Available online at www.sciencedirect.com

Journal of Applied Research and Technology

Journal of Applied Research and Technology 14 (2016) 338–344

www.jart.ccadet.unam.mx

Original

Machine Learning for benthic sand and maerl classification and coverage estimation in coastal areas around the Maltese Islands

Adam Gauci [a,b,f,*], Alan Deidun [a], John Abela [c,f], Kristian Zarb Adami [d,e,f]

[a] Department of Geosciences, Faculty of Science, University of Malta, Malta
[b] Department of Intelligent Computer System, Faculty of ICT, University of Malta, Malta
[c] Department of Computer Information Systems, Faculty of ICT, University of Malta, Malta
[d] Department of Physics, Faculty of Science, University of Malta, Malta
[e] Department of Physics (Astrophysics), University of Oxford, UK
[f] Institute of Space Sciences and Astronomy, University of Malta, Malta

Received 14 March 2016; accepted 15 August 2016
Available online 27 October 2016

Abstract

Analysis of the seabed composition over a large spatial scale is an interesting yet very challenging task. Apart from the field work involved, hours of video footage captured by cameras mounted on Remote Operated Vehicles (ROVs) have to be reviewed by an expert in order to classify the seabed topology and to identify potential anthropogenic impacts on sensitive benthic assemblages. Apart from being time consuming, such work is highly subjective and through visual inspection alone, a quantitative analysis is highly unlikely to be made. This study investigates the applicability of various Machine Learning techniques for the automatic classification of the seabed into maerl and sand regions from recorded ROV footage. ROV data collected from depths ranging between 50 m and 140 m and at 9.5 km from the northeast coastline of the Maltese Islands, is processed. Through the application of the presented technique, 5.23 GB of data corresponding to 2 h and 24 min of footage which was collected during June 2013, was initially cleaned and classified. An estimate for the percentage cover of the two benthic habitats (sandy seabed and maerl) was also computed by using artifacts encountered during the ROV survey and of known dimensions as a reference. Unlike other automatic seabed mapping techniques, the presented prototype processes video footage captured by a down-facing camera and not through acoustic backscatter. Image data is easier and much cheaper to capture. Promising results that indicate a very good degree of agreement between the true and predicted habitat type distribution values, were obtained.

Keywords: Machine Learning; Image processing; Seabed classification; Decision trees; Maerl detection; Sand detection

1. Introduction

In the last few decades there has been a dramatic upsurge in the laying of submarine cables and pipelines that are deployed mainly for communications or energy-transfer purposes. Such laying inevitably implies a number of environmental changes which warrant the conduct of environment impact assessment (EIA) studies. Moreover, pipelines are normally deployed at great depths, beyond safe SCUBA diving limits. Remotely operated vehicles (ROVs) are used to collect baseline data as well as to monitor the environmental changes. Whilst unaided human analysis of such video footage allows one to infer qualitative conclusions about the seabed type being studied, the conduction of quantitative assessments through manual means is virtually impossible and highly subjective.

In order to ensure a higher degree of security in energy supply, in 2013 the Maltese government reached an agreement with the Italian government to connect the islands with the European electrical grid. This involved the laying of two 95 km-long submarine electrical cables between Qalet Marku in Malta and Marina di Ragusa in Sicily. As shown by Borg,

* Corresponding author.
E-mail address: adam.gauci@um.edu.mt (A. Gauci).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

http://dx.doi.org/10.1016/j.jart.2016.08.003

Fig. 1. Photo of a sample of maerl taken within a laboratory.

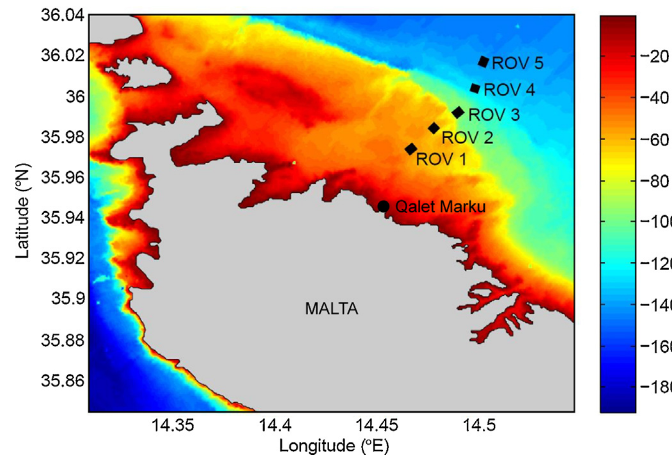

Fig. 2. Bathymetry close to the Maltese islands and the five seabed stretches of 500 m (total of 2.5 km), staggered through 1000 m-long intervals, over which ROV footage was recorded.

