Underwater 3D Object Reconstruction with Multiple Views in Video Stream via Structure from Motion

Xiao Xu, Renzheng Che, Rui Nian*, Bo He, Meimei Chen
School of Information Science and Engineering
No. 238 Songling Road
Ocean University of China, Qingdao
E-Mails: 380921653@qq.com; crz1336768325@163.com; nianrui_80@163.com

Amaury Lendasse
Arcada University of Applied Sciences, 00550 Helsinki,Finland; Department of Mechanical and Industrial Engineering and the Iowa Informatics Initiative, The University of Iowa, Iowa City, IA 52242-1527, USA;
E-Mails: amaury lendasse@uiowa.edu; lendasse@gmail.com

***Abstract*—Underwater 3D object reconstruction is one of the most essential and fundamental tasks in ocean investigations. In this paper, we try to capture the inherent geometrical variation of 3D objects at multiple visual angels with the underwater vehicles and develop a novel underwater 3D object reconstruction model for the continuous video stream. Image enhancement will be first taken into consideration by guided image filtering to acquire clearer images for object tracking which can be completed by the particle filter. And then, we use scale invariant feature transform (SIFT) and random sample consensus (RANSAC) to detect and correspond the features of objects we have tracked. From these correspondences, the 3D point cloud reconstruction will be finally developed with the help of structure from motion (SFM) and patches-based multi-view stereo (PMVS). It has been shown in the simulation experiments that the developed scheme of this paper achieves consistent performance improvements over 3D object reconstruction for underwater video stream in multiple views.**

Keywords—3D reconstruction; Structure from motion; Guided image filtering; Particle filter;

I. Introduction

3D reconstruction has achieved remarkable success for the land-based systems [1]. However, there are still a variety of difficulties in the accurate 3D reconstruction from images captured by underwater cameras due to the imaging visibility, light absorption, non-uniform illumination, etc [2]. For the underwater imaging system itself, the quality of image degrades because of the effects of light attenuation and scattering [3, 4].

Dating back to the history, most of the approaches developed for 3D object reconstruction, capture and recover the shape, structure and appearance of the real 3D objects from the motion, stereo vision and monocular cues about 2D images collected at different view angles, including hints in shading, texture, contour, shadow, defocus, vanishing points, etc [5]. Among them, SFM algorithms, which attempt to deliver a complete quantitative solution to 3D object reconstruction by the pose estimation in motion and the geometrical changes in translation, rotation, the orientation, is particularly applicable to those underwater mobile observation tasks deployed by the remotely operated vehicles (ROVs) and the autonomous underwater vehicles (AUVs).

In this paper, the stereo vision system has been set up to capture the inherent geometrical variation of 3D objects at multiple visual angels and develop a novel underwater 3D object reconstruction model for the continuous video stream combining SFM with object tracking strategies. The paper is organized as follows: section II describes the basic framework of underwater 3D object reconstruction with SFM in our system. Section III introduces the preprocessing. Section IV is about the object tracking. Section V refers to the structure from motion. Section VI is about dense point cloud reconstruction. Section VII shows the experimental results and Section VIII comes to the conclusions.

II. General Framework

In this paper, we make an attempt to explore a novel underwater 3D object reconstruction model for the continuous video stream with the help of combining SFM with object tracking strategies. A brief flow chart of underwater 3D object reconstruction with SFM is shown in Fig. 1. One guided filter [6, 7] is first introduced for image enhancement. The particle filter [8-10] is further taken as the object tracking tool in image sequence with multiple views. The corresponding key features between adjoining underwater images are then extracted in the object tracking process by SIFT [11, 12], and the false matching features could be further identified and eliminated by RANSAC [13] with one minimal subset of the key features remained in image sequence. We adopt SFM to recover and estimate the position of camera calibration and the geometry of underwater scene with sparse 3D point cloud by triangulation, iterative process, and bundle adjustment [14, 15]. At last, the patches-based multi-view stereo [16] is proposed for dense point cloud reconstruction.

