Autonomous Robot Mapping by Landmark Association

Zati Azizul (zati@um.edu.my)
Department of Artificial Intelligence, Faculty of Computer Science & IT
University of Malaya, 50603 Kuala Lumpur, Malaysia

Wai Yeap (wai.yeap@aut.ac.nz)
Centre for Artificial Intelligence Research, Auckland University of Technology
Auckland 1010, New Zealand

Abstract

This paper shows how an indoor mobile robot equipped with a laser sensor and an odometer computes its global map by associating landmarks found in the environment. The approach developed is based on the observation that humans and animals detects where they are in the surrounding by comparing their spatial relation to some known or recognized objects in the environments, i.e. landmarks. In this case, landmarks are defined as 2D surfaces detected in the robot's surroundings. They are recognised if they are detected in two successive views. From a cognitive standpoint, this work is inspired by two assumptions about the world; (a) the world is relatively stable and (2) there is a significant overlap of spatial information between successive views. In the implementation, the global map is first initialised with the robot's first view, and then updated each time landmarks are found at every two successive views. The difference here is, where most robot mapping work integrates everything they see in their update, this work takes advantage of updating only the landmarks before adding the nearby objects associated with them. By association, the map is built without error corrections and the final map produced is not metrically precise.

Keywords: inexact map; landmark association; autonomous robot

Introduction

To date, many methods have been proposed in the framework of autonomous robot navigation to construct maps. From precise geometric maps based on raw data or lines to purely topological maps using symbolic descriptions; each has its own advantages and drawbacks. From reading, cognitive scientists and roboticists have different opinions on the mapping issues (Yeap & Jefferies, 1999; Jefferies & Yeap, 2001).

On the one hand, roboticists highlighted their effort working on the mapping problem by producing metrically precise maps of the environment, else their robots would get lost while navigating or exploring. Works such as Chatila (1982), Iyengar and Elfes (1991), Kuipers (2000), Durrant-Whyte and Bailey (2006) and Thrun (2008) led the ways of using powerful sensory tools (e.g. laser and vision) for robot mapping. However, their approach must deal with the main-product; errors accumulated over time by the sensors, which is usually corrected through the use of successful probabilistic methods such as the Monte Carlo Localization (Roefel & Juengel) and the various Kalman-based filters (Caballero et al., 2008; Roumeliotis & Bekey, 2000; Nguyen et al., 2012). The requirement for precise metrical maps calls for advanced error-correction techniques which are often costly to computational complexity.

On the other hand, cognitive scientists or behavioural scientists (psychologists and geographers) took the mapping approach from totally the opposite direction; analysing humans' and animals' behaviour traversing in new environments, investigating what is being remembered most during such visits, and identifying how an individual organized conceptual knowledge gained about the environment. Included also in their discussions were landmarks which play significant role in reasoning about the environment. They also paid close observations on the use of higher-level cognitive capabilities such as the ability to identify short cuts and the ability to identify oneself in complex environment particularly when looping occurs. Such studies can be seen in these works; Gallistel and Cramer (1996), Wang and Spelke (2000), Biegler (2000), and Cheng (1986). These extensive experimental works show that robots do not need to build a metrically precise global map to navigate in the environment. Moreover, they show that inconsistent and unclear sensor data are still usable to perform path planning and achieve loop closing successfully.

It has been argued that since human live in a geometrical world, humans should be locating objects in the environment by means of reference to the geometrical features. Plenty of works have adopted this notion of frames of reference as a means to represent the location of entities in space (Wang & Spelke, 2002; Mou & McNamara, 2002; Mou et al., 2004). These researchers believed that different frame of reference is used to for different navigational activities. For instance, navigating closely spaced trees requires accurate self-to-object (egocentric) judgement else one could bump into the obstacles (Anderson et al., 1997), but planning a distant goal and maintaining a sense of orientation in large environment requires one to judge how objects are allocentrically related to one another (Loomis & Beall, 1998). Figure 1 illustrates how the two reference frames configure. Figure 1(a) and 1(b) denote the egocentric frame of reference where locations of objects in two successive views are encoded in relation to own body (e.g. left-right, front-back, or up-down) respectively. Figure 1(c)

