

Chapter 14

Underwater Multimodal Survey: Merging Optical and Acoustic Data

Pierre Drap, Djamel Merad, Jean-Marc Boï, Amine Mahiddine, Daniela Peloso, Bertrand Chemisky, Emmanuelle Seguin, Frederic Alcala and Olivier Bianchimani

Abstract ROV 3D project aims at developing innovative tools which link underwater photogrammetry and acoustic measurements from an active underwater sensor. The results will be 3D high resolution surveys of underwater sites and landscapes useful to keep in memory cultural and natural heritage. The new means and methods developed aim at reducing the investigation time in situ, and proposing comprehensive and non-intrusive measurement tools for the studied environment.

P. Drap (✉) · D. Merad · J.-M. Boï · A. Mahiddine · D. Peloso
LSIS Laboratoire des Sciences de l'Information et des Systèmes,
UMR CNRS 6168, I&M Team, Images & Models,
Aix-Marseille University Marseille, France
e-mail: Pierre.Drap@univ-amu.fr

D. Merad
e-mail: Djamel.Merad@univ-amu.fr

J.-M. Boï
e-mail: Jean-Marc.Boi@univ-amu.fr

A. Mahiddine
e-mail: Amine.Mahiddine@univ-amu.fr

D. Peloso
e-mail: Daniela.Peloso@univ-amu.fr

B. Chemisky · E. Seguin · F. Alcala
COMEX SA, 36 Boulevard des Océans,
BP 143, 13275 Marseille Cedex 9, France
e-mail: b.chemisky@comex.fr

E. Seguin
e-mail: e.seguin@comex.fr

F. Alcala
e-mail: f.alacala@comex.fr

O. Bianchimani
Septentrion Environnement,
Maison de la Mer—Plage du Prophète,
244 Corniche Kennedy, 13007 Marseille, France
e-mail: olivier.bianchimani@septentrion-env.com

In this paper, we are presenting a new method of 3D surveys which are dedicated to high resolution modeling of underwater sites. The main met constraints in situ are taken into account and this method leads to a precise 3D reconstruction. Some examples will present both the main obtained results and their limitations. We will end with the perspectives and the necessary improvements to the method, so as to automate the multimodal registration step.

Keywords Photogrammetry · 3D · Wreck · Digital photographs · Automatic Color Equalization

14.1 Introduction and Main Goals

ROV 3D project is partially funded by the European Regional Development Fund and the French Single Inter-Ministry Fund (FUI) for funding research involving both academic laboratories and industry.

The consortium consists of an organization of academic research, LSIS laboratory, and of two industrial partners, COMEX and SETP. COMEX is skilled in underwater scouting and outdoor engineering, whereas SETP is skilled in dimensional control by photogrammetry and topometry.

At the present time, and in the field of high resolution underwater surveys, there is no industrialized automatic treatment. Although, nowadays, certain private businesses propose service charge in underwater sharp metrology by photogrammetry, these offers use traditional methods of close range photogrammetry, based on automatic recognition of coded targets and bundle adjustment. These approaches, precise as they are, need human interference on the object and a site preparation (target laying in the place you want to measure), which, in the underwater context, may be a great handicap.

The project's goal is to develop automated proceedings of 3D surveys, dedicated to underwater environment, using acoustic and optic sensors. The acoustic sensor will allow acquiring a great amount of low resolution data, whereas the optic sensor (close range photogrammetry) will allow acquiring a low amount of high resolution data. From these two surveys, a high resolution numeric structure build 3D modeling of large complex scenes will be proposed to final users, with the production of different type of outputs (SIG 3D, MNT, Mosaic, . . .). These models can be analyzed and compare at any time and can be used to study sites or landscapes' evolution in time.

In practice, a 3D acoustic scanner will make a large scan of the scene to model, and an optic system will make the high resolution photogrammetric restitution of the different areas in the scene. While our software tools will do the automatic registration of both data sources, other algorithms, developed during the project, will recognize and model objects of interest. Eventually, these data will allow us to establish objects' symbolic representation and their geometry in a precise virtual facsimile, responding to the partners' need in documentation.

The ability to measure and model large underwater sites in a short time opens up many scientific challenges such as image processing, multimodal adjustment, land visualization and offers new opening to marine biology, underwater archaeology and underwater industry (offshore, harbour industry, etc.).

14.1.1 Underwater Image Processing: State of Art

The underwater image pre-processing can be addressed from two different points of view: image restoration techniques or image enhancement methods.

Fan et al. proposed a restoration method based on blind deconvolution and the theory of Wells (Fan et al. 2010). As a first step an arithmetic mean filter is used to perform image denoising, and then an iterative blind deconvolution using the filtered image is carried out. The calculation of the PSF of water is done using the following equations:

$$b = c\omega \quad (14.1)$$

$$H_{medium}(\psi, R) = \exp \left\{ -cR + bR \left[\frac{1 - \exp(-2\pi \theta_0 \psi)}{2\pi \theta_0 \psi} \right] \right\} \quad (14.2)$$

Where θ_0 is referred to the median scattering angle, ψ is the spatial frequency in cycles per radian, R is distance between sensor and object, b scattering coefficient, c attenuation coefficient and albedo ω .

