Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-23 (2016) 500–505

Towards Visual Detection, Mapping and Quantification of Posidonia Oceanica using a Lightweight AUV ⋆

Francisco Bonin-Font * **Miquel Massot Campos** * **Gabriel Oliver Codina** *

* *Systems, Robotics and Vision Group, University of the Balearic Islands, 07122 Palma de Mallorca (e-mail: francisco.bonin@uib.es, miquel.massot@uib.cat, goliver@uib.es).*

Abstract: Posidonia Oceanica (P.O.) is a Mediterranean endemic seagrass strongly related to the health of the coastal ecosystems. Monitoring the presence and state of P.O. is essential not only for safeguarding the shallow-water life diversity, but also as an indicator of the water quality. Nowadays, the control of P.O. is done by divers in successive missions of a duration limited by the capacity of the scuba tanks. This paper proposes the application of robotic and computer vision technologies to upgrade these current methods, namely: 1) employing a lightweight *Autonomous Underwater Vehicle* (AUV) equipped with cameras to survey and image marine areas, 2) the automatic discrimination of P.O. from the rest of the seafloor, using several techniques based on image texture analysis and *machine learning*, and, 3) the fast computation of 2D maps (photo-mosaics) of the surveyed areas from all the images included in the grabbed video sequences; these mosaics are extremely useful to measure the real extension of the meadows and some of the descriptors needed for a biological analysis. Experiments conducted with an AUV in several marine areas of Mallorca reveal promising results in the discrimination of different patterns of P.O. and in the construction of highly realistic photo-mosaics of the surveyed areas.

Keywords: Autonomous Underwater Vehicles, Machine Learning, Visual Surveying and Mosaicking.

1. INTRODUCTION

Posidonia Oceanica (P.O.) is an endemic low-growing seagrass of the Mediterranean that forms vast colonies with a great ecological value, playing a critical role in the development and stability of coastal ecosystems. P.O. meadows protect the shoreline against erosion, attenuate currents and wave energy, they are a source of food and refuge for numerous species and a favourable substrate for many organisms. Furthermore, they absorb great amounts of carbon and release oxygen to the water by means of the photosynthesis, increasing its quality and transparency (Diaz-Almela and Duarte, 2008).

Several biologists have studied the evolution of P.O. along and across the Mediterranean with opposite results. Some studies showed evidences of decline on a global scale (Short et al., 2012; Jorda et al., 2012) due to human activities, such as boats fuel spilt and mechanical erosion. However, other authors defend the idea that, there is not a global and general decline but a decline due to an accumulation of local impacts, which can be overcome by acting upon the corresponding local causes (Moreno et al., 2001).

The European Commission identifies, in its directive Dir 92/43/CEE, Posidonia Oceanicae as a natural habitat of priority interest requiring the delimitation of special areas of conservation. In consequence, monitoring and controlling these benthic habitats becomes a crucial and a necessary task, since, indirectly, they benefit the tourism and fishing industries, two strategic sectors in many Mediterranean resorts.

Nowadays, the control of P.O. is typically done by divers, who photograph, mark the perimeter of the meadows to see their extension, and install certain gauges inside for measuring their height. Sometimes, divers are tracked with acoustic localizers to build a georeferenced survey (Scaradozzi et al., 2009). However, this process is slow, dangerous, imprecise and limited by the oxygen capacity of the scuba diving tanks and the necessary security procedures. Therefore, new strategies arise to improve accuracy in the measurements, extend the mission duration and evade strict security protocols. Several approaches to map and control P.O. colonies exploit multispectral satellite imagery (Matarrese et al., 2008). Although they revealed to be useful to delimit meadows in shallow waters, they do not result effective in deeper areas, where the water column complicates the perception of different blue tonalities. Acoustic bathymetries performed with a *Side Scan Sonar* (SSS) attached to a vessel hull or to an underwater vehicle (Montefalcone et al., 2013) can also be used to detect and map P.O.. However, it is difficult to discriminate the P.O. from other seagrass or algae just by visual inspection, being necessary additional complex algorithms to post-process the acoustic data. Recently, first experiments to explore regions with P.O. using lightweight AUVs and to calculate their biological parameters combining some image segmentation and a side scan sonar were shown in (Vasilijevic et al., 2014). However, in this solution the P.O. is identified visually by an human operator. Thus, to improve these results there is still a lot of research to be done.

⋆ This work is partially supported by Ministry of Economy and Competitiveness under contracts TIN2014-58662-R, DPI2014-57746-C3-2-R and FEDER funds

2405-8963
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.485

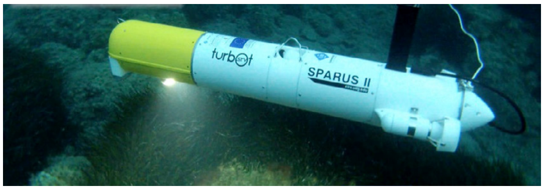

Fig. 1. SPARUS II over an area with P.O.

