A Computationally Efficient Low-bandwidth Method for Very-Large-Scale Mapping of Road Signs with Multiple Vehicles

Ashkan Amirsadri, Adrian N. Bishop, Jonghyuk Kim, Jochen Trumpf, Lars Petersson

Abstract—This paper provides a flexible solution to the problem of building and maintaining a very-large-scale map using multiple vehicles. In particular, we consider producing a map of landmarks on the scale of thousands of kilometres in an outdoor environment. The algorithm is distributed across multiple vehicles each given the task of producing and updating a local map. The vehicles are equipped with a range of sensors and selectively communicate maps to and from a central station in a bandwidth-constraint environment. The potentially overlapping local maps are asynchronously transmitted back to a central fusion centre where a global map repository is maintained. The work addresses two of the most common issues of mapping in large-scale environments, namely, computational complexity and limited communication bandwidth. The proposed communication architecture is scalable and is capable of dealing with time-varying overlapping map sizes. A general data fusion framework based on covariance intersection is proposed to tackle the problem of redundant information propagation that is caused by communicating sub-maps of arbitrary size in the network. We also provide an analysis on the applicability of covariance intersection, as compared to the optimal approach when no cross-correlation is known between estimates from different vehicles. We further analyse the solution using a number of illustrative examples.

I. INTRODUCTION

Simultaneous Localisation and Mapping (SLAM) is a popular technique [1]–[3] to address complex mapping problems under conditions of process and sensor noise and possible modeling errors. This algorithm first appeared in a seminal paper by Smith, Self and Cheeseman [4] and it has received a considerable amount of attention by the robotics community [5]–[8]. Broadly speaking, SLAM is the process of concurrently building a map of the environment and using the map to estimate the location of the robot in an unknown environment.

Although most of the initial interest in SLAM considered the problem of mapping and localisation with a single vehicle, the first decade of the twenty-first century saw a substantial interest in multi-vehicle localisation and mapping. The advantages of using multiple, cooperative, vehicles in exploration and mapping applications, compared to the single vehicle case are well known and intuitive, e.g. redundancy, improved accuracy in mapping etc. [9]–[11].

At a high-level, information fusion is the fundamental tool required for multi-vehicle SLAM as, on an abstract



Fig. 1. The red lines on the Australia map indicate the roads from which video has been captured and automatically analysed for road signs. The challenge addressed in this paper is how to further automate this task by having a fleet of ad-hoc surveying vehicles efficiently communicating their observations to a central server.

level, the problem is about combining numerous sources of information (that may be correlated) about a common parameter in order to increase ones knowledge about the parameter.

In multi-vehicle SLAM the problem of where this fusion occurs and how information is shared is a practical problem and is one that motivates much of the work in this paper (along with similar work discussed subsequently). Different multi-vehicle data fusion architectures have been suggested and implemented for tasks such as autonomous navigation [12]–[14], exploration and mapping [15]–[17] and target tracking [18], [19].

The most obvious and traditional data fusion architecture is a fully-centralised one where all the raw sensor data from multiple sources is transmitted to a central station for fusion (e.g. using a large Kalman filter). Works such as [20] and [21] provide fully-centralised approaches to the multi-vehicle SLAM problem. The main disadvantage of a centralised solution is the communication and networking complexity required. The central station also offers a single point of failure and thus centralised solutions in general are less redundant and robust. In addition, the sophistication and heavy computational load at the central station might lead to an undesirable computational bottleneck. However, centralised solutions are convenient in numerous practical applications where it is undesirable for the vehicles to communicate between themselves.

In contrast to the centralised systems, fully decentralised architectures often have no central processing station. In such systems, individual stations (e.g. individual vehicles) can perform data fusion in a fully autonomous manner, while receiving information from and transmitting information to other particular stations. In other words, fusion occurs locally

This work is supported by the Australian National University (ANU) and NICTA. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. A. Amirsadri (ashkan.amirsadri@anu.edu.au), A. N. Bishop and J. Kim are with the ANU and NICTA. J. Trumpf is with the ANU. L. Petersson is with NICTA and is an adjunct to the ANU.

at each station on the basis of local observations and the information received from neighbouring stations. Examples of decentralised SLAM can be found in [22]–[24]. Note that if the networking topology resembles a complete graph then such decentralised systems offer no advantage in terms of communication requirements. Of course, in general, decentralised solutions are more robust to failure of a given station.

Both fully-centralised and conventional decentralised architectures have proven to be effective in numerous mapping applications. However, without additional local processing it turns out that both methods fail to provide a practical and flexible solution to large-scale (millions of mapping points) mapping where limited bandwidth and processing power are a real concern. This is particularly true when the constraint of a centralised architecture is dictated by the problem. This will be discussed in more details in Sections II and III.

Despite some fundamental work (e.g. [18], [24]), the problem of selective communication has been widely neglected in the study of multi-vehicle information fusion (e.g. [25]). In large-scale, low-bandwidth mapping applications, sending all the local information to the central station is not feasible due to the limited system and communication resources present in practice. Information tailoring is necessary to avoid high communication costs and other bandwidth constraints in a distributed data collection system. Consequently, only the most valuable information should be selected and transmitted. This is the avenue that we follow in this paper.

