A Model for Learning Representations of 3D Objects Through Tactile Exploration: Effects of Object Asymmetries and Landmarks

Xiaogang Yan, Alistair Knott, and Steven Mills

University of Otago, Dunedin, New Zealand
{yanxg,alik,steven}@cs.otago.ac.nz

Abstract. In this paper, we develop a neural network model that learns representations of 3D objects via tactile exploration. The basic principle is that the hand is considered as an autonomous 'navigating agent', traveling within the 'environment' of a 3D object. We adapt a model of hippocampal place cells, which learns the structure of a 2D environment by exploiting constraints imposed by the environment's boundaries on the agent's movement, and perceptual information about landmarks in the environment. In the current paper, our focus is on 3D analogues of these 2D information sources. We systematically investigate the information about object geometry that is provided by navigation constraints in a simple cuboid, and by tactile landmarks. We find that an asymmetric cuboid conveys more information to the navigator than a symmetric cuboid (i.e., a cube) – and that landmarks convey additional information independently from asymmetry.

Keywords: Hippocampal place cells · 3D object representation · Tactile exploration · Landmarks · Recurrent self-organizing map.

1 Introduction

When a human being enters an environment, hippocampal place cells develop a cognitive map of the environment. While the person reaches one location in the environment, one place cell or multiple place cells fire simultaneously, which represents such a location in the navigation environment. The process of hippocampal cells encoding spatial locations by the integration of linear and angular self motions is called 'path integration' or 'dead reckoning' [9, 10].

Even though the exploring agent's movements are defined in an 'egocentric' reference frame, as are the perceptual stimuli it receives, the hippocampus can somehow derive from this egocentric information an 'allocentric' or 'environment-centered' representation of its location in the environment. In our current paper, we explore a 3D analogue of this navigation scenario, where the agent's hand is construed as traveling around the environment of a 3D object. Here again, information about the hand's movements and about landmarks arrive in an egocentric reference frame. We will focus on tactile information, which is more

(a) (b) (c)

Fig. 1. By executing 'egocentric' movements, an agent (here a snail) learns the 'allocentric' representation of the explored object, i.e, (a) a cube, (b) a cuboid, and (c) a cube with landmarks.

direct than visual information. The egocentric information in this case is defined in a 'hand-centered' coordinate system. From this egocentric information, the agent can construct an allocentric (i.e., object-centered) representation of the object's geometry.

For concreteness, we can visualize the 'agent' traveling around the cube as a snail, as shown in Fig. 1. The agent can move by translation (forward, back, left or right), or can change its orientation by rotating on its current plane. It can detect when it crosses onto a different plane of the cube. It can also sense tactile landmarks that it is sliding over (the colored dots). From these egocentric (snail-centered) cues, the agent can derive an environment-centered (i.e., object-centered) representation of the cube.

It is not yet understood how this is done. However, as a starting point, we can consider models of the 2D place cells system, which is one of the most studied and best understood structures in the brain [3–5]. The place cells model we will adopt is one that uses a self-organizing map (SOM) [7]: specifically, a SOM is modified to take recurrent input, called a modified SOM (MSOM) [11]. Note that we are not suggesting that hippocampal place cells are involved in haptic object exploration; there is good evidence that object representations derived from touch are developed in the parietal cortex [1, 2, 12]. However, we suggest that the parietal circuitry for learning haptic object representations might be isomorphic in some way to the hippocampal circuitry for learning 2D environment representations. Based on this assumption, we investigate what allocentric information about object geometry can be provided by constraints on hand navigation, and by tactile landmarks.

The organization of this paper is summarized as follows. Section 2 presents the background knowledge, which consists of MSOM, the relationship between constrained action sequences and the object topography and a revisit of a existing MSOM model activated by translative movements (T-MSOM) for 3D object representations shown in [13] with its drawbacks pointed out. The proposed translative and orientational movements activated MSOM (TO-MSOM) model and the landmarks together with translative and orientational movements activated MSOM (L-TO-MSOM) model are presented in Section 3. Section 4 shows simulative results of the proposed models for learning representations of two typical 3D objects. Finally, Section 5 concludes the paper with final remarks. The main contributions of this paper are highlighted as follows.

