

Non-Gaussian State Estimation in an Outdoor Decentralised Sensor Network

Ben Uproft, Bertrand Douillard, Tobias Kaupp, Matthew Ridley, Lee-Ling Ong, Suresh Kumar, Tim Bailey, Fabio Ramos, Salah Sukkarieh, and Hugh Durrant-Whyte

Abstract—This paper presents the development and demonstration of non-Gaussian, decentralised state estimation using an outdoor sensor network consisting of an autonomous air vehicle, a manual ground vehicle, and two human operators. The location and appearance of landmarks were estimated using bearing only observations from monocular cameras. We show that inclusion of visual and identity information aids validation gating for data association when geometric information alone cannot discriminate individual landmarks. The combination of geometric, appearance, and identity information provided a common description (or map) of natural features for each of the nodes in the network. We also show the final map from the live demonstration which includes position estimates and classification labels of the observed features.

I. INTRODUCTION

This paper aims to demonstrate non-Gaussian algorithms for *decentralised* state estimation using geometric information with the aid of visual and identity information for validation gating and data association. A common description (or map) of natural features in an unstructured environment was created online by an outdoor sensor network consisting of an autonomous air vehicle, a manual ground vehicle, and two human operators (Fig. 1). Unique aspects of this demonstration were (1) use of colour information from vision sensors to augment geometric information resulting in a rich world map, (2) integration of human operators as information sources, and (3) decentralised operation enabling a practical system which was robust, modular, and scalable.

Applications that benefit from multi-sensor data fusion include environmental sensing, surveillance, and search-and-rescue [1], [2], [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. The fusion approach adopted here is motivated by the need to survey and map large outdoor natural environments in which distributed sensor networks are prone to node and/or communication failure. In contrast to hierarchical and centralised distributed methods [4], [5], decentralised architectures ensure robustness to these failures while allowing scalability and modularity [6]. Three basic constraints are required to guarantee these properties: 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, and 3) Each node has



Fig. 1. The platforms used in the network: autonomous air vehicle, ground vehicle, and human operator. Close-ups of the sensor payloads are shown in the insets.

knowledge only of its immediate neighbours - there is no global knowledge of the network topology.

Previous approaches to robust decentralised data fusion have included tracking position features provided by range devices such as radar or laser [2], tracking artificial visual features with known range [7], monitoring temperature [1] or people movement [8] in an office environment. Nettleton *et al.* showed that scalable decentralised state estimation with Gaussian noise can be achieved in outdoor environments using autonomous air vehicles observing features with known dimensions [7]. 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 framework using a general algorithmic architecture applicable to many multi-sensor situations. They implemented a Mote network for temperature sensing although only Gaussian noise was considered and a tree topology over the network was enforced [1].

Ihler *et al.* demonstrated that non-parametric distributions could be used for sensor-calibration in a network with non-parametric belief propagation [9]. 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 where range and bearing observations from a laser were fused with a particle filter [2].

Unlike the above approaches, we concentrate on modelling *natural* features at three levels: *geometric*, *identity*, and *appearance* [10]. At the geometric level we have extended previous decentralised research to include bearing only, *visual* observations of natural features rather than range-bearing. Gaussian mixtures were used to model the non-Gaussian uncertainty, first demonstrated in tracking problems by Alspach and Sorenson [11], [12]. Alspach also extended these ideas to multi-target identification [13].

Observations of identity states (class labels), with corresponding geometric observations, allowed human perceptual information to be fused into the network. Probabilistic fusion of human and robotic information sources has been previously addressed only in theory or with non-probabilistic human observations [14], [15], [16]. Identity could also be inferred from appearance states which improved classification and hence data association. Real-time classification using non-Gaussian probabilistic, visual/appearance models has recently proven to be accurate [17], [18], [19]. We show that augmenting the geometric states of features with these models increases the robustness of validation gating for track-track association.

The following sections detail the algorithms and illustrate results from a live demonstration. Sections II and III describe the decentralised non-Gaussian fusion algorithms used in this work. Sec. IV describes the feature extraction and appearance modelling of landmarks observed through a monocular camera. Sec. V shows the addition of class labels to the appearance model allowing human operators to input information into the network. Finally results are shown in Sec. VI with conclusion and future work in Sec. VII.

