

MSC 65K10, 65C35, 68U10

Research Article

RAPID - A model of fast eye pupil registration and tracking by a modified metaheuristic differential evolution method based on the Verhulst-Pearl equation

Y. V. Grushko

Vitus Bering Kamchatka State University, 683032,
Petropavlovsk-Kamchatskiy, Pogranichnaya str., 4, Russia
E-mail: neuralpill@gmail.com

This paper proposes a model of fast registration and pupil tracking - "RAPID", for devices with limited computing resource (weak personal computers, smartphones, embedded systems based on ARM architecture) in order to reduce the cost of technology for individual use by people with disabilities and medical institutions. The model is based on the idea of representing the process of video oculo-graphy as a multidimensional global optimization problem and its solution by the metaheuristic method of differential evolution. The optimization problem (objective function) is formalized as a search for the region that approximates the pupil in the three-dimensional parameter space most completely - the position and approximate size of the pupil. For the considered optimization problem we propose a modification of differential evolution method based on the process of formation of genetic isolations of population of solutions in the neighborhood of all local and global extremums of the target function followed by growth of the most adapted isolation (near the global extremum) and degeneration of others according to the differential Verhulst-Pearl equation. This behavior makes the search algorithm less "greedy" and makes it possible to correctly extract the pupil from the full frame. The developed tracking model can be used in the development of software packages in the task of augmentative communication for patients with lateral amyotrophic sclerosis or diplegia syndromes, on non-specialized devices, as well as in ophthalmological complexes and infrared-pupillometers.

Key words: videooculography, differential evolution, multivariate global optimization, region of interest, Hough transform, Verhulst-Pearl model.

 DOI: 10.26117/2079-6641-2022-38-1-84-105

Original article submitted: 15.02.2022

Revision submitted: 01.03.2022

For citation. Grushko Y. V. RAPID - A model of fast eye pupil registration and tracking by a modified metaheuristic differential evolution method based on the Verhulst-Pearl equation. *Vestnik KRAUNC. Fiz.-mat. nauki.* 2022, **38**: 1, 84-105.  DOI: 10.26117/2079-6641-2022-38-1-84-105

The content is published under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/deed.ru>)

© Grushko Y. V., 2022

Funding. Scientific research work of Vitus Bering Kamchatka State University, № AAAA-A19-119072290002-9

Introduction

Videoculography, the process of recording eye movements, position and size of the pupil, is one of the most important areas of theoretical and practical research, including medical research in ophthalmology, monitoring of photoreaction in neuroresuscitation through pupillometry [1], express diagnostics of drug, alcohol and industrial intoxication [2], and in biometric identification systems by iris (where the first step is pupil search in the frame).

The most actual task is to develop accessible hardware-software complexes of alternative augmentative communication (AAC) using VOG for medical institutions, rehabilitation centers, people with motor neuron diseases and other central nervous system (CNS) diseases, such as ICD-10 G12 amyotrophic lateral sclerosis (ALS), G 80.1 diplegia, G12.1 spinal muscular atrophy, G83 locked-in syndrome and others.

Patients with the above syndromes have complete loss of speech (aphasia), paralysis of the bulbar and facial muscles, tetraplegia (loss of limb function), while consciousness is completely intact. At the same time, it is still possible to communicate with patients through eye movement.

Thus, as of 2016, a major study was completed by the Lancet journal, showing that there were ~330,918 patients with ALS alone as of 2016, and a planned increase in the incidence per 100,000 population by 4 - 5% over the past 26 years [3]. The prevalence of diplegia in Russia averages from 1.5 - 4 per 1000 newborns [4], not to mention other CNS diseases, which makes the development of tracking methods and available AAC systems, using VOG, an actual task.

The stage of pupil registration and tracking in the existing algorithms takes considerable time and computing resources, requires high-performance specialized equipment (e.g., Tobii Dynavox), which makes this technology unavailable for most medical institutions and individual use by patients.

Thus, the existing algorithms used in VOG can be divided into several classes.

The algorithms based on threshold segmentation and mathematical morphology are not very precise, but they are computationally simple and are based on the fact that the pupil becomes the darkest area in the frame when infrared (IR) illumination is placed parallel to the optical axis of a video oculo-graph (used in Neuroptics DP-2000, NPi pupillometers). The frame is binarized. As a result the largest coherent set of points is formed, the brightness of which is lower than the threshold [5]. Superfluous objects left after segmentation: eyelashes, dark regions on the cornea are removed by erosion operators (only if the eyelashes clusters do not merge with the pupil into one coherent component, and they are smaller than the pupil), and the remaining pupil area is built up by morphological dilatation with the same number of iterations and kernel. In case the glare hits the pupil boundary, the minimal convex envelopes are calculated (e.g., by the Jarvis method), then the center of mass of the contour is found [6]. In [7] adaptive binarization based on percentile function study is used. No less interesting techniques are based on the application of recursive erosion [8] and the morphological generalized distance transform. These approaches do not take into account the peculiarities of visual

apparatus functioning - processes of miosis and mydriasis of the pupil during infrared illumination of the eye, which significantly reduces their accuracy.

Methods based on boundary-step image model research: Circle Hough Transform (CHT) and modifications [9], RANSAC, Starburst [10], ElSe [11], Swirski [12]. Intuitive adaptive method based on edge segment gradient entropy study [13]. The class has good accuracy, so the CHT is robust to the loss of part of the pupil boundary. The search takes place in three-dimensional space - the center and radius of the pupil or set of its boundary points. The disadvantage of this class is high computational complexity. The method based on the construction of the optimal circular path [8] needs to know the initial point, which should be inside the pupil's area of interest.

Methods based on spatial convolution and correlation [12] [14] have a great computational complexity, but have the property of searching by a certain pattern or part of the image, which gives good accuracy and can be applied to search both face, eye, and pupil region. Such methods in most cases do not require the use of IR illumination of the eye, which is a significant advantage without the use of specialized equipment.

Using convolutional (Tiny Yolo v3) [6] or generative-adversarial neural networks (U-net, Seg-net) in video ophthalmography task [15] we can achieve high accuracy, taking into account individual features of a patient's visual system (for example, in case of iris coloboma), but high computational costs and we need to develop high-quality training and test samples for training neural networks and work of a specialized expert in machine learning, which will increase the cost of using VOG.

The problem of accelerating, and increasing the accuracy of pupil tracking and pupilometry methods in real time in conditions of limited computing resources (weak PCs, smartphones, embedded systems based on ARM architecture) to reduce the cost of technology for individual use by people with disabilities and medical institutions and rehabilitation centers is solved. This paper proposes a model of RAPID (Rapid Pupil Intelligence Detector) pupil registration and tracking. The model is based on the idea of representing the process of VOG as a multidimensional global optimization problem and its solution by the metaheuristic method of differential evolution. The optimization problem (target function) is formalized as a search for the region that most completely approximates the pupil in the three-dimensional parameter space - position and approximate size of the pupil. Consider in detail the stages of the model.

Formalizing the problem of finding a region of interest

A region of interest (ROI) is a segmented area of the frame or a set of pixels representing a significant object, in our case the pupil.

Real conditions of pupil tracking are often far from the ideal ones described in similar works. Thus, the pupil is not always the darkest and biggest area in the frame (Fig. 1a) - there may be large darkened areas on the skin epithelium and hair in case of insufficient IR-illumination of the eye. Presence of thickened eyelashes over the eye can overlap part of the pupil or merge into a single component. Bright glare on the sclera reflecting IR radiation may be present.

Correctly found pupil ROI will allow: 1) reduce the computational cost of the algorithm (including existing ones) - the computational complexity will decrease in proportion to the ratio of the frame size to the ROI size; 2) correctly determine the threshold value for adaptive segmentation and selection of connectivity components, "excluding" brightness ranges not belonging to the pupil (no need to analyze the entire frame histogram or use complex adaptive techniques); 3) exclude "noise" border points of areas not related to the pupil, which is relevant when using methods of investigating; 4) successfully use algorithms [8] for which it is necessary or desirable (Starburst, AIPF, [16]) to know the starting point in the pupil area of interest. 5) radically speed up the Hough transform by reducing the radius search area given the ROI parameters.

Since a healthy pupil is a compact elliptical/round area and contains pixels with lower intensity than the surrounding iris, and the approximate ratio of pupil size to iris size is 1/3 (mean iris size — 12 mm, pupil size — 3.25-5 mm), pattern search will help to take into account all the parameters simultaneously.

When developing the algorithm we also had to take into account the physiological features of the human visual apparatus: reactivity, miosis - pupil contraction (not more than 1.1 mm), as a reaction to light and mydriasis - its dilation (up to 8 mm).

Let a three-channel frame $F(u)$ be given, and R, G, B — red, green and blue channels of the frame respectively (matrices), then we will use the weighted method (luminosity) $f(u) = 0.299R(u) + 0.587G(u) + 0.114B(u)$ to translate $F(u)$ into grayscale $f(u)$.