Image restoration techniques need some parameters such as attenuation coefficients, scattering coefficients and depth estimation of the object in a scene. For this reason in our works, the preprocessing of underwater image is devoted to image enhancement methods, which do not require a priori knowledge of the environment.

Bazeille et al. (Bazeille et al. 2006) proposed an algorithm to enhance underwater image, this algorithm is automatic and requires no parameter adjustment to correct defects such as non-uniform illumination, low contrast and muted colors.

In this algorithm which is based on the enhancement, each disturbance is corrected sequentially. The first step is to remove the moiré effect is not applied, because in our conditions this effect is not visible. Then, a homomorphic filter or frequency is applied to remove the defects of non-uniformity of illumination and to enhance the contrast in the image.

Regarding the acquisition noise, often present in images, they applied a wavelet denoising followed by anisotropic filtering to eliminate unwanted oscillations. To finalize the processing chain, a dynamic expansion is applied to increase contrast, and equalizing the average colors in the image is being implemented to mitigate the dominant color. Fig. 14.1 shows the result of applying the Bazeille et al. algorithm.

To optimize the computation time, all treatments are applied on the component Y in $YCbCr$ space. However the use of homomorphic filter changes the geometry, which will add errors on measures after the 3D reconstruction of the scene, so we decided not to use this algorithm.

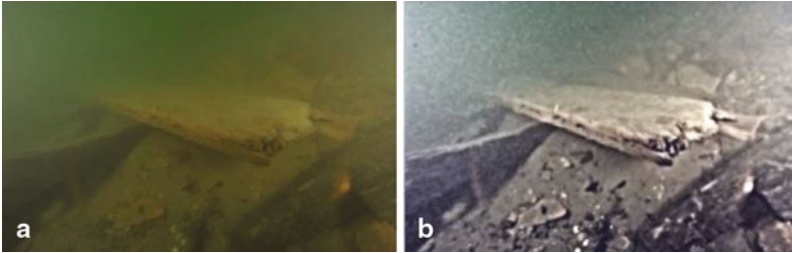
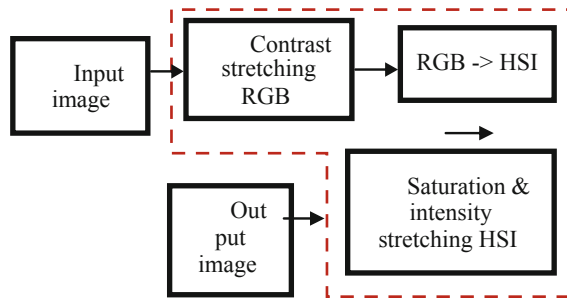


Fig. 14.1 Images before (a) and after (b) the application of the algorithm proposed by Bazeille et al. (Photo by Olivier Bianchimani (All rights reserved) on the Arle-Rhone 13 roman wreck in Arles, France)

Fig. 14.2 Algorithm proposed by Iqbal et al. (2007)



Iqbal et al. have used slide stretching algorithm both on RGB and HIS color models to enhance underwater images (Iqbal et al. 2007). There are three steps in this algorithm (see Fig. 14.2).

First of all, their method performs contrast stretching on RGB and then it converts the result from RGB to HSI color space. Finally, it deals with saturation and intensity stretching. The use of two stretching models helps to equalize the color contrast in the image and also addresses the problem of lighting.

Chambah et al. proposed a method of color correction based on the ACE model (Rizzi and Gatta 2004). ACE “Automatic Color Equalization” is based on a new calculation approach, which combines the Gray World algorithm with the Patch white algorithm, taking into account the spatial distribution of information color. The ACE is inspired by human visual system, where is able to adapt to highly variable lighting conditions, and extract visual information from the environment (Chambah et al. 2004).

This algorithm consists of two parts. The first one consists in adjusting the chromatic data where the pixels are processed with respect to the content of the image. The second part deals with the restoration and enhancement of colors in the output image (Petit 2010). The aim of improving the color is not only for better quality images, but also to see the effects of these methods on the SIFT or SURF in terms of their feature points detection. Three examples of images before and after restoration with ACE are shown in Fig. 14.3.

