Advances in the Specification and Execution of Underwater Autonomous Manipulation Tasks

Mario Prats*, Juan Carlos Garcia*, Jose Javier Fernandez*, Raul Marin* and Pedro J. Sanz*
*Computer Science and Engineering Department
Jaume-I University, Castellón, Spain
Email: {mprats, garciaju, fernandj, rmarin, sanzp}@icc.uji.es. See http://www.irs.uji.es

***Abstract*—In this paper we show how techniques that have been applied during the last years for autonomous manipulation in the air can be successfully adapted to a new challenging scenario: the underwater environment. Specifically, the developed techniques include among others: visual tracking of objects in the seabed, vision-based arm control and sensor-based grasping. Furthermore, a graphical user interface for specifying underwater manipulation actions and providing the corresponding task specification to the execution modules is also presented. This research os enclosed in the framework of our two ongoing projects in autonomous underwater intervention: RAUVI and TRIDENT.**

I. Introduction

The need for manipulation in underwater environments is significantly increasing in the last years. A large number of applications in marine environments need intervention capabilities in order to undertake specific tasks. Some examples are panel intervention (e.g. plugging a connector), search and recovery of objects, collecting marine samples from the seabed, etc.

At present, these problems are partially solved with Remote Operated Vehicles (ROVs) that are launched from oceanographic vessels, and remotely operated by expert pilots through an umbilical communications cable and complex control interfaces. These solutions present several drawbacks. Firstly, ROVs are normally large and heavy vehicles that need significant logistics for its transportation and handling. Secondly, the complex user interfaces and control methods require expert pilots for its use. These two facts significantly increase the cost of the applications. In addition, the need of an umbilical cable introduces additional problems of control and puts limits to the workspace. Another limitation is the fatigue and high stress that users of remotely operated systems normally suffer, mainly due to the master-slave control architectures used [1].

All of these points justify the need of more autonomous, cheap and easier-to-use solutions for underwater manipulation. Our IRSLab, that participates in the Spanish National project RAUVI and the European project FP7-TRIDENT, pursue these objectives, among others. In the context of these two projects, the main goal is to reach new levels of autonomy in underwater survey and intervention tasks, and to build a new lightweight intervention vehicle that can be launched from small vessels and easily programmed through a friendly user interface. These are Autonomous Underwater Vehicles for Intervention (I-AUVs), which represent a new concept of undersea robots that are not tethered to a mother ship. In fact, the history about I-AUVs is very recent, and only a few laboratories are currently trying to develop this kind of systems [2].

One of the most well-known research projects devoted to develop an I-AUV is SAUVIM [3] [4]. Along its life, this project has implemented a Graphical User Interface (GUI) combining all kind of sensor data inside a common simulation environment. Their GUI uses its own programming language and allows a high level of interaction between the user and the underwater robot in text mode. In addition, virtual reality (VR) is available within the GUI, thus showing the evolution of the complete system along the intervention mission, and assisting the user in the high-level control.

Our research group is working on this kind of underwater intervention systems in general and more concretely in (i) specific interfaces that allow an intuitive use by non-expert users, and (ii) the sensor-based control systems for performing autonomous manipulation from the previous specification. We design a two steps strategy [5], guaranteeing the ”intelligence“ in the system performance thanks to the user's inclusion in the control loop when strictly necessary, but not in a continuous way like in ROVs. This strategy is illustrated in Figure 1. In a first step, our I-AUV navigates through the underwater Region of Interest (RoI) and collects data under the control of their own internal computer system. After ending this first step, the I-AUV returns to the surface (or to an underwater docking station) where its data can be retrieved. A seafloor photo-mosaic is built, and by using a special GUI a non-expert user is able to identify the target and to select the suitable intervention task to carry out during the second step. During this second step, the I-AUV navigates again to the RoI and performs the target localization and the intervention mission autonomously.

The following sections overview the two main modules related with the specification and the execution of manipulation in our projects. These are the Graphical User Interface (Section II), and the intervention execution module (Section III), where our methods for calibration, tracking and vision-based control will be outlined. Section IV will present preliminary experiments on autonomous hooking with the real I-AUV, whereas Sections V will conclude this paper.

