A Comparison of Image Based 3D Recovery Methods for Underwater Inspections

Michael O'Byrne, Vikram Pakrashi, Franck Schoefs, Bidisha Ghosh

A COMPARISON OF IMAGE BASED 3D RECOVERY METHODS FOR UNDERWATER INSPECTIONS

Michael O'Byrne[1], Vikram Pakrashi[2], Franck Schoefs[3], Bidisha Ghosh[1]

[1] *Trinity College Dublin, Department of Civil, Structural and Environmental Engineering, Ireland*
[2] *University College Cork, Dynamical Systems and Risk Laboratory, Civil and Environmental Engineering, College Road, Cork, Ireland*
[3] *LUNAM Université, Université de Nantes-Ecole Centrale de Nantes-CNRS, GeM, Institute for Research in Civil and Mechanical Engineering, UMR 6183 & IXEAD/CAPACITES Society, University of Nantes 2, rue de la Houssinière BP 92208 44322 Nantes Cedex 3, France*

[1] obyrnemi@tcd.ie

ABSTRACT

Offshore structures can be subjected to millions of variable amplitude load cycles during their service life which is the primary cause of structural deterioration. Such fatigue loading is exacerbated by marine growth colonization which changes the surface roughness characteristics and increases the diameter of structural members. Having an accurate knowledge of these parameters is essential for analyzing the increased hydrodynamic forces acting on the structure. This paper addresses the issue of acquiring shape information by comparing two popular classes of image based shape recovery techniques; stereo photography and Structure from Motion (SfM). Stereo photography utilises a dual camera set-up to simultaneously photograph an object of interest from slightly different viewpoints, whilst SfM methods generally involve a single camera moving in a static scene. In this paper, these techniques are performed on a controlled shape in an underwater setting, as well as synthetic data which allows for an irregular shape typical of marine growth to be tested whilst still having knowledge of the exact geometrical shape. The results reveal that the self-calibrated stereo approach fared well at getting an appropriately scaled full metric reconstruction, whilst the SfM approach was more susceptible to breaking down.

KEYWORDS: *Underwater Inspections, 3D Shape Recovery, Stereo, Structure from Motion, Marine Growth.*

1 INTRODUCTION

Advances in construction materials and methods coupled with a growing appetite to exploit the bountiful offshore energy resources has lead to the installation of more and more offshore structures and at greater depths. These structures can experience extreme events in addition to unrelenting loading from wind, waves and currents. An understanding of the effects that these forces have on a structure is essential for the safe and economical design of offshore technologies. The natural process of marine growth colonization is a difficult design consideration to overlook as it causes increased screen and drag effects as well as biochemical attacks [1]. Currently, there is a limited understanding of the occurrence and behaviour of marine growth species, such as the seasonal build-up/decline, multilayer colonization caused by the interaction between various competing species, and ultimately, their contribution to the increased hydrodynamic loading on the host structure. With this in mind, there is a compelling need to track the formation and advancement of marine growth in order provide engineers with an improved insight of the process and consequently, enable decision makers to devise more informed and cost effective cleaning strategies.

Traditional approaches at estimating the shape of marine growth affected structural components rely on taking measurements are discrete points/depths, or on positioning a reference scale in the scene and using it to infer the size of nearby objects in the acquired photographs. However, the former approach only collects a sparse set of measurements, whilst the latter approach involves some guesswork when analyzing the acquired imagery and cannot be used to generate a 3D shape profile. Furthermore, the use of props (i.e. a reference scale) may be a source of inconvenience and cause delays for the diver. Given these shortcomings, there is scope to explore more sophisticated approaches for recovering the shape of structural components. For the purpose of underwater inspections, these approaches should meet some fundamental requirements:

- The system should be sufficiently portable such that a diver can easily manoeuvre underwater and visit multiple target sites.
- The acquisition of data should be conducted in a reasonable timeframe, minimizing any unnecessary time spent underwater.
- The system should be able to overcome the challenges imposed by the underwater conditions such as reduced visibility due to turbidity, illumination complexities etc.

