Groundtruth system for underwater benchmarking

Alfredo Martins, André Dias, Hugo Silva, José Almeida,
Pedro Gonçalves, Flávio Lopes, André Faria, João Ribeiro, Eduardo Silva

INESC TEC Institute for Systems and Computer Engineering of Porto
ISEP - School of Engineering, Porto Polytechnic Institute, Porto, Portugal
Email: {alfredo.martins, andre.dias, hugo.m.silva, jose.m.almeida}@inescporto.pt,
{pgoncalves,f.lopes,a.faria,jp.ribeiro@}lsa.isep.ipp.pt, eduardo.silva@inescporto.pt

***Abstract*—In this paper a vision based groundtruth system for underwater applications is presented. The proposed system as an external validation perception and localization mechanism for underwater trials in the INESC TEC / ISEP underwater robotics test tank.**

It is comprised by a stereo camera pair with external synchronization and a image processing and data recording host computer. The cameras are disposed in a rigid baseline calibrated using scenario key points.

Two target detection algorithms were tested and their results are discussed. One is based on template matching techniques allowing the tracking of arbitrary targets without particular markers and the other on color segmentation with the target vehicle equipped with light markers.

Also an example trajectory of a small ROV motion in the task is also presented.

I. Introduction

The use of external ground truth systems is extremely relevant for robotics research. The need to evaluate our developed systems by other reliable external sensor, or perception mechanism is uppermost important to the assessment of the proposed system results. For marine and underwater robotics research purposes, the use of such systems are somewhat limited by the environment characteristics.

In this paper we propose a low-cost 3D vision based external ground-truth system for application in underwater environments. Although primarily developed as an external ground truth system for experimental tests in our test tank, it is also intended to be used in real operational scenario whenever visibility conditions allows. The widely spread use by the robotics community of vision based systems for external evaluation purposes comes from the easiness one may assemble, acquire synchronized image frames, and process its results at a fraction of the cost of other more complex systems. Usually vision based external ground-truth systems are divided into two categories: commercial and custom made solutions.

Focus in the commercial category, some of the possible solutions are the Vicon[1] and Optitrac[2], which provide enough accuracy for most robotics applications, but with higher costs. More than the price itself, these systems have limitations that could make them impracticable in some scenarios, such as the pulsed infra-red light limit range to the reflective markers, unable to be used underwater scenarios and the fact that some objects to track cannot carry the markers due to the geometry[3].

Regarding custom made solution, there are a wide range of implementations that were written from the scratch and are, with some exceptions, rarely usable by anyone other than the original programmer: XVision[4],TrackIt[5]. The exceptions are the SwisTrack[6], which is a generic and flexible tool for tracking robots and insects in flocks and formation control. It is important to clarify that all solutions available are limited to 2D tracking of known markers using a vision system located above the tracking area[3].

Until now most of the related work concerning the use of external sensors on underwater environments has been primarily focused in supporting navigation of autonomous underwater vehicles(AUVs), specially using auxiliary sensors e.g LBL, and USBL acoustic navigation methods. The use of acoustic methods in the external tracking of AUVs has been disseminated, due to the need of knowing the relative vehicle position to a home base with a high degree of confidence, while conducting operations in underwater scenarios. For example one can mention the work of Cruz et al [7], that developed a low-cost external acoustic navigation system for underwater applications. The system is capable of tracking multiple vehicles sharing the same acoustic channel, and determine the vehicle position in real time without using additional sources of information. Another example for underwater vehicle tracking positioning, is the subsea position device developed by Willemenot and Morvan [8], that merges inertial and acoustic technologies. The main concept behind their approach, is the combined use of USBL sensors that have low date rate and bounded errors, with Inertial Navigation Systems(INS/GPS) that have high date rate and smooth drifts assembled together in a single mechanical mounting. The procedure allows to increase the use of normal USBL systems, extending their scope of operation to shallow or very shallow waters by improving the system range, and positioning accuracy. The use of UBSL sensors can also be employed in other type of vehicle positioning, as in Stokey et al[9], where external sensors positioning were used to perform a docking maneuver of a REMUS AUV.

The paper outline is the following: In section 2, the ground-truth architecture and hardware setup is described. In section 3 some results of calibration and experimental setup for obtaining the position of videoRay ROV are presented, and finally in section IV we discuss the conclusions and future

978-0-933957-40-4 ©2013 MTS

work.

