OBJECT-ORIENTED RECONSTRUCTION OF A SCENE USING A POINT CLOUD
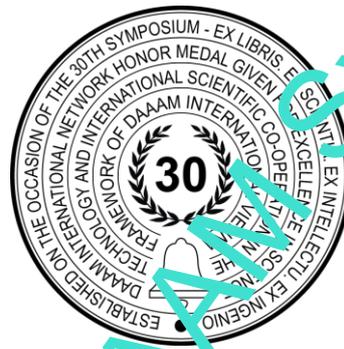
Anton Korsakov, Alexander Lopota & Ekaterina Smirnova

Abstract

Operations involving direct contact of the robot's gripping device with surrounding objects are among the most difficult in the field of robotics engineering. Therefore, minimizing the risk of tool damage and processing disruption is important. This can be achieved by using the information about the geometric parameters of the object in question and its position relative to the grip. This paper discusses approaches in reconstructing the surface of an object located in the working area of the manipulator using a point cloud, and also offers a method for describing the object by rotation surfaces using a noisy and incomplete point cloud obtained using a stereoscopic system installed on the final link of the manipulator.

Keywords: Computer Vision; Manipulator; Object-Oriented Segmentation; Object 3D modeling; Point Cloud.

1. Introduction

Operation of the robot manipulator when performing technological operations in automatic mode involves mechanical interaction with a variety of objects in the work area. At the same time, the risk of damage to the process tool and violation of the processing mode should be minimized. This is especially important when performing contact operations with an Autonomous mobile robot that operates under conditions of a priori uncertainty about the state of the work area and the nature of the object of interest. The correct positioning of the technological tool relative to the object of interest can be achieved using various sensors, but the most rational quality/cost criterion is the use of a television camera mounted on the effector (final link) of the manipulator.

It is known the solution of this particular problem by ensuring the safe interaction of the manipulator with underlying surface, for example, when taking a soil sample using a robotic manipulator complex. First the operator indicates the sampling point on the TV frame. Further, two or more images of the underlying surface in different positions of the manipulator are taken using a television camera mounted on the manipulator. The obtained images are used for stereo reconstruction of the underlying surface and determination of the metric coordinates of the sampling point using the software and algorithms developed by the authors. The three-dimensional coordinates of the sampling point and the normal vector obtained in this way are used as input data for solving the inverse kinematics problem, which results in generalized coordinates of the manipulator hinges at the soil sampling point [1].

The described approach assumes manipulator control in Supervisory mode, when the final decision on the operation is made by a human operator. The computer vision system in this case plays the role of an assistant to the operator, helping to correctly orient the manipulator effector equipped with a sampler relative to a flat surface.

A more detailed analysis of the working area is required to perform a wider range of manipulator operations with minimal operator involvement. Taking into account the technological capabilities of removable tools, it is necessary to determine the presence and location parameters at least three classes listed in table 1.

Object	Approximate object description	Processing tool
Pipe (extended, cylindrical)	Cylinder	Mechanical cutter
Rod (extended, thin)	Segment	Clipper
Chip, dangerous object (compact)	Sphere, cone	Grip of the manipulator

Table 1. Classes of objects defined in the manipulator's work area

Currently known methods for objects detecting can be subdivided in two classes: neuromorphic algorithms with pre-training and spatial-temporal filtering algorithms [2]. Use of neuromorphic algorithms is impractical due to lack of a database for training and deficiency computational resources of the robot's on-board control system.

In such a situation the object-oriented classification method should be selected for image processing. This method does not employ image pre-processing via a neural network. Segmentation is performed on the basis of both elementary (brightness and texture) features of the image as well as geometric shape of object. A similar approach has been successfully applied in solving the problem of improving the reliability of the mobile robot's visual navigation [2].

This paper describes an algorithm for reconstruction of the working area, which provides both classification of objects in accordance with table 1 for selecting a working tool, and obtaining their three-dimensional coordinates for solving the inverse problem of kinematics and control of the movement modes of the manipulator's Executive body.

2. Related works

By now a series of works addressed to development of universal methods for shape reconstruction have been published [3], [4]. So far the most common 3D modelling method in computer graphics was polygonal modelling. Delaunay triangulation algorithm [5], as well as the Ball-Pivoting Algorithm [6] have been widely spread. Based on input points the accuracy of modelling depends on the number of polygons. Polygonal modelling is characterized by its calculating speed. It is used in many automated designing systems and virtual area development due to its possibility to use many operations to define and edit objects geometry. Nevertheless, big polygonal network development requires the processing of growing amounts of data.

