

Creating 360° Underwater Virtual Tours Using an Omnidirectional Camera Integrated in an AUV

Josep Bosch*, Pere Ridao*, David Ribas* and Nuno Gracias*

* Computer Vision and Robotics Institute (VICOROB), University of Girona, 17071 Girona, Spain
j.bosch@udg.edu, pere@silver.udg.edu, dribas@udg.edu, ngracias@silver.udg.edu

Abstract—The use of omnidirectional cameras underwater is enabling many new and exciting applications in multiple fields. Among these, the creation of virtual tours from omnidirectional image surveys is expected to have a large impact in terms of science and conservation outreach. These surveys can be performed by Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs) that can cover large marine areas with precise navigation. Virtual tours are relevant in zones of special interest such as shipwrecks or underwater nature reserves for both scientists and the general public. This paper presents the first results of surveys carried out by an AUV equipped with an omnidirectional underwater camera, and explores the process of automatically creating virtual tours from the most relevant images of the datasets.

I. INTRODUCTION

In the last years, omnidirectional cameras have received increasing interest from the computer vision community in tasks such as mapping, augmented reality, visual surveillance, motion estimation and simultaneous localization and mapping (SLAM). The use of omnidirectional cameras in underwater environments opens the door to several new technological applications in fields as diverse as underwater robotics, marine science, oil and gas industries, underwater archeology and public outreach. However, due to the challenges of the underwater medium, the use of these cameras is still very limited compared with land and air applications.

This paper presents the first results of the integration of an omnidirectional underwater camera with an underwater autonomous robot. The integration of omnidirectional cameras with underwater robots is expected to have a large impact in both Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs) [1]. It will allow ROVs to be piloted directly through the images captured by the omnidirectional cameras through virtual reality headsets. This immersive experience will extend the pilot's spatial awareness and reduce the usual orientation problems during missions. For AUVs, the wide field of view of the cameras is very convenient for mapping tasks and visual SLAM [2], specially in confined or cluttered environments.

This paper focuses on the use of ROVs and AUVs as a means of creating virtual reality tours from surveys. These virtual tours are very relevant in zones of special interest like shipwrecks, diving zones, or regions of rich marine life such as nature reserves. They can be used as an attractive and innovating tool to bring archeology, marine life and environment awareness closer to the general public, or it can be a very effective way to promote a specific region or country.

The impact of acquiring omnidirectional images from an AUV is illustrated with the results of a survey over a shipwreck in the bay of Porto Pim in Horta, Azores Islands, that was carried out in September 2014 during sea trials in the framework of the MORPH EU-FP7 project [3].

A selection of the related work can be found in section II. The used equipment is described in section III. Section IV and V present the construction of the panoramas and the virtual tours. Section VI presents practical results. Finally section VII draws briefly the conclusions and future work.

II. RELATED WORK

One of the most known and used applications of omnidirectional cameras in land, is Google Street View [4]. This worldwide known technology featured in Google Maps provides panoramic views from different positions, in streets from thousands of cities around the world. The concept presented in this paper is similar to an underwater version of this product. Furthermore, it can be integrated in Google Maps to make the panoramas and tours captured available freely to all internet users.

The generation of virtual tours has evolved closely to the capabilities of creating omnidirectional panoramas. If the first omnidirectional images were captured through catadioptric systems, nowadays they can also be captured by multiple individual cameras composing an omnidirectional multi-camera system (OMS) [5] or even through the use of smartphones [6].

Many authors have explored the possibilities of generating virtual tours from omnidirectional images in land scenes. Uyttendaele et al. [7] generated virtual tours from indoor and outdoor scenes using an OMS and estimating the camera pose through vision algorithms. Taylor [8] created indoor virtual tours using a catadioptric camera system and estimating the camera poses using structure from motion. Saurer et al. [9] proposed a method that automatically computes the tour topology from the video data using structure from motion and able to find junctions and loop closures. A frame is selected for a subset of key frames only if their SIFT [10] features present a significant displacement with respect to a previous selected image. Cai et al. [11] take into account the elevation data of a known scene in order to find a limited number of representative views with the maximum possible coverage.