II. Ground truth system

A. Architecture

The proposed system (figure 1) consists in a time triggered camera (in minimum configuration) for timestamp of images, positioned a very small depth and usually mounted downward, a host image processing and recording computer and a set of predefined markers used in the underwater vehicle in order to estimate its position and pose. The 3D pose estimation is performed based on the marker structure (one sphere for z determination and 4 point markers for pose calculation). Marker monocular tracking is either performed based on color segmentation or with template matching. Although there are known limitations in color information for underwater images, we show that for the test tank considered (5m depth, 10x6m) it is still possible to obtain pose information and absolute location for small underwater vehicle such as a VideoRay Pro3 ROV up to 30 fps.

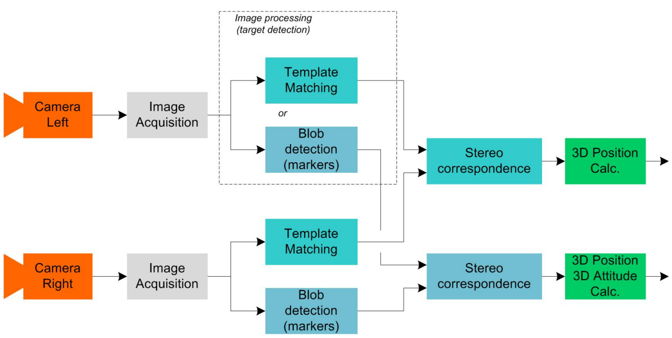

Fig. 1. System architecture

For experiments where only the required 3D localization is needed for motion targets, background subtraction techniques are used to efficiently detect relevant features in the image plane. This approach is similar to the one used in similar indoor vision based target localization scenarios [3]. Since in its minimal configuration it consists on a fixed camera, a host computer (usually a notebook) and very simple markers that can be attached to the vehicles to be tracked, it can also be deployed in mission locations both for groundtruth external validation system, or as part of a localization system for very short range cooperative underwater vehicles (obviously it is limited by water visibility conditions).

B. Setup

The system is installed in the INESC TEC / ISEP underwater robotics test tank. This tank is 10m by 6m with 5m depth and is used to perform laboratory tests and experiments for the underwater and marine robotics research performed in our lab.

It is located in the interior robotics test area of the lab and the upper 2m are above ground with a movable platform on top for easy access, see Fig.2.

Two cameras in an underwater housing were placed at the corner of the tank (in one of the smaller walls) in order to cover the largest tank area possible (Fig. 2 right). The cameras are two Basler acA1300-30gc color GigabitEthernet(GigE) digital cameras with external trigger, with maximum resolution of 1278x958 and up to 30 fps of framerate. These cameras were connected to a Gigabit switch that provided network interface with the image acquisition system running on an (intel i5 CPU running Linux operating system). Due to the GigE requirements, and camera driver restriction, the main limitation for the processing computer, besides its own computational power, is the possibility of having a Gigabit Ethernet port with jumbo frames. An external hardware trigger device was used to synchronize the cameras image capture.

The vehicle consisted in a VideoRay Pro3 small ROV. A colored light marker rig was made to test color segmentation methods and to provide 4 identifiable different points in the vehicle. The marker was constructed with 4 LED positioned in a plane with different colors and inside small diffusor white balls. A small battery was also included in order to not require any additional electrical connection to the test vehicle or system.

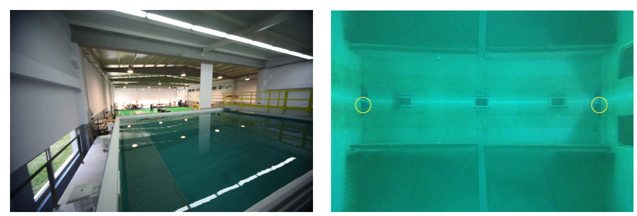

Fig. 2. Test tank (left) and camera placement marked in yellow (right)

C. Stereo Camera Calibration

The cameras calibration is performed in a two steps procedure conducted offline. First, we follow the method detailed in [10] that uses several pictures of a chessboard placed randomly inside a camera field-of-view(FOV), as illustrated in Fig.3, in order to obtain the camera intrinsic parameters(K) being: **f** the focal distance, **s** skew coefficient, and $(\mathbf{c_x}, \mathbf{c_y})$ the camera center, as defined by:

$$K = \begin{bmatrix} f & s & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Second, after obtaining cameras intrinsic parameters, there is the need to compute the geometric relation between both cameras. This procedure is denominated as obtaining the stereo camera extrinsic parameters configuration (R_{ext}, T_{ext}). To perform such calibration, it requires at least 8 extrinsic calibration control points visible in both cameras FOV, displaced in a pre-defined manner with known world (indoor pool) 3D position, as illustrated in Fig.4. Finally, using the 8-point algorithm [11] we are able to compute the Essential Matrix(E), that contains both (R_{ext}, T_{ext}) parameters, thus

