ATLANTIS PRESS
Advances in Intelligent Systems Research, volume 166
7th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2019)

The New Combined Closed-Solution for 3D Reconstruction of Environment Based on Iterative Closest Point Algorithm

Aleksander Vokhmintcev
Intellectual information technologies and systems
Research laboratory
Chelyabinsk State University
Chelyabinsk, Russia
vav@csu.ru

Andrey Melnikov
Direction of Institute
Ugra Research Institute of Information Technologies
Khanty-Mansiysk, RussiaCity
melnikovav@uriit.ru

Stepan Pachganov
Information and analytical systems cente
Ugra Research Institute of Information Technologies
Khanty-Mansiysk, Russia
s.pachganov@yandex.ru

Vladimir Burlutskii
Digital economy Institute
Yugra State University
Khanty-Mansiysk, Russia
burlutskyvv@uriit.ru

***Abstract*—This The scientific problem at solving which the present project is directed consists in the development of accurate methods for reconstruction of a three-dimensional map of the accessible of environment with requied accuracy of reconstruction. The problem of consistent aligning of 3D point data is known registration task. The most popular registration algorithm is the Iterative Closest Point algorithm. Three basic problems are characteristic for the ICP algorithm: first, the convergence of the algorithm depends strongly on the choice of the initial approximation; second, the algorithm does not take into account the local shape of the surface around each point; and, third, the search for the nearest points is of high computational complexity. In this paper a new close solutions to 3D total variation regularization will be obtained and effective algorithms for restoring 3D data will be designed. The proposed approach improves the accuracy and convergence of reconstruction methods for complex and large-scale scenes. The performance and computational complexity of the proposed RGB-D Mapping algorithm in real indoor environments is discussed.**

Keywords—simultaneous localization and mapping; a three-dimensional map; iterative closest point algorithm; orthogonal transformations; variational problem of the point-to-point

I. INTRODUCTION

Over the last decade numerous methods of SLAM (Simultaneous Location and Mapping) have been suggested. The development of an adaptive dynamic system for the reliable handling of the task of the simultaneous localization of a mobile robot and mapping of its environment in real time is one of key tasks in modern robotics and machine vision, since the creation of autonomous intelligent robotic complexes and systems is based on its fulfillment.In the field of robotics and systems of autonomous localization in unknown space over the past decade many successful methods, approaches and algorithms for building dense three-dimensional maps of the unknown envifronment developed. Historically, quick methods of handling SLAM tasks and methods that use orienting points for localization in unknown environment were the first to be worked out [1]. Subsequently, the main direction of investigations was connected with the development of intellectual SLAM methods that use various multi-touch sensors and variants of Kalman filtering for estimating the motion trajectory of a robotic system in space in order to tackle the task [2]. The known algorithms for SLAM are based on the following approaches: Particle Filter SLAM, Extended Kalman Filter (EKF), Graph-Based SLAM and Visual SLAM [3]. In addition to these approaches of the localization the following methods can be used: extreme methods of navigation based on comparison of two consecutive scans by optimizing the cross-correlation function; the method of recursive filtering; method Normal Distributions Transform is a transformation of normal distributions. Some perspective SLAM methods, apart from information about the geometrical shape of objects need to account for information about semantic properties of objects in three-dimensional space, as well as relations between them.

The performance of SLAM algorithms directly depends on the accuracy of constructing a 3D shape model by registration in the 3D point clouds obtained from depth sensors. The Iterative Closest Point (ICP) algorithm is often used as a registration algorithm [4]. A common registration algorithm solves the variational problem of finding the optimum geometric (orthogonal or affine) transformation, which best matches two clouds of points with a given correspondence between points [5]. Different functionals lead to various methods for registration of point clouds. The most commonly used methods for searching correspondence between a pair of clouds are point-to-point and point-to-plane. For the class of orthogonal transformations, the solution of the point-to-point problem was explicitly given by of Horn [6]. An exact solution of the point-to-point task is given in [7]. The variational problem of the point-to-point in the class of orthogonal transformations is usually solved using either the Levenberg-Marquardt iteration algorithm or the linearization method for small angles [8]. For the

