An audio-based strategy for the geo-referencing of visual data to support diver's monitoring missions

Riccardo Costanzi
Dipartimento di Ingegneria dell'Informazione (DII)
Centro di Ricerca "E. Piaggio"
University of Pisa
Pisa, Italy
Interuniversity Center of Integrated Systems for the Marine Environment (ISME)
Italy
riccardo.costanzi@unipi.it

Vincenzo Manzari
Naval Support and Experimentation Center (CSSN)
La Spezia, Italy
vincenzo.manzari@marina.difesa.it

Lorenzo Pollini
Dipartimento di Ingegneria dell'Informazione (DII)
Centro di Ricerca "E. Piaggio"
University of Pisa
Pisa, Italy
Interuniversity Center of Integrated Systems for the Marine Environment (ISME)
Italy
lorenzo.pollini@unipi.it

Francesco Ruscio
Dipartimento di Ingegneria dell'Informazione (DII)
University of Pisa
Pisa, Italy
francesco.ruscio@phd.unipi.it

***Abstract*—The collaboration between humans and robots represents a major topic in robotics. In the underwater domain, the employment of unmanned robots could help divers perform number of tasks such as monitoring and inspection activities. However, the underwater environment poses severe limitations that still need to be overcome to achieve autonomous operations. This research work represents an intermediate step in the human-robot switchover in underwater domain for a particular application and involves the use of ICT tools to facilitate human divers in the monitoring of Posidonia Oceanica (Po) meadows. The paper reports a strategy to geo-reference underwater visual data exploiting the synchronization between the audio track recorded by a camera and the transponder pings adopted for the acoustic positioning system. The paper also reports the exploitation of the geo-referenced data for the identification of Po meadow over the mission area using a Machine Learning algorithm. The results are very promising and can lead to an accurate geo-referenced identification of Po and the reconstruction of the surveyed area.**

Keywords—underwater vision, acoustic localization, data geo-referencing

I. Introduction

Human-robot interaction has been receiving considerable attention in the recent years, bringing benefits in terms of safety, cost and time consumption. Collaborative robotics is one of the key topics on which the activity of the University of Pisa – Dipartimento di Ingegneria dell'Informazione (UNIPI-DII) is based on. The CrossLab project is the DII development project based on the realization of interdisciplinary laboratories dedicated to the key areas of the Indusrty 4.0 paradigm. Among the laboratories, the one dedicated to Advanced Manufacturing concerns the development of robots that are capable of working effectively in real-world domains and collaborating with humans in their daily activities. The underwater domain represents a very harsh environment to deal with, in which human divers operate. The employment of unmanned robots in such environment could help human divers to perform a number of complicated tasks. Although researchers have recently opened the door to unmanned underwater robots, a long path is still necessary to pave the way to underwater applications performed in an autonomous way. Part of the CrossLab Advanced Manufacturing work is devoted to the implementation of autonomous underwater robots to support human activities such as water quality monitoring, fish farms, hull inspections, to mention but a few. One of the most frequent activities to be performed in the marine environment is represented by the assessment of the health state and the extension of Posidonia Oceanica (Po) meadows [1]. In the Italian context, the monitoring of Po is handled by the Regional Agencies for the Environmental Protection (ARPA), employing professional scuba divers who spend several days per year in such monitoring activities that are mainly based on visual surveys. The collaboration between the Ligurian ARPA (ARPAL) and DII aims at gradually employing robotics and ICT tools to support the monitoring campaigns. This can overcome the problem of assigning repetitive, dangerous, and expensive tasks to human operators. In addition, an improvement of the quality of acquired data and of the precision of their geo-referencing is expected. This research work fits into the framework discussed above and represents an intermediate step in the human-robot switchover for the specific application, involving the implementation of a visual data geo-referencing strategy based on acoustic synchronization. The visual dataset used within this research activity refers to a Po monitoring campaign performed in front of the Ligurian Coast in March 2019. The activity was performed by a diver using a Smart Dive Scooter (SDS) [2], which is a modified version of a standard dive scooter (SUEX XJ VR) equipped with a camera system and a modem in transponder configuration (Fig. 1). The acoustic positioning was accomplished using a compatible Ultra Short Baseline (USBL) device installed on a support boat. The proposed geo-referencing strategy exploits the synchronization between the audio tracks recorded by a camera and the acoustic transponder pings used to locate the position of the diver throughout the mission. This way, the visual data can be elaborated to extract useful information about the recorded area and hence used for monitoring purposes and reconstruction of the sea bottom. Furthermore, this approach is not limited to optical data [4],[5] but also to acoustic imaging devices such as sonar [6].

The work has been presented in the IEEE/MTS Oceans 2020 Global Conference [7]; here it is possible to retrieve all the details of the proposed strategy.

Figure 1 - SDS employed during the mission, acoustic transponder below the scooter, cameras on the left [2]

II. DATA SYNCHRONIZATION

Both visual and acoustic localization systems are stand-alone modules. This means that the video streams need to be synchronized with the estimated positions in order to achieve a correct geo-referencing of the collected images. The synchronization process has been tackled exploiting the strategy proposed in [2], which involves the identification of the transponder pings within the audio track of the acquired video. To this end, a processing of a video collected by one of the four cameras has been carried out. Firstly, the audio signal has been extracted from the recorded video. Then, a band pass filter in the range of 13kHz – 15kHz has been adopted to remove the environmental noise. This way, the transponder pings are clearly visible within the filtered audio track and it can be further exploited for synchronization purposes. The synchronization process has been accomplished computing the maximum of the cross-correlation function between the signal peaks present in the filtered audio track and the recorded transponder pings. The achieved time alignment of the audio track extracted by the video and the transponder pings implies the match between the video stream and the position measurements. This way, it is hence possible to associate a position information to the images collected by the camera.

