Heterogeneous Data Registration for 3D Underwater Scene Reconstruction

Thanh Huy Nguyen, Didier Gueriot, Jean-Marc Le Caillec, Christophe Sintes, Sylvie Daniel

Heterogeneous Data Registration for 3D Underwater Scene Reconstruction

Thanh Huy NGUYEN*†, Didier GUERIOT*, Jean-Marc LE CAILLEC*, Christophe SINTES*, Sylvie DANIEL†
*Institut Mines-Telecom Atlantique - UMR CNRS 6285 Lab-STICC/CID - 29238 Brest Cedex 3 - France
†Université Laval - Québec City, QC G1V 0A6 - Canada

{thanh.nguyen, didier.gueriot, jm.lecaillec, christophe.sintes}@imt-atlantique.fr,
sylvie.daniel@scg.ulaval.ca

***Abstract*—This paper addresses the heterogeneous data registration problem, which is one of the key features for any scene reconstruction and representation, especially for the underwater environment. In this study, we propose a registration method built around a 2D-to-3D feature-based approach that registers high-resolution side-scan sonar images with bathymetric data (topographic 3D point cloud) obtained by multibeam echosounder. This process enables us to achieve a global 3D mosaic of the studied underwater scene, which is informatively richer and more reliable than each individual dataset. Indeed, the interest of this data fusion representation is that it combines the benefits of using each sensor: bathymetric information provides the geometric structure of the sea-bottom, while sidescan sonar images contribute a complementary observation with better resolution of the sea-bottom reality (e.g. sedimentology, bottom-laying object, etc.)**

***Index Terms*—heterogeneous registration, data fusion, 3D scene reconstruction, sidescan sonar, multibeam echosounder.**

I. Introduction

A. Motivation

Underwater environment is more difficult to be perceived than the Earth's land surface because of its inaccessibility to ordinary sensing and localization signals (e.g. Global Positioning System), as electromagnetic waves are attenuated extremely rapidly in this environment. As a matter of fact, acoustic waves are used as the most practical means to explore the ocean floor and produce high quality observation of the scene of interest [1].

Nevertheless, the perception of an underwater environment and its resulting exploitation require the use of multiple sensors to capture specific and informatively complementary characteristics of this environment. Since each sensor has a limited range which is usually smaller than the scope of the considered environment, it is crucial to aggregate many series of sea-bottom observations in order to produce a *global* representation of the environment. This *global* term can be interpreted in two aspects: on one hand, the spatial aspect of combining spatially complementary acquisitions; on the other hand, the thematic aspect of aggregating complementary characteristics within a representation of mutual knowledges. In this paper, we focus mainly on the second topic as we propose an approach to match data acquired by two different and heterogeneous sonar systems in order to produce a rich 3D representation of the observed underwater scene. Indeed, data acquisitions consist of sidescan sonar images (providing the information on sea-bottom sediment and contents) and multibeam echosounder sets of points (giving seabed topographical maps).

The construction of sonar image mosaics without topographic knowledges may lead to some geometrical bias because of the topography simplified assumptions [2]. As a matter of fact, the integration of topographic data is critical in performing mosaic construction, not to mention 3D mosaic construction. Therefore, heterogeneous registration and data fusion become the key features for any scene reconstruction and representation framework. However, they still remain very challenging problems, and in this paper we propose a novel feature-based approach to perform the heterogeneous registration for underwater scene reconstruction.

Building a rich and reliable representation of underwater scene has always been a major interest for various critical applications, such as drone autonomous navigation [3], surveillance mission implementation and tracking [4], [5], monitoring of environments and infrastructures [6], measuring the impacts of climate change and preventions [7], etc.

B. Related work

In [2], the authors proposed an algorithm that takes into account bathymetric information (through the DTM of the observed area) provided by a multibeam echosounder in the process of building a 3D mosaic from sidescan images. Therefore, the shape of the resulting mosaic is in agreement with the topology of the sea-bottom while its textural information comes from the sidescan images.

