Constrained Initialisation for Bearing-Only SLAM

Tim Bailey
Australian Centre for Field Robotics
University of Sydney, Australia
tbailey@acfr.usyd.edu.au

***Abstract*— Simultaneous Localisation And Mapping (SLAM) is a stochastic map building method which permits consistent robot navigation without requiring an *a priori* map. The map is built incrementally as the robot observes the environment with its on-board sensors and, at the same time, is used to localise the robot. Typically, SLAM has been performed using range-bearing sensors, but the development of a SLAM implementation using only bearing measurements is desirable as it permits the use of sensors such as CCD cameras, which are small, reliable and cheap.**

However, bearing-only SLAM is hindered by the feature initialisation problem, where the estimated location of a new map landmark cannot be determined from a single measurement, and combined information from multiple measurements may be ill-conditioned.

This paper presents a solution to the feature initialisation problem called constrained initialisation, which defers the use of sensor information until initialisation becomes well-conditioned. Measurements may be used out-of-sequence and all the available information can be incorporated without inconsistency. Furthermore, this method operates within the conventional extended Kalman Filter (EKF) framework of the SLAM algorithm.

I. Introduction

While range-bearing SLAM has received a lot of attention in the recent literature (e.g., see [3], [8], [6]), little work has been presented regarding bearing-only SLAM due to the difficulty of feature initialisation. Nevertheless, bearing-only SLAM is an attractive method as it permits the use of cheap vision sensors rather than the expensive laser hardware usually required for range-bearing implementation.

Bearing-only tracking has been discussed at some length in the target-tracking literature, particularly with regard to the insufficiency of the EKF to perform track initialisation. Various solutions have been presented including using batch maximum-likelihood initialisation [11], Gaussian sum filters [1] and Monte-Carlo filters [7]. Note, target tracking requires estimation of the target velocity as well as its position, and so is a harder problem than bearing-only SLAM, which obtains velocity information from dead-reckoning sensors. Thus, SLAM tends to be stable within the EKF framework once a moderately well-conditioned initial estimate for each feature is obtained.

The problem with feature initialisation is that a single measurement does not constrain the feature location, and at least two measurements are required. However, sequential pairs of measurements tend to possess insufficient base-line to give a well-conditioned location estimate. Non-linear batch initialisation methods offer a possible solution, but a critical aspect they fail to address is *when* to perform the operation. That is, they do not specify a measure for when the result will be well-conditioned, and simply requiring a large number of measurements is not sufficient. (For example, a batch solution based on many bearing measurements taken from the same location will be incorrect regardless of the estimation technique.)

A particularly interesting non-linear batch initialisation method uses *bundle adjustment* [5] with excellent results. The claim of this paper, though, is that good results are possible for bearing-only SLAM within the conventional EKF framework, provided an appropriate "condition" measure is used to signal when to perform feature initialisation.

The technique presented in this paper is based on the *constrained initialisation* procedure presented in [14] for range-bearing SLAM. A range-only (sonar) variant of this method is presented in [10], [13], and introduces a powerful concept to defer the use of measurement information by augmenting the SLAM state-vector with past vehicle pose estimates. Thus, measurements can be postponed for indefinite periods (limited by non-linearities), which is useful if data association is uncertain or if an initial feature estimate is ill-conditioned. In addition, the information may be used out-of-sequence with other measurements.

This paper presents a variant of constrained initialisation for bearing-only data, where past vehicle pose estimates are retained in the SLAM state so that feature initialisation can be deferred until their estimates become well-conditioned. Validation of this approach is shown by a simulated SLAM implementation, which demonstrates the consistency of constrained initialisation given known data association. The data association process, in fact, dovetails well with the deferred information approach, and a batch data association method based on these deferred constraints is an area of future research.