variational problem of the point-plane for the class of affine transformations, the exact solution was proposed [9]. An approximate solution of the point-plane problem for the class of orthogonal transformations was obtained in [10]. Note that the point-plane method for the class of orthogonal transformations is more robust to noise of sensors, but for this method the explicit solution of the problem has not been found yet. This complicates the use of the method in real-time registration applications. Recently, an algorithm for dynamic registration of point clouds of deformable surfaces was proposed [11]. The method splits surfaces into sections, next each of them is processed separately, and then the results are automatically combined by minimizing a functional.

Within the framework of this project an explicit solution of the point-plane problem for the class of orthogonal transformations will be found, and aniterative registration algorithm will be designed. The ICP algorithm is characterized by two main problems: first, the algorithm does not use the local surface shape around each point, and secondly, search of nearby points has a high computational complexity [4,12]. In this respect, it is important that a shift between successive frames might be inconsiderable, all previous formations might be accurate, and all objects in a frame might be within the area of the field of vision, otherwise, the ICP method may either yield unsatisfactory results in terms of accuracy, or result in no convergence altogether [13]. In papers [14] instead of matching point to point is used the metric point to the plane. However, the approach based on point-to-plane has more constraints of the structure of the reconstructed space. The ICP algorithm (point-to-plane) has a poor convergence for scenes with a small number of geometric constraints [15]. The convergence of the ICP algorithm may be considerably improved [16]. We suggest to apply visual (color) features to significantly improve the initial point of the ICP algorithm, and the alignment between key-frames is computed by joint optimization of appearance and shape matches. [17]. The most successful feature descriptors such as SIFT (Scale Invariant Feature Transform) [18] and SURF (Speeded-Up Robust Features) [19] are robust to scale and viewpoint changes. However, other visual features could be also useful for loop-closure detection and global localization. For example, histograms of oriented gradients (HOGs) possess attractive invariance to viewpoint changes [20].

The paper is organized as follows. In Section 2, the proposed fusion close-solution algorithms for restoring 3D data are presented. Computer simulation results are provided in Section 3. Section 4 summarizes our conclusions.

II. Fusion Iterative Closest Point Algorithm for Orthogonal Transformations

In this topic, a new iterative algorithm for registration of point clouds based on the point-to-plane ICP algorithm with arbitrary affine transformations wil be presented.

Let $X = \{x_1, \dots, x_n\}$ be a source RGB-D frame, and $Y = \{y, \dots, y_m\}$ be a destination RGB-D frame in $\mathbb{R}3$. The ICP algorithm consists of the six stpes. The main step is search of a geometrical transformation between X and Y (RGB-D objects mapping):

$$Rx_i + T, \tag{1}$$

where R is a rotation matrix, t is a translation vector, $i = 1, \dots, n$.

Denote by (Y) a surface constructed from the cloud Y, by ( ) denote a tangent plane of (Y) at point y . Let (R, T) be the following function:

$$J(R) = \sum_{i=1}^{n} (\langle Rx^i - y^i n^i \rangle)^2, \tag{2}$$

where <,> is the scalar product, R is a fusion matrix of an affine-transformation: rotation and translation. Let x^i is a point from the cloud Y, n^i is the unitary normal for $T(y^i)$:

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, x^i = \begin{pmatrix} x_1^i \\ x_2^i \\ x_3^i \\ 1 \end{pmatrix}, \ n^i = \begin{pmatrix} n_1^i \\ n_2^i \\ n_3^i \\ 1 \end{pmatrix}. \tag{3}$$