There are also many research works in the literature tackling the issue of scene reconstruction; however, they have focused mainly on the homogeneous registration of 3D datasets, i.e. between two 3D point clouds [8], or two 3D meshes [9]; or on the direct computation of the third dimension from 2D sonar images. For instance, [10] proposed a method consisting in detecting ridges in sidescan sonar images and using them as primitives; or [11] performed the scene reconstruction from a sequence of images.

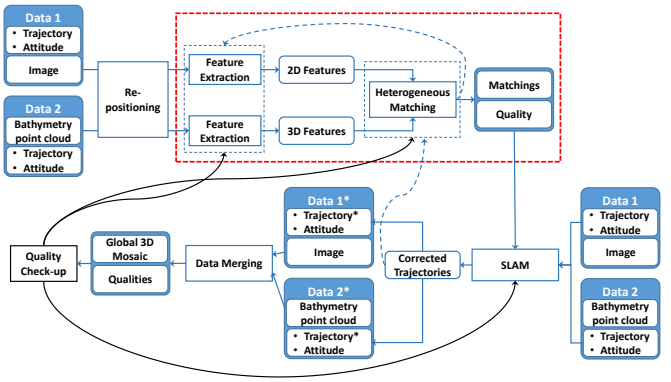


Fig. 1: Proposed framework for the underwater scene reconstruction problem. The red-dotted rectangle gives the focus of this paper within the whole framework.

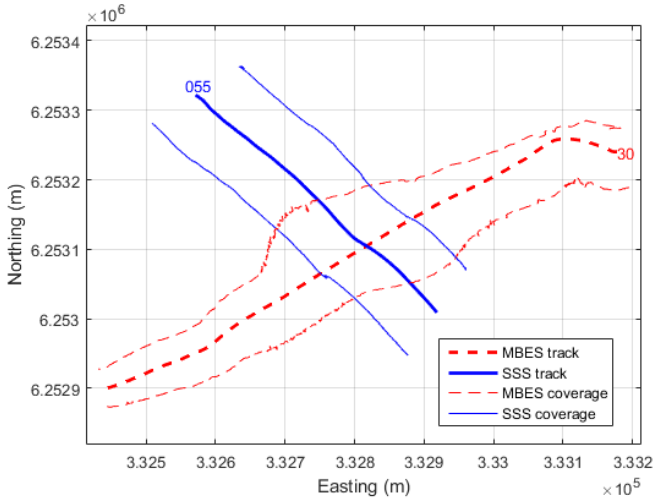


Fig. 2: Example of a sidescan sonar track (blue line) and an MBES track (red-dashed line), with their corresponding coverage which overlap partially.

C. Paper organization

The paper is organized as follows. Section II will present the global framework of underwater scene reconstruction, and emphasize the focus of our work. Then, the proposed methodology of heterogeneous registration will be carried out in Section III. In order to evaluate the accuracy and quality of proposed method, experimental results on data samples will be provided in Section IV. Finally, Section V will bring overall conclusion and perspective regarding this work.

II. Underwater Scene Reconstruction Framework

In order to deal with the complexity of underwater scene reconstruction, we propose a full framework as illustrated in Fig. 1. Five steps sum up the whole framework dynamics.

A. Repositioning of sidescan sonar's coverage within the 3D point cloud

Firstly, beginning with a sonar image on one hand and a bathymetry point cloud on the other hand, it is necessary to perform a step of coarse repositioning of the two datasets. It consists in determining the coverage overlap of the datasets and the common geographic location of the data. This step can be done by using the pre-existing navigational information of both surveys, or even using the *a priori* information of several highly remarkable features.

Indeed, in most cases, available location (provided within raw data) of sidescan sonar-carrying vehicle can be used to approximately determine its viewpoint and which part of the 3D point cloud corresponds to its observations. Figure 2 illustrates an example of the coverage overlap between a sidescan sonar image and a 3D bathymetric point cloud. Based on these geographic information, the common zone of both observations can be determined.

B. Feature extraction and heterogeneous matching

In this step, we extract the potential and remarkable features which will be the basis of our registration from both datasets on their common area, displayed in Fig. 3. Even heterogeneous, these extracted features should characterize common local behaviors of the scene impacting the recorded data from different sensors. Next, the heterogeneous matching step will measure the similarities between extracted features in order to determine whether some features pairs can be assumed and to what extent.

