3D Mapping of Benthic Habitat Using XGBoost and Structure from Motion Photogrammetry

S. Morsy 1,2 *, A. B. Yáñez Suárez 1,3, K. Robert 1

1 School of Ocean Technology, Fisheries and Marine Institute, Memorial University of Newfoundland, St. John’s, NL A1C 5R3, Canada - (Salem.Morsy, Ana-Belen.Yanez, Katleen.Robert)@mi.mun.ca
2 Public Works Department, Faculty of Engineering, Cairo University, 1 El Gamaa Street, Giza, Egypt, 12613
3 School of Fisheries, Fisheries and Marine Institute, Memorial University of Newfoundland, St. John’s, NL A1C 5R3, Canada

Commission II, WG II/7

KEYWORDS: 3D mapping, SfM, Geometric features, Machine learning, XGBoost.

ABSTRACT:

Benthic habitats mapping is essential to the management and conservation of marine ecosystems. The traditional methods of mapping benthic habitats, which involve multibeam data acquisition and manually collecting and annotating imagery data, are time-consuming. However, with technological advances, using machine learning (ML) algorithms with structure-from-motion (SfM) photogrammetry has become a promising approach for mapping benthic habitats accurately and at very high resolutions. This paper explores using SfM photogrammetry and extreme gradient boosting (XGBoost) classifier for benthic habitat 3D mapping of a vertical wall at the Charlie-Gibbs Fracture Zone in the North Atlantic Ocean. The classification workflow started with extracting frames from video footage. The SfM was then applied to reconstruct the 3D point cloud of the wall. Thereafter, nine geometric features were derived from the 3D point cloud geometry. The XGBoost classifier was then used to classify the vertical wall into rock, sponges, and corals (Case 1 - three classes). In addition, we separated the sponges class into three types of sponges: Demospongiae, Hexactinellida, and other Porifera (Case 2 - five classes). Moreover, we compared the results from XGBoost with the widely used ML classifier, random forest (RF). For Case 2, XGBoost achieved an overall accuracy (OA) of 74.45%, while RF achieved 73.10%. The OA improved by about 10% from both classifiers when the three types of sponges were combined into one class (Case 1). Results showed that the presented 3D mapping of benthic habitat has the potential to provide more detailed and accurate information about marine ecosystems.

1. INTRODUCTION

3D mapping of benthic habitats is critical in marine conservation and management. It involves creating accurate and detailed maps of the seafloor and the inhabiting organisms. These maps represent habitat types, and can be used to monitor changes over time, and inform management decisions. Traditionally, benthic habitat mapping has employed multibeam echosounder data acquisition (Brown et al., 2011; Trzcinska et al., 2020) and underwater imagery collection and annotation (Keogh et al., 2022; Mohamed et al., 2022).

The underwater imagery represents high-resolution ground-truthing data characterising the seafloor from which biological information can be extracted. The manual annotation of enormous underwater image datasets is tedious, error-prone, and time-consuming (Mahmood et al., 2020). However, the annotation of some of these images can be used to train machine learning (ML) algorithms to build full-coverage maps representing the composition of the seafloor (Brown et al., 2011). Although recent studies have considered underwater image classification using ML (e.g., Mohamed et al., 2020; Ternon et al., 2022), they lose the advantage of considering the 3D geometry of underwater habitats in the classification process.

3D point clouds provide spatial information about the seafloor, including depth, shape, and texture. One approach for mapping benthic habitats in 3D is using structure-from-motion (SfM) photogrammetry (Price et al., 2019; Bayley et al., 2020). SfM photogrammetry is a technique for creating 3D models of objects or environments based on multiple 2D images. It involves taking overlapping images of the seafloor from different angles and using software to stitch them together to create a 3D model. This can produce a more detailed and accurate representation of the seafloor at high-resolution compared to images.