The format of this paper is as follows. The next section briefly describes the basic bearing-only SLAM algorithm for the case where the features have already been initialised. Section III discusses the bearing-only

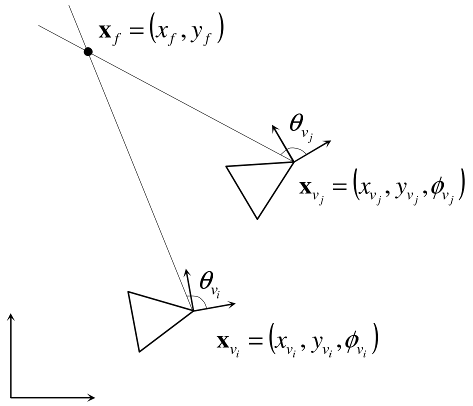

Fig. 1. Feature initialisation via the intersection of two bearing measurements.

initialisation problem in detail and Section IV presents the constrained initialisation algorithm. Section V introduces a statistic of "Gaussianness" to determine whether a feature initialisation is well-conditioned. Section VI presents simulation results using this algorithm and Section VII discusses their implications. The final section concludes with some interesting future extensions to this work.

II. Bearing-Only SLAM Algorithm with Initialised Features

The basic algorithm for range-bearing SLAM is well known and can be found, for example, in [8], [6]. With regard to notation, the augmented SLAM state $\mathbf{x}_a = [\mathbf{x}_v^T, \mathbf{x}_{f_1}^T, \ldots, \mathbf{x}_{f_n}^T]^T$ is defined by the vehicle pose $\mathbf{x}_v = [x_v, y_v, \phi_v]^T$ and the set of map features $\mathbf{x}_{f_i} = [x_{f_i}, y_{f_i}]^T$. The bearing-only algorithm, assuming the set of features is already initialised, is identical to the range-bearing algorithm except that the observation model for a feature $\mathbf{x}_{f_i}$ is simply

$$z_i = h(\mathbf{x}_v, \mathbf{x}_{f_i}) = \arctan\left(\frac{y_{f_i} - y_v}{x_{f_i} - x_v}\right) - \phi_v \qquad (1)$$

III. The Problem of Bearing-Only Feature Initialisation

The problem with bearing-only initialisation is that a single measurement is insufficient to determine the location of the feature, and at least two bearing measurements θ_{v_i} and θ_{v_j} from two different vehicle poses $\mathbf{x}_{v_i}$ and $\mathbf{x}_{v_j}$ are required. The location of the feature then is calculated as the intersection of two lines as shown in Figure 1.

$$\begin{aligned}\mathbf{x}_f &= \mathbf{g}\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right) \\ &= \begin{bmatrix} \frac{x_{v_i} s_i c_j - x_{v_j} s_j c_i + (y_{v_j} - y_{v_i}) c_i c_j}{s_i c_j - s_j c_i} \\ \frac{y_{v_j} s_i c_j - y_{v_i} s_j c_i + (x_{v_i} - x_{v_j}) s_i s_j}{s_i c_j - s_j c_i} \end{bmatrix}\end{aligned} \qquad (2)$$

where

$$s_i = \sin(\phi_{v_i} + \theta_{v_i})$$
$$c_i = \cos(\phi_{v_i} + \theta_{v_i})$$

If the bearing and pose estimates were perfectly known, the feature location would be trivial from Equation 2 but, as they are uncertain, the estimated feature location may be ill-conditioned depending on several factors: (i) the uncertainty of the pose estimates, (ii) the uncertainty of the bearing measurements, and (iii) the base-line afforded by the two vehicle locations. Furthermore, unless the correlations between the two vehicle pose estimates are included in the initialisation estimate, the estimated location may be inconsistent (i.e., the estimated uncertainty is less than its true uncertainty).

Yet another problem with bearing-only initialisation is that data association cannot be uniquely determined from less than three measurements. That is, any two bearing estimates may intersect to give a nominal feature location, and association constraint is only obtained through bearing triplets. The data association problem, and the related issue of tentative features in clutter, are not discussed further in this paper, but might also be addressed by the constrained initialisation concept described below.

IV. Bearing-Only Constrained Initialisation Algorithm

The key operation of constrained initialisation is to store past vehicle pose estimates in the SLAM state as follows.

$$\mathbf{x}_a = [\mathbf{x}_v^T, \mathbf{x}_{v_m}^T, \ldots, \mathbf{x}_{v_1}^T, \mathbf{x}_{f_1}^T, \ldots, \mathbf{x}_{f_n}^T]^T \qquad (3)$$

Each pose estimate $\mathbf{x}_{v_i}$ corresponds to the time and location where a set of measurements $\left\{\theta_{v_i}^1, \ldots, \theta_{v_i}^k\right\}$ was obtained. (It is assumed that these bearing measurements are uncorrelated to the vehicle and to each other, and so may be accumulated in a separate observation vector.)

