

## TECHNICAL VISION SYSTEM WITH ARTIFICIAL INTELLIGENCE FOR CAPTURING CYLINDRICAL OBJECTS BY ROBOT

*The article presents a variant of working with a system of technical vision, which recognizes cylindrical objects. This vision system based on artificial intelligence, which allows you to determine the circles in the image. The coordinates of the value of the circle are necessary for the exact positioning of the robot manipulator. The calculation of the gradient and threshold separation determine the gaps in the intensity of the image of the object. These methods define pixels lying on the border between the object and the background. The further process consists in connection of the segments of a contour separated by small intervals, and in association of separate short segments. Thus, contour detection algorithms accompanies by procedures for constructing object boundaries from the corresponding pixel sequences. The resulting image has sufficient information for artificial intelligence analysis to detect the circle. The software is developed and experimentally tested in operation. The operation of the technical vision system experimentally tested in the work, namely the capture of a cylindrical object. Work vision systems has been experimentally proven to work and is passionate cylindrical object. The coordinates of the value of the circle, which are necessary for the exact location of the robot manipulator were determined by artificial intelligence in 41 milliseconds. The obtained coordinates were transmitted to the microprocessor to adjust the position of the manipulator. The robot accurately captured a cylindrical object.*

*Keywords: robot, technical vision system, coordinate.*

ГОРЯЩЕНКО С.Л., ГОРЯЩЕНКО К.Л.  
Хмельницький національний університет

## СИСТЕМА ТЕХНІЧНОГО ЗОРУ З ШТУЧНИМ ІНТЕЛЕКТОМ ДЛЯ ЗАХОПЛЕННЯ ЦИЛІНДРИЧНИХ ОБ'ЄКТІВ РОБОТОМ

*У статті представлений варіант роботи із системою технічного зору, яка розпізнає циліндричні предмети. Ця система зору, заснована на штучному інтелекті, що дозволяє визначати кола на зображенні. Була розроблена функціональна діаграма роботи робота з системою технічного зору. Для цього проведено аналіз роботи системи технічного зору, визначено його можливості та задачі, що можуть виконуватись. Розроблено координатні центри для кожного елемента системи, було визначено та запропоновано їх перетворення. Система технічного зору забезпечила координацію положення циліндричного об'єкту для системи управління роботом. Визначено матрицю, що може вирішити задачу адаптивного позиціонування, також вона визначає програмні значення всіх керованих координат поточного робочого інтервалу і виробляє керуючу дію на приводи для обробки програмних значень із потрібними координатами. Також розроблено програмне забезпечення на основі штучного інтелекту, завданням якої є піксельна побудова зображення об'єкта. Отримане зображення аналізується штучним інтелектом для визначення кола та його центру. В основі методу є розрахунок градієнта, тобто поріг поділу - визначають прогаліни в інтенсивності зображення об'єкта. Вони визначають пікселі, що лежать на межі між об'єктом і фоном. Нам потрібно побудувати всі прямі, що проходять через кожну пару точок, а потім знайти всі підрядки точок, близьких до певних прямих. Таким чином, алгоритми виявлення контуру супроводжуються процедурами побудови меж об'єкта з відповідних послідовностей пікселів. Отримане зображення має достатню інформацію для аналізу штучним інтелектом на предмет виявлення кола.*

*Робота системи технічного зору була експериментально перевірена в роботі, а саме захоплені циліндричного об'єкту. Координати значення кола, які необхідні для точного розташування робота-маніпулятора були визначені штучним інтелектом за 41 мілісекунду. Отримані координати були передані до мікропроцесора для коректування положення маніпулятора. Захват робота точно захопив циліндричний об'єкт.*

*Ключові слова: робот, система технічного зору, координата.*

**Introduction.** The robots with the technical vision system (TVS) are widely and successfully in many industries in such operations as control and rejection of various parts and products in appearance, assembly of components, product packaging, installation of electronic circuits, routing of material flows in flexible production systems, management of transport mobile robots.

