

A Novel Approach in Extraction of 3D Point using Scale Invariant Feature Transform for Underwater Objects

Mukundappa B L¹, Dr. Krishna M²

¹Research Scholar, Department of Computer Science, Tumkur University Tumkur

²Centre for manufacturing Research & Technology Utilization, R V Engineering College Campus, Bangalore

Abstract: *The objective of the present paper is to design “an autonomous system to build 3D model of underwater objects” The design of the present system is one of the most critical factors which affect the quality of the software and has a major impact on the later phases, in particular testing as well as maintenance. Further, system design aims to identify the modules that should be in the system, the specification of these modules and interaction with each other to produce the expected results. At the end of the system design all the major data structures, file formats, output formats as well as major modules in the system and their specifications decide the image enhancement.*

Keywords: Underwater object, 2D model, Image enhancement, 3D Reconstruction, Epipolar geometry and homomorphic filtering

1. Introduction

In recent years generating a complete three dimensional (3D) model of an object has been a topic of interest. In computer vision and computer graphics, many techniques have been investigated till date to construct complete 3D models. Further, in the field of deep sea study, the imagery is to generate the 3D reconstruction of underwater natural scenes. and thus complete sample analyzes by a visual observation that enables the quantitative measurements [1].

Reconstruction of individual objects under controlled conditions has been researched over the last few decades [3]. A stereo vision system has been introduced to automatically generate 3D models of real objects [4]. 3D model generated is based on the merging of multi view range images obtained from a digital stereo camera [3]. Stereo images obtained from stereo cameras are rectified, and a feature-based stereo matching technique [4,5], reconstructs range images from them. A turn table is also employed to obtain multiple images into a common coordinate system automatically. The system calibrates a turntable coordinate system with respect to a camera coordinate system. After the registration of multi view range images, a 3D model was reconstructed using a volumetric integration technique. Building of 3D model in underwater objects has been used extensively to study marine objects. The developed system has been divided into six major steps, namely image enhancement, extraction of two dimensional (2D) SIFT feature points and correspondence matching of pair of images, estimation of epipolar geometry to calculate fundamental matrix to remove outliers, estimation of 3D point using camera parameters [6,7], registration of multiple views of an image, and construction of surface on point clouds removing the isolated points. As images are taken in under water conditions, images needs to be enhanced using image processing tool box of MATLAB. 2D SIFT [8] feature points were extracted from left and right images. Then extracted key points were matched between two views, called inliers. Epipolar geometry [9] has been constructed to

remove the outliers (false matches). Then fundamental matrix was recalculated. Using camera parameter which is the result of camera calibration process, estimate the 3D points of two views.

1.1 High level design

The purpose of the design is to obtain the solution of a problem specified by the requirements document. In this phase starting with what is needed, design takes us to work how to satisfy the needs. The design of the system is perhaps the most critical factor affecting the quality of the software and has a major impact on the phases testing and maintenance. System design identifies the modules in the system, the specifications of these modules and to interact with each other to produce the desired results. At the end of the system design all the major data structures, file formats, output formats as well as major modules in the system and their specifications are decided [3,4].

1.2 Design Consideration

In this section, the system development environment and related constraints are studied. This study helps in the selection of appropriate development method and the architecture.

1.2.1 Assumptions and dependencies

The GUI (Graphical User Interface) tool box of MATLAB is used to build the GUI interface of the software. Image processing tool box is used to extract and match the feature points of pair of images. Development of this software uses some standard library functions of C programming language.

1.2.2 General constraints

The system is configured to run on any Windows machine with MATLAB installed in it. A RAM with minimum of 1 GB memory is required in operational state. The images may be in the JPG or BMP format.

Volume 7 Issue 7, July 2018

www.ijsr.net

[Licensed Under Creative Commons Attribution CC BY](https://creativecommons.org/licenses/by/4.0/)

1.3 Development method

The image reconstruction process is break down into the number of modules as separate functions. The following are the descriptions of the separate functions as doing specific tasks.