Voxels-based modelling is also in use. It is usually used for visualization of technical objects and medical data, terrain representation in computer games and simulations. Voxel representation takes up a lot of space due to each voxel containing a description of the whole area of space where the 3D object is located. To reduce the amounts of data, octrees are most common in use.

Application of NURBS allows the construction of smooth complex curvy surfaces with small amounts of control vertices which leads to reducing the amount of memory needed for shape storage. NURBS-based modelling is widely used in CAD [7], animation [8], [9], organic modelling, medicine. Authors [10], [11] have pointed out such obvious advantages of NURBS surface as smoothness and good adaptation to complex surfaces. Despite all these advantages, NURBS-models are difficult to edit and it takes long time to treat.

The aforementioned methods of polygonal modelling are successfully used to reconstruct surfaces from a dense point cloud, but the results are notably worse if the input points are partially missing. According to [12], these cases of acquiring the incomplete point cloud are common in machine-building plants due to restrictions in measuring positions.

The task of surface reconstruction under condition of incomplete point cloud is solved by finding solids of revolution that satisfy the equipment and engineering machinery [13]. In most cases, their cross-sections are circles formed by intersecting a plane which is perpendicular to the central axis. Therefore, it is possible to determine their geometric dimensions and to find the central axis by dividing the surface into cross-sections. By this technique planes and cylinders are drawn from point clouds during pipe reconstruction, and the shapes of the pipes are estimated [12], [14].

The technique of dividing the surface into segments was also successfully used by Stefan Bojarovski [15] who reconstructed a coral surface by approximating the segments with a cylindrical 3D template according to [16].

Approximate description of an object as a rotation shape allows to collect information about the geometric dimensions and the shape of the object using an incomplete point cloud, meaning that it is impossible to scan the object from each angle and collect a point cloud with points from the whole object's surface. That's why the method of dividing a point cloud into primitives can handle measurement errors and has a flexible shape (variability) in terms of 3D model construction that is, the ability to model objects of complex shape. In addition approximate description of an object with its shape does not require the construction of the surface as whole that allows to reduce computational resources utilized by on board control system of the robot. Therefore, this approach was taken by us as a basis for the 3D modelling method proposed in this paper.

3. Proposed solution

Stereo-reconstruction of the workspace is performed using two or more pictures of the workspace taken in several positions relative manipulator equipped with an attached camera. This resulted in a three-dimensional point cloud, every point of which has its metric coordinates in the on-board coordinate system of the manipulator.

Further, a pre-selection (rejection) of the points from the 3D point cloud is performed. The criteria are availability of the point considering the size of the manipulator, belonging of the point to the area the object of interest, and other spatial limitations. The result of the re-projection of the point cloud onto a television frame is illustrated in Figure 1.
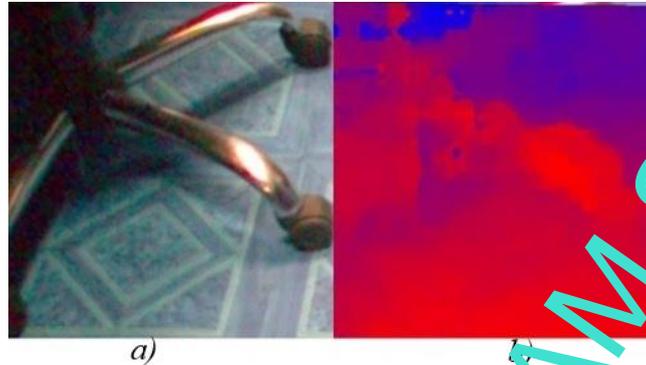

Fig. 1. Initial image (a) and the same image with the point cloud projected onto it - a depth map (b)

Considering the fact that an operator sets the area of interest, and the object is located in the foreground, it seemed to us more convenient to perform the segmentation of the point cloud obtained from the stereo pair based on the z-axis (remoteness) of the object.

For the points of the resulting point cloud, the following tasks are steadily solved:

- Formation the cross segments of the area of interest;
- Object segmentation the points of the area of interest based on z-axis;
- Calculating the center and the radius of the circumcircle of the cross segments;
- Constructing the central axis of the rotation shape;
- Calculating the radius of the rotation shape.

