PAPER

Automated Detection of Underwater Military Munitions Using Fusion of 2D and 2.5D Features From Optical Imagery

AUTHORS

A.S.M. Shihavuddin
Nuno Gracias
Rafael Garcia
Ricard Campos
Computer Vision and Robotics Group (VICOROB), University of Girona, Spain

Arthur C.R. Gleason
Department of Physics, University of Miami

Brooke Gintert
Department of Marine Geology and Geophysics, University of Miami

ABSTRACT

Technologies that can efficiently and objectively detect, identify, and map underwater military munitions are needed. The knowledge of benthic environments adjacent to underwater military munitions is crucial for remediation decisions. When attempting to identify munitions from optical imagery, tridimensional structure information obtained from the surveyed area can complement the texture information that is available in the images. In this work, we use a fusion of two-dimensional (2D) and two-and-a-half-dimensional (2.5D) features to classify munitions on the seabed from a sequence of images of an optical survey of the seabed. The 2D features respond to texture, whereas the 2.5D features respond to geometry. The 2.5D features used were coefficients of polynomial surface fitting, standard deviation, skewness, and kurtosis of the elevation, slope of principal plane, mean and standard deviation of the distance of 2.5D points to the principal plane, surface normal, curvatures, rugosity and symmetry measures. Adding the 2.5D features increased classification accuracy relative to using only 2D features when detecting discarded military munitions.

Keywords: discarded military munitions, 2.5D features, elevation map, underwater object classification, fusion of 2D and 2.5D features

Introduction

Locating and identifying navigation and safety hazards related to underwater military munitions (UWMM) is a broad and diverse problem, because historical information relating to the locations of underwater munitions is often limited and not always accurate. Furthermore, some munitions were dispersed over large areas, being dumped by vessels, dropped via aircraft, or shot as projectiles on or adjacent to live target ranges.

Currently, both wide-area searches and detailed mapping of UWMMs rely on multiple acoustic (e.g., side-scan or multibeam sonar) and/or metal detection methods. Underwater optical images of the seabed are another technology that could benefit efforts to find and characterize UWMM, because relative to the acoustic, magnetic, and electromagnetic mapping technologies, optical images have very high spatial resolution. Objects on the seabed, as small as a few centimeters in diameter, can be easily resolved with optical imagery.

The strength of underwater images of the seabed is their high spatial resolution, which is 1–2 orders of magnitude greater than other survey technologies. The weakness of these images, historically, has been the labor involved to extract quantitative information from them. Thus, optical imagery is currently used only as an inspection tool and not as a survey methodology for UWMM response (Schwartz & Brandenburg, 2009). Other communities, such as benthic ecologists, do use seabed images quantitatively, but the quantitative information is extracted using laborious, manual methods such as point counting (Kohler & Gill, 2006).

An automated process for classifying underwater optical imagery would reduce the analysis bottleneck and allow large underwater image datasets to be fully exploited for UWMM surveys. Efforts to automate underwater image classification have been made for well over a decade (see Literature Review below), but no single technique is yet widely accepted as robust. Our group has recently developed a new seabed classification algorithm that has been shown to accurately classify benthic images from coral reefs (Shihavuddin et al., 2013). Here, we tested two questions: First, can the

Shihavuddin et al. algorithm distinguish UWMM from other natural components of the seabed? Second, does the use of 2.5D (a quasi-three-dimensional system that uses only a single elevation value for any surface at any given location) features improve the discrimination of UWMM from the background relative to the use of only two-dimensional (2D) features, as currently implemented by Shihavuddin (2013).

For both natural and man-made objects, 2.5D cues provide important discriminative information for classification. To extract 2.5D features, it is necessary to generate a 3D texture model of the surface of the seabed and thereafter calculating a corresponding elevation map. The 2.5D features are calculated directly from this elevation map.

Literature Review

For the cases where the survey images contain enough overlap to allow the extraction of depth information, then 2.5D- or even 3D-based features can be important as additional features for classification. In the recent works by Pican et al. (1998), Pizarro et al. (2008), Marcos et al. (2005), Stokes and Deane (2009), Beijbom et al. (2012), Bewley et al. (2012), and Shihavuddin (2013), the issue of automated classification of underwater objects using texture information from optical imagery has been addressed. Separately, the work of Friedman et al. (2010) illustrated the importance of structural features, such as rugosity, for the classification of bathymetric maps. However, no previous work has addressed the combination of 2D texture information with 3D or 2.5D features for classification of underwater imagery.

Several works have used 3D features for classification in general computer vision applications. Knopp et al. (2010) used the Hough transform and 3D surf to generate 3D features for classifying household objects. Sandbach et al. (2012) used predefined 3D shape models for facial expression detection. Marton et al. (2011) proposed the combined use of 2D and 3D features for online robot motion applications. All of these applications were attempting to classify objects that had regular, predefined shape models such as the human body or a table. In contrast, here we were interested in classifying both man-made objects with regular, well-defined shapes as well as natural benthic objects that did not have a clearly defined shape model.

For classifying natural benthic objects, 2.5D features that can identify the characteristics of surface geometry were considered in this work. Paul and Ramesh (1986) demonstrated that mean curvature and Gaussian curvature were the fundamental second-order surface characteristics that possessed desirable invariance properties of extrinsic and intrinsic surface geometry respectively. Hetzel et al. (2001) used pixel depth, surface normal, and curvature for 3D object recognition from range images. Wang et al. (2000) used Gaussian curvature, mean surface curvature, and principal curvatures to define 3D correspondences between brain images. These features were robust to viewpoint changes and contained discriminant information.

The St. John BIOMapper (Daniel et al., 2012), which is a geographic information system tool for mapping benthic habitats, used statistics such as curvature, plan curvature, profile curvature, mean depth, variance of depth, surface rugosity, slope, and slope of the slope to characterize the complexity of the surface of the seafloor. Although not used in this paper, some of these features could be considered as potential 3D or 2.5D features for underwater object description.

