# DESIGN & DEVELOPMENT OF AUTONOMOUS SYSTEM TO BUILD 3D MODEL FOR UNDERWATER OBJECTS USING STEREO VISION TECHNIQUE

N. Satish Kumar[1], B L Mukundappa[2], Ramakanth Kumar P[1]
[1] Dept. of Information Science, R V College of Engineering, Bangalore, India
[2]Associate Prof., Dept. of Computer Science, University College of Science, Tumkur, India

### *ABSTRACT*

*The objective of the paper was to design and development of a stereo vision system to build 3D model for underwater objects. The developed algorithm first enhance the underwater image quality then construct 3D model by using iterative closest point (ICP) algorithm. From the enhanced images feature points are extracted and feature based matching was done between these pair of images. Epipolar geometry was constructed to remove the outliers between matched points and to recover the geometrical relation between cameras. Then stereo images were rectified and dense matched. Then 3D point was estimated using linear triangulation. After the registration of multi-view range images, a 3D model was constructed using a Linear Triangulation technique.*

***KEYWORDS:*** *Underwater image, ICP algorithm, 3D model*

## I. INTRODUCTION

Generating a complete 3D model of an object has been a topic of much interest in recent computer vision and computer graphics research. Many computer vision techniques have been investigated to generate complete 3D models. Underwater 3D imagery generation is still challenge due to their many unconventional parameters such as refractive index of water, light illumination, uneven background, etc. Presently there are two major approaches. First is based on merging multi-view range images into a 3D model [1-2]. The second approach is based on processing photographic images using a volumetric reconstruction technique, such as voxel coloring and shape-from-silhouettes [3]. Multi-view 3D modeling has been done by many active or passive ranging techniques. Laser range imaging and structured light techniques are the most common active techniques. These techniques project special light patterns onto the surface of a real object to measure the depth to the surface by a simple triangulation technique [4, 7]. Even though active methods are fast and accurate, they are more expensive. However, relatively less research has been done using passive techniques, such as stereo image analysis. This is mainly due to the inherent problems (e.g., mismatching and occlusion) of stereo matching. The quality of underwater images are poor as they suffered from strong attenuation and scattering of light. To overcome these problems this paper has been contributed to first enhance underwater images and applying passive method to build 3-D Model.

The work first employs image enhancing technique to reduce the effect of scattering of light and attenuation and also improves the contrast of the images. In order to remove the mismatches between pair of stereo images, this methodology computes epipolar geometry and also performs dense matching to get more features by employing rectification process. Multi-view range images are obtained using stereo cameras and turntable. The developed computer vision system has two inexpensive still cameras to capture stereo images of an object. The cameras are calibrated by a projective calibration technique. Multi-view range images are obtained by changing the viewing direction to the object. We also employ a turntable stage to rotate the object and to obtain multiple range images. Multiple range images are then registered and integrated into a single 3D model. In order to register range images automatically, we employ Iterative Closest Point (ICP) algorithm to integrate multiple range images into a single mesh model using volumetric integration technique.

Error analysis on real objects shows the accuracy of our 3D model reconstruction. Section 2 presents the problem and solution of the images in underwater conditions. Section 3 presents range image aquition methodology, Section 4 presents a 3D modeling technique of merging multi-view range images. Finally, section 5 concludes the paper.

## II. PROBLEMS IN UNDERWATER & SOLUTION

To capture the images in underwater conditions, two underwater cameras with lights enabled were mounted on a stand. Underwater imaging faces a major problem of light attenuation which limits the visibility distance and degrades the quality of the images such as blurring or lacking of structure in the regions of interest. The developed method uses efficient image enhancement algorithm. We implemented the program with Matlab. The method is comprised of three main steps:

- **Homomorphic filtering:** The homomorphic filter simultaneously increases the contrast and normalizes the brightness across the image.
- **Contrast limited adaptive histogram equalization (CLAHE):** The histogram equalization is used to enhance the contrast of the image.
- **Adaptive Noise-Removal Filtering:** A Weiner filter is implemented to remove the noise produced by equalization step.

## III. RANGE IMAGE ACQUITION AND CALIBRATION

We employ a projective camera model to calibrate our stereo camera (MINI MC-1). Calibration of the projective camera model can be considered as an estimation of a projective transformation matrix from the world coordinate system (WCS) to the camera's coordinate system (CCS). We employ a turntable to take range images while stereo cameras are stable. We set up an aquarium to take images in underwater condition. We mount MC-mini underwater cameras on a stable stand and keep an Model on a turntable. The lab setup to do experiment is as shown in fig. 2. Our system makes use of camera calibration, so we employ Tzai stereo camera calibration model to calibrate our stereo cameras. We use 8X9 check board (as shown in Fig. 1) to calibrate the cameras.

