Chapter 5

3D Surface Reconstruction of Underwater Objects Using Stereo Vision
Chapter 5

3D Surface Reconstruction of Underwater Objects Using Stereo Vision

In the previous chapters, we proposed techniques for underwater image enhancement, color correction, and stereo correspondence required to be performed for reconstruction of 3D surface of underwater objects using stereo vision and experimentally we found an efficient techniques for these steps. In this chapter, we present an approach to reconstruct 3D surface of underwater objects using these novel techniques. We employed MRF-BP for color correction (Chapter 4), and uncalibrated rectification to rectify the images. For the rectified stereo images, disparity map is estimated using Graph Cuts (Chapter 4). The depth map is computed using triangulation concept. Finally, for better visualization, 3D surface mesh is constructed using Virtual Reality Modeling Language (VRML).

5.1 Introduction

As we presented in the Chapter 1, the 3D surface reconstruction from image sequences is a widely studied topic in recent years in the community of Computer Vision. Many methods exist, but all these methods cannot be applied to underwater images. Indeed, they depend on the knowledge of the system and the environment. In underwater scenario, research is directed at exploring the use of vision, potentially in conjunction of other sensors, to automatically control unmanned submersibles, including positioning and navigation. Recent activities combine video imagery taken from multiple views of a scene to derive size and depth measurements and 3D reconstructions. These activities support (semi-) autonomous or operator-supervised missions pertaining to automatic vision guided station keeping, location finding and navigation, survey and mapping, trajectory following and online reconstruction of a composite image, search and inspection of subsea structures. These tasks require an accurate estimation of camera position, together with fast, accurate
correspondence determination, particularly for real-time registration. Common sources of error include non-planar seafloor, moving objects, illumination variations, transect superposition, positioning drift. In the field of 3D reconstruction for out-of-water applications, terrestrial applications have encouraged extensive work over the last three decades; on the other hand, a limited amount of underwater applications have been explored primarily for mapping and positioning.

The extensive research work has been carried out for reconstruction of 3D surface of underwater objects using different methods for acquisition of shape of a scene or an object. Some of the commonly used optical sensing methods are Shape from Stereopsis [147], Shape from Photometric stereo [96], Shape from Motion [72] and Active Stereo [95]. Barndou et al., [17] proposed a method for 3D reconstruction of natural underwater scenes using the stereo vision system IRIS. For acquiring an image of an underwater scene the authors expose two different ways to generate specific trajectories with the stereo vision system which depend on the capabilities offered by the underwater robot equipped with preprogrammed trajectories of the robotic arm. The authors have captured sequence of stereo images of underwater scene and rectified the stereo images using uncalibrated rectification technique proposed by Oram, [105]. The dense disparity map is generated by using the stereo matching algorithm proposed by Roy et al., [115]. Khamene and Negahdaripour incorporate cues from stereo, motion and shading flow for 3D reconstruction in underwater [72, 96]; Majidi and Negahdaripour [88] suggest the use of 3D reconstruction for global alignment of 3D sensor positions; Nicosevici et al., [101] introduce 3D reconstruction from motion video and representation of the surface topography by piecewise planar surfaces for the construction of orthomosaics. Hogue et al., [56] [55] have developed a stereo vision-inertial sensing device deployed to reconstruct complex 3D structures in both the aquatic and terrestrial domains. The sensor temporally combines 3D information, obtained using stereo vision algorithms with a 3 DOF inertial sensor. The resulting point cloud model is then converted to a volumetric representation and a textured polygonal mesh is extracted using the Marching Cubes algorithm [84].

In the non-optical sensing methods acoustic cameras are employed to reconstruct 3D mosaic proposed by Castellani et al., [20]. Josep Forest et al., [36] have adopted laser range gated imaging system for reconstruction of 3D surface of an underwater object. Dalgleish
et al., [27] have proposed a laser line scan method which significantly reduces backscatter in the raw data and enables recovery of the scenes 3D structure by means of triangulation. The similar imaging system was adopted by Carder et al., [19]. Narasimhan et al., [95] have proposed two methods, which are extensions of the method proposed by Dalgleish et al., [27]. First, unlike synchronous scanning systems, scanning is performed without any major moving parts and is instead controlled by a spatial light modulator using a digital light processing (DLP) projector. Second, compensation is made for the attenuation of the water when recovering the object radiance. The attenuation depends on the distance of each object point, where distance is recovered using triangulation.

