Primitive shape fitting in point clouds for enabling autonomous underwater grasping

D. Fornas, A. Peñalver, M. Prats, J. J. Fernández and J. Sales
Department of Computer Science and Engineering, Jaume I University
Av. de Vicent Sos Baynat, s/n 12071 Castelló de la Plana, Spain
Email: {dfornas,penalvea,mprats,fernandj,salesj}@uji.es

***Abstract*—Autonomous grasping of unknown objects by a robot is a highly challenging skill that is receiving increasing attention in the last years, and is still more challenging in underwater environments, with highly unstructured scenarios, limited availability of sensors and adverse conditions that affect the robot perception and control systems. This paper describes an approach for increase the autonomy of grasping operations on underwater primitive shaped objects from floating vehicles, in particular cylinder shaped objects like an amphora. Various sources of stereo information are used to gather 3D information in order to obtain a model of the object. Using a RANSAC primitive shape recognition algorithm the model parameters are obtained and a set of feasible grasps are computed. Then the user specifies the best one or it is specified using analytical constraints. This approach is tested with the UWSim simulator.**

Keywords - autonomous grasping, underwater, point cloud, RANSAC, shape fitting.

I. INTRODUCTION

Exploration of the oceans is attracting the interest of many companies and institutions. Remote Operated Vehicles (ROVs) are currently the most used machines but the trend is to advance towards Autonomous Underwater Vehicles (AUVs). The approach holds lots of challenges, being the autonomous grasping and manipulation tasks one of the biggest. Within this context there exist only a few projects in autonomous manipulation. The earlier achievements happened at the beginning of this century, when the project AMADEUS [1] demonstrated the coordinated control of two 7 degrees of freedom (DOF) arms submerged in a water tank. Recently, the project TRIDENT [2] demonstrated the capability to autonomously survey an area and recover an object from the seafloor, still with some interaction with a human operator. In the context of this project, a framework to grasp objects interacting with the user only in the grasp planning phase was presented [3]. This approach, focused on increasing autonomy, is further developed here.

In this paper we present a method able to perform grasping tasks more autonomously in the constrained, yet realistic, problem of grasping cylindrical objects like an amphora or a pipe.

More generally, it allows to autonomously grasp unknown objects that resemble primitive shapes in the way that the amphora can be considered a cylinder or an airplane black box resembles a cuboid for the purpose of manipulation. Grasping objects generally requires at least some partial 3D structure that can be gathered using various methods, such as stereo vision or 3D laser reconstruction [3]. Here the cheapest method is mainly used, i.e. Stereo vision, but other sources of quality stereo vision are also valid. For example, stereo vision is not useful in turbid waters, on untextured floors or in the darkness while laser reconstruction performs better in darker scenarios. The obtained 3D point cloud is then used for planning a grasp that is executed fully autonomously by the robot, given the conditions of the problem.

This article is organized as follows: next section describes the virtual scenario setup; Section III briefly outlines the 3D point cloud acquisition and RANSAC shape fitting algorithm; Section IV describes the grasp specification with the analytical model of the object; Section V shows the current results and finally, further work and conclusions are included in Section VI.

II. EXPERIMENTAL SETUP

Although future real experiments are planned, a simulated environment has been considered to develop and test the algorithms. UWSim [4] has been chosen as the simulator for these experiments because it allows working with robots in underwater environments within ROS [5]. The considered scenario is illustrated in Figure 1.The mechatronics consists of virtual model of an underwater robotic arm attached to the Girona500 AUV [6]. The arm is has 7-DOF and was developed by GraalTech (named here GT-arm) in the context of