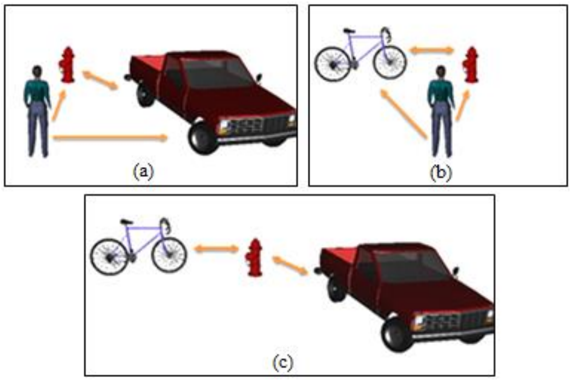

Figure 1: (a-b) The egocentric (self-to-object) spatial representation in two successive views and (c) the allocentric (object-to-object) spatial representation.

shows the allocentric frame of reference where locations of objects are encoded relative to other objects surrounding the person. The work in this paper pays attention to such approach. In particular, we are interested to grow the robot's global map by updating only the landmarks (i.e. common objects found between the robot's successive views) and then use these landmarks to associate new surfaces into the global map. The final map produced will be imprecise as a result of landmarks' association instead of views integration. The main advantage here is the mapping algorithm is relieved from complex probabilities calculation since the approach does not have to deal with the correction of accumulated sensor noise errors. The experimental setup, mapping algorithm and discussion on the final global map produced is presented.

Experimental Setup

The Robot

The robot used is a Pioneer 3DX mobile robot from MobileRobots Inc with measurement width of about 0.4m. It senses the environment using a laser source; a set of SICK LMS 200 laser rangefinder which has been mounted on its base. The sensor emits laser pulses horizontally at about 45cm from the ground with scanning range of approximate 30-32m covering 180 degrees field of view. With each laser pulse separated at half a degree from the mechanical sender, the sensor provides dense and accurate range data when used indoor. However, with two wheels for driving forward and backward and a non-driving wheel for rotation, the robot is highly vulnerable to drift errors particularly in areas where the flooring changes (e.g. tiles to carpet, carpet to cement, etc.) or when they are bumpy.

The Environment

Figure 2 shows the path (about 30x30m, shaded in yellow) traversed by robot in the experiment. The robot begins its journey from a random selected point in the office-like environment. The robot is allowed to wander on its own until commanded to stop. Since the laser data gathered is

Figure 2: The environment used for testing. Arrows showing the robot's clockwise path around the environment.

about a human's knee height, it is unavoidable for the robot to 'see' various objects scattered in the environment such as walls, table legs, chairs, boxes, cupboards, bins, space partitions, doors, pots, etc. These objects are left as it is; they are not cleared from robot's potential pathway. The only change done to the environment is the covering of glass-based walls and sliding doors with cut-out cardboards to prevent laser pulses from passing through them.

Autonomous Exploration

For exploring autonomously, the robot must decide where to go next and how to get there. In this work, we argue the robot should pick a random gap in space closest to the robot. A gap is defined as an empty space large enough for the robot to cross (i.e. > 0.6m) between two adjacent surfaces in view. Our robot calculates such a gap by finding a *minimal bounded space*; a space that contains no gap that can be covered by another gap in view. Yeap and Jefferies (1999) introduced the notion of *covering* by a gap as a space in which an individual must cross in order to reach another part of the environment that is currently in view. While they used the idea for computing the ASRs, we used it here to compute the minimal bounded space for the robot. The minimal bounded space limits the robot to 30 degrees (left or right) turn or a maximum of 3m forward drive at each interval. The limited movement ensures some parts of the view always overlap over two successive views. Algorithm for autonomous exploration is presented below. Algorithm to compute the minimal bounded space will be discussed elsewhere.