II. DECENTRALISED BAYESIAN ESTIMATION

The algorithmic structure in decentralised estimation is the same for every node in the network and is outlined in Fig. 2. Each sensor makes an observation over which a likelihood is generated. Data association is then performed with existing local tracks where either fusion or track initialisation takes place. Information can then be communicated to other nodes via the channel filter and subsequently assimilated.

The following sections describe the general Bayesian filter so as to introduce notation and from which specific GMM algorithms can be derived.

A. Local Filter

The recursive update after an observation is given by Bayes theorem:

$$p(\mathbf{x}_k | \mathbf{Z}^k) = \frac{p(\mathbf{z} = z_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{Z}^{k-1})}{p(z_k | \mathbf{Z}^{k-1})} \quad (1)$$

where \mathbf{Z}^k are observations of a state \mathbf{x}_k at time t_k , $p(\mathbf{z} | \mathbf{x}_k)$ is the likelihood model, $p(\mathbf{x}_k | \mathbf{Z}^{k-1})$ is the prediction from

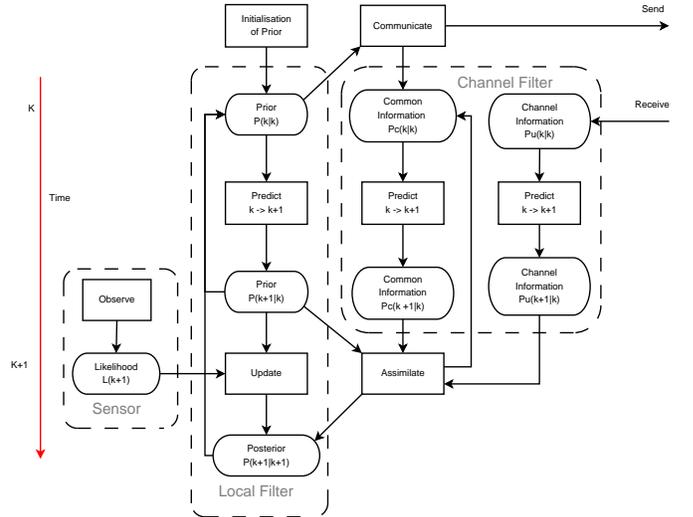


Fig. 2. Flow chart of the operations performed in distributed estimation.

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

The local prediction step is given by the Chapman-Kolmogorov equation:

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

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

B. Node-to-Node Fusion

It can be shown that fusion of the raw correlated information between nodes i and j is [20], [21]

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

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. The incorporation of redundant information in DDF systems may lead to bias, over-confidence, and divergence in estimates.

III. GAUSSIAN MIXTURE MODEL BAYESIAN FILTER

A number of representations have previously been used for multi-modal stochastic models in filtering and tracking. These include particles [22], kernel density estimates [23], and mixture models [11]. Although particle representations are powerful non-Gaussian representations, they suffer from the curse of dimensionality. Kernel density estimates also require many components for accuracy in high dimensions.

In this work, we use Gaussian mixture models (Gaussian sum approximations) which require fewer components for accurate estimation but have the disadvantage of full covariances needing to be specified. Additionally, GMMs provide a basis for analytical solutions to the general Bayesian filtering problem [11]. The following will briefly describe the algorithms for DDF with GMMs.

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

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

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. Local Filter

Substitution of GMMs into Bayes Theorem (Eq. 1) gives

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

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 $p(\mathbf{x}_k | \mathbf{z}_{k-1})$. 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.

GMMs also 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 a Gaussian of the form [24]

$$\pi \mathcal{N}(\mu_a + \mu_b, \Sigma_a^2 + \Sigma_b^2) \quad (6)$$

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

The multiplicative expansion in parameters requires a component merging technique to keep the above operations tractable. We use Salmond's algorithm [25] which results in relatively accurate estimates.