In addition, the majority of the existing multi-vehicle SLAM techniques suffer from the growing size of the local maps within individual nodes. Due to the large number of features and the rapidly increasing map size, the SLAM algorithm fails to fulfil the requirements of large-scale applications. The ramification is an immense memory and computational load on the vehicles. Consequently, appropriate strategies must be applied to limit the size of the SLAM filters in very large-scale environments. We discuss a particular pruning strategy in this paper.

The contribution of this work is the development of a multi-vehicle data fusion framework for a real-world inspired road mapping application. We introduce a hierarchical data fusion architecture and a communication scheme that allows the communication of sub-maps of arbitrary size. A practical pruning algorithm based on a measure of information gain is introduced to overcome the problem of progressively growing map sizes at individual vehicles. The communication bandwidth is reduced significantly by selectively transmitting sub-maps with the largest information contribution to a central server, where a global map repository is maintained. The proposed communication architecture is flexible in the sense that it is capable of dealing with dynamically changing mapsizes in the entire system. In addition, the fusion algorithm offered in this paper ensures that map estimates are integrated in a consistent and robust fashion.

II. PROBLEM DESCRIPTION / MOTIVATION

The motivation behind this work is a project called AutoMap where geographically located information from the



Fig. 2. Road Sign Detection and Localisation in AutoMap Project.

road scene is gathered continuously on a very large scale by a fleet of distributed vehicles such as taxis, garbage trucks, delivery vans etc and sent back to a central server where a global database is compiled (Figure 1 gives an indication of the size of the problem that we are addressing). Advanced computer vision algorithms [26] are used to automatically extract and geolocate, e.g. road signs from recorded video that are of interest to third party companies like mapping companies and road asset managers (see Figure 2) [27]. Such information is currently collected in a manual fashion and updated only every few years which is a very costly and error prone process. A setup as described in this paper enables a continuously updated database of road scene information at a fraction of the cost compared to the manual alternative. Each sensor platform, as installed in each fleet vehicle, consists of three cameras, a Global Positioning System (GPS), a 3-axis accelerometer, a 3-axis gyroscope, a 3-axis magnetometer, a processing unit and a 3G modem. Data from the sensors are continuously stored on a local hard drive, and later analysed by the local processing unit in order of importance to maximise a cost function representing the value of extracted information.

A. Resource Constraints

Analysing the vast amount of information gathered from the sensors and transmitting it back to the central server is a challenging task as the platform installed in each vehicle suffers from a number of constraints. These constraints can be categorised as 1) Communication bandwidth 2) Processing power, and 3) Memory and storage. One of the key constraints this paper sets out to address is the limited communication bandwidth provided by the 3G modem. The limited communication bandwidth not only makes it impossible to transmit all raw sensory data to a central server and analyse it there, but even the amount of extracted, symbolic information poses a challenge (see Example 1 below). Clearly, a communication architecture that allows selective communication is needed to handle this case.

Example 1. Consider a scenario with n vehicles collecting measurements and tasked at mapping a given environment. Each vehicle traverses d kilometres per day and each kilometre contains m map objects (road signs) on average. The size of each vehicle's map is given by $N = m \cdot d$ and it is assumed this map size is initialised at the start of the day. A map represented by a covariance matrix

is then assumed to require $b \cdot N^2$ bytes to transmit and the communication cost for each byte is given by c. The communication protocol requires k transmissions of $b \cdot N^2$ bytes per kilometre of road data. A simple calculation shows that the total communication cost using this method is

$$C_{total} = d^3 m^2 n k b c \tag{1}$$

per day. The communication cost (and bandwidth) in this example is proportional to the cube of the distance driven by each vehicle over a fixed period of time. Consequently, the above solution is not feasible for very-large-scale applications like AutoMap which exhibit limited system and communication resources. For example, let n = 10, d = 200, m = 10, k = 0.1, b = 8, $c = \$3 \times 10^{-8}$ (\$30 for 1GB of 3G data) and $N = md = 10 \times 200$, then $C_{total} = \$192$ per day. In this case, the cost of complete communication is prohibitive and a more efficient solution is required. Similar analysis can be done for the processing power and memory requirements.

III. THE HIERARCHICAL DATA FUSION ARCHITECTURE

This paper introduces a single-level hierarchical architecture with a central base station, called the central fusion center (CFC), to combine the local maps obtained from individual vehicles into a global map. As argued in Section II-A, by virtue of the very-large-scale nature of the problem, it is practically unrealistic to process and maintain all the map data locally at individual vehicles.

The hierarchical architecture aims to increase the processing done locally by the individual vehicles. A local SLAM algorithm is implemented in each vehicle in order to build a local map of the detected landmarks and concurrently estimate the location of the vehicle as it explores the environment (see Figure 3). Each vehicle shares a selection of its local information with the CFC (via a cellular network). The CFC is responsible for maintaining a global map and for integrating the information collected by the vehicles in a consistent fashion (see Figure 4). A feedback configuration in the system provides a route for the communication of submaps of the global map back to the local filters in individual vehicles. As such, individual vehicles indirectly have access to the information obtained by other vehicles in the system.



Fig. 3. The distributed information fusion model keeps the central map repository up-to-date.