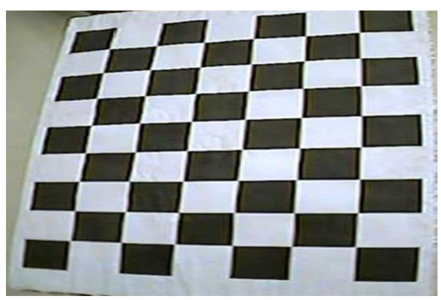

Fig.1 Check board for camera calibration

We got internal camera parameters (K1 & K2) as a result of calibration process. These internal camera parameters are useful in estimating metric 3-D reconstruction so that we can get the approximate dimensions of the object.

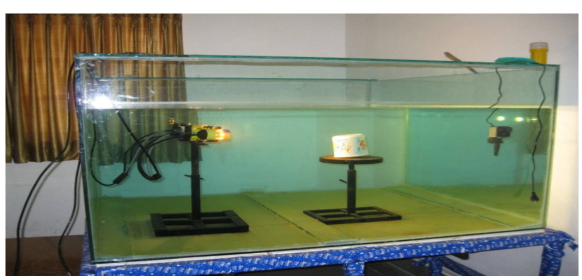

Fig. 2 Lab setup for the experimentation

# IV. 3-D MODELING METHODOLOGY

This section describes the complete overview of the developed system flow chart. The methodology is as shown in the fig. 3.

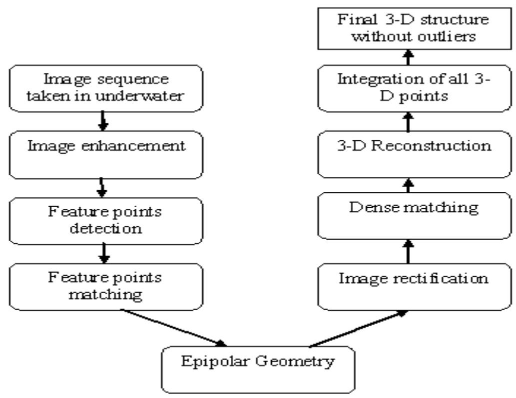


Fig.3: Developed 3-D Modeling methodology

## 4.1 Extraction of 2D feature points and Correspondence matching

The work is with large scaled underwater scenes, where illumination is frequently changing and to find a set of stable features which are useful for later stage of estimating 3-D points. Therefore, a feature based approach, namely Scale Invariant Feature Transform (SIFT) which is developed in OpenCV library is used in this work. Images are represented by a set of SIFT feature as shown in Fig 3. Although, there are some newly derived techniques that can return faster or more efficient result, the developed method chooses the SIFT, because of its invariance to image translation, scaling, partial invariance to illumination changes. Key Points after getting from SIFT are then compared between every consecutive pair of images and the matching points are used to calculate the epipolar geometry between cameras. The epipolar geometry is used to further discard false matches. The feature based approaches look for features in images that are robust under the change of view points, illumination, and occlusion. The features used can be edge element, corners, line segments, gradients, depending on the method.

## 4.2 Computation of Epipolar geometry

The epipolar geometry provides us with a constraint to reduce the complexity of correspondence matching. Instead of searching the whole image or region for a matching element, we only have to search along a line. Even when the matching is already found by other methods, epipolar geometry can be applied to verify the correct matches and remove outliers. The epipolar geometry is used for two purposes:

a) To remove false matches from SIFT matching and
b) To recover the geometrical transformation between 2 cameras from the computation of the fundamental matrix.