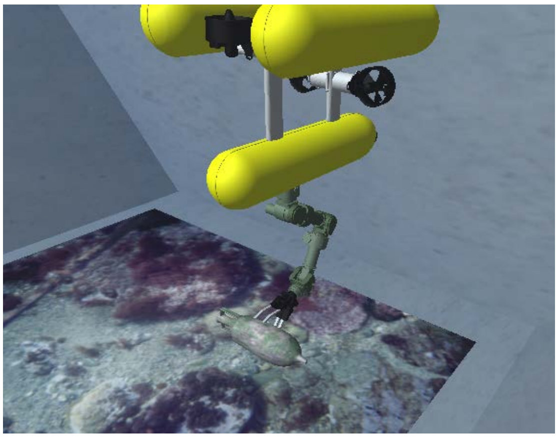


Fig. 1. Simulation environment that reproduces the CIRS pool with the GT arm attached to the Girona 500 AUV.

the TRIDENT project [2]. In this paper, the original dexterous hand (developed by University of Bologna) attached to the GT arm has been replaced by a jaw as end-effector and the vehicle is floating in the CIRS (Centre d'Investigació en Robótica Submarina, University of Girona) simulated pool. The end-effector is a jaw to allow a parallel grip. In this way, only an approach vector is needed. However, this solution has less flexibility than using a dexterous hand. The target object is an amphora lying on a textured floor.

The considered simulated vehicle-arm configuration is not physically available for experimental validation, but its features are optimal for testing purposes in simulation. Instead, the real experimental available setup that will be also considered in the future is based on a 5-DOF underwater robotic arm (in our case the CSIP Light-weight ARM 5E [7]) attached to a floating vehicle prototype that remains static (Figure 2). The vision system is provided by a Videre Stereo Camera placed near the base of the arm inside a sealed case and placed looking to the ground. The real target object is also a cylindrical object (a clay jar) lying on a planar surface surrounded with stones.

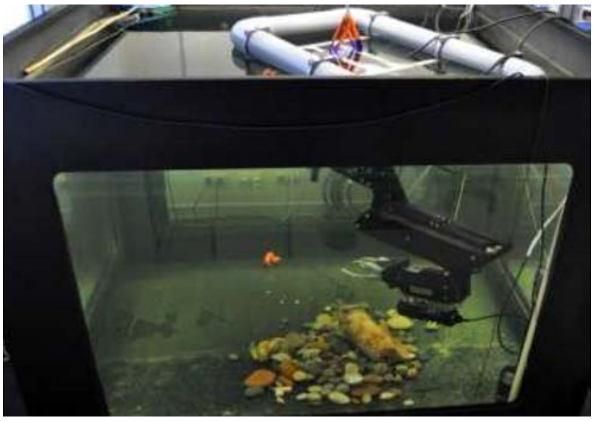


Fig. 2. Real water tank conditions used for the experiments. The ARM5E arm is attached to a floating platform that is placed in the water in a fixed position.

III. 3D RECONSTRUCTION AND SEGMENTATION

The very first step in order to grasp an object is obtain information about the environment. In this case it can be gathered either using laser stripe reconstruction or a stereo camera (real or virtual in UWSim). The algorithm has been tested with both sources but mainly stereo vision is used because is easier to perform in a simulated environment (and the resulting point clouds there are very similar). In the proposed experiments using UWSim, a virtual stereo camera is attached to the vehicle and captures a pair of images from which a 3D reconstruction with respect to a fixed frame is performed using standard ROS stereo image processing methods. A single point cloud is saved from the output of this algorithms and then it is processed using PCL [8]. This processing consists on downsampling the point cloud, in order to decrease the number of points using only the most relevant ones, and apply it an outlier filter, in order to remove the points that are potential wrong values caused by spurious particles or optical reconstruction errors. Decreasing the number of points decreases the computation time of the following steps and increases the robustness of the overall method.

