A Novel Hybrid Map Representation for DenseSLAM in Unstructured Large Environments

Juan Nieto, Jose Guivant, Eduardo Nebot
Australian Centre for Field Robotics, The University of Sydney, Australia
{j.nieto,jguivant,nebot}@acfr.usyd.edu.au

Abstract

This paper presents a new hybrid metric map representation (HMM) that combines feature maps with other dense metric sensory information. Representing non-geometric properties of the environment is one of the outstanding challenges in robotic research. The representation presented here will allow the robot to perform DenseSLAM. DenseSLAM is the process of performing SLAM whilst obtaining a dense representation of the environment. Also a new concept named *Diode Correlation* is introduced to demonstrate the consistency of the proposed maps. Initial experimental results in outdoor environments are also presented to validate the algorithm proposed.

1 Introduction

Map building is one of the most challenging problems in autonomous systems. An accurate representation of the environment is essential for any robot to operate successfully in a fully autonomous manner. When a robot performs exploration without any absolute information it needs to simultaneously build a navigation map and localise itself using this map. This problem is usually referred as Simultaneous Localisation and Mapping (SLAM) [1] or Concurrent Mapping and Localisation (CML) [2]. Most real time implementations of SLAM are based on Kalman Filters [1]. Several approaches to solve the problem of navigation in large environments have appeared in the last few years [3–5]. Most of them, however, obtain a sparse representation of the environment, which eventually could not be enough to achieve tasks such as path planning.

Map representation is of paramount importance for navigation and guidance, particularly when working in very large environments. This paper presents a new Hybrid Metric Map (HMM) representation [6] that combines feature maps with any other metric information. This approach permits the localisation of the robot and at the same time produces a dense and consistent environment representation (DenseSLAM). In particular it includes a combination of features and occupancy grid approaches to achieve efficient global localisation with features and allows *detailed navigation* in local areas by making use of OG maps. In this paper *detailed navigation* refers to the capabilities of an autonomous system to perform tasks such as path planning.
The paper is organised as follows. Section 2 presents a brief review of various map representation techniques. Section 3 introduces the HMM representation with a demonstration of consistency with the introduction of the Diode Correlation concept in section 4. Finally experimental results are presented in section 5.

2 Autonomous Mapping

Autonomous mapping is the process of constructing a representation of the world from external sensory information. A **consistent and rich** representation of the environment in large unstructured scenarios is one of the outstanding and significant research challenges in building intelligent systems. Robots use a wide variety of sensors to obtain information about the environment, however the sensors are not perfect, and the information provided is usually perturbed by some type of noise. The modelling and consistency of this uncertainty is one of the key problems in mapping. Many autonomous applications can only rely on observations of the world taken relative to the pose of the robot. The error in the robot pose estimation will introduce another source of uncertainty. Probabilistic techniques have been widely used in mapping problems since they can incorporate and propagate sensor and vehicle model errors into the rest of the system.
The next subsections present a brief review of the most common mapping techniques used in localisation and SLAM: occupancy grids, feature based maps and topological maps.

2.1 Occupancy Grid Maps

The occupancy grid mapping technique (OG) represents the environment with a grid of fixed resolution. The OG is a multidimensional grid that maintains stochastic estimates of the occupancy state of each cell. The technique was introduced by Elfes [7] and has been widely used due to the simplicity of its implementation. Each cell stores the probability of being occupied or free. One of the main problems with OG maps is that they do not take into account the correlation between cells that exists due to the robot pose uncertainty. This technique assumes that the state variables of each cell are independent, an assumption that is not true specially when working in large areas where the robot pose can potentially accumulate significant errors. OG have been used mainly for localisation given a priori map, or for map building given the robot pose. More recently some approaches of SLAM using OG have started to appear [8]. However, as stated in [9] it is of fundamental importance to have a representation of the robot pose and sensor uncertainty and their correlation to achieve convergence of the SLAM process. This is not possible using only OG techniques. In general OG techniques provide very rich representations of the environment, but they cannot provide consistent global map estimates when working in large environments.