Thus, measurements of a given feature may be accumulated over time until sufficient information is available to initialise it in the map. At this point, all the available measurement information can be constrained to give a best location estimate. These constraints might be applied simultaneously, as a batch, but this tends to perform badly in practice for the standard EKF. For a conventional EKF implementation, the quality of the solution depends on the near-linearity of the observation model which, for the highly non-linear model in Equation 1, means that the estimate residual (or innovation) must be small. In particular, bearing-only SLAM is sensitive to the order in which measurements are processed. To avoid an inconsistent update, measurements are processed in the following order:

1) Measurements of existing map features are processed first at each timestep. They are best processed

as a batch update as this appears to be a more linear calculation than if they are applied sequentially.

2) If, for any non-initialised feature, there exists within its set of accumulated measurements a well-conditioned measurement pair (see below), this pair is used to provide an initial feature estimate.
3) For each newly initialised feature, the remaining accumulated measurements are applied, again via a single batch update.

Even though these observation constraints are all available at the same time, applying them in the order described means that the estimate error is reduced at each stage for the more sensitive updates in the stages below.

In order to permit steps 2 and 3 above, it is necessary to perform two operations. The first is to augment the state vector with pose information at each timestep where measurements are obtained and deferred. Let $\hat{\mathbf{x}}_v$ be the current pose estimate and $\hat{\mathbf{x}}_m$ be the rest of the map, the pose is stored as follows.

$$[\hat{\mathbf{x}}_v^T, \hat{\mathbf{x}}_m^T]^T \longrightarrow [\hat{\mathbf{x}}_v^T, \hat{\mathbf{x}}_v^T, \hat{\mathbf{x}}_m^T]^T \quad (4)$$

$$\begin{bmatrix} \mathbf{P}_{vv} & \mathbf{P}_{vm} \\ \mathbf{P}_{vm}^T & \mathbf{P}_{mm} \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{P}_{vv} & \mathbf{P}_{vv} & \mathbf{P}_{vm} \\ \mathbf{P}_{vv} & \mathbf{P}_{vv} & \mathbf{P}_{vm} \\ \mathbf{P}_{vm}^T & \mathbf{P}_{vm}^T & \mathbf{P}_{mm} \end{bmatrix} \quad (5)$$

Thus, the vehicle pose at a given timestep is retained, along with its correlations to other stored poses and the existing map.

The second operation is to calculate whether a pair of measurements for an uninitialised feature is well-conditioned. The feature estimate for a given pair may be considered well-conditioned if the true uncertainty distribution for the feature location is well approximated by a Gaussian, which implies that the transformation from bearing-space to Cartesian-space is near-linear. If a measurement pair is found to be well-conditioned, the new feature is added to the state vector[1] via Equation 2. Subsequently, the remaining accumulated measurements for this feature can be applied as ordinary observations. And, finally, the obsolete stored pose estimates can be deleted from the state (i.e., those stored poses that no longer possess associated measurements).

V. A Statistic of Gaussianness

A new feature is considered well-conditioned if the true probability density function (PDF) of its location closely resembles the Gaussian approximation obtained from a Jacobian-based linearised transform. The statistic used to compare these two distributions is the *Kullback-Leibler distance* or relative entropy [4].

In this section, first the linearised and then the true distributions are derived, and finally the relative entropy approach to comparing them is explained.