C. Trajectory correction

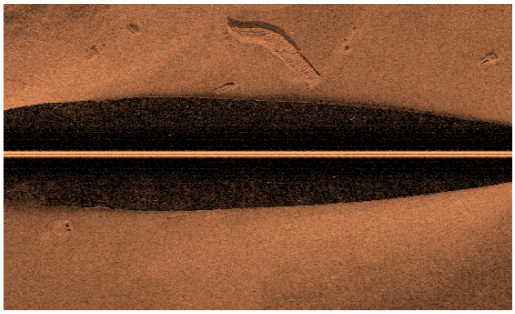
These matching results and their quality (based on similarity measures) will feed the SLAM (Simultaneous Localization And Mapping) process which estimates the best sensor trajectories allowing the previously assumed pairs of features, while maintaining the consistency of these trajectories. Corrected trajectories should then provide exact location and attitude of the acquisition sensors along the whole survey. For this step, we can reference many studies which have been effectuated on the SLAM techniques, such as [12] [13].

D. Data merging

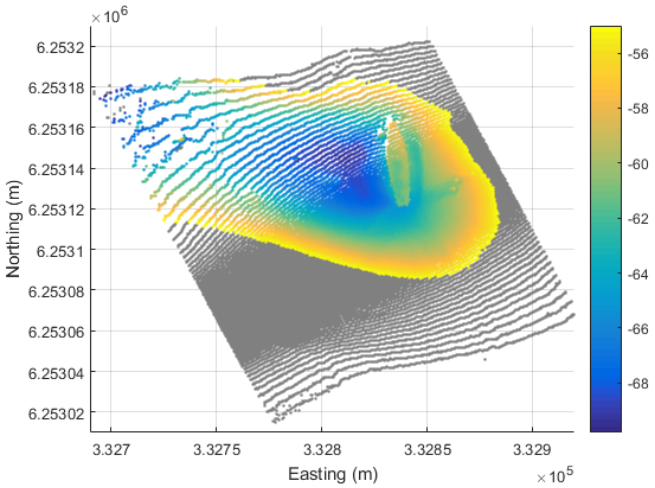
Using the corrected trajectories, every pixel from the images and every point from bathymetry data can be accurately georeferenced, allowing the data merging step. The output of this step will then be a global 3D mosaic of the considered surveys. Indeed, this mosaic will merge 2D pixel intensities from several sidescan images (where acquisitions overlap) into 3D georeferenced pixels granted by the topographical information acquired by multibeam echosounder point clouds.

E. Mosaic quality check-up and feedback

The final step is to perform a quality check-up on the resulting 3D mosaic in order to refine potential local discrepancies in the representation. Quality measurements on the produced mosaic may feedback to some of the previous steps, in order for them to adapt/adjust their behavior and improve the final 3D mosaic quality.



(a) Sidescan sonar raw image



(b) MBES point cloud (top view).

Fig. 3: Image and point cloud, respectively, acquired by sidescan sonar and multibeam echosounder in the common area between two tracks (depicted in Fig. 2). On sub-figure (b), the point cloud is unconventionally colorized in order to emphasize a shipwreck inside. Otherwise, altitude of gray points are above -55 meters

As previously said, this paper will focus mainly on the heterogeneous registration step, but knowing the whole framework process may help understand the relevancy of this specific step.

III. Proposed Heterogeneous Registration Strategy

As described, heterogeneous registration is a challenging problem, especially for underwater scene reconstruction. Sonar image resolution, range and pixel statistics heavily depend on sonar characteristics, survey settings and acquisition conditions. Moreover, different trajectories (with different viewpoints) over the same area will result in very dissimilar acquired images. Unlike multibeam echosounders, sidescan sonars are carried by cable-towed vehicles working close to the seabed. Thus, the location of these vehicles is often very inaccurate or even unavailable in some cases. Besides, as all these acquisitions are not performed simultaneously, spatial modifications of the scene will impact the acquired data. Lastly, collected data from sidescan sonar and multibeam echosounder are intrinsically different: (distance to the sonar, intensity) pair for an image pixel and (X, Y, Z) coordinates for a cloud point.