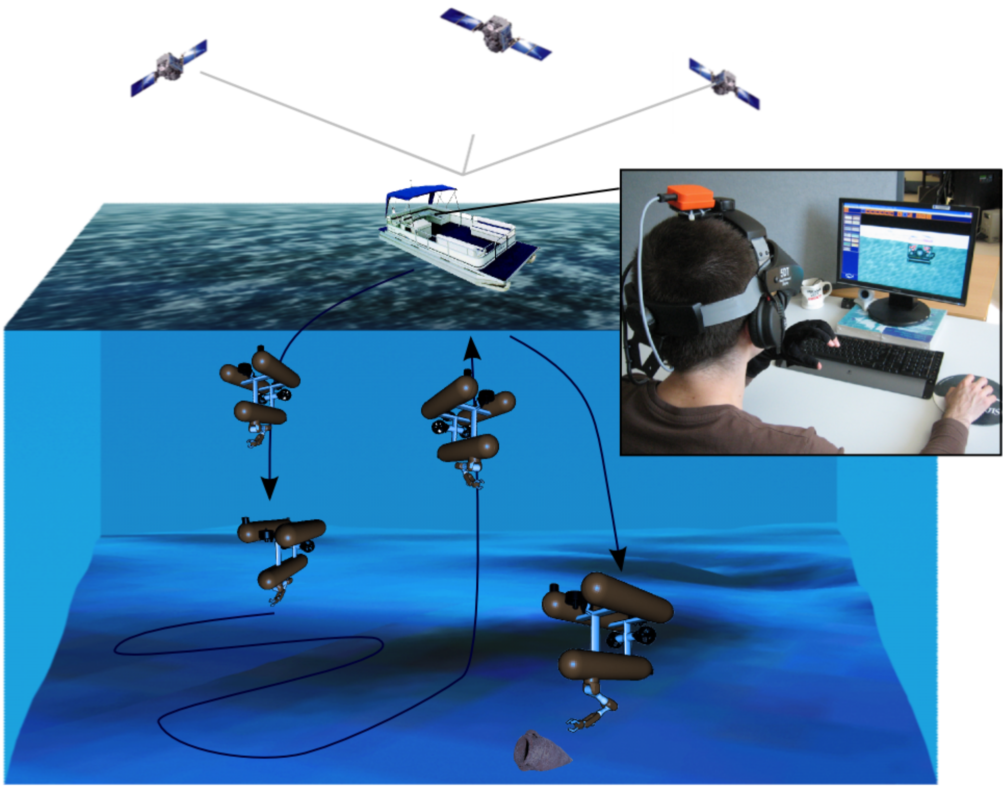


Fig. 1. An illustration of our adopted two-steps strategy for performing autonomous intervention in a shallow water context. A special user interface is used off-line only to select the target and the intervention that will be autonomously performed in the next step.

II. The User Interface: The Specification of the Intervention Mission

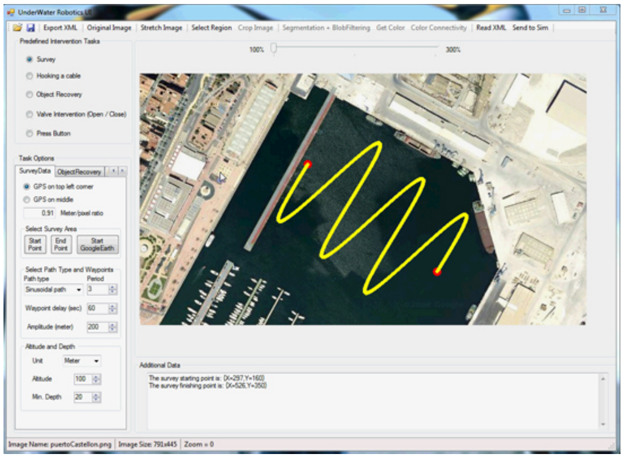
As mentioned previously, we adopt a two-stage strategy: during the first stage, the I-AUV is programmed at the surface and receives a plan for surveying a given RoI. During the survey it collects data from cameras and other sensors. At the end of this first stage, the I-AUV returns to the docking station where the data is retrieved and an image mosaic of the seabed is reconstructed [6]. The Target of Interest (ToI) is then identified on the mosaic and the intervention action is specified by means of a user interface described in this section.

The Graphical User Interface (GUI) is used to specify both the survey path and the intervention task. The former is done by loading a geo-referenced map of the area and indicating a set of waypoints (possibly using predefined sinusoidal or grid-shaped trajectories). The waypoints are sent to the vehicle control system that guides the robot through them. Figure 2a shows an example of a sinusoidal trajectory superposed on a satellite view of a harbour. Other types of survey trajectories are possible, as for example grid-shaped. After finishing the survey, and once the photo-mosaic has been built, the user first looks for the target of interest on it. After selecting the target, the intervention task is indicated by choosing between different pre-programmed actions such as grasping, hooking, pressing a button, etc.

