Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing

Роботизированные комплексы и системы. Технологии дистанционного зондирования и неразрушающего контроля

UDC 004.023
https://doi.org/10.32362/2500-316X-2023-11-1-18-30

RESEARCH ARTICLE

Continuous genetic algorithm for grasping an object of a priori unknown shape by a robotic manipulator

Andrey D. Voronkov @,
Sekou A.K. Diane

MIREA – Russian Technological University, Moscow, 119454 Russia
@ Corresponding author, e-mail: a.voronkov.rtu@yandex.ru

Abstract
Objectives. The problem of providing the interaction of a robotic manipulator with a priori unknown objects in a given workspace is of great interest both to the research community and many industries. By developing a solution to this problem, it will be possible to reduce the time taken for robots to adapt to new environments and objects therein. One of the primary stages of providing the interaction of the robotic manipulator with objects is the search for the target position of the robot gripper based on the onboard sensor subsystem, which can be carried out by a number of methods. Methods associated with machine learning and self-learning technologies may not be suitable for some applications (for example, during rescue operations) when it is necessary to quickly search for the target position of the gripper for an a priori unknown object, about which there is no relevant information in the robot database. Therefore, for this problem, heuristic approaches – for example, genetic algorithms – seem to be applicable. The objectives of this work are to implement a search based on a continuous genetic algorithm for the target position of the robot gripper including collision avoidance and study its performance under virtual simulation.
Methods. A heuristic search algorithm (continuous genetic algorithm) is used. The complex scene analysis algorithm uses classical image processing methods. In order to evaluate the effectiveness of the algorithm, virtual simulation is used.
Results. The possibility of using a continuous genetic algorithm is analyzed in the problem of grasping an object of an a priori unknown shape avoiding collisions with other objects of a static scene. A complex scene analysis algorithm and implementation of a continuous genetic algorithm are presented for finding the target position of the gripper of a Kuka LBR iiwa 7 R800 robotic control system with redundant kinematics. The results of an experimental virtual simulation of the obtained algorithm are presented.
Conclusions. The conducted research demonstrates the effectiveness of the continuous genetic algorithm in obtaining the target position of the gripper of the robotic manipulator under conditions when the static scene represents randomly located objects of various shapes.

Keywords: continuous genetic algorithm, grasping of objects of unknown shape, positioning of gripper, collision avoidance, robotic manipulator

• **Submitted:** 17.03.2022 • **Revised:** 12.04.2022 • **Accepted:** 26.10.2022

For citation: Voronkov A.D., Diane S.A.K. Continuous genetic algorithm for grasping an object of a priori unknown shape by a robotic manipulator. *Russ. Technol. J.* 2023;11(1):18–30. https://doi.org/10.32362/2500-316X-2023-11-1-18-30

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Непрерывный генетический алгоритм в задаче захвата манипуляционным роботом объекта априорно неизвестной формы

А.Д. Воронков @,
С.А.К. Диане

МИРЭА – Российский технологический университет, Москва, 119454 Россия
@ Автор для переписки, e-mail: a.voronkov.rtu@yandex.ru

Резюме
Цели. Задача взаимодействия манипуляционного робота с априорно неизвестными объектами рабочей области представляет большой интерес для научного сообщества и множества отраслей. Решение этой задачи позволит сократить время адаптации робота к новым средам и объектам в них. Один из главных этапов взаимодействия манипуляционного робота с объектами сцены – поиск целевого положения захватного устройства на основе бортовой сенсорной подсистемы – может быть осуществлен рядом методов. Методы, связанные с технологиями машинного обучения и самообучения, могут быть неподходящими для некоторых областей применения (например, во время аварийно-спасательных работ), когда требуется быстро осуществить поиск целевого положения захватного устройства для априорно неизвестного объекта, информации о котором нет в базе данных робота. Поэтому для этой задачи представляются применимыми эвристические подходы, например, генетический алгоритм. Целями работы являются реализация поиска целевого положения захватного устройства с избеганием столкновений на основе непрерывного генетического алгоритма и исследование его работоспособности в условиях виртуального моделирования.
Методы. Использован эвристический алгоритм поиска – непрерывный генетический алгоритм. В комплексном алгоритме анализа сцены использованы классические методы обработки изображения. Использовано виртуальное моделирование для оценки эффективности алгоритма.
Результаты. В работе рассмотрена возможность применения непрерывного генетического алгоритма в задаче захвата объекта априорно неизвестной формы с избеганием столкновений с другими объектами статической сцены. Представлен комплексный алгоритм анализа сцены и реализация непрерывного генетического алгоритма для решения задачи поиска целевого положения захватного устройства робота избыточной кинематики Kuka LBR iiwa 7 R800. Проведены эксперименты и приведены результаты виртуального моделирования полученного алгоритма.
Выводы. Проведенное исследование позволяет утверждать, что непрерывный генетический алгоритм достаточно эффективен в задачах поиска целевого положения захватного устройства манипуляционного робота при условиях, когда статическая сцена представляет собой хаотично расположенные объекты разной формы.

Ключевые слова: непрерывный генетический алгоритм, захват объектов неизвестной формы, позиционирование захватного устройства, избегание столкновений, манипуляционный робот

• **Поступила:** 17.03.2022 • **Доработана:** 12.04.2022 • **Принята к опубликованию:** 26.10.2022

Для цитирования: Воронков А.Д., Диане С.А.К. Непрерывный генетический алгоритм в задаче захвата манипуляционным роботом объекта априорно неизвестной формы. *Russ. Technol. J.* 2023;11(1):18–30. https://doi.org/10.32362/2500-316X-2023-11-1-18-30

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Topical tasks of modern robotic manipulators typically involve the indoor operation of robots. Under such conditions, a robot may collide with a large number of objects of a priori unknown shape, color, or texture. Images obtained using RGBD cameras constitute the source of a significant amount of initial data used for solving such robotics problems[1]. Among the tasks of manipulation robotics, for which a single RGBD camera is usually sufficient, the most relevant are: cleaning up premises, emergency rescue operations, work with products on a conveyor belt. Such tasks do not require a top-level control system to determine whether objects belong to any particular class. Thus, at the moment, the problem of interaction of robotic manipulator with objects of the workspace, whose shape, class and texture are unknown in advance, is relevant. The solution of this problem would allow robotic manipulators to more effectively interact with dozens of different objects that are found in the environment. Correct positioning of the operating tool of the robotic manipulator plays a key role in the interaction of the robot with the scene.