III. PATH ESTIMATION

A filtering process has been applied to estimate the diver's position during the mission. Given the non-linearity of the system, an Extended Kalman Filter (EKF) has been selected to address the state estimation problem. The diver was supposed to move at constant depth and at approximately constant speed. For this reason, the prediction phase is based on a first-order kinematic model, where both speed and heading have been assumed quasi-constant. On the other hand, the correction phase exploits the measurements provided by the USBL in order to refine the predicted state estimation. The possible inaccurate measurements are excluded adopting an outlier rejection phase based on the Mahalanobis distance. A constant sampling time of 0.1s has been adopted for the estimation process, so providing the estimated diver's position with a frequency of 10Hz.

IV. PO IDENTIFICATION

The achieved geo-referencing of visual data allows to extract useful information about the recorded area. In this regard, the synchronized video stream has been used to identify and describe the Po presence over the mission area. The adopted approach exploits the Machine Learning (ML) technique proposed in [3], which is capable of detecting the presence of Po within a visual dataset. In particular, the algorithm returns a binary image according to the presence of Po: black pixels if Po is detected, white otherwise. The percentage of black pixels in each frame has been selected as an indicator of the Po presence.

V. RESULTS AND CONCLUSIONS

As a validation of the presented work, Fig. 2 reports the result of the entire process: path estimation, data synchronization and Po identification. Specifically, it shows the presence of Po along the estimated diver's path. Each frame, which corresponds to a specific estimated position, is associated with a value that corresponds to the percentage of Po according to the ML image segmentation result. Such values have been used to define a chromatic scale, which

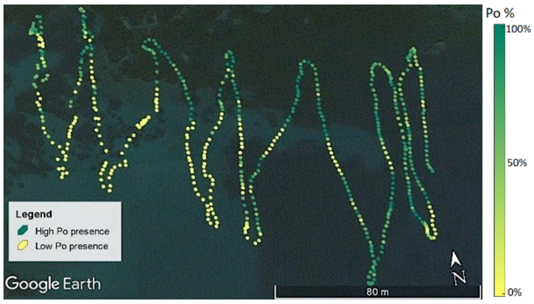

Figure 2 - Estimated diver's path with chromatic scale based on the percentage of Po detected by the ML algorithm

distinguishes the level of Po concentration in every estimated position: green color for high Po concentration, yellow color for low level of Po concentration. Each estimated position together with its Po concentration value has been plotted over a satellite image of the mission area. The Figure shows that the result in terms of Po presence is in accordance with the information retrievable from an aerial point of view. It is then possible to conclude that the proposed strategy is effective in describing the presence of the Po meadow over an area of interest and can represent a valuable approach to address the geo-referencing issue of independent data. In this regard, future developments will be related to the reconstruction of the seabed as a georeferenced 3D model. Long term objective is the integration of the visual acquisition system on an Autonomous Underwater Vehicle that can perform the monitoring campaigns autonomously, with the aim of improving the coverage performance as well as reducing the risk for human operators

ACKNOWLEDGMENT

The authors would like to thank all the SEALab members, the joint applied research laboratory between the Naval Support and Experimentation Centre (CSSN) of the Italian Navy and the Italian Interuniversity Research Center of Integrated Systems for Marine Environment (ISME), who supported the research team during the experimental activities described in this work. Work partially supported by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence)

REFERENCES

[1] M. C. Buia, M. C. Gambi, and M. Dappiano. Seagrass systems. Biologia Marina Mediterranea, 11(Suppl.1): pp.133–183, 2004.

[2] R. Costanzi et al., "Visual acquisition system for georeferenced monitoring and reconstruction of the sea bottom using audio for data synchronisation," OCEANS 2019 - Marseille, Marseille, France, 2019, pp. 1-6

[3] A. Burguera, F. Bonin-Font, J.L. Lisani and A.B. Petro. "Towards Automatic Visual Sea Grass Detection in Underwater Areas of Ecological Interest," IEEE International Conference on Emerging Technologies and Factory Automation, Berlin, 2016.

[4] S. Wirth, P.L.N. Carrasco, and G.O. Codina. (2013, June). Visual odometry for autonomous underwater vehicles. In 2013 MTS/IEEE OCEANS-Bergen (pp. 1-6). IEEE.

[5] F. Bonin-Font, G. Olivier, S. Wirth, M. Massot, P.L. Negre, and J.P. Beltran. (2015). Visual sensing for autonomous underwater exploration and intervention task. Ocean Engineering, 93, 25-44.

[6] M. Franchi, A. Ridolfi, L. Zacchini, (2018). A Forward-Looking Sonar-Based System for Underwater Mosaicing and Acoustic Odometry. 2018 IEEE OES Autonomous Underwater Vehicle Symposium (AUV 2018), Porto, Portugal, November 5-9, 2018.

[7] F. Ruscio, R. Costanzi, L. Pollini, V. Manzari, C. Barbieri, F. Gaino, (2020). Geo-referenced visual data for Posidonia Oceanica coverage using audio for data synchronization. OCEANS 2020, Gulf Coast, October 2020.