

A Biologically-Inspired Simultaneous Localization and Mapping System Based on LiDAR Sensor

Genghang Zhuang¹, Zhenshan Bing¹, Yuhong Huang¹, Kai Huang^{2,3}, and Alois Knoll¹

Abstract—Simultaneous localization and mapping (SLAM) is one of the essential techniques and functionalities used by robots to perform autonomous navigation tasks. Inspired by the rodent hippocampus, this paper presents a biologically inspired SLAM system based on a LiDAR sensor using a hippocampal model to build a cognitive map and estimate the robot pose in indoor environments. Based on the biologically inspired models mimicking boundary cells, place cells, and head direction cells, the SLAM system using LiDAR point cloud data is capable of leveraging the self-motion cues from the LiDAR odometry and the boundary cues from the LiDAR boundary cells to build a cognitive map and estimate the robot pose. Experiment results show that with the LiDAR boundary cells the proposed SLAM system greatly outperforms the camera-based brain-inspired method in both simulation and indoor environments, and is competitive with the conventional LiDAR-based SLAM methods.

I. INTRODUCTION

Simultaneous localization and mapping (SLAM), as one of the essential techniques and functionalities used by robots to perform autonomous navigation tasks [1], aims to build a map to construct a spatial representation of the unknown environment, and simultaneously locate the robot on the map being built. A large amount of research focuses on solving the SLAM problem with various types of sensors [2]–[4], including monocular and stereo cameras, RGB-D cameras, and LiDAR sensors. Compared to camera sensors, LiDAR sensors are capable of obtaining more accurate distance and depth information by actively projecting laser beams, and are less subject to illumination changes in the environment. Hence, LiDAR sensors are widely used in autonomous driving and a considerable amount of research has been performed on using LiDAR sensors to resolve the SLAM problem [3], [5], [6], and other autonomous driving tasks including object detection [7], classification [8], and tracking [9].

In regard to mammals, research findings have shown that animals such as rodents have a different navigation system. Mammals are born with instinctive abilities and skills to perform navigation and cognition tasks. Successful discoveries have been made in investigating and understanding the spatial representation and navigation system of mammalian brains. Studies from neuroscience [10] have revealed that the hippocampus and entorhinal cortex play an important role in spatial navigation by coordinating several types of neurons for different functionalities, including place cells [11], [12], head direction cells [13], grid cells [14], [15], and boundary cells [16]–[18], in which the studies demonstrate these types

of neurons are related to place recognition, boundary sensing, and path integration which is the capacity to use idiothetic cues to track the animal’s movements [19], [20].

An increasing number of studies have been inspired by the neural mechanisms to solve the navigation [21]–[23] and SLAM problem [24]. *RatSLAM* in [24] utilized place cells and head direction cells to build a hippocampal model to integrate motion data and landmarks from vision sensors when mapping. In addition, a considerable number of related studies have been published based on this navigational model, employing many types of sensors, including cameras [25]–[28] and RGB-D sensors [29].

LiDAR sensor, which corresponds to animal echolocation that exists in some bat species and odontocetes, has the intuitive advantage in detecting boundaries in the environment. In addition, with the advantages of LiDAR sensors including high accuracy and stability, significant potential exists to further improve the overall performance of the biologically inspired SLAM system by using LiDAR sensors. However, to date, little literature has been published with regard to LiDAR-based biologically inspired SLAM approaches. Considering the connection to echolocating animals, it is intriguing to explore the applications and investigate the advantages of LiDAR sensors in biologically inspired navigation. Due to the developments of neuromorphic chips [30], [31] and the emerging deployments of navigational models [32], [33], a LiDAR-based biologically-inspired model with higher performance also has the potential to benefit the navigation system for biomimetic robots. Hence, in this work we focus on leveraging LiDAR sensors in biologically-inspired SLAM.

In this paper, we present a biologically inspired SLAM system based on a LiDAR sensor to build cognitive maps for indoor environments, which mimics the mechanisms of boundary cells, place cells, and head direction cells of mammals. The system, employing a LiDAR sensor as the primary sensor, consists of three major modules, which include the LiDAR odometry, LiDAR boundary cells, and the pose cell network. The SLAM system using point cloud data from a LiDAR sensor is capable of leveraging self-motion cues from the LiDAR odometry and boundary view cues from the LiDAR boundary cells to build a cognitive map and estimate the robot pose. Experiment results show that the proposed SLAM system greatly outperforms camera-based RatSLAM in terms of accuracy for SLAM tasks in both simulation and indoor environments, and is very competitive with the conventional LiDAR-based SLAM methods. Specifically, the main contributions of this work are listed as follows:

- To fully leverage the advantages of the LiDAR sensor, a LiDAR odometry algorithm is implemented to generate

¹Department of Informatics, Technical University of Munich, Germany.

²School of Computer Science and Engineering, Sun Yat-sen University.

³Pazhou Laboratory, China.

