Literature Survey on Building 3-D Models of underwater Objects

B L Mukundappa[1] N. Satish Kumar[2] Ramakanth Kumar P[3]

[1] Associate Prof., Dept. of Computer Science, University College of Science, Tumkur, India
[2] Research Scholar, Department of CSE. R. V. College of Engineering, Bangalore, India
[3] Professor and H.O.D, Department of ISE, R V College of Engineering, Bangalore, India

Abstract:
Three-dimensional object recognition is a topic of active interest motivated by a desire to build computers with "humanlike" visual capabilities, and also by a pragmatic need to aid numerous real world applications such as robot bin-picking, autonomous navigation, automated visual inspection and parts assembly tasks. Building 3D models from multiple image sequences is a classical but challenging area in computer vision and computer graphics research. We propose a 3D modeling method that drastically improves robustness over existing approaches, and explicitly deals with underwater objects. Generated 3D models can be used to represent 3D objects in the first stage of underwater 3D target recognition system for Autonomous Underwater Vehicles (AUVs).

Keywords
AUV, ROV, Voxel, 3D Modeling, Image Registration

1. Literature Review

1.1 3D Modelling:

Soon- Yong Park et al. introduced a stereo vision system to automatically generate 3D models of real objects [1]. Calibration of the stereo camera and the turntable stage is presented. They have rectified stereo images and obtained range images of an object from multiple viewpoints. There are two major approaches achieved in this paper. The first one is based on merging multi view range images into a complete 3D model and the second one is based on processing photographic images using a volumetric reconstruction technique, such as **voxel Coloring** and **shape-from-silhouettes**.3D model generation is based on the merging of multi-view range images obtained from a digital stereo camera. The range images obtained are then automatically registered to a common coordinate system and integrated into a 3D mesh model.

1.2 Camera Calibration:

Roger Y. Tsai developed a new technique for three-dimensional (3D) camera calibration for machine vision metrology using the off-the –shelf TV cameras and lenses is described [2]. The two-stage technique is aimed at efficient computation of camera external position and orientation relative to object reference coordinate system as well as the effective focal length, radial lens distortion, and image scanning parameters. The two-stage technique has advantage in terms of accuracy, speed and versatility. The experimental results are analyzed and compared with theoretical prediction. Qurban Memont defined that Stereo-pair images from two cameras can be used to compute three-dimensional (3D) world coordinates of a point using triangulation. He proposed that, for stereo vision applications in which real-world coordinates are to be evaluated [3]. Direct linear transformation introduced by Abdel-Aziz and Karara has also been extended to incorporate distortion parameters. They also told that to compute world coordinates from stereo images requires first matching the images obtained from two different cameras to determine disparities (difference in positions of corresponding features).
Hence, camera calibration is essential to compute world coordinates from stereo-images.
Matthew Bryant et.al tracked the relative position of underwater targets [4]. They have developed a new Point identification algorithm, based on planar projective invariant indexing. They have developed algorithm which is 80% more reliable and tested a camera calibration scheme suitable for underwater computer vision applications.

1.3 Object Recognition:

Brandou et.al proposed a 3-dimension reconstruction method of small scale scenes improved by a new image acquisition method [5]. They have used stereovision system to acquire images in order to obtain several shots of an object. They have used the SIFT method (Scale Invariant Feature Transform) because the interest points are invariant to image scaling and rotation, and partially invariant to changes in illumination and 3D camera view points. Finally they have presented a complete reconstruction method of small-scale natural underwater objects.
David G.Lowe et.al has presented a new method for image feature generation called the Scale Invariant Feature Transform (SIFT) [6]. This approach transforms an image into a large collection of local feature vectors, each of

Manuscript received February 5, 2012
Manuscript revised February 20, 2012

which is invariant to image translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. In this paper the scientists achieved maximum robustness by detecting many feature types and relying on the indexing and clustering to select those that are most useful in a particular image.
Piotr jasiobedzki described about initial results of underwater experiments [7]. They have developed for recognizing and tracking natural objects. The SIFT Object Recognition and tracking (SORT) uses a set of high level natural visual features called Scale Invariant Feature Transform (SIFT) for tracking. SIFT was developed for image feature generation in object recognition applications.

1.4: Rectification

Andrea Fusilello et al. presented a linear rectification algorithm for general, unconstrained stereo rigs [8]. The algorithm takes the two perspective projection matrices of the original cameras, and computes a pair of rectifying projection matrices. The rectified images can be thought of as acquired by a new stereo rig, obtained by rotating the original cameras. The important advantage of rectification is that computing stereo correspondences is made simpler, because search is done along the horizontal lines of the rectified images. Tests proving the behavior of their method, as well as the negligible decrease of the accuracy of 3D reconstruction performed from the rectified images directly were reported.

1.5 Dense Matching on the rectified images:

Sebastien Roy et al. described a new algorithm for solving N- camera stereo correspondence problem by transforming it into a maximum-flow problem [9]. Results have shown improved depth estimation as well as better handling of depth discontinuities. A multi-resolution approach as well as local smoothness variations could be directly embedded in the graph, improving performance and depth map equality.
Vladimir kolmogorov et al. gives a more general energy minimization formulation of the problem, which allows a larger class of spatial smoothness constraints [10]. In this paper, they approached the scene reconstruction problem from the point of view of energy minimization. They build upon two algorithms that give an energy minimization formulation of the scene reconstruction problem, and then minimize the energy using graph cuts. Both of these algorithms treat the input images symmetrically, handle visibility constraints correctly, and allow spatial smoothness to be enforced. The energy can be efficiently minimized using graph cuts, and has given good experimental results.
The algorithm uses a problem formulation that is restricted to a two- camera stereo, and imposes smoothness between a pair of cameras. The energy that we minimize treats the input images symmetrically, handles visibility properly, and imposes spatial smoothness while preserving discontinuities. The algorithm can handle an arbitrary number of cameras, but imposes smoothness with respect to a single camera. [11, 12]

2. To acquire Range Image:

Peter J. Burt. et al. described a technique for image encoding in which local operators of many scales but identical shape serve as the basis functions . From a rectified stereo image pair they acquired a range image by employing a multi-resolution stereo matching technique using a Gaussian Pyramid[13].The result is a net data compression since the difference or error, image has low variance and entropy, and the low- pass filtered image may represent at reduced sample density. A further advantage of the present code is that it is well suited for many image analysis tasks as well as for image compression. In this representation, image features of various sizes are enhanced and are directly available for various image processing.

3. Implicit Representation of Object's surface:

Yi Zhang et al. implemented a fast algorithm of extracting isosurface. The fast visualization algorithm is applied to the implicit surfaces represented by Radial Basis Functions and Multi- level Partition of Unity [14]. Of all approaches, Radial Basis Functions (RBF) has some advantages, for instance, reconstructing implicit surfaces from large point-clouds and repairing incomplete meshes. Multilevel Partition of Unity (MPU) is a fast reconstruction method and it can also offer quasi Euclidean distance to the isosurface for a space point, which is necessary in our algorithm. By utilizing the mathematical properties of these implicit functions, the difficulty of calculating the shortest distance from one space point to the surface is solved. From the experiment results, the algorithm speeds up the process of extracting the isosurface.

4. Multi- view Registration:

Robert Bergevin et al. presented an algorithm that reduces significantly the level of the registration errors between all pairs in a set of range views [15]. It is an instance of a category of registration algorithms known as iterated closest- point (ICP) algorithms. Experimental results show that this refinement technique improves the calibrated registrations and the quality of the integrated model for complex multi- part objects.

Paul J. Besl et al. described a general- purpose, representation- independent method for the accurate and computationally efficient registration of 3-D shapes including free- form curves and surfaces [16]. The method handles the full six degrees of freedom and is based on the iterative closest point (ICP) algorithm. Experimental results show the capabilities of the registration algorithm on point sets, curves and surfaces.

Joachim Bauer et.al presents a method for the registration of range image pairs that were captured from different locations [17]. They presented a method for the coarse alignment of range image pairs under the condition that both scans map the same 3D plane. A robust method was used for the detection of 3D planes based on the RANSAC principle were introduced. An orthorectification process, using the extracted the 3D plane was used to produce range images without perspective distortions.

Chen and Medioni have used range images which directly provide access to three dimensional information [18]. They proposed a new approach which works on range data directly, and registers successive views with enough overlapping area to get an accurate transformation between views.

Umberto Castellani et al. proposed a technique for the three dimensional reconstruction of an underwater environment from multiple acoustic range views acquired by a remotely operated vehicle [19]. Moreover they have introduced a statistically sound thresholding to improve iterative closest point (ICP) robustness against noise and non- overlapping data.

Soon-Yong Park et al addressed refinement problem and presents an accurate and fast Point-to-(Tangent) Plane technique [20]. Point-to-Plane approach is known to be very accurate for registration refinement of partial 3D surfaces. Experimental results showed that their approach was very accurate and fast for both pair-wise registration and multi-view registration problems.

5. Integration:

Soon-Yong Park et al. presented Automatic reconstruction of a complete 3D model of a complex object. The complete 3D model is reconstructed by integrating two 3D models which are reconstructed from different poses of the object [21]. In order to reconstruct all visible surfaces of a complex object with concavities and holes, two 3D models from different poses of the object are reconstructed and integrated to obtain the complete 3D model. For each pose of the object, a 3D model is reconstructed by combining stereo image analysis, shape from silhouettes, and a volumetric integration technique. Pose integration improves the results by reconstructing a complete 3D model closer to the original object.

Anil K.Jain et. al developed a complete prototype system for automatically registering and integrating multiple views of objects from range data. The view integration scheme uses a weighted averaging technique for merging the registered views together to result in a smooth surface [22]. A smooth transformation of data can be achieved on the view boundary by calculating the depth value in the merged image using the weighted averaging algorithm.

Acknowledgment:

I owe my sincere feelings of gratitude to Dr.S.C.Sharma for his valuable guidance and suggestions which helped me a lot to write this paper. It gives us great pleasure to express my feelings of gratitude to Dr. Ramakanth Kumar for valuable guidance support and encouragement.

References

[1] Soon-Yong Park1, Murali Subbarao2A multiview 3D modeling system based on stereo vision techniques, Machine Vision and Applications (2005) 16: 148–156

[2] Roger Y Tsai A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses, IEEE JOURNAL OF ROBOTICS AND AUTOMATION, VOL. RA-3, NO. 4, AUGUST 1987

[3] Qurban Memony and Sohaib Khan Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks International Journal of Systems Science, 2001, volume 32, number 9, pages 1155 – 1159

[4] Matthew Bryantt, David Wettergreen, Samer Abdallaht, Alexander Zelinsky Robust Camera Calibration for an Autonomous Underwater Vehicle

[5] 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 Manuscript received March 30, 2007. 1 4244-0635-8/07 ©2007 IEEE, pp.1-6

[6] David G. Lowe, Object Recognition from Local Scale-Invariant Features, Proc. of the International Conference on Computer Vision, Corfu (Sept. 1999), pp.1-8

[7] Piotr Jasiobedzki, Stephen Se, Michel Bondy, Roy Jakola, Underwater 3D Mapping and Pose Estimation for ROV Operations, MDA, Space Missions, 9445 Airport Rd, Brampton, ON, L6S 4J3, Canada

[8] Andrea Fusiello1, Emanuele Trucco2, Alessandro Verri3, A compact algorithm for rectification of stereo pairs, Machine Vision and Applications (2000) 12: 16–22

[9] Skbastien Roy* Ingemar J. Cox, A Maximum-Flow Formulation of the N-camera Stereo Correspondence

Problem, NEC Research Institute 4 Independence Way Princeton, NJ 08540, U.S.A.

[10] Vladimir Kolmogorov1, Ramin Zabih1, and Steven Gortler2, Generalized Multi-camera Scene Reconstruction Using Graph Cuts, Computer Science Department, Cornell University, Ithaca, NY 148532 Computer Science Department, Harvard University, Cambridge, MA 02138

[11] David G. Lowe, Object Recognition from Local Scale-Invariant Features, Proc. of the International Conference on Computer Vision,Corfu (Sept. 1999), pp 1-8

[12] Piotr Jasiobedzki, Stephen Se, Michel Bondy, Roy Jakola, Underwater 3D Mapping and Pose Estimation for ROV Operations,

[13] Andrea Fusiello, Emanuele Trucco, Alessandro Verri, A compact algorithm for rectification of stereo pairs, Machine Vision and Applications (2000) 12: 16–22, Received: 25 February 1999 / Accepted: 2 March 2000

[14] Skbastien Roy* Ingemar J. Cox, A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem, pp 492-499

[15] Vladimir Kolmogorov1, Ramin Zabih1, and Steven Gortler2, Generalized Multi-camera Scene Reconstruction Using Graph Cuts, pp 1 -16

[16] Vladimir Kolmogorov and Ramin Zabih, Multi-camera Scene Reconstruction via Graph Cuts, pp 2-16

[17] Vladimir Kolmogorov Ramin Zabih, Computing Visual Correspondence with Occlusions via Graph Cuts, pp 1-37

[18] PETER J. BURT, MEMBER, IEEE, AND EDWARD H. ADELSON, The Laplacian Pyramid as a Compact Image Code, IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM-3l, NO. 4, APRIL 1983, pp 532-540

[19] Yi Zhang, Xin Wang, Xiao Jun Wu, Fast Visualization Algorithm for Implicit Surfaces, Proceedings of the 16th International Conference on Artificial Reality and Telexistence--Workshops (ICAT'06) 0-7695-2754-X/06

[20] Robert Bergevin, Marc Soucy, Hewe Gagnon, and Denis Laurendeau, Towards a General Multi-View Registration Technique, IEEETRANSACTIONS ON PATTERN Analysis and machine Intelligence, Vol 18 No. 5, May 1996

[21] Pual J Besl, Member IEEE, and Neil D McKay, A Method for Registration of 3D Shapes.IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 14. No. 2, Feb 1992

[22] Joachim Bauer, Konrad Karner, Andreas Klaus, Roland Perko, Robust Range Image Registration using a Common Plane, WSCG POSTERS proceedings WSCG'2004, February , 2004.

[23] Yang Chen and Gerard Medioni, Object Modeling by Registration of multiple range images, Proceedings of the 1991 IEEE International conference on Robotics and Automation, Sacramento California, pp 2724- 2729

[24] Umberto Castellani, Andrea Fusiello, and Vittorio Murino, Registration of Multiple Acoustic Range Views for Underwater Scene Reconstruction, Computer Vision and Image Understanding 87, 78–89 (2002) doi:10.1006/cviu.2002.0984, 1077-3142/02 $35.00 _c 2002 Elsevier Science (USA)

[25] Soon-Yong Park and Murali Subbarao, A Fast Point-to-Tangent Plane Technique for Multi-view Registration, Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling (3DIM'03) 0-7695-1991-1/03

[26] Soon-Yong Park and Murali Subbarao, AUTOMATIC 3D MODEL RECONSTRUCTION USING VOXEL CODING AND POSE INTEGRATION, IEEE ICIP 2002, 0-7803-7622-6/02, pp 533-536

[27] Chitra Dorai,, Anil K. Jain, Registration and Integration of Multiple Object Views for 3D Model Construction, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 20, NO. 1, JANUARY 1998, pp 83-89

[28] Xavier Armangue, Joaquim Salvi, , Overall view regarding fundamental matrix estimation, Image and Vision Computing 21 (2003) 205–220

[29] Nassir Navab, 3D Computer Vision, "Multiple View Geometry" by Hartley & Zisserman

Mukundappa B L got B.Sc Degree with Physics, Chemistry & Mathematics as major subjects from Mysore University, M.Sc degree in Chemistry & Computer Science. He has been working as Principal & Associate professor in University science college, Tumkur & has 25 Years of experience in teaching.

N. Satish Kumar, Research Scholar CSE dept. R. V. College of Engineering, Bangalore. He has received Master Degree (M.Tech) from VTU (R.V.C.E). His research areas are Digital Image processing, parallel programming.

Ramakanth Kumar P, HOD, ISE dept. R. V. College of Engineering, Bangalore. He has received PhD from Mangalore University. His research areas are Digital Image processing, Data mining, Pattern matching, Natural Language Processing.