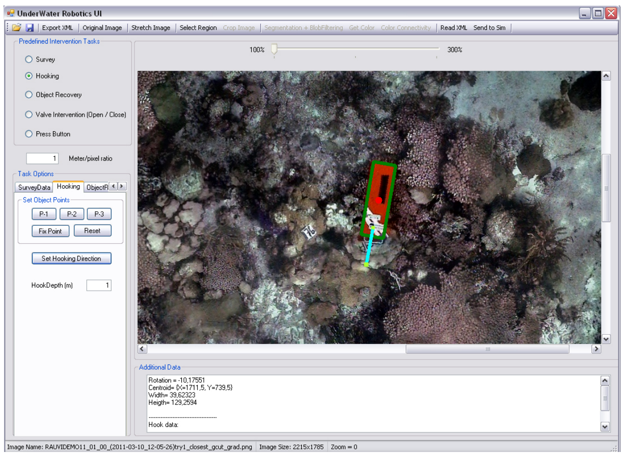
The user interface contains built-in image processing and grasp planning algorithms that automate the task specification process when possible. If automatic methods fail, the user can always specify the task parameters manually. For the experiments described in this work we consider a hooking task, which we define as enclosing the target of interest in a bounding box, and selecting the point and the direction where to attach the hook, as shown in Figure 2b.

When the specification is finished, an XML file containing the task parameters is generated. For the hooking task, this file includes:

- The image used for the specification. We assume this image is geo-referenced so that it is possible to relate pixel coordinates to meters with respect to a global frame.
- The ToI bounding box origin with respect to the image origin, and its orientation with respect to the horizontal.
- The width and height of the bounding box, both in pixels and in metric units, due to the fact that the image is geo-referenced and the camera intrinsic parameters are known, thus allowing to compute 3D dimensions from single frames.
- A hook point and direction given in pixel coordinates with respect to the bounding box origin, and also in metric units.



(a) Survey. A sinusoidal survey trajectory is specified on a satellite view of a harbour.



(b) Intervention. The object is found in the photo-mosaic and enclosed in a bounding box. Then, the task parameters are set.

Fig. 2. Mission specification in the GUI.

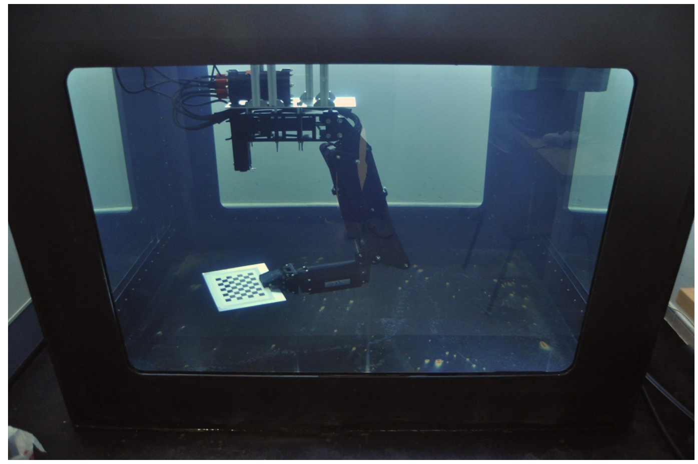


Fig. 3. In order to succeed guiding grasping actions in an autonomous way, external eye-hand calibration must be applied before

With the bounding box information, a template containing only the ToI is created and later used for object detection and tracking (see Sections III and IV). For more details on the Graphical User Interface, please refer to [7].

III. The Intervention Execution: Object Recovery

After the intervention mission has been specified, it has to be autonomously performed by the manipulator control system. In this section we describe the different techniques that we are applying to solve this problem, and some experimental results with different arms, in both air and water tank conditions.

A. Experimental Setup

We are working with the CSIP Light-weight ARM5E, which is a commercial underwater robotic manipulator that has been specially adapted for our projects. It is electrically actuated by 24V brushless DC motors. It is composed of four revolute joints, and can reach distances up to one meter. An actuated robot gripper allows for grasping small objects, and its T-shaped grooves also permit handling special tools. The arm is made of aluminium alloy partially covered with foam material in order to guarantee suitable buoyancy. The total weight in the air is about 29 kg, whereas in fresh water it decreases to 12 kg approximately. The arm is capable of lifting 12 kg at full reach, and can descend up to 300 m in water.