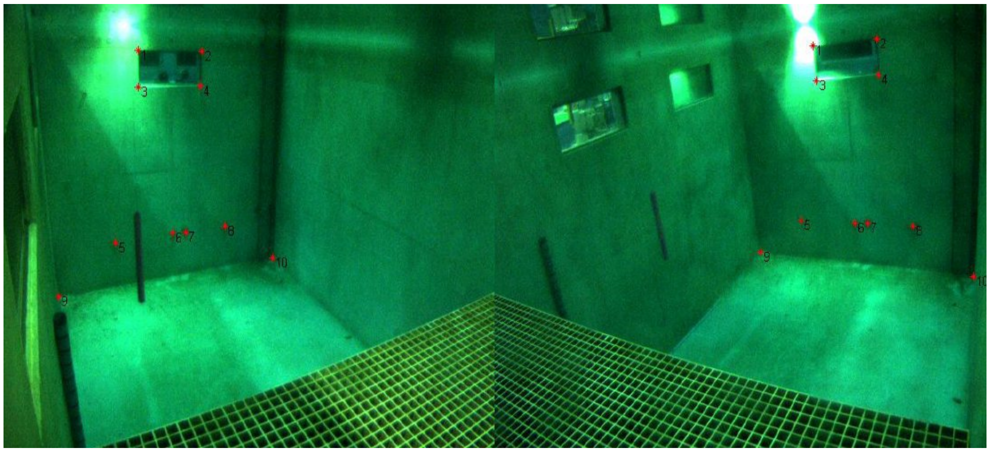

Fig. 4. Extrinsic parameter calibration with 10 (3D) position defined control points

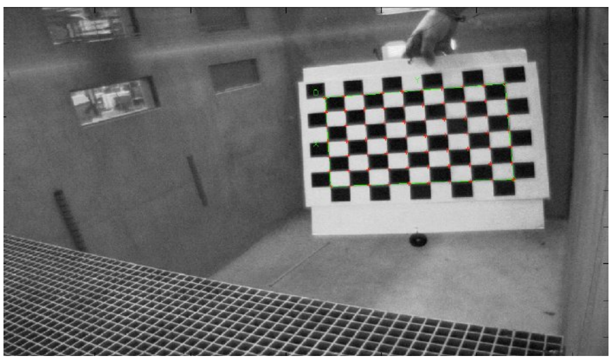

Fig. 3. Camera intrinsic parameters chessboard calibration

enabling the ground-truth system to determine 3D points, e.g when following a ROV trajectory.

D. Target detection

Since the system should provide position information from the target, it should implement at least a target detection method not requiring additional physical alterations to the vehicle to monitor. In view of this two alternative target detection methods were considered: Lucas-Kanade template matching and color segmentation.

In order to reduce the image search space for color segmentation, and since the environment is known and fixed, a prior background subtraction is performed. In this case the search is only performed in the action areas of the image.

The detectors were implemented using the OpenCV library. To perform template matching the user had to choose a template prior to the experiment, and the color segmentation process must be applied with the additional colored light markers (with previous segmentation limits definition).

E. Time synchronization

A relevant issue for ground truth motion determination is time synchronization. For localization or control analysis of underwater vehicles and in particular for AUVs ("Autonomous Underwater Vehicles") the onboard localization must be synchronized with the external ground truth in order for the comparison be meaningful. In these cases the approach taken is to synchronize the internal AUV clock with the external ground truth processing CPU clock. Since in our lab robots (and in many common research AUVs) the Linux operating system is used, we use the Chrony [12] daemon to perform multiple computer clock synchronization.

This approach is also taken in other robotics scenarios such as the land based robotic soccer one [3] for multiple robot and external cpu clock synchronization.

In addition for maritime scenarios this process also allows for further time precision taking advantage of the GPS (when available) PPS (pulse per second) signal.

For ROV experiments the data synchronization can be easily achieved using the same computer for the external image acquisition and ROV sensor processing. When that is not convenient the multiple clock synchronization approach can be used.

III. Results

In this section, we present some preliminary results that allow to evaluate our underwater ground-truth system.

A. Template Matching vs Color Based Segmentation

A set of 20 pre-determined static points in the field of view of the stereo cameras was chosen to characterize the two methods of target detection (template matching and color based segmentation). To exemplify the procedure, table I displays a comparison between the static measured points, and the template matching algorithm.