## 4.3 Fundamental matrix estimation

To estimate the fundamental matrix (F), Random Sampling Consensus (RANSAC) was used. The library of OpenCV provides functions to estimate Fundamental matrix using Lmeds and RANSAC. To estimate the fundamental matrix, equation (1) can be deduced and rewriting it in the following way

$$Uf = 0 \qquad (1)$$

$$f = \left(F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}\right)^T \qquad (2)$$

$$U = \begin{pmatrix} {}^I X_1 {}^I X'_1 & {}^I X_1 {}^I Y'_1 & {}^I X_1 & {}^I Y_1 {}^I X'_1 & {}^I Y_1 {}^I X'_1 & {}^I Y_1 & {}^I X'_1 & Y^{I\prime}_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ {}^I X_n {}^I X'_n & {}^I X_n {}^I Y'_n & {}^I X_n & {}^I Y_n {}^I X'_n & {}^I Y_n {}^I X'_n & {}^I Y_n & {}^I X'_n & Y^{I\prime}_n & 1 \end{pmatrix} \qquad (3)$$

### 4.4 Rectification & dense matching

Both stereo pairs are rectified - Rectification transforms a stereo image pairs in such a way that epipolar lines became horizontal, using the algorithm presented in [Isgrò, 1999]. This step allows an easier dense matching process. As our developed system has constructed 3-D structure of the object from multiple views, if there is more number of feature points then 3-D structure will be more accurate. So, to get more number of feature points we employ dense matching process. Depth information about the objects present in a rectified pair of images: far objects will have zero disparity and the closest objects will have maximum disparity instead. Figure 4 shows the corresponding matching features removing outliers.

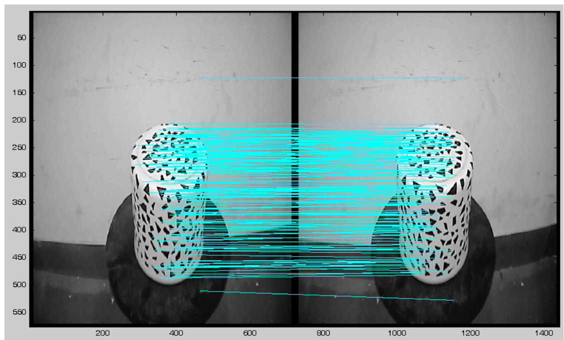

Fig. 4 Corresponding matching without outliers

After the matching image points have been discovered using dense matching, the next step is to compute the correspondent 3D object points. The method of finding the position of a third point knowing the geometry of two other known reference points is called triangulation. Since the two matching points are just the projected images of a 3D object point, then the 3D point is the intersection between two optical rays passing through the two camera centers and two matching image points. The matching points are converted into metric representation using the intrinsic parameters of camera calculated as a result of the camera calibration process. By projecting the points into 3D space and finding intersections of the visual rays, location of object points can be estimated. This process is referred as triangulation. After removing outliers, the final result is a 3D point cloud which can be interpolated to construct the 3D model of the object.

### 4.5 Outliers removal

Once the set of 3D points has been computed, the final step is to remove the isolated points, which are the points with less than 2 neighbors. A point is considered a neighbor of another if it is within a sphere of a given radius centered at that point. This final process is an effective procedure to detect any remaining outliers as outliers generally generate isolated 3D points as shown in Fig. 5.

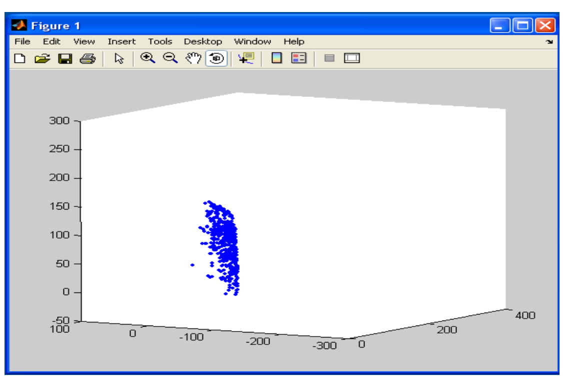


Fig. 5 Partial 3-D reconstruction of two images

We set some threshold to remove the 3-D outlier points. If any point is not having any neighbor within that threshold then that point is considered as outlier and it has to be removed from the 3-D point set. Otherwise the 3-D model is not accurate. The remaining 3D points are stored as a partial reconstruction of the surface. In this manner we calculate the remaining 3D points for the rest of the object views, and using Iterative Closest Point (ICP) algorithm, those 3D points are registered to the common coordinate system. Then those point clouds are interpolated to construct the surface. Once we get the surface of the object then that 3D model can be texture mapped so that the final 3D model of the object looks like actual object.

### 4.6 Integration of all the 3-D points

Using the above methodology, all the partial 3-D structures of the object are obtained and integrated into the common coordinate system using Iterative Closest Point (ICP) algorithm. Thus we got proper 3-D point cloud of all the views. Once we get point cloud then those points were interpolated and surface was put on the point cloud to get the 3-D model of object as shown in Fig. 6.