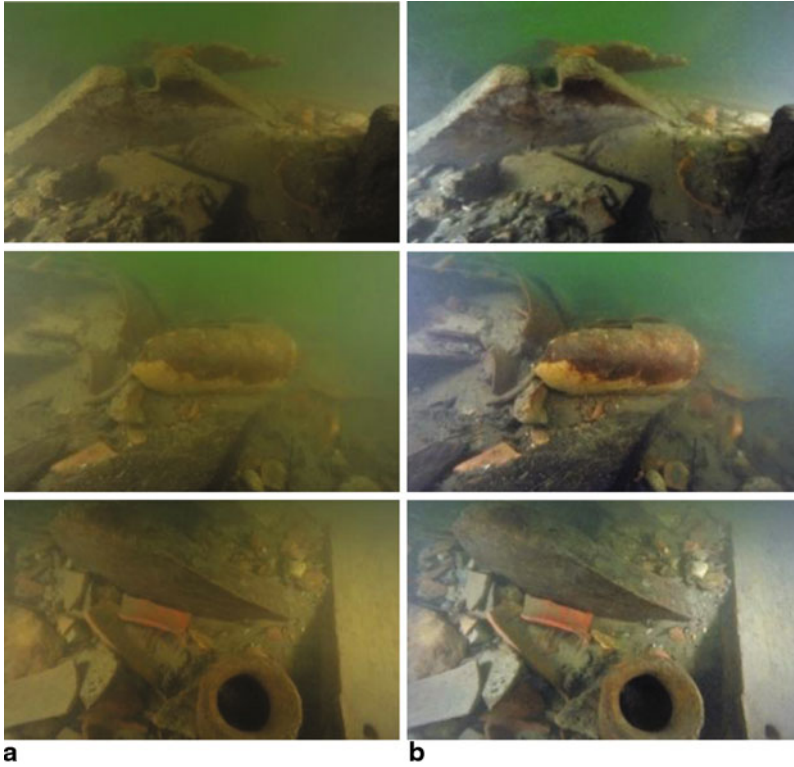


Fig. 14.3 Photographs of the wreck Arles-Rhône 13 (photo Olivier Bianchimani, all rights reserved), before (a) and after (b) the enhancement by ACE method. (Chambah et al. 2004)

Kalia et al. (Kalia et al. 2011) investigated the effects of different image pre-processing techniques which can affect or improve the performance of the SURF detector (Bay et al. 2008). And they proposed new method named IACE ‘Image Adaptive Contrast Enhancement’. They modify this technique of contrast enhancement by adapting it according to the statistics of the image intensity levels.

If P_{in} is the intensity level of an image, it is possible to calculate the modified intensity level P_{out} with Eq. (14.3).

$$P_{out} = \frac{(P_{in} - c)}{(d - c)} \times (b - a) \quad (14.3)$$

where a is the lowest intensity level in the image and equal to 0, b is its corresponding counterpart and equal to 255 and c is the lower threshold intensity level in the original image for which the number of pixels in the image is lower than 4% and d is the upper threshold intensity level for which the number of pixels is cumulatively more than 96%. These thresholds are used to eliminate the effect of outliers, and improve the intrinsic details in the image while keeping the relative contrast.



Fig. 14.4 Photographs of the wreck Arles-Rhône 13 (photo Olivier Bianchimani, all rights reserved), **a** original images, **b** results by ACE method, **c** results by IACE method “Image Adaptative Contrast Enhancement”, **d** results by the method proposed by Iqbal et al. (2007)

The results of this algorithm are very interesting. One can observe that the relative performance of IACE method is better than the method proposed by Iqbal et al. in terms of time taken for the complete detection and matching process see Fig. 14.4.

14.1.2 Underwater Object Recognition

Pattern recognition process is complex and approaches to solve it are very different, depending on whether we have an a priori knowledge of the object or not, depending on the type and the number of used sensors (one or more 2D cameras, 3D cameras, telemeter, etc.), depending on the type of object to be detected (2D, 3D, random

form, etc.). Nevertheless, there are two main families of methods to build a pattern recognition system: structural methods and statistical methods.

The first application related to our project is red coral monitoring. From its tentacular form, we are aiming at developing a structural approach which uses the objects' median (skeleton) axis as form descriptor.

There are many applications for 2D and 3D objects' skeletons in image processing (encoding, compression, . . .) and in vision in general (Merad et al. 2006; Thome et al. 2008). Indeed, we retrieve in the object's skeleton its topological structure; moreover, most of the information which are contained in the form's silhouette can be retrieved in the skeleton. Another advantage that cannot be denied is the fact that, by nature, skeleton have a graph structure. Hence, after the encoding of the coral's form under a graph structure through a 3D skeletisation process, we are going to use powerful skills from the graph theory so as to complete the matching (Shokoufandeh et al. 2005).

The second application consists in archaeologist objects recognition on an underwater site. Because of the a priori information we have, such as the type of the object (amphora, bottle, etc . . .), we will choose a statistic recognition method. (Baluja and Rowley 2005)

In an environment like wreck in 40 m deep, the vision conditions are strongly damaged. It is then necessary to free ourselves from preliminary treatments such as edge detection, line detection and other structural primitive.

Recent works showed the interest in using learning methods like adaboost, see (Freund and Schapire 1997). The advantage of this kind of methods is to only need low level descriptors such as pixels' (Baluja and Rowley 2005). LBP's (Ahonen et al. 2006), Haar's (Viola and Jones 2001), etc.

For our problem, we are going to implement an Adaboost classifier; SIFT (Lowe 2004) and/or SURF (Bay et al. 2008) will stand for weak learners. We will check the relevancy of this method by comparing its results to other standard classifiers'.

14.1.3 Merging Optical and Acoustic Data: State of Art

Related Work Optic and acoustic data fusion is an extremely promising technique for mapping underwater objects that has been receiving increasing attention over the past few years (Shortis et al. 2009). Generally, bathymetry obtained using underwater sonar is performed at a certain distance from the measured object (generally the seabed) and the obtained cloud point density is rather low in comparison with the one obtained by optical means.