It can be argued that models, methods and algorithms for grasping and transferring a priori unknown objects based on RGBD images open up wide opportunities in such areas as service robotics and special robotics for emergency rescue operations. These capabilities can be used when the robot is operating in unprepared and uncontrolled real-world environment.

CURRENT STATUS OF THE PROBLEM

The problem of grasping an a priori unknown object by a robotic manipulator is complex due to the need to process data from sensors, undertake scene analysis, and assess the planned grasp in accordance with certain criteria, bearing in mind the need to minimize the time spent on searching for options for the positions of the gripper. Much research is devoted to solving the problem of grasping a priori unknown objects while ensuring collision avoidance [1–4]. The complete solution to this problem includes the execution of a number of subtasks, among which may be included: segmentation of a priori unknown objects in an RGBD image, reconstruction of the shape of objects, determination of the position and orientation of selected a priori unknown objects, as well as synthesis of hypotheses about the optimal gripper pose with collision avoidance.

The authors of [1, 5] resort to approximation of the gripper shape using a mathematical model of a displaced and oriented cylinder, which is subsequently used to synthesize possible gripper positions and select the optimal one. It is noted in the papers that most of the existing approaches that analyze an RGBD image extract a flat surface of a table or floor on which objects are located using RANSAC[2]. By this means, the stage of segmenting the entire scene using a neural network to obtain a set of clusters of points that reflect objects on the surface can be avoided. In addition, the authors distinguish approaches to scene analysis when grasping unknown objects into global, where the selection of the gripper position is performed based on the reconstructed 3D model of the object, and local methods, which rely on the boundaries or segments of objects in the image. In this context, the approach in the present paper is for the most part global, since the search for the target position of the gripper is carried out on the basis of a selected subset of the point cloud. Working with three-dimensional data of a point cloud increases the time for finding a solution, but allows the possibility of checking for collisions of the gripper with scene objects.

Information about the segments of all visible individual objects of a static scene can form the basis for the synthesis of grasping robot configurations. For this, the authors of [6] studied the segmentation of objects without a priori knowledge of their classes. Here, a hierarchical neural network structure is used. With this approach, the neural network is able to segment the visible parts of objects and extract the predicted full segments of objects, including their hidden parts, as well as segment the "blocked" areas of objects in an RGBD image. This hierarchy, comprising the novelty of the approach, lies in the fact that each new unit of information is obtained on the basis of the previous ones. Using such a neural network, a robotic manipulator can obtain a target item from a pile of items. To do this, it removes objects blocking the target object one by one, obtaining the target position of the manipulator using the Contact-GraspNet neural network [2], which generates the position and orientation of the parallel gripper, as long as the target object has "blocked" areas. Thus, the successive grasp and transfer of objects that prevent access to the target object is achieved, followed by the grasping of the target object itself.

Reconstruction of selected objects is also the subject of a large number of studies. The reconstructed object model allows synthesizing many options for the location of the grasping device relative to the object, after which the optimal option is selected. Thus, the problem can be solved using approximation of the found subset of the point cloud with a superquadric model or a primitive-body model and through reconstruction using the

[1] An RGBD camera is a sensor widely used in robotics that provides, in addition to a color image of the environment, depth information, i.e., information about the distance from the camera to the obstacle for each pixel.

[2] Random sample consensus (RANSAC) is an iterative algorithm for estimating the parameters of a mathematical model using random samples from initial data.

machine learning algorithm that analyzes a part of the surface.

In [7], an approach is used based on replacing the selected and processed subset of the point cloud, reflecting a separate object, with a superquadric body. All possible configurations of members of the superquadric family are described by eleven parameters. During the operation of the algorithm, the plane of the table and clusters of points belonging to individual objects are selected. To ensure a robust grasping, two criteria are used: placing the gripper as close as possible to the centroid of the superquadric and placing the touching points in places with the least curvature of the surface. The grasping of an object is carried out by synthesizing the set of possible grasps and selecting the first reachable grasp.

The authors of [8] propose an approach to grasping objects of any shape by representing them as a set of primitive bodies, such as spheres, cylinders, right-angle parallelepipeds and cones. The approach is based on the logic of human behavior when trying to grasp an object. The paper also considers a set of rules that describe the starting positions of the grasping device relative to the primitive body. Thus, summarizing information about the desired location of the gripper can be embedded in the upper-level control system in advance.

In [9], the possibility of reconstructing a scene obtained based on a depth map from one camera position is studied. For the resulting incomplete representation of the scene, the "random forest" algorithm compares each 3D point obtained from the depth map with a prediction of the values of the TSDF function[3] for its neighborhood. An array of such predictions forms the resulting surface. The training sample is formed from a 3D scene model. What makes this work interesting is its lack of reliance on information about the classes of objects that the scene contains. This is of great advantage in this case due to the possibility of applying the solution to reconstruct the shape of a single object.

The continuous genetic algorithm is a classical version of the genetic algorithm, with the exception that the genes of individuals are comprised of real numbers [10]. Thus, the individual of the population is itself a vector of real numbers that contains the solution to the problem. This approach makes it possible to search for a solution in a continuous solution space, which is preferable for a number of problems. A continuous genetic algorithm is especially suited for solving the problem of finding a position and orientation in space that satisfy some criterion.

The authors of [11] propose a method for determining the position and orientation of an object with a non-deformable structure from a black-and-white image of the object and a 3D model of the object known in advance. The initial data consist in a set of key points found on a black and white image of an object detected by the SUSAN[4] algorithm, including a 3D model of the object with a set of key points marked on it. The authors used three Euler angles α, β, γ along with three projections of the transfer vector $\mathbf{T}_x$, $\mathbf{T}_y$, $\mathbf{T}_z$ as genes of an individual's chromosome. The fitness function was calculated as the average value of the distance between each key point recognized by the SUSAN algorithm and the key point of the model closest to it projected onto the image plane. As a result, the genetic algorithm selected such solutions in which the key points of the modeled object coincided with the original ones. The criterion for the end of the algorithm—i.e., the threshold value of the average distance at which the position and orientation were visually found correctly—was determined experimentally. The algorithm scheme used single-point crossing, selection by roulette wheel rotation, and elite generational reproduction—the transition of the best individual of the current generation to the next generation[5].

