Rich Probabilistic Representations for Bearing Only Decentralised Data Fusion

Ben Upcroft, Lee Ling Ong, Suresh Kumar, Matthew Ridley, Tim Bailey,
Salah Sukkarieh, and Hugh Durrant-Whyte
ARC Centre of Excellence for Autonomous Systems, Australian Centre for Field Robotics,
The Rose St. Building, J04, University of Sydney, NSW, Australia 2006
Email: b.upcroft@cas.edu.au

***Abstract*— The aim of this paper is to demonstrate the validity of using Gaussian mixture models (GMM) for representing probabilistic distributions in a decentralised data fusion (DDF) framework. GMMs are a powerful and compact stochastic representation allowing efficient communication of feature properties in large scale decentralised sensor networks. It will be shown that GMMs provide a basis for analytical solutions to the update and prediction operations for general Bayesian filtering. Furthermore, a variant on the Covariance Intersect algorithm for Gaussian mixtures will be presented ensuring a conservative update for the fusion of correlated information between two nodes in the network. In addition, purely visual sensory data will be used to show that decentralised data fusion and tracking of non-Gaussian states observed by multiple autonomous vehicles is feasible.**

I. Introduction

This paper aims to develop non-Gaussian algorithms for *decentralised*, multiple vehicle, map building. The motivation is that of multiple autonomous flight and ground vehicles cooperatively building a map of the terrain over which they are flying, each using one or more terrain sensors. Ultimately, the algorithms should scale such that any number of vehicles, with any number of payloads, in any configuration can be used.

Decentralised architectures offer a number of advantages over conventional hierarchical [1] and distributed [2], multiple fusion sensor architectures for map building tasks. Since they do not require a central resource, they are able to communicate information throughout the network in an efficient, modular, and scalable manner. There are three basic constraints to a general decentralised system: 1. There is no single central fusion centre and no node should be central to the operation of the network. 2. There is no common communications facility - communication must be kept on a strictly node-to-node basis. 3. Each node has knowledge only of its immediate neighbours - there is no global knowledge of the network topology.

The work presented here is part of the Autonomous Navigation and Sensing Experimental Research 2 (ANSER 2) project and is aimed at the development of multiple flight and ground vehicle demonstrations of general decentralised data fusion. Each vehicle will be equipped with GPS and inertial sensors and will carry a vision system payload for terrain feature tracking. Each payload processor will implement a fully decentralised data fusion algorithm. There will be no separate fusion centre on any platform and no fusion centre elsewhere on the ground.

In order to develop and test these decentralised algorithms successfully, simulations were performed on real data obtained from colour aerial images. The initial development has centred on decentralised picture compilation (or map building) using bearing-only vision sensor observations of unstructured, natural terrain features on the ground. Note that this does not involve pose estimation of the vehicles. Each platform maintains a bank of decentralised, non-Gaussian Bayesian filters for the features it observes, and transmits the information to all other platforms. The net result is that each platform maintains a complete map of all features observed by all nodes in the network. Multiple observations of the same feature, possibly by different platforms, results in an increasingly accurate estimate of the feature location for *all* nodes.

Most robust approaches to decentralised data fusion have involved tracking position features provided by range devices such as radar or laser [3], or by tracking known visual features which have been previously placed in the observed environment [4]. In this paper purely visual sensory data for natural object localisation is used. A bearing-only observation model may be used for this problem but a general probabilistic framework is needed for accurate state estimation [5], [6].

The algorithms developed here are based on the general Bayesian filter using Gaussian mixture models as the probabilistic representation. Particle representations are a common solution for nonlinear and non-Gaussian filtering but become computationally expensive in higher dimensions due to the required number of particles for accurate estimation. The problem is exacerbated if the observed feature properties are to be communicated throughout a network forcing the requirement of a compact representation. Moreover, a decentralised sensor network imposes the need for removal of common information between nodes in the network if data fusion is to be mathematically consistent. However, a particle representation does not provide this consistency without some modification to the representation.

Development of the algorithms in this paper show that Gaussian mixture models satisfy all the constraints for a general Decentralised Data Fusion (DDF) architecture while also providing analytical solutions to the operations for general Bayesian filtering. Decentralised data fusion using GMMs is demonstrated in this paper with position states of unstructured,

natural features extracted from *purely visual* sensory data obtained from separate simulated moving platforms.

II. Related Work

Applications that benefit from multi-sensor data fusion include environmental sensing, surveillance, mobile autonomous teams, and the Internet [7], [3]. In each of these problems, individual nodes of the network make local measurements or observations of the common environment and attempt to combine the measurements to produce a global estimate of the observed state.

Nettleton *et al.* showed that scalable decentralised state estimation with Gaussian noise can be achieved in outdoor environments using autonomous air vehicles observing artificial features [4]. Through the information (canonical) form of a Gaussian, it was shown that local and communicated information can be fused at any time and any order using additive information matrices. However, these additive algorithms are only valid for Gaussian representations and do not extend to general probabilistic distributions. Paskin *et al.* also demonstrated a DDF architecture using motes although only Gaussian noise was considered and a tree topology over the network was enforced [7].