Considering the underwater environment, the methods such as structured lighting or laser range finders need more equipment and provide no flexibility and they are hard to apply. Ultrasonic or sonar is widely used in underwater researches. These methods perform perfectly in long range distances. But in short range, they do not provide detailed results like cameras do. For that reason, studies are performed to combine the data extracted from these two type sensors, optical (e.g. camera) and acoustic sensors (e.g. sonar). The researchers have investigated methods for multiple-view 3D reconstruction of underwater scenes and objects that combine Dual frequency IDentification SONar (DIDSON) and stereo imagery [98]. The intent is to use the sonar to enhance reconstruction in poor visibility conditions, where visual cues become less informative. DIDSON uses high-frequency sonar (1-2 MHz) to produce range and azimuth 2D measurements that are acquired in a polar coordinate system. Even in turbid water, near optical quality 2D video can be acquired at operational ranges of 10 to 20 meters. Since the geometry of acoustic cameras differs drastically from those of pinhole cameras, the greatest challenge in combining sonar and stereo images is calibrating the system to ensure data model consistency [99, 74]. Not only do the sensors have different areas of coverage, a pixel in polar coordinates maps to a collection of pixels in the cartesian coordinate system, which further complicates searching and matching of feature points in successive images. Other challenges specific to DIDSON include limited resolution, low Signal Noise Ratio (SNR) and limited range of sight. Kim et al., [75] developed an algorithm to enhance sonar video sequences by incorporating knowledge of the target object obtained in previously observed frames. This approach involves inter-frame registration, linearization of image intensity, identification of a target object
and determining the maximum posteriori fusion of images in the video sequence. The Robert Drost et al., [31] have proposed method to reconstruct the underwater object from a sequence of 2D LIDAR (Light Detection and Ranging) images; the Shape from Silhouette method is used to extract the silhouette information from 2D LIDAR image.

In this chapter, we present an approach to reconstruct the 3D surface of underwater objects using stereo vision. The flow diagram of proposed method is shown in Figure 5.1. We performed color-correction of color stereo images using MRF-BP method (Chapter 4), which is shown as efficient method for color correction. The color corrected stereo images are rectified by using uncalibrated stereo image rectification technique proposed by Fusiello et al., [40] to reduce the 2D correspondence search problem to 1D search problem i.e. along a horizontal direction. The disparity map is computed from rectified images by using graph cuts based stereo correspondence method (Chapter 4), which is an energy minimization method. Further, the depth map is obtained from the computed disparity map and it is converted to triangular surface meshes by using Delaunay triangulation technique. Finally, the texture information is mapped onto the mesh for better visualization of 3D surface of an underwater objects.

The organization of the chapter is as follows: section 5.2, presents the rectification of stereo images using uncalibrated rectification method. The stereo correspondence method is presented in section 5.3. In section 5.4, we present triangulation technique for estimation of depth map. The mapping of texture information on 3D surface model is presented in section 5.5. The experimental results are illustrated in section 5.6. Finally, section 5.7 draws a chapter summary.

## 5.2 Uncalibrated Stereo Image Rectification

Given a pair of stereo images, rectification determines a transformation of each image plane such that pairs of conjugate epipolar lines become collinear and parallel to one of the image axes. The important advantage of rectification is that computing stereo correspondences is reduced to a 1D search problem along the horizontal raster lines of the rectified images. There are two ways of rectifying stereo images (i) calibrated stereo image rectification and (ii) uncalibrated stereo image rectification (Chapter 1). In case of calibrated stereo image
Figure 5.1: The flow diagram of proposed method for 3D Surface Reconstruction
rectification, we estimate cameras intrinsic (focal length, aspect ratio, skew and principal point) and extrinsic (rotation and translation) parameters using camera calibration toolbox with the help of calibration object (checkerboard pattern). Whereas, in uncalibrated stereo image rectification, only image features are extracted to estimate a suitable transformation.