B. Fusion - Pairwise Component Covariance Intersect

A non-optimal solution for node-to-node fusion of Gaussian representations is the Covariance Intersect (CI) filter which conservatively combines the information in two incoming channels assuming that the correlation is unknown [26]. Here, we use Gaussian mixture models (GMMs) and a variant of the Covariance Intersect algorithm [27]. However, as in the work of Ihler *et al.* [9] and Rosencrantz *et al.* [2] divergent solutions are possible.

1) *Covariance Intersect Filter*: Consider two estimates μ_a and μ_b with covariances Σ_a and Σ_b respectively. The CI algorithm computes an updated covariance matrix as a convex combination of the two initial covariance matrices in the form

$$\Sigma_c^{-1} = \omega \Sigma_a^{-1} + (1 - \omega) \Sigma_b^{-1} \quad (7)$$

$$\Sigma_c^{-1} \mu_c = \omega \Sigma_a^{-1} \mu_a + (1 - \omega) \Sigma_b^{-1} \mu_b \quad (8)$$

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

The resultant estimate is based on all possible correlations and thus removes the need for the division in Eq. 3.

2) *GMM CI*: An extension to the CI algorithm involves a pairwise CI update between each of the Gaussian components in the two mixtures that are to be fused. The weight update for each component is given by

$$\pi_c = \alpha \pi_a \pi_b \quad (9)$$

where

$$\alpha = \frac{1}{(2\pi)^{D/2} |\Sigma_\omega|^{1/2}} e^{-1/2(\mu_a - \mu_b)^T \Sigma_\omega^{-1} (\mu_a - \mu_b)} \quad (10)$$

is the scaling constant resulting from the multiplication of two Gaussians, D is the dimension of the space, and $\Sigma_\omega = \Sigma_a/\omega + \Sigma_b/(1 - \omega)$.

We have found that this update remains non-divergent for all the practical scenarios we have encountered and is always better than a straight multiplication in which the common information is not accounted for at all. Although this conservative behaviour is not guaranteed we hope that in future work convergence bounds can be defined.

IV. FEATURE EXTRACTION AND MODELLING

Monocular vision sensors were used as the only robotic information sources in our demonstrations. This resulted in difficulties discriminating between landmarks due to overlapping bearing-only observations. Thus, validation gating based only on position information had a high likelihood of failing. We show here that a visual representation of the landmarks can be maintained in addition to a position representation improving track-track association. The following sections outline the feature extraction preprocessing stage and how the features can be used to estimate position and appearance.

A. Feature Extraction

Feature extraction was performed with colour template matching [28]. Off-line, 3-D template colour histograms were created from manually selected 11×11 image patches. Online, each image was scanned for matching patches. A threshold for the value of the Bhattacharyya coefficient [29] was used to determine how closely a given patch resembled the template. The left hand side of Fig. 3 illustrates the value of the Bhattacharyya coefficient for a "tree" template and all the patches in the image on the right hand side.

Contiguous patches above a given threshold were deemed a landmark of which the centroid was assigned bearing and elevation coordinates. The geometric states could then be

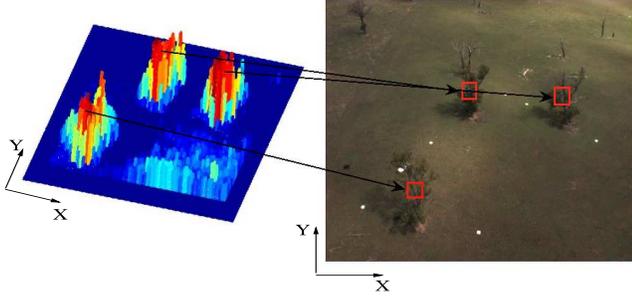


Fig. 3. Left: The value of the Bhattacharyya coefficient assigned to each patch from the image on the right using a template picked from a tree (red indicates a value of 1 while blue indicates a value of 0). Right: An example image obtained from a UAV. The arrows indicate the correspondence between the Bhattacharyya coefficient values and the extracted patches marked with squares.

estimated using a bearing-only likelihood model [30]. The model was approximated in Cartesian space with a GMM and was learnt off-line using Expectation-Maximisation (EM) [27]. Online observations only required rotation of the GMM to the specific bearing and elevation coordinates which could then be locally fused using the equations outlined in Sec. III.

B. Appearance Model

The colour histograms from the feature extraction also doubled as appearance observations \mathbf{z} . Each histogram contained 9^3 bins (729 dimensions) representing the RGB information. The high dimensionality of these observations required a compression step which could be achieved through a number of dimensionality reduction algorithms including principal components analysis (PCA) [31], multidimensional scaling (MDS) [32], Locally Linear Embedding [33], Isomap [34], and many variants. We chose Isomap as it preserves the neighbourhood of points in the low dimensional manifold and automatically estimates the number of dimensions required to retain relevant information.

However, Isomap and indeed most nonlinear dimensionality reduction algorithms are inherently deterministic. They do not provide a measure of the uncertainty of the underlying states \mathbf{x} , from noisy high dimensional observations. In addition, all the original training data must be stored to compute the embeddings for new observations. We overcome this by using the results of Isomap to train a generative probabilistic model which encapsulates the uncertainties inherent in the inference of a low dimensional state \mathbf{x} , from noisy high dimensional observations \mathbf{z} .

The probabilistic model can be used to map inputs to outputs and vice versa by computing the expected values $E[\mathbf{z}|\mathbf{x}]$ and $E[\mathbf{x}|\mathbf{z}]$. We define a joint distribution similar to a mixture of factor analysers, commonly used in the machine learning community to perform simultaneous clustering and local dimensionality reduction [35]. The only differences are that the low dimensional variable \mathbf{x} is observed, not hidden, and the Gaussian distributions $p(\mathbf{x}|\mathbf{s})$ have nonzero mean vectors ν_s , and full covariance matrices Σ_s . The graphical model in Fig. 4 depicts the assumed dependencies. The

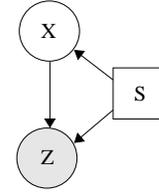


Fig. 4. Graphical model depicting the dependencies in the joint probability distribution for appearance.

discrete hidden variable \mathbf{s} introduced in the model physically represents a specific neighbourhood on the manifold over which a mixture component is representative. This representation conveniently handles highly nonlinear manifolds through the capability of modelling the local covariance structure of the data in different areas of the manifold. It can be trained with very large datasets and the computational cost of inference is independent of the number of training samples.

The following outlines the specific parameterisation for the model:

$$p(\mathbf{z}, \mathbf{x}, \mathbf{s}) = p(\mathbf{z}|\mathbf{x}, \mathbf{s})p(\mathbf{x}|\mathbf{s})p(\mathbf{s}) \quad (11)$$

$$p(\mathbf{z}|\mathbf{x}, \mathbf{s}) = \frac{1}{(2\pi)^{D/2} |\Psi_s|^{1/2}} \times e^{-\frac{1}{2}(\mathbf{z} - \Lambda_s \mathbf{x} - \mu_s)^T \Psi_s^{-1} (\mathbf{z} - \Lambda_s \mathbf{x} - \mu_s)} \quad (12)$$

$$p(\mathbf{x}|\mathbf{s}) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_s|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x} - \nu_s)^T \Sigma_s^{-1} (\mathbf{x} - \nu_s)} \quad (13)$$

where μ_s and Ψ_s are the means and covariances of the mixture describing the high dimensional space, ν_s and Σ_s are their counterparts in the low dimensional space, Λ_s are known as the loading matrices and model the transformation between the two spaces, $\mathbf{s} \in \{1, \dots, N\}$ where N is the number of components of the mixture.

Given the results of Isomap, the parameters for the joint distribution $p(\mathbf{z}, \mathbf{x}, \mathbf{s})$, can be learnt off-line using Expectation Maximisation (EM) [36], [37]. Note that learning with both low and high dimensional data \mathbf{x} and \mathbf{z} , appears to improve the solution compared to just high dimensional data [38].

Once the model has been learnt, real-time inference can be performed online. A threshold for a given probabilistic distance measure (in our case the Bhattacharyya coefficient) can then be used for validation gating.

