Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

A Method for Dealing with Assignment Ambiguity

Simon J. Julier, Jeffrey K. Uhlmann and David Nicholson

***Abstract*— Many landmark-based navigation systems suffer from the problem of assignment ambiguity: the navigation system receives a measurement from a landmark, but the identity of the landmark is not uniquely known. This uncertainty is frequently addressed by attempting to identify the landmark which caused the measurements. Two common approaches are the nearest neighbour and multiple hypothesis tracking methods. However, the nearest neighbour method is notoriously unreliable and the multiple hypothesis method can only be implemented in real time for a small number of hypotheses. In this paper we consider an alternative approach for assignment ambiguity which uses the *Covariance Union* (CU) algorithm. Every potential assignment is used to generate an estimate. CU is applied to the set of estimates. Providing the true assignment is included within the set, the unioned estimate is guaranteed to be consistent. We provide a theoretical development of CU, describe a method for computing its value, and illustrate its performance in a nonlinear navigation example.**

I. Introduction

Many navigation systems use landmarks of one form or another. It is assumed that the environment is populated by a set of landmarks and the system is provided with a map which specifies the locations of those landmarks. As the vehicle moves through the environment it uses sensors mounted on its body to detect a subset of landmarks. By matching the observed landmarks with those in the map (assigning identities to the observations), the navigation system can update the estimate of the platform's position. There are two main classes of landmarks: those which are uniquely identifiable and those which are not. An example of the first class is the system described by Hu [6]. Each landmark is a barcode which can be read by a laser scanner. An example of the second class is the system developed by Chenavier [3]. The platform uses a vision system to detect strong vertical features and associates them with a floor plan or model of the environment.

Of the two classes of features, the second are of most practical interest. The reason is that the navigation system could use naturally occurring features in the environment such as vertical edges [3], corners [9] or even features which have no set structure and only have the property that they are easy to detect and track [14]. As such, the environment does not have to be specially prepared and it allows the vehicle to operate in much more general environments. Potential applications include mobile autonomous vehicles for space exploration or navigation through a desert. However, although they are of greater practical interest, non-unique landmarks present a much greater data processing challenge. Because a landmark cannot be uniquely identified by a single observation, it can only be matched with a map through the use of contextual information such as previous estimates. In some cases this presents relatively little trouble. If the error in the platform position is small and / or the landmarks are widely separated, innovation-based gating methods [1] are sufficient to perform the assignment. However, in many practical cases these conditions do not hold true.

A number of different algorithms for handling the problem of assignment ambiguity have been developed. Almost all of them attempt to identify the correct landmark and reject information from incorrect landmarks. One of the simplest and oldest methods is to use the *nearest neighbour* algorithm [1]. Given a candidate set of landmarks, the assignment is made to the landmark with the smallest normalised innovation (Mahalanobis distance). However, this heuristic is known to be notoriously unreliable [10]. The major alternative is *multiple hypothesis tracking* (MHT) [1]. A hypothesis is constructed for each combination of assignments of measurements to targets and recursive hypothesis testing is used to identify which hypothesis is the correct one. Although the MHT method is theoretically rigorous and robust if unlimited computational resources are available, its tendency to demand the maintenance of an exponentially increasing number of hypotheses cannot generally be satisfied in practice [13]. To limit the proliferation of tentative pairings a common strategy is pruning — eliminate the hypotheses with lower probabilities and renormalise the weights on the remaining hypotheses [5]. However, by artificially limiting the set of hypotheses, this approach incurs a risk that the correct hypothesis may be rejected, thus causing the associated filter to diverge.

In this paper we describe an alternative approach to assignment ambiguity. Rather than attempt to assign the unique identity of a landmark, the approach is to try and find an estimate which yields the minimum mean squared error estimate with respect to all feasible landmarks. The approach is based on the algorithm known as Covariance Union (CU) [15]. Given a set of estimates of a state, CU finds the mean and covariance matrix which is consistent with all of the possible updates simultaneously. Assuming that the correct measurement-beacon assignment is included within the set, CU is guaranteed to produce a consistent estimate.

The structure of this paper is as follows. Section II introduces the notation and describes the problem. Section III describes the CU method for accommodating observation correlation uncertainty and Section IV describes how this algorithm can be used for the tracking assignment problem. In Section V an example application of the method is presented.

Simon J. Julier is with Naval Research Laboratory / ITT AES, 4555 Overlook Avenue SW, Washington, DC 20375, USA. Email: julier@ait.nrl.navy.mil

Jeffrey K. Uhlmann is with the Department of Computer Engineering & Computer Science, University of Missouri–Columbia, Columbia, MO 65211 USA. Email: uhlmann@cecs.missouri.edu

David Nicholson is with BAE SYSTEMS Advanced Technology Centre, FPC 267, PO Box 5, Filton, Bristol BS34 7QW, UK. Email: david.nicholson2@baesystems.com

Section VI summarises the results of the paper and draws conclusions.

