Continental Shelf Research 45 (2012) 87–97

Contents lists available at SciVerse ScienceDirect

Continental Shelf Research

journal homepage: www.elsevier.com/locate/csr

Research papers

Image-based continental shelf habitat mapping using novel automated data extraction techniques

Jan Seiler [a,b,*], Ariell Friedman [c], Daniel Steinberg [c], Neville Barrett [a,b], Alan Williams [b], Neil J. Holbrook [d]

[a] Institute for Marine and Antarctic Studies, University of Tasmania, Taroona, Tasmania 7053, Australia
[b] Wealth from Oceans National Research Flagship, CSIRO Marine and Atmospheric Research Hobart, Tasmania 7000, Australia
[c] The Australian Centre of Field Robotics, School of Aerospace, Mechanical and Mechantronic Engineering, University of Sydney, New South Wales 2006, Australia
[d] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania 7001, Australia

ARTICLE INFO

Article history:
Received 25 October 2011
Received in revised form
5 June 2012
Accepted 11 June 2012
Available online 16 June 2012

Keywords:
Habitat variability
Habitat predictors
Random forests
Monitoring
Sirius
Tasmania

ABSTRACT

We automatically mapped the distribution of temperate continental shelf rocky reef habitats with a high degree of confidence using colour, texture, rugosity and patchiness features extracted from images in conjunction with machine-learning algorithms. This demonstrated the potential of novel automation routines to expedite the complex and time-consuming process of seabed mapping. The random forests ensemble classifier outperformed other tree-based algorithms and also offered some valuable built-in model performance assessment tools. Habitat prediction using random forests performed most accurately when all 26 image-derived predictors were included in the model. This produced an overall habitat prediction accuracy of 84% (with a kappa statistic of 0.793) when compared to nine distinct habitat classes assigned by a human annotator. Predictions for three habitat classes were all within the 95% confidence intervals, indicating close agreement between observed and predicted habitat classes. Misclassified images were mostly unevenly, partially or insufficiently illuminated and came mostly from rugged terrains and during the autonomous underwater vehicle's obstacle avoidance manoeuvres. The remaining misclassified images were wrongly or inconsistently labelled by the human annotator. This study demonstrates the suitability of autonomous underwater vehicles to effectively sample benthic habitats and the ability of automated data handling techniques to extract and reliably process large volumes of seabed image data. Our methods for image feature extraction and classification are repeatable, cost-effective and well suited to studies that require non-extractive and/or co-located sampling, e.g. in marine reserves and for monitoring the recovery from physical impacts, e.g. from bottom fishing activities. The methods are transferable to other continental shelf areas and to other disciplines such as seabed geology.

1. Introduction

Habitat mapping is an essential tool to aid managers in assessing and managing the status of marine ecosystems. Currently mapping of marine habitats is principally based on two data sources, which are acoustic and optical. Both sources are acquired remotely and sampling requires no physical contact with the substrate as opposed to grab samples. Acoustic mapping technologies include multi-beam echo sounder (MBES) and side scan sonar (SSS). Optical mapping technologies include satellite and aircraft remote sensing, platform-based video camera and sediment profile camera (Rhoads and Germano, 1982). In shallow water (<100 m), the density of individual MBES soundings is generally several per square metre. In contrast, extractive sediment samples with a footprint usually $<0.25\ \mathrm{m}^2$ are generally placed several hundred metres apart. However, it is commonly the combination of the two (broad and fine-scale) that culminates in habitat maps. The latter discrete fine-scale samples are a reliable and necessary means of ground-truthing remote measurements. Visual techniques, such as digital photography and video, are also considered to work at fine scales (~ 1 m) and smaller scales. Non-extractive, image-yielding examples include investigations of Arctic habitat-forming epibenthic megabenthos (Piepenburg and Schmid, 1997) and organism–sediment relationships (Rhoads and Germano, 1982). Autonomous Underwater Vehicles (AUV) are increasingly used as carriers of high-resolution imaging sensors due to their ability to manoeuvre very close to potentially rugged terrain (Williams et al., 2010), thereby facilitating a constant image footprint. Images taken by an AUV provide two advantages: (1) the continuous photographic record yields intermediate-scale data, thereby bridging the gap between

* Corresponding author at: Institute for Marine and Antarctic Studies, University of Tasmania, Taroona, Tasmania 7053, Australia.
E-mail address: janseiler@gmx.de (J. Seiler).

0278-4343/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.csr.2012.06.003

MBES mapping and point sampling and (2) the image itself is an ideal candidate for automated data extraction.

Interrogation of digital imagery is necessary to extract qualitative and quantitative information. This task is usually carried out by a trained annotator. Whilst image capture takes only a fraction of a second, image annotation can take several minutes to tens of minutes depending on the nature and detail of information required. In fact, image interpretation and species identification is extremely time consuming and potentially subjective. Considering the various steps to produce a habitat map, annotating imagery, epitomises the proverbial bottleneck.

This study was conducted to expedite the lengthy and time-consuming process of image annotation by means of automation. Other efforts to automate the annotation process include the use of machine-learning algorithms to detect cold-water corals and sponges, as well as coverage enumeration after initial computer system training (Purser et al., 2009). It should be noted though, that this automation requires the computer system to be trained with a training set of images labelled by a human expert. This way, only a subset of the imagery is scored by a human expert and the remainder is scored (classified) by the computer system, usually with associated quantifiable error rate. Purser et al. (2009) report 45 min as the time taken to manually assess per cent coverage for dominant species (sponges and cold-water corals), where each image used 89 subsamples per image. After initial training, it took the computer system 22 s to accomplish the same task. Purser et al. (2009) used image texture features which numerically represent optical and structural attributes of corals and sponges.

