Building a hierarchy of cognitive maps from panoramic images

Aleš Štimec
University of Ljubljana,
Faculty of Computer and Information Science,
Tržaška 25, 1000 Ljubljana
e-mail: ales.stimec@fri.uni-lj.si

Abstract

This paper presents a computational model which implements formation of cognitive maps based on panoramic images captured during the exploration phase. The resulting map consists of "place cells" and topological relations between them. The formation of the cognitive map is based on the model introduced by Hafner. The use of panoramic images as inputs would result in high computational complexity of the simulation, therefore we propose to use the PCA (Principal Component Analysis) method to reduce the dimension of the input space. A physical force model is applied to extend the relatively sparse topological map with metric information. Both the computational model and the physical force model try to mimic functions performed in the mammalian brain.

1 Introduction

Robots, like people, also need to explore and navigate large-scale environments and they face some of the same obstacles (inaccurate sensor inputs, complex and dynamic environments) that humans do. However the major difference is that humans (and our ancestors) have been exploring our environment, avoiding obstacles and dangerous situations for millions of years (some more and some less successfully). In many cases navigation can be done with little or no conscious thought. Current mobile robots are still far from human-like ability to explore and navigate space. So what are humans doing right? Or rather what are robots doing wrong?

The ability to move toward a specified place within the environment is one of the most important competences for a mobile robot. Successful navigation within any environment requires the robot to determine the current position, the goal position and the required sequence of moves to move from one position to the other. Apart from the basic requirement of "staying operational" (obstacle avoidance and staying within operational limits) navigation requires the ability to build a map of the environment and to interpret it accordingly.

In 1948 Tolman introduced the idea that people create what he called *cognitive maps* [10] of their environment. The idea however, dates back to 1913 when Trowbridge [11] carried out investigations of what he called ”imaginary maps”. His primary interest was to investigate why were some people more easily confused when performing orientation tasks than others. Since then there has been an abundant amount of research in this area, especially from the psychological point of view, and many of these studies identify *hippocampus*[1] as the most likely area in the brain where a cognitive map is formed (see [6].

NOVELTY!!!

In this paper we attempt to give a detailed description of a computational model for building a cognitive map from a sequence of captured panoramic images. Images were captured using an omnidirectional camera mounted on the robot. To reduce the dimension of the input into the model we propose to use the PCA method.

STRUCTURE!!!

1.1 Related work

The model used in this paper was inspired by the discovery of *place cells* in the rat's hippocampus. O'Keefe [8] noted that certain cells in the hippocampus of a rat preferably fired when the rat was in a particular part of its environment regardless of its orientation. These cell are now called place cells and the portion of the environment where such a cell fires at a heightened rate is called a *place field*.

Many other models have been made to explain the formation of cognitive maps in the hippocampus. In the following lines we briefly introduce some of them.

One approach to model the functionality of hippocampus comes from Bousquet, Balakrishnan and Honavar [1] who modeled hippocampus as a Kalman filter. They argue that there is a metric spatial representation in the hippocampal formation which arises as a result of association between sensory inputs and dead-reckoning information generated by the animal. They use the Kalman filter for information fusion from erroneous sources.

Another interesting approach comes from Harter and Kozma [5] who use K-set model of aperiodic dynamics to explain the cognitive map formation in the hippocampus. They base their research on experiments done on trained rabits [9]. EEG recordings of rabbit brains show that chaotic dynamics is the normal state when the animal is attentive in the absence ob stimulus. But those patterns change when a familiar stimulus is presented and the animal displayed recognition of a previously stored memory. The new dynamic pattern was much more regular and ordered. The spatial pattern of this activity represented a well defined structure that was unique for each type of stimulus. Harter and Kozma applied this princliple to building a cognitive map. They tested their model with a Khepera robot in a test environment in which they placed 8 light sources that were used as salient environmental locations.

A quite different approach has been taken by Fuhs, Redish and Touretzky [3]. First they describe an algorithm that operates on real images taken from various viewing locations and returns “blob” descriptions: regions of roughly uniform intensity having rectangular or ovoid shape. Then they construct simulated place cells using radial basis functions tuned to blob parameters, and then train them by competitive learning to develop realistic place fields. The

[1] A region of the brain which, due to its shape, got its name from the Greek word for seahorse.

result is a model which takes real-world images as inputs and produces a distributed activity pattern over a set of place cells as output, from which current position can be estimated. In this respect this approach is the most similar to the model used in this paper.