Calibrating camera in underwater environment is a trivial task due to signal propagation variability which decreases the checker board image quality. This phenomenon is due to the presence of suspended particles (turbidity) which introduce reflection and refraction effects on the light beam. The refractive index represents the deviation of the light beam passing from one medium to another. This parameter is not fixed, it varies according to the environment characteristics such as water temperature, water salinity, light beam wavelength and camera depth. All these constraints affect the quality of the acquired checker images, further, it leads to lack to information for calculating camera intrinsic parameters. A water refractive index variation changes the value of some internal camera model parameters and therefore the result of rectification may be poor. Therefore, we employed uncalibrated rectification to rectify the stereo images.

In the case of uncalibrated cameras, there are more degrees of freedom in choosing the rectifying transformation and a few competing methods are present in the literature. Each aims at producing a "good" rectification by minimizing a measure of distortion, but none is clearly superior to the others, not to mention the fact that there is no agreement on what the distortion criterion should be. In this chapter, we adopt Quasi-Euclidean epipolar rectification method for uncalibrated images proposed by Andrea Fusiello et al., [40].

Geometrically, in the Euclidean frame, rectification is achieved by a suitable rotation of both image planes. The correspondent image transformation is the collineation induced by the plane at infinity. As a result, the plane at infinity is the locus of zero-disparity in the rectified stereo pair. This is signified by saying that Euclidean rectification is done with respect to the plane at infinity. In the uncalibrated case the reference plane is generic, as any plane can play the role of the infinity plane in the projective space. Our uncalibrated rectification can be seen as referred to a plane that approximates the plane at infinity.

We assume that intrinsic parameters are unknown and that a number of corresponding points \( m_1 \leftrightarrow m_2 \) are available. The method seek the collineations that make the original points satisfy the epipolar geometry of a rectified image pair.
Chapter 5. 3D Surface Reconstruction of Underwater Objects Using Stereo Vision

The fundamental matrix of a rectified pair has a very specific form, namely it is the skew-symmetric matrix associated with the cross-product by the vector \( u_1 = (1, 0, 0) \):

\[
[u_1]_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.
\] (5.1)

Let \( H_r \) and \( H_l \) be the unknown rectifying collineations. When they are applied to the corresponding tie-points \( m^r_i, m^l_i \) respectively, the transformed points must satisfy the epipolar geometry of a rectified pair, namely:

\[
(H_r m^r_i)^T [u_1]_x (H_l m^l_i) = 0.
\] (5.2)

The left-hand side of eq. (5.2) is an algebraic error, i.e., it has no geometrical meaning, so we used instead the Sampson error, that is a first order approximation of the geometric error. The matrix \( F = H_r^T [u_1]_x H_l \) can be considered as the fundamental matrix between the original images, therefore, in our case, the squared Sampson error for the \( j \)th correspondence is defined as:

\[
E_j^2 = \frac{(m^r_i F m^l_i)^2}{(F m^r_i)^2 + (F m^l_i)^2 + (m^r_i F)^2 + (m^l_i F)^2}.
\] (5.3)

where \((\cdot)_i\) is the \(i\)th component of the normalized vector.

As this equation must hold for any \( j \), one obtains a system of non-linear equations \( \{E_j = 0\} \) in the unknown \( H_r \) and \( H_l \). A least-squares solution can obtained with the Levenberg-Marquardt algorithm, but the way in which \( H_r \) and \( H_l \) are parameterized is crucial, and characterizes our approach with respect to the previous ones. We force the rectifying collineations to have the same structure as in the calibrated (Euclidean) case, i.e., to be collineations induced by the plane at infinity, namely

\[
H_r = K_{nr} R_r K_{or}^{-1}, \quad H_l = K_{nl} R_l K_{ol}^{-1}.
\] (5.4)

The old intrinsic parameters \((K_{ol}, K_{or})\) and the rotation matrices \((R_l, R_r)\) are unknown, whereas the new intrinsic parameters \((K_{nl}, K_{nr})\) can be set arbitrarily, provided that vertical
focal length and vertical coordinate of the principal point are same. Indeed, it is easy to
verify that the matrix \(K^T_{nr}[u_1]_x K_{nl}\) is equal (up to scale) to \([u_1]_x\), provided that the second
and third row of \(K_{nr}\) and \(K_{nl}\) are the same. Hence, it is not necessary to include the
matrices \(K_{nr}\) and \(K_{nl}\) in the parameterization.

Each collineation depends in principle on five (intrinsic) plus three (rotation) unknown
parameters. The rotation of one camera along its \(X\)-axis, however, can be eliminated.

Consider the matrix

\[
F = K_{or}^{-T}R_r^T[u_1]_x R_l K_{ol}^{-1}.
\]

(5.5)

Let \(R'_r\) and \(R'_l\) be the same matrices as \(R_r\) and \(R_l\) after pre-multiplying with an
arbitrary (but the same for both) rotation matrix about the \(X\)-axis. It is easy to verify that
\(R_r^T[u_1]_x R_l = R'_r^T[u_1]_x R'_l\). Geometrically, this coincide with rotating a rectified pair around
the baseline, which do not alter the rectification, but, in a real camera, it affects the portion
of the scene that is imaged. Accordingly, we set to zero the rotation around the \(X\)-axis of
the left camera.

We further reduce the number of parameters by making as educated guess on the old
intrinsic parameters: no skew, principal point in the center of the image, aspect ratio equal
to one. The only remaining unknowns are the focal lengths of both cameras. Assuming
that they are identical and equal to \(\alpha\), we get:

\[
K_{or} = K_{ol} = \begin{bmatrix}
\alpha & 0 & w/2 \\
0 & \alpha & h/2 \\
0 & 0 & 1
\end{bmatrix}.
\]

(5.6)

where \(w\) and \(h\) are width and height (in pixel) of the image.

In summary, the two collineations are parameterized by six unknowns: five angles and
the focal length \(\alpha\). Focal length is expected to vary in the interval \([1/3(w + h), 3(w + h)]\),
so we consider instead the variable \(\alpha' = \log_3(\alpha/(w + h))\) which varies in \([-1, 1]\).

The minimization of the cost function is carried out using Levenburg-Marquardt, start-
ing with all the unknown variables set to zero. When \(\alpha'\) converges outside the boundaries of
the interval \([-1, 1]\) a random restart is attempted. If the problem persists the minimization
is carried out with fixed \(\alpha' = 0\).
Finally, the new intrinsic parameters \((K_{nr} \text{ and } K_{nl})\) are set equal to the old ones: \(K_{nr} = K_{nl} = K_{ol}\), modulo a shift of the principal point, that might be necessary to center the rectified images in the customary image coordinate frame. Horizontal translation has no effect on the rectification, whereas vertical translation must be the same for both images.

## 5.3 Stereo Correspondence

The global stereo correspondence approaches transform the matching problem to a minimization of a global energy function. When applied to minimize a global cost function in stereo vision, for each pixel, all possible disparities between minimum and maximum values are considered. Energy optimization method gives excellent results, performing better in textureless areas and near depth discontinuities.

Graph Cuts techniques (detailed description is given in Chapter 4) have been recently used to solve the stereo matching problem involving global constraints. These methods transform the matching problem to a minimization of a global energy function. The minimization is achieved by finding out an optimal cut (of minimum cost) in a special graph. Different methods were proposed to construct the graph. However, when applied to minimize a global cost function in stereo vision, all of them consider for each pixel, all possible disparities between minimum and maximum values.

Greig et al. [48] were first to discover that powerful min-cut/max-flow algorithms from combinatorial optimization can be used to minimize certain important energy functions in vision. The energies addressed by Greig et al. [48] and by most later graph-based methods can be represented as:

\[
E(L) = \sum_{p \in P} D_p(L_p) + \sum_{(p,q) \in N} V_{p,q}(L_p, L_q),
\]

where \(L = \{L_p | p \in P\}\) is a labeling of image \(P\), \(D_p(\cdot)\) is a data penalty function, \(V_{p,q}(\cdot)\) is an interaction potential, and \(N\) is a set of all pairs of neighboring pixels. An example of image labeling is shown in Figure 5.2. Typically, data penalties \(D_p(\cdot)\) indicate individual label-preferences of pixels based on observed intensities and pre-specified likelihood
function. Interaction potentials $V_{p,q}$ encourage spatial coherence by penalizing discontinuities between neighboring pixels. Recently most of the papers show that, graph-based energy minimization methods provide arguably some of the most accurate solutions for the specified applications.

![Image of an image and a labeling](image)

Figure 5.2: An image in (a) is a set of pixels $P$ with observed intensities $I_p$ for each $p \in P$. A labeling $L$ is shown in (b) assigns some label $L_p \in \{0, 1, 2\}$ to each pixel $p \in P$. Such labels can represent depth (in stereo), object index (in segmentation), original intensity (in image restoration), or other pixel properties. Thick lines in (b) show labeling discontinuities between neighboring pixels. (Courtesy Yuri Boykov and Vladimir Kolmogorov)

Ishikawa [64] proves that a global minimum is indeed reachable via graph-cuts in polynomial time if the smoothness term is convex. Although this is of theoretical interest, convex smoothness terms $V_{p,q}$ over penalize large jumps in disparity and therefore those approaches tend to blur depth boundaries. However, even the simplest discontinuity preserving smoothness function, i.e. the Potts model, results in a problem formulation whose optimization is proven to be np-complete [77]. Nevertheless, Boykov et al. [15] show that a strong local optimum, which is guaranteed to lie within a known factor of the real optimum, can be calculated for non-convex smoothness terms by iterative application of their $\alpha$-expansion move. The optimal $\alpha$-expansion move is derived by computing the minimum cut on a weighted graph, which can be accomplished in almost linear time when using specialized minimum cut/maximum flow algorithms [16]. For computing the disparity, we have employed graph cuts with $\alpha$-expansion move for finding the optimum cost.
5.4 Depth Map

Depth determination serves as the most challenging part in the whole process of 3D surface reconstruction, as it calculates the 3D component missing from any given image - depth. The correspondence problem, finding matches between two images so the position of the matched elements can then be triangulated in 3D space is the key issue.

Triangulation refers to the process of determining a point in 3D space given its projections onto two, or more, images. In order to solve this problem it is necessary to know the parameters of the camera projection function from 3D to 2D for the cameras involved, in the simplest case represented by the camera matrices. Triangulation is sometimes also referred to as reconstruction. The triangulation problem is in theory trivial. Since each point in an image corresponds to a line in 3D space, all points on the line are projected to the point in the image. If a pair of corresponding points in two, or more images, can be found it must be the case that they are the projection of a common 3D point \( X \).

Figure 5.3: Representation of triangulation principle

Since, by computing the stereo correspondence between a pair of stereo images we can obtain disparity map. In order to reconstruct the 3D surface, it is necessary to compute the depth map. Thus, the depth map can be computed by triangulation with matched point
pairs and camera parameters. The mathematical equation for estimation of depth map is

\[ Z = \frac{f \times B}{d}, \]  

(5.8)

where, \( Z \) is the depth, \( f \) is focal length of the camera, \( B \) is baseline (distance between optical centers of two cameras) and \( d \) is the disparity.

## 5.5 Texture Mapping

To build a realistic 3D surface reconstruction, the depth map is converted to triangular surface meshes using the Delaunay triangulation algorithm [28]. Texture mapping was one of the first developments towards making images of three-dimensional objects more interesting and apparently more complex. Since, one of the purpose of 3D model reconstruction is to display on a computer monitor, rendering of the reconstructed model represented by only geometric information is not enough. Information about the surface texture has to be added to improve the appearance of the 3D model. By adding the texture information to the wireframe model of a object, a more realistic 3D model can be obtained. Thus, a texture mapping can be applied with relative easy and efficiency to the object. The triangular mesh makes possible to reduce the geometric complexity of a 3D surface representation. The 3D representation is then visualized in a more realistic appearance by providing the wire-frame model with texture mapping. First, a reference image is chosen as the texture map in the stereo image pair, and then, each basic triangle primitive is easily mapped with texture, since the exact position of the reference image and of the 3D structure are known. The depth map with texture mapping gives a 3D model with a good visual impression. The 3D model is stored in Virtual Reality Modeling Language (VRML) format for easy visualization and exchange of information.
5.6 Experimental Results