3.1. Formation of the cross segments of the area of interest

Initial data for the 3D modelling task consists of the point cloud in the selected area of interest, which contains the coordinates for observed points of the scene obtained by the stereo pair.

For each pixel comprising object's axis a linear equation is calculated for the lines perpendicular to the object's axis. The segment of the calculated line is formed using the width of the area in pixels (Figure 2).
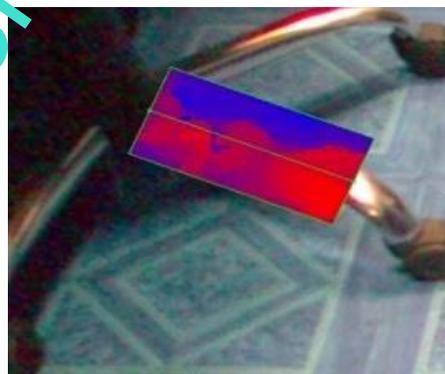

Fig. 2. Area of interest and point cloud points belonging to it

3.2. Object segmentation based on the z-axis

Z-axis-based segmentation of an object in the area of interest is performed according to the following list of actions repeated for each cross segment of the area of interest:

1. Calculating the mean z-coordinate for every pixel in the segment:

$$\bar{z} = \frac{1}{k+1}\sum_{i=0}^{k} P_{z,i}^{c,n} \tag{1}$$

where k - number of pixels in the segment, $P_{z,i}^{c,n}$- z-coordinate of the point in the point cloud corresponding to pixel i of segment n.

2. If $P_z^{c,n} > \bar{z}$ is true for a pixel, it is then removed from the set $\{P^{c,n}\}$ The value is recalculated using equation (1). The result of the calculation is illustrated by figure 3, a. The "distant" pixels are shown in navy blue, while the "close" pixels are shown in light blue.

3. Calculating the standard deviation (σ) for the z-coordinate of the "close" pixels for the area of interest looks like:

$$\sigma = \left(\frac{\sum_{i=0}^{k}(P_{z,i}^{c,n} - \bar{z})^2}{k+1}\right)^{1/2} \tag{2}$$

4. After that recalculation 1-3 is performed in order to guarantee that all points corresponding to both "distant" and "close" background are removed. The number of repetitions is chosen based on the fact that to pick out the closest out of N objects it is necessary to make N-1 repetitions.

5. If **Error! Reference source not found.** is true for a pixel, it is removed from the set $\{P^{c,n}\}$.

$$\left|P_{z,i}^{c,n} - \bar{z}\right| > 3 \cdot \sigma \tag{3}$$

6. Pixels are added to the set $\{P^{c,n}\}$ based on three-sigma criteria.
Moving from left to right:

$$(\cdot)(i,j) \in \{P^{c,n}\} \leftarrow \left|(i,j)_z - P_{z,i-1}^{c,n}\right| \le 3 \cdot \sigma \tag{4}$$

Moving from right to left:

$$(\cdot)(i,j) \in \{P^{c,n}\} \leftarrow \left|P_{z,i+1}^{c,n} - (i,j)_z\right| \le 3 \cdot \sigma \tag{5}$$

Pixels are added based on the smoothness of changing the distance of points on the surface of a smooth shape. Deleting points is performed in the longitudinal direction to cut off mistakenly added points. For example, as when part of an object and the underlying surface come into contact.

The result of the calculations according to 3-6 is illustrated in Figure 3, b. "Distant" pixels are shown in navy blue; pixels forming the set $\{P^{c,n}\}$ (from the calculations according to 5-6) are shown in yellow.

For the points of the resulting point cloud, the following tasks are steadily solved:

- The required density (T_d) of the pixels that passed the filtration procedure according to 1-6 is compared to all the points in the point cloud belonging to the longitudinal segment n of the object. $P_{cl}^{L,n}$;
- The required density (T_{de}) of the points in the point cloud belonging to the longitudinal segment n of the object $P_{sg}^{L,n}$ compared to the length of the longitudinal segment as whole.

$$\{P^{L,n}\}\neg\in \{P\} \leftarrow \left(\frac{\sum_i P_i^{L,n}}{P_{cl}^{L,n}} < T_d \ \cap \ \frac{P_{cl}^{L,n}}{P_{sg}^{L,n}} < T_{de}\right) \tag{6}$$

