Techniques for Improving Color Segmentation in the Task of Identifying Objects on Aerial Images

Denis Kasimov
Kalashnikov Izhevsk State Technical University
Izhevsk, Russia
kasden@mail.ru

Abstract—**Automatic identification of objects on aerospace images allows to increase the efficiency of making the necessary management decisions in important areas of human activity. A promising approach is the object-oriented image analysis, in which image segmentation is performed and the resulting image regions are classified into target object categories. The main problem constraining the effectiveness of this approach is the lack of segmentation accuracy. To solve this problem, the paper proposes techniques aimed at obtaining more relevant color segments of the image: partition of the image into frames with subsequent merging of color areas lying on the borders of adjacent frames, splitting color regions in relatively narrow places, as well as adaptive approximation of the edges of color areas. An experimental study of improving the quality of identification of objects as a result of the application of the developed techniques is carried out. The experiments were conducted on high resolution aerial images from a publicly available dataset. It is shown that the proposed techniques make a significant contribution to improving the efficiency of the logical approach to the identification of objects based on shape features.**

I. Introduction

Creation of methods and systems for automatic identification of objects on aerospace images is of great practical importance due to the fact that manual labeling of images is very slow and costly and is not capable of quickly providing decision makers with relevant information in such areas as emergency situations, environmental protection measures, territory development planning, etc.

Among the existing methods of identification of objects on aerospace images, there is a class of methods that utilize low-level features of image patches, for example, the angular difference feature [1], the Curvature Scale Space corner detector [2], and the morphological building index [3]. The effectiveness of these methods has a certain limit, which is fundamentally impossible to overcome without extracting and analyzing the significant high-level information contained in the image.

Another popular direction in the field of automatic identification of objects in aerial images is based on the use of artificial neural networks [4], [5], [6]. The disadvantage of machine learning techniques is the need to prepare a significant amount of training data. As a rule, neural network methods achieve good performance indicators only on the class of images for which they were originally designed and trained. In addition, neural network methods are not able to form structural and spatial descriptions of objects and scenes, while this is one of the important stages on the way to understanding the image.

Currently, a promising approach, called object-oriented image analysis (OBIA) [7], [8], is being developed, in which decision making is based on the analysis of higher-level units of information, namely image segments (superpixels). The process of identification of objects in the image is carried out according to the following scenario: image segmentation, calculation of segment features, and classification of segments into target object categories using the rule base. The works [9], [10], [11] are examples of object-oriented methods.

The advantage of object-oriented methods becomes especially obvious in relation to high-resolution images in which spatial information begins to play a significant role. On the other hand, the OBIA approach faces a number of difficulties. For example, some researchers who develop OBIA solutions note that it is not possible to effectively classify objects in images using shape features [12]. This can be explained by the distortion of the boundaries of regions extracted from the image. The unstable nature of the shape of the color areas is a consequence of the imperfection of the methods of automatic image segmentation. In the field of object-oriented image analysis, the problem of redundancy and inaccuracy of the segmentation results remains a priority [13], which significantly complicates the analysis of objects and worsens the overall quality of identification. Ignoring the shape of objects is a forced measure rather than predetermined by its objective uselessness. It is obvious that the shape of the object is very important information for decision-making; but in the existing conditions, the standard shape features do not work very well, and it is difficult to create the complex ones. The result of this is that so far the OBIA direction is not able to fully realize its significant potential.

In the work [14], the segmentation method [15] was used to obtain more accurate image segments, taking into account not only the color statistics of the image, but also the geometric characteristics of the segments being formed. This approach allows to improve the quality of the resulting image segments to a certain level, but the initial conditions for the decision-making stage still remain very difficult.

This work is carried out within the framework of the project on the development of a cognitive approach to the problems of search, analysis and description of objects in images based on context-sensitive strategies of analysis and modeling of reasoning. The approach being developed can be attributed to the object-oriented class, based on image segmentation.

The purpose of this work is to develop and study methods for increasing the relevance of areas resulting from automatic color segmentation of an image. The less distorted and redundant the extracted color areas are, the easier and more effective the further logical analysis and decision-making.

In order to extract more accurate and informative color areas, the following pre- and post-processing techniques are proposed:

- partition of the image into frames;
- merging color areas from the adjacent frames;
- splitting color regions in relatively narrow places;
- adaptive approximation of the edges of color regions.

The influence of these operations and their parameters on the quality of the result of the detection of buildings on aerial images based on shape features is investigated.

II. Image processing stages

Let's consider the main stages of image processing and analysis, paying special attention to the proposed operations of color areas refinement.

A. Partition of the image into frames

In the same picture, there are often different types of terrain (forest, industrial area, residential area, etc.), objects of incomparable size and rare texture, as well as situations where objects are obscured by others. One of the effective ways to improve the accuracy of color segmentation in this case is to increase the degree of its locality. The effect of this technique is particularly pronounced with respect to large-size images. Fig. 1 shows examples of color segmentation results for the same part of an image in the cases of global and local processing.

