International Journal of Computer Science and Applications,

Vol. 16, No. 1, pp. 17 – 37, 2019

AUTOMATED CENTROID TUNING BASED ON PARTICLE SWARM OPTIMIZATION IN THE IMAGE CLUSTERING PROCESS

IULIIA KIM

ITMO University, Lomonosov Street 9,
191002, Saint Petersburg, Russia
yulia1344@gmail.com

ILYA VIKSNIN

ITMO University, Lomonosov Street 9,
191002, Saint Petersburg, Russia
wixnin@cit.ifmo.ru

IGOR KOTENKO

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS),
4-th Liniya, 39, Saint-Petersburg, 199178, Russia
ivkote@comsec.spb.ru
http://comsec.spb.ru/kotenko/

Computer vision is one of the most prospective fields in different spheres of life. It can be embedded in devices in order to automate their working processes and make them perform pattern recognition tasks. In this paper computer vision is considered in the context of unmanned vehicles elaboration. To provide a correct work of pattern recognition, it is necessary to elaborate effective methods of image processing, and clustering is one of them. However, many existing clustering methods depend on manual tuning and do not guarantee accurate results in case of visual information distortion. Clustering method based on particle swarm optimization (PSO) was elaborated, and it provides automated centroid calculation and element distribution to clusters without human factor. Relative to PSO, particle group motion (pixel-by-pixel passage through the image), search for the best solution for the entire swarm (search for pixel with a maximum average intensity value in a certain region) were used. As for k-means method, the need of cluster amount predetermining by user was eliminated, apart from distance function minimization an operation of color function minimization was added. The developed clustering method was tested on 860 car images and 860 road signs images. Its results and contribution to the pattern recognition quality improvement were assessed in comparison with fuzzy C-means and original k-means. To assess the obtained results two measures were introduced: detection completeness and detection deviation. PSO clustering method showed the best results in them, which disclosed its high utility to the pattern recognition.

Keywords: Computer vision; image clustering; particle swarm optimization; unmanned vehicles; road traffic and territory monitoring; computer vision in bad weather conditions.

1. Introduction

During the last several decades scholars from many countries have conducted large-scale research in the area of artificial intelligence [Spector (2006)], and machine learning is quite a significant part of it. Despite this field has emerged relatively recently, machine learning has already obtained an implementation in many spheres of society, such as medical and technical diagnostics, augmented reality development, speech recognition and computer vision [Forsyth and Ponce (2003); Kochan (2002); Vernon (2004); Kozlov, *et al* (2014)].

Computer vision is a wide area of theoretical approaches and technical methods connected with object detection, object tracking and object classification. It can be embedded in devices in order to automate their working processes and make them perform pattern recognition tasks. To provide the correct work of pattern recognition, it is necessary to elaborate efficient methods of image processing; and clustering is one of them. However, many existing clustering methods depend on manual tuning and do not guarantee accurate results in case of visual information distortion.

The paper is dedicated to the ways of pattern recognition quality improvement. Relevance of the proposed paper to the computer vision area can be explained by the fact that computer vision can be embedded in cars, drones and used in order to automate such process as road traffic and territory monitoring. It is vital to provide a correct work of computer vision mechanisms, because their errors can have fatal consequences: victims, damages that consequently can cause large financial costs.

In order to avoid such consequences, the purpose to elaborate efficient methods of image processing is stated. Among them it is possible to call segmentation or its particular case - clustering. However, they do not guarantee accurate results in case of visual information distortion or high noise level. The reason is in the need of manual tuning of these methods.

This paper is the extended version of one published in proceedings of MMAP 2018 [Kim, *et al* (2018)], where clustering method based on particle swarm optimization (PSO) was developed, which analyzes image structure, determines cluster centroids and then allocates image pixels in different clusters automatically. The paper includes the image results of the improved program code: its speed was increased by analysis of neighboring regions horizontally and vertically at the same time. In addition, two measures of pattern recognition quality were introduced in order to assess contribution of clustering method: detection completeness and detection deviation.

Further structure of this paper is as follows. Section 2 provides the overview of the relevant works and sets the theoretical background. Section 3 considers the research task statement and the suggested clustering method based on particle swarm optimization. Section 4 demonstrates the experimental results, their assessment is given in Section 5. Conclusions and future research directions are outlined in Section 6.

2. Related work and theory background

2.1. *Overview of related projects*

The projects dedicated to computer vision for unmanned vehicles are being developed actively and are on the verge of integrating it into the everyday life of society. In [Rusanov and Nekrasov (2016)] a review of existing projects was conducted. Google Car is a commercial self-driving car project intended for making road traffic process safer and easier. It uses an algorithm based on involving sensors, detectors, radars, digital video recorders, which are installed on a car body and send data to an on-board computer. The computer analyzes the received information and then makes a decision about further movements. However, the self-driving cars are not well-adapted to movements in bad weather conditions such as fog or precipitations. In addition, pattern recognition algorithm not always distinguishes objects correctly, for example, crumpled sheet of paper can be classified as stone.

Another project, Spirit of Berlin, was started in 2007 at the Free University of Berlin. The project focuses on using scanners, general and omnidirectional cameras in analyzing road marking and surrounding objects. The algorithm does not work effectively, if the road marking is not clearly notable.

The purpose of developing a car model, which does not depend on road and weather conditions, is pursued by Italian company VisLab. Their algorithm tunes detector automatically according current conditions of environment. The created model passed tests successfully, however, it does not provide autonomous movement.

Self-automated cars were developed also in many others research projects, but one of the main problems is that the vision model requires manual tuning, but analysis and movement can be incorrect in case of impassibility of roads. One of the issues emerged i noise presence and blur in the transmitted images.

Example of such low quality image is represented in Fig. 1.

