

# Single Cluster PHD SLAM: Application to Autonomous Underwater Vehicles using Stereo Vision

Sharad Nagappa<sup>1</sup>, Narcís Palomeras<sup>1</sup>, Chee Sing Lee<sup>1</sup>,  
Nuno Gracias<sup>1</sup>, Daniel E. Clark<sup>2</sup> and Joaquim Salvi<sup>1</sup>

<sup>1</sup>*Computer Vision and Robotics Group, University of Girona, Spain*  
{snagappa, npalomer, cslee, ngracias, qsalvi}@eia.udg.es

<sup>2</sup>*Joint Research Institute in Signal and Image Processing, Heriot-Watt University, U.K.*  
d.e.clark@hw.ac.uk

**Abstract**—This paper considers the application of feature-based simultaneous localisation and mapping (SLAM) using a random finite sets (RFS) framework for an autonomous underwater vehicle. SLAM allows for reduction in localisation error by tracking features which provide a fixed external reference. The SLAM problem is addressed here using a single-cluster probability hypothesis density (PHD) filter. The filter uses a particle approximation for the vehicle position with a conditional Gaussian mixture PHD for the feature map. Map features are selected as unique point features generated from a stereo camera on-board the vehicle. We demonstrate the improvement in localisation applying the algorithm to a dataset obtained in an indoor test tank.

## I. INTRODUCTION

Navigation and localisation of underwater vehicles is difficult in the absence of measurements with respect to a fixed reference. Using only dead reckoning results in cumulative localisation error in the estimated position. This is an issue in cases where persistent autonomy is required, a key issue to be addressed by the PANDORA European FP7 project [1].

One of the main aims of the PANDORA project is to move toward persistent autonomy for intervention autonomous underwater vehicles (AUVs). This requires that the vehicle maintains a map of its environment and is able to localise itself accurately within that map. The use of simultaneous localisation and mapping (SLAM) provides a way to achieve this dual aim.

Early approaches to the SLAM problem involved using the extended Kalman filter (EKF) to address this as a joint state estimation problem, coupling the positions of the vehicle and landmarks detected in the environment [2], [3]. Several advances resulted in the development of unscented Kalman filter (UKF) SLAM [4] as well as particle-based FastSLAM algorithms [5], [6] over the last decade.

This work has been supported by the FP7-ICT-2011-7 project PANDORA-Persistent Autonomy through Learning, Adaptation, Observation and Re-planning (Ref 288273) funded by the European Commission and the Spanish Project ANDREA/RAIMON (Ref CTM2011-29691-C02-02) funded by the Ministry of Science and Innovation.

All these approaches, however, rely on some form of data association, i.e., features which are detected by the sensor must be associated with a particular landmark in the state using some heuristic. Accurate data association can be problematic in situations with high clutter, or false positives. This provides the motivation behind a new approach to the SLAM problem which relies on a random finite set (RFS) formulation.

RFSs provide a means of characterising the uncertainty in the number of states and measurements. This allows them to be used in cases where detection of features in the environment is unreliable and prone to false positives. The probability hypothesis density (PHD) filter [7], [8] provides a suboptimal multi-object estimation framework. By using RFSs, the PHD filter provides a principled approximation to a multi-object posterior. The filter avoids the need for data association and is able to perform well in highly cluttered environments.

Recent work by Mullane *et al.* [9]–[11] as well as Lee, Clark & Salvi [12], [13] has shown that the PHD filter can be extended to deal with SLAM scenarios. In this work, we apply the single-cluster PHD (SC-PHD) filter [12]–[14] to solve the SLAM problem in an underwater setup which uses a stereo camera to detect landmarks. Detection of features from the environment is affected by both lighting conditions, as well as inaccuracies in the camera calibration. The result is that the detected set of features is well represented using an RFS and the SC-PHD filter provides a suitable framework.

In the next section, we summarise the formulation of the SC-PHD filter. In Section III, we discuss the model for the vehicle as well as the specific implementation of the Kalman filter for dead reckoning and the SC-PHD filter for the SLAM. Section IV presents the experiments and results, with conclusions and directions for further work being presented in Section V.

## II. SINGLE CLUSTER PHD SLAM

In this section, we provide an overview of the PHD filter and its application to the SLAM problem. The reader is referred

to the work of Mullane *et al.* [9], [10] and Lee, Clark & Salvi [12], [13] for an in-depth treatment of the PHD-based SLAM methods. Additionally, the reader is referred to the works of Mahler, B.-N. Vo, B.-T. Vo, Ristic and Clark [7], [15]–[21] for further information on PHD filters.

The PHD filter provides a means of multi-object estimation by approximating the Bayes multi-object filter [7], [8], [18]. The PHD filter achieves this by representing the multi-object state and observations (including false positives) using RFSs. In any multi-object estimation problem, both the number of target states as well as their individual values are unknown and must be estimated. RFSs provide a natural means of capturing the uncertainty in the number of targets (or cardinality) as well as the spatial uncertainty.

In addressing the SLAM problem, map features in the environment can be conveniently represented using a RFS. In this case, the number of map features and their positions are *a priori* known. Additionally, sensor measurements are prone to false positives which must be accounted for by the filter. The PHD filter provides a principled framework for dealing with these uncertainties. The PHD filter achieves this while removing the need for data association, resulting in improved performance when dealing with high-clutter environments.

Formally, we represent the multi-object state at time  $k$  as

$$\mathbf{X}_k = \{\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,m_k}\} \in \mathcal{F}(\mathcal{X}), \quad (1)$$

with each state  $\mathbf{x}_{k,l}$  taking a value in the state space  $\mathcal{X}$ . In the SLAM context, these states may represent the position of different landmarks in the map. The observations are given by

$$\mathbf{Z}_k = \{\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,n_k}\} \in \mathcal{F}(\mathcal{Z}), \quad (2)$$

with the state space for each observation given by  $\mathcal{Z}$ . This observation set is the union of observations arising from targets as well as observations due to clutter (false positives).

The Bayes multi-target filter can be used to estimate the posterior distribution on the multi-target state  $\mathbf{X}_k$ ,  $p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})$ , given a sequence of observations up to time  $k$ ,  $\mathbf{Z}_{1:k}$ . The Bayes recursion is given by the combination of a one-step prediction and update:

$$p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1}) = \int p_{k-1|k-1}(\mathbf{X}|\mathbf{Z}_{k-1}) f_{k|k-1}(\mathbf{X}|\mathbf{Y}) \delta \mathbf{Y}, \quad (3)$$

$$p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k}) = \frac{p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1}) g_k(\mathbf{Z}_k|\mathbf{X})}{\int p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1}) g_k(\mathbf{Z}_k|\mathbf{X}) \delta \mathbf{X}}, \quad (4)$$

where  $f_{k|k-1}(\mathbf{X}|\mathbf{Y})$  represents the multi-target transition density,  $g_k(\mathbf{Z}|\mathbf{X})$  is the multi-target likelihood function, and the integrals in the prediction and update represent set integrals [7], [8].