Let W and H be the width and height of the original image in pixels, fW and fH be the width and height of the bulk of the target frames, fWm be the minimum width of the right frames, and fHm be the minimum height of the bottom frames. Then the original image of a large size is divided into many smaller frames:

$$IMAGE = \{\, p_{ij} \mid p_{ij} \text{ is a pixel}, i = 1..H, j = 1..W \,\},$$
$$FRAMES = \{\, F_{mn} \subseteq IMAGE \mid m = 1..M, n = 1..N \,\},$$

$$M = \begin{cases} 1, \text{if } H \le fH; \\ H \operatorname{div} fH, \text{if } H \bmod fH < fHm; \\ H \operatorname{div} fH + 1, \text{else}; \end{cases}$$

$$N = \begin{cases} 1, \text{if } W \le fW; \\ W \operatorname{div} fW, if\ W \bmod fW < fWm; \\ W \operatorname{div} fW + 1, \text{else}. \end{cases}$$

B. Color segmentation and detection of the edges of color regions

Each image frame undergoes color segmentation. The applied segmentation method is based on constructing the three-dimensional histogram in the color space HSV and searching the local maximums on the histogram by scanning the color space with a three-dimensional neighborhood analysis operator [16].

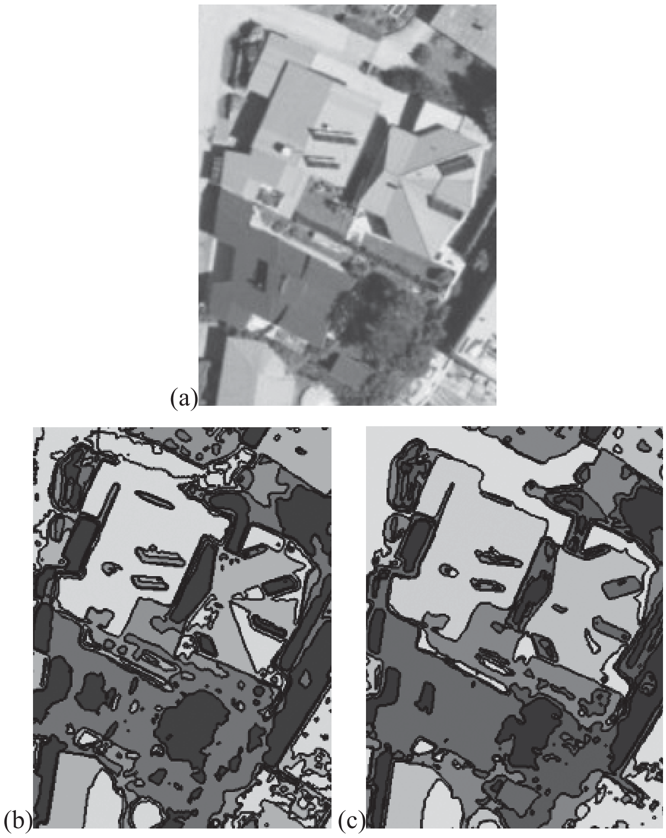


Fig. 1. Comparison of segmentation results: (a) fragment of the original image; (b) segmentation on a large frame; (c) segmentation on a smaller frame

Next, the edges of the resulting color regions are determined. A pixel belongs to the boundary of a color region if at least one of the adjacent pixels belongs to a different cluster of the color palette [16]. For convenience of further processing, the edges are represented in a vector form by cyclic lists of straight line segments, and their weak approximation is performed.

As a result of performing these operations, we have a set of color regions, each of which has a characteristic average color and an external geometric shape as a list of edge segments:

$$EdgeDetection(Segmentation(F_{mn})) = REGIONS = \{\, (A, c, E) \mid A \subset F_{mn}, c \text{ is a color value}, E = (l_1, \ldots, l_k) \,\}.$$

The resulting color areas correspond to some extent to the real objects present in the image, or their individual parts. Edge sections are ordered counterclockwise.

C. Merging areas from adjacent frames

The downside of partitioning an image into frames is the possible loss of some of the information about objects located directly on the borders of adjacent frames. The risk of missing or misclassifying targets at the decision stage increases.

Fig. 2 shows an example of the results of merging an object's areas that fall into adjacent frames.