The result of the filtration is illustrated in Figure 3, c. "Distant" pixels shown in navy blue, and the segmented object is shown in red.
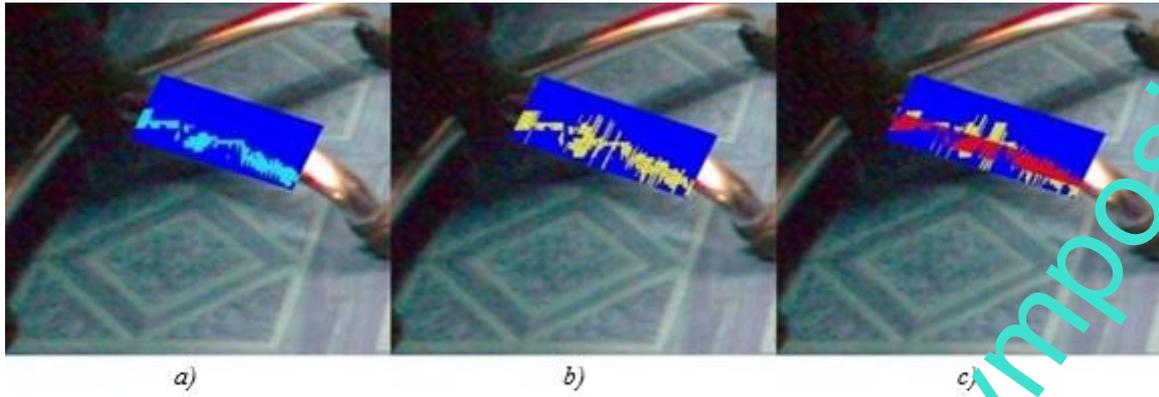
Fig. 3. Segmented object

3.3. Calculating the center and the radius of the circumcircle of the cross segments

Each cross-segment n of the object on the screen ($P^{c,n}$) corresponds to a plane in the space.
For each cross-segment $P^{c,n}$ ($n = 0.. P^{c,n}_{sg}$) the following list of actions is performed.

1. Detecting the base points for the circle construction:

- Detecting the first segmented point in the segment $S(S_x, S_y, S_z) = P^{c,n}_0$;
- Detecting the last point in the segment $F(F_x, F_y, F_z) = P^{c,n}_k$, where k - number of segmented pixels in the segment $P^{c,n}$;
- Detecting the middle point in the segment $C(C_x, C_y, C_z) = P^{c,n}_{k/2}$.

    If it is impossible to detect all three points, it is necessary proceed to the next segment.
2. Calculating the directional vectors of lines SC and CF.
3. Calculating the coefficients for the equation of plain SCF:

$$A_{SCF} \cdot x + B_{SCF} \cdot y + C_{SCF} \cdot z + D_{SCF} = 0 \ , \qquad (7)$$

    which is determined by points S, C and F.
4. Calculating the coefficients for the equations of the plains α and β, which are perpendicular to segments SC and CF and intersect them in their middle.
5. Calculating the coordinates of the intersection point (O) of the plains SCF, α, and β. The point O is the sought-out center of the circle and is found by solving the following system of linear equations:

$$\begin{cases} A_{SCF} \cdot x + B_{SCF} \cdot y + C_{SCF} \cdot z + D_{CF} = 0 \\ A_\alpha \cdot x + B_\alpha \cdot y + C_\alpha \cdot z + D_\alpha = 0 \\ A_\beta \cdot x + B_\beta \cdot y + C_\beta \cdot z + D_\beta = 0 \end{cases} \qquad (8)$$

    The radius of the circle corresponding to the cross segment n in the object is calculated using the following equation:

$$R^n = \left((O_x - S_x)^2 + (O_y - S_y)^2 + (O_z - S_z)^2\right)^{1/2} \ . \qquad (9)$$

3.4. Description the segmented object as a cylinder

Let $\{O\}$ be the set of points corresponding to the centers of the circles built from the set of points $\{P^c\}$.
To get straight line equations $y = ax + b$ for all the points from the set $\{O\}$, the method of least squares is used. Then the coefficients a and b of the equation can be found from the following equations:

$$\begin{cases} a = \left(\frac{\sum_{i=0}^{N-1} O_{x,i} \cdot O_{y,i}}{N} - \frac{\sum_{i=0}^{N-1} O_{x,i}}{N} \cdot \frac{\sum_{i=0}^{N-1} O_{y,i}}{N} \right) \times \\ \quad \times \left(\frac{\sum_{i=0}^{N-1} (O_{x,i})^2}{N} - \left(\frac{\sum_{i=0}^{N-1} O_{x,i}}{N} \right)^2 \right)^{1/2} \\ b = \frac{\sum_{i=0}^{N-1} O_{y,i}}{N} - a \cdot \frac{\sum_{i=0}^{N-1} O_{x,i}}{N} \end{cases} \tag{10}$$