The role of adaptive robots with technical vision and elements of artificial intelligence is growing even more with the expansion of the robotics process in such areas as construction, agriculture and forestry, utilities, everyday life, where the working environment is complex. TVS is needed for robots designed to study outer space and underwater, which are used to extinguish fires and other natural disasters. In such cases, the TVS should provide not only data for direct generation of feedback signals in the control of robot movements and transmission of environmental information to the human operator, but also information for the robot to form a model of the working environment for automatic action planning [1,2].

The functional tasks of TVS can divided by the level of their relative complexity. The basic tasks usually include:

- detection of the presence of the object;
- measuring the distance to the object, its linear and angular displacements;
- measurement of geometric parameters of the object (linear and angular dimensions, area, etc.);
- determination of physical characteristics of radiation from the object;
- counting the number of objects, etc.

Complex tasks are perform by the TVS which provides the control system of the manipulation robot with the information that need to capture disordered objects. Such tasks include:

- overview of the working scene (flat or spatial) to find a certain desired object - one or one of several that lying in isolation from each other;
- determining the location and orientation of the object;

At the same time, objects can differ not only in size and shape (silhouette), but also in texture, color, etc., at rest or in motion. This example expresses the fact that the relative complexity of the functional tasks of TVS characterize by a number of indicators, which will show below:

The dimension of the scene consideration is define. Spatial problems that require the analysis of three-dimensional scenes are usually much more complex than "flat" problems, in which it is enough to process images of one two-dimensional projection of the scene on the picture plane. The intermediate position occupies by quasi-plane problems, where the picture plane consider with the addition of information about the distance of visible points of the image.

Degree of isolation of objects. On this basis, the tasks of TVS are classified as follows (in ascending order of complexity): problems with a single object; problems with several isolated objects in the image; problems with images of objects that touch each other; overlapping tasks from one work scene object to another.

Homogeneity of objects. This feature characterizes the complexity of the classification problems facing TVS. If all the objects that can presented by TVS are identical, this problem does not arise at all. If there are a number of different objects, the task of classifying them is more complex, the more similar their images are, that is, the more features are needed to describe the differences of each object from the rest at once.

Stationarity of the working scene. Images of objects at rest are easier to process and analyze than images of objects moving relative to video sensors.

Number of image elements. The larger the field of view and the higher the required resolution of the TVS, the more difficult the task, because you have to process more elements of the resulting image (pixels).

The number of gradations of brightness. Problems in which it is enough to analyze the silhouettes of objects, ie to work with two-graded (binary) images, are usually solved more simply than tasks that require the analysis of multi-graded (halftone) images.

Spectral composition of the video signal. To endow a robot with the ability to perceive color images, to put it simply, is more difficult than to equip it with monochromatic vision.

**Main part.** TVS can solve the more complex problems, taking into account all the sets of described criteria and get higher the level of functional flexibility (intelligence) of robot that provide with its help [3, 4]. TVS must provide to fulfill its functional purpose in the general case as:

perception of the optical signal and image formation;  
image pre-processing in order to reduce noise exposure,  
improve contrast, correct distortion, compress information; segmentation of the image of the scene into its component parts - selection of the necessary objects, their fragments, or characteristic features;  
description of images - calculation of their geometric and other characteristics, determination of classifying features, identification of position and orientation; image analysis for pattern recognition or object classification and scene interpretation based on the problem environment; transfer of the obtained results to the robot control system.

Creating a TVS robot, corresponding to its specific functional purpose, includes the following steps:

- setting the task at the content level and creating a technical task on the TVS;
- choice of structural and functional schemes of TVS;
- mathematical formulation of problems of TVS functional blocks;
- development of algorithms for solving functional problems;
- the choice of means of implementation of algorithms and specification of technical characteristics of hardware, hardware-software and software blocks of TVS in accordance with the technical task;
- hardware and software development.

The general scheme of the robot equipped with TVS, together with sensors of the external information, show on fig.1. The robot sensory system provides the control system with information about the current situation in the external environment: the size, type, parameters, location and orientation of objects manipulated the correctness and quality of the robot's technological operations, the existence of obstacles and ways to overcome them, etc.

In accordance with the following scheme, the information from the TVS comes to different levels of the control system. For example, data on the location of obstacles is required at the highest level to build a model of the working environment, in order to plan the work of the robot. The results of object classification need at the strategic level to divide the general action plan into specific manipulation operations, setting their sequence and parameters.