Therefore, in order to overcome these difficulties, we propose an approach consisting in extracting then matching features capturing strong altitude variations or irregularities appearing in the datasets. Indeed, any remarkable and localized topographic variation (like patches of rocks) on the sea-bottom produces a shadow region on the sidescan image due to a lack of backscattered signal over a duration corresponding to the topographic height. Estimating this local seabed height will produce a series of potential hypotheses of pairing between sidescan image and bathymetric data.

At first, under the flat sea-bottom assumptions, Eq. (1) gives an estimated height of a terrain irregularity, or a bottom-lying object, according to the length of the shadow region in one line of the image.

$$\widetilde{\Delta H} = \frac{L_{shadow}}{L_{total}} H \tag{1}$$

where L_{shadow} is the acoustic shadow length (in number of pixels of the sidescan image line) and L_{total} represents the length from the first pixel of image line (i.e. the beginning of the water column) to the last pixel of the shadow. These two lengths are also illustrated on Fig. 4c. In addition, H denotes the altitude of the vehicle above the sea-bottom, and $\widetilde{\Delta H}$ is the estimated height of which causes the shadow. This relation is built straightforwardly from trigonometric considerations (illustrated by the Fig. 4a) between the vehicle's altitude H, the object's height ΔH, and the slant ranges R_A and R_B,

$$\frac{\Delta H}{H} = \frac{R_B - R_A}{R_B} = \frac{c \times (t_B - t_A)/2}{c \times t_B/2} = \frac{L_{shadow}}{L_{total}} \tag{2}$$

where c stands for underwater sound velocity, t_A and t_B, respectively, represent the round-trip time for an acoustic signal sent from sonar transmitter to reach points A and B, then back to the sonar receiver.

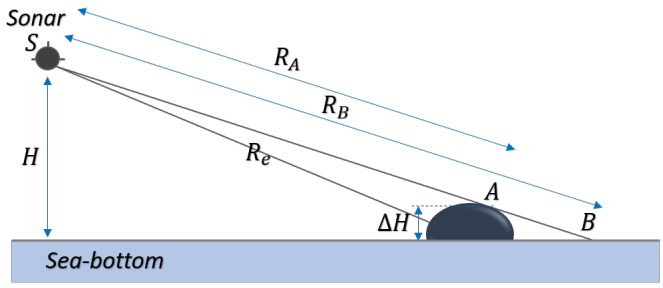
However, the sea-bottom is in general not flat, which can be considered as the altitude of ground below the vehicle is different than that of the shadow region end-point (cf. Fig. 4b). Consequently, in taking into account this altitude, which is denoted as h henceforth, Eq. (2) is extended as follows,

$$\frac{\Delta H - h}{H - h} = \frac{R_B - R_A}{R_B} = \frac{L_{shadow}}{L_{total}} \tag{3}$$

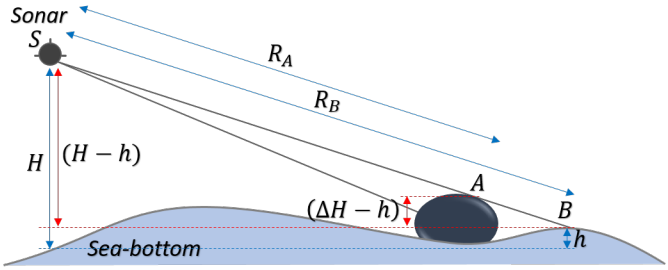
The object's height estimate then becomes,

