IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received August 21, 2021, accepted September 13, 2021, date of publication September 27, 2021, date of current version October 6, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3113311*

# Biomimetic SLAM Algorithm Based on Growing Self-Organizing Map

**YUKUN ZHANG, MENGYUAN CHEN, DEHONG TIAN, AND LINGMEI DING**
School of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China
Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Anhui Polytechnic University, Wuhu 241000, China

Corresponding author: Mengyuan Chen (mychen@ahpu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61903002, in part by the University Synergy Innovation Project of Anhui Province under Grant GXXT-2021-050, in part by Anhui Polytechnic University-Jiujiang District Industry Collaborative Innovation Special Foundation under Grant 2021cyxtb8, and in part by the Middle-Aged and Top-Notch Talent Project of Anhui Polytechnic University.

**ABSTRACT** A Biomimetic SLAM Algorithm Based on Growing Self-Organizing Map (GSOM-BSLAM), inspired by spatial cognitive mechanism of mammalian hippocampus, is proposed to resolve uncertainty problems in location identification and lack of real-time performance in simultaneous localization and mapping. The algorithm connects activation characteristics of the place cell and neurons in the output layer of the neural network to construct a topological map of space using a self-organizing growable mapping neural network. It utilizes self-motion-aware information to obtain activation response of the place cell to estimate the robot position information, improving the localization accuracy and real-time performance of the system. Meanwhile, an accurate environmental cognitive map is finally created by incorporating color-depth images for closed-loop detection and error correction for spatial cell path integration. The proposed algorithm is validated using publicly available KITTI and St. Lucia datasets. The experimental results demonstrate that the proposed algorithm outperforms RatSALM by 37.8% and 36.5% in terms of localization accuracy and real-time performance, respectively, indicating good mapping capabilities.

**INDEX TERMS** Cognitive maps, place cells, grid cells, SLAM, GSOM.

## I. INTRODUCTION

Simultaneous Location and Mapping (SLAM) technology is critical to truly autonomous robot mobility, as it focuses on modeling the robot's environment during motion and estimating its own motion state [1-3]. According to the different sensors used by mobile robots, SLAM is mainly divided into two categories: laser SLAM and vision SLAM. The early disadvantages of visual SLAM, such as large computation and weak practicality, hindered its application. However, with the rapid development of computer technology and graphics processing unit (GPU), computational difficulties have gradually been overcome, prompting visual SLAM to become the mainstream of autonomous navigation and perception. Early vision SLAM followed the filtering methods used in mobile robotics, such as Kalman filter, extended Kalman filter, and particle filter, but graph optimization methods in computer vision have become mainstream [4]. More mature vSLAM systems based on graph optimization include LSD-SLAM [5] and ORB-SLAM2 [6]. The former is mainly based on photometric invariance and uses optical flow method to track camera pose. The latter uses the neighborhood information near the corner points to calculate the feature descriptors to complete feature matching and then construct the optimization problem to solve the state variables based on the data association of the feature points. However, both methods are prone to problems, including location recognition drift and lack of real-time when encountering scenes with too large a range or high complexity. Considering that animals can methodically perform environmental cognition and navigation in complex environments, including foraging, nesting, and so on. Therefore, simulating the neural structure and cognitive mechanisms of animals to enable mobile robots to perform environmental cognition and navigation like animals has received increasing attention.

Physiological studies have revealed some neuronal cells related to navigation behavior in the hippocampus and its surrounding areas of the mammalian brain, such as Place cell [7], Grid cell [8], Head direction cell [9], Stripe cell [10], etc. When animals explore their environment, these spatial

The associate editor coordinating the review of this manuscript and approving it for publication was Luigi Biagiotti.


This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

navigation cells take on specific active forms and provide knowledge necessary for performing navigation tasks to create cognitive maps. Animals can maintain relative spatial position relationships to nests or food while moving freely in a spatial environment, and computation-ally integrate these position points and update the relationships between these spatial position points in real-time through a specific organization in the brain, a process known as "Path Integration (PI)" [11]. Numerous neurological studies have also revealed that some navigational nerve cells in the mammalian hippocampus can represent specific location points in the environment, which, when associated, can form a "cognitive map," representing the topological relationships between spatial locations [12]. In the literature [13], the team of Australian scientist Milford investigated the Rat Simultaneous Location and Mapping (RatSLAM) method based on a rodent model that centralizes information from external visual scenes into pose cells through path integration and uses a visually driven navigation system to complete the map construction task. However, this model does not incorporate the physiological properties of hippocampal structures in the rat brain and does not have high localization accuracy and stability [14]. Tian *et al.* [15] improved RatSLAM model and proposed a cognitive map construction and navigation method for mobile robots based on RGB-D sensors using CAN (Competitive Attractor Network) model.

The feature of identical grid cell firing characteristics in the same region of dorsal aspect of internal olfactory cortex and incremental changes in firing characteristics in different regions was proposed for constructing cognitive maps based on spatial location relationships at arbitrary scales. However, the model only expresses the firing activity of grid cells and generates more grid cells, making it challenging to accomplish precise localization in unknown environments and real-time map construction [16].

Kohonnen [17] proposed a Self-Organizing Map (SOM) based on the phenomenon of lateral inhibition and introduced a "winner-takes-all" competitive learning rule. The prerequisite for this Self-Organizing Map learning method to achieve topology mapping is to define its initial network structure. Unfortunately, this method cannot be adapted to non-stationary data sets. Montazeri *et al.* [18] suggested an algorithm that applies GSOM neural networks to reinforcement learning to achieve an optimal state-space representation through two growing Self-Organizing Maps. The significant advantage of this algorithm is its ability to process online data in real-time using adaptive mechanisms.

The above-mentioned approaches for autonomous localization and navigation of mobile robots have paid little attention to interactions between individual spatial navigation cells and how to exploit their properties for precise localization and map construction of mobile robots in unknown environments. Therefore, in this paper, we construct a model of each navigation cell, simulate the firing characteristics of each spatial navigation cell in the hippocampus of rat brain, and use GSOM neural network characteristics to correlate environmental information with location information characterized by the place cell to construct an environmental cognitive map and realize real-time localization and composition of the mobile robot in the environment. To summarize, the main contributions of the presented work are:

- The grid cell activity values of different spatial scales are used as the input of the GSOM neural network, and the winning neurons are obtained through its competition mechanism-synergy mechanism-adaptive mechanism. The winning neurons are associated with visual perception templates for constructing environmental cognitive maps. At the same time, the activation response of the place cell is obtained using the self-motion perception information for estimating the mobile robot position information, which improves the localization accuracy and real-time performance of the system.
- Closed-loop detection using color-depth images achieves error correction for spatial cell path integration, improving the accuracy of closed-loop detection and the global consistency of map construction.

## II. RELATED WORK

Hafting *et al.* discovered a periodic firing neuronal cell that can form a highly stable hexagonal firing field, called "Grid cell" [19], to characterize specific locations in space by changing the experimental environment and its firing characteristics, as shown in Figure 1. Typically, four parameters are calibrated for grid cells: spacing,orientation, phase (x, y), and firing field radius r. It has been demonstrated that connections between grid cells are inhibitory and that the activity value of grid cells is not related to exogenous information but only to their own motion state [20], [21].

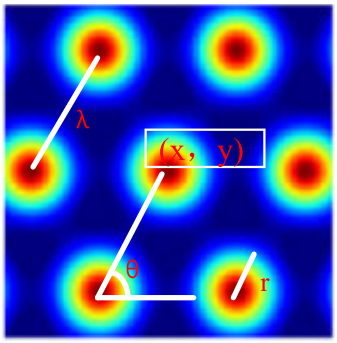



**FIGURE 1.** Grid cell firing characteristics.

Grid cells generate grid fields primarily through oscillatory interference models [22]–[25] and continuous attractor models [26], [27]. In this paper, the spatial environment to be explored is discretized, and the oscillatory interference model is used to construct multi-scale grid cells to characterize environmental location points in various directions and angles; the activity value of the entire grid cell cluster in the spatial environment is calculated to provide spatial metric information for the next step of location estimation and spatial characterization based on place cells.

