sensors

Article

A Comparative Analysis between Active and Passive Techniques for Underwater 3D Reconstruction of Close-Range Objects

Gianfranco Bianco *, Alessandro Gallo, Fabio Bruno and Maurizio Muzzupappa

Department of Mechanical Engineering, University of Calabria, Via P. Bucci 46/C–Rende, Cosenza 87036, Italy; E-Mails: alessandro.gallo@unical.it (A.G.); f.bruno@unical.it (F.B.); muzzupappa@unical.it (M.M.)

* Author to whom correspondence should be addressed; E-Mail: gianfranco.bianco@unical.it; Tel.: +39-0984-494-068; Fax: +39-0984-494-673.

Abstract: In some application fields, such as underwater archaeology or marine biology, there is the need to collect three-dimensional, close-range data from objects that cannot be removed from their site. In particular, 3D imaging techniques are widely employed for close-range acquisitions in underwater environment. In this work we have compared in water two 3D imaging techniques based on active and passive approaches, respectively, and whole-field acquisition. The comparison is performed under poor visibility conditions, produced in the laboratory by suspending different quantities of clay in a water tank. For a fair comparison, a stereo configuration has been adopted for both the techniques, using the same setup, working distance, calibration, and objects. At the moment, the proposed setup is not suitable for real world applications, but it allowed us to conduct a preliminary analysis on the performances of the two techniques and to understand their capability to acquire 3D points in presence of turbidity. The performances have been evaluated in terms of accuracy and density of the acquired 3D points. Our results can be used as a reference for further comparisons in the analysis of other 3D techniques and algorithms.

1. Introduction

The use of 3D imaging techniques in underwater applications is increasing, in fields such as the survey of submerged artefacts [1], the monitoring of marine flora and fauna [2,3] or the mapping of large areas [4]. Multi-view 3D reconstruction is gaining popularity in underwater photogrammetry [5], as it requires just an off-the-shelf camera (movie or still) to acquire a sequence of overlapped pictures of the scene under ambient or artificial lighting (passive technique). The 3D reconstruction is performed through the identification of common natural features in the image set: therefore, the accuracy of the reconstruction depends on the quality of images and textures. The main problem in underwater imaging is the poor visibility due to scattering effects [6,7], which influences the image quality and limits the application of imaging techniques in close-range measurements. In particular, the active techniques rely on a local acquisition (*i.e.*, based on a sheet or a spot) to minimize the effects of scattering. Therefore, new solutions should be developed in order to improve the 3D acquisition process. Recently, an active stereo technique based on the projection of structured light has been tested experimentally in the laboratory in turbid water [8]. The use of a whole-field structured-light technique allowed for the reconstruction of 3D scenes regardless of the texture, obtaining acceptable results in the close-range digitization of small objects. The main goal of this work is to compare the performance of this active technique with a passive multi-view technique tested in an underwater environment [9]. We want to understand if the use of structured light under water can improve the acquisition with respect to the passive approach in terms of accuracy and density of the acquired 3D points, in order to provide useful information for the design of a new underwater 3D imaging system and to gather a reference dataset to be compared with the results obtained with other algorithms and techniques. The comparison is performed under poor visibility conditions produced in the laboratory by suspending different quantities of clay in a water tank, regardless of the corrections in the optical model of image formation, which is required to deal with refractive or scattering effects. Moreover, we do not take into account other problems that occur in real world applications, like the presence of moving elements (mainly flora and fauna present in the marine environment), the colour alteration, or the instability of the imaging system that should require the synchronization of camera shutters. For a fair comparison, a stereo configuration has been adopted for both the techniques, and we have used the same setup, working distance, calibration and objects.

In the next section, titled Related Works, we provide a brief review of 3D stereo techniques usually employed in underwater environment, then in Active and Passive Stereo Techniques section we describe in detail the 3D stereo reconstruction based on both active and passive approaches. Subsequently, in the Experimentation section, the setup and the experimental tests conducted in laboratory will be described; in the section titled Results Analysis the results are discussed, while comments about the comparison between the two techniques are reported in the Discussion section. Conclusions and Remarks are reported in the last section.

2. Related Works

In this section we provide a concise review of underwater 3D optical techniques, classifying them according to their approach (passive and active) and paying particular attention to stereo systems,

based on the correspondence problem that can be solved by adopting both active and passive approaches. For a more complete review of underwater imaging techniques see also [10,11]. Underwater calibration is treated at the end of this section.

2.1. Passive Techniques

Optical imaging techniques for 3D data acquisition used in underwater environments are based on multiple acquisitions of the scene taken from different viewpoints with a movie or still camera. In these techniques, artificial lights (lamps and spotlights, for example) are used just to illuminate the scene if needed, and are not employed in the triangulation of the 3D points, which, in turn, is based on the knowledge of similar points in the image sequence, found through stereo matching algorithms.

Stereo systems use two digital calibrated cameras to capture the scene; they may be installed on underwater robots in applications like seabed mapping [12], robot localization, and tracking [13]. The use of three synchronized cameras [14] represents a good compromise between accuracy and device dimensions, where the underwater survey has to be conducted by a scuba diver. Movie or still cameras, lodged in water-proof housings, are used in seabed mapping [15] and surveys of archaeological sites [16]. Structure-from-motion techniques are used to reconstruct a scene from a sequence of overlapping images acquired by a single moving camera, for example in coral reef survey [2]. The process is based on the automatic extraction of points of interest (a sparse set of features, such as corners), the tracking of this sparse set of features across the image sequence, and the estimation of their 3D positions using multiple views.

2.2. Active Techniques

Underwater active imaging is seriously compromised by scattering and absorption in the medium, that reduces image contrast and light intensity, discouraging the use of a whole field approach in submarine applications. To reduce these limitations, active imaging systems for underwater 3D acquisition can rely on four hardware solutions: polarization, spatial, time and spectrum discrimination [17]. It is possible to improve the image contrast by using a pair of images taken through a polarizer from different orientations [18] or by reducing the backscattering volume by increasing the separation distance between the light source and the sensor [19]. Sophisticated laser-based methods employ narrow beams to reduce the degrading effect caused by the scattering component that appears in the overlapped volume of the device FOVs [20], while Laser Line Scan methods, based on triangulation, involve the optical scanning of a narrow FOV [21]. The time discrimination approach rejects the noise component with a gated receiver that confines the instantaneous scattering of a small volume of water close to the target [22]. Finally, the light absorption can be reduced by the selection of a green/blue wavelength for the light source.

Underwater 3D techniques can be based on triangulation and time-of-flight. Both of these principles adopt a local approach to scan the entire target of interest, using a sheet light or a narrow beam. The single sheet projection is useful in several applications like underwater navigation [21], pipe inspection [23] and mapping of underwater archaeological sites [24]. Finally, we have to mention the confocal imaging technique, used to improve underwater imaging in the presence of turbidity [25].

2.3. Stereo Matching

Stereo matching algorithms are generally classified based on whether they use local or global methods [26], and both categories are based on some constraints used to solve the correspondence problem: similarity constraint, epipolar constraint, uniqueness constraint, continuity constraint, and ordering constraint. Local algorithms (window based and feature based) enforce these constraints on a limited number of pixels that surround the pixel to be matched. Global algorithms enforce constraints on the whole epipolar line, passing through the pixel or on the entire image. Local algorithms are typically faster than global approaches and more suited for hardware/real-time implementation, but in most cases they are outperformed in terms of accuracy by global approaches.

A different approach involves the projection of a sequence of structured-light patterns in order to codify the scene to be acquired. In this way, the matching of the points is performed automatically, because each point object is uniquely determined by an assigned code [27]. This technique will be described in detail in the next section.

2.4. Underwater Calibration

A calibration procedure is required for the extraction of 3D metric measurements from a set of digital images. Several algorithms and methods for camera calibration, generally based on perspective or projective camera models, have been developed and discussed in the literature. The calibration of an underwater optical device must take into account the effect of refraction at the air-acrylic and acrylic-water interfaces (planer or hemispherical, so-called planar and dome port, respectively), which are present when a camera is mounted in its housing [28–30]. The refraction can be considered through the explicit modelling of the optical paths using ray tracing [13] or with the implicit incorporation of the refraction effect into camera calibration parameters [14]. In the first case, the cameras are calibrated in air, and then calibrated in water to derive the geometry of the refractive interfaces, since the principal component of both refractive effect and image distortion is radial [15]. Although the assumption on the refractive indices for the acrylic dome port and the internal air path is reliable, the changes in temperature, pressure, and salinity of the water can cause small variations in the refractive index of the media that cannot be avoided. The alternative implicit approach incorporates refractive effects of the optical components and refractive interfaces into camera calibration parameters.

3. Active and Passive Stereo Techniques

In this section, we describe the methods and algorithms employed in the two 3D imaging techniques used in the experimentation. In particular, a stereo configuration has been adopted for a fair comparison between the active technique experimented in [8] that uses structured-light and stereo acquisition, and a passive one with the same setup, working distance, calibration, objects, and lighting conditions. In this way, the acquired 3D point clouds are computed with respect to the same coordinate system; this allows us to align them automatically, limiting the errors due to the registration process.