Recently, Ihler *et al.* demonstrated that non-parametric distributions could be used for sensor-calibration in a network with an approximate communication algorithm called non-parametric belief propagation [8]. Although this algorithm converges to the true state in a number of cases, it can also result in overconfident estimates due to the fusion of common information. Rosencrantz *et al.* also showed that decentralised state estimation can be performed with non-Gaussian representations [3]. Range and bearing observations were performed using a laser with Bayesian estimation on the states achieved through the use of the particle filter. It was also demonstrated that consistent information fusion for dynamic indoor environments could only be achieved if the belief over state histories was maintained. They showed that in some cases the latest beliefs of the individual nodes could lead to false global beliefs if the histories were not included in the communication.

The work presented here also concentrates on non-Gaussian estimation but it extends previous decentralised research to include bearing only, *visual* observations of natural features rather than range-bearing observations. Gaussian mixture models are used as the probabilistic representation which was first considered in tracking problems by Alspach and Sorenson [9], [10]. Alspach also extended these ideas to multi-target identification [11].

Since static features were used throughout this paper without maintaining belief histories, the problems faced by Rosencrantz *et al.* [3] were rarely observed.

In addition, the correlations between common information of individual nodes will be accounted for through an extension of the Covariance Intersect method [12] for GMMs. Unlike the work by Rosencrantz *et al.* [3] and Ihler *et al.* [8] the generalised CI presented here ensures that common information cannot cause over-confidence of the underlying state.

Section III describes the general Bayesian filtering problem cast in a DDF framework. The main issue that arises is the mathematical consistency in which fusion is performed. GMMs are introduced in Sec. IV as a representation which satisfy all the requirements of a general DDF framework. The visual features and the position observation model of these features are described in Sec. V. Finally, DDF results of simulated air vehicles using real visual data are presented in Sec. VI. Conclusions and further work are discussed in Sec. VII.

III. Distributed Bayesian Estimation

The algorithmic structure in distributed estimation is the same for every node in the network and is outlined in Fig. 1. The essential process is:

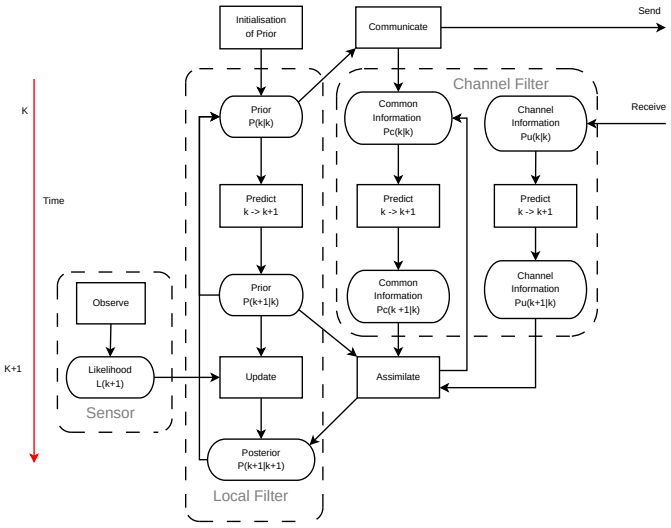

Fig. 1. Flow chart of the operations performed in distributed estimation.

1) Observations are made by a sensor.
2) A likelihood model over the state is generated using this observation.
3) Data association with existing local tracks.
4) A standard cycle of a local observation update (multiplication of prior and likelihood) and prediction (convolution of prior with process model).
5) Information is sent and received between neighbouring nodes in the network and recorded in the channel filter.
6) Data association between local tracks and incoming information from other nodes.
7) Assimilation using the generalised Covariance Intersect algorithm.
8) Repeat process.

The following sections describe the general Bayesian filter for steps 4 and 7 from which specific GMM algorithms can be derived.

A. Local Update: Bayes Theorem

Bayes theorem provides an incremental and recursive, probabilistic method for combining observations $\mathbf{Z}^k$ of a state $\mathbf{x}_k$, at time t_k, with a prior belief of the state $P(\hat{\mathbf{x}}_k|\mathbf{Z}^{k-1})$. This prior is a prediction from the posterior over the previous state $P(\mathbf{x}_{k-1}|\mathbf{Z}^{k-1})$ calculated using the Chapman-Kolmogorov equation. Observations are obtained from some sensor (modeled as a conditional probability distribution or likelihood) $L(\mathbf{z} = z_k|\mathbf{x}_k)$, and the resultant combination is a revised posterior distribution on the state:

$$P(\mathbf{x}_k|\mathbf{Z}^k) = \frac{L(\mathbf{z} = z_k|\mathbf{x}_k)P(\hat{\mathbf{x}}_k|\mathbf{Z}^{k-1})}{P(z_k|\mathbf{Z}^{k-1})} \tag{1}$$

where $\mathbf{Z}^k = \{z_k, \mathbf{Z}^{k-1}\}$ is the set of observations from all nodes in the DDF network.

B. Local Prediction: Chapman-Kolmogorov Equation

The local prediction step in a DDF network is given by the Chapman-Kolmogorov equation (also known as the Total Probability Theorem):

$$P(\hat{\mathbf{x}}_k|\mathbf{Z}^{k-1}) = \int P(\mathbf{x}_k|\mathbf{x}_{k-1})P(\mathbf{x}_{k-1}|\mathbf{Z}^{k-1}, \mathbf{x}_0)d\mathbf{x}_{k-1} \tag{2}$$

where the transition probability density $P(\mathbf{x}_k|\mathbf{x}_{k-1})$ is known as the motion model, and $P(\mathbf{x}_{k-1}|\mathbf{Z}^{k-1}, \mathbf{x}_0)$ is the updated estimate from the previous time step.