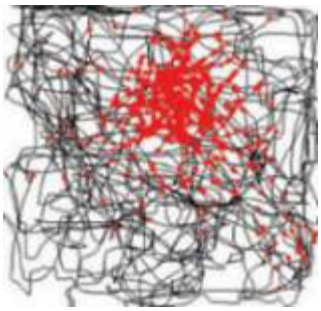


**FIGURE 2.** Place cell firing characteristics. [29].

Dostrovesky *et al.* identified "Place cells" in the hippocampus with spatial localization functions as an essential part of brain's navigation and spatial representation network, i.e., when rats are in a specific location in the spatial environment, these place cells become abnormally active, and their active range is called the "place field" [7]. Place cells can characteristically encode relative spatial locations and are the basic unit for constructing cognitive maps of intracerebral environment, and their firing activity provides a continuous and dynamic representation of spatial location. The joint response of numerous place cells produces a discrete representation of the environment, suggesting that the place cell map is essentially a landmark map [28]. The place cell firing characteristics and response map are illustrated in Figure 2, where the black line is the motion trajectory, and the red dots are the place cell response points.

In summary, grid cell firing characteristics are only correlated with animal's self-motion perception information, and a one-to-one correspondence exists between the firing fields of place cells and specific locations in the spatial environment. Herein, we propose a GSOM-BSLAM algorithm based on the interactions existing between each spatial navigation cell and GSOM neural network characteristics, solving uncertainty problems of location identification and insufficient real-time in simultaneous localization and mapping. The proposed algorithm is validated with publicly available KITTI and St. Lucia datasets, indicating a 37.8% improvement in localization accuracy and 36.5% improvement in time efficiency than RatSALM and exhibiting good composition capability.

## III. SYSTEM FRAMEWORK

The framework of GSOM-BSLAM algorithm proposed in this work is shown in Figure 3. The algorithm consists of three parts: path integration, spatial exploration learning, and environment cognitive map construction. First, the mobile robot learns to generate memory points by exploring arbitrarily in the environment and inputs real-time perceived self-motion state information (velocity and direction information) into the grid cell path integration model to construct multi-scale grid cells and calculate the activity value of grid cells. Subsequently, multi-scale grid cell activity values are used as input of GSOM neural network to characterize specific spatial location points to construct a topological map of the spatial environment. Finally, error correction is performed on the spatial cell path integral by introducing color-depth images for closed-loop detection to construct an accurate cognitive map of the environment.

## IV. MULTI-SCALE GRID CELL TO PLACE CELL MODEL BUILDING

### A. MULTI-SCALE GRID CELL PATH INTEGRATION MODEL

The mobile robot explores the experimental environment and records the self-motion perception information in real-time,

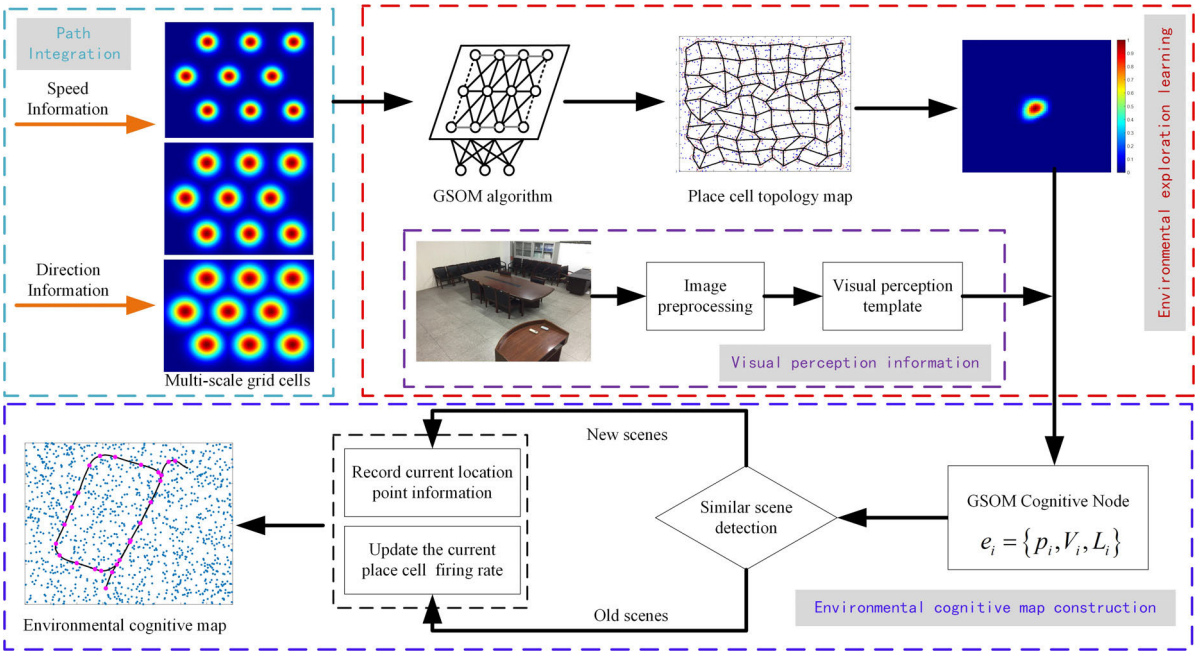



**FIGURE 3.** The framework of GSOM-BSLAM algorithm.

constructs multi-scale grid cells using an oscillating interference model [22]–[25], and calculates the grid cell activity values at different locations by equation (1) [22].

$$R_{GC}(r,\lambda,\theta,r_i) = g(\sum_{k=1}^{3}\cos(\frac{4}{\sqrt{3}\lambda}\overrightarrow{u}(\theta_k-\theta)(r-r_i))) \quad (1)$$

where $r = [x, y]$ is the current position of mobile robot in the spatial environment; $\lambda$, $\theta$ and $r_i$ are the spacing, orientation, and phase of grid cells, respectively; $i = 1, 2, \cdots, N_{GC}$, $N_{GC}$ is the total number of grid cells; and $\overrightarrow{u}(\theta_k - \theta)$ is the projection along with different directions, $k = 1, 2, 3$, $\theta_k = \{-60°, 0, 60°\}$. A Gaussian distribution function $g(x) = \exp[a(x-b)-1]$ is introduced to process the grid cells with low activity value using its localization filtering property to obtain a stable grid cell clusters, where $a = 0.3, b = -1.5$. By adjusting $\lambda$, $\theta$ and $r_i$ in Eq. 1, grid cells with different spacing, orientation, phase as well as their corresponding activity values can be obtained.

### B. GRID CELL TO PLACE CELL COMPETITION MODEL

The GSOM algorithm [30], [31] strictly adheres to the competitive learning rule, where only 1 output neuron is activated at 1 moment in the competitive learning process. In this work, the input of the GSOM algorithm is the grid cell activity values at different scales, and the output is the response values of the place cells.When several grid cell activity values at different scales are simultaneously input to the GSOM neural network, the competition among grid cells causes the connection weights between the most excitatory place cell and environmental information to be enhanced, and the connection weights between neighboring place cells and location to be weakened, thus producing place cells that respond to a specific location. The GSOM algorithm learning process is shown in Algorithm 1.