3.1. Optical Model

The pinhole camera model is used to describe the geometry of the optical devices. Considering the perspective projection of an object point w (Figure 1) with coordinates $\mathbf{X_c} = [X_c, Y_c, Z_c]^T$ in the camera coordinate system onto the image plane I with coordinates $\mathbf{x} = [x_u, y_u]^T$ we have:

$$\mathbf{x} = \frac{f}{Z_c}\begin{bmatrix} X_c \\ Y_c \end{bmatrix} \tag{1}$$

where f is the focal length. We can relate $\mathbf{X_c}$ with coordinates $\mathbf{X_w} = [X_w, Y_w, Z_w]^T$ in the world coordinate system through a rigid-body transformation:

$$\mathbf{X_c} = R\mathbf{X_w} + \mathbf{T} \tag{2}$$

where R is the rotation matrix and $\mathbf{T}$ the translation vector. By introducing the homogeneous coordinates for the vectors $\mathbf{x}$, $\mathbf{X_c}$, and $\mathbf{X_w}$ from Equations (1) and (2) in matrix notation, we get:

$$\mathbf{x} = P\mathbf{X_w} \tag{3}$$

where P is the Perspective Projection Matrix (PPM) representing the geometric model of the camera. In a real camera, the lens produces optical distortions with radial and tangential components, which must be taken into account in the camera model. This way the real (distorted) image coordinates $\mathbf{x_d} = [x_d, y_d]$ are related to the ideal (undistorted) image coordinates $\mathbf{x} = [x_u, y_u]$ through the relations:

$$x_d = x_u + k_x(x_u, y_u) \qquad y_d = y_u + k_y(x_u, y_u) \tag{4}$$

where $k_x(x_u, y_u)$ and $k_y(x_u, y_u)$ are the distortion coefficients. Finally, it must be considered that a digital camera involves measures in terms of pixels, through an affine transformation which takes into account the translation of the principal point and a scaling along u and v axis in the image plane:

$$u = s\frac{x_d}{d_x} + u_0 \qquad v = s\frac{y_d}{d_y} + v_0 \tag{5}$$

where s is a scale factor, (u_0, v_0) is the principal point location in pixel coordinates and (d_x, d_y) are the pixel dimension in the u and v direction, respectively.

Figure 1. Pinhole camera model. The image point $\mathbf{x} = (x_u, y_v)$ of a 3D point w is the intersection of the optical ray going through the optical centre O_c and the image plane I at a distance equal to the focal length f.
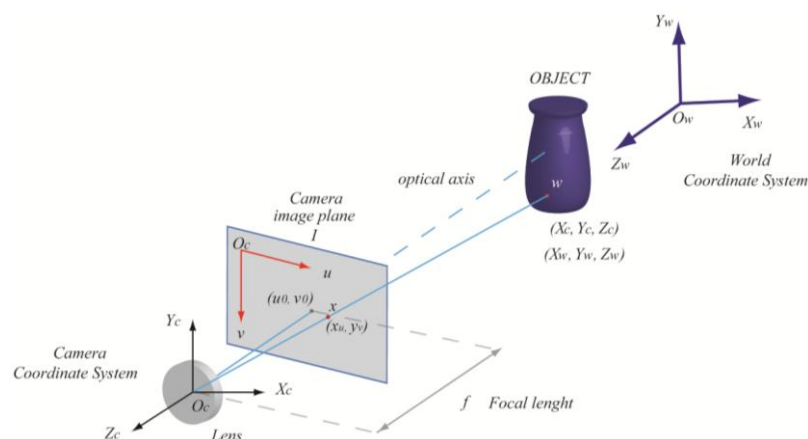
3.2. 3D Stereo Reconstruction

Given two images acquired from slightly different viewpoints, a stereo matching algorithm tries to identify the corresponding points (solving the so-called correspondence problem) in both the images related to the same scene point (Figure 2). Knowing these correspondences and the camera geometry, the 3D world coordinates can be reconstructed by triangulation [31].

Figure 2. Epipolar geometry. The points m_l and m_r, located in left and right images, respectively, are the projection on the image planes I_l and I_r of the same 3D point w of the object. The intersections of the line (O_l, O_r) (baseline) with each image plane are called epipoles (e_l and e_r). The lines l_l and l_r are called epipolar lines, intersection of the epipolar plane p (O_l, O_r, w) with the two image planes.
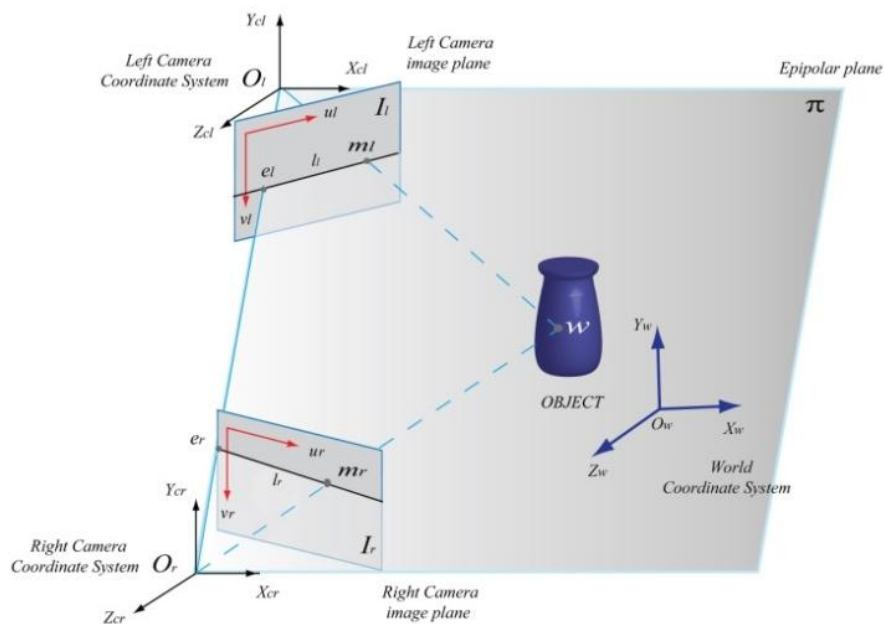
Given a point m_l in the left image, the correspondent point m_r in the right image is constrained to lie on the epipolar line (epipolar constraint). Using the perspective projection theory (see Equation (3)), the coordinates of the correspondent points ml and m_r can be defined through the relations [31]:

$$\begin{cases} m_l = P_l \mathbf{X_w} \\ m_r = P_r \mathbf{X_w} \end{cases} \tag{7}$$

where P_l and P_r are the PPMs of the left and right camera, respectively. Considering that the point m_l (conjugate of m_r) lies on the epipolar line LL, we can write the following relation (the Longuet-Higgins equation):

$$m_l^T F m_r = 0 \tag{8}$$

where F is the Fundamental Matrix, obtained from the stereo calibration process, and depends on the epipolar geometry and the two PPMs.

The last equation solves linearly F using at least eight conjugated points up to a scale factor: with more than eight points, a least-squares solution is found, so the pair of PPMs can be computed from F.

Given these two camera matrices and the pairs of points m_l and m_r that satisfy the epipolar constraint, the 3D coordinates of w are computed by triangulation. Then, finding the intersection of the optical rays corresponding to the two conjugated points m_l and m_r is a problem that can be formulated as an over-determined equation system. If we express the matrix P_l according to its rows and include it in the perspective projection Equation (3), we obtain:

$$m_l = \begin{vmatrix} \boldsymbol{p}_{l_1}^T \\ \boldsymbol{p}_{l_2}^T \\ \boldsymbol{p}_{l_3}^T \end{vmatrix} w \tag{9}$$

and then we write the coordinates of the point m_l in the left image plane I_l as:

$$\begin{cases} u_l = \dfrac{\boldsymbol{p}_{l_1}^T w}{\boldsymbol{p}_{l_3}^T w} \\[2mm] v_l = \dfrac{\boldsymbol{p}_{l_2}^T w}{\boldsymbol{p}_{l_3}^T w} \end{cases} \tag{10}$$

that represent the perspective projection Equation (3) in Cartesian coordinates. By writing the Equation (10) also for the conjugate right image point m_r, we obtain a homogeneous linear system of the type A w = 0 that can be solved by the least-squares method to compute the 3D world coordinates of the point w given the coordinates of a pair of correspondent points m_l and m_r.

To simplify the search of corresponding points, the images are commonly rectified (epipolar rectification), putting the stereo rig in a more convenient configuration in which both epipoles are located at infinity and the epipolar lines form a sheaf of parallel lines in both left and right images. In this case each pair of correspondent points are constrained to lie on the same image row (scan-line) so that the correspondence problem is reduced to a one-dimensional search along each epipolar line.

The disparity, that encodes the depth of the scene, represents the distance between x-coordinates or a pair of correspondent points in left and right (rectified) images. Finding a pair of correspondent points is not so trivial. Since the scene is acquired from different points of view, it is possible to find false correspondences due to occlusion, radiometric distortion and perspective distortion.