Fig. 4. A single-level hierarchical architecture

A. The local SLAM filter (LSF)

The local SLAM filter is a local implementation of the single-vehicle SLAM algorithm. The LSF estimates a state vector and a covariance matrix based on the observed sensor measurements and the information received from the CFC (e.g. as an initial prior). The state and covariance at the LSF in vehicle i is given by

$$\hat{\mathbf{x}}^{i}(k|k) = \begin{bmatrix} \hat{\mathbf{x}}^{i}_{v}(k|k) \\ \hat{\mathbf{x}}^{i}_{m}(k|k) \end{bmatrix}$$
(2)

$$\mathbf{P}^{i}(k|k) = \begin{bmatrix} \mathbf{P}_{vv}^{i}(k|k) & \mathbf{P}_{vm}^{i}(k|k) \\ \mathbf{P}_{vm}^{i}{}^{T}(k|k) & \mathbf{P}_{mm}^{i}(k|k) \end{bmatrix}$$
(3)

where vehicle and map components are denoted by the subscripts v and m respectively. This paper will not dwell on the details of the local SLAM filter as such algorithms have been considered numerous times in the literature [1], [28]. The dimension of $\hat{\mathbf{x}}^i(k|k)$ is different for each i as the local environment (e.g. the number of landmarks observed etc.) is different for each vehicle.

B. Map information

Given a state estimate $\hat{\mathbf{x}}(k|k)$ with covariance $\mathbf{P}(k|k)$, the so-called information vector and information matrix are defined by a bijective mapping

$$\hat{\mathbf{y}}(k|k) = \mathbf{P}^{-1}(k|k)\hat{\mathbf{x}}(k|k), \ \mathbf{Y}(k|k) = \mathbf{P}^{-1}(k|k)$$
(4)

The reason for considering the information-based representation for states and covariances is that the interpretation, communication and the fusion¹ of a group of estimates is more convenient in this form.

From Equation (3) we then define the total map information at each vehicle i as

$$\mathbf{Y}_{mm}^{i}(k|k) = \mathbf{P}_{mm}^{i}{}^{-1}(k|k) \tag{5}$$

Eric Nettleton's thesis [18] provides two important results concerning the cross-correlation between vehicle state estimates $\hat{\mathbf{x}}^i(k|k)$ under some pretty common assumptions. Suppose that the size of $\hat{\mathbf{x}}^i_m(k|k)$ is the same for all *i*; i.e. we can think of $\hat{\mathbf{x}}^i_m(k|k)$ as a local estimate of the complete global map. Also suppose that the association (i.e. ordering) amongst the elements of $\hat{\mathbf{x}}^i_m(k|k)$ is consistent between vehicles and that

$$\mathbf{E}[(\hat{\mathbf{x}}_{m}^{i}(k|k) - \mathbf{x}_{m}(k))(\hat{\mathbf{x}}_{m}^{j}(k|k) - \mathbf{x}_{m}(k))^{\top}] = \mathbf{0} \quad (6)$$

¹For brevity, we typically only discuss those equations for computing the information matrix $\mathbf{Y}(k|k)$ while the corresponding computation of $\hat{\mathbf{y}}(k|k)$ is expected to be known.

where $\mathbf{x}_m(k)$ is the actual map of the environment. Then the covariance of the best, linear unbiased, estimate of the global map is simply

$$\mathbf{Y}_{mm}(k|k) = \sum_{i} \mathbf{Y}_{mm}^{i}(k|k) \tag{7}$$

Moreover, under these assumptions

$$\mathbf{E}[(\hat{\mathbf{x}}_{v}^{i}(k|k) - \mathbf{x}_{v}^{i}(k))(\hat{\mathbf{x}}_{v}^{j}(k|k) - \mathbf{x}_{v}^{j}(k))^{\top}] = \mathbf{0}$$
(8)

where $\mathbf{x}_{v}^{i}(k)$ is the actual i^{th} vehicle location. However, in practice the assumption that

$$\mathbf{E}[(\hat{\mathbf{x}}_m^i(k|k) - \mathbf{x}_m(k))(\hat{\mathbf{x}}_m^j(k|k) - \mathbf{x}_m(k))^{\top}] = \mathbf{0}$$
(9)

is typically not justified and individual vehicle maps $\hat{\mathbf{x}}_m^i(k|k)$ may only partially overlap and be of different sizes. Therefore, the results of Nettleton above are not always applicable (as noted in much of Nettleton's own work, e.g. [24]).

C. Selective communication

Given the practical scenario envisioned for this work, it follows that limited communication bandwidth constrains the transmission of information to and from the CFC. Consequently, the accuracy of the central map should be optimised in some manner as a function of the information sent by the individual vehicles under the limited bandwidth constraints. As discussed later in Subection IV-D, we use the information gain (between the local sub-maps known at the CFC and the improved maps resulting from the local SLAM algorithm) as a measure to select the most informative sub-map within the local SLAM algorithm for communication.

IV. COMMUNICATION SCENARIO

In this section we consider the communication sequence for a single vehicle, e.g. one of the components shown in Figure 4, and discuss the process of information fusion when the shared information between the CFC and an individual vehicle is correlated and of differing sizes. The proposed communication block diagram is shown in Figure 5. Note the addition of a so-called channel filter (CHF); shown at the vehicle (but note that such a channel filter is also identically replicated at the CFC for each vehicle). The CHF maintains an information vector $\hat{\mathbf{y}}_{CH}(k|k)$ and matrix $\mathbf{Y}_{CH}(k|k)$ representing the newly acquired and shared information.