An example of a local approach in the terminology of [5] is presented in [12]. Here, the authors used a 2D depth image as input. The position and orientation of the robot's grasping device is determined based on a search for the faces of objects. The main disadvantage of this approach is the lack of collision avoidance with other obstacles, while the main the advantage is the high speed of work.

The authors of [3] and [13] also use approaches based on the applications of neural networks. In [3], based on the TSDF representation of the scene, the convolutional neural network calculates the expected grasping quality index, orientation, and width of the grasping device opening in one pass for each scene voxel. To do this, the neural network was trained on reliable data obtained in the simulator. In [13], in order to achieve the segmentation of a priori unknown objects, the neural network generates a feature map that reflects where areas with the same properties are concentrated in the image. Then, based on this map, the clustering algorithm determines the number of clusters in the image

[3] Truncated signed distance function (TSDF) is a function for representing a three-dimensional surface as a voxel grid, each voxel being marked by the distance to the nearest surface.

[4] SUSAN is an algorithm for detecting features in an image that analyzes brightness changes in the local neighborhood of the considered point.

[5] Batishchev D.I., Neimark E.A., Starostin N.V. *Application of genetic algorithms to solving discrete optimization problems.* Teaching and methodological materials for the advanced training program "Information technology and computer modeling in applied mathematics." Nizhny Novgorod: UNN; 2007. 85 p. (in Russ.).

and provides information about the resulting segments. Data on segments of a priori unknown objects can form the basis for further analysis of the scene in order to obtain the target position of the grasping device.

As part of solving the problem of analyzing a complex scene for the interaction of a robotic manipulator with a priori unknown objects of the scene, several training data sets have also been developed, such as the Object Segmentation Database. This dataset is designed to train neural networks in the task of segmenting randomly located unknown objects of various shapes and contains 111 data units, which include the RGB image of the scene, the depth map, and segmentation information. Segmented objects are assigned to only one class comprising "object"; all copies of this class are selected separately on the training images. Each element of this dataset consists of an RGB image, a depth map, and an annotated image with object segments. The scenes presented in the dataset consist of several randomly arranged objects of different shapes and textures.

Thus, the task considered in the present paper is relevant since it enables finding the required position of the grasping device for several epochs of the genetic algorithm via a heuristic approach.

APPLICATION OF A GENETIC ALGORITHM TO SEARCH THE TARGET POSITION OF A GRIPPER IN A COMPLEX SCENE

In the present work, a continuous genetic algorithm was used to find the target position of the gripper in a complex scene with randomly located objects of unknown shape. The general algorithm consists of several stages, shown in Fig. 1. As can be seen, at the beginning of the algorithm, a random scene is generated, consisting of various randomly located objects. Then the Kuka LBR iiwa 7 R800 robotic manipulator (manufactured by KUKA, Germany) moves to a predetermined configuration such that the optical axis of the RGBD camera fixed on the gripper is directed perpendicular to the scene plane. In this position, the RGBD camera takes RGB and depth shots of the scene. The resulting images are processed to produce an RGBD image, fed to the input of a neural network with the U-Net architecture, the output of which is also processed and represents a segment of the object closest to the camera, which is further considered as the target. Based on the depth map, a point cloud of the scene is constructed; this is then divided into two subsets based on information about the segment of the target object. In a subset of the point cloud of the object and obstacles, unobservable areas are eliminated using 2D Delaunay triangulation. Then, using a continuous genetic algorithm, the position of the gripper is determined in order to grasp the target object. To reach the target point, the robot first moves to the pre-grip position, then iteratively moves the gripper towards the object until the target position is reached. The grasped object is then transferred to the target area for the objects. These steps are discussed in more detail below.

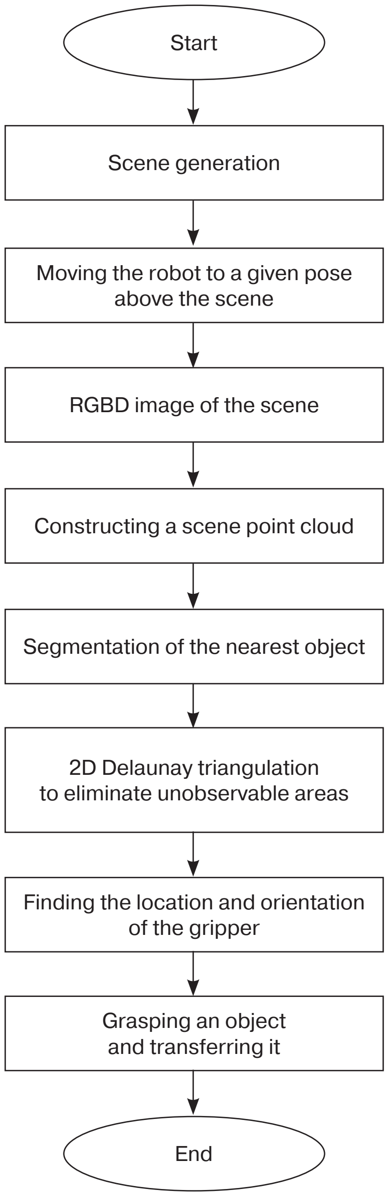


Fig. 1. General diagram of the algorithm

The initial data comprises an RGBD image from a camera attached to the robot's gripper. The result of the algorithm is a vector of generalized coordinates that describes the angles of the robot's drives at which the grasping of the target object becomes possible.

To create a chaotic scene, 20 objects, comprising bricks, stones, beams, etc., were modeled. The objects had a uniform texture and different shapes, including asymmetric, as well as the corresponding mass parameters. In the process of generating the scene, the

(a) (b)

Fig. 2. Creating a chaotic scene:
(a) models of individual scene objects; (b) generated random scene

algorithm randomly selected an object, added it to the scene, applied rotation and displacement to it, after which the algorithm waited 10 s for the object to fall and stop moving in the simulator. Thus, the presence in the desired area of the scene of a given number of randomly selected and randomly located objects was achieved. The modeled objects and generated random scene are shown in Fig. 2.