C. Fusion

It can be shown that fusion of the raw correlated information between nodes i and j is [13], [14]

$$P(\mathbf{x}|\mathbf{Z}_i \cup \mathbf{Z}_j) = \frac{1}{c}\frac{P(\mathbf{x}|\mathbf{Z}_i)P(\mathbf{x}|\mathbf{Z}_j)}{P(\mathbf{x}|\mathbf{Z}_i \cap \mathbf{Z}_j)} \tag{3}$$

where $\mathbf{Z}_{i(j)}$ are all the observations available to node i (j), $P(\mathbf{x}|\mathbf{Z}_i \cup \mathbf{Z}_j)$ is the posterior probability over the unknown state given information from both nodes, $P(\mathbf{x}|\mathbf{Z}_{i(j)})$ are the posteriors based only on locally available information, $P(\mathbf{x}|\mathbf{Z}_i \cap \mathbf{Z}_j)$ is the information the two nodes have in common, and c is a normalising constant.

Thus the problem of constructing the union $\mathbf{Z}_i \cup \mathbf{Z}_j$, reduces to finding the common information $\mathbf{Z}_i \cap \mathbf{Z}_j$ and is the *key* to the decentralised communication problem.

D. Identification of the Common Information

The incorporation of redundant information in DDF systems may lead to bias, over-confidence and divergence in estimates. Therefore this information must be removed before two nodes can freely communicate with each other.

Gaussian representations allow an analytical solution to the division which is performed by a channel filter for tree-connected networks [15]. However, for general probabilistic representations the authors have not found an analytically or numerically tractable solution to this division. Thus, even if the common information can be calculated, it cannot be removed for fusion.

In the work by Ihler *et al.* and Rosencrantz *et al.* the communication protocols do not account for common information [8], [3]. In Ihler's work, loopy belief propagation can result in overconfident estimates. The particle implementation by Rosencrantz *et al.* attempts to fuse distributions that have support on different regions of the state which in itself is mathematically inconsistent without producing a functional distribution over the individual representations. In addition, although only the most informative messages are fused there is no guarantee that the common information has been removed.

A non-optimal solution for Gaussian representations is the Covariance Intersect filter which conservatively combines the information in two incoming channels assuming that the correlation is unknown [12].

Removal of common information should be achieved in a mathematically consistent manner and therefore constrains the types of probability density representations that can be used in the DDF framework. It will be shown that the CI algorithm can be extended for the use with GMMs and for the simulations shown in this paper, provides conservative fusion updates.

IV. Gaussian Mixture Model Nonlinear Filter

Other than the particle representation, multimodal stochastic models have been largely ignored in the filtering and tracking community. However, it has been well known for some time that Gaussian mixture models (Gaussian sum approximations) provide a basis for analytical solutions to the general Bayesian filtering problem [9]. This section illustrates the Gaussian mixture model as a general representation and its advantages in the DDF framework.

A Gaussian mixture model is defined for a random variable $\mathbf{X}$ as

$$P(\mathbf{x}) = \sum_{i=1}^{N} \pi_i \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i) \tag{4}$$

where $\mathbf{x}$ are the observations of $\mathbf{X}$, π_i are positive weights with the property $\sum_{i=1}^{N} \pi_i = 1$, $\mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i)$ is a Gaussian probability density (also known as a Gaussian mixture component) with mean μ_i and full covariance Σ_i, and N is the number of mixture components.

A. Measurement Update - Bayes Theorem

Distributions that are estimated by weighted sums of Gaussians (Gaussian kernel densities, and GMMs) allow the update step involving Bayes theorem to be solved analytically which is in general not possible.

Substitution of GMMs into Eq. 1 gives

$$P(\mathbf{x}_k|\mathbf{z}_k) = A \sum_{i=1}^{M} \pi_{zi}\mathcal{N}_{zi} \sum_{j=1}^{N} \pi_{xj}\mathcal{N}_{xj} \tag{5}$$

where $A = 1/P(\mathbf{z}_k|\mathbf{z}_{k-1})$ is a normalising constant, the $\mathcal{N}_z$'s represent the likelihood distribution $P(\mathbf{z}_k|\mathbf{x}_k)$, and the $\mathcal{N}_x$'s represent the prediction $P(\hat{\mathbf{x}}_k|\mathbf{z}_{k-1})$. Similarly for π_z and π_x.

Expanding Eq. 5 results in $M \times N$ terms, each which involve a multiplication of two weighted Gaussians. Thus, the posterior distribution is represented by $M \times N$ weighted Gaussians.

B. Prediction - The Chapman-Kolmogorov Equation

As with Bayes theorem, GMMs allow an analytical solution to the Chapman-Kolmogorov equation. Substituting GMMs into Eq. 2 results in a convolution between $M \times N$ weighted Gaussians with each term resulting in a Gaussian of the form [16]

$$\pi \mathcal{N}\left(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1^2 + \boldsymbol{\Sigma}_2^2\right) \tag{6}$$

where the subscripts denote the variables for the two Gaussians and π is a constant weighting term.