Regarding the creation of underwater virtual tours, Google in a joint effort with CatlinSeaView [12], launched in 2012 a project to map coral reef zones and capture panoramic images that were later available in Google Maps. The CatlinSeaView uses an OMS camera attached to a purpose built, pole-like

structure maneuvered by divers. The advantages of using a robot to deploy an omnidirectional camera when compared to a diver-based system can be summarized as the following:

- *Depth*: The depths that robots can reach carrying cameras are greater than those that divers can reach.
- *Distance*: Robots can do longer surveys and are less time-limited.
- *Precision*: Robots can navigate more precisely than divers.
- *Safety*: Diving in difficult conditions can put in risk the health of the divers capturing the images.

III. EQUIPMENT

In order to illustrate the concept, two surveys were carried out on September 2014 and March 2015 with the Girona500 AUV and a custom omnidirectional underwater camera.

A. Girona500 AUV

The Girona500 AUV [13] (Fig. 2a) is a reconfigurable vehicle rated for 500 m depth and is hovering-capable with 5 thrusters actuated in 4 DOFs: surge, sway, heave and yaw. It is stable in Pitch and Roll. Its navigation sensor suite includes GPS, pressure sensor, a Doppler Velocity Log (DVL) and an attitude and heading reference unit (AHRS) that makes it able to navigate precisely. However, in long underwater missions, the vehicle can also be equipped with an ultra short base line (USBL) device to avoid drifting. The navigation is based on an extended Kalman filter (EKF) that combines the sensors to obtain a robust estimation of the AUV position and velocity. The vehicle has a large payload area (up to 35 liters) which allows to carry extra equipment. In the case of this paper, this area was used for the omnidirectional camera.

B. Omnidirectional camera

The camera used is an OMS, based on a Point Grey's Ladybug 3 camera [14]. The Ladybug 3 comprises 6 individual cameras and is designed for land based applications. A custom housing has been designed for the Ladybug 3 camera to make it submersible up to 60 meters. The housing is composed of a transparent poly methyl methacrylate (PMMA) dome, which contains the camera, and a body made of aluminum alloy, which contains a small form factor computer, dedicated to processing the video feed. The camera outputs 6 separate images that can be combined to create an hemispherical panorama (Fig. 1) using the calibration parameters of the camera. Due to the waterproof housing and the changes in the optics required to obtain a complete hemispherical image in water, the factory-provided calibration is not valid, and a custom calibration procedure for underwater OMS was developed.

Calibration for multi-camera systems typically covers two different sets of parameters: intrinsic parameters, concerning the image formation geometry for each individual camera, and extrinsic parameters, which describe the relative positions and orientations between cameras. In omnidirectional multi-camera systems, the calibration of the extrinsic parameters is an important challenge, due to the usual small overlap between neighboring cameras. Furthermore, for underwater cameras it must be taken into account that the direction of the rays of

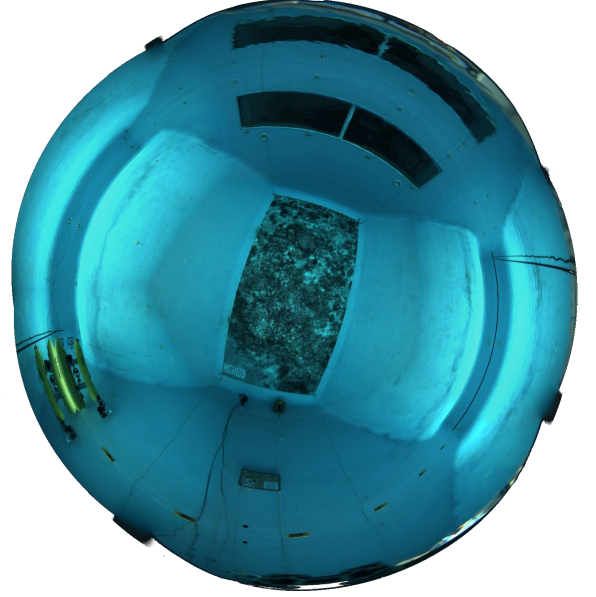


Fig. 1: Spherical projection of a testing water tank captured by the omnidirectional underwater camera.