II. Problem Statement

Suppose the system of interest is described by the discrete-time nonlinear equation:

$$\mathbf{x}(k) = \mathbf{f}\left[\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{v}(k-1), k-1\right]$$

where $\mathbf{x}(k)$ is the state vector at time step k, $\mathbf{u}(k-1)$ is the control input, $\mathbf{v}(k-1)$ is the process noise, and $\mathbf{f}[\cdot]$ is the discrete time state transition equation.

The observation model for the ith landmark is:

$$\mathbf{z}(k) = \mathbf{h}_i\left[\mathbf{x}(k), \mathbf{u}(k-1), \mathbf{w}(k), k\right],$$

where $\mathbf{z}(k)$ is the observation vector, $\mathbf{w}(k)$ is the observation noise vector and $\mathbf{h}_i[\cdot,\cdot,\cdot,\cdot]$ is the discrete time observation equation for the ith beacon. The noise vectors $\mathbf{v}(k-1)$ and $\mathbf{w}(k)$ are assumed to be zero-mean and uncorrelated with covariances $\mathbf{Q}(k)$ and $\mathbf{R}(k)$ respectively.

A. The Kalman Filter

Using the notation from [1] the estimate of $\mathbf{x}(i)$ at time step i using all observations up to time step j is $\hat{\mathbf{x}}(i \mid j)$ with covariance $\mathbf{P}(i \mid j)$. The estimate $(\hat{\mathbf{x}}(i \mid j), \mathbf{P}(i \mid j))$ is said to be *consistent* if

$$\mathbf{P}(i \mid j) - \mathrm{E}\left[\{\hat{\mathbf{x}}(i \mid j) - \mathbf{x}(i)\}\{\hat{\mathbf{x}}(i \mid j) - \mathbf{x}(i)\}^T\right] \geq \mathbf{0}.$$

The Kalman Filter (KF) proceeds according to the well-known two-step process of prediction followed by update. The prediction is given by

$$\begin{aligned}\hat{\mathbf{x}}(k \mid k-1) &= \mathbf{f}\left[\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{v}(k-1), k-1\right]\\ \mathbf{P}(k \mid k-1) &= \boldsymbol{\nabla}_x\mathbf{f}\,\mathbf{P}(k-1 \mid k-1)\,\boldsymbol{\nabla}_x^T\mathbf{f}\\ &\quad + \boldsymbol{\nabla}_v\mathbf{f}\,\mathbf{Q}(k-1)\,\boldsymbol{\nabla}_v^T\mathbf{f}\end{aligned}$$

If the beacon identity i is known or can unambiguously be determined, the update is given by standard Kalman filter equations:

$$\hat{\mathbf{x}}(k \mid k) = \hat{\mathbf{x}}(k \mid k-1) + \mathbf{W}(k)\,\boldsymbol{\nu}(k) \tag{1}$$

where

$$\begin{aligned}\boldsymbol{\nu}(k) &= \mathbf{z}(k) - \hat{\mathbf{z}}(k \mid k-1)\\ &= \mathbf{z}(k) - \mathbf{h}_i\left[\hat{\mathbf{x}}(k \mid k-1)\right], \text{ and}\end{aligned} \tag{2}$$

$$\mathbf{W}(k) = \mathbf{P}(k \mid k-1)\,\boldsymbol{\nabla}^T\mathbf{h}_i\mathbf{S}^{-1}(k) \tag{3}$$

where

$$\mathbf{S}(k) = \boldsymbol{\nabla}\mathbf{h}_i\,\mathbf{P}(k \mid k-1)\,\boldsymbol{\nabla}^T\mathbf{h}_i + \mathbf{R}(k).$$

B. Assignment Ambiguity

Suppose that the beacon cannot be unambiguously identified and at time step k a set $\mathcal{S}_k$ of s_k beacons have been identified as being potentially feasible matches with the observation (for example the landmarks are visible or lie within the detection envelope of the sensor). If an observation is incorrectly associated with a landmark in the map, the effect is to introduce unmodelled errors into the filter. For example, suppose that the filter assumes that the ith observation in the set matches with landmark i whereas in fact it matches with landmark j. Expanding (2) and using the fact that $\mathbf{x}(k) = \hat{\mathbf{x}}(k \mid k-1) + \tilde{\mathbf{x}}(k \mid k-1)$, the innovation is