C. Fusion - Generalised CI for GMMs

1) Covariance Intersect Filter: The Covariance Intersection (CI) algorithm provides a solution to the problem of combining two Gaussian random vectors in the case where the correlation between these vectors is unknown [12]. Consider two estimates μ_a and μ_b with covariances $\boldsymbol{\Sigma}_a$ and $\boldsymbol{\Sigma}_b$ respectively. The CI algorithm computes an updated covariance matrix as a convex combination of the two initial covariance matrices in the form

$$\boldsymbol{\Sigma}_c^{-1} = \omega \boldsymbol{\Sigma}_a^{-1} + (1-\omega)\boldsymbol{\Sigma}_b^{-1} \tag{7}$$

$$\boldsymbol{\Sigma}_c^{-1}\mu_c = \omega \boldsymbol{\Sigma}_a^{-1}\mu_a + (1-\omega)\boldsymbol{\Sigma}_b^{-1}\mu_b \tag{8}$$

where $0 \leq \omega \leq 1$ with ω computed so as to minimise a chosen measure for the size of the covariance matrix.

This computes the relative alignment between information matrices and produces a conservative local update based on the worst-case correlation between incoming messages. Thus using this conservative update, the fusion removes the need for the division in Eq. 3 which is often analytically intractable for general probability distributions even if the common information is known.

A number of drawbacks of the CI algorithm must be noted. One is its computational cost in optimising the free parameter ω. This can be partially improved (while sacrificing optimality) by optimising over a small discrete set of values for ω. The large error covariance bounds for the algorithm may also result in poor accuracy [17], eg. when the two estimates have the same error covariance [18]. Despite these problems, the algorithm can be extended for GMMs with positive results illustrated in the next section.

2) GMM CI: A simple extension to the CI algorithm, not previously used in any filtering applications, involves a CI between each of the Gaussian components in the two mixtures that are to be fused. This results in $M \times N$ Gaussians where M and N are the number of components in the original mixtures. Additionally, this fusion process allows cycles in the network topology while ensuring mathematical consistency.

A bearing-only tracking simulation was used to numerically verify the GMM CI algorithm. In this example, the feature exhibited a random walk within the x-y plane and was tracked by two stationary sensors. An Integrated Ornstein-Uhlenbeck process model [19] was used in addition to a bearing-only likelihood model with an uncertainty of 5°. Predictions occurred every 0.5 s and an observation was performed every 2 s. Communication occurred every 4 s.

The two nodes were arranged in three different communication topologies: 1. no communication, 2. centralised communication, 3. and decentralised communication. Since the true distribution over the tracked states was unknown, the centralised topology provided the optimal solution with maximum information content for each node. The information content can be measured using the inverse of the Renyi entropy [20], [21]

$$H = \frac{1}{1-\alpha} \log \sum_i p(x_i)^\alpha \tag{9}$$

with $\alpha = 2$.

Fig 2 illustrates the Renyi entropy for each of the nodes configured for the three topologies. It can be seen that the

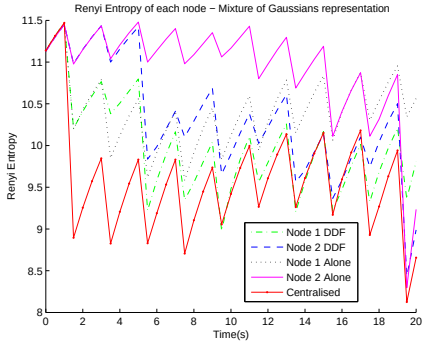

Fig. 2. The Renyi entropy for the two nodes in three communication configurations: no communication, centralised, and decentralised.

information content of the nodes performing DDF is always less (larger Renyi entropy) than the centralised solution. This indicates that even after fusion using the GMM CI algorithm that the decentralised solutions remain conservative. In addition, the decentralised solution is always better than the worst single node solution.

Note that this is only a numerical verification for one particular type of scenario and therefore does not validate the GMM CI algorithm for general applications. However, it is evidence that this algorithm can be used in a decentralised context and for this simulation provides a conservative update for correlated information.

D. Re-Parameterisation

Re-parameterisation arises in filtering problems through a variety of ways. Particle filters require resampling focussing the particles in areas that have most probability density while the multiplicative increase in parameters of mixture representations (GMMs and Parzen density estimates [22]) after each Bayesian operation must be reduced. Modifications to the naive Particle filter provide elegant resampling techniques [23], [24] but as it has been mentioned they do not lend themselves to decentralised fusion algorithms that are mathematically consistent. Parzen density estimates also provide a general probabilistic representation. However, the number of kernels required to estimate multimodal distributions with

different shaped modes can be very large especially in higher dimensions. Thus the accuracy and compactness in estimating the true distribution may be compromised.

Alternatively, it has been shown that fusion can be achieved in a consistent manner with GMMs. Sorenson *et al.* showed that merging of individual Gaussians in a mixture could be achieved while remaining within information theoretic bounds that define the similarity between the true and approximate distributions [9]. In addition, a number of existing techniques from the statistical learning field also provide methods for density estimation for this type of representation. Thus, for GMMs to be viable in the robotics and data fusion domain, these re-parameterisation techniques must be computationally fast and result in an accurate estimate of the true/original distribution.