The intractability of the full Bayes recursion when dealing with a large number of targets and clutter leads us to the

PHD filter. The PHD filter approximates the true multi-target distribution on the single-target state space using the *intensity function*, also referred to as the PHD. Analogous to moments of probability distributions, the PHD is the first-order moment of the multi-target distribution. Since the intensity function exists on the single-target state space, peaks in the intensity function indicate regions with higher likelihood of targets. Computing the integral of the intensity function provides an estimate of the number of targets present. In this manner, the PHD filter allows us to estimate both the number of targets as well as their positions.

In the SLAM problem, the state consists of both the vehicle pose as well as the map landmarks. This joint state can be modelled as a single cluster process, where the vehicle pose comprises the parent state and the map is a conditional daughter process. The joint state can be estimated using the SC-PHD filter [12]–[14]. The parent state (vehicle) is denoted by the random vector  $\mathbf{x}_k$  and the associated daughter state (map) is given by the RFS  $\mathbf{M}_k$ :

$$\mathbf{x}_k = [x_{k,1}, \dots, x_{k,n_x}], \quad (5)$$

$$\mathbf{M}_k = [\mathbf{m}_{k,1}, \dots, \mathbf{m}_{k,m_k}], \quad (6)$$

where  $\mathbf{m}_{k,l}$  denotes a single feature in the map.

We now reproduce the expressions for the prediction and update of the single-cluster PHD (SC-PHD) filter [12]–[14]. The intensity function on the joint parent-daughter state is denoted as  $D_{k-1}(\mathbf{x}, \mathbf{m})$ , which can be factored as [12]

$$D_{k-1}(\mathbf{x}, \mathbf{m}) = D_{k-1}(\mathbf{x}) \times D_{k-1}(\mathbf{m}|\mathbf{x}), \quad (7)$$

where  $D_{k-1}(\mathbf{x})$  is the intensity function of the vehicle state and  $D_{k-1}(\mathbf{m}|\mathbf{x})$  is the intensity function of the conditional map states.

The predicted intensity function is given by

$$D_{k|k-1}(\mathbf{x}, \mathbf{m}) = \int D_{k|k-1}(\mathbf{x}) p(\mathbf{x}|\mathbf{x}_{k-1}) D_{k|k-1}(\mathbf{m}|\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \quad (8)$$

where  $p(\mathbf{x}|\mathbf{x}_{k-1})$  is the Markov transition density for the parent process and  $D_{k|k-1}(\mathbf{m}|\mathbf{x}_{k-1})$  is the predicted intensity function for the conditional daughter process. The prediction for the daughter process is given by

$$D_{k|k-1}(\mathbf{m}|\mathbf{x}_{k-1}) = \gamma_{k|k-1}(\mathbf{m}|\mathbf{x}_{k-1}) + \int D_{k-1}(\mathbf{m}|\mathbf{x}_{k-1}) p_S(\mathbf{m}|\mathbf{x}_{k-1}) p_{k|k-1}(\mathbf{m}|\mathbf{m}_{k-1}, \mathbf{x}_{k-1}) d\mathbf{m}_{k-1} \quad (9)$$

where  $\gamma_{k|k-1}(\mathbf{m}|\mathbf{x}_{k-1})$  represents a birth process,  $p_S(\mathbf{m}|\mathbf{x}_{k-1})$  is the probability of survival and  $p_{k|k-1}(\mathbf{m}|\mathbf{m}_{k-1}, \mathbf{x}_{k-1})$  is the (single-target) Markov transition for map feature states. The birth process  $\gamma_{k|k-1}$  is used to initialise new targets as they appear in the sensor field of view and the probability of survival is used to model the death of targets.

Here, we assume the probability of survival to be 1, and the Markov transition density to be the identity matrix. This implies that landmarks which are currently in the map are expected to continue to exist. However, when not supported by the observations, the intensity function (subsequent to the filter update) will reflect the absence of previously confirmed landmarks.

On receiving a set of observations  $\mathbf{Z}_k$ , the factorised updated intensity function is obtained as [12], [14]

$$D_k(\mathbf{x}) = \frac{L_{\mathbf{Z}_k}(\mathbf{x})D_{k|k-1}(\mathbf{x})}{\int L_{\mathbf{Z}_k}(\mathbf{x})D_{k|k-1}(\mathbf{x})d\mathbf{x}}, \quad (10)$$

$$D_k(\mathbf{m}|\mathbf{x}) = (1 - p_D(\mathbf{m}|\mathbf{x}))D_{k|k-1}(\mathbf{m}|\mathbf{x}) + \sum_{\mathbf{z} \in \mathbf{Z}_k} \frac{p_D(\mathbf{m}|\mathbf{x})D_{k|k-1}(\mathbf{m}|\mathbf{x})g_k(\mathbf{z}|\mathbf{m}, \mathbf{x})}{\kappa_k(\mathbf{z}) + \int p_D(\mathbf{m}|\mathbf{x})D_{k|k-1}(\mathbf{m}|\mathbf{x})g_k(\mathbf{z}|\mathbf{m}, \mathbf{x})d\mathbf{m}}, \quad (11)$$

where  $L_{\mathbf{Z}_k}(\mathbf{x})$  is a multi-object measurement likelihood,  $p_D(\mathbf{m}|\mathbf{x})$  is the probability of detection of the daughter state (map feature) conditioned on the parent state (vehicle position),  $g_k(\mathbf{z}|\mathbf{m}, \mathbf{x})$  is the conditional single object measurement likelihood function, and  $\kappa_k(\mathbf{z})$  is the clutter intensity. The multi-object measurement likelihood  $L_{\mathbf{Z}_k}(\mathbf{x})$  is given by

$$L_{\mathbf{Z}_k}(\mathbf{x}) = \exp\left(-\int p_D(\mathbf{m}|\mathbf{x})D_{k|k-1}(\mathbf{m}|\mathbf{x})d\mathbf{m}\right) \times \prod_{\mathbf{z} \in \mathbf{Z}_k} \left(\kappa_k(\mathbf{z}) + \int p_D(\mathbf{m}|\mathbf{x})D_{k|k-1}(\mathbf{m}|\mathbf{x})g_k(\mathbf{z}|\mathbf{m}, \mathbf{x})d\mathbf{m}\right). \quad (12)$$

The clutter intensity function  $\kappa_k(\mathbf{z})$  is assumed to have a Poisson distribution on the number of points such that

$$\kappa_k(\mathbf{z}) = \lambda_c p(\mathbf{z}) \quad (13)$$

where  $\lambda_c$  is the mean of the cardinality distribution and  $p(\mathbf{z})$  is the spatial distribution of the clutter.

### III. APPLICATION TO UNDERWATER ROBOTICS

In this paper, we consider application of the SC-PHD SLAM algorithm to an underwater scenario using the Girona 500 AUV. The Girona 500 AUV (Fig. 1) is an underwater robotic vehicle developed by the Research Centre in Underwater Robotics (CIRS) at the University of Girona. The vehicle utilises a triple-hull design which provides significant stability in roll and pitch. In the configuration used here, the vehicle is equipped with 5 thrusters providing 4 degrees of freedom (DoF) – translation in  $x, y, z$  and rotation in the yaw angle. Maximum operational depth is rated at 500 m.