– Based on the authors' knowledge, this is the first time to present a neural network model for learning representations of 3D object via tactile exploration by executing both translative and orientational movements.
– Simulative results based on a 3D cube and cuboid demonstrate the effectiveness of the proposed models for learning representations of 3D objects. More importantly, the statistics and systematic analysis verify that the models are more accurate to learn a representation of a cuboid than a cube, which is owing to the contributive asymmetrical topography of a cuboid.
– The positive effect of landmarks is verified by the statistics analysis of simulative results of the models representing the cube and cuboid.

2 Background

In this section, we present the background knowledge for the proposed models. Specifically, the detailed description of MSOM algorithm is firstly presented. Then, the constraint of object topographies placed on action sequences for exploration is identified. After that, for comparison and for showing the contribution of this paper, drawbacks of the existing T-MSOM model are pointed out.

2.1 Modified Self-Organizing Map (MSOM)

Owing to the added previous state input, MSOM comes to learn frequently occurring input *sequences*, which is different from SOM learning the frequently occurring input *patterns* [3]. Regarding an input $x(t) \in \mathcal{R}^m$ at time instance t, the activity of unit i of a MSOM $\mathcal{M} \in \mathcal{R}^{n\times n}$ at that time instance is defined as

$$a_i(t) = \exp(-\eta d_i(t)), \tag{1}$$

where $i \in 1, 2, \cdots, n^2$, $\eta > 0$ is a design parameter, and $d_i(t)$ is a distance function, which is defined as a weighted sum of two parts. The first part is $\|x(t) - w_i(t)\|_2^2$ with $\|\cdot\|_2$ denoting the 2-norm of a matrix or vector, which is to evaluate the distance between the input $x(t)$ and the weight $w_i(t)$ of unit i (for simplicity, we name it as regular weight); and the second part is $\|c(t) - c_i(t)\|_2^2$, which is to evaluate the distance between the context weight $c(t)$ for the map $\mathcal{M}$ at time instance t and the individual context weight $c_i(t)$ of unit i. By introducing a weight factor $\xi \in (0, 1)$ to adjust the effect of such two parts on $d_i(t)$, the distance function $d_i(t)$ is formulated as

$$d_i(t) = (1 - \xi)\|x(t) - w_i(t)\|_2^2 + \xi\|c(t) - c_i(t)\|_2^2. \tag{2}$$

The context weight $c(t)$ for the map $\mathcal{M}$ in (2) is defined as

$$c(t) = (1 - \kappa)w^*(t - 1) + \kappa c^*(t - 1), \kappa \in (0, 1), \tag{3}$$

Algorithm 1 MSOM

Input: Input data $x(t) \in \mathcal{R}^m$

Output: A convergent MSOM $\mathcal{M} \in \mathcal{R}^{n \times n}$

1: Randomly initialize all m-dimensional regular weights $w_i(0) \in (0,1)$ and set all context weights $c_i(0) = \mathbf{0} \in \mathcal{R}^m$, $i = 1, 2, \dots n^2$

2: **while** feature map is not convergent **do**

3: Sampling: draw sample input $x(t) \in \mathcal{R}^m$

4: Competition: find best matching unit based on a distance discriminant function:

$$f(x(t)) = \underset{i}{\operatorname{argmin}} \ (1-\xi)\|x(t) - w_i(t)\|_2^2 + \xi\|c(t) - c_i(t)\|_2^2,$$

where $c(t) = (1-\kappa)w_{f(x(t-1))}(t-1) + \kappa c_{f(x(t-1))}(t-1)$, $\xi \in (0,1)$ $\kappa \in (0,1)$

5: Cooperation: select $f(x(t))$ neuron's neighbourhood neurons defined by a time-varying decreasing neighbourhood function $\mathcal{H}(i, f(x))(t)$

6: Adaptation: update regular weights and context weights of all selected neurons:

$$w_i(t+1) = w_i(t) + \mathcal{L}(t)\mathcal{H}(i, f(x(t)))(t)(x(t) - w_i(t)),$$

$$c_i(t+1) = c_i(t) + \mathcal{L}(t)\mathcal{H}(i, f(x))(t)(c(t) - c_i(t)),$$

where $\mathcal{L}(t)$ is a time-varying decreasing learning rate function.