Then new project solutions to 3D total variation regularization will be obtained. The function J(R) can be rewritten:

$$J(R) = \sum_{i=1}^{n} (\langle Rx^i - y^i, n^i \rangle)^2 = \sum_{i=1}^{n} (\langle Rx^i, n^i \rangle - \langle y^i, n^i \rangle)^2 = \sum_{i=1}^{n} (\langle Rx^i, n^i \rangle)^2 - 2\langle Rx^i, n^i \rangle \langle y^i, n^i \rangle + (\langle y^i, n^i \rangle)^2. \tag{4}$$

The term $<y^i, n^i>$ = const with respect to R, so the function J(R) (4) takes the form (5):

$$\sum_{i=1}^{n} \left(\langle Rx^i, n^i \rangle\right)^2 - 2\langle Rx^i, n^i \rangle \langle y^i, n^i \rangle + \left(\langle y^i, n^i \rangle\right)^2 \longrightarrow \min_R. \tag{5}$$

Let's solve the variational problem J(R) with respect to t_{ij}, i=1,2,3.

$$\frac{\partial J}{\partial t_j} = \sum_{i=1}^{n} 2\langle Rx^i, n^i \rangle n_j^i - 2n_j^i \langle y^i, n^i \rangle = 0, \tag{6}$$

where j=1,2,3.

Then

$$\sum_{i=1}^{n} \langle Rx^i, n^i \rangle n_j^i = c_j, j = 1, \dots, 3. \tag{7}$$

Denote by the $(XN)^i$ the following matrix:

$$(XN)^i = \begin{pmatrix} x_1^i n_1^i & x_1^i n_2^i & x_1^i n_3^i & 0 \\ x_2^i n_1^i & x_2^i n_2^i & x_2^i n_3^i & 0 \\ x_3^i n_1^i & x_3^i n_2^i & x_3^i n_3^i & 0 \\ n_1^i & n_2^i & n_3^i & 0 \end{pmatrix}. \tag{8}$$

Let us consider the following expression:

$$\langle Rx^i, n^i \rangle = Tr(R \cdot (XN)^i). \tag{9}$$

Taking into account (7) and (9) can be rewritten as

$$\sum_{i=1}^{n} n_j^i Tr(R \cdot (XN)^i) = c_j, j = 1,2,3. \quad (10)$$

Denote by $(n_jXN)^i$ the following matrices. Denote that

$$n_j^i Tr(R \cdot (XN)^i) = Tr\left(R \cdot \left(n_j XN\right)^i\right). \quad (11)$$

Taking into account (10) and (11) can be rewritten as

$$\sum_{i=1}^{n} Tr\left(R \cdot \left(n_j XN\right)^i\right) = c_j. \quad (12)$$

Denote that

$$\sum_{i=1}^{n} Tr\left(R \cdot \left(n_j XN\right)^i\right) = Tr\left(\sum_{i=1}^{n} R \cdot \left(n_j XN\right)^i\right) = Tr\left(R\left(\sum_{i=1}^{n}\left(n_j XN\right)^i\right)\right) = c_j. \quad (13)$$

Let's solve the variational problem J(R) with respect to r_{ij}, j=1,2,3:

$$\frac{\partial J}{\partial r_{ij}} = \sum_{k=1}^{n} 2\langle Rx^k, n^k\rangle x_j^k n_i^k - 2\, x_j^k n_i^k \langle y^k, n^k\rangle = 0, \quad (14)$$

where j=1,2,3, k=1,…,n.

Then

$$\sum_{i=1}^{n}\langle Rx^k, n^k\rangle x_j^k n_i^k - x_j^k n_i^k \langle y^k, n^k\rangle = \sum_{i=1}^{n}\langle Rx^k, n^k\rangle x_j^k n_i^k - \sum_{i=1}^{n} x_j^k n_i^k \langle y^k, n^k\rangle = 0. \quad (15)$$