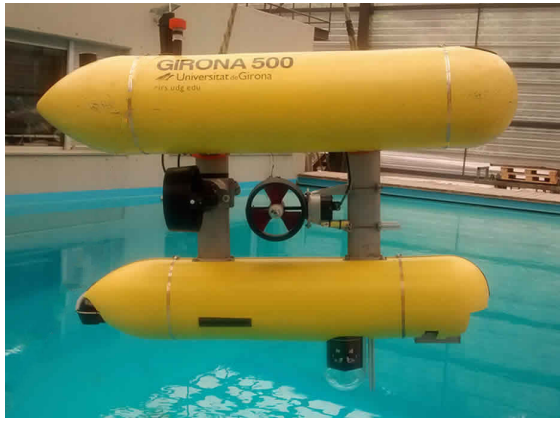
light changes in every medium transition found along the path from a point in water to the imaging sensor inside the camera. In order to model accurately the distortion due to this effect, it becomes essential to explicitly model and simulate the intersection of each light ray with the different media. Due to this phenomenon, any small variation in the estimated relative position of the housing can significantly affect the final direction of the rays and end up generating projection errors.

The calibration procedure was done in three different stages. The first consisted of the estimation of the intrinsic parameters, which was done separately for each single camera in air and without the waterproof housing. The second stage consisted of the estimation of the extrinsic parameters, also done in the same conditions as the first step. Finally, the last stage took place underwater and estimated the camera pose with respect to the waterproof housing. The full design and calibration of the camera can be found in [15].

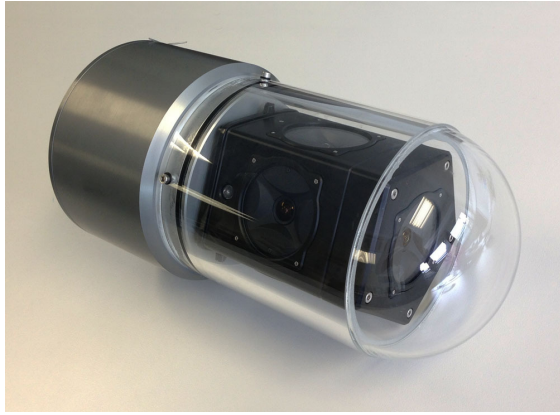
IV. PANORAMA COMPOSITION

Making use of the calibration of the cameras, each pixel of the output images can be associated with a 3D ray in the space. Except for the small area where there is image overlap, it is not possible to estimate the distance to the objects from just a set of images acquired at a single location. For this reason, for visualization purposes, the world around the camera is assumed to be a sphere, where all of the points sensed by the camera are at a constant preselected distance. Once the sphere radius is defined, a spherical point cloud is quick to compute, and it can be easily loaded in a computer 3D viewer or re-projected into a 2D image, being the equirectangular and spherical projections the most commonly used.

Given a world point in Cartesian coordinates $Q =$



(a) The omnidirectional camera integrated with the Girona500 AUV.



(b) Omnidirectional underwater camera used for the survey.

Fig. 2: Used equipment for data collection.

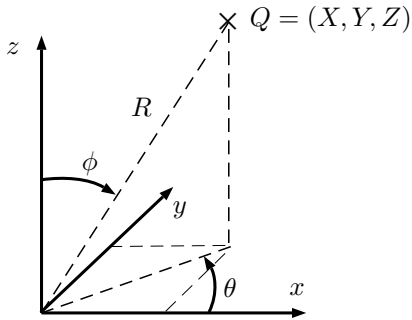


Fig. 3: Conversion from Cartesian to spherical coordinates.