Friedman et al. (2012) presented a method for calculating the rugosity, slope, and aspect features of the Delaunay triangulated surface mesh of the seabed terrain by projecting areas onto the plane of best fit using principal component analysis (PCA). They used a set of 3D statistics (rugosity, slope, and aspect) to define the characteristics of the seabed terrain for scientific ecological applications. Friedman et al. (2010) used these same statistics derived from fine-scale bathymetric reconstructions (created using georeferenced stereo imagery) for classification. Given their generality, we also use these 3D statistics (Friedman et al., 2010) as features for classification, but in the context of munitions detection.

In this work, we combined a suite of 2.5D features with the state-of-the-art 2D features for classification. The 2D features and general classification algorithm used the proposed framework by Shihavuddin et al. (2013). The 2.5D features were polynomial surface coefficients, elevation statistics, slope of surface (Friedman et al., 2010), curvature (Wang et al., 2000), surface normal (Hetzel et al., 2001), rugosity (Friedman et al., 2010), and symmetry (Kovesi, 1997) extracted from the elevation map. The following sections describe the dataset, the generation of the elevation map, the proposed 2.5D features, and the obtained results followed by concluding remarks.

Ordnance Reef Dataset

The Ordnance Reef dataset was collected by the National Oceanic

and Atmospheric Administration at the "Ordnance Reef" site off of Waianae, Hawaii, by divers using a hand-held high-definition video camera. Divers swam a series of parallel transects over an area of approximately 12 × 12 m collecting video with the camera pointed vertically down at the seabed. During processing, individual frames were grabbed from the video stream and subsampled by removing every other scan line to remove interlacing effects because the camera, a Sony HDR-XR520V, did not have progressive scan capability. The resulting frames were then composited into a mosaic using the technique described by Lirman et al. (2007). The resolution of the Ordnance Reef mosaic (Figure 1) was 7,949 × 8,444 pixels at .635 mm per pixel.

The Ordnance Reef mosaic was used to test the performance of the 2.5D features. This dataset had mainly two main classes: 5-inch shells and background. There are 9,899 patches of background class and 782 patches of 5-inch shell class in the Ordnance Reef dataset (randomly extracted from the Ordnance Reef mosaic). The size of the patch is 128 by 128 pixels which is approximately 66.1 cm^2 area. 50% of these patches were used for training and 50% of them were used for testing. The reason the Ordnance Reef dataset was used for this test is that, although the 5-inch shells tended to have a distinctive 3D shape, which presumably would translate to distinctive 2.5D features, there was also coral growth over the shells, thereby resulting in similar textures between the shells and the background.

Elevation Map Generation

Our approach to creating 3D textured models of the seabed using a structure from motion technique from a moving platform is an extension of existing 2D (Gracias et al., 2003; Lirman et al., 2007) and 3D mosaicing algorithm (Nicosevici et al., 2010). The estimation of 3D structure began by identifying common points across pairs of images using a variant of the Scale Invariant Feature Transform (Lowe, 2004). Next, a robust sampling technique was used to identify and eliminate false point matches among images, resulting in multiple image observations of the same set of 3D points (Fischler & Bolles, 1981). Finally, sparse bundle adjustment (Triggs et al., 2000) was used to obtain a set of 3D points that best complied with the image observations. Figure 2 illustrates an example of 3D reconstruction of a portion of the Ordnance dataset.

The geometry 2.5D features were extracted from an elevation map of the surface, which was created from the 3D texture model. In order to create an elevation map, a fixed reference plane was needed, where the unconstrained geometry of the reconstructed 3D surface could be projected orthogonally. The center of the plane was found using the centroid of the vertices of the calculated triangles in 3D reconstruction with a least squares fitting. The normal of this reference plane was computed using PCA regression. Starting with the covariance matrix constructed from the vertices of the surface triangular mesh, we estimated the normal by computing its third smallest eigenvector (Bair et al., 2006).

For each vertex in the surface, we then defined a height value with respect to this initial plane. In this plane, we then sampled an image using a regular grid of the user defined size and resolution. For each of these discrete positions to sample, we generated a ray passing through the sample point and perpendicular to the plane. Figure 3 illustrates the steps of creating elevation map.

Due to the large number of queries, we speed up the ray-surface intersection test using axis-aligned bounding

FIGURE 1

Mosaic image with corresponding elevation map created from Ordnance dataset. (a) Ordnance mosaic (b) Elevation map of Ordnance mosaic. The resolution of the mosaic is 7,949 × 8,444 pixels at 0.635 mm per pixel. Each mosaic encompasses an area of 5.4 by 5.4 meter squares.

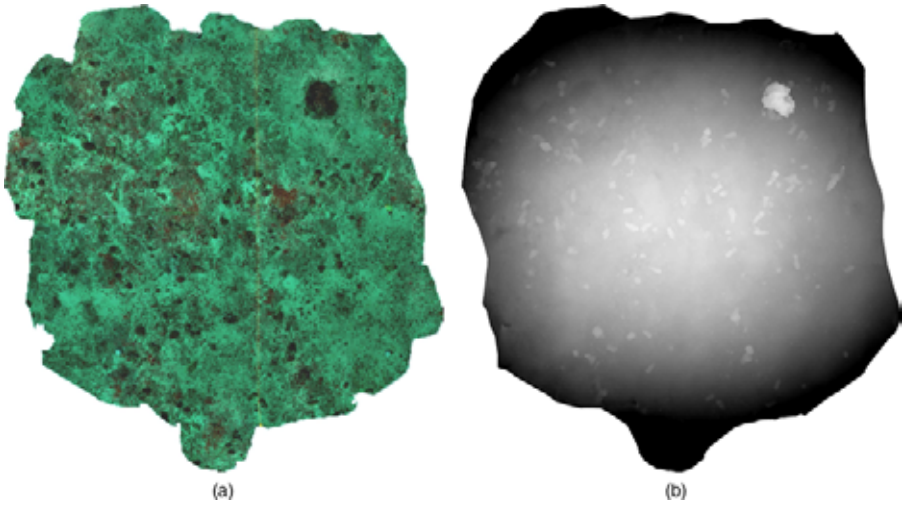
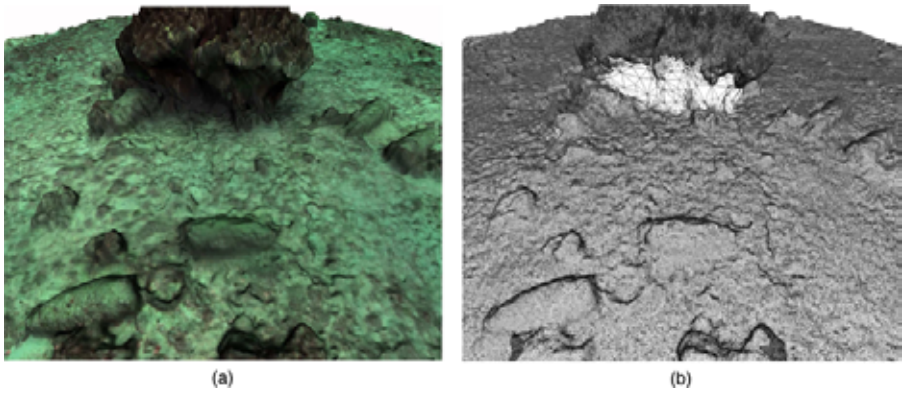