2.2 Feature Maps

Feature map techniques [10] represent the environment with parametric features such as points, lines, cylinders, corners, etc. A feature is a distinctive part of the environment that can be easily discriminated by some type of sensor and described with a set of parameters. This method requires a model for each type of feature considered in the environment. Although only the location of the features is used to represent and maintain the map, other information such as geometry, color, etc., can also be used in the identification stage or during the data association process. In particularly this representation has been widely used in SLAM with Kalman filters [3, 11]. With this approach the state vector containing the robot pose is augmented when new features are observed and validated. Since the features are observed with respect to the robot position the estimated robot pose and map will be correlated due to the robot pose uncertainty. At the limit, when many observations are incorporated, the map becomes fully correlated. This means that a theoretically perfect map can be achieved if one feature position is known with absolute accuracy. The feature technique will give a sparse representation of the environment, which in most cases will be of limited resolution and will not be able to provide information to perform *detailed navigation.*

2.3 Topological Maps

Topological approaches [12] represent the robot environment by using graphs. With metric approaches such as feature maps or occupancy grids, the objects' locations are defined in a cartesian coordinates frame. With topological maps the environment is represented by nodes and arcs that correspond to possible travel paths. Each node represents a *distinctive place* in the environment, and the edges represent the relative position between adjacent nodes. With pure topological maps it is only necessary to represent the connectivity between nodes and no information about the absolute position of the nodes is required. Topological maps are attractive due to the compactness of the representation. They are efficient representations for tasks like path planning and are appealing in indoor applications where clear distinctive places can be found between areas where navigation can be performed aided by very basic information such as wall following. One of the main problems using this representation in more complex environments is place recognition. If the robot travels between two places that look alike, the lack of metric information makes the discrimination of these two places very unreliable. In this case the logic of the topological map will be broken and the robot will not be able to evaluate its position. This is probably the main reason why most of the approaches use metric information over topological maps [4, 12]. Finally another limitation of this technique is that it does force the planner to follow specific trajectories to pass through or very near distinctive places.

3 Hybrid Metric Maps

In the past few years various studies presented implementations of maps combining different approaches [13]. These range from integrating topological with metric maps [4, 5, 14] to the use of multi-resolution OG approaches [8]. In [4] and [5], the authors introduce a hybrid approach that uses topological and feature maps to perform large-scale SLAM. These methods scale very well to large environments using feature maps only. In [15] a hybrid OG / landmark map is used, however these two maps are considered independent, as such this will give inconsistent results, particularly when working in large areas.

This section presents a new mapping framework called *Hybrid Metric Maps* (HMM). The HMM combines feature maps with any other metric representation. The novel issue that this representation introduces is the incorporation of the correlation between the landmarks and the metric map. The approach allows the robot to perform DenseSLAM. This will enable the robot to afford precise path control and more demanding tasks such as global path planning. Moreover the HMM representation presents significant advantages to solve the data association problem in a very robust manner using point data correlation techniques [16].

When working with feature based maps, a set of features can be used to partition the region covered by the map. One example of these partitions is shown in Fig. (1). In this case, we used triangular regions that will be referred as *local triangular regions* (LTR). Each LTR is defined by the position of three landmarks called vertex as shown in Fig. (2). Any point that belongs to a LTR can be characterised by a convex linear combination of the three vertex points associated with this sub-region. In Fig. (2) a LTR Ω_i is defined by the vertex points

$\{L_{i,1}, L_{i,2}, L_{i,3}\}$. A local coordinate frame is defined based on the three vertex points according to the next equation:

$$\begin{aligned} \overrightarrow{a_i} &= L_{i,2} - L_{i,1} = (a_x, a_y) \\ \overrightarrow{b_i} &= L_{i,3} - L_{i,1} = (b_x, b_y) \end{aligned} \tag{1}$$

Any point that belongs to Ω_i can be expressed as:

$$\begin{gathered} X = L_{i,1} + \alpha \cdot \overrightarrow{a_i} + \beta \cdot \overrightarrow{b_i} = \\ = (1 - \alpha - \beta) \cdot L_{i,1} + \alpha \cdot L_{i,2} + \beta \cdot L_{i,3} \\ \\ \alpha > 0, \ \ \beta > 0 \\ \alpha + \ \beta \ \leq 1 \\ \\ \forall \ \ X \ \ \backslash \ \ X \in \Omega_i \end{gathered} \tag{2}$$

Furthermore any function of the global position X can be locally defined as a function of the local representation of X.

$$z = f(X) = f(L_{i,1} + \alpha \cdot \overrightarrow{a_i} + \beta \cdot \overrightarrow{b_i}) = g(\alpha, \beta) \tag{3}$$

