Analysis of Three Dimensional Object Reconstruction Method for Underwater Images

Rashmi Gokhale, Dr. Ramakanth Kumar P

Abstract: The aim of this paper is to design and develop a three dimensional reconstruction system using a feature detection technique called as speeded up robust features (SURF). First the feature points are detected using SURF and matched between pair of underwater stereo images. Further epipolar geometry is utilized to remove the unfavourable feature points and to derive geometrical relation. Linear triangulation is used to compute 3d points, which are further dense matched and mapped with the appropriate texture to resemble the original object. The proposed design is tested on a set of underwater images, using the information about the camera calibration parameters. The result of testing shows that the designed method can be utilized to reconstruct a partial 3D approximation of the underwater objects.

Index Terms: 3D reconstruction, Speeded up Robust Features, SURF, epipolar geometry, underwater stereo images, linear triangulation, camera calibration, feature point detection

————————————◆————————————

1 INTRODUCTION

In computer vision, three dimensional object reconstruction is the procedure to recreate the appearance and shape of the objects to create the effect of original object. This procedure of reconstructing three dimensional view of objects in general and for objects living and occurring below the surface of the water in specific, is gaining more and more importance because of its use in computer vision and image processing. The above stated technique is an essential contributor to application areas such as image analysis of medical objects, robot bin picking [4], simulating man-made objects to use in manipulated real time environments. Video surveillance, study of objects under sea, video simultaneous and localisation and mapping, are other important applications which can be added to the above list. The general consensus about underwater environment is that it throws up difficult situations. These under-water objects seem to be unclear and distorted due to the density differences of water and air medium, which causes light to refract (bend). The structure less environment is one more inbuilt property of underwater environment [1]. Methods commonly used for 3d reformation are researched using different view-points. Active method reconstructs the 3d view of an object in a continuous feedback mechanism, using radiometric inputs or mechanical inputs. Examples include imaging using laser range, where an object is scanned and the appearance of an object is computed, which results in a reconstructed 3d object [10]. Structured light technique, in which special light patterns are projected onto the surface of a real object to measure the depth of the surface by triangulation techniques [3].

Passive techniques don't interfere constantly with the reconstruction procedure. For example, the stereo images are segmented use stereo image analysis. There are two main approaches used for stereo image analysis. One approach is based on merging multi view range images into a complete 3D model and the second approach is based on processing photographic images using a volumetric reconstruction technique, such as voxel Colouring and shape-from-silhouettes. This paper mainly focuses on passive 3d modelling approach, using a feature detection technique called as Speeded up Robust Features (SURF). SURF is a feature detector which is known to be robust and is been researched and analysed more and more for object recognition problems and three dimensional reconstruction problems. SURF selects feature points of an image from the salient features of its linear box-space [9]. Section 2describesa three dimensional reconstruction system using SURF, section 3 represents conclusion and future work.

2 3D RECONSTRUCTION SYSTEM USING SURF

The following section describes the detailed 3d image reconstruction steps. The Figure 1 depicts the diagrammatic representation of the 3d reconstruction system, followed by step by step description of each of the blocks in the system.

- Rashmi Gokhale is currently pursuing M.Tech program in Information Technology, from R.V.College of Engineering, under Visweswaraya Technological University Belgaum, India.
  E-mail: bg.rashmi@gmail.com
- Dr. Ramakanth Kumar P is currently professor and Head of the Department, Information Science and Engineering, in R.V.College of Engineering, Bangalore, India. E-mail: ramakanthkp@rvce.edu.in