2 The Cognitive Map Model

In this section we review the computational model used to create a cognitive map of the explored environment. This model is largely based on the model presented by Hafner [4].

In out model, the map of the environment cbonsists of J nodes $\mathbf{o} = (o_1, o_2, ..., o_J)$ representing place cells, which are fully interconnected by connection weights α_{jk} and angles ρ_{jk}, $(j, k \in \{1, 2, ..., J\})$. Each of these map nodes is also connected with the input layer $\mathbf{f} = (f_1, f_2, ..., f_I)$ (I denotes the size of the input vector) via weight vectors $\mathbf{r}_j$, $(j, k \in \{1, 2, ..., J\})$. A preprocessed input vector is fed into the input layer and triggers a "winner takes all" process in the map layer. In each time step t the activation a_k^t of each map node o_k is determined using the sum x_k^t of he feature similarity s_k^t, the connectedness c_k^t and the movement value l_k^t

$$x_k^t = s_k^t + c_k^t + l_k^t - \Theta \ , \tag{1}$$

of in a sigmoid function $a_k^t = g\left(x_k^t\right) = \left(1 + e^{-x_k^t}\right)$. A fixed threshold Θ is subtracted to shift the input of the sigmoid function toward the negative values to achieve a greater variety of possible activation values. The map node o_j with the highest activation value is the winner node. Its weight vector, as well as its connectedness and movement value to the previous winner node, are updated in each discrete time step.

The feature similarity s_k^t is a measure of resemblance between the received input and the an already perceived view in this map node. It is calculated as a product between the input vector $\mathbf{f}$ and the weight vector $\mathbf{r}_k$ connecting $\mathbf{f}$ to the node o_j in the map layer, $s_k^t = \sum_{i=1}^{I}\left(f_i^t r_{ik}^t\right)$. Since both $\mathbf{f}$ and $\mathbf{r}_k$ are normalized, it reaches maximum when $\mathbf{f} = \mathbf{r_k}$. The weight vector $\mathbf{r}_k$ of the winner node is updated according to the Kohonen learning rule

$$\mathbf{r}_k^t = \mathbf{r}_k^{t-1} + \delta\left(\mathbf{f}^t - \mathbf{r}_k^{t-1}\right) \ , \tag{2}$$

where δ is a learning constant. The weight vector is then normalized. This rule gradually aligns the weight vectors $\mathbf{r}_k$ of the winning node o_k with the direction of the normalized input $\mathbf{f}$.

By introducing connectedness c_k^t we take into account in which place field the model assumes it was in the previous step as well as the (un)certainty of connections between place fields (represented by map nodes). To calculate the connectedness value c_k^t of each map node o_k to other map nodes o_j we weigh the connection weight α_{jk} with the last activation a_j^{t-1} of the map nodes

$$c_k^t = \sum_{j=1}^{J} a_j^{t-1}\alpha_{jk}^t \ . \tag{3}$$

The connection value between two consecutive winner map nodes o_k and o_k increases according to the following rule:

$$\alpha_{jk}^t = \alpha_{jk}^{t-1} + \gamma\left(\alpha_{max} - \alpha_{jk}^{t-1}\right) \ , \tag{4}$$

where γ is a learning constant and α_{max} is a constant denoting maximum connection weight.

In addition to the preprocessed image which serves as input into this model we also capture the direction in which the robot was moving in the time interval $(t-1, t]$ using an electronic compass mounted on the robot. Since we cannot estimate the length of the traveled path in this time interval we assume the distances between the consecutive places to be the same.

The movement value m_{jk} between nodes o_j and o_k is calculated as

$$m_{jk} = cos\left(|\rho_{jk} - \rho_{inp}|\right) \quad , \tag{5}$$

where ρ_{jk}^t is the stored movement angle and ρ_{inp}^t is the input angle. The value of $|\rho_{jk} - \rho_{inp}|$ lies between 0 and 2π therefore m_{jk} lies somewhere between -1 and 1.

In activation function, the movement value m_{jk} between nodes o_j and o_k is weighed with the last activation a_j^{t-1} of the node o_j:

$$l_k^t = \sum_{j=1}^{J} a_j^{t-1} m_{jk}^t \quad . \tag{6}$$

The stored angle ρ_{jk}^t is updated each time when o_k is the winner node after o_j by bisecting the stored angle ρ_{jk}^{t-1} and the input angle ρ_{inp}^t.

