Brain-Like Navigation Study Partial Notes

1. Biological background:

Navigating in three dimensional environments is a critical capability for many current or prospective robotic tasks. Inspired by the ability of mammals to navigate to hidden target locations via direct routes that may traverse previously unvisited terrain. It is thought that mammals are able to use internal spatial representations to navigate directly to a target location without following explicit sensory cues or a sequence of learned actions. According to current advances in brain neuroscience and life sciences, place cells, head direction cells, grid cells, etc., play an important role in biological navigation.

1. **Place cells :** place cells usually have a single spatial receptive field and fire when the animal enters a specific area.
2. **Head direction cells :** Head direction cells are a type of neurons in the mammalian brain. They can discharge when the animal is oriented in a particular direction.
3. **Grid cells :**grid cells exhibit several properties that afford large-scale vector navigation. Grid cells also show stable spatial firing correlates but with multiple firing fields distributed in a regular hexagonal array that covers all environments visited by the animal.

2. Architecture and Detailed model:

2.1.1 3D grid cells:
Represent 3D position and metric information for 3D path integration.

2.1.2 Head direction cells :
A multilayered head direction cell model is used to represent the robotic orientation.

2.1.3 Multidimensional continuous attractor network :
The MD-CAN has many recurrent connections which cause the network to converge over time to certain stable states (attractors, activity packets or bumps) in the absence of external input.The 2D MD-CAN and 3D MD-CAN are used to represent the multilayered head direction cell model and 3D grid cell model respectively.

2.2Architecture:

The NeuroSLAM system architecture is in Fig. 1. In 3D, the robot's 4DoF pose (x, y, z, yaw) is jointly represented by 3D grid and head direction cell networks. The pose cell network does path integration via self-motion cues and calibration with local visual cues. The multilayer graphical experience map's creation and relaxation rely on local view cells, pose cells, and 3D visual odometry.

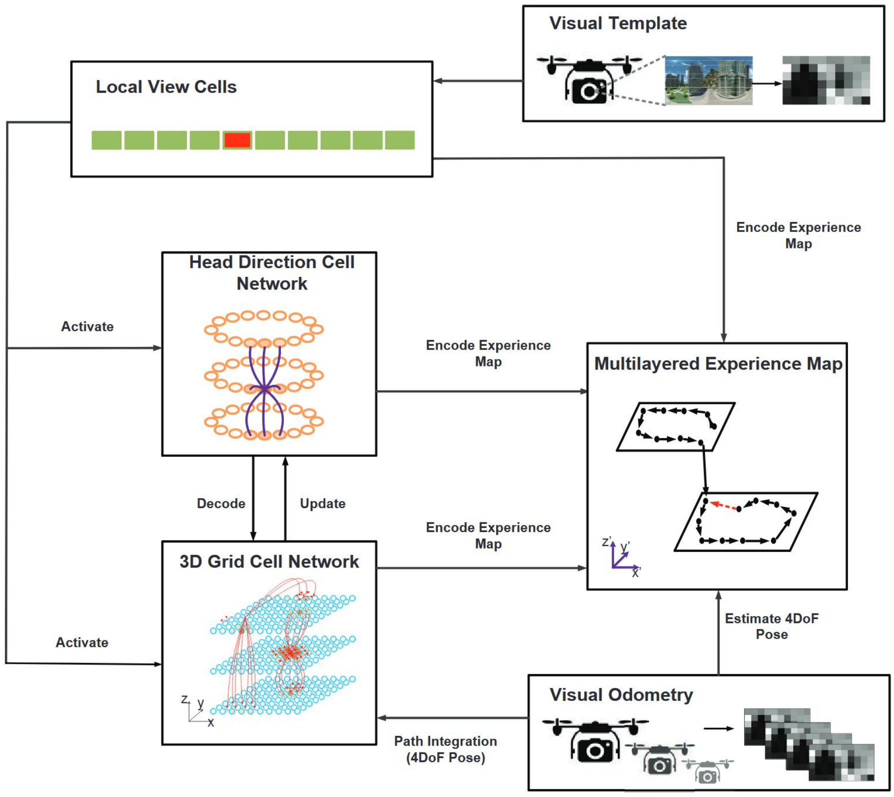

Fig. 1 The NeuroSLAM architecture. The system consists of conjunctive pose cells combining the 3D grid cells and multilayered head direction cells, the multilayered experience map and vision modules. The conjunctive pose cell network performs path integration based on the local view cues and self-motion information. The distinct scenes are encoded by the local view cells. The output from the local view cells, the 3D grid cells, the multilayered head direction cells and the 3D visual odometry drives the creation of a 3D multilayered experience map, which is a hybrid spatial representation with a topological and locally metric 3D graphical map of the 3D environment.

