UNDERSTANDING UNDERWATER ARCHAEOLOGICAL SCENARIOS BY OPTICAL AND ACOUSTIC SENSING TOOLS

MARCO REGGIANNINI

Signals and Images Lab, Institute of Information Science and Technologies
National Research Council of Italy
Via G. Moruzzi 1, 56124, Pisa

ABSTRACT

Underwater seafloors represent a huge archive of manmade artefacts, a cultural heritage whose abandon process started since man has been able to travel by the waters. In order to put this heritage under safeguard and preservation archaeologists require and request for technological support provided by the scientific community. Automated procedures tailored for the purposes of manmade object recognition in the underwater scenario represent the main topic discussed in this work. In particular the authors propose a set of procedures that enables to extract meaningful insights about the inspected environment and that can be exploited to assign a label of interest, in terms of cultural significance, to the surveyed areas. Furthermore the authors introduce a sketch of a framework according to which the identification of interesting areas on the seafloor may be implemented in terms of a Bayes decision system.

Index Terms— Underwater Archaeology, Multi-sensor data analysis, Robust object recognition, Underwater scene understanding, Data integration.

1. INTRODUCTION

Oceans represent unlimited source of knowledge for the men. The sea is the natural habitat of an extremely huge biological diversity and besides playing the role of functional medium for the propagation of human traffic and market, it also represents the most ancient archive of cultural heritage. The huge amount of wrecks lying all over the seafloors ($6 \cdot 10^6$ according to UNESCO) introduces the issue of protecting and preserving this common heritage from the actual threat of criminal misappropriation and illegal removing from the sites. To the purpose of tackling these issues the authors participated to national (THESAURUS [1]) and international (ARROWS [2]) research projects. The main goal of these multidisciplinary partnership projects has been to develop dedicated robotic survey platforms, basically underwater vehicles properly adapted to the projects' requirements, equipped with suitable sensing devices and endowed with skills in order to perform underwater missions, either under human supervision or in autonomous modality, to map a seafloor area, detect and identify those regions that potentially host meaningful archaeological sites.

The paper is organized as follows: section 2 concerns a brief discussion of the state of the art about payload data processing and analysis in the underwater mapping field, section 3 describes the development of data analysis algorithms to extract meaningful features from the collected data, section 4 concerns high level data processing to generate overview mosaics and 3D reconstruction of the seafloor environment, section 5 concerns the description of a robust object recognition procedure based on the collection and exploitation of heterogeneous data while the final section 6 is dedicated to the conclusions.

2. STATE OF THE ART

To a certain extent underwater surveying for object recognition purposes represents still a novel research field, resulting form the intersection between engineering, computer vision and historical-cultural expertises. The technology supporting the mission of the underwater archaeologist actually derives from military experience cumulated during the second half of the twentieth century. Many scientists devoted their effort to develop procedures in order to provide insights about the inspected seafloor environments. More in detail engineers focused on the implementation of procedures aiming at i) the enhancement of the signal to noise ratio of the captured data, ii) the generation of large scale overviews of the scene by means of mosaicking algorithms and iii) providing further information about the seafloor morphology by performing 3D reconstruction.

The pre-processing techniques we come up with are inspired by the schemes proposed and discussed in [3] and [4] as far as acoustic sensors are concerned, and in [5] for optical sensors. Although a relevant part of the work described in this paper has concerned the design and implementation of suitable methods for the preliminary enhancement of the raw data, in this context major focus has been devoted to the discussion of the authors' approach concerning the main core of the scene understanding process and the related data analysis methods. For further details about the implemented data pre-processing procedures the reader may refer to [6].

For valuable discussions concerning geometry detection issues, the reader is referred to [7] and [8]. On the other hand, image classification and segmentation methods are thoroughly discussed in [9]. Methods to process the collected data for large scale maps generation and 3D reconstruction of the seafloor environment can be found in [10] and [11].

A relevant byproduct of these processing stages is the availability of maps, mosaics, 3D models and more, that may be exploited as educational material for dissemination purposes, as in [12],[13]. To the best of the author's knowledge a few approaches have been devoted to the joint exploitation of the huge and heterogeneous amount of information gathered during the underwater archaeology mission.

The remaining sections of this paper will be devoted to the discussion of procedures for the manipulation of the gathered data set to the purpose of building a robust framework for object recognition purposes, based on the simultaneous integration of multiple layers of information, represented by raw data and data processing results.