2.1 Adding metric information

Studies show (see [8]) that the cognitive map formed in the hippocampus represents space in a metric form. However the layout of the cognitive map this model builds is not in correspondence with the metric space, since spatial realations between nodes do not reflect in the overall structure of the map. Therefore we present a physical force model that extends the resulting topological map with metric information.

To get a better idea of the map, it is useful to arrange it as a graph $G = (V, E)$, where V is the set of map nodes and E is the set of edges between them. Since we do not know the distance between connected map nodes, the only information that can be used for this purpose is the angle information ρ_{jk} stored in the cognitive map. Let us now assume that the map nodes are repulsive charges and that edges between nodes are springs. Let $\nu\left(v, w\right) \in \mathbb{R}^2$ be the distance vector between nodes v and w and $d\left(v, w\right) = \|\nu\left(v, w\right)\|$ its Euclidean length. Between each pair of nodes we assume a force

$$F\left(v, w\right) = F_{rep}\left(v, w\right) + F_{attr}\left(v, w\right) \quad , \tag{7}$$

the sum of a repulsive caused by charges at each node and an attractive force caused by springs between nodes. If two nodes are not connected then $F_{attr}\left(v, w\right) = 0$.

Now let k be a constant denoting the desired distance between map nodes. The attractive force between nodes v and w is then calculated as

$$F_{attr}\left(v, w\right) = \frac{d^2\left(v, w\right)\nu\left(v, w\right)}{k} \quad , \tag{8}$$

and the repulsive force as

$$F_{rep}(v,w) = \frac{-k^2 \nu(v,w)}{d(v,w)} \quad . \tag{9}$$

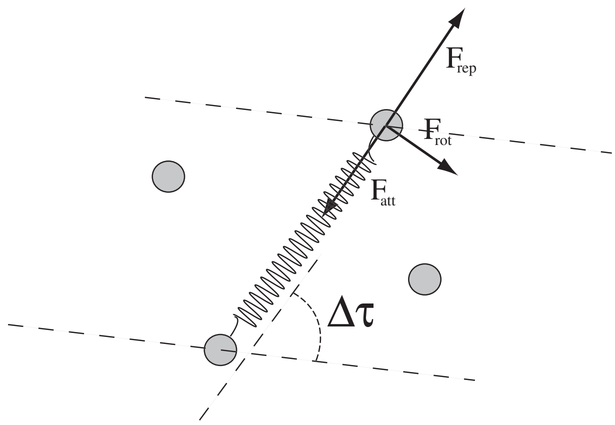
The task now is to find such a configuration of nodes in which the graph is balanced. That means that the sum of forces between all nodes

$$\sum_{v,w \in V, v \neq w} \left(F_{rep}(v,w) + F_{attr}(v,w) \right) \quad , \tag{10}$$

is minimal and therefore the potential energy of the system is minimal. An algorithm for a fast and stable solution for this problem comes from Fruchteman and Reingold [2]. Their algorithm however does not solve the question of edge orientation therefor we introduce rotational forces $F_{rot}(v,w)$ into the algorithm. These forces take into account the angle information between connected map nodes

$$F_{rot}(v,w) = \nu_{\perp}(v,w)\, d(v,w)\, \Delta\tau(v,w) \quad , \tag{11}$$

where $\Delta\tau(v,w)$ denotes the difference between the current and stored edge orientation and $\nu_{\perp}(v,w)$ denotes the unit vector normal to $\nu(v,w)$ whose cross product with the preferred edge direction vector does not contain any negative components. Since $\Delta\tau(v,w) = \pi - \Delta\tau(w,v)$ it follows that $F_{rot}(v,w) = -F_{rot}(w,v)$. All forces affecting a node in the map layer are shown in the picture 2.1.



To find a stable solution to this problem we used simulated annealing [7]. Forces on each of the nodes are calculated in each step and nodes are moved in the direction of the sum of the forces affecting each node. The distance by which each node is moved is limited by the temperature of that node. The overall temperature of the system is gradually reduced until a stable solution emerges.

2.2 Adapting place field sizes

At this point we have a cognitive map extended with metric information. Using this information we can deduce which part of environment are represented by place field of a certain map node. A question we will address now is the question of place field sizes.

Overly large place fields lack spatial information which could be used for navigation; on the other hand, overly small place fields require too many map nodes to represent the entire area of the environment. Therefore it is important to control the learning parameters of the aforementioned model to normalize the sizes of place fields.

