Real Time Visual SLAM for underwater robotic inspection

Magda Meireles, Renato Lourenço, André Dias, José Miguel Almeida, Hugo Silva, Alfredo Martins

INESC TEC Institute for Systems and Computer Engineering of Porto
ISEP - School of Engineering, Porto Polytechnic Institute, Porto, Portugal
Email: magda.meireles, renato.lourenco @lsa.isep.ipp.pt
andre.dias, jose.m.almeida, hugo.m.silva, aom @inesctec.pt

Abstract—**This paper addresses the development of an underwater visual navigation system for a Remotely Operated Vehicle (ROV) based on Real-Time Simultaneous Localization and Mapping method using natural landmarks. Our proposed approach was tested in an indoor tank, where field experiments were performed to obtain 3D vehicle (VIDEORAY Pro3 ROV) trajectory, and results validated using an external stereo vision "ground-truth" system.**

Index Terms—**Underwater navigation system, SLAM, stereo vision, ROV.**

I. Introduction

In recent years, mobile robotics community has been putting a significant amount of effort into the development of environment based navigation systems. One of the most utilized technique for mobile robots navigation is denoted as Simultaneous Localization and Mapping (SLAM). The SLAM method consists on the robot ability to incrementally build a consist map of its surrounding environment, while simultaneously being able to determine its location within the map [1]. Although SLAM has been already subject of extensive research for ground mobile robotics applications [2] its use on underwater robotics has been up until now somewhat limited, specially if one considers real-time applications.

The underwater environment is a challenging scenario for SLAM, due to its limited sensorial possibilities, since Global Positioning Systems (GPS) and odometry sensors commonly used in SLAM approaches are unavailable. Even for visual sensors, the lack of image texture, environment being inherently noisier combined with blur makes the track of reliable features extremely difficult.

The development of visual navigation methods for underwater robots is essential for using this type of vehicles in hazard applications that until recently had been performed by humans or not even conducted at all. Underwater vehicles are currently being used in critical water infrastructures inspection tasks e.g. dams, harbors, oil-gas offshore platforms [3], and SLAM methods such as the ones described in [4], [5], [6] are being used to perform navigation of underwater mobile robots, in partially structured environments based on feature information extracted from acoustic data produced by imaging sonars. In [4] a mixed solution combining visual and sonar information is used to compute SLAM in water man-made structures. First, imaging sonar information is used to obtain the location of man-made vertical planar structures present in the environment. Afterwards, a voting algorithm based on the

Hough transform [7], uses line features information combined with uncertainty extracted from the sonar flow, and integrates the obtained information in a SLAM Extended Kalman filter (EKF) framework. Other work presents an approach of the unstructured underwater environment [8], where contextual information is obtained by segmenting the image into background and regions of interest. Features extracted from points of interest are then computed by using common robust feature descriptors as SIFT [9] and SURF [10].

Usually underwater SLAM applications are conducted offline, but nowadays novel marine robotics applications require the withdraw of the human-in-the-loop urging the need for real-time algorithms to be put in place. MonoSLAM [11] is a real-time algorithm which can recover the 3D trajectory of a monocular camera, moving rapidly through a previously unknown scene. This implementation was the first implementation of SLAM robotics method into the vision domain of a single uncontrolled camera. Other real-time Visual SLAM algorithms [12], overcome some of the specific challenges associated with underwater visual SLAM e.g. limited field-of-view (FOV) imagery, feature-poor regions by using an online bag-of-words measure for intra and inter-image saliency that proved to be useful for image key-frame selection, information-gain based link hypotheses and novelty detection. Others, discusses the possibility of building in real-time a mosaic of the seafloor relying on a SLAM framework [13]. The goal is to provide an unmanned underwater vehicle with a relatively rough visual map of the seafloor that supports basic navigation and context awareness. To solve the problem of autonomous ship hull inspection a vision based simultaneous localization and mapping is used [14]. The method denoted as Pose-graph SLAM, uses a combination of SIFT descriptors and Harris corners [15] within a pairwise image registration framework to provide camera-derived relative-pose constraints.