We consider the following sequence of steps:

- 1) Communicating the CFC information to the vehicle
- 2) Updating the channel filter using the map information from the CFC
- 3) Updating the local SLAM filter
- 4) Selecting the local vehicle sub-map to communicate to the CFC
- 5) Updating the channel filter using the selected sub-map from the LSF
- 6) Updating the global map using the communicated information from the local vehicles

In this paper, the first three steps and the last three steps are referred to as downlink and uplink respectively. We outline the downlink in detail and note the uplink is essentially equivalent modulo semantics.



Fig. 5. Single-vehicle Information Communication Block Diagram

A. Communicating the CFC information to the vehicle

All the landmarks² within the global map held by the CFC that are in a pre-defined radius around the vehicle are transmitted to the vehicle³.

This so-called regional map that is sent from the CFC to the i^{th} vehicle is denoted by $\mathbf{M}_{R}^{i}(\hat{\mathbf{y}}_{R}^{i}, \mathbf{Y}_{R}^{i})$. This information will be received at the communication channel filter (CHF) of the local vehicle.

B. Updating the channel filter using the map information from the CFC

Vehicle information is never communicated and consequently the channel filter will never maintain any states other than map states.

Let's assume that the communicated regional information map and the existing information map in the channel filter are given by $\mathbf{M}_{R}^{i}(\hat{\mathbf{y}}_{R}^{i}, \mathbf{Y}_{R}^{i})$ and $\mathbf{M}_{CH}^{i}(\hat{\mathbf{y}}_{CH}^{i}, \mathbf{Y}_{Ch}^{i})$ respectively. Under the independence assumptions discussed previously by Nettleton, and where $\hat{\mathbf{y}}_{CH}^{i}$ and $\hat{\mathbf{y}}_{R}^{i}$ represent complete and overlapping maps then

$$\begin{aligned} \mathbf{Y}_{CH}^{i}(k|k) &= \mathbf{Y}_{CH}^{i}(k|k-1) \\ &+ \left[\mathbf{Y}_{R}^{i} - \mathbf{Y}_{CH}^{i}(k|k-1)\right] \\ &= \mathbf{Y}_{R}^{i} \end{aligned} \tag{10}$$

where $\mathbf{Y}_{CH}^{i}(k|k)$ denotes the *i*th channel's information matrix at time k given the updated information at time k from the regional sub-map. However, if the channel map information and the transmitted map have different sizes and/or there is some cross-correlation between the shared information and the existing data in the channel then this approach may lead to inconsistent estimates of the common information between nodes. A method to overcome inconsistency is to employ covariance intersection (CI) [29] to calculate the common information between two nodes⁴.

We now drop the superscript i where there is no danger of confusion (and in this section we consider only the

²When we say information is transmitted it is typically meant that the corresponding state vector (or information space representation) and the corresponding marginalised covariance (or information space equivalent) is transmitted.

 $^{^{3}\}mbox{We}$ assume that the CFC can access the global coordinates of the sensor platforms on demand.

⁴See the appendix for a result concerning estimation consistency, CI and fusion while neglecting cross-correlations (as a formal justification for the use of CI in the subsequent discussions). We are not aware of a similar formal argument along the lines given in the appendix for justifying CI (and we show that for some cross-correlations simply neglecting the crosscorrelation will outperform CI and remain consistent).

communication between the CFC and a single vehicle *i*). The CI algorithm requires both information matrices to be of the same size. Thus, define the map domain \mathcal{M}_F as the union of the landmarks in the channel $\mathbf{M}_C(\hat{\mathbf{y}}_{CH}, \mathbf{Y}_{CH})$ and the regional map $\mathbf{M}_R(\hat{\mathbf{y}}_R, \mathbf{Y}_R)$ as shown in Figure 6.



Fig. 6. The map domain $\mathcal{M}_{\rm F}$ is the union map obtained by combining the regional map domain sent from the CFC and the existing map domain in the channel filter.

Two projection matrices are defined \mathbf{G}_{R2F} and \mathbf{G}_{C2F} and consist of 0 and 1 elements. These matrices inflate $\hat{\mathbf{y}}_R$ and $\hat{\mathbf{y}}_{CH}$ to match the cardinality of the union \mathcal{M}_F by padding those components in each respective vector by zero when the landmark indexed by that component is present only in the other vector. The CI algorithm is then given by

$$\mathbf{Y}_{CH}(k|k) = \omega [\mathbf{G}_{C2F}(k|k)\mathbf{Y}_{CH}(k|k-1)\mathbf{G}_{C2F}^{T}(k|k)] + (1-\omega)[\mathbf{G}_{R2F}(k|k)\mathbf{Y}_{R}(k|k)\mathbf{G}_{R2F}^{T}(k|k)]$$
(11)

where $\mathbf{Y}_{CH}(k|k)$ denotes the *i*th channel's information matrix at time k given the updated information at time k from the regional sub-map.

The new information received from the CFC is given by

$$\mathbf{I}^{*}(k|k) = \mathbf{Y}_{CH}(k|k) - \mathbf{G}_{C2F}\mathbf{Y}_{CH}(k|k-1)\mathbf{G}_{C2F}^{T} \quad (12)$$

The information increment is sent to the LSF, e.g. see Figure 5 to be combined with the locally running SLAM filter. Computing the increment prevents double counting of information in the LSF as discussed next.