#### 1) COMPETITION MECHANISM OF GSOM ALGORITHM

The competition principle of GSOM algorithm is shown by equation (2) [30].

$$O_{i(g^*)}(g^*) = \min\left\|g^* - \vartheta_j^T\right\| \quad (2)$$

The place cell $O_{i(g^*)}$ with the minimum Euclidean distance from stimulus $g^*$ is the winning cell of stimulus $g^*$. Also, the winning GSOM neuron $O_{i(g^*)}$ is the excitation center of GSOM map when stimulated by $g^*$, where stimulus $g^*$ represents the grid cell firing rate input and $\vartheta_j$ is the feedforward connection relation of the jth winning cell.

#### 2) SYNERGISTIC MECHANISM OF GSOM ALGORITHM

The winning neuron $O_{i(g^*)}$ establishes a self-centered topological neighborhood in which the neurons within a certain neighborhood excite together. The domain function is commonly expressed in Gaussian function equation (3).

$$\vartheta_{i(g^*)_j} = \exp\left\{-\frac{d^2_{i(g^*)_j}}{2\rho^2}\right\} \quad (3)$$

$d^2_{i(g^*)_j}$ denotes the distance between node $i(g^*)$ and node j. The synergistic neighborhood of GSOM algorithm shrinks with time, i.e., the effective radius $\rho$ of synergistic neighborhood decays exponentially with time t.

$$\rho(t) = \sigma_0 \exp(-\kappa_\sigma t) \quad (4)$$

where $\sigma_0$ is the decay factor, set to 0.8 to adjust the topological connection relationship between nodes. The time adjustment constant $\kappa_\sigma$, set to 0.5, prevents the topological connection relationship between place cells from becoming weaker when the speed of mobile robot slows down, resulting in a prolonged time.

**Algorithm 1** GSOM Learning Algorithm

Input: Grid cell activity values at different scales
Output: Place cells response values
Initialize: N (Number of GSOM neurons),
$\omega$(Connection weights between GSOM neurons)
T (Maximum number of iterations)。
for t = 1:1:T do
for Sample S (Sample vector x) do
Step1 Competition Stage

$$O_{i(g^*)}(g^*) = \min\left\|g^* - \vartheta_j^T\right\|$$

Step2 Synergy Stage

$$\vartheta_{i(g^*)_j} = \exp\left\{-\frac{d^2_{i(g^*)_j}}{2\rho^2}\right\}$$

Step3 Adaptive stage

$$\begin{cases}\omega_{ij}(t) = \alpha_0\left(g^* - \vartheta_j(t)\right)\vartheta_{i(g^*)_j}(t)\\ \vartheta_j(t+1) = \vartheta_j(t) + \Delta\vartheta_j(t) \quad (j = 1, 2, \cdots N)\end{cases}$$

end for

$$\rho(t) = \sigma_0 \exp(-\kappa\sigma t)$$

end for

#### 3) GSOM ALGORITHM ADAPTIVE MECHANISM

The adaptive mechanism of GSOM algorithm is synaptic adaptive mechanism. The self-organization process of regulating synaptic connection weights $\omega_{ij}(i = 1, 2, \cdots n; j = 1, 2, \cdots N)$ of GSOM neurons, $\omega_{ij}$ denotes the connection weight of ith grid cell input to the jth place cell.

The adaptive law of GSOM is as follows[31]:

$$\begin{cases}\omega_{ij}(t) = \alpha_0(g^* - \vartheta_j(t))\vartheta_{i(g^*)_j}(t)\\ \vartheta_j(t+1) = \vartheta_j(t) + \Delta\vartheta_j(t)(j = 1, 2, \cdots N)\end{cases} \quad (5)$$

where $\alpha_0$ is the learning rate adjustment constant, set to 0.9 to adjust the learning rate of GSOM algorithm.

### C. VISUAL IMAGE PROCESSING AND CLOSED-LOOP DETECTION

Path integration based on Self-motion aware information is subjected to cumulative error over a wide range of motion space, and studies have indicated that rats correct for this error when they encounter familiar scenes. The proposed algorithm corrects for the cumulative error of path integration by introducing color-depth images to construct a visual perceptual template for closed-loop detection.

We herein employ an image matching algorithm to achieve closed-loop detection during motion, and image matching is performed using intensity distribution of scan lines in color and depth images. The scan line intensity distribution is a one-dimensional vector resulted from summing and normalizing the intensities of all pixel columns of a grayscale image. Figures 4(a) and 4(b) display the color image and depth image of the same scene, respectively.

The average absolute intensity difference between intensity distributions of image scan lines is called the intensity offset and is denoted by $g(c)$.

$$g\left(c, I^j, I^k\right) = \frac{1}{b-|c|}\left(\sum_{i=1}^{b-|c|}\left|I^j_{i+\max(c,0)} - I^j_{i-\min(c,0)}\right|\right) \tag{6}$$

where $I^j$ and $I^k$ are intensity distributions of image scan lines being compared, $c$ is the distribution offset, and $b$ is the image width.

In order to reduce the influence of illumination on image acquisition and matching, and to improve the matching accuracy of images in different environments, we use both color and depth image to match simultaneously to determine the absolute position. Since the actual spatial environment varies in light intensity at different time periods, this paper assigns different weights to the difference between offsets of depth and color images to obtain the matching degree measure $G$ of the image.

$$G = \mu_R\left|g_{iR}(c) - g(c)\right| + \mu_D\left|g_{iD}(c) - g(c)\right| \tag{7}$$

where $\mu_R$ and $\mu_D$ are the weights of color and depth images, respectively, and $\mu_R + \mu_D = 1$. $g_{iR}(c)$ and $g_{iD}(c)$ are the scanned intensity distributions of the color and depth images, respectively. The minimum offset $c_m$ of $I^j$ and $I^k$ pixels in a continuous image is the minimum value of the matching degree measure for the two images.

$$c_m = \min_{c\in[\rho-b,b-\rho]}(G) \tag{8}$$

where the bias $\rho$ ensures that the two images exhibit an overlap amount. The similarity threshold of the image is set to $c_t$, when $c_m < c_t$, the current view is a new view and saved to the visual template set $\{V_i\}$, and when $c_m \geq c_t$, it is considered to return to a familiar scene.

## V. ENVIRONMENTAL COGNITIVE MAP CONSTRUCTION

An environmental cognitive map is constructed by associating competing winning place cells with visual perceptual templates according to GSOM neural network, providing a spatial coordinate system for carrier motion. Each cognitive node $e$ with a topological relationship contains the activity value $p$ of a place cell, the visual perceptual template is associated with the place cell C, and the topological relationship $L$ is between cognitive nodes. A single cognitive node is defined as:

$$e_i = \{p_i, V_i, L_i\} \tag{9}$$

A new cognitive node is created on the map when the location metric of current cognitive node exceeds the cognitive threshold or when a new visual perceptual template appears. The new cognitive node is represented as:

$$e_i = \left\{p_i, V_i, L_i + \Delta L_{ij}\right\} \tag{10}$$

where $\Delta L$ is the spacing between the current cognitive node and the previous cognitive node. When the visual template detects a return to the familiar scene, a closed-loop verification is performed using principles of temporal and geometric consistencies to ensure the accuracy of map construction, and all experiences are updated at the closed-loop when the verification is successful[13].

$$\Delta L_{ij} = \rho\left[\sum_{j=1}^{N_f}\left(L_j - L_i - \Delta L_{ij}\right) + \sum_{k=1}^{N_t}\left(L_k - L_i - \Delta L_{ki}\right)\right] \tag{11}$$

where $\rho = 0.6$ is the cognitive speed constant that determines the learning speed of mobile robots to the environment. $N_f$ is the number of current cognitive nodes connected to other nodes in the cognitive map. $N_t$ is the number of connections of other cognitive nodes to the current node $e_i$. The process of environmental cognitive map construction is shown in Figure 5.