$$\begin{aligned}\boldsymbol{\nu}(k) &= \mathbf{w}(k) + \mathbf{h}_i\left[\mathbf{x}(k-1)\right] - \mathbf{h}_j\left[\hat{\mathbf{x}}(k \mid k-1)\right]\\ &= \mathbf{w}(k) + \mathbf{h}_i\left[\hat{\mathbf{x}}(k \mid k-1)\right] + \boldsymbol{\nabla}\mathbf{h}_i\,\tilde{\mathbf{x}}(k \mid k-1)\\ &\quad - \mathbf{h}_j\left[\hat{\mathbf{x}}(k \mid k-1)\right]\\ &= \mathbf{w}(k) + \boldsymbol{\nabla}\mathbf{h}_i\,\tilde{\mathbf{x}}(k \mid k-1)\\ &\quad + \left(\mathbf{h}_i\left[\hat{\mathbf{x}}(k \mid k-1)\right] - \mathbf{h}_j\left[\hat{\mathbf{x}}(k \mid k-1)\right]\right).\end{aligned}$$

Therefore, from (1) the effect of associating a measurement with the incorrect beacon is to add the error term

$$\mathbf{W}(k)\Big(\mathbf{h}_i\left[\hat{\mathbf{x}}(k \mid k-1)\right] - \mathbf{h}_j\left[\hat{\mathbf{x}}(k \mid k-1)\right]\Big) \tag{4}$$

to the estimate.

By far the most common way to overcome this problem is to attempt to identify the true identify of the landmark. The nearest neighbour algorithm chooses the beacon with the smallest normalised innovation. The normalised innovation $q(k)$ is given by

$$q(k) = \boldsymbol{\nu}^T(k)\mathbf{S}^{-1}(k)\,\boldsymbol{\nu}(k)$$

where $\boldsymbol{\nu}(k)$ and $\mathbf{S}(k)$ are calculated under the assumption that landmark i is observed. However, this approach is notoriously unreliable. It takes no account of the previous history of measurements and can be readily corrupted by large noise values within a single time step.

A more refined approach is to use multiple hypothesis tracking (MHT) [1, 7]. For each feasible assignment of $\mathcal{S}_k$ to the map, a hypothesis is created and a filter is initialised using the previous estimate and the pattern of assignments associated with the hypothesis. Each filter is run independently and in parallel and the probability that a given hypothesis is correct at any given time can be calculated from the likelihood. For example, under the assumption that the state and observation errors are Gaussian, the likelihood that an observation is associated with landmark i is

$$p_i(k) = \frac{1}{(2\pi)^{m/2}|\mathbf{S}(k)|}\exp\left\{\frac{1}{2}\boldsymbol{\nu}^T(k)\mathbf{S}^{-1}(k)\,\boldsymbol{\nu}(k)\right\}$$

where m is the dimension of the observations. For multiple measurements the joint probability can be calculated. Given the set of hypothesised estimates, a single statistically summarised estimate can be formed by weighting them as follows [2]:

$$\hat{\mathbf{x}}(k \mid k) = \sum_{i=1}^{s_k} p_i(k)\hat{\mathbf{x}}_i(k \mid k)$$

$$\mathbf{P}(k \mid k) = \sum_{i=1}^{s_k} p_i(k)\left[\mathbf{P}_i(k \mid k) + \{\hat{\mathbf{x}}(k \mid k) - \hat{\mathbf{x}}_i(k \mid k)\}\{\hat{\mathbf{x}}(k \mid k) - \hat{\mathbf{x}}_i(k \mid k)\}^T\right]$$

However, there are two difficulties with MHT. The first is that the number of hypotheses is potentially unlimited. Suppose that, at each time step, there are p possible hypotheses. After k time steps there will be p^k hypotheses which must be maintained. Therefore, various computational techniques, such as pruning, must be used to reduce the number of hypotheses. However, as explained above there is a chance that the correct assignment might be rejected. Because each filter assumes that it has the correct gating, it does not include the error term from (4) and so it will become inconsistent. This can often lead to catastrophic filter failure.

A suboptimal, but more practical, alternative to MHT is offered by the Covariance Union (CU) data fusion algorithm [15]. The rationale behind CU is to calculate a single estimate that is consistent with *all* possible assignments. However, it is unclear how much performance is sacrificed in order to achieve consistency. The main purpose of this paper is to generate some initial results by inserting CU into a simulator and comparing its output with that from a variety of baseline solutions over multiple time-steps of a nonlinear navigation problem.

