8th IFAC Conference on Control Applications in Marine Systems
Rostock-Warnemünde, Germany
September 15-17, 2010

Submapping SLAM based on acoustic data from a 6-DOF AUV ⋆

Josep Aulinas * **Chee Sing Lee** * **Joaquim Salvi** *
Yvan R. Petillot **

* *Computer Vision and Robotics group (ViCoRob), University of Girona, 17071 Girona, Spain (e-mail: {cslee,jaulinas,qsalvi}@eia.udg.edu)*
** *Oceans Systems Lab, School of EPS, Heriot-Watt University, EH144AS Edinburgh, United Kindgom (e-mail: Y.R.Petillot@hw.ac.uk)*

Abstract: Autonomous Underwater Vehicles (AUVs) need positioning systems besides the Global Positioning System (GPS), since GPS does not work in underwater scenarios. Possible solutions are the Simultaneous Localization and Mapping (SLAM) algorithms. SLAM algorithms aim to build a map while simultaneously localizing the vehicle within this map. However, they offer limited performance when faced with large scale scenarios. For instance, they do not create consistent maps for large areas, mainly because uncertainties increase with the scale of the scenario. In addition, the computational cost increases with the map size. The use of local maps reduces computational cost and improves map consistency. Following this idea, in this paper we propose a new SLAM approach that uses independent local maps together with a global level stochastic map. The global level contains the relative transformations between local maps. These local maps are updated once a new loop is detected. Local maps that are sharing a high number of features are updated through fusion, maintaining the correlation between landmarks and vehicle. Experimental results on real data obtained from the REMUS-100 AUV show that our approach is able to obtain large map areas consistently.

Keywords: Autnonomous vehicles, robotics, navigation

1. INTRODUCTION

Mapping and localization techniques are necessary for many underwater applications. Some examples of these applications are underwater cartography, geological mapping, off-shore structures inspection, studies of biodiversity or deep-water archaeology. Different underwater vehicles have been developed in order to explore completely unknown underwater regions, for instance the so called Autonomous Underwater Vehicles (AUVs). An AUV is equipped with onboard sensors, which provide information about the vehicle, such as speeds, orientations or accelerations, and about the environment, such as 3D clouds of points from the sea floor or the relative location of salient features with respect to the vehicle. This information is very valuable to calculate the approximate position of the vehicle.

Terrestrial and aerial vehicles can localize themselves with Global Positioning System (GPS). However, underwater, GPS can not be used because electromagnetic waves are strongly attenuated through the medium of water. A standard for bounded `xyz` navigational position measurements for underwater vehicles is the long-baseline (LBL) acoustic transponder system (Hunt et al. (1974)). LBL operates on the principle of time-of-flight and it is been proven to operate up to a range of 10 km (Whitcomb et al. (1999)). The main drawback of LBL is that it requires two or more acoustic transponder beacons to be tethered to the sea floor. Short-baseline (SBL) systems provide more accurate positioning information, but suffer from the same drawbacks than the LBL. Internal sensors, such as the Inertial Measurement Unit (IMU) and the Doppler Velocity Log (DVL) do not give absolute localization, therefore the localization problem suffers from drift due to odometric noise. Furthermore, the detection of salient features in the environment is a complex task due to measurement noise. These noises makes the mapping and localization a difficult challenge. Simultaneous Localization and Mapping (SLAM), also known as Concurrent Mapping and Localization (CML), is one of the fundamental challenges of robotics (Durrant-Whyte and Bailey (2006)). The SLAM problem involves a joint task of simultaneously estimating the map and localizing the vehicle inside this map.

A well known and widely used SLAM approach is the Extended Kalman Filter SLAM (EKF-SLAM) (Smith et al. (1988)). EKF-SLAM represents the vehicle's pose and the location of a set of environment features in a joint state vector. This vector is estimated and updated by the EKF. The EKF provides a suboptimal solution due to

⋆ The authors also acknowledge the support of the research project DPI-2007-66796-C03-02 funded by the Spanish Ministry of Science and Innovation. J. Aulinas holds a University of Girona scholarship (BR-07/03). C.S. Lee holds a grant from the Catalan government (IUE/2365/2009)

several approximations and assumptions, which result in divergences (Castellanos et al. (2007)). In large areas, EKF complexity grows with the number of landmarks, because each landmark is correlated to all other landmarks. This means that EKF memory complexity is $O(n^2)$ and a time complexity of $O(n^2)$ per step, where n is the total number of features stored in the map.