From the measured points vs vision based 3D estimation the following results where obtained:

From table II, one can easily note the higher mean error for the color segmentation method. This occurs since the effective target size is much smaller in the image due to the inherent loss of color information underwater.

In Fig. 5, one can observe the VideoRay ROV during the color based segmentation experiments.

TABLE I
RESULTS COMPARISON OF 20 POINT EXTRACTED STATIC POINTS AND THE LOCALIZATION MEASUREMENTS PERFORMED USING TEMPLATE MATCHING ALGORITHM

X(m)			Y(m)			Z(m)		
St	Tm	err(%)	St	Tm	err(%)	St	Tm	err(%)
5.48	5.50	0.36	0.00	0.04	-	2.33	2.39	2.58
3.91	3.88	0.77	0.00	0.03	-	0.64	0.58	9.38
0.78	0.80	2.56	0.00	0.03	-	0.28	0.31	10.7
1.00	1.03	3.00	0.00	0.04	-	1.00	1.05	5.00
2.00	2.02	1.00	0.00	0.03	-	2.00	2.08	4.00
2.00	2.01	0.50	0.00	0.04	-	1.00	1.07	7.00
5.00	5.05	1.00	1.00	1.06	6.00	1.00	1.09	9.00
3.50	3.54	1.14	1.00	1.05	5.00	1.00	1.08	8.00
4.06	4.08	0.49	4.22	4.24	0.47	2.00	2.05	2.50
2.00	2.02	1.00	4.22	4.27	1.18	2.00	2.08	4.00
3.00	3.03	1.00	4.22	4.24	0.47	2.00	2.08	4.00
3.00	3.03	1.00	4.22	4.25	0.71	3.00	3.09	3.00
3.00	3.02	0.67	4.22	4.23	0.24	2.50	2.55	2.00
1.00	1.02	2.00	4.50	4.52	0.44	2.00	2.07	3.50
2.00	2.02	1.00	4.50	4.52	0.44	2.00	2.04	2.00
4.00	4.05	1.25	4.50	4.59	2.00	2.00	2.10	5.00
3.00	3.03	1.00	5.00	5.08	1.60	1.00	1.09	9.00
3.00	3.02	0.67	5.00	5.07	1.40	2.00	2.08	4.00
2.00	2.03	1.50	5.23	5.26	0.57	1.00	1.06	6.00
4.06	4.03	0.74	5.23	5.25	0.38	1.00	1.05	5.00

TABLE II
MEAN ERROR AND STD DEVIATION FOR BOTH DETECTION METHODS

	Temp. Match		Color Seg.	
	$\bar{x}(cm)$	$\sigma(cm)$	$\bar{x}$	σ
x	2.57	0.9	4.77	2.08
y	3.8	1.9	2.08	3.15
z	7.7	2.3	12.7	2.29

Fig. 5. Tracking VideoRay ROV trajectory with color based segmentation

The RMS epipolar line error obtained from the number of extrinsic parameter calibration points was calculated to be 1.9 pixels. This error quantifies the quality of the stereo rectification.

B. ROV motion

A teleoperated ROV motion was also performed and the vehicle trajectory was determined by the system using template matchin, as displayed in Fig. 6.

In figure 7 is presented the 3D trajectory executed by the ROV, detected with template matching.

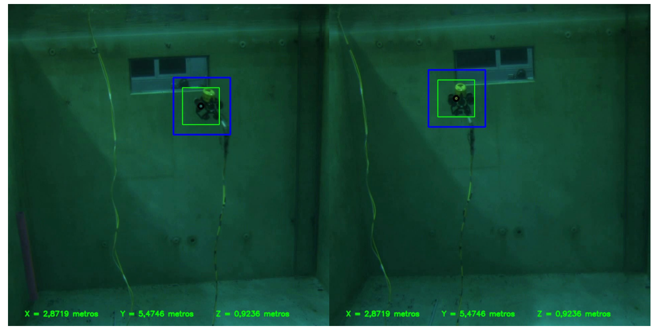

Fig. 6. Template matching visual stereo based target localization in test tank with a shorter baseline

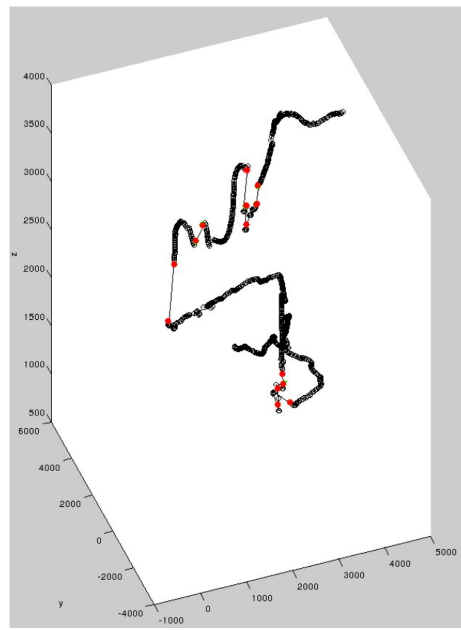

Fig. 7. Rov 3D trajectory

IV. CONCLUSIONS

In this paper we presented an external groundtruth target localization system for underwater applications.