The vehicle is equipped with a variety of sensors for navigation and mapping. An inertial measurement unit (IMU) provides vehicle orientation and angular velocity; velocity and depth are measured using a Doppler velocity log (DVL)

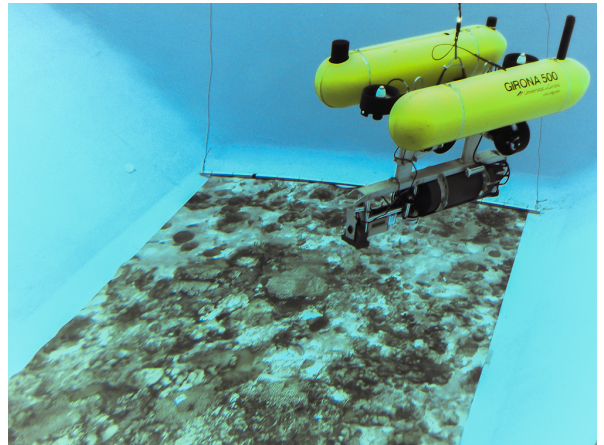


Fig. 1. The Girona 500 AUV in the test tank the stereo camera viewing the poster.

and depth sensor respectively. A global positioning system (GPS) provides vehicle position at the surface and is not used in the experiments here. In addition to this, the vehicle is equipped with a downward looking calibrated stereo camera for surveying and mapping of the sea-floor.

The vehicle is teleoperated within a test tank measuring 16 m × 8 m × 5 m. The floor of the tank is fitted with a known pattern created from a sea-floor mosaic. The camera is used to detect features from this pattern for use with SLAM. Images from the camera are also compared against the known pattern to estimate the ground-truth trajectory of the vehicle. In the next section, we describe the motion and observation model used by the vehicle. This is followed by a description of the implementation of the SLAM algorithm, feature detection method and the experimental results.

#### A. Motion Model

The vehicle state consists of the position and linear velocities at time  $t_k$  given by

$$\mathbf{x}_k = [x_k, y_k, z_k, \dot{x}_k, \dot{y}_k, \dot{z}_k]^T. \quad (14)$$

Vehicle orientation and angular velocity  $[\phi, \theta, \psi, \dot{\phi}, \dot{\theta}, \dot{\psi}]$  are measured by an IMU and not estimated.

Using a constant velocity model, the predicted state at time  $k$  is given by

$$\mathbf{x}_{k|k-1} = f(\mathbf{x}_{k-1}, \mathbf{u}_k, \Delta_t) + \mathbf{v}_k, \quad (15)$$

where  $f(\mathbf{x}_{k-1}, \mathbf{u}_k, \Delta_t)$  is the prediction function,  $\mathbf{u}_k = [\phi_k, \theta_k, \psi_k]$  is the control input comprising the vehicle orientation,  $\Delta_t = t_k - t_{k-1}$  and  $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \Sigma_v)$  is a zero mean white Gaussian process noise. For the model used here, the

prediction function is given by

$$f(\mathbf{x}_k, \mathbf{u}_k, \Delta_t) = \begin{bmatrix} \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} + \mathbf{R}_k(\phi_k, \theta_k, \psi_k) \Delta_t \begin{bmatrix} \dot{x}_k \\ \dot{y}_k \\ \dot{z}_k \end{bmatrix} \\ \dot{x}_k \\ \dot{y}_k \\ \dot{z}_k \end{bmatrix}, \quad (16)$$

where  $\mathbf{R}_k(\phi_k, \theta_k, \psi_k)$  denotes the  $3 \times 3$  rotation matrix.

Sensor observations received from the DVL and IMU are transformed to the vehicle coordinate system. In the vehicle's reference frame, the DVL measurements consist of velocities along the three axes  $\mathbf{z}_{k,\nu} = [\dot{x}_{k,\nu}, \dot{y}_{k,\nu}, \dot{z}_{k,\nu}]$ . For a given state, the observation equation is given by

$$\mathbf{z}_{k,\nu} = \mathbf{H}_{k,\nu} \mathbf{x}_{k|k} + \mathbf{w}_{k,\nu}, \quad (17)$$

$$\mathbf{H}_{k,\nu} = \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{I}_3 \end{bmatrix}, \quad (18)$$

where  $\mathbf{0}_{p \times q}$  represents a  $p \times q$  zero matrix,  $\mathbf{I}_p$  is the  $p \times p$  identity matrix and  $\mathbf{w}_{k,\nu} \sim \mathcal{N}(\mathbf{0}, \Sigma_\nu)$  is zero mean white Gaussian noise associated with the DVL measurement.

Observations from the depth sensor are used to update the vehicle depth. The observation equation for the depth measurement  $\mathbf{z}_{k,d}$  is given by

$$\mathbf{z}_{k,d} = \mathbf{H}_{k,d} \mathbf{x}_k + \mathbf{w}_{k,d}, \quad (19)$$

$$\mathbf{H}_{k,d} = [0 \ 0 \ 1 \ 0 \ 0 \ 0], \quad (20)$$

where  $\mathbf{w}_{k,d} \sim \mathcal{N}(\mathbf{0}, \Sigma_d)$  is zero mean white Gaussian noise associated with the depth measurement.

### B. Localisation using the EKF

In the absence of SLAM, vehicle localisation is performed using the above model and an extended Kalman filter (EKF). The EKF estimates the filtered distribution of the vehicle state as a Gaussian distribution with mean and covariance given by  $(\mathbf{x}_k, \mathbf{P}_k)$ . The predicted mean and covariance are given by

$$\mathbf{x}_{k|k-1} = f(\mathbf{x}_k, \mathbf{u}_k, \Delta_t), \quad (21)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1} \mathbf{P}_{k-1} \mathbf{F}_{k-1}^T + \Sigma_v, \quad (22)$$

where  $\mathbf{F}_{k-1}$  is the Jacobian of the prediction function (equation (16)).

For a given observation  $\mathbf{z}_k$ , the updated mean and covariance of the distribution are given by

$$\mathbf{x}_k = \mathbf{x}_{k|k-1} + \mathbf{K}_k \hat{\mathbf{y}}_k, \quad (23)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}, \quad (24)$$

where  $\mathbf{K}_k$  represents the Kalman filter gain, and

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \Sigma_w, \quad (25)$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}, \quad (26)$$

$$\hat{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}, \quad (27)$$

with  $\mathbf{H}_k = \mathbf{H}_{k,\nu}$ ,  $\Sigma_w = \Sigma_\nu$  for DVL measurements and  $\mathbf{H}_k = \mathbf{H}_{k,d}$ ,  $\Sigma_w = \Sigma_d$  for depth measurements. These prediction and update equations are used to estimate the position of the vehicle.

The sensor observations comprise velocity measurements from the DVL with the result that the uncertainty on the vehicle position is unbounded. In the next section, we illustrate how measurements corresponding to static landmarks in the environment are used to improve the vehicle localisation.