$$\widetilde{\Delta H} = (H - h)\frac{L_{shadow}}{L_{total}} + h \tag{4}$$

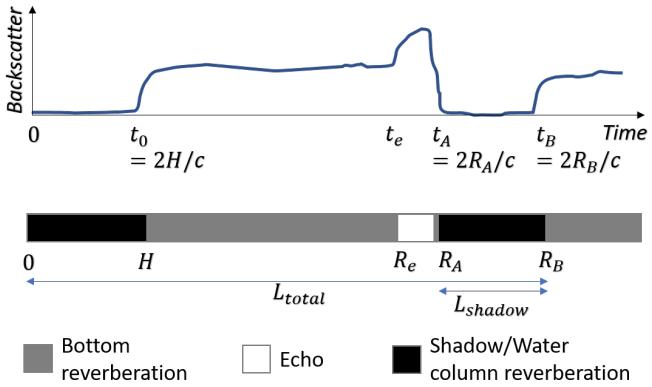
Accordingly, a detection of water column and shadow region on sonar image based on pixel intensity, and the prior knowledge of altitude H of the sonar-carrying vehicle (provided within sonar data) will enable us to estimate the height of the sea-bottom irregularity or object. Furthermore, sea-bottom remarkable features often impact more than one line of images. Thus, such a pointwise estimation may be carried out along all



(a) Acquisition geometry of sidescan sonar under flat bottom assumption.



(b) Acquisition geometry of sidescan sonar in the case of non-flat bottom.



(c) Illustration of a backscattered signal from sea-bottom with an object, and the corresponding sidescan sonar image line in an ideal gray level representation.

Fig. 4: Transversal view of a sidescan sonar acquisition (i.e. in the perpendicular plane to the towfish track) and sidescan sonar image formation.

the consecutive pings containing the studied shadow in order to build a descriptor (for instance, a cross-track elevation profile) representing a terrain elevation profile, which should have a detectable impact in the bathymetry dataset.

IV. Experimental Results

In this study, we work with the datasets acquired in the Sydney bay and part of the "Common Dataset Set" designed for the "Shallow Survey '99" conference [14], in which the image comes from a Klein 5400 sidescan sonar, whereas the bathymetry is provided by a RESON Seabat 8101 multibeam echosounder (MBES). As the datasets are composed of many regions with different properties, we particularly concentrate

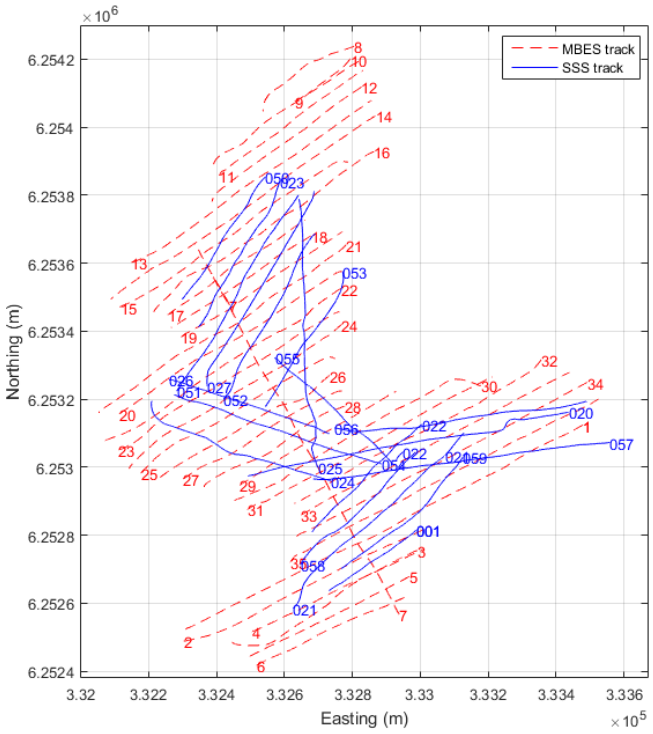


Fig. 5: Western survey trajectories of datasets used in this paper, where blue lines represent sidescan tracks (providing images) and red-dashed lines stand for MBES tracks (providing bathymetric data).

on the one with many interesting features such as shipwrecks or piles from a bridge, namely the Western region. Figure 5 depicts the trajectory of surveys from both systems on this region.

A. Height estimation from sonar image features

In order to validate the proposed algorithm, we also effectuate a processing chain to detect shadow regions in sidescan sonar images. More specifically, an image denoising is performed to remove speckle effect on sidescan sonar images. This *despecklization* relies on a specific wavelet decomposition associated with a soft thresholding on wavelet coefficients [15], [16]. Then, consistent shadow regions can be detected on images through an adaptive thresholding on image intensity, followed by a morphological opening operator that improves the smoothness of the detected shadow shape. Once the shadows are detected, the local elevation profiles will be estimated. Figure 6 demonstrates the detection and estimation results on a shipwreck observed by sidescan sonar on the common area of both sensors' coverage.

