Scan-SLAM: Combining EKF-SLAM and Scan Correlation

Juan Nieto, Tim Bailey and Eduardo Nebot

1

ARC Centre of Excellence for Autonomous Systems (CAS) The University of Sydney, NSW, Australia {j.nieto,tbailey,nebot}@acfr.usyd.edu.au

Summary. This paper presents 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 with scan correlation. Instead of geometric models, landmarks are defined by templates composed of raw sensed data, and scan correlation is shown to produce landmark observations compatible with the standard EKF-SLAM framework. The resulting *Scan-SLAM* combines the general applicability of scan correlation with the established advantages of an EKF implementation: recursive data fusion that produces a convergent map of landmarks and maintains an estimate of uncertainties and correlations. Experimental results are presented which validate the algorithm.

Key words: Simultaneous localisation and mapping (SLAM), EKF-SLAM, scan correlation, Sum of Gaussians (SoG), observation model

1.1 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) 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 [11]. 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. 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 [6]. One problem with EKF-SLAM is that requires the sensed data to be modelled as geometric shapes, which limits the approach to environments suited to such models.

This paper presents a new approach to SLAM which is based on the integration of scan correlation methods with the EKF-SLAM framework. The map is constructed as an on-line 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 geometric models, but rather are defined by a template 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* that uses raw data to represent landmarks and scan matching to produce landmark observations. In essence, this approach presents a new way to define generic observation 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 1.3 describes a *sum of Gaussians* (SoG) representation for scans of range-bearing data, based on the work presented in [1, Chapter 4]. Section 1.4 presents a method for obtaining a Gaussian likelihood function from the scan correlation procedure, which produces an observation in a form compatible with EKF-SLAM. Section 1.5 describes the generic observation model that is applied to all feature observations obtained by scan correlation. Section 1.6 presents *Scan-SLAM*, which uses the developments of Sections 1.4 and 1.5 to implement a scan correlation based observation update step within the EKF-SLAM framework. The method is validated with experimental results. Finally, conclusions are presented in Section 1.8.

1.2 Related Work

A significant issue with EKF-SLAM [4] 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 observation 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 geometric 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 primarily been used as a localisation mechanism from an *a priori* map [13, 8, 3, 9], with the *iterated closest point* (ICP) algorithm [2, 10] and occupancy grid correlation [5] being the most popular correlation methods.

Two significant methods have been presented that perform scan correlation based SLAM. The first [12] 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 [7] accumulates a selected history of scans, and aligns them as a network. This approach is based on the algorithm presented in [10].

The main concern with these 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 EKF-SLAM with scan correlation methods.

1.3 Scan Matching using Gaussian Sum Representation

A set of point measurements may be represented as a sum of Gaussians. This representation permits efficient correlation of two scans of data, and has a Bayesian justification which ensures that, under certain conditions, the scan alignment estimate is consistent (see [1]). SoG correlation also avoids limitations inherent to occupancy grid and ICP correlation methods; these being fixed-scale granularity and point-to-point data associations, respectively.

For a set of range-bearing measurements, such as a range-laser scan, the measurements and their uncertainties are first converted to sensor-centric Cartesian space. That is, a range-bearing measurement $\mathbf{z}_i = (r_i, \theta_i)$ with Gaussian uncertainty \mathbf{R}_i , is converted to Cartesian coordinates

$$\mathbf{x}_{i} = \mathbf{f} \left(\mathbf{z}_{i} \right) = \begin{bmatrix} r_{i} \cos \theta_{i} \\ r_{i} \sin \theta_{i} \end{bmatrix}$$
$$\mathbf{P}_{i} = \nabla \mathbf{f}_{\mathbf{z}_{i}} \mathbf{R}_{i} \nabla \mathbf{f}_{\mathbf{z}_{i}}^{T}$$

where the Jacobian $\nabla \mathbf{f}_{\mathbf{z}_i} = \frac{\partial \mathbf{f}}{\partial \mathbf{z}_i}$. We define an *n*-dimensional Gaussian as

$$g(\mathbf{x}; \bar{\mathbf{p}}, \mathbf{P}) \triangleq \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{p}})^T \mathbf{P}^{-1}(\mathbf{x} - \bar{\mathbf{p}})\right)$$

where $\bar{\mathbf{p}}$ and \mathbf{P} are the mean and covariance, respectively. An *n*-dimensional sum of Gaussians (SoG) is defined as the sum of k scaled Gaussians.

$$G(x) \triangleq \sum_{i=1}^{k} \alpha_i g(\mathbf{x}; \bar{\mathbf{p}}_i, \mathbf{P}_i)$$



Fig. 1.1. The left-hand figure shows the set of raw range-laser data points transformed to a sensor-centric coordinate frame. The right-hand figure shows the SoG representation of this scan.

where, for a *normalised* SoG, the sum of the scaling factors α_i is one. However, normalisation is not necessary for correlation purposes—only relative scale is important—and it is more convenient to work with non-normalised SoGs. An example of a SoG produced from a range-laser scan is shown in Fig. 1.1.