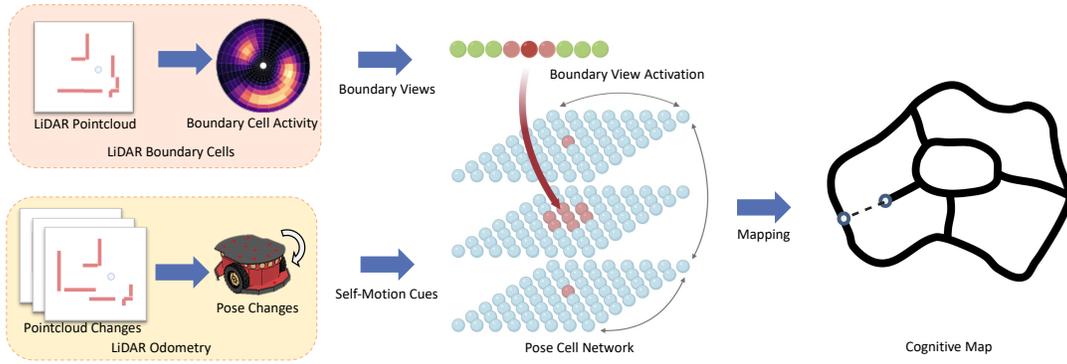


Fig. 1. System framework and major modules of the LiDAR-based biologically inspired SLAM system. The system is composed of the LiDAR odometry, LiDAR boundary cells, and the pose cell network.

pseudo-odometry data as the self-motion cues for the robot, to reduce the usage of physical odometric sensors such as IMU and wheel encoders that are prone to internal errors and noise.

- The egocentric boundary cell model for LiDAR is implemented to provide boundary view cues to perform place recognition and position calibration. A two-stage boundary view matching approach is proposed to reduce the computational complexity of place recognition.
- The pose cell network model is adapted for the LiDAR sensor to implement the spacial location representation for the robot, and perform path integration and position calibration with the LiDAR self-motion cues and LiDAR boundary view cues.

II. RELATED WORK

A. Conventional LiDAR SLAM

Error accumulation in point cloud matching and odometry drift is one of the main challenges of LiDAR-based SLAM. A number of studies have attempted to tackle this problem. Kohlbrecher et al. [5] proposed a scan-to-map matching algorithm with multi-resolution maps to reduce the matching errors. A LiDAR-based SLAM algorithm proposed by Grisetti et al. [34] utilized particle filters to reduce the odometry drift while mapping. However, these algorithms without loop closure detection, which is a technique to recognize previously visited locations, are still subject to cumulative sensor errors over long time mapping. Hess et al. [6] proposed the Cartographer SLAM system, which could perform loop closure detection for LiDAR SLAM to further improve the accuracy of mapping and localization. To detect loop closures, the system regularly runs the global pose optimization, in which the LiDAR scan is matched with all collected submaps to find the closest submap and update the pose estimate. It requires the loop closure detection to complete before the new scan being received.

B. Brain-Inspired Navigation

A number of studies were investigated to solve the SLAM problem with biologically inspired methods. Milford et al. proposed *RatSLAM* [24] to utilize the place cell and head direction cell model to perform loop closure detection and

solve the SLAM problem based on a camera sensor. Experimental results in [27], [28] also show the applicability of their proposed system in large-scale scenarios. Zhou et al. [26] employed the ORB features for the camera sensor to improve loop closure detection when mapping. Tian et al. [29] proposed an RGB-D sensor-based SLAM method with the path integration model, and *BatSLAM* [35] combined the mapping module of *RatSLAM* and a sonar sensor to solve the SLAM tasks. However, in the related literature, it is rare that studies have focused on biologically inspired SLAM approaches using LiDAR sensors.

III. METHODOLOGY

The proposed biologically inspired SLAM system uses a LiDAR as the main sensor and data source. The architecture of the SLAM system is illustrated in Fig. 1. LiDAR odometry plays an important role in generating motion data for the robot. The boundary cells provides boundary view cues that are processed and integrated based on the LiDAR observation. The pose cell network estimates the robot 3-DoF pose $\mathbf{p} = (x, y, \theta)^T$ by performing path integration and loop closure based on the self-motion data from LiDAR odometry and boundary view cues from the boundary cell module.

A. LiDAR Odometry

The LiDAR odometry in the SLAM system provides odometric motion data for the pose cells to update the network state and perform path integration. Receiving consecutive point clouds from the LiDAR sensor, the aim of the LiDAR odometry is to estimate the pose change $\Delta\mathbf{p} = (\Delta x, \Delta y, \Delta\theta)$ during the time period of the robot movement.

The LiDAR sensor rotates in the horizontal plane and scans the surrounding objects with a fixed angle increment based on the angular resolution of the LiDAR sensor. A point cloud P from a complete scan can be defined as a sequence of endpoint distances as follows:

$$P = (d_i, i = 1, 2, \dots, N), \quad (1)$$

where d_i is the distance of the i -th scan endpoint, and N is the number of points in a complete scan, which is determined by the scanning angular resolution of the LiDAR sensor.

To estimate the pose change $\Delta\mathbf{p}$ with a given new input point cloud, the points are first transformed into the form of

\mathbb{R}^2 Cartesian coordinate $\mathbf{e} = (x, y)^T$ as follows:

$$\mathbf{e}_i = \begin{pmatrix} d_i \cdot \cos\alpha_i \\ d_i \cdot \sin\alpha_i \end{pmatrix}, \quad (2)$$

where α_i is the scanning angle of the i -th endpoint.