7: **end while**

8: **return** $\mathcal{M}$

where $w^*(t-1)$ and $c^*(t-1)$ denote the regular weight and context weight of the unit in MSOM with the maximal activity $a_i(t)$ at previous time instance $t-1$, respectively. By norming the activities of all MSOM units shown in (1),

$$p_i(t) = \frac{a_i(t)}{\sum_{j=1}^{n^2} a_j(t)}, \tag{4}$$

which denotes the activity probability of unit i for the current input at time instance t. During training, the regular weight $w_i(t)$ is updated as

$$w_i(t+1) = w_i(t) + \mathcal{L}(t)\mathcal{H}(i, f(x(t)))(t)(x(t) - w_i(t)), \tag{5}$$

and the individual context weight $c_i(t)$ is changed as

$$c_i(t+1) = c_i(t) + \mathcal{L}(t)\mathcal{H}(i, f(x))(t)(c(t) - c_i(t)), \tag{6}$$

where $\mathcal{L}(t)$ and $\mathcal{H}(i, f(x(t)))(t)$ are a time-varying decreasing learning rate function and neighbourhood function respectively with $f(x)$ denoting the index of the unit in MSOM with the maximal activity for the current input $x(t)$. At the beginning of training, the regular weight $w_i(0) \in (0,1)$ is randomly selected and the context weight $c_i(0) = 0$. The process of MSOM is shown in Algorithm 1.

2.2 Action Sequences Constrained by Object Topographies

To lay a basis for further investigation, in this subsection, we present the relationship (more specifically, the constraint relationship) between navigation action sequences and object's topographies. Regarding the constraint on action

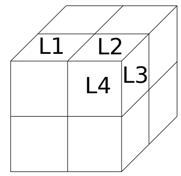


Fig. 2. Geometrical description of a cube with four locations L1, L2, L3 and L4.

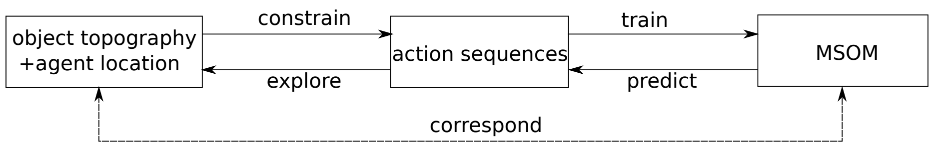


Fig. 3. Schematic of relationships among object topography and agent location, action sequences and the MSOM.

sequences played by object topographies, we can refer to a cube shown in Fig. 2. Assuming a navigation agent starts in location 'L1' facing Right, after moving directly forward, it reaches location 'L2' facing Right. Then, the agent could reach location 'L3' by moving forward over the edge or could get location 'L4' by moving right over the edge. Thus, from the same starting exploration position and orientation, different action sequences lead the navigation agent to different locations. Different object topographies support different exploration action sequences and thus, constrained action sequences implicitly contain object topography information. Relationships among action sequences, the object topography plus navigation location and the MSOM are illustrated in Fig. 3.

Without performing orientational movements, starting from 'L2' facing Right, after moving forward over the edge to reach 'L3', moving right over the edge to reach 'L4' and moving back over the edge to go to 'L2', the agent is back in its starting location – but importantly, it is now facing a different direction than it did when it started. This highlights an important geometrical property of navigation in 3D space – the 'non-commutativity of rotations' (a good discussion is given in [6]). For our purposes, the key point about this property is that our navigating agent needs the ability to rotate in its current plane, as well as to translate, to make the task of returning to a given state tractable. We begin by presenting a model with translative movements but no rotational movements (i.e., T-MSOM model), and then introduce a model including orientational movements as well (i.e., TO-MSOM model).

2.3 T-MSOM

A basic model T-MSOM to learn representations of 3D objects based on translative movements and surface information is presented in the previous work [13, 14]. Since orientational movements of a navigation agent are normally performed

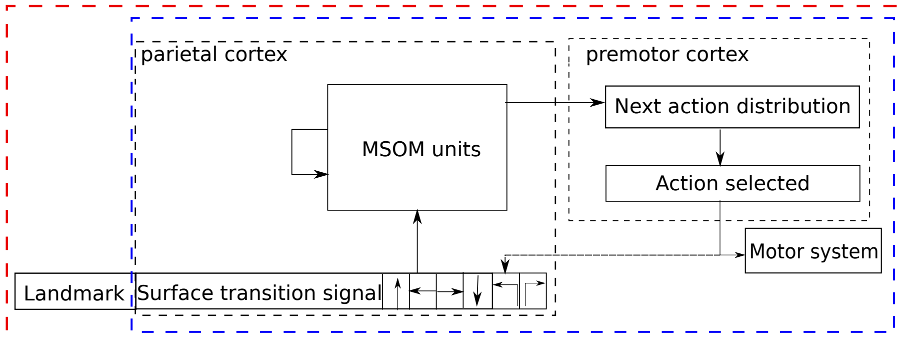


Fig. 4. Architectures of TO-MSOM model and L-TO-MSOM model for learning to represent 3D objects via translative movements (↑: move forward; ←: move left; →: move right and ↓: move back) and orientational movements (↰: rotate 90° counterclockwise and ↱: rotate 90° clockwise) together with the surface transition signal, where the blue frame shows the architecture of TO-MSOM model and the red frame illustrates the architecture of L-TO-MSOM model.