To test the use of environment features in real-time visual slam underwater robotic inspection operations, we developed a modified implementation of the RT-SLAM approach [16]. The method is a novel openSLAM framework, that copes with already existing EKF-SLAM [17] or FASTSLAM [18] implementations within a ROS [19] framework. The framework allows to estimate an underwater remote operated vehicle trajectory (ROV) in real-time (at 30 fps). The proposed underwater visual SLAM system uses a monocular camera setup (VideoRayROV), and is characterized in a laboratory test tank
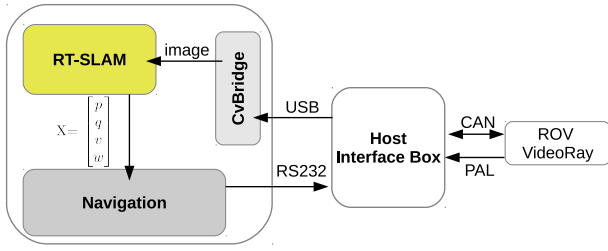
Fig. 1. Implemented architecture for underwater visual SLAM

setup with an external stereo visual ground truth system [20].

This article is organized as follows: Section II describes the underwater navigation system, the RT-SLAM algorithm and the system components. In section III, we present our external ground-truth vision system for underwater applications. The experimental results obtained in a controlled field environment are illustrated in Section IV. Finally, the conclusions and further work are given in Section V.

II. UNDERWATER NAVIGATION SYSTEM

In this section we detail the method applied to perform underwater visual SLAM as well as the hardware and software implementation issues. The underwater navigation implemented technique is based on monocular visual SLAM denoted by RT-SLAM [16]. The proposed architecture is detailed in Fig.1 and in principle is designed to use a monocular camera scheme, although the architecture can be extended to support a stereo camera setup.

A. RT-SLAM Algorithm

There are several SLAM techniques that could be applied to the estimation of the position and attitude. The most important ones are the EKF-SLAM [17] and the FastSLAM [18]. The FastSLAM is based on a recursive Monte Carlo sampling where the motion and observation model are modelled by a non-Gaussian pose distribution. Went compared with the EKF-SLAM techniques this is viewed as an improvement, due to the linearisation process of non-linear models, although the EKF has the advantage of faster processing and the ability to combined sub-mapping techniques to reduce the computational complexity. Therefore, based on this issues, we propose to evaluate the performance of the RT-SLAM implementation as a solution to support the navigation layer by providing the estimation of the position and attitude of the ROV while performing underwater navigation in structure environments. The RT-SLAM[16] is a standard formulation of EKF-SLAM, based on visual features, such as Harris Corner[15], with the ability to runs up to 60 Hz with 640x480 images. The internal architecture of the RT-SLAM is composed by the following layers: maps, robots, sensors, landmarks and observations. This internal organization provides the required structure to ensure a flexible, generic and efficient development tool to add new functionalities such as new methods of extracting landmarks and new sensors more feasible in underwater scenarios, such as an echo sounder. The RT-SLAM kinematic model to

do the filter prediction is based on a constant velocity model as expressed by:

$$
X = \begin{bmatrix} p \\ q \\ v \\ w \end{bmatrix} \tag{1}
$$

where p and q are respectively the position and orientation, and v and w the linear and angular velocities. This model does not require additional hardware setup and is designed for a monocular approach which means a limitation regarding scale factor observation.

The SLAM algorithm integrates an EKF filter to data associate for features correspondence. This is a C++ implementation, and it runs up to 60Hz with 640x480 images [16]. The SLAM architecture consists of five main objects, maps, robots, sensors, landmarks and observations. Robots have several sensors, providing observations of landmarks. States of robots, sensors and landmarks are stored in the stochastic map. The prediction model of underwater robot is represented by a simple kinematic model (constant velocity, constant acceleration), as defined by:

B. System Components

The implemented architecture as been performed with a VideoRay PRO 3E ROV, in Fig.2, with 3-DOF - surge, heave and yaw. This vehicle is designed for underwater exploration at depths up to 152 meters. The vehicle has three control thrusters, two for horizontal movements and one for vertical movement. Additionally, the vehicle is equipped with a two cameras without overlap view, one forward and one rear-facing, a pressure sensor, compass and altimeter.
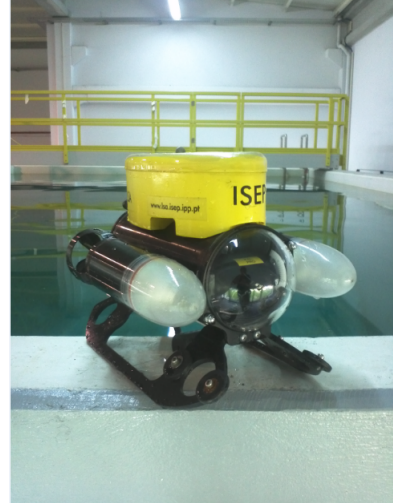

Fig. 2. VideoRay PRO 3E ROV.

The ROV is connected to a host interface box through an umbilical cable, which carries the ROV's power supply, the PAL video signal and the data between the vehicle and the host station over a CAN interface. The host interface box
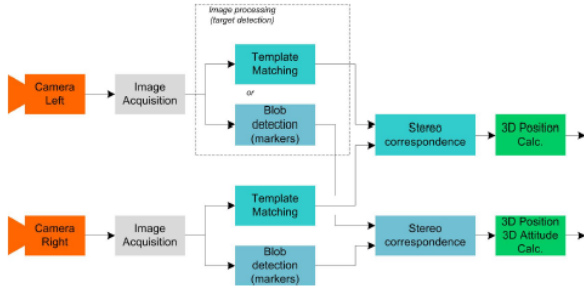
Fig. 3. Groundtruth system architecture [20].

is plugged to 240 VAC and is also responsible for enabling the interaction with the ROV through serial communication (RS232). This interaction is carried out by a PC, that also receives the video signal after being digitized by a frame grabber (Pinnacle DVC100) to a resolution of 640x480.

C. Implementation

The RT-SLAM algorithm as been integrated under Robot Operating System (ROS)[19] framework. This open-source framework provides a structured communications layer based on topics and was designed to be as distributed and modular. The modularity of the framework allow us to integrate and developed nodes, such as the RT-SLAM, in a straightforward manner. Therefore, based on this modularity we have a node responsible for the acquisition of the image through the frame grabber and publisher as a topic message the image that is received by the ROS RT-SLAM node.

III. GROUND-TRUTH SYSTEM

As an exogenous system to validate the RT-SLAM implementation in underwater scenarios, we used the INESC TEC/ISEP indoor ground-truth system [22] and extended to underwater scenarios in [20]. The ground-truth system is an external validation stereo based visual perception system for robot localization in underwater trials. The referred system is installed in the INESC TEC/ ISEP underwater robotic test tank (5m depth, 10x6m).

The ground-truth architecture, see Fig.3, consists of a time triggered stereo camera setup for image acquisition positioned at a very small depth mounted downwards. The cameras were placed at the corner of the tank, two Basler acA1300-30gc color Gigabit Ethernet (GigE) digital cameras with external trigger, with maximum resolution of 1278x958 and up to 30 fps of framerate. The ground-truth system setup for image processing also consists on recording computer and a set of predefined markers attached in to the underwater vehicle in order to be able to estimate its position.

The 3D position estimation is performed based on the marker structure (one sphere for z axis determination and 4 point markers for position calculation). Marker monocular tracking is either performed based on color segmentation or with template matching.

IV. EXPERIMENTAL RESULTS

This section reports the reliability of the presented algorithm, we prove this using our test tank to compare with our ground-truth approach described in section III.

The VideoRay trajectory was performed by remote control to obtain experimental data from monocular camera. The environment features for operational conditions are given by underwater structures in a controlled field environment. During trajectory, the vehicle gathered a measurements from the frontal camera. These are used to extract visual features that allow obtain vehicle localization and map building. The visual landmarks detected by camera data are presented in Fig.4.

Results for ROV trajectory using our SLAM approach are presented in Fig.5, all the software implementation was performed using MATLAB in a Intel I5 Dual Core 2.5GHz processor. In Fig.6, we can observe the ROV heading, the obtained roll, pitch and yaw angles for the performed trajectory.