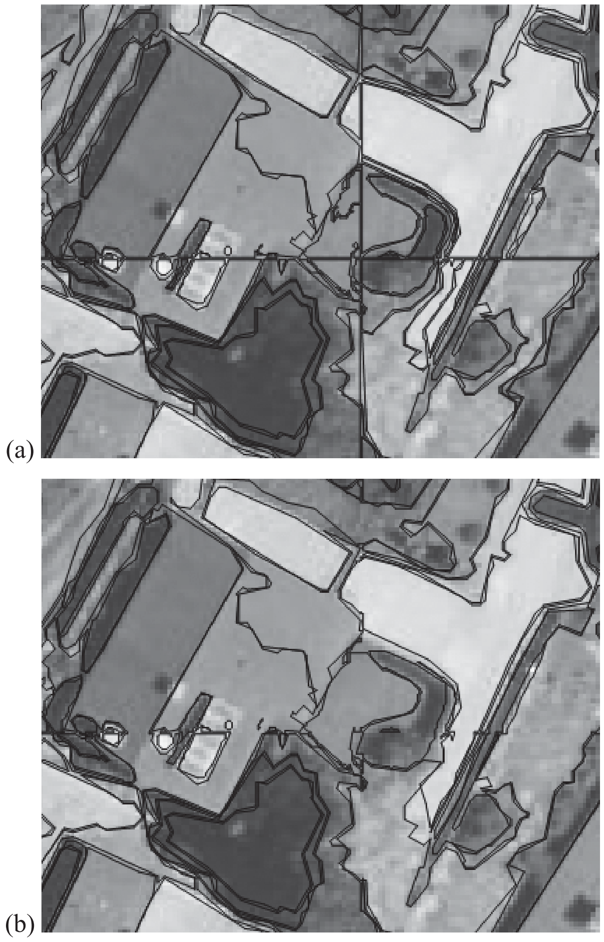


Fig. 2. Example of merging regions: (a) source areas; (b) result areas

Areas are merged by the developed Algorithm 1.

Algorithm 1 Merging image regions

1) Identify the areas $AdjRegions \subseteq REGIONS$ that have vertical and horizontal edge segments at the borders of the frames. The proximity of a segment to a frame's border is determined taking into account the tolerance threshold $\delta = 2$ px, compensating for deviations that could occur after the weak approximation of the regions' edges.

2) For each $R_1 = (A_1, c_1, E_1) \in AdjRegions$, $R_2 = (A_2, c_2, E_2) \in AdjRegions \setminus \{R_1\}$ lying in adjacent frames and having a similar color ($c_1 \cong c_2$) perform:

2.1) In the lists E_1 and E_2, identify the alternating sequence $K = (n_1, s_2, n_3, s_4, \ldots, n_k)$, $k \geq 1$ (see Fig. 3) of non-adjacent and adjacent sections located along the corresponding frame border, based on the operation of projecting the end points of the line segments onto each other. If one area exactly flows into another, the sections n_1 and n_k will be fictitious (of zero length).

2.2) If there are no adjacent sections, do not merge the regions.

2.3) Loop by adjacent sections (s_i):

2.3.1) Sample the pixel bands P_1 and P_2 located near the section s_i: $P_1 \subseteq A_1$, $Near(P_1, s_i)$, $P_2 \subseteq A_2$, $Near(P_2, s_i)$.

2.3.2) If the colors of the pixels at the borders of the frames are not similar: $AvgColor(P_1) \neq AvgColor(P_2)$, then do not merge the regions.

2.4) Perform the merge $R' = R_1 \cup R_2$, excluding the edges' fragments that assigned to the sections of the sequence K under indices from 2 to $k\text{–}1$.

2.5) Loop over all n_i such that $1 < i < k$:

2.5.1) Create a new area R'' by merging fragments of the edges E_1 and E_2, corresponding to the section n_i. Mark that the region R'' is nested in the region R'.

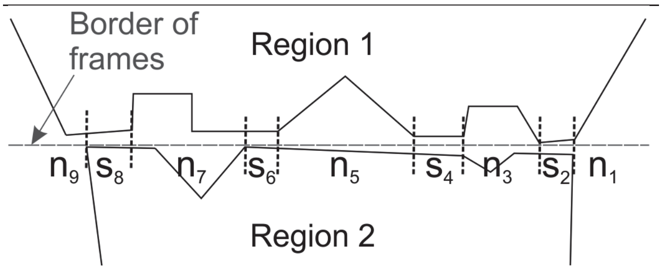


Fig. 3. Example of a sequence of non-adjacent and adjacent segments

The execution of the above algorithm is repeated as long as there are any changes.

To check the similarity of color values, complex color difference models can be used, for example, [17], [18]. However, in the considered problems the introduction of several simple and natural similarity rules is quite a workable way. For example, colors are considered similar if at least one of the following statements regarding their HSV components is true:

- both colors are dark ($V_1 \leq 27 \wedge V_2 \leq 27$);
- both colors are gray and of close brightness ($S_1 \leq 17 \wedge S_2 \leq 17 \wedge \Delta V \leq 20$);
- the difference between the components' values is small ($\Delta H \leq 30 \wedge \Delta S \leq 15 \wedge \Delta V \leq 15$).

The proposed merge algorithm may be useful not only for combining regions from different frames of the same image, but also for solving the problem of merging images of adjacent territories.

D. Splitting areas in narrow places

Due to the fact that the task of automatic color segmentation, by definition, is the assignment of pixels to groups by the criterion of color similarity, we are faced with the well-known problem of semantic gap [19]. One of its manifestations is the appearance of color areas that capture objects of different semantic categories due to the fact that these objects are close to each other and their pixels have similar color values. However, it can be seen that there are many regions that connect different objects through narrow "isthmuses". As a rule, such "isthmuses" are not inherent in real objects. Thus, another method of refinement of color areas on the basis of geometric splitting of the figure into parts in narrow places suggests itself.

The developed Algorithm 2 cuts the region *Edge* into the subregions *NewEdges* in optimal places.