Taking into account the expression $\sum_{k=1}^{n} x_j^k n_i^k \langle y^k, n^k\rangle$ is constant (15) can be rewritten as

$$\sum_{k=1}^{n} x_j^k n_i^k Tr(R \cdot (XN)^k) = c_{ij}. \quad (16)$$

Denote the set of matrices $x_j^k n_i^k XN^k$. The expression (16) can be rewritten as

$$\sum_{k=1}^{n} Tr\left(R \cdot x_j^k n_i^k XN^k\right) = c_{ij}. \quad (17)$$

We get the set of 12 linear equations with 12 variables r_{ij}, t_k, i,j,k=1,2,3. We can write this system as the matrix equality:

$$Z \cdot \mathrm{P} = C, \quad (18)$$

where matrix Z has size 9 × 12, C is the column consist of 12 numbers c_i and c_{ij}, and the column H: $\mathrm{P} = (r_{11}, r_{12}, r_{13}, t_1, r_{21},\ r, r_{23}, t_2, r_{31}, r_{32}, r, t_3)$.

III. Robot Mapping Algorithm

In this project we use approach to the implementation of the task is based on the application of EKF. We use the term “visual feature” and term “semantic landmark” synonymously. The robot platform is equipped with multi sensor cameras (Kinect and Beward camera) which take measurements of the relative location between any semantic landmark and the robot platform itself. The absolute locations of the semantic tags are not available. The implementation of the suggested visual SLAM algorithm uses nonlinear kinetic models of robot platform and nonlinear asynchronous observation models in controlled/uncontrolled indoor conditions). The state of the system of interest consists of the position and orientation of the robot platform together with the position of all semantic tags. General algorithm of robot mapping is shown in Fig.1.

The updating of a history of camera positions and the robot's movements, as well as the configuration of surveyed characteristics are carried out at a certain interval of time. As a rule, data are added not at every step, but in case the position of the camera (the robot) has significantly changed, for example, when a considerable displacement or rotation has taken place as compared to the previous memorized position. A mathematical model of the scene will be saved in the form of a graph the vertices of which correspond to certain moments of time [21]. A new combined adaptive method was developed in the present work for the generating a three-dimensional combined dense map of an accessible environment and determining the position of the robot in a relative coordinate system based on the history of camera positions, on symbolic (semantic) tags, on the robot's movements and on the matching of obtained three-dimensional depth maps which account for the accuracy of their superimposition, as well as geometric relationships between various images of one scene[22].

The variational problem can be solved numerically by various iterative methods. The matching between pairs of points from two clouds is determined by finding geometrically homothetic elements of objects. Such matching is invariant under rotation, rescaling, and parallel translation. A comparative analysis of exact and approximate solutions of the variational problem was performed using reference databases. Specifically, experiments were conducted using the image databases ALOI (Amsterdam Library of Object Images) and Indoor Segmentation and Support Inference (ECCV) [23]. It was found that the approximate solution methods yield fairly accurate results when applied to the reconstruction of three-dimensional environmental scenes. Thus, we can conclude that visually associated characteristics used for solving the variational problem in the ICP algorithm make it possible to obtain a closed-form solution of the variational registration problem for a given three-dimensional scene.

A comparative analysis of the reliability of detection and the accuracy of localization of geometrically distorted objects in a scene was performed for linear and nonlinear detectors, descriptor methods, and the present method [4, 13]. More specifically, the accuracy of three-dimensional reconstruction was studied as a function of the number and spatial distribution of singular points in a frame. The minimum number of singular points ensuring prescribed accuracy of threedimensional reconstruction was determined for each studied descriptor/method. Guaranteed estimates for the accuracy and computational complexity of the proposed method were substantiated using an adaptive approach. It was found that the reconstruction accuracy is a nonlinear function of the number of singular points in a frame; for descriptors of all types, this function has a single pronounced peak. It was shown in [24] that the registration method proposed for three-dimensional objects is the best in terms of missing errors and false alarms for a wide range of white noise distortions (from 50 to 10 dB).