2.3 Conjunctive pose cell model :The conjunctive pose cells consisting of 3D grid cells and multilayered head direction cells to represent a 4DoF pose (x, y, z, yaw).

2.3.1 3D grid cell model :

The 3D grid cell network is a 3D MD-CAN. It mimics the 3D spatial neural representation in the mammalian brain, as shown in Fig. 2.The 3D grid cells represent the absolute location (x, y, z) in 3D space. The activity matrix P^{gc} describes the activity in the 3D grid cells.Updating activities based on three key processes.

1. The activity is updated by the attractor dynamics with excitation and inhibition.
2. The activity packets are shifted by 3D path integration based on the translational and rotational velocity provided by 3D visual odometry.
3. When the robot sees familiar scenes, the activity is updated by local view calibration.

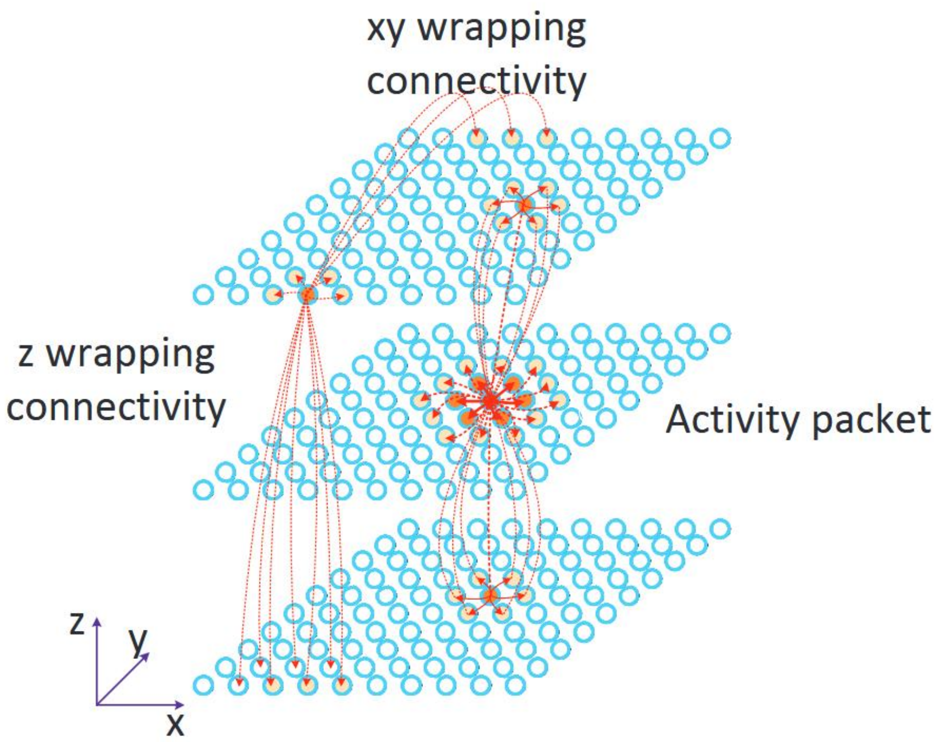

Fig. 2 3D grid cell network model. A 3D MD-CAN (3D cube) represents the 3D grid cells with a stable activity packet. The warp connections of neural units can represent 3D location periodically. Each 3D grid cell is activated when the robot moves in a large scale space.

Attractor dynamics:

The process of the internal attractor dynamics in the 3D MD-CAN includes three stages:

1. Parts of 3D grid cells are excited by the local excitatory process.
2. All 3D grid cells are inhibited by the global inhibition process.
3. The activity of the 3D grid cells is normalized.

Local excitation

We create the excitatory weight matrix $\varepsilon_{u,v,w}^{gc}$ using a 3D Gaussian distribution. The distance indexes between units are represented by u, v and w. The weight is calculated by

$$\varepsilon_{u,v,w}^{gc} = \frac{1}{(\delta_x\sqrt{2\pi})} e^{-u^2/(2\delta_x^2)} \cdot \frac{1}{(\delta_y\sqrt{2\pi})} e^{-v^2/(2\delta_y^2)} \cdot \frac{1}{(\delta_z\sqrt{2\pi})} e^{-w^2/(2\delta_z^2)}$$

where δ_x, δ_y and δ_z are the constants of variance for 3D spatial distributions. The activity change in a 3D grid cell is calculated by

$$\Delta P^{gc}_{x,y,z} = \sum_{i}^{n_x}\sum_{j}^{n_y}\sum_{k}^{n_z} P^{gc}_{i,j,k}\varepsilon^{gc}_{u,v,w}$$

where the three dimensions of the matrix are n_x, n_y, n_z. The distance indexes are calculated by

$$u = (x - i)(\mathrm{mod}\ n_x)$$
$$v = (y - j)(\mathrm{mod}\ n_y)$$
$$w = (z - k)(\mathrm{mod}\ n_z)$$