## VI. EXPERIMENTAL RESULTS AND ANALYSIS

The robot learns and memorizes the 100∗100 unknown spatial environment, generating the corresponding multi-scale grid cell firing field covering the whole two-dimensional environment. The running speed of the running body is no more than 1 m/s, and the position update period is 2s. The speed of the running body remains the same in the same cycle, and the speed changes randomly in different cycles. When the runner reaches the set spatial boundary, it runs in the direction of mirror reflection from the original direction. The detailed simulation experiment parameters are shown in Table 1.

### A. ANALYSIS OF MULTI-SCALE GRID CELL COMPUTATIONAL MODELS

From a biological point of view, we simulate the free movement of rats in the environment and generate multi-scale grid cells covering the whole environment by integrating self-motion information to learn and remember the spatial environment information. Figure 5 depicts the firing of some grid cells after normalization, where the maximum firing rate

(a) Color image. (b) Depth image.

**FIGURE 4.** Color image and depth image of the same scene.

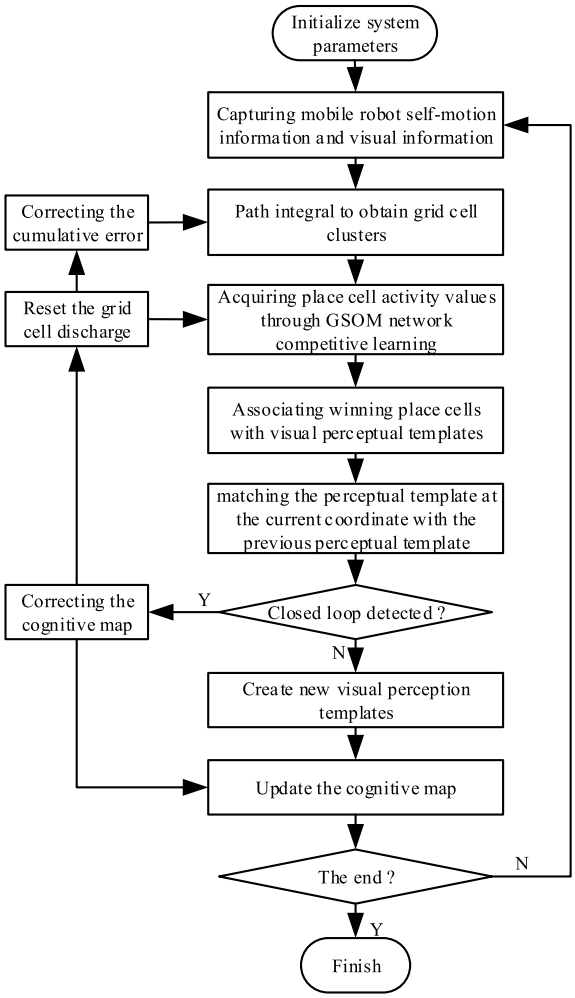



**FIGURE 5.** Solution for constructing environmental cognitive maps.

is located at the center of the circle, and the lowest firing rate is at the circle edge. The simulation results indicate that the input of different parameters of self-motion information can generate grid cells of different scales, and the firing pattern demonstrates a periodic grid-like structure, manifesting similar results with firing characteristics of biological grid cells.

**TABLE 1.** Simulation experiment parameters setting.

| Description | Parameter | Numeric value/range |
|---|---|---|
| Speed Information | $v$ | 0.1~1 m/s |
| Direction Information | $\theta$ | $[0, 2\pi]$ |
| Grid cell spacing | $\lambda$ | $[1m, 10m]$ |
| Speed balance factor | $\gamma$ | 0.01 |
| Gain constants | $\xi$ | 0.5 |
| Activate initial weights | $\omega_{ij}$ | 0.01 |
| Learning rate | $\varepsilon$ | 0.01 |
| Activating regulators | $\sigma^2$ | 0.05 |
| Training samples | $N_{sample}$ | 900 |
| Maximum distance | $d_{\max}$ | 0.2 |

### B. ANALYSIS OF TOPOLOGICAL MAP CONSTRUCTION BASED ON GSOM

Vision sensors collect environmental information in the location space area and obtain information about the location of the robot's passable area. The projection position of the camera on the ground at the starting point of the mobile robot is used as the world coordinate origin. During the robot's exploration of the environment, the coordinates of the robot's current position are transmitted back to the robot at the same time intervals as the training sample points of the GSOM network.

Topology maps of the environment structure were built using both a learning-based method (fused GSOM) and a non-learning method (non-fused GSOM algorithm). The performance evaluation index of the experiment is the localization error and localization time. Figures 8(a) and 8(c) show the topological map construction process of the