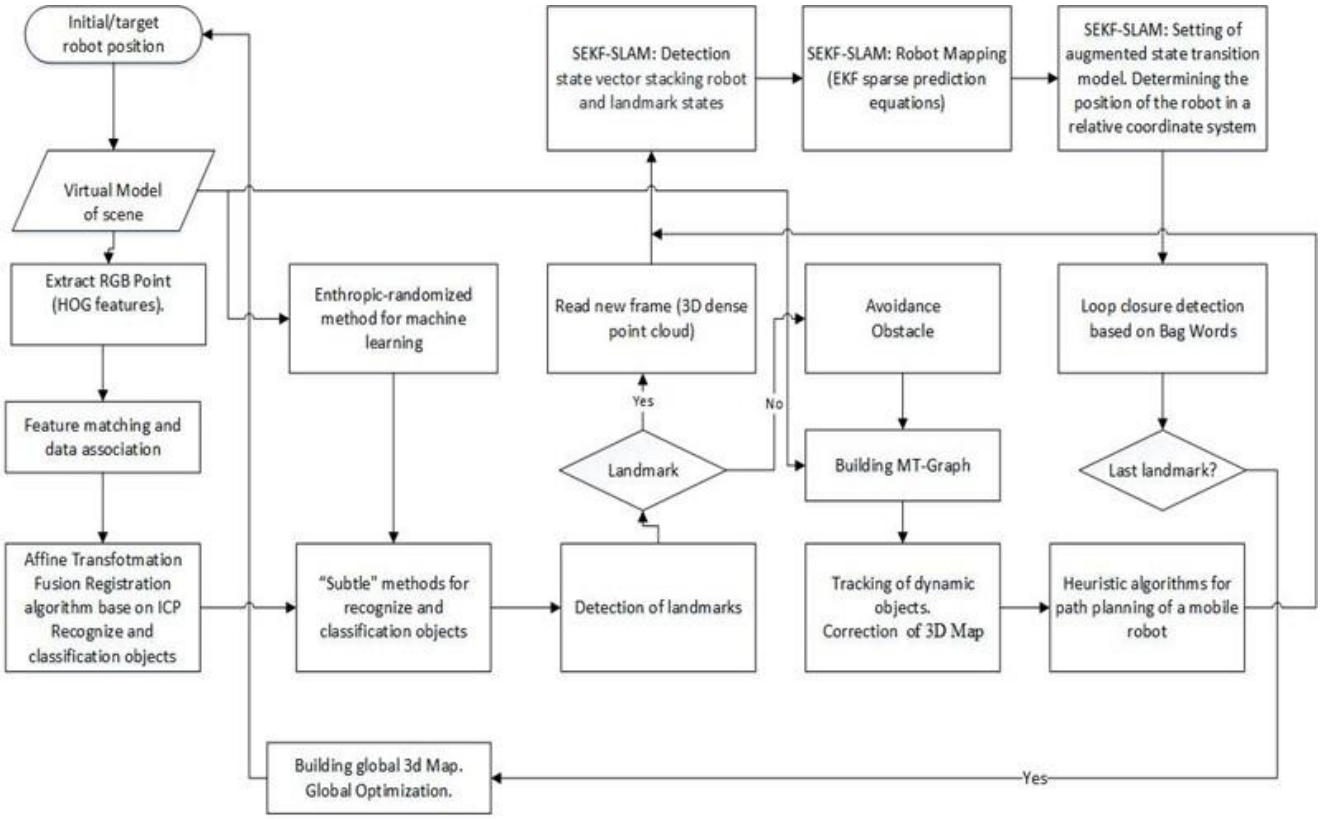

Fig. 1. The block diagram of the proposed robot mapping algorithm

The basic shortcomings of the classical ICP method are associated with its restricted convergence domain and high computational complexity. It was shown in [16] that projection methods can reduce the computational complexity of the ICP registration method from O(NS log (NT)) for the ICP method with a k-D tree to O(NS) for the ICP method with a spherical or triangular constraint [25]. The computational complexity of the registration method can be estimated as , where k is the number of steps in the registration algorithm, O1 is the computational complexity of the first step of the algorithm, and F is a parameter determining a partition of the cloud of points into elementary units containing raster element values in three-dimensional space. The results suggest that the proposed method is an efficient tool for computing closed-form solutions of the given variational problem and can be used in real time.