To determine the parameters of the region of interest of the pupil $f_{roi}(u)$ ($u = (u_1, u_2)$ — a point in two-dimensional space in an image) located in the $f(u)$ of size $N \times M$, it is necessary to create a parametric cube $\bar{f}(u)$, in which the points of the pupil center and its size will be clearly and unambiguously distinguished from the background of other points, for example, have a maximum value (Fig. 1).

The target function $\bar{f}(\tau)$ can be explained in terms of convolution with isotropic Haar wavelets. Let a wavelet w , described by a point τ (τ_1, τ_2 — the approximate center of the pupil, τ_3 — the size), which consists of two regions in the ratio $1/k$ — the outer light boundary $w' = \{\kappa | \kappa = 1, \kappa \in w\}$ and the inner dark region $w'' = \{\kappa | \kappa = -1, \kappa \in w\}$, by default, given the above features of the visual system $k = 3$.

Since the values at points of w belong to only two levels i.e. $w : \mathbb{N}_0 \rightarrow \{-1, 1\}$, the convolution process can be greatly accelerated using the integral matrix. Let us form an integral matrix $\delta_{int}(u)$ from $f(u)$. Each element of $\delta_{int}(u)$ contains the sum of pixel intensities up to the point (u_1, u_2) , i.e. $\delta_{int}(u) = \sum_{i=0, j=0}^{i \leq u_1, j \leq u_2} f(u)$. The integral matrix quickly

calculates the sums of point intensities in arbitrary rectangular regions (Regardless of the size of Haar-sign w , 3 arithmetic operations are performed to calculate convolution).

For correct calculation of the target function in three-dimensional space, it is necessary to replicate beforehand (before $\delta_{int}(u)$) the outer indentations of the image $f(u)$ with dimensions equal to $\sigma = R_{max}/2$, where R_{max} — the maximal allowable area of the pupil search, and R_{min} — minimal. Replication also makes it possible to calculate $\delta_{int}(u)$ only once for the whole target function, which significantly reduces the computational complexity of the algorithm.

Then the integral value of any rectangular wavelet domain \bar{R} can be calculated in a similar way:

$$\mathbf{d} = \left(\tau_1 - \frac{\tau_3}{2}, \tau_2 - \frac{\tau_3}{2}\right), \mathbf{p} = (\mathbf{d}_1 + \tau_3, \mathbf{d}_2), \mathbf{q} = (\mathbf{d}_1, \mathbf{d}_2 + \tau_3), \mathbf{b} = (\mathbf{d}_1 + \tau_3, \mathbf{d}_2 + \tau_3),$$

$$\ell_{\bar{R}} = \delta_{\text{int}}(\mathbf{d}) + \delta_{\text{int}}(\mathbf{b}) - \delta_{\text{int}}(\mathbf{p}) - \delta_{\text{int}}(\mathbf{q}). \quad (1)$$

Calculating (1) over the region $w' = \{\kappa | \kappa = 1, \kappa \in w\}$ as $\ell_{w'}$ and the region $w'' = \{\kappa | \kappa = -1, \kappa \in w\}$ as $\ell_{w''}$, we formalize the problem of finding the region $f_{\text{roi}}(u)$ that most completely approximates the pupil given by the parameters τ^* as:

$$\tau^* : \bar{f}(\tau^*) = \max_{\tau} \bar{f}(\tau), \quad \bar{f}(\tau) = k^2 \left(\frac{\ell_{w'}}{(k^2 - 1) \tau_3^2} - \frac{\ell_{w''}}{\tau_3^2} \right). \quad (2)$$

Fig. 1 shows the projections of the parametric cube along 3 planes in the vicinity of the maximum, as well as a three-dimensional model of the target function. All images were prepared in the ImageJ 1.53 image analysis software package and matplotlib library.

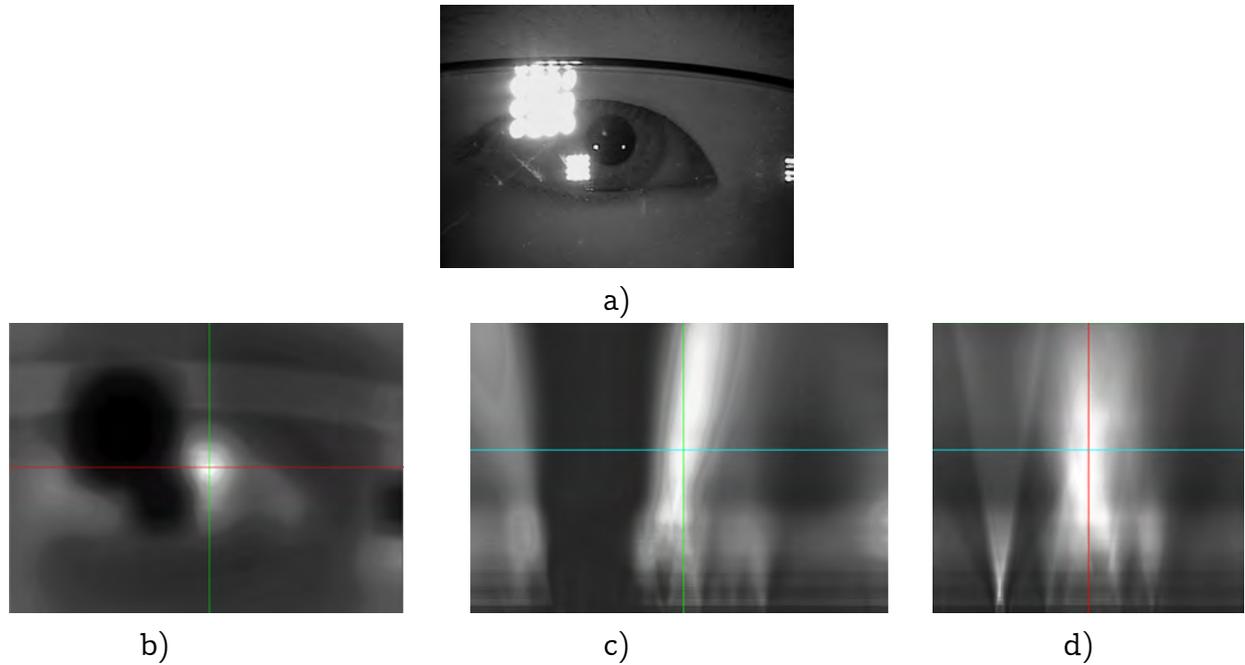


Fig. 1. Projections of the problem in question (parametric cube) along the planes (in extremum point): red — XZ, blue — XY, green — YZ. a) original frame; b) YX projection; c) XZ projection; d) YZ projection;

To solve the problem, we analyzed the target function (Table 1), on eye images, taken from the CASIA-IrisV4 database [17] of the Institute of Automation of the Chinese Academy of Sciences and prepared in the KamGU laboratory of student eye images. Various metaheuristic methods for global optimization were studied and tested: differential evolution, annealing simulation method, genetic algorithms, and swarm intelligence method. The works [18, 19] in which a comparative analysis of the above methods was given, the multidimensionality of the problem and the possibility of paralleling the method were taken into account. As a result, the best method for solving problem (2) - the method of differential evolution (DE) was chosen.

Table 1

| Target function (2) characteristic | |
|---|--|
| Characteristic | Interpretation in the problem in question |
| Nonlinear, Multidimensional, Non-differentiable Discrete | The relationship between the input $f(u)$ and the result $\bar{f}(\tau)$, cannot be expressed in linear form (changes in lighting, size and position of the pupil, noise in the frame, etc.). |
| Multimodal | Represented by a set of pixels (points) with a finite set of values. |
| Small space of acceptable solutions | A large number of local extremes are formed in the coagulation process: in the areas of eyelash thickening, shadows on the skin epithelium and iris. |
| | The set of points belonging to the extremum (the neighborhood of the pupil) is very small compared to the size of the original function $f(u)$. |

Searching for a region of interest by differential evolution

The method of differential evolution - a stochastic method of multivariate global optimization that models the processes of mutation, crossover and selection of individuals in a population [20].

In the process of investigation of the method it was possible to reveal the main advantages, due to which the method can be used in the problem under study (2): speed; robustness; reliability in the study of non-linear, multimodal functions, without knowledge of the search space (the target function depends on the input image).

Let it be required to solve (2), the initial population of solutions consists of n individuals, $n \geq 3$, and the search space is 3-dimensional ($D = 3$). Then the population t can be described by the set $\zeta(t) = \{\tau_1, \tau_2, \dots, \tau_n\}$, where τ_i is a possible solution characterized by the set of genes $\tau_i = (\tau_{i,1}^t, \tau_{i,2}^t, \tau_{i,3}^t)$.

The DE algorithm at each iteration involves three evolutionary processes: mutation, crossingover, and selection of the best individual. In mutation, for each target τ_i^t из $\zeta(t)$, 2 random individuals are selected τ_{r1}^t, τ_{r2}^t , such that $\tau_{r1}^t \neq \tau_{r2}^t$. Using the target and two random vectors, a new mutation vector v_i^t .

For practical eyetracking applications, the method must find a solution with probability close to 1. In the case of DE, the probability of convergence to the global maximum increases with increasing n . On the other hand population increase slows down the convergence speed and consequently region finding, that can not be allowed for devices with weak computational resource (weak PCs, microprocessors architectures).

To get high probability of finding true extremum on multimodal function keeping acceptable convergence, new models of mutation, crossingover and selection are proposed. Let us denote the proposed modification as Strategy for Finding the Most Adapted Genetic Isolation (SFAGi) and formally set it as system (3) using differential equation of the restricted population growth model - Verhulst-Pearl [21]:

$$\begin{cases} v_i^t = \left(1 - \frac{P(t)}{n}\right) \tau_i^t + \frac{P(t)\tau_{\text{best}}^t}{n} + \xi(\tau_{r1}^t - \tau_{r2}^t) \\ \frac{dP(t)}{dt} = \gamma P(t) \left(1 - \frac{P(t)}{n}\right), P(t_0) = P_0, \end{cases} \quad (3)$$

where P – the isolation size at time t , P_0 – the initial best isolation size. The analytical solution of the equation from (3) is found by separating the variables and integrating:

$$\int \frac{dP(t)}{P(1 - P(t)/n)} = \int \gamma dt$$

Because left side can be presented as

$$\frac{1}{P(1 - P/n)} = \frac{n}{P(n - P)} = \frac{1}{P} + \frac{1}{n - P}$$

Hence

$$\begin{aligned} \int \frac{dP}{P} + \int \frac{dP}{n - P} &= \int \gamma dt \\ \ln|P| - \ln|n - P| &= \gamma t + C \\ \ln \left| \frac{n - P}{P} \right| &= -\gamma t - C \\ \frac{n - P}{P} &= A e^{-\gamma t} \quad (A = \pm e^{-C}) \end{aligned}$$

From here we get:

$$P(t) = \frac{n}{1 + \left(\frac{n - P_0}{P_0}\right) e^{-\gamma t}}. \quad (4)$$

Then the final view of the mutation model used in the video oculoigraphy software package, taking into account (3) and (4), is written as follows:

$$\begin{aligned} v_i^t &= \left(1 - \frac{1}{1 + e^{-\gamma(t-\eta)}}\right) \tau_i^t + \frac{\tau_{\text{best}}^t}{1 + e^{-\gamma(t-\eta)}} + \xi(\tau_{r1}^t - \tau_{r2}^t) = \\ &= \frac{\tau_{\text{best}}^t + \tau_i^t e^{\gamma(\eta-t)}}{1 + e^{\gamma(\eta-t)}} + \xi(\tau_{r1}^t - \tau_{r2}^t), \quad (5) \end{aligned}$$

where η – special included coefficient which represents the evolution time of each isolate, γ – the specific growth rate of the best genetic isolate (Malthusian parameter), τ_{best}^t – the best found position of the whole population during t iterations, ξ – the mutation strength, the amplitude of perturbations introduced into the gene pool.

In the proposed model, because of its stochastic nature (random distribution of individuals over the decision space), the size of the best isolation is specified as a transfer coefficient in the range $[0; 1]$, where 1 is the maximum achieved best isolation capacity equal to n , and 0 is the complete absence (or few individuals) of individuals in best isolation.

The first stage of the mutation model (5), the exploratory search, is responsible for the appearance of genetic isolations depending on the landscape of the function, i.e. isolations are formed around all extremes of the function (the so-called landscape of adaptability), both local and global. Over time η , each of the isolations dynamically models the features of the function landscape, adjusting the distribution of the "internal" entropy source $\xi(\tau_{r1}^t - \tau_{r2}^t)$, to them, in the process finding the best solution for each of the isolations, which eliminates greediness of the algorithm and convergence to local solutions.

The second step (5) is responsible for the dynamics of growth γ of the most fitted genetic isolation and degeneration of others. The isolation that has the highest value of the target function begins to grow over time, while the other, less adapted ones, degenerate. Increasing the population of solutions in the most adapted isolation gives a more detailed search across the landscape, in the neighborhood of $\max_{\tau} \bar{f}(\tau)$, and hence a more accurate convergence, in contrast to the classical interpretation of DE, which is relevant when the initial population size is small.

The mutation model behaves nonlinearly (Fig. 2a, b, Fig. 4).

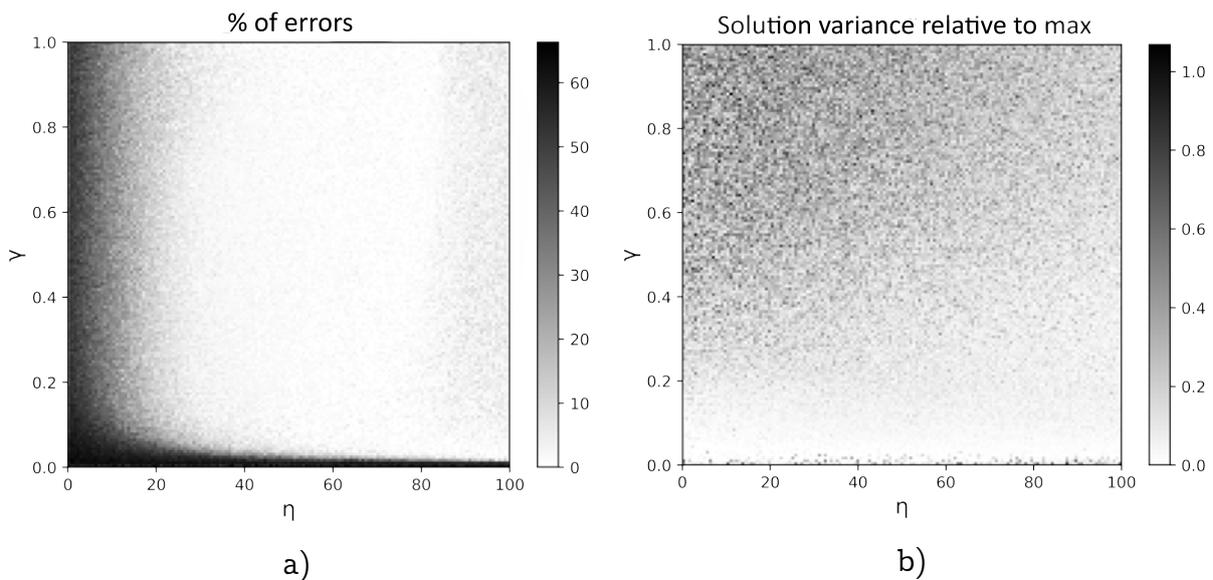


Fig. 2. Dependence of the accuracy of the solutions of the modified mutation model on input parameters: a) scatter of error rate (in %); b) scatter of candidate solutions relative to the maximum

During the experiments, it was noticed that increasing the evolution time of isolations η guarantees the most accurate finding of the region in which the global maximum is present, and decreasing the specific growth rate γ , adapted isolation, gives the most accurate approximation to $\max_{\tau} \bar{f}(\tau)$, i.e. the minimum dispersion of solutions relative to the max.

At the end of the mutation operation, perform a crossover operator on the target and mutant vector, resulting in a trial vector $m_{i,j}^t = (m_{i,1}^t \dots m_{i,D}^t)$, where each component is defined as:

$$m_{i,j}^t = \begin{cases} v_{i,j}^t, & \text{if } Rd \geq C_r, \quad i = 1..n, \quad j = 1..D \\ \tau_{i,j}^t, & \text{otherwise,} \end{cases} \quad (6)$$

where C_r — the probability of crossing, Rd — a function that returns a random number (uniform or Gaussian distribution) in the interval $[0; 1]$.

To select the individual with the best value of the target function, we use the following operator:

$$\tau_{i,j}^{t+1} = \begin{cases} m_{i,j}^t, & \text{if } \bar{f}(m_{i,j}^t) \geq \bar{f}(\tau_{i,j}^t) \\ \tau_{i,j}^t, & \text{otherwise.} \end{cases} \quad (7)$$

To improve the characteristics of the method, it was decided to use the polyhybrid crossover rule, instead of classic (6), (7). To do this, let us compose a set H , consisting of mutation vector v_i^t and recombinant individuals using the expression $\prod_{\mu=1}^D (\tau_{i,\mu}^t + v_{i,\mu}^t)$, where D is the number of measurements of the problem to be solved. Then, each single term of the fully expanded expression can be represented as a vector-individual of the H set (it is important to preserve the order of genes by the index μ).

The selection operator selects the first individual from H , satisfying the condition $\bar{f}(h) > \bar{f}(\tau_i^t)$, where $h \in H$, and completes its work on the individual i . This behavior of operators makes the method less "greedy" and more deterministic, and at each iteration guarantees fast convergence to an extremum.

As a stopping criterion, the following scheme is proposed for the pupil tracking problem:

$$\sum_{i=1}^n \bar{f}(\tau_i^t) - \bar{f}(\tau_{best}^t) \rightarrow 0. \quad (8)$$

Due to the fact that \bar{f} is nonlinear and varies from frame to frame, taking different values and positions of the global extremum, the proposed stopping scheme shows better results compared to a fixed number of iterations or reaching a certain value of the optimum (that we do not know).

The stopping criterion correlates directly with the time parameters η and the Malthusian parameter γ - the method is guaranteed to converge in finite time using (8).

When the stopping condition is reached, we will get the non-normalized coordinates τ_{best}^t of the pupil region $f_{roi}(u)$ — its center and size, and normalized coordinates: $j = 1..2 \quad \tau_j^* = \tau_{best,j}^t - \sigma \quad \tau_3^* = \tau_{best,3}^t$.

In Fig. 3 and Fig. 4 present a comparative analysis of optimization methods in solution accuracy and convergence plots. In this paper, the study was conducted for the two best-known schemes [20, 22], the recently proposed modification of the FSDE - Forced Differential Evolution Strategy described in [23], and the particle swarm optimization method.

The parameters used are: $\xi = 0.5$, $C_r = 0.5$, $\gamma = 0.27$, $\eta = 35$. Python 3.8, Matplotlib, NumPy were used for scientific calculations and plotting.

Test functions:

1. f1 — (2), $n = 150$, accuracy ± 3 px. $D = 3$, search area is $\tau_1 \in [0, 480]$, $\tau_2 \in [0, 640]$, $\tau_3 \in [30, 240]$;
2. f2 — Rastrigin, $n = 15$, accuracy ± 0.0001 , $D = 2$, search area is $\tau_i \in [-5.12, 5.12]$.

Test databases: CASIA:Lamp (key features: blink on the pupil, overlap of the pupil with eyelids and eyelashes), CASIA:Thousand (key features: blinks, glasses, lens flare, small pupil size, large dark areas from lack of IR illumination); images from the video sequences of the Vitus Bering KamGU lab.

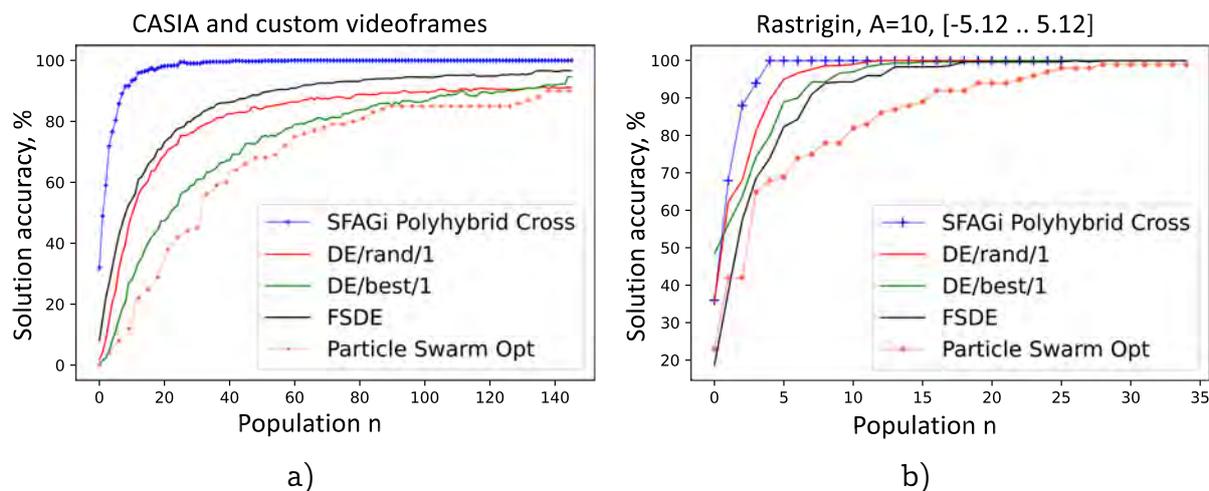


Fig. 3. Dependences of the % accuracy of the methods on the population size: a) selection of the optimal number of population for f1; b) optimal number of population for f2

Table 2 shows the results obtained by different numerical methods. Since algorithms are stochastic, and must be executed in real time (long-term), 3000 independent program runs were implemented for each method and function, in order to find the most accurate competitiveness.

The proposed modification of the differential evolution method allows to use a small initial population size (~ 35 -40 pcs. for (2) in detection mode, 5-11 pcs. for tracking mode) and to find the exact value of extremum in complex multimodal functions even in the case where the set of admissible solutions of the function is very small, as in the case of (2) or Easom function.

Features of the RAPID model in tracking mode

The steps described above can be used to process frames/medical images - let's call it detection mode. In tracking mode, the RAPID model is accelerated significantly without losing accuracy. For this purpose, some techniques described below are used. Let $f^i(u)$ be the current frame, i be the index of the current frame, and $\tau^{**} = \arg \max_{\tau} \bar{f}^{i-1}(\tau)$ be the parameters of the pupil region of interest $f_{\text{roi}}^{i-1}(u)$ of the previous frame.

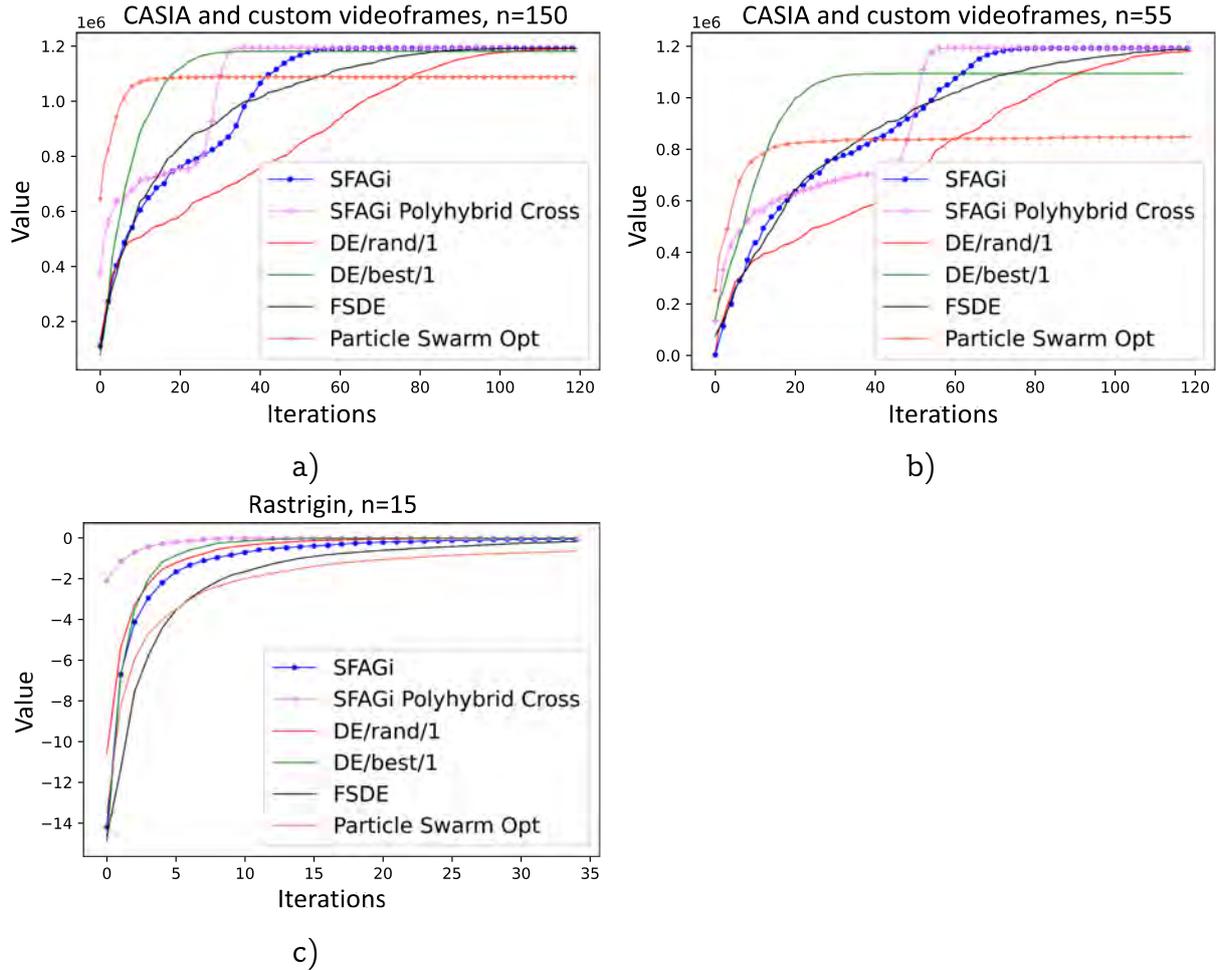


Fig. 4. Graph of convergence to a global extremum: a, b) for the considered problem (2), with different population size; c) for the Rastrigin test function

1. Calculation of replicated paddings and integral matrix δ_{int} is performed not for the whole input frame, but for the region f_{sub} , which is defined by the extremum point of the previous frame as $(\tau_1^{**}, \tau_2^{**}, R_{\text{max}})$ and its intersection (overlay) with $f^i(u)$ — safe region. Think of f_{sub} as a separate frame from which all unnecessary information has been removed (white square in Fig. 5 or 6b). So in fact we process not a frame of 640×480 size but a matrix of maximum possible size of 120×120 pixels containing only the pupil (without losing quality of the object of interest unlike making a pyramid of images).
2. Paddings size of f_{sub} can be left as constant $\sigma = \frac{R_{\text{max}}}{2}$ (this variant is more stable to changes of average illumination level of the frame) or adjust them depending on ROI size found on the previous frame i.e. $\sigma = \frac{\tau_3^{**}}{2} + \varepsilon$, where ε is an error leveling changes of illumination level (i.e. ROI size of current frame can be more or less than previous one).
3. The initial population is generated by the region f_{sub} . The projections are constrained by the region f_{sub} for each coordinate. This behavior allowed the method

Table 2

Comparative analysis of metaheuristic methods in three-dimensional parameter space

| Numerical Method | Dispersion by coordinates | | error rate (%) | | Iterations | |
|----------------------------------|---|----------------------------------|----------------|-------|------------|----|
| | f1 | f2 | f1 | f2 | f1 | f2 |
| SFAGi polyhybrid crossover | 0.0 | 1.39270408e-18 2.78978045e-18 | 0.0% | 0.0% | 31 | 11 |
| SFAGi | 0.0 | 2.81984127e-18 1.09386561e-17 | 0.0% | 0.0% | 62 | 22 |
| DE/rand/1 | 0.4888645 0.53323579 1.37237616 | 1.08026786e-16 7.65302185e-15 | 6.3% | 0.18% | 111 | 18 |
| DE/best/1 | 0.0 | 1.41516561e-18 2.58613739e-18 | 5.2% | 1.4% | 38 | 11 |
| FSDE | 0.20081199 0.11025189 1.28534807 | 8.33460518e-10 7.30447736e-11 | 3.7% | 2.1% | 92 | 40 |
| Particle Swarm Optimization | 0.60483871 0.70447451 17.28121748 | 2.22924456e-05 2.37325195e-05 | 9.3% | 7.3% | 18 | 61 |

to ignore local extremums located far from the center of the pupil (in fact, the region f_{sub} contains only the pupil), as a result the initial population size is additionally reduced to 5 - 11 pcs., and the accuracy is increased. As a super fast pseudorandom number generator, Xorshift128 (G. Marsaglia), a shift register generator of 128 bit version, is used.

4. Since the processing occurs in f_{sub} , it is possible to reduce the investigation time η of the target function to 10 - 20 and increase the Malthusian parameter γ to 0.5 - 0.7, which automatically reduced the number of convergence iterations of the algorithm to 30 - 35, without loss of accuracy.
5. After finding the best non-normalized point of the current frame $f^i(u)$ i.e. τ_{best} , it is necessary to perform normalization in the sense of the whole frame (taking into account the fact that the search was conducted in the matrix f_{sub}): $j = 1..2$ $\tau_j^* = \tau_j^{**} - \frac{R_{\text{max}}}{2} + (\tau_j^{**} - \sigma)$, $\tau_3^* = \tau_{\text{best},3}$, where τ^* - normalized coordinates of the current frame's ROI $f_{\text{roi}}(u)$ (a black square inside white square in Fig. 6).
6. The model allows for adaptive segmentation of the found $f_{\text{roi}}(u)$. Since RAPID allows to find the region that approximates pupil most completely - it occupies $\sim 1/2$ to $1/3$ of $f_{\text{roi}}(\tau)$, the following variant of segmentation with an adaptive threshold is proposed:

$$d = (\tau_{\text{best},1} - \frac{\tau_3^*}{2}, \tau_{\text{best},2} - \frac{\tau_3^*}{2}), p = (d_1 + \frac{\tau_3^*}{2}, d_2), q = (d_1, d_2 + \frac{\tau_3^*}{2}), b = (d_1 + \frac{\tau_3^*}{2}, d_2 + \frac{\tau_3^*}{2})$$

$$\ell_{\bar{R}} = \delta_{\text{int}}(d) + \delta_{\text{int}}(b) - \delta_{\text{int}}(p) - \delta_{\text{int}}(q),$$

$$[\cdot]: \mathbb{B} \rightarrow \{0, 1\}, \quad \beta(u) = \left[f_{\text{roi}}(u) \leq \frac{\ell_{\bar{R}}}{(\tau_3^*/s)^2} \right] \quad (9)$$

where $[\cdot]$ is the Iverson notation defined on the logical set, $\beta(u)$ is the segmented region, $s = 2$ or 3 .

Since the center of found $f_{\text{roi}}(u)$ is uniquely located in the neighborhood of the center of the pupil, it is possible to carry out a significant optimization of the algorithm - to perform segmentation and search for boundaries of the connectivity components simultaneously (excluding components not belonging to the pupil, but which would remain in the case of the standard threshold binarization process), using the breadth-first search rule and (9), modifying it to preserve the set of boundary points.

Let us represent $f_{\text{roi}}(u)$ as a planar undirected graph with the center in the point u^* , then the algorithm for segmentation and border extraction can be described using pseudocode (Listing 1).

Algorithm 1 Boundary extraction of connectivity components by adaptive threshold.

Require: u^* , where $u^* = (\tau_1^*, \tau_2^*)$, $f_{\text{roi}}(u)$

$B \leftarrow u^*$, ▷ Initially empty FIFO-queue

$E \leftarrow \{ \}$ ▷ initially empty set of pupil edge points

while $B \neq \emptyset$ **do**

$u \leftarrow B_{\text{head}}$

for $p \in R_8(u)$, where $R_8(u) = \{(u_1 + i, u_2 + j) \mid i, j = -1..1\}$ **do**

if $p \in f_{\text{roi}}(u)$ **and not** p^z **then**

if $\left[f_{\text{roi}}(p) \leq \frac{\ell_{\bar{R}}}{(\tau_3^*/s)^2} \right]$, where $[\cdot]: \mathbb{B} \rightarrow \{0, 1\}$ **then**

$B \leftarrow p$

else

$E \leftarrow p$

end if

$p \leftarrow p^z$ ▷ The point p is marked as passed during segmentation

end if

end for

end while

return E

A simplified block diagram of the pupil detection and tracking model RAPID based on the modified differential evolution method is shown in Fig. 5.

Fig. 6a shows the process of pupil registration in the $f(u)$, the dots show the evolution of the population distribution of the decision space from green to white (along the t -axis). The white clusters are the formed isolations in the neighborhood of all extremums (the largest clustering of points is located near the global extremum).

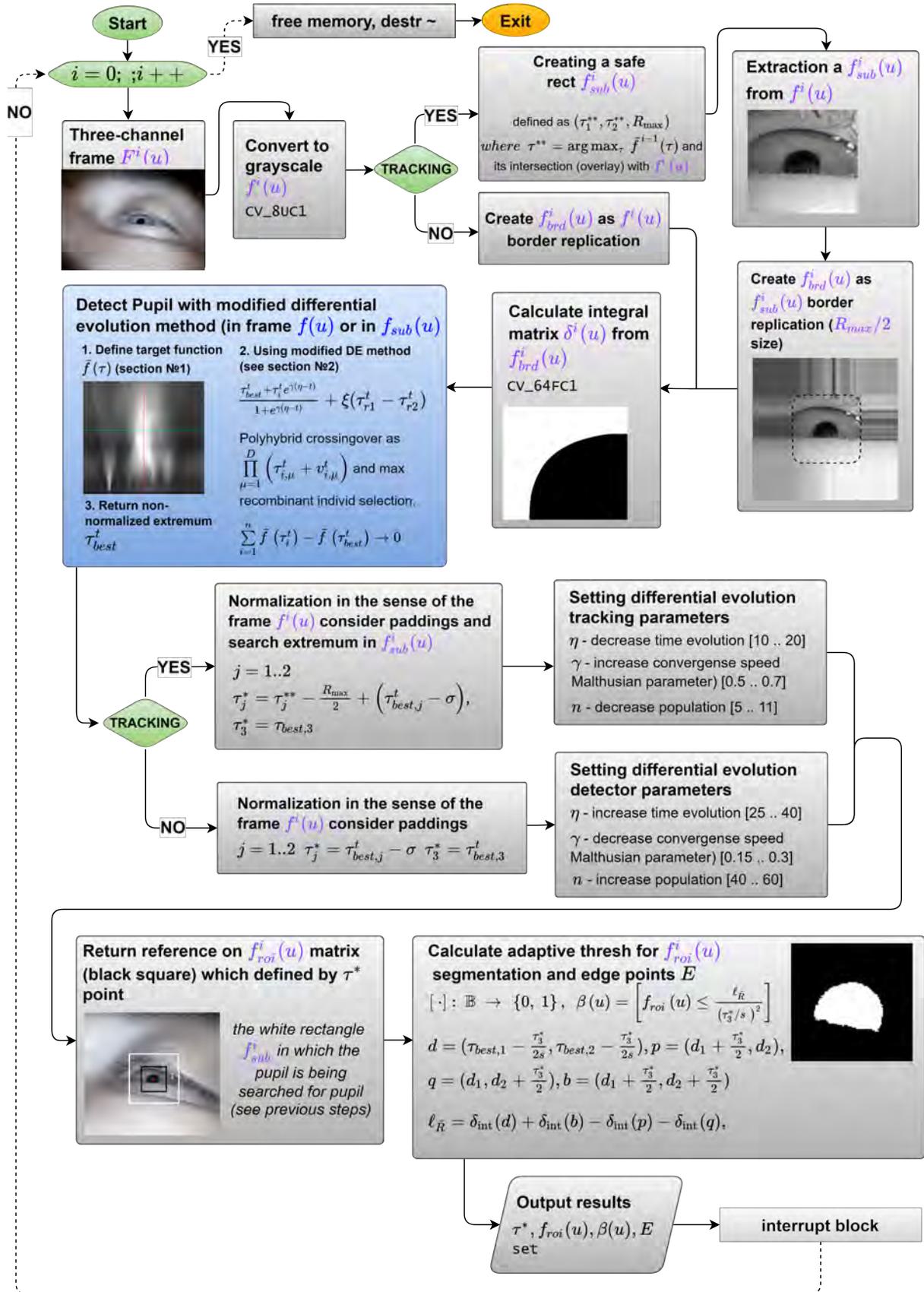


Fig. 5. Simplified block diagram of the RAPID model

Fig. 6 b,c shows the process of pupil tracking, the search for the pupil region of interest $f_{roi}(u)$ (black square) takes place in the $f_{sub}(u)$ region (white square), which significantly increases the accuracy and speed of tracking.

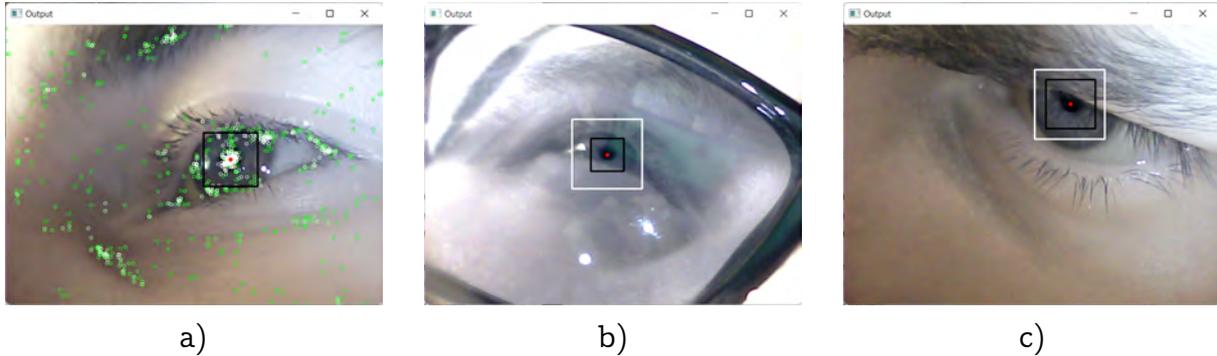


Fig. 6. An example of the RAPID model in some cases (black region - found pupil ROI): a) detection mode in a whole frame; b) pupil tracking in the sub-region when the user is wearing glasses c) tracking of a partially occluded pupil

Using the cross-platform framework Qt and C++ the library RAPID.h and the application for modeling (Fig. 7) and tuning parameters for the specific task and the future algorithm (pupillometry, communication, identification, etc.) were developed, it is possible to adjust the parameters of modified differential evolution, search area, pupil registration mode, as well as obtain the coordinates of pupil ROI, adaptive binarization threshold, estimate detection accuracy, time, number of iterations of the algorithm, etc.

Example of using the model in parametric pupil identification

To perform accurate parametric identification of the pupil, it is necessary to radically reduce the amount of processed information in the found region of interest $f_{roi}(u)$, performing segmentation by threshold using (9) and Listing 1.

If the optical axis of the eye (during VOG procedure), is parallel to the camera axis, the pupil component will be a dark circle (or part of it) whose boundary points can be described using the circle equation $r^2 = \sum_{i=1}^n (u_i^* - u_i)^2$. For parametric identification of the pupil in the $f_{roi}(u)$, we will use the Circular Hough Transform.

All the boundary points of the set E are bypassed. There is a voting procedure: each boundary point $\varepsilon \in E$ is assigned to a set of points of accumulative space A which can be described by a circle of radius r in the center of ε , the accumulative value in these points is incremented.

After the voting is finished, the point which receives the largest number of votes, i.e. $\arg \max A(u)$, will be responsible for the best position of the pupil, in $f_{roi}(u)$. The $f_{roi}(u)$ and its parameters τ^* allows to significantly reduce the time and area of search for the pupil radius, and to exclude most of the false circles. Thus the search area is narrowed to $r \in \left[\frac{\tau_3^*}{3}, \frac{\tau_3^*}{2} \right]$. In Fig. 8 shows the result of the algorithm. Table 5 shows a comparative analysis of Hough transform with preprocessing for the whole frame and when used together with the proposed model.

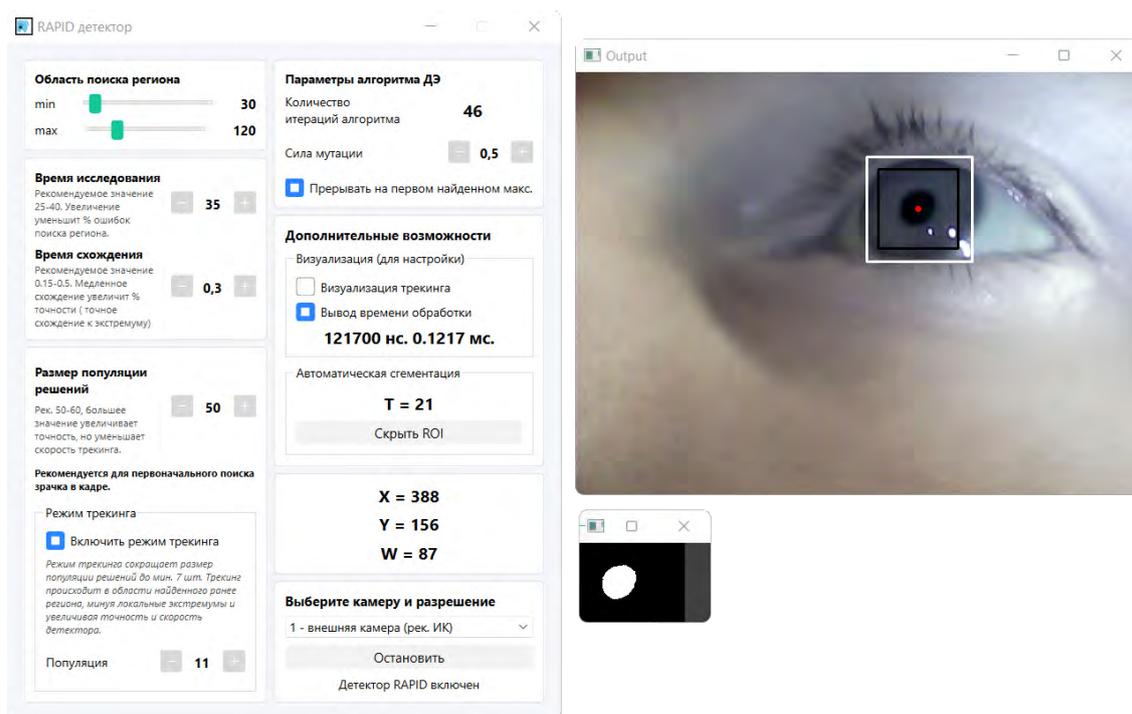


Fig. 7. Application for modeling and tuning of ROI search parameters for a specific task and video oculography algorithm

Experimental results

A numerical experiment on eye images was performed. Videoculograph was a webcam with set pixel resolution 640×480 , remote IR filter and built-in backlight for the eye - two SMD IR LEDs, with a peak wavelength of ~ 940 nm. (near-infrared), connected in parallel through resistors rated 120 Ohms, to the power supply of the camera. The choice of NIR is due to the highest safety according to Russian National Standard 62471-2013, in contrast to the far-IR, which can cause burns during prolonged exposure. Since the human visual system is not susceptible to IR radiation during video oculography there is no pupil miosis as a protective reaction of the retina. To reduce the harm of IR radiation on the retina, white LED illumination is used to reduce the diameter of the pupil.

The programming language was C++ in conjunction with the OpenCV 4.5.3 library, data structures were handled by C-style and dynamic arrays, the library `<Chrono.h>` was used to measure the execution time. Benchmark performance of the model on different architectures is shown in Table 3.

A comparative analysis was also performed in Table 4 with a widely known method for finding a region of interest - Haar Cascades, with a specially trained model for eye detection.

The proposed model allows you to approach the process of pupil registration step by step, first searching for the region that most fully approximates the pupil, removing unnecessary objects that interfere with the correct search, then locally applying the method based on the study of the boundary model of the image or brightness/ morpho-

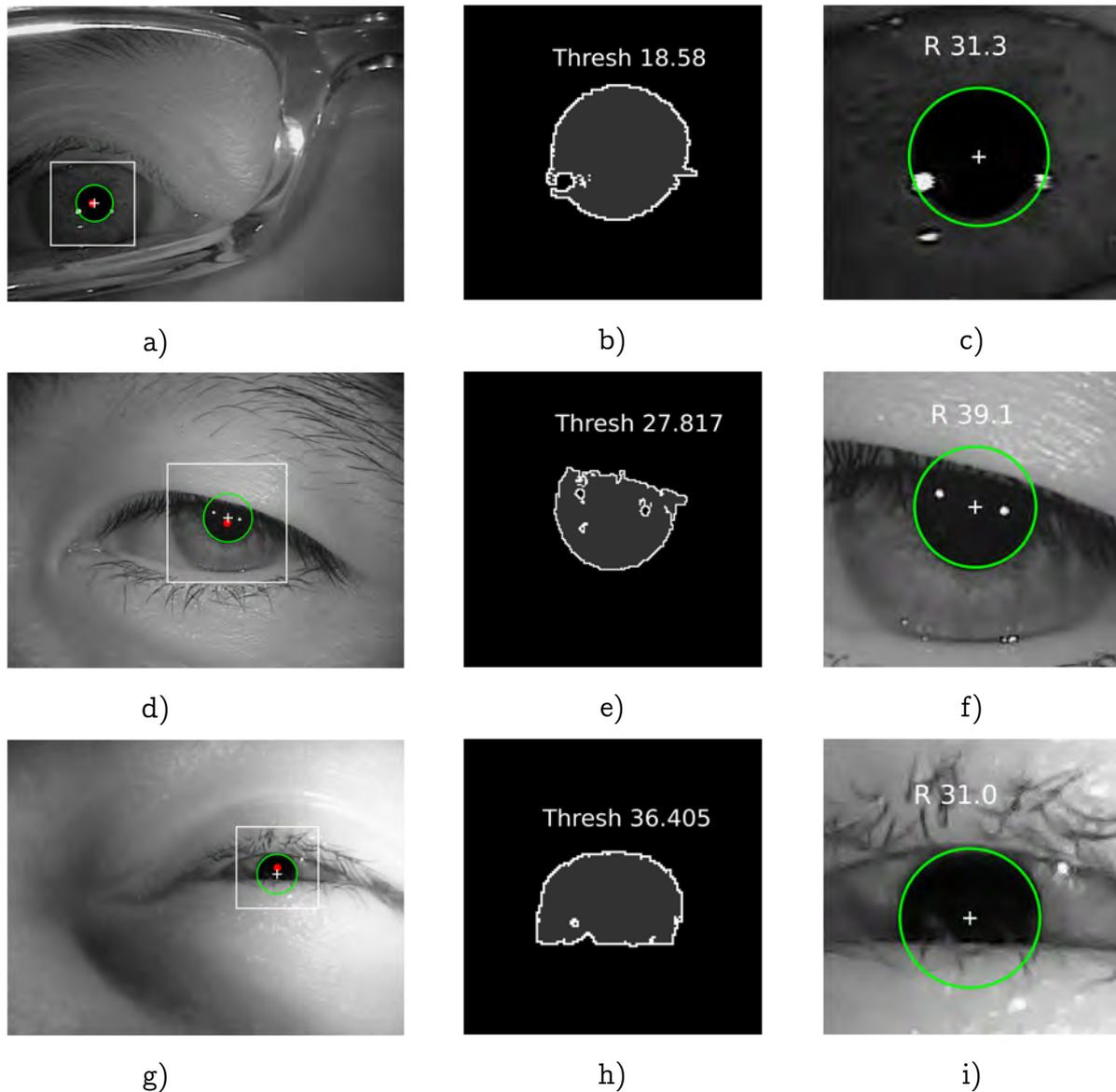


Fig. 8. An example of an algorithm using the RAPID model (detection mode), adaptive segmentation using (9), and CHT: a, d, g) the white square - the $f_{\text{roi}}(u)$, the red dot - center of $f_{\text{roi}}(u)$, the green circle - the pupil, and white cross - center; b, e, h) adaptively segmented $f_{\text{roi}}(u)$, white borders — E set; c, f, i) found pupil center and its size in px.

logical. An example of using the model with different techniques is presented in Table 5.

Conclusion

The proposed eye pupil tracking model, based on the search for a region that maximally approximates the pupil using the differential evolution method, was successfully applied in the problem of video oculography in order to reduce computational costs and increase accuracy, which will make it possible to apply algorithms on non-specialized devices and microcontroller architectures, and also reduce the cost of video oculography

Table 3

Processing time of one frame by RAPID model in tracking and detection mode on different computing architectures.

| | | | | |
|--------------------------------|--|---|---|--|
| Architecture | Intel(R) Core(TM) i7-9750H, 2.59 GHz, x64 bit | Inte(R) Core(TM) i7-3517U, 1.7 GHz, x64 bit | Broadcom BCM2837, ARM Cortex-A53 1200 MHz, , x32 bit RAM 1024 Mb | BCM2835, ARM1176JZF-S, 700 MHz, x32 bit RAM 512 Mb |
| Compiler, flags, vectorization | MSVC 2019, optimization flag /O2, c++14, (Visual Studio, Qt) vectorization support SSE2, AVX2. | | G++, optimization flag, /O2, c++14, vectorization support - NEON instructions | |
| Tracking mode n = 11 | 0,0898 ms 89800 ns | 0,4938 ms 493800 ns | 1,9181 ms 1918139 ns | 8,928845 ms 8928845 ns |
| Detection mode n = 50 | 0,8814 ms 881400 ns | 3,24 ms 3242600 ns | 20,49026 ms 20490260 ns | 79,033488 ms 79033488 ns |

Table 4

Comparison of eye region detection

| Using architecture - Broadcom BCM2837, ARM Cortex-A53, 1200 MHz, RAM 1024 Mb | Average time frame processing | Accuracy |
|--|-------------------------------|----------|
| RAPID model (tracking mode) | 1.91 ms. | 96.854% |
| RAPID model (detection mode) | 20,49 ms. | 93.490% |
| Haar-Cascade model | 64,85 ms. | 89.516% |

Table 5

Comparative analysis of some pupil registration methods

| Method and its characteristic (with /O2 optimization and NEON instructions) | Accuracy | Average time |
|--|----------|--|
| Median filter (5 x 5) + Canny + Hough transform (OpenCV) | 96.8 % | 50.59 ms. search radius area [30..50] |
| RAPID (tracking mode) + median + boundary search Listing 1 + Hough transform (OpenCV) | 98.7% | 11.09 ms. search radius area $\left[\frac{\tau_3^*/2}{3}, \frac{\tau_3^*/2}{2}\right]$ |
| Morphological Closing (remove blink) + Recursive erode (OpenCV) + find the most distant point in a frame | 88.5% | 370.50 ms. |
| RAPID (tracking mode) + recursive morphological erode (OpenCV) + find distant point in a ROI | 93.1% | 18.87 ms. |

technology, making it more accessible to people with disabilities and medical institutions.

Since the differential evolution method used in the RAPID belongs to the class of stochastic ones, it was necessary to reduce the percentage of false solutions as a result of descending to local optimum and reduce the probability of exact solution to unity. As a result, a modified numerical differential evolution method was proposed

based on the formation of genetic isolates in the vicinity of local and global extremums of the target function and the subsequent growth of the most adapted one - in the vicinity of the true extremum, according to the Verhulst-Pearl equation. The crossover and selection operators were revised, which made the model more deterministic, and allowed for guaranteed movement in the direction of the best value at each iteration. The stopping condition of the differential evolution method for the video oculography problem was determined, taking into account the peculiarities of the target function. The modification surpassed existing numerical methods in the problem (2).

In the future, it is planned to use the architecture of a multilayer perceptron in problem (2) to analyze the response of the integrated brightness of the pattern parts: the upper, lower and central parts, which will increase the accuracy of searching for the region of interest in complex images; to implement the developed method in the [6].

The proposed model can be used to detect different objects (letters, faces, figures etc.) by adjusting the search binary pattern (grid) for the target function.

Competing interests. The author declares that there are no conflicts of interest with respect to authorship and publication.

Contribution and responsibility. The author contributed to the writing of the article and is solely responsible for submitting the final version of the article to the press. The final version of the manuscript was approved by the author.

Acknowledgments. I express my gratitude to the supervisor, Doctor of Physical and Mathematical Sciences R.I. Parovik for a number of comments that contributed to the improvement of the presented work.

