Vision-Based Underwater SLAM for the SPARUS AUV

Josep Aulinas, Institute of Informatics and Applications, Girona/Spain, jaulinas@eia.udg.edu
Yvan R. Petillot, Oceans Systems Laboratory, Edinburgh/UK, Y.R.Petillot@hw.ac.uk
Xavier Lladó, Institute of Informatics and Applications, Girona/Spain, llado@eia.udg.edu
Joaquim Salvi, Institute of Informatics and Applications, Girona/Spain, qsalvi@eia.udg.edu
Rafael Garcia, Institute of Informatics and Applications, Girona/Spain, rafa@eia.udg.edu

Abstract

An overview of underwater SLAM implementations as well as submapping SLAM approaches is given in this paper. Besides, the implementation of the so called selective submap joining SLAM on the SPARUS AUV is presented. SPARUS carries a down-looking optical camera. The information gather by this camera is run through SLAM, together with on-board navigation sensors, producing a precise localization of the vehicle and a consistent final map. Experimental validation on a real dataset is described, showing a promising performance of our implementation.

1. Introduction

In the last decade, different underwater vehicles have been developed in order to explore underwater regions, especially those of difficult access for humans. The use of Remotely Operated underwater Vehicles (ROVs) is very common, however, ROVs require a link, i.e. a tether, to the ship in order to be operated by a person aboard of the ship. The tether is a group of cables that carry electrical power, video and data signals back and forth between the operator and the vehicle. In order to avoid the need for a tether, several research groups and developers focused on developing Autonomous Underwater Vehicles (AUVs). AUVs are equipped with on-board sensors, which provide valuable information about the vehicle state and the environment. Combining this information with control algorithms makes the vehicle fully autonomous.

Some examples of these AUVs are the ICTINEU (see Fig. 1) and the SPARUS (see Fig. 2) developed by VICOROB research group at the University of Girona. Both vehicles were developed with the purpose to participate in the Student Autonomous Underwater Challenge – Europe (SAUC-E) competition. ICTINEU won the 2006 SAUC-E edition, while SPARUS is a more recent development, torpedo shaped AUV, which won the 2010 SAUC-E edition. SPARUS is the vehicle we used to test the method presented in this work.

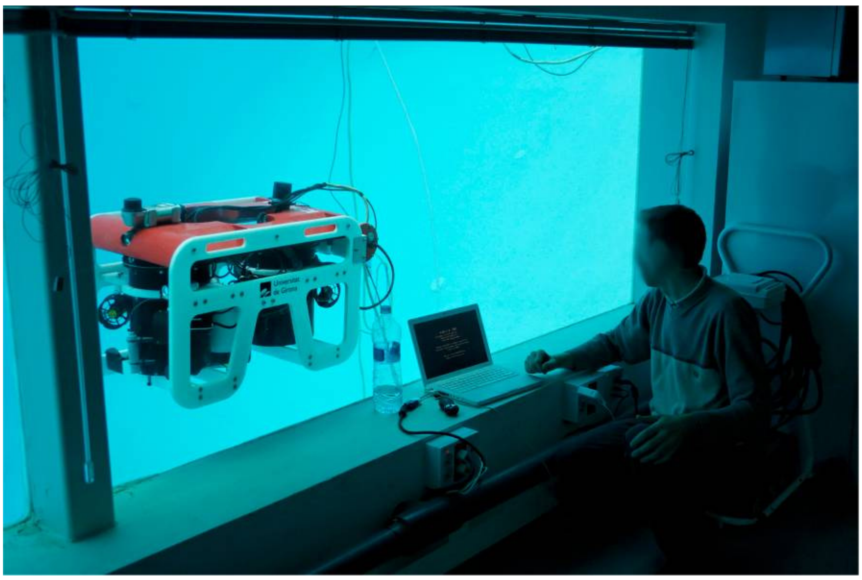

Fig.1: Picture of ICTINEU AUV operating inside the test water tank at the Underwater Robotics Research Center (CIRS – Girona, Spain).

Fig.2: Picture of SPARUS AUV operating open water under the supervision of a diver.

Some widely used sensors for land and aerial robots do not work or are not precise enough underwater. For instance, the use of cameras is difficult due to the lack of visibility and scattering; the laser range finders are imprecise working in these scenarios because of light attenuation; and GPS signal does not work underwater. The sensors used on SPARUS AUV are the Inertial Measurement Unit (IMU) and the Doppler Velocity Log (DVL) to measure navigation data, while a down looking camera is used to gather data from the environment. The IMU and the DVL do not give absolute localization, therefore if the vehicle is wrongly localized, nor the IMU neither the DVL will provide useful information to recover the right position. In addition, as the positioning is relative to past information, the localization problem is biased and the measurement noise produces drift. On the other hand, the detection of salient features in this environment is a complex task, since camera images are noisy. Noise together with lack of other navigation aids makes the task of mapping and localization a difficult challenge.