IV. Conclusions

In this paper, a effective algorithm for restoring 3D data combining visual features and depth information was proposed. The results were discussed and compared based on the results of computer simulation. The proposed approach is based on visual slam based on ICP is in line with the basic trends of development of modern methods and algorithms of simultaneous localization and mapping in an unknown space. In the proposed formulation of the problem we plan to find a solution of variational problem based on the combination of data on feature points (the color scenes), and a dense threedimensional point cloud (depth data). The considered functionals are composed of terms that measure the average of the squares of the distances to visial-associated characteristic points with the normalizing factor (the variation of metric characteristics of a function of two variables), terms that measure the average of the squares of the distances for a dense point cloud based on metrics "point-to-plane" and terms that presenet various generalizations of such functionals. The solution of the variational problem will be obtained using various iterative methods. Within the framework of the project, the variational problem of the ICP algorithm is extended to a class of affine transformations. The computational experiments carried out showed a significant increase in the accuracy of the modified ICP algorithm.

Acknowledgment

This work was supported by the RFBR (project no. 18-37-20032).

References

[1] C. Hertzberg, R. Wagner, and O. Birbach, "Experiences in building a visual slam system from open source components," Proc. 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, pp. 2644-2651, May 2011.

[2] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments," Proc. 12th International Symposium on Experimental Robotics, Delhi, India, vol.79. pp. 477-491, December 2014.

[3] F. Endres, J. Hess, N. Engelhard, and J. Sturm "An evaluation of the RGB-D SLAM system," Proc. IEEE International Conference on Robotics and Automation, Taipei, Taiwan, pp. 1691-1696, October 2012.

[4] P. Besl, and N. McKay "A method for registration of 3D shapes," Journal of IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.14(2), pp. 239-256, March 1992.

(https://www.researchgate.net/publication/3191994_A_method_for_registration_of_3-D_shapes_IEEE_Trans_Pattern_Anal_Mach_Intell)

[5] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse "MonoSLAM: Real-Time single camera SLAM," Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.7, pp. 1052-1067, June 2007. (https://dl.acm.org/citation.cfm?id=1263144.1263479)

[6] B.K.P. Horn, H.M. Hilden, and S. Negahdaripour "Closed-form solution of absolute orientation using orthonormal matrices," Journal of the Optical Society of America, vol. 5 (7), pp. 1127-1135, 1988.

(https://www.osapublishing.org/josaa/abstract.cfm?uri=josaa-5-7-1127)

[7] G.K. Tam, Z.Q. Cheng, Y.K. Lai, F.C. Langbein, Y. Liu, D. Marshall, R.R. Martin, X.F. Sun, and P.L. Rosin "Registration of 3D point clouds and meshes: A survey from rigid to nonrigid," Journal of IEEE Transactions on Visualization and Computer Graphics, vol. 19 (7), pp. 1199–1217, July 2013. (https://ieeexplore.ieee.org/document/6361384)

[8] K. L. Low "Linear least-squares optimization for point-to-plane ICP surface registration," Technical Report TR04-004, Department of Computer Science, University of North Carolina at Chapel Hill, 2004.

[9] S. Du, N. Zheng, S. Ying, and J. Liu "Affine iterative closest point algorithm for point set registration," Journal of Pattern Recognition Letters, vol. 31(9), pp. 791–799, July 2010.

(https://dl.acm.org/citation.cfm?id=1774990)

[10] K. Khoshelham "Closed-form solutions for estimating a rigid motion from plane correspondences extracted from point clouds," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 114, pp. 78–91, April 2016.

