



Cognitive Mapping Based on Conjunctive Representations of Space and Movement

Taiping Zeng^{1,2} and Bailu Si^{1*}

¹State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China,

²University of Chinese Academy of Sciences, Beijing, China

It is a challenge to build robust simultaneous localization and mapping (SLAM) system in dynamical large-scale environments. Inspired by recent findings in the entorhinal–hippocampal neuronal circuits, we propose a cognitive mapping model that includes continuous attractor networks of head-direction cells and conjunctive grid cells to integrate velocity information by conjunctive encodings of space and movement. Visual inputs from the local view cells in the model provide feedback cues to correct drifting errors of the attractors caused by the noisy velocity inputs. We demonstrate the mapping performance of the proposed cognitive mapping model on an open-source dataset of 66 km car journey in a 3 km × 1.6 km urban area. Experimental results show that the proposed model is robust in building a coherent semi-metric topological map of the entire urban area using a monocular camera, even though the image inputs contain various changes caused by different light conditions and terrains. The results in this study could inspire both neuroscience and robotic research to better understand the neural computational mechanisms of spatial cognition and to build robust robotic navigation systems in large-scale environments.

OPEN ACCESS

Edited by:

Shuai Li,

Hong Kong Polytechnic University,
Hong Kong

Reviewed by:

Yang Zhan,

Shenzhen Institutes of Advanced
Technology (CAS), China

Lin Xiao,

Jishou University, China

*Correspondence:

Bailu Si

sibailu@sia.ac.cn

Received: 05 January 2017

Accepted: 19 October 2017

Published: 22 November 2017

Citation:

Zeng T and Si B (2017) Cognitive Mapping Based on Conjunctive Representations of Space and Movement. *Front. Neurorobot.* 11:61. doi: 10.3389/fnbot.2017.00061

1. INTRODUCTION

One of the fundamental challenges to integrate robots into our daily life is to endow with the ability of spatial cognition. Over the past two decades, many efforts have been exerted to develop truly autonomous robot navigation systems that are able to simultaneously localize, map, and navigate without human interventions in domestic home and workplace. This line of research has been recognized as the problem of Simultaneous Localization and Mapping (SLAM) (Smith and Cheeseman, 1986; Bailey and Durrant-Whyte, 2006; Durrant-Whyte and Bailey, 2006; Thrun and Leonard, 2008) and constitutes a significant area of robotic research. The approaches to solving SLAM problem fall into two classes. The classical approach is the probabilistic one that uses the Extended Kalman Filters, the Particle Filters, etc., as the basis for robot SLAM (Thrun et al., 2005). Numerous variants of filtering algorithms have been proposed to estimate the position and the heading direction of the robot in large environments. However, most of these algorithms require costly sensors, significant computational resources, and the assumption of static environment. Although filtering algorithms provide beautiful mathematical frameworks for SLAM, they are in lack of robustness, since bad data association (i.e., the matching between perception and internal representation) can often lead them into irreversible divergence, impairing their application

to dynamic and large-scale environments. For a good review, one can refer to Thrun and Leonard (2008); Dudek and Jenkin (2010).

The other approach of SLAM research is sparked by the inspirations from neuroscience and aims to build neural network models that work in analogy to the neural circuits of spatial cognition (Zeno et al., 2016). Animals show amazing navigation ability (Klaassen et al., 2011; Tsoar et al., 2011), which has long been hypothesized to result from internal map-like representation in neural space, i.e., cognitive map of the environment (Tolman, 1948). The idea of cognitive map finds embodiment, thanks to the discovery of place cells (PCs) in the hippocampus of mammalian brains (O'Keefe and Dostrovsky, 1971; O'Keefe and Conway, 1978). A random set of place cells are recruited to encode positions in an environment and fire selectively in some restricted portions of the environment. The firing locations of place cells are randomly redistributed across environments and do not conserve metric relation. Place cells are considered to form topological maps of environments. Head direction cells (HD cells), first found in the post-subiculum (Taube et al., 1990), fire when the animal's head points to a specific direction relative to an allocentric reference frame. Head directions cells exist in various brain regions (Taube, 2007) and provide directional information like a compass. The grid cells discovered in the dorsolateral band of the medial entorhinal cortex (dMEC) fire in multiple locations that collectively define triangular grids spanning the whole environment explored by the animal (Hafting et al., 2005). The population codes of grid cells can accurately encode animal's spatial positions in an environment (Mathis et al., 2012). The relative spatial phases of nearby grid cells in dMEC are kept fixed across environments, up to a coherent rotation and shift (Fyhn et al., 2007). Grid cells are thought to be an important ingredient of the inner GPS in mammalian brains (Knierim, 2015). Grid cells show a rich variety of response properties. Pure positional grid cells, abundant in layer II of MEC, do not show strong modulation by movements. Conjunctive grid cells, residing mainly in layer III, V, and VI of MEC, are selective also to running directions (Sargolini et al., 2006). More recently, speed cells found in the MEC are characterized by a context-invariant positive, linear response to running speed, which is a key component of the dynamic representation of the animal's self location in MEC (Kropff et al., 2015).

Sitting between sensory cortex and hippocampus, MEC is a converging area where grid cells, HD cells, and speed cells work together to integrate movement and sensory information. Computational models of grid cells reveal possible information processing mechanisms of MEC in forming spatial representations and provide building blocks for cognitive mapping systems (Zilli, 2012). The large majority of grid cells can be divided into three classes: oscillatory interference (OI) models, continuous attractor network models (CAN) (Fuhs and Touretzky, 2006; McNaughton et al., 2006; Burak and Fiete, 2009; Knierim and Zhang, 2012; Si et al., 2014), and self-organization (SO) models (Kropff and Treves, 2008; Si and Treves, 2013; Stella et al., 2013; Stella and Treves, 2015). OI models propose that the hexagonal grid pattern arises from the interference between velocity modulated oscillators and a theta oscillator (Burgess et al., 2007; Welday et al., 2011). The OI models are susceptible to drifting error and lack of robustness (Remme et al., 2010) and may not be robust enough for robot

navigation systems. The CAN models use structured connections between cells to generate hexagonal grid firing patterns in the population level (Fuhs and Touretzky, 2006; McNaughton et al., 2006; Burak and Fiete, 2009). The grid firing pattern is a stable network state arising from the collective behavior of the neural population and lies in a continuous manifold of steady states. If coupled by velocity inputs, a continuous attractor network model of grid cells is able to propagate its network state along the manifold, perform accurate path integration, and, therefore produce single cell hexagonal grid firing patterns in the environment (Burak and Fiete, 2009). The conjunctive grid cells observed in layer III and deeper layers of MEC are modeled by a CAN that encodes position and velocity explicitly (Si et al., 2014). The conjunctive position-by-velocity representation produces intrinsically moving patterns that can be selected by velocity inputs. The movement of the pattern is not dependent on the amplitude of the velocity signal or on the firing properties of the cells and, therefore, is able to achieve robust path integration even if the connections are contaminated by noise. SO models of grid cells assume that grid cells receive broadly tuned spatial inputs, e.g., feedforward inputs from visual areas or the hippocampus. Single cell firing rate adaptation could result in triangular firing patterns (Kropff and Treves, 2008). Since the feedforward inputs function as error correcting signals, the anchoring of grid maps in the environment is strongly influenced by the inputs experienced by the animal (Si et al., 2012; Stensola et al., 2015).

Although numerous computational models of grid cells have been proposed to account for the mechanisms of integrating movement and sensory information into spatial representations, so far, only a few robot navigation systems utilize computational mechanisms of spatial cognition. In the model of RatSLAM, pose cells were developed to encode positions in large-scale environment for long-term robotic navigation tasks (Milford and Wyeth, 2008). Pose cell is a kind of abstract cell, which updates its activity by displacing a copy of the activity packet, rather than performing path integration according to the dynamics of the network. In Jauffret et al. (2015), a modulo projection of the path integration is used to generate the grid cell firing patterns. However, grid cells aid neither localization nor mapping in the mobile robot system. There is a sparse number of robot navigation systems, which use grid cell networks to do path integration and represent the pose of a robot (Sheynikhovich et al., 2009; Yuan et al., 2015; Mulas et al., 2016). Sheynikhovich et al. (2009) built a biological plausible model including visual, medial entorhinal cortex, hippocampus, and striatum to perform navigation in a simulated environment. Their model successfully reproduced many experimental results, such as rescaling of firing fields of place cells and grid cells and solved water maze task by reinforcement learning based on spatial representations. Both (Yuan et al., 2015) and (Mulas et al., 2016) used a model of pure positional grid cells in layer II of MEC (Burak and Fiete, 2009) to track the position of a robot and used visual inputs to anchor grids in the environment. They tested their system in indoor environments of dozens of meters in side. The robustness is yet to be verified in large-scale environments.

The theory of neural dynamics has been extensively adopted to model brain functions, such as memory (Treves, 1990; Mi et al., 2017), navigation (Gerstner and Abbott, 1997), and sensory

integration (Zhang et al., 2016). The theory of neural dynamics has become popular in robotics to enhance the cognitive abilities of robots (Erlhagen and Bicho, 2006), such as obstacle avoidance (Xiao et al., 2017), coordinated path tracking (Xiao and Zhang, 2016), motion tracking (Ding et al., 2017), and grasping (Knips et al., 2017).

In this work, we follow the research line of neural dynamics and developed a cognitive mapping model for mobile robots, taking advantage of the coding strategies of the spatial memory circuits in mammalian brains. The key components of the proposed model include HD cells, conjunctive grid cells, and local view cells. Both HD cells and conjunctive grid cells are modeled by continuous attractor networks that operate on the same principles. More specifically, HD cells in the model represent arbitrary conjunctions of head directions and rotations of the animals. Due to the asymmetric recurrent connections and the network dynamics, intrinsically rotating patterns spontaneously emerge in the network. Angular velocity inputs to the network pick up patterns rotating with appropriate velocities and track the head direction of the robot. Similarly, conjunctive grid cells in the model encode conjunctions of positions and translations in the two-dimensional environment. The inputs from HD cells and speed cells activate a subset of the conjunctive grid cells that produce triangular patterns moving intrinsically in the neural tissue with velocity proportional to the running velocity of the robot. Both HD cells and conjunctive grid cells get inputs from local view cells, which are activated once the robot moves to a familiar scene and which provide anchoring cues. The mapping performance of the proposed model is demonstrated on a 66 km long urban car journey. Our cognitive mapping model generates a coherent map of the outdoor environment using single camera image data. The experimental results show that our model is stable and robust in large-scale outdoor environment of variable light conditions and terrains.