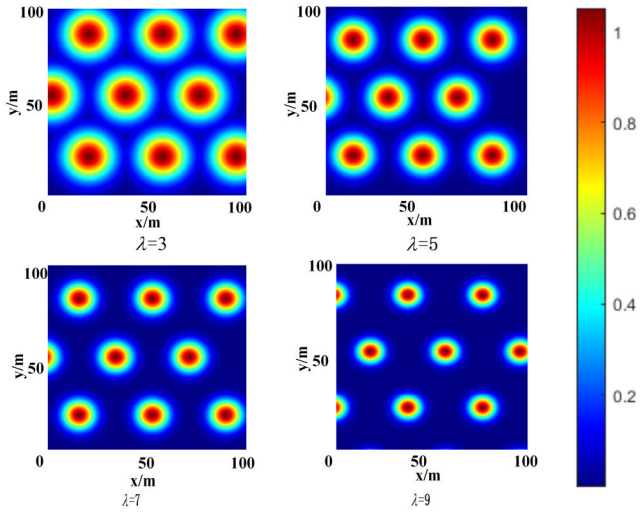



**FIGURE 6.** Partial grid cell firing rate at $\theta = 60°$.

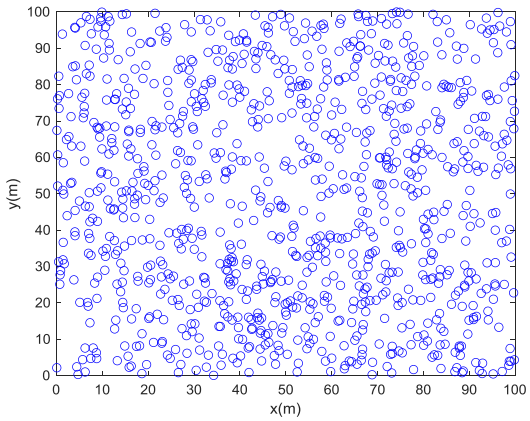



**FIGURE 7.** Environmental sample distribution map.

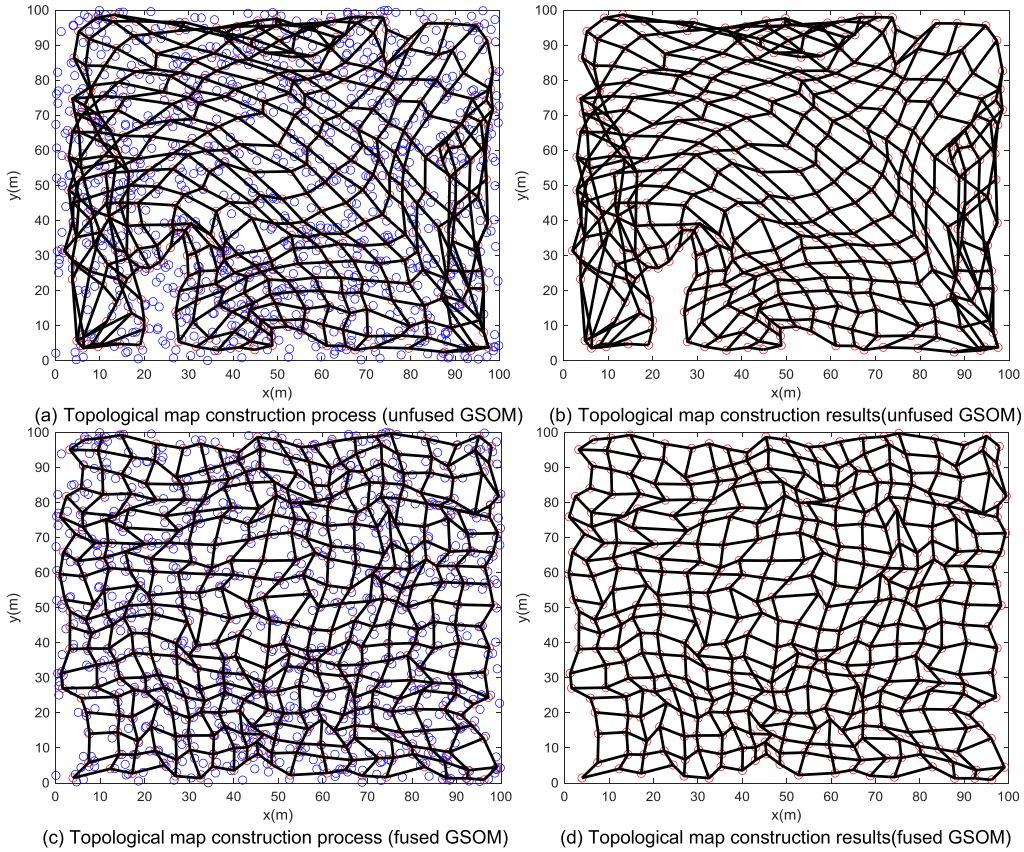
unfused GSOM algorithm (non-learning method) and the fused GSOM algorithm (learning-based method), respectively, where the red circles indicate the competing winning neurons, the blue circles are the environmental sample collection points. The solid black lines stipulate the connection weights between each winning neuron. Figures 8(b) and 8(d) correspond to the topological map construction results of the unfused GSOM algorithm and the fused GSOM algorithm. The comparison reveals that the topological map construction result after incorporating the GSOM algorithm is closer to the sample spatial distribution. This phenomenon is due to the fact that the GSOM algorithm as a learning-based method can adjust the topological relationship between cognitive points using the coordinated action of competition mechanism-cooperative mechanism-adaptive mechanism during the learning process, which gives the algorithm a flexible structure to process data and non-fixed data sets online. In contrast, the traditional non-learning-based topology map construction method does not have dynamic learning capability, and its topology map construction time increases with the exploration time of the mobile robot. At the same time, the newly recorded data will affect the topological structure of the built map, resulting in a relatively poor representation of the spatial environment in the final topological map.

Table 2 gives the measured localization time and localization error during environmental learning of the mobile robot for the above two schemes. It can be seen that the learning-based method which fuses GSOM has a greater advantage over the non-learning-based topological map construction method (unfused GSOM) in both localization time and localization accuracy after training competitive learning. The data comparison shows that the localization time after fusing the GSOM algorithm is reduced by about 17.2% and the localization error is reduced by about 19.8% compared to that without the GSOM algorithm. The reason for this progress is that the fused GOSM algorithm allows each winning neuron to better characterize a specific spatial location in the environment, which provides accurate location metric information for the construction of environmental cognitive maps. At the same time, the GOSM algorithm has the ability to characterize the current spatial information by using a few place cell firing responses, which improves the computational efficiency of the algorithm to a certain extent, so that the computational time consumption of the mobile robot in the process of environment learning can still be effectively controlled even if the number of samples gradually increases.

### C. EXPERIMENTAL ANALYSIS OF ENVIRONMENTAL COGNITIVE MAPPING BASED ON GSOM ALGORITHM

The results of environmental cognitive map construction based on GSOM competitive learning network are depicted in Figure 9. The experimental scenario is publicly available KITTI dataset 06 sequence, where Figure 9 (a), (b), and (c) show the environmental cognitive maps constructed with scale taking values of 2, 5, and 8, respectively, where the red nodes are place cell response points, the blue lines are detected closed-loop scenes, and the green lines are detected false match scenes. From the comparison of the three figures, when the scale is small, it is easy to confuse node matching due to numerous constructed place cells, resulting in more false matches during closed-loop detection. When the scale is too large, the constructed place cells are insufficient to express the current scene, while when the scale is 5, the constructed place cells are more evenly distributed and can better express the current environmental scene. Consequently, the environmental cognitive map constructed with a scale value of 5 exhibits fewer false matches than the other two.

Figure 10 presents a comparison of localization error at different scales, and Figure 11 indicates closed-loop precision-recall analysis at different scales. From Figures 10 and 11, when the scale takes the value of 5, the corresponding localization error is the lowest, and the accuracy rate is the highest, so the scale value is set to 5 in this work. The experimental results demonstrate that setting appropriate scale values, using GSOM algorithm to characterize spatial environment information, outputting in the form of response value of place cells, and activating corresponding place cells by visual

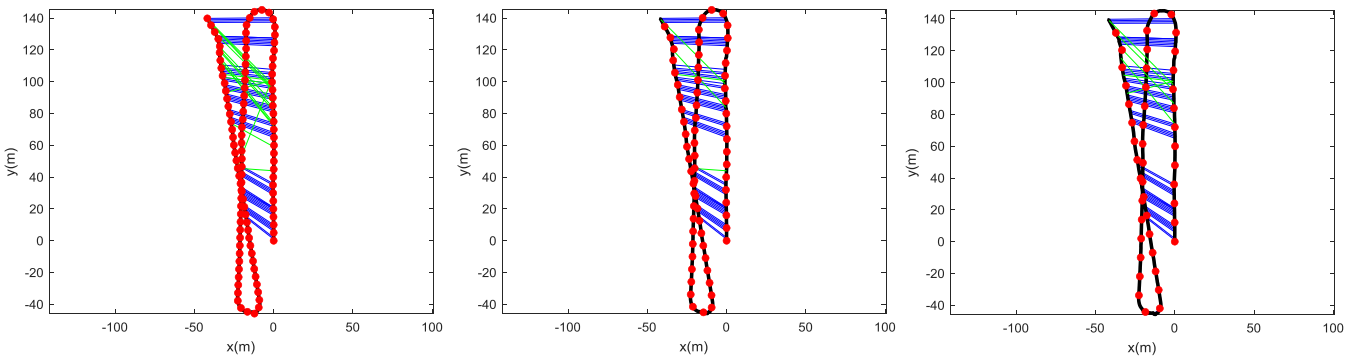



(a) Topological map construction process (unfused GSOM)
(b) Topological map construction results(unfused GSOM)
(c) Topological map construction process (fused GSOM)
(d) Topological map construction results(fused GSOM)

**FIGURE 8.** Topology map construction process and results.

**TABLE 2.** Experimental results of spatial characterization.

| Number of samples/pcs | localization run time/s | | localization error/m | |
|---|---|---|---|---|
| | Non-learning-based method (unfused GSOM) | Learning-based method (fused GSOM) | Non-learning-based method (unfused GSOM) | Learning-based method (fused GSOM) |
| 100 | 16.56 | 14.08 | 3.44 | 3.01 |
| 300 | 18.45 | 16.23 | 3.21 | 2.65 |
| 500 | 19.23 | 17.35 | 1.45 | 1.12 |
| 700 | 21.35 | 19.35 | 2.35 | 1.83 |
| 900 | 25.24 | 21.62 | 2.67 | 2.24 |




(a) scale value=2
(b) scale value=5
(c) scale value=8

**FIGURE 9.** Environmental cognitive maps at different scales.

cues to correct grid cell firing patterns, thus correcting path integration, can effectively solve the problem of cumulative errors occurring during motion, making the cognitive paths more consistent with learning trajectories and ensuring the accuracy and stability of environmental cognitive maps.