(https://www.sciencedirect.com/science/article/abs/pii/S0924271616000253?via%3Dihub)

[11] S. Cheng, I. Marras, and S. Zafeiriou "Active nonrigid ICP algorithm," Proc. 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, Ljubljana, Slovenia, pp. 1- 8, May 2015.

[12] Y. Chen, and G. Medioni "Object modeling by registration of multiple range images," Proc. IEEE International Conference on Robotics and Automation, Sacramento, USA, vol. 10(3), pp. 2724-2729, April 1991.

[13] C. Langis, M. Greenspan, and G. Godin "The parallel Iterative closest point algorithm," Proc. IEEE 3-rd International Conference 3D Digital Imaging and Modeling, Quebec City, Canada, pp. 195-204, May 2001.

[14] S. Thrun, W. Burgard, and D. Fox "A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping," Proc. IEEE International Conference on Robotics and Automation, San Francisco, USA, pp. 321–328, April 2000.

[15] A. Segal, D. Haehnel, and S. Thrun "Generalized-ICP," Proc. of Robotics: Science and Systems, Seattle, USA, vol. 2(4), pp. 645-653, June 2009.

[16] S. Rusinkiewicz, and M. Levoy "Efficient variants of the Iterative closest pont algorithm," Proc. Int. Conference on 3D Digital Imaging and Modeling, Quebec City, Canada, pp.145–152, May-June 2001.

[17] M. Fischler, and R. Bolles "Random Sample Consensus: A Paradigm for model fitting with applications to image Analysis and automated cartography," Journal of Graphics and Image Processing, pp. 381–395, June 1981. (https://dl.acm.org/citation.cfm?doid=358669.358692)

[18] D. G. Lowe "Object recognition from local scale invariant features," Proc. 7th International conference on Computer Vision, Kerkyra, Greece, vol. 2, pp. 1150–1157, September 1999.

[19] D. Miramontes-Jaramillo, V. Kober, V. H. Diaz-Ramirez, and V. Karnaukhov "A novel image matching algorithm based on sliding histograms of oriented gradients," Journal of Communications Technology and Electronics, vol. 59(12), pp. 1446–1450, December 2014. (https://link.springer.com/article/10.1134/S1064226914120146)

[20] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool "SURF: speeded up robust features," Journal of Computer Vision and Image Understanding, vol. 110, pp. 346–359, June 2008. (https://www.sciencedirect.com/science/article/pii/S1077314207001555)

[21] A. Vokhmintsev, and K. Yakovlev "A real-time algorithm for mobile robot mapping based on rotation invariant descriptors and ICP," Proc. 5th Analysis of Images, Social Networks and Texts, Communications in Computer and Information Science, Yekaterinburg, Russia, vol. 661, pp. 357-369, April 2017.

[22] A. Vokmintsev, T. Botova, I. Sochenkov, A. Sochenkova, and A. Makovetskii "Robot mapping algorithm based on kalman filtering and symbolic tags," in Proc. of the SPIE, Applications of Digital Image Processing XL, San Diego, United States, vol. 10396, pp. 103962I, August 2017.

[23] A. Vokmintsev, M. Timchenko, A. Melnikov, A. Kozko, and A. Makovetskii "Robot path planning algorithm based on symbolic tags in dynamic environment," in Proc. of the SPIE, Applications of Digital Image Processing XL, San Diogo, United States, vol. 10396, pp. 103962E, August 2017.

[24] A. Vokhmintcev, I. Sochenkov, V. Kuznetsov, and D. Tikhonkikh "Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram," Journal of Doklady Mathematics, vol. 93(1), pp. 37-41, January 2016. (https://link.springer.com/article/10.1134/S1064562416010178)

[25] A. V. Vokhmintcev, A. V. Melnikov, K. V. Mironov, and V. V. Burlutskiy "Reconstruction of three-dimensional map based on closed form solution of variational problem of multi-sensor data registration," Journal of Doklady Mathematics, vol. 99(1), pp. 108–112., April 2019. (https://link.springer.com/journal/11472)