The major contributions in this paper are twofold. First, the neural network framework presented in this paper is based on recent experimental studies on the hippocampal–entorhinal

circuits and aims to model layer III and deep layers of the MEC and the hippocampus. In the system, conjunctive HD-by-velocity cells and conjunctive grid-by-velocity cells work hand in hand to integrate movement and sensory information and build a large-scale map. Second, the neural dynamics of the hippocampal–entorhinal circuits is modeled. Provided with the inputs from local view cells, the neural dynamics of the system functions as a general mechanism for error correction and pattern completion.

This paper proceeds as follows. In Section 2, we describe the HD cell model, the conjunctive grid cell model, and other core components of the robot navigation model. The experimental setup and the resulting neural activities, rate maps, and the semi-metric topological map are presented in Section 3. Section 4 discusses the results and points to future directions, with a brief conclusion in Section 5.

2. MATERIALS AND METHODS

In this study, we propose a cognitive mapping model that has similar regions and cell types as in the entorhinal–hippocampal circuits of mammalian brains. Localization is the key function of the proposed model and is performed by the HD cells and the conjunctive grid cells. The HD cells and the conjunctive grid cells in the model correspond to layer III, V, and VI of the MEC of mammalian brains. They integrate rotational or translational velocity and form directional or positional codes, respectively. The local view cells in the model correspond to the retrosplenial cortex or visual cortex and relay visual inputs to the MEC. Cognitive map in the system corresponds to the hippocampus.

2.1. Neural Network Architecture

The architecture of the proposed model is shown in **Figure 1**. The HD cells in the model form a ring attractor in the neural space. A bump of activity emerges in the network. The phase of the bump represents the robot's real angle of the head direction in the physical environment up to a constant shift. Tuned angular velocity inputs active a subset of the HD neurons, and the bump

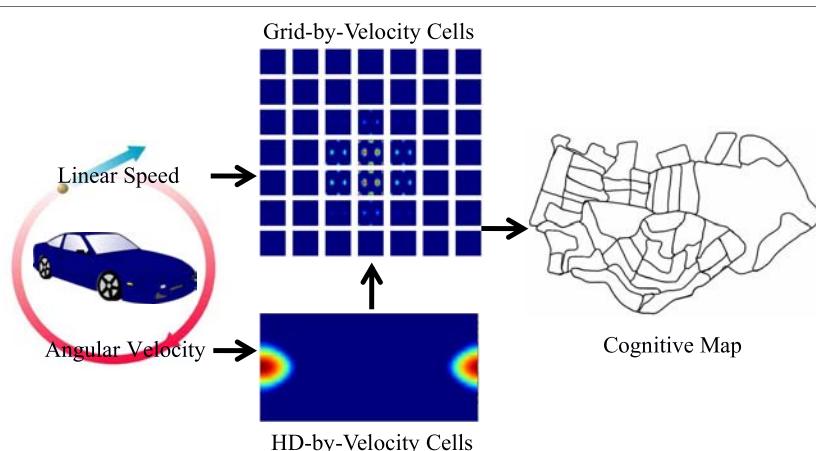


FIGURE 1 | The neural network architecture of the model and the diagram of information flow. The HD-by-velocity cells, updated by angular velocity, represent head directions. The grid-by-velocity cells receiving translational velocity, converted from linear speed and the head direction representations, provide positional representation, which is in turn utilized to build a cognitive map. The activities of the HD-by-velocity cells and the grid-by-velocity cells are encoded by heat maps with red for high activity and blue for no activity.

rotates intrinsically with the same velocity as the angular velocity of the robot.

The conjunctive grid cells in the model form a torus attractor due to the periodic boundary conditions of the neural space. The phase of the grid pattern encodes the position of the robot in the physical environment up to modulo operation. The conjunctive grid cells are activated by the head direction cells and the speed cells, which are both available in the MEC of rodents. The grid pattern in the network moves with a velocity proportional to the movement of the animal in the physical environment. Thus, the HD cells and the conjunctive grid cells work hand in hand and form population codes of the head direction and the position of the robot in the environment.

The conjunctive representations of space and movement allow the networks to have stable states for all movement conditions. This coding property endows the network robustness regardless the movement conditions.

2.2. HD-by-Velocity Cell Model

The HD-by-velocity cells integrate angular velocity and represent one-dimensional head direction. This is achieved by conjunctive encoding of head direction and rotation.

2.2.1. Neural Representation of Head Direction and Angular Velocity

Each unit in the network is labeled by its coordinate (θ, ν) on a two-dimensional manifold. $\theta \in [0, 2\pi]$ is the internal representation of head directions in the environment. Note that θ has periodic boundary condition. $\nu \in [-L_r, L_r]$ encodes angular velocity. Both ν and θ are dimensionless quantities in the neural space. They reflect the real angular velocity and the head direction of the robot. We design the connection weights between units to sustain a single bump of activity both in the direction dimension θ and in the rotation dimension ν . The strength of the connection from a presynaptic unit (θ', ν') to a postsynaptic unit (θ, ν) is written as

$$J(\theta, \nu | \theta', \nu') = J_0 + J_1 \cos(\theta - \theta' - \nu') \cos(\lambda(\nu - \nu')), \quad (1)$$

where $J_0 < 0$ is a uniform inhibition, $J_1 > 0$ describes the interaction strength, and λ defines the spread of velocity tuning. The connection weight from the unit at θ' to the unit at θ in the direction dimension is asymmetric, since the postsynaptic unit that is maximally activated by θ' is not centered at θ' but at $\theta' + \nu'$. The asymmetry causes the bump to move along the direction dimension with a velocity determined by ν' .

2.2.2. Network Dynamics

We consider the activity of the network driven by velocity and sensory inputs. The firing rate $m(\theta, \nu)$ of the unit at coordinate (θ, ν) is governed by the differential equation

$$\begin{aligned} \tau \dot{m}(\theta, \nu) = & -m(\theta, \nu) + f \left(\iint D\theta D\nu J(\theta, \nu | \theta', \nu') m(\theta', \nu') \right. \\ & \left. + I_\nu + I_{view} \right), \end{aligned} \quad (2)$$

where I_ν and I_{view} are the velocity tuned input and the calibration current injected by local view cells, respectively, which we will

explain in detail below. τ is the time constant set to 10 ms. $f(x)$ is a threshold-linear function: $f(x) \equiv [x]_+ = x$ when $x > 0$ and 0 otherwise. The shorthand notations are used $\int D\theta = \frac{1}{2\pi} \int_0^{2\pi} d\theta$, and $\int D\nu = \frac{1}{2L_r} \int_{-L_r}^{L_r} d\nu$.

2.2.3. Angular Velocity Inputs

In order to integrate angular velocity, the activity bump in the HD network should be placed at appropriate position in the velocity dimension of the neural space. The desired bump location in the ν dimension corresponding to a given external angular velocity V is (Si et al., 2014)

$$u(V) = \arctan(\tau V). \quad (3)$$

Here, τ is the time constant defined in equation (2).

The velocity input to the HD-by-velocity units is simply modeled by a tuning function of Gaussian shape

$$I_\nu(\nu | V) = I_r \left[1 - \epsilon + \epsilon \exp \left(-\frac{(\nu - u(V))^2}{2\sigma_r^2} \right) \right]. \quad (4)$$

Here, I_r is the amplitude of the rotational velocity input, ϵ defines the strength of the velocity tuning, and σ_r is the sharpness of the velocity tuning.

2.2.4. Estimation of the Head Direction

The center of the activity bump on the θ axis and the ν axis encodes the head direction and angular velocity of the robot. We define Fourier transformations to recover the head direction and the angular velocity of the robot from the network state

$$\psi = \angle \left(\iint m(\theta, \nu) \exp(i\theta) D\theta D\nu \right), \quad (5)$$

$$\phi = \frac{\angle (\iint m(\theta, \nu) \exp(i\lambda\nu) D\theta D\nu)}{\lambda}. \quad (6)$$

Here, i is the imaginary unit, and function $\angle(Z)$ takes the angle of a complex number Z . $\psi \in [0, 2\pi]$ is the estimated phase of the bump in the direction axis of the neural space and corresponds to the head direction of the robot in the physical space. $\phi \in (-L_r, L_r)$ is the estimated phase of the bump in the velocity axis of the neural space and can be mapped to the angular velocity of the robot in the physical space by inverting equation (3)

$$V = \frac{\tan(\phi)}{\tau}. \quad (7)$$

Note that L_r should be chosen large enough, so that the recovered velocity V is able to represent all possible angular velocities of the robot.

2.3. Grid-by-Velocity Cell Model

Now, we expand our HD-by-velocity cell model to do path integration in two-dimensional environments. To this end, grid-by-velocity cells need represent two-dimensional spatial locations and two-dimensional velocities. The units in the network are wired with appropriate connection profiles, so that hexagonal grid firing pattern is generated and translated in the spatial dimension of the neural space.

2.3.1. Neural Representation of Position and Velocity

Units in the grid-by-velocity network is labeled by coordinates $(\vec{\theta}, \vec{\nu})$ in a four-dimensional neural space. $\vec{\theta} = (\theta_x, \theta_y)$ represents two-dimensional positions in the environment. $\vec{\nu} = (\nu_x, \nu_y)$ encodes the velocity components in the environment. We assume $\vec{\theta}$ has periodic boundary conditions, i.e., $\theta_x, \theta_y \in [0, 2\pi]$. ν_x and ν_y are chosen in $[-L_t, L_t]$.