Algorithm 2 Splitting an area in narrow places

NewEdges ← ∅
Queue ← { *Edge* }
foreach *E* ∈ *Queue* **do**
    m ← +∞
    if *Area*(*E*) is **not** *VerySmall* **then**
        for *i* = 1 .. |*E*| **do**
            p_i ← *StartPoint*(e_i)
            for *j* = 1 .. |*E*|, *j* ≠ *i* **do**
                foreach p_j ∈ { *StartPoint*(e_j), *Project*(p_i, e_j) }
                    v ← *LineSegment*(p_i, p_j)
                    if ∃*k* (*k* ≠ *i* ∧ *k* ≠ *j* ∧ *Intersect*(*v*, e_k)) **then**
                        continue
                    end
                    r_1 ← |*v*| / *Perimeter*(*SubChain*(*E*, p_i, p_j))
                    r_2 ← |*v*| / *Perimeter*(*SubChain*(*E*, p_j, p_i))
                    if (r_1 and r_2 are **not** *Large*) ∧ ($r_1 + r_2 < m$)
                    then
                        $m \leftarrow r_1 + r_2$
                        $p_1 \leftarrow p_i$
                        $p_2 \leftarrow p_j$
                    end
                end
            end
        end
    end
    if *m* = +∞ **then**
        NewEdges ← *NewEdges* ∪ { *E* }
    else
        S_1 ← *SubChain*(*E*, p_1, p_2)
        S_2 ← *SubChain*(*E*, p_2, p_1)
        Queue ← *Queue* ∪ { S_1, S_2 }
    end
end

The function *Project*(*p*, *e*) returns a point that is the result of projecting the point *p* onto the line segment *e*, if one exists.

The function *Intersect*(*v*, *e*) returns the true value if the line segments *v* and *e* intersect.

The function *SubChain*(*E*, p_1, p_2) returns the subchain *S* of the chain *E*, which corresponds to a counterclockwise traversal from p_1 to p_2. *S* is closed by a line segment between p_2 and p_1.

The algorithm searches for the place of the chain cut, in which the ratio of the section length to the perimeter of the resulting subchain is minimal. The area is splitted only if it is not very small and there is a satisfactory (rather narrow in relative terms) cutting place.

It is advisable not to input areas belonging to the category of "Greenery", because they rarely capture other categories of objects and therefore do not require splitting. Such filtering is relatively easy to implement and saves computational resources. Another way to optimize the algorithm is to consider only those cutting places in which there is a negative inflection of the area edge.

Fig. 4 shows an example of the result of splitting a color area into parts.

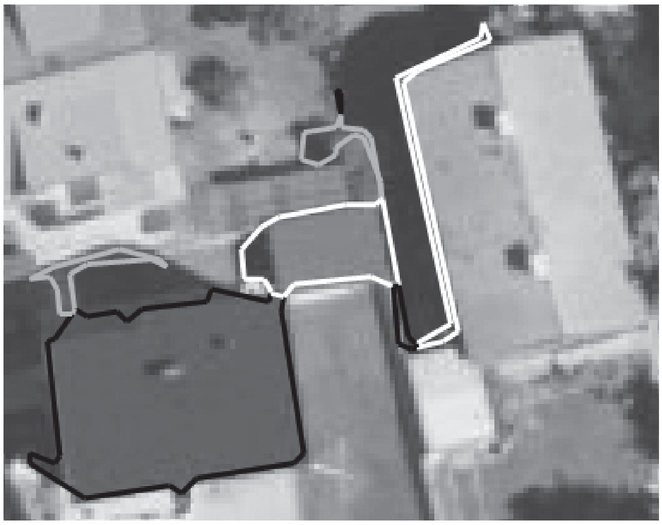


Fig. 4. Superregion was divided into 7 subareas

E. Adaptive approximation of edges

Depending on the hypothesis being tested about the category of the object, different degrees of approximation of its shape features may be required. For example, the characteristic of straightness of the edge is in the first row among the distinctive features of roads. The straightness of the road's edge may deteriorate either due to the influence of other objects (for example, due to the presence of a car parked at the border of the road or a tree overhanging the road), and due to the fact that during automatic color segmentation of the image, the road area can be falsely replenished with fragments of the adjacent territory (sidewalks, curbs) that are similar in color or, on the contrary, contaminated or defective fragments having different colors may be falsely subtracted from it. As a result, winding fragments appear in the shape of the edge, reducing the straightness index and making it difficult to identify the road. The task is even more complicated by the fact that straight sections need to be identified at a long distance, because roads have a significant length in comparison with other categories of objects.

This problem can be solved by making a stronger approximation of the edges of potential roads, as shown in Fig. 5.

The developed approximation method, which is considered in detail in the paper [16], forms straight line segments and circular arcs by recursively augmenting the approximated chains.

The threshold of maximal deviation of the resulting chain from the original one is dependent on the size of the area in order to adapt to objects of different sizes.