B. Taking into account the non-flat bottom impact

However, in the case of non-flat sea-bottom, the height estimate from Eq. (2) becomes irrelevant. In considering the altitude varying sea-bottom as in reality, the height estimation based on Eq. (4) can be carried out using two prior altitude

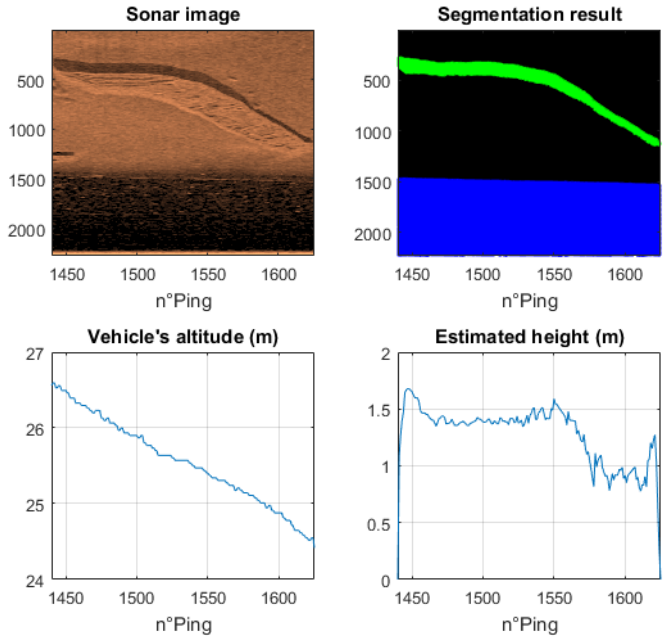


Fig. 6: Segmentation highlighting the shadow region (green) and the water column (blue), and estimated height as a function of ping number of a shipwreck, using Eq. (1) without topographic knowledge.

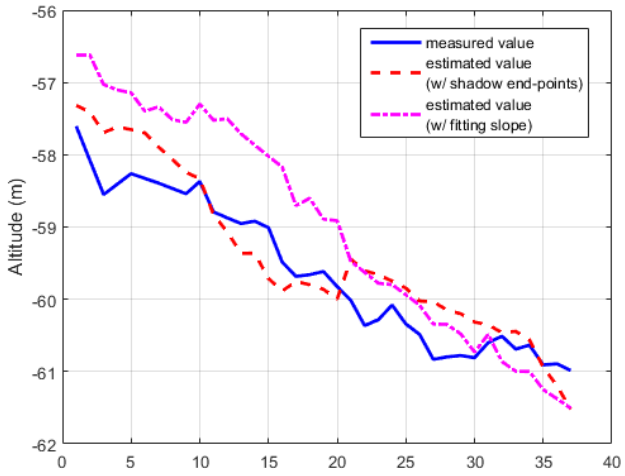


Fig. 7: Local elevation profile measured from the bathymetric 3D point cloud and estimated from sonar image (either with additional altitude information of the shadow end-point or from the fitting slope).

values h of the shadow end-point B (as illustrated in Fig. 4b) interpreted based on the bathymetry of the studied region. The first one is computed according to the form of the sea-bottom where the shipwreck lies, more specifically, by a slope fitting approach on the observed sea-bottom. For the second one, we simulate the viewpoint of sidescan sonar on the bathymetric data, then determine the ground point where the acoustic wave reaches the bottom right after hitting the object. This ground point is the end-point of the shadow region caused by the object on the sea-bottom, and its altitude is the value h of interest. Figure 7 displays the estimated elevation profiles of the shipwreck, which is displayed on Fig. 3 and examined on Fig. 6, in taking into account these two h values alongside with the local elevation profiles measured directly from the bathymetric data.