Given two vehicle pose estimates $\mathbf{x}_{v_i}$ and $\mathbf{x}_{v_j}$, and their associated bearing measurements θ_{v_i} and θ_{v_j} of a single landmark, the transform to initialise the feature estimate is as follows.

$$\begin{bmatrix} \mathbf{x}_{v_i} \\ \mathbf{x}_{v_j} \\ \mathbf{x}_f \end{bmatrix} = \mathbf{f}\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right) = \begin{bmatrix} \mathbf{x}_{v_i} \\ \mathbf{x}_{v_j} \\ \mathbf{g}\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right) \end{bmatrix} \quad (6)$$

where $\mathbf{g}\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right)$ is given in Equation 2. The Jacobian of this transform $\nabla\mathbf{f}_\mathbf{x} = \frac{\partial \mathbf{f}}{\partial(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j})}$ permits a linearised transformation of the covariance matrix of the form $\mathbf{P}_t = \nabla\mathbf{f}_\mathbf{x}\mathbf{P}\nabla\mathbf{f}_\mathbf{x}^T$, where $\mathbf{P}$ represents the covariance matrix of the vehicle states and bearing measurements, and $\mathbf{P}_t$ is of the vehicle states and feature location.

The true (algebraic) density transformation for the non-linear function in Equation 6 is given by

$$p\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \mathbf{x}_f\right) = p\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right) |\det(\nabla\mathbf{h}_{\mathbf{x}_f})| \quad (7)$$

where $p(\cdot)$ represents the joint PDF of the enclosed parameters, and the Jacobian $\nabla\mathbf{h}_{\mathbf{x}_f} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_f}$ is the partial derivative matrix of Equation 8 with respect to the feature location $\mathbf{x}_f$.

$$\begin{bmatrix} \theta_{v_i} \\ \theta_{v_j} \end{bmatrix} = \mathbf{h}\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \mathbf{x}_f\right) = \begin{bmatrix} \arctan\left(\frac{y_f - y_{v_i}}{x_f - x_{v_i}}\right) - \phi_{v_i} \\ \arctan\left(\frac{y_f - y_{v_j}}{x_f - x_{v_j}}\right) - \phi_{v_j} \end{bmatrix} \quad (8)$$

Note, the PDF $p\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right)$ is an 8-dimensional Gaussian for the two vehicle states and two bearing measurements, but the θ terms are substituted with the functions in Equation 8 such that the distribution becomes a non-Gaussian function of $(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \mathbf{x}_f)$. This non-Gaussian function is then scaled by

$$|\det(\nabla\mathbf{h}_{\mathbf{x}_f})| = \left|\frac{x_{v_i}(y_f - y_{v_j}) + x_{v_j}(y_{v_i} - y_f) + x_f(y_{v_j} - y_{v_i})}{[(x_f - x_{v_i})^2 + (y_f - y_{v_i})^2][(x_f - x_{v_j})^2 + (y_f - y_{v_j})^2]}\right| \quad (9)$$

to produce the joint PDF of the vehicle states and initial Cartesian feature location.

The relative entropy $D(f||g)$ "is a measure of the inefficiency of assuming that the distribution is $g(\mathbf{x})$ when the true distribution is $f(\mathbf{x})$" [4, page 18]. It is not a

[1] Further information on how to initialise a feature using Jacobians, so as to determine its covariance and correlations with the existing state, can be found in [2, Section 2.2.4]. This reference describes state augmentation for range-bearing measurements, but the same approach applies to bearing-bearing initialisation.

true distance metric, as it is not symmetric and does not satisfy the triangle inequality, but can still provide a measure of how close two distributions are, such that a threshold might be applied for when they are considered near enough. Relative entropy is defined as follows

$$D(f||g) = \int_{-\infty}^{\infty} f(\mathbf{x}) \ln \frac{f(\mathbf{x})}{g(\mathbf{x})} d\mathbf{x} \tag{10}$$

which has the property that $D(f||g) \geq 0$, with equality to zero if and only if the two distributions are identical.

The relative entropy is used here to compare the analytical transformed PDF from Equation 7 with the linearised Gaussian approximation. Presently, a closed-form solution to this calculation is not known and, instead, a Monte Carlo solution is used. That is, the *sample relative entropy* is obtained by first sampling $\{\mathbf{x}^1, \ldots, \mathbf{x}^n\}$ from the true distribution $f(\mathbf{x})$ and calculating the approximate relative entropy as

$$D(f||g) \approx \frac{1}{n} \sum_{k=1}^{n} \left[\ln f(\mathbf{x}^k) - \ln g(\mathbf{x}^k)\right] \tag{11}$$

For the analytical PDF from Equation 7, and its linearised Gaussian approximation, the sample relative entropy is calculated via the following steps.