Information about the location and orientation of objects is necessary for the formation at the tactical level of the movements of the working body and the parts of the manipulator, which in turn build the program laws of coordinated change of the respective degrees of mobility. Information about the deviation of the actual trajectory from the programmed one used to generate control signals on the drives. It is need for during program processing in order to correct the movement of the working body of the robot.

It is often difficult to separate the tasks of the described levels of control in real robotic systems. The functions of the high levels take over by the human operator. The lower level of the control system physically implement directly in the power units of the actuators, and other levels built on universal or specialized measuring units.

The general functional scheme that show in fig. 1 corresponds to the case of connecting the TVS to existing software control devices of the robot, which are divided into devices of cyclic, positional, contour and position-contour control.

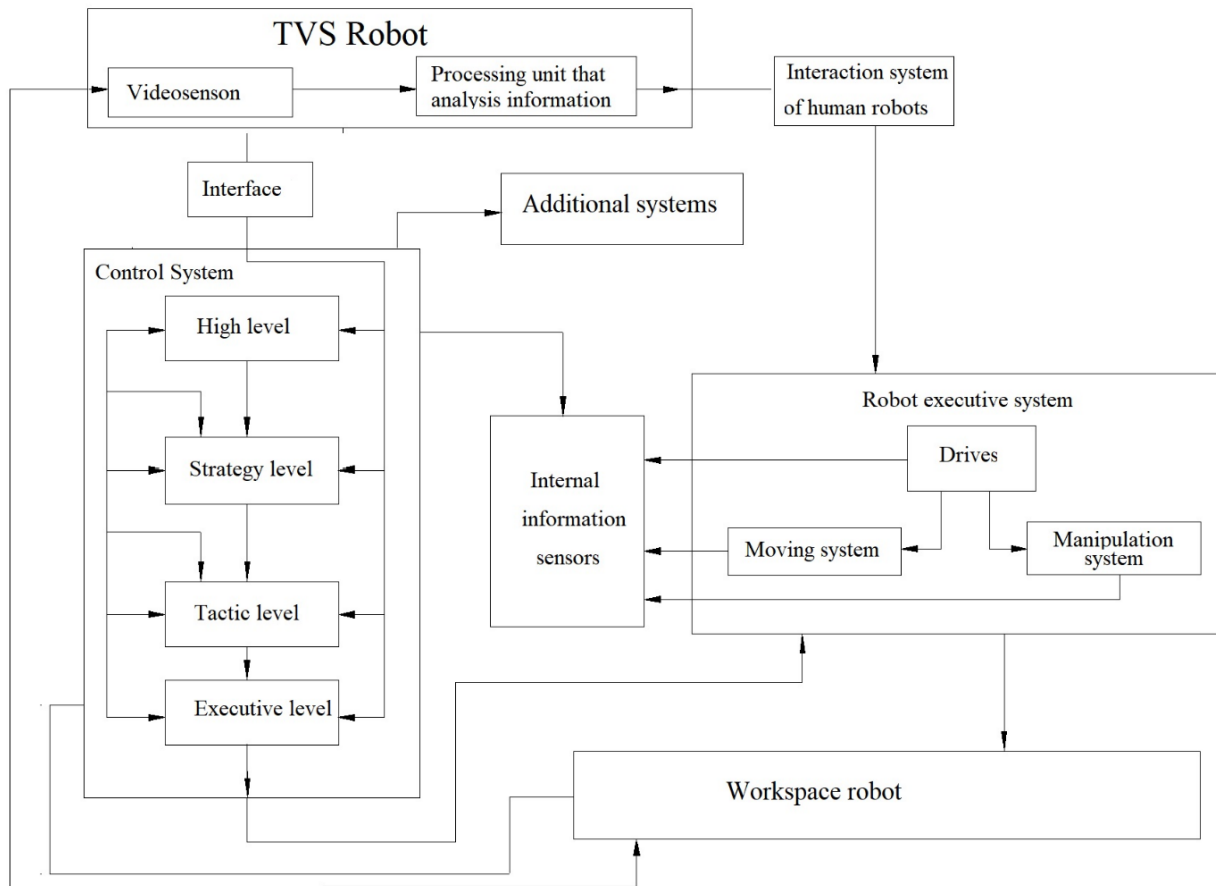


Fig. 1. Functional diagram of robot with TVS

Statement of the problem of adaptive positioning of the manipulator when capturing a non-oriented part lying on the working plane.