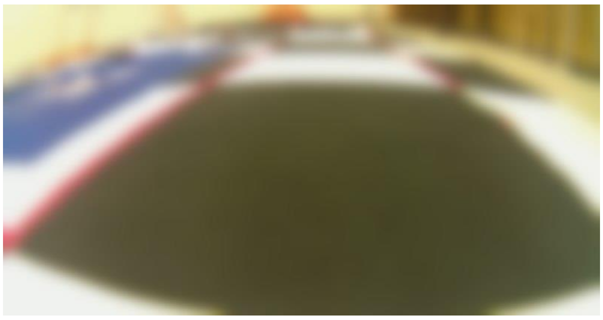

Fig. 1. Example of transmitted low quality image in self-automated car.

According to the enumerated examples, it is possible to make a conclusion that during developing unmanned vehicles it is vital to focus on pattern recognition quality, specifically on error percentage reduction. In this paper the pattern recognition error is a situation when the needed object in the image either is not detected or is detected incorrectly.

2.2. *Illumination of pattern recognition process*

Pattern recognition process consists in three basic stages:

- filtering and image preparation for the analysis; the image preparation includes image compression, selection of needed regions, ridding of noise;
- logical processing of filtering results, which is responsible for object detection in the filtered image;
- decision-making based on the results of the logical processing, which implies classification of the detected objects.

Image segmentation [Aly, *et al* (2011)] is related to the ways of increasing the accuracy of recognition. It is a process of dividing a digital image into a set of its constituent regions in order to select objects and their boundaries. As there is no general solution for the image segmentation task, different methods and algorithms were created, which are oriented to the determined categories of images. Depending on the category of the image, its priority properties and then the ways of grouping them are chosen.

The methods listed below can serve as examples:

- active contour method [Xiang, *et al* (2005)] (deformation of original image contours to the boundaries of the specified objects);
- topological alignment method (matching two consecutive frames in the image stream);
- watershed method [Li, *et al* (2007)] (establishment of boundary watersheds between different segments according to the determined rule).

As for semantic part of the visual information, the following kinds of image segmentation can be distinguished:

- semantic segmentation based on the use of Fully Convolutional Networks (pixel-to-pixel mapping without prior allocation of specific areas);
- weakly supervised semantic segmentation (with use of bounding frames and special labels on the images);
- region-based semantic segmentation (region allocation based on predetermined grouping rules).

One of the segmentation methods is clustering. Image clustering is a division of pixels into several non-intersecting groups (clusters) in such a way that pixels from the same group have similar features, meanwhile the features of pixels from different groups vary significantly from each other.

Let suppose that there are: X – a set, which consists of N objects; C – a set, which consists of M identifiers, such as a number, a name or a label; $d(x, x')$ – the distance function between objects, where x and x' are two objects in the image. The distance function is represented in (1):

$$d(x, x') = \sqrt{(x - x')^2} \,. \tag{1}$$

It is necessary to divide the set of objects X into M non-intersecting subsets (clusters) in such a way that each cluster was represented as an aggregate of objects from the set X, whose distance function values d are closed to each other.

In addition, the following conditions must be fulfilled:

- each cluster is assigned a cluster identifier C_j, $j \in [1; M]$ (number, name, label);
- each object x_i, $i \in [1; N]$ can belong to one and only one cluster.

There are plenty of different clustering methods. Some of the most spread examples of these methods are *k*-means and fuzzy *C*-means. The idea of the *k*-means algorithm is to minimize the distance between objects in the clusters. The algorithm stops working when the further minimization becomes impossible.

The main steps of the *k*-means algorithm are as follows:

1) At the beginning of the algorithm the quantity of clusters is set and then, according to the determined rule, centroids are allocated (centers of mass of clusters). The minimizing function is represented in (2):

$$J = \sum_{i=1}^{N} \sum_{k=1}^{M} d(x_i, c_k) \,, \tag{2}$$

where X – a set of clustering objects, $x_i \in X$ a clustering object, $i \in [1; N]$ *i* ∈ *[1; N]*, C – a set of clusters, $c_k \in C$ – centroid, $k \in [1; M]$, M – an amount of clusters, N – an amount of objects, d – a value of distance function between object and centroid.

2) Each object correlates with the determined cluster by calculating the value of the distance function between this object and each center of mass and then selecting the least one among the calculated values. After that the centers of mass of clusters are recalculated, as in (3):

$$c_j = \frac{\sum_{t=1}^{T} x_t}{T} \,, \tag{3}$$

where $x_t \in C_j$, $t \in [1; T]$; T – an amount of objects in the cluster C_j; j – a cluster number, $j \in [1; M]$; M – an amount of clusters.

3) If $c_j = c_j - 1$, it means that object clustering is completed, otherwise it is necessary to return to the second step and recalculate centroids again.

Fuzzy *C*-means algorithm is based on the fuzzy logic, i.e. on the assumptions that each clustering object from the set X, which consists of N objects to some extent belongs to a particular cluster from the set of clusters C.

The main steps of the fuzzy *C*-means algorithm:

1) As input values, there are: M – an amount of clusters, $1 < m < \infty$ – a measure of accuracy, $0 < \varepsilon < 1$ – a criterion of the end, $U_0 = u_{ij}(x_i, c_j)$: $x_i \in X$, $c_j \in C$ – a weighting matrix of belonging of the clustering object $x_i \in X$, $i \in [1; N]$ to the cluster $C_j \in C$, $j \in [1; M]$; $0 < u_{ij} < 1$.

The minimizing function is shown in (4):

$$J = \sum_{i=1}^{N} \sum_{k=1}^{M} u_{ij}^{m} \cdot d(x_i, c_k) \,. \tag{4}$$

2) The centroids are calculated, as in (5):

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \,. \tag{5}$$

3) After this the weights are recalculated, as in (6):

$$u_{ij} = \frac{1}{\sum_{k=1}^{M}\left(\frac{d(c_j,x_i)}{d(c_k,x_i)}\right)^{\frac{2}{m-1}}}. \quad (6)$$

4) The next step is to compare $|U_k - U_{k-1}|$ with the value of ε. In case the first value is less than the second one, the algorithm is finished, otherwise it is necessary to return to the step 2 and recalculate the centroids.