FIGURE 2

3D reconstruction of a portion of the Ordnance Reef dataset displayed as a textured surface (a) and as a triangular mesh (b). The scale varies in this oblique view with the lower (closer) edge of the image corresponding to ~2 m on the seabed.



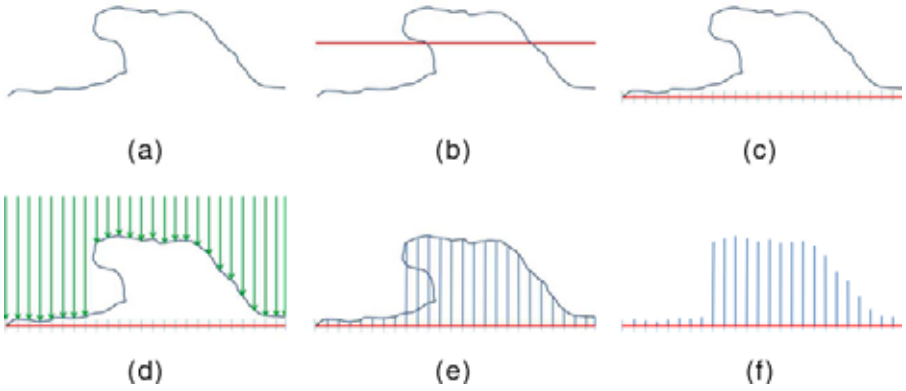
boxes (AABB; Zuiderveld, 1994). The AABB tree (Pierre et al., 2013) provided a static data structure that made it easier to perform efficient intersection and distance queries on sets of finite 3D geometric objects. After finding an intersection point for each grid sample, the height was computed relative to the distance from the reference plane. Note that the ray intersection queries were computed from top to bottom, while the computation of the height value is from bottom to top (Figure 3).

2.5D Features Formulation

Our method used a patch-based classification framework, where a sliding window was used to sample "patches" around regularly spaced center pixels and classify them based on available 2D and 2.5D features. The size of the patches was dictated by the sizes of the important object classes. Generally, the size of the patch should be as small as possible while still being large enough to encapsulate distinct features of an object.

FIGURE 3

Step-by-step illustration of the process of creating the 2.5D elevation map from the 3D mesh. (a) Input 3D mesh (b) PCA (c) Reference + image sampling (d) Ray-intersection queries (e) Height computation (f) Final result.



For the Ordnance Reef dataset, we chose a patch size of 128 × 128 pixels for both 2D and 2.5D features.

Coefficients of Polynomial Surface Fitting (Keren & Gotsman, 1999)

The elevation map of a patch was approximated as a 2D polynomial surface modeled by a polynomial equation of third degree (Keren & Gotsman, 1999). Figure 4 shows the example of an estimated polynomial surface of a patch from the Ordnance Reef dataset.

$$f = p_1 + p_2x + p_3y + p_4x^2 + p_5xy + p_6y^2 + p_7x^2y + p_8xy^2 + p_9y^3$$

The coefficients of the polynomials were extracted through least square fitting and used as features. Each polynomial in each patch provided nine 2.5D features.

Standard Deviation, Skewness, and Kurtosis of the Elevation (Hetzel et al., 2001)

From the elevation map, we calculated the standard deviation, skewness, and kurtosis using the altitude values of each pixel in the patch, resulting in three additional 2.5D features. For calculating skewness and kurtosis, the following equations were used where z_i represents the elevation of a 2.5D data point, z_m and S are the mean and the standard deviation of the 2.5D points, and N is the number of data points.

$$S = \frac{\sum_{i=1}^{N} (z_i - z_m)^2}{N}$$

$$Skewness = \frac{\sum_{i=1}^{N} (z_i - z_m)^3}{(N-1)S^3}$$

$$Kurtosis = \frac{\sum_{i=1}^{N} (z_i - z_m)^4}{(N-1)S^4}$$

Slope of the Principal Plane of the Patch (Friedman et al., 2010)

A plane in 3D space was fitted as the principal plane of each patch as described above, in the elevation map generation section. The angle of this plane with respect to the ground plane (the plane where the elevation is fixed at 0 value) was used as another feature.

Distance of 3D Points to the Principal Plane (Friedman et al., 2010)

Two more 2.5D features were given by the mean and standard deviation of the distance of the points on the elevation map to the principal plane. These distances were measured from the points on the elevation map to the corresponding projected points on the principal plane of the surface.

