

# Visible Structures Highlighting Model Analysis Aimed at Object Image Detection Problem

I R Saifudinov<sup>1</sup>, V V Mokshin<sup>1</sup>, P I Tutubalin<sup>1</sup>, L M Sharnin<sup>1</sup> and D G Hohlov<sup>1</sup>

<sup>1</sup>Kazan National Research Technical University named after A. N. Tupolev - KAI, Karl Marx str. 10, Kazan, Russia, 420111

**Abstract.** The research considers an approach to solving the problem of reducing the data processed in mobile platforms oriented video-analytic systems. Different models of human visual attention has been analyzed and also classified according to the image segmentation. The results obtained are presented in the form of the method used for isolating the borders of the most significant object in the image. They are based on the optimization the length and curvature values of the object borders. This approach allows to filter significant structures in the image and process them with the help of image segmentation techniques. The method was evaluated according to usage the accuracy criterion along with other methods in delineation of boundaries: threshold value, morphological processing and watershed. Software and hardware system for registering dump trucks has been improved. It gives the possibility to automatize the process of registration dump trucks, training them during the construction of roads.

## 1. Introduction

Automated video surveillance systems (video analytics) used in many spheres of human activity, likewise state institutions and manufacturing enterprises are extremely popular nowadays. The systems usage allows to create options for events automatic alarm, efficiency in employees labour improvement by means of direct control of the performing activity. Herewith, because of the large number of tasks solved in video analytics, as well as their multi-criteria, autonomy tendency, and indistinct nature of tasks, there is a certain necessity to find the effective approach for highlighting image visible structures that allows to work on the basis of mobile platforms. The attempt for analysis of various tasks has been made in the research. [1], [3], [6-8].

The method of highlighting of visible structures in an image based on visibility measuring of length and curvature of the curve that is similar to the concept suggested by Lowe [4] has been taken into consideration. The research deals with the measure of visibility according to the criterion of false positive detection in comparison to the method of delimiting contour of object [2]. In addition to the evaluation of the method, it is compared with the segmentation and accuracy of other methods of optimization of significant information.

## 2. Visibility network construction

Orientation elements are the basic computational elements of the network [9]. Each element  $p_i$  is connected with a processor that normally can perform certain calculations based on the conditions and those with  $k$  denotation performed by the neighbour processor. This defines a single network containing  $kn^2$  processing blocks with the local communication. In the current implementation,  $k$  is equal to 48, which provides a reasonable angular resolution. Let's appeal to the associated orientation

sequence of the elements  $p_i, \dots, p_{i+N}$ , where each element is a linear segment or interval, like length of a curve  $N$ . (curves can be continuous or with any number of intervals). The optimization task is formulated to maximize the value of  $\Phi_N$  above overall length of a curve  $N$ , starting with  $p_i$ .

$$\max_{(p_{i+1}, \dots, p_{i+N}) \in \delta^N(p_i)} \Phi_N(p_i, \dots, p_{i+N})$$

where  $\delta^N(p_i)$  is a number of all possible length of a curve  $N$ , starting with  $p_i$ .

For a certain of measures class  $\Phi(\cdot)$ , the calculation of  $\Phi_N$  can be obtained by simple local computations made repeatedly. To illustrate, let's take the first curves three elements long only. In this case:

$$\max_{(p_{i+1}, p_{i+2}) \in \delta^2(p_i)} \Phi_2(p_i, p_{i+1}, p_{i+2})$$

where  $p_i$  is defined by  $p_{i+1}$  (one of  $p_i$ 's  $k$  of neighbours) and  $p_{i+2}$  ( $p_{i+1}$  neighbour) for a given element so that the rate  $\Phi_2(p_i, p_{i+1}, p_{i+2})$  will be in its maximum. Simple approach (brute-force method) in the analysis of  $k^2$  value in different curves will be required anew. Suppose, however, that  $\Phi_2$  corresponds to the condition of:

$$\max_{\delta^2(p_i)} \Phi_2(p_i, p_{i+1}, p_{i+2}) = \max_{p_{i+1}} \Phi_1(p_i, \max_{p_{i+2}} \Phi_1(p_{i+1}, p_{i+2}))$$

In this case, the maximization of the rate  $\Phi_2$  can be achieved by application of  $\Phi_1$  used repeatedly over shorter curves. The general approach can be formulated the same way, e.i:

$$\max_{\delta^N(p_i)} \Phi_N(p_i, \dots, p_{i+N}) = \max_{p_{i+1} \in \delta(p_i)} \Phi_1(p_i, \max_{\delta^{N-1}(p_{i+1})} \Phi_{N-1}(p_{i+1}, \dots, p_{i+N})) \quad (1)$$

where  $\delta(p_i)$  is equal to  $\delta^1(p_i)$ . Thus, the searching area required for each length of a curve  $N$  is being reduced, from  $p_i$  to  $kN$ , instead of  $k^N$  which is essential for a brute-force method approach. The concept (1) is related to the optimality principle, underlining for all multistage decision-making processes. This is a special case in dynamic programming in particular. It refers to the family of functions that follows concept (1) of extensible functions.