The developed system is capable of detecting and tracking moving vehicles in our underwater test tank. Visual target information is gathered from a stereo camera pair and synchronized with the target vehicle internal clocks thus providing correlated data for vehicle trajectory and algorithm analysis.

The system can be used to track both marked or unmarked vehicles. A light emitting marker can be used to provide additional information about the vehicle pose.

The preliminary experiments show promising results particulary when the vehicle moves relatively close to the baseline.

In future work we will increase the area coverage by adding two more cameras on opposite pool side. We also like to improve the global tank illumination to allow the use of other computer vision techniques for underwater autonomous vehicle tracking. Sub-pixel precision matching algorithms using normalized cross correlation are being study to increase the accuracy in the estimation of the extrinsic control points, thus diminishing the overall stereo calibration error. Since we followed an open development approach, in future plans the system will be integrated into a ROS infrastructure, allowing further developments and a simple integration within the research community. We also plan to conduct more tests with a new-redesigned marker, less bulky thus helping the color based segmentation process. Finally, a generic setup and calibration process will be developed for enabling the use of the system in field experiences.

ACKNOWLEDGMENTS

The authors acknowledge the support given by the ISEP-IPP and INESC TEC to this project. This work is co-financed by the North Portugal Regional Operational Programme (ON.2 O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF) and by the ERDF - European Regional Development Fund through the COMPETE Programme and by National Funds through the FCT - Portuguese Foundation for Science and Technology within project FCOMP-01-0124-FEDER-022701.

REFERENCES

[1] G. H. Lee, M. Achtelik, F. Fraundorfer, M. Pollefeys, and R. Siegwart, “A benchmarking tool for mav visual pose estimation,” in *Proc. 11th Int Control Automation Robotics & Vision (ICARCV) Conf*, 2010, pp. 1541–1546.

[2] S. Augenstein, “Monocular pose and shape estimation of moving targets, for autonomous rendezvous and docking,” Ph.D. dissertation, Stanford University, 2011.

[3] J. Almeida, A. Martins, and E. Silva, “Real-Time Visual Ground-Truth System,” in *IbPRIA Iberian Pattern Recognition and Image Analysis*, Funchal, Portugal, Jun. 2013, pp. 304–313.

[4] P. M. S. N. Fry, M. Bichsel and D. Robert, “Tracking of flying insects using pan-tilt cameras,” *Journal of Neuroscience Methods*, vol. 101, pp. 59 – 76, August 2000.

[5] C. Interaction and R. lab John Hopkins University, “A brief tour of xvision [online],” http://www.cs.jhu.edu/CIPS/xvision, 2008.

[6] T. Lochmatter, P. Roduit, C. Cianci, N. Correll, J. Jacot, and A. Martinoli, “Swistrack - a flexible open source tracking software for multi-agent systems,” in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems IROS 2008*, 2008, pp. 4004–4010.

[7] N. Cruz, L. Madureira, A. Matos, and F. Pereira, “A Versatile Acoustic Beacon for Navigation and remote Tracking of Multiple Underwater Vehicles,” in *OCEANS01*, 2001.

[8] E. Willemenot, P.-Y. Morvan, H. Pelletier, and A. Hoof, “Subsea positioning by merging inertial and acoustic technologies,” in *OCEANS 2009-EUROPE*. IEEE, May 2009, pp. 1–8.

[9] R. Stokey, B. Allen, T. Austin, R. Goldsborough, N. Forrester, M. Purcell, and C. von Alt, “Enabling technologies for REMUS docking: an integral component of an autonomous ocean-sampling network,” *IEEE Journal of Oceanic Engineering*, vol. 26, no. 4, pp. 487–497, 2001.

[10] J. Y. Bouget, “Camera calibrationn toolbox for matlab.” [Online]. Available: http://www.vision.caltech.edu/bougetj/calib_doc/index.html/

[11] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press, 2003.

[12] “Chrony.” [Online]. Available: http://chrony.tuxfamily.org/