III. Preprocessing

Since the poor visibility is one major limitation in the underwater vision. One guided filter is first introduced for

978-1-4673-9724-7/16/$31.00 ©2016 IEEE

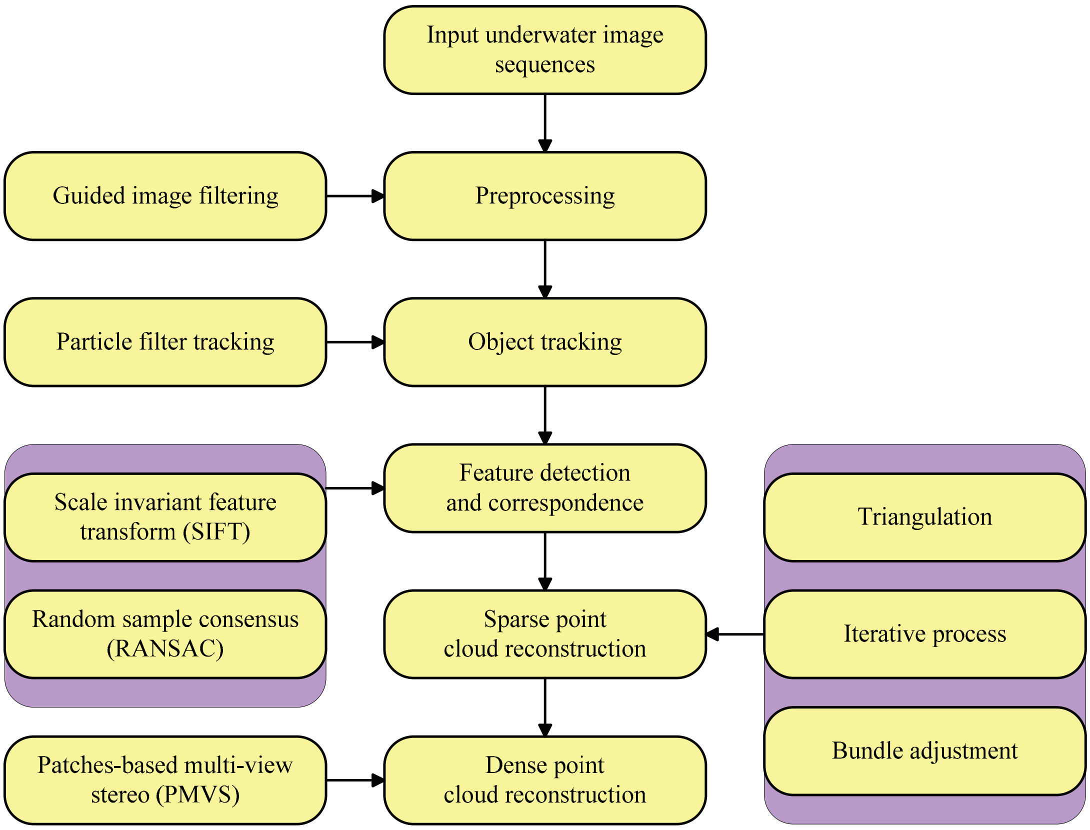


Fig. 1. The flow chart of underwater 3D Object Reconstruction with SFM

those original underwater images in the video stream to perform a fast and non-approximate edge preservation and promote the details like dehazing and feathering, regardless of the kernel size and the intensity range.

Let $\{s_1, \cdots s_n, \cdots s_N\}$ be N input underwater images collected in our system, I_n be the n-th guided underwater image and q_n be the n-th output underwater image. The local linear transform [6] is,

$$q_n^i = a^j I_n^i + c^j, \forall i \in \omega^j \tag{1}$$

Where i and j are pixel indexes and $\left(a^j, c^j\right)$ are some linear coefficients assumed to be constant in the window ω^j centered at the pixel j.

To determine the linear coefficients, we minimize following cost function in the window,

$$E\left(a^j, c^j\right) = \sum_{i \in \omega^j} \left(\left(a^j I_n + c^j - p_n \right)^2 + \varepsilon \left(a^j \right)^2 \right) \tag{2}$$

Where ε is a regularization parameter preventing a^j from being too large.

IV. Object Tracking

The particle filter is further taken as the object tracking tool in image sequences with multiple views to focus on the motion trajectories of underwater 3D objects all the time.