In some applications a function can be defined locally by an observer that has its position well defined with respect to the vertices of the related LTR. This means that the position uncertainty of the observer will be low since it is expressed with respect to a local frame. Then any information gathered from this location and associated with position can be accurately represented in the local frame. Due to the structure of the map, the vertex points and any interior point of the LTR are highly correlated. High uncertainty in the vertex points will not affect the quality of any property defined as a function of the observer position (local). This is true if the observer measures certain property z of points that are inside the LTR and it is well localised with respect to the vertex points of this LTR. The property z is expressed as a function $g(\alpha, \beta)$, where (α, β) are local coordinates. Any improvement in the estimation of the position of the vertex points will imply an improvement of z expressed in a global coordinate frame. If the position of the vertex points is exact (in respect to the global frame) then a linear coordinate transformation will provide the conversion from the local representation $g(\alpha, \beta)$ to the global frame version $f(X)$.
A typical application of this concept is when a robot is concurrently doing SLAM and measuring a property z. The property z does not necessarily have to be used for the robot localisation process.
Assume a vehicle simultaneously doing SLAM and measuring three properties: soil salinity, humidity and terrain occupancy. These properties can be locally represented in each LTR by using occupancy grid techniques. Fig. (3) illustrates one example of different internal properties in a given LTR. In some cases, due to the high correlation between the vertex points and the vehicle pose states, it will be possible to represent a particular property as a deterministic function of the local coordinates (α, β) inside the LTR. In addition one of the properties (e.g. terrain occupancy) could be used in the data association stage of the SLAM process or to do path planning.

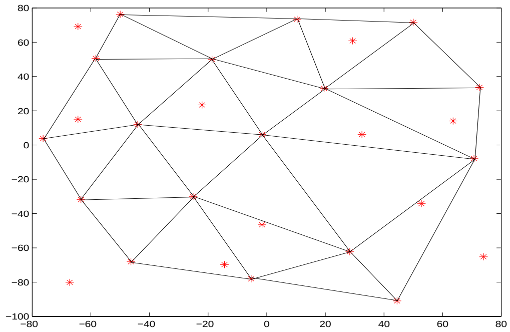

Figure 1: Landmarks map ('*') and a particular partition on triangular sub-regions (LTRs). As it can be seen, not all the landmarks are needed as vertex points in the LTRs definition.

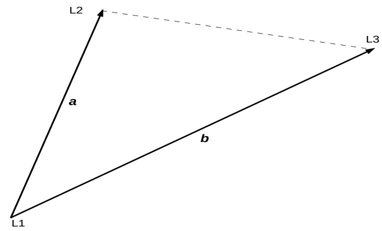

Figure 2: Detail of an individual LTR defined by three vertex points, $\{L_{i,1}, L_{i,2}, L_{i,3}\}$ and the direction vectors, $\{\overline{a_i}, \overline{b_i}\}$.

If a set of observed objects are geographically close from the vehicle viewpoint, then the error due to the vehicle pose uncertainty will be a common component of these estimated landmark positions. This is a typical case in SLAM where the vehicle accumulates uncertainty in its estimated position and incorporates observations that are used to synthesise a map. Due to this fact the estimates of landmarks that are geographically close will present similar uncertainties and high cross-correlations. Any update in a particular landmark will imply a similar update on any landmark sufficiently close to the first one.
The local representation defined by Eq. (2) takes advantages of this particular fact. It can be seen that

$$\begin{aligned} X \to L_{i,1} &\Rightarrow \alpha \to 0, \ \ \beta \to 0, \ \ (1 - \alpha - \beta) \to 1 \\ X \to L_{i,2} &\Rightarrow \alpha \to 1, \ \ \beta \to 0, \ \ (1 - \alpha - \beta) \to 0 \\ X \to L_{i,3} &\Rightarrow \alpha \to 0, \ \ \beta \to 1, \ \ (1 - \alpha - \beta) \to 0 \end{aligned} \tag{4}$$

This means that if a relative estimated point X, is close to $L_{i,1}$ and the estimation process generates changes in

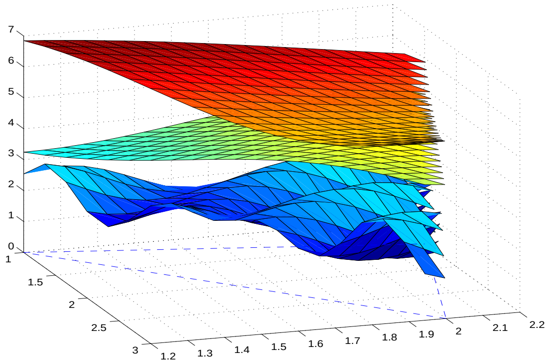

Figure 3: A set of properties can be defined as a function of the local coordinate variables. The triangular LTR's shape can be seen on the plane xy (dashed lines).

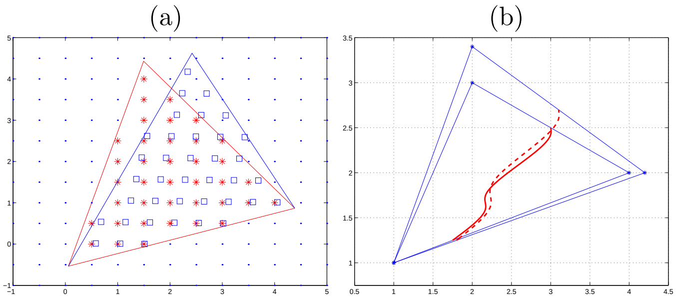

Figure 4: Effect of moving the base landmarks position. (a) Over a grid. '*' before the update and 'square' grid after the update (b) In a curve.

the base landmarks $L_{i,2}, L_{i,3}$ but no change is introduced in $L_{i,1}$ then a very small change will be made over the estimate of X. This can be seen by analysing the variation of an internal point X when there is a change in the LTR landmarks position.

$$\begin{aligned} \delta X &= (1-\alpha-\beta)\cdot\delta L_{i,1} + \alpha\cdot\delta L_{i,2} + \beta\cdot\delta L_{i,3} \\ &= \alpha\cdot\delta L_{i,2} + \beta\cdot\delta L_{i,3} \end{aligned} \tag{5}$$

From (5) it can be seen that $\delta X \to 0$ when $\delta L_{i,1} = 0$ and $\alpha = \beta = 0$. Fig. (4) shows the effects of changing the position of the base landmarks in a LTR. In (a) the concept is illustrated in a grid, and in (b) in a curve. It is evident that there are small changes in the points close to the static landmarks, and large changes in the points close to the landmarks that moved.

The HMM can also be applied to represent $3D$ maps. The $2D$ triangular LTRs can be extended to $3D$ tetrahedral LTRs including an additional vertex point. The vectors to define the local frame are now three:

$$\begin{aligned} \overrightarrow{a_i} &= L_{i,2} - L_{i,1} \\ \overrightarrow{b_i} &= L_{i,3} - L_{i,1} \\ \overrightarrow{c_i} &= L_{i,4} - L_{i,1} \end{aligned} \tag{6}$$

and any point that belongs to Ω_i can be expressed as:

$$\begin{gathered} X = L_{i,1} + \alpha\cdot\overrightarrow{a_i} + \beta\cdot\overrightarrow{b_i} + \gamma\cdot\overrightarrow{c_i} \\ \alpha > 0,\quad \beta > 0,\quad \gamma > 0 \\ \alpha + \ \beta + \gamma \le 1 \\ \forall\ \ X\ \setminus\ \ X \in \Omega_i \end{gathered} \tag{7}$$

As mentioned before, the occupancy grid technique requires accurate determination of the robot position. However, in real applications the robot pose is always known with some uncertainty. Previous hybrid maps approaches based on OG neglected this uncertainty and therefore they generate inconsistent results, in particular when applied to large environments. The approach presented here uses the features included in the stochastic map to define the triangular boundaries of a grid map. By using a feature based representation it is possible to estimate the vehicle position simultaneously with the map, however the map obtained will be a sparse representation of the environment. Augmenting the map with the OG representation provides additional information that will give a denser and consistent representation of the environment.

The HMM representation can combine feature based maps with OG. The approach uses the best of both worlds, the consistency of the feature maps, that will enable the robot to localise itself and the rich representation that can be get with OG maps.