The use of submaps has been shown to address the problems of consistency and computational complexity. An early example of this strategy is the Decoupled Stochastic Map (DSM) (Leonard and Feder (2000)). The DSM uses non-statistically independent submaps. Therefore the correlations are broken introducing inconsistency in the map. Similar inconsistencies were seen in (Aulinas et al. (2010)), where submaps were assumed to be independent but still shared information. Different techniques, such as the Constrained Local Submap Filter (CLSF) (Williams et al. (2002)) or Local Map Joining (MJS) (Tardós et al. (2002)) produce efficient global maps by consistently combining completely independent local maps. The Divide and Conquer SLAM (DCS) (Paz et al. (2008)) is capable to recover the global map in approximately $O(n)$ time. The Constant Time SLAM (CTS) (Leonard and Newman (2003)), the Atlas approach (Bosse et al. (2004)), and the Hierarchical SLAM (HS) (Estrada et al. (2005)) store the link between local maps by means of an adjacency graph. The HS imposes loop constraints on the adjacency graph, producing a better estimation of the global level map. The Conditionally Independent Local Maps (CILM) (Piniés and Tardós (2008)), is based on sharing information between consecutive submaps. This way, a new local map is initialized considering the *a-priori* knowledge.

These submapping techniques demonstrate that using submaps, both linearization errors and computational cost can be addressed at the same time, improving the consistency of EKF-SLAM (Castellanos et al. (2007)). Only few of them have been tested on underwater scenarios (Williams (2001); Roman and Singh (2007)), where some extra constraints have to be taken into account. Firstly, the terrain sensing is limited to either acoustics (Ribas (2008)) or near-field vision (Eustice (2005)), because electromagnetic waves are strongly attenuated in the water. Secondly, underwater scenarios are in general unstructured and require 3D navigation (6-DOF motion), while most current SLAM solutions are used on man-made (geometrically simple) indoor spaces, where a 2D map representation is sufficient. Therefore, the use of SLAM on AUV navigations requires further testing and improvements.

The main contribution of our approach is the strategy used to decide whether to fuse the submaps. This decision is made on the basis that fusing two maps that share many landmarks will produce a better update than fusing two maps that only share a few landmarks. The experiments done with real data show a a bounded effect of the linearization error and also a precise reconstruction of the map since the drift suffered in shorter distances is smaller, and the data association can be more robustly solved as compared to other state of the art techniques.

The rest of the paper is structured as follows: Section 2 describes the novelty of our SLAM approach. The standard EKF and the map fusion approach are also presented. Section 3 describes the experimental setup and the results obtained using a 6-DOF vehicle. Finally, conclusions and future work are presented in Section 4.

Algorithm I: Selective Submap Joining SLAM

```
begin mission
while navigating do
    x̂_i, P̂_i = EKF SLAM() ← (Build submap M_i)
    x̂_G, P̂_G = build_global_map(x̂_i, P̂_i)
    H_Loop = check_possible_loops(x̂_G, P̂_G)
    for j = H_Loop do
        refer M_i and M_j to a common base reference
        H_ij = data_association(x̂_i, x̂_j, P̂_i, P̂_j)
        if H_ij > threshold then
            x̂_ij, P̂_ij = map_fusion(x̂_i, P̂_i, x̂_j, P̂_j, H_ij)
            x̂_G, P̂_G = update_global_map(x̂_ij, P̂_ij)
        endif
    endfor
endwhile
```

2. SELECTIVE SUBMAP JOINING BASED SLAM

The basis of the Selective Submap Joining SLAM (SSJS) (see Algorithm I) lies in the EKF-based SLAM. A sequence of EKF-based submaps is built, as explained in Subsection 2.1. The size of these submaps is predefined by the total number of features per map and by the uncertainty boundaries. The links between local maps are stored in a global level map, as described in Subsection 2.2. This graph information allows checking whether a loop closing event is occurring, following the strategy presented in Subsection 2.3. The main novelty of our approach lies in the fact that upon loop closure, we decide to fuse two maps or to keep them independent depending on the number of common landmarks, in contrast to other approaches that fuse maps regardless of the information they share (Williams (2001); Tardós et al. (2002); Estrada et al. (2005); Paz et al. (2008)).

2.1 Map Building

A map is built using a standard EKF algorithm (Algorithm II). The EKF estimates the state, at a certain time step k, of a dynamic non-linear system from a series of incomplete and noisy measurements, as its mean x_k and the covariance P_k. The algorithm iterates continously through three steps: prediction, observation and update. The prediction stage uses the motion model f to estimate the current state $\widehat{\mathbf{x}}_k$ from the previous time step $\mathbf{x}_{k-1}$, and control inputs (i.e. odometry) $\mathbf{u}_k$, if available. (see (1)). The hat notation denotes an estimate based only on this prediction, before corrections from sensor input.

$$\widehat{\mathbf{x}}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) \qquad \widehat{\mathbf{P}}_k = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T \quad (1)$$

In general, the motion model is a non-linear function, which requires the following linearizations for predicting the state covariance at time k:

$$\mathbf{F}_k = \frac{\partial f}{\partial \mathbf{x}} |_{\mathbf{x}_{k-1}} \qquad \mathbf{G}_k = \frac{\partial f}{\partial \mathbf{x}} |_{\mathbf{u}_k} \quad (2)$$