# Recursive Scan-Matching SLAM

Juan Nieto, Tim Bailey and Eduardo Nebot

*ARC Centre of Excellence for Autonomous Systems (CAS)*
*The University of Sydney, NSW, Australia*

**Abstract**

This paper presents *Scan-SLAM*, a new generalisation of simultaneous localisation and mapping (SLAM). SLAM implementations based on *extended Kalman filter* (EKF) data fusion have traditionally relied on simple geometric models for defining landmarks. This limits EKF-SLAM to environments suited to such models and tends to discard much potentially useful data. The approach presented in this paper is a marriage of EKF-SLAM and scan correlation. Landmarks are no longer defined by analytical models; instead they are defined by templates composed of raw sensed data. These templates can be augmented as more data becomes available so that the landmark definition improves with time. A new generic observation model is derived that is generated by scan correlation, and this permits stochastic location estimation for landmarks with arbitrary shape within the Kalman filter framework. The statistical advantages of an EKF representation are augmented with the general applicability of scan matching. Scan matching also serves to enhance data association reliability by providing a shape metric for landmark disambiguation. Experimental results in an outdoor environment are presented which validate the algorithm.

*Key words:* Simultaneous localisation and mapping (SLAM), EKF-SLAM, scan correlation.

## 1 Introduction

A mobile robot must know where it is within an environment in order to navigate autonomously and intelligently. Self-location and knowing the location of other objects requires the existence of a map, and this basic requirement has lead to the development of the *simultaneous localisation and mapping* (SLAM)

*Email address:* {j.nieto,tbailey,nebot}@acfr.usyd.edu.au (Juan Nieto, Tim Bailey and Eduardo Nebot).

algorithm over the past two decades, where the robot builds a map piece-wise as it explores the environment. The predominant form of SLAM to date is *stochastic SLAM* as introduced by Smith, Self and Cheeseman [17]. Stochastic SLAM explicitly accounts for the errors that occur in sensed measurements: measurement errors introduce uncertainties in the location estimates of map landmarks which, in turn, incur uncertainty in the robot location estimate, and so the landmark and robot pose estimates are dependent. Most practical implementations of stochastic SLAM represent these uncertainties and correlations with a Gaussian probability density function (PDF), and propagate the uncertainties using an *extended Kalman filter* (EKF). This form of SLAM is known as EKF-SLAM [8]. One problem with EKF-SLAM is that it requires geometric-shaped landmark models to account for the sensed data, which limits the approach to environments suited to such models.

This paper presents an approach to SLAM which integrates scan correlation methods with the EKF-SLAM framework. The map is constructed as an online data fusion problem and maintains an estimate of uncertainties in the robot pose and landmark locations. There is no requirement to accumulate a scan history. Unlike previous EKF-SLAM implementations, landmarks are not represented by simplistic analytical models, but rather are defined by templates of raw sensor data. This way the feature models are not environment specific and good data is not thrown away. The result is *Scan-SLAM*, which uses raw data to represent landmarks and scan matching to generate landmark observations. In essence, this approach presents a new way to define arbitrary shaped landmark models, and in all other respects Scan-SLAM behaves in the manner of conventional EKF-SLAM.

The format of this paper is as follows. The next section presents a review of related work. Section 3 presents *Scan-SLAM*, which uses the method presented in Section 4.2 to implement a scan correlation based observation update step within the EKF-SLAM framework. Section 4 discusses scan segmentation and scan alignment. The Iterative Closest Point (ICP) algorithm is described in Section 4.2. The method is validated with experimental results presented in Section 5. Finally conclusions are presented in Section 6.

## 2 Related Work

A significant issue with EKF-SLAM [6] is the design of the observation model. Current implementations require landmark observations to be modelled as geometric shapes, such as lines or circles. Measurements must fit into one of the available geometric categories in order to be classified as a feature, and non-conforming data is ignored. The chief problem with geometric landmark models is that they tend to be environment specific, so that a model suited

to one type of environment might not work well in another and, in any case, a lot of useful data is thrown away.

An alternative to analytical feature models is a procedure called *scan correlation*, which computes a maximum likelihood alignment between two sets of raw sensor data. Thus, given a set of observation data and a reference map composed similarly of unprocessed data points, a robot can locate itself without converting the measurements to any sort of geometric primitive. The observations are simply aligned with the map data so as to maximise a correlation measure. Scan correlation has been used as a localisation mechanism from an *a priori* map [21,10,5,11], with the *iterated closest point* (ICP) algorithm [4,13] and occupancy grid correlation [7] being the most popular correlation methods.

Two important methods have been presented to perform SLAM via scan correlation. The first [19] uses *expectation maximisation* (EM) to maximise the correlation between scans, which results in a set of robot pose estimates that give an "optimal" alignment between all scans. The second method, called *consistent pose estimation* (CPE) [9], accumulates a selected history of scans, and aligns them as a network; this approach is based on the algorithm presented in [13].

The main concern with existing scan-correlation SLAM methods is that they do not perform data fusion, instead requiring a (selected) history of raw scans to be stored, and they are not compatible with the traditional EKF-SLAM formulation. This paper presents a new algorithm that combines the EKF-SLAM and scan correlation methods.

## 3 Scan-SLAM

Scan-SLAM is identical to conventional EKF-SLAM except for the definition of the landmark appearance model. Landmarks are defined by a template of raw sensor data; landmarks are observed by a process of scan-matching. This process gives rise to a generic observation model, which is the location of a local coordinate frame embedded in the landmark template. The landmark templates also facilitate an augmented data association strategy.

### *3.1 Algorithm Overview*

A landmark definition template is created by extracting a cluster of data points from a measurement scan and transforming them to a local coordinate frame.

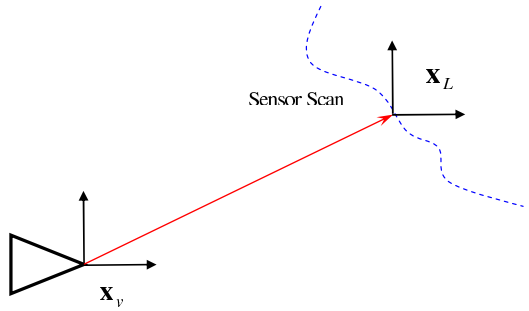



Fig. 1. The local axis position is defined as the centroid of the template points and orientation equal to the vehicle heading.

While there is no inherent restriction as to *where* this local axis is defined, it is more intuitive to locate it somewhere close to the landmark data-points and, in this paper, the local coordinate frame is defined as the centroid of the template points. Figure 1 illustrates the concept. A new landmark is added to the SLAM map by adding the global pose of its coordinate frame $\mathbf{x}_L$ to the SLAM state vector. Note that the landmark description template is not added to the SLAM state and is stored in a separate data structure.