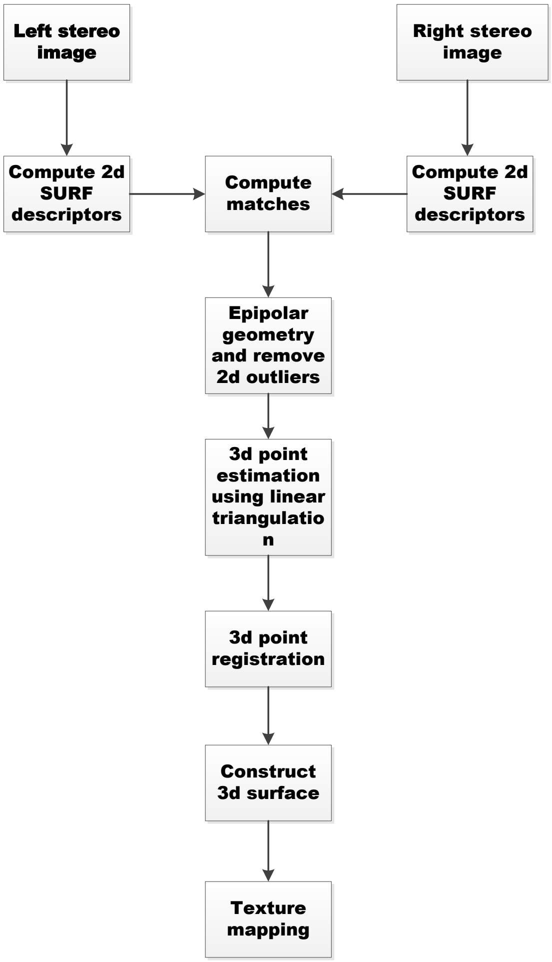

Figure.1 3d Reconstruction system using SURF

Pair of underwater stereo images that is left and right side images is fed as input to the three dimensional reconstruction system.

2.1 Compute 2d SURF descriptors

The first task is to compute the feature descriptors from the stereo images using SURF technique. Local invariant features perform well in pattern recognition problems due to their robust nature, uniqueness and predictability characteristics. Herbert Bay et. al., [6] introduced the local invariant interest points' detector-descriptor Speeded up Robust Features abbreviated as SURF. SURF is invariant to common image transformations, rotation, scale change, illumination change and small change in viewpoint. SURF uses integral images (summed area tables), which are intermediate representations for the image and contain the sum of grey scale pixel values of image, to reduce computation time. The detector is based on Hessian matrix to make use of its good performance in computation time and accuracy.

Given a point x= (x, y) in an image I, H(x, σ) is the Hessian matrix in x at scale σ defined as:

$$H(x, \sigma) = \begin{bmatrix} I_{xx}(x,\sigma) & I_{xy}(x,\sigma) \\ I_{xy}(x,\sigma) & I_{yy}(x,\sigma) \end{bmatrix},$$

Where $I_{xx}(x,\sigma)$, $I_{xy}(x,\sigma)$ and $I_{yy}(x,\sigma)$ represent the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with image I in point x. The descriptor makes use of Haar-wavelet responses within the interest points of vicinity. SURF descriptor first, identifies a reproducible orientation based on information from a circular region around the point of interest. Then, it builds a square region aligned to the selected orientation and extracts its SURF descriptor.

2.2 Compute Matches

The task of finding correspondence between images of the same scene or object is essential in many computer vision application areas. This can be achieved by three main steps, detection, description and matching [6]. In detection step, interest points are selected from distinctive locations in an image such as corners and blobs. These interest points should distinctive and repeatable, that is they should be detectable even under poor viewing conditions. In description step, the neighbourhood of each interest point is represented by a feature vector. The description procedure should be invariant to noise, detection errors and geometric and photometric deformations. The last step which comprises of matching feature vectors of different images. This is usually done based on the distance between features vectors, e.g. Euclidean distance for example. Key Points derived from SURF technique are then compared between every consecutive pair of images and the matching points are used to calculate the epipolar geometry between set of stereo images. The epipolar geometry is used to further discard false matches. Figure 2 depicts the pair of images with the surf descriptors detected and matched.

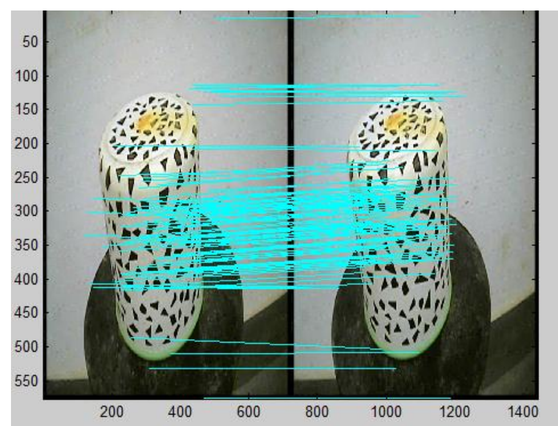

Figure 2Surf descriptor matches with outliers

2.3 Epipolar Geometry and Remove 2d Outliers

The correspondence problem is to find a set of points in one image which can be identified as the same points in the second [7]. Finding matching feature points in computer vision, is a standard problem is complex to solve due the change of viewpoints, illumination, camera settings, image noises. The epipolar geometry helps to reduce the complexity of correspondence matching. Instead of searching the whole image or region for a matching element, there is a need to search along a line. If the matching is already found by the other methods, epipolar geometry can be applied to verify the correct matches and remove outliers. In this proposed work, the epipolar geometry is used to remove false matches and to recover the geometrical transformation between 2 images captured from different viewpoints, by calculating the fundamental matrix. Figure 3, depicts the feature point matching, between pair of images with the outliers removed.

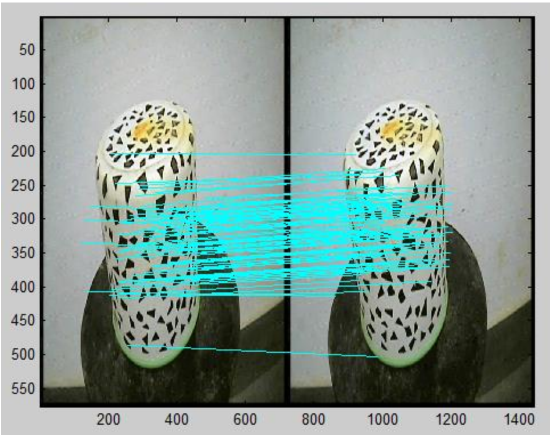

Figure.3 Surf descriptor matches with outliers removed

As the focus of the work is to build 3-D structure of the object from multiple views, if there are more number of feature points then 3-D structure will be more accurate. Therefore, to get more number of feature points dense matching process is employed.

2.4 3d Point Estimation Using Linear Triangulation and Registration

After matching image feature points, the next step is to compute the corresponding projected points in three dimensional view. The method of finding the position of a third point knowing the geometry of two other known reference points is called triangulation [8]. In two dimensions, given two points at known locations A and B and the third point to estimate X . The distance between A and B is d . The viewing angle from A to X is α and from B to X is β. Hence the distance z from C to the base line d can be calculated by the law of sines as per the formula $z = \frac{\sin\alpha \sin\beta}{\sin(\alpha+\beta)}$.

Refer Figure 4 for the diagrammatic representation of the problem. The above concept of 2D triangulation is applied to stereo 3D reconstruction. Knowing the positions of two image points and camera calibration parameters and the transformation between 2 images captured from different viewpoints, the position of the corresponding 3D point is computed.

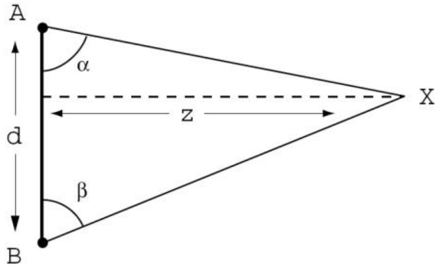

Figure.4 Triangulation Problem

Figure 5, depicts the projected 3d point cloudcomputed using the triangulation. The next step of 3D reconstruction is to register all the 3D points found with every pair of images into one single world coordinate. Since the transformation between very two images are known, 3D points from one coordinate system can be transform into other by multiplying with the corresponding transformation matrix [R, t]. Once the 3D points are formed, the next step is to remove the isolated points, which are the points with less than 2 neighbours. A point is considered a neighbour of another if it is within a sphere of a given radius centered at that point. This process will detect any remaining outliers.

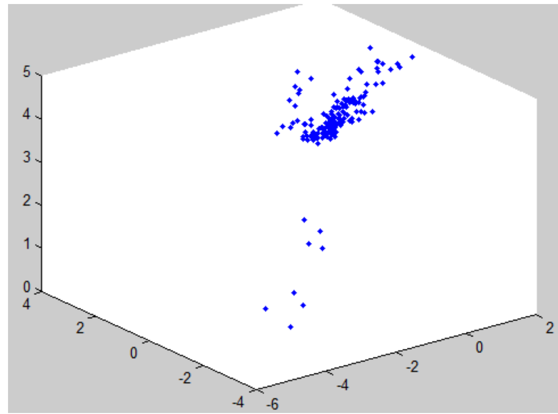

Figure.5 3d Point cloud

2.5 Surface Reconstruction

All the partial 3-D point clouds of the object are obtained and integrated into the common coordinate system and surface reconstructed using Iterative Closest Point (ICP) algorithm. This algorithm is employed to minimize the difference between two clouds of points. It iteratively revises the translation and rotation collectively called as transformation, needed to minimize the distance between the two point clouds. Figure 6 depicts the three dimensional surface reconstructed using the 3d point clouds.