In this work, local LiDAR mapping on an occupancy grid map is adopted to compute the pose change. A local occupancy grid map \mathbf{M} is employed to downsample the accumulation of LiDAR observations and reduce the effect of the sensor noise and dynamic obstacles [36]. The point cloud from LiDAR observation is matched against the local map to reduce the cumulative drift in scan-to-scan matching [5].

To estimate the pose change $\Delta\mathbf{p}$, a nonlinear optimization is constructed and performed to find the objective pose change $\Delta\mathbf{p}'$, where the current LiDAR point cloud can best match the local map \mathbf{M} after being projected. The optimization objective is defined as follows:

$$\Delta\mathbf{p} = \underset{\Delta\mathbf{p}'}{\operatorname{argmax}} \sum_i^N \mathbf{M}^* \left(\mathbf{R}_{\Delta\mathbf{p}'} \cdot (\mathbf{R}_{\mathbf{p}} \cdot \mathbf{e}_i + \mathbf{T}_{\mathbf{p}}) + \mathbf{T}_{\Delta\mathbf{p}'} \right), \quad (3)$$

where \mathbf{p} is the current pose of the robot. \mathbf{M}^* is an upsampled local map of \mathbf{M} to provide a continuous occupancy function for the discrete grid map \mathbf{M} [5]. $\mathbf{R}_{\mathbf{p}}$ and $\mathbf{T}_{\mathbf{p}}$ are the corresponding rotation and translation transforms for a given pose $\mathbf{p} = (x, y, \theta)^T$:

$$\mathbf{R}_{\mathbf{p}} \cdot \mathbf{e} + \mathbf{T}_{\mathbf{p}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \mathbf{e} + \begin{pmatrix} x \\ y \end{pmatrix}. \quad (4)$$

The input point cloud is initially transformed to the local map with the current pose \mathbf{p} . The objective $\Delta\mathbf{p}'$ is optimized to yield the pose change estimate $\Delta\mathbf{p} = (\Delta x, \Delta y, \Delta\theta)$ as the output of LiDAR odometry, which is subsequently used to update the activity of pose cells and perform path integration.

B. LiDAR Boundary Cells

With the LiDAR observation inputs, the LiDAR boundary cells module detects and processes the surrounding boundaries into boundary cell activity, which can be considered as abstracted boundary view information for the egocentric scene. Egocentric boundary views are used to provide the boundary cues to the pose cells to perform loop closure.

The computational model of boundary cells was first predicted in [16], [37], and the existence of boundary cells was discovered and proposed in [17], [18]. Each neuron of the boundary cells responds to a boundary in proximity in a certain area, which is denoted as the receptive field. In the LiDAR boundary cell model, the receptive fields of the cells are arranged around the robot in concentric rings at different distances, as illustrated in Fig. 2. Each boundary cell is tuned to fire maximally with the input at the center of its receptive field. Given a LiDAR point input $\mathbf{e} = (d, \alpha)$ in the polar coordinate system, the neuron activity update Δb_i for the i -th boundary cell, whose receptive field center is at (d_i, α_i) , is defined as follow:

$$\Delta b_i = \frac{1}{d} \cdot e^{-\left(\frac{d-d_i}{\sigma_{d_i}}\right)^2} \cdot e^{-\left(\frac{\alpha-\alpha_i}{\sigma_{\alpha}}\right)^2}, \quad (5)$$

where σ_{d_i} and σ_{α} are the parameters to determine the area

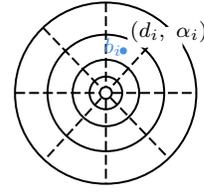


Fig. 2. Receptive fields of boundary cells.

sensitivity of the receptive field by tuning the variances for angle and distance. σ_{d_i} is a variable linear to d_i in order to expand the receptive field when distance increases [17], while σ_{α} is a constant determined by the number of boundary cells in a ring. The activity level of boundary cells is inversely proportional to the distance to corresponds the property that the boundary cells' firing rate gradually increases with approaching closer to the boundary [17].

The real-time LiDAR boundary cell activity will be compared with the previously learned boundary views to generate a pose calibration activity if a matched previous view is found, or to learn a new boundary view if none is matched. To reduce the matching complexity, a two-stage boundary view matching approach is proposed in this work. In the first stage, coarse prematching is performed to boost the real-time performance and reduce the computational complexity. A coarse view feature for the LiDAR boundary cells is proposed to roughly describe and index the boundary views by utilizing the aggregate of the boundary cell firing activity. The coarse view feature h is defined as follows:

$$h = \left\lfloor 10^{-d_s} \cdot \sum_i^M b_i \right\rfloor, \quad (6)$$

where d_s is a constant downscaling factor to control the preciseness for the coarse matching, and M is the size of the boundary cells. The coarse view feature h can be considered a hash value for the LiDAR boundary view, which is leveraged to reduce unnecessary matching processes in the second stage. By combining the coarse feature, a LiDAR boundary view v is constructed as a tuple of the coarse view feature h and the boundary cell activity B :

$$v = \langle h, B = \{b_i, i = 1, 2, \dots, M\} \rangle. \quad (7)$$

In the second stage of the boundary view matching process, for the boundary views with the equal coarse view feature h , the corresponding views are subsequently compared to estimate the similarity $s = S(B_i, B)$, which is evaluated by computing the mean squared differences between the two boundary views:

$$S(B_1, B_2) = \frac{1}{M} \sum_i^M (B_1(i) - B_2(i))^2, \quad (8)$$

The similarity s is then used to determine whether the learned boundary view v_i is matched with the current boundary view v . The boundary views are subsequently associated with the boundary view vector V , which is defined as a vector of matching activity levels of the boundary views. To maintain the matching levels over time for each boundary view, the activity of each node in the vector will be updated as follows:

$$V_i = 1 - \frac{\min(s_t, S(B_i, B))}{s_t}, \quad (9)$$

where s_t is the matching threshold for similarity. The activity level is calculated based on the matching error between the boundary view $v = \langle h, B \rangle$ for the current LiDAR observation and the i -th learned boundary view $v_i = \langle h_i, B_i \rangle$, in which the error is clamped, inverted, and scaled to $[0, 1]$. In the event that no present boundary view is matched, the vector V will be extended to V_{i+1} to associate the new boundary view. Eventually, the activity V is output to the pose cell network to perform pose association for a new boundary view or loop closure for learned boundary views.

C. Pose Cell Network

In this work, we adapt and implement the 3D pose cell network that combines 2D place cells and 1D head direction cells to maintain the pose representation and integrate self-motion cues from LiDAR odometry and the boundary view cues from LiDAR boundary cells, which is designed to reduce odometry drift and solve boundary view ambiguity in the process of mapping. Leveraging the pose cell network enables the proposed SLAM system to build a cognitive map by performing path integration based on self-motion cues. In addition, with the LiDAR view cues, the pose cell network can calibrate the estimated pose and the online cognitive map by performing loop closure to reduce the accumulated errors and drifts by LiDAR odometry.

The pose cell network is a 3D continuous attractor network (3D-CAN) [38], which can be represented as a 3D matrix of the activity: $\mathbf{PC}_{x',y',\theta'}$. The three dimensions of the pose cell network represent the three degrees of the 3-DoF pose $\mathbf{p} = (x, y, \theta)^T$, respectively, where $x', y', \theta' \in \mathbb{Z}$ are the discrete representation of $x, y, \theta \in \mathbb{R}$. Each pose cell unit in the pose cell network is connected with its neighbor units with excitatory and inhibitory connections, which wrap across the boundaries of the network in three dimensions to enable the pose cell network to represent an unbounded space with a limited number of pose cells [24]. The pose cell network enrolls local excitation and global inhibition activities, which are based on a three-dimensional Gaussian distribution, to self-update the pose cell network dynamics over time [28]. The stable state of the pose cell network, in which the activated cells are clustered, encodes the estimate of pose \mathbf{p} as the centroid of the activity packet [39].

Driven by the robot motion, the path integration updates the activity of the pose cells based on the self-motion cues from the LiDAR odometry detailed in SECTION III-A. The activity of each pose cell is shifted along with the translational and rotational movement based on the pose change $\Delta \mathbf{p} = (\Delta x, \Delta y, \Delta \theta)$. The activity update for each pose cell is defined as follows:

$$\Delta \mathbf{PC}_{x',y',\theta'} = \sum_{i=\delta_x}^{\delta_x+1} \sum_{j=\delta_y}^{\delta_y+1} \sum_{k=\delta_\theta}^{\delta_\theta+1} r_{ijk} \cdot \mathbf{PC}_{(x'+i),(y'+j),(\theta'+k)} \quad (10)$$

$$\delta_x = \lfloor k_x \Delta x \rfloor, \delta_y = \lfloor k_y \Delta y \rfloor, \delta_\theta = \lfloor k_\theta \Delta \theta \rfloor,$$

where r_{ijk} is a residue based on the fractional part of the pose changes spread over the $2 \times 2 \times 2$ cube to reduce the precision loss by quantization [24], k_x, k_y, k_θ are the constant scaling factors for the three dimensions.

Given the LiDAR boundary view cues from the boundary cells, a calibrating activity is injected into the pose cell network to perform further loop closure and re-localization. The activity of each pose cell is updated based on the activity of boundary cells V , as defined as follows:

$$\Delta \mathbf{PC}_{x',y',\theta'} = k_V \cdot \sum_i \mathbf{A}_{i,x',y',\theta'} \cdot V_i, \quad (11)$$

where k_V is the constant calibration rate, and $\mathbf{A}_{i,x',y',\theta'}$ is an adjacency matrix for the connections from the boundary cells to the pose cells. When a new boundary view V_i is learned, an excitatory link from V_i to the current state of the pose cell network (x', y', θ') is established, where $\mathbf{A}_{i,x',y',\theta'}$ is accordingly set to 1. To solve the boundary view ambiguity, only a sequence of updates by consecutive boundary views can shift the main activity packet of the pose cells.