The robot must take an object, the position and orientation of which on the work plane are determined by the TVS from the information coming from the webcam, which is fixed above the stage (Fig. 3.2). We introduce the following coordinate systems:  $O_a X_a Y_a Z_a$  - absolute,  $O_o X_o Y_o Z_o$  - associated with the object,  $O_p X_p Y_p Z_p$  - associated with the basis of the work,  $O_3 X_3 Y_3 Z_3$  - associated with the capture,  $O_r X_r Y_r Z_r$  - associated with the camera, more precisely with its image plane, on which the image of a working field is displayed (Fig.2).

Convenient mathematical apparatus of homogeneous coordinates when calculating transitions from one coordinate system to another. Homogeneous coordinates of a point in three-dimensional space are called four numbers  $h_1, h_2, h_3, h_4$ , which are not equal to zero and are associated with its Cartesian coordinates  $x, y, z$  ratios  $x = h_1/h_4, y = h_2/h_4, z = h_3/h_4$ .

For example, the point  $(x, y, z)$  corresponds to the vector of homogeneous coordinates  $(h_1, h_2, h_3, 1)$ . Using homogeneous coordinates, we describe the transition from some coordinate system  $O_o XYZ$  to the system  $O1X'Y'Z'$ , shifted and rotated relative to the first:

$$h' = T_0^1 \cdot h, \quad \dots \dots \dots \quad \dots \dots (1)$$

where  $h'$  is the vector of homogeneous coordinates in the new system;  
 $h$  is a vector of homogeneous coordinates in the old system, and the transition matrix has the form

$$T_0^1 = \begin{pmatrix} x_{X'} & y_{X'} & z_{X'} & 0_{OX'} \\ x_{Y'} & y_{Y'} & z_{Y'} & 0_{OY'} \\ x_{Z'} & y_{Z'} & z_{Z'} & 0_{OZ'} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2)$$