Since photogrammetry requires working on a large scale, it therefore makes it possible to obtain dense 3D models. The merging of photogrammetric and acoustic models is similar to the fusion of data gathered by a terrestrial laser and photogrammetry. The fusion of optical and acoustic data involves the fusion of 3D models of very different densities—a task which requires specific precautions (Drap and Long 2005; Hurtós et al. 2010).

Only a few laboratories worldwide have produced groundbreaking work on optical/acoustic data fusion in an underwater environment. See for example (Singh et al. 2000) and (Fusiello and Murino 2004) where the authors describe the use of techniques that allow the overlaying of photo mosaics on bathymetric 3D digital terrain maps (Nicosevici et al. 2009). In this case we have important qualitative information coming from photos, but the geometric definition of the digital terrain map comes from sonar measurements.

Optical and acoustic surveys can also be merged using structured light and high frequency sonar as by Chris Roman and his team (Singh et al. 2007). This approach is very robust and accurate in low visibility conditions but does not carry over qualitative information.

Merging Cloud of Points Merging between different sources is also one of the basic problems in computer vision, pattern recognition and robotics.

The fact that the correspondence points between two point clouds are unknown a priori makes the task of merging difficult and interesting in the same time. These points of correspondence are useful to estimate the position and orientation of one point cloud compared to another in the same coordinate system.

The methods used in this topic are often variants of the ICP (Iterative Closest Point) which has been proposed by Besl and McKay (Besl and McKay 1992) and which remains the most used in the majority of software for automatic registration between two models.

This method converges to the first local minimum which is due to the outliers of the matching. Several solutions have been implemented to solve this problem. Chen and Medioni (1991) suggest replacing the measurement of distance between points, which is used in the original method, by measuring the distance between a point and a tangential plane which makes the algorithm less sensitive to local minima.

Rusinkiewicz and Levoy (2001) provide a comparison between several variants of the standard algorithm in terms of convergence time. The authors also proposed an optimized method. The idea behind this method is to classify points in the direction of their normal, then sampling on each class and reject the outliers.

The objective of these methods is to calculate the rigid transformation between two partially overlapping point clouds. Their purpose can be decomposed into two parts; the first part is the matching between points of two clouds. The second part is the estimation of the transformation 3D.

Assuming that we have a set of matching points: $\{P_i\}$ and $\{P'_i\}$ with $i = 1, 2, \dots, N$.

$$\{P'_i\} = R \times P_i + t$$

where R is the rotation matrix and t is the translation.

To find the correct transformation between the two point clouds, we must find a solution that minimizes the least squares error.

$$err = \sum_{i=1}^N R P_i + t - P'_i{}^2$$

To solve this problem we can use the singular value decomposition (SVD) of the covariance matrix C , which is time-efficient and especially easy to use.

$$C = \sum_{i=1}^N P_i - \text{centroid}_p \times (P'_i - \text{centroid}_{p'})^T$$

$$U, S, V = \text{SVD}(C)$$

$$R = VU^T$$

$$t = -R \times \text{centroid}_p + \text{centroid}_{p'}$$

With knowledge of how to compute the transformation between the two sets of matching points, the problem of merging data from different sources is summarized in the detection of matching points.

Many techniques have been proposed in literature to find the signature of an object at the scene or description of each point compared to its neighborhood.

Sehgal et al. (2010) proposed to use the SIFT feature detector (Lowe 2004) on 3D data, knowing that the use of this algorithm was reserved only for the detection of interest points in 2D images. They projected each point on a plane to form an image and the intensity of each pixel being the distance between the point and the plane. This method requires dense pixel information to use SIFT whereas the points from data are sparsely distributed.

Johnson and Herbert (Johnson 1997) introduced the notion of spin images to represent surfaces. Each spin image is a local descriptor of a surface to a point defined by its position and normal. This method requires that the two point clouds be of uniform resolution. Mian et al. (2006) used tensors to describe partial surfaces. They showed that this method is more efficient than spin images in terms of recognition rate and efficiency.

Sahillioglu and Yemez (2010) proposed an automatic technique to find correspondences between isometric shapes. They divided the data source and target into surface patches of equal area with each patch represented by the point at its center. This method needs an initial correspondence and the data must be isometric and represented as manifold meshes, which is not the case for data acquired by optical or acoustic sensors.

Rusu et al. (2008) introduced a new point signature which described the local 3D geometry. The signature was presented like a histogram with its invariant to position, orientation and point cloud density. First they estimated the surface normal at each point. Then they computed the histogram using the method proposed in Wahl et al. (2003).

Most of methods in literature use data taken from the same sensors or from CAD tools, but the main issues encountered in our topic are that data which come from different sources (optical and acoustic sensors) have different scales and different resolutions. That is why we try to find a method that will be less susceptible to these issues.

14.2 Photogrammetry: Dense Map

To model the environment by photogrammetry in an unsupervised way it is first necessary to automatically orient a set of unordered images. This orientation phase, which is crucial in photogrammetry, as computer vision, has seen the past three years a great boom. The problem was first solved in the case of ordered series of photographs, for example, made a circle around an object and recently in the case of photographs unordered (Barazzetti et al. 2010; Shunyi et al. 2007). Once all the photographs are oriented several methods are proposed for producing a dense cloud of 3D points to represent the area photographed.