An underwater camera can be mounted either on the arm wrist or on the base link in order to provide a top view of the manipulation area. It is a Bowtech DIVECAM-550C-AL high-resolution color CCD camera, rated up to 100 m depth. The current configuration of the arm and camera is shown in Figure 3. The most suitable area for manipulation is around 80 cm below the arm base link. This area guarantees the highest distance to the workspace limits and is also free of arm singularities. At this moment the camera is placed next to the arm base link and faced downwards. This configuration guarantees that there is an intersection between the camera field of view and the arm workspace that allow to visually control the arm during execution of the task. In addition, this view is similar to the one obtained during the survey and used in the GUI for task specification, thus simplifying the matching process.

B. Eye-Hand calibration

In order to transform 3D coordinates of objects detected by the camera into the manipulator reference system, the camera must be externally calibrated with respect to the manipulator base link (known as eye-hand calibration). We do this by attaching a checkerboard pattern (whose model is known) to the gripper in such a way that the pattern origin is in a known pose with respect to the arm end-effector. This allows computing the whole kinematic chain from the arm base link to the pattern origin.

Then, the pattern is moved to different poses in the camera field of view, and an image is captured for each of them. Next, the intrinsic and extrinsic calibration parameters are computed following standard calibration techniques (using the Matlab Calibration Toolbox in this case). With the extrinsic parameters for each view, and the corresponding base-pattern

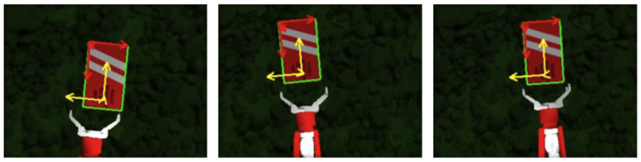


Fig. 4. Visual tracking and servoing experiments in simulation

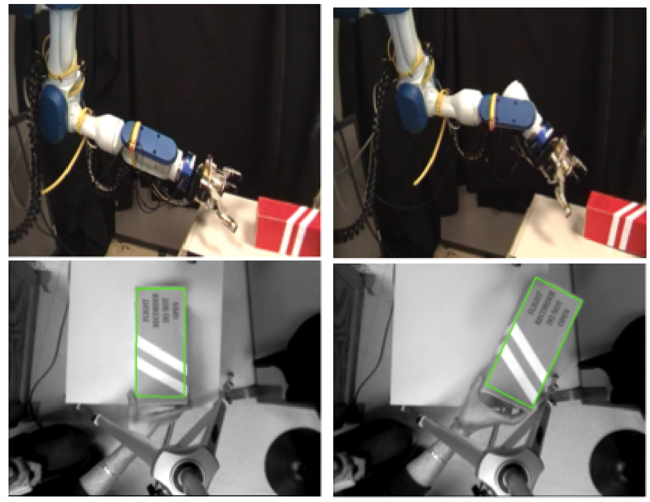


Fig. 5. In-the-air experiments on tracking and visual servoing for grasping. The complete sequence can be accessed at http://goo.gl/fPJle

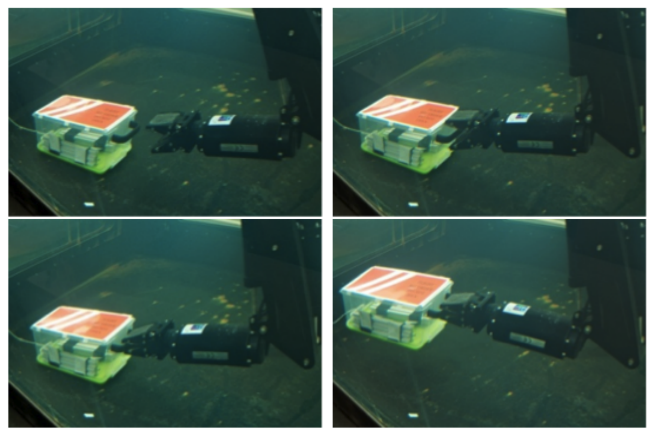


Fig. 6. Water tank experiments on autonomous grasping using visual servoing techniques

transformation described previously, the transformation between the camera and the arm base frame can be estimated. This procedure has to be performed only once, and the calibration will remain valid as long as the relative arm-camera position is not changed.

C. Tracking and Visual Servoing for grasping

As it cannot be assumed that the vehicle can accurately maintain its position on top of the object to manipulate (due to underwater currents and other dynamic effects), it is necessary to continuously track the object and visually guide the arm motion accordingly.