3. HUNTING FOR REGULARITIES

The Autonomous Underwater Vehicle is one of the most popular platform exploited for oceans' survey purposes, an unmanned robot tool that can be programmed to perform missions without the requirement of human supervision. The survey platform must be equipped with proper tools in or- der to suitably sense the environment. To this aim optical and acoustic sensors have been installed aboard the vehicle. The selected sensors represent an optimal choice concerning mapping purposes, since they feature complementary properties in terms of large scale mapping performances (by the acoustic sensor, for a starting overview of the scene) and close range missions (by optical cameras, for the collection of high resolution detailed data).

Within the framework of the aforementioned research projects the crucial goal was to find man made artefacts located within the inspected environmnent. Since manmade objects usually exhibit regularity features in their external appearance the main approach adopted to the project purposes has been to detect salient features concerning meaningful regularity of the scene (geometrical regularity, textural regularity) and exploit the corresponding descriptor to define a label identifying quantitatively the interest ranking of the scene.

3.1. Geometrical regularity detection

As far as the detection of geometrical regularities within the inspected environment is concerned, a dedicated procedure has been implemented based on a statistical approach. The proposed algorithm takes inspiration from the work by [8], which is in turn based on the Gestalt principles modeling the human perception. Concerning these theories the main statement providing inspiration within the circumstances discussed in this paper, asserts that there must not be any perception of "structuredness" in an observation of pure noise. Based on that an algorithm has been implemented, designed as the cascade of two main steps: a candidate selection followed by a validation step. The candidate selection consists of the identification of groups of pixels in the image that, provided specific collinearity constraints are fulfilled, can be considered as potential structures featuring geometrical regularities (fragments of lines or ellipses). The selected pixels are later fed as input to the validation stage, according to which a candidate is considered to be valid if the observation of an event as structured as the considered one, is too unlikely to take place in a realization of pure noise ([6]).

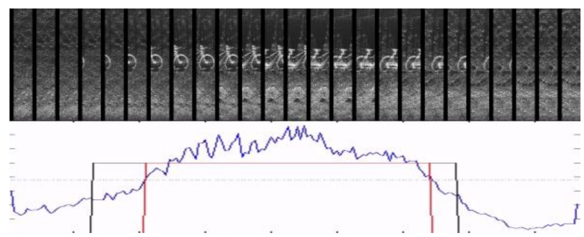

Fig. 1. Scene attentive analysis based on the geometry descriptor.

Under these assumptions the geometry detector can be efficiently exploited to build an attentive analysis system. Once a geometry descriptor has been defined for a map (for example as the average number of the detected curves, each one weighted according to their relevance) the evolution of the geometry descriptor over time can be considered as the the temporal assessment of the scene interest ranking. The mentioned procedure is illustrated in figure 1, where a clear correlation between the growth of the geometry descriptor and the presence of interesting objects can be observed.

3.2. Textural regularity detection

As mentioned above regularity perceived in the appearance of the objects' surfaces may be exploited to perform a discrimination and to discern between regions of the collected maps belonging to different classes. This can be immediately understood considering regions of the seabed that exhibit recognizable features such as sweeps of sand ripples. In that case it is easy to understand that the peculiar feature of the quasi periodical trend of the seabed could be exploited as an identifying trait of that region. Based on these considerations a procedure for the classification of seafloor areas by means of the frequency analysis of the captured maps has been implemented. In a nutshell the proposed algorithm, which takes inspiration from [9], performs an analysis of the frequency spectrum of the captured maps. This is done by convolving the 2D signal with a set of Gabor wavelet filters, properly selected in terms of center frequency and orientation. The convolution takes place in a small window centered on each pixel in the map and allows to extract a set of filter responses that constitutes a descriptive bag of features assigned to that pixel. The set of responses describes the frequency signature for the corresponding pixel and its neighbourhood, and can be exploited for the classification purpose. Indeed every image pixel is considered for a clustering step, aiming at grouping pixels according to a similarity criterion (for example the Eculidean distance minimization in the K-means algorithm) applied to the corresponding pixel descriptor, its corresponding set of Gabor convolution responses.

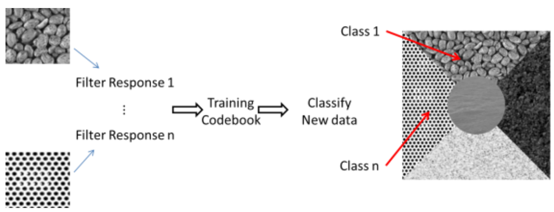

Fig. 2. Classification based on texture analysis.

The clustering step allows to assign a specific category label to each pixel, therefore it represents the final stage of a region segmentation and classification process. The classification procedure (an example in figure 2) could be further improved by introducing prior information related to previously captured data, that could be exploited to train the segmentation algorithm.