Following the creation of the random scene, the robot was moved to an initial position specified by a predefined vector of generalized coordinates. On the basis of RGB and depth maps of the scene, a point cloud was constructed and data were generated to be delivered to the input of a neural network with the U-Net architecture [14]. The neural network was preliminarily trained to segment the nearest object in an RGBD image using a simulator-collected and annotated training set. Here, it was necessary to process the data prior to delivery to the input of the neural network separately from the data received from the output of the neural network. The corresponding processing stages are shown in Fig. 3. As can be seen from the figure, pre-processing consisted in finding a point in the depth map with a minimum distance value, cropping the image section with

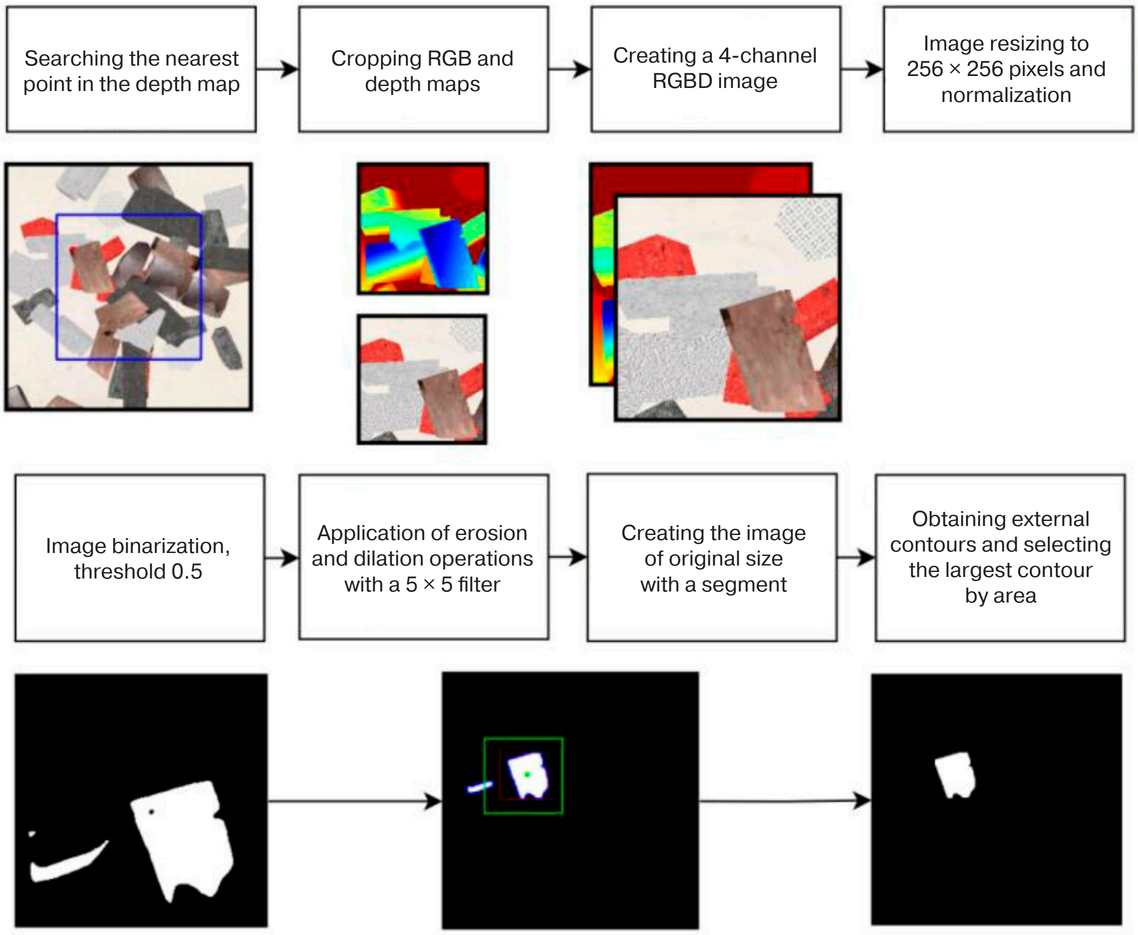


Fig. 3. Preliminary and final data processing for a segmenting neural network with U-Net architecture

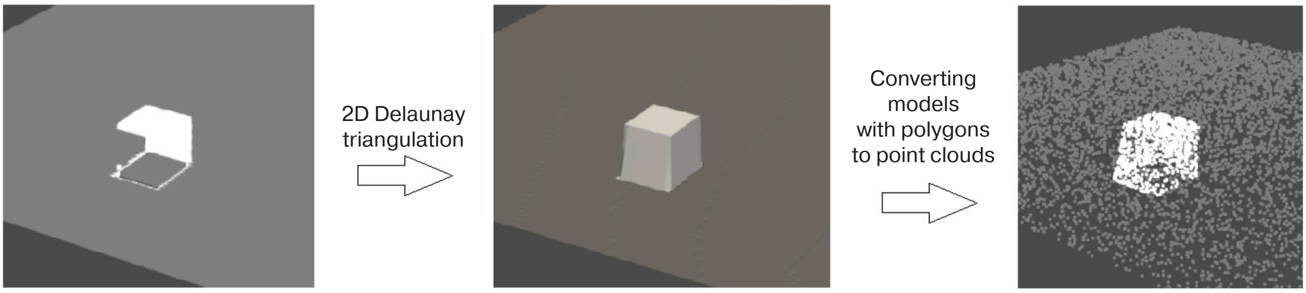


Fig. 4. Transformation of the point cloud in order to eliminate unobservable areas

this point in the center, forming a 4-channel RGBD image from RGB and depth maps, scaling to the size of the neural network input layer 256 × 256, and normalization. Post-processing was required to eliminate the false segments detected by the neural network. To do this, the output of the neural network with a threshold $t = 0.5$ was transformed into a binary mask, with morphological dilation and erosion operators used to eliminate small erroneously segmented areas, and the segment with the largest area being taken as the true segment.