3.3. System Calibration

A calibration procedure is needed to compute the intrinsic parameters of each camera (focal length, coordinates of the principal points, radial and tangential distortions, pixel size) and the extrinsic parameters (translation and rotation with respect to a world coordinate system) of camera and stereo system. These parameters are computed using the well-known Camera Calibration Toolbox for Matlab [32]. The intrinsic and extrinsic parameters of each camera are obtained by correlating the coordinates of known points located on a calibration sample (*i.e.*, a checkerboard) with the corresponding coordinates on the image plane. The next step is to compute the extrinsic parameters of the system (stereo calibration), relating each camera frame to a unique world coordinate system. In this way a relationship between the world coordinate system and the left and right camera coordinate systems is established.

3.4. Passive Approach

A Patch-based Multi-View Stereo (PMVS) algorithm, widely used in air applications, has been employed to solve the correspondence problem on a calibrated image pair [33]. Recently, PMVS has also been used for 3D reconstructions of archaeological artifacts in deep water [9]. This open source software, typically employed for multiple-views 3D reconstruction, has been used in a stereo configuration in the laboratory tests, in order to conduct a fair comparison with the active technique. The algorithm estimates the surface orientation while enforcing/strengthening the local photometric consistency, which is important to obtain accurate models from low textured objects or from images affected by blur due to turbidity in underwater environment. PMVS outputs a dense collection of small oriented rectangular patches (a local tangent plane approximation of a surface), and consists of a three-step procedure (matching, expanding, filtering). In the first one, a sparse matching algorithm based on Harris and DoG (Difference-of-Gaussian) operators, detects and matches a collection of reliable point features, sparsely distributed in both images, that satisfy the epipolar constraint. In the expansion step, these initial matches are propagated to the neighboring pixels, in order to obtain a dense collection of patches. Finally, in the last step, false matches are deleted using the visibility constraint. The number of 3D points calculated is strictly related to image resolution, image quality (contrast, focus, *etc*.) and object surface properties, *i.e.*, a low textured object can generate false correspondences and a lesser number of points.

3.5. Active Approach

The active stereo technique used in this work is based on the codification of a black/white pattern set projected on the object through a digital projector. In particular, the grey-code technique generates a more effective matching between correspondent points in stereo pairs [27]. The object is illuminated by a set of n temporally encoded patterns of black/white bands, with a progressively halved width, so that n images are captured by each camera. A binary code (n bit) is assigned to each point of the image, and the values 0 and 1 are associated to intensity levels, *i.e.*, 0 = black and 1 = white (see Figure 3). This procedure allows to codify $2n - 1$ lines, defined as crossing zones between white and black bands. Moreover, coded patterns with a bandwidth of four pixels, shifted in steps of one pixel for a total of four pattern positions, are used to exploit the minimum resolution of the projector. By projecting both horizontally and vertically striped coded patterns, a double code is assigned to the intersection points through horizontal and vertical lines. This procedure allows us to codify automatically each point of the object surface. In this work a set of eight vertical and eight horizontal patterns (8-bit code) is used for grey-code. Other 4 + 4 patterns are projected for vertical and horizontal code shifting, respectively, with a bandwidth of four pixels. A projector resolution of 800×600 points allows to codify 799 lines $\times$ 599 lines = 478,601 points. The projector is only used to establish the correspondences and is not involved in the triangulation, so the calibration of its optics is not necessary. In contrast to traditional passive approaches, this technique does not rely on images with consistent textures, because each point on the object surface is precisely identified by a double binary code.

Figure 3. (**Left**): triangulation of a stereo configuration composed by two cameras and a projector. The points m_L and m_R are the projection of the same 3D point object w on the image planes of the two cameras. (**Right**): examples of binary patterns and code shifting.
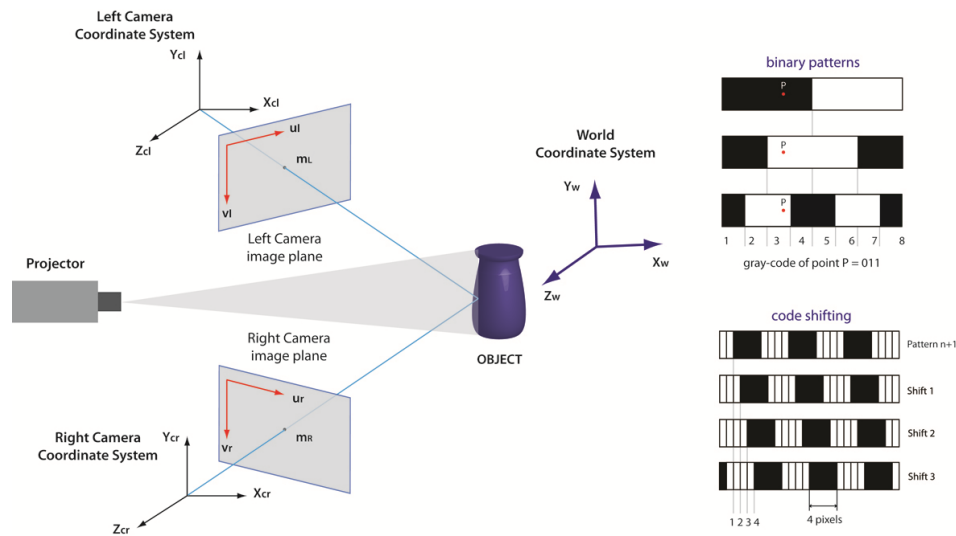
3.6. Calibration

A calibration procedure is needed to compute the intrinsic parameters of each camera (focal length, coordinates of the principal points, radial and tangential distortions, pixel size) and the extrinsic parameters (translation and rotation with respect to a world coordinate system) of camera and stereo system.

The camera calibration procedure is based on an implicit approach. The intrinsic and extrinsic parameters of each camera are obtained by correlating the coordinates of n known points w_i located on a calibration sample with the corresponding coordinates $m_i = [u_i, v_i]$ on the image plane. By writing the Equation (10) in matrix form for this set of n known points:

$$\begin{bmatrix} w_i^T & 0 & -u_i w_i^T \\ 0 & -w_i^T & v_i w_i^T \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \boldsymbol{p}_3 \end{bmatrix} = 0_{2x1} \tag{11}$$

we obtain a set of 2n homogeneous linear equations that can be solved via the least-squares method. In theory, six non-coplanar points are sufficient to compute P, but in practice more points are necessary to compensate measure errors due to image noise, for example. This method is the well-known Direct Linear Transformation (DLT) [34], a quite fast process that minimizes an algebraic error using a simplified camera model, leading to less accurate results.

To obtain a better accuracy, non-linear methods have been developed [35] These approaches minimize a geometrical error, obtaining a non-linear objective function:

$$\varepsilon(P) = \sum_{i=1}^{n} \left(\frac{\boldsymbol{p}_1^T w_i}{\boldsymbol{p}_3^T w_i} - u_i \right)^2 + \left(\frac{\boldsymbol{p}_2^T w_i}{\boldsymbol{p}_3^T w_i} - v_i \right)^2 \tag{12}$$

This minimizing function represents the distance between the image points mi and the projection of the points w_i in the image, the so-called re-projection error. The next step is to compute the extrinsic parameters of the system (stereo calibration), relating each camera frame to a unique world coordinate

system. (see Figure 2). The purpose of the stereo calibration is to find the relationship between the world coordinate system and the left and right camera coordinate systems:

$$X_{cl} = M_l \mathbf{X_w} \tag{13}$$

$$X_{cr} = M_r \mathbf{X_w} \tag{14}$$

where $\mathbf{M_l} = [R_{cl}, t_{cl}]$ is the transformation matrix between the left camera coordinate system and the world coordinate system, $\mathbf{M_r} = [R_{cr}, t_{cr}]$ is the transformation matrix between the right camera coordinate system and the world coordinate system, and $\mathbf{X_{cl}}$, $\mathbf{X_{cr}}$, $\mathbf{X_w}$ represent the coordinates of the object point w in left camera, right camera and world coordinate systems respectively. The transformation matrices $\mathbf{M_l}$ and $\mathbf{M_r}$ represent rotations (R_{cl}, R_{cr}) and translations (t_{cl}, t_{cr}) between the world coordinate system and the left and right camera coordinate systems. The extrinsic parameters can be finally obtained through the same iterative process used to estimate the intrinsic parameters of each camera. It is based on the matching of correspondent points of the calibration sample in the image pair and consists of an overall optimization problem solving, that minimizes the re-projection error between camera planes.

4. Experimentation

Underwater imaging is seriously compromised by the turbidity of the medium, which decreases image contrast and attenuates light intensity, resulting in loss of details and colour alteration. Therefore, in this preliminary analysis, we have decided to conduct experimental tests in the laboratory using a water tank in order to control the lighting conditions and to ensure the repeatability and the reliability of measurements. The laboratory tests aim to analyse the effects of optical distortions, refraction and turbidity in close-range 3D reconstructions of objects for both the proposed techniques. In order to worsen the transparency of water (turbidity) and create a scattering media, different quantities of clay (with particles of about few micrometres in size) have been suspended in the water tank (2×3 m). In this way, along with the increase of the turbidity level, the loss of visibility increases and the ratio of light reflected by the objects decreases progressively. A calibration of a volume of about $400 \times 300 \times 200$ mm has been performed at a distance of 1 m and two ceramic objects have been acquired. The object distance has been fixed by assessing the best compromise between overlapping of FOVs (cameras and projector) and brightness of the projector. The projector has been used as the only light source, so we have measured the illuminance with a luxmeter at the working distance in front of the system, as reported in Table 1, in order to provide a measurement of the lighting conditions at each different turbidity level. We have fixed the maximum value of the clay concentration at 20 mg/L, which required an exposure time of less than 10 s.