(a) (b)

Fig. 5. Example of the result of approximation of the edges: (a) initial boundary obtained in the course of weak approximation at the initial (general) stage of processing; (b) re-approximated boundary

F. *Feature calculation and object classification*

In the process of object detection, various geometric and textural features and relations between color regions are analyzed; hypotheses about categories of objects are put forward and verified by logical inference. In contrast to the existing methods [12], in the approach being developed a special role is given to the shape features and the external environment of the areas as elements on the basis of which it is possible to model reasoning, thereby overcoming the semantic gap and obtaining more reasonable identification results.

At the current stage of the project development, the following features are automatically calculated and used in decision-making: significant elements of the edge; straightness; the presence of three sides of a rectangle; the presence of corners of a rectangle; area; average width; elongation; squareness; circleness; tortuosity; average color; and contour point density.

Below is an example of a classification rule that deals with the situation of partial hiding of a building by trees:

$$\exists a,b \in REGIONS\ (Trees(a) \wedge Adjacent(a, b) \wedge HasRectAngles(b) \wedge Tortuosity(CommonBorder(a, b))\ \text{is}\ High) \rightarrow Building(b).$$

During the logical analysis of the image, the inference tree is built.

A detailed consideration of the stage at which the classification of objects takes place is beyond the scope of this paper. The development of strategies and methods of analysis and decision making applied at the classification stage is the subject of a separate study.

III. Experimental study

An experimental study of the presented techniques was made on images from the publicly available IAIL benchmark dataset [20], [21]. The images have dimensions of 5000×5000 and high spatial resolution.

The aim of the experiments was to assess how strongly the proposed techniques of improving the color areas can affect the quality of the results of object identification. In the experiments, the target objects were buildings.

The indicators *Recall*, *Precision* and *F1*-measure, widely used in the problems of semantic image labeling [22], were applied as evaluation metrics. These indicators are calculated as follows:

$$Recall = \frac{TP}{TP + FN},$$

$$Precision = \frac{TP}{TP + FP},$$

$$F1 = \frac{2TP}{2TP + FN + FP},$$

TP being the number of ground truth objects that the system has successfully identified, *FN* being the number of ground truth objects that the system has missed, and *FP* being the number of false objects present at the system's output.

F1-measure is an integrated representation of the *Recall* and *Precision* metrics. The calculation of the indicators was based on the ground truth data available in the benchmark.

Table I presents the obtained values of the performance indicators for different sizes of cutting an image into frames.

TABLE I. Dependence of identification quality on the image partition size

No	Frame size	Recall	Precision	F1
1	1000×1000	0.611	0.723	0.663
2	750×750	0.620	0.738	0.674
3	500×500	0.651	0.724	0.686
4	400×400	0.666	0.747	0.704
5	300×300	0.700	0.803	**0.748**
6	250×250	0.671	0.808	0.733
7	200×200	0.723	0.774	**0.747**
8	150×150	0.654	0.774	0.709
9	125×125	0.671	0.825	0.740

Thus, the optimal frame size is in the range of 200 to 300 pixels. With an excessive reduction in the frame size, the subsequent merging of color areas becomes difficult and the number of errors increases. In addition, the image processing speed is noticeably reduced.

When the technique of merging color areas from neighboring frames is not used, the value of *F1* decreases by 4% with a frame size of 300×300 and by 6.7% with a frame size of 200×200. The degradation of the quality of the final result also takes place when the operation of splitting areas into parts in narrow places is turned off: overall deterioration by 11.2% and 9.1% respectively due to a noticeable decrease in recall.

Table II shows the effect of changing the degree of approximation of the edges of regions. Best results are obtained with a sufficiently large threshold of relative deviation (2.5% of the region size).

TABLE II. IMPACT OF APPROXIMATION ON THE RESULT

No	Relative deviation threshold	Recall	Precision	F1
1	0 (without approximation)	0.363	0.799	0.499
2	0.005	0.380	0.911	0.536
3	0.010	0.426	0.887	0.575
4	0.015	0.509	0.864	0.640
5	0.020	0.617	0.818	0.704
6	0.025	**0.700**	**0.803**	**0.748**
7	0.030	0.734	0.732	0.733
8	0.035	0.777	0.690	0.731
9	0.040	0.800	0.648	0.716
10	0.045	0.823	0.630	0.713
11	0.050	0.834	0.606	0.702

Thus, the approximation has a significant impact on the final result of the classification of image objects.

Since the comparison of the quality of color segmentation per se with other methods presents a certain difficulty, the author provides comparative indicators of the overall performance of building identification. Detection was carried out on the basis of shape features and logical rules, as mentioned in Section II.F.

Table III contains numerical estimates of the quality of the work of alternative methods and the proposed method, obtained on images of West Tyrol from the IAIL benchmark dataset. In our case, image processing was performed with optimal parameter values determined in the above experiments. The alternative methods are based on neural networks; their performance evaluations are taken from [20].