One advantage of SfM-generated 3D points over other types of 3D data (e.g., bathymetric LiDAR) is that RGB values of points can be added to the geometric features traditionally employed to build full coverage classification maps. Geometric features (e.g., linearity, planarity, scattering) are based on eigenvalues and eigenvectors of neighbourhood points and are provided as input to ML classifiers (Chehata et al., 2009; Weinmann et al., 2015; Morsy and Shaker, 2022) such as XGBoost (Ghatkar et al., 2019; Nemani et al., 2022).

Extreme Gradient Boosting (XGBoost) is an ML algorithm that uses decision trees to classify data. It is particularly effective at handling large, complex datasets and has been used successfully in various applications for images and 3D point cloud classification (Zhang et al., 2021). By training the algorithm on a dataset produced from SfM photogrammetry and corresponding ground-truth data, it is possible to create accurate maps of benthic habitats.

One of the key advantages of using XGBoost and SfM photogrammetry for 3D mapping of benthic habitats is that it provides high-resolution maps that can be used to identify...
small-scale features such as coral reefs, seagrass beds, and rocky outcrops. These maps can identify different habitat types, monitor their changes over time, and evaluate management strategies.

2. RELATED WORK

Underwater video surveys have recently gained significant interest for benthic habitat identification and classification (Keogh et al., 2022; Mohamed et al., 2022; Ternon et al., 2022). Videos are usually recorded using high-definition (HD) cameras mounted on remotely operated vehicles (ROVs) (Robert et al., 2017; Keogh et al., 2022). These videos, in turn, are converted into 2D images to generate 3D point clouds using SFM photogrammetry (Casella et al., 2017; Price et al., 2021; Ventura et al., 2022). These point clouds are further classified for benthic habitat classification using different ML algorithms (Pierce et al., 2021; De Oliveira et al., 2021; De Oliveira et al., 2022).

So far, very few studies have considered using 3D point clouds in benthic habitat mapping. For instance, Pierce et al. (2021) labelled 10,000 image patches into seven classes of a coral reef area near Cheeca Rocks in the Florida Keys National Marine Sanctuary. Then, they assigned those labels to either the point cloud or mesh during the SFM process to create fully reference versions for image-based or point-based classification. A convolutional neural network was used as an image-based classification method with different thresholds and reached 94.10% overall accuracy. In addition, two point-based classification methods were evaluated, namely fast multilevel semantic segmentation and the fully convolutional network. They achieved an overall accuracy of 88.50% and 90.00%, respectively. Ternon et al. (2022) used SFM to generate point clouds from underwater images and then create a DSM and RGB Ortho-mosaic image. Six spatial predictors were extracted from the DSM, including slope, aspect, profile convexity, plan convexity, maximum curvature, and change rate of the bathymetry. These predictors were combined with the DSM and the RGB Ortho-mosaic image, representing input layers for a maximum likelihood classifier (MLC). Two rocky reef sites at St Malo Bay in Brittany, France, were used for evaluation. The MLC achieved an average overall accuracy of 82.20% for eight classes.

In the study of De Oliveira et al. (2021), they evaluated two 3D point-based classification methods for cold-water coral reef identification in the southwest of Ireland. The first was Support Vector Machines (SVM) which used dimensionality as a parameter for classification. Six datasets in the study area were tested for coral separation from the seabed. The highest, lowest, and average overall accuracies were 90.00%, 46.40%, and 68.20%, respectively. The second method was based on eight geometrical features derived from the point cloud. Then, a Gradient Boosting Trees (GBT) algorithm was applied to label each point as coral or seabed. The same datasets were tested and demonstrated 94.60%, 9.30%, and 68.00% as the highest, lowest, and average overall accuracies, respectively.