There are two factors that are essential for visibility measure. The first one is related to the length of a curve, and the second factor is related to its shape. The length of a curve is determined by the number of its elements that have a factual curve (rather than an interval) passing through these elements. They are called active elements. Whereas elements that are associated with intervals are referred to as virtual elements where local visibility  $\sigma_i$  corresponds to  $p_i$ . If  $p_i$  is the active element, then  $\sigma_i$  has positive value, which is equal to 1 and 0 for the virtual element  $\sigma_i$ . A measure associated with the length of a curve  $p_i, \dots, p_{i+N}$  is determined by the equation:

$$\sum_{j=i}^{i+N} \sigma_j \quad (2)$$

The measure rate above(2) presents the sum of the local values of the visibility of active elements along the curve.

Then, the attenuation function associated with the curve  $p_i, \dots, p_j$  is defined as follows:

$$\rho_{i,j} = \prod_{k=i+1}^j \rho_k$$

where  $\rho_{i,i} = 1$ . The measure in (2) is modified by the attenuation coefficients is:

$$\sum_{j=i}^{i+N} \rho_{i,j} \sigma_j \quad (3)$$

The measure rate in equation (3) is weighted contribution of the local visibility values  $\sigma_j$  along a curve, that are in reverse dependency to a number of virtual elements along  $p_i, \dots, p_j$ . To measure the shape of a curve, measure that in reverse dependency to the total curvature of a curve is used. The total curvature of  $\gamma$  is defined as  $\int_{\gamma} \left(\frac{d\theta}{ds}\right)^2 ds$ , where  $\theta(s)$  is a slope along a curve and  $\frac{d\theta}{ds}$  at the point  $P$  is known as the local curvature at this point (the reciprocal value of the radius of curvature  $R$ ). It is necessary to use the total curvature to obtain a measure that is limited and in the position of reverse dependency to the total curvature. The next measure is relevant to the following:

$$\exp^{-\int_{\gamma} \left(\frac{d\theta}{ds}\right)^2 ds} \quad (4)$$

In order to obtain a discrete approximation to the measure in (4), we denote  $a_k$  to indicate orientation difference between the  $k$ -th element and its successor and  $\Delta s$  as length of the orientation element. A discrete approximation to full curvature of a measure along  $p_i, \dots, p_j$ , will be:

$$C_{i,j} = \prod_{k=i}^{j-1} f_{k,k+1}$$

where

$$f_{k,k+1} = \exp^{-\frac{2a_k \operatorname{tg} \frac{a_k}{2}}{\Delta s}} \quad (5)$$

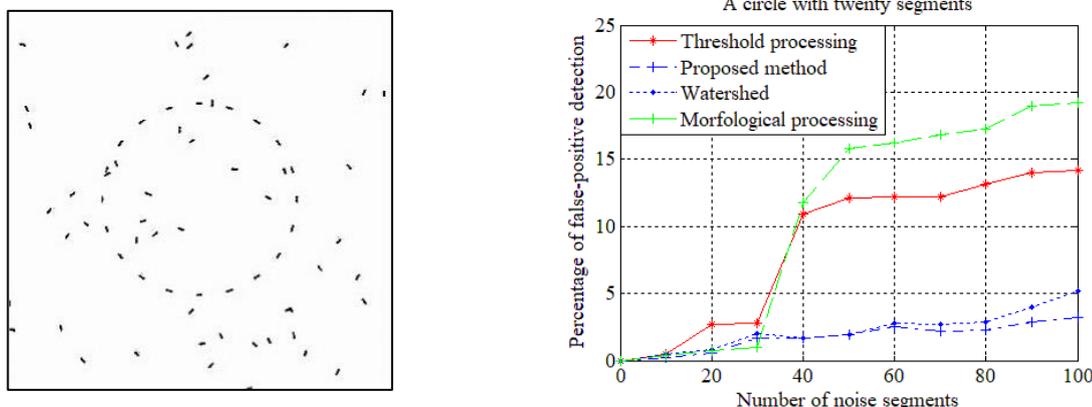
$C_{i,j}$  is the weight of each value of local visibility  $\sigma_j$  along a curve. A measure that shows a high rate for long curves with low overall curvature is now defined as:

$$\sum_{j=1}^{i+N} C_{i,j} \rho_{i,j} \sigma_j \quad (6)$$

The measure in equation (6) is weighted contribution of the local visibility values  $\sigma_j$  along a curve. The curves that will receive a high measure on (6) are long curves, more straight with the least number of intervals.

### 3. Analysis of the visibility network: discussion and results

To analyse visibility network we compared the percentage of false positive detections. Test samples located around the perimeter of circle at equal intervals consist of short, oriented segments in the field of segments with random position and orientation as it shown in figure 1.

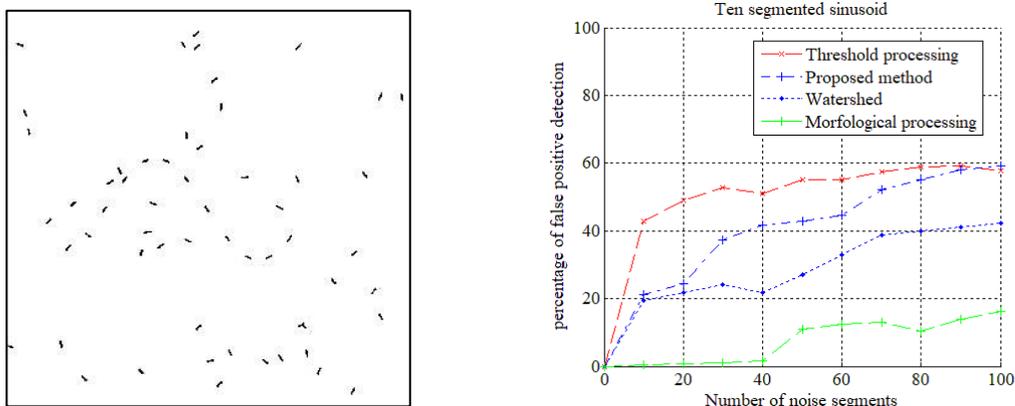


**Figure 1.** A circle with twenty segments.

The methods that are used for the calculation visibility forms and noise segments are: the threshold value [10], the watershed [11,13], the morphology [12], and the proposed approach. Segments were sorted in ascending order according to their visibility of the most ( $\phi_1$ ) and less ( $\phi_n$ ) noticeable segments. For given  $m$  form segments, false-positive are defined as noise segments, which are assigned a visibility greater than  $\phi_{m+1}$ . A false positive estimate for each method was calculated for samples consisting of different numbers of shape and noise segments rate. A false positive rate for each combination (for example, 20 shape segments and 70 noise segments) was estimated by averaging false positive value by means of more than ten attempts with different noise samples. The picture in the right side of Figure 2 is the graph of false positive rate percentage for a circle with twenty segments.

Each method can be accomplished well enough (less than 10% of false positives) at a low noise level (40 noise segments or less). The results of methods applied begin to diverge at higher noise levels. It should be noted that threshold processing is superior to morphological processing while watershed is relative to the approach suggested, although the latter is more expensive to calculate.

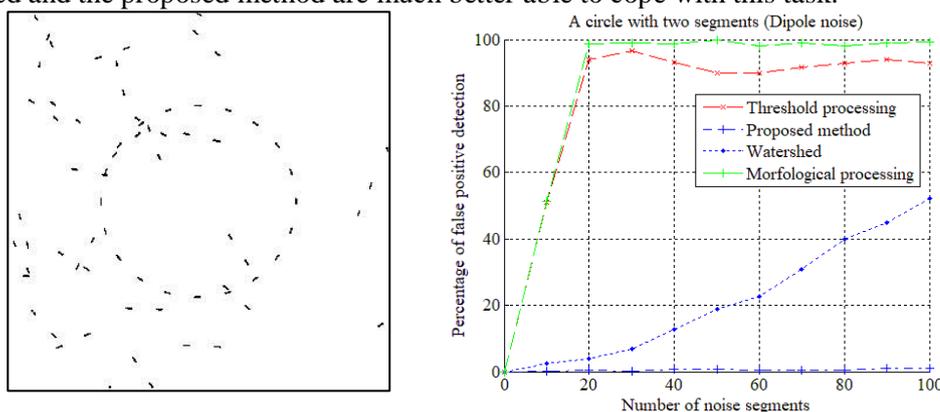
And, finally, at lower signal-to-noise ratios watershed and the proposed approach have significantly lower false-positive estimates. The next comparison was identical to the first one, but shape segments are formed by an unlimited sinusoid (Figure 2).



**Figure 2.** A circle with a ten segment sinusoid.

In the right half of Figure 2 is a graph of the percentage of false positives relative to the number of noise segments for a ten segment sinusoidal curve. Comparatively low indicators of the proposed method compared with other methods can be attributed to its apparent dependence on closure. However, it still outperforms the threshold processing for higher signal-to-noise ratios and has an error rate comparable to threshold processing (i.e., within 5%) at lower signal-to-noise ratios.

In the third comparison a field consisting of correlated noises (i.e., a dipole) was used (Figure 3). The dipole consists of two collinear segments separated by an interval equal to the distance between neighboring segments of the circle. Since the two segments forming the dipole are collinear, the degree of closeness between the segments forming the dipole is greater than between the adjacent segments of the circle. Therefore, it is impossible to distinguish noise segments from the shape segments using only a local measurement. From the graph it can be seen that the methods of threshold processing and morphological processing have almost a 100% false positive level, whereas the watershed and the proposed method are much better able to cope with this task.

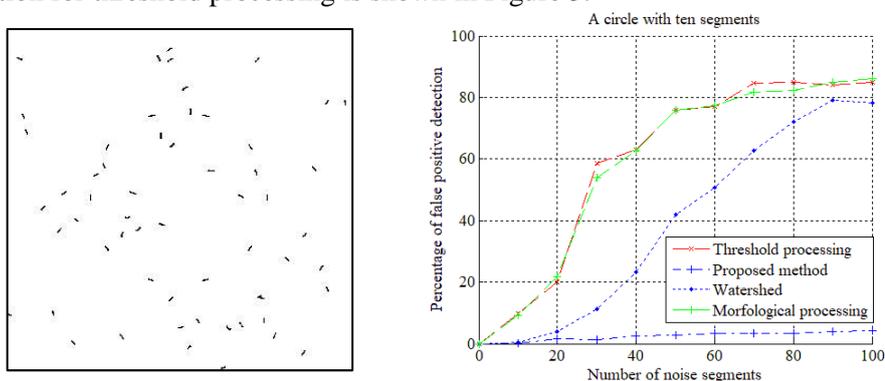


**Figure 3.** A circle with twenty segments. Dipole noise.

In the fourth comparison (Figure 4), ten segments are used. This is a complex picture, because the sampling frequency is so small, so that there is only one segment at 36 degrees of the circle. Most of methods do not work well even at relatively high signal-to-noise ratios. For noise level 80, threshold processing and the morphological processing method are performed at 90% false positive level. The watershed method is performed a bit better, with a false-positive level of 70%. In contrast, the false-positive level for the proposed method is less than 5%.

Let's consider practical use examples of the specified methods in the decision of the object segmentation problem on the image in the video-analytical system focused on mobile platforms for the

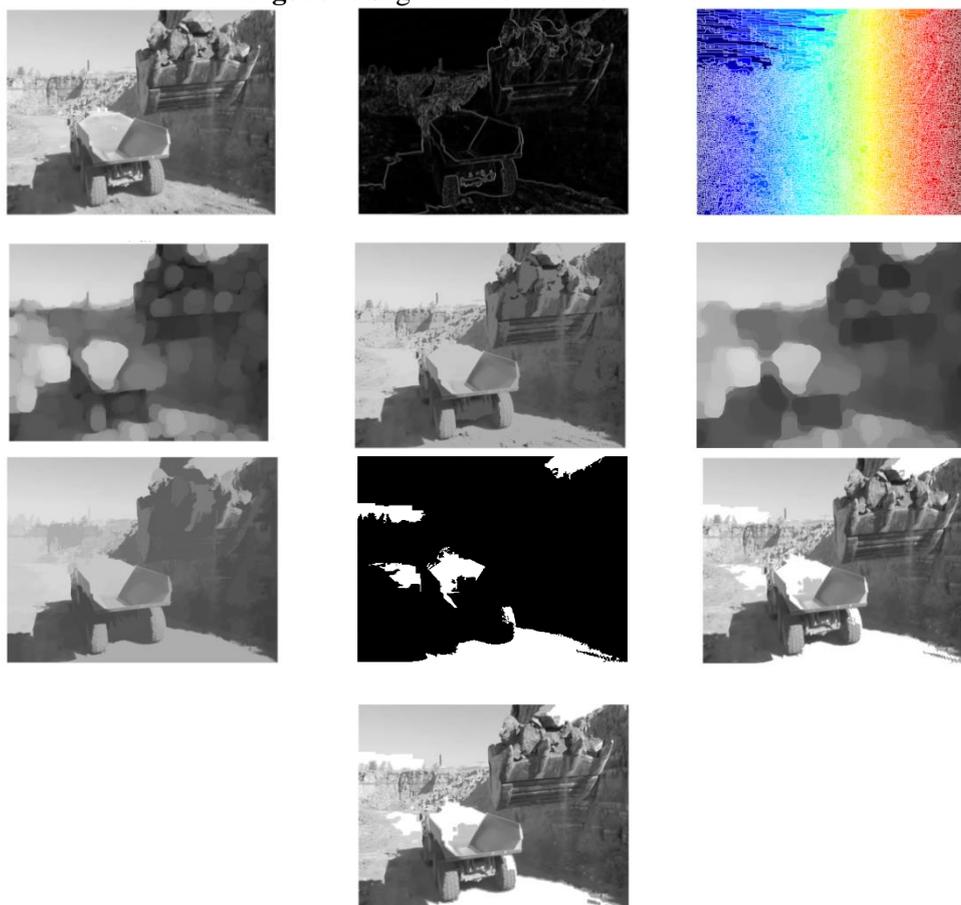
account number of transportations dump trucks. The image segmentation results are shown below. Segmentation for threshold processing is shown in Figure 5.



**Figure 4.** A circle with ten segments.

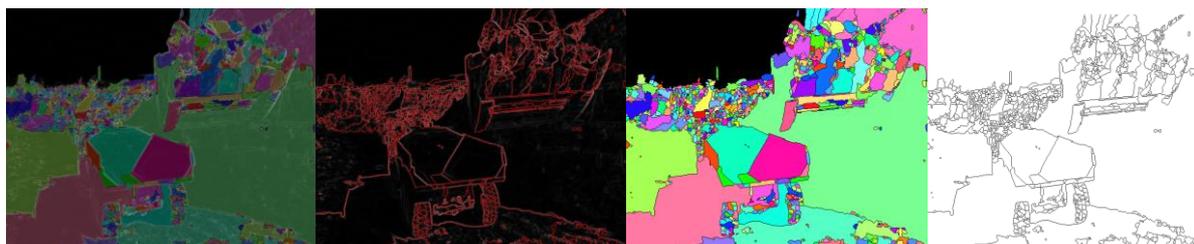


**Figure 5.** Segmentation. Threshold value.



**Figure 6.** Stages of watershed segmentation. From left to right, from top to bottom.

Segmentation for the watershed is shown in Figure 6. The results of morphological processing segmentation are shown in Figure 7.



**Figure 7.** Stages of morphological processing segmentation. From left to right.

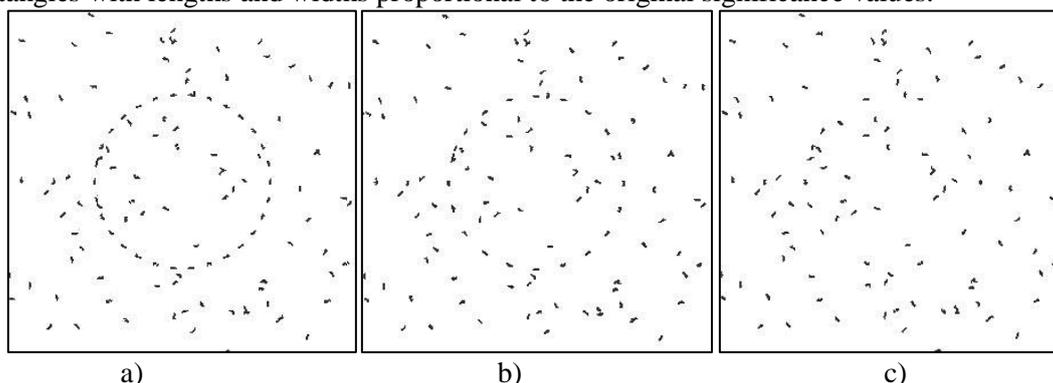
The results of the proposed method segmentation for isolating significant structures are shown in Figure 8.



**Figure 8.** Left. Preprocessing an image using the Sobel operator. The middle. Map of the importance of the image. On right. The selected area of the image.

The proposed method of identifying significant structures was compared with other measures that calculate the optimal structure: contrast discrimination, a combined approach to optimization [18]; quantitative indicators of changes based on the organization of functions: eigenvalues and eigenvectors [19] and stochastic areas of completion: a neural model of illusory shape of the contour and significance [20-23]. Let us consider examples of practical use of these methods in solving the problem of segmentation of important structures in the image.

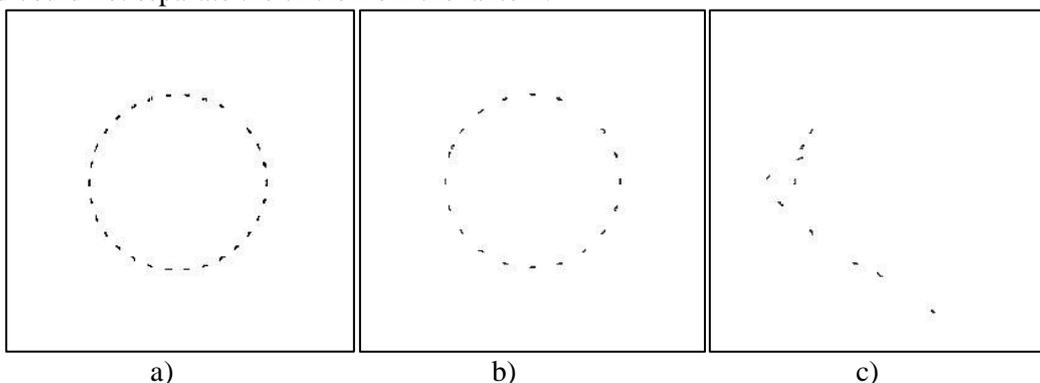
To get an idea of the strengths and weaknesses of boundary selection measures, it is useful to apply them to a simple test scheme consisting of edges from a circle (thirty, twenty, or ten evenly spaced samples) against a background of one hundred edges of random position and orientation. Three test patterns are shown in figure 9. To visualize the values assigned to each edge, the edges are displayed as rectangles with lengths and widths proportional to the original significance values.



**Figure 9.** (a) thirty edge circle against the background of a hundred edges of noise. (b) twenty edge circle. (C) ten edge circle.

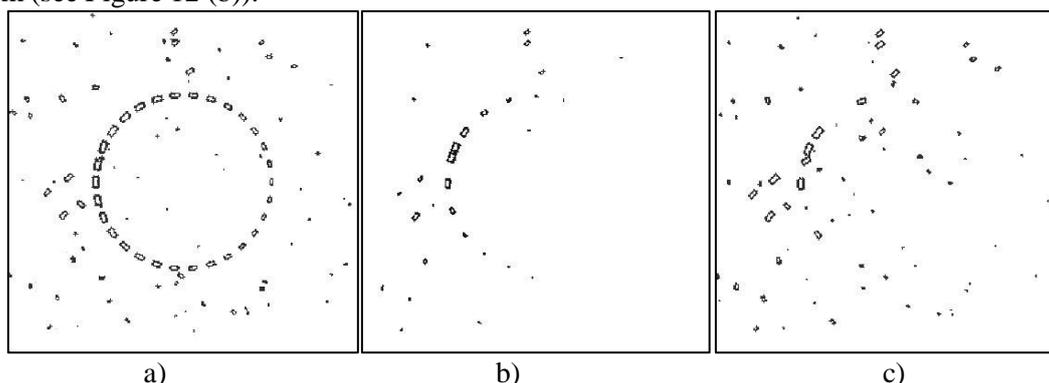
First, let us consider the results of measuring the method of contrast discrimination. It is important to note that the previously described modified optimization problem is solved. That is,  $y^T A y$  is optimized for all vectors,  $y \in \{0,1\}^n$  and  $|y| = m$  where  $n$  - is the total number of edges and  $m$  is the number of edges. Since the method is given the number of edges, it has a significant advantage over

other measures, and these results should be interpreted accordingly. Segmentation for thirty and twenty edge circles is shown in Fig.10 (a) and (b). With the exception of excluding one edge of the circle in the 1 hour orientation, and including a false 10-hour edge, the method calculates perfect segmentation. However, the results are in ten regional districts (see figure 10 (c)) show that the method could not separate the circle from the lantern.



**Figure 10.** Edges of edges, (a) computed by contrast discrimination for thirty edge circle, (b) for twenty edge circle, and (C) for ten edge circle.

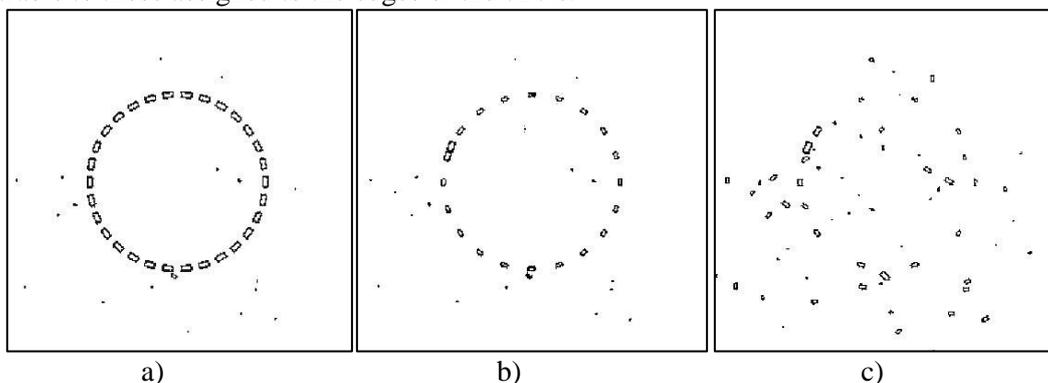
To better visualize the large range of significance values calculated by the quantitative change measure method, the length and width of the rectangles are drawn proportionally to the  $\log(1.0 + 10^6 \cdot x_i)$ , where  $x_i$  - is the significance of the edge  $i$ . In Fig. 11 (a) results for the thirty edge circle calculated using the quantitative change method are shown. In General, the edges of the circle are set to a larger value than the edges of the background. However, it is observed that the significance values in the upper left part of the circle are much larger than the values in the lower right corner. If the eigenvector with the highest positive real eigenvalue is interpreted as a limited distribution of random walks between edges, it is seen that this distribution is dominated by random walks (with reversals in the direction) through the parasitic edge at the 10 o'clock position. Because the measure does not provide tangential continuity or closure, the effect of one inconveniently positioned edge can be profound. For Fig.11 (b) and (c) the asymmetry becomes very pronounced as the circle sample becomes less frequent. The consequence of this is that the measure has failed to isolate the circle from its base, even in the case where a very simple method such as stochastic completion areas has a small problem (see Figure 12 (b)).