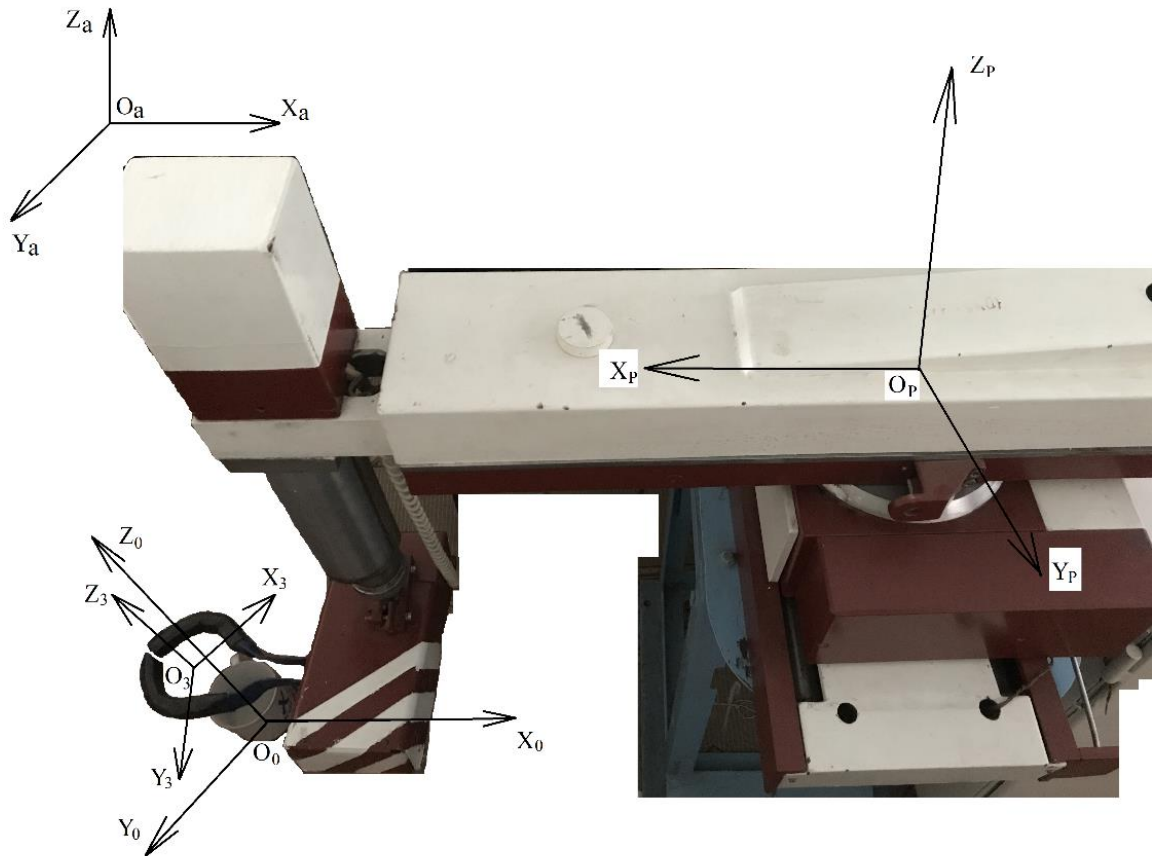


Fig. 2. The coordinate system of the robot and the TVS

Here, the last column is a vector of homogeneous coordinates of the origin of the original system in the new coordinate system, and the 3x3 matrix describes the rotation of the old system relative to the new one.

The robot must bring the gripper to a position where the object would be inside, and their orientation provided a secure grip on the object to perform the task. The position and orientation of the gripping device describe as a 4x4-transition matrix  $T^R$ , which determines the rotation and parallel transfer of the system  $O_3X_3Y_3Z_3$  relative to any selected coordinate system, in which it is most convenient to set the desired state of the gripping device.

We can say that the position of the system  $O_0X_0Y_0Z_0$  in the picture plane of the webcam describe using a matrix:

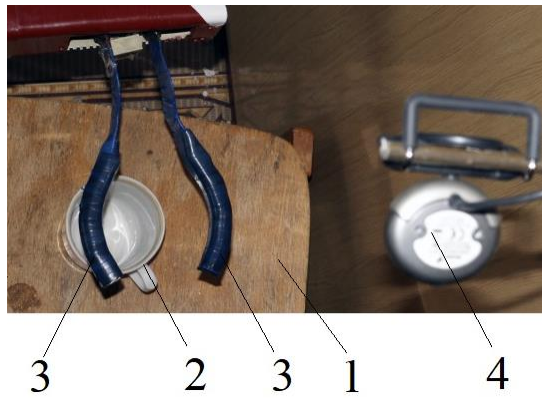
$$T_0^K = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 & x_K \\ \sin\varphi & \cos\varphi & 0 & y_K \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \dots\dots (3)$$

where  $\varphi$  is the angle of orientation of the object in the working plane, which can be found by the TVS in real time, as well as the coordinates of the geometric center of the silhouette of the object.

Note that you can immediately draw the caliber of the camera, so that bypassing the intermediate transitions to the absolute coordinate system. We obtain a matrix that connects the systems  $O_KX_KY_KZ_K$  and  $O_PX_PY_PZ_P$ . Using this matrix, the calculation of the software state of the gripping device on the next portion of information about the coordinates  $x_K$ ,  $y_K$  and orientation  $\varphi$  of the object coming from TVS is slightly simplified.

$$(T^R)_P = T_K^P T_0^K T_3^0, \quad (4)$$

TVS provided targeting to the robot control system, which on the matrix can solve the problem of adaptive positioning, and it determine the program values of all controlled coordinates of the current working interval and produce a control action on actuators to process program values with the desired orientation (Fig.3).



- 1 – Base;
- 2 – Cylindrical object;
- 3 – Manipulator finger;
- 4 – Video camera.

Fig. 3. Robot is preparing for capture cylindrical object in zone of TVS

Segmentation is the process of dividing a scene into its constituent parts or objects. Segmentation is one of the main elements of the automated system of technical vision, because at this stage of processing objects are selected from the stage for further recognition and analysis. Segmentation algorithms usually base on two fundamental principles: discontinuity and similarity. In the first case, the main approach base on the definition of contours, and in the second - on the definition of the threshold level and the expansion of the area. These concepts are applicable to both static and dynamic (time-dependent) scenes. In the latter case, motion can be a powerful tool for improving the performance of segmentation algorithms.