Let $\mathbf{b}_n$ be the state vectors of underwater 3D objects in n-th underwater image q_n, $\mathbf{z}_n$ be all observation vectors of underwater 3D objects from the first n underwater images $\{q_1, \cdots q_n\}$ and M be the number of particles or samples. The realization steps of particle filter are described as follows.

a) Initialization: Sample the $\left\{\mathbf{b}_1^k\right\}_{k=1}^M$ particle set at q_1, the particle weight $\left\{G_1^k\right\}_{k=1}^M$ is the average value $1/M$.

b) Sampling: Set underwater image as q_n. The $\left\{\mathbf{b}_n^k\right\}_{k=1}^M \sim f\left(\mathbf{b}_n \middle| \mathbf{b}_{0:n-1}, \mathbf{z}_n\right)$ is obtained by sampling from the importance density function $f\left(\mathbf{b}_{0:n} \middle| \mathbf{z}_n\right)$.

c) Weight updating: Weight is calculated and normalized the each particle when the observation value coming,

$$G_n^k \infty G_{n-1}^k \frac{p\left(\mathbf{z}_n \middle| \mathbf{b}_n^k\right) p\left(\mathbf{b}_n^k \middle| \mathbf{b}_{n-1}^k\right)}{f\left(\mathbf{b}_n^k \middle| \mathbf{b}_{0:n-1}^k, \mathbf{z}_n\right)}, \tilde{G}_n^k = \frac{G_n^k}{\sum_{k=1}^M G_n^k} \tag{3}$$

Posterior probability density is approximately defined $p\left(\mathbf{b}_n \middle| \mathbf{z}_n\right) = \sum_{k=1}^{M} G_n^k \delta\left(\mathbf{b}_n - \mathbf{b}_n^k\right)$, where δ is the Dirac delta function.

d) Resampling: According to $\left\{\tilde{G}_n^k\right\}_{k=1}^{M}$ of the particle samples $\left\{\mathbf{b}_0^k\right\}_{k=1}^{M}$, renewed obtain the particle samples $\left\{\tilde{\mathbf{b}}_n^k\right\}_{k=1}^{M}$ and $\left\{\tilde{G}_n^k\right\}_{k=1}^{M}$ should be averaged. Then, the posterior probability density is approximately defined $\hat{p}\left(\mathbf{b}_n \middle| \mathbf{z}_n\right) = \frac{1}{M}\sum_{k=1}^{M} \delta\left(\mathbf{b}_n - \tilde{\mathbf{b}}_n^k\right)$.

e) Output: Output the estimate state $\tilde{\mathbf{b}}_n = \sum_{k=1}^{M} \tilde{G}_n^k \mathbf{b}_n^k$ and N underwater 3D objects region images $\left\{B_1, \cdots B_n, \cdots B_N\right\}$.

V. STRUCTURE FROM MOTION

SFM is applied here to recover and estimate the position of camera calibration and the geometry of underwater scene with sparse 3D point cloud.

A. Triangulation

The location of one unknown 3D point in the underwater environment is first determined by measuring its angle from three known points at either end of a fixed baseline for triangulation until a mesh of triangles is established at the largest appropriate scale. In this way, those unknown 3D points inside the triangles can all be accurately located with reference to the particular 3D point in the underwater environment.

After feature detection and correspondence by SIFT and RANSAC, we obtain sets of image point correspondences $\mathbf{u}'$ and $\mathbf{u}''$. From these correspondences, we take the eight-point algorithm [17] to calculate the fundamental matrix F. The essential matrix E can be analytically determined through fundamental matrix F and the known internal camera parameters K and K', $E = K'^T FK$. We acquire relative position information (rotation matrix R and translation matrix T) by resolving E, then the two camera matrix $\mathbf{c}_1$ and $\mathbf{c}_2$ corresponding to underwater images B_1 and B_2 can be computed. We suppose that the b-th spatial point $\mathbf{x}_{12}^b$ in the real world space is visible in B_1 and B_2. Let $\mathbf{u}_{12}^{b\prime}$ and $\mathbf{u}_{12}^{b\prime\prime}$ be projections of the point $\mathbf{x}_{12}^b$ in B_1 and B_2.

Suppose that $\mathbf{u}_{12}^{b\prime} = \mathbf{c}_1 \mathbf{x}_{12}^b$. We write in homogeneous coordinates,

$$\mathbf{u}_{12}^{b\prime} = w(u_{12}^{b\prime}, v_{12}^{b\prime}, 1)^T \tag{4}$$

where $(u_{12}^{b\prime}, v_{12}^{b\prime}, 1)$ is the homogeneous coordinates of $\mathbf{u}_{12}^{b\prime}$ corresponds to $\mathbf{c}_1$, w is an unknown scale factor and $\mathbf{c}_{1l}^T$ is the l-th row of the matrix $\mathbf{c}_1$. Now, the equation $\mathbf{u}_{12}^{b\prime} = \mathbf{c}_1 \mathbf{x}_{12}^b$ may be written as,

$$\left\{\begin{array}{l} wu_{12}^{b\prime} = \mathbf{c}_{1,1}^T \mathbf{x}_{12}^b \\ wv_{12}^{b\prime} = \mathbf{c}_{1,2}^T \mathbf{x}_{12}^b \\ w = \mathbf{c}_{1,3}^T \mathbf{x}_{12}^b \end{array}\right\} \tag{5}$$

Eliminating w using the third equation, we arrive at,

$$\begin{array}{l} (u_{12}^{b\prime} \mathbf{c}_{1,3}^T - \mathbf{c}_{1,1}^T) \mathbf{x}_{12}^b = 0 \\ (v_{12}^{b\prime} \mathbf{c}_{1,3}^T - \mathbf{c}_{1,2}^T) \mathbf{x}_{12}^b = 0 \end{array} \tag{6}$$

It may be written in the form $A\mathbf{x}_{12}^b = 0$ for a suitable 4×4 matrix, A. These equations define $\mathbf{x}_{12}^b$ only up to an indeterminant scale factor, and we seek a nonzero solution for $\mathbf{x}_{12}^b$. One finds $\mathbf{x}_{12}^b$ to minimize $\left\|A\mathbf{x}_{12}^b\right\|$ subject to the condition $\left\|\mathbf{x}_{12}^b\right\| = 1$. The solution is the unit eigenvector corresponding to the smallest eigenvalue of the matrix $A^T A$. It can be solved using the singular value decomposition or Jacobi's method for finding eigenvalues of symmetric matrix [18,19].

B. Iterative Process

Enough 3D sparse points in underwater image sequences are then selected for a reliable 3D reconstruction performance in an iterative process.

C. Bundle Adjustment

After iterative process, we get m spatial points $\{\mathbf{x}^1 \cdots, \mathbf{x}^b \cdots, \mathbf{x}^m\}$. Bundle adjustment is further taken to boil down the sparse point cloud by minimizing the reprojection error between the locations of the observed and predicted 3D points. This optimization problem is usually formulated as a non-linear least squares problem, where the error is the squared L_2 norm of the difference between the observed feature location and the projection of the corresponding 3D point on the image plane of the camera.

Let o_{bn} denote the binary variables that equal 1 if point $\mathbf{x}^b$ is visible in underwater image B_n and 0 otherwise. Assume also that each underwater image B_n is parameterized by a vector B_n' and each 3D point by a vector $\mathbf{x}^{b\prime}$. Let $\mathbf{u}_n^b$ be the projection of the b-th point on B_n, $\mathbf{e}$ be a vector of parameters and $h(\mathbf{e})$ be the vector of residuals/reprojection errors for a 3D reconstruction,

$$h(\mathbf{e}) = \sum_{b=1}^{m}\sum_{n=1}^{N} o_{bn} J(K(B_n', \mathbf{x}^{b'}), \mathbf{u}_n^b)^2 \tag{7}$$

where $K(B_n', \mathbf{x}^{b'})$ is the predicted projection of point $\mathbf{x}^b$ on underwater image B_n and $J(\cdots)$ denotes the Euclidean distance. Then the optimization problem we wish to solve is the non-linear least squares problem,