Suppose we have an existing SLAM estimate of map landmarks and a set of associated template models. As new scans become available, the SLAM estimate can be updated by the following process. First, the location of a map landmark relative to the vehicle is predicted to determine whether the landmark template $\mathbf{S}_L$ is in the vicinity of the current observed scan $\mathbf{S}_o$. This vehicle-relative landmark pose is the predicted observation $\hat{\mathbf{z}} = [\hat{x}_\delta, \hat{y}_\delta, \hat{\phi}_\delta]^T$ according to the observation model (see Equation 4 below). If there is a successful data association between the landmark template and the current scan, the landmark template is aligned via the scan matching algorithm, using $\hat{\mathbf{z}}$ as an initial guess. This process is illustrated in Figure 2.

At a higher level of abstraction, the result of this algorithm can be described by the following pseudocode function interface.

$$\texttt{[z, R] = scan\_align(}\mathbf{S}_L\texttt{, }\mathbf{S}_o\texttt{, }\hat{\mathbf{z}}\texttt{)}$$

That is, scan alignment results in an observation $\mathbf{z}$, with uncertainty $\mathbf{R}$, which is the pose of the landmark template frame relative to the current vehicle pose. The form of $\mathbf{z}$ is described in the next section. Having obtained the observation $\mathbf{z}$ and $\mathbf{R}$, the SLAM state is updated in the usual manner of EKF-SLAM.

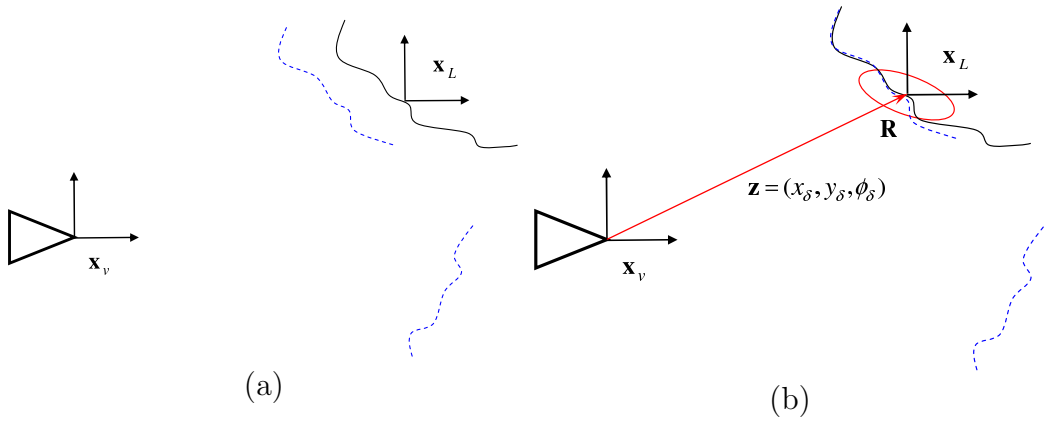



Fig. 2. Figure (a) shows a stored scan landmark template (solid line) positioned at the predicted pose $\hat{\mathbf{z}}$ relative to the vehicle, and a new observed scan (dashed line). Figure (b) shows the scan alignment evaluated with the scan correlation algorithm from which the observation vector $\mathbf{z}$ is obtained.

### *3.2 Generic Observation Model*

The data points representing a scan landmark template are stored in a data structure separate from the actual stochastic SLAM estimate, and is used by the scan alignment algorithm to obtain *landmark location observations*. The stochastic SLAM state is composed of the global locations of landmark coordinate frames. Thus, the stochastic observation model for *all* landmarks is the measurement of a global landmark frame as seen from the global vehicle pose (see Figure 3).

The SLAM state vector

$$\mathbf{x}_{joint} = \begin{bmatrix} \mathbf{x}_v \\ \mathbf{x}_L \end{bmatrix} \tag{1}$$

consists of the vehicle pose

$$\mathbf{x}_v = [x_v, y_v, \phi_v]^T \tag{2}$$

and a set of landmark frame locations

$$\mathbf{x}_L = [x_{L_1}, y_{L_1}, \phi_{L_1}, \ldots, x_{L_i}, y_{L_i}, \phi_{L_i}, \ldots, x_{L_n}, y_{L_n}, \phi_{L_n}]^T \tag{3}$$

The generic observation model for the pose of a landmark coordinate frame

with respect to the vehicle is as follows.

$$\begin{aligned} \mathbf{z} = [x_\delta, y_\delta, \phi_\delta]^T &= \mathbf{h}\left(\mathbf{x}_L, \mathbf{x}_v\right) \\ &= \begin{bmatrix} (x_L - x_v)\cos\phi_v + (y_L - y_v)\sin\phi_v \\ -(x_L - x_v)\sin\phi_v + (y_L - y_v)\cos\phi_v \\ \phi_L - \phi_v \end{bmatrix} \end{aligned} \tag{4}$$

### *3.3 Landmark Representation*

Scan-SLAM defines the landmarks by a template of points in a local landmark coordinate frame. This local representation allows to *enrich* the landmark-templates with the information collected from different sensors, such as 3D-lasers or cameras. The addition of more information allows to build *high-dimensional landmarks* which makes them more distinctive and makes the association process very robust [15]. An example is shown in Figure 4. Figure (a) shows the information collected from a laser in an outdoor environment. Figure (b) shows an image of the same part of the environment and (c) shows the laser information superimposed with the video image. The laser was calibrated with the camera using the algorithm presented in [22] [1] . The laser and video information and any other sensed data can be stored in the landmark local coordinate frame to represent a high-dimensional landmark.

[1] Thanks to Fabio Ramos for the images.

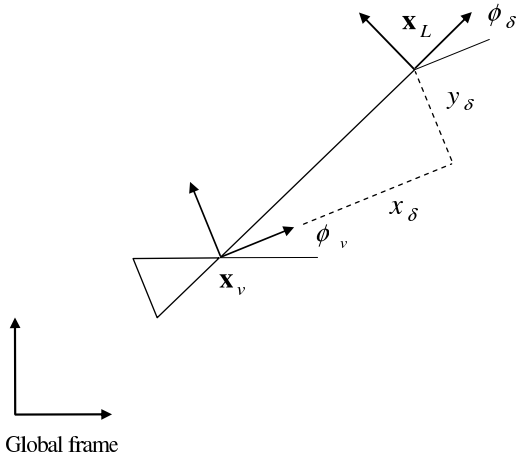
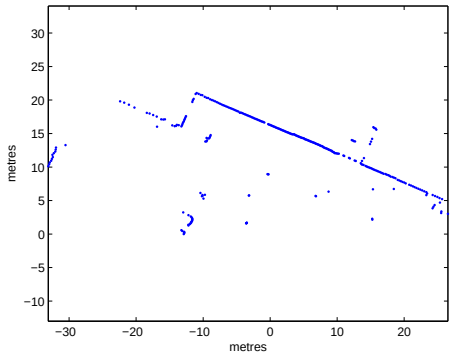



Fig. 3. All features are represented in the SLAM map as a global pose identifying the location of the landmark coordinate frame. The generic observation model for these features is a measurement of the global landmark pose $\mathbf{x}_L$ with respect to the global vehicle pose $\mathbf{x}_v$. The vehicle-relative observation is $\mathbf{z} = [x_\delta, y_\delta, \phi_\delta]^T$.