4. OVERVIEWS AND PROFILES OF THE SEAFLOOR

Although they represent consolidated literature topics, large scale map generation and 3D reconstruction methods still deserve emphasis in the scene understanding process, even more so when dealing with the issues of underwater environment inspection. This is especially true as far as the underwater archaeologist demands are pursued, e.g. to be provided with an overview of the inspected environment and to have available tridimensional information in order to put forward hypothesis on the site under exploration.

4.1. Mosaics

Mosaicking algorithms [11] represent a popular solution to the fast degradation of the optical perturbation in the water medium. Mosaicking exploits the persistence of spatial keypoints captured in multiple consecutive maps.

Fig. 3. Mosaicking output results on data collected during the ARROWS experimental mission in Levanzo Island, Sicily (Italy).

Given two images of the same scene captured under different perspectives the algorithm consists in detecting robust features (such as SIFT features [14]) and in identifying those features that match between the two maps. Given that a sufficient number of robust matchings is provided, these points may be exploited to estimate the geometrical transforma- tion that maps one image to the other. Once the maps have been transformed to a common projective space they can be stitched together to return the final enlarged map (figure 3).

4.2. 3D reconstruction

3D reconstruction represents a valuable tool towards a complete understanding of the underwater scenario, both for what concerns navigation purposes as well as the correct interpretation of the inspected site from an historical-archaeological perspective. The estimation of the 3D profile can be carried out indirectly by processing the individual frames of an optical video stream. Based on a Structure From Motion [15] approach, the 3D reconstruction algorithm starts from a preliminary processing that shares similarities with the mosaicking algorithm. Indeed the procedure concerns primarily the detection and matchings of salient features detected in multiple consecutive frames.

Fig. 4. 3D reconstruction of an amphora wreck by means of a Structure from Motion based procedure.

Then the main goal is to estimate the projective transform that maps points in the space onto different camera planes, according to epipolar geometry constraints [15]. Once the camera projections are known the 3D coordinates of the points in the space can be estimated, resulting in a point cloud (figure 4) that can be further processed to return a densified mesh representing realistically and metrically the surveyed scene.

5. DATA INTEGRATION

So far it has been presented a batch of procedures, each returning an individual description of the environment and highlighting specific features of the inspected scenario, as they are perceived by each individual sensor perspective. To the best of the author's knowledge previous research in this field suffered from a scarcity of consideration concerning the implementation of object recognition procedures based on the simultaneous manipulation of the overall available information. Actually every underwater mission deals with the is- sue of gathering huge amounts of data, referring to payload data together with the vehicle's attitude data provided by ancillary sensors suites, specifically dedicated to the navigation and positioning measurements. This amount of information is further enriched by the output results of processing algorithms applied at different levels, for example those described in the previous sections.

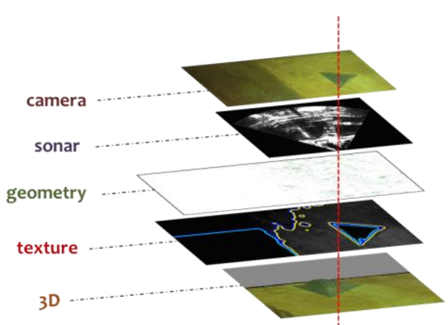

Fig. 5. Multidimensional multisensor data map.

In the framework of ARROWS and THESAURUS the authors devoted considerable effort to the investigation of novel methodologies aiming at the exploitation of the overall amount of gathered data, for object recognition purposes. More in detail a multidimensional map consisting of a layered structure of raw and processed data has been introduced. A point in that map (figure 5) returns all the available information (optical and acoustic backscatter, geometry and texture, 3D estimated coordinates, etc...) about the corresponding point in the tridimensional space.

A more rigorous framework based on a statistical approach is provided in the following. Here it is worth to remind the main goal of the underwater survey experience, that is to capture meaningful payload data and decide to which one of the two possible categories, scene with inter- esting object (*state 1*) or scene without interesting object (*state 0*), the observed state of nature corresponds to. Under these assumptions it is possible to reformulate the starting object recognition problem within a Bayes decision frame- work, where the main goal is to provide an estimate of the posterior probability of observing state 1 or state 0, given the additional information provided by the measurement of the descriptive feature f (e.g. geometry or texture). In the proposed frame the main parameters that are supposed as known are the prior probabilities of observing state 1 and state 0 and the two class-conditional probability densities of the measured feature, given that the state of nature is known. Finally, by inverting the Bayes theorem equality we obtain a numerical decision rule that allows to decide which one of the two possible states of nature is being observed (figure 6):

$$\mathit{state}\ 1 \text{ if } P(\mathit{state}\ 1|f) \geq P(\mathit{state}\ 0|f) \text{ o.w. } \mathit{state}\ 0 \quad (1)$$