A common and powerful method is the expectation-maximisation (EM) algorithm which provides a general approach to the problem of maximum likelihood (ML) parameter estimation in statistical models with variables that are not observed [25]. An example of such hidden variables are the underlying mixture components in a GMM. However, EM is sensitive to parameter initialisation and can converge to a local maximum rather than the true value for the maximum likelihood. In addition, the computational complexity of the EM algorithm for GMMs is $O(i \times ND^2)$ where i is the number of iterations performed, N is the number of samples, and D is the dimensionality of the state. Thus, the convergence can be very slow if the initial parameters are particularly bad compared to the true values. However, it has been previously shown [26] that the X-means algorithm [27], [28] (a fast implementation of k-means [29]) results in a reasonable parameter initialisation for the EM algorithm. Thus, ensuring only a few iterations are needed before convergence to the ML is achieved.

E. Data Association

Data association in distributed systems is a complex problem. The reason for this is that hard association decisions made locally, in an optimal manner with respect to local observations, may not be optimal at the global level when all sensor information is made available. Further, an incorrect association decision is almost impossible to undo once data has been fused into a track.

The concept of divergence or distance between two densities underlies measures for the differences between two distributions. The Kullback-Leibler distance and the mutual information are common distances used in data association. However, exact solutions to them are often intractable both analytically and numerically for general probability distributions due to the required division for calculating the distance. Numerical instabilities arise when samples from the denominator have very small values but can be alleviated by Laplace smoothing [30].

An alternative distance is the Bhattacharya coefficient which provides a measure that is numerically tractable:

$$D_B(P_1(\mathbf{x})||P_2(\mathbf{x})) = \int \sqrt{P_1(\mathbf{x})P_2(\mathbf{x})}\, d\mathbf{x} \qquad (10)$$

and is equal to one when the two distributions are the same and zero when there is an infinite distance between them.

V. Natural Visual Feature Selection

The generality of the data fusion techniques presented in this paper ensure that there is flexibility in the feature selection scheme that can be used. The natural feature selection approach adopted in this work is aimed at being simple enough that it validates the DDF algorithms that have been developed in this paper for unstructured environments. The intention is not to demonstrate the actual performance of the extraction itself since this approach can be modified in a number of ways to improve robustness.

The information content of noisy sensory data is assumed to be inversely proportional to the probability of occurrence [31]. Thus less frequent states of a random variable provide greater information than more likely ones *i.e.* they are more unique for that particular set of data. Note that this assumption is purely heuristic in the context of feature extraction but for the following simulations proved to be adequate.

The information content within image regions is computed through visual cue histograms involving colour, hue and texture. Feature selection is subsequently performed by explicitly extracting the least frequent (maximally informative) image pixels. The feature selection is very attractive as the information content in natural imagery can be computed in near real time.

Any feature extraction scheme identifies specific regions in an image that exceed a general information threshold. Each such feature is comprised of several pixels in general, and each pixel can be described by the raw color intensities, multi-scale texture and other visual cues (e.g. intensity gradient, brightness gradient, texture gradient).

In a practical implementation, each extracted feature (a dark red contiguous area in the left image of Fig. 3) is sub-divided into image patches of a fixed size (e.g. 11×11). The right image of Fig. 3 shows the particular patches representing the centroids of the extracted features.

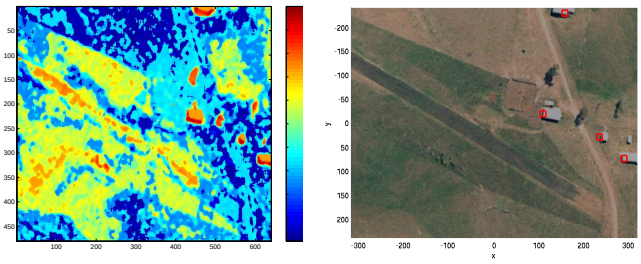

Fig. 3. Left: Information content in hue and texture space. A two dimensional colour histogram of the raw red and green intensities in the image was generated. Subsequently, the information content of each pixel was computed using the formula $log(\frac{1}{p})$, where p is the probability of occurrence. Image regions colour coded red in this figure are maximally informative. Right: The feature patches that are subsequently extracted from the centroid of contiguous highly informative areas.

Texture and colour information were used in the feature extraction process [32] by convolving 11×11 pixel patches in

the RGB colour space with a bank of Gabor wavelets [33] at 2 scales and 2 orientations. This results in an 847 description vector in the observation space.

A. Kinematic States of the Visual Features

To produce a map for navigational purposes, the position of the visual features must be described. Using a single visual sensor requires the range of a static feature to be inferred from two images taken from different positions. In this paper a three dimensional bearing only likelihood model was used for position estimation [19].

To ensure that the likelihood model fits into the framework described throughout this paper, a GMM was learnt offline to approximate the true distribution. This was achieved by sampling from the true likelihood model over a range of 350m with a bearing uncertainty of 1°. To ensure accuracy of the resultant model, the cutoff at 350m was smoothed using a Gaussian falloff. The GMM was then learnt using the EM algorithm with 20 Gaussian components and initial means spaced equally over the range. This model was then used without modification in the online implementation.