(a)

(b)

(c)

Fig. 4. Figure (a) shows the information collected from a laser in an outdoor environment. Figure (b) shows an image of the same part of the environment and (c) shows the laser information superimposed with the video image. The scan points are represented with '+' in the video image.

### *3.4 Data association*

A difficult problem in SLAM is data association, largely because the traditional criterion used to distinguish one landmark from another has been solely based on the *innovation gate* distance metric. Exclusive use of this criterion does not exploit other landmark attributes that may be useful for disambiguation. As a byproduct of scan-alignment, we obtain a measure of how well a scan segment matches the shape of a landmark template and we use this as an additional data association metric.

#### *3.4.1 Innovation Gate*

The traditional data association metric in target tracking is the *normalised innovation squared* [2], which is the error between predicted and actual observation normalised by error covariance. A threshold on this quantity is known as the *innovation gate*.

The innovation gate is also used for data association in stochastic SLAM, and permits landmarks to be distinguished by their position. However, landmarks that are too close in position give rise to ambiguous associations. Whether two landmarks are too close to disambiguate is a function of their position uncertainty, the vehicle pose uncertainty and the measurement noise. Batch association methods [14,1], where multiple measurements are considered simultaneously, greatly reduce this uncertainty but ambiguous associations cannot be completely eliminated.

#### *3.4.2 Shape Matching*

Landmark attributes can greatly augment position-based association reliability, and an attribute that is available from the landmark templates is shape. If an aligned scan segment does not have the same shape as its landmark template, the association is rejected. Thus, data association is strengthened by accepting only when both shape and innovation gate criteria are met.

## 4 Implementation

In this section we describe four operations that are required to apply scan matching to the EKF-SLAM algorithm. While these four operations are required to implement the algorithm, the best solution for them is still an open problem. We stress here that our current implementations are essentially "proof of concept", and do not constitute our primary contribution, which is the concepts described in the previous section.

The first operation described is scan segmentation whereby a scan of data points is divided into clusters and *reliable* clusters are selected to represent landmarks. The second is scan correlation, aligning the data of multiple scans so that data representing the same object overlap. The third is the generation of a suitable measure of alignment uncertainty or variance. The fourth is a shape matching criteria that determines whether an alignment is sufficiently good to be accepted as a valid observation. This last operation is not essential but serves to assist the conventional data association mechanism used in SLAM.

The methods in this section are by no means the only way of accomplishing the underlying operations. The main thrust of this paper is the generic observation model described in Section 3, and the implementations developed here are but practical realisations to achieve this goal. For segmentation we use a distance metric and saliency score; for scan correlation we use ICP; for variance estimation we implement a sampling algorithm; and for shape matching

we employ a simple distance gate.

### *4.1 Scan Segmentation*

When a new scan is received it is first broken into clusters. This is done by comparing consecutive range measurements, which arrive in order of sweep angle. For two successive measurements, if the difference in either range or angle is greater than a threshold, the latter measurement forms the beginning of a new cluster. This algorithm groups together clusters of near points.

Having segmented the scan, it remains to determine which clusters will make a reliable landmarks. Two criteria are used. The first is that a cluster must contain more than a minimum number of points, and the second is a measure of how informative a cluster is for pose estimation.

A number of measures for cluster informativeness appear in the literature. A parameter called *typical boundary error* is introduced in [16] and applied to 3-D images. The averaged RMS error for the vertices gives an estimate of point alignment accuracy. Similarly, in [18] another parameter is defined called *registration index*. The registration index gives an indication of how well two surfaces may be registered. In [20] a parameter called *object saliency score* is defined. This parameter is defined as the inverse of the covariance matrix trace, obtained after aligning two shapes. The larger the object saliency score, the more certain the pose estimate from the registration process.

These parameters can then be used to predict the accuracy of the registration process with a particular segment and then, to assist in deciding whether the segment will be incorporated as a new landmark or not. In this paper we use the *object saliency score* which is defined in [20] as

$$\mathbb{S} = \frac{1}{trace(\mathbf{R})} \tag{5}$$

where $\mathbf{R}$ is the scan matching covariance matrix obtained as explained in Section 4.3. Since the matrix $\mathbf{R}$ is the covariance of the observation which includes position and orientation, if the object saliency score is applied as in Equation 5 we would be adding values with different dimensions (e.g. metres for position and radians for orientation). For that reason, here we obtain the saliency for position and orientation separately.

The segmentation procedure is as follows: when a new frame is received it is first segmented into clusters using the method explained above. The autocovariance of a cluster is obtained by calculating the covariance of the alignment between the cluster and a copy of itself. The covariance of the alignment is

| Object | Wall | Corner | Arbitrary Shape |
|---|---|---|---|
| $\mu_{\mathbb{S}}$ | 1.14 | 124 | 592 |
| $\sigma^2_{\mathbb{S}}$ | 0.85 | 0.53 | 0.7 |

Table 1
Object Saliency Score for three different objects. The first row shows the mean of the saliency score obtained with 100 trials. The second row shows the variance of the saliency score.

calculated using the method presented in Section 4.3. This method rotates and translates a copy of the cluster and then applies scan matching to align them back. This process is done with $N$ copies of the cluster, and the autocovariance is calculated based on the result of the alignments. Section 4.3 explains this process in detail. Following, the object saliency score is calculated and based on the result the segment is rejected or accepted as a new landmark.

To illustrate the concept, the object saliency score was calculated for the three objects presented in Figure 6. In order to analyse the stability of the index, the saliency score was calculated several times for each object. An observation of each object was taken and the saliency score was evaluated 100 times using different sensor pose samples for the covariance estimation (see Section 4.3). Table 1 summarises the result. Based on the outcome of the 100 trials, we obtained a mean and variance values for the saliency score of each object. As expected, the wall has the lowest value with mean equal to 1.14. On the other hand the corner and the arbitrary shape object possess much higher score index mean, 124 and 592 respectively. The second row in the table shows the variance of these values over the 100 trials. As seen, the index presents very small variance for all cases. The saliency score presented in the table was calculated using the variance in position. In all the cases, the variance in orientation was very small (less than 0.3 degrees).

Figure 7 shows an example of the saliency score obtained with two segments extracted from scans taken with a Sick laser. The segment showed in (a) was obtained by observing a wall. This segment possess very low saliency score equal to 2.7. (b) illustrates the observations obtained from the corner of a building. The saliency score of the corner is 112.