Two major families of methods exist. Those that use solid models as the voxels (Furukawa and Ponce 2010; Zeng Gang et al. 2005) are based on the discretization of space into cells and the goal is to discriminate between full and empty cells to define the boundary between them. The advantage of this method is to use lots of photographs taken from arbitrary viewpoints. In contrast the delicacy of the final model depends on the resolution of the voxel grid can be RAM consuming. On the other hand methods using meshes could adapt their resolution to better reconstruct the details of the scene (Morris and Kanade 2000).

Also the work of Furukawa and Ponce on the dense map generation (Agarwal et al. 2010; Furukawa and Ponce 2010) have also resulted in open source publications. We use this work, merged with our own photogrammetric software in order to use calibrated camera and some constraints on the bundle adjustment, for several months and some examples are presented in this paper.

These developments were coupled with a software bundle adjustment of operating on an unordered set of photographs are based on an implementation of SIFT on GPU due to Changchang Wu, University of Washington (<http://cs.unc.edu/ccwu/siftgpu>) and implantation of developments PhotoTourism (Agarwal et al. 2010; Snavely et al. 2010). The bundle adjustment used is based on the Sparse Bundle Adjustment of Lourakis (Lourakis and Argyros 2009) and a version of the bundle adjustment of a set of unordered photographs is available in open source: Bundler software (Snavely et al. 2010).

On the other hand, since 2007 IGN (Institut Géographique National, in France) has decided to publish in open source the APERO MICMAC software, dedicated to the automatic orientation of an unordered set of photographs and the calculation of the automatic mapping on a set of photographs oriented. (<http://www.micmac.ign.fr>) (Pierrot-Deseilligny and Cléry 2011). Their approach is more rigorous from a photogrammetric point of view and allows using calibrated camera.

14.3 Acoustic Survey

The acoustic survey is mainly useful to produce a 3D model of the complete submarine site with a good precision in an absolute coordinate system.

To perform this survey an acoustic scanner 3D is used which can produce multiple 3D points with a same couple (x, y) coordinates, in opposite of standards bathymetric echo sounder.

This particularity is a essential to create a model of complex structures like walls, caves, overhangs . . .

The BlueView BV5000 system used is composed of a high frequency sonar (1,35 MHz), mounted on a pan & tilt system. Both Sonar and pan & tilt are managed from the surface by the dedicated software.

The system works by mechanically scanning a thin vertical sonar slice around the selected area. At each direction, a profile of the surface is taken and added to the other direction profiles to create a final 3D point cloud.

Through mechanical rotation of the sonar head, the BV5000 is capable of producing 3D points from a stationary location

To cover all the structure with enough density of points, multiple stations can be realized around the site and merged after by algorithms like ICP (iterative closest points).

Some reference points can be located with high accuracy in absolute coordinate system to georeferenced the final model.

The sonar head is composed of 256 beams for an aperture of 45 °. The beam width is 1 ° per 1 ° and time resolution is 3 cm. The maximum range is 30 m to be optimal between 1 and 20 m.

14.4 Experimentations

We present here the first experimentations of the ROV-3D project. The first one is merging acoustic data with photogrammetry in a cave close to Marseille (Fig. 14.7). And the second one is a survey of a modern wreck, “le Liban” also close to Marseille. This second survey is done only with photogrammetry with more than 1,00 photographs.

14.4.1 *Close Range Acoustic and Optical Survey*

Under the ROV-3D project we experimented with the fusion of acoustic data from two sensors, high precision, high frequency acoustic camera sold by Blue View and the photogrammetry system automatically MENCISOCIETY (see Fig. 14.5). An experiment was done on the Cave of the Imperial land in Marseilles.

The 3D Scanner is a BlueView BV5000 of active acoustic system that provides a point cloud of 3D high-resolution imagery of underwater sites, see the scanner during the acquisition phase in Fig. 14.6.

Unlike conventional bathymetric measurement systems that retain only the high points, the 3D sonar, installed near the bottom, can acquire and maintain multiple points of elevation Z for a given pair of coordinates (X, Y).

This system “3D Scanner” opens up new possibilities for constructing 3D models of complex structures such as drop offs, overhangs, or even caves.

The “3D scanner” is mounted on a hinged support along 2 axes (horizontal and vertical) allowing rotation from top to bottom and from right to left.

Fig. 14.5 The acoustic camera in situ, just in front of the cave—see next image—(photo Bertrand Chemisky, all rights reserved)



Fig. 14.6 The underwater cave close to Marseille, “L’Impérial de terre” 30 m depth (photo Pierre Drap, all rights reserved)



With an acoustic aperture of 45° of the scanner itself and the system is capable of measuring an area of 45° or 360° on a spherical surface comprising the whole environment surrounding it on the scanner and a range up to **30 m**. In the latter case, the rotation along the vertical axis that accumulated along the horizontal axis allows 4 or 5 scans cover the entire hemisphere to be measured.

Each scan or set of scans are performed in a fixed position and using such a tripod. To obtain a sufficiently dense cloud of points on the stage and studied by size, and multiple stations can be performed and the results merged.