The connection weight from unit $(\vec{\theta}', \vec{\nu}')$ to $(\vec{\theta}, \vec{\nu})$ is described as

$$J(\vec{\theta}, \vec{\nu} | \vec{\theta}', \vec{\nu}') = J_0 + J_k \cos \left(k \sqrt{\sum_{j \in \{x,y\}} \|\theta_j - \theta'_j - \nu'_j\|^2} \right) \cos \left(\lambda \sqrt{\sum_{j \in \{x,y\}} (\nu_j - \nu'_j)^2} \right), \quad (8)$$

where the integer $k=2$, is chosen so that the network accommodates two bumps both in θ_x axis and in θ_y axis. There is only one bump in each of the velocity dimensions, however. $\|d\|$ is the distance on a circle: $\|d\| = \text{mod}(d + \pi, 2\pi) - \pi$, and $\text{mod}(x, y) \in [0, y)$ gives x modulo y.

2.3.2. Network Dynamics

Although grid-by-velocity units are organized with different manifold structure from the HD-by-velocity units, they share the same intrinsic dynamics as in equation (2)

$$\tau \dot{m}(\vec{\theta}, \vec{\nu}) = -m(\vec{\theta}, \vec{\nu}) + f \left(\iint D\vec{\theta} D\vec{\nu} J(\vec{\theta}, \vec{\nu} | \vec{\theta}', \vec{\nu}') m(\vec{\theta}', \vec{\nu}') + I_\nu + I_{view} \right). \quad (9)$$

Note that $\int D\vec{\theta} = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} d\theta_x d\theta_y$, and $\int D\vec{\nu} = \frac{1}{4L_t^2} \int_{-L_t}^{L_t} \int_{-L_t}^{L_t} d\nu_x d\nu_y$.

2.3.3. Translational Velocity Inputs

In order to perform accurate path integration, the velocity of the moving bumps in the neural space should always be kept proportional to the velocity of the robot in the physical space. The network requires tuned velocity input to pin the activity bumps at appropriate positions on the velocity axes in the neural space, so that the bumps move with the desired velocity.

The translational velocity $\vec{V} = (V_x, V_y)$ of the robot is obtained by projecting the running speed to the axes of the reference frame using the HD estimated from the HD-by-velocity units (equation (5)). In the brain, running speed is encoded by the speed cells in the MEC. Given the translational velocity of the robot, the desired positions on the velocity axes in the neural space is given by $\vec{u}(\vec{V})$ (Si et al., 2014)

$$\vec{u}(\vec{V}) = \frac{1}{k} \arctan \left(\frac{2\pi\tau\vec{V}}{S} \right), \quad (10)$$

where the function arctan operates on each dimension of \vec{V} . S is a scaling factor between the external velocity of the robot in the

physical environment and the velocity of the bumps in the neural space. S determines the spacing between the neighboring fields of a grid firing pattern in the environment.

The velocity-tuned inputs to the grid-by-velocity units take a Gaussian form for simplicity

$$I_\nu(\vec{\nu} | \vec{V}) = I_t \left[1 - \epsilon + \epsilon \exp \left(-\frac{|\vec{\nu} - \vec{u}(\vec{V})|^2}{2\sigma_t^2} \right) \right], \quad (11)$$

where $|\cdot|$ is the Euclidean norm of a vector. I_t is the amplitude of the translational velocity input. ϵ is the strength of velocity tuning as in equation (4), and σ_t the width of the translational velocity tuning.

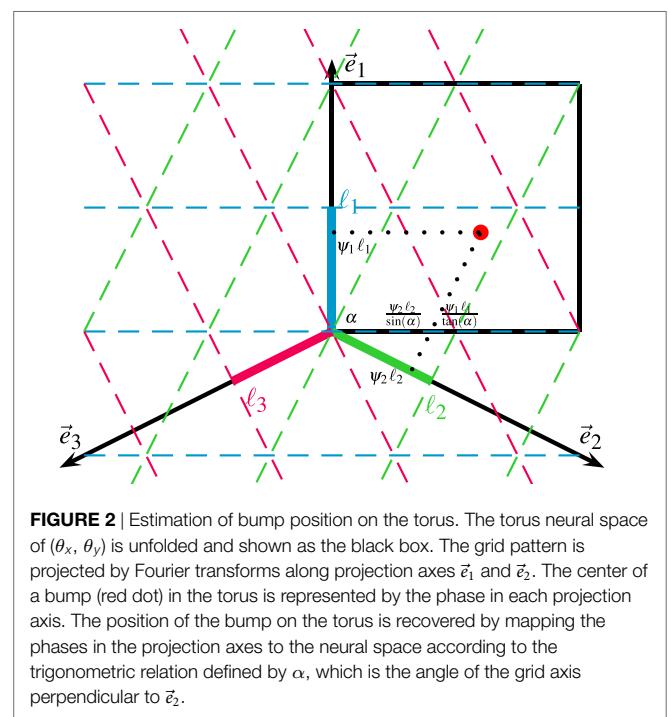
2.3.4. Estimation of the Bump Position and Localization

The center of the bumps on the neural manifold encodes spatial position and translational velocity. We use Fourier transformations to estimate the center of the bumps on the four-dimensional neural manifold.

First, we project the firing activity to the projection axes defined by the grid pattern in the spatial dimensions of the neural manifold and obtain the phases of the grid pattern along the projection axes

$$\psi_j = \frac{\angle \left(\iint D\vec{\theta} D\vec{\nu} m(\vec{\theta}, \vec{\nu}) \exp \left(ik \frac{\vec{\theta}^T \vec{e}_j}{\ell_j} \right) \right)}{k}, \quad (12)$$

where i is the imaginary unit, \vec{e}_j the unit direction of the j -th projection axis, namely $\vec{e}_1 = [0, 1]^T$, $\vec{e}_2 = [\sin\alpha, -\cos\alpha]^T$, $\vec{e}_3 = [-\sin\alpha, -\cos\alpha]^T$ (see Figure 2). α , determined by the grid pattern in the torus, is the angle of the second grid axis. In our case, $\alpha = \arctan(2)$ due to the square shape of the torus. ℓ_j is the



respective wavelength of the grid pattern along the projection axis j , i.e., $\ell_1 = 1$, $\ell_2 = \ell_3 = \sin \alpha$.

Second, we estimate the center of the bumps in the spatial dimensions of the neural manifold according to the trigonometric relation between the projection axes and the spatial axes. As shown in **Figure 2**, the center of the bumps in the θ_y dimension is simply

$$\hat{\theta}_y = \psi_1 \ell_1. \quad (13)$$

By mapping the scaled phase $\psi_1 \ell_1$ and $\psi_2 \ell_2$ to the θ_x direction, the center of the bumps in the θ_x dimension is given by

$$\hat{\theta}_x = \frac{\psi_2 \ell_2}{\sin \alpha} + \frac{\psi_1 \ell_1}{\tan \alpha}. \quad (14)$$

The phases estimated in equation (12) are constrained by the periodic boundary conditions. In order to estimate the true position of the robot in the environment, the phases should be unwrapped by considering the phases of two consecutive time steps. After unwrapping, the cumulative phases are used in equations (13) and (14) to estimate the position of the robot

$$\hat{x}(t) = \frac{S}{2\pi/k} \hat{\theta}_x(t), \quad (15)$$

$$\hat{y}(t) = \frac{S}{2\pi/k} \hat{\theta}_y(t), \quad (16)$$

where $\hat{\theta}_x(t)$ and $\hat{\theta}_y(t)$ are the cumulative phases after unwrapping at time t . The scaling factor $\frac{S}{2\pi/k}$ is the ratio between the velocity of the robot in the physical environment and the velocity of bumps in the network.

To estimate the center of the bumps in the velocity dimensions, the same method as equation (6) is used in the following form

$$\phi_j = \frac{\angle \left(\iint D\vec{\theta} D\vec{\nu} m(\vec{\theta}, \vec{\nu}) \exp(i\lambda\nu_j) \right)}{\lambda}, \quad (17)$$

where $j \in \{x, y\}$. $\phi_j \in (-L_t, L_t)$ is the position of the bump on the axes ν_j .

According to the mapping between the velocity of the bumps in the neural space and the velocity of the robot, the velocity of the robot in the environment is given by

$$V_j = \frac{\text{Stan}(k\phi_j)}{2\pi\tau}. \quad (18)$$

Note that L_t should be chosen large enough to cover the maximal translation velocity experienced by the robot.

2.4. Estimating Angular and Translational Velocity from Visual Inputs

The angular velocity and the translational speed of the robot are estimated by matching two consecutive frames from the camera, according to the method described in Milford and Wyeth (2008). The estimated angular velocity is utilized by the HD-by-velocity units to form HD representations (equation (4)). The translational velocity is obtained by combining the HD estimation from the HD-by-velocity network and the translational speed estimated from the image sequence, and is fed to the grid-by-velocity network to perform path integration (equation (11)).

2.5. Calibration from Local View Cells

Since movement estimation is subject to noise, both the head direction cell network and the grid cell network need calibration by visual inputs. Following Milford and Wyeth (2008), we extract local view templates from images to encode the scenes observed from the camera. If a view is sufficiently different from previously observed views, a new view cell is added to the system, together with the view template and the estimated bump positions ψ , $\hat{\theta}_x$ and $\hat{\theta}_y$. If the robot comes back to a previously visited location, and a familiar view reappears, the corresponding local view cell is activated and provides inputs to the networks. We model the current injected from a view cell to the HD cell network as

$$I_{\text{view}}(\theta) = I_d \exp \left(-\frac{\|\theta - \psi\|^2}{2\sigma_d^2} \right), \quad (19)$$

where I_d is the amplitude of the injected current to the HD cell network, ψ is the associated phase with the view template, σ_d is sharpness of the Gaussian tuning.

For the conjunctive grid cell network, the input from a view cell is given by

$$I_{\text{view}}(\vec{\theta}) = I_p \left(\frac{1}{3} \sum_{j=1}^3 \cos \left(k \frac{(\vec{\theta} - \hat{\vec{\theta}})^T \vec{e}_j}{\ell_j} \right) + C \right), \quad (20)$$

where I_p is the amplitude of the injected current, C is a constant to adjust the baseline, $\hat{\vec{\theta}}$ is the phase associated with the local view template (see equations (13) and (14)), and other parameters are the same as in equation (12). Note that the injected current from a local view cell depends only on $\vec{\theta}$, and is the same for different $\vec{\nu}$.