1) Sample from the 8-D Gaussian distribution $p\left(\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}\right)$, and record the likelihood of the Gaussian at each sample point. Thus, we have a set of samples $\{\mathbf{x}^1, \ldots, \mathbf{x}^n\}$, where each $\mathbf{x}^k = [x^k_{v_i}, y^k_{v_i}, \phi^k_{v_i}, x^k_{v_j}, y^k_{v_j}, \phi^k_{v_j}, \theta^k_{v_i}, \theta^k_{v_j}]^T$, and a set of weights $\{w_1, \ldots, w_n\}$.
2) Transform each sample $\mathbf{x}^k$ using Equation 6.
3) Scale each sample weight using Equation 9 such that $w_k = w_k \cdot |\det(\nabla \mathbf{h}_{\mathbf{x}_f})|$. Note, for each w_k, the Jacobian $\nabla \mathbf{h}_{\mathbf{x}_f}$ is evaluated at the sample value, which means that Equation 9 is evaluated as a function of $(\mathbf{x}^k_{v_i}, \mathbf{x}^k_{v_j}, \mathbf{x}^k_f)$.
4) Compute the linearised Gaussian, with covariance matrix $\mathbf{P}_t = \nabla \mathbf{f}_\mathbf{x} \mathbf{P} \nabla \mathbf{f}_\mathbf{x}^T$, from the known mean and covariance of $[\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \theta_{v_i}, \theta_{v_j}]^T$ and Equation 6.
5) For each sample $\mathbf{x}^k = [\mathbf{x}^k_{v_i}, \mathbf{x}^k_{v_j}, \mathbf{x}^k_f]^T$ calculate the weight v_k for that point on the Gaussian distribution.
6) Compute the sample relative as $D(true||gauss) \approx \frac{1}{n}\sum_{k=1}^{n}(\ln w_k - \ln v_k)$.

Computing the sample relative entropy in 8-D is expensive and imprecise but, while not ideal, it appears from the simulation results below to provide an adequate measure for when feature initialisation becomes well-conditioned.

VI. Simulation

For this problem, assuming known data association, simulation is in fact more instructive than real data results. Simulation provides ground truth and true knowledge of the noise statistics.

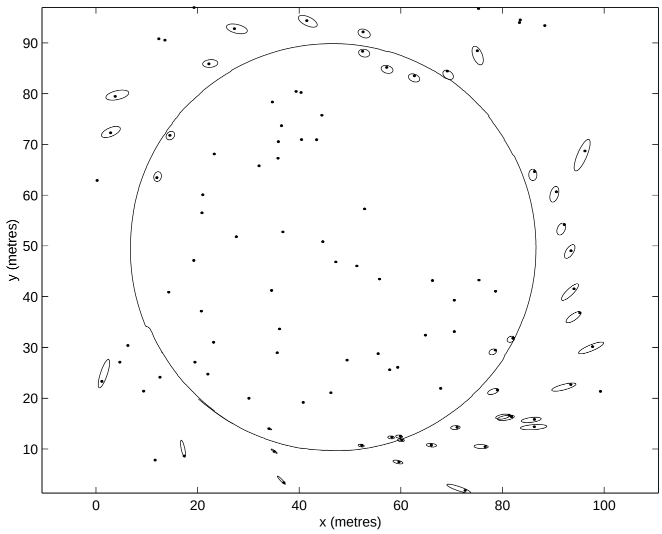

Fig. 2. Simulation results showing estimated trajectory, true feature locations and 3σ covariance ellipses of the estimated feature locations.

In this experiment, the sensor field-of-view is ± 30 degrees[2] with a maximum range of 30 metres. For simplicity, it is assumed that the bearing measurement uncertainty is constant with $1/2$ a degree standard deviation. The simulated environment consists of 100 features randomly distributed within a 100 by 100 metre region. The vehicle starts at pose $(20m, 20m, -0.8rad)$ and travels with nominal speed and steer angle of $2m/s$ and $0.05rad$, respectively. These nominal control values are corrupted with Gaussian noise with standard deviations $0.2m/s$ and $0.01rad$, respectively, for each $0.1s$ time interval.