Whilst Purser et al. (2009) quantifies the percentage of seabed covered by two organisms within an image, our study applies the machine-learning algorithm random forests (Breiman, 2001) to automate the process of assigning habitat classes to an entire image of the seafloor. The novelty in our approach is the use of geo-referenced stereo imagery from AUV mounted digital cameras to generate a centimetre-scale bathymetric reconstruction in the form of a triangulated irregular network. This results in a rugosity value for the overlapping footprint area of each image pair. Usually multiple features are required to describe a habitat comprehensively. We therefore used additional descriptors such as image texture (Local Binary Patterns, LBP), image colour (hue–saturation–values, HSV) and patchiness (Patch-Gap summaries, PG) to increase the accuracy of semi-automated habitat prediction. LBP and HSV are well-established methods in industrial machine vision applications (Ojala et al., 2002). In order to reliably employ these methods in an industrial setting, conditions such as lighting are constant and machine tasks are simple i.e. separating red and green apples. Applying the above-mentioned methods to imagery collected in the field with variable lighting regimes and complex machine tasks is a challenging proposition. Our study explores this challenge by testing the applicability of machine-learning algorithms to automate habitat classification in a practical application, using AUV derived images acquired on Tasmanian deep-water rocky reefs. Existing maps of Tasmania's inshore marine habitats are based on scientific echo sounder data and manually annotated video footage for ground-truthing and are restricted to depths <40 m (Barrett et al., 2001). With the exception of a multibeam sonar mapping trial in this region (Nichol et al., 2009) in which the AUV imagery was acquired as a means of ground-truthing, no other studies in this area exist. The study focused on highly complex rocky reef habitats below 40 m depth, which are difficult to efficiently sample using extractive methods such as Agassiz trawl or grab sampler. Due to the geology of Tasmania's south–east coast, our study site exemplifies deep-water rocky reef environments in this area.

The specific aim of this paper is to develop a novel analytical method to automate the process of assigning habitat classes to images of the seafloor, by automatically extracting colour, texture, rugosity and patchiness values from typical field acquired images and therefore curtail image processing time. To assess the success of this process, we evaluate the error rate of misclassifying images and sources of error. Two new processing techniques are developed to extract fine-scale bathymetry from stereo image pairs to calculate a common complexity measure, rugosity and extract fine-scale habitat distributions to calculate multivariate measures of 'patchiness'. We also discuss the relevance of this repeatable and cost-effective method to process the large volumes of image data needed to document the largely unknown fine-scale variability in habitat distributions.

2. Methods

2.1. Study area

The study area is situated immediately to the east of O'Hara Bluff, eastern Tasman Peninsula, Tasmania, Australia (Fig. 1). It forms part of the 'Peninsula Mapping Region' (Barrett et al., 2001) which has a dominantly easterly aspect, high vertical cliffs, deepwater reefs (to 100 m depth) and medium to high wave exposure. Geologically, the coastline is composed of dolerite, sedimentary rock and, to a lesser extent, granite (Barrett et al., 2001). This study uses data from a 4.6 km transect over the deep-water rocky reef of O'Hara Bluff and its offshore extension and transition zones between hard and soft substrate in 34–77 m depths. The traverse took just over 3 h (vehicle speed = 0.4 m/s).

2.2. Data acquisition

The Autonomous Underwater Vehicle (AUV) *Sirius* operated by the Australian Centre for Field Robotics at the University of Sydney sampled benthic habitats using a pair of downward-looking Pixelfly HiRes (1360×1024 pixels) digital cameras. Two strobes synchronously illuminated the field of view. The AUV was able to maintain a virtually constant altitude of 2 m above the seafloor, which equates to an image footprint of 1.6×1.3 m^2. Image acquisition at a 1 s interval with a speed over ground of $\sim$0.4 m/s provided an unbroken photographic record.

Sirius is a modified version of the SeaBED AUV (Singh et al., 2004a), built by the Woods Hole Oceanographic Institution, designed to be passively stable in pitch and roll. Yaw, forward and backward movement is controlled by a pair of aft-facing thrusters. Vertical (depth) movement of the positively buoyant vehicle is accomplished by one vertical thruster. Geographical vehicle positioning on the surface was accomplished using GPS. Navigation underwater is achieved using a Doppler velocity log, inertial measurement unit, ultra-short baseline acoustic positioning system, pressure sensor and a compass. To further reduce positional error introduced by dead-reckoning and sensor inaccuracies, the simultaneous localisation and mapping (SLAM) technique was used to re-navigate the estimated vehicle trajectories (Williams et al., 2008). Consequently, the intersecting survey pattern (Fig. 1 bottom panel) was necessary to maintain high spatial accuracy using SLAM.

2.3. Automated feature extraction

Colour, shape and texture features were used to characterise benthic habitats in each image. Stereo-photogrammetry was used to construct micro-topography for each stereo image pair to provide a measure of terrain complexity or 'rugosity' where the more complex surfaces had higher rugosity values. The three sets of features used were first and second order statistics of