V. HUMAN OPERATORS

Human operators are considered an integral part of the sensor network. They can play several roles when interacting with the network [10]. Here, we regard the operator as an information source contributing both geometric and visual observations [39]. Two methods for location entry were used: 1) Operators entered observations on a graphical user interface (GUI) displaying an aerial map obtained prior to the actual demonstration and 2) by estimating the range and bearing of a feature relative to their own position. Raw range and bearing observations were converted using a human

sensor model learnt off-line and thus both location entry methods resulted in a probabilistic estimate of the same form as the robotic representation.

Entry of visual information ultimately influenced the inference in the joint probabilistic model. Visual information was entered in the form of identity state observations represented by a random variable o , as shown in Fig. 5. The sensor model $P(o|s)$ was specified manually [39]. Here we show

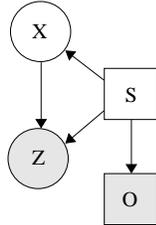


Fig. 5. Graphical model with the additional identity observations.

a correspondence between mixture components and feature identity leading to the interpretation of the s node as an identity label. This enabled us to classify a subset of the data *a priori* with human interpretable labels. Maximum Likelihood could then be used to learn the parts of the model with labelled data while EM could still be used for unlabelled data. With the addition of identity labels, human operators could input meaningful observations aiding inference for appearance.

Figure 6 shows low-dimensional data points for outdoor features from Isomap with the learnt covariances from the above model. Each 729-dimensional patch z , was reduced to

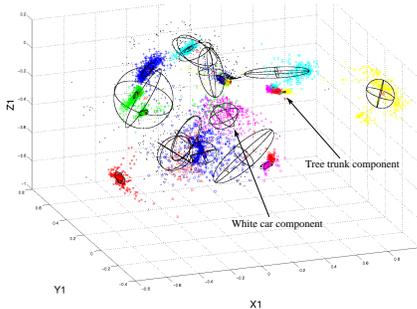


Fig. 6. The low dimensional embedding from Isomap with the learnt covariances from the probabilistic model. The data contained 12388 patches from outdoor features. The Isomap residual also indicated that 3 dimensions would be adequate for retaining the relevant information.

a 3-dimensional representation x as indicated by the residual of Isomap. The number of components was chosen to be 27 with 17 of them manually assigned a label. Two of the labelled components are shown in the figure.

VI. RESULTS

The algorithms were demonstrated using a four node network with three of the platforms illustrated in Fig. 1. Both vehicles were equipped with Global Positioning System (GPS) and Inertial Measurement Unit (IMU) sensors in

addition to a single vision sensor. The two human operators were able to input and receive information using tablet PCs with attached hand-held GPS units. Physical communication between nodes was achieved with standard IEEE 802.11b wireless network adaptors using the UDP protocol. The system architecture as a whole was developed using the Active Sensor Network framework [40] while the software implementation of the system adopted the component paradigm of the Orca robotics project¹. This ensured a modular approach to development and ease of software component interaction and communication which is essential in such a system. Decentralised communication and testing of various probabilistic representations also required the development of a low-level communication software library² and a general library for probabilistic algorithms³.

The experiments were performed at an outdoor test facility over an area of a few square kilometres. The position of each of the vehicles and the fused estimates at each of the nodes could be monitored with an online GUI overlaying a geo-referenced aerial image taken prior to the demonstration. A number of objects such as trees, sheds, and cars were surveyed using differential GPS measurements allowing comparisons to ground truth.

Each platform maintained a bank of decentralised, non-Gaussian Bayesian filters for the features it observed, and transmitted the information to all other platforms. The net result was that each platform maintained a complete map of all features observed by all nodes in the network. Multiple observations of the same feature, possibly by different platforms, resulted in an increasingly accurate estimate of the feature location for *all* nodes.

A. Accuracy of Position Estimates

Fig. 7 illustrates GUI screenshots of a sequence of map updates for observations of a tree from a ground vehicle. The camera was mounted sideways so forward movement automatically increased the baseline between observations. As the vehicle moved past the tree, the updated estimate increased in accuracy and converged to the true location of the object (indicated by the white cross). Note that the label, representing the component with the highest probability inferred from the appearance model, also correctly identified the object.