Table 1. Measurements of the turbidity conditions.

Turbidity Level	Clay Used (mg/L)	Illuminance (lux)
Clear Water	0	96.7
Turbidity level 1(T1)	5	80.4
Turbidity level 2 (T2)	10	70.5
Turbidity level 3 (T3)	15	44.6
Turbidity level 4 (T4)	20	21.5

4.1. Underwater Setup

Two Nikon D200 cameras, with a Charge Coupled Device (CCD) sensor size of 23.6 × 15.8 mm and a resolution of 3,872 × 2,592 pixels, equipped with an AF-Nikkor 35 mm lens, have been used to acquire the scene. The projector used in our experimentation is a Mitsubishi PK20, characterized by a very small size (123 × 48 × 97 mm), an acceptable resolution of 800 × 600 pixels with a brightness of 25 lumens and a low power consumption that allows for the use of a battery pack as power source. The three optical devices are protected by acrylic waterproof housings and fixed on an aluminium support; the optical setup is depicted in Figure 4. A dome port is mounted on the camera housing to reduce the refractive effects at the air-water interface [36].

Figure 4. Optical setup of the underwater 3D acquisition system: underwater calibration (**Left**) and acquisition (**Right**).
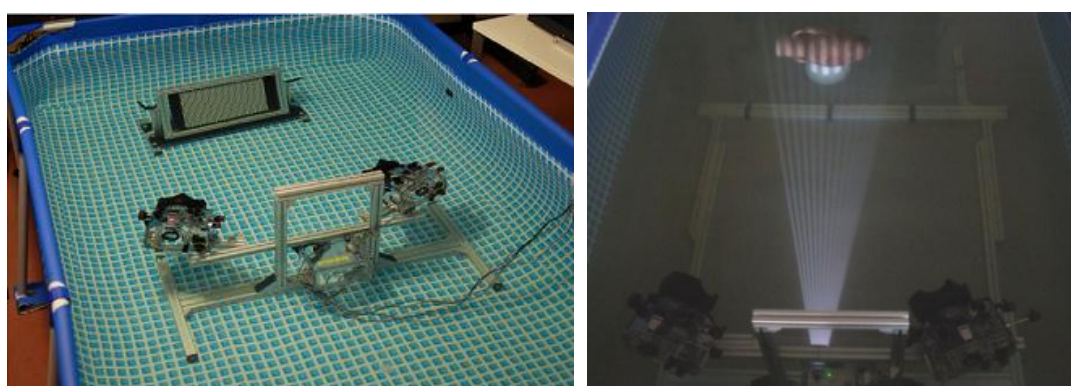

The system dimensions are 950 × 530 × 560 mm, with a weight of about 17 kg in air. The two cameras are mounted on the same bar, in order to have a 620 mm baseline (the distance between the optical centres of the cameras) and a 36 degrees angle between the optical axes. The projector is fixed on an independent bar in a lower position, in order to take into account the differences between the Fields-Of-View (FOVs) of the cameras and the projector (Figure 5). The latter generates an oblique off-axis projection, while the cameras present a symmetric FOV with respect to the optical axis. We chose this setup in order to allow for the acquisition of an object of 30–40 cm at a working distance of 1 m, maintaining a reduced scattering volume due to the overlap of the FOVs of the imaging devices.

The devices should be adjusted (rotation and vertical translation) in order to set up the acquisition at the chosen working distance, allowing the overlapping of the FOVs of cameras and projector as shown in Figure 5. The two cameras are remotely controlled by a PC through a USB interface connection, and the projection of the image sequence is controlled with an appropriate cable in order to optimize the acquisition time, manage the camera setting parameters and analyse the captured images in real time. The optical configuration of the system does not change for the active and passive techniques, therefore we can compare more effectively the results of 3D acquisitions obtained with the same camera positions, optical devices, calibration parameters and working distance. In particular, the same calibration procedure is performed in water for both techniques, computing the intrinsic and extrinsic parameters without changing the configuration of the cameras. The use of the same setup can ensure a

better reconstruction accuracy without generating systematic errors due to multiple calibration procedures. In this way we can conduct a fair comparison between the two techniques.

Figure 5. The optical configuration of the underwater 3D system can be adjusted by rotating and translating the cameras, in order to ensure the overlapping of the FOVs for the required working distance.
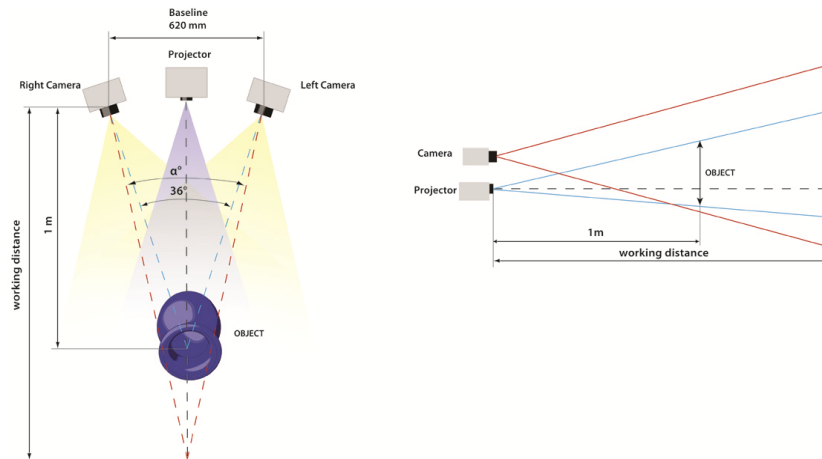
An amphora and a mask (Aeolus) have been acquired in air and water in this experimentation (Figure 6). For the passive stereo technique, only one image pair has been acquired by projecting a white pattern, while the active stereo technique needs at least 24 images to work. The images have been acquired in RAW format to ensure high quality and to extract each colour channel directly from RAW camera data. A 35 mm lens has been used to enclose the whole object in the camera FOV at a working distance of about 1 m. During the acquisitions, the aperture of the cameras was fixed to f/11 for a depth of field of about 100 mm, and the exposure time ranged from 0.8 s for the most luminous patterns, to 10 s for the darkest ones. The passive technique requires instead just 0.8 s for the projection of a white pattern to illuminate the scene. These long acquisition times are due to the use of the projector as the only light source, and to its low brightness.

Figure 6. Images of the two objects with a projected grey-code pattern, acquired at different turbidity levels.
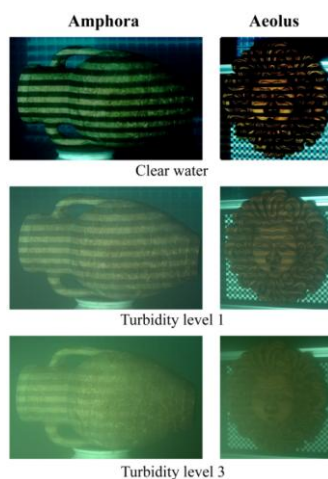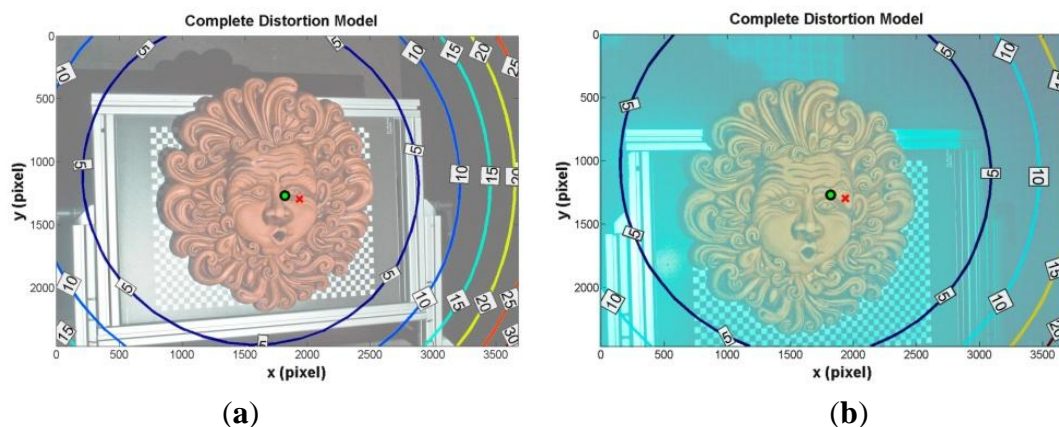
4.2. Calibration

The calibration is performed in air and in water tank, in order to compare the results of intrinsic and extrinsic parameters obtained in both conditions and to evaluate its impact on 3D reconstructions. The calibration is conducted with a planar black and white checkerboard pattern (Figure 4) composed by 32 $\times$ 42 squares with an 8 mm side. This calibration sample has been placed in several poses and acquired by the two cameras in RAW RGB format in order to extract each channel (Red, Green, Blue) and perform a single colour calibration.