3.1 SLAM using the HMM: *Diode Correlation*

This section will introduce the concept of *Diode Correlation* that will be used to demonstrate the consistency of HMM. The SLAM problem using the HMM is now introduced using Bayes formulation. Solving the SLAM problem requires the evaluation of the following probability distribution:

$$\begin{aligned} &P(x_k, f, m | Z^k, U^k, x_0) \\ &\quad \propto \quad P(z_k | x_k, f, m) P(X_k, m, f | Z^{k-1}, U^k, x_0) \end{aligned} \tag{8}$$

where x_k is the robot pose vector at time k, f is the features map, m is the metric map, Z^k is the set of observations received until time k, z_k are the observations at time k, U^k are the control inputs and x_0 the initial conditions.

If the data association problem can be solved then the observation can be divided into two groups: the observations belonging to the features map and the ones belonging to the metric map,

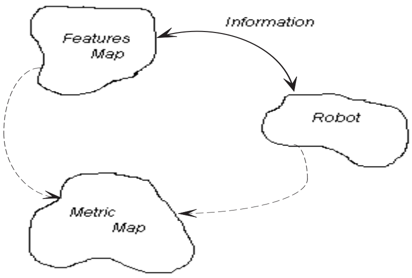

Figure 5: *Diode Correlation* effect in the HMM: The information between feature map and robot is bidirectional while the information between robot and metric map is unidirectional

$$P(z_k|x_k, f, m) = P(z_k^f|x_k, f)P(z_k^m|x_k, m) \qquad (9)$$

Now assume that the metric map does not provide any information to the robot pose and the features map then,

$$P(x_k, f|Z^k, m, U^k) = P(x_k, f|Z_f^k, U^k) \qquad (10)$$

This concept is illustrated in Fig. 5. There is information going in both directions between the robot and the landmarks map, and there is information going from the robot and features map to the metric map, but not from the metric map to the rest of the system. This assumption is justified since in a local area (LTR) the metric and the feature map will be highly correlated. In the limit when they are fully correlated, the information accumulated in one map is the same as in both maps. Then one map could be used for a SLAM process (features map) and the other to maximize (or accumulate) all the information gathered from the sensor, (e.g. to get an OG map). We will call to this effect *Diode Correlation* because the information goes in only one direction. The maps are still consistent between them due to the high correlation present in geographically close objects. The HMM uses this fact to get a consistent hybrid map.

Applying Eq. 9 and 10 to 8 it is possible to get the next result:

$$\begin{aligned} &P(x_k, f, m|Z^k, U^k, x_0) \\ &\quad \propto \quad P(z_k^f|x_k, f)P(x_k, f|Z_f^{k-1}, U^k, x_0) \\ &\qquad\quad P(z_k^m|x_k, m)P(m|Z^{k-1}) \end{aligned} \qquad (11)$$

The first factor in Eq. (11) is a features bayesian SLAM implementation, and the second factor is the metric bayesian map estimation. This equation demonstrates that by virtue of the *Diode Correlation* property the SLAM problem can be decoupled into two different problems: features based SLAM and the estimation of the metric map.

Figure 6: The utility car used for the experiments is equipped with a Sick laser range and bearing sensor, linear variable differential transformer sensor for the steering mechanism and back wheel velocity encoder.

From the practical point of view the *Diode Correlation* can be seen in the following way: Consider a landmarks map where a subset of landmarks are estimated and used in a SLAM process. The rest of the landmarks are mapped but not used for the SLAM. If each element of the second map is observed only once, then it is not necessary to maintain the cross-correlation between these new landmarks. Furthermore, if some landmarks of the second map are strongly correlated to geographically close landmarks of the SLAM map and in addition they are defined relative to them then the cross-correlation with respect to other landmarks of the SLAM map will be very weak and could be ignored [17]. This means that if for example the SLAM map has $N1$ landmarks (geographically close to the second map), and the second map has $N2$ landmarks, then for a full filter implementation would be necessary to maintain a covariance matrix of $(N1 + N2) * (N1 + N2)$. By considering the diode correlation property only a matrix of $N1 * (N1 + N2)$ will be required to be maintained. If $N1 << N2$, then the second update will be much cheaper than the full update. The second map can be thought now as the metric map in the HMM representation.

4 Experimental Results

This section presents experimental results in an outdoor environment. In the experiment a standard utility vehicle was fitted with dead reckoning sensors and laser range sensors as shown in Fig. (6).

The experimental test was done in the car park close to the ACFR building. The dataset can be found in [18].

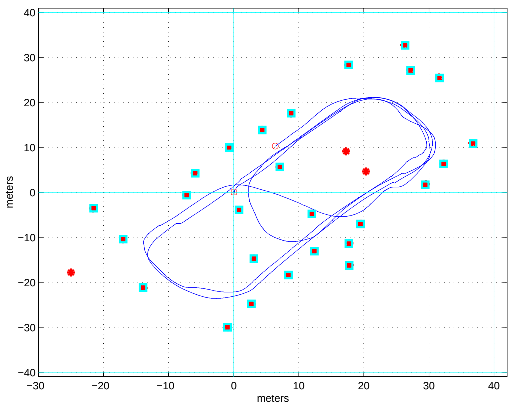

Figure 7: Estimated trajectory and landmark positions using SLAM

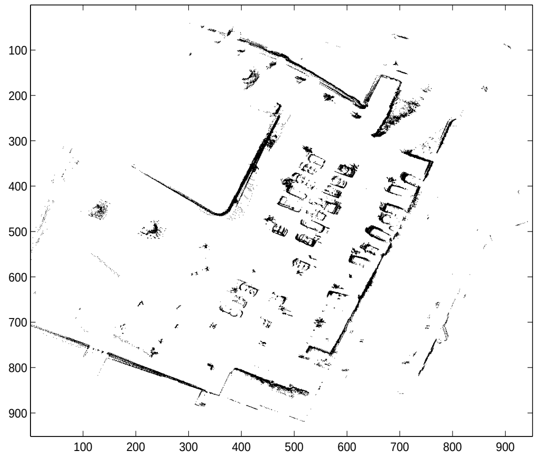

Figure 8: Car park map obtained with SLAM using laser range sensor.

(9) shows a satellite picture of the test place. No reliable GPS ground truth was available due to lack of satellite availability in this type of environment. Fig. (7) shows the trajectory and landmark positions estimated with the SLAM algorithm.

Fig. (8) shows a map of the environment obtained with the laser sensor. This map presents a plot of the laser observations after the errors in the landmarks' estimation settled to a constant value. A satellite image of the environment is also displayed in Fig. (9) to compare with the map obtained in Fig. (8).

The first step to implement the HMM approach is to divide the map into LTRs as shown in Fig. (10). The criterion used here to divide the regions was based on the distance between the base landmarks. In a robust implementation the correlation between the base landmarks should also be checked, however because this is not a big area and the landmarks were closed enough, this was not necessary. Part (a) of this figure shows the LTRs in

Figure 9: Car park satellite picture.

(a) (b)

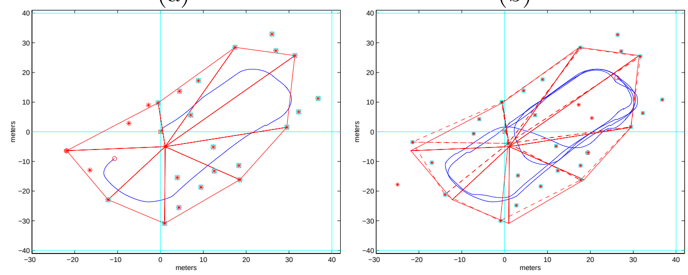

Figure 10: The LTRs in the outdoor experiment. (a) In the first lap (b) After some laps the landmark positions change. '-' triangles position in the first lap, '–' after some laps.

the first lap, after all the landmarks were incorporated and the regions were already formed. Part (b) shows the LTRs after a few laps. In (b) the landmark positions were updated changing the triangles' position and shape. It is evident that the right part of the map had a much smaller change than the left part. This is due to the fact that the vehicle started at location (0,0) and circulated clockwise. The map in the right has less uncertainty than the left part since the vehicle will usually accumulate error when performing exploration and SLAM. It is noteworthy that in general the LTRs will rotate or translate, but the shape will be preserved due to the correlation between the landmarks defining the vertices of a LTR.

Finally Fig. (11) shows a detailed view of the LTRs with the landmarks and OG maps at the beginning of the process and after a few laps.

