Seafloor Analysis and Understanding for Underwater Archaeology

Marco Reggiannini[1], Ovidio Salvetti[1]

Via G. Moruzzi 1, Pisa

ISTI, CNR[1]

Abstract

Surveying the oceans' floors represents at the same time a demanding and relevant task to operators concerned with marine biology, engineering or sunken cultural heritage preservation. Scientific researchers and concerned persons combine their effort to pursue optimized solutions aiming at the mapping of underwater areas, the detection of interesting objects and, in case of archaeological survey mission, the safeguard of the detected sites. Among the typical tools exploited to perform the cited operations the Autonomous Underwater Vehicles (AUVs) represent a validated and reliable technology. These vehicles are typically equipped with properly selected sensors that collect data from the surveyed environment. This data can be employed to detect and recognize targets of interest, such as manmade artefacts located on the seabed, both in an online or offline modality. The adopted approach consists in laying emphasis on the amount of regularity contained in the data, referring to the content of geometrical shapes or textural surface patterns. These features can be used to label the environment in terms of more or less interesting areas, where more interesting refers to higher chances of detecting the sought objects (such as man-made objects) in the surveyed area. This paper describes the methods developed to fulfill the purposes of mapping and object detection in the underwater scenario and presents some of the experimental results obtained by the implementation

URL: www.si.isti.cnr.it (ISTI, CNR)

[1]Institute of Science and Information Technologies - National Research Council of Italy.

of the discussed techniques in the underwater archaeology field.

Keywords: Underwater Object detection, Optical and acoustic data processing, Mosaicking, 3D reconstruction, Shape recognition, Texture analysis

1. Introduction

Mapping the oceans' floors represents an extremely demanding task to the man. The peculiar environmental setting is for the most part out of reach to human operators because of the hard environmental conditions that make the survey complex and dangerous. Nevertheless the sea waters cover approximately the 72% of the planet's surface and mapping the seafloors is still a relevant task of typical concern to many involved operators such as biologists, engineers and archaeologists.

On the other hand it is known that the oceans' floors host large amounts of cultural heritage (more than 3 millions of wrecks according to the latest UNESCO reports) as a consequence of shipwreckages that took place during the past ages. This fostered the combined commitment between cultural institutions and scientific researchers to pursue a solution towards the safeguard of this collective patrimony. In this framework several ventures have been started, based on the effort of either national (THESAURUS - TecnicHe per lEsplorazione Sottomarina Archeologica mediante lUtilizzo di Robot aUtonomi in Sciami, PAR FAS 2007-2013 Regione Toscana) as well as international (ARROWS - ARchaeological RObot Systems for the World's Seas, European FP7 project) cooperating consortia ([1], [2] and [3]). These projects have been focused on the main purposes of mapping, diagnosing, cleaning and securing of underwater and coastal archaeological sites. To perform all the cited operations a marine vehicle, such as an autonomous underwater vehicle (AUV), can be profitably exploited. The vehicle can be equipped with properly selected sensors, in order to collect data from the surveyed environment in an optimal way.

Typical sensors that turn out to be useful in this frame are optical cameras coupled with acoustic sensors, like sidescan sonars or multibeam sonars. The data collected by the AUVs during the mission campaigns can be processed in order to detect targets of interest located on the seabed. The main approach adopted in the processing procedure is to emphasize the amount of regularity detected in the captured data, hence highlighting fragments of geometrical shapes, such as primitive curves, or homogeneous areas exhibiting similar textural patterns. A strong and persistent presence of this regularities is considered a clue for the presence of man made targets on the seafloor.

The features are computed by processing the optical and acoustic collected data. The output result of the overall signal processing chain consists in the labeling of the represented environment in terms of more or less interesting scenarios. The term interesting usually refers to a quantitative index which numerically expresses our confidence about the presence of some specific sought object inside the environment. Hence it could be a score which, based on the number and relevance of the detected features, could indicate the likely presence of an interesting object. Given the generality of the proposed approach the object to be detected can be represented by a large variety of targets, here including archaeological wrecks as well as flora specimens, posidonia prairies or corals or even underwater industry structures like oil and gas pipelines.

The methods developed to fulfill the cited purposes will be described in detail in the remaining part of the paper, which is organized as follows: section 2 concerns a brief summary of the existing commercial solutions for underwater vehicles, in section 3 the exploited sensors and their main features are discussed, section 4 represents the paper's core and concerns the description of the main approach and the multiple techniques developed to the purpose of understanding and representing the underwater scenarios, finally section 5 concludes the paper by discussing potential future prospects in the field of the underwater optical and acoustic signal analysis.