a) Get a scan of the environment
b) Identify gaps in view
c) Compute the minimal bounded space
d) Select a gap as target
e) Move towards the gap and stop
f) Repeat

Data Acquisition

At each scan, the 2D range data obtained are processed using line segmentation algorithm to generate planar

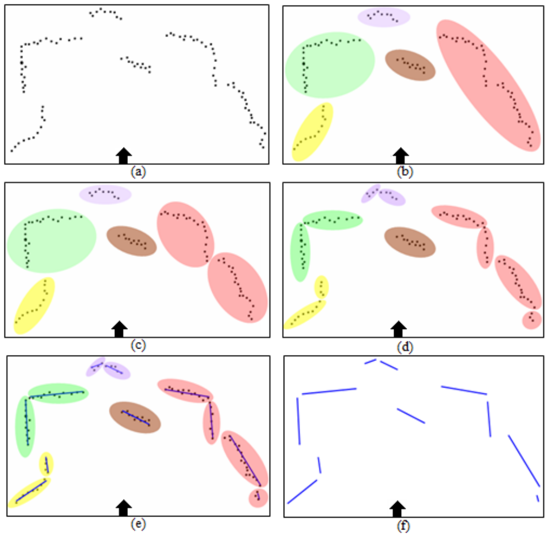

Figure 3: Processes involved in extracting surfaces from the environment.

surfaces so they would correspond to the geometrical properties scanned from the environment. There are many sophisticated algorithms such as the popular split-and-merges, line regressions and Hough transforms to extract line from points; all interested in providing an accurate polygonal model of the environment. However, since we do not need to build an exact map, precision is not of utmost important. A straightforward method for computing lines from laser points is thus implemented. First, the laser points are grouped into different clusters. This is done by going through the laser readings one after another in a clockwise manner and calculating the Euclidean distance between them. If the distance between them exceeds a set of threshold (currently set at 1.2m), a new cluster is formed. Second, for each cluster, the exact shapes of the lines in it are recursively computed using the average gradient descent between neighbouring points. Points on the same slope are grouped as a line representing a surface (see Figure 3). Note that for simplicity, small surfaces (defined as < 500mm) in view are simply ignored.

Computing the Global Map

Map Initialisation and Surface Tagging

A robot's global map is traditionally a structure built from integrating robot's successive views based on correcting the cumulative errors collected as the robot explores its environment. Here, we will show how we use landmark association to compute the map. Same as in traditional approach, it begins with initializing the map with the robot's first view. The processes from here on are a little different. First, we remove tiny surfaces (defined as surfaces smaller than 50cm) when we generate each view so anything larger than that are used. We made the assumption that only the larger ones are regarded with importance since they have the highest change to be the walls or part of the walls or some

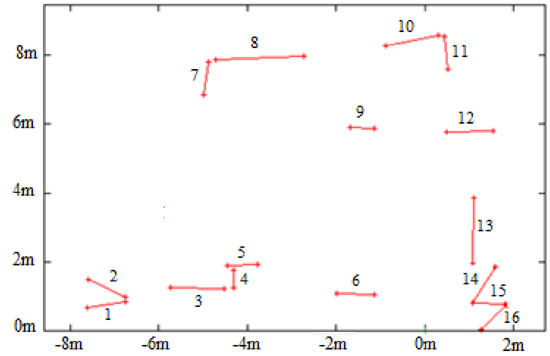

Figure 4: The global map initialized with surfaces from the first view. ID distinct one surface to the other.

major obstacles to avoid during exploration. Tiny surfaces computed may not be as useful to the robot and are dismissed as junks in the implementation. Then, the surfaces from the robot view are registered to (1) a frame of reference which acts like a buffer or a short term memory to track common surfaces or landmarks between every two successive views and (2) the global map. Each time a surface enters the global map, it will be tagged with an ID or a numbering marker. The increment of the ID numbers is proportionate to the increment of the number of surfaces entering the global map. Similarly these IDs are duplicated onto its counterpart in the frame of reference. Note that for initialisation the surfaces from the robot's first view are registered into the global map without any coordinate change. Figure 4 shows the global map initialised with surfaces from the robot's first view.