5 Conclusions

The paper presented a new map representation that combines features with any other metric information in a consistent manner. The paper also introduced a new concept, the *Diode Correlation*, which allows the robot to build a denser map doing SLAM, in particular using the HMM

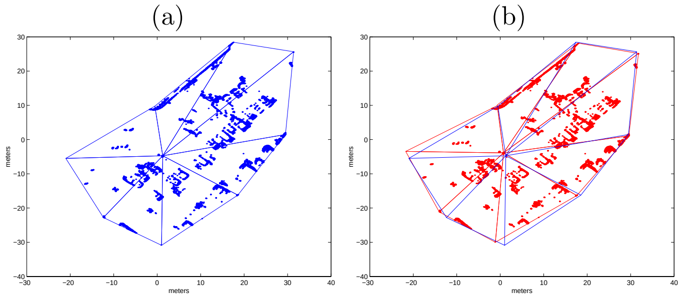

Figure 11: All the LTRs in the outdoor experiment. (a) In the first lap (b) After some laps, the local grid maps change because the landmark positions change.

representation. Experimental results in an outdoor environments were presented to demonstrate the algorithms. Among the areas of current and future research are:

- Application of the HMM to dynamic environments.
- Building a hierarchical map by integrating different sensors.
- Use of the representation to solve data association problems.

References

[1] J. Guivant, E. Nebot, and H. Durrant-Whyte, "Simultaneous localization and map building using natural features in outdoor environments," in *Proc. IAS-6 Intelligent Autonomous Systems*, July 2000, pp. 581–586.

[2] J. J. Leonard and H. J. S. Feder, "A computationally efficient method for large-scale concurrent mapping and localization," in *Ninth International Symposium on Robotics Research*, October 1999, pp. 316–321.

[3] J.E. Guivant and E.M. Nebot, "Optimization of the simultaneous localization and map building algorithm for real time implementation," *IEEE Transaction on Robotics and Automation*, vol. 17, no. 3, pp. 242–257, June 2001.

[4] T. Bailey, *Mobile Robot Localisation and Mapping in Extensive Outdoor Environments*, Phd thesis, University of Sydney, Australian Centre for Field Robotics, 2002.

[5] M. Bosse, P. M. Newman, J. J. Leonard, and S. Teller, "An atlas framework for scalable mapping," MIT Marine Robotics Laboratory Technical memorandum 2002-04, http://oe.mit.edu/%7Ejleonard/, 2002.

[6] J. Guivant, J. Nieto, F. Masson, and E. Nebot, "Navigation and mapping in large unstructured environments," Sent to publication, http://www.acfr.usyd.edu.au/publications/downloads/2003/Guivant186/ijrr_ltr.pdf, 2002.

[7] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *IEEE Computer*, vol. 22, pp. 46–57, June 1989.

[8] A.C. Schultz, W. Adams, B. Yamauchi, and M. Jones, "Unifying exploration, localization, navigation and planning through a common representation," in *IEEE International Conference on Robotics and Automation*, 1999, vol. 4, pp. 2651–2658.

[9] J.A. Castellanos, J. D. Tardos J.D, and Schmid, "Building a global map of the environment of a mobile robot: the importance of correlations," in *Proc. IEEE Robotics and Automation*, May 1997, vol. 2, pp. 1053–1059.

[10] J. Leonard and H. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Transaction on Robotics and Automation*, vol. 7, pp. 376–382, 1991.

[11] S.B Williams, P. Newman, G. Dissanayake, and H. Durrant-Whyte, "Autonomous underwater simultaneous localisation and map building," in *IEEE International Conference on Robotics and Automation*, 2000, vol. 2, pp. 1792–1798.

[12] B.Kuipers and Y.T.Byun, "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations," *Journal of Robotics and Autonomous Systems*, , no. 8, pp. 47–63, 1991.

[13] S. Thrun, "Robotic mapping: A survey," Tech. Rep., School of Computer Science, Carnegie Mellon University, 2002.

[14] S. Thrun and A. Bucken, "Integrating grid-based and topological maps for mobile robot navigation," in *Thirteenth National Conference on Artificial Intelligence*, August 1996.

[15] A. Makarenko, F. Bourgault, S.B. Williams, B. Grocholsky, and H.F. Durrant-Whyte, "An experiment in integrated exploration," in *IEEE International Conference on Intelligent Robots and Systems*, 2002.

[16] A. Elfes, *Occupancy Grids: A probabilistic framework for robot perception and navigation*, Phd thesis, Department of Electrical Engineering, Carnegie Mellon University, 1989.

[17] J. Guivant, *Efficient Simultaneous Localization and Mapping in Large Environments*, Phd thesis, University of Sydney, Australian Centre for Field Robotics, 2002.

[18] Australian Centre for Field Robotics University of Sydney, "Experimental outdoor data," http://www.acfr.usyd.edu.au/homepages/academic/enebot/dataset.htm.