With this relevant point cloud, a RANSAC algorithm, described in [9], is used twice to separate the object from the background. First, the background plane is detected with a RANSAC plane fitting algorithm and the resulting parameters are used to remove the plane inliers from the original point cloud. In the next step, other RANSAC algorithm is used to obtain the cylinder parameters associated to the object (that is supposed to be an amphora). These algorithms are parameterized to allow fitting quality and performance control. The result of this steps is a set of inliers that represent the detected amphora points and the analytical parameters: a point in the obtained model axis, the axis direction and the cylinder radius. This process has been tested separately from the final execution using clouds extracted from the real stereo camera (experimental setup), the virtual stereo camera (UWSim) and the laser stripe reconstruction technique with different degree of success. In general, stereo cameras perform better with good light conditions such as the experimental water tank or the simulation scene while laser reconstruction benefits from darker conditions where the contrast with the green ray is higher (like in deep sea operations or the water tank with lights off and windows closed). Different segmentations from virtual and real stereo cameras is shown in Figure 3.

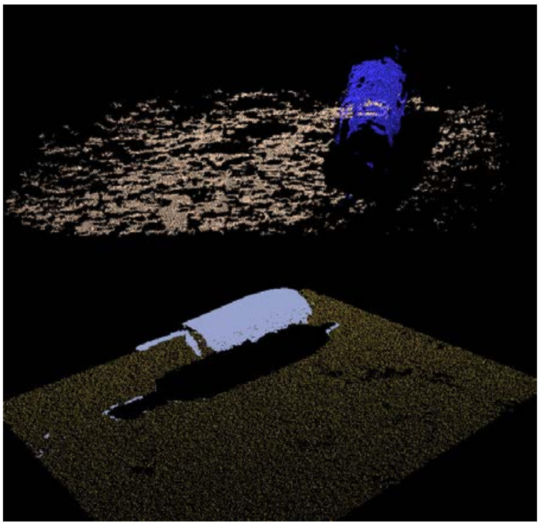


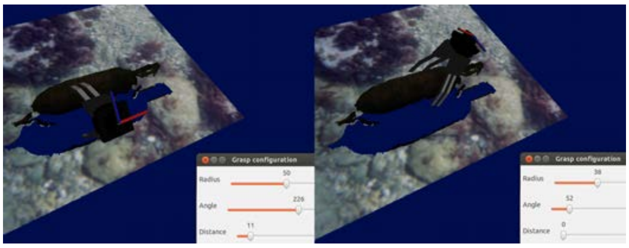
Fig. 3. Segmentation of the object from the background using real camera (top) and virtual camera (bottom) images.

IV. GRASP SPECIFICATION

Using the cylinder model obtained, a grasp can be specified. To avoid errors, the grasp point is computed using the most significant points of the cylinder inliers (the 90% of the points nearer to the centre). The middle point of the cylinder axis is used as a starting grasp point. Then taking into account the amphora radius and desired approach distance and angle, the grasping end-effector frame is moved away from the starting position, with this free variables allowing computing different grasp frames around the cylinder centre axis.

This freedom allows two different possibilities: use this variables to maximize grasp characteristics such as angle with the floor and stability or use them to allow the end-user to set up a grasp with an easy and quick interface. While the first approach is really appealing sometimes the loss of autonomy of the second becomes a huge robustness increase, as the user is who decides whether a grasp is good enough or not. This differs from [3] in the sense that the user is not giving a pair of 2D points in an image but he is modifying the grasp in a 3D space instead. This could be a quite difficult task if the interface allows to set the grasp in a completely free 3D space. For this reason, an interface using the obtained analytical model to set the grasp approach vector allows the user to move the end-effector around the cylinder axis to place it in the desired pose really quickly. Two possible grasp configurations are shown in Figure 4.

After the grasp to perform has been specified, it is necessary to check whether it is feasible or not. This can be done by computing the inverse kinematics of the whole arm kinematic chain and calculating its reachability. Our approach is to adopt a classical iterative inverse jacobian method where the jacobian is computed in order to exploit the kinematic redundancy of the whole system. This process can be done immediately after any movement on the grasp posture, notifying the user immediately after the posture has been chosen.



Radius 50
Angle 226
Distance 11

Radius 38
Angle 52
Distance 0

Fig. 4. Grasp specification interface showing the two types of grasp: side grasp (left) and front grasp (right).