B. Improved Track-Track Association using Appearance

The ground vehicle shown in Fig. 1 circled around the tree and the red car displayed in Fig. 8(a). Fig. 8 illustrates the difficulty in track to track association when position observations are the only source of information. As can be seen in Fig 8(b), observations of the tree and the red car overlap. Based on position information only, all the observations were fused into one single estimate (Fig. 8(c)). However, when the visual states of the tree and the car were included, two different tracks were maintained (Fig. 8(d)).

¹<http://orca-robotics.sf.net>

²<http://crud.sf.net>

³<http://spasm.sf.net>

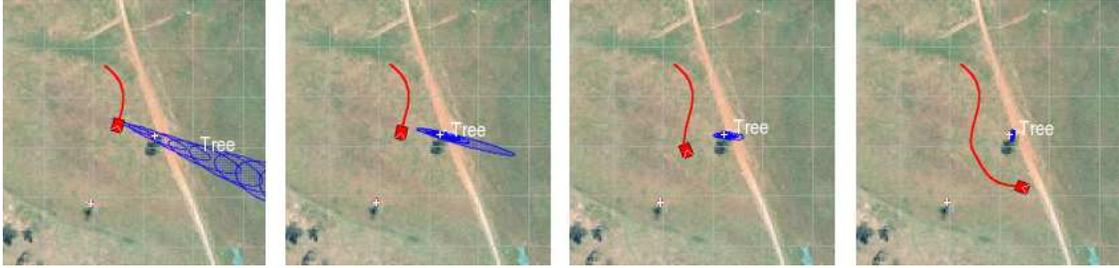


Fig. 7. Surveyed differential GPS locations are marked as white crosses on the aerial photograph. Ellipses of the same colour are 2-D projections of a Gaussian mixture representing the position estimate of a single feature. Labels indicate the component with the highest probability inferred from the appearance model. The ellipses and labels combine to make up the entire map. The red line represents the trajectory taken by the vehicle.

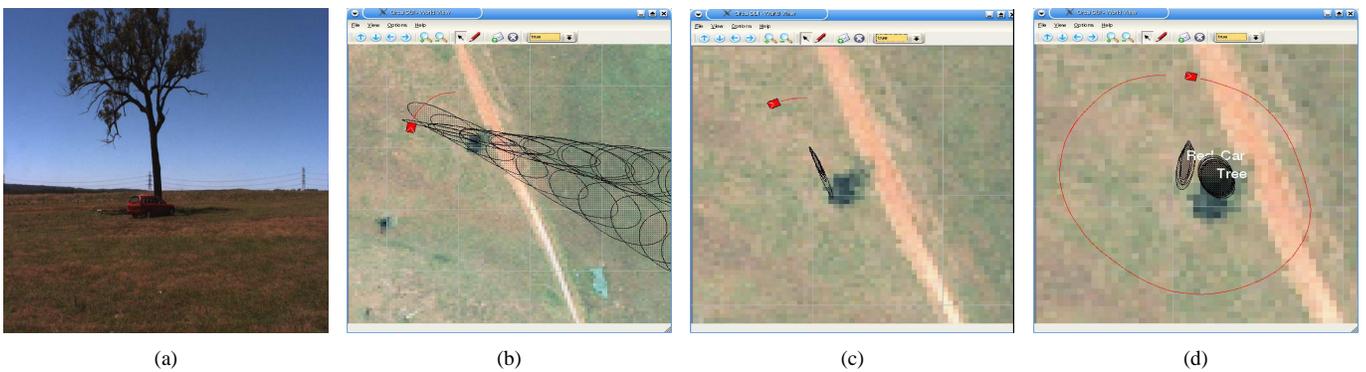


Fig. 8. (a) The two landmarks: a tree and a red car. (b) bearing only observations of these landmarks. (c) Position estimate of the two landmarks without appearance discrimination. Note that only one track was initialised. (d) Position estimate with appearance discrimination. Two tracks were maintained although the locations of the landmarks were very close. The ground vehicle is represented by a red rectangle and its trajectory indicated by the red curve.