$$\mathbf{e}^* = \arg\min_{\mathbf{e}} \|h(\mathbf{e})\| \tag{8}$$

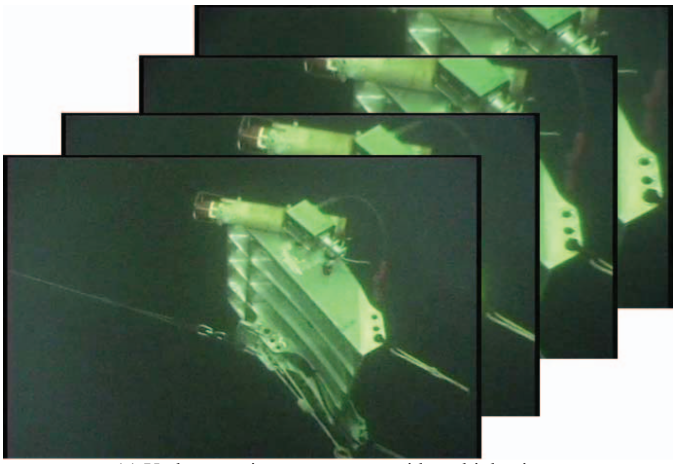
The Levenberg-Marquardt (LM) algorithm [20] is the most popular algorithm for solving non-linear least squares problems, and is the algorithm of choice for bundle adjustment. LM operates by solving a series of regularized linear approximations to the original nonlinear problem.

VI. Dense Point Cloud Reconstruction

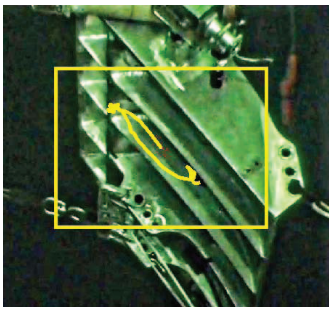
The patches-based multi-view stereo (PMVS) algorithm is at last taken for a dense point cloud reconstruction for underwater 3D objects to spread the initial sparse point matching to those nearby pixels in the underwater image sequence and eliminate incorrect point matching lying either in front of or behind the observed underwater 3D object surface with a series of the visibility constraints.

VII. Simulation Experiment

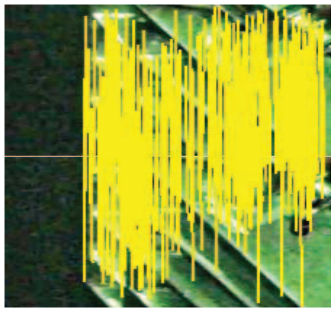
In the experiment simulations, we set up a Myring streamline AUV system with CCD camera with the horizontal resolution 480 TVL/PH and the minimum scene illumination 0.28 lux on board for 3D object reconstruction tasks. Fig. 2 lists the 3D reconstruction performance for an example underwater image sequence by our proposed approach. It is shown from our simulation experiments that the developed scheme of this paper achieves consistent performance improvements over 3D object reconstruction for underwater video stream in multiple views.



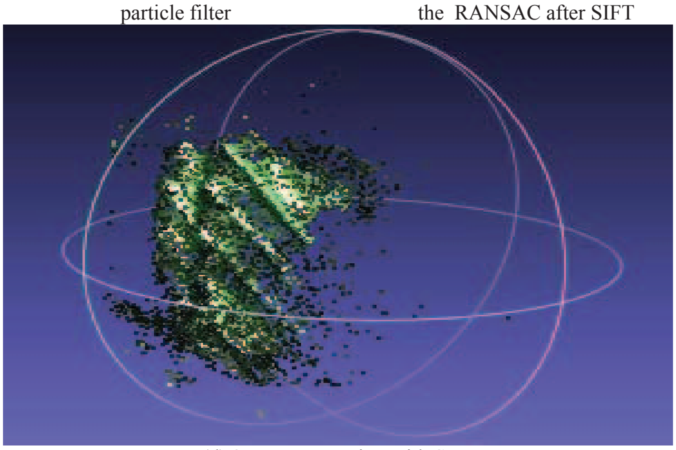
(a) Underwater image sequence with multiple views