The system has been first calibrated using the RAW camera files, extracting de-mosaicked grey scale images. The comparison between calibration parameters in air and water (Table 2) shows that the acrylic interface affects only the principal point coordinates that differ by 1 mm and 0.5 mm in the x and y directions, respectively, while focal length variation is equal to just 0.37 mm. This is due to a non-perfect alignment between the optical axes of camera lens and dome port, and to asymmetric components like the dome port shape. The comparison between the values in air and in water shows a difference on the focal length due to the refractive effect of 4.87 mm, which corresponds to an increase of about 13%. This variation of focal length implies a decrease of the FOV proportional to the medium index, because the CCD size is constant. Finally, the analysis of the principal point coordinates shows negligible variations. The complete distortion model, defined as the sum of radial and tangential components for each camera as depicted in Figure 7, shows the displacement of the principal point with respect to the image centre in x and y directions. The maps of optical distortions show that deviations are very small in the image centre where the object to be captured is located (maximum deviation at object border of about three pixels).

**Table 2.** Intrinsic parameters (mean values for left and right camera) in air and water (RGB images).

Parameter (mm)	Air	Water
Focal length	37.19	42.43
Principal point px	12.23	11.04
Principal point py	8.22	7.42

Figure 7. Complete distortion model (radial plus tangential distortions) in air and clear water. The cross and the circle indicate the image centre and the principal point, respectively. The contours represent the optical distortions values (in pixel), (**a**) Air; (**b**) Water.

(**a**)   (**b**)

To evaluate the reliability of the underwater calibration, we considered the re-projection error for each corner point of the calibration checkerboard, defined as the distance between the points of the actual image, and those computed from the back projection of the ideal checkerboard. The error dispersions are Gaussian-distributed and the values of mean and standard deviation obtained in water, computed on both x and y directions, are comparable with the ones obtained in air (see Figure 8). This comparison allows us to consider that the assumed linear model of calibration in air can be used also in water, without taking into account the refraction effects.

Figure 8. Distribution of the re-projection error (in pixels) in air and clear water (RGB images); (**a**) Air; (**b**) Water.

(**a**) (**b**)
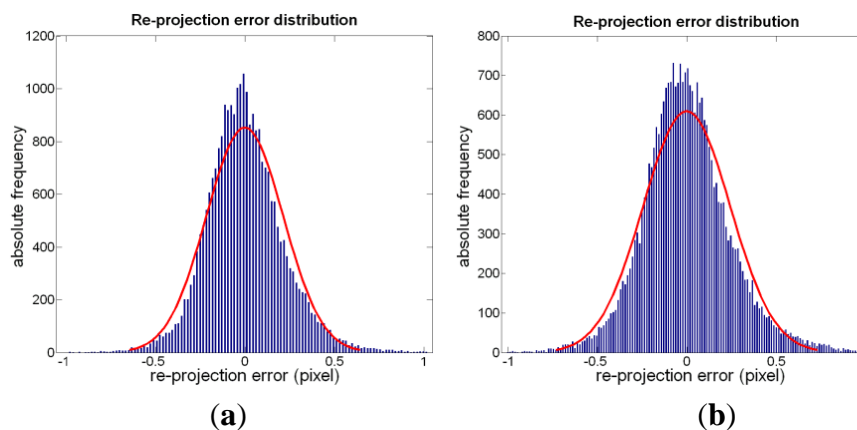
Finally, in order to assess the effect of chromatic aberration, we have estimated the calibration parameters in the Red, Green and Blue colour channels. This optical phenomenon can cause blur in the images due to a degraded focus quality (longitudinal chromatic aberration) and a misregistration of the colour channels (transversal chromatic aberration), two factors that are difficult to reduce [37]. Despite this issue, the results have not shown significant differences in the values obtained by performing the camera calibration in water for each colour channel.

4.3. 3D Reconstruction—Passive Stereo

The dense stereo matching algorithm PVMS is used for the 3D reconstruction with the passive stereo technique. The algorithm inputs are an undistorted image pair and the 3×4 camera projection matrix, and the output is a coloured 3D point cloud. The camera projection matrix is computed from the same calibration parameters (intrinsic parameter matrix, rotation matrix and translation vector between left and right camera) of the stereo rig used for 3D reconstruction with the active technique. In order to correct the distortion of the image pair, we used a tool provided with the Camera Calibration Toolbox [32]. The PMVS parameters used for the fine tuning of the 3D reconstructions are the size of the correlation window and the level in the internal image pyramid used for the computation.

In our experimentation, we used a fixed correlation window with a size of 7×7 pixels, while the image resolution (image pyramid level) has been chosen accordingly to the turbidity level. A series of preliminary tests have been conducted in order to properly set up the algorithm parameters. In particular, for each turbidity condition, the value of the pyramid level k has been progressively reduced (the image is resized by a factor 2 k) from 1 to 3, maintaining the correlation window fixed to 7×7 pixels.

Furthermore, the image pyramid level is fixed to 1, increasing the correlation window size from 7×7 pixels to 21×21 pixels. The results showed that as the turbidity increases, a wider correlation window is needed to overcome the loss of contrast due to the scattering effect. Otherwise, the same results can be obtained by using a fixed correlation window on a smaller image (with a higher pyramid level). A larger correlation window means higher computation time, hence, in order to preserve the reconstruction quality and save time, we used a small window of 7×7 pixels with k = 3 for the highest turbidity level, k = 2 for the medium turbidity level and k = 1 for low turbidity level.

4.4. 3D reconstruction—Active Stereo

The different image sets of the two objects have been processed to obtain the point clouds. In the active case, the software for 3D reconstruction is provided by Scansystems [38], so the 3D point cloud of the scene is computed from the 50 image pairs and the calibration data.

The 24 projected patterns allow us to exploit the maximum resolution of the projector (800×600) in order to obtain the maximum number of codified patches and, subsequently, the maximum number of 3D points. The lowest width of the bands of the projected pattern is four pixels, so the window of the codified patch is made up by about 20 pixels, if we refer to the used working distance. The reduction of the window size requires more patterns and a projector with a greater resolution, but this latter suffers also from a greater sensibility to the scattering.

On the basis of our preliminary investigations, where we reduced the width of the projected bands, we can assume that a minimum band of four pixels is a good compromise between the reduction of the scattering effect and the acquisition of a sufficient number of 3D points. In fact, we have verified that, if the reduction of the width of bands is too strong, the threshold algorithm applied to detect and separate the black and white bands on the object image in presence of turbidity (medium and high turbidity conditions) is not able to work properly.

5. Result Analysis

The point clouds computed through the active and passive techniques must be further elaborated to obtain the final 3D models, because many outlier points are present. Rapidform® software has been used to edit the point clouds, while we used the open source software Cloud Compare [39] to perform measurements and comparisons on the 3D data. In this analysis, we aim to quantify the performance of both techniques by computing the density of 3D points, the percentage of outliers and the accuracy for the two objects. Moreover, a multi-channel analysis has been performed to evaluate the quality of 3D reconstructions by elaborating the images separately in each of the three colour channels.

5.1. Point Cloud Editing

The first step consists in the manual deletion of outlier points that do not belong to the object surface (Figure 9a–c). This is the most time-consuming process, as it is highly dependent upon the surface complexity of the object to be processed. Whereas for a simple object, such as the amphora, only few steps with the lasso selection tool are required, a more complex surface requires manual and accurate intervention. In particular, once that a view of the 3D point cloud is chosen, a cursor is moved

with the mouse on the screen to sketch a lasso (an arbitrary closed line) that allows for selecting a group of 3D points. The second cleaning step is the application of the noise filtering functions implemented in the software (Figure 9d), that delete automatically the noisy points close to the object surface. Figure 10 shows the points clouds acquired in different turbidity conditions with both the techniques after the editing procedure.

Figure 9. Point cloud cleaning procedure. Outlier point manual selection (**a**), selected outlier points (**b**), cleaned point cloud after outlier deletion (**c**) and final point cloud after noise filtering application (**d**).
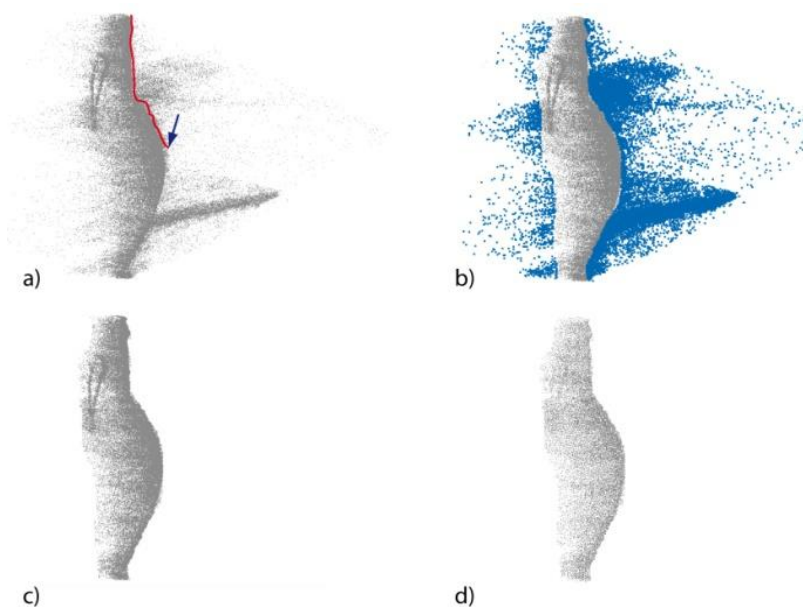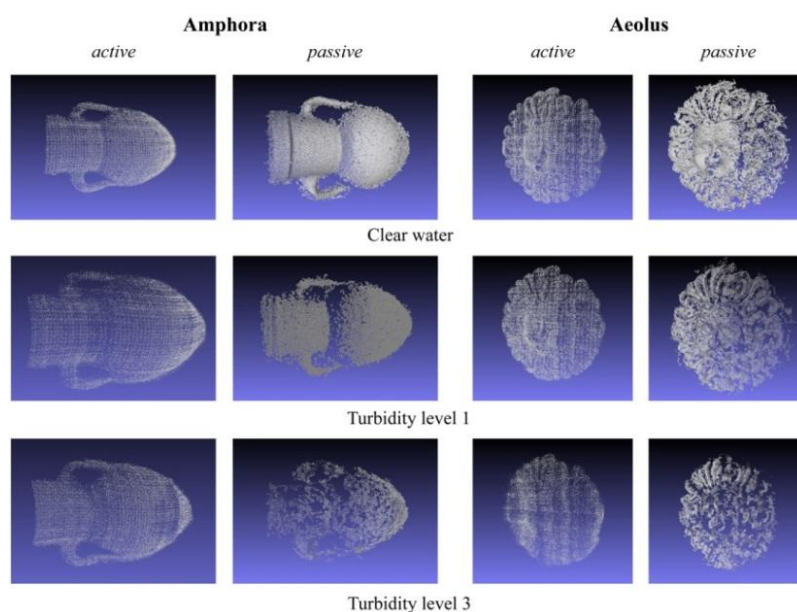

Figure 10. Point clouds obtained after editing procedure in different environment conditions for both active and passive techniques.

5.2. 3D Point Density

The density of 3D points acquired is computed by considering the ratio between the number of 3D points obtained after the editing procedure and the number of pixels that constitute the object in the image. The number of 3D points per 100 pixels (Np, Na for passive and active, respectively) are reported in Table 3 for both techniques. By comparing Np and Na in air, we found that the active technique returns a point density about three times lower than Np, as it depends on the projector resolution (800×600 in our case), while the passive technique depends on the image resolution. However, in medium turbidity (T2) conditions, this difference is reduced, because the passive technique is more affected by the turbidity which causes blur in the images. In fact, as we can see in Table 3, in the passive case the increase of turbidity causes a drastic decrease of the acquired points, while in active stereo the reduction of points is less noticeable, because this latter is less dependent on textures.

Table 3. Acquired 3D points per 100 pixel (Np for passive and Na for active stereo) and percentage values of deleted 3D points (Npc% for passive Nac% for active stereo) for Amphora and Aeolus.

Condition	Amphora				Aeolus			
	Np	Npc%	Na	Nac%	Np	Npc%	Na	Nac%
Air	13.0	1.93	3.4	4.52	11.9	0.98	3.2	3.17
CW	14.3	2.79	3.9	17.50	13.8	0.94	3.5	22.61
T1	14.2	4.23	4.1	23.96	3.8	3.53	3.2	24.93
T2	3.8	5.07	3.3	34.58	2.8	9.93	2.9	20.50
T3	2.9	6.35	3.5	30.63	2.5	3.21	1.7	42.81
T4	0.2	20.42	2.2	47.99	-	-	-	-

Moreover, it is important to analyse the percentage of points deleted during the preliminary editing procedure. The percentage values of deleted points (Npc% and Nac%) reported in Table 3 show that the active stereo technique is more affected by scattering, even under low turbidity conditions. Therefore, a time-consuming process is needed to manually delete the 3D points before using the noise filter, which works only on points that are very close to the object surface. The passive technique requires a lower noise reduction, but the scattering effect causes a loss of image contrast that compromises the matching of correspondent points and may generate holes in the reconstructed point clouds.

5.3. Multi-Channel Analysis

It is common knowledge that in the underwater environment the absorption of light components varies with the depth, in particular the red component is completely absorbed at about 10 m, so the colours in the images of underwater scenes are altered and present some variations. However, in close-range capture this problem can be solved by using artificial light sources as strobes or lamps, and a multi-channel analysis can be performed in order to evaluate the quality of 3D reconstructions in the three separate colour channels. In fact, the contrast of the acquired images can be very different in the three channels, depending on the illumination and the object texture, so we had to check if the channel

separation of light results in a spectral discrimination of component noise. The acquired RGB images used in the 3D reconstruction have been separated into the Red, Green and Blue channels and then processed. In the passive case, no significant differences have been observed, while in the active case substantial differences in the point density have been registered. As we can see in Figure 11, the outlier points are highly reduced when using the Green channel. The percentage of deleted points at turbidity level T1 has been reduced by about 66%, hence reducing the need for manual cleaning operation. Table 4 summarizes the results for the amphora object at each turbidity level, showing that the improvement ranges from about 40% (turbidity level T4) to about 70% (turbidity level T2). The results for the Aeolus mask present similar improvements, so they have been omitted.

Figure 11. Acquired images and unclean point clouds at a light turbidity level (T1) in each colour channel: the blue points are due to scattering effects, which are highly reduced in the green channel.
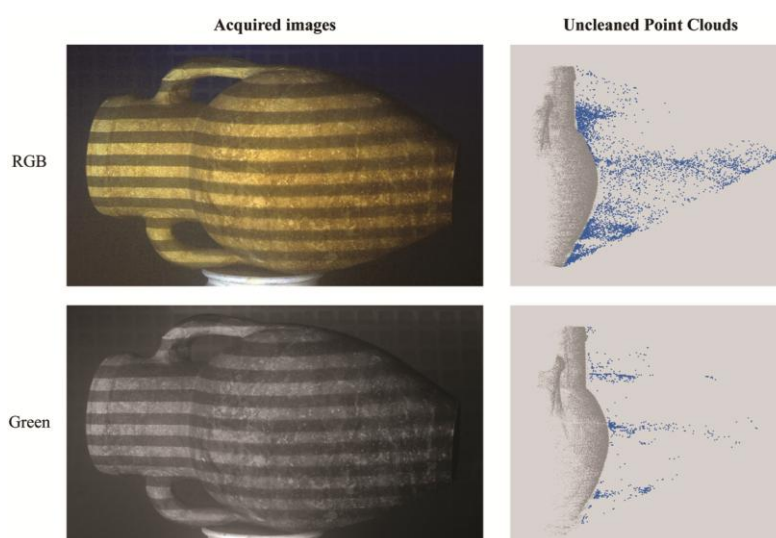

Table 4. Percentage values of deleted 3D points for active stereo technique in each colour channel, Red (Nac%-R), Green (Nac%-G) and Blue (Nac%-B), compared to the values computed from RGB images (Nac%-RGB). The table is related to the object Amphora.

Condition	Na RGB	Na R	Na G	Na B	Nac %-RGB	Nac %-R	Nac %-G	Nac %-B
Air	3.4	3.3	3.3	2.8	4.52	3.49	2.67	17.02
CW	3.9	3.8	4.4	4.0	17.50	12.33	6.45	11.72
T1	4.1	4.4	5.2	3.9	23.96	11.33	8.09	16.09
T2	3.3	3.0	2.4	2.0	34.58	15.84	9.26	16.11
T3	3.5	3.6	2.3	1.5	30.63	23.55	16.56	36.75
T4	2.2	2.1	1.2	0.5	47.99	29.03	28.10	58.56

5.4. Accuracy Evaluation

To evaluate the accuracy of the 3D reconstruction in water for both techniques, a series of experiments based on a statistical approach have been planned. First, two rectified samples, a plane

(297 × 420 mm) and a cylinder (height 400 mm, radius 150 mm), have been acquired in clear water. These two samples allow us to quantify and compare the displacements of the acquired 3D points with respect to two known surfaces taken as references. The accuracy has been evaluated by comparing the 3D point clouds acquired in clear water with a cloud obtained in air and digitized through the active technique, which ensures high-accuracy reconstruction in air [38]. To measure the accuracy we have computed the mean distance μ (mm) among points over the entire cloud (distance between a point on one cloud and the closest point on the other cloud), using Cloud Compare software.

Table 5 contains the mean distance for the two samples and the related standard deviation σ (mm) obtained with the two techniques. The results lead us to consider that the two techniques show a very similar accuracy in water: we can note that the error is greater for a curved surface like the cylinder, given a mean distance of 0.29 mm and 0.20 mm for active and passive techniques, respectively. We can consider the values obtained for the flat surface (0.09 mm and 0.11 mm) as the maximum accuracy obtainable in water with the two techniques, respectively.

Table 5. Accuracy evaluation on a planar and cylindrical sample in clear water (mm).

	Plane Sample		Cylinder Sample	
	Passive	**Active**	**Passive**	**Active**
μ	0.11	0.09	0.20	0.29
σ	0.08	0.07	0.13	0.25

In the second test, we have evaluated the accuracy at different turbidity levels for the Aeolus mask and amphora and compared the results with the ones obtained in air. A map of the mean distance distribution is shown in Figure 12, while the values of μ and σ are reported in Tables 6 and 7. There is a slight increase in the standard deviation in presence of low and medium turbidity. When the environmental conditions are the same, the discriminant value is represented by the material properties: Aeolus shows a higher standard deviation caused by its dark texture. When comparing the μ data, we can assume that the accuracy depends on the texture of the objects and the turbidity conditions: we can conclude that the two techniques give comparable results, obtaining a geometrical error in high turbidity conditions (T3) lesser than 2.5 mm, a value that can be acceptable for many real world applications.

Figure 12. Map of mean distance distribution for the Aeolus mask.
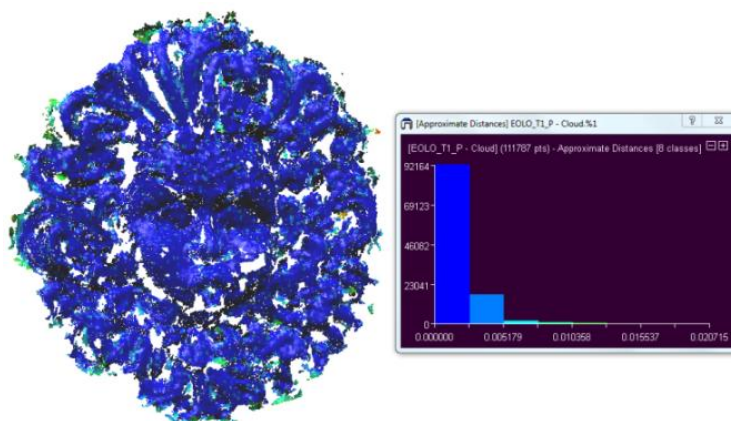
By comparing the results in clear water of Tables 5–7, it is apparent that the accuracy of the 3D reconstruction for both the techniques depends on the object shape; in particular, it increases considering a planar surface, a regular curved surface (Amphora) and an irregular surface (Aeolus).

Table 6. Amphora: accuracy evaluation (mean distance μ in mm and standard deviation σ).

Condition	Active		Passive	
	μ	σ	μ	σ
CW	0.40	0.15	0.22	0.55
T1	0.44	0.23	0.52	0.65
T2	0.47	0.23	0.69	0.78
T3	0.75	0.66	0.85	0.87
T4	0.86	0.88	2.23	1.86

Table 7. Aeolus: accuracy evaluation (mean distance μ in mm and standard deviation σ).

Condition	Active		Passive	
	μ	σ	μ	σ
CW	0.75	0.81	0.70	1.06
T1	1.61	1.26	1.53	0.90
T2	1.98	1.56	2.08	1.49
T3	2.32	1.98	2.44	2.00
T4	2.83	2.81	-	-

In [40] the authors have analysed the refraction effects on the 3D reconstruction by multi-view acquisition, obtaining a minimum accuracy of 0.39 mm in clear water. This value can be compared with our results: with regards to the active technique, the value is the same, while for the passive case it is greater, even if included in the error range of the standard deviation.

5.5. 3D Point Cloud Integration

By analysing the acquired point clouds (see Figure 10), we can confirm that the active technique is able to reconstruct the entire surface of the object in each condition and the point density is more or less constant over the surface as the 3D points are regularly distributed, while passive 3D reconstructions return a variable point density on the surface and are not able to reconstruct some portions. It is apparent that the active technique computes in each case the 3D point coordinates of the codified scene, while the passive one computes only the 3D coordinates of matched points in the image areas where features are found. In particular, as shown in Figure 13, it is difficult to find features in dark and homogeneous areas with the passive technique, while the codification of image points may be improved by projecting structured-light patterns. Obviously, in absence of other information, void parts must be reconstructed in post-processing, with the inevitable possibility of errors. For this reason we propose an integration of 3D point clouds acquired with the two approaches, exploiting the advantages of each technique: the better point density of the passive technique and the better coverage of the object surface of the active technique.

Figure 13. Dark and homogeneous areas present in the Aeolus image (**Left**) are better detected using structured light (**Right**), in the presence of turbidity.

A comparison of two point clouds acquired with the two techniques is shown in Figure 14, where it is possible to recognize the areas that are not reconstructed by the passive technique. The union of the two point clouds is performed with Rapidform software (Combine tool) in order to obtain a single point cloud. Then we apply a redundancy filter in order to delete the closest points by setting an automatic threshold value, and finally a triangular mesh can be created (Figure 14). This resulting point cloud has been compared with the one acquired in air (as in Table 7) for the T3 case and we have obtained a mean distance of 1.95 mm with a standard deviation of 1.75 mm. The point density is increased at 4.7 points per 100 pixels. With respect to each individual technique, the error has decreased as a greater number of closer points have been found.

Figure14. (**Left**), comparison between the two point clouds of the Aeolus mask obtained with passive (blue points) and active technique (red points), respectively; (**Right**), triangular mesh of the Aeolus mask obtained from the integrated point cloud.
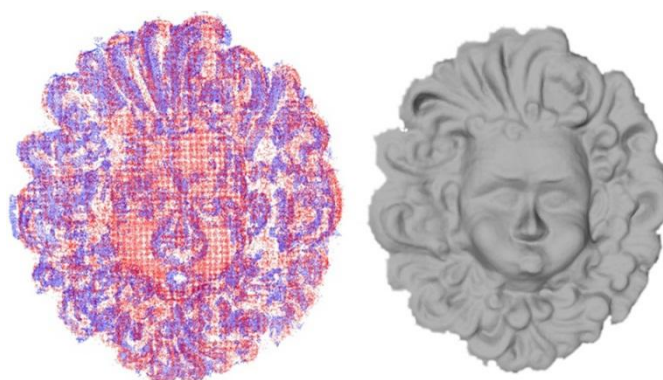
6. Discussion

The experimental setup proposed in this work is not suitable for tests in real marine conditions and applications (handling by a scuba diver or installation on a ROV). In fact, a series of additional considerations must be taken into consideration with regards to the presence of moving elements (flora and fauna), the colour alteration, the corrosion of the object surface, the instability of the imaging system and the synchronization of the camera shutters. The purpose of our system is to perform a fair

comparison between two 3D techniques based on image acquisition that use two different approaches to reconstruct 3D models. Moreover, some hardware modifications must be taken into account for the lighting sources (wavelengths and illuminance) in order to improve the image acquisition phase and reduce the required time. We neglect the variation of absorption of colour spectrum in marine environments, because we have conducted the acquisitions in the laboratory using just an artificial light source.

Therefore, a deep analysis of the real condition effects must be done and a new design of experimentations in real cases is required—we hope to present them in a future work. We can consider the results of this experimentation as reference data which we will compare to the ones to be conducted in a true marine environment. For example, we could understand if the calibration parameters can be computed in laboratory with a water tank avoiding a calibration procedure *in situ*, or if it is necessary to include refraction and scattering effects in the optical model of image formation. However, we have found that the two techniques gave comparable results and the 3D data integration can be a solution to improve the 3D acquisition quality in terms of coverage of the acquired scene and stability in turbidity conditions. In particular, the projection of patterns on the scene can help the reconstruction of areas with scarce contrast and natural features. For this purpose, a series of possible solutions to facilitate the application in real cases can be proposed as follows:

- using the fringe projection technique that requires the projection of just one pattern, considerably reducing the acquisition time;
- projecting a coloured pattern to increase the image contrast of the acquired pattern;
- projecting a fringe pattern using a grating fixed on the flash lamp handled by a scuba diver;
- employing a more powerful light source;
- studying, with a multi-channel approach, the separation of the projected pattern from the texture acquired simultaneously with one shot, in order to apply active and passive techniques and perform the 3D data integration.

7. Conclusions

We have proposed a preliminary comparison between two 3D imaging techniques based on active and passive approaches. The experimentation has been conducted in the laboratory in a water tank, controlling the lighting conditions in order to test the performances of both techniques in terms of accuracy and density of the acquired 3D points. 3D acquisitions have performed under poor visibility conditions, obtained by suspending different concentrations of clay in water, which allow us to take into account the scattering effects, *i.e.*, the main limitation in underwater imaging. For a fair comparison, we have adopted a 3D system in stereo configuration, composed by two cameras and a projector lodged in waterproof housings. Therefore, the acquisition process is performed for both the techniques with the same experimental setup, working distance, lighting conditions and calibration parameters. This setup allows us to align automatically the acquired point clouds, because these are computed with respect to the same coordinate system.