Point clouds generated that were acquired from fixed stations, algorithms such as ICP (Iterative Closest Point) can assemble the sets of points against each other. It then remains to position the merged point cloud in an absolute reference, using reference points whose coordinates can be determined by acoustic positioning systems such USBL (Ultra Short Baseline).

Two dives were devoted to the acquisition of 3D data with the acoustic scanner.

Two more dives were used achieve the photogrammetry survey, see Figs. 14.7 and 14.8. This system consists of a hardware part, three cameras on a linear, calibrated.

Fig. 14.7 The photogrammetric survey using three digital camera synchronized (photo Pierre Drap, all rights reserved)



Fig. 14.8 Merging optical and acoustic underwater data



The shots are processed by a synchronous software house, which on one hand trying to calculate beam by adjusting the largest block possible, i.e. the block containing the most triplets as possible, then once the block is calculated, the triplets are used to obtain a depth map of the central apparatus producing a dense cloud of 3D point. The adjustment of the beam is therefore only there to ensure the cohesion of these triplets, the 3D points to them being calculated using a single triplet.

The set of photographs were processed using the software from Menci company and also the pipeline defined by a SIFT version, the Bundler (Snavely et al. 2010) and finally the dense map process proposer by Furukawa and Ponce.

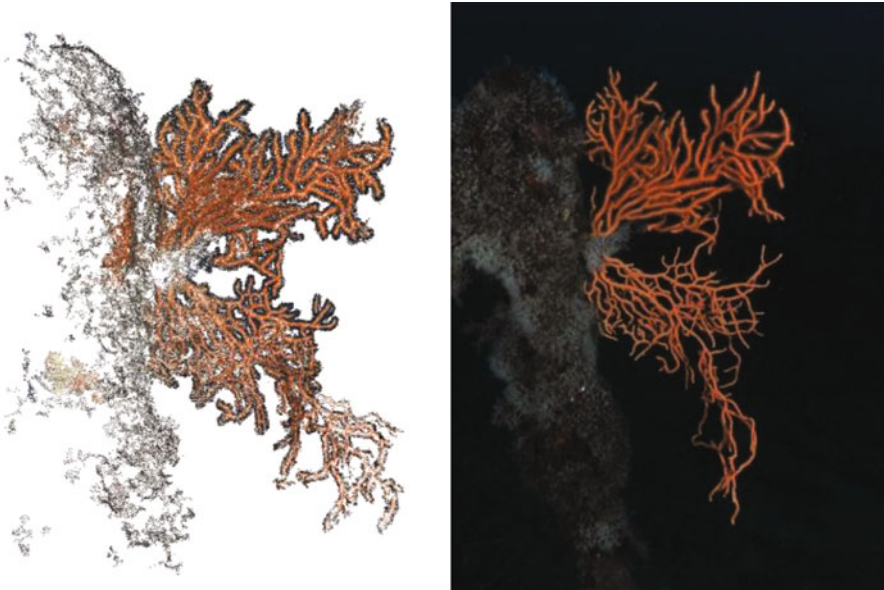


Fig. 14.9 On the right, one of the 64 digital photographs taken (photo Olivier Bianchimani, all rights reserved), on the left, the point cloud automatically computed by the Furukawa method

14.4.2 Large Scale Detail by Photogrammetry

We also tested as part of this mission approach automatic mapping proposed by Furukawa and Ponce (Furukawa and Ponce 2010). We tested this approach on large-scale details, Gorgonaria whose ends were slightly in motion because of the current, see figure below (Fig. 14.9).

One camera was used and sixty photographs were taken for each of the tests that follow. The study of the accuracy and the percentage of coverage was not yet done, that the study dating from April 2011.

14.4.3 The “Liban” Wreck

The Liban is a ship built in 1882 in Glasgow (Scotland), measuring 91 m long and 11 wide. It was equipped with a steam engine. On June 7, 1903 at noon, the Liban left the port of Marseille and less than 1 h later sunk after a big collision with another ship.

Today, the Liban is a very attractive dive site close to Marseilles at 25 m deep. We chose to survey the bow in order to develop and test our approach. Three dives and almost 2,000 photos were necessary to obtain the 3D point cloud visible in the next figure (Fig. 14.10 and 14.11).



Fig. 14.10 On the right, one of the 1,221 digital photographs taken (photo Olivier Bianchimani, All right reserved), on the left, The “Liban” wreck.: Cloud of 3D points measured by photogrammetry using Furukawa method

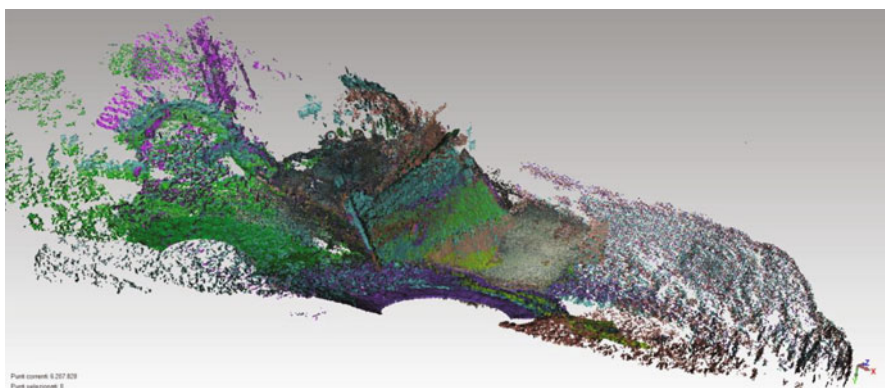


Fig. 14.11 An acoustic survey was done on the same wreck in order to scale the photogrammetric model. Each scan from its own station is represented in a different color

The acoustic survey was done in 1 day and six stations. All survey from these stations were merged using ICP algorithm with an RMS of 0.045 m.