C. Updating the local SLAM filter

When the local SLAM filter receives the information increment from the channel filter it uses this information to update its estimates. For this purpose, proper projection matrices G_{N2H} and G_{L2H} are defined as previously discussed in order to inflate the information increment vector $\mathbf{i}(k|k)$ and the local information vector $\hat{\mathbf{y}}(k|k)$ to the size of the union domain \mathcal{M}_H . In constructing the former projection matrices, in addition to padding the respective vectors with zeros at those elements corresponding to the non-overlapping landmarks, we must also pad components into $\mathbf{i}^*(k|k)$ with zero to correspond with the vehicle components in $\hat{\mathbf{y}}(k|k)$. Recall no vehicle state is communicated. The update is done according to:

$$\mathbf{Y}(k|k) = \mathbf{G}_{L2H}\mathbf{Y}(k|k-1)\mathbf{G}_{L2H}^{T} + \mathbf{G}_{N2H}\mathbf{I}^{*}(k|k)\mathbf{G}_{N2H}^{T}$$
(13)

and as the LSF is typically executed in the standard state space it follows that $\mathbf{x}(k|k) = \mathbf{Y}^{-1}(k|k)\hat{\mathbf{y}}(k|k)$ and $\mathbf{P}(k|k) = \mathbf{Y}^{-1}(k|k)$ as before.

D. Selecting the local vehicle sub-map to communicate to the CFC

This algorithm is motivated by the AutoMap practical application where the primary objective is to construct and maintain a high-quality global map at a centralised station using information collected (and pre-processed to some degree) at local vehicles. Since the communication bandwidth is limited as previously noted, the 'most informative' sub-map⁵ needs to be selected and transmitted back to the CFC.

There are numerous measures of informativeness. The simplest method involves selecting a sub-map based on the measured information gain. In this application, the information gain is computed by taking the information matrix of the available local landmarks (in the LSF) and comparing this with the existing channel information (all the information transmitted from the LSF previously). We define the information gain of the local map according to:

$$\mathbf{I}(k|k) = \mathbf{Y}_{mm}(k|k) - \mathbf{G}_{C2M}(k|k)\mathbf{Y}_{CH}(k|k)\mathbf{G}_{C2M}^{T}(k|k)$$
(14)

where $\mathbf{Y}_{mm}(k|k) = \mathbf{P}_{mm}^{-1}(k|k)$ and an appropriate (as in previous arguments) inflation matrix \mathbf{G}_{C2M} has been used.

Assume that I(k|k) encodes the information gain regarding a total number of p landmarks. Due to the existing communication constraints, the information of q landmark (q < p) will be transmitted where q is determined by the available bandwidth or an allocated communication budget for time k. The q landmarks with the highest information gain will be selected for transmission. The simple method which is used here is done by picking up the q landmarks with the largest diagonal elements in the I(k|k) matrix. The selected information sub-map for communication to the CFC will be denoted by $\hat{y}_{mm}^*(k|k)$ and $Y_{mm}^*(k|k)$. This information sub-map is sent to the channel filter prior to transmission to the CFC (see Figure 5).

E. Pruning the local SLAM filter

As mentioned before in this paper, in large-scale mapping applications, it is imperative to prevent the size of the local map within the individual vehicles from growing unboundedly. To achieve this, pruning algorithm is implemented at each communication time to limit the size of the SLAM filter. Landmarks with the lowest information gain are eliminated from both the LSF and CHF of each vehicle to reduce the size of the local map to a pre-defined constant n_{pr} , without comprising the integrity of the system.

The salient point here is that, this pruning method is distinct from the standard computationally efficient solutions to the SLAM problem, in a sense that the information (and cross-information) of the discarded landmarks is not lost, due to the previous communication of information to the CFC. This information can be restored locally at any time by downloading the map information from the server.

⁵The 'most-informative' sub-map is necessarily ambiguous. Intuitively one would like to consider the available communication resources and subject to this constraint then select those landmarks in the local vehicle's map that will reduce the uncertainty in any resulting global map constructed at the CFC.



Fig. 7. The SLAM estimates from individual vehicles (no communication). No map information is communicated between the vehicles and the CFC. The 3σ uncertainty ellipses are shown in the figure.

V. SIMULATION AND RESULTS

Simulations were conducted to evaluate the performance of the communication algorithm proposed in this work. The simulation consisted of 3 vehicles driving around overlapping circular trajectories in an environment of 100 landmarks. A non-linear unicycle motion model was implemented to estimate the 2D position and orientation of each vehicle. In addition to the motion sensors, vehicles were also equipped with a range/bearing sensor which provided observations to landmarks along their line of sight.⁶

The first part of the results concentrates on the landmark localisation performance. Two separate runs were performed using identical system configurations, noise parameters and observations⁷. During the first run, none of the vehicles communicated any information to the server and each vehicle simply constructed a local map of its observed landmarks. In the second run, the vehicles communicated their map information at regular intervals with the CFC, according to the bandwidth efficient algorithm described in Section IV. At each communication interval, in addition to the information of newly discovered landmarks, the map information of up to q = 10 landmarks with the highest information gain (See Equation 14) was also transmitted to the CFC. In order to limit the size of the local map within the individual vehicles, a pruning algorithm with $n_{pr} = 15$ was implemented to cut the number of the LSF landmarks to 15 (see Subsection IV-E). Subsequently, the available regional map information from the CFC was communicated to each vehicle for fusion, as described in Subsection IV-A.