Besides, Table I summarizes three statistical results on the similarity between these estimates, which are Mean Squared Error (MSE), and correlation coefficient between each estimated curve and the measured curve, with associated p-value (associating with the probability of null hypothesis, i.e. the hypothesis that there is no relationship between two elevation values, is true). A low p-value (e.g. less than 0.05) implies that the corresponding correlation is significant, and also a low probability of observing the null hypothesis between estimated and measured curves.

$$\begin{aligned} \mathrm{MSE}(X,Y) &= \frac{1}{N}\sum_{i=1}^{N}(X_i - Y_i)^2 \\ \rho_{XY} &= \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \end{aligned} \tag{5}$$

TABLE I: Similarity measure between the measured and estimated elevation profiles.

	MSE	**Correlation coefficient ρ**	***p*-value**
Slope-fitting	0.7408	0.9763	7.5E-25
Shadow end-point	0.2118	0.9268	1.9E-16

However, as these two altitude values are computed based on the simulation of sidescan sonar's location and viewpoint on the bathymetry, there are two main reasons making this simulation become less accurate. Yet they explain partially why both estimated heights are slightly dissimilar from the measured profile.

- Location of the sonar-carrying vehicle is uncertain, as it is underwater and towed to the vessel by cable;
- Low spatial resolution of the bathymetric 3D point cloud makes it difficult to have a good measurement of h. As a matter of fact, for shipwreck displayed on Fig. 6, we have only 37 points along its length (or in other words, an average of 1.15 meters between two points).

In addition, Fig. 8 comparing two members of Eq. (3) highlights another source of difference between local elevation profiles. Indeed, the quality of the processing chain impacts the sonar image and, as such, plays an important role in the height estimation accuracy. Similarity measures between ratios L_{shadow}/L_{total} and $(\Delta H - h)/(H - h)$ respectively are also shown on Table II.

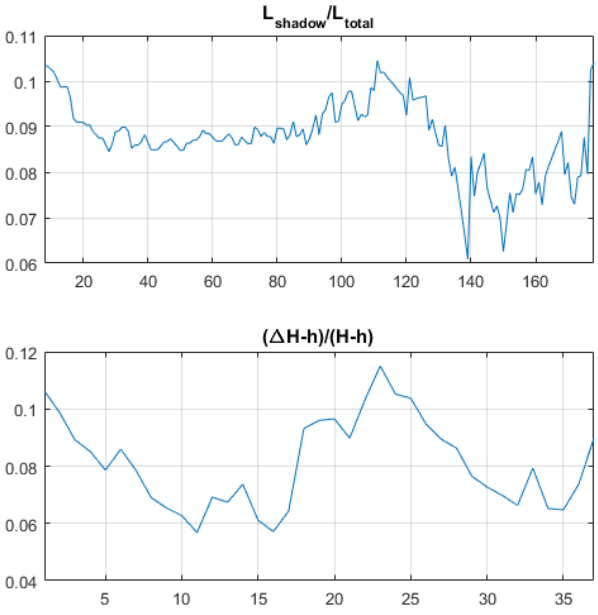


Fig. 8: Comparison between $^{L_{shadow}}/_{L_{total}}$ and $^{(\Delta H-h)}/_{(H-h)}$ calculated from the sonar image and the bathymetric point cloud, respectively.

TABLE II: Similarity measure between two ratios $^{L_{shadow}}/_{L_{total}}$ and $^{(\Delta H-h)}/_{(H-h)}$.

MSE	**Correlation coefficient** ρ	***p*-value**
3.73E-4	0.6559	1.05E-5

V. Conclusion and Perspective

In order to take advantage of the strength of two heterogeneous sensors in the process of reconstructing the underwater scene, this paper has proposed a method to register data from these sensors, which are intrinsically different and also consist of spatial and temporal variabilities. The ability to retrieve the altitude information from sonar image features (especially shadow regions) has also been demonstrated, as well as indicating their coherence with the measured elevation profile. Furthermore, despite being incompletely demonstrated in this paper, the proposed framework has shown that its main step has potential with very promising results.

In this paper, we have performed an algorithm to retrieve the height of bottom-lying object or terrain irregularity based on the characteristics of its shadow on sonar image. In other words, we have used the information from the region behind the object in the insonification direction away from sidescan sonar. Besides, this pointwise estimated height relates only to the highest point of object that causes the shadow; hence, it is unable to provide the object's shape. In the future, our algorithm may be supplemented by a Shape-From-Shading [17] algorithm which yields altitude information in accumulating the relief from the forefront of the object.