A solution to the lack of GPS signal and the presence of noise are the Simultaneous Localization and Mapping (SLAM) algorithms. SLAM algorithms aim to build an approximate map of the area and calculate the approximate position of the vehicle within this map. In order to do so, SLAM algorithms combine the information coming from all sensors. Our SLAM approach, called the selective submapping SLAM, uses navigation readings to improve vehicle localization, and the map through its correlation with the vehicle position. To have a SLAM algorithm working properly, we need to select robust landmarks, i.e. objects, rocks and other salient elements. These robust landmarks must be easy to observe when seen for a second time, and easy to associate with previous observations. This procedure is important to close a loop, i.e. revisiting an area, because closing a loop means a reduction on the uncertainty and a more consistent final map.

In what follows, a background on underwater SLAM implementations is first given in Section 2. Afterwards, a summary on existing SLAM algorithms working with submaps is presented. Section 3 describes the implementation of our approach on SPARUS. Section 4 presents the experimental validation, while Section 5 gives the conclusions.

2. Background

This section surveys main existing SLAM implementations for underwater applications, focusing on the filtering technique used to handle noise and drift uncertainty, the main sensor used to gather data from the scenario, and the type of feature used to build the map. Afterwards, submapping SLAM approaches are summarized.

2.1. Underwater SLAM

Several approaches tackle the localization problem on known scenarios. Some approaches use GPS-aided localization, *Caiti et al. (2005), Erol et al. (2007)*, but the attenuation of electromagnetic waves through the medium of water limits the application of GPS to near surface activities, or otherwise forces the vehicle to visit often the surface to recover its position. A standard for bounded xyz navigational position measurements for underwater vehicles is the Long-BaseLine (LBL) acoustic transponder system, *Hunt et al. (1974), Olson et al. (2006).*

The equivalents to GPS underwater are the acoustic transponders, such as LBL or Short-BaseLine (SBL). These positioning systems have limited range, accuracy and an associated cost of deployment. 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. SBL systems provide more accurate positioning information, but suffer from the same drawbacks than the LBL. Recently, several AUVs use Ultra Short-Baseline (USBL) technology, which consists of a transceiver, usually placed on the surface, on a pole under the vessel, and a transponder mounted on the AUV. This technology is more accurate than LBL and SBL. Another set of approaches avoid the use of external devices by using computer algorithms. For instance, the use of particle filters for AUV localization presented in *Maurelli et al. (2008)*. This approach is shown to work with high performance. However, it only works when the map is known a-priori.

When the map is unknown, SLAM is conducted. Underwater scenarios are still one of the most challenging scenarios for SLAM because of reduced sensory possibilities. Underwater SLAM approaches have many problems due to the unstructured nature of the seabed and the difficulty to identify reliable features. Many underwater features are scale dependant, sensitive to viewing angle and scale. A SLAM proposal tackles the problem using point features, *Williams et al. (2004)*. This approach proposed to fuse information from the vehicle's on-board sonar and vision systems. They use EKF based SLAM combined with sonar and vision to obtain 3D structure and texture (see Fig. 3).

Leonard et al. (2001) and *Newman et al. (2003)* also used point features. The former implemented the decoupled stochastic mapping and performed tests on a water tank, while the later proposed the constant time SLAM and used LBL information to help on the localization. On the other hand, non-feature based approaches to SLAM using bathymetric information were presented by *Barkby et al. (2009)* and *Roman et al. (2007).*

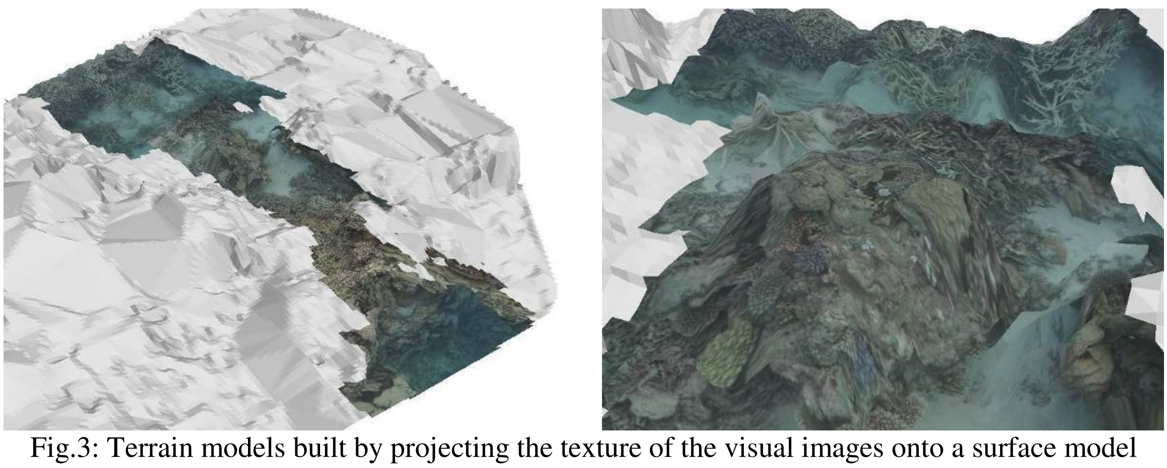

Fig.3: Terrain models built by projecting the texture of the visual images onto a surface model generated by sonar data, *Williams (2004)*