(X, Y, Z) , it can be converted to spherical coordinates (Figure 3) $Q = (\theta, \phi, R)$ through Equations 1–3.

$$R = \sqrt{X^2 + Y^2 + Z^2} \quad (1)$$

$$\theta = \text{atan2}(Y, X), -\pi \leq \theta \leq \pi \quad (2)$$

$$\phi = \text{acos}\left(\frac{Z}{R}\right), 0 \leq \phi \leq \pi \quad (3)$$

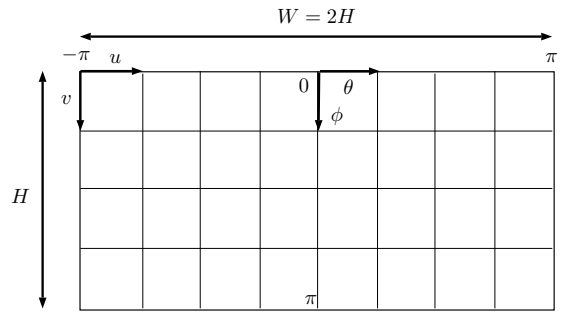


Fig. 4: Scheme and notation for the equirectangular projection.

The equirectangular projection projects a given point Q to a cylinder (Figure 4) through Equations 4 and 5:

$$u = \frac{\theta + \pi}{2\pi} \cdot W \quad (4)$$

$$v = \frac{\phi}{\pi} \cdot H \quad (5)$$

The inverse equations are:

$$\theta = \frac{u \cdot 2\pi}{W} - \pi \quad (6)$$

$$\phi = \frac{v \cdot \pi}{H} \quad (7)$$

For the spherical projection (Fig. 5) the equations are:

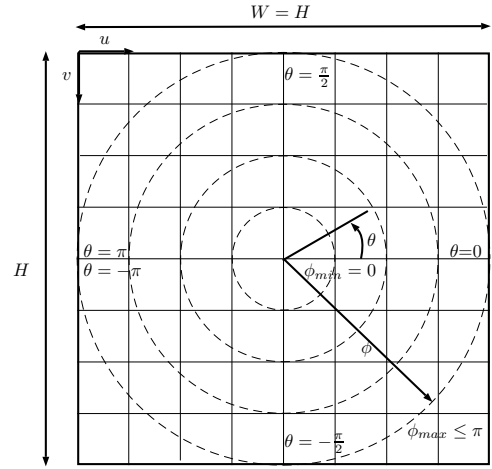


Fig. 5: Scheme and notation for the spherical projection.

$$u = \frac{\phi \cdot \cos(\theta) + \phi_{max}}{2\phi_{max}} \cdot W \quad (8)$$

$$v = \frac{\phi_{max} - \phi \cdot \sin(\theta)}{2\phi_{max}} \cdot H \quad (9)$$

The inverse equations are:

$$\theta = \text{atan2}\left(\left(\frac{H}{2} - v\right), \left(-\frac{W}{2} + u\right)\right) \quad (10)$$

$$\phi = \sqrt{\left(-\frac{W}{2} + u\right)^2 + \left(\frac{H}{2} - v\right)^2} \cdot \frac{2\phi_{max}}{H} \quad (11)$$

The first step when composing a panorama is choosing its parameters: projection type, projection distance and final size. For every pixel of the panorama, the 3D point it represents is computed according to the inverse equations of its projection (Equations 6, 7 or 10, 11). This 3D point is then projected to each one of the six cameras according to the calibration. If the point is only in the FOV of one camera, we give to the pixel of the panorama the same intensity values as the pixel corresponding to the projection of the 3D point into the camera. In the case of overlapping regions, a blending criterion is needed, to establish the value of the panorama pixel. The blending criteria chosen for the panoramas displayed in this paper is known as gradient blending. Further details of this technique can be found in [16].

V. VIRTUAL TOUR CREATION

Thanks to the time synchronization between the camera and the AUV and using its precise navigation, every panoramic image can be automatically tagged with its GPS location, heading, depth and altitude. From all the images captured, only a reduced group is selected to create a virtual tour of the survey. Different criteria are applied in this step depending on the purposes of the survey and the scope of the virtual tour.