Landmark Identification

At each step (after robot move), the frame of reference will contain two views; the existing one from the previous step (with the surfaces tagged), and, a copy of the current view (with the surfaces untagged). At this point, both views are in their own coordinate systems. In order to compare two successive views for the robot, the mapping algorithm must describe surfaces in both views under the same coordinate system in the frame of reference. To do this, we transformed the previous view onto the current's coordinate by rotating it using the turn angle parameter then translating it using the move distance recorded. The following is the standard coordinate transformation formula used in the implementation:

$$x' = x\cos\theta - y\sin\theta + \delta x$$
$$y' = x\sin\theta + y\cos\theta + \delta y$$

Where

x' is the transformed x-coordinates
y' is the transformed y-coordinates
θ is the robot's turn angle
δx is the translation in x-direction, and
δy is the translation in y-direction

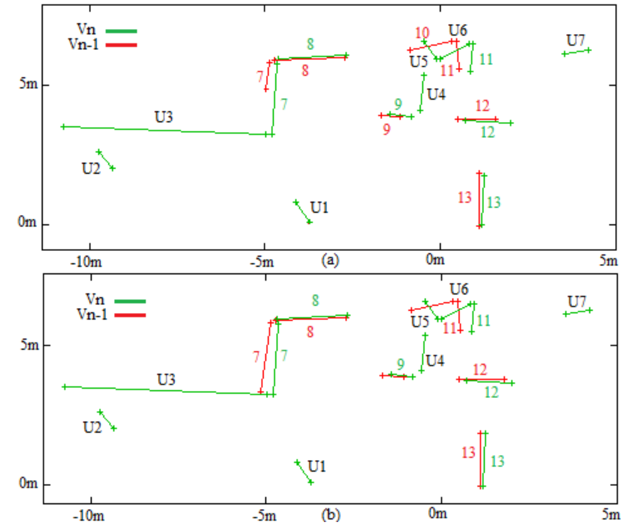

Figure 5: Comparison of overlapping surfaces over two successive views. (a) Matching surfaces inherit similar ID, and (b) all matching surfaces are normalised.

As mentioned, the main-product of using views integration is the measurement errors denoted by δx and δy which causes major distortions in the map computed if they are accumulated over time. However in this work, the errors are over only two successive views which make them trivial to the computation. Figure 5 depicts the comparison between surfaces in two successive views after the robot drives 2m forward. Vn denotes the robot's current view (in green) and Vn-1 the robot's previous view (in red). Note that only matching surfaces from the transformed Vn-1 are kept for comparison with Vn therefore surfaces 1-6 and 14-16 are deleted from the frame of reference. The principles applied to determine a match is to calculate the orientation between two surfaces that are close together. Two close surfaces are considered to be of the same orientation if their orientation does not differ by more than 10 degrees. This is a liable threshold due to the turning or forward driving at each interval is limited by the robot's minimal bounded space, consequently deriving some odometry drift, however not too bad drift that the overlapping bits are too disoriented or too far apart over two successive views. In the case the matching algorithm produces more than one candidate as matching surfaces, the surface that has the most similar orientation would be chosen as the matched surface or the landmarks. Surfaces from the current view which do not match any of the surfaces from the previous view are labelled as unknown (see U1-U7 in Figure 5(a)) and will be mapped as new surfaces in the map. To normalise the landmarks, the shorter end-points between both surfaces are lengthened to match the longer end-points so both surfaces are identical in length. Figure 5(b) shows the two views after all landmarks (7-9, 11-13) are normalised. Similarly, existing surfaces with similar IDs in the global map are normalised as well.