### *4.2 Scan Alignment via ICP*

Unprocessed data correlation, also called scan alignment or range-image registration, is the process of aligning an observed set of (2-D or 3-D) points with a reference point set. ICP [4,23] is arguably the most commonly used range-image registration technique, with its popularity due mainly to its simplicity and efficiency. The basic algorithm works as follows. Let $P_o = \{p_1, ..., p_m\}$ rep-

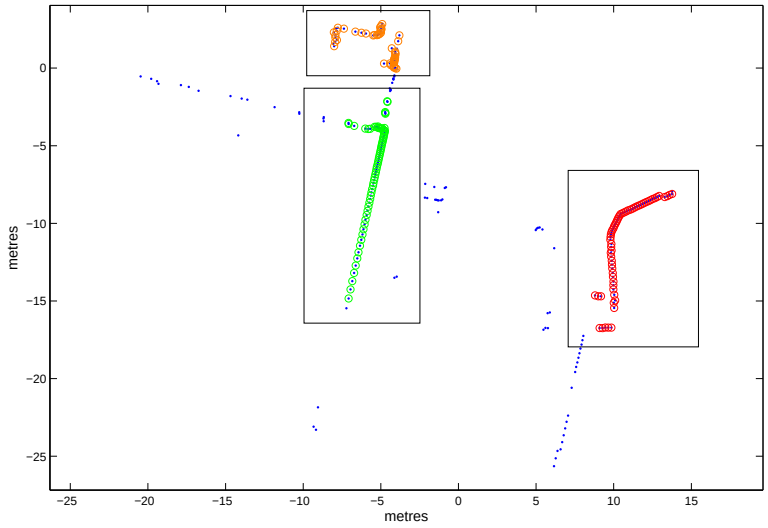
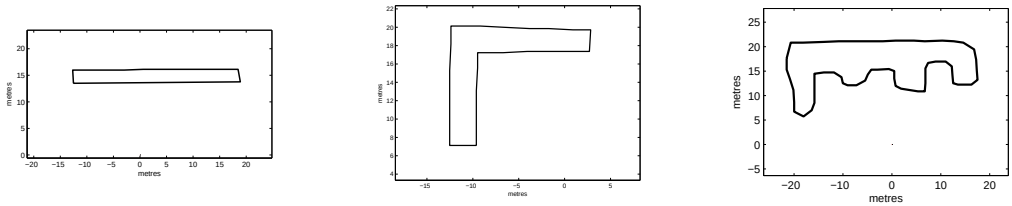


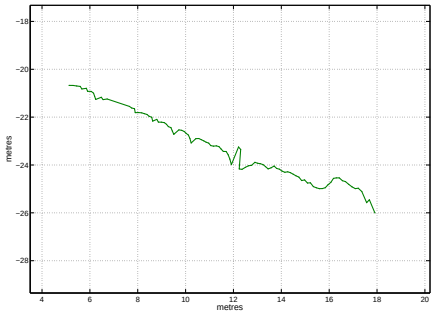

Fig. 5. Scan segmentation based on distance. The '.' represent the laser scan. Three cluster of points were selected by the algorithm. The individual clusters are represented with different marker styles and colours and enclosed by rectangles.



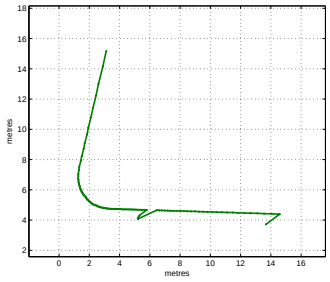

(a) (b) (c)

Fig. 6. The figure shows three different shapes used to test the saliency score. Figure (a) represents a simulated wall, (b) a corner and (c) an object with arbitrary shape.

resent the observation point set and $P_r = \{p_1, ..., p_n\}$ be the reference point set. The aim of the algorithm is to find a geometric transformation to align the observed points $P_o$ to the reference point set $P_r$. This transformation is composed by a translation and a rotation. ICP is an iterative algorithm, consisting in two steps. The first step involves finding the correspondence between the points in the observed and reference scan. This is done using a nearest neighbour algorithm. The second step is to obtain a rotation and translation between the two scans, minimising the *mean square error* (MSE). The algorithm is initialised with an initial pose guess and, until the estimated pose satisfies some convergence criterion, it is iteratively refined by a process of point-to-point data association and least-squares transformation. Each point $p \in P_o$ is first transformed to the reference coordinate frame using the cur-




(a)




(b)

Fig. 7. Segments extracted from experimental data using a Sick Laser. The segment in (a) has very low saliency score equal to 2.7 because it is very linear; no orthogonal parts. The segment in (b) has a saliency score equal to 112.

rent pose estimate, and then associated to its nearest neighbour in $P_r$. The original point $p$ and its associate $q$ are added to an association set $E$. Finally, the pairs in $E$ are used to calculate the relative pose that minimises the least-mean-squared error between the associated points. For the experiments in this paper we use a quaternion based method as presented in [4]. An example of ICP alignment is shown in Figure 8.

### *4.3 Scan Alignment Variance*

In order to fuse the scan alignment results with other sensor measurements, it is necessary to obtain a covariance estimate of the alignment.

In [12], the square error function is linearised and the covariance matrix is derived analytically using linear regression theory. One of the problems of ICP is the heuristic way the algorithm finds corresponding pair of points. The method to obtain the covariance matrix presented in [12] will give good results only if corresponding points are correctly associated. The method does not consider the error in point correspondence that will happen when observing objects with symmetry (e.g. a wall in a long corridor) and thus tends to give very optimistic covariances.

The method presented in [3] obtains the covariance matrix by estimating the Hessian matrix of the error function minimised by ICP. It is claimed that by using the Hessian matrix, the method does not suffer the corresponding problem as in [12]. Unfortunately the method still tends to be very optimistic and fails to obtain realistic covariances where points correspondence cannot be guarantee. In addition this method does not take measurement noise into account, it assumes equal noise for all the points.

Another technique to estimate the covariance of scan alignment was presented in [20]. Instead of using one initial relative transformation, the registration process is run $N$ times with random initial relative transformations. The distribution of the $N$ samples after the alignment process gives an estimate of the alignment covariance. The covariance estimates using the sampling method give a good approximation to the actual distribution. In order to include the measurement noise into the covariance estimates, the method transforms the sensor scans into occupancy grid maps and uses grid correlation [11] to weight the samples. The samples with low correlation response are filtered out.

The technique used in this paper is also based on a sampling approach. First, the scan alignment algorithm is run to align the observation with the reference scan. Then, $N$ samples of the vehicle pose are generated around the ICP solution. The registration process is run $N$ times using the samples as initial relative transformations. A covariance matrix is obtained for each sample by linearizing the square error and deriving the covariance matrix analytically as in [12]. The $N$ Gaussians are then converted into one Gaussian by computing the first and second moment of the mixture.

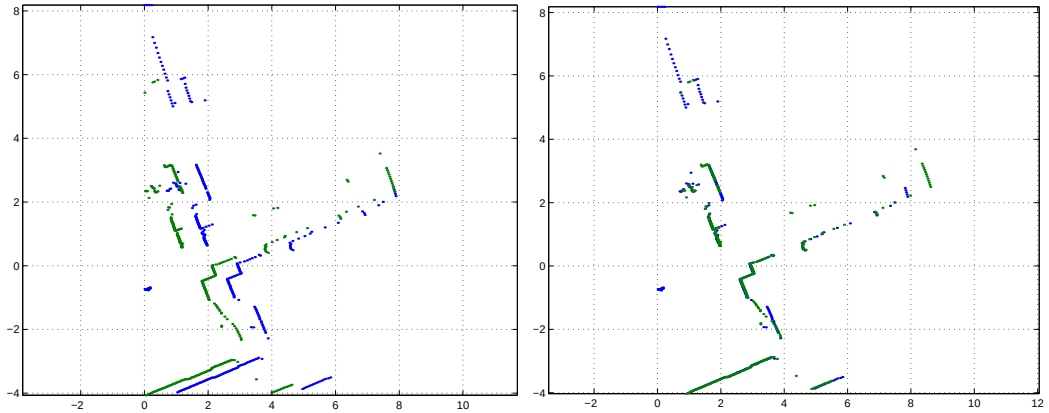
$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i}^{N}\mathbf{x}_i \tag{6}$$

$$\mathbf{P} = \frac{1}{N}\sum_{i}^{N}\left(\mathbf{P}_i + (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T\right) \tag{7}$$

Using this sampling method, the algorithm eliminates the problem of optimistic results due to erroneous data associations between points, since each sample of the robot pose will make different point-to-point associations.