where $(O_{x,i}, O_{y,i})$ - the coordinate of the point i from the set $\{O\}$, N - number of points in the set $\{O\}$.
Calculating the distance $O_{d,n}$ from each point O_n belonging to the set $\{O\}$ using the resulting linear equation:

$$O_{d,n} = \frac{|a \cdot O_{x,n} - O_{y,n} + b|}{(a^2 + 1)^{1/2}} \tag{11}$$

The mean distance (O_d) from each point $\{O\}$ to the line is calculated as follows:

$$O_d = \frac{1}{N} \cdot \sum_{i=0}^{N-1} O_{d,i} \tag{12}$$

In the case $O_{d,n} \geq O_d$ are true for this concrete a point, it is then removed from the set $\{O\}$. The same operations are then repeated for the projections of the points onto the plane Oxz.

Calculating the radius of the cylinder R is performed according to the following list of actions:

1. Calculating the mean radius of the circles constructed from $\{P\}$.

$$\bar{R} = \frac{1}{N} \cdot \sum_{i=0}^{N-1} R_i \tag{13}$$

where N - size of the set of all R^n ($\{R\}$) (9).

2. Calculating the mean linear deviation for each R:

$$\rho = \frac{1}{N} \sum_{i=0}^{N-1} |\bar{R} - R_i| \tag{14}$$

3. If $R_n > \rho$ is true for a point, it is removed from the set $\{R\}$.
4. Calculating the cylinder radius:

$$R = \frac{1}{N} \sum_{i=0}^{N-1} R_i \tag{15}$$

3.5. Description the segmented object as a sphere

There are two approaches to description an arbitrary object as a sphere. The first one consists in constructing the central axis and calculating the radiuses R_n for each resulting cross section of the object, that has been illustrated in the previous sections. The resulting values of radiuses are then circumscribed into the circle using the method of least squares.

The second approach is universal and can be used for objects of any shape. It is, in essence, an addition to the calculations from the previous sections. In place of the equation (15) the sought-out sphere radius is calculated as half of its axis' length:

$$R = \frac{1}{2} \left((O_{x,0} - O_{x,N-1})^2 + (O_{y,0} - O_{y,N-1})^2 + (O_{z,0} - O_{z,N-1})^2 \right)^{1/2} \tag{16}$$

where N - size of $\{O\}$.

4. Experimental results

A mathematical model described in section 3 is realized using the $C++$ programing language. The initial data consists of an image from the stereoscopic system and a previously filtered point cloud. To visualize the resulting mathematical model, a re-projection of a spatial shape onto the initial image is performed.

The re-projection of the bases of the constructed cylinders and a diametric cross sections on the constructed sphere are illustrated in Figure 4.
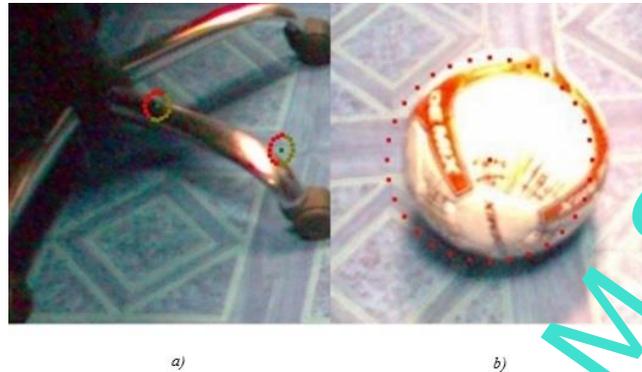

Fig. 4. Re-projection of the bases of the constructed cylinder (a) and a diametric cross section of the constructed sphere (a) onto the initial image

The experiment is repeated for wires illustrated in Figure 5.

Fig. 5. Initial image of the wires (a), the initial image with a point cloud projection (b), with a re-projection (c)

The numerical results of the experiments are shown in Table 2.