in practice and without such kind of movements, an agent could not go back to the start position with the same orientation, an improved and more practical model with orientational movements considered is of significance. Meanwhile, because objects generally do not have the differences among surfaces, to be more realistic, the surface information included in T-MSOM model should be left out. What's more, [13] presented the informal one test result about the effect of object asymmetry on the model's performance, while in the current paper, we present a statistics and more systematic study of the effect and extend the analysis to consider the effect of tactile landmarks on the object's surface.

3 Proposed Models

In this section, the proposed TO-MSOM and L-TO-MSOM model are presented.

3.1 TO-MSOM Model

By deleting the not generally-existed surface information in T-MSOM model and considering widely-performed orientational movements, TO-MSOM model is developed and its architecture is shown in a blue frame in Fig. 4. As illustrated in the figure, TO-MSOM model mainly consists of four parts: the input, MSOM units, next action distribution and action selected. The input to MSOM units is to simulate the circuit of object representations from the somatosensory cortex to the parietal cortex, and the next action distribution to action selected is to imitate the circuit from the premotor cortex to the motor system. The details of such four parts are illustrated as below.

Algorithm 2 TO-MSOM model

Input: Constrained action sequences of the object to be explored and represented
Output: A representation of the object explored

```
1: Randomly initialize the exploration starting position and orientation of the agent
2: while training steps are not finished do
3:     Input: input the executed constrained action
4:     MSOM units: activate MSOM units to be responsive to the current input
5:     Next action distribution: predict next possible actions allowed by the object
6:     Action selected: select the most possible action allowed by the object and perform
7: end while
8: return
```

Input The input of TO-MSOM model is composed of the constrained action sequences, which is comprised of translative and orientational movements. Note that the bit of surface translation signal is to encode the difference between the movement of moving directly (that is, moving forward, left, right, back directly) and moving over the edge (that is, moving forward, left, right, back over the edge). The input part is to encode and simulate the obtained sensorimotor information from the peripheral sensors.

MSOM units The units in MSOM are driven to learn the frequently occurring action *sequences*, which are constrained by the object's topography. As pointed out above, starting from the same location and orientation, the navigation agent can lead to different locations and/or orientation by executing different action sequences. Therefore, with regard to one starting exploration location, each unit in MSOM comes to be responsive to one/many particular location(s) on the object via learning constrained action sequences. After training, given a particular MSOM activity pattern, the learning model could reconstruct or say predict the navigation agent's position owing to the learnt representation of such an object. Note that this MSOM units part aims to imitate neurons involved in the circuit of object representations fulfilled in the parietal cortex.

Next action distribution Regarding each input at one time instance, there is an activity pattern in the MSOM, which denotes one particular location on the object. Based on the learnt representation of such an object, the model attempts to predict the next action possibly available to be preformed. In this model, the MSOM activity pattern is the input to a network, which is implemented by a multiple layer perceptron (MLP), and the output of MLP is the probability distribution of all actions predicted to be possibly performed.

Action selected After obtaining the possible action's probability distribution, the next action to be executed is then selected, which is based on the Boltzmann selection. Note that the selection procedure can be regulated by setting the selection decision policy involved in the Boltzmann selection. Regarding this model, if the navigation agent fails to execute one action to reach a new location for perceiving object's information, the probability of such an action is set to be zero, which implies that it cannot be selected for the next action to be performed. Moreover, in the model, for expediting the learning process, the navigation agent is commanded to find the boundary of the object as quickly as possible and therefore, the probability of the moving forward action is increased by a positive reinforced bias number. This part is to simulate the circuit of action selection performed in the premotor cortex.

After the next action to be performed is selected, an encoded signal of such an action is transferred to the motor system to perform. The performed action then leads the navigation agent to a different location, which gives rise to an update of the sensory information about the object and contributes to represent such an object. The flow diagram of the model is shown in Algorithm 2.