3 Experimental results

In some experiments we have tried to show how the model builds a cognitive map of the environment and how the model modifies the created map as additional information about the environment becomes available (e.g. robot revisits a part of the environment which it has already mapped). For this purpose we captured a few image sequences using an omnidirectional camera mounted on an ATRV-Mini robot. Image sequences were captured in an outdoor environment as the robot was led around a triangular obstacle (a flowerbed in front of the faculty building). The robot was led around the obstacle six times and each time along a slightly different path (1368 images were captured during this phase). Therefore it is expected that the ”quality” of the created the cognitive map will increase as the additional information about the environment is added.

4 Conclusions

Under construction...

Acknowledgments

This research has been supported in part by the following funds: Research program Computer Vision P2-0214 (RS), the EU FP6-004250-IP project CoSy, the ECVision, EU (IST-2002-35454) research network, CONEX project, and SI-A project.

References

[1] O. Bousquet, K. Balakrishnan, and V. Honavar. Is the hippocampus a kalman filter? In *Proceedings of the Pacific Symposium on Biocomputing*, pages 655–666, 1998.

[2] T. M. J. Fruchteman and E. M. Reingold. *Graph-drawing by force directed placement, Software - Practise and Experience*, pages 21:1129–1164. 1991.

[3] M. C. Fuhs, A. D. Redish, and D. S. Touretzky. A visually driven hippocampal place cell model. In J. Bower, editor, *Computational Neuroscience: Trends in Research*, pages 379–384. Plenum Publishing, New York, 1998.

[4] V. V. Hafner. *Cognitive Maps for Navigation in Open Environments*. Master’s thesis, University of Sussex, UK, September 1999.

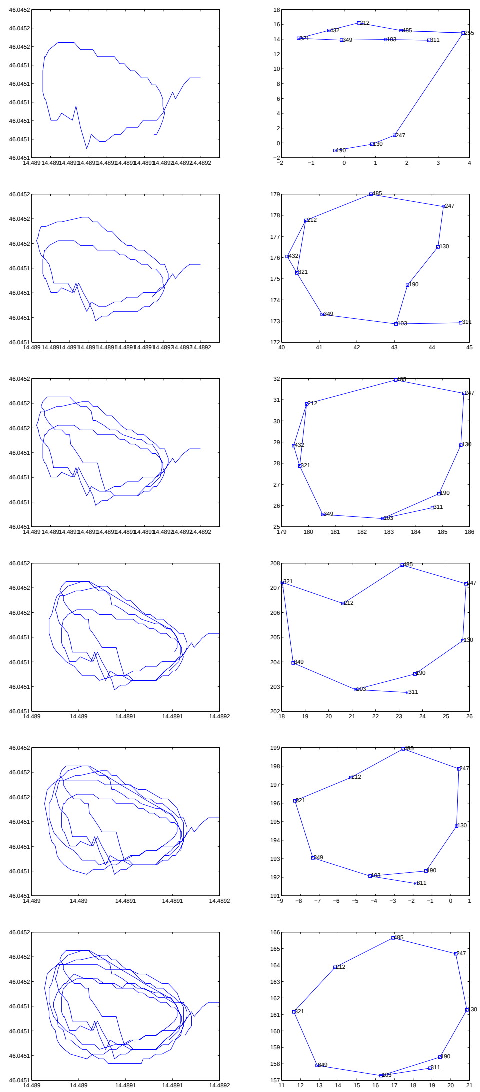


Figure 1: Images on the left show the path.

[5] D. Harter and R. Kozma. Navigation and cognitive map formation using aperiodic neurodynamics. In *Proceedings of the 8th International Conference on the Simulation of Adaptive Behaviour*, 2004.

[6] E. A. Maguire, D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. J. Frackowiak, and C. D. Frith. Navigation-related structural changes in the hippocampi of taxi drivers. In *Proceedings of the National Academy of Sciences of the United States of America*, volume 97, pages 4398–4403, April 2000.

[7] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys*, 21(6):1087–1092, 1953.

[8] J. O'Keefe and Lynn Nadel. *The hippocampus as a cognitive map*. Oxford University Press, 1978.

[9] C. A. Skarda and W. J. Freeman. How brains make chaos in order to make sense of the world. 10:161–195, 1987.

[10] E. C. Tolman. Cognitive maps in rats and man. *The Psychological Review*, 55:189–208, 1948.

[11] C. C. Trowbridge. Fundamental methods of orientation and imaginary maps. *Science*, 38:888–897, 1913.