A particle filter is used to handle the uncertainty in the navigation solution provided by the vehicle, *Fairfield et al. (2008)*. This approach was successful in minimizing the navigation error during a deep sea mapping mission. The method was capable of providing real-time localization, with comparable results to the ones given by SBL and USBL. A vision-based localization approach for an underwater robot in a structured environment was presented in *Carreras et al. (2003)*. The system was based on a coded pattern placed on the bottom of a water tank and an on-board down-looking camera. The system provided three-dimensional position and orientation of the vehicle along with its velocity. Another vision-based algorithm, *Eustice et al. (2008)*, used inertial sensors together with the typical low overlap imagery constraints of underwater imagery. Their strategy consisted on solving a sparse system of linear equations in order to maintain consistent covariance bound within a SLAM information filter. The main limitation on vision-based techniques is that they are limited to near field vision (1-5m), and also deep water mission will require higher amounts of energy for lighting purposes. In a previous works, *Eustice et al. (2005)* and *Eustice et al. (2006)*, they presented the reconstruction of the RMS Titanic from a set of images and using IF. Using Sparse Extended Information filter (SEIF) and forward-looking sonar, *Walter et al. (2008)* presented a SLAM approach to inspect ship hull.

Instead of vision, *Ribas et al. (2008)* used mechanically scanned imaging sonar to obtained information about the location of vertical planar structures present in partially structured environments. In this approach, the authors extracted line features from sonar data, by means of a robust voting algorithm (see Fig. 4). These line features were used in the EKF base SLAM.

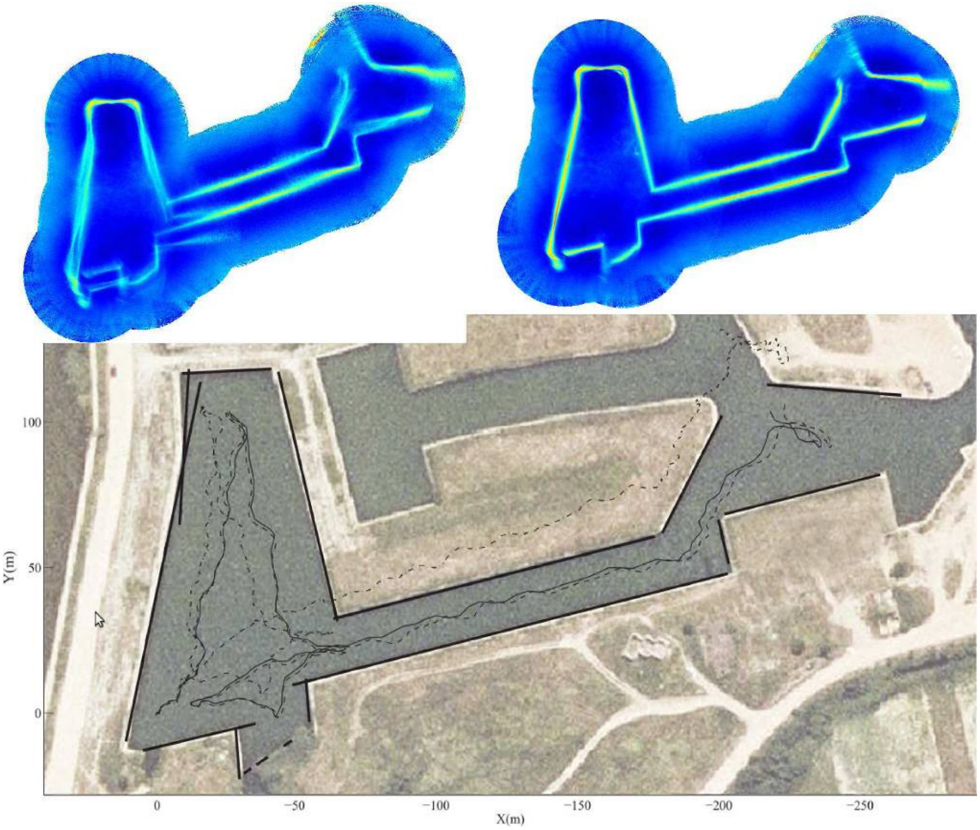

Fig.4: Abandoned marina SLAM example, using imaging sonar, image extracted from *Ribas et al. (2008)*. Top left plot shows the superposition of imaging sonar readings, based on dead reckoning trajectory. Top right plot is the same superposition but in this case after filtering the trajectory through EKF base SLAM. Bottom picture shows a satellite image of the abandoned marina, with the lines representing the boundary between water and land. The trajectory of the vehicle is plotted, using dead reckoning estimates (dashed line) and using SLAM algorithms (solid line).

In *Tena-Ruiz et al. (2004)* side-scan sonar was used to sense the environment. The returns from the sonar were used to detect landmarks in the vehicle's vicinity. Observing these landmarks allows correcting the map and vehicle location; however, after long distances the drift is too large to allow associating landmarks with current observations. For this reason, they proposed a method that combines a forward stochastic map in conjunction with a backward Rauch-Tung-Striebel (RTS) filter to smooth the trajectory.