### C. Single Cluster PHD SLAM

The vehicle position estimated by the EKF can be improved by using a SLAM framework. To implement the SC-PHD filter, we use a particle representation for the vehicle position and a Gaussian mixture (GM) representation for the per-particle feature map. This follows a similar implementation to that derived by Lee *et al.* [12]. A Rao-Blackwellised particle filter (RBPF) is used to estimate the vehicle state with the Kalman filter used to estimate the velocity and a particle representation for the vehicle position and heading. Here, we include the heading angle in the state space since we are able to correct erroneous heading angles obtained using a magnetic compass. The vehicle state is then composed of  $\mathbf{x}_k = \{\boldsymbol{\mu}_k, \dot{\boldsymbol{\mu}}_k\}$  where  $\boldsymbol{\mu}_k = [x_k, y_k, z_k, \psi_k]$  represents the position and heading of the vehicle, and  $\dot{\boldsymbol{\mu}}_k = [\dot{x}_k, \dot{y}_k, \dot{z}_k]$  represents the velocities. The intensity function on the vehicle state is approximated using a set of weighted particles  $\{(\boldsymbol{\mu}_k^{(i)}, \dot{\boldsymbol{\mu}}_k^{(i)}, \Sigma_k^{(i)}), w_k^{(i)}\}$  as

$$D_k(\mathbf{x}) = \sum_{n=1}^N w_k^{(n)} \delta_{\mathbf{x}_k^{(n)}}(\mathbf{x}). \quad (28)$$

Associated with the  $i^{\text{th}}$  particle is a representation of the map features given by a weighted Gaussian mixture with  $N_{k,m}^{(i)}$  weights, means and covariances, given by  $\{w_{k,m}^{(j|i)}, \mathbf{m}_k^{(j|i)}, \mathbf{P}_k^{(j|i)}\}, j \in \{1, \dots, N_{k,m}^{(i)}\}$ :

$$D_k(\mathbf{m}|\mathbf{x}^{(i)}) = \sum_{j=1}^{N_{k,m}^{(i)}} w_{k,m}^{(j|i)} \mathcal{N}(\mathbf{m}_k^{(j|i)}, \mathbf{P}_k^{(j|i)}), \quad (29)$$

where  $j|i$  indicates the conditioning of the map state on the parent state, and the subscript  $m$  is used to refer to the map landmarks.

1) *Prediction*: The predicted intensity function for the vehicle state at time  $k$  is obtained by sampling from the Markov transition given by equation (15) with  $f = f([\mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{k-1}^{(i)}, \mathbf{z}_{k-1}^{(i)}], [\phi_k, \theta_k, \psi_k^{(i)}], \Delta_t)$ .

The predicted intensity function for the daughter process is used to predict the state of the map as well as add new landmarks to the map. In this paper, we deal with static landmarks; the predicted map weights and states which exist at time  $k-1$  are thus unchanged at time  $k$ . The prediction step is used to initialise new landmarks in the map. A

good choice is to follow the strategy of Ristic *et. al.* [21] which initialises new targets in regions corresponding to the observations. It should be noted that adopting this approach results in a slightly different form of the PHD update given by equation (11). The birth function is conditioned on the parent process and, for the  $i^{\text{th}}$  parent, is given by

$$\gamma_{k|k-1}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}) = \sum_{j=1}^{N_{k,b}} w_{k,b}^{(j|i)} \mathcal{N}(\mathbf{m}_{k,b}^{(j|i)}, \mathbf{P}_{k,b}^{(j|i)}). \quad (30)$$

The formulation of the birth function is described in greater detail in Section III-D. The predicted intensity function for the daughter process is then the union of the current landmarks and those introduced through the birth process:

$$D_{k|k-1}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}) = w_{k|k-1,m}^{(j|i)} \mathcal{N}(\mathbf{m}_{k|k-1,m}^{(j|i)}, \mathbf{P}_{k|k-1,m}^{(j|i)}) + \sum_{j=1}^{N_{k,b}} w_{k,b}^{(j|i)} \mathcal{N}(\mathbf{m}_{k,b}^{(j|i)}, \mathbf{P}_{k,b}^{(j|i)}), \quad (31)$$

with

$$w_{k|k-1,m}^{(j|i)} = w_{k-1,m}^{(j|i)}, \quad (32)$$

$$\mathbf{m}_{k|k-1,m}^{(j|i)} = \mathbf{m}_{k-1,m}^{(j|i)}, \quad (33)$$

$$\mathbf{P}_{k|k-1,m}^{(j|i)} = \mathbf{P}_{k-1,m}^{(j|i)}. \quad (34)$$

The prediction takes this formulation since the probability of survival  $p_S$  is 1, the Markov transition is the identity matrix and there is no process noise.

2) *DVL and Depth Sensor Updates:* Updates occur when either DVL or depth measurements are received. The observation function for the DVL velocities is linear and Gaussian. We thus estimate the velocity state using the Kalman filter with mean and covariance given by  $(\hat{\boldsymbol{\mu}}_k^{(i)}, \Sigma_k^{(i)})$ . For observations from the DVL  $\mathbf{z}_{k,\nu}$ , we estimate the vehicle state as

$$p_k(\boldsymbol{\mu}_k, \hat{\boldsymbol{\mu}}_k | \mathbf{z}_{k,\nu}) = p(\hat{\boldsymbol{\mu}}_k | \mathbf{z}_{k,\nu}, \boldsymbol{\mu}_k) p(\boldsymbol{\mu}_k | \mathbf{Z}_{k,\nu}) \quad (35)$$

where the distribution  $p(\hat{\boldsymbol{\mu}}_k | \mathbf{z}_{k,\nu}, \mathbf{x}_k)$  is estimated using the Kalman filter equations (equations 21–27). The updated weights are then obtained as

$$\hat{w}_k^{(i)} = w_{k-1}^{(i)} \mathcal{N}(\mathbf{z}_{k,\nu}; \mathbf{H}_k \boldsymbol{\mu}_k^{(i)}, \mathbf{S}_k^{(i)}), \quad (36)$$

$$w_k^{(i)} = \frac{\hat{w}_k^{(i)}}{\sum_{j=1}^N \hat{w}_k^{(j)}}, \quad (37)$$

$$\mathbf{S}_k^{(i)} = \mathbf{H}_k \Sigma_{k|k-1}^{(i)} \mathbf{H}_k^T + \Sigma_\nu. \quad (38)$$

Updates from the depth sensor take the form of a standard particle filter update with the likelihood function given by

$$\hat{w}_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{z}_{k,d} | \mathbf{x}_k^{(i)}) \quad (39)$$

$$p(\mathbf{z}_{k,d} | \mathbf{x}_k^{(i)}) = \mathcal{N}(\mathbf{z}_{k,d}; z_k^{(i)}, \Sigma_d). \quad (40)$$

3) *Landmark Updates:* On receiving a set of landmarks, the SC-PHD update is applied. The set of observed landmarks  $\mathbf{Z}_{k,m}$  are used to correct the position of the vehicle using equation (11). The update for the daughter process is conditioned on the parent state, but is otherwise identical to the standard Gaussian mixture PHD update. The predicted intensity function is given by  $D_{k|k-1}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)})$  (equation (31)).