The results of the mentioned runs are illustrated in Figures 7 and 8 respectively. The resulting mean and the 3σ covari-

⁶The open source SLAM simulation software by Tim Bailey (available from http://www-personal.acfr.usyd.edu.au/tbailey) was modified and extended to multiple vehicles for use in simulations in this work.

⁷The exact value of all the system parameters are available from the authors on request.



Fig. 8. The multi-vehicle SLAM estimates obtained by selectively communicating map information to the CFC (using q = 10). The 3σ uncertainty ellipses are shown in the figure.

ance ellipses of the estimated landmark positions are shown in each figure alongside the true location of the landmarks. Figure 9 shows the overlaid, enlarged view of the small rectangles shown in Figures 7 and 8. As can be seen from the figures, the landmark localisation accuracy of the multivehicle SLAM with periodic communication outperforms the case where no communication exists.

Figure 10 shows the uncertainty of the obtained map for different values of maximum communicated map size (q). The trace of the map covariance matrix was used here as a commonly accepted measure of uncertainty estimation, cf. [30]. The results suggest that, although the overall uncertainty decreases with the number of communicated local landmarks, the performance quickly converges towards a threshold, as the maximum communicated sub-map size increases. This threshold corresponds to the communication of the entire local maps. This is due to the fact that the algorithm dynamically selects the most informative sub-maps (landmarks with the highest information gain) to transmit. Consequently, if communicated, the landmarks which have not been observed recently and have an insignificant information contribution will have a very small effect on the quality of the global map.

Figure 11 shows the changes in the size of the global and



Fig. 9. The enlarged view of the map estimates. The small rectangular boxes of Figures 7 and 8 are overlaid together. The 3σ uncertainty ellipses are shown in the figure.

local maps for the duration of the exploration, with $n_{pr} = 15$. As can be seen, the pruning algorithm effectively prevents the local map from growing unboundedly. Preliminary evidence suggest that the simulation is comparable in terms of performance and at some points the landmark localisation accuracy improves by applying the pruning algorithm.

We conclude this section by providing a cost analysis for the solution provided in this work and a conventional communication algorithm with no selective communication (e.g. [25]). For this purpose, the values provided in the scenario explained in Example 1 from Section II are used.⁸ These values can be found in Table I. Table II provides a summary of the calculated data regarding each strategy. Although it is assumed that our communication scenario communicates more frequently (20 times more often), the total communication cost is substantially smaller than that of [25]. To reflect the actual network costs, the total communication cost is also compared after considering an overhead cost due to the carrier's flag fall fee.



Fig. 10. The performance of the algorithm in terms of the uncertainty of the obtained global map at the CFC for different values of maximum communicated sub-map sizes.



Fig. 11. The local and global map sizes for $n_{pr} = 15$. The results are shown in the case where the vehicles close their loops twice. The vertical dashed lines demonstrate the communication times.

TABLE I VALUES FROM EXAMPLE 1, SECTION II.

n	d	m	k	b	c
$10 \ vehicles$	200 km	10signs/km	0.1com./km	8 bytes	30\$/GB

⁸Please note that the values used in this example are not related to the values used in the simulation.

VI. CONCLUSION

This paper presented an efficient data fusion framework for the problem of multi-vehicle SLAM for very-large-scale road mapping applications. The solution is efficient in terms of both computational complexity and communication bandwidth. A practical pruning algorithm based on information gain was applied to overcome the problem of growing map sizes at the local nodes. An analysis on the applicability of covariance intersection was also provided to justify the use of this algorithm in the paper. The proposed communication architecture is capable of dealing with dynamically changing map-sizes in the system and is able to consistently fuse this map information in order to build a global map. The mapping solution is potentially scalable to environments with thousands of vehicles and many millions of landmarks.

REFERENCES

- H. Durrant-Whyte and T. Bailey, "Simultaneous localisation and mapping (SLAM): Part i the essential algorithms," *Robotics and Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.
- [2] J. Kim and S. Sukkarieh, "Airborne simultaneous localisation and map building," in *Robotics and Automation*, 2003. Proceedings. ICRA'03. IEEE International Conference on, vol. 1. IEEE, 2003, pp. 406–411.
- [3] M. Dissanayake, P. Newman, H. Durrant-Whyte, S. Clark, and M. Csorba, "An experimental and theoretical investigation into simultaneous localisation and map building," *Experimental Robotics VI*, pp. 265–274, 2000.
- [4] R. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *The international journal of Robotics Research*, vol. 5, no. 4, p. 56, 1986.
- [5] J. Castellanos, J. Montiel, J. Neira, and J. Tardós, "Sensor influence in the performance of simultaneous mobile robot localization and map building," *Experimental Robotics VI*, pp. 287–296, 2000.
- [6] H. Feder, J. Leonard, and C. Smith, "Adaptive mobile robot navigation and mapping," *The International Journal of Robotics Research*, vol. 18, no. 7, p. 650, 1999.
- [7] E. Menegatti, A. Zanella, S. Zilli, F. Zorzi, and E. Pagello, "Rangeonly SLAM with a mobile robot and a wireless sensor networks," in *Robotics and Automation*, 2009. ICRA'09. IEEE International Conference on. IEEE, 2009, pp. 8–14.
- [8] M. Lázaro and J. Castellanos, "Localization of probabilistic robot formations in SLAM," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on.* IEEE, 2010, pp. 3179–3184.
- [9] Y. Cao, A. Fukunaga, and A. Kahng, "Cooperative mobile robotics: Antecedents and directions," *Autonomous robots*, vol. 4, no. 1, pp. 7–27, 1997.
- [10] D. Foxf, W. Burgard, H. Kruppa, and S. Thiun, "Collaborative multirobot localization," KI-99: advances in artificial intelligence: 23rd Annual German Coference [sic] on Artificial Intelligence, Bonn, Germany, September 13-15, 1999: proceedings, p. 255, 1999.
- [11] W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun, "Collaborative multi-robot exploration," in *Robotics and Automation*, 2000. *Proceedings. ICRA'00. IEEE International Conference on*, vol. 1. IEEE, 2000, pp. 476–481.
- [12] M. Bryson and S. Sukkarieh, "An information-theoretic approach to autonomous navigation and guidance of an uninhabited aerial vehicle in unknown environments," in *Intelligent Robots and Systems*, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on. IEEE, 2005, pp. 3770–3775.
- [13] J. Kim, "Autonomous navigation for airborne applications," Ph.D. dissertation, Dept. of Aerospace, Mechanical and Mechatronic Engineering, Graduate School of Engineering, University of Sydney, 2004.
- [14] S. Williams, "Efficient solutions to autonomous mapping and navigation problems," Ph.D. dissertation, Citeseer, 2001.
- [15] E. Nerurkar, S. Roumeliotis, and A. Martinelli, "Distributed maximum a posteriori estimation for multi-robot cooperative localization," in *Robotics and Automation*, 2009. ICRA'09. IEEE International Conference on. IEEE, 2009, pp. 1402–1409.