TABLE III. COMPARATIVE NUMERICAL EVALUATIONS OF THE OVERALL PERFORMANCE OF BUILDING IDENTIFICATION

Method	IoU	Accuracy
FCN [20]	46.86	95.83
Skip [20]	54.91	96.47
MLP [20]	57.95	96.66
Proposed OBIA-method	53.13	85.10

The pixel-oriented metrics *IoU* and *Accuracy* were used here, which are also popular in the field of semantic classification. The *IoU* index considers the sets of target object pixels in the prediction and in the ground truth and is calculated as the ratio of the cardinality of Intersection of these sets to the cardinality of their Union. The *Accuracy* indicator shows the proportion of true pixels in the prediction [22]. These indicators are characterized by the fact that they register the slightest deviations of the results produced by the program relative to the ground truth, which, of course, is of no small importance for use in cartography tasks. This feature explains the fact that the *IoU* and *Accuracy* indicators obtained by the author are smaller in comparison with the above estimates of *Recall* and *Precision*, which allow a certain degree of deviation of the detected object boundaries from the ground truth.

The comparative assessments show that the quality of identification is at the same level as shown by machine learning methods based on neural networks. It should be noted that our results were obtained without using any training data. As is known, the preparation of training datasets requires significant costs. The proposed approach allows automating this process and simplifying it by at least half.

Fig. 6 shows an example of the result of the detection of buildings in the image.

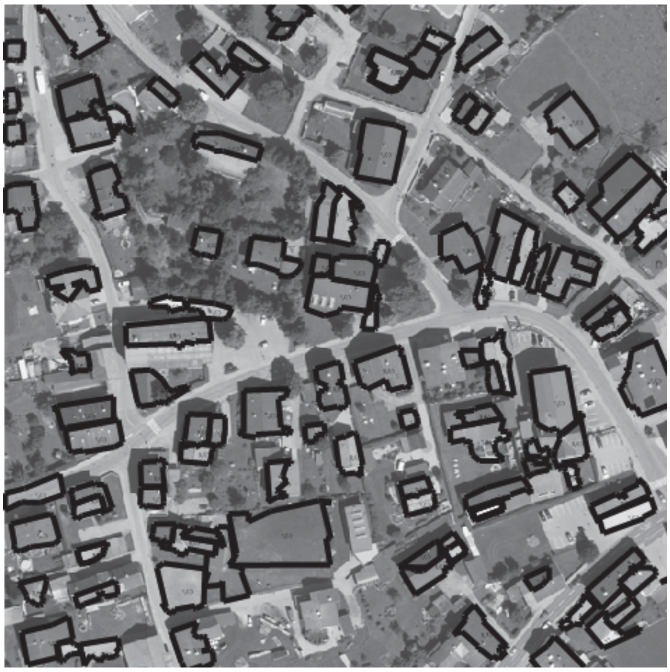


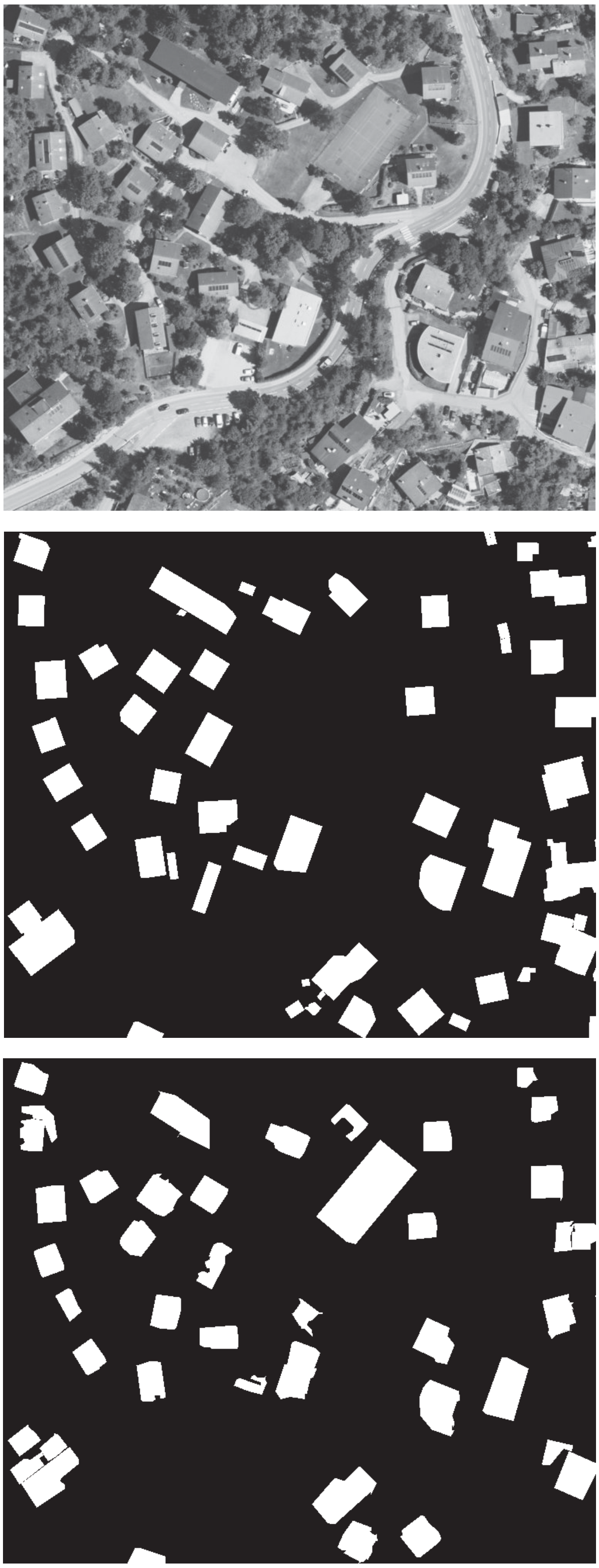
Fig. 6. Automatically identified buildings

Fig. 7 provides an example for visual comparison between detection results and reference data.