Illustration of the accuracy in which a GMM can represent a 3D bearing only likelihood can be achieved through sampling from both the original likelihood model and comparing it to samples from the model learnt through EM (Fig. 4). The

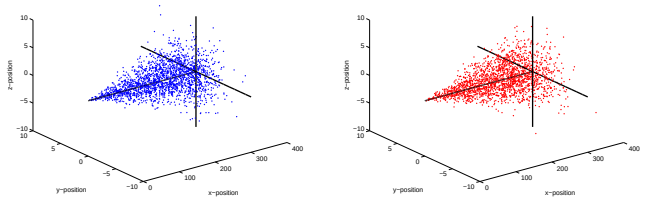

Fig. 4. Comparison of samples from the true bearing-only likelihood model (left), with a range cutoff at 350m and a bearing uncertainty of 1°, and a GMM approximation (right) learnt offline using EM. The GMM model consisted of 20 Gaussian components.

Bhattacharya distance between these two sets of samples is 0.95 illustrating that the GMM is a good approximation to the true model.

VI. Results

To demonstrate the feasibility of including natural visual features in standard filtering methods and in the DDF architecture described earlier, a set of simulated experiments were performed on real visual data obtained from a geo-referenced aerial photograph of the Australian Centre for Field Robotics' test site in Marulan, Australia (Fig. 5). The intention for future work is to demonstrate this on actual autonomous air vehicles and thus sequences of image frames were extracted from the aerial image simulating two simultaneous flights. The altitude of the flights was 200m above ground level, however this information was not included in the filtering algorithms. Thus

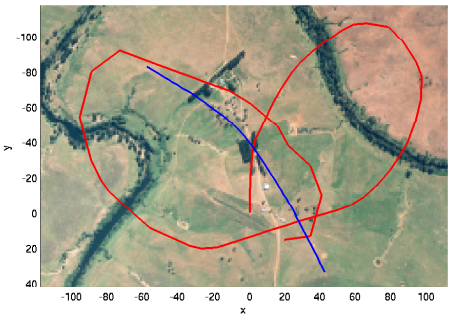

Fig. 5. Geo-referenced aerial image of the Australian Centre for Field Robotics' test site in Marulan, Australia. The two lines are the simulated flight trajectories for two air vehicles. This figure and subsequent figures are best viewed in colour.

localisation of features was purely achieved through a bearing-only model.

The velocity of each vehicle was approximately 100 km.hr^{-1}. The frame rate of the cameras was 20 Hz with a field of view of 60°. The resolution of each image was 640×480 for an area of $300\,\text{m} \times 250\,\text{m}$. Thus one pixel covered an area of approximately $5\,\text{m}^2$.

A. Single Vehicle Results

A typical image from one of the simulated vehicles is shown in Fig. 6. Overlayed on the left most image are samples taken from the likelihood distribution of the observation for one of the patches shown in Fig. 3. The samples from the likelihood model of the observation from the next image for the same patch is also shown. On the right, side views of the same observations from ground level are shown. Note that the samples from the two observations cross at ground level indicated by a height of zero here.

Using the Bhattacharya distance, data association is performed between these two observations, and after the match is confirmed an update is performed with the result shown in Fig. 7. It can be seen that the uncertainty in the range has now

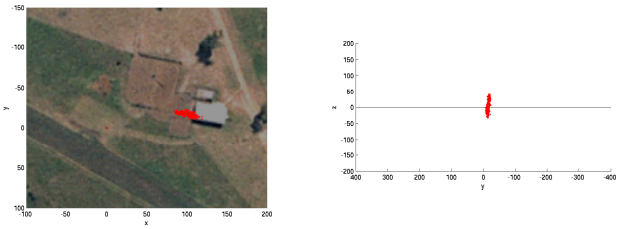

Fig. 7. Top and side view of the samples from the probability distribution over the feature patch after an update was performed between two observations.

significantly reduced. Further observations of this patch will also result in a data association match and after updating, the estimate in range will have converged further.

It was found that the threshold for defining a match during data association was very low and is actually the case for all

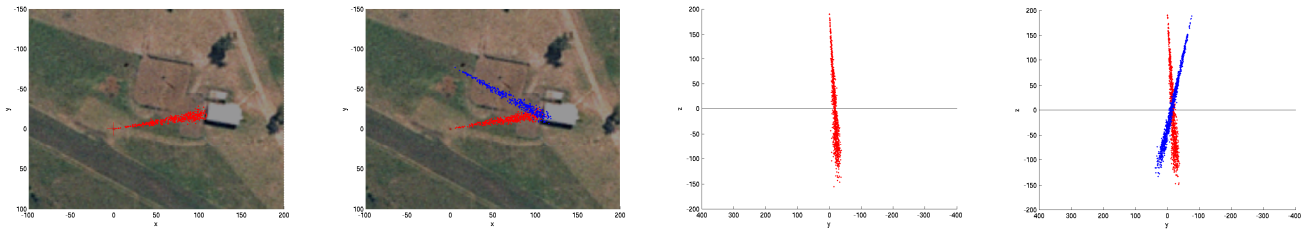

Fig. 6. Left: Samples from a bearing-only likelihood observation of a feature patch extracted from a shed. Inner left: Second observation of the feature after the vehicle has moved. Right and inner right: Side view of the same observations. Note the conical shape of the distribution.

information measures. The reason for this was that there is actually little similarity between the large likelihood and the updated distribution. A geometrical data association method which defines the amount of overlap and not just the similarity may help to resolve this issue.

