Methods of Semantic Integrity Preservation in the Pattern Recognition Process

Iuliia Kim, Anastasiia Matveeva, Ilya Viksnin, Roman Patrikeev
Saint Petersburg National Research University of Information Technologies, Mechanics and Optics
Saint Petersburg, Russian Federation
yulia1344@gmail.com, anastasiamatveevaitmo@gmail.com, Wixnin@cit.ifmo.ru, null_horizon@mail.ru

***Abstract*–Computer vision is a wide area of theoretical research and technical methods connected with object detection, object tracking and object classification. In this article computer vision is considered in context of embedding it into automobiles in order to automate the road traffic process through video stream analysis. During road traffic it is vital to detect objects quickly and correctly, so the authors pay attention to the pattern recognition quality, especially to the visual information semantic integrity preservation. Their main purpose is to find the ways of its possible improvement respectively to three basic stages of the pattern recognition process. To avoid semantic integrity violations of information in the initial stage of the image analysis the authors propose normalization; in the second stage new clustering method was developed based on particle swarm optimization and k-means algorithm; in the final stage of the pattern recognition process the Haar cascade classifier was used with normalized training samples. The obtained image processing algorithm was implemented in case of blurred and noisy images and proved its effectiveness respectively to the visual information semantic integrity preservation.**

I. INTRODUCTION

A. Computer vision implementation

Lately artificial intelligence obtained a wide spread in the world [1]. The authors of this article are going to illuminate an integrated part of artificial intelligence, computer vision, namely pattern recognition and problems connected with its quality improvement.

Questions related to the computer vision were covered repeatedly in books, research articles. For instance, in 2003 a fundamental work was published [2]. There were described such aspects as geometric and analytical image properties, filtration methods, segmentation, classification, object detection and search for extremums in the image.

The authors consider computer vision implementation in the context of embedding it in automobiles to automate road traffic by analyzing video stream, that is why its incorrect work can cause fatal consequences. Authors of various scientific works agree that one of the main purposes in the area of computer vision is the error percentage reduction in the pattern recognition [3], [4]. The pattern recognition error is understood as situation when the needed object either is not detected or is detected incorrectly.

B. Research purpose statement

Information integrity [5] is a category, which excludes the possibility of unauthorized operations with information during various operations. This aspect can be subdivided into three constituents: pragmatic, syntactic, semantic.

Pragmatic integrity – feature, which provides expounding facts based on reliable evidence. Syntactic integrity – information integrity category that guarantees information presentation using correct expressions and structures. Semantic integrity – category, which does not allow information context violations [6]. During the pattern recognition process there is a probability that an object will be detected incorrectly, and this will cause analysis errors, that is why the authors affirm semantic integrity in the area of computer vision plays the most significant role.

Based on studied sources, the research purpose was stated: elaboration of algorithm providing semantic integrity preservation and improving pattern recognition process quality.

Purpose realization required to perform the following tasks:

1) illumination of stages of the pattern recognition process;

There are many methods to improve pattern recognition quality, but in most cases they were offered separately from each other, that is why the authors want to propose a sequence of methods, which is possible to use from the beginning to the end of image processing. Thus, the next task is:

2) search for methods providing visual information semantic integrity preservation for each stage of the pattern recognition process.

II. THEORY AND LITERATURE REVIEW

A. Pattern recognition process

The pattern recognition process [7] consists in three basic stages:

1) image preparation for the analysis;

2) processing of the image preparation results;

ISSN 2305-7254

3) object detection and classification based on the results of the previous step.

Violation of visual information semantic integrity can happen due to the following reasons [6]:

- noise presence in the image;
- image blur;
- incorrect training samples for the classifier;
- insufficient training samples for the classifier.

B. Image preparation for the analysis

In the initial stage of the image analysis in [8] Laplacian of Gaussian filter is proposed. However, for video stream analyzing there is need in faster performing. Among other image preprocessing methods binarization can be named.

Binarization - image conversion to bicolor, black and white. It is based on calculation of intensity threshold. To do this, there are several methods, among them Otsu's method, Niblack's method, Aikewell's method [9].

C. Processing of the image preparation results

During the second stage of image processing to minimize the recognition error, in the article [10] ideas about necessity of using segmentation were expressed. Among segmentation methods it is possible to name: inversion method, active contour method [11], watershed method [12].

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; weakly supervised semantic segmentation; region-based semantic segmentation [13].