For this, we have implemented a template tracking method [8] [9]. During the task specification on the seabed mosaic, after the target has been selected, its bounding box is used for extracting a template of the object. This template is matched at each iteration on the video stream during the intervention phase. A homography between the template and its match in the current image is computed and used for transforming the grasp points to its actual pose. From the tracked grasp points, given in the camera coordinates, a position-based control approach is executed in order to guide the robot hand towards the grasp position. The controlled hand frame is moved in a straight line towards the grasp points, taking into account the kinematic redundancy of the robot. For performing this action, the eye-hand calibration described in a previous section is used.

Figures 4, 5 and 6 show, respectively, visual servoing experiments performed in simulation, in the air with a 7 DOF PA10 Arm, and in a water tank with the 4 DOF Light-weight ARM5E.

IV. Experiments with the real I-AUV

The first integration experiments have been carried out at the CIRS water tank at the University of Girona. In a first integration effort, the arm was mechanically attached to the GIRONA 500 AUV, the buoyancy of both subsystems was suitably adapted, and the first tele-operation results were obtained in the task of recovering an amphora, as shown in Figure 7.

In a second integration effort, the autonomous capabilities of the system were demonstrated in the task of searching and recovering a Flight Data Recorder from the floor of the water tank (see Figure 8)

During the intervention stage, after finding the target of interest, the vehicle kept its position and attitude with visual feedback from the tracked target. Vision-based station keeping was performed with two degrees of freedom: the horizontal motion of the vehicle was controlled in order to keep the tracked template origin close to a desired position in the current view. Vertical motion was controlled with the altimeter feedback in order to keep a suitable distance to the floor of around one meter, measured from the base of the arm. While keeping station, the arm was able to autonomously retrieve the object in different trials.

V. Conclusion and future work

In this paper, our current research on autonomous underwater manipulation within the RAUVI and FP7-TRIDENT projects has been presented. For the specification of manipulation actions we have described a Graphical User Interface that loads seabed images and provides automatic methods for object detection and task specification. The main motivation behind this research line is to improve the complex and tedious tele-operation units currently used for piloting a ROV.

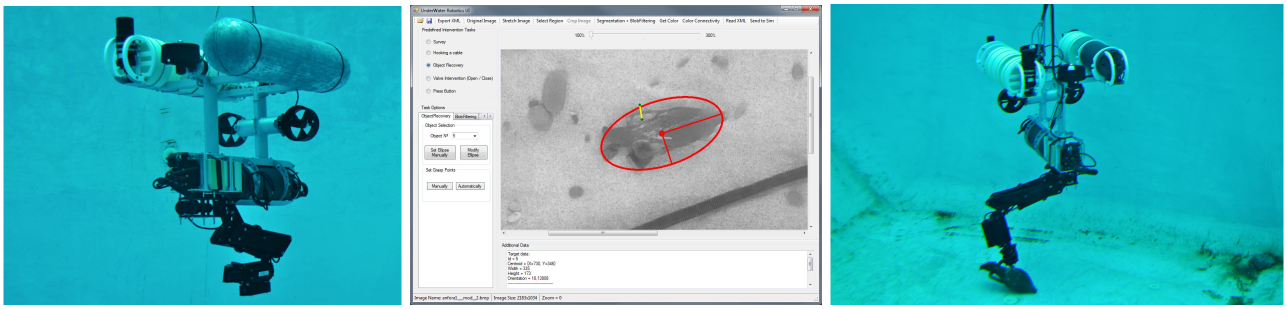


Fig. 7. Firts experiments, in water tank conditions, after integration of the arm system in the final I-AUV prototype. (a) The I-AUV prototype, integrating the GIRONA 500 AUV and the hand-arm system; (b) GUI displaying the target characterization for grasping; (c) The grasping execution of an amphora in tele-operated way