Fig. 8 illustrates the extracted features from a larger area represented as yellow squares for vehicle 1 and cyan circles for vehicle 2. The updated distributions for individual feature

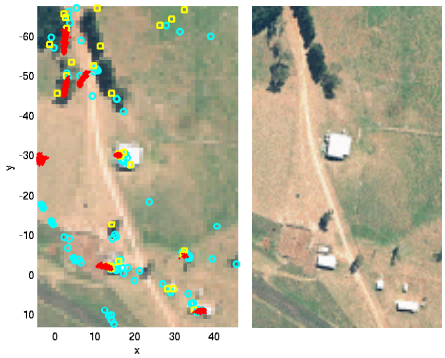

Fig. 8. Part of the area traversed by both nodes with features represented as yellow squares for node 1 and cyan circles for node 2. Samples from the distributions representing features that were observed more than once are displayed in red. Some features are not initiated as tracks due to the small number of times they are observed. Note that the persistent features are from objects that are quite distinct such as the sheds and bush.

tracks of the two vehicles are represented as red samples. Features that have been observed less frequently are not initiated as individual tracks, resulting in a relatively sparse set of features used for filtering.

The updated distributions actually have very accurate bearing uncertainty and converge quickly to the size of the 11×11 patch. For the size of the actual objects of interest, these uncertainties are too small and future improvements to the visual feature representation [32] will hopefully account for this problem.

In addition, there are some spurious features that are initiated as tracks. This is due to poor data association and should be improved with more sophisticated probabilistic visual state models that can be incorporated in a Bayesian filtering framework [32]. Pattern matching from frame to frame should also help local data association problems.

B. DDF Results

The demonstration of fusion of information from two nodes that have observed different features are shown in the following figures. Fig. 9 illustrates the features that have been initiated as tracks for each of the vehicles (red samples for vehicle 1 and blue samples for vehicle 2) before any communication was performed. Notice that each of the vehicles have observed features in different areas. The right hand side pictures illustrate the features tracked by the individual nodes after they have communicated. It can be seen that they both have the exact same posterior representation of the area once communication has been performed. Thus, even if one of the vehicles fail or communication breaks down, they both have a map of the global area without relying on a centralised fusion processor.

These results demonstrate that DDF can be performed using general probabilistic models and visual states of natural features rather than pure range observations. Due to space constraints, analysis and accuracy of this generalised DDF architecture will be performed in a separate technical article.

VII. Conclusion

A scalable, robust, and non-Gaussian decentralised data fusion architecture has been presented. In contrast to previous research, only visual sensory data of natural, unstructured features was used. Gaussian mixture models allowed the position states of the visual features to be incorporated in a general Bayesian estimation framework. Unlike particle representations, GMMs allow the common information to be accounted for using a generalised Covariance Intersect method. Numerical results for these simulations indicate that the fusion between nodes is mathematically consistent but further detailed analysis must be performed.

Further research will concentrate on improvements to computational speed of the re-parameterisation. Possible avenues to consider would be variational Bayesian EM techniques [34] or the use of heuristics to improve the initialisation of EM. Comparisons with component reduction algorithms such as those described by Sorenson *et al.* [9] also need to be performed.

There is still much to incorporate in the visual feature representation such as performing Bayesian filtering on the actual visual states such as colour and texture [32]. Combined with

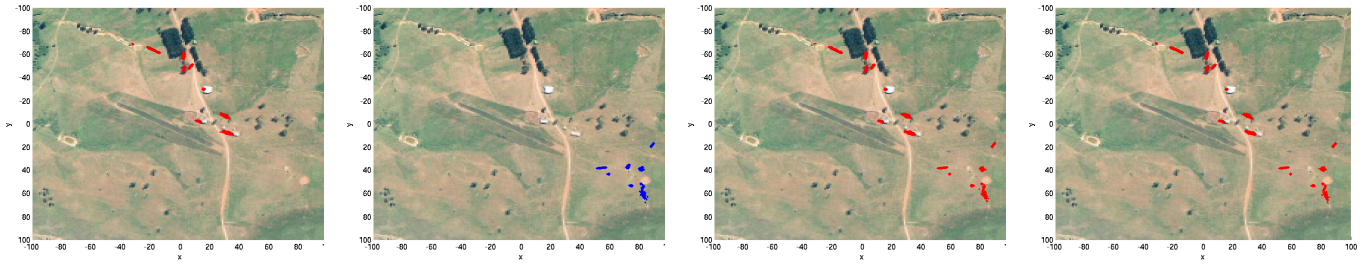

Fig. 9. Left: Samples from the posterior distributions over observed features of node 1 before communication (plotted in red). Inner left: Samples from the posterior distributions over observed features of node 2 (plotted in blue). Right and inner right: Samples from fused, *i.e.* after communication, posteriors from both nodes.

frame to frame pattern matching techniques, major advances in data association is expected.