**TABLE 3.** Error analysis table for the trajectories of the three algorithms.

| Sequence | ORB-SLAM2 | | | | RatSLAM | | | | GSOM-BSLAM | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | median | RMSE | STD | mean | median | RMSE | STD | mean | median | RMSE | STD |
| 00 | 0.213 | 0.148 | 0.116 | 0.098 | 0.427 | 0.532 | 0.206 | 0.325 | **0.034** | **0.021** | **0.076** | **0.012** |
| 01 | 0.221 | 0.153 | 0.155 | **0.015** | 0.158 | 0.187 | 0.172 | 0.112 | **0.046** | **0.032** | **0.082** | 0.021 |
| 02 | 0.156 | 0.087 | 0.102 | 0.032 | 0.189 | 0.201 | 0.145 | 0.156 | **0.034** | **0.034** | **0.061** | **0.015** |
| 03 | 0.287 | 0.235 | 0.136 | 0.123 | 0.156 | 0.186 | 0.161 | 0.094 | **0.094** | **0.104** | **0.116** | **0.086** |
| 04 | 0.235 | 0.186 | 0.135 | 0.087 | 0.122 | 0.155 | 0.142 | 0.034 | **0.051** | **0.043** | **0.092** | **0.032** |
| 05 | 0.136 | 0.248 | 0.106 | 0.112 | 0.213 | 0.354 | 0.133 | 0.126 | **0.083** | **0.085** | **0.085** | **0.065** |
| 06 | 0.135 | 0.097 | 0.065 | **0.035** | 0.352 | 0.341 | 0.084 | 0.121 | **0.065** | **0.065** | **0.062** | 0.056 |
| 07 | 0.231 | 0.154 | 0.112 | 0.113 | 0.124 | 0.235 | 0.146 | 0.089 | **0.078** | **0.063** | **0.084** | **0.035** |
| 08 | 0.182 | 0.135 | 0.102 | 0.086 | 0.156 | 0.234 | 0.145 | 0.112 | **0.069** | **0.067** | **0.079** | **0.051** |
| 09 | 0.321 | 0.215 | **0.023** | 0.125 | 0.532 | 0.641 | 0.079 | 0.356 | **0.053** | **0.049** | 0.035 | **0.025** |
| 10 | 0.084 | **0.011** | **0.021** | **0.020** | 0.156 | 0.257 | 0.063 | 0.130 | **0.043** | 0.015 | 0.041 | 0.027 |
| Average | 0.200 | 0.152 | 0.098 | 0.077 | 0.235 | 0.302 | 0.134 | 0.150 | **0.059** | **0.053** | **0.074** | **0.038** |

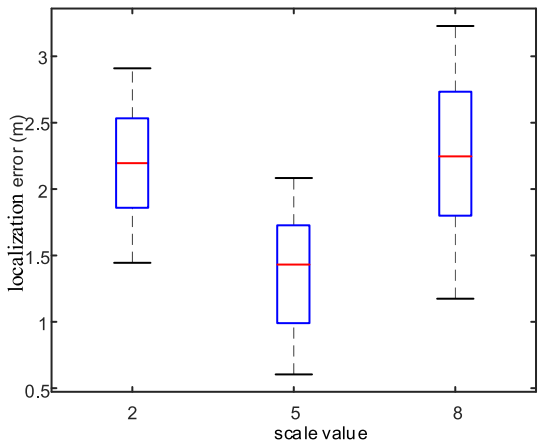



**FIGURE 10.** Plot of localization error at different scales.

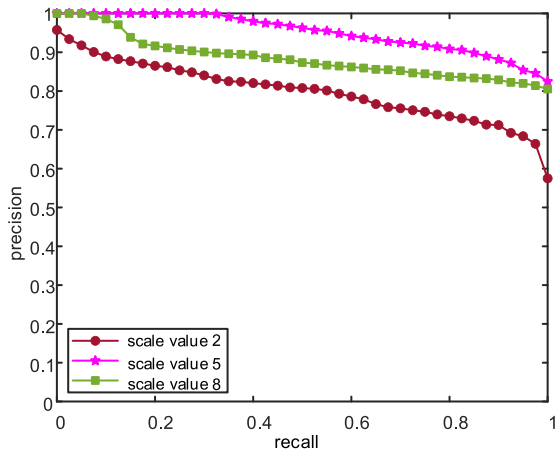



**FIGURE 11.** Precision-recall at different scales.

### D. EXPERIMENTAL COMPARISON ANALYSIS WITH OTHER ALGORITHMS

#### 1) KITTI DATASET EVALUATION

In this paper, the proposed GSOM-SLAM algorithm is evaluated and tested with ORB-SLAM2 and RatSLAM algorithms on publicly available KITTI datasets 00, 02, 05, and 08 sequences. Ground truth is available as the reference trajectory in KITTI dataset. The results of trajectory runs of the three algorithms in the four sequences are displayed in Figure 12. The figure manifests that trajectory maps obtained by a proposed algorithm are closer to real trajectories. Table 3 shows the mean error, root mean square error, median, and standard deviation analysis of trajectory localization obtained by the three compared algorithms. The table also shows that the proposed algorithm outperforms ORB-SLAM2 and RatSLAM in most sequences, with RatSLAM having the worst overall performance. This is because RatSLAM algorithm compresses the image before performing image matching (converting the acquired color image to grayscale image for dimensionality reduction). The RatSLAM method of image matching by absolute difference summation is greatly affected by illumination and has low closed-loop accuracy. The ORB-SLAM2 algorithm is the mainstream visual SLAM algorithm. However, it also suffers from using only point features for pose update, which is susceptible to sparse environmental texture features, too fast motion, and curved motion, resulting in low localization accuracy and difficult pose estimation. Our algorithm uses color-depth images for closed-loop detection, which effectively reduces light intensity impact on matching and improves closed-loop accuracy. Concurrently, we correlate activation characteristics of place cell with neuron response of output layer of neural network, construct a topological map by GSOM algorithm, and activate the place cell firing response using self-motion perception information to estimate mobile robot position information. From the experimental results, the mobile root means square error of algorithm is reduced by 24.5% compared with ORB-SLAM2 and 44.8% compared with RatSLAM, which significantly improves the localization accuracy of robot in an unknown environment and demonstrates a good composition capability.

#### 2) ST. LUCIA DATASET EVALUATION

To further evaluate the performance of GSOM-BSLAM algorithm, simulation experiments were conducted using publicly available dataset St. Lucia, which was collected from cars driving on the track around the University of Queensland St. Lucia campus. Figure 13(a) shows a plan view of the data