Clustering is considered a particular case of segmentation. One of the most used clustering methods is k-means. The idea of the k-means [14] algorithm is to minimize the distance between objects in the clusters. The algorithm stops working when the further minimization becomes.

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).

2) Each object correlates with determined cluster by calculating the value of distance function between this object and each center of mass and then selecting the least one among calculated values. After that, the centers of mass of clusters are recalculated, as in (1):

$$c_j = \frac{\sum_{t=1}^{T} x_t}{T}, \quad (1)$$

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

However, the k-means does not provide uniquely information semantic integrity preservation, because the user has to specify the amount of clusters beforehand. In general, the majority of clustering methods possess the following drawbacks:

- sensitivity to outliers (values that are allocated from the total sample);
- uncertainty of actions with objects that possess the properties of two different clusters simultaneously;
- need for the user to specify clustering parameters.

D. Object detection and classification

For classification it is possible to use Naive Bayes classifier, support vector machine, Haar cascades. Classifying with Haar cascades is admitted as fast and simple method for analyzing video in real time. This method is based on Haar-like features [15]. The high accuracy of object description requires the presence of big number of features. That is why for complicated object detection the Haar-like features are combined into cascade classifier. The work with cascade classifier includes two steps. The first one consists in tuning classifier with training samples. The second stage consists in using the tuned classifier. The image is represented in the form of matrix, as in (2):

$$II(x, y) = \sum_{i=0, j=0}^{i \le x, j \le y} I(i, j), \quad (2)$$

where I is pixel intensity of the source image.

Each matrix element is the sum of intensity values in the rectangle from the point (0,0) to the point *(x, y).* The training procedure happens during T iterations, and the result is cascade, which consists of T weak classifiers. The work of the trained classifier occurs as follows: an image with size $W \times H$ is given as an input. The algorithm scans it on eleven scales: the scanning window size is 24×24 pixels, and each next level is 1,25 times more than previous one.

III. Algorithm for Visual Information Semantic Integrity Preservation

According to the material listed before, an algorithm that provides information integrity preservation and thus, improvement of pattern recognition quality was elaborated.

A. Image preparation for the analysis

As the authors of this article are interested in computer vision implementation in the process of road traffic, it is impossible to neglect the color image components. Thus, for the further image analysis it was proposed to use normalization. It allows making an image insensitive to the light changes, ridding it of unnecessary noise. Normalization is achieved by dividing the RGB vector components of each pixel by the length of this vector, as in (3):

$$[r', g', b'] = \left[\frac{r}{\sqrt{r^2+g^2+b^2}}, \frac{g}{\sqrt{r^2+g^2+b^2}}, \frac{b}{\sqrt{r^2+g^2+b^2}}\right], \quad (3)$$

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 implementation of the normalization for the source image represented in the Fig. 1 illustrates the Fig. 2:

Fig. 1. Source image

Fig. 2. Normalized image

B. Processing of the image preparation results

Taking into account the enumerated shortcomings of existing clustering methods, a new one was developed, which allows establishing border between foreground and background with higher accuracy. It combines the elements of particle swarm optimization method [16] and k-means method.

Particle swarm optimization method is widely used in data processing, robotic systems [17] and consists in the following steps:

1) There is a swarm of particles, and each of them is assigned a coordinate $x_i \in \boldsymbol{R^n}$ and a velocity $v_i \in \boldsymbol{R^n}$; $f\colon \boldsymbol{R^n} \to \boldsymbol{R}$ is an objective function that needs to be minimized; p_i – the best known position of i^{th} particle; g – the best known state of the entire swarm.

2) For each particle it is needed to:

- generate an initial position using a random vector;
- assign to the best known position of 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 a velocity value v_i for the particle.

3) It is required to repeat the following sequencing for each i^{th} particle until the determined stop 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 (4):

$$v_i = w \cdot v_i + \varphi_p \cdot r_p \times (p_i - x_i) + \varphi_g \cdot r_g \times (g - x_i), \quad (4)$$

  where × is vector product operation; w, φ_p, φ_g are the parameters specified by user;
- change the particle position according to (5):

$$x_i = x_i + v_i, \quad (5)$$

- 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.