3. Clustering method based on particle swarm optimization

The algorithms mentioned above have quite substantial drawbacks, such as:

- sensitivity to outliers (values that are allocated from the total sample);
- need for the user to specify the amount of clusters beforehand;
- presence of some uncertainty degree in the threshold parameter definition;
- uncertainty of actions with objects that possess the properties of two different clusters simultaneously;
- need for the user to specify clustering parameters.

Due to these shortcomings of clustering algorithms, there is such a situation when the clustering task can be solved incorrectly, namely: for the clustering objects $x_i \in X$, $i \in [1; N]$, and the clusters $C_j \in C$, $j \in [1; M]$ with the centroids c_j, $\exists t \neq j, g_{x_i} \neq g_{c_t}, g_{x_i} = g_{c_j}, x_i \in C_t, x_i \notin C_j$, where g_{x_i} – a label of belonging of the i^{th} object to the cluster, g_{c_j} – a label of belonging of the centroid to the j^{th} cluster.

In this way, authors of this paper have the following research task: it is necessary to find such centroids c_j for the $C_j \in C$ that $\forall x_i \in X\, \exists!\, C_j$, $g_{x_i} = g_{c_j}, x_i \in C_j$, $g_{x_i} \neq g_{c_a}, x_i \notin C_a, g_{x_i} \neq g_{c_b}, x_i \notin C_b, a \in [1; j-1], b \in [j+1; M]$.

To solve this pattern recognition quality improvement task, it is essential to exclude the user's direct participation in the process of specifying the amount of clusters and a rule (or a set of rules) during the initial centroid allocation.

Methods providing automated image segmentation and clustering were proposed in many scientific works. In the [Woźniak and Połap (2018)] processing of image segments for fruit peel defects detection was proposed. The source image is divided into segments, which are given as an input to Adaptive Artificial Neural Network (AANN). The image processing is automated, because AANN adapts to each input and its features.

Segmentation based on color maps was developed in [Li, *et al* (2018)]. Its authors rely on four-color labeling theorem. They improved existed four-color labeling method, which used random initial distribution and developed heuristic four-color labeling. This iterative algorithm generates more reasonable color maps with a global view of the whole image and provides better results in case of images with clutters and complicated structures.

In [GeethaRamani (2018)] image clustering is used in order to provide automated retinal screening. In the described method the heuristic based clustering is included. Initial centers allocated according to measures defining statistical data distribution are incorporated in the proposed methodology.

Another effective clustering method was proposed for heterogeneous disease expression data [Huang, *et al* (2013)]. Recursive *K*-means spectral clustering method (ReKS) was developed, which was found to be superior to the hierarchical clustering method and much faster than k-means.

In [Lu, *et al* (2018)] a novel data clustering algorithm was elaborated. It is based on heuristic rules, which are built according to k-nearest chain, and give an opportunity to get rid of the need in specifying the number of clusters. *K*-Nearest Neighbors Chain (KNNC) serves as a basis for proposing two heuristic rules to find initial clusters and their proper amount. The first heuristic rule reflects the degree of separation of clusters and the second rule determines the inner compactness of a cluster.

In order to resolve the issue of arbitrary choices on clustering parameters, the authors decided to use some elements of particle swarm [Karpenko and Seliverstov (2009)] optimization.

Particle swarm optimization [Wang and Wang (2008)] method consists in the following steps:

1) There is a swarm of S particles, and each of them is assigned a coordinate $x_i \in \boldsymbol{R_n}$ and a velocity $v_i \in \boldsymbol{R_n}$; $f: \boldsymbol{R_n} \to \boldsymbol{R}$ is an objective function that needs to be minimized; p_i – the best known position of the its particle (in the context of solving the given optimization problem); g – the best known state of the entire swarm.

2) For each particle $s_i \in S$, $i \in [1; S]$ it is needed to:

- generate an initial position using a random vector in the range from r_{min} to r_{max}, these values are lower and upper boundaries of the search-space, respectively;
- assign to the best known position of the particle p_i its initial value x_i;
- in case $f(p_i) < f(g)$, there is a necessity to update the value from g to p_i;
- generate velocity value of the particle v_i, which belongs to the interval from $-(r_{max} - r_{min})\_max - r\_min)$.

3) It is required to repeat the following sequencing for each i^{th} particle until the predetermined stopping criterion is fulfilled:

- generate random vectors r_p and r_g, which have a range of admissible values in the interval between 0 and 1;
- update the velocity value of the particle, as in (7):

$$v_i = w \cdot v_i + \varphi_p \cdot r_p \times (p_i - x_i) + \varphi_g \cdot r_g \times (g - x_i), \tag{7}$$

where $\times$ is a vector product operation; w, φ_p, φ_g are the parameters specified by user;

- change the particle position according to (8):

$$x_i = x_i + v_i\ ; \tag{8}$$

- compare the values of $f(x_i)$ and $f(p_i)$; if the first value is less than the second one, it is needful to update the best known position of the particle from p_i to x_i and then in case $f(p_i) < f(g)$ it is necessary to change the value of the parameter g to the value of the parameter p_i.

4) As a result of the operations above, the parameter g will contain the best solution.

For pattern recognition quality improvement the authors of the paper developed a clustering method which combines some elements from particle swarm optimization (numerical optimization method) and from *k*-means algorithm (cluster analysis method). From each algorithm such operations were selected that do not require random parameter settings and do not take into account user's subjective opinion (user has just an observing role). In this clustering method, all calculations will happen automatically, and the user no longer needs to generate manually any input values.

It is necessary to normalize the source image before using the developed algorithm. Normalization allows making an image insensitive to the light changes, ridding it of unnecessary noise.

It is achieved by dividing the RGB vector components of each pixel by the length of this vector, as in (9):

$$(r', g', b') = \left(\frac{r}{\sqrt{r^2+g^2+b^2}}, \frac{g}{\sqrt{r^2+g^2+b^2}}, \frac{b}{\sqrt{r^2+g^2+b^2}}\right), \tag{9}$$

where r, g, b are the initial values of pixel's RGB vector; r', g', b' are the normalized values of pixel's RGB vector.

The algorithm of the developed clustering method consists in the following seven procedures:

1) Rounding W' pixels horizontally and H' pixels vertically to the nearest values of W and H, respectively, which are multiple of 10.

2) Sequential selection of 10 by 10 regions (clusters) in the image and search for a pixel with a maximum average intensity value in each region – these pixels will be the centers of mass c_j; $j \in \left[1; W \cdot \frac{H}{100}\right]$ (in case there are more than one pixel with a maximum average intensity value in the region, it is possible to choose any of them). The general formula of pixel's average intensity calculation is represented in (10):

$$I_{av} = \frac{r'+g'+b'}{3}, \tag{10}$$

where r', g', b' are the normalized values of pixel's RGB vector.

Thus, the condition of centroid determination can be expressed as $c_j = x_q$ $\forall s \neq q,\ I_{av_q} > I_{av_s},\ I_{av_q} = \frac{r'_{x_q}+g'_{x_q}+b'_{x_q}}{3},\ I_{av_s} = \frac{r'_{x_s}+g'_{x_s}+b'_{x_s}}{3};\ s, q \in [1; P]$, where P is the total amount of pixels in the j^{th} cluster.

3) Comparison of the rounded average intensity values for elements with maximum average intensity values from neighboring regions. Comparison goes vertically and horizontally relatively to each pixel. It was found empirically by authors that the most effective rounding is to two decimal places. If the rounded average intensity values are equal to each other, two neighboring clusters are combined into one. In the new cluster the centroid is the pixel with a maximum average intensity value. It is necessary to repeat this step until there are M clusters c_j, $j \in [1; M]$, with the pairwise distinct rounded average intensity values of the centers of mass.

4) Calculation of two parameters for each pixel x_i, $i \in [1; W \cdot H]$, relative to each centroid - the distance function d and the so-called color function f. The color function is represented in (11):

$$f(x_i, c_j) = \sqrt{\left(r'_{x_i} - r'_{c_j}\right)^2 + \left(g'_{x_i} - g'_{c_j}\right)^2 + \left(b'_{x_i} - b'_{c_j}\right)^2}\ , \tag{11}$$

where r'_{x_i}, g'_{x_i}, b'_{x_i} are the normalized [Xie, *et al* (2013)] values of the RGB vector of the pixel x_i; c_j is the centroid of the cluster C_j; r'_{c_j}, g'_{c_j}, b'_{c_j} are the normalized values of the RGB vector of the centroid c_j.

5) Then for the pixel x_i it is necessary to find a centroid c_a, $a \in [1; M]$, relative to which the square root of distance function value will be minimum and a centroid c_b, $b \in [1; M]$, relative to which the value of color function value will be minimum. After that the following differences need to be calculated, as it is represented in (12) and (13):

$$d_{diff} = |d(x_i, c_a) - d(x_i, c_b)|\ , \tag{12}$$

$$f_{diff} = |f(x_i, c_a) - f(x_i, c_b)|\ . \tag{13}$$

6) The function, whose difference was less, is chosen as a priority function (in case d_{diff} equals f_{diff}, the distance function obtains a priority, because pixels that are closer to each other more likely belong to the same object than the ones that have similar colors). The allocation of pixels to clusters is realized according to the priority function, i.e. the pixel x_i will be assigned to a cluster, if the priority function value between this pixel and this cluster's centroid is minimal. Thus, the objective function of the elaborated clustering method takes the form, as it is represented in (14):

$$\begin{cases} \sum_{k=1}^{M} \sum_{i=1}^{N} d(x_i, c_k) \to min \\ \sum_{k=1}^{M} \sum_{i=1}^{N} f(x_i, c_k) \to min \end{cases} . \tag{14}$$

7) Ridding the image of noise. For this purpose the authors chose a non-local means method. It is illustrated in (15):

$$u(p) = \frac{1}{C(p)} \int_{\Omega} v(q) f(p, q) dq \tag{15}$$

where $u(p)$ is the filtered intensity value of the pixel color component at the point p, $v(q)v(q)$ is the unfiltered intensity value of pixel color component at point q; $f(p, q)$ – the weighting function, $C(p)$ – the normalizing factor.

As the weighting function the Gaussian function is used, it is shown in (16):

$$f(p, q) = e^{\frac{-|B(q) - B(p)|^2}{h^2}} \tag{16}$$

where h is the filter parameter (in general, for the RGB color images $h = 3$), $B(p)$ is the local average intensity value of color components of pixels around the point p, $B(q)$ is the local average intensity value of color components of pixels around the point q.

Normalizing factor $C(p)$ is calculated, as in (17):

$$C(p) = \int_{\Omega} f(p, q) dq \tag{17}$$

The developed clustering method uses the following elements borrowed from the particle swarm optimization method: particle group motion (pixel-by-pixel passage through the image), search for the best solution for the entire swarm (search for a pixel with a maximum average intensity value in a certain region).

At the same time the main differences from the original algorithm are the next points: each particle has a fixed velocity value which excludes the necessity of its manual recalculating by the user, the initial particle parameters are not specified randomly.

As for the *k*-means cluster analysis method [Qi, *et al* (2016)], the next aspects were improved: the need of cluster amount predetermining by the user was eliminated, apart from distance function minimization an operation of color function minimization was added, which gave an opportunity to increase the probability that pixels will be assigned to clusters correctly.

4. Experiments

To check the effectiveness of the clustering method based on particle swarm optimization 860 test images of cars and 860 images of road signs were picked and normalized.

For this purpose the mixture of manual photos, images provided by Stanford University laboratory [Cars Dataset (2019)] and images from Russian Traffic Sign Dataset [Russian Traffic Sign Dataset (2019)] (both datasets are publicly available) were used.