III. Covariance Union (CU)

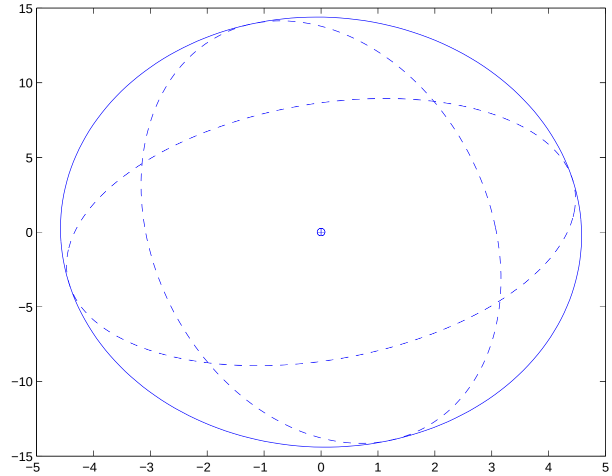
CU considers the following problem: suppose a filtering algorithm is provided with two observations with means and covariances $(\mathbf{m}_1, \mathbf{M}_1)$ and $(\mathbf{m}_2, \mathbf{M}_2)$ respectively. It is known that one observation corresponds to a correct association, and the other to an incorrect association. However, the identity of the consistent estimate is unknown and cannot be determined. In this circumstance, the only way the KF can be guaranteed to give a consistent estimate is if it updates with an observation which is consistent with respect to *both* measurements[1]. This unioned estimate has a mean and covariance $(\mathbf{u}, \mathbf{U})$ and obeys the property

$$\mathbf{U} \geq \mathbf{M}_1 + (\mathbf{u} - \mathbf{m}_1)(\mathbf{u} - \mathbf{m}_1)^T \quad (5)$$

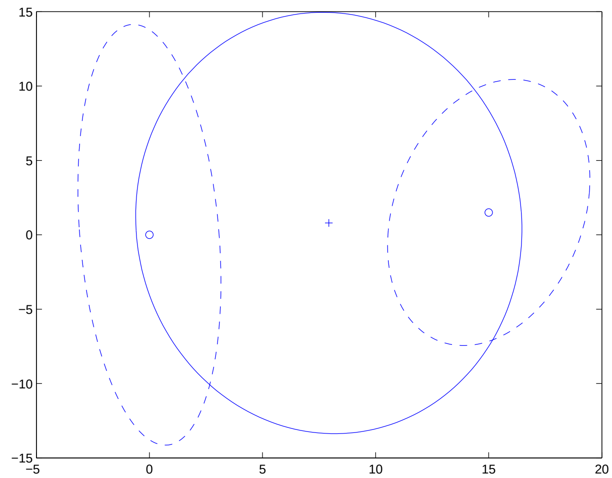
$$\mathbf{U} \geq \mathbf{M}_2 + (\mathbf{u} - \mathbf{m}_2)(\mathbf{u} - \mathbf{m}_2)^T \quad (6)$$

where some measure of the size of $\mathbf{U}$ (e.g., determinant) is minimized. Such a unioned estimate can be constructed by applying convex or semidefinite optimization methods. Given this Covariance Union (CU) of the two measurements, the KF can be applied directly to update the prediction $(\hat{\mathbf{x}}(k \mid k-1), \mathbf{P}(k \mid k-1))$ with the CU-derived observation $(\mathbf{u}, \mathbf{U})$. In other words, the above equations simply say that if the estimate $(\mathbf{m}_1, \mathbf{M}_1)$ is consistent, then the translation of the vector $\mathbf{m}_1$ to $\mathbf{u}$ will require its covariance to be enlarged by the addition of a matrix at least as large as the outer product of $(\mathbf{u} - \mathbf{m}_1)$ in order to be consistent. The same reasoning applies if the estimate $(\mathbf{m}_2, \mathbf{M}_2)$, is consistent. Covariance Union therefore determines the smallest covariance $\mathbf{U}$ that is large enough to guarantee consistency regardless of which of the two given estimates is consistent.

[1]This is direct consequence of the result by Nishimura [11]. Nishimura considered the problem of modelling errors and showed that, even if the process and observation models contain errors, a consistent estimate can be achieved by increasing the process and observation noise by a sufficient amount.



(a) Coincident means.



(b) Non-coincident means.

Fig. 1. Illustration of CU for estimates with coincident and non-coincident means. The original means and covariances are shown as dashed ellipses with circles. The minimum determinant union is shown as a solid ellipse with mean denoted by a cross.