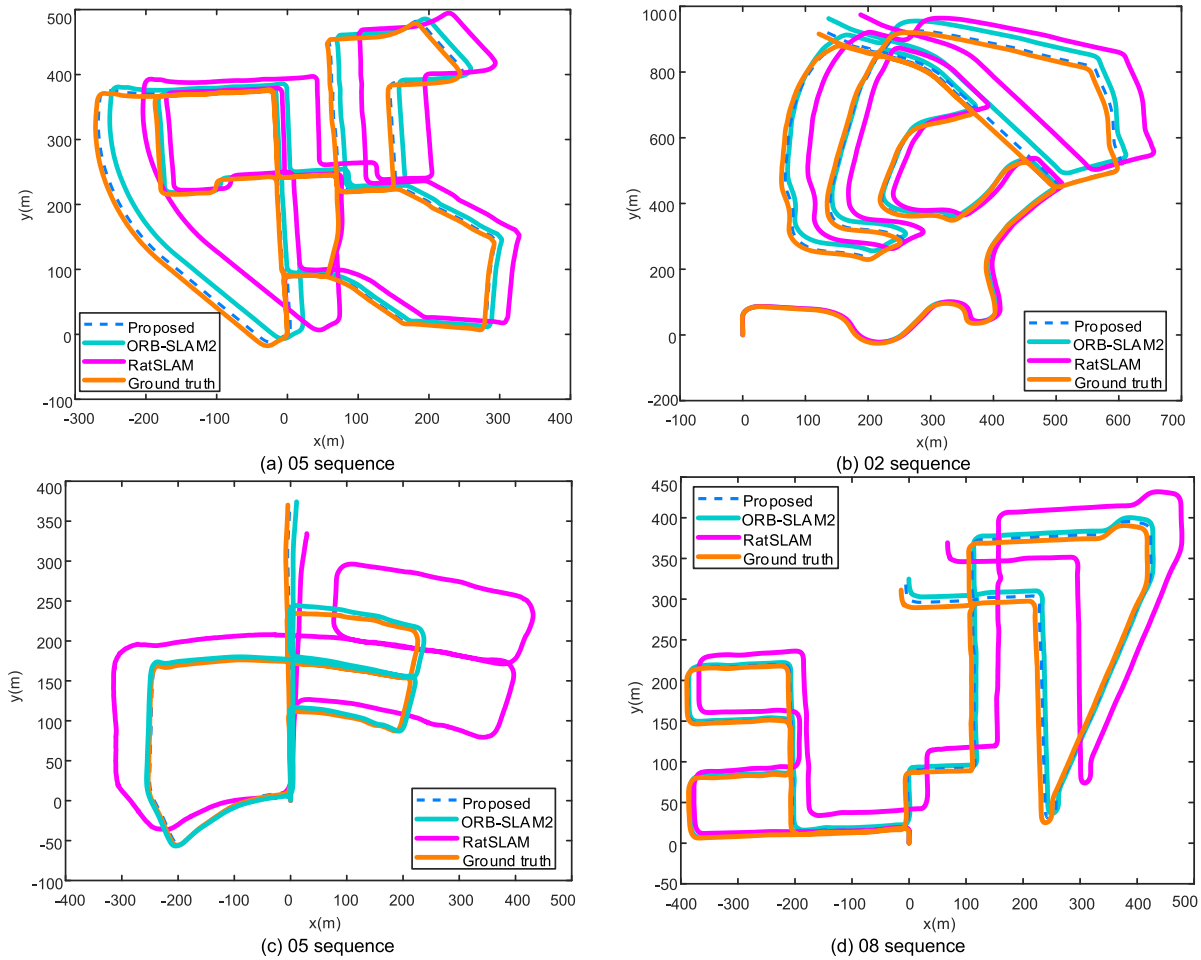
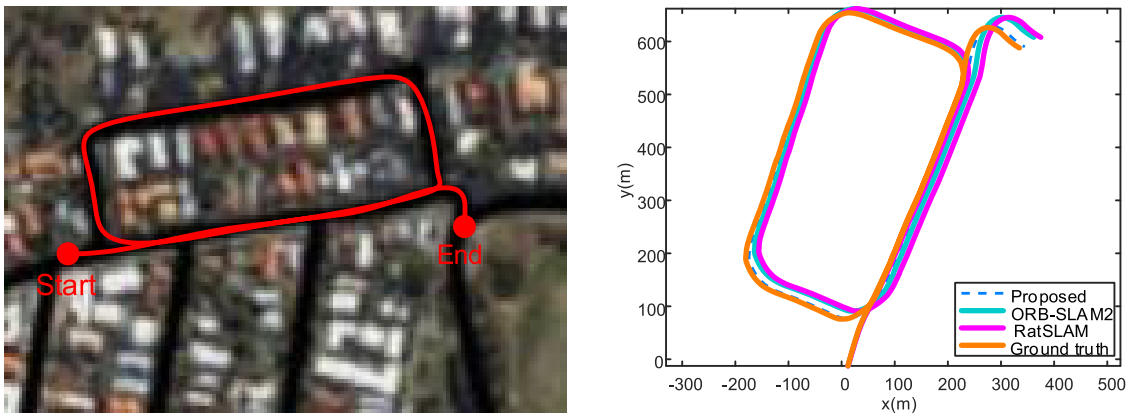



**FIGURE 12.** Comparison of the trajectory maps of the three algorithms under KITTI dataset.




(a)Plan view of the data collection site for the St. Lucia dataset

(b) Environmental cognitive map constructed by three algorithms

**FIGURE 13.** Environmental cognitive map construction results.

collection site of the public dataset St. Lucia. Figure 13(b) displays the trajectory maps acquired by GSOM-BSLAM, ORB-SLAM2, and RatSLAM algorithms on the St. Lucia dataset. Table 4 shows the performance comparison of the three algorithms on the St. Lucia dataset regarding map building results. As Figure 18 demonstrates, the trajectory map obtained by GSOM-BSLAM algorithm is closer to the real trajectory, while the running time is reduced by 36.5% compared with RatSLAM and 10.8% compared with ORB-SLAM2. This result is attributed to GSOM algorithm by selecting a place cell position field with the maximum activity value as the spatial position information and using

**TABLE 4. Comparative performance analysis of Environmental cognitive map construction results.**

| Algorithm | Time consumption /min | Absolute trajectory error /m | Closed-loop accuracy /% |
|---|---|---|---|
| RatSLAM | 52 | 0.74 | 76.5 |
| ORB-SLAM2 | 37 | 0.59 | 85.4 |
| GSOM-BSLAM | 33 | 0.46 | 87.9 |

a few place cell firing information characterizing the current spatial information, improving the computational efficiency of algorithm. Meanwhile, since our algorithm introduces color depth images for closed-loop detection to correct the errors caused by spatial grid cell path integration and place cell position estimation, the absolute trajectory error of GSOM-BSLAM is reduced by 22.0% compared with ORB-SLAM2 and 37.8% compared with RatSLAM, ensuring the accuracy of environmental cognitive map construction.

## VII. CONCLUSION

Based on the knowledge of mammalian spatial navigation cell path integration and autonomous navigation and localization mechanism, a GSOM-BSLAM algorithm is proposed inspired by the spatial cognitive mechanism of mammalian hippocampus. The algorithm uses the training process of GSOM competitive neural network to simulate the firing characteristics of individual cells when rats explore their environment and derives the mapping relationship from grid cells to place cells. An accurate environmental cognitive map is finally constructed by introducing color depth images for closed-loop detection and error correction for spatial cell path integration. The experimental results indicate that GSOM-BSLAM algorithm proposed exhibits greater advantages regarding localization accuracy and real-time performance than RatSLAM and ORB-SLAM2 algorithms. The GSOM-BSLAM algorithm proposed builds on the cognitive mechanism of biological navigation and extends the application of biological spatial navigation cells in mobile robot navigation and localization. However, there are still some areas for improvement in this work. For instance, when the mobile robot turns or moves too fast, causing the camera to shake or encounter sparse scene environments (glass or white walls), it is difficult to achieve effective position capture with a simple vision sensor system. Accordingly, the subsequent use of IMU inertial and vision information data fusion is considered to achieve the precise position of mobile robots in environment.

## REFERENCES

[1] L. Cui and C. Ma, "SDF-SLAM: Semantic depth filter SLAM for dynamic environments," *IEEE Access*, vol. 8, pp. 95301–95311, 2020.

[2] Y. Ai, T. Rui, M. Lu, L. Fu, S. Liu, and S. Wang, "DDL-SLAM: A robust RGB-D SLAM in dynamic environments combined with deep learning," *IEEE Access*, vol. 8, pp. 162335–162342, 2020.

[3] H. Bavle, P. De La Puente, J. P. How, and P. Campoy, "VPS-SLAM: Visual planar semantic SLAM for aerial robotic systems," *IEEE Access*, vol. 8, pp. 60704–60718, 2020.