This paper compares two classes of image based shape recovery techniques, namely stereo photography and Structure from Motion (SfM). Such a comparison is warranted by the extensive effort and expense associated with undertaking underwater inspections which calls for a thorough review of viable methods. Stereo photography employs a dual camera set-up to simultaneously photograph a specimen of interest from slightly different viewpoints. 3D information can then be extracted by examining the relative positions of objects captured by each camera. SfM offers an attractive way of recovering the shape of underwater infrastructural elements from a data acquisition perspective as it only relies on a single unconstraint imaging device. The trajectory of prominent features in the scene are tracked over time are then used to reconstruct their 3D positions.

In this paper, these techniques are performed on a controlled shape in an underwater setting, as well as synthetic data which has the advantage of eliminating instrumental errors and allows for more irregular shapes typical of marine growth to be tested whilst still having complete knowledge of their geometrical shape.

The following section provides a background to shape recovery techniques. Section 3 details the methodology of stereo photography and SfM, whilst Section 4 evaluates the performance of each approach when applied to the real and synthetic data. The final section concludes the paper.

2 BACKGROUND

Practical methods for obtaining shape information for short ranges underwater can be partitioned into two main groups: laser and image based methods. Underwater laser methods are becoming an increasingly popular choice for high precision underwater inspections. Three laser based methods have potential in an underwater environment: laser radar, which uses time-of-flight [2]; amplitude-modulation [3], which use phase-comparison; and triangulation methods [4]. Laser radar systems typically employ a laser timed-pulse distance-measurement system. A drawback of laser radar systems is that their range is severely restricted in underwater environment due to attenuation. At such short ranges the transit time is very small which leads to significant measurement errors. Additionally, refraction and scattering make this method hard to implement for high-resolution underwater ranging. Laser triangulation methods employ the well-known triangulation relationship. They are capable of producing high spatial and range resolution at short ranges. However, like all underwater laser based systems, they remain costly and often require extensive calibration procedures. Furthermore, concerns exist about their portability given their size, power requirements, and data-transmission limitations.

Common image based approaches suitable for underwater application include structured lighting, stereo photography, Structure from Motion (SfM), shading and depth from defocus. Structured lighting techniques follow a similar principle to laser scanning. They utilize a light projector and a camera to project a light pattern onto an object and capture how it interacts with the shape [5]. Given the reliance on a light source to encode depth information, the success of this approach is particularly susceptible to absorption and scattering. Structured lighting is an example of an active approach which means it interferes with the scene through the projected light. The projected light may mask the object's natural colour which could be an important factor when analyzing the imagery for other purposes (e.g. identifying marine growth species).

3D recovery approaches that only use standard cameras are an attractive option given that the equipment is inexpensive and readily available, require minimal training on the part of the diver, and cameras are almost always included as part of an inspection routine anyways. Passive image based approaches include depth from defocus, depth from shading, stereo and SfM.

Depth from shading [6] exploits shadows for the recovery of an object's 3D shape. This approach can only recover depth up to an unknown scale factor and the results are not as reliable as other methods. Depth from focus/defocus [7] is the problem of estimating the 3D surface of a scene from a set of two or more images, taken from the same position. The set of images is typically obtained by adjusting the focal setting. Depth from defocus is only effective for small camera-object distances. Furthermore, the requirement to remain static is often not possible in an underwater setting.

Stereo photography [8], also known as stereo vision, involves simultaneously photographing a scene from two vantage points and then examining the relative positions of objects in the images from each camera. The primary challenge faced in stereo photography is the correspondence problem which seeks to locate the same object in both images. A wide range of stereo matching algorithms have been proposed to solve this problem [9]. Another issue is that the cameras must be carefully synchronised. The advantage of stereo vision is that calibrated systems can provide properly scaled fully Euclidean reconstructions with a good degree of accuracy. The stereo system can be pre-calibrated using a well-known checkerboard procedure [10] or through self-calibration using the static scene as a constraint on the camera parameters [11]. Furthermore, useful depth information can be extracted from a single stereo image pair. This aspect means that a stereo rig is less affected by non-rigid scene situations. For real world underwater scenes, non-rigidity is ubiquitous to some extent - possibly originating from deforming surfaces, floating particulate, illumination changes or any combination thereof. Finally, stereo photography can be incorporated into a multi-view stereo framework in order to recover the full 3D shape of the subject.

Finding structure from motion presents a similar problem as finding the structure/shape from stereo vision. In some sense both are geometrically equivalent. In both cases, the correspondence between images and the reconstruction of 3D object needs to be found. In SfM, features such as corner points or other stand-out features are tracked from one image to the next. The feature trajectories over time are then used to reconstruct their 3D positions and the camera's motion. Although SfM and stereo photography share some similar principles, there is a stark difference in terms of practical implementation which has a direct affect on the 3D reconstruction results that can be obtained. For SfM, the input imagery is an image sequence/video acquired from a single moving camera. SfM approaches applied underwater include [12]. A fundamental limitation of classical SfM algorithms is that they cannot be applied to non-rigid scenes, although recent work [13] has allowed for some scene evolution to occur. SfM is capable of generating a fully Euclidean reconstruction up to an unknown global scale factor. Additional information is required to determine this scale factor. Possible ways for finding this scale factor include incorporating other data sources (e.g. odometry) or having an object of known dimensions in the scene. However, odometry sensors are prone to drifting over time and positioning the known object in the scene can be cumbersome for diver.

3 METHODOLOGY

This section outlines the general methodology of stereo photography and SfM, as well as presenting the test imagery used in this paper. The methodology of stereo and SfM is illustrated in Figure 1.

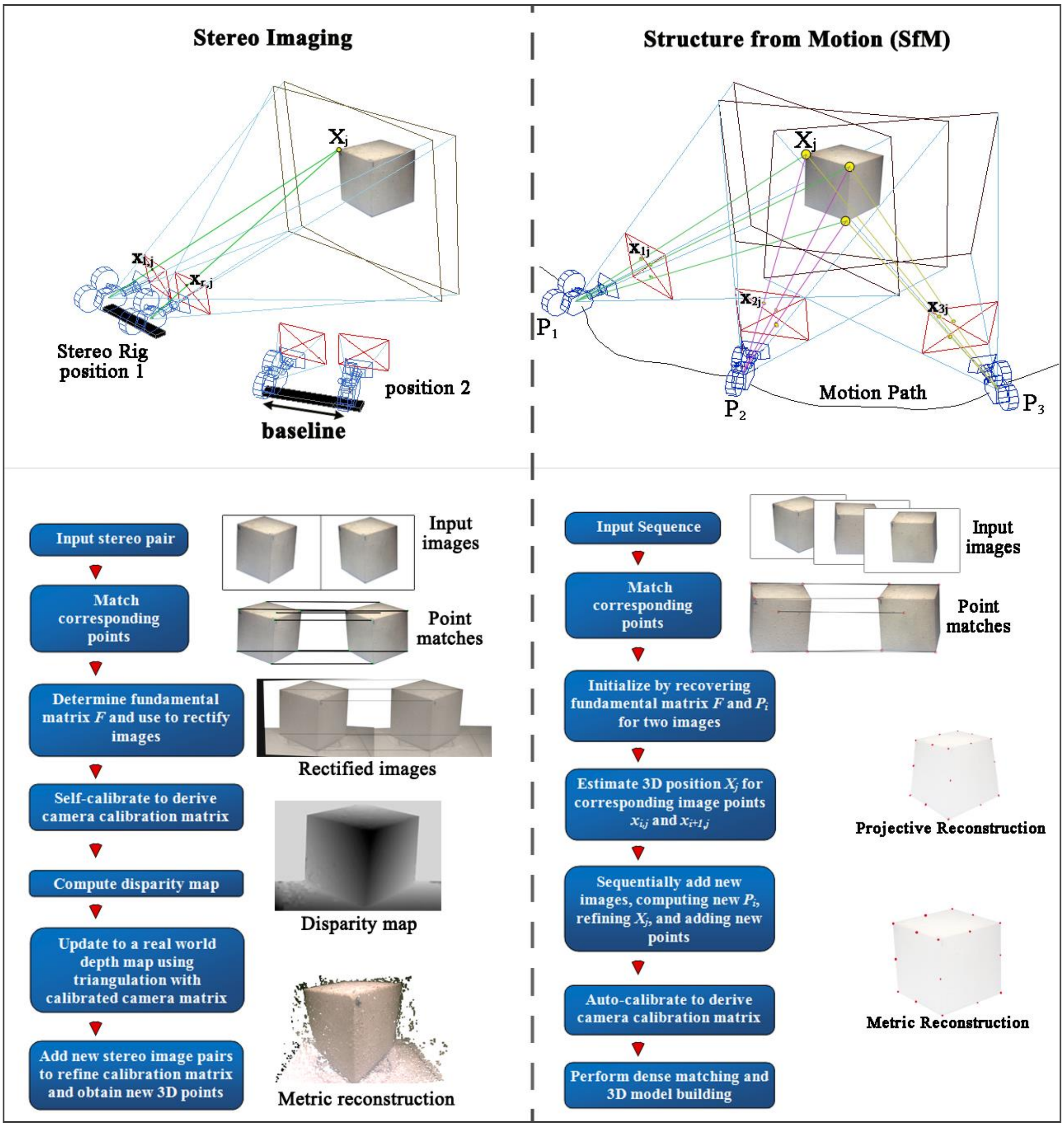


Figure 1: Methodology of Stereo Imaging and Structure from Motion (SfM)

3.1 Stereo Imaging

The first step of the stereo process involves rectifying the input images. This facilitates the correspondence problem as it confines the search space along a horizontal line. Rectification is achieved by applying a transformation to both images which is computed using the fundamental matrix, F. Putative matches must be found in both images (e.g. through a feature detector such as SIFT [14] in order to compute F using the equation:

$$x_r F x_l = 0 \ . \tag{1}$$

where the points x_l and x_r are a pair of matching points in the left and right image respectively. Once F is calculated, and the images are rectified, dense matching can occur. The method used for this paper adopts a normalized cross-correlation to similarity measure within a belief propagation Markov Random Field (MRF) framework to produce a disparity map, which is inversely related to the scene depth. The depth is obtained by firstly calibrating the stereo system and then computing the essential matrix which can be decomposed to give the relative position and orientation of the two cameras, and subsequently allow triangulation of the viewing rays from each camera to get a set of 3D points.

The calibration must be done under the same conditions as that of the image acquisition stage. Self-calibration, or auto-calibration offers an attractive way of determining the intrinsic and extrinsic camera parameters. Self-calibration refers to the process of obtaining a calibrated camera matrix using the constraints in the scene. The stereo method used in this paper adopts a self-calibration procedure as outlined in [11]. In 3D scenes and general motions, each pair of views provides two constraints on the 5 degree-of-freedom calibration. Therefore, three views are the minimum needed for full calibration with fixed intrinsic parameters between views [15]. Quality modern imaging sensors and optics may also provide further prior constraints on the calibration such as zero skew and unity aspect ratio. Integrating these priors will reduce the minimal number of images needed to two. The only additional information required is the baseline distance which enables a properly scaled reconstruction. The accuracy of the cameras parameters obtained through self-calibration is usually lower than that of conventional checkerboard based pre-calibration procedures, however, the practical advantages often outweigh this reduction in accuracy. Moreover, errant camera parameters do not always translate to an appreciably errant reconstruction.

3.2 Structure from Motion (SfM)

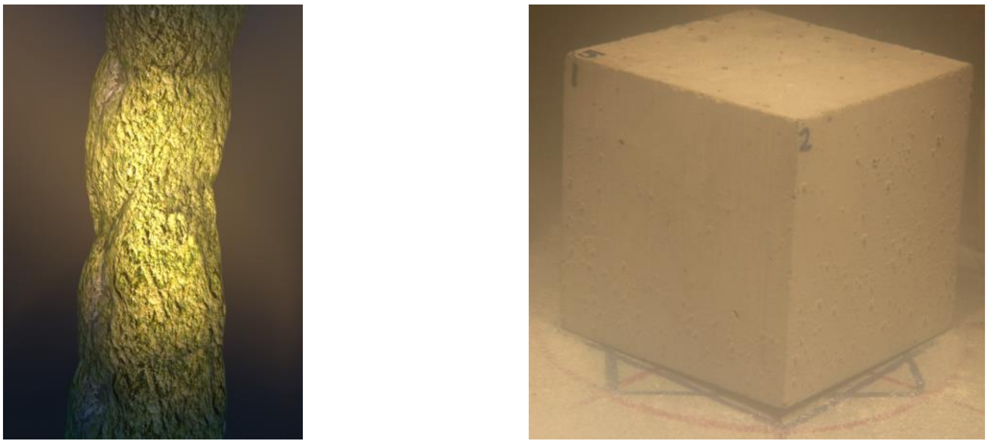
In a typical SfM system, a primitive reconstruction is obtained by firstly matching feature points between two input images. A 3D model is then initialized from two good reconstructions, followed by repeatedly adding matched images, triangulating feature matches, and bundle-adjusting the structure and motion. As in the case of the stereo system, SIFT [14] was used as the feature detector. The success of SfM is heavily dependent on the use of a good feature tracker. This paper employs the SfM method described in [16] as part of the VisualSFM suite, where specific details on its implementation can be found.

Whilst SfM is algorithmically more complex than stereo methods, it is able to combine the benefits of narrow-baseline matching and wide-baseline matching from a video sequence. Narrow-baseline matching occurs when the change in camera position and/or orientation is small, meaning that points of interest will look similar in nearby video frames. This facilitates feature matching, however, the depth computation is quite sensitive to image coordinate measurement noise. On the other hand, wide-baseline matching can suffer from substantial change of scale, different degrees of foreshortening, increased occlusion, and large disparities. All of these factors make it much more difficult to determine correct correspondences automatically, however the depth computation is more accurate.

3.3 Test Data

The stereo and SfM methods are applied to a controlled shape in an underwater setting and synthetic data. The controlled shape is a cube of known dimensions (10 cm x 10 cm x 10 cm). The orthogonal angles and flat surfaces of the cube make it easy visualize the extent of errors in the reconstructed shape as an ill-calibrated camera will not preserve angles and planar surfaces.

Synthetic data was used to imitate the textures and irregular shapes typical of marine growth, whilst still retaining full control of the exact geometrical shape. The synthetic data and real data are illustrated in Figure 2a and 2b respectively.



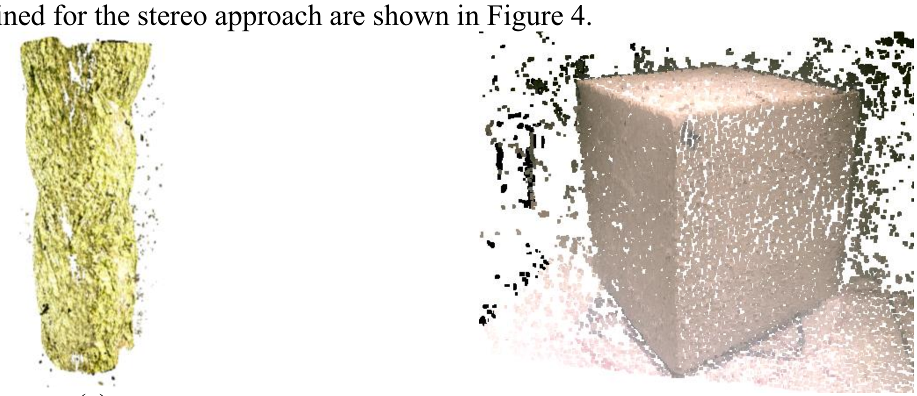
(a) **(b)**

Figure 2: Test imagery: (a) synthetic data, and (b) real underwater concrete cube

4 RESULTS

4.1 Stereo Results

The results obtained for the stereo approach are shown in Figure 4.



(a) (b)

Figure 3: Stereo reconstruction of (a) synthetic data, and (b) concrete cube in an underwater setting

The performance of the stereo approach is measured by comparing some known dimensions against the corresponding dimensions from the reconstructed shapes. These values are provided in Table 1.

Table 1: Stereo Reconstruction.

Specimen	Reconstructed Dimension (cm)	Actual Dimension (cm)
Synthetic Data		
Max diameter	46.5	48.1
Min diameter	35.9	34.5
Cube		
Side Length	9.6	10

Some of the errors encountered from the stereo process, including those due to the self-calibration phase, are conveyed in Figure 4 which depicts a reconstruction from one side of the cube.

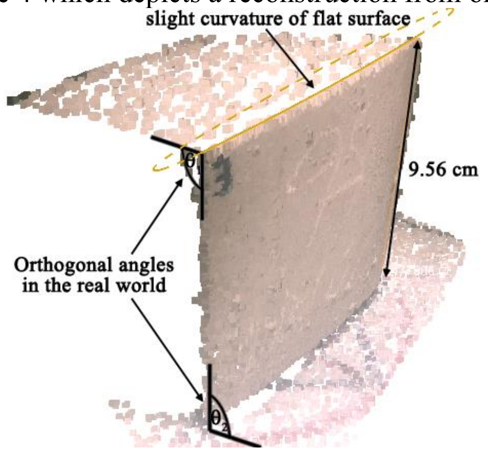


Figure 4: Errors in the stereo process

It may be observed that from Figure 4 that the errors include: 1) a slight curvature of what should be a flat surface, 2) the orthogonal angles in the real world are slightly less in the reconstructed shape (~85degrees), and 3) the scale of the reconstructed cube is slight less than that of the real cube. However, these errors are quite mild and overall it may be considered that the stereo approach performs quite well.

4.2 SfM Results

For the SfM approach, a meaningful reconstruction was only produced for the synthetic data (Figure 5). In this case, the reconstructed synthetic data appears to agree quite well with the original model, however given that SfM systems are incapable of determining a scale factor by themselves, the resulting reconstruction remains scale ambiguous.

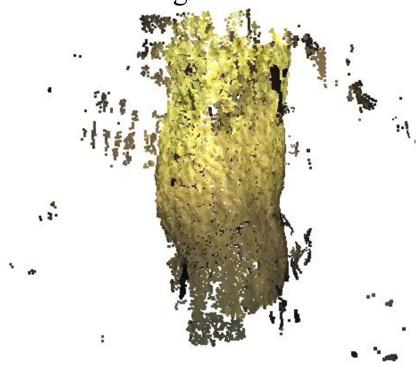


Figure 5: SfM reconstruction of synthetic data

As is the nature of image processing and computer vision, techniques can fail at numerous points along their pipeline due to a host of factors such as scene complexities, sub-optimal image acquisition, and the inability of computers to fully interpret a scene in the same way that a human can. Under slightly different circumstances (e.g. a more textured specimen, different camera setting etc) a satisfactory reconstruction could have been obtained for the underwater concrete cube case. However, this example does that demonstrate that SfM is prone to breakdown and question marks exist over its reliability in an underwater setting.