Although there is much investigation still to be performed in this area, it has been demonstrated that rich visual features can be used in general decentralised data fusion and it is hoped that further research is encouraged in developing natural visual feature representations for Bayesian estimation.

REFERENCES

[1] H. Hashemipour, S. Roy, and A. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Trans. Automatic Control*, vol. 33, pp. 88–93, 1988.

[2] H. Nii, "Blackboard systems," *AI Magazine*, 1986.

[3] M. Rosencrantz, G. Gordon, and S. Thrun, "Decentralized sensor fusion with distributed particle filters," *In Proceedings of the Conference on Uncertainty in AI (UAI), Acapulco, Mexico*, 2003.

[4] E. Nettleton, P. Gibbens, and H. Durrant-Whyte, "Closed form solutions to the multiple platform simultaneous localisation and map building (slam) problem," *Sensor Fusion: Architectures, Algorithms, and Applications IV*, pp. 428–437, 2000.

[5] P. Maybeck, "Stochastic models, estimation, and control, volume 3," *Academic Press Inc, New York*, 1979.

[6] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings F*, vol. 140, pp. 107–113, 1993.

[7] M. A. Paskin and C. E. Guestrin, "Robust probabilistic inference in distributed systems," *In Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence (UAI-04)*, 2004.

[8] A. Ihler, J. Fisher III, R. Moses, and A. Willsky, "Nonparametric belief propagation for self-calibration in sensor networks," *Information Processing in Sensor Networks, Berkeley, California, USA*, 2004.

[9] H. Sorenson and D. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, p. 465, 1971.

[10] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. AC-17, p. 439, 1972.

[11] D. Alspach, "A Gaussian sum approach to the multi-target identification-tracking problem," *Automatica*, vol. 11, p. 285, 1975.

[12] S. Julier and J. Uhlmann, "A nondivergent estimation algorithm in the presence of unknown correlations," *In Proceedings of the American Control Conference*, vol. 4, pp. 2369–2373, 1997.

[13] Y. Bar-Shalom, Ed., *Multitarget-multisensor tracking: advanced applications*. Artech House, 1990.

[14] Y. Bar-Shalom and X. Li, *Multitarget-multisensor tracking: principles and techniques*, 1995.

[15] S. Grime and H. Durrant-Whyte, "Data fusion in decentralised sensor networks," *Control Engineering Practice*, vol. 2, pp. 849–863, 1994.

[16] R. Muirhead, *Aspects of multivariate statistical theory*. Wiley, 1982.

[17] D.Nicholson, S.Julier, and J.Uhlmann, "DDF: an evaluation of covariance intersection," *In Proceedings of the Conference on Information Fusion 2001*, 2001.

[18] C. Chong and S. Mori, "Convex combination and covariance intersection algorithms in distributed fusion," *In Proceedings of Information Fusion 2001*, 2001.

[19] L. Stone, C. Barlow, and T. Corwin, "Bayesian multiple target tracking," *Artech House, Norwood*, 1999.

[20] A. Renyi, "On measures of entropy and information," *In Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, p. 547, 1961.

[21] J. Kapur, *Measures of information and their applications*. Wiley, 1994.

[22] E. Parzen, "On estimation of a probability density function and mode," *Computational Statistics and Data Analysis*, vol. 3, p. 1065, 1962.

[23] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, p. 174R188, 2002.

[24] A. Doucet, N. Freitas, and N. Gordon, "An introduction to sequential monte carlo methods," *Sequential Monte Carlo Methods in Practice, Springer-Verlag*, p. 3R14, 2001.

[25] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B*, vol. 39, pp. 1–37, 1977.

[26] B. Upcroft, S. Kumar, M. Ridley, L. L. Ong, and H. Durrant-Whyte, "Fast re-parameterisation of Gaussian mixture models for robotics applications," *In Australasian Conference on Robotics and Automation 2004 (ACRA '04), Canberra, Australia*, 2004.

[27] D. Pelleg and A. Moore, "X-means: Extending k-means with efficient estimation of the number of clusters," *In Proceedings of the 17th International Conference on Machine Learning*, p. 727, 2000.

[28] ——, "Accelerating exact k-means algorithms with geometric reasoning," *In Proceedings of the 5th International Conference on Knowledge Discovery in Databases*, p. 277, 1999.

[29] D. MacKay, "Information theory, inference, and learning algorithms," *http://www.inference.phy.cam.ac.uk/mackay/itila/*, 2003.

[30] M. Rosencrantz, G. Gordon, and S. Thrun, "Learning low dimensional predictive representations," *In Proceedings of the Twenty-First International Conference on Machine Learning, Banff, Alberta, Canada*, 2004.

[31] T. Cover and J. Thomas, *Elements of Information Theory*. J. Wiley and Sons, 1991.

[32] S. Kumar, F. Ramos, B. Upcroft, M. Ridley, L. L. Ong, S. Sukkarieh, and H. Durrant-Whyte, "A stochastic model for natural feature representation," *In IEEE Conference Proceedings on Information Fusion 2005*, 2005.

[33] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America*, vol. 4, p. 2379R2394, 1987.

[34] Z. Ghahramani and M. Beal, "Variational inference for Bayesian mixtures of factor analysers," *In Advances in Neural Information Processing Systems*, vol. 12, pp. 449–455, 2000.