Figure 9 shows an example of covariance estimation using the sampling based approach. The object observed is depicted with a solid line in Figure (a). The covariance obtained is represented with an ellipse. Figure (b) shows a zoom of the estimated covariance. The standard deviations obtained with the sampling based approach were $\sigma$ = [0.89 m, 0.85 m, 0.36 deg] and applying linearization and evaluating the analytical solution as in [12] the standard deviations obtained were $\sigma$ = [0.025 m, 0.027 m, 0.13 deg]. Clearly the second solution is very optimistic since is giving accuracy of 2 cm even when the object observed possesses symmetry. It can be also seen in Figure 9 (b) that the covariance estimated with the sampling approach is consistent with the actual error.

Figure 10 shows the samples used to calculate the covariance. The left-hand figure in (a) shows the samples from the robot position (x-y). The samples were obtained from a uniform distribution around the position obtained with ICP. The right-hand figure in (a) shows the samples distribution after applying



(a) (b)

Fig. 8. Example of scan alignment using ICP: (a) shows two scans taken from different position before alignment. (b) shows the scans aligned using ICP.

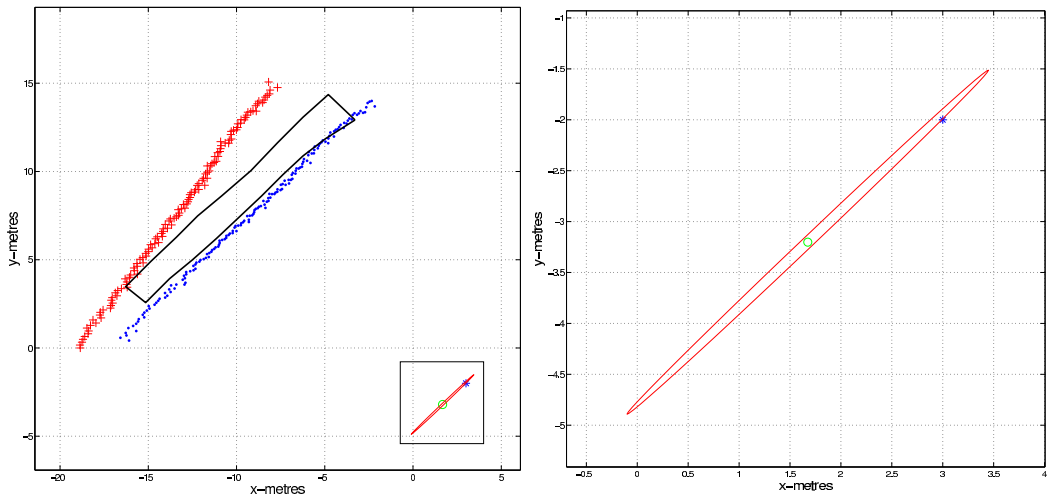
ICP. The scan matching algorithm aligns the samples in the direction parallel to the object observed. Finally, Figure (b) shows the samples for the vehicle heading where it can be seen that all the samples converge to the same value. The observation was taken from [3 m, -2 m, -10 deg].

The variation of the covariance estimates with the number of samples was analysed. A simulated wall parallel to the x axis was used for the analysis. Figure 11 (a) shows the standard deviation in x-y. As expected, the uncertainty in x is much larger since the object is parallel to this axis. Each value was obtained as the mean of 10 trials. Figure (b) shows the variation of the standard deviation in the vehicle orientation. As it can be observed the standard deviations tend to be very stable with the number of samples, which means that even for cases such as the one shown in the example where the covariance of the alignment is large, a low number of samples gives good results.

### *4.4 Shape Validation*

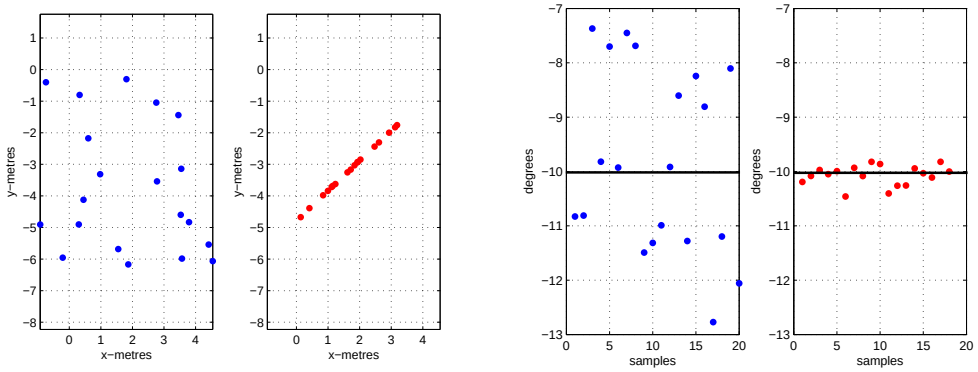
Once two scans have been aligned, it is possible to determine how well their data-points fit together, for data association purposes. We wish for this shape-matching criterion to be immune to view-point variation, so only those points that comprise an overlapping region are compared in the algorithm.

The basic procedure is as follows. First, the points in each scan that form part of a common region are extracted. If there are too few points in this overlapping region for either scan, then the match is rejected. Next, each extracted point from scan one is mapped to its nearest-neighbour in scan two, and their distances are computed. If an excess percentage of neighbours are greater than a threshold away, the match is rejected. Otherwise the scans are




(a) (b)

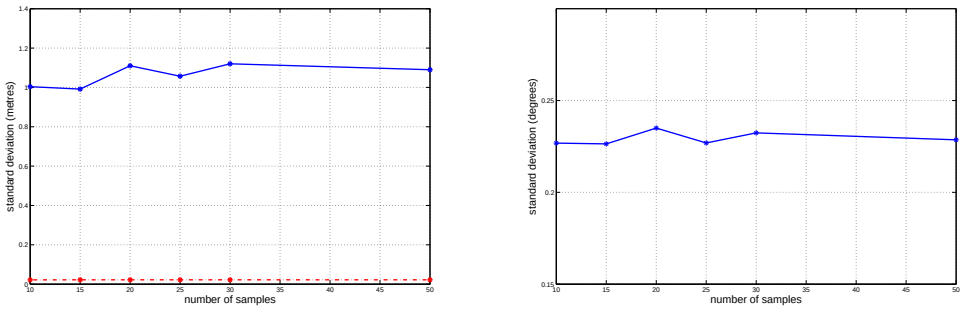
Fig. 9. Figure (a) shows the covariance estimated when observing a simulated wall which is depicted with solid line. The '.' represents the reference scan and the '+' the observation. The $2\sigma$ covariance bound is represented with an ellipse. Figure (b) shows a zoom of the estimated covariance. The 'o' denotes the vehicle position estimated with ICP and the '*' the actual observation position.