When the grasp frame is selected and reachable it is executed in the UWSim simulator. In Figure 5 two possible approaches can be seen. With the 7-DOF GT arm there are enough degrees of freedom to reach different positions with completely different orientations while the use of the ARM5E arm in real experiments will constrain the possible approach vectors in a great extent, making almost impossible to make orientation variations by the user without including the vehicle in the kinematic chain. Finally, the grasp is executed moving the end-effector directly towards the object. The grasp simulation capabilities of UWSim consist of a range sensor that attaches the object to the inner part of the gripper when a given threshold is reached assuming that if the hand reaches the object without colliding with the jaws it will be able to grasp it. Although it doesn't use physics yet to perform the grasping, it lets us to visualize approximately how this grasp would perform in a real scenario. It is not the goal of this paper to increase the capabilities of UWSim in this aspect.

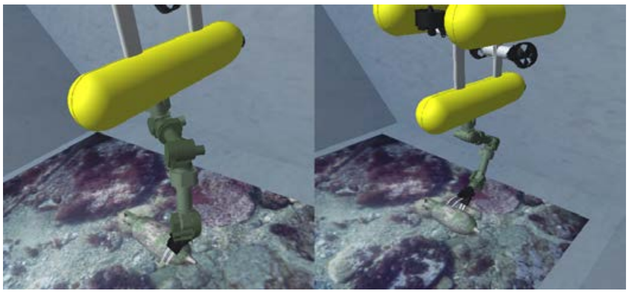


Fig. 5. UWSim showing the GT arm in the obtained grasp poses from two grasp specification processes.

V. RESULTS

The grasp specification algorithm presented in this paper has been executed with the described vehicle in the simulation environment. This has allowed to test the first steps of point cloud processing and the grasp specification interface successfully. The use of the simulator has made possible using an unavailable 7-DOF arm that can exploit the algorithms better than using an arm with less DOFs. Moreover, this decision has also prevented optical errors to appear because the light conditions of the simulated environment are almost ideal. For this reason, with which most simulations must cope, the results are good but the robustness of this method has not been fully tested yet.

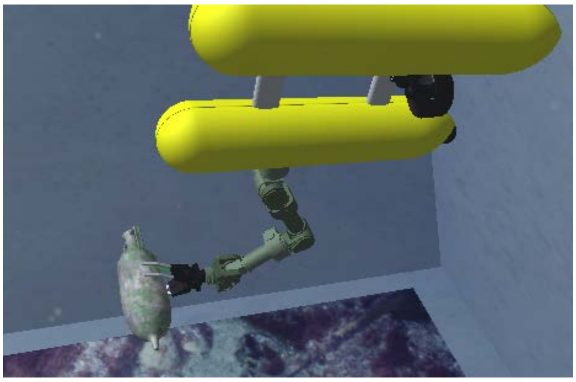


Fig. 6. GT arm carrying the amphora.

The segmentation and grasp specification stages have also been tested with real images using the Videre stereo camera. The point cloud from it (Figure 7) has been used successfully to validate the grasp specification phase. This step is critical for the final grasp posture quality and the tests show that this method performs better in flat surfaces. However, the cylindrical shape segmentation method has demonstrated to be robust.

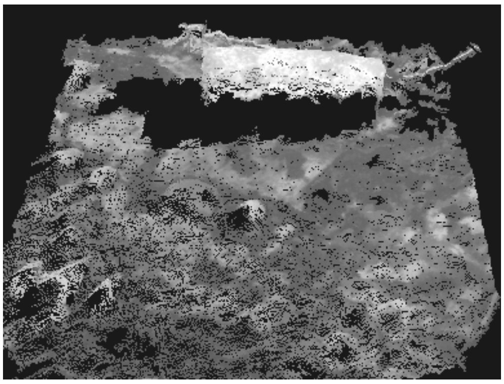


Fig. 7. Point cloud obtained with the Videre stereo camera of the real scenario with the amphora.