In the mapping process, the information from the LiDAR odometry, LiDAR boundary cells, and the pose cells are combined and accumulated to estimate the robot pose $\mathbf{p} = (x, y, \theta)^T$ in the \mathbb{R}^3 space, and build a cognitive map as a topological graph of robot movement experiences. A node in the cognitive map is an experience node defined as a tuple of the states of the pose cells \mathbf{PC} , the boundary cells V , and the pose estimate \mathbf{p} :

$$e_i = \langle \mathbf{PC}^i, V^i, \mathbf{p}^i \rangle. \quad (12)$$

After a period of robot movement, when a new boundary view is learned, a new experience node e^j , separate from the previous node e^i , is created based on the pose change $\Delta \mathbf{p}^{ij}$. Then the two experience nodes are connected with a directed edge from e^i to e^j based on the pose transition l_{ij} , which is defined as follows:

$$e_j = \langle \mathbf{PC}^j, V^j, \mathbf{p}^i + \Delta \mathbf{p}^{ij} \rangle, \quad (13)$$

$$l_{ij} = \langle \Delta \mathbf{p}^{ij}, \Delta t^{ij} \rangle.$$

When a learned boundary view is observed and loop closure is detected, a new transition between two existing experience nodes is established, which results in a closure in the cognitive map. The pose transition $\Delta \mathbf{p}$ of each node and edge is accordingly updated as follows to distribute the accumulated odometry error over the loop trail, and update the pose estimate $\mathbf{p} = (x, y, \theta)^T$ of the robot.

$$\Delta \mathbf{p}^i = a \left[\sum_{j=1}^{N_f} (\mathbf{p}^j - \mathbf{p}^i - \Delta \mathbf{p}^{ij}) + \sum_{k=1}^{N_t} (\mathbf{p}^k - \mathbf{p}^i - \Delta \mathbf{p}^{ki}) \right], \quad (14)$$

where a is a constant correction factor set to 0.5, N_f and N_t are the number of the outgoing and incoming edges of e_i . Hence, the cognitive map and the pose estimate \mathbf{p} are obtained and updated online as the output of the proposed LiDAR-based biologically inspired SLAM system.

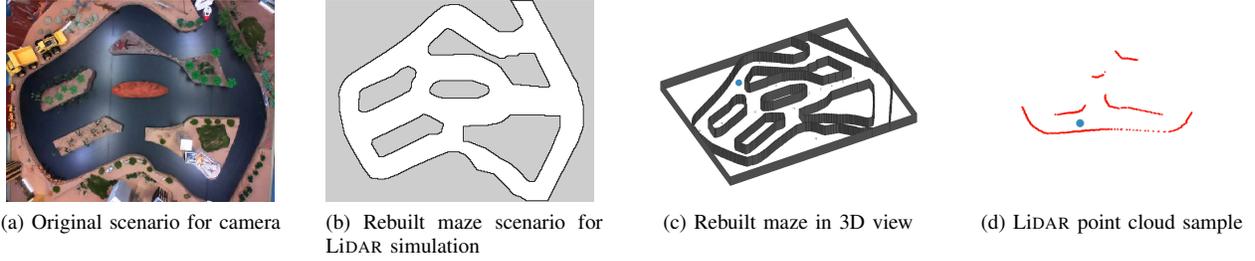


Fig. 3. Overview of the maze simulation scenario.

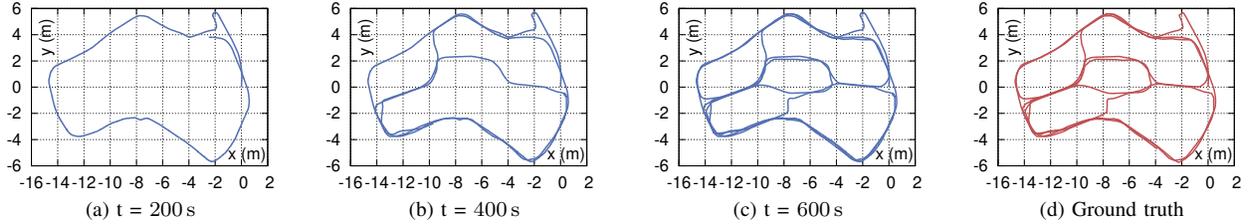


Fig. 4. Mapping progress in the maze simulation scenario.

IV. EXPERIMENTS

In this section, we present the experimental results for the proposed approach. To evaluate the LiDAR-based biologically inspired SLAM system, experiments are conducted in both simulation and real indoor scenarios to investigate the mapping and localization performance. The proposed SLAM system is implemented and tested on the *Robot Operating System (ROS)* [40]. We first report the experimental results of a simulation experiment to evaluate the mapping performance as well as the localization accuracy. Second, a mapping experiment in an indoor environment based on a ground robot is detailed. The results of the indoor experiment are illustrated and analyzed to demonstrate the applicability and accuracy of the proposed system in the real world.

A. Simulation Experiment

In the simulation experiments, a maze scenario is ported from the *iRat 2011 Australia* dataset [39], which is originally designed for the visual *RatSLAM*. Since the original dataset does not include LiDAR sensor data, it is not feasible to carry out the LiDAR SLAM experiments using the dataset. Therefore, we rebuild the maze scenario in the simulator to make it suited for LiDAR-based simulations.

Fig. 3 gives an overview of the maze scenario. Utilizing the *Stage* simulator in ROS, the point cloud data for a robot spawned in the maze scenario are obtained from a configurable virtual LiDAR in the simulator. At the same time, the ground truth information of the poses and movements of robots are also collected for evaluation purposes, but the odometry data is not involved in the SLAM process. Only the LiDAR point cloud data are used in the proposed method.