Fig. 1 shows the CU estimate of two pairs of estimates which were calculated using the algorithm described in the appendix. In the first example the means are coincident and the results are

$$\mathbf{m}_1 = \begin{pmatrix}0\\0\end{pmatrix} \qquad \mathbf{M}_1 = \begin{bmatrix}10 & -10\\-10 & 20\end{bmatrix}$$
$$\mathbf{m}_2 = \begin{pmatrix}0\\0\end{pmatrix} \qquad \mathbf{M}_2 = \begin{bmatrix}20 & 10\\10 & 80\end{bmatrix}$$
$$\mathbf{u} = \begin{pmatrix}0\\0\end{pmatrix} \qquad \mathbf{U} = \begin{bmatrix}20.97 & -1.08\\-1.08 & 207.25\end{bmatrix}$$

In this case $\mathbf{u}$ tightly encloses both $\mathbf{M}_1$ and $\mathbf{M}_2$. In the second example the means are not coincident and

$$\mathbf{m}_1 = \begin{pmatrix}0\\0\end{pmatrix} \qquad \mathbf{M}_1 = \begin{bmatrix}10 & -10\\-10 & 20\end{bmatrix}$$
$$\mathbf{m}_2 = \begin{pmatrix}15\\1.5\end{pmatrix} \qquad \mathbf{M}_2 = \begin{bmatrix}20 & 10\\10 & 80\end{bmatrix}$$
$$\mathbf{u} = \begin{pmatrix}7.93\\0.79\end{pmatrix} \qquad \mathbf{U} = \begin{bmatrix}72.89 & -3.71\\-3.71 & 200.63\end{bmatrix}$$

The covariance matrix does not enclose the covariance matrices of the two estimates. This illustrates that CU does *not* assume that the estimates are bounded and the results from set membership theory (such as those described in [12]) do not apply when the means are not coincident.

Although general optimization algorithms can be applied to compute CU results, greater computational efficiency may be achieved by transforming the CU conditions, e.g., for more direct application of specialized convex optimization algorithms. Using the substitution, $\mathbf{u} = \omega\mathbf{m}_1 + (1-\omega)\mathbf{m}_2$, for $0 \leq \omega \leq 1$, the CU conditions for two estimates can be specialized to:

$$\begin{aligned}\mathbf{U} &\geq \mathbf{M}_1 + (1-\omega)^2(\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T\\ \mathbf{U} &\geq \mathbf{M}_2 + \omega^2(\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T\end{aligned}$$

where it is assumed that the optimal vector $\mathbf{u}$ is a convex combination of the given mean vectors. It turns out that this assumption is valid for the minimization of the determinant (or any matrix norm) of $\mathbf{U}$ because the addition of any vector component that is orthogonal to $\omega\mathbf{m}_1 + (1-\omega)\mathbf{m}_2$ introduces an additional nonzero component to the covariance $\mathbf{U}$.

More generally, for any set of n measurements in the same coordinate frame, e.g., $(\mathbf{m}_1, \mathbf{m}_1)$, $(\mathbf{m}_2, \mathbf{M}_2)$, ..., $(\mathbf{m}_n, \mathbf{M}_n)$, in which one or more elements of the set is a measurement of a system of interest, whose state is maintained as the estimate $(\mathbf{x}, \mathbf{P})$, it is possible to construct a unioned measurement $(\mathbf{u}, \mathbf{U})$ that is consistent with respect to each element of the set of measurements. In particular, $(\mathbf{u}, \mathbf{U})$ is defined by the following constraints:

$$\begin{aligned}\mathbf{U} &\geq \mathbf{m}_1 + (\mathbf{u} - \mathbf{m}_1)(\mathbf{u} - \mathbf{m}_1)^T\\ \mathbf{U} &\geq \mathbf{M}_2 + (\mathbf{u} - \mathbf{m}_2)(\mathbf{u} - \mathbf{m}_2)^T\\ &\vdots\\ \mathbf{U} &\geq \mathbf{M}_n + (\mathbf{u} - \mathbf{m}_n)(\mathbf{u} - \mathbf{m}_n)^T\end{aligned}$$

This unioned estimate $(\mathbf{u}, \mathbf{U})$ can then be used to update the system estimate $(\mathbf{x}, \mathbf{P})$. The updated estimate is guaranteed to be consistent as long as the estimate $(\mathbf{x}, \mathbf{P})$ and at least one element of the set of measurements are consistent with respect to the system of interest.

In fact it is possible to extract a little more accuracy from the CU method by applying it to the separate estimates obtained from updating the system state with each measurement. The unioned estimate is guaranteed to be consistent because one of these updates is itself consistent, and its covariance may be relatively smaller because the set of estimates to which it is applied will tend to have smaller covariances and smaller deviations in their means due to the updates.