Surface Normal (Hetzel et al., 2001)

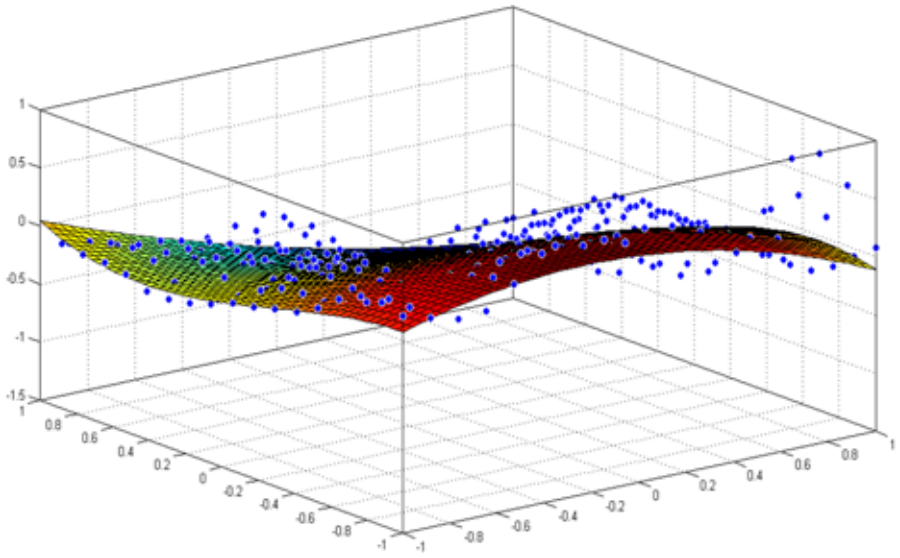
The surface normal of the 3D surface was defined as the slope and aspect of each facet within the analysis patch. The surface normal for each pixel in the 2.5D elevation map was calculated using normalized first derivatives along x, y, and z termed as n_x, n_y, and n_z, respectively. The pair of angles (α, β) of the surface normal was calculated using the following two equations (Hetzel et al., 2001). From each patch, taking the mean and standard deviation of the surface normal at all pixels generated four 2.5D features for classification.

$$\alpha = \arctan\left(\frac{n_z}{n_y}\right)$$

$$\beta = \arctan\frac{\sqrt{\left(n_y^2 + n_z^2\right)}}{n_z}$$

FIGURE 4

Example of polynomial surface fitting of a patch from Ordnance Reef dataset using sparse points.



Curvatures (Wang et al., 2000)

The Gaussian curvature (G) and the mean surface curvature (M) were calculated from the partial derivatives of the elevation map (E) at each pixel. Here $\Omega = \{(x,y,z), (y,z,x), (z,x,y)\}$ is the set of circular shifts of (x,y,z).

$$G = \frac{\sum_{(x,y,z)\in\Omega}\left[E_x^2\left(E_{yy}E_{zz} - E_{yz}^2\right) + 2E_xE_y\left(E_{xz}E_{yz} - E_{xy}E_{zz}\right)\right]}{\left(E_x^2 + E_y^2 + E_z^2\right)^2}$$

$$M = \frac{\sum_{(x,y,z)\in\Omega}\left[E_z^2\left(E_{xx} + E_{yy}\right) - 2E_xE_yE_{xy}\right]}{2\left(E_x^2 + E_y^2 + E_z^2\right)^{1.5}}$$

Two principal curvatures (k_1 and k_2) were extracted using G and M. Also, from k_1 and k_2, a shape index (S), which represents the type of curvature, and a curvedness index (C), which represents the degree of curvature, were calculated. The following four equations were used for calculating k_1 and k_2, S and C (Wang et al., 2000). Using the mean and the standard deviation of G, M, k_1, k_2, S and C, twelve 2.5D features were extracted from the elevation map representing surface curvature characteristics for each patch.

$$k_1 = M + \sqrt{M^2 - G}$$

$$k_2 = M - \sqrt{M^2 - G}$$

$$S = \frac{2}{\pi}\arctan\left[(k_2 + k_1)/(k_2 - k_1)\right]$$

$$C = \sqrt{(k_1^2 + k_2^2)/2}$$

Rugosity (Friedman et al., 2010)

Rugosity is commonly used in marine science to characterize seafloor habitats. It relates to small-scale amplitude variations in the height of a surface and is considered as an important 2.5D or 3D feature (Lundblad et al., 2004). Rugosity, r, of a patch was calculated by dividing the area, A_s, of the contoured surface bounded by the patch, by the area, A_p, of the orthogonal projection of the surface onto the principal patch plane. Figure 5 illustrates an example of surface area and area on the principal plane of the patch.

The area of the surface A_s was calculated using the Friedman's method (Friedman et al., 2010). In this method, the surface was considered as a set of triangles created by sampling points. By calculating the area of each triangle and summing up all, the total area of the surface was calculated. The principal plane was considered a rectangular plane in 3D space where four vertices were the four points at each corner of the patch. Using these four 3D points, the area of the principal plane of the patch was calculated. This principal plane decouples the slope effect, which otherwise might result in a higher rugosity value on a slope containing a relatively smooth surface.

$$Rugosity = \frac{A_s}{A_p}$$

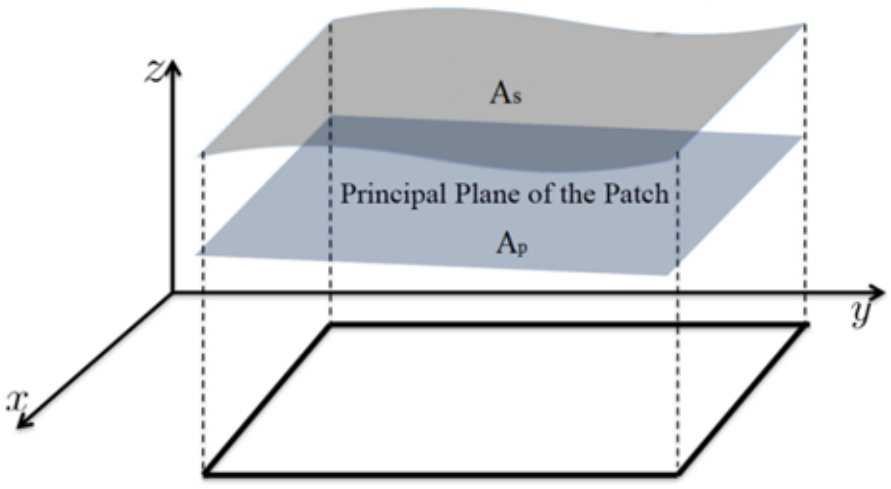
Symmetry (Kovesi, 1997)

Symmetry is an effective indicator to detect the structure of man-made objects and certain living structures (Kovesi, 1997). Man-made objects usually contain higher recognizable symmetry or partial symmetries. The symmetry image was created using the method by Kovesi (1997). In this method, a Log Gabor filter (David, 1987) was used to measure the phase symmetry at each pixel of the patch. The mean and variance of a symmetry image of a patch of elevation map produced two 2.5D features. The phase symmetry was notably different between munitions and background when applied to patches from the elevation map of the Ordnance Reef dataset (Figure 6).