3.2 L-TO-MSOM Model

Landmarks in an environment are reported to have an effect on a navigation agent for exploring the environment, such as leading to remapping of the same environment and speeding up finishing a navigation task [8]. To investigate the effect of tactile landmarks on object representations, L-TO-MSOM model is developed with its architecture illustrated in the red frame in Fig. 4. Differing from TO-MSOM model, L-TO-MSOM model is not only activated by the constrained action sequences but also the landmarks on the object to be explored. The landmark in L-TO-MSOM model mainly denotes tactile landmarks, such as the temperature and texture differences among locations on the object. Since the implementation flow diagram of L-TO-MSOM model is similar to that of TO-MSOM model, it is omitted in the paper.

4 Simulation Results and Comparisons

To evaluate the effectivenesses of the proposed models and investigate effects of object asymmetries and landmarks on representing objects, two typical 3D objects–a $2 \times 2 \times 2$ cube and a $3 \times 2 \times 1$ cuboid–are assigned to be represented.

4.1 Effects of Object Asymmetries

To validate the effectiveness of TO-MSOM model for representing 3D objects as well as investigate the effect of encoding approach for the translative movements over the edge on the model, three kinds of TO-MSOM models, named TO-MSOM-1, TO-MSOM-2 and TO-MSOM-3, are assigned to explore the cube and cuboid with a random initial exploration position. Specifically, TO-MSOM-1 is the model using one bit of surface transition signal together with the four directly translative movement bits to denote the four translative over the edge movements (i.e., moving forward, left, right and back over the edge); TO-MSOM-2 is the model by utilizing four independent bits to denote such four translative over the

Table 1. Probability distribution of $P_{\max}$, represented as $P_{\max} \sim \mathcal{N}(\mu, \sigma^2)$ with μ and σ denoting the mean and standard deviation respectively, for TO-MSOM-1, TO-MSOM-2 and TO-MSOM-3 models when representing a cube and cuboid

Model	TO-MSOM-1	TO-MSOM-2	TO-MSOM-3
#Cube	$(6.59\%, 0.90\%^2)$	$(6.47\%, 0.66\%^2)$	$(6.63\%, 0.72\%^2)$
#Cuboid	$(11.19\%, 1.10\%^2)$	$(11.72\%, 1.20\%^2)$	$(11.26\%, 1.58\%^2)$

Table 2. 95% confidence interval of $P_{\max}$ for TO-MSOM-1, TO-MSOM-2 and TO-MSOM-3 models when representing a cube and cuboid

Model	TO-MSOM-1	TO-MSOM-2	TO-MSOM-3
#Cube	$[4.75\%, 8.43\%]$	$[5.12\%, 7.82\%]$	$[5.16\%, 8.10\%]$
#Cuboid	$[8.94\%, 13.44\%]$	$[9.27\%, 14.17\%]$	$[8.03\%, 14.49\%]$

edge movements, and TO-MSOM-3 is the model by using four bits of surface translation signal and four directly translative movement bits. Note that each model is to explore such two three objects for 30 sampled random tests/paths. Each test has 20 epochs and each epoch contains 100 exploration steps. The following results are based on statistics analysis of sampled 30 tests.

To evaluate effectivenesses of the proposed model, a criterion $P_{\max}$ is introduced, which is defined as

$$P_{\max} = \frac{T(\alpha = \beta)}{\phi}, \tag{7}$$

where ϕ denotes a fixed size of a sliding window, $T(\cdot)$ denotes how many times a given event happened in a given window in the sliding window series, α denotes the probability of the actual agent position in the reconstructed probability distribution and β denotes the maximal value in the reconstructed probability distribution. Furthermore, another criterion D_{geodesic} is also developed and defined as

$$D_{\text{geodesic}} = \sum_{i=1}^{m} \sum_{j=1}^{4} g(i, \delta) p_{ij}, \tag{8}$$

where $g(i, \delta)$ denotes the geodesic distance between position i and the actual agent position δ, m denotes the number of available exploration positions on the object; $j = 1, 2, 3, 4$ is used to respectively denote North South East West orientations; and p_{ij} denotes the probability of the agent being in location i and with the particular j orientation in the reconstruction distribution. Regarding details about such two criteria, please refer to [13, 14].