V. CONCLUSIONS AND FUTURE WORK

An algorithm to perform real time visual SLAM for underwater environments has been presented. This system is characterized in a laboratory test tank setup with an external stereo visual ground truth system. The navigation system for underwater robotic inspection produced good and accurate results with a VideoRay Pro3 ROV up to 30 fps. ROV trajectory is obtained using environment features of underwater structures.The data obtained for roll, yaw and pith are consistent with the path described by ROV and corroborated by ground-truth system.

Concerning SLAM applications there are several expectations for future work such as detect other kind of landmarks, since RT-SLAM is modular and can easily be adapted to integrate additional types of features like horizontal and vertical lines. One other alternative that can be used in our RT-SLAM implementation is the introduction of inertial navigation systems to improve the underwater vehicle pose estimation. The fusion of inertial and visual SLAM data can provide estimates for obtaining the unknown scale parameter.

REFERENCES

[1] H. Durrant-Whyte, "T.Bailey,Simultaneous Localisation and Mapping (SLAM):Part I The Essential Algorithms", Australian Centre for Field Robotics (ACFR) J04, The University of Sydney, Sydney NSW 2006, Australia.

[2] Francisco Bonin-Font, Alberto Ortiz, and Gabriel Oliver. 2008. Visual Navigation for Mobile Robots: A Survey. J. Intell. Robotics Syst. 53, 3 (November 2008), 263-296. DOI=10.1007/s10846-008-9235-4