VI. CONCLUSIONS AND FUTURE WORK

This paper presents a new framework to perform autonomous manipulation of unknown structured objects in underwater environments. The experimental validation has been focused on grasp tasks in the constrained, yet realistic, problem of grasping unknown cylindrical objects like an amphora or a pipe. This framework can be further developed to recognize other primitive shapes. These shapes could be spheres, cuboids or even more complex models that can be approximated with a set of primitive shapes as shown in [10]. With that flexibility, the system could be capable of specify the grasp of other objects without the need of knowing the exact model of the object. However this flexibility will have the drawback of more computation time in these time sensitive tasks. The steps to obtain a model, specify a grasp and execute it are described and should be further validated in a real and practical scenario. The real testbed for the experiments has been shown in Figure 2. This will demonstrate the use of this specification procedure, although it will not use all its flexibility due to the fact that it is not possible to reach all possible orientations with the ARM5E arm, as it has only 5-DOF. The real object to grasp, with cylindrical shape, is shown in Figure 8. With this experiments, other issues are expected such as the proper calibration of the optical devices and the grasp execution control using tactile sensor feedback.

Fig. 8. Real object to grasp.

ACKNOWLEDGMENT

This research was partly supported by Spanish Ministry of Research and Innovation DPI2011-27977-C03 (TRITON Project), by the European Commission's Seventh Framework Programme FP7/2007-2013 under Grant agreement 248497 (TRIDENT Project), by Foundation Caixa Castell-Bancaixa PI.1B2011-17, and by Generalitat Valenciana ACOMP/2012/252.

REFERENCES

[1] G. Marani, D. Angeleti, G. Cannata and G. Casalino, *On the functional and algorithmic control architecture of the amadeus dual arm robotic workcell*, In WAC 2000, Hawaii (USA), June 2000.

[2] P. J. Sanz, P. Ridao, G. Oliver, C. Insaurralde, G. Casalino, C. Silvestre, C. Melchiorri, A. Turetta. *TRIDENT: Recent Improvements about Underwater Intervention Missions*, in IFAC Workshop on Navigation, Guidance and Control of Underwater Vehicles (NGCUV 2012), Porto(Portugal), April 2012.

[3] M. Prats, J. J. Fernández, and P. J. Sanz, *Combining Template Tracking and Laser Peak Detection for 3D Reconstruction and Grasping in Underwater Environments*, in International Conference on Intelligent Robots and Systems IROS 2012, Algarve(Portugal), October 2011.

[4] M. Prats, J. Pérez, J. J. Fernández, and P. J. Sanz, *An Open Source Tool for Simulation and Supervision of Underwater Intervention Missions*, in International Conference on Intelligent Robots and Systems IROS 2012, Algarve(Portugal), October 2011.

[5] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Y. Ng, *ROS: an open-source Robot Operating System*, in International Conference on Robotics and Automation (ICRA), Kobe(Japan), May 2009.

[6] D. Ribas, P. Ridao, LL. Magí, N. Palomeras, and M. Carreras. *The Girona 500, a multipurpose autonomous underwater vehicle.* In Proceedings of the Oceans IEEE, Santander, Spain, June 2011.

[7] CSIP. Light-weight arm 5e datasheet. *http://www.ecarobotics.com/ftp/ecatalogue/520/Fiche_LightWeight_5E.pdf.*

[8] R. B. Rusu and S. Cousins, *3D is here: Point Cloud Library (PCL)*, in IEEE International Conference on Robotics and Automation (ICRA), Shanghai(China), May 2011.

[9] R. Schnabel, R. Wahl, and R. Klein *Efficient RANSAC for Point-Cloud Shape Detection*, in Computer Graphics Forum (Vol. 26, No. 2, pp. 214-226). Blackwell Publishing Ltd., June 2007.

[10] S. Garcia, *Fitting primitive shapes to point clouds for robotic grasping*, Doctoral dissertation, Master Thesis, Royal Institute of Technology, Sweden.