To increase the speed of the algorithm, only a part of the point cloud was constructed near the selected segment. Due to the depth map obtained from a single camera position being the only source of data for constructing a point cloud, there were unobservable areas in the constructed cloud. This could prevent the genetic algorithm from correctly assessing the fitness of an individual and searching for the correct position of the gripper. To solve this problem, the point cloud was divided into two subsets based on the selected segment:

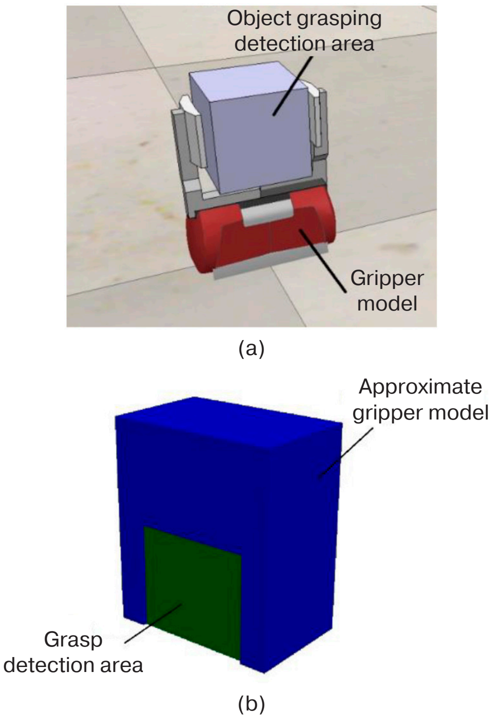


Fig. 5. Gripping device models:
(a) original model; (b) approximated model

one subset of points belonging to the obstacle and another subset belonging to the target object. Further, using two-dimensional Delaunay triangulation, the polygons between the points were completed. The resulting models with polygons were then sent to the input of an algorithm that converted the 3D model into a point cloud. The resulting 2 point clouds, which did not contain unobservable areas, were consequently suitable for the operation of a continuous genetic algorithm (Fig. 4).

The position and orientation of the working tool of the robotic manipulator relative to the base coordinate system are uniquely described by a six-parameter vector: three projections of the transfer vector and three Tait-Bryan angles describing successive rotations of the object in the *ZYX* order; these six real numbers make up the chromosome of an individual. For the operation of the genetic algorithm, a 3D model of the gripper was used. The detection area was set within which the target object for grasping was to be located. To reduce the time spent on processing the 3D model of the gripper, it was approximated by a low-poly model having the same geometry as the original model (Fig. 5).

To calculate the fitness of an individual, the calculation of the number of points of the target object and the obstacle located inside the gripper model and the grasp detection area is used. To do this, we used an algorithm based on the ray tracing method, which determines the fact that a point is located inside a closed three-dimensional surface. In order for the resulting position of the working body to be oriented vertically, the desired intervals for the angles of rotation relative to the *X* and *Y* axes were set: $-20° < \alpha_g < 20°$, $-20° < \beta_g < 20°$. When calculating the fitness function, going beyond the boundaries of these intervals was taken into account. The fitness of an individual was calculated by the formula:

$$F(p)=\begin{cases}1000+x, \\ \text{if } O_{\text{inside gr}}>0 \text{ or } I_{\text{inside gr}}>0, \\ 500+x, \\ \text{if } O_{\text{inside gr}}=0 \text{ and } I_{\text{inside gr}}=0 \text{ and } I_{\text{inside vol}}=0, \\ \dfrac{1}{I_{\text{inside vol}}}+x, \\ \text{if } I_{\text{inside vol}}>0 \text{ and } O_{\text{inside gr}}=0 \text{ and } I_{\text{inside gr}}=0,\end{cases}$$

where x is the sum of the absolute values of the excess of the selected angles α and β over the given intervals; $O_{\text{inside gr}}$ is the number of points in the obstacle point cloud inside the gripper model; $I_{\text{inside vol}}$ is the number of object points inside the detection area; $I_{\text{inside gr}}$ is the number of object points inside the gripper model.

The genetic crossover and mutation operators used are shown in Fig. 6. As can be seen from the figure, the crossover operator is an elementwise weighted sum of the genes of the two original parental chromosomes, while the weight coefficient is a random number. The mutation operator, designed to introduce random changes in the chromosome of an individual, is presented in three variations: introducing a random change in each gene of the original chromosome with a probability of 50%, introducing a change in a randomly selected gene of each logical subgroup of genes, and initiating a change in one randomly selected gene. The choice of the mutation operator is made with equal probability.

The following scheme was used in the implemented continuous genetic algorithm:

- construction of the initial population near the central point of the segmented object;
- population size: 34 individuals;
- 16 individuals are selected for crossover: 10 best and 6 random individuals;
- pairs for crossover are random, with each pair of individuals interbreeding twice, giving four offsprings;
- offsprings undergo mutation with a probability of 30%, while the type of mutation is chosen with equal probability;
- elite generational reproduction: the 2 most adapted individuals of the previous generation are copied to the next generation and do not undergo mutation;
- criterion for terminating the algorithm's work: reaching a given number of epochs or reaching a certain value of the fitness function.

The following conditions were chosen as criteria for stopping the search for a solution by the genetic algorithm:

- the number of object points located inside the grasping detection area is more than 40% of their total number;
- more than 50 epochs have passed.

The criteria for stopping the search were chosen based on the analysis of the simulation results of the robotic manipulator. During a series of runs of the algorithm with different target objects, it was noticed that the threshold value of the proportion of object points located in the grasping detection area of the total number of points, equal to 40%, in most cases ensures successful grasp of the object. At the same time, the selected ultimate number of epochs of the algorithm, equal to 50, provides a balance between the speed of calculation and the probability of finding a solution; this is despite the fact that in most cases the search for the position of the gripper to successfully grasp the object took much less than 50 epochs.