Fig. 2 and Fig. 3 outline examples from this set of test images.

Fig. 2. Source image of a car.

Fig. 3. Source image of a road sign.

The normalization [Choudhury, *et al* (2017)] results of these source images are shown in Fig. 3 and Fig. 4.

Fig. 4. Normalized image of the car.

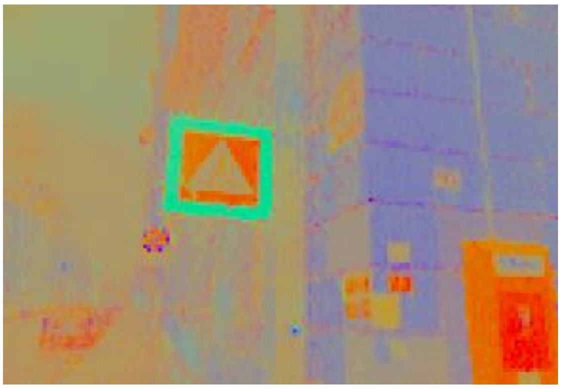

Fig. 5. Normalized image of the road sign.

Fig. 6 and Fig. 7 illustrate the work results of the clustering method based on particle swarm optimization.

Fig. 6. Clustered image of the car by dint of PSO.

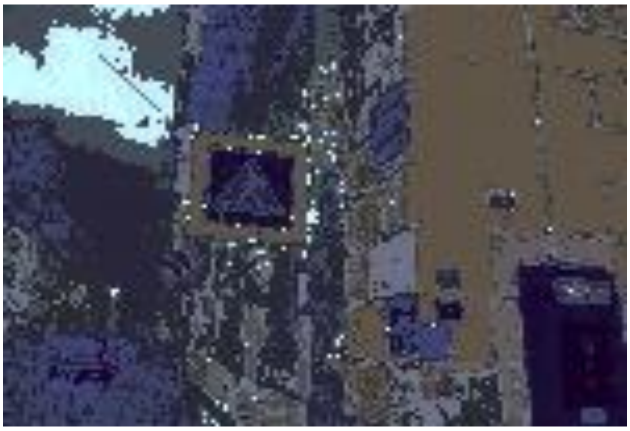

Fig. 7. Clustered image of the road sign by dint of PSO.

5. Assessment of clustering based on PSO

During the experiment the following shortcomings of the elaborated clustering method were disclosed: there is no consistent pattern in determining the amount of decimal places in the rounded average intensity value. The optimal rounding of this parameter for cluster combining was empirically established by the authors equal to 2 and it is not universal.

Such drawback may provoke two opposite situations:

1) there are extra clusters in the output image, especially in the places of glare;

2) in the output image several different objects are merged into one cluster.

The examples of work of the clustering method based on PSO with reduced number of decimal places (1 decimal place) in the rounded average intensity value are represented in Fig. 8 and Fig. 9. All the objects were combined into one on the car image. As for the road sign picture, despite its clear sign representing, the image loses in the color constituent.

Fig. 8. Example of clustering quality worsening because of reduced number of decimal places in the rounded average intensity value (car image).

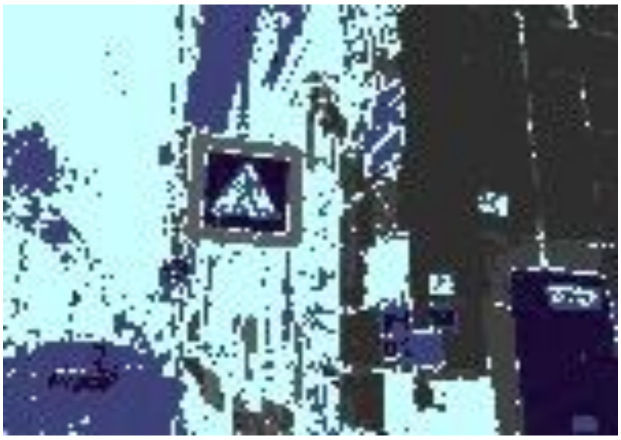

Fig. 9. Example of clustering quality worsening because of reduced number of decimal places in the rounded average intensity value (road image).

The examples of work of the clustering method based on PSO with increased number of decimal places (3 decimal places) in the rounded average intensity value are represented in Fig. 10 and Fig. 11. In the given image there are extra detected regions: piece of land, parts of sky, building.

Fig. 10. Example of clustering quality worsening because of increased number of decimal places in the rounded average intensity value (car image).

Another significant shortcoming of the PSO clustering method consists in the fact that the method depends on the input image size: the more the number of the analyzed pixels, the greater is the risk that more clusters will be detected, and due to this the time of work will be enlarged. This causes the necessity of proportional reducing the source image size or increasing the initial cluster size. To perform clustering, the authors reduced proportionally the image size. Currently the size of reduced image needs to be established empirically. Particularly, for the images represented in this paper the maximum size of 150 pixels wide and tall was chosen.

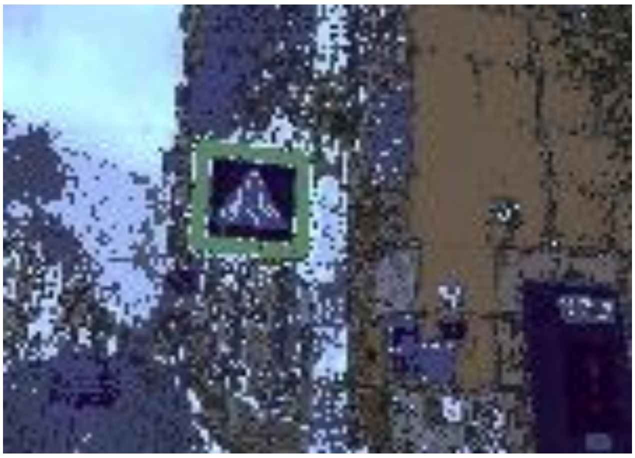

Fig. 11. Example of clustering quality worsening because of increased number of decimal places in the rounded average intensity value (road sign image).