As noted, our approach focuses on the features of the shape of objects. The experiments confirmed the viability of this strategy, but also revealed some difficulties that need to be overcome to achieve higher performance:

- A rectilinear rectangular shape does not always correspond to buildings only. Some other man-made objects (sports grounds, yard areas, parking lots, lawn areas, etc.) may also have such a shape. To eliminate misclassifications of this kind, it is necessary to introduce additional spectral and texture features, as well as to strengthen the analysis of the scene using case-based reasoning.
- Unfortunately, the proposed techniques for improving color segmentation do not completely eliminate the problem of the redundancy of color regions. There are still cases when one real object is represented by two or more color segments that have not very pronounced and distinctive shape. This leads to missing a part of targets. To solve this problem, an additional technique

can be used, which involves the construction of a non-convex hull around a group of color regions in order to restore information about the shape of a fragmented object.



(a) (b) (c)

Fig. 7. Results in comparison with the ground truth labeling: (a) input image; (b) reference data; (c) results of the program

- The shape of artificial objects may be distorted if they are partially overlapped by other objects (for example, trees or the shadow of large buildings). In this case, it makes sense to identify common parts of the edges of adjacent objects and exclude them from the consideration during the calculation of the shape features so that the winding sections do not mix with the straight fragments of the object's edge. This will eliminate the averaging effect.

It should be noted that the achieved values of the quality indicators are preliminary. In particular, no attempts have been made to refine the detected objects, for which special techniques and mechanisms of changing the degree of abstraction are needed. It is expected that after creating adaptive context-sensitive strategies of modeling of reasoning on spatially distributed image data, it will be possible to significantly improve the overall effectiveness of the system.

In general, the experiments show that the proposed techniques can improve the results of color segmentation. The correctness of the latter has a direct impact on the final quality of the identification of image objects.

IV. Conclusion

The paper proposes a number of techniques to improve the degree of correspondence between the color areas obtained as a result of color segmentation and the real objects present in the image.

Partitioning the image into frames increases the locality of color segmentation. As the size of the segmentation window decreases, the accuracy of the area edges increases, the number of areas covering different objects decreases, and the level of detail of the extracted information increases.

Using the created algorithm for merging areas lying on the borders of adjacent frames, it is possible to eliminate information gaps between individual image frames. The degree of fragmentation of the frame-by-frame segmentation results is reduced.

The proposed algorithm for splitting areas in narrow places serves the purpose of separating objects of different categories from each other. A color area that captures two or more objects has false shape characteristics that can randomly meet the requirements of matching the wrong object category. Accordingly, eliminating superregions at an early stage reduces the risk of misclassification at the decision-making stage.

Adaptive approximation of the edges of color areas allows to successfully identify the key parameters of the shape even in conditions of strong distortion of the areas. The creation of complex shape features is simplified. Calculated features become more stable and can be used intensively in further high-level image analysis.

In general, the proposed techniques can improve the accuracy and informativeness of the extracted color areas, reduce the load on the decision-making module and simplify it in some aspects. The work contributes to the improvement of the direction of object-oriented image analysis.

At this point, the proposed object-oriented approach shows approximately the same performance as methods based on artificial neural networks: the *IoU* index is at the level of 53%. This assessment proves that the approach is workable, while showing the need for further strengthening of the high-level analysis stage. Unlike analogues, the results were obtained without training on labeled datasets, which indicates a higher degree of universality.

Naturally, the spatial resolution of images affects the accuracy of shape analysis. If, at lower resolutions, the clear corners of artificial objects are preserved, then on such images the proposed approach remains quite workable. When the corners are rounded, the object of interest becomes difficult to distinguish from other objects.

Noises are usually removed at the first stages of image processing, which is a certain difficulty. In the proposed approach, the noise is eliminated at the upper levels, with the help of logical analysis.

One of the main advantages of the proposed approach is the expansion of shape features. As an additional effect, the possibility of automating the preparation of training datasets can potentially be considered.

The main direction of further work is the development of new features of objects with greater descriptive ability and strengthening the classification stage by creating flexible strategies of context-sensitive analysis and modeling of reasoning, as well as active application of case-based decision-making.

ACKNOWLEDGMENT

This work is supported by the Russian Science Foundation under grant No. 18-71-00109.

REFERENCES

[1] X. Huang, H. Chen, and J. Gong, "Angular difference feature extraction for urban scene classification using ZY-3 multi-angle high-resolution satellite imagery", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 135, 2018, pp. 127-141.