IV. Resolving Beacon Ambiguity with CU

In the navigation context, CU is triggered by the presence of more than one beacon measurement falling inside the system's validation gate. Specifically, the Mahalanobis distance is calculated between each measurement and the system state. If more than one distance value is smaller than some pre-specified threshold, CU is invoked to calculate a consistent update between the measurements and the state. Consequently, the beacon ambiguity problem is solved in a consistent manner, albeit at the expense of a conservative update. The dynamics of the navigation filter is scenario dependent. For example, the swollen covariance of a CU update may accommodate even more ambiguous assignments at the next time step. Nevertheless, a consistent update is guaranteed. On the other hand, if there is only a single assignment to be made, a standard Kalman update would be triggered and this would restore some accuracy to the system. The next section describes an example that is aimed at investigating the performance of CU over multiple time steps of a nonlinear navigation problem.

V. Example

The CU algorithm was tested in the scenario which is illustrated in Fig. 2 — a vehicle drives between regularly spaced blocks. A bearings-only sensor is fixed to the front of the vehicle and can detect targets which are rigidly fixed to the surface of the blocks. This is an abstraction of a vehicle driving through an urban environment and using a single camera to detect and track features (such as edges of windows) which are on the walls of the building.

The state space of the system consists of the (x, y) position and orientation θ of the vehicle. The control inputs for the vehicle are front wheel speed $V(k)$ and front steer angle $\delta(k)$. The process model for the vehicle is

$$x(k) = x(k-1) + V(k)\Delta t \cos\left[\theta(k) + \delta(k)\right] \quad (7)$$
$$y(k) = y(k-1) + V(k)\Delta t \sin\left[\theta(k) + \delta(k)\right] \quad (8)$$
$$\theta(k) = \theta(k-1) + V(k)\Delta t \frac{\sin\left[\delta(k)\right]}{B} \quad (9)$$

where B is the wheel base of the vehicle. The observation model is for a bearings-only measurement. Therefore, if beacon i is observed the observation model is

$$\mathbf{z}(k) = \tan^{-1}\left[\frac{y_i - y(k)}{x_i - x(k)}\right] - \theta. \quad (10)$$

The truth model identifies all beacons which lie within the FOV, are oriented such that the angle between the beacon

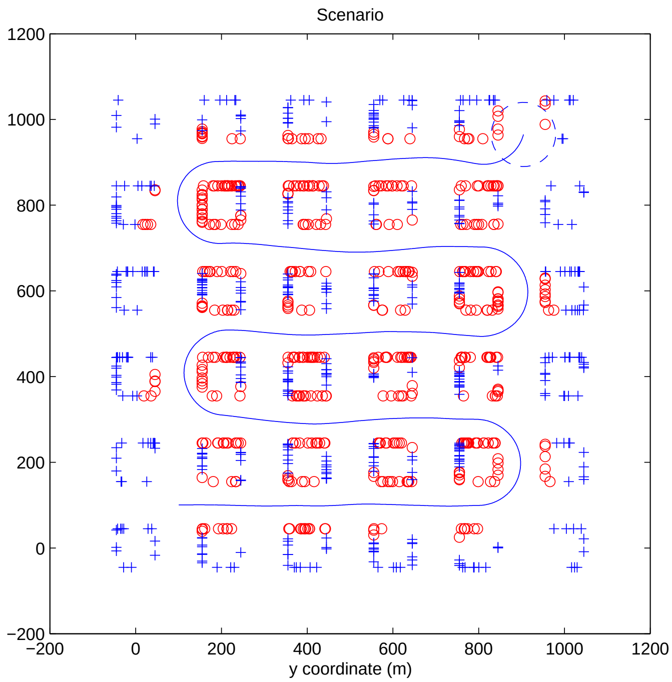


Fig. 2. The experimental scenario. A vehicle follows a criss-cross path through a series of "streets" between "buildings". Each "building" has a set of targets on it (observed one circles, unobserved ones crosses). The sensor can only detect the relative bearing to those targets and no ID information exists with each target. The maximum range at which a sensor can be detected is 75m (circle at the end of the true vehicle path line).

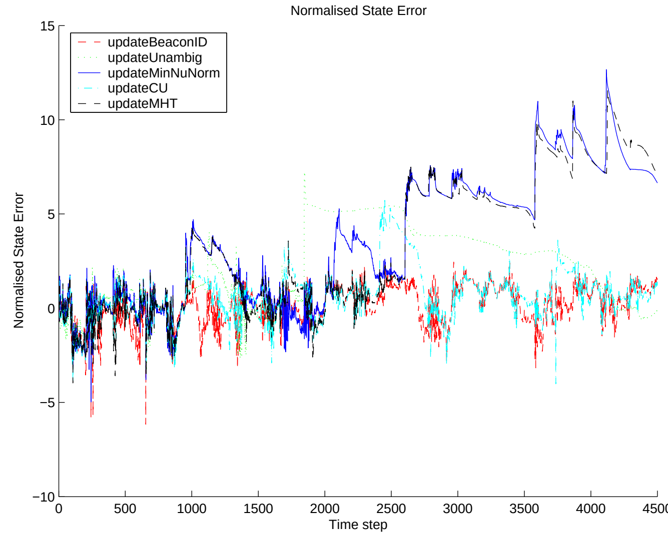


Fig. 3. The log of the normalised error in the estimates. If the estimate is consistent, the average value should be log(3)=1.1.