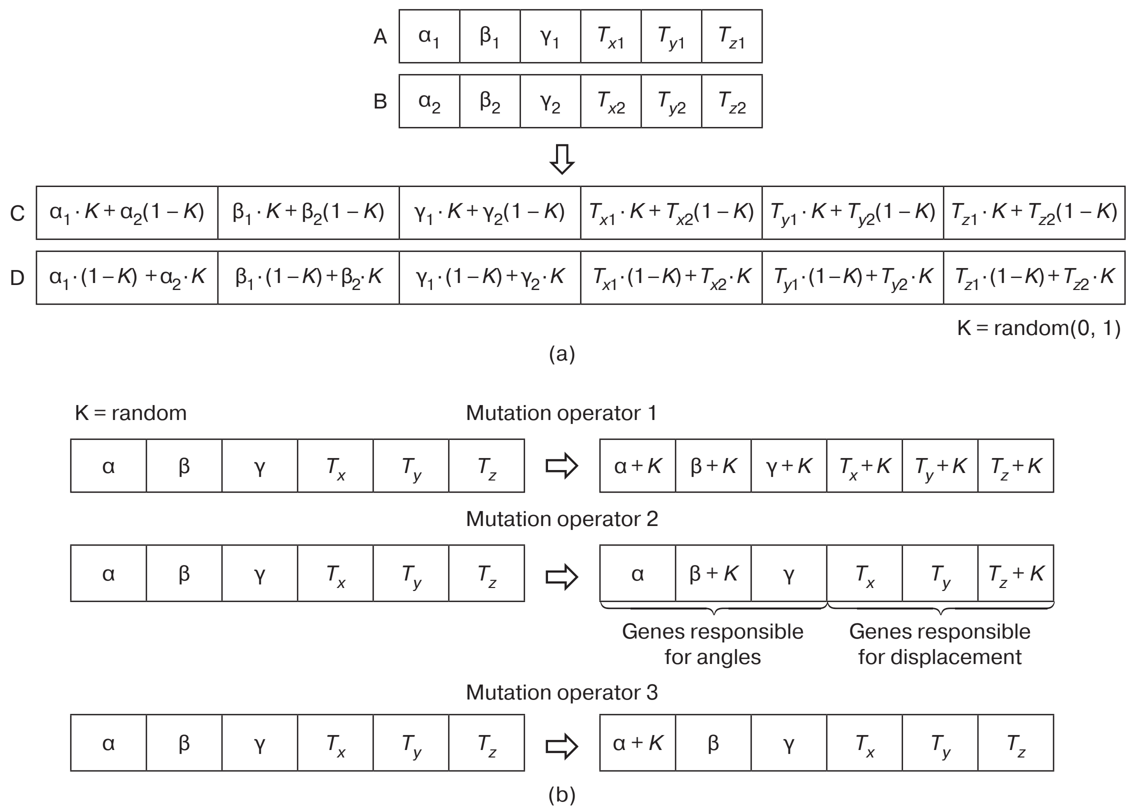


Fig. 6. Genetic operators: (a) crossover operator; (b) set of mutation operators

The target position and orientation of the gripper are determined by the individual with the highest degree of fitness (minimal $F(p)$ value) at the end of the genetic algorithm. Based on these data, the pre-grasp position is calculated by shifting the obtained position along the OZ axis by −0.1 m. Then the drive angles for ten intermediate positions of the gripper between the pre-grasp position and the target position are calculated based on solving the inverse kinematics problem. To grasp the object, the robot approaches the gripper by sequentially setting the drive angles for each intermediate position. In this way, approximation along the vector is achieved to avoid collisions with other objects in the scene. The robot with the grasped object then moves to a position above the target area using a predetermined generalized coordinate vector and the gripper opens. The least squares procedure of the SciPy Python library was used to solve the inverse problem of kinematics. This procedure implements the nonlinear least squares method, in which the search for the minimum of the objective function of several variables is carried out:

$$F(\mathbf{Q}) = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{4}\left(A_{ij}^{08}(\mathbf{Q}) - A_{ij}^{08\_\mathrm{goal}}\right)^2,$$

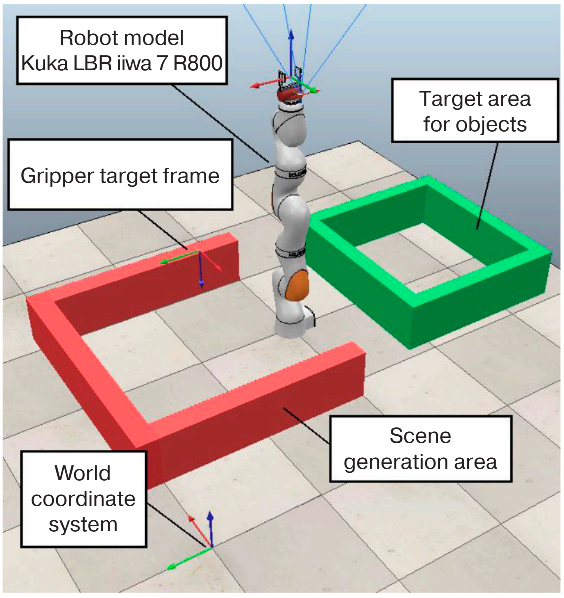
where $\mathbf{A}^{08}(\mathbf{Q})$ is a homogeneous transformation matrix describing the transition from the world coordinate system to the coordinate system fixed on the gripper, calculated based on the generalized coordinate vector $\mathbf{Q}$ found by the numerical method and representing a matrix of 1 column and 7 rows, the elements of which are the angles of the joints of a robotic manipulator; $\mathbf{A}^{08\_\mathrm{goal}}$ is the target uniform transformation matrix computed using the given position and orientation.

Thus, it is possible to determine the assumptions made during the development of the algorithm. Here, it is assumed that the scene objects have a uniform texture and are not deformable, that one of the dimensions of each scene object is less than the maximum opening width of the grippers, that all objects are in the working area of the robot, and that the approximate gripper model has the maximum opening width.

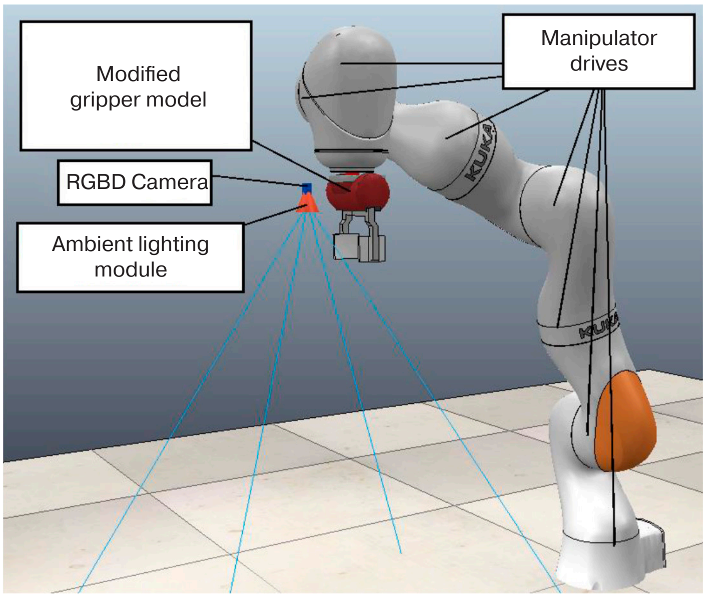
EXPERIMENTAL EVALUATION OF THE METHOD

In order to simulate the interaction of a robotic manipulator with objects of a complex visual scene, a virtual scene was developed in the CoppeliaSim simulator. Its original form is shown in Fig. 7. The scene comprises a model of the Kuka LBR iiwa 7 R800 robotic manipulator with a 512 × 512-pixel RGBD camera mounted on a gripper and an ambient lighting module designed to make the RGB image uniform by eliminating shadows. There is a source zone for scene generation and a target zone.