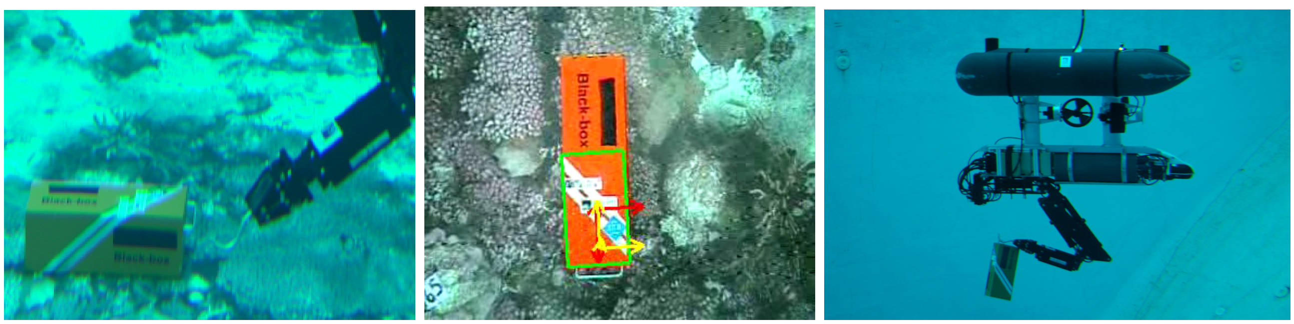


Fig. 8. Latest experiments on autonomous hooking. (a) The arm is approaching to the target while the vehicle performs station keeping; (b) the associated video sequence displaying the visual tracking algorithm in execution; (c) After retrieval of the target, the I-AUV returns to the surface

Template-based techniques for searching and visually detecting the target, and monocular vision-based pose estimation methods have been outlined. These allow to compute the relative pose of the object with respect to the manipulator, and, therefore, controlling the arm towards the contact points. Some of the experiments carried out in water tank conditions demonstrate the viability of these techniques. Concretely, we have first performed grasping actions with a fixed underwater manipulator in a water tank. Then, we have integrated the robot arm into an AUV and have performed autonomous experiments on hooking objects placed on the floor. For future work, further improvements in manipulation can be made by generating smooth velocity and acceleration trajectories, and by implementing error recovery actions when the manipulation action fails. We further plan to integrate the GUI into a 3D simulator package, and to apply augmented reality techniques in order to improve the interaction with the user and to assist with the specification and supervision of the intervention mission. Finally, these promising results encourage us to follow with the next step: a shallow water test of the system by the end of 2011.

Acknowledgment

This research was partly supported by the European Commission Seventh Framework Programme FP7/2007-2013 under grant agreement 248497 (TRIDENT Project), by Spanish Ministry of Research and Innovation DPI2008-06548-C03 (RAUVI Project), and by Fundació Caixa Castelló-Bancaixa P1-1B2009-50

References

[1] B. Boyle, R. McMaster, and J. Mixon, "Teleoperation of an underwater robotic repair system using graphical simulation," in *IEE Colloquium on Control of Remotely Operated Systems: Teleassistance and Telepresence*, May 1995, pp. 2/1–2/4.

[2] J. Yuh, "Design and control of autonomous underwater robots: A survey," *International Journal of Autonomous Robots*, vol. 8, pp. 7–24, 2000.

[3] J. Yuh, S. Choi, C. Ikehara, G. Kim, G. McMurty, M. Ghasemi-Nejhad, N. Sarkar, and K. Sugihara, "Design of a semi-autonomous underwater vehicle for intervention missions (sauvim)," in *Proceedings of the International Symposium on Underwater Technology*, Apr. 1998, pp. 63 –68.

[4] G. Marani and S. Choi, "Underwater target localization," *IEEE Robotics & Automation Magazine*, vol. 17, no. 1, p. 18, March 2010.

[5] P. Sanz, P. Ridao, G. Oliver, C. Melchiorri, G. Casalino, C. Silvestre, Y. Petillot, and A. Turetta, "TRIDENT: A framework for autonomous underwater intervention missions with dexterous manipulation capabilities," in *7th Symposium on Intelligent Autonomous Vehicles IAV-2010*. IFAC, 2010.

[6] T. Nicosevici, N. Gracias, S. Negahdaripour, and R. Garcia, "Efficient three-dimensional scene modeling and mosaicing," *J. Field Robot.*, vol. 26, pp. 759–788, October 2009.

[7] J. C. García, J. F. Fresneda, P. Sanz, and R. Marín, "Increasing autonomy within underwater intervention scenarios: The user interface approach," in *2010 IEEE International Systems Conference*, San Diego, CA, USA, 4 2010, pp. 71–75.

[8] M. Prats, J. C. García, J. J. Fernández, R. Marín, and P. J. Sanz, "Towards specification, planning and sensor-based control of autonomous underwater intervention," in *IFAC 2011*, Milano, Italy, Aug. 2011, to be published.

[9] S. Benhimane and E. Malis, "Real-time image-based tracking of planes using efficient second-order minimization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004, pp. 943–948.