The examples of work of the clustering method based on PSO with increased number of resultant cluster quantity due to absence of preliminary image size reduction are represented in Fig. 12 and Fig. 13. It is notable that in the resultant image many extra clusters, distorting contours structure, were detected that during further analysis can provoke different issues.

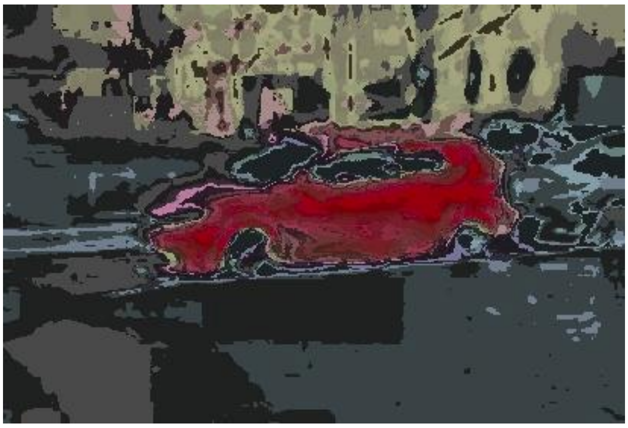

Fig. 12. Example of clustering quality worsening because of a large size of the source image (car image).

To improve the effectiveness of this method, it is planned to disclose dependencies between image features (pixel quantity, histogram of gradients) and such aspects as: number of decimal places in rounded average intensity value, and initial cluster size.

To perform comparison, apart from the clustering method based on PSO, the test images were clustered by the fuzzy C-means and the k-means (as an input value, the number of clusters obtained in the PSO clustering method was used; maximum iteration quantity – 25).

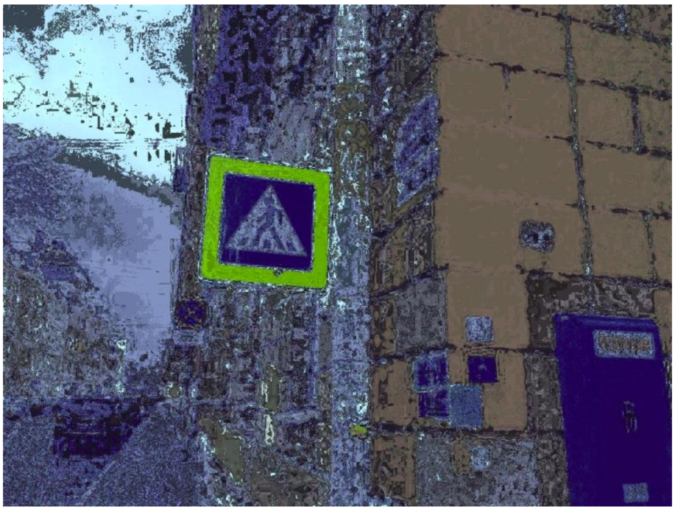

Fig. 13. Example of clustering quality worsening because of a large size of the source image (road sign image).

Fig. 14 and Fig. 15 represent the worked results of the fuzzy C-means clustering method. Fig. 16 and Fig. 17 depict the worked results of the k-means clustering method. With k-means and fuzzy C-means clustering methods it became possible to separate objects from background. However, the algorithm, proposed by the authors, marks details not only legibly, but also without marking extra glares.

With k-means and fuzzy C-means clustering methods it became possible to separate objects from background. However, the algorithm proposed by the authors marks details not only legibly, but also without marking extra glares.

To illuminate influence of the represented clustering methods on the pattern recognition quality the authors implemented Haar cascade classifier [Choudhury, *et al* (2017)] on the given example to detect the car or road sign. It is possible to see the results in Fig. 18-23.

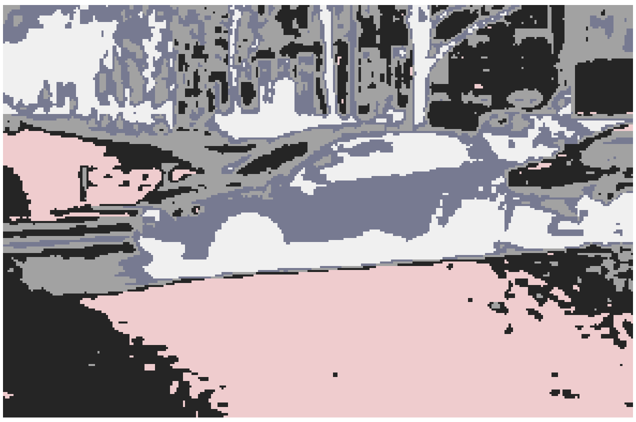

Fig. 14. Clustered image of the car by fuzzy C-means clustering method (car image).

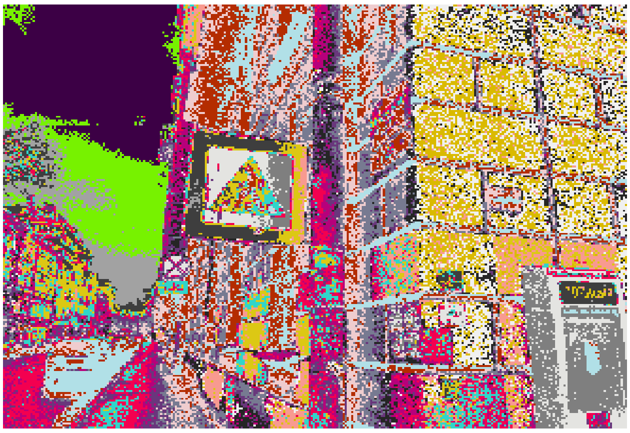

Fig. 15. Clustered image of the car by fuzzy C-means clustering method (road sign image).

Fig. 16. Clustered image of the car by k-means clustering method (car image).