Moreover, as the SLAM process needs a reduced number of pairs of features, this initial iteration of feature extraction should be sufficient at first. Then, on the following iterations, more compatible profiles should be progressively introduced to this process. When lacking of robust pairs of matched profiles, additional features can be selected for the registration process (e.g. altitude information using Shape-From-Shading algorithm guided by our previously assumed pairs of pointwise features).

References

[1] X. Lurton, *An introduction to underwater acoustics: principles and applications*. Springer Science & Business Media, 2002.

[2] D. Gueriot, "Bathymetric and side-scan data fusion for sea-bottom 3d mosaicing," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No.00CH37158)*, vol. 3, 2000, pp. 1663–1668 vol.3.

[3] J. J. Leonard and A. Bahr, "Autonomous underwater vehicle navigation," in *Springer Handbook of Ocean Engineering*. Springer, 2016, pp. 341–358.

[4] A. Gad and M. Farooq, "Data fusion architecture for maritime surveillance," in *Information Fusion, 2002. Proceedings of the Fifth International Conference on*, vol. 1. IEEE, 2002, pp. 448–455.

[5] D. Smith and S. Singh, "Approaches to multisensor data fusion in target tracking: A survey," *IEEE transactions on knowledge and data engineering*, vol. 18, no. 12, pp. 1696–1710, 2006.

[6] R. van Overmeeren, J. Craeymeersch, J. van Dalfsen, F. Fey, S. van Heteren, and E. Meesters, "Acoustic habitat and shellfish mapping and monitoring in shallow coastal water–sidescan sonar experiences in the netherlands," *Estuarine, Coastal and Shelf Science*, vol. 85, no. 3, pp. 437–448, 2009.

[7] A. C. Baker, P. W. Glynn, and B. Riegl, "Climate change and coral reef bleaching: An ecological assessment of long-term impacts, recovery trends and future outlook," *Estuarine, coastal and shelf science*, vol. 80, no. 4, pp. 435–471, 2008.

[8] V. Murino, A. Fusiello, N. Iuretigh, and E. Puppo, "3D mosaicing for environment reconstruction," in *Proceedings of the International Conference Pattern Recognition*, vol. 3, 2000, pp. 358–362.

[9] M. Johnson-Roberson, O. Pizarro, and S. Williams, "Towards three-dimensional heterogeneous Imaging sensor correspondence and registration for visualization," in *Oceans 2007 - Europe, Vols 1-3*, 2007, pp. 906–911\r1581.

[10] E. Coiras, Y. Petillot, and D. M. Lane, "Multiresolution 3-D reconstruction from side-scan sonar images," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 382–390, 2007.

[11] N. Brahim, D. Guriot, S. Daniel, and B. Solaiman, "3d reconstruction of underwater scenes using didson acoustic sonar image sequences through evolutionary algorithms," in *OCEANS 2011 IEEE - Spain*, June 2011, pp. 1–6.

[12] M. Kaess, A. Ranganathan, and F. Dellaert, "isam: Incremental smoothing and mapping," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1365–1378, 2008.

[13] S. Reed, I. Tena Ruiz, C. Capus, and Y. Petillot, "The fusion of large scale classified side-scan sonar image mosaics," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2049–2060, 2006.

[14] "DSTO Australia Defense Science & Technology Organization Common data set," Shallow Survey 99 - International Conference on High Resolution Surveys in Shallow Water, Oct. 1999.

[15] D. L. Donoho, "De-noising by soft-thresholding," *IEEE transactions on information theory*, vol. 41, no. 3, pp. 613–627, 1995.

[16] C. Sintes, R. Garello, and D. Gueriot, "Interferometric signal denoised by wavelets," in *OCEANS 2006*, Sept 2006, pp. 1–6.

[17] E. Prados and O. Faugeras, "Shape from shading," in *Handbook of mathematical models in computer vision*. Springer, 2006, pp. 375–388.