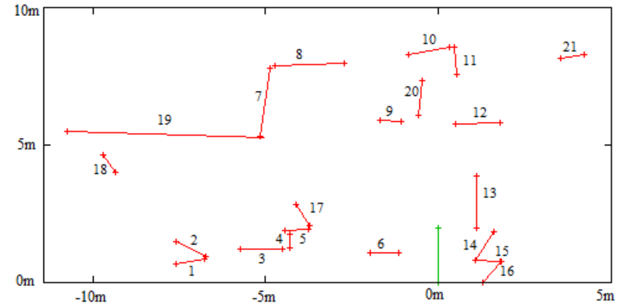

Figure 6: New surfaces (17-21) transferred into the map via their nearest landmarks. Green line depicts the robot current position in the map after the 2m forward drive.

Landmark Association for Update

Once the landmarks are identified and normalised inside the frame of reference, and the same landmarks are also normalised inside the global map, update is done by transferring new (unknown) surfaces from the frame of reference into the global map via the landmarks. When transferring a new (unknown) surface into the map, one uses its position with respect to its nearest landmark in the frame of reference. This is significant because if errors were introduced in the matching calculation, choosing the nearest landmark would suppress the errors to a minimum. For this reason, U1-U3 in Figure 5(b) is transferred into the global map by landmark 7, U4 by landmark 9 and finally U5-U7 by landmark 11. Note that not all surfaces transferred into the map are new to the map thus it is necessary to check if an incoming surface is already known in the map. To perform the check, the incoming surface is compared with existing surfaces in the map to see if they intersect one another. An intersection indicates a cluttered area in the map thus there is no need to transfer the incoming surface. If the incoming surface is positioned close to another surface in the map, the two could possibly be the same surface. In this case, the incoming surface inherits the ID already assigned to the surface inside the map. However these corresponding surfaces may not be of the same length so they are normalised since surfaces having the same ID must be of the same length. Any successful insertion of surfaces into the global map will be registered with an ID and this is done by increasing the last ID in the map by 1. The final step is to also update Vn in the frame of reference with the ID tags from the global map, before forgetting Vn-1 (deleting it from memory) so only Vn is brought forward for the successive comparison. Figure 6 shows the transfer result.

The Map Produced

This work is aimed to demonstrate that the mapping algorithm is robust, at least for mapping in a reasonably large office environment. It is also crucial to show that the final map produced is imprecise yet of sensible shape in comparison to the physical environment (see Figure 1). In the experiment, the robot is let to wander on its own where it computes its global map in real-time. Over 130m were

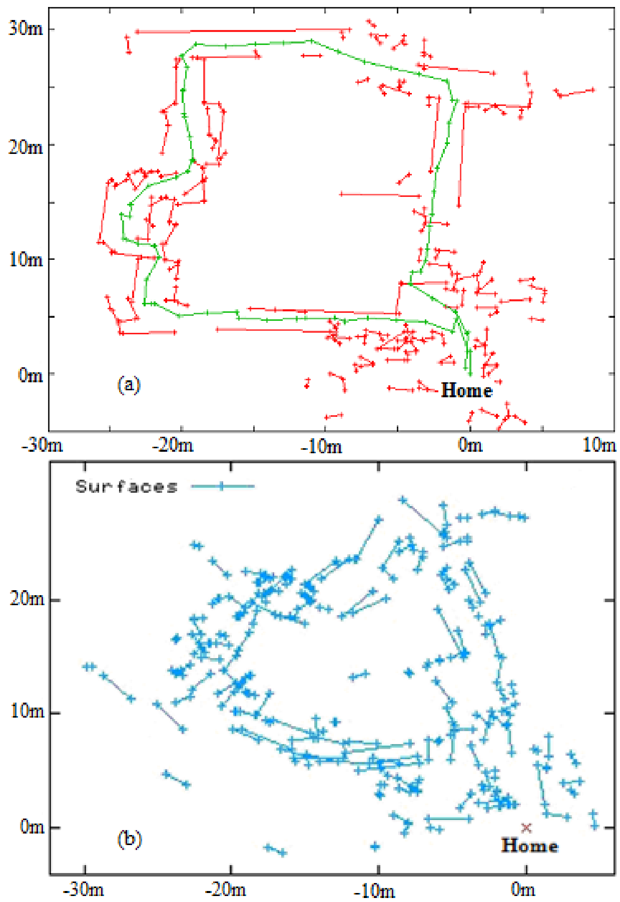

Figure 7: (a) Final map produced via our approach and (b) map produced via traditional approach without error correction