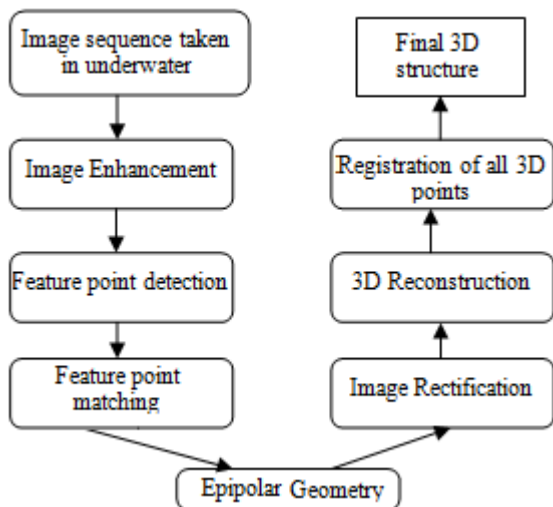


Figure 1: Methodology for the development and reconstruction of underwater objects

Figure 1 show the various steps involved in 3D reconstruction of underwater objects.

1.3.1 Image enhancement

As the images taken in underwater condition by stereo cameras have lot of noise due the light scattering. Those images quality needs to be improved by filtering and the overall contrast of the images need to be improved. To capture the images in underwater conditions two underwater cameras with lights enabled mounted on a stand. Underwater imaging faces a major problem of light attenuation which limits the visibility distance and degrades the quality of images such as blurring or lacking of structure in the regions of interest. The developed method uses efficient image enhancement algorithm comprised of three 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.

1.3.2 Extraction of 2D feature points and its correspondence matching

In 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 3D points. A feature based approach, namely Scale Invariant Feature Transform (SIFT) developed in C language is used. Here, the images are represented by a set of SIFT features. Although, there are some newly derived techniques that can return faster results, the developed method chooses the SIFT, because of its invariance to image translation, scaling, partially invariance

to illumination changes. 2D SIFT descriptors of feature points are computed from enhanced images. Extraction of key points of the object is demonstrated and shown in figure 2.

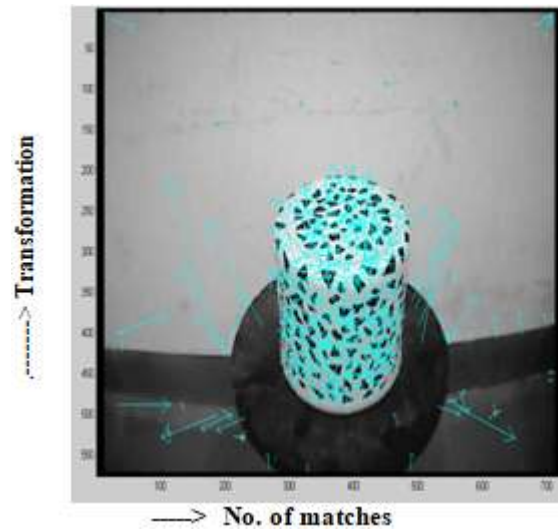


Figure 2: Extraction of 2D key points using SIFT

1.3.3 Matching between pair of images

Key points from SIFT are compared between every consecutive pair of images and the matching points are used to estimate 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 corners, line segments, gradients, depending on the method. The results of this step may yield some false matches also; hence to get only correct matches Epipolar geometry is calculated.

1.3.4 Epipolar geometry

The Epipolar geometry provides the constraint to reduce the complexity of correspondence matching. Instead of searching the whole image for a matching element, we 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 developed system uses the Epipolar geometry for two purposes: To remove false matches from SIFT and to recover the geometrical transformation between two cameras from the computation of the fundamental matrix.