Moving on, De Oliveira et al. (2022) used SFM-derived 3D point cloud and ML algorithms for 3D mapping of cold-water coral reefs. They evaluated six ML algorithms, namely, SVM, Random Forest (RF), GBT, k-Nearest Neighbours (KNN), Logistic Regression, and Multilayer Perceptron (MLP). In order to evaluate accuracy variation between ML algorithms, they trained them based on different sample sizes (i.e., 1,000 samples and 10,000 samples) with different parameters. As a result, eighteen models of ML were created and tested. The Piddington Mound area, located in the southwest of Ireland, was used for 3D point cloud reconstruction from HD video data acquired with an ROV. The 3D point clouds were classified into four classes: live coral, dead coral, coral rubble, and sediment and dropstones. The results showed that four models yielded F1-scores of more than 90.00% and could distinguish between the four classes. The highest model was the GBT classifier, with an average F1-score of 95.10%, followed by RF, MLP, and KNN, with an average F1-score of 94.20%, 92.30%, and 91.60%, respectively. It should be noted that the four models were trained on 10,000 samples. Thus, increasing the training samples has improved the classification accuracy significantly.

3. STUDY AREA

The study area of this research is the Charlie-Gibbs Fracture Zone (CGFZ), which is located in the North Atlantic Ocean (Figure 1). It is roughly halfway between Greenland and the Azores and extends from the Mid-Atlantic Ridge to the Wyville-Thomson Ridge. The area is a major pathway for the flow of deep water in the Atlantic, which brings nutrients to the surface. The fracture zone is also associated with several seamounts and other undersea features, which are essential habitats for a diverse range of marine organisms such as cold-water corals and sponges, which provide habitat for several associated species.

Figure 1. The Charlie-Gibbs Fracture Zone and study area locations. The TOSCA survey area is marked with the red box (upper panel), and the location of Dive 9 with the vertical wall is highlighted in green (lower panel).
3.1 Data Acquisition

In June 2018, the CGFZ was surveyed during the Tectonic Ocean Spreading of the Charlie-Gibbs Fracture Zone’s (TOSCA) expedition on-board the research vessel Celtic Explorer (CE18008). A HD oblique-facing camera, Kongsberg Maritime OE14-502a HDTV, was mounted on the ROV Holland I to record videos (1080 resolution at 25 frames per second) of benthic habitats. The position of the ROV was continuously recorded using Ultra Short Baseline (USBL) systems (IXSEA GAPS USBL and Sonardyne Ranger 2 USBL).

3.2 Data Processing

For this paper, we extracted images from a vertical wall from Dive 9 (Figure 1). Dive 9 exhibited a high abundance of Demospongiae and Hexactinellida sponges with a scatter of corals from the order Scleractiniana. A total of 135 images were extracted from the HD videos at a rate of one frame per second using Blender 2.92 software. The coordinates of each frame were obtained from the USBL data and exported as CSV files. The images and coordinates files were imported into Agisoft Metashape 1.6.1 software, where the coordinates of the images were projected to UTM zone 25N. The 3D point clouds were then reconstructed and exported as XYZ files (Figure 2). The vertical wall has 17,264,337 points in total, with a dimension of 3.45 x 4.0 m and water depths ranging from 1,874 m to 1,889 m.

Points for three categories (i.e., sponges, corals, and rock) were manually annotated and geotagged within the point clouds using the Agisoft Metashape software (Figure 3). Approximately, 15% (2,648,414) of the total points were labelled as reference data (i.e., ground truth). Table 1 provides the class breakdown of the reference data for the vertical wall.

<table>
<thead>
<tr>
<th>Class</th>
<th>Number of points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>1,466,955</td>
</tr>
<tr>
<td>Demospongiae</td>
<td>363,992</td>
</tr>
</tbody>
</table>

Table 1: Breakdown of the reference data.

4. METHODOLOGY

After the point cloud reconstruction and reference data annotation, nine geometric features were derived from the 3D point cloud geometry (i.e., coordinates) based on a spherical neighbourhood (Mohamed et al., 2021). These features included eight covariance features and verticality, as listed in Table 2. XGBoost was then used to classify the vertical wall into rock, sponges, and corals (Case 1). The sponges class was further separated into three types: Demospongiae, Hexactinellida, and other Porifera. Then, we evaluated the XGBoost performance for classifying the point cloud of the vertical wall into five categories rock, corals, and three types of sponges (Case 2). Moreover, we compared the results from the XGBoost classifier with the widely used ML classifier, RF. The resulting accuracy was assessed using the overall accuracy (OA), precision, recall, and F1-score.