(a) (b)

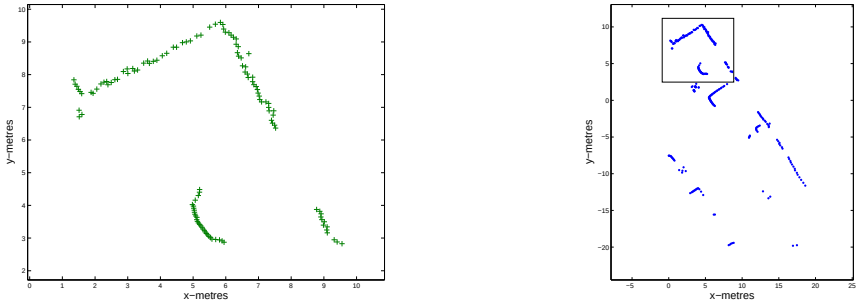
Fig. 10. The left-hand figure in (a) shows the samples from the robot position (x-y) used to obtain the covariance. The samples were obtained from a uniform distribution. The right-hand figure shows the samples distribution after applying ICP. (b) shows the samples for the vehicle heading. The observation was taken from [3m , -2m, $-10^o$]

taken to possess similar shape. Figures 12 and 13 show an example of the shape validation procedure. Figure 12 (a) shows a scan landmark and (b) an example of a scan that contains that scan landmark. Figure 13 (a) illustrates a case of a positive shape validation result and (b) a case where the alignment is rejected by the shape validation procedure.



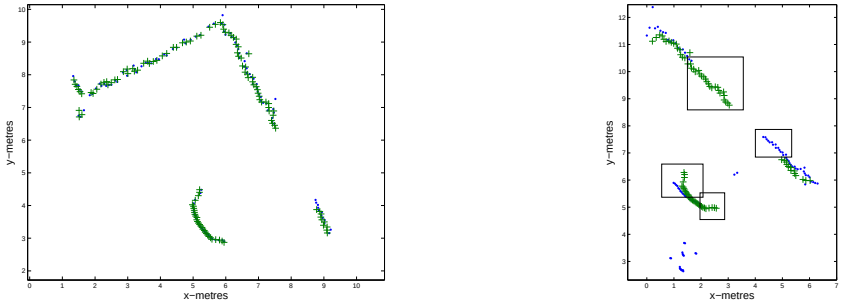

(a) (b)

Fig. 11. The figures show the standard deviation of the covariances obtained using different number of samples. Each value was obtained as the mean of 10 trials. A simulated wall parallel to the x-axis was used. (a) shows in dashed line the standard deviation in the y direction and in solid line for x. As expected, the uncertainty in x is much larger since the object is parallel to this axis. (b) shows the variation of the standard deviation in the vehicle orientation.




(a) (b)

Fig. 12. Figure (a) shows a landmark template. (b) shows a laser scan which includes the scan-landmark showed in (a). The rectangle enclose the overlapping region.




(a) (b)

Fig. 13. Figure (a) shows an example of a match accepted by the algorithm. The points of the scan landmark are depicted by '+' and the '.' represents the whole laser scan. The figure shows only the points that form part of a common region. (b) shows an example where the algorithm rejected the match. The rectangles enclose areas where the two scans do not overlap.

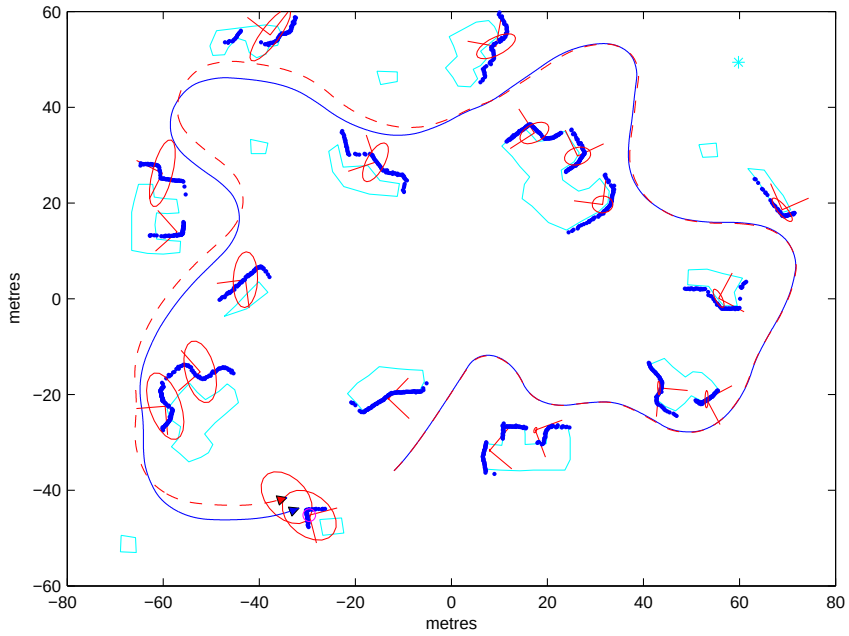
## 5 Results

This section shows simulation and experimental results of the algorithm presented. The importance of the simulation results is in the possibility to compare the actual objects position with the estimated by Scan-SLAM.

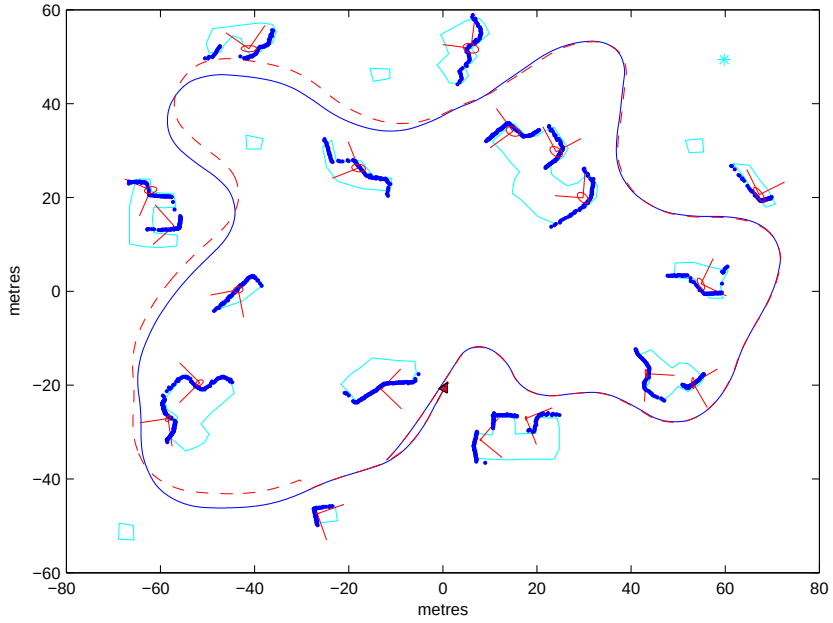
Figure 14 shows the simulation environment. The experiment was done in an area of 180 by 160 metres with a sensor field of view of 40 metres. The vehicle travels at a constant speed of 3 m/s. The sensor observations are corrupted with Gaussian noise with standard deviations of 0.1 metres in range and 1.5 degrees in bearing. The simulation map consists of objects with different geometry and size. The segmentation was done as explained in Section 4.1. An algorithm searches for clusters of neighbour points which contain a minimum number of points. Due to the shape of the objects it was not necessary the use of the saliency index. All the clusters were incorporated as landmarks.