Global inhibition

Each 3D grid cell inhibits nearby cells by a local inhibitory process. We create an inhibitory weight matrix $\psi^{gc}_{u,v,w}$ to update the activity during the local inhibitory process. Then the activity of all 3D grid cells is updated by the global inhibition φ equally. The processes of the local inhibition and the global inhibition are calculated by

$$\Delta \mathbf{P}^{\mathbf{gc}}_{\mathbf{x,y,z}} = \sum_{\mathbf{i}}^{\mathbf{n_x}}\sum_{\mathbf{j}}^{\mathbf{n_y}}\sum_{\mathbf{k}}^{\mathbf{n_z}} \mathbf{P}^{\mathbf{gc}}_{\mathbf{i,j,k}}\boldsymbol{\psi}^{\boldsymbol{gc}}_{\boldsymbol{u,v,w}} - \boldsymbol{\varphi}$$

where we control all values in P^{gc} to nonnegative values.

Activity normalisation

The total activity in 3D grid cells is kept one by activity normalisation. The activity $\mathbf{P}^{gc'}_{x,y,z}$ is calculated by

$$\mathbf{P}^{gc'}_{x,y,z} = \frac{\mathbf{P}^{gc}_{x,y,z}}{\sum_{i}^{n_x}\sum_{j}^{n_y}\sum_{k}^{n_z}\mathbf{P}^{gc}_{i,j,k}}$$

In the following sections, we describe the update process of the activity in 3D grid cells by 3D path integration and local view calibration.

Path integration

The path integration projects the 3D grid cells activity into nearby cells. The activity is shifted in x, y plane and z dimension according to the translational velocity V and height change velocity Vh along x, y, z axis respectively under current head direction in yaw(θ). The activity change ΔU^{gc}_{lmn} is calculated by

$$\Delta U^{gc}_{lmn} = \sum_{x=\delta_{x_0}}^{\delta_{x_0}+1}\sum_{y=\delta_{y_0}}^{\delta_{x_0}+1}\sum_{z=\delta_{z_0}}^{\delta_{x_0}+1} \gamma\, U^{gc}_{(l+x)(m+y)(n+z)}$$

The amount of activity injected is determined by two inputs One is from the product of the sending unit. U^{gc} .Another is from the residue, γ. The residue is calculated according to the fractional parts of the offsets, $\delta_{x_f}, \delta_{y_f}, \delta_{z_f}$.

$$\begin{bmatrix}\delta_{x_0}\\ \delta_{y_0}\\ \delta_{z_0}\end{bmatrix} = \begin{bmatrix}\lfloor k_x v\cos\theta\rfloor\\ \lfloor k_y v\sin\theta\rfloor\\ \lfloor k_z v_h\rfloor\end{bmatrix}$$

$$\begin{bmatrix}\delta_{x_f}\\ \delta_{y_f}\\ \delta_{z_f}\end{bmatrix} = \begin{bmatrix}\lfloor k_x v\cos\theta - \delta x_0\rfloor\\ \lfloor k_y v\sin\theta - \delta y_0\rfloor\\ \lfloor k_z v_h - \delta z_0\rfloor\end{bmatrix}$$

where k_x, $k_y$and k_z are constants for 3D path integration The γ is calculated by

$$\gamma = f\left(\delta_{x_f}, x - \delta_{x_0}\right) f\left(\delta_{y_f}, y - \delta_{y_0}\right) f\left(\delta_{z_f}, z - \delta_{z_0}\right)$$

$$f(a, b) = \begin{cases} a, & \text{if } b = 1; \\ 1 - a, & \text{if } b = 0. \end{cases}$$

Local view calibration

The local view calibration resets the accumulative errors in path integration based on the translational and rotational velocity provided by 3D visual odometry. The local view cells are associated with the 3D grid cells and the multilayered head direction cells. When the robot sees familiar view, the prior associations are recalled. We use a vector V to represent the activity of the local view cells. If the current view is similar to the previous view, the associated local view cell is active. The connection matrix Ψ stores the learned connections among the local view cell vector, the 3D grid cell matrix and the multilayered head direction cell matrix. We use a modified version of Hebbs law for learning connections. The connection between the $P_{x,y,z,\theta}$ and V_i is calculated by

$$\Psi^{t+1}_{i,x,y,z,\theta} = \max\left(\tau V_i P_{x,y,z,\theta}, \Psi^{t}_{i,x,y,z,\theta}\right)$$

The τ is a learning rate. The activity change in 3D grid cells and multilayered head direction cells is calculated by

$$\triangle P_{x,y,z,\theta} = \frac{\delta}{n_{\text{act}}} \sum_i V_i \Psi_{i,x,y,z,\theta}$$

where the constant δ controls the strength of local view calibration. The n_{act} is the number of active local view cells.