Underwater SLAM implementations have some points in common, for instance, imaging sonar is widely used, the most common filtering technique is the Extended Kalman Filter (EKF) and point features are commonly used to represent the map. Some approaches use side-scan sonar or optical cameras, which seems to become more important as technology advances. The use of EKF based SLAM has been shown to handle uncertainties properly; however, the computational cost associated with EKF grows with the size of the map. In addition, linearization errors accumulate in long missions, increasing the chance of producing inconsistent mapping solutions.

2.2. Submapping SLAM

The use of submaps has been shown to address both issues, linearization errors and computational costs, at the same time, thereby improving the consistency of EKF based SLAM, *Castellanos et al. (2007)*. An early example of this strategy is the decoupled stochastic mapping, *Leonard et al. (2001)*, which uses non-statistically independent submaps. As a result, correlations are broken and inconsistency is introduced into the map. The constant time SLAM, *Newman et al. (2003)*, uses multi overlapping local submaps with the frame referenced to one of the features in the submap. This technique maintains a single active map and computes a partial solution, independently. However in non-linear cases the consistency is not proven.

Different techniques, such as the constrained local submap filter, *Williams et al. (2004)* or the local map joining, *Tardós et al. (2002)*, produce efficient global maps by consistently combining completely independent local maps. The main idea behind this approach is to build maps of limited size and then, once completed, merge these small maps to a global one. The so called atlas SLAM, *Bosse et al. (2004)*, consists of a hierarchical strategy that achieves efficient mapping of large-scale environments. They used a graph of coordinate frames, with each vertex in the graph representing a local frame, and each edge representing the transformation between adjacent frames. In each frame, they build a map that captures the local environment and the current robot pose along with the associated uncertainties. The divide and conquer SLAM, *Paz et al. (2008)*, uses the divide and conquer strategy from fundamental graph theory. The hierarchical SLAM, *Estrada et al. (2005)*, consists on a lower (or local) map level, which is composed of a set of local maps that are guaranteed to be statistically independent, and the upper (or global) level, which is an adjacency graph whose arcs are labeled with the relative location between local maps. An estimate of these relative locations is maintained at this level in a relative stochastic map. Every time the vehicle closes a loop a global level optimization is performed, producing a better estimate of the whole map. Conditionally independent SLAM, *Piniés et al. (2008)*, is based on sharing information between consecutive submaps so that, a new local map is initialized with a-priori knowledge.

3. Implementation on SPARUS AUV

SPARUS is equipped with several sensing devices: Doppler velocity log (DVL), inertial measurement unit (IMU), down-looking camera, forward-looking camera, imaging sonar and GPS (see Fig. 5). In this work, only DVL, IMU and down-looking camera are used, producing information about velocities, orientations and about the sea floor. The SLAM approach used here is the so called selective submap joining algorithm, *Aulinas et al. (2010)*. The main idea of this approach is to use EKF based SLAM to build local maps ($\mathbf{x_i}$, $\mathbf{P_i}$), where $\mathbf{x_i}$ is the state vector describing vehicle's pose, vehicle's velocities and the map, while $\mathbf{P_i}$ is its associated uncertainty. The size of these local maps is bounded by the total number of features and by the level of uncertainty. The relative topological relationship between consecutive local maps is stored in a global level map ($\mathbf{x_G}$, $\mathbf{P_G}$). The global level is used to search for loop closure ($\mathcal{H}_{Loop}$), i.e. the vehicle is revisiting a region. The loop closing strategy involves a decision on whether to fuse local maps depending on the amount of found correspondences between submaps. The whole process is presented in *Algorithm I*, and detailed in *Aulinas et al. (2010*).

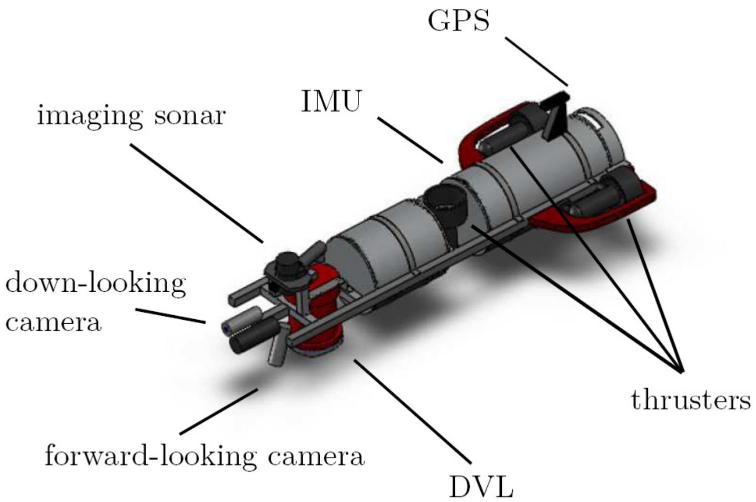

Fig.5: SPARUS 3D model with its sensors.