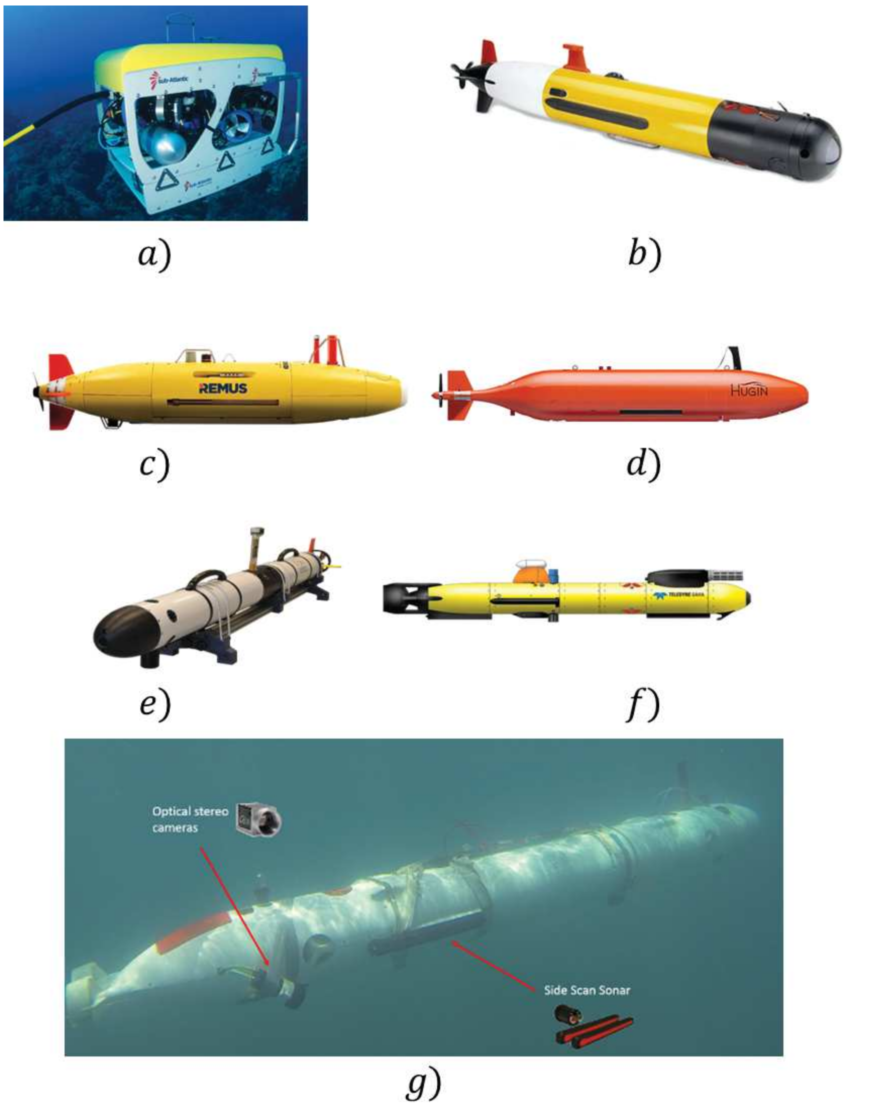

Figure 1: Sub-Atlantic Mohawk (a), an example of commercial ROV vehicle (picture available at http://flowergarden.noaa.gov/). Remus 100 (b), an example of commercial AUV (picture available at http://hydroid.com/remus-100-marine). Popular Commercial AUVs (c-f) and TifOne (g), the research AUV designed and implemented within the THESAURUS project. Details of the optical and acoustic payload mounting on TifOne are displayed.

2. Vehicle Platforms for Underwater Surveying

Among the technological systems employed to survey the underwater setting it is appropriate to spend a few considerations about the mobile platforms that are typically employed for maritime survey purposes and provide a brief presentation of their main properties and features. These mobile platforms can be roughly grouped in two classes: the Remotely Operated Vehicles (ROV, figure 1-a), that can be directly maneuvered by human operators thanks to a wired connection from a control platform to the vehicle itself, and vehicles that are designed to perform underwater missions without human supervision, i.e. the already mentioned AUVs (figure 1-b). In the following the latter category will be considered for further analysis since AUV has been chosen as the reference mobile platform for the marine survey tasks in THESAURUS and ARROWS. Among the most relevant outcomes of the mentioned projects several solutions of AUVs specifically dedicated to the archaeological exploration task have been designed, implemented and experimentally tested.

AUVs are programmable robotic vehicles that, depending on their design, can drive or glide through the ocean without the requirement of a real-time control by human operators. When needed AUV's control and localization may be performed by means of acoustic communication channels towards surface platforms or by exploiting underwater positioning methods based on networks of acoustic beacons distributed on the seafloor, such as Long, Short or Ultrashort Baseline technologies ([4]). Otherwise, once the mission has started the vehicle performs the planned tasks autonomously, without interacting with human operators until the end of the mission.

Most of the available AUVs feature a torpedo-like shape. They are often employed as multi-purpose platforms for oceanographic experiments since they can be quite easily deployed in the marine environment and activated to perform specific measurements. The main components that make up an AUV are: i) an essential system resulting from the combination of the chassis structure and all the engineering systems governing the mechanical behaviour (propellers, actuators, etc.), here including a primary processing unit assigned to the implementation and control of basic tasks (navigation and attitude control, internal humidity measurement and generic vehicle diagnostics), ii) a data capture system resulting from the integration of hardware tools (payload sensors) and software modules (ad hoc implemented data processing algorithms) dedicated to the tasks of surveying and understanding the environment. Installation of payload sensors can be adapted according to different mission scenarios, from physical parameters measurement such as CTD probing (conductivity, temperature and depth/pressure profiling), to inspections of the environment by means of optical and acoustic mapping sensors (TV cameras and side scan sonar).

Existing AUVs feature strong pressure resistance, an important property that allows to reach relevant depths, typically from hundreds to thousands of meters. This enables the system to perform large scale mapping of vast areas as well as close-up inspections of the seafloor.

So far a large number of commercial solutions have been proposed by companies operating in the maritime field. REMUS (figure 1-b,c) represents one of the most popular typologies of AUV. It is produced by Hydroid, which is part of Kongsberg Maritime Company (Norway). Its length varies from 1.6 m (REMUS 100) to 3.84 m (REMUS 6000) and it is rated for a maximum operating depth of 100 m (REMUS 100) to 6000 m (REMUS 6000). Its battery ensures a mission duration of 10 hours (REMUS 100) to 22 hours (REMUS 6000) and a maximum speed of 2.3 m/s.

Hydroid is also the manufacturer of HUGIN (figure 1-d), another commercial AUV which is suitable for underwater inspection and mapping purposes. Its specifications are very similar to the REMUS ones, differing only for its larger dimensions (up to 6.4 m) and better performances in terms of battery endurance (up to 74 hours).

The mentioned vehicles represent interesting solutions given their versatility for applications to a wide range of underwater operations, from pipeline monitoring to mine countermeasure missions, here including mapping for biological inspection and geological assessment purposes. The main drawbacks of the described systems are represented by their relevant cost (from 100K€ to 1M€) and by their voluminous dimensions (HUGIN may weigh up to 1550 Kg), which make them impracticable for operations carried out by small groups of people and implies the rental of expensive dedicated machinery and supporting ships in order to perform a safe and proper deployment.

Many companies provide solutions that can be considered relatively comfortable in terms of transport and deploy, such as the OceanServer IVER-3 (figure 1-e) or the modular Teledyne GAVIA (figure 1-f). Both the vehicles have an approximating length of 2 meters and a weight of about 50 Kg. The additional modularity feature of the Teledyne GAVIA enables the user to adapt the vehicle to the experimental requirements, by installing onboard the proper payload sensor modules. This represents a crucial requirement, that inspired the mechanical design activity within both projects, THESAURUS and ARROWS. Indeed modularity represents a desirable condition in order to make the sensor platform versatile and adaptive to different mission scenarios, such as the strongly varying altitude parameter (distance of the vehicle from the seafloor) that typically differentiates between large scale survey missions (e.g. 40 m altitude) or close range observations of localized spots (few meters altitudes).

The following sections concern a detailed description of the engineering and information technology results achieved within the research projects previously introduced, representing the scientists' answer to the lack of current market solutions dedicated to the fulfillment of archaeologists' desiderata. This is true for what concerns the lack of embedded software systems for the processing, integration and understanding of the captured payload sensor data, aiming at the discovery of currently undetected underwater archaeological sites.