normal and the view angle is less than 90 degrees, and are within a certain distance (L) of the platform. From this set a single beacon is drawn at random and used in each update step. Given the observation, each estimator has a list of available beacons. It uses the same set of criteria (distance, normal angle, FOV) to determine a list of potential candidates. In order to detect potential ambiguities a χ^2 test was performed: if the normalised innovation is less than 4 (the average should be 1), the beacon is considered to be a potential candidate.

Given the list of candidates, five different data fusion algorithms were tested:

1) **updateBeaconID:** Kalman filter, given the true beacon ID. This is a best case baseline and is used to compare the performance of all of the other filters.
2) **updateUnambig:** Kalman filter, which only updates if the beacon can be unambiguously identified.
3) **updateMinNuNorm:** Kalman filter, using the beacon with the smallest normalised innovation (i.e., the nearest neighbour solution).
4) **updateCU:** CU. For each gated beacon, perform a Kalman update. Use CU to calculate the union of all of the different updates (this was implemented in a computationally simpler but sub-optimal pairwise manner).
5) **updateMHT:** Partial MHT. This uses the algorithm described in Section II. However, only the single best hypothesis is maintained and is used to update the estimate.

Each filter, being nonlinear, was implemented using the scaled unscented transformation [8] with $\alpha = 10^{-3}$ and $\beta = 2$.

A total of 50 Monte Carlo runs was performed. In each run, the beacon distribution, the initial vehicle position, and the process noise and observation noise, were drawn from a normal distribution. The navigation filter was tuned so that the truth filter (with beacon IDs) gave consistent estimates. Fig. 3 plots the log of the normalised mean squared error in the estimate from each filter calculated over five runs. The normalised mean squared error is given by

$$q(k) = \left(\hat{\mathbf{x}}(k \mid k) - \mathbf{x}(k)\right) \mathbf{P}^{-1}(k \mid k) \left(\hat{\mathbf{x}}(k \mid k) - \mathbf{x}(k)\right)^T$$

where $\mathbf{x}(k)$ is the true value at time step k.

The results suggest that the normalised error in the CU filter is, on average, smaller than that of the other filters (apart from the baseline with access to true beacon IDs). This difference is particularly pronounced after about time step 2500. However, it should be noted that *all* of the filters (apart from 1, which was always provided with the true beacon ID) lost track in different runs. To quantify this track loss, the following criteria was tested: if a filter receives measurements for 10 successive runs but does not gate any of them with the map, then it is declared lost. Table V lists the loss counts for the different filter implementations. It shows that the CU filter has the smallest number of losses and only using unambiguous measurements has the largest number of losses. The poor performance of the partial MHT would be improved if hypotheses were propagated over multiple time steps.

VI. Conclusions

This paper has considered the problem of assignment ambiguity for landmark-based tracking systems. We have examined a novel algorithmic approach, Covariance Union (CU), which accommodates assignment ambiguity by computing

Scheme	Loss count
updateBeaconID	0
updateUnambig	18
updateMinNuNorm	13
updateCU	7
updateMHT	11

TABLE I

THE TOTAL LOSS COUNT FOR EACH FILTER FROM 50 SEPARATE RUNS.

an updated estimate that is consistent with all the feasible estimates. Initial experiments show that the performance of CU is comparable or better than that of other techniques, but its computational demands are potentially much less because CU essentially summarizes many hypotheses in a single unioned state estimate. Although this paper has presented CU as an alternative to MHT, we believe that CU represents an effective mechanism to complement MHT by replacing pruned hypotheses with CU estimates. Further work will focus on detailed and systematic simulation studies, but perhaps more importantly on the evaluation of CU in current hardware-based robotic simultaneous localization and mapping systems, e.g., [4].

REFERENCES

[1] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, New York NY, USA, 1988.

[2] I.P. Bottlik and S. S. Blackman. Co-ordinated Presentation of Multiple Hypotheses in Multi-Target Tracking. In *Signal and Data Processing of Small Targets*, volume 1096. SPIE, 1989.

[3] F. Chenavier and J. Crowley. Position Estimation for a Mobile Robot Using Vision and Odometry. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 2588–2593, Nice, France, 12–14 May 1992.

[4] A.R. Cooke, P. Greenway, M. le Maistre, R. Manley, D. Mullin, D.Nicholson, D. Valachis. Application of Simultaneous Localisation and Mapping in Autonomous Exploration of Urban Environments. In *Signal Processing, Sensor Fusion and Target Recognition XII*, volume 5096. SPIE, 2003.