References

1. Oshorov A. V., Aleksandrova E. V., Muradyan K. R., Sosnovskaya O. Y., Sokolova E. Y., Savin I. A. Pupillometry as a method of monitoring photoreactivity in neuroresuscitation, *Voprosy neirohirurgii imeni N.N. Burdenko*, 2021. vol. 85, no. 3, pp. 117–123 (In Russian).
2. Kucalo A. L., Cimbali M. V., Homich D. S., Varenikov M. G., SHtejnberg N. V. Dynamic Pupillometry as a Screening Diagnostic Method for Industrial Toxicant Poisoning, *Medicina ekstremal'nyh situacij*, 2018. vol. 20, pp. 487–493, (In Russian).
3. Logroscino Giancarlo, Piccininni Marco, Marin Benoît, Nichols Emma, Abd-Allah Foad, Abdelalim Ahmed, Alahdab Fares, Asgedom Solomon, Awasthi Ashish, Chaiah Yazan, Daryani Ahmad, Do Phuc Huyen, Dubey Manisha, Elbaz Alexis, Eskandarieh Sharareh, Farhadi Farzaneh, Farzadfar Farshad, Fereshtehnejad Seyed-Mohammad, Fernandes Eduarda, Lorkowski Stefan. Global, regional, and national burden of motor neuron diseases 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016, *The Lancet Neurology*, 2018. vol. 17, pp. 1083–1097 (In Russian).
4. Tkachenko Ekaterina Sergeevna, Goleva Ol'ga Petrovna, SHCHerbakov Denis Viktorovich, Halikova Adel' Ravil'evna. Cerebral Palsy: State of the Study of the Problem (Review), *Mat' i ditya v Kuzbasse*, 2019. vol. 2, pp. 4–9 (In Russian).
5. Durna Y., Ari F. Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images, *Applied Sciences*, 2017. vol. 7, pp. 498.
6. Grushko Y. V. Hardware-software complex of augmentative communication system based on eyetracking technology, *Vestnik KRAUNC. Fiz.-Mat. Nauki*, 2019. vol. 27, no. 2, pp. 55–73 (In Russian).
7. Bonteanu P., Cracan A., Bonteanu G., Bozomitu R. A. Robust Pupil Detection Algorithm Based on a New Adaptive Thresholding Procedure / *IEEE International Conference on e-Health and Bioengineering EHB (2019)*, 2019, pp. 276.
8. Matveev I. A. Methods and algorithms for automatic processing of images of the iris of the eye. Dissertation for the degree of Doctor of Technical Sciences, 2014 (In Russian).
9. Alkuzaay M., Alshemmary E. Towards Accurate Pupil Detection Based on Morphology and Hough Transform, *Baghdad Science Journal*, 2020. vol. 17, no. 2, pp. 583–590.

10. Dongheng L., Winfield D., Parkhurst D. J. A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches / *Paper Presented at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2005)*, 3, 2005, pp. 79.
11. Fuhl W., Santini T., Kübler T., Kasneci E. ElSe: ellipse selection for robust pupil detection in real-world environments / *The Ninth Biennial ACM Symposium*, 2016, pp. 123–130.
12. Swirski. L. Bulling. A. Dodgson. N. Robust real-time pupil tracking in highly off-axis images / *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA)*, 2012, pp. 173–176.
13. Topal Cihan, ÇAKIR Halil, Akinlar Cuneyt. An Adaptive Algorithm for Precise Pupil Boundary Detection using Entropy of Contour Gradients, 2017.
14. Yang Z. Intelligent Evaluation of Strabismus in Videos Based on an Automated Cover Test, *Applied Sciences*, 2019, pp. 59.
15. Fuhl Wolfgang, Geisler David, Rosenstiel Wolfgang, Kasneci Enkelejda. The Applicability of Cycle GANs for Pupil and Eyelid Segmentation, *Data Generation and Image Refinement*, 2019, pp. 4406–4415.
16. Grushko Y. V., Parovik R. I. Fast Pupil Tracking based on the Study of a Boundary-stepped Image Model and Multidimensional Optimization Hook-Jives Method, *Informatika i automatizacija – Informatics and automation*, 2021. vol. 2, pp. 435–462 (In Russian).
17. Chinese Academy of Sciences Institute of Automation. Iris image database, version 4, 2021 <http://www.cbsr.ia.ac.cn/china/Iris20Databases20CH.asp>.
18. Kovalevich, A. A., Jakimov A. I., Albkeirat D. M. Study of stochastic optimization algorithms for use in simulation of systems, *Informacionnye tehnologii – Information technologies*, 2011. vol. 8, pp. 55–60 (In Russian).
19. Pupkov K. A., Feoktistov V. A. Algorithm «Differential evolution» for the problem of technical design, *Informacionnye tehnologii – Information technologies*, 2004. vol. 8, pp. 25–31 (In Russian).
20. R. Storn, K. Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization*, 1997. vol. 11, no. 4, pp. 341–359.
21. Tsoularis A. N., Wallace James. Analysis of Logistic Growth Models, *Mathematical biosciences*, 2002. vol. 179, pp. 21–55.
22. Jeyakumar Gurusamy, C. Shanmugavelayutham. Convergence Analysis of Differential Evolution Variants on Unconstrained Global Optimization Functions, *International Journal of Artificial Intelligence and Applications*, 2011.
23. Meera Ramadas, Ajith Abraham, Sushil Kumar. FSDE-Forced Strategy Differential Evolution used for data clustering, *Journal of King Saud University - Computer and Information Sciences*, 2019. vol. 31, no. 1, pp. 52–61.



Grushko Yuriy Vasilyevich ✉ – PhD student of the Fac. of Phys. & Math., Vitus Bering Kamchatka State University, Petropavlovsk-Kamchatskiy, Russia,  ORCID 0000-0002-3663-0018.

УДК 004.93, 004.94, 519.6

Научная статья

RAPID — модель быстрой регистрации и трекинга зрачка глаза с помощью модифицированного метаэвристического метода дифференциальной эволюции на основе уравнения Ферхюльста-Пирла

Ю. В. Грушко

Камчатский государственный университет имени Витуса Беринга,
683032, Петропавловск-Камчатский, ул. Пограничная, 4, Россия

E-mail: neuralpill@gmail.com

В работе предлагается модель быстрой регистрации и трекинга зрачка «RAPID» для устройств с ограниченным вычислительным ресурсом (слабые персональные компьютеры, смартфоны, встраиваемые системы на базе архитектуры ARM) с целью снижения стоимости технологии для индивидуального использования людьми с ограниченными возможностями и медицинскими учреждениями. В основу модели легла идея представления процесса видеоокулографии, как задачи многомерной глобальной оптимизации и ее решение метаэвристическим методом дифференциальной эволюции. Задача оптимизации (целевая функция) формализована, как поиск региона, наиболее полно аппроксимирующего зрачок в трёхмерном пространстве параметров – положение и приблизительный размер зрачка. Для рассматриваемой задачи оптимизации предложена модификация метода дифференциальной эволюции, в основе которого лежит процесс формирования генетических изоляций популяции решений в окрестностях всех локальных и глобальных экстремумов целевой функции, с последующим ростом наиболее приспособленной изоляции (рядом с глобальным экстремумом) и вырождением иных, в соответствии с дифференциальным уравнением Ферхюльста-Пирла. Данное поведение делает алгоритм поиска менее «жадным» и дает возможность корректно выделять зрачок из полного кадра. Разработанная модель трекинга может быть использована при разработке программных комплексов в задаче аугментативной коммуникации для пациентов с синдромами латерального-амиотрофического склероза или диплегии, на неспециализированных устройствах, а также в офтальмологических комплексах и инфракрасных-пупиллометрах.

Ключевые слова видеоокулография, дифференциальная эволюция, многомерная оптимизация, регион интереса, преобразование Хафа, модель Ферхюльста-Пирла.

 DOI: 10.26117/2079-6641-2022-38-1-84-105

Финансирование. НИР КамГУ им. Витуса Беринга, № АААА-А19-119072290002-9

Поступила в редакцию: 15.02.2022

В окончательном варианте: 01.03.2022

Для цитирования. Grushko Y.V. RAPID - A model of fast eye pupil registration and tracking by a modified metaheuristic differential evolution method based on the Verhulst-Pearl equation // *Вестник КРАУНЦ. Физ.-мат. науки.* 2022. Т. 38. № 1. С. 84-105.  DOI: 10.26117/2079-6641-2022-38-1-84-105

Конкурирующие интересы. Конфликтов интересов в отношении авторства и публикации нет.

Авторский вклад и ответственность. Автор участвовал в написании статьи и полностью несет ответственность за предоставление окончательной версии статьи в печать.

Контент публикуется на условиях лицензии Creative Commons Attribution 4.0 International (<https://creativecommons.org/licenses/by/4.0/deed.ru>)

© Грушко Ю. В., 2022



Грушко Юрий Васильевич  – аспирант физико-математического факультета, Камчатский государственный университет имени Витуса Беринга, Петропавловск-Камчатский, Россия,  ORCID 0000-0002-3663-0018.