The results for the constrained initialisation SLAM are shown in Figure 2. Here, the circle depicts the estimated vehicle path, the points represent the true feature locations (including non-observed features), and the ellipses indicate the mean and 3σ uncertainty bounds of each feature estimate after completing the loop. The errors in the vehicle pose estimate are indicated in Figure 3 which shows a plot of the error in the x-axis pose estimate and its 2σ uncertainty estimate. Notice the sporadic dips in the variance estimate, which are the result of deferred feature initialisation and subsequent update with all accumulated measurements.

Without the deferred measurements and check for well-conditioned initialisation, the SLAM algorithm diverges as soon as an initial mean is obtained that is sufficiently distant from the true feature location. This occurs almost immediately as sequential measurement pairs tend to be

[2]The limited field-of-view of a monocular camera is a significant handicap compared to, say, panoramic vision as the maximum information from bearing measurements occurs orthogonal to the vehicle motion. This would indicate that a monocular camera is best installed offset to the forward direction or on a pan-tilt head (active vision).

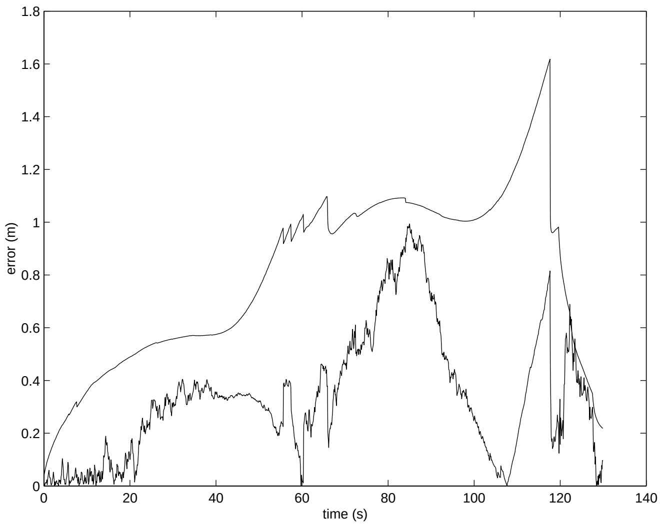

Fig. 3. Absolute error between the estimated x-coordinate location and the true value, shown with the 2σ uncertainty estimate.

very ill-conditioned and the chance of a distant mean estimate is very high. Constrained initialisation, on the other hand, performs appropriately and also retains the information from all intervening measurements.

The quality of the relative entropy statistic for initialisation, however, is questionable. Considerable tuning of the "near-Gaussian" threshold was required to obtain good initial feature values without failing to initialise the features at all. In the result shown in Figure 2, a number of features in the latter quarter of the loop were not initialised because no bearing-bearing pairs reached the required threshold. This indicates that the threshold should perhaps be a function of the total uncertainty for the PDF of $[\mathbf{x}_{v_i}, \mathbf{x}_{v_j}, \mathbf{x}_f]^T$.

VII. Discussion

Constrained initialisation is clearly a powerful technique for deferring measurement information and applying it at a later time. The simulation results show that bearing data can be deferred and incorporated out-of-sequence in a consistent manner, without loss of information, due to the maintenance of correlations between past vehicle states.

The key to consistent bearing-only SLAM, however, is not information theoretic, but depends upon the feature initialisation procedure reasonably meeting the EKF assumptions of small-error, near-Gaussian estimates. Essentially, the initial feature estimate must be "near enough" to the true feature location. In practice, the duration of postponement is not critical, provided dead-reckoning is reasonably accurate, as all information is recovered up to the limits imposed by non-linearity. The critical issue is to ensure that initialisation is not performed too early and is ill-conditioned. For non-linear batch methods, this may require waiting until a combined set of measurements is sufficient to define a well-conditioned estimate. For the experiment in this paper, it involves delaying initialisation until a single bearing-bearing pair yields a stable solution.