In the simulation experiment, the robot is controlled manually to navigate the maze for 600 seconds to evaluate the mapping performance with loop closure during long-term mapping. Fig. 4 illustrates the progressive cognitive maps built by the SLAM system in the simulation experiment and the ground truth path provided by the ROS Stage simulator.

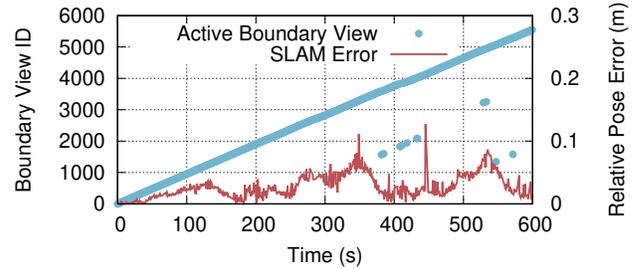


Fig. 5. Active boundary cell and relative pose error over time in simulation.

As shown in the mapping progress over time, the error in the cognitive map remains low over time. The mapping result of high similarity to the ground truth path demonstrates that SLAM performance is sufficiently accurate.

Fig. 5 reveals the activity of the boundary cells, which further reflects the loop closure detection and calibration of the cognitive map and the robot pose estimate. The number of boundary cells increases as the robot keeps moving to unexplored scenes, which are learned as the new boundary views. This is consistent with Fig. 5, in which the increase of the boundary view ID composes the main part of the scatter plot. The figure also reveals the loop closure detection during the mapping progress. The dramatic decrease in the active boundary view ID indicates that a number of previously learned boundary views are matched with the current LiDAR observation. Combining Fig. 4 and Fig. 5, it is obvious that during the mapping process, at around 380s the proposed SLAM managed to detect the previously learned views and performed loop closure calibration to reduce the accumulated error in the map, which demonstrates the ability of the boundary cells to detect loop closure in the mapping process.

In Fig. 5 we also report the relative pose error (RPE) [41] during the mapping process in the simulation experiment. The localization RPE computed based on the ground truth data is up to 0.127m at around 440s, with a mean value of 0.0284m for the 600 seconds of the mapping process in the $16.8\text{m} \times 12.6\text{m}$ space. As shown in the graph, the

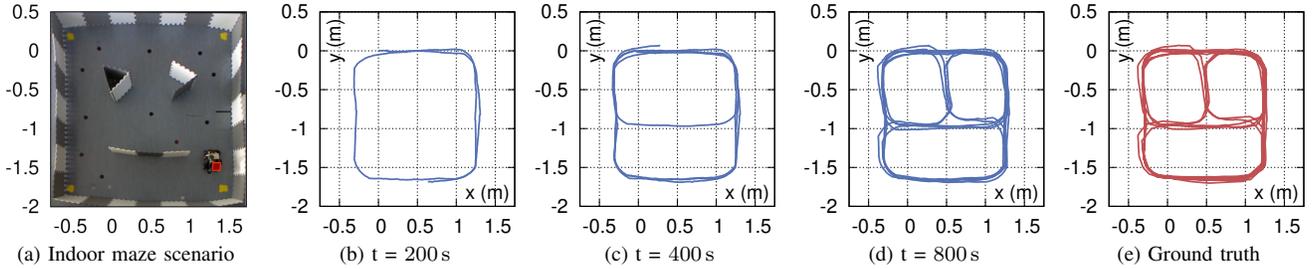


Fig. 6. Mapping progress in the real indoor maze environment.

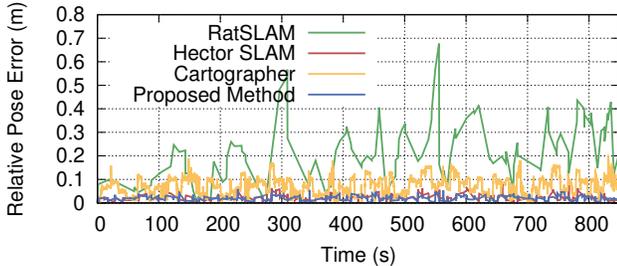


Fig. 7. Relative pose errors over time in the maze mapping process.

error remains at a low level for the whole mapping process. For several periods of time around 250 s to 350 s, and 480 s to 540 s, the position error started to accumulate but was soon corrected and remained low for a period of time. The transition of the accumulated error shows the process of loop closure detection and demonstrates the high performance of the proposed SLAM system for mapping and localization.

B. Indoor Experiment

To further evaluate the performance of the proposed SLAM system, experiments were conducted in a real indoor scenario, based on a ground robot platform. The ground robot is driven by four differential wheels which are controlled by an embedded *Raspberry Pi 3* single-board computer. A 2D *Hokuyo UTM-30LX-EW* LiDAR sensor is installed on the robot to obtain LiDAR point cloud for mapping experiments.

To specifically evaluate the mapping accuracy in the experiments, as shown in Fig. 6a, we built a maze experimental environment with an external localization and tracking system based on a bird view camera. Due to the limitations of the field of view and installation height for the camera, the size of the maze is relatively limited. In the indoor maze experiment, the robot was controlled to continuously navigate the maze for a specified period of time, while the bird view camera kept tracking the movements of the robot in real time to provide the ground truth for localization evaluation in parallel.