The main novelity in this implementation as compared to the one presented in *Aulinas et al (2010)* is the use of an optical system as the main environment sensor unit. Therefore, the observation model is redefined in order to match with a camera model. In this case, the inverse depth parametrization is used, *Civera et al. (2008).*

Algorithm I: Selective Submap Joining SLAM

begin mission
<u>**while**</u> navigating <u>**do**</u>
    $\widehat{\mathbf{x}}_i, \widehat{\mathbf{P}}_i$ = EKF SLAM() $\leftarrow$ (*Build submap* $\mathcal{M}_i$)
    $\widehat{\mathbf{x}}_G, \widehat{\mathbf{P}}_G$ = build global map($\widehat{\mathbf{x}}_i$, $\widehat{\mathbf{P}}_i$)
    $\mathcal{H}_{Loop}$ = check possible loops($\widehat{\mathbf{x}}_G$, $\widehat{\mathbf{P}}_G$)
    <u>**for**</u> $j = \mathcal{H}_{Loop}$ <u>**do**</u>
        refer $\mathcal{M}_i$ and $\mathcal{M}_j$ to a common base reference
        $\mathcal{H}_{ij}$ = data association($\widehat{\mathbf{x}}_i, \widehat{\mathbf{x}}_j, \widehat{\mathbf{P}}_i, \widehat{\mathbf{P}}_j$)
        <u>**if**</u> $\mathcal{H}_{ij} > threshold$ <u>**then**</u>
            $\widehat{\mathbf{x}}_{ij}, \widehat{\mathbf{P}}_{ij}$ = map fusion($\widehat{\mathbf{x}}_i$, $\widehat{\mathbf{P}}_i$, $\widehat{\mathbf{x}}_j$, $\widehat{\mathbf{P}}_j$, $\mathcal{H}_{ij}$)
            $\widehat{\mathbf{x}}_G, \widehat{\mathbf{P}}_G$ = update global map($\widehat{\mathbf{x}}_{ij}, \widehat{\mathbf{P}}_{ij}$)
        <u>**endif**</u>
    <u>**endfor**</u>
<u>**endwhile**</u>

4. Experimental validation

Experimental validation was done through the data acquired by SPARUS during a survey mission. The mission consisted of navigating an area of about 20m×20m, in a grid of 5m×5m. Vehicle's depth was almost constant around 17 meters. The total navigation time was about 17 minutes. The vehicle carried a down-looking camera that acquired a total of 3199 images, Fig. 6. Experimental results obtained with SLAM show that there is a significant improvement on trajectory estimate as compared to dead reckoning, Fig. 7.