The photogrammetric model, with a high dense cloud of point was scaled and georeferenced (on a local reference system) using the acoustic survey.

In this case, both of the survey see the more or less the same part of the site, the only big difference between them is the resolution and so, the general accuracy.

Merging photogrammetric data on the full acoustic model was done with an RMS of 0.032 m.

We are still working on merging partial photogrammetric survey on acoustic data in an automatic way.

14.5 Conclusion

The ROV3D is an ambitious project, partially funded by the European Regional Development Fund and the French Single Inter-Ministry Fund (FUI) for funding research involving both academic laboratories and industry and also by French regional structure as “Conseil Régional PACA”, “Conseil Général des Bouches du Rhône” and “Marseille Provence Métropole”. It takes benefit from a strong collaboration between a research laboratory and two private companies in order to be able to test and improve methods and algorithms.

The project aims to produce a complete set of tools and methods for underwater survey in complex and varied environment, for example, real 3D sites as caves, wrecks, walls where a simple type of terrain modelling as DTM is not enough.

Moreover the interest of this project is to produce accurate 3D models with texture information and thanks to the combination between the acoustic and the optical approaches developing specific image processing filters in order to correct photo illumination in underwater conditions.

The project is now quite mature and close to be fully operational. We are still working on marine integration in a small ROV and also on a real-time process in order to have continuous feedback on-board of the survey performed by the ROV. A draft mosaic and a 3D model can be computed on the fly using synchronized video cameras. We are working on a hybrid system, merging high- and low-resolution cameras in order to be able to process results in real-time as well as to be able to process results off-line with a high quality, as presented in Sect. IV of this paper.

References

- Agarwal S, Furukawa Y, Snavely N, Curless B, Seitz SM, Szeliski R (2010) Reconstructing Rome. *Computer* 43:40–47 (isbn/issn:0018–9162)
- Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. *IEEE Trans Pattern Anal Mach Intell* 28(12):2037–2041
- Baluja S, Rowley HA (2005) Boosting sex identification performance. *AAAI* 1508–1513
- Barazzetti L, Scaioni M, Remondino F (2010) Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation. *Photogramm Rec* 25:356–381 (Blackwell Publishing Ltd(Pub.), isbn/issn:1477–9730).
- Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). *Comput Vis Image Underst* 110:346–359 (isbn/issn:1077–3142)

- Bazeille S, Quidu I, Jaulin L, Malkasse J-P (2006, 16–19 Oct) Automatic underwater image pre-processing. Paper presented at the CMM'06—CHARACTERISATION DU MILIEU MARIN
- Besl PJ, McKay ND (1992) A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* 14:239–256 (IEEE Computer Society(Pub.), isbn/issn:0162–8828)
- Chambah M, Semani D, Renouf A, Courtellemont, P, Rizzi A (2004) Underwater color constancy: enhancement of automatic live fish recognition. Paper presented at the 16th Annual symposium on electronic imaging, United States
- Chen Y, Medioni G (1991) Object modeling by Reg strat ion of multiple range images. *Int Conf Robotic Autom* 2724–2729
- Drap P, Long L (2005, 16 and 17 juin) *Relevés optiques et systèmes d'information pour l'archéologie sous-marine*. Paper presented at the Des images au 3D: Extraction automatique et assistée de modèles 3D, ENSTA, Paris
- Fan F, Kecheng Y, Bo F, Min X, Wei Z (2010) Application of blind deconvolution approach with image quality metric in underwater image restoration. Paper presented at the image analysis and signal processing (IASP), 2010 international conference on 9–11 Apr 2010, pp 236–239
- Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci* 55:119–139 (isbn/issn:0022–0000)
- Furukawa Y, Ponce J (2010) Accurate, dense, and robust multiview stereopsis. *IEEE Trans Pattern Anal Mach Intell* 32:1362–1376 (isbn/issn:0162–8828)
- Fusiello A, Murino V (2004) Augmented scene modeling and visualization by optical and acoustic sensor integration. *IEEE Trans Vis Comput Graph* 10:625–636 (isbn/issn:1077–2626)
- Hurtós N, Cufí X, Salvi J (2010) Calibration of optical camera coupled to acoustic multibeam for underwater 3D scene reconstruction, Sydney, NSW (isbn/issn:9781424452217 (ISBN))
- Iqbal K, Abdul Salam R, Osman A, Talib AZ (2007) Underwater image enhancement using an integrated colour model. *IAENG Int J Comput Sci* 34:239–244
- Johnson AE (1997) Spin-images: a representation for 3-D surface matching. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania
- Kalia R, Keun-Dong L, Samir BVR, Je S-K, Oh W-G (2011, 9–11 Fev). An analysis of the effect of different image preprocessing techniques on the performance of SURF: speeded up robust feature. Paper presented at the 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), pp 1–6
- Lourakis M, Argyros A (2009) SBA: a software package for generic sparse bundle adjustment. *ACM Trans Math Softw* 36:1–30
- Lowe D (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60:91–110
- Merad D, Didier JY, Scuturici M (2006) Tracking 3D free form object in video sequence. Third Canadian conference on computer and robot vision
- Mian AS, Bennamoun M, Owens R (2006) Three-dimensional model-based object recognition and segmentation in cluttered scenes. *IEEE Trans Pattern Anal Mach Intell* 28:1584–1601 (isbn/issn:0162–8828)
- Morris D, Kanade T (2000) Image-consistent surface triangulation. *IEEE Trans Vis Comput Graph* 1:332–338
- Nicosevici T, Gracias N, Negahdaripour S, Garcia R (2009) Efficient three-dimensional scene modeling and mosaicing. *J Field Robot* 26:759–788 (Wiley Ltd.(Pub.), isbn/issn:1556–4959)
- Petit F (2010) Traitement et analyse d'images couleur sous-marines: modèles physiques et représentation quaternionique. Doctorat, Sciences et Ingénierie pour l'Information, Poitier
- Pierrot-Deseilligny M, Cléry I (2011, 2–4 Mar) APERO, an open source bundle adjustment software for automatic calibration and orientation of a set of images. Paper presented at the 3D Arch 2011, Trento, Italy. ISPRS Commission V Symposium, Image Engineering and Vision Metrology (Pub.)
- Rizzi A, Gatta C (2004) From retinex to automatic color equalization: issues in developing a new algorithm for unsupervised color equalization. *J Electron Imaging* 13:75–84