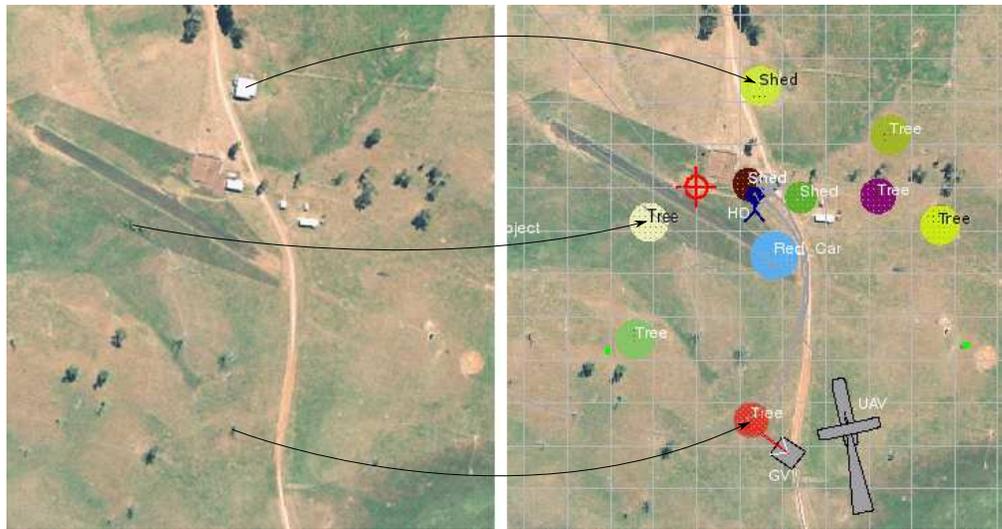


Fig. 9. RHS: The belief of one of the nodes in the network. Each set of coloured ellipses corresponds to a particular feature and the labels represent the identity state with highest probability. The icons represent each of the different nodes in the network; UAV = air vehicle, GV = ground vehicle, HO = human operator. LHS: the original aerial image with arrows highlighting a few of the correspondences with the belief of the node.

C. Decentralised Fusion Map

The map shown in Fig. 9 is a live screenshot of the belief of one of the platforms after multiple nodes entered the network. Each set of coloured ellipses with a corresponding label represents a different feature. Qualitatively, it can be seen that an accurate map with correct feature classification was achieved.

VII. CONCLUSION

We have shown decentralised state estimation with Gaussian mixture models in combination with appearance models to aid in validation gating for data association. The algorithms were demonstrated on a network involving heterogeneous platforms in a large outdoor area. In the future we hope to continue investigating conservative fusion techniques for non-Gaussian representations and define non-divergent bounds for these methods.