traversed and 103 robot views were collected and used throughout the exploration. Figure 7(a) depicts the final map produced using our approach after the robot loop the environment in a clockwise fashion. Since we argue that our landmark association approach does not require error corrections, we reproduced a map using views integration without one for simple comparison (see Figure 7(b)). Without error corrections, the same environment traversed by the robot would produce a heavily distorted map if the errors accumulated by views integration are not corrected.

Discussion and Future Directions

From a robotics perspective, the map shown in Figure 7(a) is considered imprecise in the sense it is not metrically accurate and has missing surfaces. However, when compared to the physical world (Figure 1), it can be seen that the overall shape of the environment experienced is captured and well maintained by said map. The approach therefore can be considered successful, at least on a laser mobile robot. The present implementation shows that one can utilize recognized objects i.e. landmarks between successive egocentric views to represent allocentrically other objects within one's surroundings. The key hypothesis in this approach is the premise that the world is generally stable enough; that the objects in the environment is there however one reorients and views them. Consequently, there is also significance overlap of information in our successive views, more if we consider taking smaller steps or limits our orientation while moving, letting us know what lies immediately behind us and what may appear in front as we continue our journey.

It can be argued that compared to views integration, our approach offers a simpler and less computationally expensive method for computing a laser robot's global map. This is mainly due the robot not having to deal with accumulated errors while integrating views. While there are other works, notably Steinhage and Schoner (1997) that constantly recalibrates from one error prone local reference frame to the next, and memorising different vantage point of views of the home base for homing, they are by principles still limited to errors due to the need to integrate multiple sources of information. In our case, recalibration is based on recognising some landmarks between two successive views and homing is performed by simply recognising some landmarks registered in the allocentric global map.

The implementation using landmark association also shows how a robot is able to produce an imprecise global map. This means the algorithm developed here may shed some light on how human cognitive mapping process work. Rough overall shape of the environment (imprecise and incomplete map) accords to two key features of the human and animal cognitive mapping process, namely; (a) human and animal do not remember everything they experienced in their journey, and (b) what they actually remember is an abstract representation of objects in relation to other objects in the environment.

As exciting as the current result may be, the approach developed here is not restricted to a mobile robot equipped with laser and odometry sensors. We believe it should also work well or even better with visual robots. This is due to the fact that vision allows a richer description of the environment, which in consequence improves landmark recognitions. For this reason, heading towards the utility of vision would be an important future research. It would also be interesting to extent the current work into incorporating local spaces concept and the notion of exits (Yeap & Jefferies, 1999) to reason about the global and the immediate spaces computed by the robot. Continue refinement of the algorithm and testing in larger environment would also ensure the approach is ready for practical robot applications. Finally, it would also be interesting to consider conducting some human studies by showing the results from the implementation and ask the human subjects to sketch their own map or answer some basic questions about the landmark locations captured by the robot.

Conclusion

A new approach to build a mobile robot's map of the environment is presented which shows how a global map is computed using landmark association and not views integration. The interesting finding from this work is how a frame of reference is utilised to compare and track landmark across two successive views of the robot. The approach is supported by numerous observations on how human and animal perceive the stable world particularly in how they use recognized objects (landmarks) to estimate and relate approximately the positions of other objects in their immediate surroundings. The implementation of the approach shows the map computed does not have to be metrically precise or complete for the robot to successfully close loops and maintains a good overall shape of the environment traversed.