[5] I. J. Cox and J. J. Leonard. Modeling a Dynamic Environment Using a Multiple Hypothesis Approach. *Artificial Intelligence*, 66(2):311–344, April 1994.

[6] H. Hu, D. Gu and J. M. Brady. Outdoor Navigation of a Mobile Robot with Multiple Sensors. In *SPIE Intelligent Control Systems*, Pittsburgh, PA, USA, 14–17 October 1997.

[7] C. Hue, J.-P. Le Cadre and P. Perez. Tracking multiple objects with particle filtering. Technical Report 1361, IRISA, Rennes, France, 2000.

[8] S. J. Julier, and J. K. Uhlmann. The Scaled Unscented Transformation. In *Proceedings of the IEEE American Control Conference*, pages 4555–4559, Anchorage AK, USA, 8–10 May 2002. IEEE.

[9] J. Leonard and H. F. Durrant-Whyte. *Directed Sonar Sensing for Mobile Robot Navigation*. Kluwer Academic Press, Boston MA, USA, 1991.

[10] 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, December 2000.

[11] T. Nishimura. On the a priori information in sequential estimation problems. *IEEE Transactions on Automatic Control*, TAC-11(4):197–204, April 1966.

[12] F. C. Schweppe. *Uncertain Dynamic Systems*. Prentice Hall, Englewood Cliffs, NJ, USA, 1973.

[13] C. M. Smith. *Integrated Mapping and Navigation*. PhD thesis, MIT, 1998.

[14] C. Tomasi and T. Kanade. Detection and tracking of point features. Cmu-cs-91-132, CMU, Pittsburgh, PA, USA, April 1991.

[15] J. K. Uhlmann. Covariance Consistency Methods for Fault-Tolerant Distributed Data Fusion. *Information Fusion*, 4(3):201–215, 2002.

APPENDIX

Given the prior observations $(\mathbf{m}_1, \mathbf{M}_1)$ and $(\mathbf{m}_2, \mathbf{M}_2)$ and a candidate mean $\mathbf{u}$, this appendix describes a method for calculating $\mathbf{U}$ which is guaranteed to be consistent.

Define

$$\mathbf{U}_1 = \mathbf{M}_1 + \{\mathbf{u} - \mathbf{m}_1\}\{\mathbf{u} - \mathbf{m}_1\}^T$$
$$\mathbf{U}_2 = \mathbf{M}_2 + \{\mathbf{u} - \mathbf{m}_2\}\{\mathbf{u} - \mathbf{m}_2\}^T$$

Let

$$\mathbf{S} = \sqrt{[\mathbf{U}_2]}$$

where $\mathbf{U}_2 = \mathbf{S}^T\mathbf{S}$.

Let $\mathbf{V}$ and $\mathbf{D}$ be the matrices of eigenvectors and eigenvalues of $\left(\mathbf{S}^{-1}\right)^T \mathbf{U}_1\mathbf{S}^{-1}$.

Theorem 1: Given the prior observations $(\mathbf{m}_1, \mathbf{M}_1)$ and $(\mathbf{m}_2, \mathbf{M}_2)$ and a candidate mean $\mathbf{u}$, a covariance matrix $\mathbf{U}$ which obeys conditions (5) and (6) is given by

$$\mathbf{U} = \mathbf{S}^T\mathbf{V}\max(\mathbf{D}, \mathbf{I})\,\mathbf{V}^T\mathbf{S}$$

where $\max(\mathbf{A}, \mathbf{B})$ calculates the matrix which is the element-wise maximum of the matrices $\mathbf{A}$ and $\mathbf{B}$.

Proof: From the definitions of $\mathbf{V}$ and $\mathbf{D}$,

$$\mathbf{V}\mathbf{D}\mathbf{V}^T = \left(\mathbf{S}^{-1}\right)^T \mathbf{U}_1\mathbf{S}^{-1}.$$

Therefore,

$$\mathbf{U}_1 = \mathbf{S}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{S}.$$

Similarly, because $\mathbf{V}$ is orthonormal,

$$\mathbf{U}_2 = \mathbf{S}^T\mathbf{V}\mathbf{I}\mathbf{V}^T\mathbf{S}.$$

Therefore, any matrix which can be written in the form

$$\mathbf{U} = \mathbf{S}^T\mathbf{V}\mathbf{C}\mathbf{V}^T\mathbf{S}$$

where $\mathbf{C} - \max(\mathbf{D}, \mathbf{I}) \geq \mathbf{0}$ will be consistent with respect to both $\mathbf{U}_1$ and $\mathbf{U}_2$.

■