CONCLUSION

Having an accurate 3D shape reconstruction of offshore infrastructural elements is of great practical importance when analysing the forces exercised by the waves, winds and currents. Image based systems can achieve 3D shape reconstructions in an economical and accessible manner. This paper presents a comparison between two of the most viable classes of image based 3D shape recovery techniques, namely, stereo photography and SfM. These classes were chosen for comparison as the extensive effort and expense associated with undertaking underwater inspections necessitates a review of the most practical approaches. In this paper, details are given about the practical implementation, methodology, additional considerations, and the advantages and disadvantages of each approach, allowing inspectors to decide on which approach best suits their needs.

The two considered classes are performed on a controlled shape in an underwater setting, as well as synthetic data. The results indicate that stereo can provide a reliable solution and allow depth to be extracted from only a single stereo pair, while SfM is more susceptible to breaking down. Therefore, stereo photography is a better option for scenes with some uncertainty. While both of these methods are not necessarily reflective of all stereo and SfM methods, the comparison does serve to underline how each approach may typically respond in an underwater environment.

ACKNOWLEDGMENTS

This work is funded by the Irish Research Council for Science, Engineering and Technology (IRCSET) and CAPACITES/IXEAD society under the Enterprise Partnership Scheme (EPS).

REFERENCES

[1] F. Schoefs, C. Guillo, M. Boukinda, A. Rouhan. Fatigue of Jacket Platforms: Effect of Marine Growth Modeling, Proceedings of the 24th Offshore Mechanics and Arctic Engineering, June 2005.

[2] D. McLeod, J. Jacobson, M. Hardy, C. Embry. Autonomous inspection using an underwater 3D LiDAR. *San Diego: An Ocean in Common*, September 2013.

[3] L. Mullen, A. Laux, B. Concannon, E. P. Zege, I. L. Katsev, A. S. Prikhach. Amplitude-modulated laser imager. *Applied Optics*, 43(19): 3874-92, 2004.

[4] F. Chen, G.M. Brown, S. Mumin. Overview of three-dimensional shape measurement using optical methods. *Optical Engineering*, 39(1): 10-22, 2000.

[5] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, A. V. Razionale. Experimentation of structured light and stereo vision for underwater 3D reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4): 508-518, 2011.

[6] Z. Shaomin, S. Negahdaripour. Recovery of 3D depth map from image shading for underwater applications. *Oceans '97 MTS/IEEE New York Conference Proceedings*, October 1997.

[7] S.K. Nayar, M. Watanabe, M. Noguchi. Real-time focus range sensor. *IEEE International Conference on Computer Vision*, 20-23, June 1995.

[8] V. Brandou, A. G. Allais, M. Perrier, E. Malis, P. Rives, J. Sarrazin, P. M. Sarradin. 3D reconstruction of natural underwater scenes using the stereovision system IRIS. OCEANS 2007 - Europe, June 2007.

[9] A. Olofsson. *Modern stereo correspondence algorithms: investigation and evaluation*. Linköping, 2010.

[10] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330-1334, 2000.

[11] Z. Zhang, Q. T. Luong, O. Faugeras. Motion of an uncalibrated stereo rig: self-calibration and metric reconstruction. *12th International Conference on Pattern Recognition*, October. 1994.

[12] S. Cocito, S. Sgorbini, A. Peirano, M. Valle. 3-D reconstruction of biological objects using underwater video technique and image processing. *Experimental Marine Biology and Ecology*, 297(1): 57-70, 2003.

[13] Y. Dai, H. Li, M. He. A simple prior-free method for non-rigid structure-from-motion factorization. *2012 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2012*, June, 2012.

[14] D.G. Lowe, Object recognition from local scale-invariant features. *Seventh IEEE International Conference on Computer Vision*, September, 1999.

[15] O. D. Faugeras, Q. T. Luong, S. J. Maybank, Camera self-calibration: Theory and experiments, *Computer Vision — ECCV'92*, 321-334, May 1992.

[16] Changchang, W. *Towards Linear-Time Incremental Structure from Motion*. *2013 International Conference on 3D Vision (3DV)*, June 2013.