- 1) *Exhaustive selection*: For small extension surveys, where the region of interest (ROI) is well known *a priori*, the most representative images can be manually selected. While this technique ensures the best selection, according to the purpose of the virtual tour promoters, it is slow and human-dependant.
- 2) *Time-based sampling*: For long surveys without delimited ROI, images can be sub-sampled automatically with a fixed time interval, t , between captures. In the case of a constant camera frame rate, r , this is equivalent to selecting a sample every $t \times r$ images. This method is especially indicated when the purpose is the replay of the AUV's mission for further analysis.
- 3) *Distance-based sampling*: Similarly to the previous case, images can be sub-sampled by distance. An algorithm creates initially a subset of selected images containing only the first image of the sequence, and analyses each one of the images of the dataset in chronological order. Only those images whose distance to any of the subset of selected images is larger than a threshold are copied to the selected subset. In contrast with the previous technique, this method avoids the presence of multiple similar images when the vehicle speed is low.
- 4) *Feature-based selection*: For missions where the ROI cannot be precisely defined *a priori* or for missions covering a wide area with a large amount of irrelevant images, we have developed a technique to select *a posteriori* the ROI over the navigation plot and automatically extract its most interesting images. The algorithm works in the following way: The operator

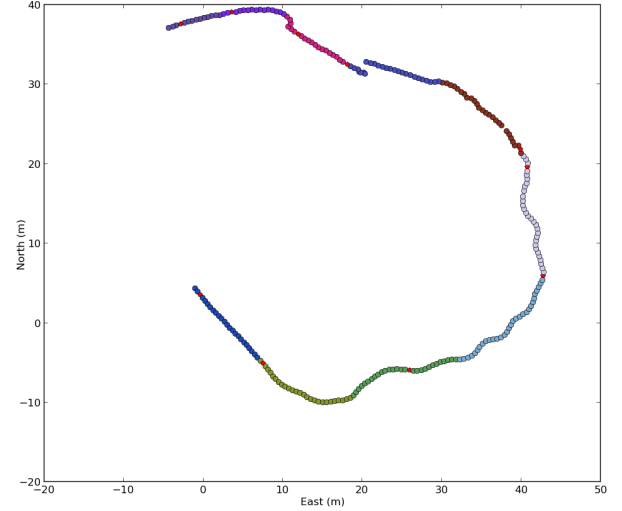


Fig. 6: Feature-based selection of panoramic images with 10 regions of a survey of the harbour in St. Feliu de Guíxols, Spain in March 2015. The different colors represent each one of the regions, while the red dots represent the selected panoramas.

selects the ROI over a navigation plot and chooses the number of images, n , to be extracted from that area. All the panoramic images pertaining to the area are selected and a k-means algorithm [17] is run in order to divide the panoramic images in n regions. The most representative image of each region will be chosen for the virtual tour. In order to choose the most representative image, a feature detection algorithm is run for all the images of the region, and the panorama with a higher number of features is selected. This technique has been successfully tested with SURF [18] and SIFT [10] detectors. As it can be observed in Table I images with low visual interest have a small number of features, while the number of features increases accordingly to the visual interest of the images.