Combining all of the 2.5D features created a feature vector of 34 components. Table 3 summarizes the features used in our method.

FIGURE 5

Example of surface area, A_s, and area on the principal plane, A_p, of the patch.



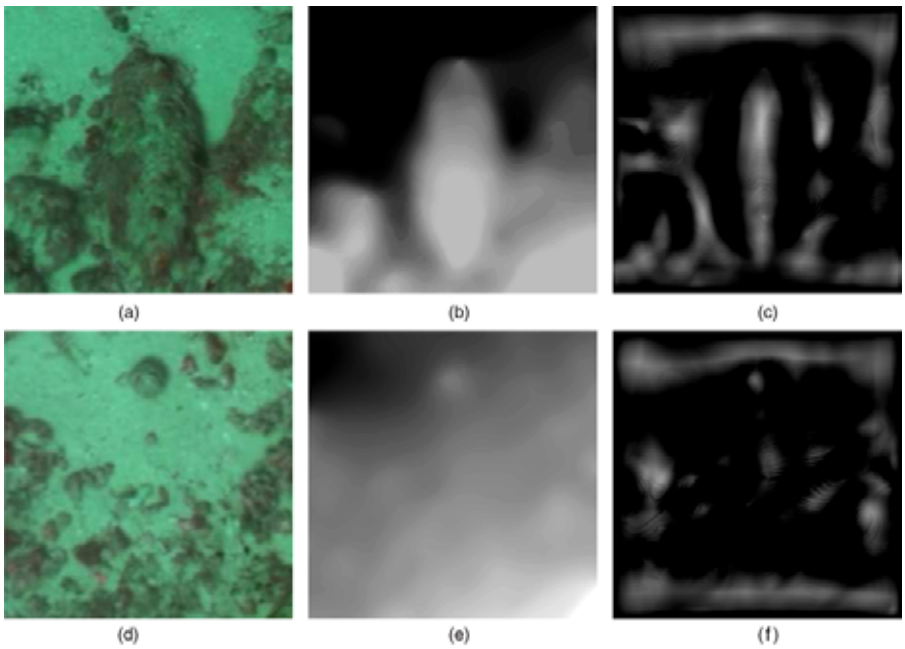
Implementation

The classification framework is implemented to perform thematic mapping on the Ordnance mosaic and is tested on a part of Ordnance dataset used as the validation set (50% of the dataset). Thematic mapping was defined as the segmentation of an image (in this case, the Ordnance Reef mosaic) into several different classes that keep a spatial coherence in order to obtain a diagrammatic representation of an area on the sea floor. The implemented framework contained the following steps.

1. Part of the mosaic was randomly selected as a training set and the other part as a validation set. The training set and the validation set were mutually exclusive. Each sample was a 128 × 128 pixel patch of the mosaic. The Ordnance Reef mosaic had 9,899 patches of background class and 782 patches of 5-inch shell class in total dataset. From this set we create training set using parts of it and validation using the rest. For 10 crossfold validation, in each iteration a new training set and a new validation set were randomly chosen from this collection of patches in the dataset. Note that the classes were not equally probable: the 5-inch shell class occupied only 7.32% of the dataset.
2. We used the Ordnance Reef dataset characteristics, like dataset size, color availability, class type, resolution of the patch to select the best options from among those described by (Shihavuddin et al., 2013). Table 1 describes the selected configuration.

FIGURE 6

Examples of phase symmetry on elevation map images containing 5-inch shells and background. (a) 5-inch shell, (b) elevation map, (c) phase symmetry image, (d) background, (e) elevation map, (f) phase symmetry image.



Three sets of runs were conducted. One set used only 2D features, another used only 2.5D features, and a third used both 2D and 2.5D features. For the runs with 2D features, we used only the first four substeps in the feature extraction step; for 2.5D, we used only the fifth substep; for the 2D + 2.5D, we used all the substeps in the feature extraction step (Table 1).

3. The training set was used to find the boundaries among classes and the validation set to access accuracy. We also created the thematic mapping on the mosaic and access the performance comparing with ground truth. Ground truth was created using manual segmentation and annotation from the part of the Ordnance mosaic.

Overall accuracy (Story & Congalton, 1986), average precision, kappa (Jacob, 1960), and average mutual information (AMI; Finn, 1993) were used to evaluate the performance of these three sets of classification runs. Overall accuracy was the ratio of correctly classified units by the total number of units. Precision was defined as the fraction of retrieved instances that are relevant, while recall was the fraction of relevant instances that are retrieved. Average precision summarized the precision-recall curve by measuring the area under the curve. Kappa was the measurement of how the classification result related to the reference data discarding the chance agreements. AMI measured the information gain after classification by comparing it with the reference data. The

TABLE 1

Configuration of the implemented method.

Steps	Methods
Preprocessing	CLAHS (Zuiderveld, 1994)
Segmentation	TurboPixels (Levinshtein et al., 2009)
Feature extraction	1. Gabor filter response (Porter & Canagarajah, 1997)
	2. GLCM (Shihavuddin et al., 2013)
	3. CLBP (Ojala et al., 1996)
	4. Hue and Opponent histograms (Van de Weijer & Schmid, 2006)
	5. 2.5D features
Feature modification	1. Chi-square and Hellinger kernel (Soh & Tsatsoulis, 1999), PCA
	2. L1 normalization
Classification	Probability density weighted mean distance (PDWMD) (Stokes & Deane, 2009) with equal prior probability
Thematic mapping	morphological filtering (Shihavuddin et al., 2013)

reference data or the ground truth of the Ordnance Reef mosaic were created manually by experts.