Fig. 17. Clustered image of the car by k-means clustering method (road sign image).

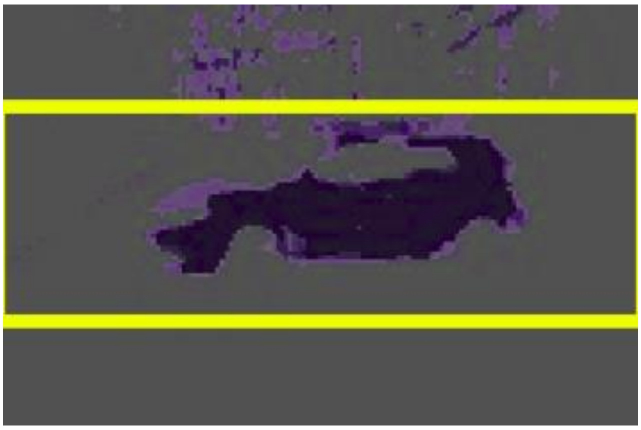

Fig. 18. Result of classifying (PSO clustered image of car).

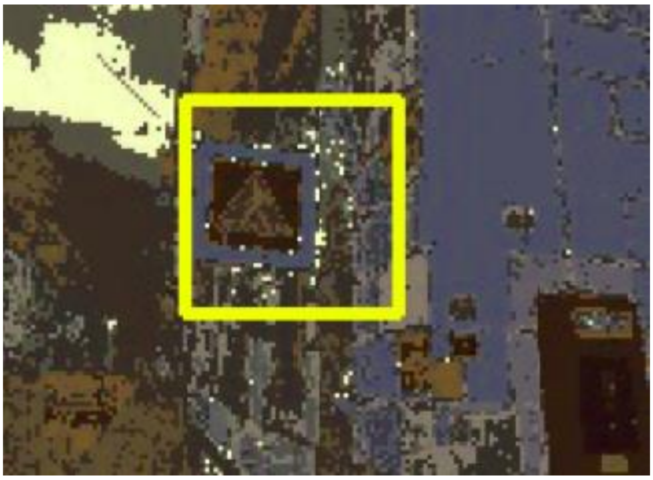

Fig. 19. Result of classifying (PSO clustered image of road sign).

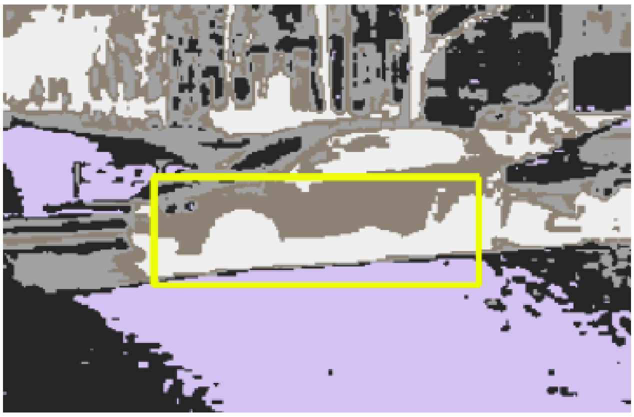

Fig. 20. Result of classifying (fuzzy C-means clustered image of car).

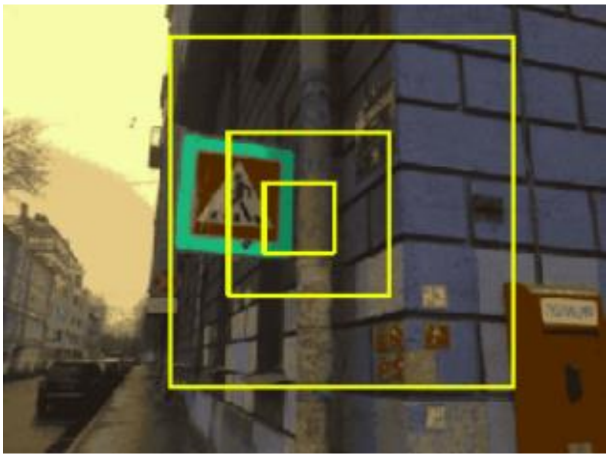

Fig. 21. Result of classifying (fuzzy C-means clustered image of road sign).

Fig. 22. Result of classifying (k-means clustered image of car).

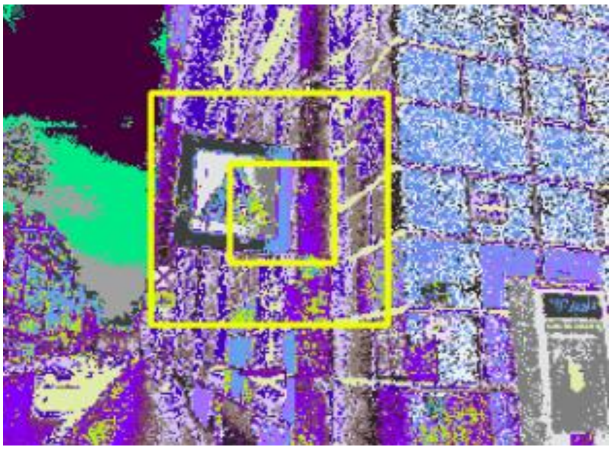

Fig. 23. Result of classifying (k-means clustered image of road sign).

In order to conduct the assessment of work effectiveness for three enumerated clustering methods in the context of object detection, two criterions were introduced: (1) detection completeness and (2) in case of first criterion presence - deviation of the detected region from the real object size. Detection completeness means that the given object is completely included in the marked area without losing any details ("+" or "-"). Results relatively to this parameter are represented in Table 1.

Table 1. Comparison table of detection completeness for different clustering methods.

	PSO	**Fuzzy *C*-means**	***k*-means**
Car	+	-	+
Road sign	+	+	+

For the cases, where detection completeness is positive, the deviation of the marked area size relatively to the given object real size was calculated. To simplify this task, objects are considered as the areas of rectangular form (the area is determined according to the extreme points of the object in the image), and its square value is counted. The results are shown in Table 2 in the relation format of the detected area square to the object area square. In the situations, when several regions were detected by Haar classifier, the largest area was taken into account.