2.6. Neural Network Parameters

The HD-by-velocity units are regularly arranged on a rectangular grid in the two-dimensional manifold of direction and rotation, with 51 units in the direction dimension and 25 units in the rotation dimension. All together, there are 1,275 HD-by-velocity units. $L_r = 0.0095$ is sufficient to represent the maximal angular velocity of the vehicle.

The grid-by-velocity units are regularly arranged on a rectangular grid in the four-dimensional neuronal manifold of position and velocity. There are 15 units in each of the positional dimensions and 7 units in each of the velocity dimensions, resulting in 11,025 grid-by-velocity units. To represent the maximal experienced translational velocity of the robot, L_t is chosen to be 0.3.

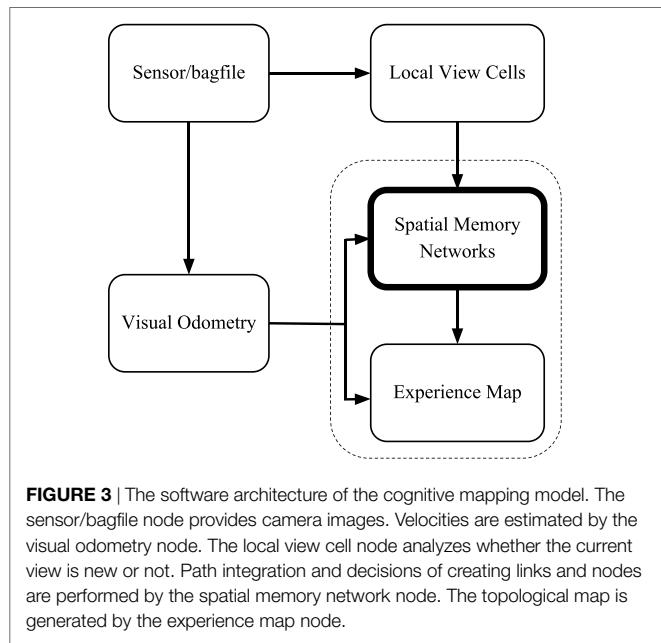
Table 1 summarizes the parameters used in the model.

2.7. Map Representation

A map of the environment can be formed by reading out the spatial codes in the grid-by-velocity network. We could resort to place cell models (Solstad et al., 2006; Si and Treves, 2009; Cheng and Frank, 2011) to build cognitive maps. In this study, we simply adopt the experience map representation (Milford and Wyeth, 2008), which is a topological map storing the spatial positions and their transitions measured from visual odometry. A graph relaxation algorithm (Duckett et al., 2002) is applied to optimize the topological map when the robot comes back to a familiar location and closes a loop (Milford and Wyeth, 2008).

TABLE 1 | Values of the parameters used in the system.

Parameter	J_0	J_1	λ	L_r	I_r	ϵ	σ_r	J_k	k	L_t	S	I_t	σ_t	I_d	σ_d	I_p	C
Value	-60	50	0.8	0.0095	50	0.8	0.012	50	2	0.3	30	60	0.1	60	2.19	200	0.5



2.8. Implementation of the Cognitive Mapping Model

The proposed cognitive mapping model is implemented in the C++ language and is run in the Robot Operating System (ROS) Indigo on Ubuntu 14.04 LTS (Trusty). We organize the software architecture of our cognitive mapping model into five nodes (**Figure 3**), similarly as that of the open RatSLAM system (Ball et al., 2013), for the ease of reusing their visual processing algorithm and comparing with their results.

The visual odometry node measures the angular velocity and the translational speed according to the changes of visual scenes. It receives frames from the ROS either from a camera or from stored data in a bagfile.

The local view cell node determines whether the current scene is new or is a familiar one. It provides calibration current to the networks in the spatial memory node.

The spatial memory network node includes the head-direction cell network and the conjunctive grid cell network. This node receives two types of ROS messages as inputs: odometry and view templates. As shown in Section 2.2.3, 2.3.3, and 2.5, the HD cell network and the conjunctive grid cell network integrate velocity information and visual information to form neural codes of the pose of the robot. The spatial memory network node also makes decisions about the creation of nodes and links in the experience map and sends ROS messages of graph operations to the experience map node.

The experience map node reads out the neural codes of the conjunctive grid units and represents the key locations of the environment as the vertices in a topological graph. A vertex stores

both the position estimated from the conjunctive grid units and the odometry relative to its previous vertex in the graph. On loop closure, a topological map relaxation method is used to find the minimum disagreement between the odometric transition information and the absolute position in the topological map space. When the position encoded in the conjunctive grid cell network is far enough from the position of the previous vertex, the spatial memory network node would inform the experience map to create a new vertex and a new edge linking to the previous vertex.

We write python scripts to visualize the live state of our robot navigation system. The key components of the system are shown, including the real time neural activity of the HD cells and the conjunctive grid cells (**Figures 4A,B**), the image of the scene and the local view templates (**Figure 4C**), as well as the current experience map (**Figure 4D**).

2.9. St Lucia Suburb Dataset

The St Lucia Suburb dataset was first used in Milford and Wyeth (2008). The dataset was gathered in the suburb area of St Lucia in Brisbane, Australia by a vehicle. The vehicle was driven around the road network, visiting every street at least once. The trajectory is 66 km long, finished within about 100 min at typical driving speed up to 60 km/h. The trajectory spans an area of 3 km by 1.6 km in East–West and North–South direction, respectively. The street views were recorded in 10 frames/s into a video of resolution 640×480 pixels by a laptop's built-in web-cam mounted on the roof of the vehicle. The GPS information of the vehicle was not collected, however.

St Lucia is a very challenging environment to map due to its big variability and fairly large scale. There are 51 inner loops of varying size and shape, and more than 80 intersections in the road network (**Figure 5A**). The roads vary from busy multilane streets to small tracks, the terrains range from flat plains to steep hills, and the light conditions change a lot as well (**Figure 5B**).

3. RESULTS

We tested our cognitive mapping model on the St Lucia dataset (Ball et al., 2013). Our cognitive mapping model is run on a personal computer with 3.4 GHz six-core Intel i7 processor and 64 GB memory. The program is paralleled using OpenMP. Video S1 in Supplementary Material shows the mapping process.

3.1. Neural Representation

Figures 6A,C show the activity of the HD-by-velocity units and the conjunctive grid-by-velocity units in the beginning of the experiment, when the robot is stationary. In the HD-by-velocity network, a single bump of activity emerges. Since the angular velocity is zero at this time, the bump is centered in the velocity dimension (red arrow in **Figure 6A**). The position of the

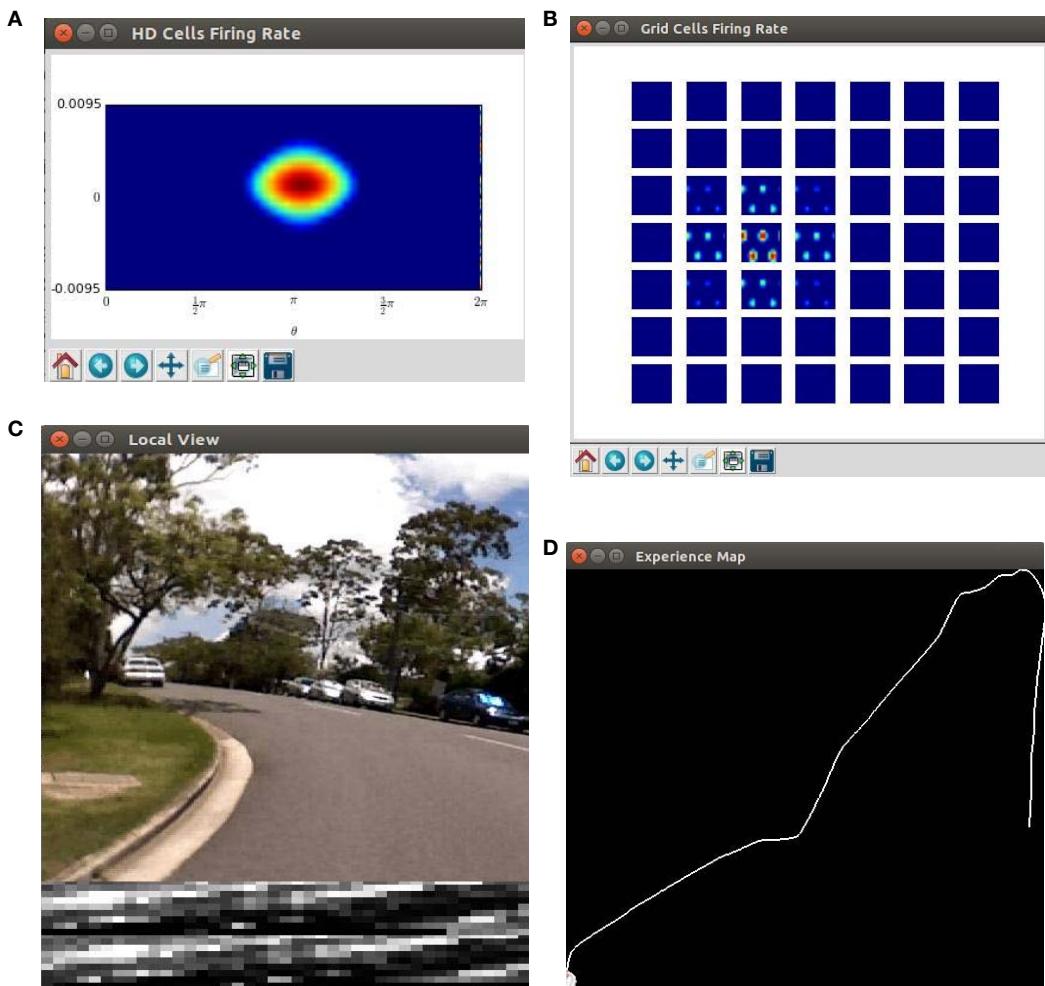


FIGURE 4 | Screenshots of the cognitive mapping model. **(A)** The neural activity of the HD-by-velocity units is depicted as a heat map, with warm colors for high activities and cool colors for activities diminishing to zero; **(B)** the activity of the conjunctive grid-by-velocity units is shown as heat maps, and each panel corresponds to a subspace of the neural manifold with the same velocity label. The same color code as in **(A)** is used; **(C)** the video frame (top) is the visual input to the robot. A visual feature map is constructed to represent the current scene (bottom). The best matched template is shown in the middle; **(D)** experience map showing paths with nodes and links.

bump in the direction dimension (black arrow in **Figure 6A**) encodes the direction of the robot in the environment. At this moment, the activity of the grid-by-velocity units is also centered in the center of the velocity dimensions (red arrows in **Figure 6C**) due to the zero translational velocity input. In spatial dimensions, the same grid pattern sustains, with peak activity decays gradually when the velocity labels of the units are away from the center. The phase of the grid pattern (short black arrows in **Figure 6C**) encodes the position of the robot in the environment.