4.1 Geometric Features Extraction

Based on the spatial distribution of the 3D points within the local neighbourhood, the respective 3D covariance matrix was calculated for each 3D point (Jutzi and Gross, 2009). The eigenvalues of the covariance matrix were directly used to describe the local 3D structure or derive features that express unique geometric properties (Mallet et al., 2011). For describing the local dimensionality, the features of linearity (L), planarity (P), and scattering (S) provided information about the presence of a linear 1D structure, a planar 2D structure, or a volumetric 3D structure. Further measures were provided by omnivariance (O), anisotropy (A), eigenentropy (E), the change of curvature (C), and the sum of eigenvalues (S). In addition, the verticality (V) feature was considered, which was derived from the normal vector's vertical component (Demantké et al., 2012).

4.2 Machine Learning Classifiers

The XGBoost was applied with its standard parameters settings in the XGBoost python library, such as the number of trees (n_estimators=100) and features during a fit (n_features_in_=4). The objective function was set to account for multiclass labelling (objective=multi:softprob). The RF parameters were set to default as in the scikit-learn library, with the number of trees set to 300 and the class weight to “balanced” to consider
the percentage of classes in the training dataset while creating the model.

<table>
<thead>
<tr>
<th>Covariance Features</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity</td>
<td>$L = \frac{\lambda_1 - \lambda_2}{\lambda_1} \cdot \frac{\lambda_1}{\lambda_3}$</td>
</tr>
<tr>
<td>Planarity</td>
<td>$P = \frac{\lambda_2 - \lambda_3}{\lambda_1}$</td>
</tr>
<tr>
<td>Scattering</td>
<td>$S = \frac{\lambda_3}{\lambda_1}$</td>
</tr>
<tr>
<td>Omnivariance</td>
<td>$O = \frac{3}{\lambda_1 \lambda_2 \lambda_3}$</td>
</tr>
<tr>
<td>Anisotropy</td>
<td>$A = \frac{\lambda_1 - \lambda_2}{\lambda_3}$</td>
</tr>
<tr>
<td>Eigen entropy</td>
<td>$E = -\sum_{i=1}^{3} \lambda_i \ln(\lambda_i)$</td>
</tr>
<tr>
<td>Change of curvature</td>
<td>$C = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}$</td>
</tr>
<tr>
<td>Sum</td>
<td>$\sum = \lambda_1 + \lambda_2 + \lambda_3$</td>
</tr>
<tr>
<td>Verticality</td>
<td>$V = 1 - ((0,0,1), v_3)$</td>
</tr>
</tbody>
</table>

Table 2. Geometric features’ mathematical expressions.

where $\lambda_1, \lambda_2, \lambda_3 =$ eigenvalues  
v_3 = the third eigenvector

4.3 Training and Testing Datasets

Previous studies have randomly selected ML classifiers’ training/testing datasets (Zavalas et al., 2014; Mohamed et al., 2018; Mohamed et al., 2020; Letard et al., 2021). However, this provided biased results because of the high spatial correlation between training and testing datasets. Therefore, this research divided the reference data into 75% for training and 25% for testing, and all training points within a sphere of a 5 cm radius around the testing points were removed to limit the effect of the spatial correlation (Letard et al., 2022). The training dataset was used to train the XBoost and RF models, and the testing dataset was used to evaluate the models.

The data division, point cloud classification, and accuracy assessment were implemented using Python programming language, mainly scikit-learn and XGBoost libraries, on a Dell machine with Intel® Xeon® W-2123 processor, 3.60 GHz, and 32 GB RAM. The results were visualized using the CloudCompare software.