When exploring the cube and cuboid, the models' statistics probability distributions of $P_{\max} \sim \mathcal{N}(\mu, \sigma^2)$, with μ and σ denoting the mean and standard

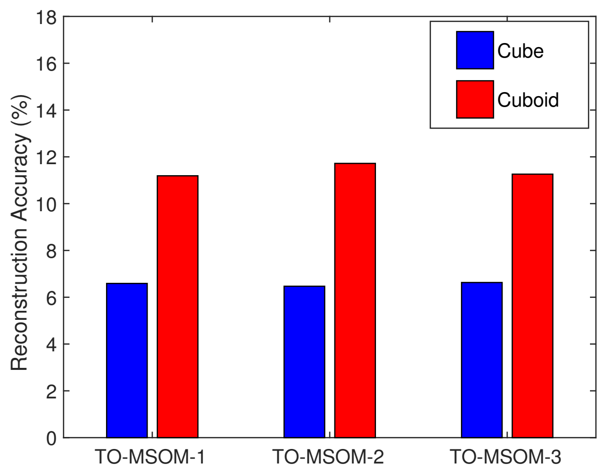
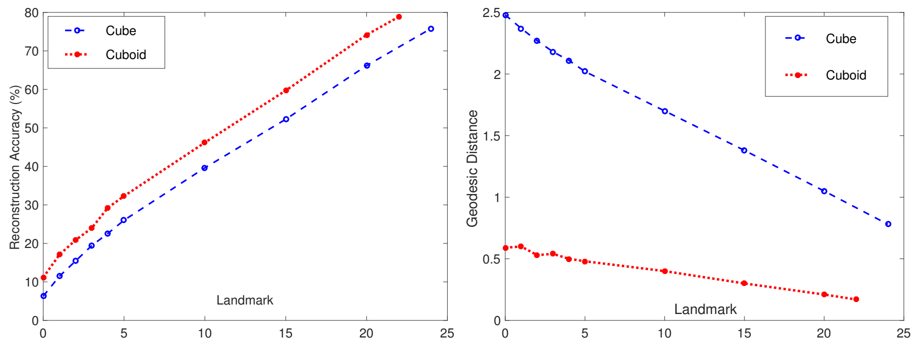


Fig. 5. Statistics means of $P_{\max}$ of TO-MSOM-1, TO-MSOM-2 and TO-MSOM-3 models when representing a cube and cuboid.

Table 3. Probability distribution of D_{geodesic}, represented as $D_{\text{geodesic}} \sim \mathcal{N}(\mu, \sigma^2)$ with μ and σ denoting the mean and standard deviation respectively, for TO-MSOM-1, TO-MSOM-2 and TO-MSOM-3 models when representing a cube and cuboid

Model	TO-MSOM-1	TO-MSOM-2	TO-MSOM-3
#Cube	$(2.48, 1.45\%^2)$	$(2.48, 1.18\%^2)$	$(2.48, 1.39\%^2)$
#Cuboid	$(2.39, 1.73\%^2)$	$(2.39, 1.81\%^2)$	$(2.38, 2.02\%^2)$

deviation respectively, are illustrated in Table 1. From the table, we can see that 1) the models have advantages in representing the cuboid over the cube, which is owing to the additional asymmetry topography information of the cuboid; 2) the models are effective on representing 3D objects; and 3) the encoding technique for translating over the edge movements does not make a difference on the model's representing ability. Corresponding t-based 95% confidence intervals and statistics means of $P_{\max}$ are shown in Table 2 and Fig. 5 respectively, which suggests the positive effect of asymmetry topography on TO-MSOM model's representing ability. Corresponding statistics probability distributions of D_{geodesic} for TO-MSOM models are illustrated in Table 3, which further verifies the contributive effect of the asymmetry geometry. Note that the accuracy of TO-MSOM model could be improved by adding other perceptual information, such as tactile landmarks (discussed later) and surface information (presented in [13]) or by making the navigation agent articulated, which is more like mammals' hands and can perceive and detect information on different surfaces of the object in parallel.



(a) (b)

Fig. 6. Statistics means of P_{max} and D_{geodesic} for L-TO-MSOM model when representing a cube and cuboid.

4.2 Effects of Landmarks and Object Asymmetries

To investigate the effect of landmarks, L-TO-MSOM model is commanded to explore and represent the cube and cuboid. Similarly, each model with different numbers of landmarks explores such two 3D objects for 30 random tests/paths and the following results are based on the statistics analysis.