[2] A. Maarir, B. Bouikhalene, and Y. Chajri, "Building Detection from Satellite Images based on Curvature Scale Space Method", *Walailak Journal of Science and Technology*, vol.14, no.6, 2017, pp. 517-525.

[3] L.I. Jimenez, J. Plaza, and A. Plaza, "Efficient Implementation of Morphological Index for Building/Shadow Extraction from Remotely Sensed Images", *The Journal of Supercomputing*, vol.73, is.1, 2017, pp. 482-494.

[4] S. Golovanov, R. Kurbanov, A. Artamonov, A. Davydow, and S. Nikolenko, "Building Detection from Satellite Imagery Using a Composite Loss Function", *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2018, pp. 229-232.

[5] A.K. Goswami, S. Gakhar, and H. Kaur, "Automatic Object Recognition from Satellite Images using Artificial Neural Network", *International Journal of Computer Applications*, vol.95, no.10, 2014, pp. 33-39.

[6] Z. Guo, Q. Chen, G. Wu, Y. Xu, R. Shibasaki, and X. Shao, "Village Building Identification Based on Ensemble Convolutional Neural Networks", *Sensors*, vol.17, is.11, 2017, 2487.

[7] T. Blaschke, G.J. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addink, R.Q. Feitosa, F. Meer, H. Werff, F. Coillie, and D. Tiede, "Geographic Object-Based Image Analysis – Towards a new paradigm", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.87, 2014, pp. 180-191.

[8] G. Chen, Q. Weng, G.J. Hay, and Y. He, "Geographic Object-based Image Analysis (GEOBIA): Emerging trends and future opportunities", *GIScience & Remote Sensing*, 2018.

[9] S. Khatriker and M. Kumar, "Building Footprint Extraction from High Resolution Satellite Imagery Using Segmentation", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol.XLII-5, 2018, pp. 123-128.

[10] Y. He, X. Zhang, and L. Hua, "Object-Based Distinction between Building Shadow and Water in High-Resolution Imagery Using Fuzzy-Rule Classification and Artificial Bee Colony Optimization", *Journal of Sensors*, vol.2016, 2016, 2385039, 10 p.

[11] M. Janalipour and A. Mohammadzadeh, "Building Damage Detection Using Object-Based Image Analysis and ANFIS From High-Resolution Image (Case Study: BAM Earthquake, Iran)", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.9, is.5, 2016, pp. 1937-1945.

[12] J. Jia, N. Yang, C. Zhang, A. Yue, J. Yang, and D. Zhu, "Object-oriented feature selection of high spatial resolution images using an improved Relief algorithm", *Mathematical and Computer Modelling*, vol.58, 2013, pp. 619-626.

[13] S. Lhomme, D.C. He, C. Weber, and D. Morin, "A new approach to building identification from very-high-spatial-resolution images", *International Journal of Remote Sensing,* vol. 30, is.5, 2009, pp. 1341-1354.

[14] R.R. Antunes, E.S. Bias, G.A.O.P. Costa, and R.S. Brites, "Object-Based Analysis For Urban Land Cover Mapping Using The InterIMAGE And The SIPINA Free Software Packages", *Bulletin of Geodetic Sciences*, vol. 24, is. 1, 2018, pp. 1-17.

[15] M. Baatz and A. Schäpe, "Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation", *In: XII Angewandte Geographische Informationsverarbeitung, AGIT Symposium. Proceedings*, 2000, pp. 12-23.

[16] D.R. Kasimov, A.V. Kuchuganov, V.N. Kuchuganov, and P.P. Oskolkov, "Approximation of Color Images Based on the Clusterization of the Color Palette and Smoothing Boundaries by Splines and Arcs", *Programming and Computer Software*, vol.44, no. 5, 2018, pp. 295-302.

[17] G. Sharma, W. Wu, and E.N. Dalal, "The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations", *Color Research & Application*, vol.30, no.1, 2005, pp. 21-30.

[18] J. Lv and J. Fang, "A Color Distance Model Based on Visual Recognition", *Mathematical Problems in Engineering*, vol.2018, 2018, 4652526, 7 p.

[19] J. Eakins and M. Graham, *Content-based image retrieval : Technical report, JISC Technology Applications Programme*. Newcastle: University of Northumbria at Newcastle, 1999, 59 p.

[20] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark", *IEEE International Geo-science and Remote Sensing Symposium (IGARSS)*, 2017.

[21] B. Huang, K. Lu, N. Audebert, A. Khalel, Y. Tarabalka, J. Malof, A. Boulch, B.L. Saux, L. Collins, K. Bradbury, S. Lefever, and M. El-Saban, "Large-scale semantic classification: outcome of the first year of Inria aerial image labeling benchmark", *GARSS2018 - IEEE International Geoscience and Remote Sensing Symposium*, Jul 2018, pp. 1-4.

[22] E. Fernandez-Moral, R. Martins, D. Wolf, and P. Rives, "A new metric for evaluating semantic segmentation: leveraging global and contour accuracy", *Workshop on Planning, Perception and Navigation for Intelligent Vehicles, PPNIV17*, 2017.