Results

The Ordnance Reef dataset contained examples of 5-inch shells, which are man-made objects appropriate for testing the performance of including 2.5D features for classification and thematic mapping. Table 2 presents the summary of experimental results for Ordnance dataset using 2D, 2.5D, and 2D + 2.5D features for classification and thematic mapping.

When using only 2D features, the overall accuracy was 96.23% and average precision was 78.11%. Using only 2.5D features, the overall accuracy was 94.36% and average precision was 66.31%. Using both 2D and 2.5D features together provided better results than 2D or 2.5D individually. In this case, the overall accuracy was 96.85% and average precision was 79.82%.

The classification accuracy for this dataset was high in general, as there were only two classes and most of the samples belong to the background class. More comprehensive insight on the classification performance can be obtained from other indicators, namely the average precision for the validation set and the kappa statistics for the thematic mapping. Figure 7 illustrates the precision-recall curve of background and 5-inch shell classes on the validation set when 2D + 2.5D features are used.

For thematic mapping, we obtained kappa values of 66.51% and 55.78% using individually 2D or 2.5D features, respectively, whereas using 2D + 2.5D features resulted in a kappa value of 71.82%, which is 5.31% higher than using only 2D features (Table 2). The Ordnance Reef dataset had the majority of patches from the background class. Therefore, kappa was more sensitive than the other classification performance metrics to the performance gain of our method since it discards the chance agreement. A similar pattern of results was observed on the validation set, where the average precision was the highest (79.82%) using 2D + 2.5D features or 1.71% higher than the sole use of 2D features. Also, the AMI value for 2D + 2.5D features was highest compared with individual use of 2D or 2.5D features.

All of the presented 2.5D features were not equally effective on this dataset. Some of them had more discriminative information than the others. To evaluate their relative performance, we used them individually and assessed the accuracy using kappa and average precision (Table 3). The coefficients of polynomial surface fitting, curvatures, and symmetry features performed better than the others. Nevertheless, elevation statistics, slope, and rugosity also performed moderately well in terms of kappa and average precision.

Figure 8 illustrates the thematic mappings created using 2D, 2.5D, and 2D + 2.5D features. Figure 9 shows the entire Ordnance Reef mosaic classified into 5-inch shell and background classes using 2D + 2.5D features. In

FIGURE 7

Precision-recall curves for both (a) background and (b) 5-inch shells on the validation set of the Ordnance dataset.

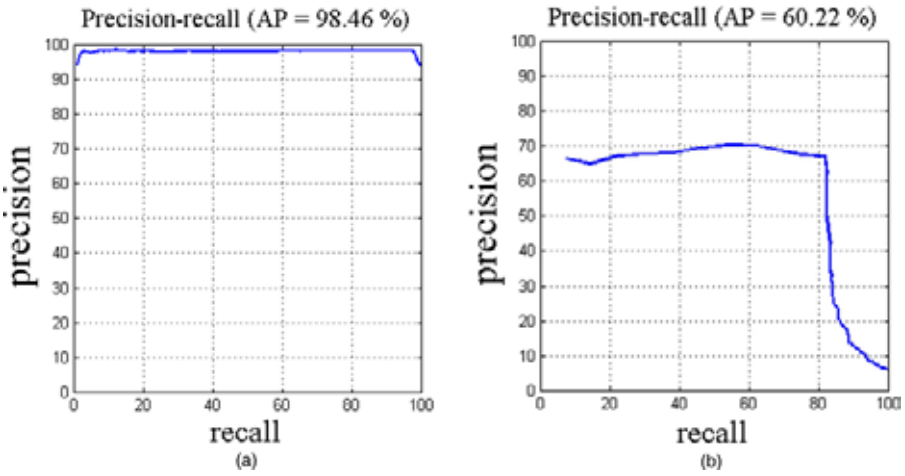


TABLE 2

Summary and comparison of classification results using 2D, 2.5D, and 2D + 2.5D features for Ordnance dataset.

		2D	2.5D	2D + 2.5D
Thematic mapping	Accuracy	96.01%	94.59%	96.95%
	Kappa	66.51%	57.58%	71.82%
	AMI	0.146	0.107	0.165
Validation set	Accuracy	96.23%	94.36%	96.85%
	Average precision	78.11%	66.31%	79.82%

TABLE 3

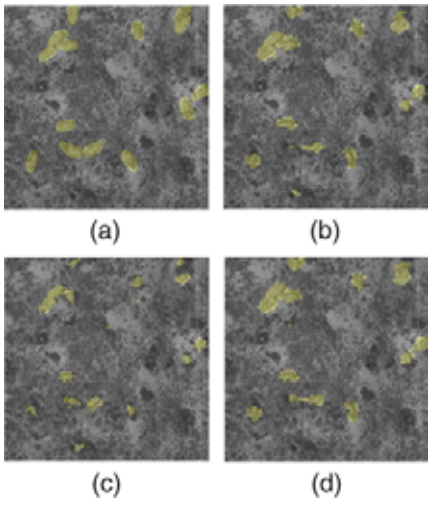
Summary of the 2.5D features used in this experiment and their individual performance on the Ordnance Reef dataset.

Type of the 2.5D features	Number of features	Kappa	Average precision
Coefficients of polynomial surface fitting	9	44.82	62.94
Standard deviation, skewness and kurtosis of the elevation of the surface	3	26.45	55.70
Slope of the principal plane of the patch	1	29.23	56.63
Mean and standard deviation of the distance of the 3D points to the principal plane	2	22.43	54.50
Surface normal (mean and standard deviation of the slope and aspects of each facet within the analysis patch)	4	21.19	53.11
Curvatures	12	36.19	59.32
Rugosity	1	27.68	49.78
Symmetry	2	32.24	57.71

both figures, the yellow color represents the segmented 5-inch shells. By comparing Figures 8(a) and 8(d), we can see that, using the combination of 2D and 2.5D features, all the 5-inch shells were detected even though not very well segmented. In case of the only 2D features, one of the 5-inch shells is not detected (Figure 8b).