Table 2. Comparison table of detected region deviation for different clustering methods.

	PSO	**Fuzzy *C*-means**	***k*-means**
Car	2,04	-	1,14
Road sign	2,14	3,27	7,09

PSO clustering method gives the least deviation value except the case with k-means clustering of the road sign image. However, unlike k-means, PSO clustering does not detect the image extra regions nested in each other.

6. Conclusion

The paper proposed and investigated the clustering method based on particle swarm optimization. The developed method was tested for a car vision system and its results and contribution to the pattern recognition quality improvement were evaluated.

It was noted that sometimes it can detect and mark the bigger region in the image than the real size of the object in this region. This aspect can be corrected by improving training samples for classifiers. The advantages of the proposed method are as follows:

(1) use of particle swarm optimization helped to eliminate the necessity for user of cluster amount predetermining - one of the main reasons of incorrect clustering in the majority of cluster analysis methods;

(2) the authors liquidated the need for users of calculating and specifying the threshold parameters; this need often led to a situation when one pixel could be assigned to several clusters at the same time which contradicts with the clustering task;

(3) there are rules that regulate grouping of objects, which possess the features of the different clusters; thanks to this the uncertainty in correlation of objects and clusters is minimized and, as a result, the pattern recognition error probability is reduced;

(4) clustering parameters are predetermined automatically and do not require user's intervention.

However, currently, this method has the following disadvantages: (1) absence of clear dependencies between way of rounding the average intensity values and image features; (2) absence of clear dependencies between size of initial cluster and image features.

In further research it is planned to liquidate the mentioned drawbacks of the developed clustering method, propose the ways of its improvement and to finalize the embedded vision system intended for unmanned vehicles (cars and drones).

Acknowledgments

This research is being partially supported by the grants of the RFBR (projects No. 16-29-09482, 18-07-01488 and 18-29-22034), by the budget (the project No. 0073-2019-0002), and by Government of the Russian Federation (Grant 08-08).

References

Aly, A.; Deris, S.; Zaki, N. (2011): Research review for digital image segmentation techniques. IJCSIT, 3(5), pp. 99-106.

Cars Dataset (2019): URL: https://ai.stanford.edu/~jkrause/cars/car_dataset.html.

Choudhury, S.; Chattopadhyay, S. P.; Hazra, T. K. (2017): Vehicle detection and counting using haar feature-based classifier. The 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON), pp. 106-109.

Forsyth, D.; Ponce, J. (2003): *Computer Vision: A Modern Approach*. Saddle, River, Nj, U, A: Prentice Hall.

GeethaRamani, R. (2018): Macula segmentation and fovea localization employing image processing and heuristic based clustering for automated retinal screening. Computer methods and programs in biomedicine, vol. 160, pp. 153-163.

Huang, G.T.; Cunningham, K.I.; Benos, P.V.; Chennubhotla, C.S. (2013): Spectral clustering strategies for heterogeneous disease expression data. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, pp. 212-223.

Karpenko, A.; Seliverstov, E. (2009): Overview of the particle swarm methods for the global optimization problem. Science and Education: a scientific edition of the Bauman Moscow State Technical University, 3, pp. 1-26.

Kim, I.; Matveeva, A.; Viksnin, I.; Kotenko I. (2018): Image Clustering Method based on Particle Swarm Optimization. Proceedings of the 2018 Federated Conference on Computer Science and Information Systems.

Kochan, A. (2002): Machine vision guides the automotive industry. Sensor Review 22(2), pp. 119-124.

Kozlov, S.; Neretin, E.; Kukolkina, V. (2014): Machine vision application in digital dermoscopy for suspected melanoma of the skin. Saratov Journal of Medical Scientific Research, 10(2), pp. 281-285.

Li, K., Tao, W., Liu, X., Liu, L. (2018): Iterative image segmentation with feature driven heuristic four-color labeling. Pattern Recognition, vol. 76, pp. 69-79.

Li, N.; Liu, M.; Li, Y. (2007): Image Segmentation Algorithm using Watershed Transform and Level Set Method. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP ’07), pp. 613-616.

Lu, J.; Zhu, Q.; Wu, Q. (2018): A novel data clustering algorithm using heuristic rules based on k-nearest neighbors chain. Engineering Applications of Artificial Intelligence, vol. 72, pp. 213-227.

Qi, J.; Yu, Y.; Wang, L.; Liu, J. (2016): K*-Means: An Effective and Efficient K-Means Clustering Algorithm. IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom), pp. 242-249.

Rusanov, A.; Nekrasov, D. (2016): Review of the working principles and algorithms for recognition of environmental objects in unmanned cars. New information technologies in automated systems, 19, pp. 323-329.

Russian Traffic Sign Dataset (2019): URL: http://graphics.cs.msu.ru/en/research/projects/rtsd.

Spector, L. (2006): Evolution of artificial intelligence. Artificial Intelligence, 170(18), pp. 1251–1253.

Vernon, D. (2004): *Machine Vision in the Electronics and PCB Inspection Industry. The Current Position and Future Directions*. Maynooth College Ireland.

Wang, J.; Wang, D. (2008): Particle swarm optimization with a leader and followers. Progress in Natural Science, 18(11), pp. 1437-1443.

Woźniak, M.; Połap, D. (2018): Adaptive neuro-heuristic hybrid model for fruit peel defects detection. Neural Networks, vol. 98, pp. 16–33.

Xiang, Y.; Chung, S.; Ye, J. (2005): A new active contour method based on elastic interaction. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), vol. 1, pp. 452-457.

Xie, L.; Tian, Q.; Zhang, B. (2013): Feature normalization for part-based image classification. IEEE International Conference on Image Processing, Melbourne, pp. 2607-2611.