- Rusinkiewicz S, Levoy M (2001) Efficient variants of the ICP algorithm. Third international conference on 3D digital imaging and modeling (3DIM)
- Rusu RB, Blodow N, Marton ZC, Beetz M (2008) Aligning point cloud views using persistent feature histograms, *Intelligent Robots and Systems. IROS 2008. IEEE/RSJ*, 22–26 Sept. 2008, Nice, France. pp 3384–3391. ISBN 978-1-4244-2057-5
- Sahillioglu Y, Yemez Y (2010) 3D shape correspondence by isometry-driven greedy optimization. In the twenty-third IEEE conference on computer vision and pattern recognition, CVPR 2010, San Francisco, CA, USA, pp 453–458
- Sehgal A, Cernea D, Makaveeva M (2010) Real-time scale invariant 3D range point cloud registration. International conference on image analysis and recognition (ICIAR 2010), Vol. 6111 of LNCS, Povoá de Varzim, Portugal, pp 220–229
- Shokoufandeh A, Macrini D, Dickinson S, Siddiqi K, Zucker SW (2005) Indexing Hierarchical structures using graph spectra. *IEEE Trans Pattern Anal Mach Intell* 27(7):1125–1140
- Shortis M, Harvey E, Abdo D (2009) A review of underwater stereo-image measurement for marine biology and ecology applications. *Oceanogr Mar Biol* 47:257–292 (Crc Press-Taylor & Francis Group(Pub.), isbn/issn:0078–3218)
- Shunyi Z, Ruifang Z, Zuxun Z (2007) Generation of 3D surface model of complex objects based on non-metric camera. Paper presented at the image processing, 2007. ICIP 2007. IEEE international conference on Sept 16–Oct 19 2007. Vol 3, pp III–89–III–92, isbn/issn:1522–4880
- Singh H, Roman C, Whitcomb L, Yoerger D (2000) Advances in fusion of high resolution underwater optical and acoustic data. Paper presented at the Proceedings of the 2000 International Symposium on Underwater Technology, Tokyo Japan
- Singh H, Roman C, Pizarro O, Eustice R, Can A (2007) Towards high-resolution imaging from underwater vehicles. *Int J Robot Res* 26:55–74 (Sage Publications, Inc.(Pub.), isbn/issn:0278–3649)
- Snavely N, Simon I, Goesele M, Szeliski R, Seitz SM (2010) Scene reconstruction and visualization from community photo collections. *Proc IEEE* 98:1370–1390 (isbn/issn:0018–9219)
- Thome N, Merad D, Miguët S (2008) Learning articulated appearance models for tracking humans: a spectral graph matching approach. *Image Comm* 23(10):769–787
- Viola PA, Jones MJ (2001) Robust real-time object detection. *Int J Comput Vis* 57(2):747
- Wahl E, Hillenbrand U, Hirzinger G (2003) Surflet-pair-relation histograms: a statistical 3D-shape representation for rapid classification. In proceedings of the fourth international conference on 3-D digital imaging and modeling, pp 474–481
- Zeng Gang S, Quan PL, Sillion F (2005) Progressive surface reconstruction from images using a local prior. Paper presented at the Computer Vision, 2005. ICCV 2005. Tenth IEEE international conference on 17–21 Oct 2005. Vol. 2, pp. 1230–1237 Vol. 1232, isbn/issn:1550–5499