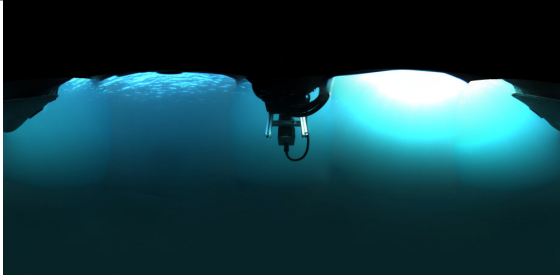
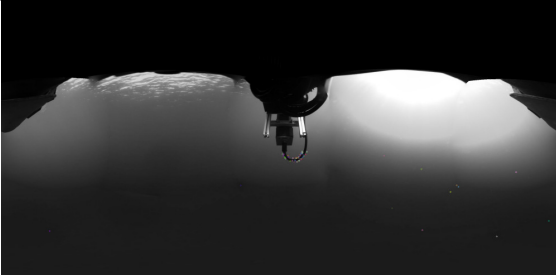
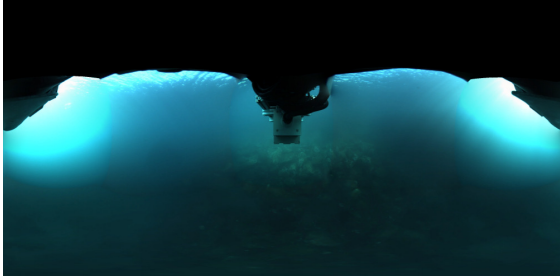
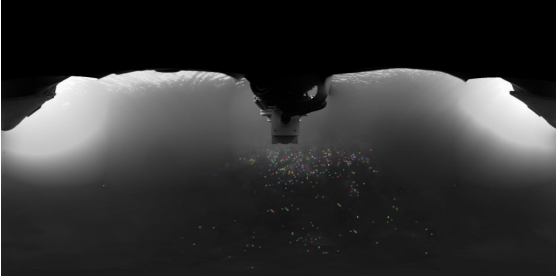
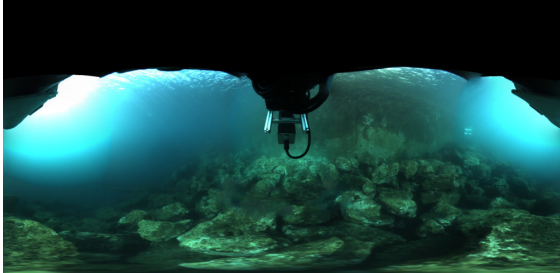
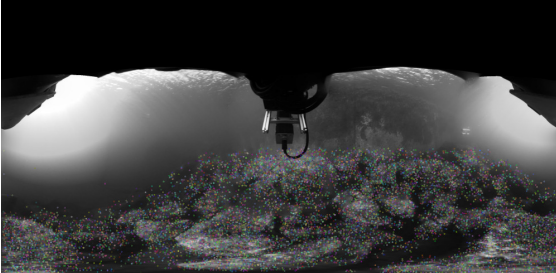
At this step, the selected panoramic images are ready to be loaded into any software capable of displaying 360° virtual tours. The last step is to choose the connections between the images composing the tour. Most platforms propose automatically connections between images and the author only needs to check and modify as he wishes.

Once published the virtual tour, the final users will be able to choose a panorama, look into the scene in all directions and change the location of the view moving to adjacent panoramas. Apart from navigating through a friendly web platform, there exists the possibility to use a virtual reality headset, such as the Oculus Rift, to achieve a full immersive experience.

VI. RESULTS

The impact of acquiring omnidirectional images from an AUV is illustrated with the results of a survey over a ship-

TABLE I: Comparison of features found in panoramic images from the same sequence but with different levels of content relevance. It exists a clear correlation between the visual interest of the panorama for a final user and the number of features found. The colored circles in the second column of images represent the features found by SURF feature detector. Features were only searched in the lower half of the image to avoid the detection of features related to surface reflections and the body of the AUV. The dataset is from a survey conducted in the harbour of St. Feliu de Guíxols, Spain in March 2015.

Equirectangular Image	Features found	Number of features
		Low: 62
		Medium: 416
		High: 10970

wreck in the bay of Porto Pim in Horta, Azores Islands, that was carried out in September 2014 during the trials in the framework of the MORPH EU-FP7 project [3]. The Girona500 AUV was teleoperated from a land-base by a pilot who was operating the robot through the preview images acquired from the omnidirectional camera. After the recovery of the vehicle, the data from the camera was processed. The panoramic images from the survey were created from the individual images of each camera composing the OMS and the calibration information, and tagged with a GPS location according to the AUV's navigation (Figure 7).

The different sampling methods detailed in section V were tested, as seen in figure 8. Due to the fact that the survey covered a small area and was already centered over the shipwreck region, the preferred solution was an operator-based selection. Fig. 9 shows the final solution adopted for this survey, with its location in a real map, the panoramas composing the virtual tour and the links between them, and the view from the web-based navigation tool. This virtual tour is publicly available in Google Maps.

VII. CONCLUSIONS AND FUTURE WORK

This paper makes a two-fold contribution to the field by (1) providing the first results on integrating and deploying an omnidirectional camera on a low cost autonomous vehicle, and (2) by exploring the concept of automatically created virtual tours using a widely accessible navigation platform on the web.