The progress of the mapping experiment for 900 s in the maze settings is illustrated in Fig. 6 from Fig. 6b to Fig. 6d, in which each cognitive map was taken at the annotated intervals during the mapping process. To further quantify the localization and mapping performance, in the indoor experiments, we additionally introduce two related SLAM algorithms for comparison purposes. To date, since it is rare that studies have focused on LiDAR-based biologically inspired SLAM algorithms, we evaluated *RatSLAM* [39],

which is based on a camera sensor and is one of the most commonly used open-source biologically-inspired SLAM system. In addition, the experiments also include two conventional LiDAR-based SLAM systems named *Hector SLAM* [5] and *Cartographer* [6], which are ones of the state-of-art indoor SLAM methods for indoor scenarios [3], [4]. We mainly evaluated the SLAM accuracy for the involved SLAM methods in the same experimental settings. In the experiments, relative pose errors (RPE) are computed based on the estimated poses given by the SLAM algorithms and the real-time ground-truth poses given by the bird view camera. Fig. 7 shows the position RPE curves of the proposed method, *RatSLAM*, *Hector SLAM*, and *Cartographer*, during the mapping process of the indoor maze environment. As shown in the figure, the error curve of *RatSLAM* fluctuated more dramatically than that of the proposed biologically-inspired method and the others. The position error of the proposed SLAM system is much stabler and lower than *RatSLAM* and is close to the state-of-art conventional SLAM methods. Interestingly, in this particular small-scale scenario, the error of *Cartographer* is slightly higher than *Hector SLAM*. The Root Mean Square Error (RMSE) for RPE of the proposed method for the estimated poses in the indoor experiments is 2.23 cm. In comparison, the RMSE of *RatSLAM*, *Hector SLAM*, *Cartographer* are 25.38 cm, 2.59 cm, and 7.97 cm. The results demonstrate that the accuracy of the proposed LiDAR-based biologically-inspired SLAM method outperforms camera-based *RatSLAM* and the performance is close to and slightly higher than that of the conventional LiDAR SLAM method *Hector SLAM*.

V. CONCLUSION

In this paper, we present a novel biologically-inspired SLAM system that mimics the mechanisms of boundary cells, place cells, and head direction cells, and leverages the LiDAR sensor to perform cognitive mapping and localization for indoor environments. With the self-motion cues and the boundary view cues from the LiDAR odometry and LiDAR boundary cells, the proposed SLAM system performs path integration and loop closure in the pose cell network to maintain and calibrate the robot pose estimates and build a cognitive map. Experimental results in simulation and the realistic indoor environment show that the proposed SLAM system is highly applicable and accurate in both scenarios and greatly outperforms *RatSLAM*, and is competitive with the state-of-art conventional LiDAR-based SLAM methods.

ACKNOWLEDGMENTS

This research has received funding from the Pazhou Laboratory Grants with No. PZL2021KF0020.

REFERENCES

- [1] S. Thrun, "Probabilistic robotics," *Communications of the ACM*, vol. 45, no. 3, pp. 52–57, 2002.
- [2] T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual slam algorithms: a survey from 2010 to 2016," *IPSP Transactions on Computer Vision and Applications*, vol. 9, no. 1, p. 16, 2017.
- [3] J. M. Santos, D. Portugal, and R. P. Rocha, "An evaluation of 2d slam techniques available in robot operating system," in *2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 2013, pp. 1–6.
- [4] M. Filipenko and I. Afanasyev, "Comparison of various slam systems for mobile robot in an indoor environment," in *2018 International Conference on Intelligent Systems (IS)*. IEEE, 2018, pp. 400–407.
- [5] S. Kohlbrecher, O. Von Stryk, J. Meyer, and U. Klingauf, "A flexible and scalable slam system with full 3d motion estimation," in *2011 IEEE international symposium on safety, security, and rescue robotics*. IEEE, 2011, pp. 155–160.
- [6] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2d lidar slam," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 1271–1278.
- [7] B. Yang, W. Luo, and R. Urtasun, "Pixor: Real-time 3d object detection from point clouds," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2018, pp. 7652–7660.
- [8] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.
- [9] A. Asvadi, P. Girão, P. Peixoto, and U. Nunes, "3d object tracking using rgb and lidar data," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 1255–1260.
- [10] E. I. Moser, E. Kropff, and M.-B. Moser, "Place cells, grid cells, and the brain's spatial representation system," *Annu. Rev. Neurosci.*, vol. 31, pp. 69–89, 2008.
- [11] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat," *Brain research*, 1971.
- [12] J. O'Keefe, "Place units in the hippocampus of the freely moving rat," *Experimental neurology*, vol. 51, no. 1, pp. 78–109, 1976.
- [13] J. S. Taube, R. U. Muller, and J. B. Ranck, "Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis," *Journal of Neuroscience*, vol. 10, no. 2, pp. 420–435, 1990.
- [14] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, "Microstructure of a spatial map in the entorhinal cortex," *Nature*, vol. 436, no. 7052, pp. 801–806, 2005.
- [15] M. E. Hasselmo, "Grid cell mechanisms and function: contributions of entorhinal persistent spiking and phase resetting," *Hippocampus*, vol. 18, no. 12, pp. 1213–1229, 2008.
- [16] T. Hartley, N. Burgess, C. Lever, F. Cacucci, and J. O'Keefe, "Modeling place fields in terms of the cortical inputs to the hippocampus," *Hippocampus*, vol. 10, no. 4, pp. 369–379, 2000.
- [17] C. Lever, S. Burton, A. Jeewajee, J. O'Keefe, and N. Burgess, "Boundary vector cells in the subiculum of the hippocampal formation," *Journal of Neuroscience*, vol. 29, no. 31, pp. 9771–9777, 2009.
- [18] J. R. Hinman, G. W. Chapman, and M. E. Hasselmo, "Neuronal representation of environmental boundaries in egocentric coordinates," *Nature communications*, vol. 10, no. 1, pp. 1–8, 2019.
- [19] Z. Bing, A. E. Sewisy, G. Zhuang, F. Walter, F. O. Morin, K. Huang, and A. Knoll, "Toward cognitive navigation: Design and implementation of a biologically inspired head direction cell network," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
- [20] J. H. Byrne, *Learning and memory: a comprehensive reference*. Academic Press, 2017.
- [21] Z. Bing, C. Meschede, K. Huang, G. Chen, F. Rohrbein, M. Akl, and A. Knoll, "End to end learning of spiking neural network based on r-stdp for a lane keeping vehicle," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 4725–4732.
- [22] M. Lechner, R. Hasani, A. Amini, T. A. Henzinger, D. Rus, and R. Grosu, "Neural circuit policies enabling auditable autonomy," *Nature Machine Intelligence*, vol. 2, no. 10, pp. 642–652, 2020.
- [23] Z. Bing, C. Meschede, G. Chen, A. Knoll, and K. Huang, "Indirect and direct training of spiking neural networks for end-to-end control of a lane-keeping vehicle," *Neural Networks*, vol. 121, pp. 21–36, 2020.
- [24] M. J. Milford, G. F. Wyeth, and D. Prasser, "Ratslam: a hippocampal model for simultaneous localization and mapping," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 1. IEEE, 2004, pp. 403–408.
- [25] M. J. Milford and A. Jacobson, "Brain-inspired sensor fusion for navigating robots," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 2906–2913.
- [26] S.-C. Zhou, R. Yan, J.-X. Li, Y.-K. Chen, and H. Tang, "A brain-inspired slam system based on orb features," *International Journal of Automation and Computing*, vol. 14, no. 5, pp. 564–575, 2017.
- [27] F. Yu, J. Shang, Y. Hu, and M. Milford, "Neuroslam: a brain-inspired slam system for 3d environments," *Biological Cybernetics*, vol. 113, no. 5–6, pp. 515–545, 2019.
- [28] M. J. Milford and G. F. Wyeth, "Mapping a suburb with a single camera using a biologically inspired slam system," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1038–1053, 2008.
- [29] B. Tian, V. A. Shim, M. Yuan, C. Srinivasan, H. Tang, and H. Li, "Rgb-d based cognitive map building and navigation," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 1562–1567.
- [30] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *Ieee Micro*, vol. 38, no. 1, pp. 82–99, 2018.
- [31] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, "The spinnaker project," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 652–665, 2014.
- [32] R. Kreiser, A. Renner, Y. Sandamirskaya, and P. Pienroj, "Pose estimation and map formation with spiking neural networks: towards neuromorphic slam," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2159–2166.
- [33] R. Kreiser, M. Cartiglia, J. N. Martel, J. Conradt, and Y. Sandamirskaya, "A neuromorphic approach to path integration: a head-direction spiking neural network with vision-driven reset," in *2018 IEEE international symposium on circuits and systems (ISCAS)*. IEEE, 2018, pp. 1–5.
- [34] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE transactions on Robotics*, vol. 23, no. 1, pp. 34–46, 2007.
- [35] J. Steckel and H. Peremans, "Batslam: Simultaneous localization and mapping using biomimetic sonar," *PloS one*, vol. 8, no. 1, p. e54076, 2013.
- [36] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, 1989.
- [37] C. Barry, C. Lever, R. Hayman, T. Hartley, S. Burton, J. O'Keefe, K. Jeffery, and N. Burgess, "The boundary vector cell model of place cell firing and spatial memory," *Reviews in the Neurosciences*, vol. 17, no. 1–2, pp. 71–98, 2006.
- [38] A. Samsonovich and B. L. McNaughton, "Path integration and cognitive mapping in a continuous attractor neural network model," *Journal of Neuroscience*, vol. 17, no. 15, pp. 5900–5920, 1997.
- [39] D. Ball, S. Heath, J. Wiles, G. Wyeth, P. Corke, and M. Milford, "Openratslam: an open source brain-based slam system," *Autonomous Robots*, vol. 34, no. 3, pp. 149–176, 2013.
- [40] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, 2009, p. 5.
- [41] R. Kümmerle, B. Steder, C. Dornhege, M. Ruhnke, G. Grisetti, C. Stachniss, and A. Kleiner, "On measuring the accuracy of slam algorithms," *Autonomous Robots*, vol. 27, no. 4, pp. 387–407, 2009.