3. AUV Sensor Equipment for Optimal Data Acquisition

In the circumstances described in this paper the AUV is equipped with sonars and optical cameras in order to map the seafloors. More in detail the sensing system features two digital cameras in stereo configuration with a sonar device, which can be a multibeam or a side scan sonar. The choice of the acoustic payload sensors to be employed during the missions depends on the specific purpose of the mission itself and on the environmental scenario that has to be surveyed. The side scan sonar and the multibeam forward looking sonar return large scale maps of the seafloor that are typically processed to detect obstacles, objects or areas of the seafloor showing interesting features. On the other hand the multibeam echosounder returns detailed 3D bathymetry maps of the inspected area in the form of point cloud data.

Typically optical and acoustic sensors are installed on the same AUV, they share the vehicle reference frame and capture co-located data, that is they return a multi-sensor description of the environment from a common perspective. Considering the object detection as a primary goal for this work, the selected sensor suite turns out to be the optimal choice. Indeed optical and acoustic devices feature complementary properties in terms of resolution and best operating conditions.

An example of a research-oriented AUV, equipped with optical and acoustic payload sensors, is represented in figure 1-f, displaying the TifOne, an AUV that has been developed by the Mechatronics and Dynamic Modelling Laboratory of the University of Florence in the framework of the ARROWS project.

Due to the complex and noisy environment in which the survey operations are performed, the data collected by the payload sensors must undergo multiple processing stages, starting from the enhancement of the raw output signal, ending with specific computer vision and image processing techniques with the purpose of extracting as much information as possible from the collected data. A preliminary step in the manipulation of the data consists in the restoration of the signal and in the enhancement of the relevant properties, whose integrity is crucial for the correct understanding of the scenario. In our specific case we are concerned with correcting those systematic geometrical distortions introduced in the signal due to the peculiar perception of the environment, which is intrinsic in every employed sensor.

As an example consider the geometrical distortions affecting the side scan

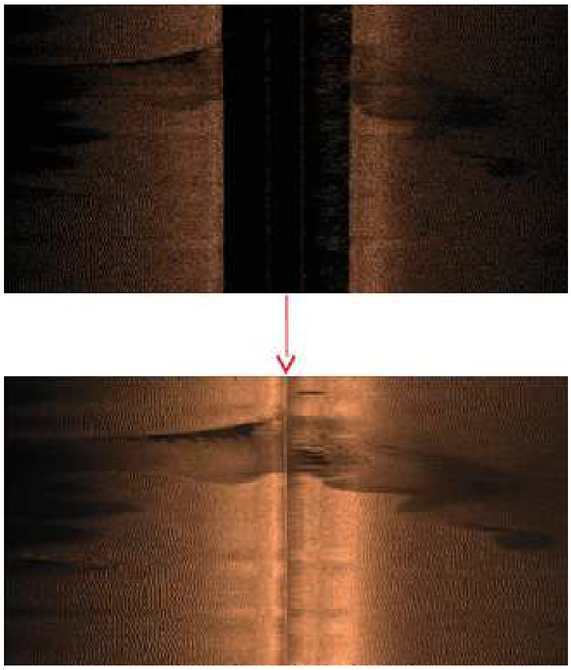

Figure 2: Side scan sonar map represented by Slant (up) and Ground (down) corrected coordinates.

sonar mapping: the measured time of the backscatter echo is not linearly proportional to the seabed ground range. Indeed, within a row of a Side Scan Sonar map, different segments of pixels with the same given length, starting from different positions of the row, correspond to spatial segments with different lengths on the horizontal (ground) range axis. To correctly represent the data in the seabed frame the *Slant-to-Ground* transformation indicated below is required:

$$y_{i,j} = \sqrt{\frac{c^2 t_{i,j}^2}{4} - H_i^2} \tag{1}$$

where $y_{i,j}$ is the $j-th$ horizontal range sample in the $i-th$ ping, c is the sound velocity in the sea, $t_{i,j}$ is the $j-th$ slant range (time) sample in the $i-th$ ping and H_i is the first echo return of the $i-th$ ping, i.e. the sensor $i-th$ altitude value. Besides the transform in equation (1) allows to get rid of the central black stripe in the raw output sonogram, which is caused by the acoustic wave propagation through the water column (figure 2).

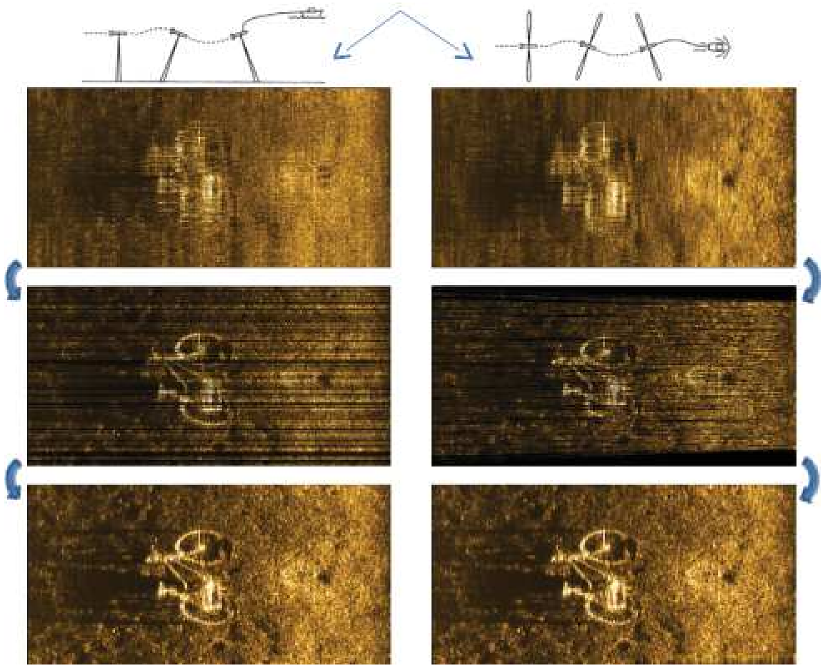

Figure 3: Sidescan sonar map: altitude distortions and corresponding restoration. Original side scan map available on http://www.jfishers.com/.

Relevant sources of image distortion come from the random distortions in the sonogram formation, caused by the unpredictable fluctuations in the AUV attitude (figure 3).

In both the mentioned examples auxiliary sensor devices, such as inertial measurement units, gyrocompass, DVL, etc., can be exploited to provide the vehicle's position and pose values during the payload acquisition. This additional information can be profitably used to restore the corrupted data (see for example [5]). Results of the restoration are presented in the lower part of figure 3.