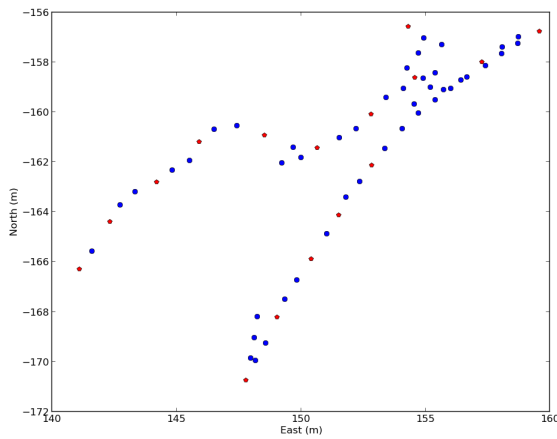
As future work, new approaches for selecting the most relevant images from a survey dataset will be explored. We will focus on techniques for reducing the processing time in order to decrease the amount of time required for large datasets.

ACKNOWLEDGMENTS

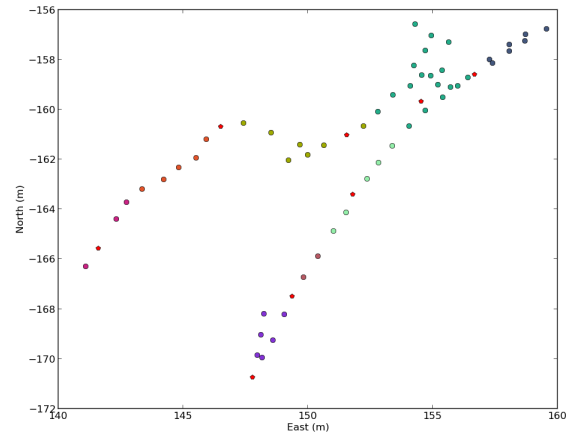
This research was supported by the MORPH EU FP7-Project under the Grant agreement FP7-ICT-2011-7-288704, the Spanish National Project OMNIUS under the agreement CTM2013-46718-R, the Generalitat de Catalunya through the ACCIO/TecnoSpring program (TECSPR14-1-0050), "la Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya", and the University of Girona under a grant for the formation of researchers.



Fig. 7: Equirectangular projection of a shipwreck survey in Horta, Azores Islands in September 2014.



(a) Selection of panoramic images sequence applying a distance-based sampling with a minimum distance of 2m. Red dots represent the selected panoramas, while blue dots are the discarded panoramas.



(b) Feature-based selection of panoramic images with 8 regions. The different colors represent each one of the regions, while the red dots represent the selected panoramas.

Fig. 8: Comparison between distance-based and features-based sampling techniques applied in a survey carried out in Horta, Azores Islands in September 2014.

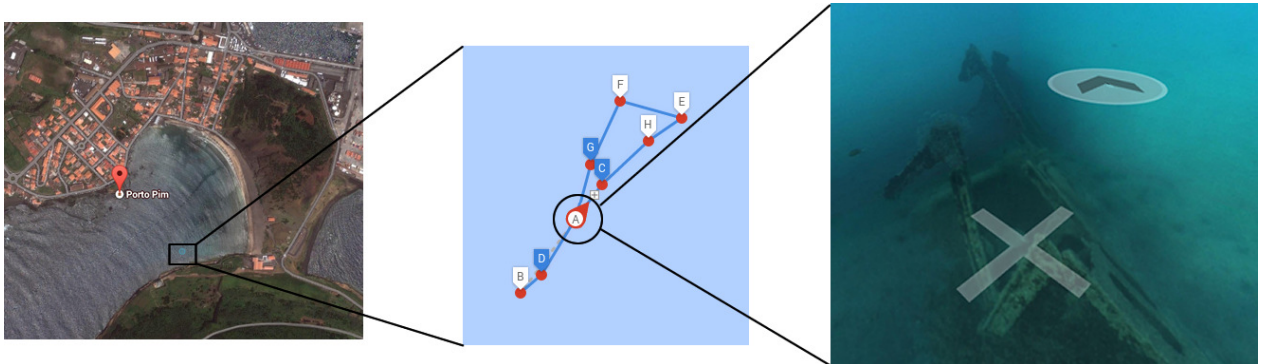


Fig. 9: Location and sample image of a shipwreck virtual tour in Google Maps. **Left:** Location of the shipwreck in Porto Pim bay in Horta, Azores Islands. **Middle:** Different panoramas and connections available in the virtual tour. **Right:** Sample view of the shipwreck from A.