5. RESULTS AND DISCUSSION

The classified point cloud from XGBoost and RF for Case 1 (three classes) and Case 2 (five classes) are shown in Figures 4 and 5, respectively. The distribution of different classes based on visual interpretation shows that dense sponge aggregations are located in the upper and lower areas of the vertical wall, with sparse corals observed.

XGBoost achieved an OA of 74.45%, while that of RF was 73.10%. The OA improved by ~10% when the three types of sponges were combined into one class using both classifiers, where XGBoost and RF demonstrated an OA of 84.35% and 83.46%, respectively. Although RF achieved a close OA to XGBoost, the F1-score from XGBoost was superior for all individual classes (Tables 3 and 4). It should be noted for Case 1 that the most frequent classes represented in the reference data have the highest classification scores (Rock, Sponges, and Corals, respectively). This was not observed for Case 2 with the separation of the Sponges class, where Rock had the highest F1-score, followed by Corals, then the different types of sponges.

Table 3. F1-scores (%) of the individual classes for Case 1 (three classes).

<table>
<thead>
<tr>
<th>Class</th>
<th>RF</th>
<th>XGBoost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>87.33</td>
<td>87.18</td>
</tr>
<tr>
<td>Sponges</td>
<td>80.54</td>
<td>81.11</td>
</tr>
<tr>
<td>Corals</td>
<td>66.39</td>
<td>80.78</td>
</tr>
<tr>
<td>Average</td>
<td>78.09</td>
<td>83.03</td>
</tr>
</tbody>
</table>

Table 4. F1-scores (%) of the individual classes for Case 2 (five classes).

<table>
<thead>
<tr>
<th>Class</th>
<th>RF</th>
<th>XGBoost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>88.86</td>
<td>89.21</td>
</tr>
<tr>
<td>Demospongiae</td>
<td>40.62</td>
<td>43.79</td>
</tr>
<tr>
<td>Hexactinellid</td>
<td>63.69</td>
<td>64.70</td>
</tr>
<tr>
<td>Other Porifera</td>
<td>11.69</td>
<td>16.23</td>
</tr>
<tr>
<td>Corals</td>
<td>70.46</td>
<td>81.90</td>
</tr>
<tr>
<td>Average</td>
<td>55.06</td>
<td>59.16</td>
</tr>
</tbody>
</table>

XGBoost and RF are both ensemble methods that fit several decision trees on various sub-samples of the dataset. Although both could effectively predict the composition of the seafloor,
the effectiveness of these classifiers depends on the quality and quantity of the input data and the specific characteristics of the habitat being studied.

There are some challenges associated with using this technique. For example, it is not easy to collect data in areas with strong currents. It can also be challenging to process large datasets of SfM photogrammetry data, particularly if the data is collected over a large area. In addition, if there are few labelled datasets available for training ML models. Despite these challenges, using XGBoost and SfM photogrammetry for 3D mapping benthic habitats is a promising approach that can improve our understanding of marine ecosystems and inform management decisions. As technology improves and datasets become more extensive and comprehensive, this technique will likely become an increasingly important tool for marine scientists and conservationists.

6. CONCLUSIONS

This paper explored using SfM photogrammetry and XGBoost classifier for benthic habitat 3D mapping at very high resolutions. The combination of XGBoost and SfM photogrammetry allows for the collection of large amounts of data in a relatively short time, which can be used to create high-resolution maps of benthic habitats. It also reduces the need for manual data processing, which is time-consuming. Moreover, it allows for identifying different types of benthic classes based on their 3D structure, providing more detailed information than traditional methods that rely on 2D images. By using these tools to identify critical habitats, monitor changes over time, evaluate management strategies, and educate the public, we can work towards protecting and conserving marine ecosystems for future generations.

ACKNOWLEDGEMENTS

We would like to thank the deck and scientific crew of the TOSCA survey and ROV Holland 1 aboard the RV Celtic Explorer. Funding for the analysis came from a Canada Research Chair in Ocean Mapping as well as the Ocean Frontier Institute funded Benthic Ecosystem Mapping & Engagement (BECoME) project.

REFERENCES