The results for the Scan-SLAM algorithm are shown in Figure 14 (a) and (b). Here the solid line depicts the ground truth for the robot pose and the dashed line the estimated vehicle path. The actual object positions are represented by the light solid line and the segment positions by the dark points. The local axis pose for each scan landmark is also shown and the ellipses indicate the $3\sigma$ uncertainty bound of each scan landmark. The local axis position was defined equal to the average position of the raw points included in the segment, and the orientation equal to the vehicle orientation. Figure 15 shows a zoom of the left top area of the simulation which depicts in more detail the sensory information that is added to the representation obtained by the algorithm. Figure 14 (b) shows the result after the vehicle closes the loop and the SLAM algorithm updates the map. The good alignment between the actual object positions and the estimates by the algorithm illustrates the accuracy of the results obtained by the approach.

The algorithm was also tested using experimental data. In the experiment a standard utility vehicle was fitted with dead reckoning and laser range sensors. The testing environment was the car park area near the university building. Figure 16 (a) shows the vehicle used for the experiments and (b) shows a satellite picture of the experimental area. The environment is mainly dominated by buildings and trees. Figure 17 (a) illustrates the result obtained with the algorithm. The solid line denotes the trajectory estimated. The light points represent a laser-image obtained using feature-based SLAM and GPS, which can be used as a reference for the objects position. The dark points represent the template scans and the ellipses the $1\sigma$ covariance bounds. The local axes for the scan landmarks were also drawn in the figure. The same segmentation algorithm applied for the simulation was used here. After the scan is segmented into clusters of points, the saliency score index is calculated




(a)




(b)

Fig. 14. Figure (a) shows the simulation environment. The solid line depicts the ground truth for the robot pose and the dashed line the estimated vehicle path. The actual object positions are represented by the light solid line and the segment positions by the dark points. The ellipses indicate the $3\sigma$ uncertainty bound of each scan landmark. Figure (b) illustrates the result after closing the loop.

for each cluster in order to decide whether they are incorporated as a new landmark. The saliency score threshold used for acceptance was set to 100. Five scan landmarks were incorporated and used for the SLAM. The accuracy of the result can be seen by comparing the scan landmarks position with the

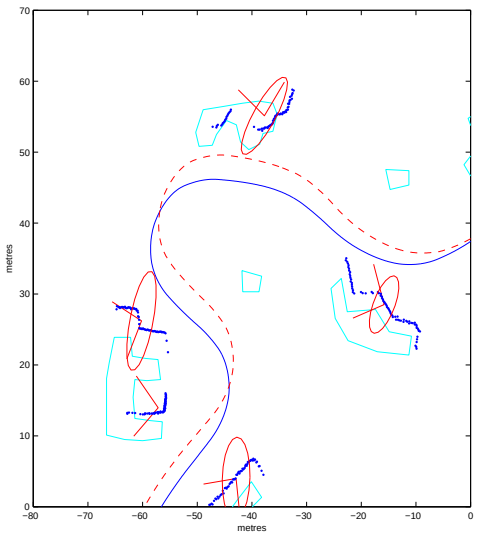



Fig. 15. Zoom of the left top area of the simulation environment. The figure shows the actual object poses with plain lines. The landmark templates are denote with '.'. The $3\sigma$ ellipses for the local axis representing the scan landmarks are also shown.

(a) (b)

Fig. 16. Figure (a) shows the experimental platform used to collect data. The utility car is equipped with Sick lasers, linear variable differential transformer sensor for the steering mechanism, back wheel velocity encoder and inertial unit. Figure (b) illustrates a satellite picture of the testing environment.

superimposed laser-image. Figure 17 (b) shows an example of the alignment between a scan-landmark and a new sensor frame. The figure shows one of the scan-landmarks and a laser scan at the moment these two are being aligned by the scan-matching algorithm. Figure 18 illustrates the final map with the local coordinate systems, the landmark template points and its object saliency score. As mentioned, a minimum of 100 in saliency was required for a segment to be accepted as a landmark.

## 6 Conclusions

EKF-SLAM is currently the most commonly used filter to solve the stochastic SLAM problem. An important issue with EKF-SLAM is that it requires sensory information to be modelled as geometric shapes and the information that does not fit in any of the geometric models is usually rejected. On the other hand, scan correlation methods use raw data and are not restricted to geometric models. Scan correlation methods have mainly been used for localisation given an *a priori* map. Some algorithms that perform scan correlation based SLAM have appeared, but they do not perform data fusion and they require storage of a history of raw scans.

The Scan-SLAM algorithm presented in this paper combines scan correlation with EKF-SLAM. The hybrid approach uses the best of both paradigms; it incorporates raw data into the map representation and so does not require geometric models, and estimates the map in a recursive manner without the need to store the scan history. It works as an EKF-SLAM that uses arbitrary shaped landmark models and utilises scan correlation algorithms to produce landmark observations. Experimental results in an outdoor environment were presented that showed the efficacy of the algorithm.

## Acknowledgments

This work is supported by the ARC Centre of Excellence programme, funded by the Australia Research Council (ARC) and the New South Wales State Government.

## References

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

[2] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley and Sons, 2001.

[3] O. Bengtsson and A.J. Baerveldt. Localization in changing environments - estimation of a covariance matrix for the idc algorithm. In *IEEE International Conference on Intelligent Robots and Systems*, volume 4, pages 1931–1937, 2001.

[4] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256,

1992.

[5] W. Burgard, A. Derr, D. Fox, and A.B. Cremers. Integrating global position estimation and position tracking for mobile robots: The dynamic markov localization approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1998.

[6] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3):229–241, 2001.

[7] A. Elfes. Occupancy grids: A stochastic spatial representation for active robot perception. In *Sixth Conference on Uncertainty in AI*, 1990.

[8] J. Guivant and E. Nebot. Optimization of the simultaneous localization and map building algorithm for real time implementation. *IEEE Transactions on Robotics and Automation*, 17(3):242–257, 2001.

[9] J.S. Gutmann and K. Konolige. Incremental mapping of large cyclic environments. In *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pages 318–325, 1999.

[10] J.S. Gutmann and C. Schlegel. Amos: Comparison of scan matching approaches for self-localization in indoor environments. In *1st Euromicro Workshop on Advanced Mobile Robots (Eurobot'96)*, pages 61–67, 1996.