Acknowledgments

This work is partially funded by the Bantuan Kecil Penyelidikan (BK058-2014) University of Malaya, Kuala Lumpur, Malaysia

References

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual review of neuroscience*, *20*(1), 303-330.

Azizul, A. (2013). *Robot mapping without a precise map* (Doctoral dissertation, Auckland University of Technology).

Biegler, R. (2000) Possible uses of path integration in animal navigation. Animal Learning and Behaviour, 28(3): 257-277.

Caballero, F., Merino, L., Gil, P., Maza, I. & Ollero, A. (2008). A probabilistic framework for entire WSN localization using a mobile robot. *Journal of Robotics and Autonomous Systems.*

Chatila, R. & Laumond, J. P. (1985). Position referencing and consistent world modeling for mobile robots. *In Proc. IEEE Int. Conf. Robot. Automat.*, pp. 138–143.

Cheng, K. (1986) A purely geometric module in the rat's spatial representation. Cognition, 23: 149-163.

Cussins, A. (1998). Subjectivity, objectivity and frames of reference in Evans's theory of thought. *Electronic Journal of Analytic Philosophy*, December.

Dongbo, L. I. U., Zhongxiao, H., Peng, X. I. A. O., & Changsong, L. I. U. (2013). Mobile Robot Monte Carlo Localization Based on Improved Local Invariant Feature Extraction Algorithm. *International Journal of Applied Mathematics and Statistics™*, *51*(21), 91-100.

Durrant-Whyte, H., & Bailey, T. (2006). Simultaneous localization and mapping: part I. *Robotics & Automation Magazine, IEEE*, *13*(2), 99-110.

Gallistel, C. R. and Cramer, A. E. (1996) Computations on metric maps in mammals: Getting oriented and choosing a multi-destination route. Journal of Experimental Biology 199:211-117.

Iyengar, S. S., & Elfes, A. (1991). Autonomous Mobile Robots: *Perception*. IEEE Computer Society Press.

Jefferies, M. E. & Yeap, W. K. (2001). The utility of global representations in a cognitive map. In *Conference on Spatial Information Theory*, pp 233-246, Morro Bay, USA.

Kuipers, B. J. (2000) The spatial semantic hierarchy. Artificial Intelligence, 19:191-233.

Loomis, J. M., & Beall, A. C. (1998). Visually controlled locomotion: Its dependence on optic flow, three-dimensional space perception, and cognition. *Ecological Psychology*, *10*(3-4), 271-285.

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 162.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 142.

Nguyen, T. T., Phung, M. D., Tran, T. H., & Tran, Q. V. (2012). Mobile robot localization using fuzzy neural network based extended Kalman filter. In *Control System, Computing and Engineering (ICCSCE), 2012 IEEE International Conference on* (pp. 416-421). IEEE.

Roefer, T. & Juengel, M. (2003). Vision-based fast and reactive Monte-Carlo Localization. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA-2003)*, pp. 856-861, Taipei, Taiwan.

Roumeliotis, S. I. & Bekey, G. A. (2000). Collective localization: A distributed Kalman filter approach to localization of groups of mobile robots. In *Proc. IEEE International Conference on Robotics and Automation*, pp. 2958-2965, San Francisco, California, April 22-28.

Steinhage, A., & Schöner, G. (1997). Self-calibration based on invariant view recognition: Dynamic approach to navigation. *Robotics and Autonomous Systems*, *20*(2), 133-156.

Thrun, S. (2008). Simultaneous Localization and Mapping. In *Robotics and Cognitive Approaches to Spatial Mapping* (pp. 13-41). Springer Berlin Heidelberg.

Wang, R. F. and Spelke, E. S. (2000) Updating egocentric representations in human navigation. Cognition, 77: 215-250.

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in cognitive sciences*, *6*(9), 376-382.

Yeap, W. K. & Jefferies, M. E. (1999). Computing a representation of the local environment. *Artificial Intelligence*, **107:**265-301.