The updated intensity function is given by a combination of the missed-detection term  $D_{k,M}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)})$  and the detection term  $D_{k,D}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)})$  as [12]

$$D_k(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}) = D_{k,M}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}) + D_{k,D}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}). \quad (41)$$

The missed detection term is given by the weighted Gaussian mixture

$$D_{k,M}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}) = \sum_{j=1}^{N_{k,m}^{(i)}} (1 - p_D(\mathbf{m}_{k|k-1}^{(j|i)})) w_{k|k-1,m}^{(j|i)} \mathcal{N}(\mathbf{m}_{k|k-1}^{(j|i)}, \mathbf{P}_{k|k-1}^{(j|i)}), \quad (42)$$

and the detection term is given by the mixture

$$D_{k,D}(\mathbf{m}|\mathbf{x}_{k|k-1}^{(i)}) = \sum_{\mathbf{z} \in \mathbf{Z}_{k,m}} w_{k,m,\mathbf{z}}^{(j|i)} \mathcal{N}(\mathbf{m}_{k,\mathbf{z}}^{(j|i)}, \mathbf{P}_{k,\mathbf{z}}^{(j|i)}), \quad (43)$$

where the updated weights and states  $(w_{k,m,\mathbf{z}}^{(j|i)}, \mathbf{m}_{k,\mathbf{z}}^{(j|i)}, \mathbf{P}_{k,\mathbf{z}}^{(j|i)})$  are obtained according to a Kalman filter update [18]:

$$w_{k,m,\mathbf{z}}^{(j|i)} = \frac{p_D(\mathbf{m}_{k|k-1}^{(j|i)}) w_{k|k-1,m}^{(j|i)} g_{\mathbf{z}}(\mathbf{m}_{k|k-1}^{(j|i)})}{\kappa_k(\mathbf{z}) + \sum_{l=1}^{N_{k-1,m}} p_D(\mathbf{m}_{k|k-1}^{(l|i)}) w_{k|k-1,m}^{(l|i)} g_{\mathbf{z}}(\mathbf{m}_{k|k-1}^{(l|i)})}, \quad (44)$$

$$g_{\mathbf{z}}(\mathbf{m}_{k|k-1}^{(j|i)}) = \mathcal{N}(\mathbf{z}; h(\mathbf{m}_{k|k-1}^{(j|i)}), \mathbf{S}_{k,m}^{(j|i)}) \quad (45)$$

$$\mathbf{m}_{k,\mathbf{z}}^{(j|i)} = \mathbf{m}_{k|k-1}^{(j|i)} + \mathbf{K}_{k,m} \hat{\mathbf{y}}_{k,m,\mathbf{z}}, \quad (46)$$

$$\mathbf{P}_{k,\mathbf{z}}^{(j|i)} = (\mathbf{I} - \mathbf{K}_{k,m} \mathbf{H}_{k,m}^{(j|i)}) \mathbf{P}_{k|k-1,m}^{(j|i)}, \quad (47)$$

$$\mathbf{S}_{k,m}^{(j|i)} = \mathbf{H}_{k,m}^{(j|i)} \mathbf{P}_{k|k-1,m}^{(j|i)} (\mathbf{H}_{k,m}^{(j|i)})^T + \Sigma_m, \quad (48)$$

$$\mathbf{K}_{k,m} = \mathbf{P}_{k|k-1,m}^{(j|i)} (\mathbf{H}_{k,m}^{(j|i)})^T (\mathbf{S}_{k,m}^{(j|i)})^{-1}, \quad (49)$$

$$\hat{\mathbf{y}}_{k,m,\mathbf{z}} = \mathbf{z} - h(\mathbf{m}_{k|k-1}^{(j|i)}). \quad (50)$$

In the above equations,  $h(\mathbf{m}_{k|k-1}^{(j|i)})$  denotes the observation function which may be nonlinear and  $\mathbf{H}_{k,m}^{(j|i)}$  denotes the corresponding Jacobian. As an alternative to linearising the update, this may be performed using the unscented Kalman filter [22], thus avoiding the need to evaluate the Jacobian. In the absence of prior information, we assume the spatial distribution of the clutter to be uniform over the entire observation space:  $\kappa_k(\mathbf{z}) = \text{constant}$ . Without any management, the number of terms in the filter grows exponentially. Several terms in the mixture may have low weight, thus making

negligible contribution to the intensity function. There will also be terms which are similar to each other and could be approximated by a single Gaussian component without loss of accuracy. To manage the growth of terms, we adopt the pruning and merging procedure described by Vo and Ma [18].

In the following section, we will discuss the observation model for the landmarks and define the observation function  $h(\mathbf{m}_{k|k-1}^{(j|i)})$  and probability of detection  $p_D(\mathbf{m}_{k|k-1}^{(j|i)})$ .

The estimate of the vehicle state is obtained by computing the weighted mean of the intensity function on the parent state. The estimated landmarks are selected according to the map associated with the highest weight particle:

$$\hat{\mathbf{x}}_k = \sum_{i=1}^N w_k^{(i)} \mathbf{x}_k^{(i)}, \quad (51)$$

$$\hat{i}_k = \arg_i \max w_k^{(i)}, \quad (52)$$

$$\hat{\mathbf{M}}_k = \mathbf{M}_k^{(\hat{i}_k)}. \quad (53)$$

Since the estimated map  $\hat{\mathbf{M}}_k$  also contains unconfirmed landmarks, we select landmarks which have weights  $w_{k,m}^{(j|i)}$  greater than 0.5.

#### D. Feature Observation Model

In the experiments conducted here, a downward looking stereo camera (Fig. 2) is used to estimate point features of interest in the environment. Rectified stereo images are obtained from the camera and features of interest are detected using the speeded-up robust features (SURF) [23] extractor, a blob detection algorithm that identifies robust features in the images.

Landmarks in the map exist in the world coordinate system, while observations are made in the coordinate system of the camera. It is thus necessary to evaluate the transformation between the the camera and the vehicle, and the vehicle and the world coordinate systems. For simplicity, we assume in the following equations that the

By comparing features between the two rectified camera images (Fig. 3), we obtain a set of matched observations with each observation taking the form  $[u_l, v_l, u_r, v_r]$  with the subscripts  $l$  and  $r$  denoting the pixel values in the left and right images respectively. The set of  $L_k$  observations is then defined as

$$\mathbf{Z}_{k,m} = \{\mathbf{z}_{k,m}^{(1)}, \dots, \mathbf{z}_{k,m}^{(L_k)}\}, \quad (54)$$

$$\mathbf{z}_{k,m}^{(i)} = [u_l^{(i)}, v_l^{(i)}, d^{(i)}]^T, \quad (55)$$

$$d^{(i)} = u_l^{(i)} - u_r^{(i)}. \quad (56)$$

For a given state in the map,  $\mathbf{m}_k^{(j|i)} = [x_{k,m}^{(j|i)}, y_{k,m}^{(j|i)}, z_{k,m}^{(j|i)}]^T$ , we can obtain its projection on to the camera image plane

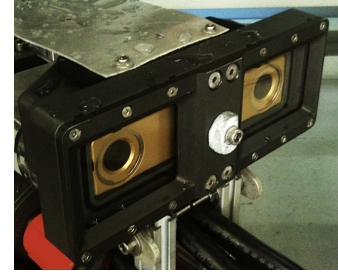


Fig. 2. The stereo camera inside its waterproof housing.