In order to study the performance of the continuous genetic algorithm when searching for the target position of the grasping device, five experimental runs of the algorithm were carried out on the generated scene. The process of grasping one object is shown in Fig. 8. Convergence graphs of the continuous genetic algorithm are shown in Fig. 9. As can be seen from the graphs, the algorithm converged to a satisfactory solution in less than 31 epochs in all experimental runs. The initial fitness of the best individual in most cases was high due to individuals of the initial population being created in the vicinity of the central point of the segmented object, along with a maximal width of the opening of the grasping device in the approximated model. This increased the probability that a part of the point cloud



(a)



(b)

Fig. 7. Scene for simulation: (a) composition of the scene; (b) robot model

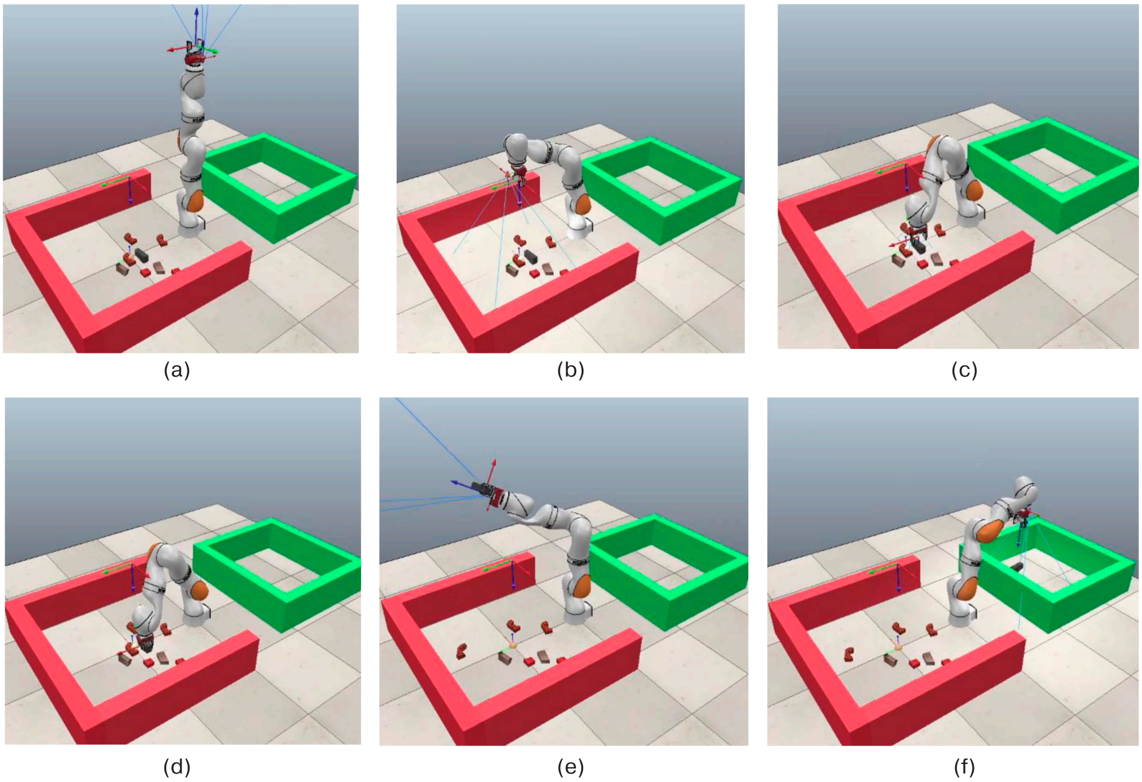


Fig. 8. The process of grasping one object: (a) the configuration of the robot at the start of the program; (b) robot configuration for obtaining RGB and depth maps of the scene; (c) pre-grasp configuration; (d) last intermediate configuration; (e) move to end position; (f) configuration for dropping the object to the target area

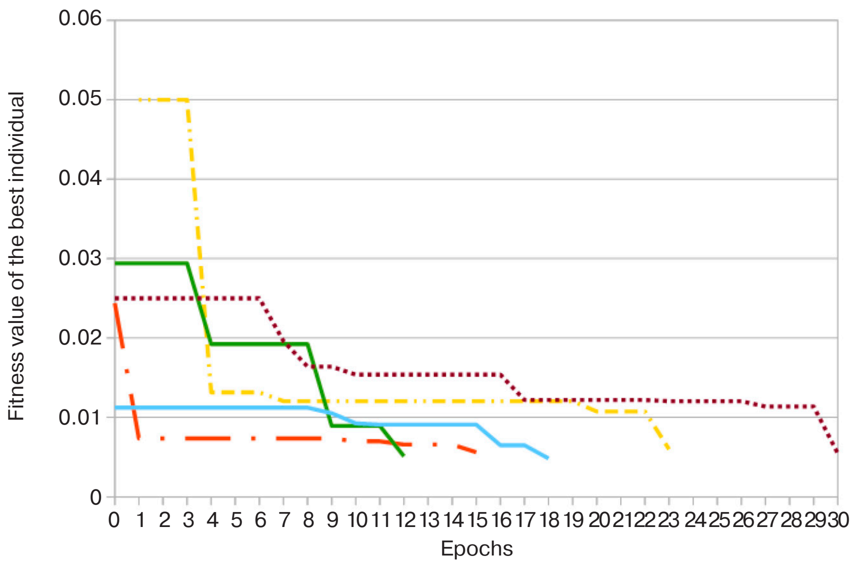


Fig. 9. Charts of convergence

of the object in the created individual was in the grasping detection area, and the gripper did not collide with the object. Since in this case the genetic algorithm needs more iterations to find the position of the gripper without colliding with the object's point cloud, an increase in the convergence time of the algorithm is associated with an increase in the size of the target object. The average convergence time was 1.9 s on a computer equipped with an AMD Ryzen 5 3500U processor; the GPU was not involved.

CONCLUSIONS

On the basis of the developed software for implementing the analysis of the RGBD image of the scene, the genetic algorithm is confirmed as suitable for solving the problem of obtaining the target position of the grasping device. The advantages of the developed solution are the ability to work using a point cloud obtained from one position of the camera, as well as the capability to eliminate unobservable areas. An important role is played by genetic operators and the implementation of the fitness function since their calculation directly affects the average time of finding a solution, which should therefore be minimized. Therefore, measures were taken to increase the speed of the algorithm: use of a subset of the original point cloud and approximation of the gripper with a primitive low-poly model.