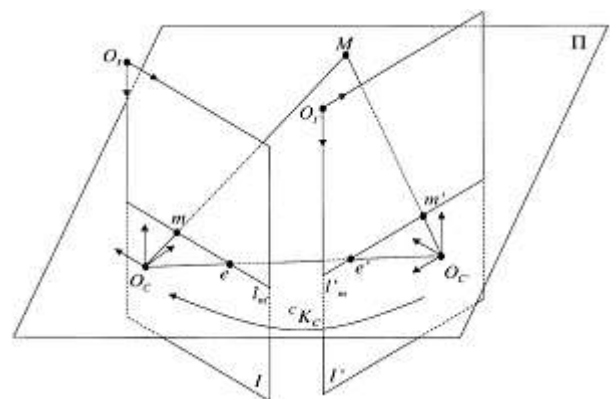


Figure 3: The geometric relation between two pinhole cameras

1.3.5 Fundamental matrix estimation

To estimate the fundamental matrix (F), there are many methods. The library of Open CV provides functions to estimate Fundamental matrix using Lmeds (Least Median Squares) and RANSAC (Random Sampling Consensus). To estimate the fundamental matrix, equation (4) can be deduced and rewriting it in the following way

$$Uf = 0 \quad \dots\dots\dots (1)$$

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

$$U = \begin{pmatrix} x_1'x_1 & x_1'y_1 & x_1' & y_1'x_1 & y_1'y_1 & y_1' & x_1' & y_1' & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_2'x_2 & x_2'y_2 & x_2' & y_2'x_2 & y_2'y_2 & y_2' & x_2' & y_2' & 1 \end{pmatrix} \quad (3)$$

The project uses RANSAC since both techniques are quite similar. RANSAC algorithm calculates for each F the number of inliers. Once the outliers are removed, F is recalculated to find better result. One advantage of this algorithm is its ability to find robust estimation of the model parameters leading to accuracy, despite the presence of the outliers. The RANSAC algorithm uses 7 point method to calculate the fundamental matrix. Once the fundamental matrix is estimated, for every feature points in one image we can calculate its corresponding Epipolar line. The matching points lies on the corresponding Epipolar line, hence helps the detection and removal of false matches.

1.3.6 Metric reconstruction

After the matching image points have been discovered, the next step is to compute the correspondent 3D object points. The method of finding the position of third point by knowing the geometry of two other 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 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. After removing outliers, the final result is a 3D point cloud which can be interpolated to construct the 3D model of the object.

1.3.6.1 Camera calibration

In order to perform triangulation, the positions of two cameras and the geometric relation between them must be known. Hence, before addressing the problem of estimating 3D points using triangulation method, we need to start with the camera calibration process. The stereo camera system acquires several images of the calibration pattern, which is a 9 x 7 checkerboard pattern (figure 4 from different positions and orientations. After that, the calibration toolbox is implemented using MATLAB uses the images to estimate the intrinsic matrices (K1 left camera and K2 – right camera - that determine the optics and internal geometry of both cameras). Since the cameras are static, only the object is rotated using turntable.

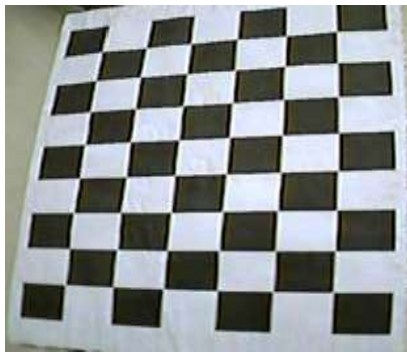


Figure 4: Check board for camera calibration

Hence, the calibration is done only once and set offline. Moreover, the extrinsic matrix [R t] can also be computed from the fundamental matrix F that we have estimated in the previous step. Therefore, in the case where the transformation between cameras is uncertain, we can use the following method to calculate Rotation matrix R and translation vector t.

1.3.6.2 Estimation of [R, t] from fundamental matrix

The system obtained the matrix F from the matching points and two intrinsic matrices K1 and K2 from the calibration process.

$$F = K^T [R \ t] K^{-T} \dots\dots\dots (4)$$

Essential matrix E is given by, E=[R t] then

$$E = K^T F K^{-T} \dots\dots\dots (5)$$

Using method developed by Hartley and Zisserman, calculate the Single Value Decomposition (SVD) of E as follows.

$$[U \ S \ V]^T E \dots\dots\dots (6)$$

Then the Rotation matrix R and translation t vector are given by.

$$R = [U \ W \ V] \ T \text{ or } R = [U \ W^T \ V^T] \dots\dots\dots (7)$$

If the two cameras are identical then K1=K2=K. from Equation (7)

$$t = U_3 \text{ or } t = -U_3 \dots\dots\dots (8)$$

This method returns two values of t and two values of R hence there are four solutions, but in the end only one correct pair of [R, t] describes the correct transformation between the two cameras. These solutions create different geometrical interpretations between two cameras as shown in equation 8. Among the 4 geometries, a unique one shown in figure 5(a) describes the correct geometry because the object point is in front of both cameras. Therefore, to determine correct geometry, the following steps are performed to select a pair of test image points.