Given two scans of Cartesian data points, where each point has a mean and variance, the respective scans may be represented by two SoGs.

$$G_1(\mathbf{x}) = \sum_{i=1}^{k_1} \alpha_i g(\mathbf{x}; \bar{\mathbf{p}}_i, \mathbf{P}_i)$$
$$G_2(\mathbf{x}) = \sum_{i=1}^{k_2} \beta_i g(\mathbf{x}; \bar{\mathbf{q}}_i, \mathbf{Q}_i)$$

A likelihood function for the correlation of these two SoGs is given by their cross-correlation.

$$\begin{split} \Lambda(\mathbf{x}) &= G_1(\mathbf{x}) \star G_2(\mathbf{x}) \\ &= \int \sum_{i=1}^{k_1} \alpha_i g(\mathbf{u} - \mathbf{x}; \bar{\mathbf{p}}_i, \mathbf{P}_i) \sum_{j=1}^{k_2} \beta_j g(\mathbf{u}; \bar{\mathbf{q}}_j, \mathbf{Q}_j) d\mathbf{u} \\ &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \alpha_i \beta_j \gamma_{ij}(\mathbf{x}) \end{split}$$
(1.1)

where $\gamma_{ij}(\mathbf{x})$ is the cross-correlation of two Gaussians $g(\mathbf{x}; \bar{\mathbf{p}}_i, \mathbf{P}_i)$ and $g(\mathbf{x}; \bar{\mathbf{q}}_j, \mathbf{Q}_j)$.

$$\gamma_{ij}\left(\mathbf{x}\right) = \frac{1}{\sqrt{(2\pi)^{n} \left|\mathbf{\Sigma}\right|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mu})^{T} \mathbf{\Sigma}^{-1}(\mathbf{x}-\bar{\mu})\right)$$

$$ar{\mu} = ar{\mathbf{p}}_i - ar{\mathbf{q}}_j$$
 $\boldsymbol{\Sigma} = \mathbf{P}_i + \mathbf{Q}_i$

The result is a likelihood function that provides a measure of scan alignment, and a maximum-likelihood alignment can be obtained as

$$\mathbf{x}_M = \arg\max \Lambda(\mathbf{x}) \tag{1.2}$$

where pose \mathbf{x}_M is the maximum-likelihood location of scan 1 with respect to scan 2. More precisely, \mathbf{x}_M is the location of the coordinate frame of scan 1 with respect to the coordinate frame of scan 2.

Full details of Gaussian sum correlation can be found in [1, Section 4.4]. In particular, it describes SoG scaling factors, SoG correlation in a plane (with alignment over position *and* orientation), and various alternatives for efficient implementation.

1.4 Scan Correlation Variance

For scan correlation to be compatible with EKF-SLAM, it is necessary to approximate the correlation likelihood function in (1.1) by a Gaussian. This section presents a method to compute a mean and variance for scan correlation based on the shape of the likelihood function in the vicinity of the point of maximum-likelihood. The resulting approximation is reasonable because the likelihood function tends to be Gaussian in shape in the region close to the maximum-likelihood.

The first step in deriving this approximation is to compute the variance of a Gaussian function given only a set of point evaluations of the function. Given a Gaussian PDF

$$g(\mathbf{x}; \bar{\mathbf{p}}, \mathbf{P}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{p}})^T \mathbf{P}^{-1}(\mathbf{x} - \bar{\mathbf{p}})\right)$$

the maximum-likelihood value is found at its mean

$$g(\bar{\mathbf{p}}; \bar{\mathbf{p}}, \mathbf{P}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}|}} = C_M$$

Any other sample \mathbf{x}_i from this distribution will have the value

$$g(\mathbf{x}_i; \bar{\mathbf{p}}, \mathbf{P}) = C_M \exp\left(-\frac{1}{2}(\mathbf{x}_i - \bar{\mathbf{p}})^T \mathbf{P}^{-1}(\mathbf{x}_i - \bar{\mathbf{p}})\right)$$
$$= C_i$$

By taking logs and rearranging terms, we get

$$(\mathbf{x}_i - \bar{\mathbf{p}})^T \mathbf{P}^{-1} (\mathbf{x}_i - \bar{\mathbf{p}}) = -2(\ln C_i - \ln C_M)$$
(1.3)

Thus, given a set of samples $\{\mathbf{x}_i\}$ and their associated function evaluations $\{C_i\}$, along with the maximum-likelihood parameters $\bar{\mathbf{p}}$ and C_M , then the inverse covariance matrix \mathbf{P}^{-1} (and hence \mathbf{P}) can be evaluated. The only requirement is that the number of samples equals the number of unknown elements in \mathbf{P}^{-1} .