Methods - calculation of the gradient, threshold separation - determine the gaps in the intensity of the image of the object. Ideally, these methods define pixels lying on the border between the object and the background. In practice, this series of pixels rarely fully characterizes the border that have noise, gaps at the border due to uneven lighting and other effects that lead to image blur. Thus, contour detection algorithms accompanies by procedures for constructing object boundaries from the corresponding pixel sequences.

One of the simplest approaches to connecting contour points is to analyze the characteristics of pixels in a small area (for example, in an area of 3X3 or 5X5) of each point  $(x, y)$  of an image that has already undergone a contour detection procedure. All points that are similar (the definition of the similarity criterion is given below) are connected, forming a boundary of pixels with some common properties.

In this analysis, to establish the similarity of the contour pixels, it is necessary to determine:

- the magnitude of the gradient required to build a contour pixel;
- direction of the gradient.

The first characteristic is denoted by the value  $G[f(x, y)]$ .

$$G[f(x, y)] = [G_x^2 + G_y^2]^{1/2} = \left[ \left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2 \right]^{1/2},$$

$$G[f(x, y)] \cong |G_x| + |G_y|. \quad (5)$$

Thus, the contour pixel with coordinates  $(x', y')$  is similar in magnitude in the previously defined neighborhood  $(x, y)$  of the pixel with coordinates  $(x, y)$ , if the inequality is true

$$|G[f(x, y)] - G[f(x', y')]| \leq T, \quad (6)$$

where  $T$  is the threshold value.

The direction of the gradient is set by the angle of the gradient vector defined in the equation

$$G[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix}, \quad (7)$$

$$\theta = \arctg \begin{bmatrix} G_x \\ G_y \end{bmatrix}, \quad (8)$$

where  $\theta$  is the angle (relative to the x-axis) along which the rate of change is most important. Then we can say that the pixel angle of the contour with coordinates  $(x', y')$  in some area  $(x, y)$  is similar to the pixel angle with coordinates  $(x, y)$  when performing the following inequality:

$$|\theta - \theta'| < A, \quad (9)$$

where  $A$  is the threshold value of the angle.

The direction of the contour at the point  $(x, y)$  is actually perpendicular to the direction of the gradient vector at this point. However, to compare the directions, the inequality gives equivalent results.

Based on these assumptions, we connect a point in some plane  $(x, y)$  with a pixel that has coordinates  $(x, y)$ , if the criteria for magnitude and direction met. Moving from pixel to pixel and representing each joining point as the center of the plane, the process repeat for each point of the image. A standard library procedure use to match light intensity levels to contour pixel sequences.

The goal is to determine the size of the rectangles with which you can build a quality image. Construction of such rectangles is carried out as a result of definition of strictly horizontal and vertical contours. The further process consists in connection of the segments of a contour separated by small intervals, and in association of separate short segments.

Consider the method of connecting boundary points by determining their location on the curve of a special kind. First, assuming that  $n$  points give on the x-plane of the image, it is required to find subsequences of points lying on straight lines. One possible solution is to construct all the lines passing through each pair of points, and then to find all the subsequences of points close to certain lines. The task associated with this procedure is to find the  $n(n-1)/2 \sim n^2$  lines and then make comparisons  $n(n-1)/2 \sim n^2$  of each point with all the lines. This process is time consuming from a computational point of view, except for the simplest applications.

This problem solve in another way, using the approach proposed by Hoag and called the Hague transformation. Consider the point  $(x_i, y_i)$  and the general equation of a straight line  $y = ax_i + b_i$ . There are an infinite number of lines passing through the point  $(x_i, y_i)$ , but they all satisfy the equations  $y = ax_i + b_i$ . at different values of  $a$  and  $b$ . However, if we write this equation in the form  $b = -ax_i + y_i$  and consider the plane  $ab$  (parameter space), then we have the equation of one line for a fixed pair of numbers, associated with a point  $(x_i, y_i)$  at a point  $(a', b')$ . Where the values of  $a'$  and  $b'$  are the parameters of the line on which the points  $(x_i, y_i)$  and points  $(x_j, y_j)$  are located in the  $xy$  plane. Virtually all points located on this line in the parameter space and have intersection lines at the point  $(a', b')$ .