The statistics means of P_{max} and D_{geodesic} when L-TO-MSOM represents a cube and cuboid are illustrated in Fig. 6. As we can see from the figure, we can draw the conclusion that the simulation result validates 1) the effectiveness of the model for representing 3D objects; 2) the positive effect of landmarks on the model's learning representations ability; and 3) the superiority of a cuboid to a cube for representation due to the asymmetry topography of the cuboid.

5 Conclusion and Future Work

In the paper, TO-MSOM model activated by translative and orientational movements has been proposed to learn representations of 3D objects. To investigate the effect of landmarks as well as object asymmetries, L-TO-MSOM model activated by landmarks together with translative and orientational movements has been developed. Statistics simulative results of TO-MSOM model and L-TO-MSOM model for learning representations of two typical 3D objects– a $2 \times 2 \times 2$ cube and a $3 \times 2 \times 1$ cuboid–demonstrate that 1) the proposed models are effective on learning representations of 3D objects; 2) object asymmetries have a positive effect on representations; 3) landmarks also positively contribute to learn representations. Future work is to design a more realistic and practical model to learn representations of 3D objects with an articulated agent, which is to simulate human beings' hands and consists of multiple independently moving 'fingers' to compete and coordinate for achieving a task. Another interesting direction is about investigating the model's representation ability for a more complicated 3D object, such a cup with curved surfaces, and a situation, such as a cup on a desk.

6 Acknowledgment

The authors would like to thank Martin Takac for the earlier work on a SOM-based navigation model.

References

1. Behrmann, M., Tipper, S.P.: Object-based attentional mechanisms: Evidence from patients with unilateral neglect. In: Conscious and nonconscious information processing. pp. 351–375. Cambridge: The MIT Press (1994)
2. Chafee, M.V., Averbeck, B.B., Crowe, D.A.: Representing spatial relationships in posterior parietal cortex: single neurons code object-referenced position. Cerebral Cortex **17**(12), 2914–2932 (2007)
3. Dar, H., Knott, A.: Learning and representing the spatial properties of objects via tactile exploration. Technical Report OUCS-2018-01 (2018)
4. Grossberg, S., Pilly, P.K.: How entorhinal grid cells may learn multiple spatial scales from a dorsoventral gradient of cell response rates in a self-organizing map. PLoS computational biology **8**(10), e1002648 (2012)
5. Grossberg, S., Pilly, P.K.: Coordinated learning of grid cell and place cell spatial and temporal properties: multiple scales, attention and oscillations. Phil. Trans. R. Soc. B **369**(1635), 20120524 (2014)
6. Jeffery, K.J., Wilson, J.J., Casali, G., Hayman, R.M.: Neural encoding of large-scale three-dimensional space—properties and constraints. Frontiers in psychology **6**, 927 (2015)
7. Kohonen, T.: Self-organized formation of topologically correct feature maps. Biological cybernetics **43**(1), 59–69 (1982)
8. Lozano, Y.R., Page, H., Jacob, P.Y., Lomi, E., Street, J., Jeffery, K.: Retrosplenial and postsubicular head direction cells compared during visual landmark discrimination. Brain and Neuroscience Advances **1**, 2398212817721859 (2017)
9. McNaughton, B., Chen, L., Markus, E.: "dead reckoning," landmark learning, and the sense of direction: a neurophysiological and computational hypothesis. Journal of Cognitive Neuroscience **3**(2), 190–202 (1991)
10. McNaughton, B.L., Battaglia, F.P., Jensen, O., Moser, E.I., Moser, M.B.: Path integration and the neural basis of the 'cognitive map'. Nature Reviews Neuroscience **7**(8), 663–678 (2006)
11. Strickert, M., Hammer, B.: Merge som for temporal data. Neurocomputing **64**, 39–71 (2005)
12. Uchimura, M., Nakano, T., Morito, Y., Ando, H., Kitazawa, S.: Automatic representation of a visual stimulus relative to a background in the right precuneus. European Journal of Neuroscience **42**(1), 1651–1659 (2015)
13. Yan, X., Knott, A., Mills, S.: A neural network model for learning to represent 3d objects via tactile exploration. In: Proceedings of the 40th annual conference of the cognitive science society (2018)
14. Yan, X., Knott, A., Mills, S.: A neural network model for learning to represent 3d objects via tactile exploration: technical appendix. Technical Report OUCS-2018-05 (2018)