[4] L. Yuntian, M. Rongjun, and S. Yongzhi, "A survey of visual SLAM in unmanned systems," *Control Decis.*, vol. 36, no. 3, pp. 513–522, 2012.

[5] J. Engel, T. Schöps, and D. Cremers, "LSD-SLAM: Large-scale direct monocular SLAM," *Proc. Eur. Conf. Comput. Vis.* Zurich, Switzerland: Springer, 2014, pp. 834–849.

[6] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.

[7] J. O'keefe and J. DOSLROVSKV, "The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat," *Brain Res.*, vol. 34, no. 1, pp. 171–175, 1971, doi: 10.1016/0006-8993(71)90358-1.

[8] M. Chen and D. Tian, "Bionic SLAM algorithm based on multi-scale grid cell to place cell," *J. Comput.-Aided Des. Comput. Graph.*, vol. 33, no. 5, pp. 712–723, May 2021.

[9] D. Bush, C. Barry, D. Manson, and N. Burgess, "Using grid cells for navigation," *Neuron*, vol. 87, no. 3, pp. 507–520, Aug. 2015, doi: 10.1016/j.neuron.2015.07.006.

[10] J. Krupic, N. Burgess, and J. O'keefe, "Neural represent-tations of location composed of spatially periodic bands," *Science*, vol. 337, no. 6096, pp. 853–857, 2012.

[11] A. T. Keinath, R. A. Epstein, and V. Balasubramanian, "Environmental deformations dynamically shift the grid cell spatial metric," *eLife*, vol. 7, pp. 1–22, Oct. 2018, doi: 10.7554/eLife.38169.

[12] M. Yuan, B. Tian, V. A. Shim, H. Tang, and H. Li, "An entorhinal-hippocampal model for simultaneous cognitive map building," in *Proc. 29th AAAI Conf. Artif. Intell.* Reston, VA, USA, 2015, pp. 586–592.

[13] M. J. Milford, G. F. Wyeth, and D. Prasser, "RatSLAM: A hippocampal model for simultaneous localization and mapping," in *Proc. IEEE Int. Conf. Robot. Automat.*, New York, NY, USA, May 2004, pp. 403–408.

[14] M. Milford, R. Schulz, D. Prasser, G. Wyeth, and J. Wiles, "Learning spatial concepts from RatSLAM representations," *Robot. Auto. Syst.*, vol. 55, no. 5, pp. 403–410, May 2007.

[15] B. Tian, V. A. Shim, M. Yuan, and C. Srinivasan, "RGB-D based cognitive map building and navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Piscataway, PA, USA, Apr. 2013, pp. 1562–1567.

[16] C. Barry and N. Burgess, "Neural mechanisms of self-location," *Current Biol.*, vol. 24, no. 8, pp. R330–R339, Apr. 2014.

[17] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990, doi: 10.1109/5.58.325.

[18] H Montazeri, S Moradi, and R Safabakhsh, "Continuous state/action reinforcement learning: A growing self-organizing map approach," *Neurocomputing*, vol. 74, no. 7, pp. 1069–1082, 2011, doi: 10.1016./j.neucom.2010.11.012.

[19] R. M. Grieves, S. Jedidi-Ayoub, K. Mishchanchuk, A. Liu, S. Renaudineau, É. Duvelle, and K. J. Jeffery, "Irregular distribution of grid cell firing fields in rats exploring a 3D volumetric space," *Nature Neurosci.*, vol. 4, pp. 11–17, Aug. 2021.

[20] A. L. Juavinett and E. M. Callaway, "'Pattern and component motion responses in mouse visual cortical areas," *Current Biol.*, vol. 25, no. 13, pp. 1759–1764, 2015.

[21] J. J. Couey, A. Witoelar, S. J. Zhang, K. Zheng, and J. Ye, "Recurrent inhibitory circuitry as a mechanism for grid formation," *Nature Neurosci.*, vol. 16, pp. 318–324, Mar. 2013.

[22] N. Burgess, C. Barry, and J. O'Keefe, "An oscillatory interference model of grid cell firing," *Hippocampus*, vol. 17, no. 9, pp. 801–812, 2007.

[23] M. E. Hasselmo, L. M. Giocomo, and E. A. Zilli, "Grid cell firing may arise from interference of theta frequency membrane potential oscillations in single neurons," *Hippocampus*, vol. 17, no. 12, pp. 1252–1271, Dec. 2007.

[24] W. Wang, "The broadband power shifts in entorhinal EEG are related to the firing of grid cells," *Heliyon*, vol. 7, no. 1, Jan. 2021, Art. no. e06087.

[25] U. M. Erdem and M. Hasselmo, "A goal-directed spatial navigation model using forward trajectory planning based on grid cells," *Eur. J. Neurosci.*, vol. 35, no. 6, pp. 916–931, Mar. 2012.

[26] E. I. Moser, Y. Roudi, M. P. Witter, C. Kentros, T. Bonhoefier, and M. B. Moser, "Grid cells and cortical representation," *Nature Rev. Neurosci.*, vol. 15, no. 7, pp. 466–481, 2014.
[27] 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, no. 1, pp. 69–89, Jul. 2008.
[28] C. Fuhs and D. S. Touretzky, "A spin glass model of path integration in rat medial entorhinal cortex," *J. Neurosci.*, vol. 26, no. 16, pp. 4266–4276, 2006.
[29] E. B. Knudsen and J. D. Wallis, "Hippocampal neurons construct a map of an abstract value space," *Cell*, vol. 184, no. 18, pp. 4640–4650.e10, Sep. 2021.
[30] C. Mengyuan and X. Minghui, "Research on mobile robot bionic location algorithm based on growing self-organizing map," *J. Electron. Inf. Technol.*, vol. 43, no. 4, pp. 1003–1013, 2021.
[31] K. Soman, V. Muralidharan, and V. S. Chakravarthy, "A unified hierarchical oscillatory network model of head direction cells, spatially periodic cells, and place cells," *Eur. J. Neurosci.*, vol. 47, no. 10, pp. 1266–1281, May 2018.

**YUKUN ZHANG** was born in Bozhou, Anhui, China, in 1995. He received the B.S. degree from the School of Mechanical and Electrical Engineering, Huainan Normal University, Huainan, China, in 2019. He is currently pursuing the M.S. degree with the School of Electrical and Electronic Engineering, Anhui Polytechnic University, China.

His current research interests include SLAM and navigation for mobile robotics.

**MENGYUAN CHEN** received the Ph.D. degree in robotics from the University of Science and Technology of China, in 2019.

He is currently an Associate Professor of robotics with the Department of Electrical and Electronic Engineering, Anhui Polytechnic University, China. He is also a Master's Supervisor with Anhui Polytechnic University. He presided over the National Natural Science Foundation of China, Natural Science Foundation of Anhui, China, and Anhui Key Research and Development Programs (Foreign Scientific and Technological Cooperation). He has published more than 30 papers in peer-reviewed journals, international conferences, and book chapter. His research interests include cognitive robotics and SLAM, especially on the robot's navigation.

**DEHONG TIAN** was born in Fuyang, Anhui, China, in 1995. He received the B.S. degree from the School of Mechanical and Electrical Engineering, Huainan Normal University, Huainan, China, in 2017. He is currently pursuing the M.S. degree with the School of Electrical and Electronic Engineering, Anhui Polytechnic University, China.

His current research interests include SLAM and navigation for mobile robotics.

**LINGMEI DING** was born in Taizhou, Jiangsu, China, in 1994. She received the B.S. degree from the Department of Information Engineering, College of Mechanical and Electrical Engineering, Anhui Polytechnic University, Wuhu, China, in 2019. She is currently pursuing the M.S. degree with the School of Electrical and Electronic Engineering, Anhui Polytechnic University.

Her current research interests include SLAM and navigation for mobile robotics.