The problem with representing a straight line by an equation  $y = ax_i + b_i$  is that both parameters  $a$  and  $b$  tend to infinity if the line assumes a vertical position. To eliminate this difficulty, a normal representation of a straight line in the form is used

$$x \cos \theta + y \sin \theta = \beta. \quad (10)$$

This representation is used to construct a table of collecting elements in the same way as the method described above, but instead of straight lines, we have sinusoidal curves in the plane  $\theta\rho$ . As before, the  $M$  points lying on the line  $x \cos \theta + y \sin \theta = \beta$  and correspond to the  $M$  sinusoidal curves that intersect at the point  $(\theta, \rho)$  of the parameter spaces. If the method of growth  $\theta$  is used and the corresponding  $\rho$  is found for it, the procedure gives  $M$  points in the collecting element  $A(i, j)$  connected with the point  $(\theta, \rho)$ .

Based on the above, a system of technical vision developed according with [6, 7].

The TVS consists of a webcam connected to a computer. Special software "Detector" reads the image from the camera and it digitize. The software uses artificial intelligence that based on a neural network. The artificial intelligence has been trained in 18,000 images of circles and cylindrical objects. The circle is determined in a few milliseconds. The coordinates of the center of the circle and the values of the radius transmit to the robot control unit. The obtained data adjust the position of the manipulator to capture the object. The robot captures a cylindrical object accord by algorithm that create and download early.

The general view of the program "Detector" show in Fig.4.

**Conclusions.** Registration of the presence of the object in the field of view of the sensor of external video information (video sensor); allowed to count the number of objects in the field of view or those that passed in front of the video sensor. Reading and decryption of labels, as well as the detection of obstacles to the movement of the robot allows you to calibrate the program code, as well as calibration of the geometric parameters of the manipulator, linking it to the coordinate system of the workspace. This allows us to measure the geometric and physical parameters of objects, to control the products for visible defects, completeness, and compliance with standards. Determining the position and orientation of objects in the work area of the robot depends on measuring the speed of moving objects and visual control of the correctness of the robot.