TABLE II

A numerical example to compare the algorithm in [25] and the solution provided in this paper. The example uses the values provided in Table I and assumes q = 10 and $n_{pr} = 20$ for the selective communication method. For the definition of different variables please see Example 1 in Section II.

	Total distance	Total number of	Landmarks in the LSF	Total number of	Communicated	Total bytes sent by each	Total communication cost per
	per day	landmarks per day	of each vehicle	communications per day	sub-map size	vehicle per day	day (including overhead)
Bryson &	$n \cdot d$	$n \cdot m \cdot d$	$m \cdot d$	$n \cdot d \cdot k$	$m \cdot d$	$b \cdot d \cdot k \cdot (m \cdot d)^2$	$c \cdot n \cdot b \cdot d \cdot k \cdot (m \cdot d)^2$
Sukkarieh [25]	2000	20000	2000	200 (k=0.1)	2000	6.4×10^{8}	\$192 (\$192)
Algorithm	$n \cdot d$	$n \cdot m \cdot d$	n_{pr}	$n \cdot d \cdot k$	q	$b \cdot d \cdot k \cdot q^2$	$c \cdot n \cdot b \cdot d \cdot k \cdot q^2$
from Section IV	2000	20000	20	4000 (k=2)	10	320000	\$0.0096 (\$7.68)

- [16] J. Stipes, R. Hawthorne, D. Scheidt, and D. Pacifico, "Cooperative localization and mapping," in *Networking, Sensing and Control, 2006. ICNSC'06. Proceedings of the 2006 IEEE International Conference* on. IEEE, 2006, pp. 596–601.
- [17] S. Thrun, "A probabilistic on-line mapping algorithm for teams of mobile robots," *The International Journal of Robotics Research*, vol. 20, no. 5, p. 335, 2001.
- [18] E. Nettleton, "Decentralised architectures for tracking and navigation with multiple flight vehicles," 2003.
- [19] M. Liggins, C. Chong, I. Kadar, M. Alford, V. Vannicola, S. Thomopoulos *et al.*, "Distributed fusion architectures and algorithms for target tracking," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 95–107, 1997.
- [20] J. Fenwick, P. Newman, and J. Leonard, "Cooperative concurrent mapping and localization," in *Robotics and Automation*, 2002. Proceedings. ICRA'02. IEEE International Conference on, vol. 2. IEEE, 2002, pp. 1810–1817.
- [21] S. Thrun and Y. Liu, "Multi-robot SLAM with sparse extended information filers," *Robotics Research*, pp. 254–266, 2005.
- [22] L. Sharon, L. Ling et al., "Six dof decentralised SLAM," 2003.
- [23] L. Ong, M. Ridley, J. Kim, E. Nettleton, and S. Sukkarieh, "Six dof decentralised SLAM," in *Proceedings of the Australasian Conference* on Robotics and Automation, 2003, pp. 10–16.
- [24] E. Nettleton, S. Thrun, H. Durrant-Whyte, and S. Sukkarieh, "Decentralised SLAM with low-bandwidth communication for teams of vehicles," in *Field and Service Robotics*. Springer, 2006, pp. 179–188.
- [25] M. Bryson and S. Sukkarieh, "Co-operative localisation and mapping for multiple uavs in unknown environments," in *Aerospace Conference*, 2007 IEEE. IEEE, 2007, pp. 1–12.
- [26] G. Overett, L. Petersson, L. Andersson, and N. Pettersson, "Boosting a heterogeneous pool of fast hog features for pedestrian and sign detection," in *Intelligent Vehicles Symposium*, 2009 IEEE. IEEE, 2009, pp. 584–590.
- [27] L. Petersson. (2012, 3) NICTA AutoMap. [Online]. Available: http://www.nicta.com.au/research/projects/AutoMap
- [28] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *Robotics and Automation, IEEE Transactions* on, vol. 17, no. 3, pp. 229–241, 2001.
- [29] S. Julier and J. Uhlmann, "General decentralized data fusion with covariance intersection (ci)," 2001.
- [30] Y. Bar-Shalom, X. Li, T. Kirubarajan, and J. Wiley, *Estimation with applications to tracking and navigation*. Wiley Online Library, 2001.