This paper presents a preliminary investigation of a statistic for well-conditioned bearing-only feature initialisation. The statistic chosen is the relative entropy between the analytical PDF of the feature location and its linearised Gaussian approximation, based on the concept that a close match between these distributions indicates a near-linear transformation from measurement-space to feature-space. This statistic gave reasonable results over a limited set of conditions, but possessed several problems. First, it requires numerical Monte Carlo solution, which is imprecise and computationally expensive. Second, the interpretation of what is actually meant by imposing a threshold on relative entropy is not clear. Typically, in the literature, relative entropy is used to maximise the fit of a model to a distribution, not to provide a threshold for nearness of two PDFs. Also, from the simulation experiments, it appears that the optimal threshold may not be constant, but may need to increase when the analytical PDF has larger uncertainty. Plainly, as a statistic for bearing-only initialisation, the relative entropy requires substantial refinement in terms of computation and interpretation of the threshold.

VIII. Conclusions and Future Work

The constrained initialisation method presented in this paper permits reliable and consistent bearing-only SLAM within the statistical framework of the EKF. The key contribution of this approach is the ability to consistently defer information so that measurements may be used at a later time and out-of-sequence. The accumulated information may also be used to calculate whether a set of measurements are sufficiently well-conditioned to permit feature initialisation.

The relative entropy as a statistic of feature initialisation consistency appears to effectively gate out very ill-conditioned bearing-pairs, but is difficult to fine-tune for marginally well-conditioned pairs. Future research may examine the relationship between an optimal gate threshold and the uncertainty of the PDFs being compared. A further problem with the relative entropy statistic is its computation, which preferably would have a closed-form solution, or some more efficient numerical solution.

A reliable statistic for feature initialisation consistency is crucial for any recursive formulation of bearing-only SLAM, including those calculating initial estimates via non-linear optimisation. It is hoped that this paper will encourage investigation of alternative statistics that may produce better results than the relative entropy approach presented here. Some possible solutions might include: (i) sensitivity analysis of the feature location estimate due to variations in the vehicle pose values and bearing measurements, or (ii) examining clusterings of feature locations obtained from sets of bearing measurements from different poses. It is important to note that some traditional measures may not work due to the invalidity of near-Gaussian assumptions. For example, a test examined during this research was to compute the variance of the range estimates obtained from a bearing-pair. This test failed to detect very uncertain range estimates, even when using a non-linear approximation of the transformed distribution such as the unscented transform [9], as the range PDFs were highly non-Gaussian (and, thus, not well characterised by their second moment).

Finally, this paper presents simulation results with known data association. An important quality of the constrained initialisation method is its suitability to batch data association methods (e.g., [12], [2]) as it retains all correlation information to facilitate highly constrained joint-likelihood associations. A significant challenge for bearing-only SLAM will be to determine how to appropriately gate associations to non-initialised (deferred) features, as they will possess non-Gaussian distributions and will not obey the traditional χ^2 probability of acceptance models.

IX. REFERENCES

[1] D.L. Alspach and H.W. Sorenson. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, 17(4):439–448, 1972.

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

[3] J.A. Castellanos, J.D. Tardós, and G. Schmidt. Building a global map of the environment of a mobile robot: The importance of correlations. In *IEEE International Conference on Robotics and Automation*, pages 1053–1059, 1997.

[4] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley and Sons, 1991.

[5] M. Deans and M. Hebert. Experimental comparison of techniques for localization and mapping using a bearing-only sensor. In *International Symposium On Experimental Robotics*, 2000.

[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] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140(2):107–113, 1993.

[8] J. Guivant, E. Nebot, and H.F. Durrant-Whyte. Simultaneous localization and map bulding using natural features in outdoor environments. In *Sixth International Conference on Intelligent Autonomous Systems*, volume 1, pages 581–588, 2000.

[9] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.

[10] J.J. Leonard and R.J. Rikoski. Incorporation of delayed decision making into stochastic mapping. In *International Symposium On Experimental Robotics*, 2000.

[11] D. Lerro and Y. Bar-Shalom. Bearing-only tracking for maneuvering targets in clutter. *International Federation of Automatic Control*, pages 267–272, 1996.

[12] 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.

[13] R.J. Rikoski, J.J. Leonard, and Newman P.M. Stochastic mapping frameworks. In *IEEE International Conference on Robotics and Automation*, pages 426–433, 2002.

[14] S. Williams, G. Dissanayake, and H. Durrant-Whyte. Constrained initialisation of the simultaneous localisation and mapping algorithm. In *International Conference on Field and Service Robotics*, pages 315–330, 2001.