FIGURE 8

Thematic mapping using 2D, 2.5D, and 2D + 2.5D features. Here the yellow color represents the segmented 5-inch shells. (a) Ground truth (b) 2D features (Kappa 66.51%) (c) 2.5D features (Kappa 57.58%) (d) 2D + 2.5D features (Kappa 71.82%).

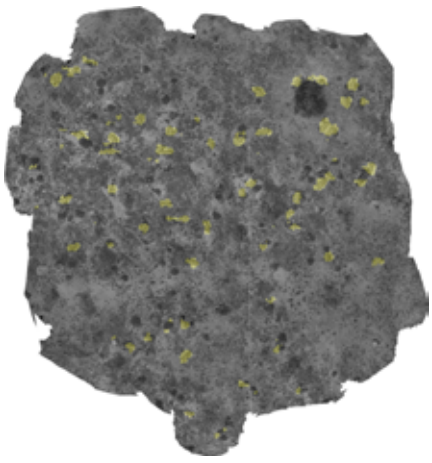


Conclusion

Automated classification of underwater objects using optical imagery can be very beneficial for the scientific community to assess different aspects of the underwater world. In this work, for automated UWMM detection, we proposed the use of 2.5D features to complement other 2D features existing in the literature. The novel part of this work was the combined use of 2D and 2.5D features from optical imagery for benthic object classification. Coefficients of polynomial surface fitting, standard deviation, skewness and kurtosis of the elevation, slope of principal plane, mean and standard deviation of the distance of 2.5D points to the principal plane, surface normal, curvatures, rugosity and symmetry measures were extracted from the elevation map of the surface and used as 2.5D features in this work. We were able to achieve higher classification performance on the test dataset with the proposed method in terms of overall accuracy, average precision, average mutual information and kappa statistics. Using the fusion of 2D and 2.5D features, we obtained 5.31% gain in kappa for thematic mapping on the Ordnance Reef mosaic. This increase in kappa provided the best insight on the performance gain of our method because it discarded chance agreement, which is large for a two-class dataset particularly when the prior probabilities of the classes were nonequal. We also analyzed the individual performance of 2.5D features and found that the coefficients of polynomial surface fitting, curvatures,

FIGURE 9

Thematic mapping performed on the Ordnance mosaic with the use of both 2D and 2.5D features. Here the yellow color represents the segmented 5-inch shells.



and symmetry were the most important features for this dataset. As future work, we would like to test the proposed method on other datasets and generate more sophisticated 2.5D features able to work in general for varied and challenging object classes.

Acknowledgments

This work was partially funded by the Spanish MICINN under grant CTM2010-15216 (MuMap), by the EU Project FP7-ICT-2011-288704 (MORPH), and by the U.S. Department of Defense SERDP program (project MR-2414). A.S.M. Shihavuddin was supported by the MICINN under the FI program. Nuno Gracias was funded by EU Project FP7-INFRASTRUCTURES-2012-312762 (EUROFLEETS2). We would like to thank the authors of Weka, CLBP, Prtools, and Vlfeat toolboxes.

Corresponding Author:

A.S.M. Shihavuddin
Computer Vision and
Robotics Group (VICOROB),
University of Girona, Spain
Email: shihav407@yahoo.com

References

Bair, E., Hastie, T., Paul, D., & Tibshirani, R 2006. Prediction by supervised principal components. J Am Stat Assoc. 101(473):119-37. http://dx.doi.org/10.1198/016214505000000628.

Beijbom, O., Edmunds, P., Kline, D., Mitchell, B, & Kriegman, D. 2012. Automated annotation of coral reef survey images. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, Rhode Island: IEEE.

Bewley, M.S., Douillard, B., Nourani-Vatani, N., Friedman, A.L., Pizarro, O., & Williams, S.B. 2012. Automated species detection: An experimental approach to kelp. The Australasian Conference on Robotics and Automation (ACRA). IEEE.

Daniel, L., Lloyd, F., Queen, P., & Charles, R. 2012. Introduction to Data Analysis Using Geographic Information Systems. St. Paul, MN: University of Minnesota Extension Service. Available at http://conservancy.umn.edu/bitstream/93734/1/5740.pdf.

David, J. 1987. Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc Am B. 4:2379-94. http://dx.doi.org/10.1364/JOSAA.4.002379.

Finn, J.T., 1993. Use of the average mutual information index in evaluating classification error and consistency. Int J Geogr Inf Syst. 7(4):349-66. http://dx.doi.org/10.1080/02693799308901966.

Fischler, M., & Bolles, R. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM. 24(6):381-95. http://dx.doi.org/10.1145/358669.358692.

Friedman, A., Pizarro, O., & Williams, S. 2010. Rugosity, slope and aspect from bathymetric stereo image reconstructions. Sydney: OCEANS IEEE. pp. 1-9.

Friedman, A., Pizarro, O., Williams, S., & Johnson-Roberson, M. 2012. Multi-scale measures of rugosity, slope and aspect from benthic stereo image reconstructions. PloS One. 7(12):e50440. doi:10.1371/journal.pone.0050440.

Gracias, N., van der Zwaan, S., Bernardino, A., & Santos-Victor, J. 2003. Mosaic based navigation for autonomous underwater vehicles. IEEE J Oceanic Eng. 28(4):609-24. doi:10.1109/JOE.2003.819156

Hetzel, G., Leibe, B., Levi, P., & Schiele, B. 2001. 3D object recognition from range images using local feature histograms. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2:394-9.

Jacob, C. 1960. A coefficient of agreement for nominal scales. Educ Psychol Meas. 20(1):37-46. http://dx.doi.org/10.1177/001316446002000104.