The results of calibration show that a linear model can be employed in 3D underwater acquisition, without taking into account the refractive index of media, with an acceptable accuracy. The point clouds obtained with the active technique give more stable results in presence of turbidity, due to the

use of coded patterns to compute the 3D point coordinates, but they are affected by a noticeable noise level due to scattering even in low turbidity, thus requiring some manual operations to clean the point cloud. To reduce this problem we have investigated the use of a single colour channel for the 3D reconstruction, finding that the acquisition of underwater images in the Green channel may give good improvements in terms of noise reduction and surface accuracy. The passive technique gives a cleaner and denser point cloud in clear water with low turbidity levels, but the results are strongly related to the surface texture of the acquired object. In fact, in low-textured areas, the loss of contrast due to turbidity results in holes and missing areas. Moreover, we propose an integration of 3D point clouds acquired with the two approaches, exploiting the advantages of each technique: the better point density of the passive technique and the better coverage of the object surface of the active technique. The 3D point cloud resulting by the integration allows for the improvement of accuracy and density.

Finally, the proposed 3D system has been designed for laboratory tests and at the moment is not suitable for real world applications. In future works, we will design a new experimentation layout to test the two techniques also in a marine environment, taking into account more factors such as the variable lighting conditions, the surface corrosion, the system instability, the synchronisation of the cameras and more. For the real tests, we will probably employ the fringe projection technique, because it allows for decreasing drastically the required time for the acquisition. Our results can be used as a reference for further comparisons in the analysis of other 3D techniques and algorithms.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

References

1. Gallo, A.; Angilica, A.; Bianco, G.; de Filippo, F.; Muzzupappa, M.; Davidde, B.; Bruno, F. 3D Reconstruction and Virtual Exploration of Submerged Structures: A Case Study in the Underwater Archaeological Site of Baia (Italy). In Proceedings of 13th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2012), Brighton, UK, 19–21 November 2012; pp. 121–128.
2. Bythell, J.; Pan, P.; Lee, J. Three-dimensional morphometric measurements of reef corals using underwater photogrammetry techniques. *Coral Reefs* **2001**, *20*, 193–199.
3. Bräger, S.; Chong, A.K. An application of close range photogrammetry in dolphin studies. *Photogramm. Record* **1999**, *16*, 503–517.
4. Pizarro, O.; Eustice, R.; Singh, H. Large area 3-D reconstructions from underwater optical surveys. *IEEE J. Ocean. Eng.* **2009**, *34*, 150–169.
5. Schulein, R.; Javidi, B. Underwater multi-view three-dimensional imaging. *IEEE J. Disp. Technol.* **2008**, *4*, 351–353.

6.   Jaffe, J.S.; Dunn, C. A model-based comparison of underwater imaging systems. *SPIE Proc.* **1988**, *0925*, 344–350.

7.   Cho, M.; Javidi, B. Three-dimensional visualization of objects in turbid water using integral imaging. *IEEE J.Disp. Technol.* **2010**, *6*, 544–547.

8.   Bruno, F.; Bianco, G.; Muzzupappa, M.; Barone, S.; Razionale, A.V. Experimentation of structured light and stereo vision for underwater 3D reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 508–518.

9.   Skarlatosa, D.; Demestihab, S.; Kiparissia, S. An 'open' method for 3D modelling and mapping in underwater archaeological sites. *Int. J. Herit. Digital Era* **2012**, *1*, 2–24.

10. Kocak, D.M.; Caimi, F.M. The current art of underwater imaging—with a glimpse of the past and vision of the future. *Mar. Technol. Soc. J.* **2005**, *39*, 5–26.

11. Jaffe, J.S.; Moore, K.D.; McLean, J.; Strand, M.P. Underwater optical imaging: Status and prospects. *Oceanography* **2001**, *14*, 64–75.

12. Allais, A.G.; Brandou, V.; Dentrecolas, S.; Gilliotte, J.P.; Perrier, M. Iris—A Vision System to Reconstruct Natural Deep-Sea Scenes in 3D. In Proceedings of the Seventeenth International Offshore and Polar Engineering Conference, Lisbon, Portugal, 1–6 July 2007; pp. 111–118.

13. Corke, P.; Detweiler, C.; Dunbabin, M.; Hamilton, M.; Rus, D.; Vasilescu, I. Experiments with Underwater Robot Localization and Tracking. In Proceedings of the IEEE International Conference on Robotics and Automation, Roma, Italy, 10–14 April 2007.

14. Menci. Available online: http://www.menci.com (accessed on 10 June 2013).

15. Leone, A.; Diraco, G.; Distante, C. Stereoscopic System for 3D Seabed Mosaic Reconstruction. In Proceedings of International Conference on Image Processing (ICIP), Atlanta, GA, USA, 8–11 October 2006; pp. II541–II544.

16. Drap, P.; Scaradozzi, D.; Gambogi, P.; Long, L.; Gauche, F. Photogrammetry for Virtual Exploration of Underwater Archeological Sites. In Proceedings of the XXI International Scientific Committee for Documentation of Cultural Heritage (CIPA) Symposium, Athens, Greece, 1–6 October 2007.

17. He, D.M.; Seet, G.G.L. Divergent-beam lidar imaging in turbid water. *Opt. Laser. Eng.* **2004**, *41*, 217–231.

18. Schechner, Y.Y.; Karpel, N. Clear Underwater Vision. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; pp. I536–I543.

19. Narasimhan, S.G.; Nayar, S.K.; Sun, B.; Koppal, S.J. Structured Light in Scattering Media. In Proceedings of SIGGRAPH Asia '08, Suntec, Singapore, 11–13 December 2008; pp. 1–8.

20. Moore, K.D. Intercalibration method for underwater three-dimensional mapping laser line scan systems. *Appl. Opt.* **2001**, *40*, 5991–6004.

21. Dalgleish, F.R.; Caimi, F.M.; Mazel, C.H.; Glynn, J.M. Extended Range Underwater Optical Imaging Architecture. In Proceedings of IEEE Oceans 2006, Boston, MA, USA, 15–21 September 2006.

22. Tan, C.; Seet, G.; Sluzek, A.; He, D. A novel application of range-gated underwater laser imaging system (ULIS) in near-target turbid medium. *Opt. Lasers Eng.* **2005**, *43*, 995–1009.

23. Rives, P.; Borrelly, J.J. Underwater Pipe Inspection Task using Visual Servoing Techniques. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Grenoble, France, 11 September 1997; pp. 63–68.

24. Roman, C.; Inglis, G.; Rutter, J. Application of Structured Light Imaging for High Resolution Mapping of Underwater Archaeological Sites. In Proceedings IEEE Oceans, Sydney, Australia, 24–27 May 2010.

25. Levoy, M.; Chen, B.; Vaish, V.; Horowitz, M.; McDowall, I.; Bolas, M. Synthetic Aperture Confocal Imaging. In SIGGRAPH; Los Angeles, CA, USA, 8–12 August 2004.

26. Scharstein, D.; Szeliski, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision* **2002**, *47*, 7–42.

27. Barone, S.; Razionale, A.V. A reverse engineering methodology to capture complex shapes. In Proceedings of XVI International Congress of Engineering Graphics (INGEGRAF), Zaragoza, Spain, 2–4 June 2004; pp. 1–10.

28. Lavest, J.M.; Rives, G.; Lapresté, J.T. Underwater Camera Calibration. In Proceedings of 6th European Conference on Computer Vision—Part II (ECCV 2000), Dublin, Ireland, 26 June–1 July 2000; pp. 654–668.

29. Treibitz, T.; Schechner, Y.; Kunz, C.; Singh, H. Flat refractive geometry. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 51–65.

30. Kunz, C.; Singh, H. Hemispherical Refraction and Camera Calibration in Underwater Vision. In Proceedings IEEE Oceans, Quebec, Canada, 15–18 September 2008.

31. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2000; p. 655.

32. Bouguet. Available online: http://www.vision.caltech.edu/bouguetj/ (accessed on 10 June 2013).

33. Furukawa, Y.; Ponce, J. Accurate, Dense, and Robust Multi-View Stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1362–1376.

34. Abdel-Aziz, Y.I.; Karara, H.M. Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. In Proceedings of the Symposium on Close-Range Photogrammetry, Falls Church, VA, USA, 1971; pp. 1–18.

35. Salvi, J.; Armangué, X.; Batlle, J.A. Comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognit.* **2004**, *35*, 1617–1635.

36. Edge, M. *The Underwater Photographer*; Focal Press: Boca Raton, FL, USA, 2006; Chapter 5, pp. 314–317.

37. Cronk, S.; Fraser, C.S.; Hanley, H.B. Automatic metric calibration of colour digital cameras. *Photogramm. Record* **2006**, *21*, 355–372.

38. Scansystems. Available online: http://www.scansystems.it (accessed on 10 June 2013).

39. CloudCompare. Available online: http://www.danielgm.net/cc/ (accessed on 10 June 2013).

40. Kang, L.; Wu, L.; Yang, Y.H. Experimental study of the influence of refraction on underwater three-dimensional reconstruction using the SVP camera model. *Appl. Opt.* **2012**, *51*, 7591–7603.