One example of the network states in the middle of the experiment is displayed in **Figures 6B,D**. The activity bump of the HD-by-velocity network is centered at (3.25, 0.0069) (**Figure 6B**), which means that the current HD angle of the robot is encoded as 186.21°, and meanwhile, the robot is rotating counterclockwise at 0.69 rad/s in the physical environment. At this moment, the activity pattern of the grid-by-velocity network is centered at $\nu_x = -0.0889$ and $\nu_y = -0.0045$ in the velocity dimensions,

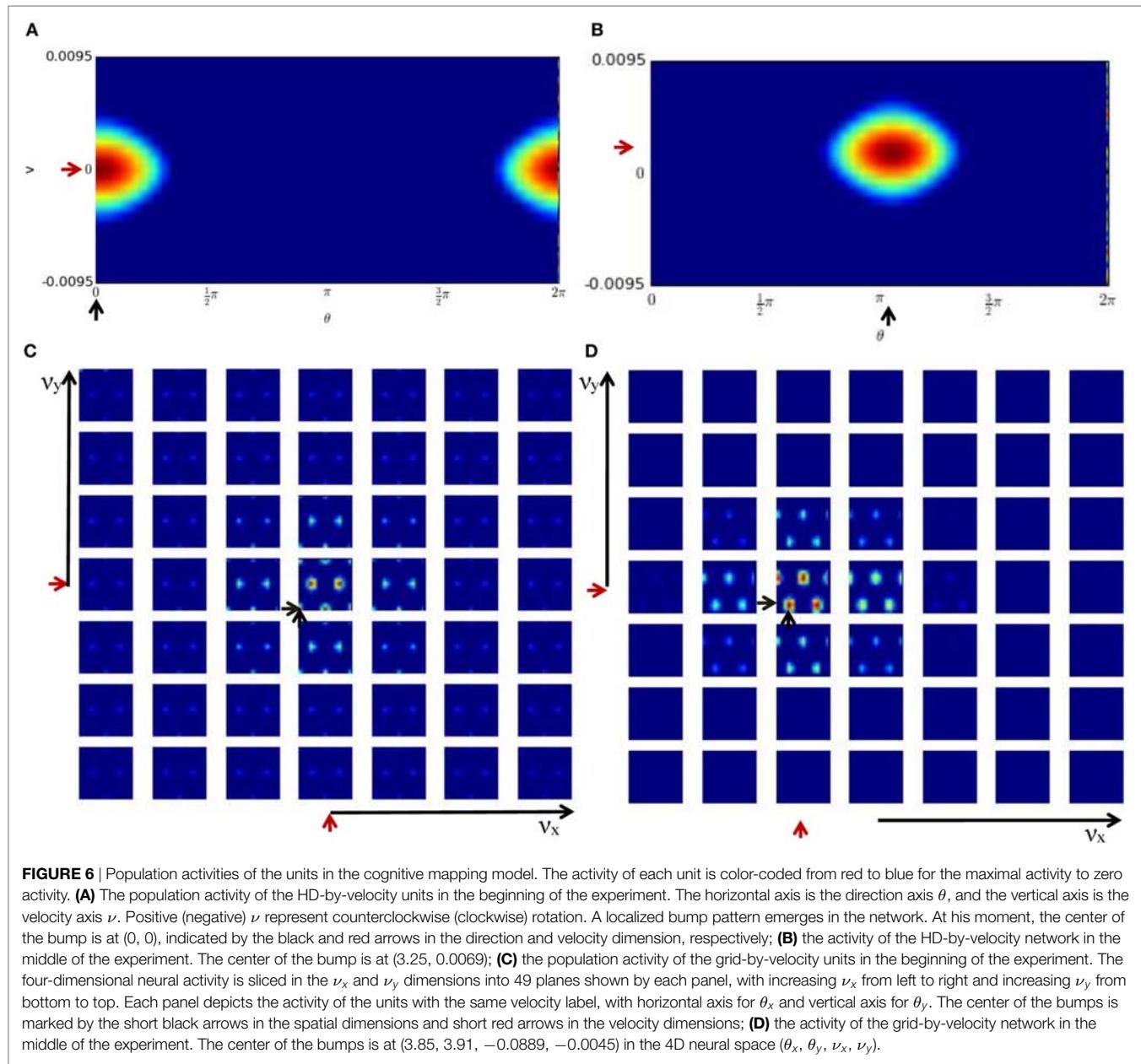
corresponding to the movement of the robot in $V_x = -11.43$ and $V_y = -0.57$ m/s.

3.2. Pattern Completion with Visual Feedback

The network states are calibrated by the cues provided by the local view cells. The calibration is achieved by the pattern completion mechanism of the attractor networks. **Figure 7** and Video S2 in Supplementary Materials show an example of pattern completion process during loop closure. In **Figures 7A,B**, the robot traverses a path for the first time, while the networks integrate velocity inputs, forming localized bumps that encode the head direction and the position of the robot. When the robot closes a loop and revisits familiar places (**Figures 7C,D**), local view cells get activated and provide strong cues to the networks. Weak bumps emerge at the phases associated to the local view cells and quickly become the dominating bumps in the network, retrieving phases similar to



FIGURE 5 | The complex environment of St Lucia. **(A)** The map of St Lucia from a bird's eye view. Black lines are the major road network in St Lucia urban area. The area is approximately 3.0 km by 1.6 km in size; **(B)** example scenes experienced by the robot in St Lucia. Panel **(A)** reused from <http://www.itee.uq.edu.au/think//filething/get/246/MichaelFig3.png>.



the phases when the robots first visited this place (Figures 7A vs. 7D). Once the robot moves to unfamiliar areas, local view cells become inactive, and the bumps revert to their localized shapes (Figures 7E,F).

3.3. Cognitive Map

The cognitive map created by the mapping system is shown in Figure 8. The thick green line is composed of the vertices of the topological graph. The positions of the vertices are determined by the conjunctive grid cells and graph relaxation. The link between two related vertices is shown by the fine blue line. Since the physical distance between two topological vertices is considered, the experience map is actually a semi-metric topological map, which can be compared visually with the ground truth map (Figure 5A) by naked eyes. The cognitive map conserves the overall layout of

the road network of the environment. The cognitive map correctly represents all loop closures and intersections, although the orientation and length of the path are slightly different from the ground truth map. Overall, the map built by our method is consistent with the true map of the environment, and of similar quality as compared with that obtained in Ball et al. (2013).

3.4. Firing Rate Maps

To see the responses of single units, we show the activities of example units on top of the cognitive map. Figure 9A depicts the firing rate map of the HD-by-velocity unit with label (0, 0) in the neural manifold. It fires strongly when the robot moves in broad directions close to southwest. The unit responds to a wide range of directions, since the width of the bump in the HD-by-velocity network is quite big, covering about 70° in head

