	770279811
Absolute number of false positives	29340306
Absolute number of false negatives	1204203
Fractional number of correctly classified pixels	0.9618586645820147
Fractional number of false positives	0.03663763108493009
Fractional number of false negatives	0.00150370433305522
Variance of right classified pixels	1.9063211494840804E-4
Standart deviation of right classified pixels	0.013806958931944719
Total time of teaching and recognition, milliseconds	1515482

Table 3. Assessment of objects detection accuracy, measured for pixels.

## 7 CONCLUSION

The framework, offered in article, allows to create analysis systems for single images and video data rather easily. We outline that program implementation of processing algorithms can be replaced by arbitrary, while saving interfaces that allows to use third-party libraries, in particular – widely known library OpenCV (OpenCV). The received results can be applied in development of image and video processing systems of and video for smart cities. Further development of the offered prototype requires development of full

ontology, including the expanded description of subject data (Turusova, 2007), in particular processing of sound information (for voice recognition) and geospatial information.

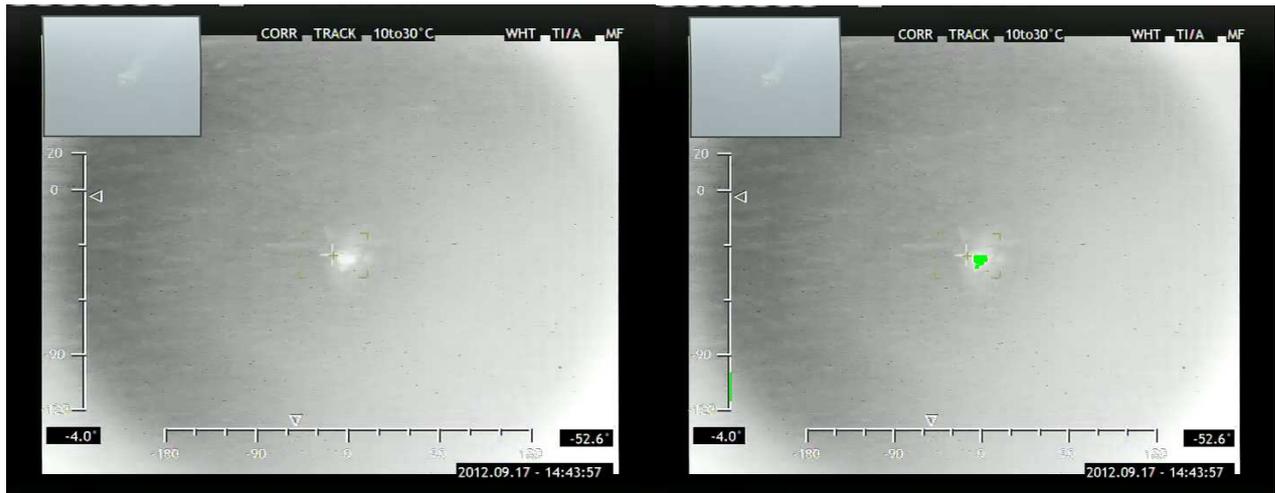


Fig 2. Analysis of data from the Camcopter S-100 UAV. At the left – input data, on the right – result of selection of marker segments (they are marked by green). Despite poor quality of a picture the required object is selected with an accuracy of 98%. False operation in the lower left corner is caused by presence of additional technical information and can easily be removed, if necessary.

In addition, an important direction for further development of the proposed approach and the prototype is the usage of inference engine for creating a chain of analysis algorithms. When operating in this mode for train the system to recognize a certain class of objects one need to set only their formalized description and the sequence of steps for objects detection will be generated automatically. The proposed approach allows to implement the idea of a "strong learning" where the learning process consist not only in changing the number values (weights of elements, training samples representation, etc.), but the structure of the system (Potapov, 2007). This opens up opportunities for the use of the system in various subject domains in fully automated mode.

## 8 REFERENCES

- GALIANO, F.R.: Classification algorithm for section of Earth's surface based on the singular value decomposition of matrices. In: Information technologies. Vol.12, pp. 35-37. Saint Petersburg, 2009.
- GALIANO P.R., Pelevin M.S., Ermolaev V.I., Thvetkov M.V. The fast computation of hierarchic image segmentation for unaccumulative segments union criterions while processing ice situation data. In: SPIIRAS Proceedings. Vol. 30. Saint Petersburg, 2013.
- ZHUKOVA N.A., Pankin A.V.: Unmanned underwater vehicles as a component of decision support system. In: Proceedings of Fifth All-Russian Scientific and Technical Conference "Technical problems of reclamation of the world's oceans". Vladivostok, 2013. GONZALEZ RAFAEL C., Woods, Richard E.: Digital Image Processing. USA, 2007.
- KHARINOV, M. V.: Storage and Adaptive Processing of Digital Image Information. St. Petersburg, 2006. Massachusetts Institute of Technology, <http://cities.media.mit.edu/about/cities>.
- NOCK, R., Nielsen, F.: Statistical Region Merging. In : IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, pp. 1–7, 2004.
- OpenCV, <http://opencv.org/>.
- POPOVICH, V. V. and others.: Intelligent geographic information systems for monitoring the maritime environment. Sankt-Peterburg, 2013.
- POPOVICH, V., Voronin, M.: Data harmonization, integration and fusion: three sources and three major components of geo information technologies. In: Proceedings of IF&GIS, International workshop, pp. 41-46. Saint-Petersburg, 2005.
- POTAPOV A. S.: Pattern recognition and machine perception: A general approach based on the principle of minimum description length. St.-Petersburg, 2007.
- PRATT, WILLIAM K.: Digital image processing. New York/Chichester/Brisbane/Toront, 1978.
- DUDA, RICHARD O., Hart, Peter E., Stork, David G.: Pattern Classification. November 2000.
- ROBINSON, D. J., Redding, N. J., Crisp, D. J.: Implementation of a fast algorithm for segmenting SAR imagery. Australia, 2002. SCANEX Research and Development Center, <http://www.scanex.ru/en/index.html>.
- SHAPIRO, LINDA G., Stockman, George C.: Computer Vision. USA, 2001.
- SLEATOR D.D., Tarjan R.E.: Self-Adjusting Binary Search Trees. In: Journal of the ACM. Vol. 32, pp. 652–686. USA, 1985.
- TARJAN R.E.: Efficiency of a Good But Not Linear Set Union Algorithm. In: Journal of the ACM. Vol.22, pp. 215–225. USA, 1975.
- TRUSOVA, YU.: Image analysis thesaurus. Version 1.0. In: Pattern Recognition and Information Processing. Proceedings of the 9th International Conference. - Vol.2, pp. 210-214. Minsk, 2007.
- VERBA, V.S., Neronskiy, L.B., Osipov, I.G., Turuk, V.E.: Space-borne Earth Surveillance Radar Systems. Moscow, 2010.
- Web Services Business Process Execution Language, Version 2.0 OASIS Standard. <http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.pdf>, 2007.