In the underwater mapping field the signal pre-processing stage represents an important step to be performed before applying algorithms aiming at a high level understanding of the surveyed scenario. Indeed the performance of an automatic understanding system, which involves algorithms borrowed from the computer vision, machine learning and image processing background, strongly depends on the quality of the captured data. Actually we have to be confident that the data, either considered as a straight raw output of the sensor or as the result of the preprocessing stages, exhibit the highest achievable quality. This is an auspicious precondition, that should be pursued in order to allow the dedicated processing units to automatically detect those features and attributes of conspicuous importance for the understanding process.

4. Analysis of the Captured Data and Underwater Scene Understanding

A mixture of unfavourable factors makes the collection of reliable data in the underwater setting a hard task: due to the perturbations' spherical spreading the optical and acoustic energy collected by the sensors decreases proportionally to the squared inverse of the travelled distance.

Moreover the underwater medium heavily affects the spectrum of the optical signal by filtering out a large percentage of the visible frequency range ([6] and [7]). This reduces the maximum operational range of the optical sensor to few meters. Underwater imaging is also corrupted by typical hazing effects that may strongly reduce the visibility in the image.

4.1. Mosaicking

The described sensing limitations can be tackled in case a large set of data relating to the same scenario is available. By exploiting computer vision algorithms that perform the alignment and the integration of multiple maps it is possible to generate a representation of the entire surveyed environment. This can be obtained by exploiting mosaicking procedures. These techniques start from the hypothesis that the surveyed environment features a planar morphology or can be approximated as planar from the camera point of view. Let's consider a point x in multiple consecutive maps resulting from the projections onto the camera plane of the same 3D world point. Under the hypothesis that the different images are related by a projective transformation (homography in case of planar projection), the relation between the point coordinates in image i and j can be formally expressed as:

$$\mathbf{x}_i \sim \mathbf{H}_{i,j}\mathbf{x}_j \tag{2}$$

where $H_{i,j}$ represents the homography transform that maps points of image j on image i. In case of projections of $3D$ points on the camera plane the homography is usually represented by a 3×3 matrix with 8 degrees of freedom.

The homography can be estimated directly from the captured images by considering 3 conditions for every interest point, defined by equation (2). Given n points we have $3n$ equations and $8 + n$ unknown (scale unkown has to be considered if we perform the estimation exploiting points from the captured cameras), so we must have at least 4 points to correctly estimate $H_{i,j}$.

The x points are usually selected based on the detection of reliable features in the data, such as SIFT features ([8]). The features that are detected, matched and exploited to estimate $H_{i,j}$ correspond to those features appearing in the overlapping areas of consecutive maps. The transformed maps are finally stitched together and eventually processed by blending techniques to generate a final seamless mosaic map.

In the unfavourable case that the number of detected reliable features is too low due to bad quality or very noisy data the feature based mosaicking may become an unfeasible operation.

Under that condition the additional data collected by auxiliary sensors measuring the vehicle attitude may help towards the estimation of the required transformations.

In the framework of the ARROWS project the mosaicking procedure based on a feature matching approach has been tested on a set of acoustic and optical data captured during experimental campaigns performed in Sicily and Israel.

Figure 4: Optical mosaic map. The image results from the stitching process of 61 frames captured by a Basler ACE GigE camera during an experimental campaign performed at the Cala Minnola site, in front of the Levanzo Island, Sicily.

During a mission at the Levanzo Island, in Sicily a set of optical images has been captured at the Cala Minnola site and have been processed to return an

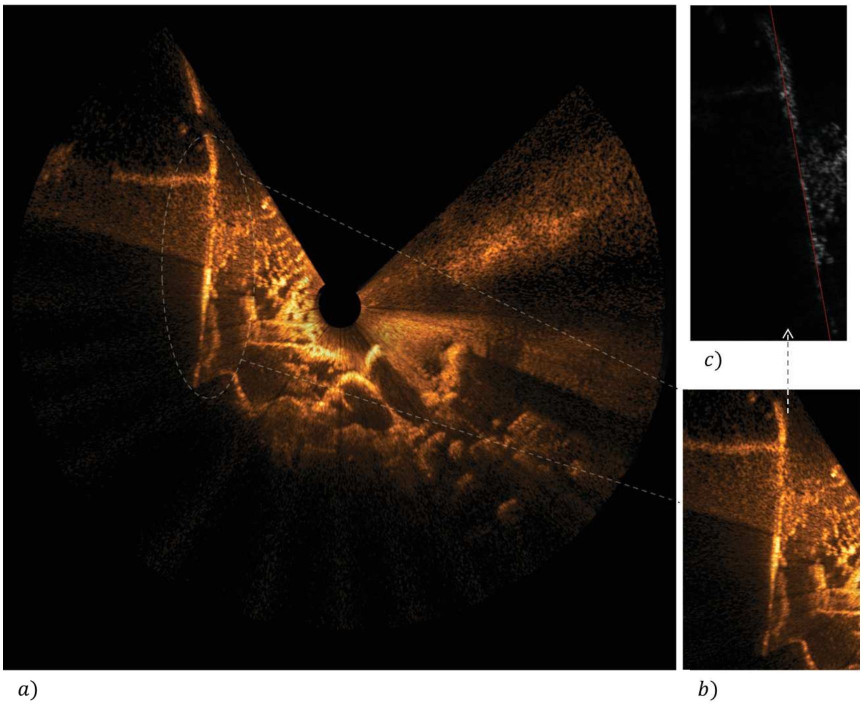

Figure 5: Image a): Forward Looking Sonar mosaic map. The image results from the stitching process of 17 frames captured by a Blueview MB P900 sonar during an experimental campaign performed in Israel, at the Caesarea ancient harbour. Image b): Linear detail, belonging to a pier wall structure, of the mosaic map on the left. Image c): The linear structure represented in image b) has been detected by means of the geometry detection techniques described in this paper.

overall map of the archaeological scene, as illustrated in figure 4. The same procedure has been tested during the Israel campaign. In the latter circumstance a multibeam forward looking sonar sensor (Blueview MB-P900) has been employed to survey a pier wall in the nearness of the Caesarea harbor. Part of the captured data has been processed to obtain a large scale map of that area, as illustrated in figure 5-a.