- Back project them to 3D location.
- Calculate the depth of 3D points to two cameras.

Choose the geometry with both depths is positive.

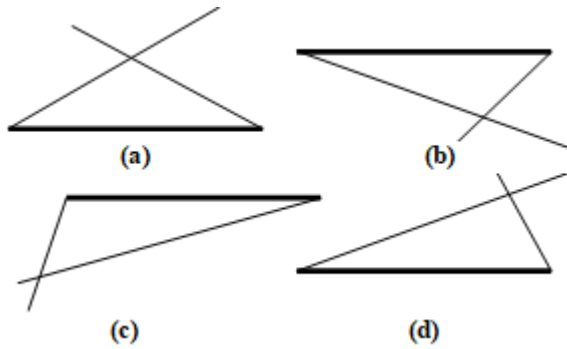


Figure 5: Geometrical interpretation of the 3D point estimation

1.3.7 3D 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 two 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. 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.

1.4 Data flow diagrams

Data flow diagram of “Autonomous System to build 3D model of underwater objects” has been spitted into three levels level 0, level 1 and level 2.

1.4.1 DFD level 0

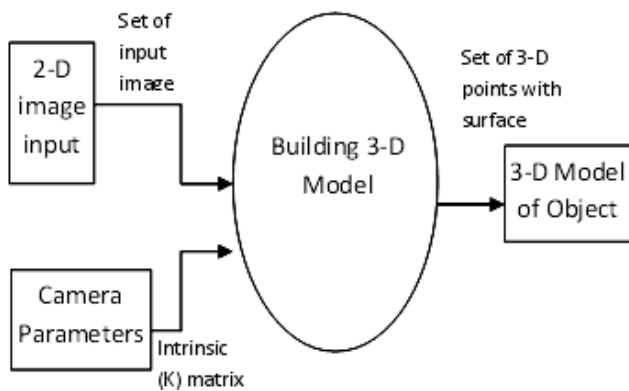


Figure 6: Level 0 DFD

The level 0 DFD describes the main flow of the system. Images taken at all the angles of the object will be given as input to the main system (building 3D model) along with camera parameters. The 3D point cloud is plotted on a graph and then the points are interpolated to have a surface and then texture mapped to get the actual appearance of 3D object.

1.4.2 DFD level 1

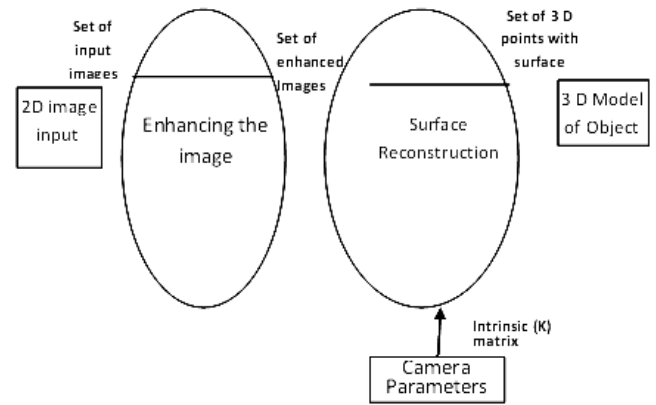


Figure 7: Level 1 DFD

The level 1 DFD includes both enhancing and surface reconstruction process. These two processes are explained in the next DFD level.