Lanfranco, Mifsud, Rizzo, and Schembri (1999), Dimech, Borg, and Schembri (2004), Sciberras et al. (2009), and Agnesi et al. (2009), the planned transect intersected with maerl assemblages, which consist of accumulations of calcareous rhodophyte thalli belonging mainly to Corallinaceae and marginally to Peyssonneliaceae (Fig. 1).

In this study, the applicability of various Machine Learning (ML) techniques for the automatic classification of the seabed into maerl and sand regions from recorded ROV footage, was investigated. The developed prototype contributes by providing an alternative or to supplement costly and highly labor-intensive acoustic benthic mapping techniques. The manual analyses of collected video footage is also elevated and a non-subjective approach to quantify the seabed type is made available.

Examples of artificial intelligence (AI) methods used for seabed classification can be found in literature; however, most studies consider acoustic backscatter data. Stephens and Diesing (2014) used multi-beam echo-sounder data collected from the North Sea to classify the substrate types. In this case, the ground truth was determined through the collected samples. Moškon, Žibert, and Kavšek (2015) also demonstrated how classifiers can be applied to raw multi-beam acoustic data for seabed mapping. Similar studies were carried out by Coiras and Williams (2009) and Landmark, Solberg, Austeng, and Hansen (2014). Investigations of seabed classification from image data were not found. As suggested by Stephens and Diesing (2014), reliable automated approaches that provide quantitative results and which can be included in monitoring programs are still relatively novel. While backscattered acoustic signals highlight the differences in the seabed clearly, surveys making use of such signals require a lot of planning, a lengthy permitting procedure, and are very expensive to run. In this study, camera footage captured by an ROV is used. The required data can also be recorded by a towed down-facing camera and is independent of the water transparency and of differences in the lightning conditions.

The following section provides further details on the video footage and how the frames were extracted. Details about the classification methods used and the obtained results are discussed in Sections 3 and 4, respectively. Planned future work and concluding remarks are given in Section 5.

2. Image data

For this study, 2 h and 24 min of video footage which were recorded by the down-facing camera of the ROV over five seabed stretches of 500 m were processed (Fig. 2). The ROV transects followed the path proposed for the submarine interconnector cable. The RAW video was captured at 29 frames per second and recorded on digital media in MPEG2 format. The processing was carried out on frames of 720×480 pixels. Since the velocity of the ship was kept constant during data capture, the coordinates of each frame could be computed from the corresponding timestamps. In Table 1, the starting and ending coordinates (in decimal degrees), the average depth, the average vessel speed and the number of frames extracted for every stretch are summarized.

The resolution of the raster set was estimated with frames that showed artifacts of known dimensions. In particular, reference was made to a bomb shell dating from World War II which was captured in a number of frames (Fig. 3). As documented in Boyd (2009), the body length (from the tip without the tail fins) and the diameter at the widest part of such 500-lb general purpose bombs are of 90.67 cm and 30.02 cm, respectively. The number of pixels representing these lengths were obtained from five different frames and the averages were found to be 301.84 and 92.87 pixels. By using these two measurements, an estimate for the physical dimension of individual pixels could be calculated as being 0.003004 m and 0.003233 m, with an average of 0.003118 m. Since the frames consisted of 720×480 pixels, each scene represented an area of 2.2451 m × 1.4968 m.

Some of the frames were found to contain periodic noise that contaminated the signal. Such frames were projected onto the Fourier domain and any high frequency components away from the central axes were masked out to remove this periodic pattern. Figure 4 depicts the noise removal process on the blue channel in both the image and Fourier spaces. Through this enhancement, no extra information or artifacts was added to the image data and the luminous values remained unchanged.