Despite the success of the resulting solution, there are opportunities for its further improvement. In particular, the study conducted in [5] clearly indicates that fixing a camera on the end of the robot offers an advantage over a camera fixed above the stage, since it then becomes possible to generate a more detailed point cloud of the scene taken from several positions along a predetermined scanning path. Generating a point cloud based on RGBD images from several camera positions allows it to be constructed more informatively to avoid excessive use of Delaunay triangulation to complete the cloud. Also, in a number of studies [4, 15] parallelism of actions is used; this involves the use of calculations to obtain the target gripper position during the movement of the robot to some intermediate position above the scene. Finally, it is of interest to study the optimal scheme of the genetic algorithm in terms of the number of individuals in the population for obtaining the most efficient crossover and mutation operator.

The novelty of the results obtained lies in the study of the applicability of a continuous genetic algorithm in the problem of positioning the grasping tool of a robotic manipulator. The results of the study can be used to solve the problems of positioning a robotic manipulator in an environment with a priori unknown objects. Examples of such tasks include cleaning premises, performing warehouse operations, and clearing rubble.

Authors' contributions

A.D. Voronkov participated in the development of the concept of using a continuous genetic algorithm in the task of positioning the gripper of the manipulator; developed algorithmic and software implementing a continuous genetic algorithm and a complex algorithm for scene analysis; participated in the drafting of the text and making edits to the article.

S.A.K. Diane participated in the development of the concept of the application of the continuous genetic algorithm; proposed the scenario of the final experiment; provided scientific guidance; participated in the drafting of the text and making edits to the article.

REFERENCES

1. Lei Q., Wisse M. Fast grasping of unknown objects using cylinder searching on a single point cloud. In: *Ninth International Conference on Machine Vision* (*ICMV 2016*). 2016. V. 10341. https://doi.org/10.1117/12.2268422
2. Sundermeyer M., Mousavian A., Triebel R., Fox D. *Contact-GraspNet: efficient 6-DoF grasp generation in cluttered scenes*. arXiv [Preprint]. 2021. 7 p. Available from URL: https://arxiv.org/abs/2103.14127
3. Breyer M., Chung J., Ott L., Siegwart R., Nieto J. Volumetric grasping network: Real-time 6 DOF grasp detection in clutter. In: *4th Conference on Robot Learning* (*CoRL 2020*). P. 1602–1611. Available from URL: https://proceedings.mlr.press/v155/breyer21a/breyer21a.pdf
4. Lippiello V., Ruggiero F., Siciliano B., Luigi V. Visual grasp planning for unknown objects using a multifingered robotic hand. *IEEE/ASME Transactions on Mechatronics*. 2013;18(3):1050–1059. https://doi.org/10.1109/TMECH.2012.2195500
5. Lei Q., Meijer J., Wisse M. A survey of unknown object grasping and our fast grasping algorithm-C shape grasping. In: *2017 3rd International Conference on Control, Automation and Robotics* (*ICCAR*). P. 150–157. https://doi.org/10.1109/ICCAR.2017.7942677
6. Back S., Lee J., Kim T., Noh S., Kang R., Bak S., Lee K. *Unseen object amodal instance segmentation via hierarchical occlusion modeling*. arXiv [Preprint]. 2021. 8 p. Available from URL: https://arxiv.org/abs/2109.11103
7. Abhijit M., Federico T., Perez-Gracia A. Grasping unknown objects in clutter by superquadric representation. In: *2018 Second IEEE International Conference on Robotic Computing* (*IRC*). P. 292–299. https://doi.org/10.1109/IRC.2018.00062
8. Miller A., Knoop S., Christensen H., Allen P. Automatic grasp planning using shape primitives. In: *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*. P. 1824–1829. https://doi.org/10.1109/ROBOT.2003.1241860
9. Firman M., Aodha O., Julier S., Gabriel J. Structured prediction of unobserved voxels from a single depth image. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). P. 5431–5440. https://doi.org/10.1109/CVPR.2016.586
10. Gridin V.N., Solodovnikov V.I. A continuous genetic data preprocessing algorithm searching coefficients of the approximating function. *Novye informatsionnye tekhnologii v avtomatizirovannykh sistemakh = New Information Technologies in Automated Systems*. 2018;21:302–306 (in Russ.).
11. Rossi C., Abderrahim M., Diaz J. EvoPose: A model-based pose estimation algorithm with correspondences determination. *2005 IEEE International Conference Mechatronics and Automation.* V. 3. P. 1551–1556. https://doi.org/10.1109/ICMA.2005.1626786
12. Jabalameli A., Behal A. From single 2D depth image to gripper 6D pose estimation: A fast and robust algorithm for grabbing objects in cluttered scenes. *Robotics.* 2019;8(3):63. https://doi.org/10.3390/robotics8030063
13. Xiang Y., Xie C., Mousavian A., Fox D. Learning RGB-D feature embeddings for unseen object instance segmentation. In: *4th Conference on Robot Learning* (*CoRL 2020*). P. 461–470. Available from URL: https://proceedings.mlr.press/v155/xiang21a/xiang21a.pdf
14. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2015. P. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

15. Dune C., Marchand E., Collowet C., Leroux C. Active rough shape estimation of unknown objects. In: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems.* P. 3622–3627. https://doi.org/10.1109/IROS.2008.4651005

About the authors

Andrey D. Voronkov, Postgraduate Student, Department of Control Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: a.voronkov.rtu@yandex.ru. https://orcid.org/0000-0003-4688-9346

Sekou A.K. Diane, Cand. Sci. (Eng.), Assistant Professor, Department of Control Problems, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). Scopus Author ID 57188548666, ResearcherID T-5560-2017, https://orcid.org/0000-0002-8690-6422

Об авторах

Воронков Андрей Дадашевич, аспирант кафедры проблем управления Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: a.voronkov.rtu@yandex.ru. https://orcid.org/0000-0003-4688-9346

Диане Секу Абдель Кадер, к.т.н., доцент кафедры проблем управления Института искусственного интеллекта ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). Scopus Author ID 57188548666, ResearcherID T-5560-2017, https://orcid.org/0000-0002-8690-6422

Translated from Russian into English by Evgenii I. Shklovskii
Edited for English language and spelling by Thomas A. Beavitt