Object in question	Actual diameter, mm	Diameter obtained experimentally, mm	Error, %
Cylindrical chair leg	36	32	11
Soccer ball	216	192	11
Street wire	25	22	12

Table 2. Comparison of the actual dimensions of the objects with the ones obtained experimentally

5. Conclusion

In this paper, we propose an algorithm developed by the authors for approximate description of a point cloud collected from a stereoscopic system with a three-dimensional surface by dividing it into transverse segments. An approximate description of an object with a rotation shape using a point cloud allows you to get information about the geometric dimensions and shape of an object with a noisy and incomplete point cloud. This requires less computing power than is required for a robotic on-board control system. A mathematical description of the method is given and tested using a digital model. Using the proposed method, the dimensions of some objects located in the working space of the manipulator are obtained. Comparison of actual and calculated dimensions provided only a moderate match with an error of 11-12%, which can be explained by inaccuracies in the calibration of the stereoscopic system and the algorithm for calculating the centre and radius of the circle described in paragraph 3. Presumably, the error could be reduced if the calculations were performed using methods for estimating model parameters using sampling (the least squares method, the RANSAC

method). As a limitation of the method, we can note the approximation of an object by a rotation figure, which can cause significant errors when describing objects of complex shape. The further direction of research should be the description of objects of complex shape using modelling in parts.

6. Acknowledgments

7. References

[1] Gromoshinskii D.A. et al. (2018). Providing safe ground sampling inside the working zone of a manipulator with computer vision, Proceedings of the International Scientific and Technological Conference Extreme Robotics And Conversion Tendencies, Saint-Petersburg, pp. 185-194, 2018

[2] Smirnova E.Y., Stepanov D.N. and Goryunov V.V. (2016). A technique of natural visual landmarks detection and description for mobile robot cognitive navigation, Proceedings of the 26th International DAAAM Symposium 2016, DAAAM, pp. 0905-0911, 2016

[3] Alliez P. et al. (2007). Voronoi-based Variational Reconstruction of Unoriented Point Sets, ACM Trans. Graph., Michael Garland, pp. 39-48, 2007

[4] Kazhdan M. and Hoppe H. (2013). Screened poisson surface reconstruction, ACM Trans. Graph. 2013, 32, № 3, 2013

[5] De Loera, Jesús A.; Rambau, Jörg; Santos, Francisco. (2010). Triangulations, Structures for Algorithms and Applications, Algorithms and Computation in Mathematics. 25. Springer.2, 2010

[6] Bernardini F. et al. (1999). The Ball-Pivoting Algorithm for Surface Reconstruction, EEE Transactions On Visualization And Computer Graphics, 5, 4, 1999

[7] Behshad A. and Ghasemil M.R. (2014). Nurbs-iga-based modelling: Analysis and optimization of laminated plates, Tehnički vjesnik, 21, 4(2014), pp. 789-797. 2014

[8] Cai J. et al. (2016). Skeletal Animation with Anisotropic Materials, Graphical Simulation of Deformable Models. Springer International Publishing, pp. 85-104, 2016

[9] Yu L., Yu J. and Wang Z. (2017). A realistic 3D articulatory animation system for emotional visual pronunciation, Multimed. Tools Appl. Springer New York LLC, 76, №18, pp. 19241-19262, 2017

[10] Xiong H. et al. (2015). A method for accurate reconstructions of the upper airway using magnetic resonance images, PLoS One, 10, №6, pp. 1-14, 2015

[11] Galassi F. et al. (2018). 3D reconstruction of coronary arteries from 2D angiographic projections using nonuniform rational basis splines (NURBS) for accurate modelling of coronary stenoses, PLoS One, 13, №1, pp. 1-23, 2018

[12] Midorikawa Y. and Masuda H. (2018). Extraction of surfaces using section curves for engineering plants, Proceedings of the ASME Design Engineering Technical Conference, 1B, 2018

[13] Schnabel, R., Wahl, R. and Klein, R. (2007). Efficient RANSAC for Point-Cloud Shape Detection, Computer Graphics Forum, 26, pp. 214-226, 2007

[14] Kawashima K., Kanai S. and Date H. (2014). As-built modelling of piping system from terrestrial laser-scanned point clouds using normal-based region growing, J. Comput. Des. Eng. Elsevier Masson SAS, 1, №1, pp. 13-26, 2014

[15] Bojarovski S. (2014). Robust tracking-based skeleton reconstruction of cold-water corals from computer tomography images, (in Russian: Saint-Petersburg State University), 2014

[16] Yureidini A., Kerrien E. and Cotin S. (2012). Robust RANSAC-based blood vessel segmentation, SPIE Medical Imaging, 2012