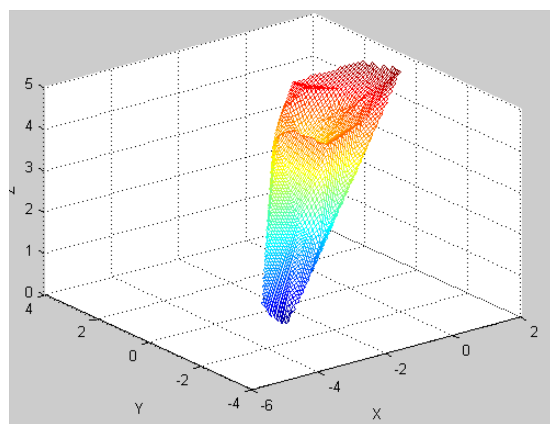

Figure.6 Reconstructed Surface

2.6 Texture Mapping

Once this point cloud is formed the surface is superimposed on the point cloud to get the 3-D model of object. Further the texture of the original object has been mapped onto the 3-D model so that it closely resembles the original object. Figure 7, depicts the reconstructed surface with the relevant texture mapped onto it, to get the original object appearance.

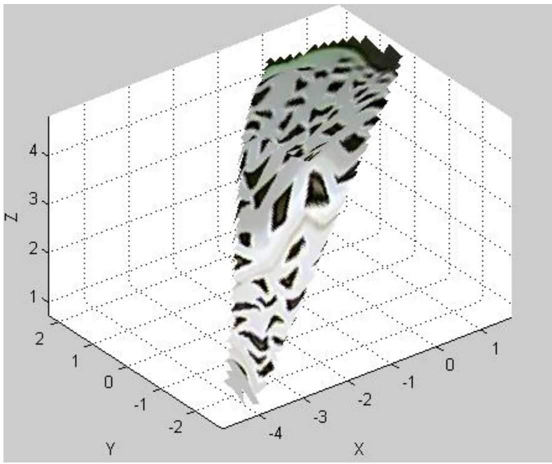

Figure.7 Reconstructed Surface with Texture Mapped

3 CONCLUSION & FUTURE WORK

The designed and developed system is a solution explored in the domain of passive techniques used for three dimensional object reconstruction. The sample result of the 3d reconstructed model, with an input set of 4 stereo images is shown Figure 7. These images are processed with the information about camera calibration parameters. The result of testing shows that the designed method can reconstruct a partial 3D approximation of the underwater objects. Future work can be an enhancement of this system, by processing a set of eight or more stereo pair images, so that a complete or near to complete reconstructed object can be achieved.

REFERENCES

[1]. Murino V, Trucco A, “Three-Dimensional Image Generation and Processing in Underwater Acoustic Vision” Proceedings of the IEEE, vol. 88, Issue: 12, pp. 1903– 1948, 2000.

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

[3]. Rafael Garcia, Tudor Nicosevici and Xevi Cufí, “On the Way to Solve Lighting Problems in Underwater Imaging”

[4]. B L Mukundappa, N. Satish Kumar, Ramakanth Kumar P3 Literature Survey on Building 3-D Models of underwater Objects.

[5]. Y. Petillot, J. Salvi, B. Batlle, “3D large scale seabed reconstruction for UUV Simultaneous localization and mapping,” IFAC Workshop on Navigation, Guidance and Control of Underwater Vehicles, NGCUV'08, Killaloe (Ireland): 2008.

[6]. H. Bay, T. Tuytelaars and L. Van Gool, “SURF: Speeded Up Robust Features”, in: 9th European Conference on Computer Vision, (2006) May 7-13.

[7]. Wikipedia, “Correspondence problem,” 2008

[8]. G. Chou, “Large scale 3d reconstruction: a triangular based approach,” 2000.

[9]. Edouard Oyallon1, Julien Rabin, “An analysis and implementation of the SURF method, and its comparison to SIFT”, published in Image Processing On Line, 12-Feb-2013

[10]. Siddharth Jain, “A survey of Laser Range Finding”, published in CiteSeerX, in 2003.