In the context of the ARSEA national Spanish project TIN2014-58662-R, we propose to upgrade the prevalent methodology used to visualize, map and control P.O. meadows. The idea consists in 3 main points: 1) to survey marine areas with an AUV equipped with a stereo camera looking downwards and programmed to navigate at a constant altitude; this permits to extend the duration, to automatize the missions over geo-referenced areas, to increase the precision in the computation of the camera motion and to gain flexibility in the depth coverage avoiding any human risk, 2) to employ several computer vision technologies based on texture analysis and machine learning to discriminate automatically P.O. from the rest of the sea bottom; 3) to apply an innovative and fast photo-mosaicing approach (García et al., 2016) to build 2D real-coloured precise mosaics of the surveyed areas and binarized mosaics where the P.O. is segmented from the rest; these models are geo-referenced to absolute locations (in the basis of the AUV navigation data and a GPS) and can be used to control its state and evolution, to study the spatial structure of the meadow, and to measure automatically some of the descriptors suitable for assessing its ecological status.

2. EXPERIMENTAL SETUP

This paper presents an essentially experimental work. A SPARUS II AUV (Carreras et al., 2013) (see figure 1), property of the University of the Balearic Islands, was programmed to move in 5 different locations colonized with P.O. of the west and north-west coast of Mallorca, with different levels of tourist exploitation and with different environmental conditions (illumination and turbidity). The vehicle was equipped with a *Doppler Velocity Log* (DVL), a pressure sensor, an *Inertial Measurement Unit* (IMU), a GPS to be geo-referenced in the water surface, and a Point Grey Bumblebee stereo rig, grabbing at 10fps with its lens axis perpendicular to the seafloor. The vehicle estimated its displacement from the DVL data and a stereo visual odometer, running additionally an advanced visual *Simultaneous Localization and Mapping* (SLAM) (Negre et al., 2016) algorithm to correct periodically its estimated position.

Five additional datasets were recorded by a diver holding a GoPro Hero 4 with its lens axis also perpendicular to the bottom. The vehicle and the diver speeds were approximately 0.5 nautical knots which resulted in an adequate overlap between successive images of 50%, in average. The origin of all these aforementioned video sequences was geo-referenced with a GPS hold at the water surface.

Some of the datasets grabbed by SPARUS II were recorded during the sunset thanks to the spotlights attached to the vehicle. The original resolution of all the images grabbed with the Bumblebee was 1024×768 and the resolution of the images grabbed with the GoPro was 1920×1080 pixels.

Figure 2 shows several images extracted from video sequences grabbed in different areas of Santa Ponsa and Port of Valldemossa. P.O. corresponds to the darker areas.

3. VISUAL DISCRIMINATION OF POSIDONIA

The first objective of this novel procedure was to discriminate automatically the alive P.O. in the images from the background (rocks, algae, sand, or dead matte). The P.O. visual detection is based on (Massot-Campos et al., 2013), and it runs in 3 main stages: a) the image texture description, b) the training phase, and c) the automatic classification. However, the training stage has several differences with respect to (Massot-Campos et al., 2013): 1) the number of training images and its distributed origin, 2) the size of the patches and their overlap, and 3) the number and type of the trained classifiers.

3.1 Image Description and The Training Dataset

Images were divided in 50% overlapping patches of 32 pixels each one. In the same way as (Massot-Campos et al., 2013), each patch was described by 168 attributes, computed from the Law's energy filters and the Co-ocurrence matrix. In short: 1) The basic Law's vectors designed to detect levels ($L5=[1,4,6,4,1]$), edges ($E5=[-1,-2,0,2,1]$), spots ($S5=[-1,0,2,0,-1]$), waves ($W5=[-1,2,0-2,1]$) and ripples ($R5=[1,-4,6,-4,1]$), are convolved with themselves to obtain the analogous 9 and 17 component vectors (L9, E9, S9, W9, R9, L17, E17, S17, W17 and R17). Then, the original vectors are multiplied all with all to obtain 25 5×5 2D kernels, and analogously, one can obtain 25 9×9 and 25 17×17 2D kernels. 2) These 75 kernels are convolved with each image patch; the average (μ) and the standard deviation (σ) calculated from each resulting convolution are the first 2×75 descriptors. 3) 18 additional descriptors are calculated from the Gray Level Co-occurrence matrix of each image patch: average (μ), variance (σ^2), standard deviation (σ), contrast (Con), entropy (Ent), homogeneity (Hom), angular second moment (ASM), energy (E), skew (Ske), kurtosis (Kur), maximum probability by rows and by columns (max_i, max_j), GLCM mean by rows and by columns (μ_i, μ_j), GLCM variance by rows and by columns (σ_i^2 and σ_j^2) and GLCM correlation (σ_{ij}^2).

A dataset composed by 69 images was used to train a classifier model based on machine learning and data mining. Contrarily to (Massot-Campos et al., 2013) and in order to supply the classifier trainer with a wide range of patch samples, images selected for the training dataset were taken from 6 different video sequences recorded from both cameras, the GoPro and the Bumblebee, with different types of P.O. textures and different illumination conditions. Patches with P.O. were hand-labelled as *class 1* and patches with no P.O. were labelled as *class 0*. One third of the training images had only P.O., another third had no-P.O., and the last third contained patches with P.O. and patches without. All these images were hand-binarized to define the ground truth of each patch-class. Patches not completely full of P.O. but labelled as *class 1* introduced slight errors in the classification training model. Figure 3 shows some samples included in the training dataset with its hand-labelled ground truth.