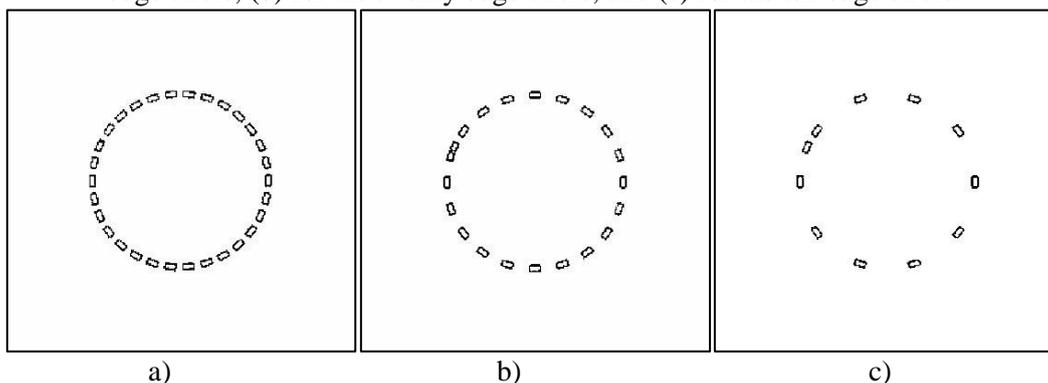
**Figure 11.** The values of significance, (a) calculated by the quantitative change in the indicators for the thirty edge circle, (b) for the twenty edge circle, and (C) for the ten edge circle.

For figure 12 (a) and (b) significance maps for thirty and twenty edge circles calculated by the stochastic completion regions are shown. With the exception of the false 10-hour edge, all edges of the background are assigned gaps with a small value. The contrast between the values calculated by this method and the estimated quantitative indicators of changes is quite impressive. The difference in the power of discrimination is due solely to the use of directivity and the multiplication of vector  $x$  by

vector  $x$  in this method. This multiplication imposes a restriction that the edge must form a bridge between the other two. The result is intervals with a larger range of values than those with only  $x$  or  $x$  components. However, this limitation is not enough to distinguish the edges of the ten edge circles from the background borders. A significance map calculated as the stochastic completion regions for the ten edge circle is shown in figure 12 (c). A significant number of background edges have gaps comparable to those assigned to the edges of the circle.



**Figure 12.** The significance values calculated by the stochastic completion regions, (a) for the thirty edge circle, (b) for the twenty edge circle, and (c) for the ten edge circle.



**Figure 13.** The values of significance calculated by the proposed approach, (a) for the thirty edge circle, (b) for the twenty edge circle, and (c) for the ten edge circle.

Finally, figure 13 (a-c) shows significance maps calculated using the proposed approach. With the exception of the false 10-hour edge, all edges of the background are assigned intervals with a small value. This is true even for ten edge circles, which no other measure could segment from the background.

#### 4. Conclusions

In addition to comparing false positive detection, Table 1 shows the time required to run the algorithms on the Android 5.0 operating system. With a resolution of 640x480 pixels. It can be seen from the table that the proposed method for identifying significant structures shows the fastest and least resource-consuming results. Thus, the proposed model for significant structures identifying is an effective tool for primary segmentation of images in video-analytical systems oriented at mobile platforms. As results have shown, the method of isolating notable structures showed the most accurate results. The method of selecting notable structures is based on the measure of visibility calculation that includes applies the principle of dynamic programming [5, 15]. Also it would be used queueing system to analyse such kind models [16-17]. Two internal properties of the network of significance, such as extensibility and geometric convergence, allow us to optimize the measure of significance and effectively restore optimal curves (in polynomial time).