REFERENCES

- [1] 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.
- [2] 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.
- [3] F. Bourgalt and H. Durrant-Whyte, "Communication in general decentralized filters and the coordinated search strategy," *In Proceedings of The 7th International Conference on Information Fusion*, 2004.
- [4] H. Hashemipour, S. Roy, and A. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Trans. Automatic Control*, vol. 33, pp. 88–93, 1988.
- [5] H. Nii, "Blackboard systems," *AI Magazine*, 1986.
- [6] J. Manyika and H. Durrant-Whyte, *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. Ellis Horwood, 1994.
- [7] 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.
- [8] A. Brooks, A. Makarenko, T. Kaupp, S. Williams, and H. Durrant-Whyte, "Implementation of an indoor active sensor network," *Proceedings of the 9th International Symposium on Experimental Robotics*, 2004.
- [9] 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.
- [10] T. Kaupp, A. Makarenko, S. Kumar, B. Uprocft, and S. Williams, "Operators as information sources in sensor networks," *In IEEE/RSJ International Conference Proceedings on Intelligent Robots and Systems*, 2005.
- [11] H. Sorenson and D. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, p. 465, 1971.
- [12] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. AC-17, p. 439, 1972.
- [13] D. Alspach, "A Gaussian sum approach to the multi-target identification-tracking problem," *Automatica*, vol. 11, p. 285, 1975.
- [14] G. Pavlin, P. de Oude, and J. Nunnink, "A mas approach to fusion of heterogeneous information," *In IEEE/WIC/ACM Joint Conference on Intelligent Agent and Web Technology*, 2005.
- [15] K. Ross, R. Chaney, G. Cybenko, D. Burroughs, and A. Willsky, "Mobile agents in adaptive hierarchical bayesian networks for global awareness," *In IEEE International Conference on Systems, Man, and Cybernetics*, vol. 3, pp. 2207–2212, 1998.
- [16] D. J. Bruemmer and M. C. Walton, "Collaborative tools for mixed teams of humans and robots," in *NRL Workshop on Multi-Robot Systems*, Washington, DC, USA, 2003, pp. 219–229.
- [17] A. Torralba, K. Murphy, W. Freeman, and M. A. Rubin, "Context-based vision system for place and object recognition," *Proceedings of the IEEE International Conference on Computer Vision, ICCV*, 2003.
- [18] A. Torralba, K. Murphy, and W. Freeman, "Contextual models for object detection using boosted random fields," *In Advances in Neural Information Processing Systems*, 2004.
- [19] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories," *CVPR*, 2004.
- [20] Y. Bar-Shalom, Ed., *Multitarget-multisensor tracking: advanced applications*. Artech House, 1990.
- [21] Y. Bar-Shalom and X. Li, *Multitarget-multisensor tracking: principles and techniques*, 1995.
- [22] 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.
- [23] E. Parzen, "On estimation of a probability density function and mode," *Computational Statistics and Data Analysis*, vol. 3, p. 1065, 1962.
- [24] R. Muirhead, *Aspects of multivariate statistical theory*. Wiley, 1982.
- [25] D. Salmond, D. Atherton, and J. Bather, "Mixture reduction algorithms for uncertain tracking," *IFAC Proceedings Series*, vol. 2, p. 775, 1989.
- [26] 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.
- [27] B. Uprocft, L. L. Ong, S. Kumar, M. Ridley, T. Bailey, S. Sukkarieh, and H. Durrant-Whyte, "Rich probabilistic representations for bearing-only decentralised data fusion," *In IEEE Conference Proceedings on Information Fusion 2005*, 2005.
- [28] M. Nixon and A. Aguado, *Feature Extraction and Image Processing*. Elsevier, 2001.
- [29] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, p. 564, 2003.
- [30] L. Stone, C. Barlow, and T. Corwin, "Bayesian multiple target tracking," *Artech House, Norwood*, 1999.
- [31] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. Wiley-Interscience, New York, 2001.
- [32] T. Cox and M. Cox, *Multidimensional Scaling*. Chapman and Hall, 1994.
- [33] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by Locally Linear Embedding," *Science*, vol. 290, pp. 2323–2326, 2000.
- [34] J. Tenenbaum, V. DeSilva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, 2000.
- [35] Z. Ghahramani and G. Hinton, "The EM algorithm for mixtures of factor analysers," *University of Toronto Technical Report CRG-TR-96-1*, 1996.
- [36] 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.
- [37] F. Ramos, S. Kumar, B. Uprocft, and H. Durrant-Whyte, "Representing natural objects in unstructured environments," *In Advances in Neural Information Processing Systems, Workshop on Machine Learning Based Robotics in Unstructured Environments*, 2005.
- [38] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of nonlinear manifolds," *Technical Report, University of Pennsylvania, Dept. of Computer Science*, 2002.
- [39] T. Kaupp, B. Douillard, B. Uprocft, and A. Makarenko, "Hierarchical model for fusing information from human operators and robots," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'06)*, Beijing, China, 2006.
- [40] A. Makarenko, A. Brooks, S. Williams, H. Durrant-Whyte, and B. Grocholsky, "A decentralized architecture for active sensor networks," *In IEEE International Conference on Robotics and Automation*, 2004.