Keren, D., & Gotsman, C. 1999. Fitting curves and surfaces with constrained implicit polynomials. IEEE T Pattern Anal. 21(1):31-41. http://dx.doi.org/10.1109/34.745731.

Knopp, J., Prasad, M., Willems, G., Timofte, R., & Gool, L. 2010. Hough transform and 3d SURF for robust three dimensional classifications. European Conference on Computer Vision (ECCV). 6316:589-602.

Kohler, K.E., & Gill, S.M. 2006. Coral Point Count with Excel extensions (CPCe): A Visual Basic program for the determination of coral and substrate coverage using random point count methodology. Comput Geosci. 32(9):1259-69. http://dx.doi.org/10.1016/j.cageo.2005.11.009.

Kovesi, P. 1997. Symmetry and asymmetry from local phase. Tenth Australian Joint Convergence on Artificial Intelligence. pp. 2-4.

Levinshtein, A., Stere, A., Kutulakos, K., Fleet, D., Dickinson, S., & Siddiqi, K. 2009. Turbopixels: Fast superpixels using geometric flows. IEEE T Pattern Anal. 31(12):2290-7. http://dx.doi.org/10.1109/TPAMI.2009.96.

Lirman, D., Gracias, N., Gintert, B., Gleason, A., Reid, R.P., Negahdaripour, S., & Kramer, P. 2007. Development and application of a video-mosaic survey technology to document the status of coral reef communities. Environ Monit and Assess. 125:59-73. http://dx.doi.org/10.1007/s10661-006-9239-0.

Lowe, D.G. 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision (IJCV). 60(2):91-110.

Lundblad, E., Wright, D., Naar, D., Donahue, B., Miller, J., Larkin, E., & Rinehart, R. 2004. Classifying deep water benthic habitats around Tutuila. American Samoa. In: 24th Annual ESRI User Conference. San Diego, CA: Environmental Systems Research Institute (ESRI). Paper 1208.

Marcos, M., Soriano, M., & Saloma, C. 2005. Classification of coral reef images from underwater video using neural networks. Optics Express. 13:8766-71. http://dx.doi.org/10.1364/OPEX.13.008766.

Marton, Z., Pangercic, D., Blodow, N., & Beetz, M. 2011. Combined 2d–3d categorization and classification for multimodal perception systems. Int J Robot Res. 30(11):1378-402. http://dx.doi.org/10.1177/0278364911415897.

Nicosevici, T., Gracias, N., Negahdaripour, S., & Garcia, R. 2010. Efficient three-dimensional scene modeling and mosaicing. J Field Robot. 16(10):759-88.

Ojala, T., Pietikainen, M., & Harwood, D. 1996. A comparative study of texture measures with classification based on featured distributions. Pattern Recogn. 29(1):51-9. http://dx.doi.org/10.1016/0031-3203(95)00067-4.

Paul, J.B., & Ramesh, C.J. 1986. Invariant surface characteristics for 3D object recognition in range images, computer vision, graphics. Image Processing. 33(1):33-80. http://dx.doi.org/10.1016/0734-189X(86)90220-3.

Pican, N., Trucco, E., Ross, M., Lane, D., Petillot, Y., & Tena Ruiz, I. 1998. Texture analysis for seabed classification: co-occurrence matrices vs. self-organizing maps. OCEANS '98 Conference Proceedings. 1:424-8.

Pierre, A., Stephane, T., & Camille, W. 2013. 3D fast intersection and distance computation (AABB tree). CGAL User and Reference Manual. CGAL Editorial Board, 4.3 edition.

Pizarro, O., Rigby, P., Johnson-Roberson, M., Williams, S., & Colquhoun, J. 2008. Towards image based marine habitat classification. Quebec: OCEANS'08 MTS/IEEE.

Porter, R., & Canagarajah, N. 1997. Robust rotation invariant texture classification: wavelet, gabor filter and gmrf based schemes. Vision, Image and Signal Processing, IEE Proceedings. 144(3):180-8.

Sandbach, G., Zafeiriou, S., Pantic, M., & Rueckert, D. 2012. Recognition of 3d facial expression dynamics. Image Vision Comput. 30(10):762-73. http://dx.doi.org/10.1016/j.imavis.2012.01.006.

Schwartz, A., & Brandenburg, E. 2009. An overview of underwater technologies for operations involving underwater munitions. Mar Technol Soc J. 43(4):62-75. doi:10.4031/MTSJ.43.4.12.

Shihavuddin, A., Gracias, N., Garcia, R., Gleason, A., & Gintert, B. 2013. Image-based coral reef classification and thematic mapping. Remote Sensing. 5(4):1809-41. http://dx.doi.org/10.3390/rs5041809.

Soh, L., & Tsatsoulis, C. 1999. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. IEEE T Geosci Remote. 37(2):780-95. http://dx.doi.org/10.1109/36.752194.

Stokes, M., & Deane, G. 2009. Automated processing of coral reef benthic images. Limnol Oceanogr-Meth. 7:157-68.

Story, M., & Congalton, R.G. 1986. Accuracy assessment—A user's perspective. Photogramm Eng Rem S. 52(3):397-9.

Triggs, B., McLauchlan, P., Hartley, R., & Fitzgibbon, A. 2000. Bundle adjustment—A modern synthesis. In: Proceedings of the International Workshop on Vision Algorithms: Theory & Practice. Lecture Notes in Computer Science. London: Springer-Verlag. 1883:298-372.

Van de Weijer, J., & Schmid, C. 2006. Coloring local feature extraction. In: Proceedings of the 9th European Conference on Computer Vision (ECCV). Volume Part II, Berlin, Heidelberg: Springer. Lecture Notes in Computer Science. 3952, pp. 334-48.

Wang, Y., Peterson, B.S., & Staib, L.H. 2000. Shape-based 3D surface correspondence using geodesics and local geometry. IEEE Conference on Computer Vision and Pattern Recognition. Proceedings. 2:644-51.

Zuiderveld, K. 1994. Contrast limited adaptive histogram equalization. San Diego, CA, USA: Academic Press Professional, Inc Graphics gems IV. pp. 474-85.