We have conducted experiments using our underwater database to reconstruct the 3D surface of underwater objects. The experiments are carried out for three pairs of underwater color stereo images (dataset1, dataset2 and dataset3). Since, we are unable to obtain ground truth 3D surface model for our datasets due to some practical limitations, the proposed method cannot be evaluated quantitatively. We evaluated the proposed approach by visual inspection. The proposed reconstruction method consists of following steps: color correction using MRF-BP method, rectification using uncalibrated stereo image rectification technique, graph cuts for estimating disparity map, computation of depth map using triangulation principle, creation of triangular mesh using delaunay triangulation for estimated depth values, and finally mapping of texture information on to surface of the mesh.

5.6.1 Color Correction using MRF-BP

Since, the underwater images shows large radiometric variations, it is necessary to render the color values prior to applying the stereo correspondence methods in order to find accurate point correspondence between underwater stereo image pairs. In the previous chapter (Chapter 4), we concluded that MRF-BP based color correction improves the performance of stereo correspondence method. Figure 5.4 shows the result of MRF-BP based color correction applied on the datasets considered.

5.6.2 Uncalibrated Image Rectification

We rectify the color corrected stereo images by using uncalibrated stereo image rectification technique proposed by Fusiello et al., [40]. The important advantage of rectification in stereo vision is that, it reduces the stereo correspondence search problem from 2D to 1D i.e. along horizontal scanline direction. The uncalibration rectified stereo images are shown in Figure 5.5.
Figure 5.4: Color corrected underwater stereo images using MRF-BP for dataset1, dataset2 and dataset3 respectively

Figure 5.5: first and second column: Underwater rectified stereo images for dataset1, dataset2 and dataset3 respectively
5.6.3 Stereo Correspondence

In the previous chapter (Chapter 4), we experimentally showed that graph cuts based stereo correspondence approach works better at textureless regions and near depth discontinuities and yields smooth disparity map. Therefore, we estimate disparity map by using graph cuts based stereo correspondence approach. Figure 5.6 shows the result of disparity maps obtained using graph cuts for dataset1, dataset2 and dataset3 respectively.

![Figure 5.6: Disparity map computed using graph cuts for dataset1, dataset2 and dataset3 respectively](image)

5.6.4 Depth Map

To reconstruct the 3D surface of objects it is necessary to have depth information of an object of interest. The depth map is constructed from the disparity map using the mathematical eq. (5.8). In order to estimate depth map, it is necessary to have information above focal length \( f \) of the camera and the baseline \( B \) distance (i.e. distance between optical center of two cameras). These parameters are empirically calculated and substituted in the eq. (5.8). The focal length of the camera for dataset1, dataset2 and dataset3 is 798.7885, and the baseline is 10 cm. The Figure 5.7 shows the result of depth maps for dataset1, dataset2 and dataset3 respectively.

5.6.5 Triangulated Mesh

We computed the triangular meshes for the estimated depth map using Delaunay triangulation technique in the MATLAB environment. Figure 5.8 shows the result of triangular surface mesh for dataset1, dataset2 and dataset3 respectively.
Figure 5.7: Textured depth maps of dataset1, dataset2 and dataset3 respectively

Figure 5.8: Triangular mesh created for smooth depth map of dataset1, dataset2 and dataset3 respectively

5.6.6 Texture Mapped Meshes

Finally, for better visualization, the triangular surface mesh is mapped with the texture information extracted from one of the stereo image and converted into VRML format. Figure 5.9 shows the result of texture mapped reconstructed 3D surface for dataset1, dataset2 and dataset3 respectively.

Figure 5.9: 3D surface model with texture for dataset1, dataset2 and dataset3 respectively
5.6.7 Comparative Study of Our Techniques

In this section, we present comparative study of our techniques for reconstructing the 3D surface of underwater objects. The main objective of this comparative study is to find suitable technique for reconstructing 3D surface of underwater objects. We conducted experiments on the stereo images in two scenarios.