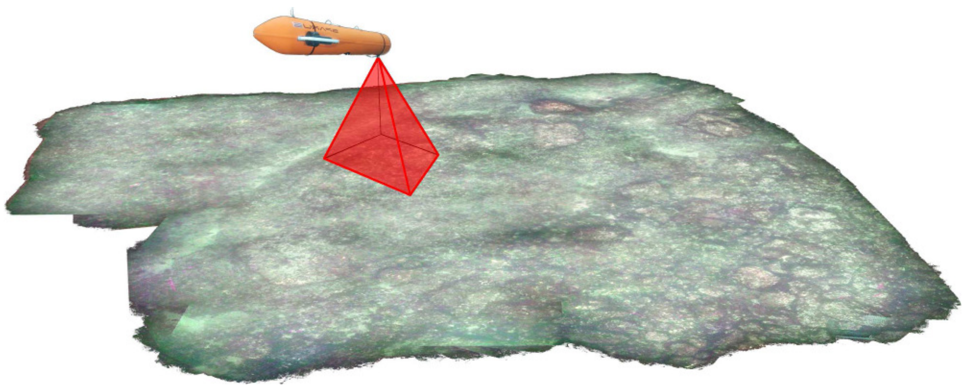

Fig.6: Working principle for the SPARUS down-looking camera

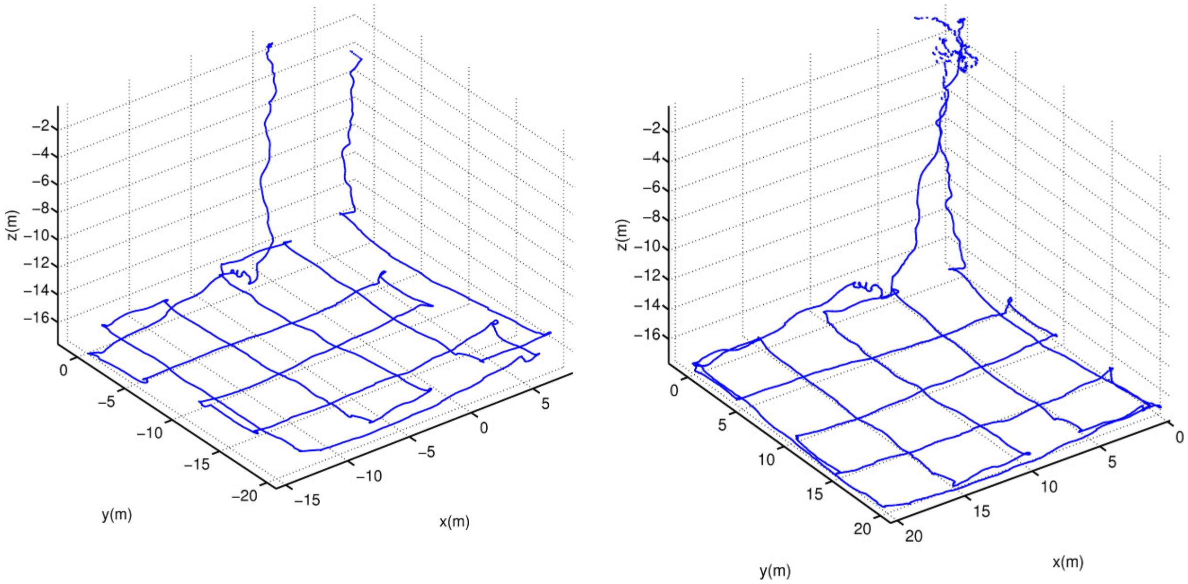

Fig.7: 3D view of vehicle's trajectory. Drift during mission with ending point far from the starting point (left) and drift corrected by using SLAM (right)

In order to test the performance of SLAM using submaps, a subset of random 2D points where extracted from a mosaic of the scene, *Garcia et al. (2006)*. These 2D points where then back referred to the image they belonged. This subset of points was used, instead of automatically detecting features. The performance of our SLAM implementation using this set of points is shown in Fig. 8. This figure shows a sequence of 5 frames containing first one landmark, and later on two landmarks. In addition, the uncertainty projected on the image plane is drawn, decreasing consistently after being observed for the second time. Fig. 9 presents a top and a frontal view of the resulting map and trajectory. In these views, one can see vehicle's trajectory corrected with SLAM and the landmark location, as well as its associated uncertainties. Finally, Fig. 10 shows a 3D plot of these results.

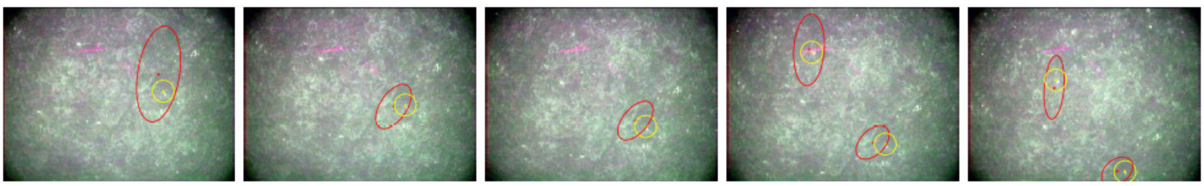

Fig.8: Sequence of down-looking camera frames. New observations and their associated measurement uncertainties are drawn, together with the prediction of a landmark that was already in the map, and the projection of its associated uncertainty onto the image plane. One can observe a reduction on uncertainty with new observations.