4.2. Geometry presence assessment

A relevant activity has been developed bearing in mind the primary goal of an archaeological mission, that is the detection of potential structures related to human made objects. The automated system that should perform this task, must recognize specific features exhibited by the selected candidates and put forward an hypothesis on the manmade object's nature. Hence a set of proper criteria has to be chosen in order to grab the most relevant objects' attributes. Within the multiple possible choices we oriented our approach towards assessing the presence of regularity features contained in the captured data. These regularity attributes may refer to the geometric shapes that define the contours of objects as perceived by the sensor device, hence fragments of primitive curves such as lines, circles or ellipses. Starting from the hypothesis that a high concentration of regular curves is a marker for the presence of man-made objects or shipwrecks, we focused our work on the automatic detection of elementary geometric features (line segments, elliptical arcs) in images. This represents a classical computer vision issue which has been thoroughly tackled and discussed by the scientific community (see for example [9] and [10]). The current procedures for geometric features recognition can be roughly classified into Hough-based and region chaining methods.

The Hough-based algorithms make use of implementative variations of the Hough transform. These methods ensure that pixels belonging to the same geometric structure are mapped to the same point into a parameter space whose dimensionality is given by the number of parameters. This typology of algorithms implies elevated computational costs since the procedure complexity grows proportionally with the number of the curve parameters. Therefore it is a good choice to detect lines, but not to detect circular or elliptical shapes. In the authors experience Hough based procedures have been useful to detect the presence of structures featuring mostly linear shapes, such as architectural elements or remains of ancient walls. An example comes from the Israel mission mentioned in section 4.1: the captured acoustic dataset has been processed to detect the presence of primitive curves and the result is illustrated in figure 5-c, where the red line identifies a pier wall structure (detailed in figure 5-b) detected by the algorithm.

The second class of detection methods relies on region growing and chaining techniques. This method exploit the geometric properties of the shapes directly assessed from the images, such as straightness for line segments or curvature properties for ellipses.

These algorithms usually start with a seed pixel or a group of pixels. Then additional pixels are added, provided that they obey some geometric properties of the candidate shape. For example a pixel with coordinates $\mathbf{p}$ and intensity $x(\mathbf{p})$ can be considered aligned to an elliptical arc a if the angle θ formed by the normal $\mathbf{n}_a(\mathbf{p})$ to the arc and the image gradient falls below a predetermined threshold θ_{th}, as expressed by equation (3) and illustrated in figure 6:

$$\theta\left(\nabla x(\mathbf{p}), \mathbf{n}_a(\mathbf{p})\right) \leq \theta_{th} \tag{3}$$

Starting from the work presented in [11] we implemented a procedure for primitive curves recognition purposes. The Ellipse and Line Segment Detector (ELSD) algorithm described in that paper is based on the Gestalt theory, whose applications to computer vision issues are thoroughly discussed in [12]. In a nutshell this method is based on the so called *Helmoltz's perception principle*

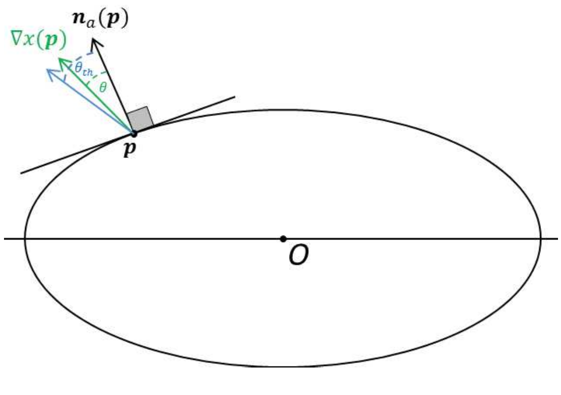

Figure 6: Elliptical collinearity condition.

which formally states that in an unstructured image, only a very small number of detections (false alarms) should take place. The decision about a meaningful candidate curve is based on the probability of observing candidates as structured as the considered one: the smaller this probability value is, the more meaningful the candidate curve is to be considered.

More details about the implemented curve recognition procedure and its application to archaeological sites detection can be found in [13], [14]. This technique can also be exploited on sonar maps to perform attentive analysis of the data (figure 7). By applying the curve recognition algorithm to the new maps, as soon as they become available during the mission, an istantaneous label of interest is assigned to the surveyed regions. Based on that the system can autonomously decide in real time whether a specific area is worth of more detailed inspection or not.

Once more it is worth reminding that a correct restoration of the signal, as illustrated in figure 3, may affect critically the curve recognition process. This is even more important in case the surveyed environment features large varieties of shapes and contours, such as archaeological sites including amphoras, plates and complex wrecks.

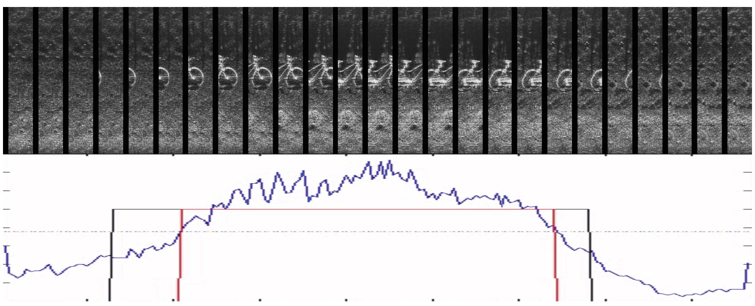

Figure 7: Attentive analysis procedure based on geometry detection applied to the side scan sonar map of figure 3. The blue curve growth is proportional to the number of detected curves, the red line represents the threshold over which the detected number of curves is to be considered relevant and the black line represents a ground truth reference.

4.3. Texture analysis

As previously mentioned, this work is mainly concerned about detecting regularity features in the captured data. Besides shape information, the regularity in terms of *textural* features can be exploited to group pixels of an image in classes that share similar patterns. In this sense the surface appearance of the objects located in the surveyed environment is the attribute that can be exploited to perform image segmentation and classification.

More in particular the appearance of an object's surface can be represented in the spatial frequency domain and the surface patterns may be classified by exploiting their specific frequency content, which plays therefore the role of a discriminative signature. A mathematical tool that has been succesfully employed to perform this operation is the 2D Gabor wavelet function, whose mathematical definition is expressed by equation 4.

$$h(x, y) = \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]\cos(2\pi u_0 x) \tag{4}$$

The wavelet orientation can be adjusted by properly rotating the coordinate system, as in equation 6.

$$x' = x\cos\theta_0 + y\sin\theta_0 \tag{5}$$

$$y' = -x\sin\theta_0 + y\cos\theta_0 \tag{6}$$

In the Fourier domain the transformed Gabor wavelet is represented as a 2D Gaussian function centered at the specific u_0 radial frequency value. From a signal processing point of view the Gabor wavelet can be considered a bandpass filter centered on the specific wavelet band.