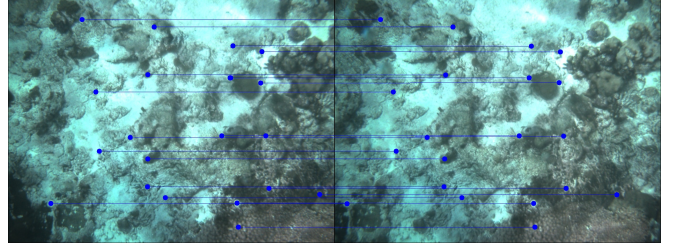


Fig. 3. Features from the left and right rectified images are matched. Along with the disparity, each pair of pixels is an element of the observation set. The 3D point corresponding to the pair of pixels can be computed by triangulation.

as

$$\begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} = g(\mathbf{m}_k^{(j|i)}, \mathbf{P}_l) \quad (57)$$

$$\begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = g(\mathbf{m}_k^{(j|i)}, \mathbf{P}_r) \quad (58)$$

where

$$g(\mathbf{m}_k^{(j|i)}, \mathbf{P}) = \mathbf{P} \left[ \begin{array}{c|c} \mathbf{R}(-\phi_k, -\theta_k, -\psi_k^{(i)}) & -\mathbf{T}_k^{(i)} \\ \mathbf{0}_{1 \times 3} & 1 \end{array} \right] \begin{bmatrix} \mathbf{m}_k^{(j|i)} \\ 1 \end{bmatrix}, \quad (59)$$

$\mathbf{P}_l$  and  $\mathbf{P}_r$  represent the  $3 \times 4$  left and right camera projection matrices, and  $[\phi_k, \theta_k, \psi_k^{(i)}]$  and  $\mathbf{T}_k^{(i)}$  are respectively the orientation and  $3 \times 1$  translation vector of the left camera. For simplicity, we assume in the above equation that the orientation of the camera is identical to that of the vehicle. When this is not the case, the rotation matrix must be calculated using the orientation of the camera rather than the vehicle. The Jacobian required in the Kalman filter update in equations (48) and (49) is obtained by computing the partial derivatives of this observation function.

Points which lie outside the view frustum of the stereo camera will not produce a valid projection on the image plane. The field of view (FOV) of the camera thus defines the visibility of map landmarks and thus the probability of detection  $p_D(\mathbf{m}_k^{(j|i)})$ . Variations in lighting and focus also have an effect on the detection of features, however, we

ignore this and define

$$p_D(\mathbf{m}_k^{(j|i)}) = \begin{cases} p_D & \text{if } \mathbf{m}_k^{(j|i)} \in V(\mathbf{m}_k^{(j|i)}), \\ 0 & \text{otherwise,} \end{cases} \quad (60)$$

where  $V(\mathbf{m}_k^{(j|i)})$  denotes the FOV of the camera corresponding to the  $i^{\text{th}}$  parent.

The inverse of the observation function  $g(\cdot)$  is required to introduce new landmarks in the filter. Given the projection matrices  $\mathbf{P}_l$  and  $\mathbf{P}_r$ , and a matched pair of points  $[u_l, v_l]$  and  $[u_r, v_r]$ , it is possible to perform triangulation [24] and obtain the coordinates of the point in 3D space  $[x, y, z]$  (the projection of this point gives us back the matched points). We utilise the OpenCV library [25] to perform this triangulation and define  $\mathbf{m}_c = [x_c, y_c, z_c]^T$  as the coordinates of the triangulated point in the camera coordinate system:

$$\mathbf{m}_c^{(j)} = g^{-1}(\mathbf{P}_l, \mathbf{P}_r, [u_l^{(j)}, v_l^{(j)}], [u_r^{(j)}, v_r^{(j)}]), \quad j \in [1, L_k]. \quad (61)$$

New landmarks are introduced into the map in the world coordinate system rather than the camera coordinate system. This requires a transformation which is conditioned on the parent, obtained as

$$\begin{bmatrix} \mathbf{m}_{k,b}^{(j|i)} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}(\phi_k, \theta_k, \psi_k^{(i)}) & \mathbf{T}_k^{(i)} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_c^{(j)} \\ 1 \end{bmatrix}. \quad (62)$$

In the experiment conducted here, we initialise the covariance associated with the new landmarks to a predefined constant  $\mathbf{P}_b$  and initialise the birth weight  $w_{k,b}^{(j|i)}$  to a small constant in equation (30).

## IV. EXPERIMENTS AND RESULTS

### A. Experimental Setup

The SC-PHD SLAM algorithm is tested with a dataset collected using the Girona 500 AUV [26]. The vehicle is equipped with a DVL which provides velocity measurements, an IMU which provides orientation and angular velocity, a depth sensor, and the downward looking stereo camera which provides images at 10 Hz with a resolution of  $1024 \times 768$ . The vehicle is teleoperated in an indoor tank of  $16 \text{ m} \times 8 \text{ m} \times 5 \text{ m}$  in a lawnmower pattern. The bottom of the tank is fitted with a poster simulating a sea-floor. By comparing images from the camera against this known mosaic, it is possible to generate a ground-truth trajectory of the vehicle. Fig. 1 shows the Girona 500 AUV in the test tank with the stereo camera used to detect the features from the poster at the bottom of the tank.

The recorded sensor observations are processed using both the EKF and SC-PHD SLAM summarised previously, with all processing being performed offline. The SC-PHD SLAM is run using 400 particles with a maximum of 75 features detected in any pair of images from the camera. The probability of detection is set as  $p_D = 0.975$ , and we use  $\lambda_c = 20$  for the mean of the clutter cardinality distribution.

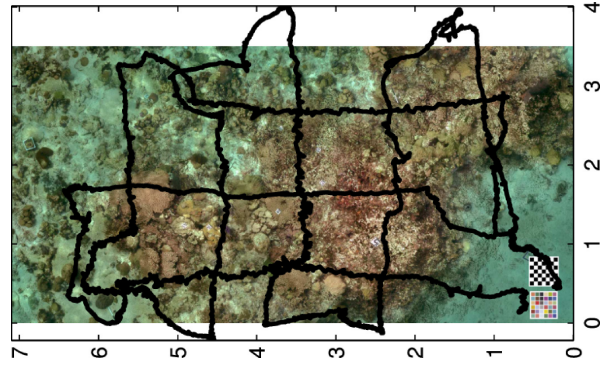


Fig. 4. Ground-truth trajectory superimposed on the mosaic. Images from the camera are matched against the mosaic to estimate the true position.

### B. Computation of the Ground-Truth

The ground-truth trajectory is obtained by registering images from the camera to the original image of the poster mosaic. The estimation of the camera poses is parametrised on the 6 DoF translation and rotation wrt the local reference frame, denoted by  $\Theta$ .