VII. APPENDIX

The following theorem justifies the use of covariance intersection in the presence of non-zero but unknown correlation. It indicates that there are cases where consistency can be achieved by simply neglecting (setting to zero) the unknown correlation. Such an approach will outperform covariance intersection in these cases. However, the theorem also shows that neglecting the unknown correlation will cause inconsistency in the general case. We have not seen a formal statement of such a result and thus we included it for completeness. Consider two estimates $\mathbf{a} \sim \mathcal{N}(\mathbf{c}^*, \mathbf{P}_{aa})$ and $\mathbf{b} \sim \mathcal{N}(\mathbf{c}^*, \mathbf{P}_{bb})$ of some fixed parameter \mathbf{c}^* . Let $\mathbf{c} \sim \mathcal{N}(\mathbf{c}^*, \mathbf{P}_{cc})$ denote the best linear unbiased estimate of \mathbf{c}^* obtained via a linear combination of \mathbf{a} and \mathbf{b} . We are mainly interested in the construction of \mathbf{P}_{cc} or its estimate. All matrices thus-far are strictly positive-definite.

If $\mathbf{a} \sim \mathcal{N}(\mathbf{c}^*, \mathbf{P}_{aa})$ is consistent then $\mathbf{P}_{aa} \geq \mathbf{P}_{cc}$. If $\mathbf{P}_{ab} = \mathrm{E}[(\mathbf{a} - \mathbf{c}^*)(\mathbf{b} - \mathbf{c}^*)^\top] \geq \mathbf{0}$ is the correlation between the two estimators (\mathbf{P}_{ab} may be zero) then

$$\mathbf{P}_{cc}^{-1} = \mathbf{P}_{aa}^{-1} + (\mathbf{P}_{aa}^{-1}\mathbf{P}_{ab} - \mathbf{I})(\mathbf{P}_{bb} - \mathbf{P}_{ab}^{\top}\mathbf{P}_{aa}^{-1}\mathbf{P}_{ab})^{-1} \cdot (\mathbf{P}_{ab}^{\top}\mathbf{P}_{aa}^{-1} - \mathbf{I})$$
(15)

is exact. If $\mathbf{P}_{ab} = \mathbf{0}$ then $\mathbf{P}_{cc}^{-1} = \mathbf{P}_{aa}^{-1} + \mathbf{P}_{bb}^{-1}$.

The covariance intersection algorithm presumes $P_{ab} > 0$ but that P_{ab} is unknown and computes

$$\widetilde{\mathbf{P}}_{cc}^{-1} = \omega \mathbf{P}_{aa}^{-1} + (1 - \omega) \mathbf{P}_{bb}^{-1}$$
(16)

where $\omega \in [0, 1]$. Typically, ω is chosen to minimise some measure of $\widetilde{\mathbf{P}}_{cc}$. Note it is known that $\widetilde{\mathbf{P}}_{cc} \geq \mathbf{P}_{cc}$; i.e. $\widetilde{c} \sim \mathcal{N}(\mathbf{c}^*, \widetilde{\mathbf{P}}_{cc})$ is consistent.

Let $\widehat{\mathbf{P}}_{cc}^{-1} = \mathbf{P}_{aa}^{-1} + \mathbf{P}_{bb}^{-1}$ where $\widehat{\mathbf{P}}_{cc}$ is computed simply by neglecting the existence of a non-zero \mathbf{P}_{ab} . Obviously, $\mathbf{P}_{cc} = \widehat{\mathbf{P}}_{cc}$ when $\mathbf{P}_{ab} = \mathbf{0}$. Then $\widehat{\mathbf{P}}_{cc} \leq \widetilde{\mathbf{P}}_{cc}$.

The main result of this section is summed up in the following theorem.

Theorem 1. For some \mathbf{P}_{aa} and \mathbf{P}_{bb} there exists a choice of $\mathbf{P}_{ab} > \mathbf{0}$ such that $\mathbf{P}_{cc} < \widehat{\mathbf{P}}_{cc} < \widetilde{\mathbf{P}}_{cc}$ holds with strict inequality. However, for all \mathbf{P}_{aa} and \mathbf{P}_{bb} there exists a different choice of $\mathbf{P}_{ab} > \mathbf{0}$ such that $\widehat{\mathbf{P}}_{cc} < \mathbf{P}_{cc}$ holds with strict inequality.

For reasons of brevity, a proof of this theorem and surrounding analysis will be presented in a future article.

This theorem states that neglecting (defining to be zero) the cross-correlation when combining two unbiased estimates may lead to a consistent solution that out-performs the solution given by covariance intersection. However, obviously this is dependent on the particular cross-correlation and only holds for some values of the individual estimator covariances. Moreover, for any individual estimator covariances there are values of cross-correlation such that if one were to simply neglect it than the corresponding solution will be inconsistent. Thus, if one suspects a non-zero cross-correlation between two estimators then using covariance intersection is the safest route to achieving a consistent estimate (as it is guaranteed to be consistent) even though it is more conservative than simply neglecting (defining to be zero) the cross-correlation.