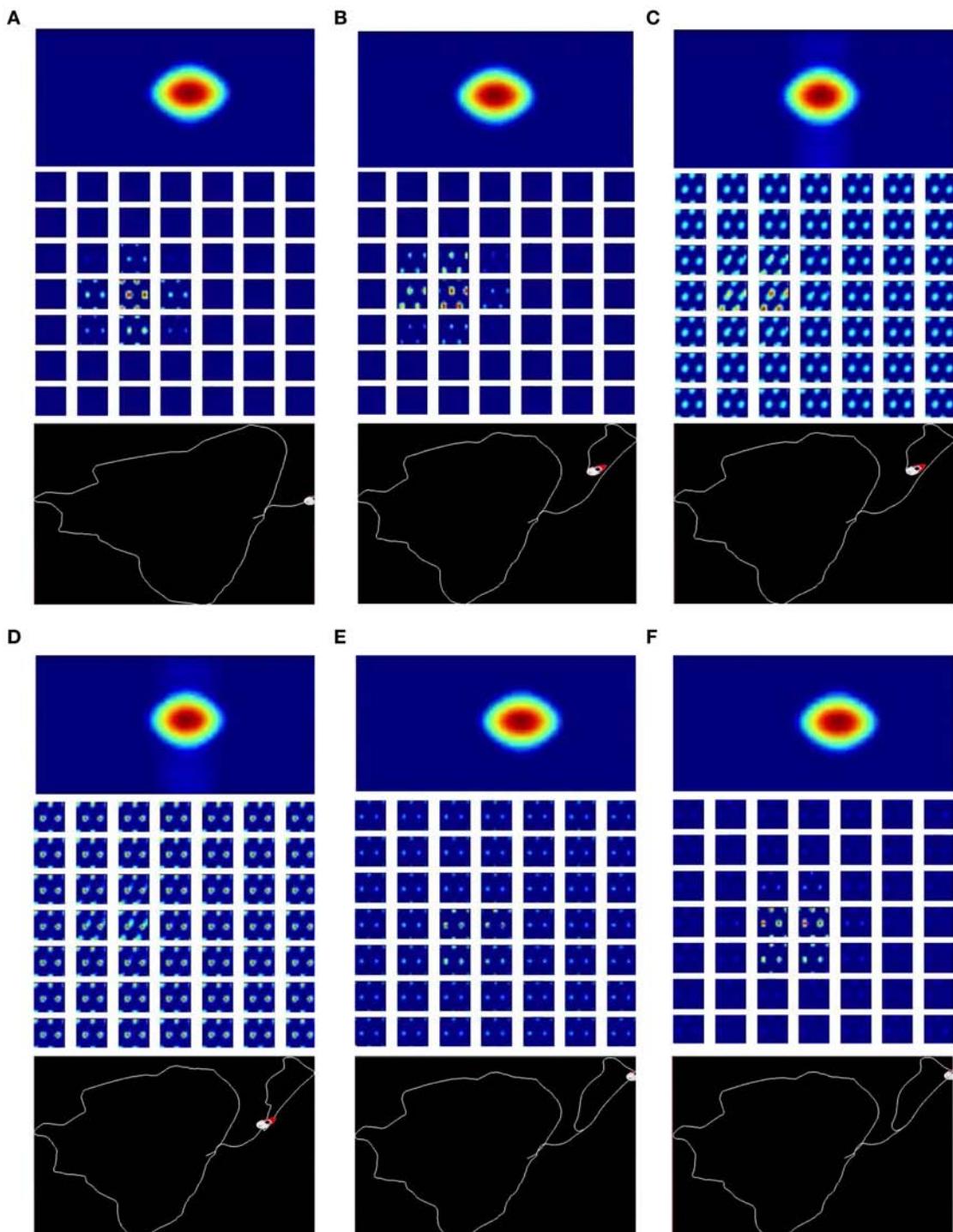


FIGURE 7 | Calibration of the network states during loop closure by the help of visual feedback from the local view cells. There are three parts in each panel: the activity of the HD-by-velocity network (top), the activity of the grid-by-velocity network (middle), and the corresponding experience map at the time (bottom). **(A)** The robot drives along a path for the first time. The activity bump of the HD-by-velocity network is centered at (3.3763, 0.0023) and the bumps of the grid-by-velocity network are at (1.1781, 5.7298, -0.1027, -0.0191); **(B)** before the robot revisits a familiar place, the center of the bump of the HD-by-velocity network moves to (3.0973, 0.0021) and the center of the bumps of the grid-by-velocity network (bottom) is at (2.2176, 2.3562, -0.1272, -0.0158); **(C)** The robot revisits the same place as in **(A)** and starts to close a loop. The local view cells inject currents to the HD-by-velocity network and the grid-by-velocity network, and create weaker bumps at the same phases as the network states in **(A)**; **(D)** after multiple continuous current injections, the networks retrieve similar phases as those when the robot first visited the place; **(E)** once the robot enters a new area, the local view cells become inactive, and the network states start to sharpen; **(F)** After loop closures, the network states are modulated by the velocity inputs. The bumps revert to their localized shapes.

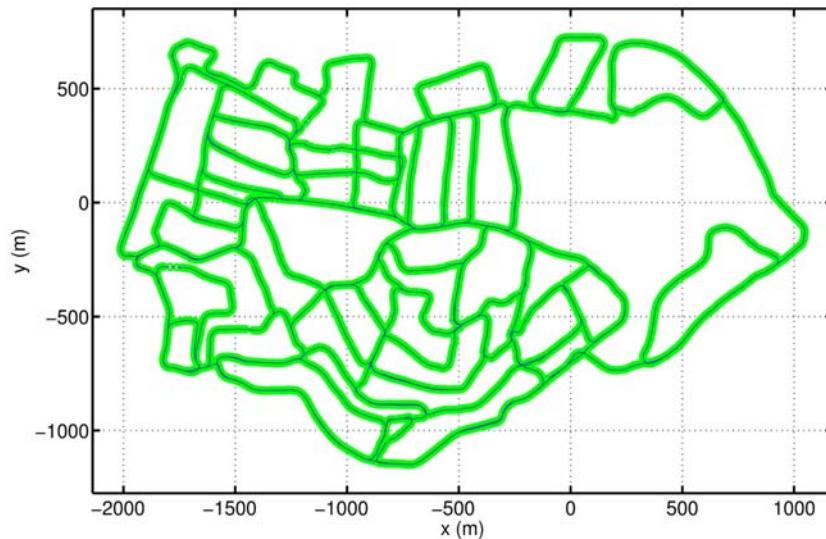


FIGURE 8 | The semi-metric cognitive map of St Lucia formed in the cognitive mapping model. The green thick line comprises topological graph vertices, and the blue thin line consists of links between connected vertices.

directions. The unit does not fire on many of the roads parallel to the southwest direction, since on those roads the robot moves northeast, i.e., opposite to the preferred firing direction of the unit. This can be confirmed by the firing rate map of a unit with opposite preferred head direction, shown in **Figure 9B**. This unit has the label $(\pi, 0.0024)$. This unit fires in broad directions close to northeast, and its firing rate map does not overlap with that of the unit in **Figure 9A**, meaning that it keeps silent when the unit in **Figure 9A** fires. This unit prefers a counterclockwise rotational velocity of 0.24 rad/s . Therefore, on bending paths, when the robot turns counterclockwise at about northeast directions, its firing rate always increases first and then decreases, which means that the head direction and the angular velocity of the robot become close to and away from the preferred conjunction of direction and rotational velocity of this unit. **Figure 9C** shows the firing rate map of the HD units with label 0 in the dimension θ . The firing pattern in **Figure 9C** indicates these units encode quite large range of directions centered at southwest. In local regions, these cells conserve consistent preference for directions.

Due to the periodic boundary conditions of the grid pattern in the conjunctive grid-by-velocity cell network, each grid unit fires at multiple distinct locations in the environment. **Figures 10A,B** give the firing rate maps of two example grid units in the network with the same spatial label but different velocity labels. In each of these firing fields, the firing rate always gradually increases when the robot moves closer to the center of the firing field and decreases bit by bit when the robot leaves the field center. The firing fields of the two example grid units overlap a lot, since they have the same spatial preference. The difference in firing field comes from the fact that the grid unit in **Figure 10B** prefers high translational velocity, however, the grid unit in **Figure 10A** prefers low translational velocity. Therefore, the grid units in **Figure 10A** fires in many locations close to the turns of the road, where the robot has to slow down significantly. The total

activity of the grid units, which share the same spatial preference as the units in **Figures 10A,B** is shown in **Figure 10C**. The firing map is composed of many distributed firing fields, but they are lack of global grid structure. This is similar to the fragmented firing pattern observed in hairpin maze (Deridkman et al., 2009), where the animal also runs on linear paths. Note that the distance between the grid fields of the units in the network is much larger than the typical spacing observed in animal experiments. This is due to the fact that the velocity input to the model is defined in the visual space of the input images and is different from the actual velocity of the robot.

4. DISCUSSION

We proposed a robust cognitive mapping model, which integrates movement and sensory information and forms stable spatial representations of the environment. We demonstrated that our cognitive mapping model using only visual sensory information could successfully build a coherent topological map of large-scale outdoor environment from an open-source dataset of 66 km urban car journey (see Video S1 in Supplementary Materials).

There are two major contributions in this study. First, the network architecture of our cognitive mapping model is consistent with the synaptic organization of the MEC. Our model is built upon layer III and deep layers of the MEC. Layer III and deep layers of the MEC may be the main site to perform key computations for navigation, and the layer II of the MEC may merely relays information to the hippocampus and functions as the output of the MEC. Supporting experimental evidence shows that there exist strong projections from deep layers to superficial layers of the MEC. HD cells, conjunctive grid-by-head-direction cells, and speed cells are abundant in layer III and deep layers of the MEC, and encode direction and speed information needed in path integration. However, layer II of the MEC contains mainly pure

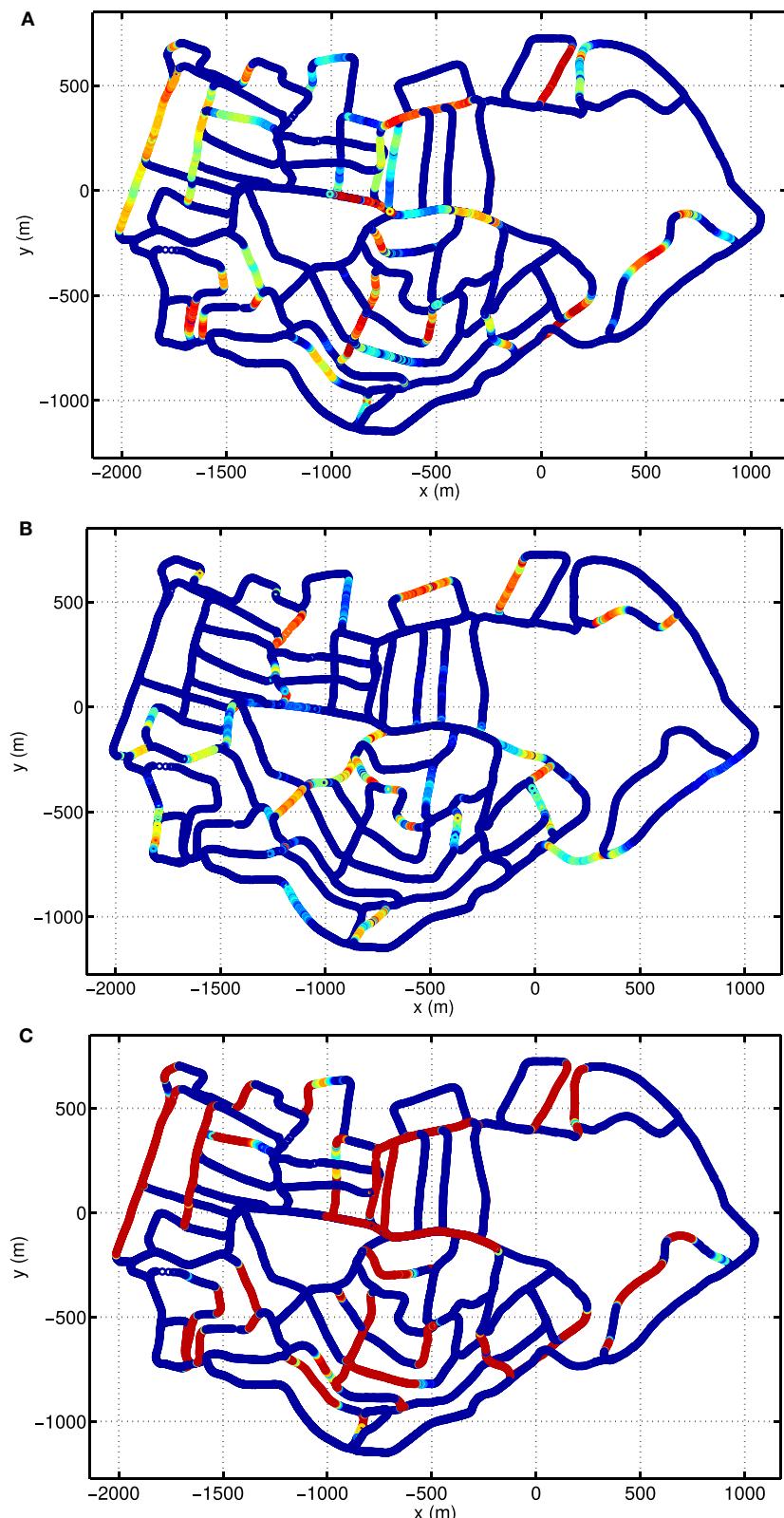


FIGURE 9 | Firing rate maps of example HD-by-velocity units. In each panel, the firing rate is plotted at the locations in the experience map. Firing rate is color-coded by the same jet colormap, i.e., red for high firing rate and blue for zero firing rate. **(A)** The firing rate map of the unit at $(0, 0)$ in the (θ, ν) manifold; **(B)** The firing rate map of the unit at $(\pi, 0.0024)$, which has the opposite preference in head direction as the unit in **(A)**; **(C)** The total activity of all the units with the same label $\theta = 0$, i.e., summing firing rates of all the units along the dimension ν at $\theta = 0$.