Estimation of the ground-truth pose relies on the identification of point correspondences using feature-based robust matching [27]. The outcome of the matching are two lists of correspondences  $\mathbf{x}^i$  and  $\mathbf{x}^m$  of points in the camera image and the poster image respectively, related by the observation equation

$$\mathbf{x}^i = \mathbf{Q}(\mathbf{x}^m, \Theta) + \varepsilon, \quad (63)$$

where  $\mathbf{Q}$  is a projection function that takes into account the pose, the camera calibration and the scale of the mosaic poster, and  $\varepsilon$  is assumed to be Gaussian random noise.

The maximum likelihood estimate of  $\Theta$  is

$$\Theta_{ML} = \arg \min_{\Theta} |\mathbf{x}^i - \mathbf{Q}(\mathbf{x}^m, \Theta)|^2. \quad (64)$$

The ground-truth pose is then obtained by carrying out a minimization using a non-linear least squares algorithm [28], [29].

### C. Results

We summarise the results from the experiment in this section. Fig. 4 shows the ground-truth trajectory followed by the vehicle superimposed on the sea-floor mosaic. Figures 5 and 6 show the trajectories from the EKF dead reckoning and SC-PHD SLAM algorithms. The dead reckoning trajectory suffers from significant deviation, which we show to be the result of inaccurate heading measurements. On the other hand, the trajectory obtained from the SLAM does not suffer the same inconsistencies and provides a closer match to the ground-truth.

Fig. 7 shows the squared error in the position  $(x, y)$  estimated by the dead reckoning and SLAM. The dead reckoning

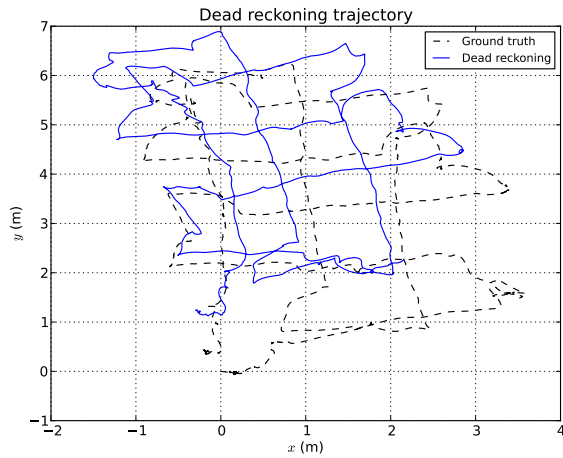


Fig. 5. Trajectory estimated from the dead reckoning. The trajectory shows large deviation due to inaccuracy in the heading angle.

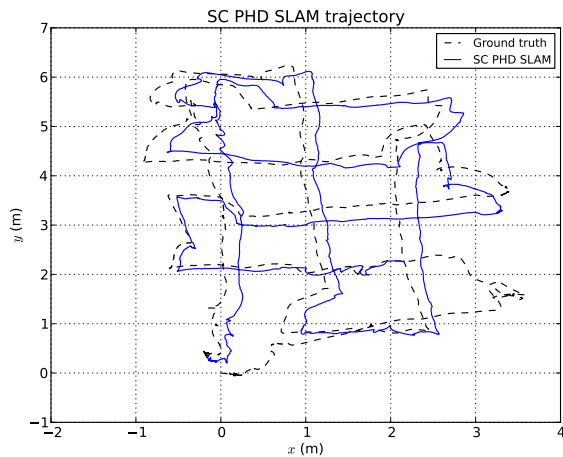


Fig. 6. Trajectory estimated from the SC-PHD SLAM. Errors in the heading angle can be corrected resulting in lower deviation compared to the dead reckoning.

exhibits a much larger error which we attribute to inaccurate vehicle heading obtained from the magnetic compass. Fig. 8 illustrates the error in heading estimated by the compass and the ground-truth (with an added offset such that the average error is zero). The figure shows that the compass is prone to large time-varying deviations between  $-15^\circ$  and  $+15^\circ$ . In the absence of a fixed bias, the trajectory estimated by the dead reckoning suffers from the time-varying nature of the estimated heading, resulting in a poor estimate of the trajectory. On the other hand, incorporating the heading into the state space allows it to be estimated by the SLAM algorithm and results in a closer match to the ground-truth.

In addition to improved localisation, the SLAM also produces an estimate of the map features. Figures 9 and 10 illustrate the confirmed landmarks selected from the map intensity function. The features lie in a plane corresponding to the bottom of the tank with a standard deviation of 6 cm. Errors in the camera calibration result in larger uncertainty in the 3D position of the landmarks.

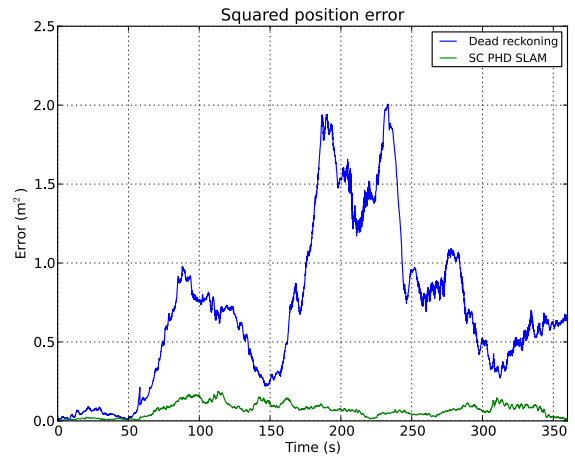


Fig. 7. Squared position error for the dead reckoning and SLAM.

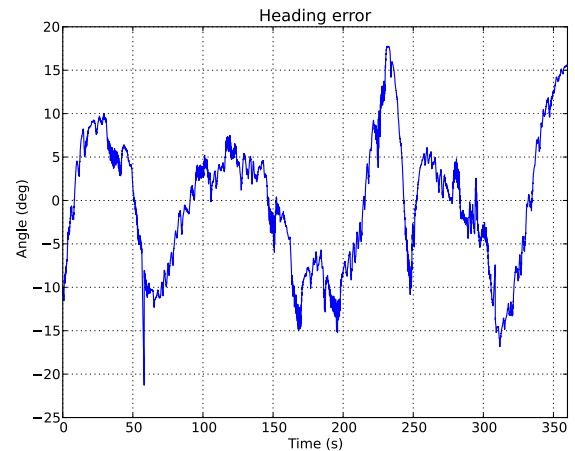


Fig. 8. Squared position error for the dead reckoning and SLAM.

## V. CONCLUSIONS

This paper has presented an application of the single-cluster PHD SLAM algorithm in underwater robotics. The Girona 500 AUV is equipped with a downward looking stereo camera and features extracted from the camera are used by the algorithm to improve the vehicle localisation. The trajectory from the SC-PHD SLAM algorithm is compared with the dead reckoning from an extended Kalman filter. The results show a significant reduction in the localisation error of the vehicle trajectory in the case of the SLAM algorithm.