From these two methods such operations were selected that do not require manual parameter settings. From the particle swarm optimization method were borrowed such operations as 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). However, unlike the original method, each particle has a fixed velocity value which excludes the necessity of its manual recalculating by user, initial particle parameters are not specified randomly. As for k-means cluster analysis method, its next aspects were improved: the need of cluster amount predetermining by user was eliminated, apart from distance function minimization an operation of color function minimization was added. This aspect makes the greatest contribution to the semantic integrity provision, because a particular object in the image (in some cases – the part of object) is characterized with relative color uniformity and sharp change absence, so if the distance function d for the pixels a and b of one object tends to a minimum, color function f also will tend to a minimum.

The algorithm of the developed clustering method for the image $W \cdot H$ consists in the following procedures:

1) Sequential selection of regions (clusters) of the determined size in the image and search for a pixel with a maximum average intensity value in each region – these pixels will be centers of mass c_j. The formula of pixel's average intensity calculation is represented in (6):

$$I_{av} = \frac{r' + g' + b'}{3}, \quad (6)$$

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

2) Comparison of the centroid average intensity values from neighbor regions in order to unite clusters, where average intensity values of centers of mass diverse insignificantly.

3) Calculation of two parameters for each pixel x_i, $I \in [1; W \cdot H]$ relative to each centroid: the square root of the distance function d and the color function f. The formulas for distance function and color function, respectively, are represented in (7) and (8):

$$d(x_i, c_t) = (x_i - c_t)^2 \quad (7)$$

$$f(x_i, c_t) = (r'_{x_i} - r'_{c_t})^2 + (g'_{x_i} - g'_{c_t})^2 + (b'_{x_i} - b'_{c_t})^2, \quad (8)$$

where r'_{xi}, g'_{xi}, b'_{xi} are the normalized values of RGB vector of pixel x_i; c_t is the centroid of the cluster C_t; r'_{ct}, g'_{ct}, b'_{ct} are the normalized values of RGB vector of centroid c_t.

4) Then for pixel x_i it is necessary to find a centroid c_a, relatively to which the distance function value will be minimum and a centroid c_b, relatively to which the color function value will be minimum. After this, the following differences need to be calculated, as it is represented in (9) and (10):

$$d_{diff} = | d(x_i, c_a) - d(x_i, c_b) | \quad (9)$$

$$f_{diff} = | f(x_i, c_a) - f(x_i, c_b) | \quad (10)$$

5) 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). The allocation of pixels to clusters is realized according to the priority function, i.e. pixel x_i will be assigned to a cluster, if the priority function value between this pixel and this cluster's centroid is minimal.

6) Ridding the image of noise. For this purpose the authors chose non-local means method [18]. It is illustrated in (11):

$$u(p) = \frac{1}{C(p)} \int_{\Omega} v(q) f(p,q) dq, \quad (11)$$

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

As the weighting function Gaussian function is used, it is shown in (12):

$$f(p,q) = e^{-\frac{|B(q)-B(p)|^2}{h^2}}, \quad (12)$$

where *h* is the filter parameter (in general, for 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 (13):

$$C(p) = \int_{\Omega} f(p,q) dq \quad (13)$$

The result of clustering the source normalized image is represented in the Fig. 3:

Fig. 3. Result of clustering the source image after the stage of normalization

C. Object detection and classification

In order to avoid object classifying errors, the authors decided to normalize the training samples. This not only improves the pattern recognition quality due to unification, but also gives a possibility to provide less amount of training images. The result of Haar classifier work on car detection is represented in the Fig. 4:

Fig. 4. Result of Haar classifier work on car detection

IV. EXPERIMENT IN CASE OF SEMANTIC INTEGRITY VIOLATION

Image defects can be caused because of camera problems, bad weather conditions, compression operations, etc. To assess the effectiveness of the mentioned sequence of image analysis actions, the authors blurred the source image and added there more noise. The result of only Haar classifier implementation with non-normalized training samples is showed in the Fig. 5:

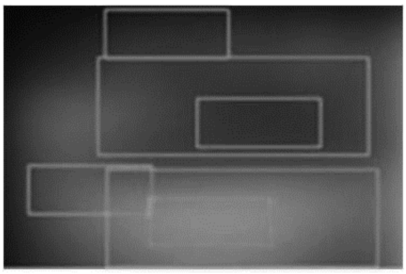

Fig. 5. Car detection using only Haar classifier with non-normalized training samples

Several additional recognized regions appeared that do not contain parts of the car. However, through consistent implementation of normalization, clustering based on particle swarm optimization, Haar cascade classifier with normalized training samples it became possible to eliminate emerged visual information semantic integrity violation. The recognition result is represented in the Fig. 6:

Fig. 6. Result of pattern recognition through the implementation of the image analysis algorithm proposed by the authors (projected onto the original image)

V. CONCLUSION

Three basic stages of pattern recognition were enumerated, and the authors found effective ways to preserve visual information semantic integrity respectively to each stage: in the beginning – image normalization, then clustering based on particle swarm optimization, finally – using of Haar cascade classifier with normalized training samples. The proposed sequence of methods gave the encouraging results, and in prospect the authors plan to implement found image analysis algorithm in practice, namely in the field of automated vehicle management in conditions of real road traffic.