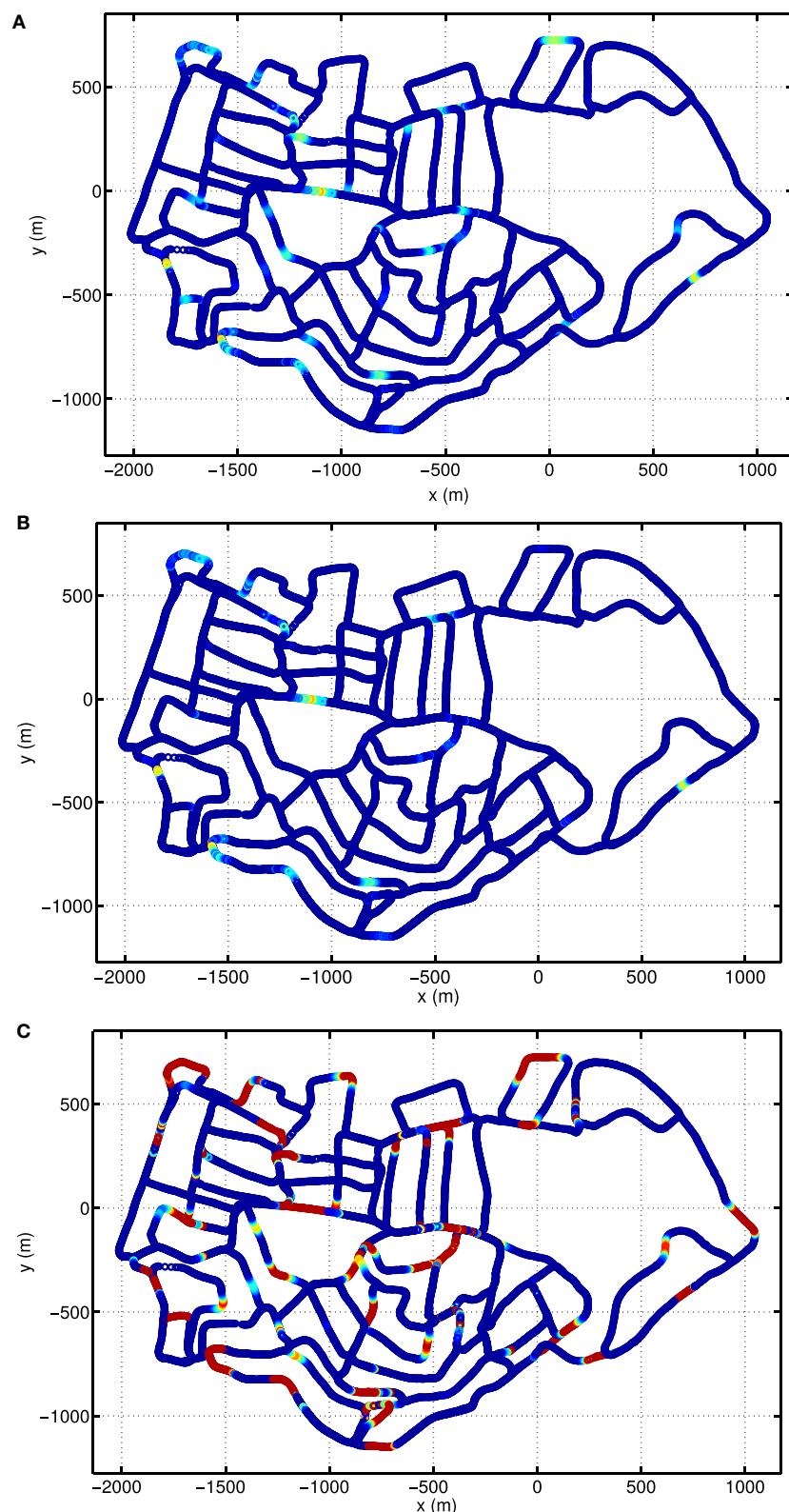


FIGURE 10 | Firing rate maps of example conjunctive grid-by-velocity units. In each panel, firing rate is plotted at the locations in the experience map. For all panels, the same colormap is used to code firing rate, red for high firing rate and gradually changing to blue for firing rate decreasing to zero. **(A)** The firing rate map of the unit at $(\pi, \pi, 0, 0)$ in the $(\theta_x, \theta_y, \nu_x, \nu_y)$ manifold; **(B)** the firing rate map of the unit at $(\pi, \pi, 0.1, 0.1)$; **(C)** the map of the total activity of the units at (π, π) in the (θ_x, θ_y) manifold, which sums the firing rates of the units along the dimensions ν_x and ν_y .

positional grid cells. Second, our cognitive mapping model provides coherent robust mechanisms to perform path integration. Both the HD cells and the conjunctive grid cells in the model integrate motion and perception information using the same principle. The performance of the model is manifested by experiments in large-scale natural environment. Compared with the pose cell network in Ball et al. (2013), our model achieves similar mapping results as those shown in Ball et al. (2013). Although our model reuses some of the vision and map operation methods of Ball et al. (2013), the core components of the HD cell network and the conjunctive grid cell network in our system are fundamentally different from the pose cell network. In our model, path integration is performed by the intrinsic attractor dynamics driven by velocity inputs, rather than simply displacing the copy of the activity package. In the proposed model, the convergence process of local view calibration is also performed by the dynamics of the attractors, instead of global inhibition.

The grid cells in our model does not express grid-like firing pattern in large-scale environment (**Figure 10**). Meanwhile, the HD cells in the model fire at a broad range of directions (**Figure 9**). This is due to the fact that local view cells anchor grids to local cues and alters the firing patterns of the units. It has been shown that in environments with abundant local cues, such like corridors, grids cells show fragmented firing maps (Derdikman et al., 2009). It is much harder for animals to form a globally consistent metric map in large than in small environments. Indeed, as shown in Carpenter et al. (2015), grid cells in rats acquire globally coherent representations after repeated exploration, while initially grid maps are only locally consistent.

Several potential limitations exist in the current study. First, the number of units in the system is large and requires substantial amount of computational resources. Second, place cell network is not included in our system. The function of place cells is not to provide a metric map. Experiments showed that in large-scale environments, place cell has multiple irregularly spaced place fields (Rich et al., 2014; Liu et al., 2015). A metric map is necessary if direct interpretation is preferred. For robotic applications, it would suffice to have distributed neural representations if the robots are able to distinguish different locations of the environment.

REFERENCES

- Bailey, T., and Durrant-Whyte, H. (2006). Simultaneous localization and mapping (SLAM): part II. *IEEE Robot. Autom. Mag.* 13, 108–117. doi:10.1109/MRA.2006.1678144
- Ball, D., Heath, S., Wiles, J., Wyeth, G., Corke, P., and Milford, M. (2013). OpenRAT-SLAM: an open source brain-based SLAM system. *Auton. Robots* 34, 149–176. doi:10.1007/s10514-012-9317-9
- Burak, Y., and Fiete, I. R. (2009). Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput. Biol.* 5:e1000291. doi:10.1371/journal.pcbi.1000291
- Burgess, N., Barry, C., and O'Keefe, J. (2007). An oscillatory interference model of grid cell firing. *Hippocampus* 17, 801–812. doi:10.1002/hipo.20327
- Carpenter, F., Manson, D., Jeffery, K., Burgess, N., and Barry, C. (2015). Grid cells form a global representation of connected environments. *Curr. Biol.* 25, 1176–1182. doi:10.1016/j.cub.2015.02.037
- Cheng, S., and Frank, L. M. (2011). The structure of networks that produce the transformation from grid cells to place cells. *Neuroscience* 197, 293–306. doi:10.1016/j.neuroscience.2011.09.002
- Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-B., et al. (2009). Fragmentation of grid cell maps in a multicompartment environment. *Nat. Neurosci.* 12, 1325–1332. doi:10.1038/nn.2396
- Ding, L., Xiao, L., Liao, B., Lu, R., and Peng, H. (2017). An improved recurrent neural network for complex-valued systems of linear equation and its application to robotic motion tracking. *Front. Neurorobot.* 11:45. doi:10.3389/fnbot.2017.00045
- Ducket, T., Marsland, S., and Shapiro, J. (2002). Fast, on-line learning of globally consistent maps. *Auton. Robots* 12, 287–300. doi:10.1023/A:1015269615729
- Dudek, G., and Jenkin, M. (2010). *Computational Principles of Mobile Robotics*. New York, NY: Cambridge University Press.
- Durrant-Whyte, H., and Bailey, T. (2006). Simultaneous localization and mapping: part I. *IEEE Robot. Autom. Mag.* 13, 99–110. doi:10.1109/MRA.2006.1638022
- Erlhagen, W., and Bicho, E. (2006). The dynamic neural field approach to cognitive robotics. *J. Neural Eng.* 3, R36. doi:10.1088/1741-2560/3/3/R02
- Fuhs, M. C., and Touretzky, D. S. (2006). A spin glass model of path integration in rat medial entorhinal cortex. *J. Neurosci.* 26, 4266–4276. doi:10.1523/JNEUROSCI.4353-05.2006

For future research, we plan to include a place cell network, which integrates multiple grid cell modules with diversity of spacings and orientations, and investigate its function in path planning and reward-based learning. We will also investigate neurobiologically inspired algorithms to improve the computational efficiency of the system for robot navigation.