[3] Lazinica, Aleksandar, "Mobile Robots: Towards New Applications", Germany, Pro Literatur Verlag, edition 2006, 309-315
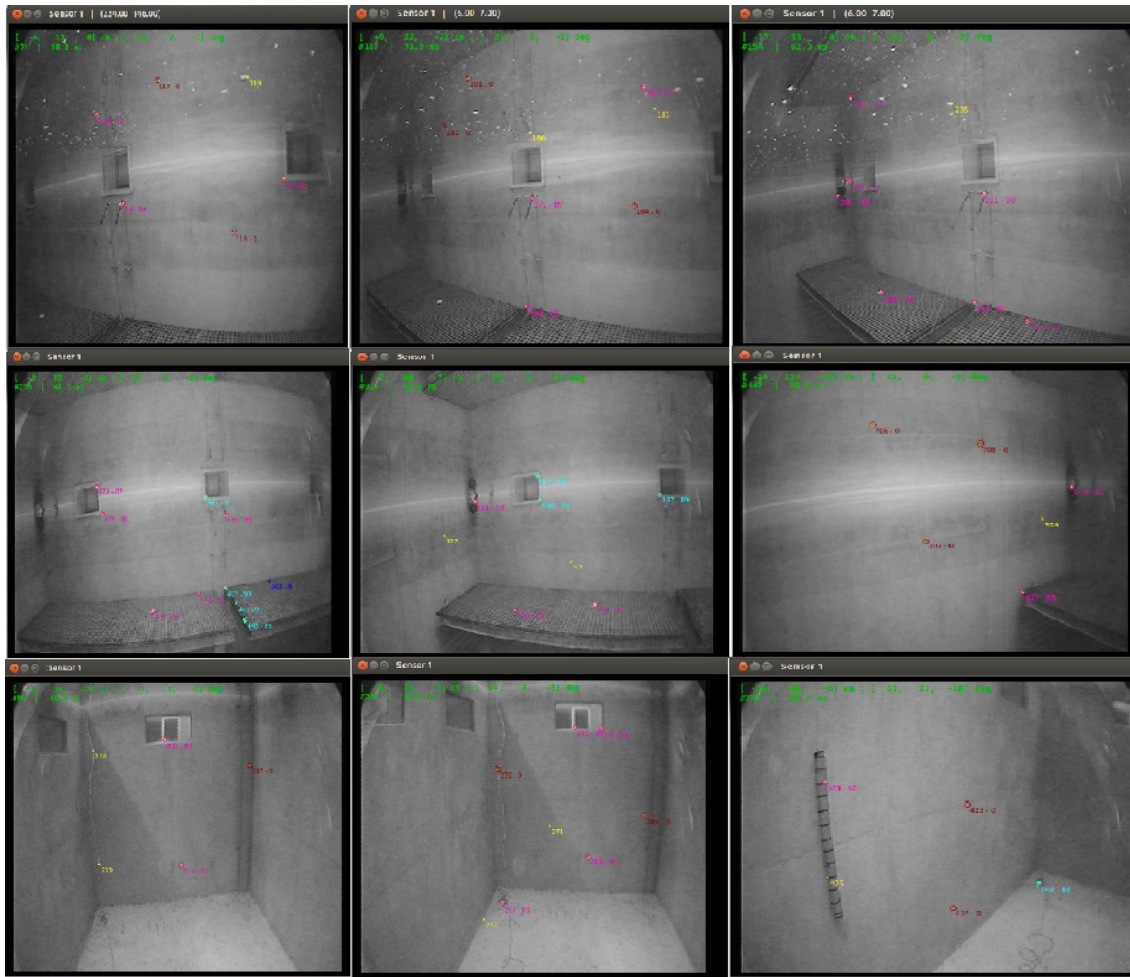
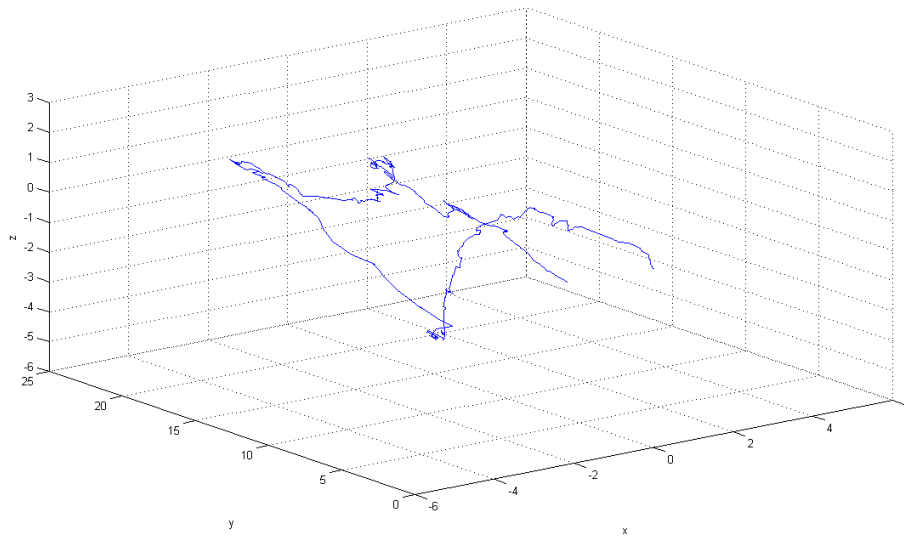Fig. 4.   Key Point detection (Harris corners) for underwater RT- SLAM

Fig. 5.   Video-Ray performed Trajectory

[4] Ribas, D., Ridao, P., Tards, J. D. and Neira, J. (2008), Underwater SLAM in man-made structured environments. J. Field Robotics, 25: pages 898921.

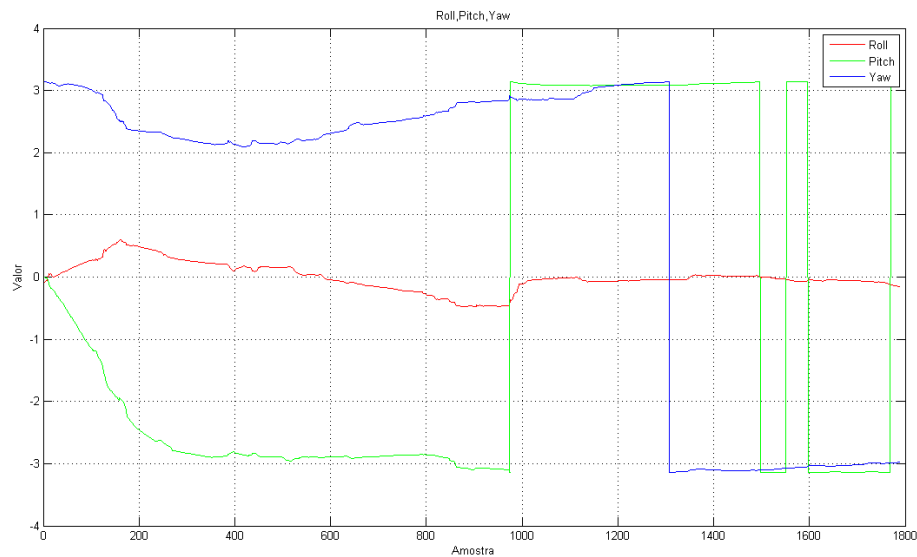[5] D.Ribas,P.Ridao, J.D. Tardos and J. Neira, "Underater SLAM in a