REFERENCES

[1] Dunjko and H.J. Briegel, "Machine learning & artificial intelligence in the quantum domain: a review of recent progress", Reports on Progress in Physics., in press.

[2] David A. Forsyth. J. Ponce *Computer Vision: A Modern Approach*. Saddle, River, Nj, U, A: Prentice Hall, 2003.

[3] U.B. Neto and E.R. Dougherty, *Error estimation for pattern recognition.* New York: Wiley-IEEE Press, 2015.

[4] M. V. Akinin, A. V. Akinina, A. V. Sokolov and A. S. Tarasov, "Application of EM algorithm in problems of pattern recognition on satellite images", *2017 6th Mediterranean Conference on Embedded Computing (MECO).,* Jul. 2017, pp. 1-4.

[5] I. Jovanov and M. Pajic, "Sporadic data integrity for secure state estimation", *2017 IEEE 56th Annual Conference on Decision and Control (CDC).*, Dec. 2017, pp. 163-169.

[6] P. Santra, A. Roy, and K. Majumder, "A Comparative Analysis of Cloud Forensic Techniques in IaaS", *Advances in Computer and Computational Sciences. Advances in Intelligent Systems and Computing.,* vol. 554, Sept. 2017, pp. 207-215.

[7] S. Singh and M. Sinha, "Pattern recognition based on specific weights", *International Journal of Applied Pattern Recognition.,* vol. 5, Jan. 2018, pp. 1-10.

[8] B. Ganeshan, S. Abaleke, R. C. D. Young, C.R. Chatwin, and K. Miles. "Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage", *Cancer imaging: the official publication of the International Cancer Imaging Society.,* Jul. 2010, pp. 137-143.

[9] H.S. Israfilov, "Research of methods for binarization of images", *Herald of Science and Education.,* Jun. 2017, pp. 43-50.

[10] A. Aly, S. B. Deris, and N. Zaki, "Research review for digital image segmentation techniques", *International Journal of Computer Science & Information Technology (IJCSIT).,* vol. 3, Oct. 2011, pp. 99-105.

[11] Y. Xiang, A.C.S. Chung, and J. Ye, "A new active contour method based on elastic interaction", *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).*, vol. 1, Jul. 2005, pp. 452-457.

[12] N. Li, M. Liu, and Y. Li, "Image Segmentation Algorithm using Watershed Transform and Level Set Method", *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, Honolulu, H..*, Jun. 2007, pp. 613-616.

[13] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks", *International Journal of Multimedia Information Retrieval.*, vol. 6, Nov. 2017, pp. 1-7.

[14] Erwin, M. Fachrurrozi, A. Fiqih, B.R. Saputra, R. Algani, and A. Primanita, "Content based image retrieval for multi-objects fruits recognition using k-means and k-nearest neighbor", *2017 International Conference on Data and Software Engineering (ICoDSE).*, Nov. 2017, pp. 1-6.

[15] R. Ke, Z. Li, J. Tang, Z. Pan, and Y. Wang, "Real-Time Traffic Flow Parameter Estimation From UAV Video Based on Ensemble Classifier and Optical Flow", *IEEE Transactions on Intelligent Transportation Systems.*, vol. PP, Mar. 2018, pp. 1-11.

[16] A. P. Karpenko and E. Y. Seliverstov, "Overview of the particle swarm methods for the global optimization problem", *Science and Education: a scientific edition of the Bauman Moscow State Technical University.*, vol. 3, Mar. 2009, pp. 1-26.

[17] I. I. Viksnin, A. L. Drannik, R. A. Iureva, and I. I. Komarov, "Flocking factors' assessment in case of destructive impact on swarm robotic systems", *2016 18th Conference of Open Innovations Association and Seminar on Information Security and Protection of Information Technology (FRUCT-ISPIT).*, Sept. 2016, pp. 357-363.

[18] K. Leng, "An improved non-local means algorithm for image denoising", *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP).,* Dec. 2017, pp. 149-153.