5.6.8 Scenario - I

In this scenario, initially, we perform color correction on the color stereo images using MRF-BP method, the color corrected stereo images are rectified using uncalibrated stereo image rectification technique. Further, we compute disparity map by employing adaptive weight-based cross-correlation (AWCC) method. By using the disparity values, depth map is estimated using triangulation technique and the depth map is converted into triangular surface mesh. Finally, the texture information is mapped onto the mesh. Figure 5.11 shows the reconstructed 3D surface model for dataset1, dataset2 and dataset3 respectively.
5.6.9 Scenario - II

In this scenario, we perform color-correction on the color stereo images using MRF-BP based color-correction technique; the color-corrected stereo images are rectified using un-calibrated stereo image rectification technique, then we compute disparity map by employing energy minimization method i.e. graph cuts. Further, the depth map is estimated by using the disparity values based on triangulation principle. The computed depth map is converted into mesh using Delaunay triangulation and finally, the texture information is mapped onto the mesh. Figure 5.12 shows the reconstructed 3D surface model for dataset1, dataset2 and dataset3 respectively.

We evaluated the results using qualitative analysis i.e. by visual inspection of obtained 3D surface model of each scenario. The experimental results conducted for Scenario I and Scenario II shows that the 3D surface model of underwater objects obtained using Scenario II (MRF-BP with graph cuts) yields better 3D surface model compared to Scenario I (MRF-BP with AWCC). This is due to the fact that the MRF-BP corrects the color values of the neighboring pixels of given point accurately in both the images. The color correction done in corresponding points in both images helps the graph cuts method to find
Table 5.1: Comparison of processing time for combination of Scenario I and Scenario II

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Scenario I (minutes)</th>
<th>Scenario II (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dataset1</td>
<td>4.5</td>
<td>3.8</td>
</tr>
<tr>
<td>dataset2</td>
<td>3.7</td>
<td>3.2</td>
</tr>
<tr>
<td>dataset3</td>
<td>3.6</td>
<td>2.9</td>
</tr>
</tbody>
</table>

corresponding points accurately, in addition, graph cuts transform the matching problem to a minimization of a global energy function. When applied to minimize a global cost function, for each pixel, all possible disparities between minimum and maximum values are considered. We observed that the the 3D surface model reconstructed using disparity map obtained based on AWCC method has some discontinuity across neighboring regions. The AWCC converts the color corrected values to adaptive weighted values for estimation of disparity, this leads to loss of color information in depth discontinuity regions. Therefore, the reconstructed 3D surface model has some depth discontinuity. However, the the graph cuts helps in obtaining smooth disparity map in discontinuity regions, because disparity map is estimated directly using color values unlike weights computed in AWCC, which retains color values at depth discontinuity regions. Therefore, the reconstructed 3D surface model has smooth surface.

5.6.10 Processing Time

The comparison of processing time is carried out for Scenario I and Scenario II. The results were obtained on an Intel Pentium Core i5-2500 at the speed 3.30 GHz and 4 GB of RAM. The algorithms were implemented in MATLAB environment. From the Table 5.1, it is observed that the Scenario II (MRF-BP with graph cuts) exhibits very low computational cost compared to Scenario I (MRF-BP with AWCC) and which delivers better results for reconstructing the 3D surface of an underwater objects.
5.7 Chapter Summary

In this chapter, we have present an approach to reconstruct the 3D surface of underwater objects using stereo vision. We have reconstructed 3D surface model for three pairs of color stereo images, which are captured in various turbidity levels. We employed MRF-BP method for color correction. In order to reduce the stereo correspondence problem from 2D to 1D, we performed uncalibrated stereo image rectification on the color corrected images. Further, we computed disparity map on these rectified image using graph cuts based approach. The depth map is computed by using triangulation technique and finally, 3D surface is constructed using VRML. The experimental results shows that the combination MRF-BP with graph cuts yields better 3D surface model with less computational cost compared to the combination of MRF-BP with AWCC. Since, we don’t have ground truth 3D surface model for underwater objects, all the experimental results are compared by visual inspection.