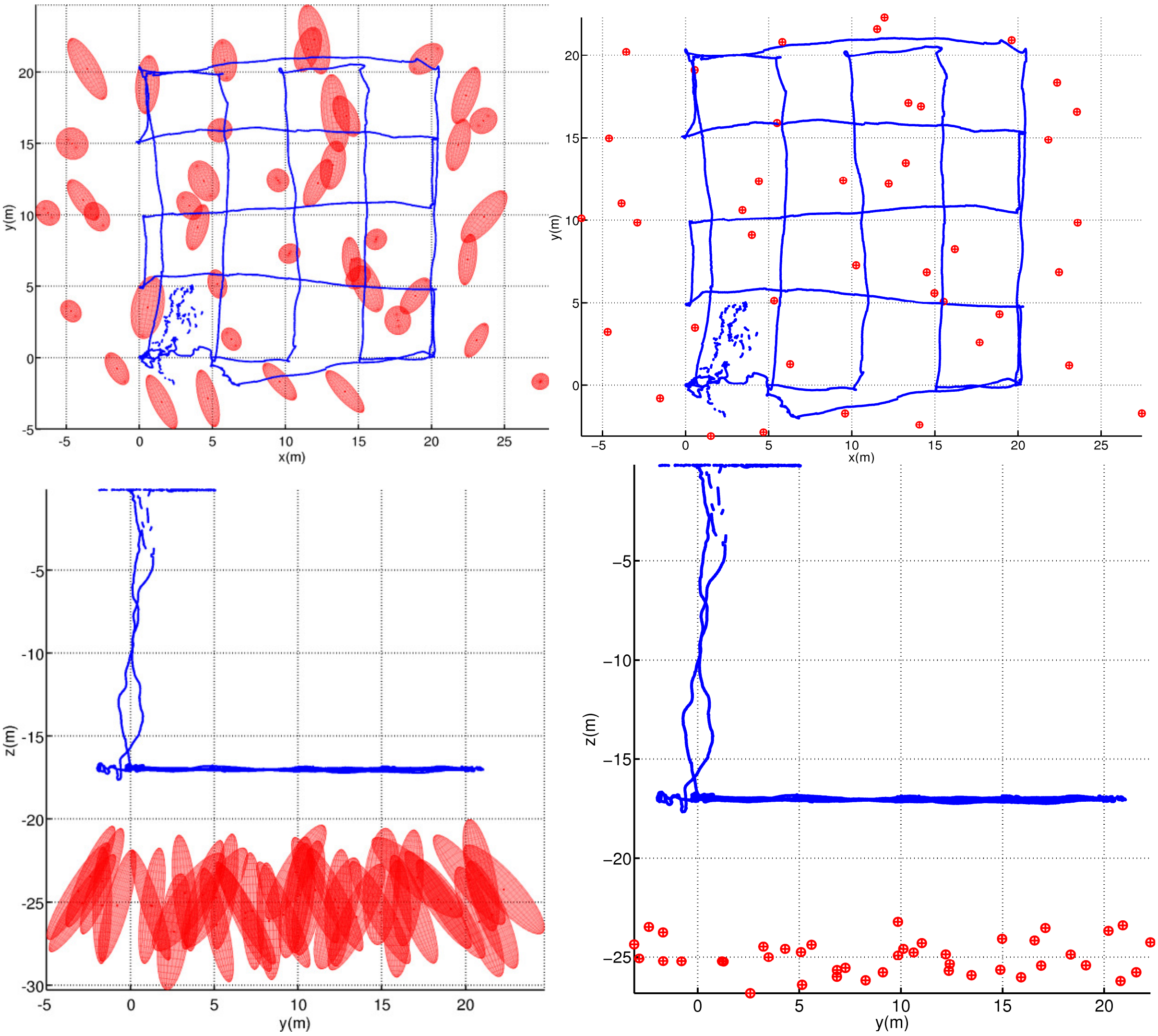

Fig.9: Different views of the results produced by SLAM. On the top, presents a top view of the execution, while on the bottom a frontal view is presented. On the left, landmark uncertainties are drawn, while on the right, only landmarks are shown. In all plots, vehicle trajectory is drawn.

5. Conclusions

The main contribution of this paper is a SLAM implementation for an underwater vehicle, SPARUS AUV. First, the most representative underwater SLAM implementations were surveyed, reaching the conclusion that Extended Kalman filter is widely used for this sort of applications. However, Extended Kalman filter suffers several limitations that can be addressed by using submaps. For this reason, a summary of the state-of-the-art on submapping approaches was presented. A SLAM algorithm is then briefly introduced, and adapted for its use on the SPARUS AUV. Experiments done with real data show a bounded effect of the linearization error, a precise trajectory estimates, and a three-dimensional map reconstruction. Besides, the observation model for a down-looking optical camera was introduced. This model was based on inverse depth parameterization. Experiments conducted in a real unstructured environment demonstrated that SLAM improves vehicle trajectory in comparison to dead reckoning. Moreover, SLAM combined with inverse depth parameterization was capable of producing a consistent map.

Acknowledgments

This work was partially funded through the Spanish Ministry of Education and Science (MCINN) under grant CTM2010-15216 and the EU under grant FP7-ICT-2009-248497.

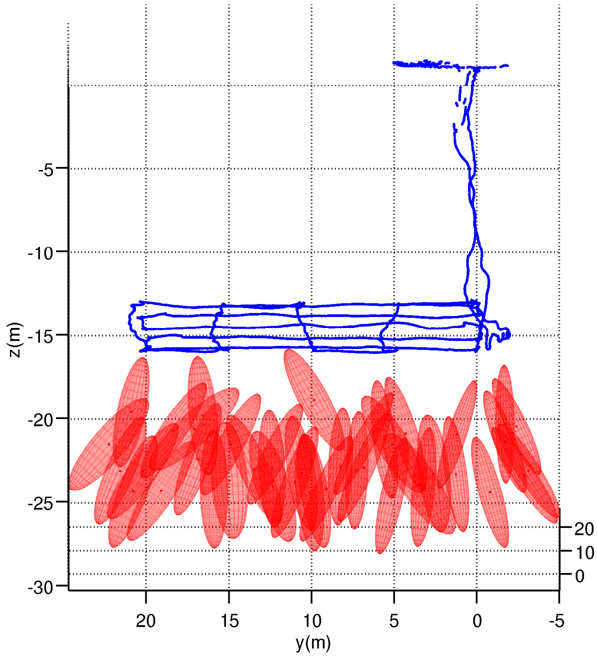

Fig.10: 3D plot of the SLAM solution

References

AULINAS, J.; LLADÓ, X.; SALVI, J.; PETILLOT, Y.R. (2010), *Selective submap joining for underwater large scale 6-DOF SLAM*. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), pp.2552-2557, Taipei.

BARKBY, S.; WILLIAMS, S.B.; PIZARRO, O.; JAKUBA, M. (2009), *Incorporating prior bathymetric maps with distributed particle bathymetric SLAM for improved AUV navigation and mapping*. MTS/IEEE Oceans Conference and Exhibition, vol. 2, Biloxi, pp.1-7, Biloxi.

BOSSE, M.; NEWMAN, P.M.; LEONARD, J.J.; TELLER, S. (2004), *SLAM in large scale cyclic environments using the atlas framework*. Int. J. Robotics Research 23/12, pp.1113-1139.