Fig. 6. Roll, pitch an yaw of Video-Ray

Marina Environment" IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007, San diego, USA.

[6] David Ribas, Pere Ridao, José Neira, Juan D. Tards: SLAM using an Imaging Sonar for Partially Structured Underwater Environments.IROS 2006: 5040-5045

[7] Duda, R. O. and P. E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," Comm. ACM, Vol. 15, pp. 11-15 (January, 1972).

[8] Aulinas, J.Carreras, M., Llado, X., Salvi, J., Garcia, R.,Prados, R., Petillot, Y. "Feature extraction for underwater visual SLAM." IEE OEANS 2011, Spain.

[9] David G. Lowe. 2004. Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vision 60, 2 (November 2004).

[10] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. 2008. Speeded-Up Robust Features (SURF). Comput. Vis. Image Underst. 110, 3 (June 2008).

[11] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM:Real-time single camera SLAM,"IEEE Transactions on Pattern Analysis and Machine Intelligence", vol. 29, no. 6, pp. 1052 1067, 2007

[12] Kim.A, Eustice "R.M Real-Time Visual SLAM for Autonomous Underwater Hull Inspection using Visual Saliency" IEEE Transactions on Robotics 2013.

[13] Fausto Ferreira, Gianmarco Veruggio, Massimo Caccia, and Gabriele Bruzzone. 2012. Real-time optical SLAM-based mosaicking for unmanned underwater vehicles Journal of Intelligent Service Robots.

[14] Kim, A.; Eustice, R. (2009). "Pose-graph Visual SLAM with Geometric Model Selection for Autonomous Underwater Ship Hull Inspection." IEEE/RSJ International Conference on Intelligent Robots and Systems: 1559-1565

[15] C. Harris and M. Stephens (1988). "A combined corner and edge detector". Proceedings of the 4th Alvey Vision Conference. pp. 147-151.

[16] C. Roussillon, A. Gonzalez, J. Sol, J. Codol, N. Mansard, S. Lacroix, and M. Devy, "RT-SLAM: A Generic and Real-Time Visual SLAM Implementation". ICVS 2011, volume 6962 of Lecture Notes in Computer Science, page 31-40. Springer.

[17] Tim Bailey, Juan I. Nieto, JosÃ© E. Guivant, Michael Stevens, and Eduardo Mario Nebot. IROS, page 3562-3568. IEEE, (2006)

[18] Montemerlo, M.; Thrun, S.; Koller, D.; Wegbreit, B. (2002) "FastSLAM: A factored solution to the simultaneous localization and mapping problem".Proceedings of the AAAI National Conference on Artificial Intelligence. pp.593?598.

[19] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source Robot Operating System," in ICRA Workshop on Open Source Software, 2009.

[20] A. Martins, A. Dias, H. Silva, J. Almeida, P. Goncalves, F. Lopes, A.Faria, J. Ribeiro and E. SIlva, "GroundTruth System for Underwater Benchmarking", IEEE OCEAN's 2013. San Diego, September 2013

[21] Costa, M.J. , Goncalves, P. , Martins, A., Silva, E. ,"Vision-Based Assisted Teleoperation for Inspection Tasks with a Small ROV", OCEANS 2012 MTS/IEEE HAMPTON ROADS conference.

[22] A. Dias, J. Almeida, A. Martins and E. Silva, "Real-Time Visual Ground-Truth System for Indoor Applications", IbPRIA 2013: