Nonlinear Dynamics and Neuroscience

Izvestiya Vysshikh Uchebnykh Zavedeniy. Applied Nonlinear Dynamics. 2024;32(5)

Article DOI: 10.18500/0869-6632-003107

Biomorphic navigation system version*

Y. A. Malichev[1] ✉, *V. G. Yakhno*[1,2]

[1]National Research Lobachevsky State University of Nizhny Novgorod, Russia
[2]Federal Research Center A. V. Gaponov-Grekhov Institute of Applied Physics of the RAS, Nizhny Novgorod, Russia
E-mail: ✉smmalisheva@gmail.com, yakhno@ipfran.ru

Received 7.11.2023, accepted 13.02.2024, available online 27.05.2024, published 30.09.2024

Abstract. The *purpose* of this work is to create and study the dynamics of the functioning of a biorelevant visual navigation system. *Methods.* The work uses simultaneous navigation and mapping systems RatSLAM and Orb-SLAM. The RatSLAM system is a biorelevant model of visual navigation in the rodent hippocampus. The Orb-SLAM system is a simultaneous navigation and mapping system that works on the principle of searching and tracking changes in the position of key points in the image. *Results.* The article presents a version of a modified visual navigation system. The system consists of a visual odometry module based on the Orb-SLAM system, as well as a mapping and loop closure module based on the RatSLAM system. This allows you to combine the localization accuracy of systems operating on the principle of tracking key points in the image and neural filtering of biorelevant systems. Using the constructed system, location estimates were obtained on public and new data sets. *Conclusion.* The constructed visual navigation system determines the location of the subject (video camera) in space, which is in good agreement with the ground truth location data.

Keywords: simultaneous localization and mapping systems, path integration, landmarks, neural networks.

Acknowledgements. The study has been supported in the frames of the Governmental Project of the Institute of Applied Physics RAS (Project #FFUF-2024-0037).

For citation: Malichev YA, Yakhno VG. Biomorphic navigation system version. Izvestiya VUZ. Applied Nonlinear Dynamics. 2024;32(5):606–624. DOI: 10.18500/0869-6632-003107

Introduction

The ability to navigate in space and determine the position of things is important for the survival of humans and animals. In recent decades, research in the field of spatial perception has been particularly successful and has become a subject of interest in neurobiology [1–10]. These studies have allowed us to understand some of the mechanisms used by animals in the navigation process and have identified a set of cell types responsible for processing spatial information. This

*The paper presents materials of a talk given at the conference "Nonlinear dynamics in cognitive research — 2023".

allows us to come closer to understanding the functioning of the neural networks underlying this fundamental cognitive ability.

To study the mechanisms of navigation in neural systems and test hypotheses about their functioning, it seems necessary to create computational models. Such models exist and are called biomorphic models of navigation systems [11]. These models accept the same data as biological navigation systems as input, and at output they provide an estimate of the subject's position in space relative to other objects. The output data of these systems can be studied using statistical methods, obtaining metrics of the quality of the systems' operation for their comparison. This work is devoted to the development of one of the versions of such a system and the adjustment of its operating modes.

Visual navigation in biological systems. Edward Tolman proposed and demonstrated that animals have a so-called cognitive map – a mental picture of the environment that contains information about the location of various key landmarks and their relationships to each other. This picture supports orientation in a complex, changing environment [1]. It has been shown that animals rely on distance calculations to determine location. Thus, animals can move between two locations in the absence of information about the environment, for example in the dark, as they integrate their internal signals about changes in location [2]. Examples of such signals are signals from the vestibular system, which monitors movement, and proprioception (the sense of one's own posture in space).

Animals have several navigational strategies. The interaction of these strategies can reduce localization error and exploit new capabilities. Many animals can locate themselves using path integration [3], in which the animal continuously monitors its current direction and distance from some reference point as it moves away from it. Knowing its current location, the animal can determine the relative positions of other locations whose path integration coordinates it knows with an accuracy proportional to the accuracy with which it estimates its position using path integration. Furthermore, within familiar terrain, the inevitable error associated with path integration can be reduced by information gained from landmark recognition. Etienne et al [4] reported the first behavioral evidence for this hypothesis, showing that hamsters use visual cues to reset the errors of their internal path integrator.

An important aspect of the study is understanding the biological basis, the elements of the system. The main physiological areas of environmental cognition and navigation in the brain are the hippocampus and its surroundings [12–14]. John O'Keefe, in a series of experiments with freely moving rodents, during which extracellular recording of activity was carried out, found that the activity of some cells in the Ca1 and Ca3 areas of the hippocampus was almost exactly predicted by the spatial position of the animals. These neurons were called place cells [5].

Place cells normally have low activity but increase their activity strongly when the animal is in the region of space that contains the neuron's activation region. Different place cells are sensitive to different regions of the surrounding space, so that only a small subset of them are active at any one location, thus providing a precise encoding of the animal's location. At the population level, place cells provide a kind of "map" of the environment, similar to the cognitive map proposed by Tolman [1]. For one location, the activation of place neurons is constant over time, allowing key landmarks to remain constant. However, in another region, place neurons may change their location of activity or stop firing altogether, a process called remapping. Thus, for any region, a place neuron will have a specific representation of space. However, over larger spaces, a place neuron may encode several spatial regions [15]. Place neurons are able to rely on information obtained through path integration. In addition, place neurons also play a role in episodic memory [16].

Head direction cells were the second class of spatially responsive neurons discovered in rodents. These neurons were first reported by Rank [8], who was one of the first neuroscientists to successfully record single neurons from the brains of freely moving animals. Motivated by O'Keefe's recent discovery of place cells in the hippocampus, he examined the parts of the brain afferent to the hippocampus and found in the postsubiculum (also called the dorsal presubiculum) a predominance of neurons whose firing was markedly enhanced when the animal faced a particular direction. In 1984, he published an abstract reporting their discovery and tasked his graduate student Taube with characterizing these neurons [6]. Taube and colleagues then followed the connections of the postsubicular head direction neurons and found a broad and complex circuit involving both low-level brainstem structures and high-level cortical structures [7].

Each head-direction cell has its own preferred direction of activation. Together, the population spans a full 360 degrees of horizontal space [7]. A remarkable feature of the system is that the neurons' signal has one of the sharpest tuning curves of any neuron in the brain and is highly coherent, meaning that the activation of neurons at any point in the network is consistent with which way the animal is facing at any given time. Any manipulation of the environment or the animal itself that distorts the processing of direction affects all the cells together, and it is impossible to change the direction of activation of one cell without changing the direction of activation of all the others. This high degree of coherence has led to the suggestion that the network is interconnected in such a way that each cell is driven not only by incoming sensory signals but also by the activity of other head-direction neurons in the network. This type of network has been called a "continuous ring attractor network" [8]. Such a network provides the ability to update the signal when the animal's head turns in a different direction — all that is needed is a mechanism that links the turn in a certain direction with a shift in the activity of the attractor network by a certain angle [8].

In 2005, May-Britt and Edvard Moser discovered another type of neuron that is involved in processing spatial information, the grid cell [9]. Like place neurons, this type of neuron was activated when a specific area of space was present. However, these cells were activated over the entire area in a repeating triangular pattern, like a "tiling". Based on their regular and repeating pattern of activation, these neurons were named [9]. Grid cells are considered to be among the most numerous cells in the superficial layers of the middle entorhinal cortex, although they are also found in deeper layers. A grid cell can be described by three coordinates: period (the distance between adjacent activation fields), orientation (relative to some reference direction), and phase (the two-dimensional displacement of the lattice axes toward an external reference point). Grid cells are anatomically organized into modules that have similar periods and orientations, but their phases are shifted by different values. The phase of the activation signal of such a neuron can change depending on the environment in which the animal finds itself, but, as with head direction neurons, they can be activated in all environments [10]. It is now known that grid cells are most likely a substrate for path integration and influence the activity of place neurons.

Thus, various biological modules are known to perform functions related to navigation, but the specific mechanism of information processing is still poorly understood. The hippocampus can integrate the path and guide animals to a goal along an unknown route. How the hippocampus performs these calculations is still unclear [17–19]. Current research in neurobiology focuses on the details of the mechanism of information transfer from grid cells to place cells, but does not consider the operation of the entire navigation system as a whole [20–24]. However, the dynamics of the entire system is also of interest. To study it, it seems reasonable to create a functional model of the biological navigation system based on the described biological facts.

1. Technical implementation of the visual navigation system

First of all, we note the very promising universal models proposed by V. D. Tsukeman [25–29]. In them, relational neural networks with even cyclic inhibition are used to describe navigation operations in various spaces. However, in order to optimize the development carried out at this stage and to more clearly compare it with the test results of previously developed technical systems [11,30], an approach called the simultaneous localization and mapping (SLAM) system was chosen in the development described here. These technical systems, similar to visual navigation systems in living prototypes and operating on the basis of image processing, have been rapidly developing in recent years. Most approaches are based on the detection of common visual features such as SIFT or SURF extracted from relatively high-resolution monocular or stereo images [30].

Among these systems, there are also more biologically relevant systems, such as RatSLAM [11] — an alternative SLAM system based on the neural processes underlying navigation in the rodent brain. The system's neural filtering, which builds hypotheses about localization by accumulating sensory data, allows it to function even when its perception is ambiguous.

1.1. RatSLAM system. The system consists of three main modules – pose cells, local view cells and experience maps, and it is also possible to add a fourth module – visual odometry. This system operates as part of the Robot operating system (ROS) [31]. This allows for a modular architecture of programs, facilitates multitasking and allows for standardized methods of exchanging information (topics and messages) between program modules. The structure of the system is shown in Fig. 1.

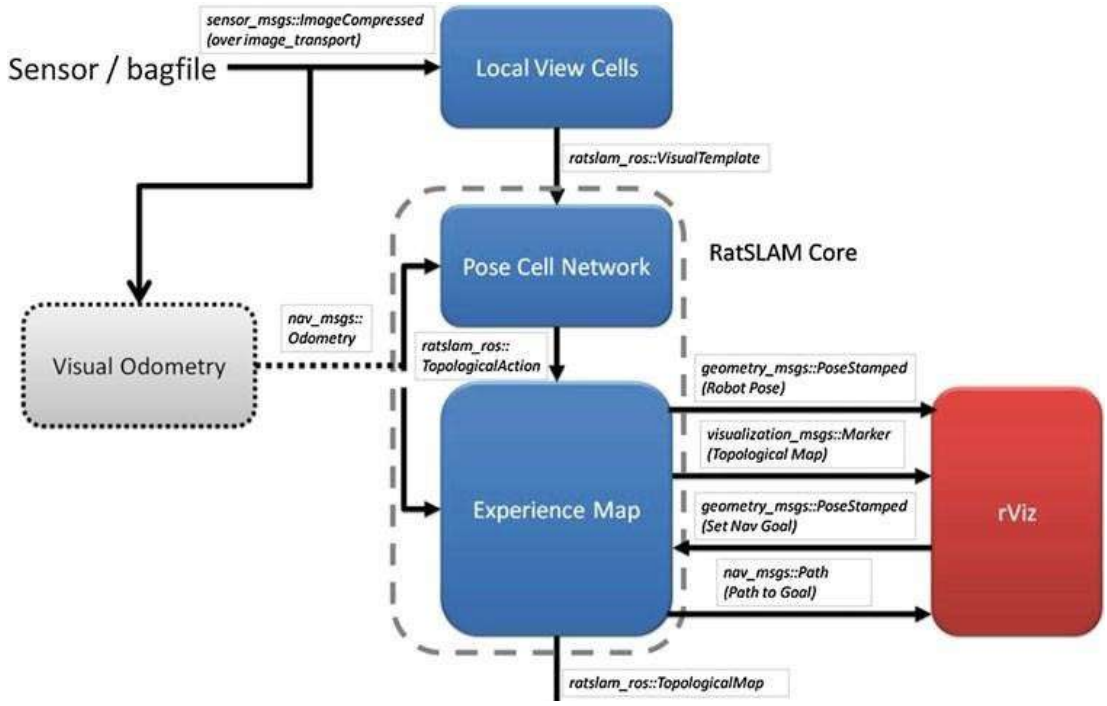

Fig. 1. Structure of nodes and messages for OpenRatSLAM [32]. If odometry is already provided by the dataset or robot, the Visual Odometry node is not required, as shown by the dotted lines

1.2. Pose cells. The pose cells are a continuous attractor network (CAN) of modules [33] connected by excitatory and inhibitory connections. In its characteristics, it is similar to a navigation neuron found in many mammals and called a grid cell [9]. The network has a three-dimensional prism configuration. Its cells are connected to neighboring cells by excitatory connections that pass through all the boundaries of the network. The network is shown in Fig. 2.

The coordinates of the cell array nominally correspond to the three-dimensional position of the ground robot — x, y, and θ. The dynamics of the cell network in pose are such that the stable state is a single cluster of activated cells, called an activity packet or energy packet. The centroid of this packet encodes the best internal estimate of the current pose of the robot. This dynamic behavior is achieved by locally excitatory and globally inhibitory connections described by the distribution ε:

$$\varepsilon_{a,b,c} = e^{-(a^2+b^2)/k_p^{exc}} e^{-c^2/k_d^{exc}} - e^{-(a^2+b^2)/k_p^{inh}} e^{-c^2/k_d^{inh}}, \tag{1}$$

where k_p and k_d are the dispersion constants for the location and direction, respectively, and a, b, and c are the distances between cells in the x, y, and θ coordinates, respectively. The dispersion constants are fixed by the setup and should not be changed. The connections span all six faces of the pose cell network, as shown by the longer arrows in Fig. 2. The change in the activity level of a cell ΔP due to its internal dynamics is given by:

$$\Delta P_{x',y',\theta'} = \sum_{i=0}^{S_{xy}-1} \sum_{j=0}^{S_{xy}-1} \sum_{k=0}^{S_z-1} P_{i,j,k} \varepsilon_{a,b,c} - \varphi, \tag{2}$$

where S_{xy} is the length of the side of the square (x, y) of the pose cell network plane, S_z is the height of the network, and ϕ is the magnitude of the global drag.

The proper motion information provided by the odometric input biases the activity in the pose cell network to represent the robot's motion based on the nominal spatial scale for each pose cell. Excitatory connections from the local view cells provide the mechanism for performing loop closure.

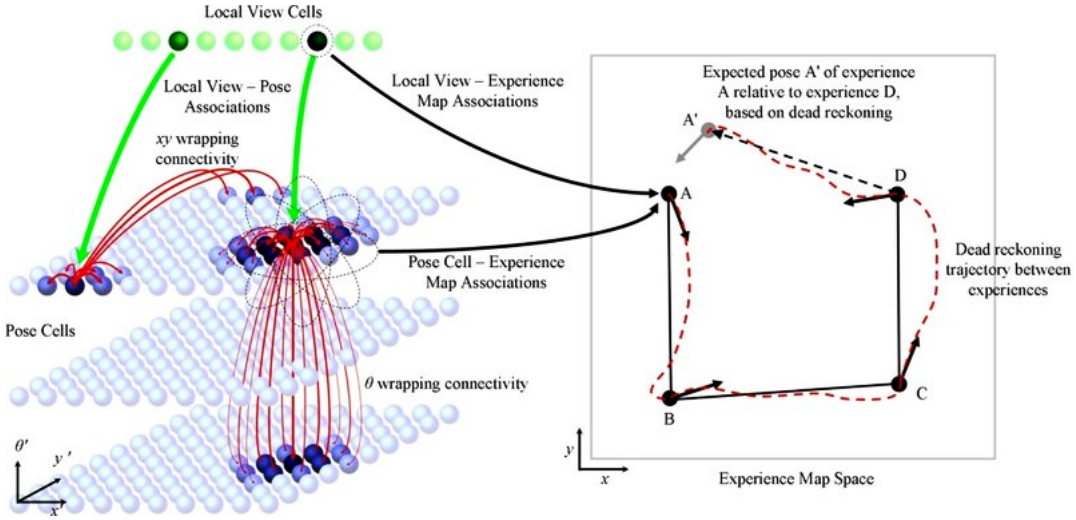

Fig. 2. Basic modules of the RatSLAM system [32]

1.3. Local view cells. Local view cells are an expandable array of units, each corresponding to a distinct visual scene in the environment. When a new visual scene is encountered, a new local view cell is created and associated with the pixel data of that scene. An excitatory connection β (single-shot learning) is created between this local view cell and the centroid of the dominant burst of activity in the pose cell network at that moment in time. When the system sees that scene again, the local view cell is activated and increases the activity of the corresponding pose cell through this excitatory connection:

$$\Delta P_{x',y',\theta'} = \delta \sum_i \beta_{i,j,k} V_i \,, \tag{3}$$

where the constant δ determines the strength of the influence of visual cues on the robot's pose estimation.

The saturation process ensures that each visual template can only inject activity for a short period of time to avoid false relocalization when the robot is stationary. The activation of local view cells associated with a particular visual template is a nonlinear process. Thus, the change in the representation of the subject's location in the pose cell network induced by visual templates is also a nonlinear process. If a sufficiently long sequence of familiar visual scenes occurs in the correct order, the constant excitation of a pose cell leads to relocalization. That is, the dominant burst of activity moves to the same position as when the scene was first presented.

After preliminary processing of the image coming from the camera (cropping of uninformative areas of the frame, brightness normalization, conversion to monochrome and compression), the local view cell module compares the resulting visual template representing the current image from the camera with all previously learned templates. A similarity measure is calculated based on the sum of absolute differences (SAD) between the current visual template and each previously learned visual template. If the smallest difference is less than the threshold, then the corresponding existing template is selected. Otherwise, the current visual template is added to the template database. The module's operation diagram is shown in Fig. 3.

This operation is nonlinear because it has a threshold value. The network of pose cells has a finite size, but the connection of opposite edges of the network means that theoretically an infinite region of space can be mapped by the network. This implies that some pose cells correspond to multiple points in space.

An experience map is a graphical map that estimates a unique pose estimate for a robot by combining information from pose cells and local view cells. Each point in the experience map can be defined by three variables:

$$e_i = \left\{P^i, V^i, p^i\right\}, \tag{4}$$

where P^i and V^i are the states of activity in the pose and local view cells, respectively, at the time of experience formation, and p^i is the location of the experience in the experience map space (the space in which the error distribution over the graph is performed). A new experience is created when the current state of activity in the pose P^i and local view V^i cells does not match the state associated with any of the existing experiences. To compare how closely the current pose and local view states match the states associated with each experience, a metric S is used, defined as follows:

$$S^i = \mu_p \left|P^i - P\right| + \mu_v \left|V^i - V\right|, \tag{5}$$

where μ_p and μ_v are the weights of the corresponding contributions of the pose and local view codes to the match score. If $\min(S) \geqslant S_{\max}$, a new experience is created, determined by the current activity states of the pose and local view cells. The error distribution algorithm distributes

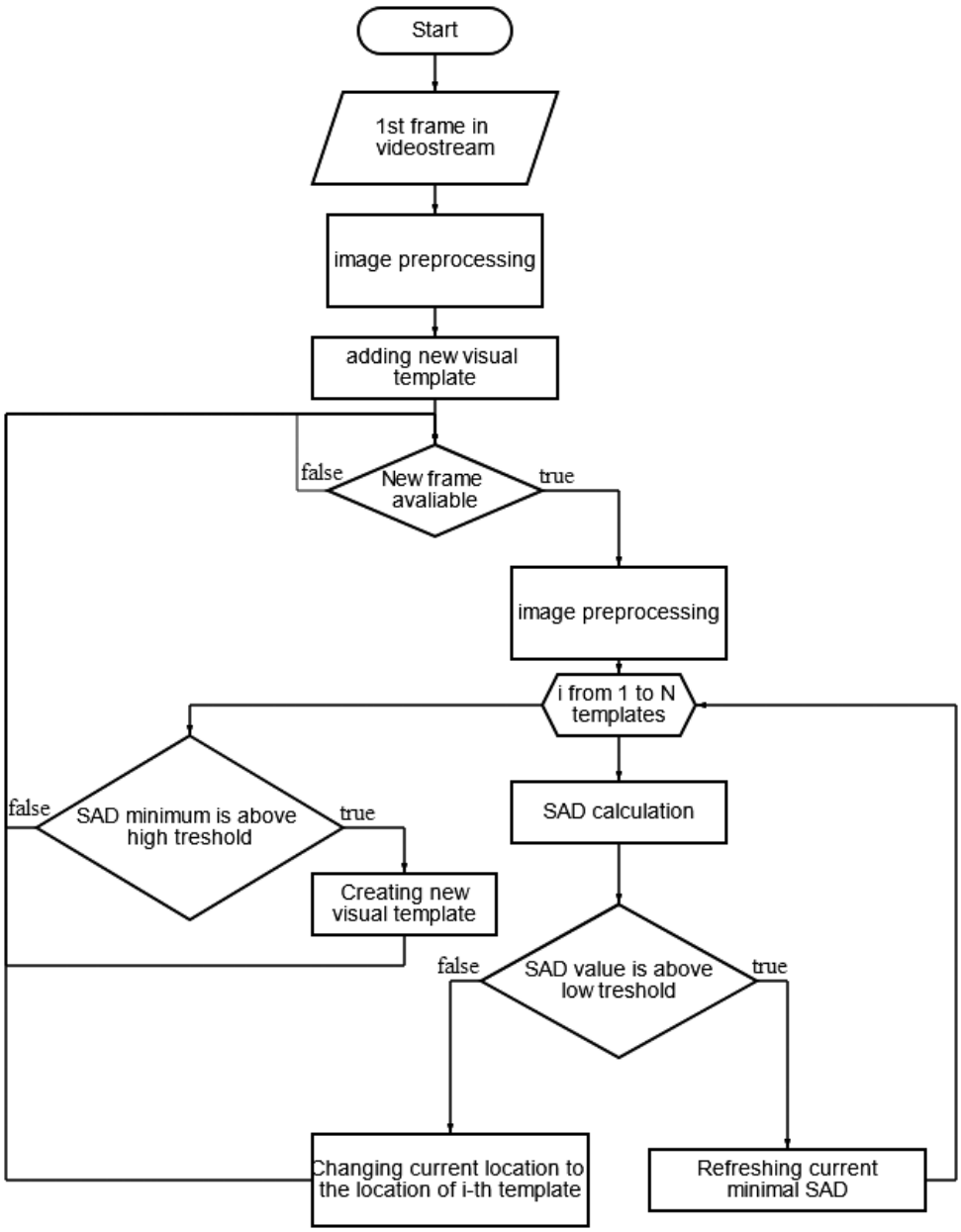

Fig. 3. Block diagram of video data processing in a local view cells module

the odometric error across the graph, creating a map of the object's environment that can be easily interpreted by humans. The change in the location of an experience is determined by

$$\Delta p^i = \alpha \left[\sum_{j=1}^{N_f} (p^j - p^i - \Delta p^{ij}) + \sum_{k=1}^{N_t} (p^k - p^i - \Delta p^{ki}) \right], \tag{6}$$

where α is the rate constant of correction equal to 0.5, N_f is the number of connections from experiment p_i to other experiments, and N_t is the number of connections from other experiments to experiment p_i.

1.4. Visual odometry. The visual odometry module determines the camera movement by comparing successive images. The rotation speed is estimated by determining the relative horizontal displacement of two successive image intensity profiles with the minimum average absolute difference between the two profiles. The intensity profiles are created by summing the intensities of the image pixels in the vertical direction. The translation speed is estimated by multiplying the minimum difference by a scale factor and is limited to a maximum value to prevent distortion of the results with large illumination changes.

1.5. Visual Odometry Add-on Module. The RatSLAM system includes a visual odometry module. However, this module has low accuracy of position determination. Using this module as a source of odometry information is one of the options for the system. It is also possible to use external odometry sources, such as a speed signal from the robot's wheels. Thus, it is possible to increase the accuracy of the system by using external visual odometry based on more modern technologies.

The ORB-SLAM2 system can be used as a source of high-precision visual odometry. Its operation is based on searching for special points on the image from a video camera, tracking their movement between frames of the video stream and constructing their supposed position and the position of the subject in three-dimensional space [30]. This system uses FAST special points with ORB descriptors [34].

The Bundle Adjustment algorithm is used to calculate the movement of the video camera between frames of the video stream, and DBoW is used to close the loops. This system is a full-fledged navigation and map construction system, i.e. it creates a global map and performs loop closure (resetting the path integration error when entering known areas of space). Therefore, to use this system as a visual odometry system, it is sufficient to disable the loop closure system and transform (differentiate with respect to time) the output data of the ORB-SLAM system.

Thus, in the modified system, visual odometry is determined by a block representing the ORB-SLAM system, and the mapping of the terrain (in the form of visual templates corresponding to certain points in space) and the closing of loops are performed by the RatSLAM system. This allows combining the metric accuracy of ORB-SLAM and the neural processing of RatSLAM.

2. Testing the developed system

The output of the SLAM algorithm is an estimated camera trajectory along with an estimate of the resulting map. Although it is possible to estimate the quality of the resulting map, obtaining accurate spatial maps is problematic due to the influence of various uncontrollable factors. For example, algorithms that work with keypoints in an image generate a map of the location of keypoints in space. However, finding the actual points corresponding to the special points of the map can be difficult. Therefore, the analysis is based mainly on the quality of the estimated trajectory obtained from the RGB image sequence. For the evaluation, we assume that the algorithm output is a sequence of positions from the estimated trajectory $P_1, ..., P_n \in SE$ and from the actual trajectory $Q_1, ..., Q_n \in SE$.

2.1. Absolute trajectory error (ATE). The absolute trajectory error (ATE) is estimated by comparing the absolute distances between the predicted trajectory and the true trajectory. Since both trajectories can be defined in arbitrary coordinate systems, they must first be aligned. This is done in closed form using Horn's method [37], which defines a rigid body transformation

S corresponding to a least-squares solution that maps the predicted trajectory $P_1 : n$ to the true trajectory $Q1 : n$. The absolute trajectory error at time step i can be calculated as

$$F_i = Q_i^{-1} S P_i \,. \tag{7}$$

Estimating the root mean square error (RMSE) for all time indices of translation components, we obtain

$$\mathrm{RMSE}\,(F_{1:n}) = \left(\frac{1}{n} \sum_{i=1}^{n} \|\mathrm{trans}\,(F_i)\|^2 \right)^{1/2}. \tag{8}$$

2.2. Relative position error (RPE). Relative Position Error (RPE) measures the local accuracy of a trajectory over a fixed time interval Δt. Thus, the relative position error corresponds to the trajectory drift, which is useful for evaluating visual odometry systems. We define the relative pose error at time step i as

$$E_i = \left(Q_i^{-1} Q_{i+\Delta}\right)^{-1} \left(P_i^{-1} P_{i+\Delta}\right)^{-1}. \tag{9}$$

From a sequence of n camera poses, $m = n - \Delta t$ (where Δt is the number of frames) of individual relative pose error values within the sequence are obtained. Based on these values, the root mean square error (RMSE) is calculated over all time indices of the motion component:

$$\mathrm{RMSE}\,(E_{1:n}, \Delta) = \left(\frac{1}{m} \sum_{i=1}^{m} \|\mathrm{trans}\,(E_i)\|^2 \right)^{1/2}, \tag{10}$$

where $\mathrm{trans}\,(E_i)$ refers to the translation component of the relative pose error E_i. It is possible to estimate the mean error, since it is more robust to outliers. Some use the median instead of the mean, which attributes even less influence to outliers. For visual odometry systems that fit successive frames, the time parameter $\Delta t = 1$, which is an intuitive choice; $\mathrm{RMSE}\,(E_{1:n})$ then determines the drift per frame. For systems that use more than one previous frame, larger values of Δ may also be appropriate. It therefore makes sense to average over all possible time intervals Δt, that is, to compute

$$\mathrm{RMSE}\,(E_{1:n}) = \frac{1}{n} \sum_{\Delta=1}^{n} \mathrm{RMSE}\,(E_{1:n}, \Delta). \tag{11}$$

RPE can be used to estimate the global trajectory error by averaging over all possible time intervals. RPE estimates both translational and rotational errors, while ATE estimates only translational errors. Thus, the RPE metric allows us to combine rotational and translational errors into a single measure. However, rotational errors are also indirectly captured by ATE, as they manifest themselves in incorrect translations. From a practical point of view, ATE has an intuitive visualization that facilitates visual inspection. However, these two metrics are highly correlated. Since the studied system uses exclusively monocular (using a single camera) algorithms, determining the exact map scale without additional information is impossible. Therefore, calculating the scale error is impossible, and all metrics will be obtained for a normalized map scale.

2.3. Datasets used for testing. The following datasets are used for testing:

1) KITTI Dataset [38];
2) NN sequences (video stream filmed by the authors).

The KITTI dataset is a sequence of images acquired by a stereo camera (in this paper, only images from one lens are used) mounted on a car, so all sequences were taken outdoors. The cameras are calibrated, and there is accurate information about the positions of the frames obtained by LiDAR and GPS sensors. No other information was used. All sequences were taken at 10 FPS. Sequences 0, 2 [38] were used for testing.

2.4. Recording the data set "NN Sequences". The system's operation was also tested on new test sequences recorded in Nizhny Novgorod from a smartphone camera installed behind the car's windshield. The resulting images are 1280×960 pixels in size, and several loop closures occur during the sequences.

To create a data set for testing visual navigation systems, it is necessary that the images from this set be of sufficiently high resolution and that there are no objects in the camera's field of view that interfere with the view of the space (when placed behind the windshield of a car — glare on the glass, car structural parts, etc.). It is also necessary to meet requirements specific to visual navigation systems.

To be able to evaluate the accuracy of the system, it is necessary to record the true trajectory of movement. For this system, it is possible to use a signal from GPS/GLONASS, since the error of the location signal received from them is less than the error of determining the location of this system. Since the system includes a module that works with key points obtained from images, it was necessary to additionally eliminate image distortion for the correct operation of this module. To eliminate distortion, it is necessary to know the parameters of the camera distortion, that is, it was necessary to calibrate the camera.

The visual odometry module used, which works with key points obtained from images, does not allow the use of cameras with automatic focus adjustment. Therefore, the images were recorded without automatic focus adjustment.

Since in this system the landmarks on the route map are visual templates formed directly from the pixel data of the image, it is desirable to have the greatest variability of the image during the sequence: different types of buildings, different types of streets, etc.

Thus, it is possible to use a smartphone camera provided that it is pre-calibrated, and the automatic focal length adjustment function is disabled, and the GPS-determined location is recorded in parallel. These requirements were met using the open source software products Android camera calibration tool and Android dataset recorder, created by the Robot Perception and Navigation Group, University of Delaware. These programs are used to calibrate the camera and record the dataset, respectively. The recorded dataset was then converted to the ROS (Robot Operating System) message format.

3. Results of the setup and their discussion

3.1. Visual navigation system. The implementation of the RatSLAM system in the OpenRatSLAM library was chosen as the main system [32]. This system was assembled and tested on a data set from the system developers, as well as on test data sets. The correct display of the camera position and the operation of the loop closure system were confirmed. During the work, the ability to work with images in the rgb8 bit format, output of the trajectory after the end of the work, calculation and output of the odometry message in ORB-SLAM2 were added, and the viewer of visual templates and the current view was improved. The ability to skip loop closure events was added to the ORB-SLAM2 system. A screenshot of the system in operation is shown in Fig. 4.

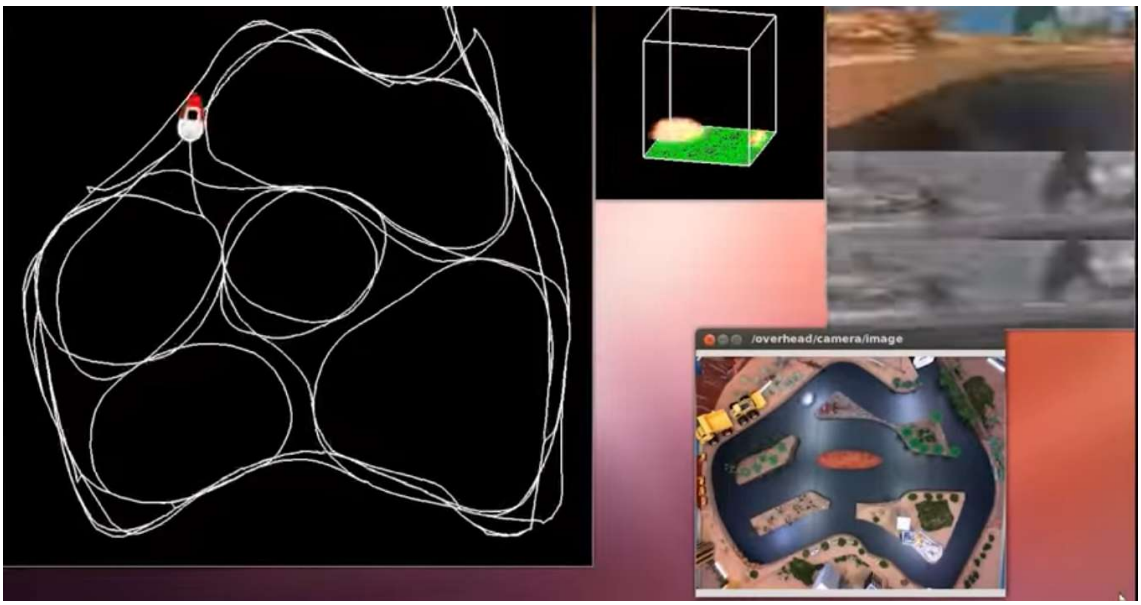

Fig. 4. Screenshot of the OpenRatSLAM system [32]. Top row from left to right: map built by the system; visualization of activation processes in the network of pose cells; input image. Bottom image is a top view of the test site (color online)

3.2. System Performance Metrics. When testing the system's operation on the KITTI data set, the metrics presented in the Table were obtained.

Model performance metrics (m)

Sequence	Metrics	Parameters	System	
			Ratslam	Modified system
Kitti-00	APE	Max	98.51	42.17
		Mean	39.81	19.68
		Median	35.81	19.90
		Min	0.57	6.03
	RPE	Max	5.38	8.17
		Mean	1.75	2.83
		Median	1.58	2.57
		Min	0.05	0.17

Fig. 5 shows visualizations of the distribution of the absolute localization error of systems on the trajectory of motion (on the left — the original Ratslam system, on the right — for the modified system).

Graphs of the dependence of the average value of the position determination error on the strength of the excitatory connection from the cells of the local type, which determines the threshold of the nonlinear process — the event of closing the cycle, as well as on the spatial discreteness of the system operation were constructed. These graphs are presented in Fig. 6, 7.

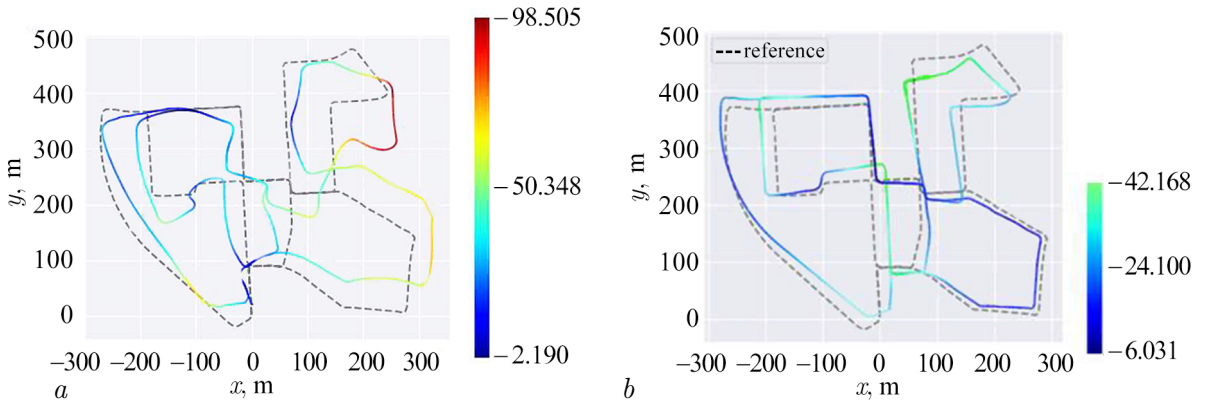

Fig. 5. Visualizations of the error distribution of systems along the motion trajectory: *a* — original Ratslam system, *b* — modified system (color online)

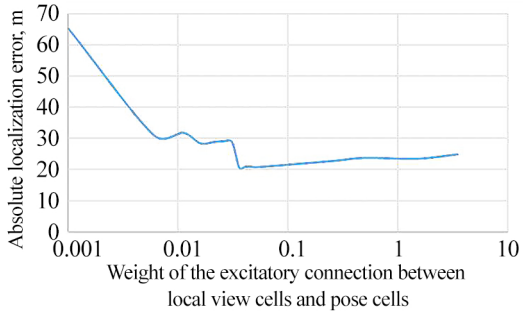

Fig. 6. Dependence of the average absolute localization error on the weight of the excitatory connection between local view cells and pose cells

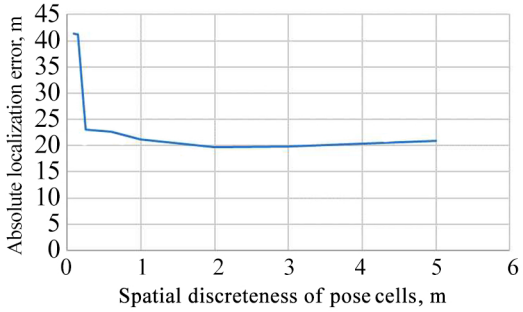

Fig. 7. Dependence of the average absolute localization error on the weight of the spatial discreteness of the system operation

Conclusion

In accordance with the stated goal, possible ways of creating a version of a complex nonlinear control system that performs biologically relevant visual navigation operations were considered. The necessary metrics evaluating the accuracy of this system were selected for the assembled prototype. Based on these metrics, some parameters of the prototype system were adjusted, which allowed optimizing its functioning.

The resulting system implements navigation strategies similar to those used in biological prototypes. In the network of pose cells, functionally similar to the networks of grid cells in the mammalian brain, path integration is performed, allowing one to determine the location in unfamiliar areas of space. This is consistent with the experiments [3]. During the process of path integration, an integration error accumulates, which, as in biological systems, is reset when entering familiar areas of space, when recognizing familiar places [4].

In the resulting system, such recognition occurs when the visible image and a familiar visual template match. However, the methods used to compare images are weakly relevant to the proposed biological analogues. Biological methods are more stable when changing illumination, viewing angle, and over long time intervals. Therefore, in the future, it is planned to use more biologically relevant methods [39]. In addition, the resulting map is equally detailed in frequently and rarely visited areas, which does not correspond to modern concepts [40]. Therefore, in the future, it is planned to use a model of space that expands with experience [40]. Also, in future studies, it is possible to use odometry methods that are more relevant to biological analogues.

References

1. Tolman EC. Cognitive maps in rats and men. Psychological Review. 1948;55(4):189–208. DOI: 10.1037/h0061626.
2. McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser M-B. Path integration and the neural basis of the "cognitive map". Nature Reviews Neuroscience. 2006;7:663–678. DOI: 10.1038/nrn1932.
3. Mittelstaedt H, Mittelstaedt M-L. Homing by Path Integration. In: Papi F, Wallraff HG, editors. Avian Navigation. Proceedings in Life Sciences. Berlin, Heidelberg: Springer; 1982. P. 290–297. DOI: 10.1007/978-3-642-68616-0_29.
4. Etienne AS, Maurer R, Boulens V, Levy A, Rowe T. Resetting the path integrator: a basic condition for route-based navigation. Journal of Experimental Biology. 2004;207(9):1491–508. DOI: 10.1242/jeb.00906.
5. O'Keefe J, Dostrovsky J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Research. 1971;34(1):171–175. DOI: 10.1016/0006-8993(71)90358-1.
6. Muller RU, Ranck JB Jr, Taube JS. Head direction cells: properties and functional significance. Current Opinion in Neurobiology. 1996;6(2):196–206. DOI: 10.1016/s0959-4388(96)80073-0.
7. Taube JS. The head direction signal: Origins and sensory-motor integration. Annual Review of Neuroscience. 2007;30:181–207. DOI: 10.1146/annurev.neuro.29.051605.112854.
8. *Amit DJ.* Modeling Brain Function: The World of Attractor Neural Networks. New York, NY: Cambridge University Press; 1989. DOI: 10.1017/cbo9780511623257.004.
9. Moser E, Roudi Y, Witter MP, Kentros C, Bonhoeffer T, Moser M-B. Grid cells and cortical representation. Nature Reviews Neuroscience. 2014;15:466–481. DOI: 10.1038/nrn3766
10. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. Microstructure of a spatial map in the entorhinal cortex. Nature. 2005;436:801–806. DOI: 10.1038/nature03721.
11. Milford MJ. Robot Navigation from Nature: Simultaneous Localisation, Mapping, and Path Planning Based on Hippocampal Models. Vol. 41. Springer Tracts in Advanced Robotics. Berlin, Heidelberg: Springer; 2008. 196 p. DOI: 10.1007/978-3-540-77520-1.
12. Eichenbaum H. The role of the hippocampus in navigation is memory. Journal of Neurophysiology. 2017;117(4):1785–1796. DOI: 10.1152/jn.00005.2017.
13. Gu Y, Lewallen S., Kinkhabwala AA, Domnisoru C, Yoon K, Gauthier JL, Fiete IR, Tank DW. A map-like micro-organization of grid cells in the medial entorhinal cortex. Cell. 2018;175(3):736–750. DOI: 10.1016/j.cell.2018.08.066.
14. Poulter S, Hartley T, Lever C. The neurobiology of mammalian navigation. Current Biology. 2018;28(17):R1023–R1042. DOI: 10.1016/j.cub.2018.05.050.
15. Park E, Dvorak D, Fenton AA. Ensemble place codes in hippocampus: CA1, CA3, and dentate gyrus place cells have multiple place fields in large environments. PLoS One. 2011;6(7):e22349. DOI: 10.1371/journal.pone.0022349.
16. Goode TD, Tanaka KZ, Sahay A, McHugh TJ. An integrated index: Engrams, place cells, and hippocampal memory. Neuron. 2020;107(5):805–820. DOI: 10.1016/j.neuron.2020.07.011.
17. Epstein RA, Patai EZ, Julian JB, Spiers HJ. The cognitive map in humans: spatial navigation and beyond. Nature Neuroscience. 2017;20(11):1504–1513. DOI: 10.1038/nn.4656.
18. Sarel A, Finkelstein A, Las L, Ulanovsky N. Vectorial representation of spatial goals in the hippocampus of bats. Science. 2017;355(6321):176–180. DOI: 10.1126/science.aak9589.
19. Savelli F, Knierim JJ. Origin and role of path integration in the cognitive representations of the hippocampus: computational insights into open questions. Journal of Experimental Biology. 2019;222(1):jeb188912. DOI: 10.1242/jeb.188912.

20. Rolls ET, Stringer SM, Elliot T. Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. Network. 2006;17(4):447–65. DOI: 10.1080/09548980601064846.
21. Si B, Treves A. The role of competitive learning in the generation of DG fields from EC inputs. Cognitive Neurodynamics. 2009;3(2):177–87. DOI: 10.1007/s11571-009-9079-z.
22. Savelli F, Knierim JJ. Hebbian analysis of the transformation of medial entorhinal grid-cell inputs to hippocampal place fields. Journal of Neurophysiology. 2010;103(6):3167–83. DOI: 10.1152/jn.00932.2009.
23. Danjo T, Toyoizumi T, Fujisawa S. Spatial representations of self and other in the hippocampus. Science. 2018;359(6372):213-218. DOI: 10.1126/science.aao3898.
24. Lian Y, Burkitt AN. Learning spatiotemporal properties of hippocampal place cells. eNeuro. 2022;9(4):ENEURO.0519-21.2022. DOI: 10.1523/ENEURO.0519-21.2022.
25. Tsukerman VD, Cheshkov GN. Fundamentals of nonlinear dynamics of sensory perception. I. Phase coding in oscillatory networks. Neurocomputers: development, application. 2002; (7–8): 65–72.
26. Tsukerman VD. Mathematical model of phase coding of events in the brain. Mathematical biology and bioinformatics. 2006;1(1):97–107. DOI: 10.17537/2006.1.97.
27. Tsukerman VD, Eremenko ZS, Karimova OV, Kulakov SV, Sazykin AA. Cognitive neurodynamics of two strategies of navigation behavior of organisms. Izvestiya VUZ. Applied Nonlinear Dynamics. 2011;19(6):96–108. DOI: 10.18500/0869-6632-2011-19-6-96-108.
28. Tsukerman VD, Kharybina ZS, Kulakov SV. Mathematical model of spatial encoding in the hippocampal formation. II. Neurodynamic correlates of mental trajectories and the problem of decision making. Mathematical biology and bioinformatics. 2014;9(1):216–256. DOI: 10.17537/2014.9.216.
29. Tsukerman VD. Towards creative cognition, the creative principles of relational neural networks with even cyclic inhibition. In: Proceedings of the VII All-Russian Conference. Nonlinear dynamics in cognitive research 2021, Nizhny Novgorod, September 20–24, 2021. Nizhny Novgorod: IPFRAN; 2021. P. 186–189.
30. Mur-Artal R, Montiel JMM, Tardós JD. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Transactions on Robotics. 2015;31(5):11471163. DOI: 10.1109/TRO.2015.2463671.
31. Quigley M, Conley K, Gerkey B, Faust J, Foote T, Leibs J, Wheeler R, Ng A. ROS: an open-source Robot Operating System. ICRA Workshop on Open Source Software. 2009;3(3.2):5.
32. Ball D, Heath S, Wiles J, Wyeth G, Corke P, Milford M. OpenRatSLAM: an open source brain-based SLAM system. Autonomous Robots. 2013;34:149–176. DOI: 10.1007/s10514-012-9317-9.
33. Samsonovich A, McNaughton BL. Path integration and cognitive mapping in a continuous attractor neural network model. The Journal of Neuroscience. 1997;17(15):5900–5920. DOI: 10.1523/jneurosci.17-15-05900.1997.
34. Rublee E, Rabaud V, Konolige K, Bradski GR. ORB: An efficient alternative to SIFT or SURF. In: Proceedings of the 2011 International Conference on Computer Vision, ICCV 2011. Barcelona, Spain, November 6–13, 2011. P. 564–2571.
35. Triggs B, McLauchlan PF, Hartley RI, Fitzgibbon AW. Bundle Adjustment — A Modern Synthesis. In: Triggs, B., Zisserman, A., Szeliski, R. (eds) Vision Algorithms: Theory and Practice. IWVA 1999. Lecture Notes in Computer Science, vol. 1883. Berlin, Heidelberg: Springer; 2000. P. 298–372. DOI: 10.1007/3-540-44480-7_21.
36. Galvez-López D, Tardos JD. Bags of Binary Words for Fast Place Recognition in Image Sequences. IEEE Transactions on Robotics. 2012;28(5). P. 1188–1197. DOI: 10.1109/TRO.2012.2197158.

37. Horn BK. Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A. 1987;4(4):629–641. DOI: 10.1364/josaa.4.000629.
38. Geiger A, Lenz P, Stiller C, and Urtasun R. Vision meets robotics: the KITTI dataset. The International Journal of Robotics Research. 2013;32(11):1231–1237. DOI: 10.1177/02783 64913491297.
39. Yu F, Wu Y, Ma S, Xu M, Li H, Qu H, Song C, Wang T, Zhao R, Shi L. Brain-inspired multimodal hybrid neural network for robot place recognition. Sci Robot. 2023;8(78). DOI: 10.1126/scirobotics.abm6996.
40. Zhang H, Rich PD, Lee AK, Sharpee TO. Hippocampal spatial representations exhibit a hyperbolic geometry that expands with experience. Nature Neuroscience. 2023;26(1):131–139. DOI: 10.1038/s41593-022-01212-4.