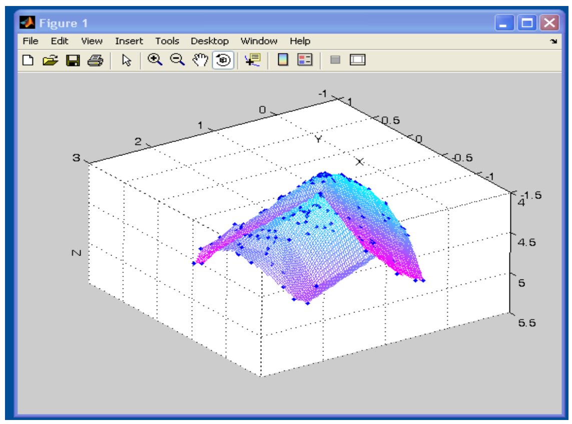


Fig 6 3-D model with surface from 4 views

### 4.7 Texture Mapping

After obtaining the 3-D model of an object, texture of the original object has been mapped onto the 3-D model so that it looks same as the object. The result of the texture mapped 3-D model is as shown in fig. 7

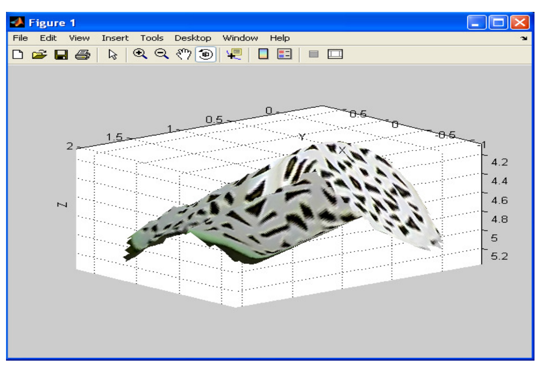


Fig 7 3-D model with texture mapped

## V. Conclusion

The system consists of an inexpensive underwater stereo camera, a turn table and personal computer. Developed autonomous System to build 3-D Model of underwater objects is easy to use and robust under illumination changes as this system extracts SIFT features rather than intensity values of the images.

The images are enhanced and feature points of those images are extracted and matched between pair of stereo images. Final 3D reconstruction is optimized and improved in a post processing stage. Geometrical 3Dconstruction obtained with natural images collected during the experiment out to be very efficient and promising. The estimation of dimensions of the object is also nearly accurate.

## ACKNOWLEDGMENT

I owe my sincere feelings of gratitude to NAVAL Research Board, New Delhi for their supporting, guidance and suggestions which helped us a lot to write the paper.

## REFERENCES

[1] Oscar Pizarro, Ryan Eustice and Hanumant Singh, "Large Area 3D Reconstructions from Underwater Surveys"
[2] Soon-Yong Park, Murali Subbarao, "A multiview 3D modeling system based on stereo vision techniques"
[3] St´ephane Bazeille(1), Isabelle Quidu(1), Luc Jaulin(1), Jean-Phillipe Malkasse, "Automatic Underwater Image Pre-Processing"
[4] Rafael Garcia, Tudor Nicosevici and Xevi Cufí, "On the Way to Solve Lighting Problems in Underwater Imaging"
[5] S. M. Christie and F. Kvasnik, "Contrast enhancement of underwater images with coherent optical image processors"
[6] Kashif Iqbal, Rosalina Abdul Salam, Azam Osman and Abdullah Zawawi Talib, "Underwater Image Enhancement Using an Integrated Colour Model"
[7] Roger Y Tsai, "A Versatile Camera Calibration Techniaue for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses"
[8] Qurban Memony and Sohaib Khanz, "Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks"
[9] Matthew Bryantt, David Wettergreen, Samer Abdallaht, Alexander Zelinsky, "Robust Camera Calibration for an Autonomous Underwater Vehicle"
[10] H. Rob, "Rob Hess - School of EECS @ Oregon State University"; http://web.engr.oregonstate.edu/~hess/index.html.
[11] Y. Bougett, Camera calibration toolbox for Matlab.; http: //www.vision.caltech.edu/bouguetj/calib doc/.
[12] R. Hartley and A. Zisserman, Multiple view geometry in computer vision, Cambridge, UK; New York: Cambridge University Press, 2000.
[13] G. Chou, "Large scale 3d reconstruction: a triangular based approach," 2000.
[14] D.G. Lowe, "Distinctive Image Features from Scale Invariant Feature Points (SIFT), University of British Columbia" 2004

### Authors

**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.

**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.

**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.