In this paper, we are concerned with 3-dimensional Gaussians, (i.e., to represent the distribution of a landmark pose $[x_L, y_L, \phi_L]^T$), and so we present the full derivation of variance evaluation for this case. We define the following variables

$$C'_{i} = -2(\ln C_{i} - \ln C_{M})$$
$$\mathbf{x}_{i} - \bar{\mathbf{p}} = [x_{i}, y_{i}, z_{i}]^{T}$$
$$\mathbf{P}^{-1} = \begin{bmatrix} a \ b \ c \\ b \ d \ e \\ c \ e \ f \end{bmatrix}$$

Substituting these into (1.3) and expanding terms gives

$$x_i^2 a + 2x_i y_i b + 2x_i z_i c + y_i^2 d + 2y_i z_i e + z_i^2 f = C_i^{\prime}$$
(1.4)

The result is an equation with six unknowns (a, \ldots, f) , and so a solution can be found given six samples from the Gaussian. This is posed as a matrix equation of the form

$$Ax = b$$

where the *i*-th row of **A** is $[x_i^2, 2x_iy_i, 2x_iz_i, y_i^2, 2y_iz_i, z_i^2]$, **x** is the unknowns $[a, b, c, d, e, f]^T$, and **b** is the set of solutions $\{C'_i\}$. For a Gaussian function, the solution of this system of equations gives the exact covariance matrix of the function.

Since the scan correlation likelihood function is not exactly Gaussian (although we presume it has approximately Gaussian shape near the maximum likelihood location), different sets of samples will produce different values for **P**. To reduce this variation, we evaluate more than the minimum number of samples and compute a least-squares solution using *singular value decomposition* (SVD), which results in a much more stable covariance estimate.

In summation, two SoGs are aligned according to a maximum likelihood correlation, to give the pose \mathbf{x}_M between scan coordinate frames. A number of samples, N > 6, from the region near \mathbf{x}_M are evaluated and the alignment variance \mathbf{P}_M is computed by SVD. At a higher level of abstraction, the result of this algorithm can be described by the following pseudocode function interface

$$[\mathbf{x}_M, \mathbf{P}_M]$$
 = scan_align($G_1(\mathbf{x}), G_2(\mathbf{x}), \mathbf{x}_0)$

where \mathbf{x}_0 is an initial guess of the pose of $G_1(\mathbf{x})$ relative to $G_2(\mathbf{x})$. A good initial pose is required to promote reliable convergence.

1.5 Generic Observation Model

We define a landmark by a SoG in a local landmark coordinate frame, and the Scan-SLAM map stores a global pose estimate of this coordinate frame in its state vector (see Section 1.6 for details). Thus, all observations of landmarks obtained by scan matching can be modelled as the measurement of a global landmark frame \mathbf{x}_L as seen from the global vehicle pose \mathbf{x}_v (see Fig. 1.2).

The generic observation model for the pose of a landmark coordinate frame with respect to the vehicle is as follows.

$$\mathbf{z} = [x_{\delta}, y_{\delta}, \phi_{\delta}]^{T} = \mathbf{h} (\mathbf{x}_{L}, \mathbf{x}_{v})$$

$$= \begin{bmatrix} (x_{L} - x_{v}) \cos \phi_{v} + (y_{L} - y_{v}) \sin \phi_{v} \\ -(x_{L} - x_{v}) \sin \phi_{v} + (y_{L} - y_{v}) \cos \phi_{v} \\ \phi_{L} - \phi_{v} \end{bmatrix}$$
(1.5)

1.6 Scan-SLAM Update Step

When an object is observed for first time, a new landmark is created. A landmark definition template is created by extracting from the current scan the set of measurements that observe the object. These measurements form a SoG, which is transformed to a coordinate frame local to the landmark. 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, we define the local coordinate frame as the centroid of the template SoG. A new landmark is added to the SLAM map by adding the global pose of its coordinate frame 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.