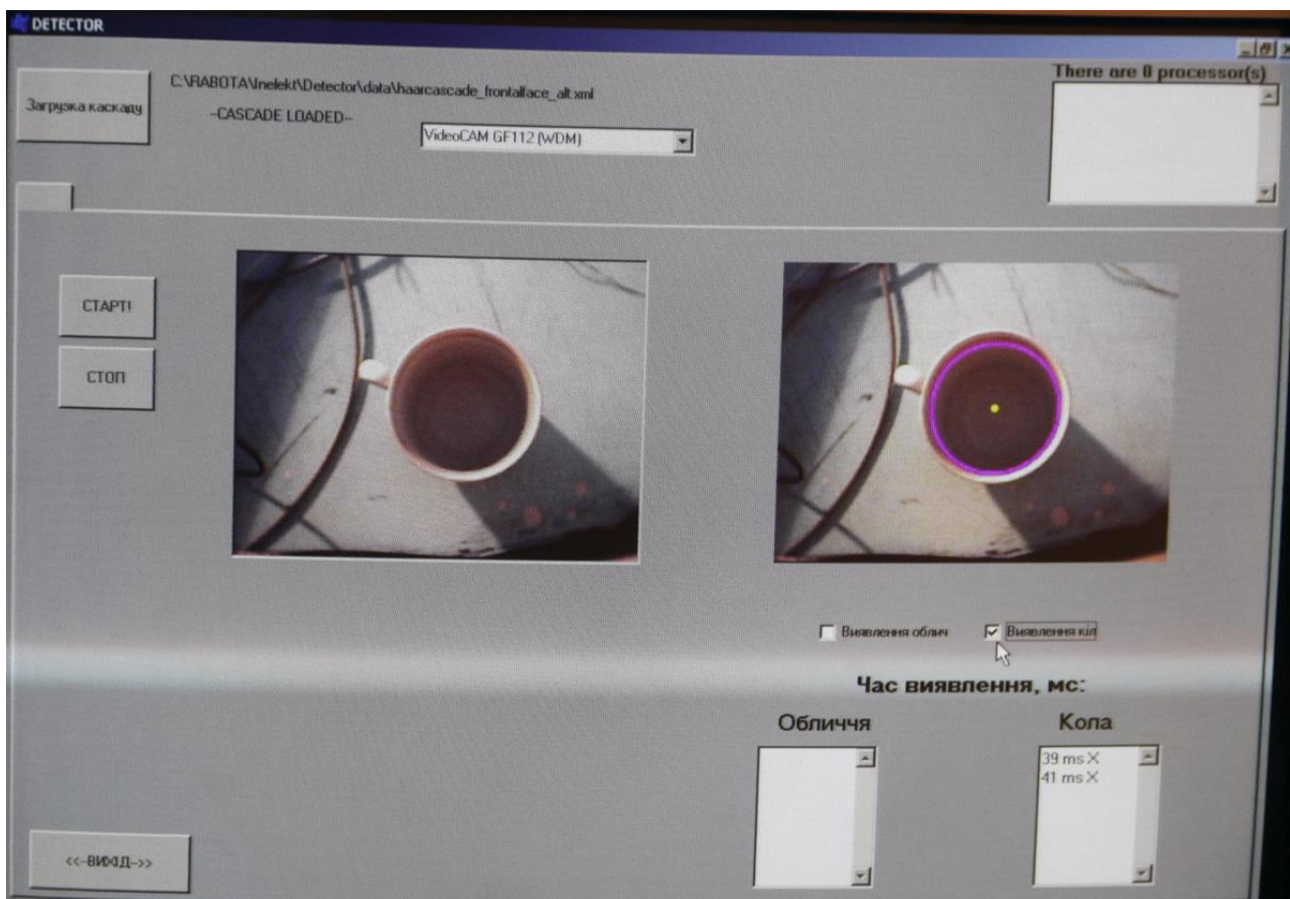


Fig. 4. Operation of the technical vision system program

The software is developed and experimentally tested in operation. The operation of the technical vision system experimentally tested in the work, namely the capture of a cylindrical object. Work vision systems worked experimentally and is passionate cylindrical object. The coordinates of the value of the circle, which are necessary for the exact location of the robot manipulator determined by artificial intelligence in 41 milliseconds. The obtained coordinates transmitted to the microprocessor to adjust the position of the manipulator. The robot accurately captured a cylindrical object.

#### References

1. Customize Your Ur Robot With Cutting-Edge Products [Electronic resource] // 1. – 2018. – Access mode: <https://www.universal-robots.com/plus/>.
2. Technical Data GRIPKIT [Electronic resource] // 1. – 2018. – Access mode: <https://www.weiss-robotics.com/gripkit/en/>.
3. Nanshu Lu. Flexible and Stretchable Electronics Paving the Way for Soft Robotics / Nanshu Lu, Dae-Hyeong Kim. // 1. – 2013. – P. 51.
4. DMOS Microstepping Driver with Translator And Overcurrent Protection, 2018. – (5). – (4988-DS; вып. 5).
5. ABBs 6 axis robot – for flexible and compact production [Electronic resource] // 1. – 2018. – Access mode: <https://new.abb.com/products/robotics/industrial-robots/irb-120>.
6. Horiashchenko S.L. Trends in the development of modern power element base for control of light industry devices [in Ukrainian] / S.L. Horiashchenko, K.L. Horiashchenko - Visnyk of Khmelnytsky National University, Technical sciences. - 2004, №.5, - p. 174-177
7. Horiashchenko K.L. The use of self-programmed systems in the design of electronic means [in Ukrainian] / Measuring and computing in technological processes, 2008, №.1, - p.194-196
8. Horiashchenko S. L. Development model of household robot/ S. L. Horiashchenko, V. V. Nadopta, D. M. Nesterchuk // Visnyk of Khmelnytsky National University. Technical sciences. - 2014. - №.1. - P. 50-52.

Надійшла / Paper received: 21.07.2020  
Надрукована / Paper Printed : 02.09.2020