5. CONCLUSION

In summary, conjunctive space-by-movement attractor network models are proposed in this paper to achieve cognitive mapping in a large-scale natural environment. Head direction cells and conjunctive grid cells work on the same principle, and represent positions, head directions, and velocity at the same time. Our model facilitates to reach a better understanding of the neural mechanisms of spatial cognition and to develop more accurate models of the brain. Furthermore, it inspires innovative high performance cognitive mapping systems, which are able to function in dynamical natural environment with long-term autonomy.

AUTHOR CONTRIBUTIONS

TZ and BS conceived the model, performed experiments, analyzed the results, and wrote the paper.

FUNDING

This work is supported by the Distinguished Young Scholar Project of the Thousand Talents Program of China (grant No. Y5A1370101) and the State Key Laboratory of Robotics of China (grant No. Y5A1203901).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://www.frontiersin.org/article/10.3389/fnbot.2017.00061/full#supplementary-material>.

VIDEO S1 | Mapping process.

VIDEO S2 | An example of pattern completion process.

- Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., and Moser, E. (2007). Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446, 190–194. doi:10.1038/nature05601
- Gerstner, W., and Abbott, L. (1997). Learning navigational maps through potentiation and modulation of hippocampal place cells. *J. Comput. Neurosci.* 4, 79–94. doi:10.1023/A:1008820728122
- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801–806. doi:10.1038/nature03721
- Jauffret, A., Cuperlier, N., and Gaussier, P. (2015). From grid cells and visual place cells to multimodal place cell: a new robotic architecture. *Front. Neurorobot.* 9:1. doi:10.3389/fnbot.2015.00001
- Klaassen, R. H., Alerstam, T., Carlsson, P., Fox, J. W., and Lindström, Å (2011). Great flights by great snipes: long and fast non-stop migration over benign habitats. *Biol. Lett.* 7, 833–835. doi:10.1098/rsbl.2011.0343
- Knierim, J. J. (2015). From the gps to hm: place cells, grid cells, and memory. *Hippocampus* 25, 719–725. doi:10.1002/hipo.22453
- Knierim, J. J., and Zhang, K. (2012). Attractor dynamics of spatially correlated neural activity in the limbic system. *Annu. Rev. Neurosci.* 35, 267–285. doi:10.1146/annurev-neuro-062111-150351
- Knips, G., Zibner, S. K., Reimann, H., and Schöner, G. (2017). A neural dynamic architecture for reaching and grasping integrates perception and movement generation and enables on-line updating. *Front. Neurorobot.* 11:9. doi:10.3389/fnbot.2017.00009
- Kropff, E., Carmichael, J. E., Moser, M.-B., and Moser, E. I. (2015). Speed cells in the medial entorhinal cortex. *Nature* 523, 419–424. doi:10.1038/nature14622
- Kropff, E., and Treves, A. (2008). The emergence of grid cells: intelligent design or just adaptation? *Hippocampus* 18, 1256–1269. doi:10.1002/hipo.20520
- Liu, S., Si, B., and Lin, Y. (2015). “Self-organization of hippocampal representations in large environments,” in *Proceedings of the 2015 International Joint Conference on Neural Networks*, Killarney.
- Mathis, A., Herz, A. V., and Stemmler, M. (2012). Optimal population codes for space: grid cells outperform place cells. *Neural Comput.* 24, 2280–2317. doi:10.1162/NECO_a_00319
- McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M.-B. (2006). Path integration and the neural basis of the ‘cognitive map’. *Nat. Rev. Neurosci.* 7, 663–678. doi:10.1038/nrn1932
- Mi, Y., Katkov, M., and Tsodyks, M. (2017). Synaptic correlates of working memory capacity. *Neuron* 93, 323–330. doi:10.1016/j.neuron.2016.12.004
- Milford, M. J., and Wyeth, G. F. (2008). Mapping a suburb with a single camera using a biologically inspired SLAM system. *IEEE Trans. Robot.* 24, 1038–1053. doi:10.1109/TRO.2008.2004520
- Mulas, M., Wanek, N., and Conradt, J. (2016). Hebbian plasticity realigns grid cell activity with external sensory cues in continuous attractor models. *Front. Comput. Neurosci.* 10:13. doi:10.3389/fncom.2016.00013
- O’Keefe, J., and Conway, D. (1978). Hippocampal place units in the freely moving rat: why they fire where they fire. *Exp. Brain Res.* 31, 573–590.
- O’Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175. doi:10.1016/0006-8993(71)90358-1
- Remme, M. W., Lengyel, M., and Gutkin, B. S. (2010). Democracy-independence trade-off in oscillating dendrites and its implications for grid cells. *Neuron* 66, 429–437. doi:10.1016/j.neuron.2010.04.027
- Rich, P. D., Liaw, H.-P., and Lee, A. K. (2014). Large environments reveal the statistical structure governing hippocampal representations. *Science* 345, 814–817. doi:10.1126/science.1255635
- Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M.-B., et al. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312, 758–762. doi:10.1126/science.1125572
- Sheynikhovich, D., Chavarriaga, R., Strösslin, T., Arleo, A., and Gerstner, W. (2009). Is there a geometric module for spatial orientation? Insights from a rodent navigation model. *Psychol. Rev.* 116, 540–566. doi:10.1037/a0016170
- Si, B., Kropff, E., and Treves, A. (2012). Grid alignment in entorhinal cortex. *Biol. Cybern.* 106, 483–506. doi:10.1007/s00422-012-0513-7
- Si, B., Romani, S., and Tsodyks, M. (2014). Continuous attractor network model for conjunctive position-by-velocity tuning of grid cells. *PLoS Comput. Biol.* 10:e1003558. doi:10.1371/journal.pcbi.1003558
- Si, B., and Treves, A. (2009). The role of competitive learning in the generation of DG fields from EC inputs. *Cogn. Neurodyn.* 3, 177–187. doi:10.1007/s11571-009-9079-z
- Si, B., and Treves, A. (2013). A model for the differentiation between grid and conjunctive units in medial entorhinal cortex. *Hippocampus* 23, 1410–1424. doi:10.1002/hipo.22194
- Smith, R. C., and Cheeseman, P. (1986). On the representation and estimation of spatial uncertainty. *Int. J. Rob. Res.* 5, 56–68. doi:10.1177/027836498600500404
- Solstad, T., Moser, E. I., and Einevoll, G. T. (2006). From grid cells to place cells: a mathematical model. *Hippocampus* 16, 1026–1031. doi:10.1002/hipo.20244
- Stella, F., Si, B., Kropff, E., and Treves, A. (2013). Grid maps for space-flight, anyone? They are for free! *Behav. Brain Sci.* 36, 566–567. doi:10.1017/S014025X13000575
- Stella, F., and Treves, A. (2015). The self-organization of grid cells in 3d. *Elife* 4, e05913. doi:10.7554/eLife.05913
- Stensola, T., Stensola, H., Moser, M.-B., and Moser, E. I. (2015). Shearing-induced asymmetry in entorhinal grid cells. *Nature* 518, 207–212. doi:10.1038/nature14151
- Taube, J. S. (2007). The head direction signal: origins and sensory-motor integration. *Annu. Rev. Neurosci.* 30, 181–207. doi:10.1146/annurev.neuro.29.051605.112854
- Taube, J. S., Muller, R. U., and Ranck, J. B. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. Neurosci.* 10, 420–435.
- Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents Series)*. Cambridge, MA: The MIT Press.
- Thrun, S., and Leonard, J. J. (2008). “Simultaneous localization and mapping,” in *Springer Handbook of Robotics*, eds M. E. Jefferies and W.-K. Yeap (Berlin, Heidelberg: Springer), 871–889.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189. doi:10.1037/h0061626
- Treves, A. (1990). Graded-response neurons and information encodings in autoassociative memories. *Phys. Rev. A* 42, 2418. doi:10.1103/PhysRevA.42.2418
- Tsoar, A., Nathan, R., Bartan, Y., Vyssotski, A., DellOmo, G., and Ulanovsky, N. (2011). Large-scale navigational map in a mammal. *Proc. Natl. Acad. Sci. U.S.A.* 108, E718–E724. doi:10.1073/pnas.1107365108
- Welday, A. C., Shlifer, I. G., Bloom, M. L., Zhang, K., and Blair, H. T. (2011). Cosine directional tuning of theta cell burst frequencies: evidence for spatial coding by oscillatory interference. *J. Neurosci.* 31, 16157–16176. doi:10.1523/JNEUROSCI.0712-11.2011
- Xiao, L., and Zhang, Y. (2016). Dynamic design, numerical solution and effective verification of acceleration-level obstacle-avoidance scheme for robot manipulators. *Int. J. Syst. Sci.* 47, 932–945. doi:10.1080/00207721.2014.909971
- Xiao, L., Zhang, Y., Liao, B., Zhang, Z., Ding, L., and Jin, L. (2017). A velocity-level bi-criteria optimization scheme for coordinated path tracking of dual robot manipulators using recurrent neural network. *Front. Neurorobot.* 11:47. doi:10.3389/fnbot.2017.00047
- Yuan, M., Tian, B., Shim, V. A., Tang, H., and Li, H. (2015). “An Entorhinal-Hippocampal Model for simultaneous cognitive map building,” in *Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin Texas.
- Zeno, P. J., Patel, S., and Sobh, T. M. (2016). Review of neurobiologically based mobile robot navigation system research performed since 2000. *J. Rob.* 2016:8637251. doi:10.1155/2016/8637251
- Zhang, W.-H., Chen, A., Rasch, M. J., and Wu, S. (2016). Decentralized multisensory information integration in neural systems. *J. Neurosci.* 36, 532–547. doi:10.1523/JNEUROSCI.0578-15.2016
- Zilli, E. A. (2012). Models of grid cell spatial firing published 2005–2011. *Front. Neural Circuits* 6:16. doi:10.3389/fncir.2012.00016

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Zeng and Si. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.