6. CONCLUSIONS

This work consists of a summary description of the authors' research activity concerning underwater scene understanding issues, with applications to the mapping and preservation of underwater cultural heritage. The discussed topics represent cutting edge research that has been, to the best of the authors' knowledge, scarcely considered in the previous scientific literature. In particular a set of procedures returning meaningful description and valuable details for the understanding of the underwater scenario is provided. In addition to popular procedures concerning the processing of the data to improve the signal to noise ratio, to correct systematic geometry distortions, generate large overviews of the scene and return estimated 3D profiles, specific attention has been devoted to the implementation of methods for detecting features that carry relevant information to the purpose of manmade object detection, such as geometry and textural markers. Worthy of note is

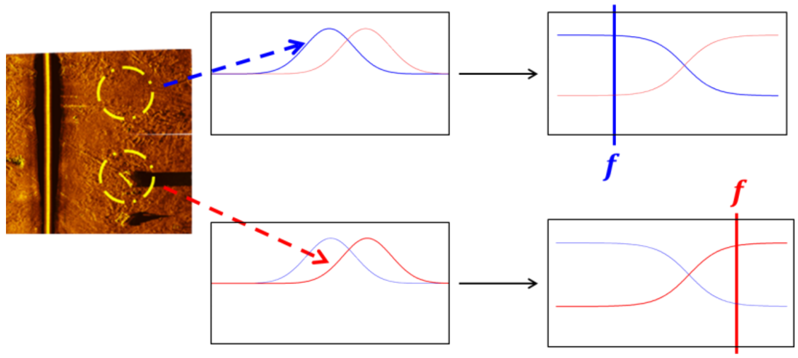

Fig. 6. Bayesian decision framework for object recognition purposes.

the introduction of a Bayes statistical framework, considered as a method to include and integrate the huge amount of captured information and the exploitation of such a tool to fulfill the identification task of interesting sites in terms of manmade artefacts presence. This work has been partially supported by the European ARROWS project, that has received fund- ing from the EU-FP7 Programme, under grant agreement no. 308724.

7. REFERENCES

[1] B. Allotta et al., " Thesaurus Project: Design of New Autonomous Underwater Vehicles for Documentation and Protection of Underwater Archaeological Sites," in *Progress in Cultural Heritage Preservation*, vol. 7616 of *Lecture Notes in Computer Science*, pp. 486–493. Springer Berlin Heidelberg, 2012.

[2] B. Allotta et al., " The ARROWS project for underwater archaeology," in *Strategie e Programmazione della Conservazione e Trasmissibilità del Patrimonio Culturale*, Italian Heritage Award 2013, pp. 80–87. Roma: Fidei Signa Edizioni Scientifiche, 2015.

[3] P. Blondel, *The Handbook of Sidescan Sonar*, Springer Praxis Book, 1st edition, 2009.

[4] D.T. Cobra, A.V. Oppenheim, and J.F. Jaffe, " Geometric distortions in side-scan sonar images: A procedure for their estimation and correction," *IEEE Journal of Oceanic Engineering*, vol. 17, pp. 252–268, 1992.

[5] Z. Zhang, " A Flexible New Technique for Camera Calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[6] M. Reggiannini, *A Cooperative Approach For Pat- tern Recognition in Underwater Scene Understanding by Multi-sensor Data Integration*, PhD dissertation, University of Pisa, 2016.

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

[8] V. Pătrăucean, P. Gurdjos, and R.G. von Gioi, " A Parameterless Line Segment and Elliptical Arc Detector with Enhanced Ellipse Fitting," in *Computer Vision ECCV 2012*, Lecture Notes in Computer Science, pp. 572–585. Springer Berlin Heidelberg, 2012.

[9] A.K. Jain and F. Farrokhnia, " Unsupervised texture segmentation using Gabor filters," in *IEEE International Conference on Systems, Man and Cybernetics*, Nov 1990, pp. 14–19.

[10] T. Nicosevici and R. Garcia, *Efficient 3D Scene Modeling and Mosaicing*, Springer Publishing Company, Incorporated, 2013.

[11] O. Pizarro, *Large scale structure from motion for autonomous underwater vehicle surveys*, PhD dissertation, MIT, 2004.

[12] M. Magrini, D. Moroni, M.A. Pascali, M. Reggiannini, O. Salvetti, and M. Tampucci, " Virtual environment as a tool to access the marine abysses," in *OCEANS 2015 - Genova*, May 2015, pp. 1–5.

[13] M. Magrini, M.A. Pascali, M. Reggiannini, O. Salvetti, and M. Tampucci, " Virtual Immersive Environments for Underwater Archaeological Exploration," in *IMTA-5 Berlin, Germany*. 2015, SciTePress.

[14] D. G. Lowe, " Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[15] R.I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, 2004.