1.4.3 DFD level 2

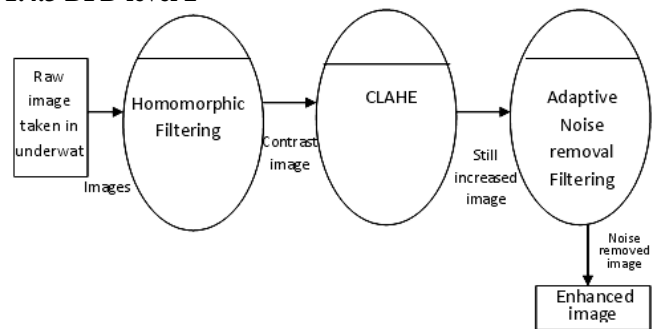


Figure 8: Level 0 DFD

The image enhancement process is made as another module for level 2 DFD. The raw image is passed to the homomorphic filter, and then histogram equalization is used to enhance the contrast of the image. A Weiner filter is implemented to remove the noise produced by equalization step.

1.4.3.1 DFD level 2

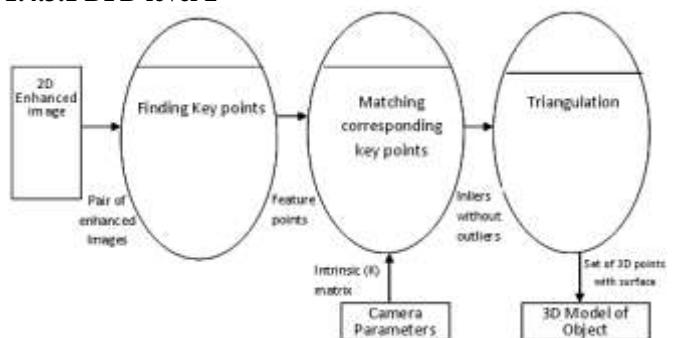


Figure 9: Level 2 DFD

The 2D descriptor finding process is made one of the modules for level 2 DFD. The next step is matching corresponding key points of pair of images thus enables us to find inliers with removing the outliers and also fundamental matrix. With camera parameters and inliers we are constructing the triangulation.

2. Conclusion

In the present work, an underwater stereo camera, along with a turn table and personal computer has been used for the enhancement of objects. This developed autonomous system to build 3D model of underwater objects is robust and easy to use. 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 3D constructions obtained with natural images collected during the experiment are very efficient and promising. The estimation of dimensions of the object is also nearly accurate.

Dr. M. Krishna, Professor and Director – R&D, Centre for Manufacturing Research & Technology, Utilization, R V College of Engineering, Bangalore.

References

- [1] Oscar Pizarro, et. al, "Large Area 3D Reconstructions from Underwater Surveys", IEEE, 2004, pp: 678 – 687.
- [2] Soon-Yong Park and Murali Subbarao (2003) "A Fast Point-to-Tangent Plane Technique for Multi-view Registration", Proceedings of the Fourth International Conference on 3D Digital Imaging and Modeling", IEEE, 2003, pp: 968-981.
- [3] Qurban Memont and Sohaib Khan, "Camera calibration and 3D world reconstruction of stereo-vision using neural networks", International Journal of Systems Science, 2001, volume 32, pp: 1155 – 1159.
- [4] St'ephane Bazeille, Isabelle Quidu, Luc Jaulin, Jean-Phillipe Malkasse "Automatic Underwater Image Pre-Processing", CMM'06 Characterisation du milieu marin , Oct. 2006, pp: 16-19.
- [5] S.M. Christie and F. Kvasnik, "Contrast enhancement of underwater images with coherent optical image processors," APPLIED OPTICS, vol. 35, 1996, pp. 817-825.
- [6] K. Ibal, I. Salam, and A. Osman, "Underwater Image Enhancement Using an Integrated Colour Model," IAENG International Journal of Computer Science, vol-34, 2007. pp:45-48
- [7] R Tsai., 1987, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses". IEEE Journal on Robotics and Automation, pp: 323- 344.
- [8] Qurban Memont 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, pp: 1155 – 1159.
- [9] Matthew Bryant, et.al, "Robust Camera Calibration for an Autonomous Underwater Vehicle", Carnegie Mellon University. <http://www.syseng.anu.edu.au/rsl/>. Carnegie mellon university, December 2000 pp: 111-116.

Author Profile

Sri B.L Mukundappa, Associate Professor, Dept of Computer Science, University College of Science, Tumkur university, Tumkur.