In particular, we have shown that the SC-PHD SLAM algorithm is able to compensate for errors in the heading angle obtained from a magnetic compass as well as reduce the drift due to measurement noise from the DVL. The SC-PHD SLAM algorithm produces a geometric map of detected features. In the scenario tested here, this consists of points lying on a plane. Further work will consider the sonar-based mapping as well as the closed form implementation using a Gaussian approximation for the vehicle position.



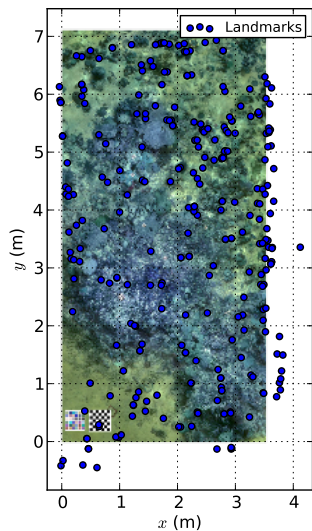


Fig. 9. Confirmed landmarks from the SLAM overlaid on the mosaic.

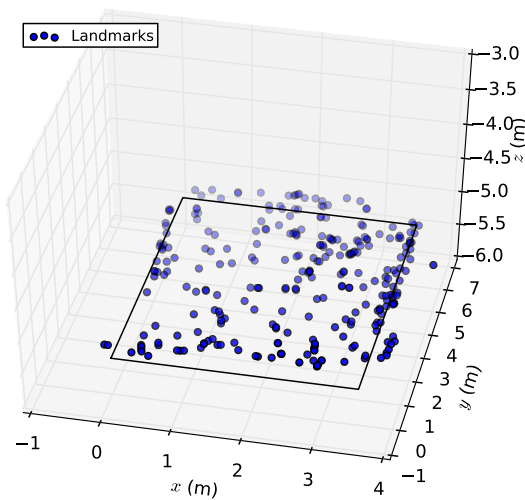


Fig. 10. 3D coordinates of the confirmed landmarks. The rectangle corresponds to the mosaic in the test tank.

## REFERENCES

- [1] "PANDORA FP7 project," April 2013. [Online]. Available: <http://persistentautonomy.com/>
- [2] R. Smith, M. Self, and P. Cheeseman, "Autonomous robot vehicles," I. J. Cox and G. T. Wilfong, Eds. New York, NY, USA: Springer-Verlag New York, Inc., 1990, ch. Estimating uncertain spatial relationships in robotics, pp. 167–193. [Online]. Available: <http://dl.acm.org/citation.cfm?id=93002.93291>
- [3] J. Leonard and H. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," in *Intelligent Robots and Systems '91. Intelligence for Mechanical Systems, Proceedings IROS '91. IEEE/RSJ International Workshop on*, 1991, pp. 1442–1447 vol.3.
- [4] R. Martinez-Cantin and J. Castellanos, "Unscented SLAM for large-scale outdoor environments," in *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, 2005, pp. 3427–3432.
- [5] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem," in *In Proceedings of the AAAI National Conference on Artificial Intelligence*. AAAI, 2002, pp. 593–598.
- [6] M. Montemerlo, S. Thrun, D. Roller, and B. Wegbreit, "FastSLAM 2.0: an improved particle filtering algorithm for simultaneous localization and mapping that provably converges," in *Proceedings of the 18th international joint conference on Artificial intelligence*, ser. IJCAI'03. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2003, pp. 1151–1156.
- [7] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, October 2003.
- [8] R. P. S. Mahler, *Statistical Multisource Multitarget Information Fusion*. Artech House, 2007.
- [9] J. Mullane, B.-N. Vo, M. D. Adams, and W. S. Wijesoma, "A random set formulation for bayesian slam," in *IROS*, 2008, pp. 1043–1049.
- [10] J. Mullane, B.-N. Vo, M. Adams, and B.-T. Vo, "A random-finite-set approach to bayesian slam," *Robotics, IEEE Transactions on*, vol. 27, no. 2, pp. 268 –282, april 2011.
- [11] D. Moratuwage, B.-N. Vo, D. Wang, and H. Wang, "Extending Bayesian RFS SLAM to multi-vehicle SLAM," in *Control Automation Robotics Vision (ICARCV), 2012 12th International Conference on*, 2012, pp. 638–643.
- [12] C. S. Lee, D. Clark, and J. Salvi, "Slam with single cluster phd filters," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, may 2012, pp. 2096 –2101.
- [13] C. S. Lee, D. E. Clark, and J. Salvi, "SLAM With Dynamic Targets via Single-Cluster PHD Filtering," *IEEE J. Sel. Topics Signal Process. (accepted)*, vol. PP, pp. 1–1, 2013.
- [14] A. Swain and D. Clark, "The single-group phd filter: An analytic solution," in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, july 2011, pp. 1 –8.
- [15] B.-N. Vo, S. Singh, and A. Doucet, "Sequential monte carlo implementation of the phd filter for multi-target tracking," in *Proceedings of the Sixth International Conference of Information Fusion*, vol. 2, 2003, pp. 792 – 799.
- [16] T. Zajic and R. P. S. Mahler, "Particle-systems implementation of the phd multitarget-tracking filter," in *Signal Processing, Sensor Fusion, and Target Recognition XII*, I. Kadar, Ed., vol. 5096, no. 1. SPIE, 2003, pp. 291–299.
- [17] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, October 2005.
- [18] B.-N. Vo and W.-K. Ma, "The Gaussian Mixture Probability Hypothesis Density Filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, November 2006.
- [19] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic Implementations of the Cardinalized Probability Hypothesis Density Filter," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3553–3567, July 2007.
- [20] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1523–1543, October 2007.
- [21] B. Ristic, D. E. Clark, B.-N. Vo, and B.-T. Vo, "Adaptive target birth intensity in PHD and CPHD filters," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1656–1668, 2012.
- [22] E. A. Wan and R. V. D. Merwe, "The Unscented Kalman Filter," in *Kalman Filtering and Neural Networks*, S. Haykin, Ed. John Wiley & Sons Inc, 2001, pp. 221–280.
- [23] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, Jun. 2008. [Online]. Available: <http://dx.doi.org/10.1016/j.cviu.2007.09.014>
- [24] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.
- [25] G. Bradski, "Open source computer vision library," <http://opencv.willowgarage.com/wiki/>, 2006. [Online; accessed April-2013]. [Online]. Available: <http://opencv.willowgarage.com/wiki/>
- [26] D. Ribas, N. Palomeras, P. Ridao, M. Carreras, and A. Mallios, "Girona 500 AUV, from survey to intervention," *IEEE/ASME Transactions on Mechatronics*, vol. 17, no. 1, p. 4653, February 2012.
- [27] N. Gracias, "Mosaic-based Visual Navigation for Autonomous Underwater Vehicles," Ph.D. dissertation, Instituto Superior Técnico, Lisbon, Portugal, June 2003. [Online]. Available: [vislab.isr.ist.utl.pt](http://vislab.isr.ist.utl.pt)
- [28] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988.
- [29] N. Gracias and J. Santos-Victor, "Trajectory reconstruction with uncertainty estimation using mosaic registration," *Robotics and Autonomous Systems*, vol. 35, pp. 163–177, July 2001.