**Table 1.** Comparison of the time spent on segmentation and resource intensity.

Method	Calculation time (seconds)	RAM Consumption (MB)
Threshold processing	1.36	70
Watershed	2.03	105
Morphological treatment	1.78	85
Allocation of significant structures	1.09	45
Contrast discrimination	1.66	77
Stochastic completion areas	1.83	98
Quantitative indicators of changes	1.88	95

At the same time, they restrict the range of possible functions that can be used as a measure of significance, so the method has some limitations with scale invariance, fusion of curves, and grouping in the presence of compounds. In addition, overcoming overlapping problems will require asymptotically increasing the complexity of the method, since the discretization is closely related with using dynamic programming to effectively optimize the chosen measure.

## 5. References

- [1] Mokshin V V, Saifudinov I R, Kirpichnikov A P and Sharnin L M 2016 Vehicle recognition based on heuristic data and machine learning *Bull. Kazan Techn. Univ.* **19(5)** 130-137 (in Russian)
- [2] Saifudinov I R, Mokshin V V and Kirpichnikov A P 2017 Grouping contours of objects of structural images based on the network of element visibility *Bull. Kazan Techn. Univ.* **20(9)** 120-123 (in Russian)
- [3] Treisman A 1982 Perceptual Grouping and Attention in Visual Search for Features and for Objects *Journal of Experimental Psychology: Human Perception and Performance* **8(2)** 194-214
- [4] Lowe D G 1985 *Perceptual Organization and Visual Recognition* (Kluwer Academic Publishers, Boston)
- [5] Kormen T, Leiserson Ch, Rivest R and Shtaine K 2005 *Dynamic Programming Algorithms: construction and analysis (Introduction to Algorithms)* (Moscow: Williams) p 1296 (in Russia)
- [6] Gorilik A L and Skripkin V A 1989 *Recognition methods* (Higher School Publishing) p 232 (in Russia)
- [7] Vapnik V N and Chervonenkis A I 1974 *Theory of pattern recognition* (Moscow: Science) p 416 (in Russia)
- [8] Nesteruk V F and Sokolova V A 1980 Questions of the theory of perception of subject images and a quantitative assessment of their contrast *Optoelectronic Industry* **5** 11-13 (in Russia)
- [9] Shashua A and Ullman S 1988 Structural saliency: The detection of globally salient structures using a locally connected network *2nd Intl. Conf. on Computer Vision (ICCV '88), Clearwater, FL*
- [10] Laurent H and Radu H 1993 Figure-Ground Discrimination: a Combinatorial Optimization Approach *IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers* **15(9)** 899-914

- [11] Sarkar S and Boyer K L 1996 Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors *Computer Vision and Pattern Recognition* **40** DOI: 10.1109/CVPR.1996.517115
- [12] Williams L R and Jacobs D W 1997 Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Saliency *Neural Computation* **9(4)** 837-858
- [13] Shakawat M, Sarker Z, Tan W H and Logeswaran R 2007 Morphological based technique for image segmentation *International Journal of Information Technology* **14(1)** 55-80
- [14] Sahoo P K 1988 A survey of thresholding techniques *Computer Vision, Graphics and Image Processing* **41** 233-260
- [15] Rutkowski W S 1979 Shape completion *Comput Vis Graph Image Process* **9** 89-101
- [16] Yakimov I, Kirpichnikov A, Mokshin V, Yakhina Z and Gainullin R 2017 The comparison of structured modeling and simulation modeling of queueing systems *Communications in Computer and Information Science (CCIS)* **800** DOI: 10.1007/978-3-319-68069-9\_21
- [17] Tutubalin P I and Mokshin V V 2017 The Evaluation of the cryptographic strength of asymmetric encryption algorithms *IEEE Second Russia and Pacific Conference on Computer Technology and Applications (RPC)* 180-183 DOI: 10.1109/RPC.2017.8168094
- [18] Laurent H and Radu H 1993 Figure-Ground Discrimination: a Combinatorial Optimization Approach *IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers* **15(9)** 899-914
- [19] Sarkar S and Boyer K L 1998 Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors *Computer Vision and Pattern Recognition* **7(1)** 110-136
- [20] Williams L R, Jacobs D W 1997 Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Saliency *Neural Computation* **9(4)** 837-858
- [21] Borisova I V, Ilegkii V N and Kravec S A 2017 The application of the brightness gradient orientation for automatic object tracking systems *Computer Optics* **41(6)** 931-937 DOI: 10.18287/2412-6179-2017-41-6-931-937
- [22] Myasnikov E V 2017 Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches *Computer Optics* **41(4)** 564-572 DOI: 10.18287/2412-6179-2017-41-4-564-572
- [23] Goshin Y V and Kotov A P 2017 Parallel implementation of a multi-view image segmentation algorithm using the hough transform *Computer Optics* **41(4)** 588-591 DOI: 10.18287/2412-6179-2017-41-4-588-591