[11] K. Konolige and K. Chou. Markov localization using correlation. In *International Joint Conference on Articial Intelligence*, pages 1154–1159, 1999.

[12] F. Lu. *Shape Registration Using Optimization for Mobile Robot Navigation*. PhD thesis, Department of Computer Science, University of Toronto, 1995.

[13] F. Lu and E. Milios. Robot pose estimation in unknown environments by matching 2D range scans. *Journal of Intelligent and Robotic Systems*, 18:249–275, 1997.

[14] J. Neira and J.D. Tardós. Data association in stochastic mapping using the joint compatibility test. *IEEE Transactions on Robotics and Automation*, 17(6):890–897, 2001.

[15] J. Nieto. *Detailed Environment Representation for the SLAM Problem*. PhD thesis, University of Sydney, Australian Centre for Field Robotics, 2005.

[16] X. Pennec and J. P. Thirion. A framework for uncertainty and validation of 3d registration methods based on points and frames. *International Journal of Computer Vision*, 25(3):203–229, 1997.

[17] R. Smith, M. Self, and P. Cheeseman. A stochastic map for uncertain spatial relationships. In *Fourth International Symposium of Robotics Research*, pages 467–474, 1987.

[18] A.J. Stoddart, S. Lemke, A. Hilton, and T. Renn. Estimating pose uncertainty for surface registration. *Image and Vision Computing*, 16(2):111–120, 1998.

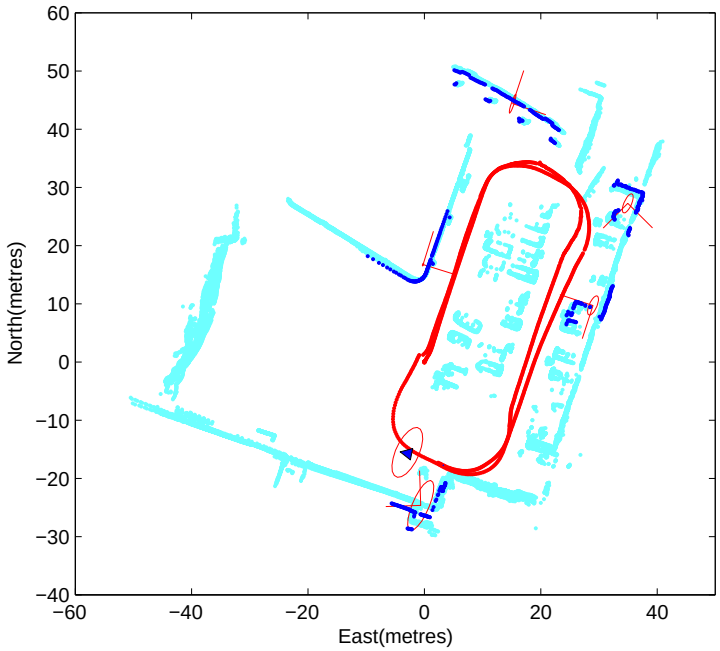
[19] S. Thrun, W. Bugard, and D. Fox. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. In *International Conference on Robotics and Automation*, pages 321–328, 2000.

[20] Chieh-Chih Wang. *Simultaneous Localization, Mapping and Moving Object Tracking*. PhD thesis, Robotics Institute, Carnegie Mellon University, 2004.

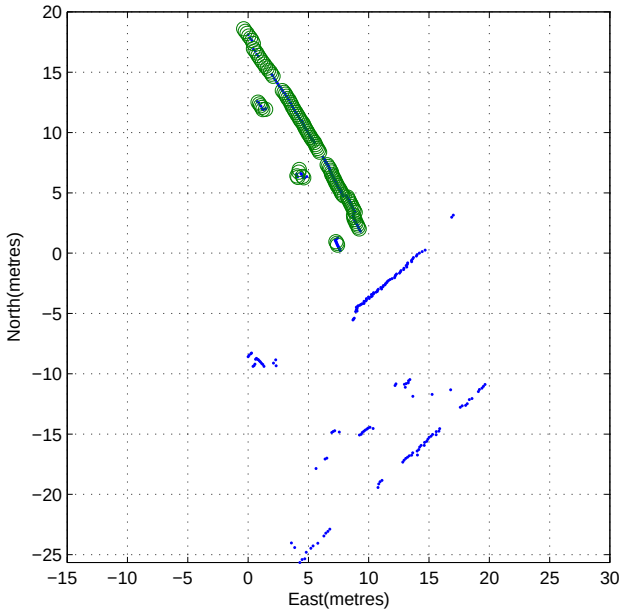
[21] G. Weiß, C. Wetzler, and E. Puttkamer. Keeping track of position and orientation of moving indoor systems by correlation of range-finder scans. In *International Conference on Intelligent Robots and Systems*, pages 595–601, 1994.

[22] Q. Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder (improves camera calibration). In *IEEE International Conference on Intelligent Robots and Systems*, 2004.

[23] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.




(a)




(b)

Fig. 17. Figure (a) shows the Scan-SLAM result obtained in the car park area. The solid line denotes the trajectory estimated. The light points represent a laser-image obtained using feature-based SLAM and GPS. The dark points represent the template scans position and the ellipses the $1\sigma$ covariance bounds. (b) shows a result of the alignment between a scan-landmark and a new laser scan. The scan landmark template is denoted by 'o' and the laser points by '.'.

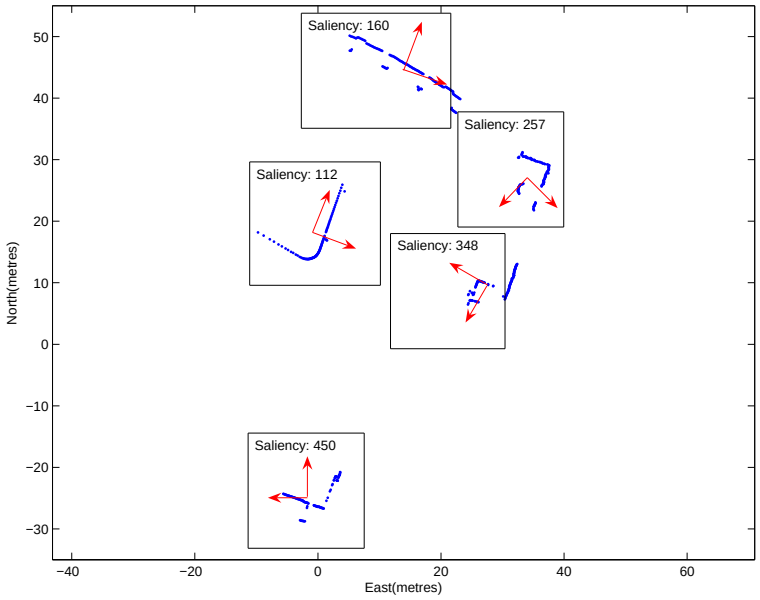



Fig. 18. The figure shows the local coordinate systems, the landmark template and its saliency score for the scan-landmarks incorporated. A minimum saliency of 100 was required to accepted a segment as a landmark.