CAITI, A.; GARULLI, A.; LIVIDE, F.; PRATTICHIZZO, D. (2005), *Localization of autonomous underwater vehicles by floating acoustic buoys: A set membership approach*. IEEE J. of Oceanic Engineering 30/1, pp.140-152.

CARRERAS, M.; RIDAO, P.; GARCIA, R.; NICOSEVICI, T. (2003), *Vision-based localization of an underwater robot in a structured environment*. IEEE Int. Conf. Robotics and Automation, vol. 1, pp.971-976.

CASTELLANOS, J. A.; MARTÍNEZ-CANTIN, R.; TARDÓS, J. D.; NEIRA, J. (2007), *Robocentric map joining: Improving the consistency of EKF-SLAM*. Robotics and Autonomous Systems 55/1, pp.21-29.

CIVERA, J.; DAVISON, A.J.; MONTIEL, J.M.M. (2008), *Inverse depth parametrization for monocular SLAM*. IEEE Transactions on Robotics 24/5, pp.932-945.

EROL, M.; VIEIRA, L.; GERLA, M. (2007), *AUV-aided localization for underwater sensor networks.* Int. Conf. Wireless Algorithms, Systems and Applications (WASA 2007), pp.44-54.

ESTRADA, C.; NEIRA, J.; TARDÓS, J.D. (2005), *Hierarchical SLAM: real-time accurate mapping* *of large environments*. IEEE Transactions on Robotics 21/4, pp.588-596.

EUSTICE, R.M.; SINGH, H.; LEONARD, J.; WALTER, M.; BALLARD, R. (2005), *Visually navigating the RMS Titanic with SLAM information filters*. Robotics Science and Systems, pp.57-64.

EUSTICE, R.M.; SINGH, H.; LEONARD, J.; WALTER, M. (2006), *Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters*. Int. J. of Robotics Research 25/12, pp.1223-1242.

EUSTICE, R.M.; SINGH, H.; LEONARD, J.; WALTER, M. (2008), *Visually augmented navigation for autonomous underwater vehicles.* IEEE J. of Oceanic Engineering 33/2, pp.103-122.

FAIRFIELD, N.; WETTERGREEN, D. (2008), *Active localization on the ocean floor with multibeam sonar.* MTS/IEEE OCEANS, Quebec, pp.1-10.

GARCIA, R.; CUFÍ, X.; RIDAO, P.; CARRERAS, M. (2006), *Constructing photo-mosaics to assist UUV navigation and station-keeping.* Robotics and Automation in the Maritime Industries, Ch. 9, pp.195-234.

HUNT, M.M.; MARQUET, W.M.; MOLLER, D.A.; PEAL, K.R.; SMITH, W.K.; SPINDEL, R.C. (1974), *An acoustic navigation system*, Technical Report WHOI-74-6. Woods Hole Oceanographic Institution.

MAURELLI, F.; KRUPINSKI, S.; PETILLOT, Y.; SALVI, J. (2008), *A particle filter approach for AUV localization.* MTS/IEEE OCEANS, pp.1-7.

OLSON, E.; LEONARD, J.; TELLER, S. (2006), *Robust range-only beacon localization*. J. Oceanic Engineering 31/4, pp.949-958.

PAZ, L.M.; TARDÓS, J.D.; NEIRA, J. (2008), *Divide and Conquer: EKF SLAM in O(n).* IEEE Transactions on Robotics 24/5, pp.1107-1120.

PINIÉS, P.; TARDÓS, J.D. (2008), *Large scale SLAM building conditionally independent local maps: Application to monocular vision.* IEEE Trans. Robotics 24/5, pp.1094-1106.

RIBAS, D.; RIDAO, P.; TARDÓS, J.D.; NEIRA, J. (2008), *Underwater SLAM in man made structured environments*, J. Field Robotics 25/11-12, pp.898-921.

ROMAN, C.N.; SINGH, H. (2007), *A Self-consistent bathymetric mapping algorithm.* J. Field Robotics 24/1-2, pp.23-50.

TARDÓS, J.D.; NEIRA, J.; NEWMAN, P.M.; LEONARD, J.J. (2002), *Robust mapping and localization in indoor environments using sonar data.* Int. J. of Robotics Research 21/4, pp.311-330.

TENA-RUIZ, I.; RAUCOURT, S.; PETILLOT, Y.; LANE, D. (2004), *Concurrent mapping and localization using side-scan sonar.* IEEE J. Oceanic Engineering 29/2, pp.442-456.

WALTER, M.; HOVER, F.; LEONARD, J. (2008), *SLAM for ship hull inspection using exactly sparse extended information filters*. Int. Conf. Robotics and Automation, Pasadena, pp. 1463-1470.

WHITCOMB, L.L.; YOERGER, D.R.; SINGH, H. (1999), *Combined Doppler/LBL based navigation of underwater vehicles*. Int. Symp.Unmanned Untethered Submersible Technology.

WILLIAMS, S.; MAHON, I. (2004), *Simultaneous localisation and mapping on the Great Barrier Reef.* IEEE Int. Conf. Robotics and Automation 2, New Orleans, pp.1771-1776