(b) Object tracking trajectory with particle filter



(c) Feature correspondence with the RANSAC after SIFT



(d) 3D reconstruction with SFM

Fig. 2. 3D reconstruction of underwater images in sequence with multiple views

VIII. Conclution

In this paper, we have put forward a novel approach of the underwater 3D object reconstruction model for the continuous video stream combining SFM with object tracking strategies. The particle filter has been introduced as the object tracking tool in image sequence with multiple views to focus on the motion trajectories of underwater 3D objects all the time. SFM is then established with triangulation, iterative process, and bundle adjustment, to recover and estimate the position of camera calibration and the geometry of underwater scene with sparse 3D point cloud. The simulation results have shown the effectiveness and feasibility of the proposed approach, which is comparable to the direct 3D object reconstruction in underwater image representation.

Acknowledgment

This work is partially supported by the National High-Tech R&D 863 Program (2014AA093410), the Natural Science Foundation of P. R. China (31202036), the National Program of International S&T Cooperation (2015DFG32180), the National Science & Technology Pillar Program (2012BAD28B05), and the Natural Science Foundation of P. R. China (41376140).

References

[1] W. Abbeloos, "Real-Time Stereo Vision", Karel de Grote-Hogeschool University College, Belgium, 2010.

[2] W. Hou, S. Woods, E. Jarosz, W. Goode, A. Weideman, “Optical turbulence on underwater image degradation in natural environments”, Appl. Opt. 51, 678–2686, 2012.

[3] R. Nian, F. Liu, B. He, “Early underwater artificial vision model in ocean investigations via independent component analysis”, Sensors, 13(7): 9104-9131, 2013.

[4] R. Nian, B. He, J. Yu, Z. M. Bao, Y.F Wang, “ROV-based underwater vision system for intelligent fish ethology research”, International Journal of Advanced Robotic Systems, 2013.

[5] R. Nian, G. Ji, W. Zhao, C. Feng, “3D Object Recognition from 2D Invariant View Sequence under Translation, Rotation and Scale by Means of ANN Ensemble”, Proceedings 2005 IEEE International Workshop on VLSI Design and Video Technology, 2005.

[6] K. He, J. Sun, and X. Tang, “Guided image filtering”, Computer Vision–ECCV 2010, Springer Berlin Heidelberg, 1-14, 2010.

[7] J. Zhang, G. Lin, L. Wu, Y. Cheng, “Speckle filtering of medical ultrasonic images using wavelet and guided filter”, Ultrasonics, 2016

[8] P. Fearnhead, “Sequential Monte Carlo methods in filter theory”, Diss, University of Oxford, 1998.

[9] N. Sun, R. Nian, B. He, T. Yan, “Consistent fish tracking via multiple underwater cameras”, OCEANS 2014-TAIPEI. IEEE, 2014.

[10] H. Liang, F. Kang, “Tracking UUV based on interacting multiple model unscented particle filter with multi-sensor information fusion”, Optik-International Journal for Light and Electron Optics ,2015.

[11] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, 2004.

[12] M. Chen, R. Nian, B. He, S. Qiu, X. Liu, T. Yan, “Underwater Image Stitching based on SIFT and Wavelet Fusion”, OCEANS’15 MTS/IEEE Genova, 2015.

[13] M. A. Fischler, R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”, Commun, ACM, 24(6): 381-395, 1981.

[14] B. Triggs,P. F. McLauchlan, R. I. Hartley, “Bundle adjustment—a modern synthesis”, Vision algorithms: theory and practice, Springer Berlin Heidelberg, 298-372, 1999.

[15] C. Wu, S. Agarwal, B. Curless, “Multicore bundle adjustment”, Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011.

[16] Y. Furukawa, J. Ponce, “Accurate, dense, and robust multiview stereopsis”, Pattern Analysis and Machine Intelligence, IEEE Transactions on 32(8): 1362-1376, 2010.

[17] R. I. Hartley, “In defense of the eight-point algorithm”. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1997.

[18] K. E. Atkinson, “An Introduction to Numerical Analysis”, John Wiley & Sons, 2008.

[19] W. H. Press, “Numerical Recipes in C: The Art of Scientific Computing”, Cambridge University Press, Cambridge, 2007.

[20] J. Nocedal, S. Wright, “Numerical optimization”, Springer Science & Business Media, 2006.