The convolution with the Gabor wavelet, which is performed on small windows centered on the image pixels, becomes a multiplication between the transformed functions in the Fourier domain and this results in the emphasis of the common frequency components. This operation is repeated by varying u_0 and θ_0, then, for every convolution result, specific features are computed, such as

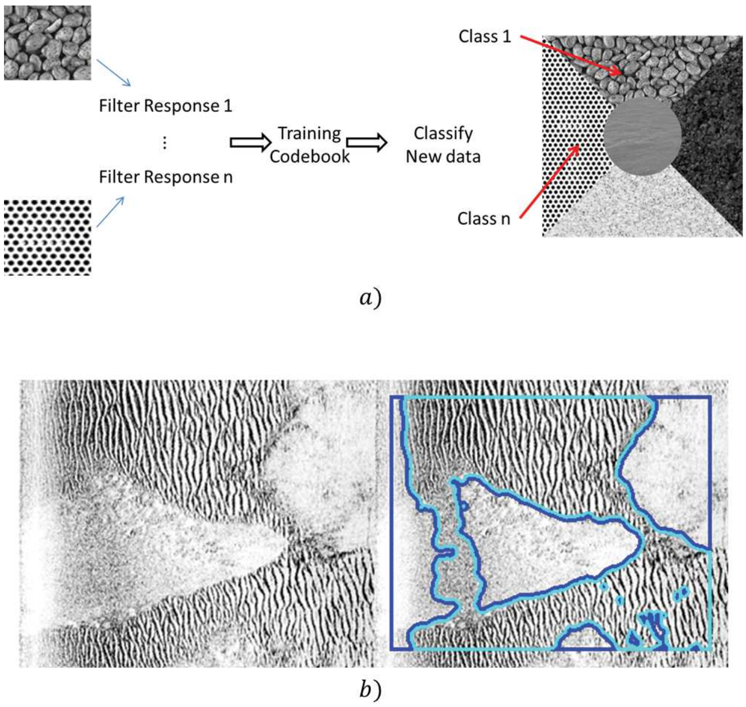

Figure 8: Image a): Preliminary training for the texture classification algorithm (original image taken from http://note.sonots.com/). Image b): Texture classification of side scan sonar maps by means of Gabor filtering (original image taken from http://www.ise.bc.ca/).

the energy of the filtered image, as in [15]. Hence for every image pixel we obtain a set of features describing the pixel frequency content. The similarity between pixels is assessed by comparing the computed feature vectors by means of a proper proximity criterion. To this aim popular clustering algorithms, like K-means (see for example [16]), can be successfully employed. The final segmentation of the map is performed by repeating the described operation for every pixel.

The algorithm can be executed as an unsupervised procedure where the cluster centroids are estimated iteratively from the data, by applying techniques as the cited K-means. An advanced implementation is based on a preliminary training stage in which a set of known classified patterns are processed with different Gabor wavelets in order to provide the main spectral features for various classes, such as rock, sand, mud, posidonia, etc. The results obtained this way are then used to train the algorithm, which is later employed to classify areas in new captured data, based on the increased a priori knowledge. A conceptual sketch of this procedure is represented in figure 8-a and a result of the segmentation process applied on real side scan sonar data is represented in figure 8-b.

This way it is expected that the resulting process may be employed also for online purposes, aiming at a fast preliminary classification of the environment to quickly identify the interesting spots.

4.4. 3D Reconstruction

The techniques described so far allow to perform a large scale mapping of the seafloor. This is an important step since in the first instances of the survey operations the main goal is that of detecting those regions of the seabed that feature high probabilities for the presence of archaeological objects. Once a certain area has been identified a dedicated survey on the localized spot must be performed. The data captured during this close-range survey can be employed for an offline stage of signal processing, aiming at the generation of 3D models of the interesting targets.

Accurate 3D models of the objects can be obtained by processing optical data by advanced photogrammetry methods, such as Structure From Motion (see [17] and [18]). This method is based on the estimation of the 3D coordinates of a point $\mathbf{X}_i$ from the projection $\mathbf{x}_{ij}$ of the point itself on the multiple image planes, defined by the subscript j. The link between $\mathbf{X}_i$ and the point $\mathbf{x}_{ij}$, identifying the projection on the j-th image, is expressed by:

$$\lambda_j \mathbf{x}_{ij} = \mathbf{P}_j \mathbf{X}_i \tag{7}$$

where the *camera matrix* $\mathbf{P}_j$ represents our knowledge about the camera intrinsic parameters (aspect ratio, skew, focal length and camera center coordinates) and the camera extrinsic parameters (the rotation and translation transforms that define the camera pose and position with respect to a global world reference frame). λ_j is a scale factor which accounts for the representation of the image point $\mathbf{x}_{ij}$ in homogeneous coordinates.

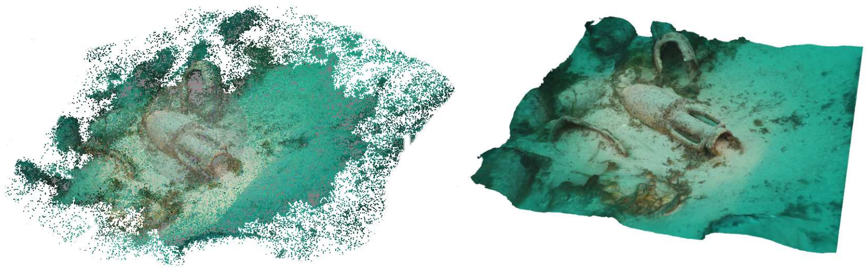

Figure 9: 3D sparse point cloud (left) and refined mesh (right) of an amphora generated from optical data. The image displays a detail of the Cala Minnola wreck site, an underwater archaeological site that has been surveyed in the experimental mission performed in the ARROWS project framework.

It is presumed that the projections of $\mathbf{x}_i$ on different images are recognized as generated from the same spatial point. This can be performed by methods for the detection and matching of salient features, as the SIFT method touched in the 4.1 paragraph. The computation is performed by exploiting multiple constraints on the $\mathbf{x}_i$, resulting from equation (7), expressed for different camera poses j. The estimation of the point coordinates and camera matrices results in the generation of a 3D point cloud, such as the one on the left side of figure 9.

The result can be further refined by estimating the dense cloud and, later, the mesh surface fitting the point cloud. This can be performed by exploiting popular open source softwares such as Meshlab ([19]). An example of point cloud processing is illustrated on the right side of figure 9.

Another typical way of obtaining 3D bathymetric data in the underwater environment is by employing acoustic multi-beam sonars. The output data returned by the multibeam sonar consists of a set of 3D points lying in the intersection between the acoustic beams and the seafloor plane. The depth values and the direction of arrival of the scatterers are hence the direct output

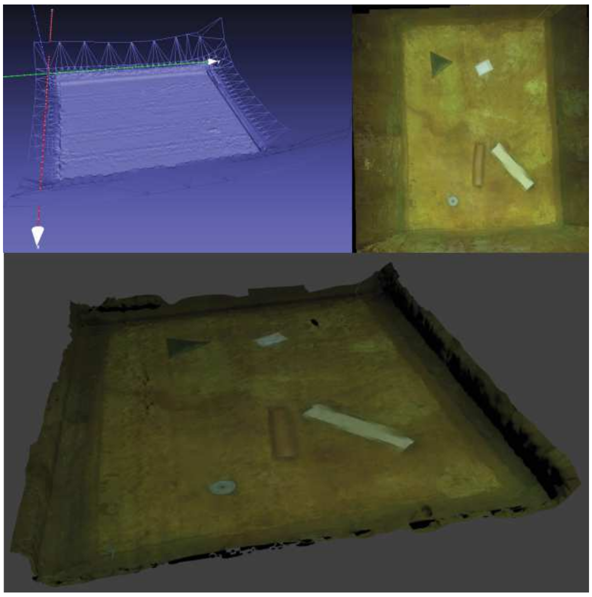

Figure 10: Heterogenous dataset captured during an experimental session at the small pool of the Ocean Systems Lab, Heriot Watt Universtity (Edinburgh). In the left upper part the pool bathymetry map captured by the Blueview MB2250 is illustrated. The right upper part represents the optical mosaic obtained by stitching the GoPro images of the pool floor while the lower part represents the integration between the bathymetric map and the optical mosaic.

of this sensor. This information can be integrated with additional sensor pose and position measurements. This enables the generation of 3D point cloud of the environment. The mapping performance of a Blueview MB-2250 multibeam sonar has been tested by collecting an experimental dataset within the small pool environment of the Ocean Systems Laboratory, Heriot Watt University of Edinburgh (Scotland). The collected maps have been processed to extract the linear subset of interest from the data, corresponding to the intersection between the sensor beams and the pool floor. Later, by integrating these data with the pose measurements of the sensor, it was possible to perform the 3D alignment of the point cloud. The point cloud obtained this way has been further processed to generate the 3D mesh of the terrain (figure 10, upper left), and the result has been integrated with the optical mosaic of the floor (figure 10, upper right). The optical dataset has been obtained by employing a GoPro camera, collecting data simultaneously with the MB-2250. The result of the fusion between bathymetry and textural information is represented in the lower part of figure 10.

4.5. Data Integration

During the survey missions each of the employed payload sensors will provide an individual description of the environment. To the purpose of robustly recognizing archaeological objects it is useful to introduce a synthesis structure summarizing all the informative content related to a seabed area.

This can be formally defined as a multi-dimensional map, made up of multiple layers, each of which refers to a specific category of information. A point in this map returns the entire available information for the corresponding 3D point in the world ([20]). This information may refer to (i) the raw captured data, (ii) the output results of data analysis algorithms, (iii) the bathymetry data collected by dedicated sensors or estimated by computer vision procedures. A point in the map $\mathbf{p}$ can be formally defined as an n-D vector, where n represents the dimensionality of the collected information, which varies with the number of employed payload sensors and the implemented algorithms for data analysis:

$$\mathbf{p}(x,y) = \{p_1(x,y), ..., p_n(x,y)\} \tag{8}$$

According to the previous definition, an example of the multi dimensional map that would realistically represent the output of an AUV archaeological mission, could be built as in the following:

$$\begin{aligned} p_1(x,y) &= \mathit{Optical\ Map\ Intensity\ value\ in}\ (x,y) \\ p_2(x,y) &= \mathit{Acoustic\ Map\ Intensity\ value\ in}\ (x,y) \\ &... \\ p_n(x,y) &= \mathit{Bathymetry\ Map\ value\ in}\ (x,y) \end{aligned} \tag{9}$$

A conceptual sketch of the described fusion map is illustrated in figure 11. In comparison with object recognition procedures based on an individual data typology, it is expected that considering the whole set of available information can be a promising way to perform an efficient object recognition task, reliable with respect to false alarms rejection. A preliminary example of data integration, results from the stitching of camera images mosaic on the multibeam bathymetry map (figure 10). Future work will involve enlarged sets of heterogeneous data and it will be performed on the real data that will be captured during the final experimental surveys of the ARROWS project.

5. Conclusions

The robotic and automation technology presented in this paper will make easier the underwater archaeologists' work, carried out in a hostile and complex environment.

The many implemented procedures aim at providing the archaeologists with methods to perform a thorough analysis of the large and heterogenous amount of data returned by the payload sensors. The proposed processing options aim at the fulfillment of all the archaeologist's requirements and enable him to indirectly perform measurements and formulate historical interpretations on the findings. Moreover, in order to disseminate knowledge regarding the underwater cultural heritage and to increase the sensitivity for its preservation, the devel-

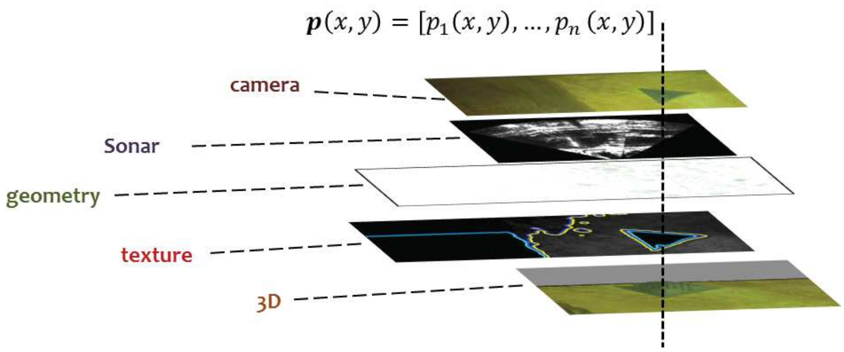

Figure 11: Conceputal diagram illustrating the fusion map. A point in the multilayer structure summarizes all the available information about the surveyed environment.

oped tools allow to address different audiences, including the general public. In particular, one of the purposes of this work is to devise new dissemination channels making use of 3D immersive environments to make more attractive the collected information. The developed methodology has been tested by organizing specific campaigns in relevant European sites, such as the Egadi Archipelagos in Italy, or the Estonian area of the Baltic sea. Most of the presented results, including the collection of the data, its processing using the reported methods, the 3D reconstructions and the virtual scenarios developed with the aim of replicating the experience of wreck exploration and survey, have been made possible in the framework of the European FP7 project ARROWS.

References

[1] B. Allotta, S. Bargagliotti, L. Botarelli, A. Caiti, V. Calabrò, G. Casa, M. Cocco, S. Colantonio, C. Colombo, S. Costa, M. Fanfani, L. Franchi, P. Gambogi, L. Gualdesi, D. La Monica, M. Magrini, M. Martinelli, D. Moroni, A. Munafò, G. Pace, C. Papa, M. Pascali, G. Pieri, M. Reggiannini, M. Righi, O. Salvetti, M. Tampucci, Thesaurus project: Design of new autonomous underwater vehicles for documentation and protection of underwater archaeological sites, in: M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, R. Caffo (Eds.), Progress in Cultural Heritage Preservation, Vol. 7616 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, pp. 486–493. doi:10.1007/978-3-642-34234-9_50.
URL http://dx.doi.org/10.1007/978-3-642-34234-9_50

[2] B. Allotta, R. Costanzi, M. Magrini, N. Monni, D. Moroni, M. A. Pascali, M. Reggiannini, A. Ridolfi, O. Salvetti, M. Tampucci, Towards a robust system helping underwater archaeologists through the acquisition of geo-referenced optical and acoustic data, in: L. Nalpantidis, V. Krger, J.-O. Eklundh, A. Gasteratos (Eds.), Computer Vision Systems, Vol. 9163 of Lecture Notes in Computer Science, Springer International Publishing,

2015, pp. 253–262. doi:10.1007/978-3-319-20904-3_24.
URL http://dx.doi.org/10.1007/978-3-319-20904-3_24

[3] B. Allotta, R. Costanzi, A. Ridolfi, C. Colombo, F. B. M. Fanfani, F. Pazzaglia, O. Salvetti, D. Moroni, M. A. Pascali, M. Reggiannini, M. Kruusmaa, T. Salumäe, G. Frost, N. Tsiogkas, D. M. Lane, M. Cocco, L. Gualdesi, D. Roig, H. T. Gündogdu, E. I. Tekdemir, M. I. C. Dede, S. Baines, F. Agneto, P. Selvaggio, S. Tusa, S. Zangara, U. Dresen, P. Lätti, T. Saar, W. Daviddi, The arrows project: adapting and developing robotics technologies for underwater archaeology, in: IFAC Workshop on Navigation, Guidance and Control of Underwater Vehicles, (NGCUV'2015), At Girona, Spain, 2015.
URL http://hdl.handle.net/2158/1001621

[4] X. Lurton, An introduction to underwater acoustics : principles and applications, 1st Edition, Springer, 2002.
URL http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/3540429670

[5] D. T. Cobra, A. V. Oppenheim, J. F. Jaffe, Geometric distortions in side-scan sonar images: A procedure for their estimation and correction, IEEE Journal of Oceanic Engineering 17 (1992) 252–268. doi:10.1109/48.153442.

[6] M. Massot-Campos, G. Oliver-Codina, One-shot underwater 3d reconstruction, in: Emerging Technology and Factory Automation (ETFA), 2014 IEEE, 2014, pp. 1–4. doi:10.1109/ETFA.2014.7005282.

[7] R. C. Smith, K. S. Baker, Optical properties of the clearest natural waters (200–800 nm), Appl. Opt. 20 (2) (1981) 177–184. doi:10.1364/AO.20.000177.
URL http://ao.osa.org/abstract.cfm?URI=ao-20-2-177

[8] D. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110. doi:10.1023/B:

VISI.0000029664.99615.94.
URL http://dx.doi.org/10.1023/B%3AVISI.0000029664.99615.94

[9] P. V. C. Hough, Machine Analysis of Bubble Chamber Pictures, in: International Conference on High Energy Accelerators and Instrumentation, CERN, 1959.

[10] D. H. Ballard, Readings in computer vision: Issues, problems, principles, and paradigms, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987, Ch. Generalizing the Hough Transform to Detect Arbitrary Shapes, pp. 714–725.
URL http://dl.acm.org/citation.cfm?id=33517.33574

[11] V. Pătrăucean, P. Gurdjos, R. von Gioi, A parameterless line segment and elliptical arc detector with enhanced ellipse fitting, in: A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, C. Schmid (Eds.), Computer Vision ECCV 2012, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, pp. 572–585. doi:10.1007/978-3-642-33709-3_41.
URL http://dx.doi.org/10.1007/978-3-642-33709-3_41

[12] A. Desolneux, L. Moisan, J.-M. Morel, Gestalt theory and computer vision, in: A. Carsetti (Ed.), Seeing, Thinking and Knowing, Vol. 38 of Theory and Decision Library A:, Springer Netherlands, 2004, pp. 71–101. doi:10.1007/1-4020-2081-3_4.
URL http://dx.doi.org/10.1007/1-4020-2081-3_4

[13] D. Moroni, M. Pascali, M. Reggiannini, O. Salvetti, Underwater scene understanding by optical and acoustic data integration, Proceedings of Meetings on Acoustics (POMA) 17. doi:10.1121/1.4792225.

[14] D. Moroni, M. Pascali, M. Reggiannini, O. Salvetti, Underwater man-made and archaeological object detection in optical and acoustic data, Pattern Recognition and Image Analysis. Advance in Mathematical Theory and Applications 24 (2014) 310–317. doi:10.1134/S1054661814020138.

[15] N. Mittal, D. Mital, K. L. Chan, Features for texture segmentation using gabor filters, in: Seventh International Conference on Image Processing And Its Applications 1999 (Conf. Publ. No. 465), Vol. 1, 1999, pp. 353–357 vol.1. doi:10.1049/cp:19990342.

[16] A. Jain, F. Farrokhnia, Unsupervised texture segmentation using gabor filters, in: IEEE International Conference on Systems, Man and Cybernetics, 1990, pp. 14–19. doi:10.1109/ICSMC.1990.142050.

[17] R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press, ISBN: 0521540518, 2004.

[18] M. VARGA, Practical Image Processing And Computer Vision, Halsted Press, New York, NY, USA, 2005.

[19] V. C. Lab, Meshlab: an open source 3d mesh processing system. (2008). URL http://meshlab.sourceforge.net/

[20] D. Moroni, M. A. Pascali, M. Reggiannini, O. Salvetti, Curve recognition for underwater wrecks and handmade artefacts, in: Proceedings of the 4th International Workshop on Image Mining. Theory and Applications, 2013, pp. 14–21. doi:10.5220/0004392300140021.