REFERENCES

- [1] S. Negahdaripour, H. Zhang, P. Firoozfam, and J. Oles, "Utilizing panoramic views for visually guided tasks in underwater robotics applications," in *OCEANS, 2001. MTS/IEEE Conference and Exhibition*, vol. 4, 2001, pp. 2593–2600 vol.4.
- [2] A. Rituerto, L. Puig, and J. Guerrero, "Visual slam with an omnidirectional camera," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, Aug 2010, pp. 348–351.
- [3] J. Kalwa, M. Carreiro-Silva, F. Tempera, J. Fontes, R. Santos, M.-C. Fabri, L. Brignone, P. Ridao, A. Birk, T. Glotzbach, M. Caccia, J. Alves, and A. Pascoal, "The MORPH concept and its application in marine research," in *OCEANS - Bergen, 2013 MTS/IEEE*, June 2013, pp. 1–8.
- [4] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, and J. Weaver, "Google street view: Capturing the world at street level," *Computer*, no. 6, pp. 32–38, 2010.
- [5] L. Puig, J. Bermúdez, P. Sturm, and J. Guerrero, "Calibration of omnidirectional cameras in practice: A comparison of methods," *Computer Vision and Image Understanding*, vol. 116, no. 1, pp. 120–137, Jan. 2012.
- [6] A. Sankar and S. Seitz, "Capturing indoor scenes with smartphones," in *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '12. New York, NY, USA: ACM, 2012, pp. 403–412. [Online]. Available: <http://doi.acm.org/10.1145/2380116.2380168>
- [7] M. Uyttendaele, A. Criminisi, S. B. Kang, S. Winder, R. Hartley, and R. Szeliski, "High-quality image-based interactive exploration of real-world environments," *IEEE Computer Graphics & Applications (CG&A)*, vol. 24, no. 3, pp. 52–63, May 2004.
- [8] C. Taylor, "Videoplus: a method for capturing the structure and appearance of immersive environments," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 8, no. 2, pp. 171–182, Apr 2002.
- [9] O. Saurer, F. Fraundorfer, and M. Pollefeys, "Omnitour: Semi-automatic generation of interactive virtual tours from omnidirectional video," in *3DPVT10*, 2010.
- [10] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Key-points," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [11] H. Cai and J. Y. Zheng, "Key views for visualizing large spaces," *J. Vis. Comun. Image Represent.*, vol. 20, no. 6, pp. 420–427, Aug. 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.jvcir.2009.04.005>
- [12] M. González-Rivero, P. Bongaerts, O. Beijbom, O. Pizarro, A. Friedman, A. Rodríguez-Ramírez, B. Upcroft, D. Laffoley, D. Kline, C. Bailhache *et al.*, "The catlin seaview survey—kilometre-scale seascape assessment, and monitoring of coral reef ecosystems," *Aquatic Conservation: Marine and Freshwater Ecosystems*, vol. 24, no. S2, pp. 184–198, 2014.
- [13] D. Ribas, N. Palomeras, P. Ridao, M. Carreras, and A. Mallios, "Girona 500 AUV: From Survey to Intervention," *Mechatronics, IEEE/ASME Transactions on*, vol. 17, no. 1, pp. 46–53, Feb 2012.
- [14] I. Point Grey Research, "360 Spherical - Ladybug3 - Firewire Camera," <http://www.ptgrey.com/products/ladybug3/>.
- [15] J. Bosch, N. Gracias, P. Ridao, and D. Ribas, "Omnidirectional underwater camera design and calibration," *Sensors*, vol. 15, no. 3, pp. 6033–6065, 2015. [Online]. Available: <http://www.mdpi.com/1424-8220/15/3/6033>
- [16] R. Prados, R. García, N. Gracias, J. Escartin, and L. Neumann, "A novel blending technique for underwater giga-mosaicing," *IEEE Journal of Oceanic Engineering*, vol. 37, no. 4, pp. 626–644, 2012.
- [17] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. pp. 100–108, 1979.
- [18] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.