As new scans become available, the SLAM estimate of existing map landmarks can be updated by the following process. First, the location of a map



Fig. 1.2. All SoG 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 $\mathbf{z} = [x_{\delta}, y_{\delta}, \phi_{\delta}]^T$.

landmark relative to the vehicle is predicted to determine whether the landmark template SoG $G_L(\mathbf{x})$ is in the vicinity of the current scan SoG $G_o(\mathbf{x})$. This vehicle-relative landmark pose is the predicted observation $\hat{\mathbf{z}}$ according to (1.5). If the predicted location is sufficiently close to the current scan, the landmark template is aligned with the scan using the SVD correlation algorithm, using $\hat{\mathbf{z}}$ as an initial guess (see Fig. 1.3).

$$[\mathbf{z}, \mathbf{R}] = \operatorname{scan}_{\operatorname{align}}(G_L(\mathbf{x}), G_o(\mathbf{x}), \hat{\mathbf{z}})$$

The result of scan alignment gives the pose of the landmark template frame with respect to the current scan coordinate frame, which is defined by the current vehicle pose. This is the new landmark observation \mathbf{z} with uncertainty \mathbf{R} , in accordance with the generic observation model in (1.5). Having obtained the observation \mathbf{z} and \mathbf{R} , the SLAM state is updated in the usual manner of EKF-SLAM.



Fig. 1.3. The top figure shows a stored scan landmark template (solid line) and a new observed scan (dashed line). The bottom figure shows the scan alignment evaluated with the scan correlation algorithm from which the observation vector z is obtained.

1.7 Results

This section presents 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.

Fig. 1.4 shows the simulation environment. The experiment was done in a large area of 180 by 160 metres with a sensor field of view of 30 metres. The vehicle travels at a constant speed of 3m/s. The sensor observations are corrupted with Gaussian noise with standard deviations 0.1 metres in range and 1.5 degrees in bearing. The simulation map consists of objects with different geometry and size. In order to select the segments to be added into the navigation map, a basic segmentation algorithm was implemented that selects sensor segments that contain a minimum number of neighbour points.

The results for the Scan-SLAM algorithm are shown in Fig. 1.4. Here the solid line depicts the ground truth for the robot pose and the dashed line the estimated vehicle path. The actual object's position is represented by the light solid line and the segment's position by the dark points. The local axis pose for each scan landmark is also shown and the ellipses indicate the 3σ 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. Fig. 1.5 shows the result after the vehicle closes the loop and the EKF-SLAM updates the map. The alignment between the actual object's position and that estimated by the algorithm after closing the loop can be observed.



Fig. 1.4. The figure 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 3σ uncertainty bound of each scan landmark.

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 near the ACFR building. The environment is mainly dominated by buildings and trees. Fig. (1.6) illustrates the result obtained with the algorithm. The solid line denotes the trajec-



Fig. 1.5. Simulation result after closing the loop.

tory estimated. The light points is a laser-image obtained using feature-based SLAM and GPS, which can be used as a reference. The dark points represent the template scans and the ellipses the 1σ covariance bound. The local axis for the scan landmarks were also drawn in the figure. The segmentation criterion was also based on distance. Seven scan landmarks were incorporated and used for the SLAM.

1.8 Conclusions and Future Work

EKF-SLAM is currently the most popular filter used to solve stochastic SLAM. 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 be 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



Fig. 1.6. Scan-SLAM result obtained in the car park area. The solid line denotes the trajectory estimated. The light points is a laser-image obtained using feature-based SLAM and GPS. The dark points represent the template scans and the ellipses the 1σ covariance bound.

need to store the scan history. It works as an EKF-SLAM that uses raw data as landmarks and utilises scan correlation algorithms to produce landmark observations. Finally experimental results were presented that showed the efficacy of the new algorithm.

In terms of future research, there is a lot of scope for developing the data association capabilities that arise from combining scan correlation with EKF-SLAM metric constraints. The ability of batch data association within an EKF framework to reject spurious data is well developed (e.g., [1, Chapter 3]). Scan correlation has the potential to strengthen the rejection of outliers by matching consecutive sequences of scans and removing points that are not reobserved. Also, scan correlation provides a measure of how well the *shape* of one scan fits another and can reject associations that have compatible metric constraints but misfitting shape.

A second area for future work is the development of continuously improving landmark templates. As a landmark is reobserved, perhaps from different view-points, the template model that describes it can be refined to better represent the object.

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

- T. Bailey. Mobile Robot Localisation and Mapping in Extensive Outdoor Environments. PhD thesis, University of Sydney, Australian Centre for Field Robotics, 2002.
- 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.
- 3. 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.
- 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.
- 5. A. Elfes. Occupancy grids: A stochastic spatial representation for active robot perception. In Sixth Conference on Uncertainty in AI, 1990.
- 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.
- 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.
- 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.
- K. Konolige and K. Chou. Markov localization using correlation. In International Joint Conference on Articial Intelligence, pages 1154–1159, 1999.
- 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.
- 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.
- 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.
- G. Weiß, C. Wetzler, and E. Puttkamer. Keeping track of position and orientation of moving indoor systems by correlation of range-finder scans. In *Interna*tional Conference on Intelligent Robots and Systems, pages 595–601, 1994.