



Performance evaluation of underwater image pre-processing algorithms for the improvement of multi-view 3D reconstruction

Alessandro Gallo¹, Fabio Bruno¹, Loris Barbieri¹, Antonio Lagudi¹, Maurizio Muzzupappa¹

¹ Department of Mechanical, Energy and Management Engineering (DIMEG), University of Calabria, Via P. Bucci 46C, 87036 Rende, Italy

ABSTRACT

3D models of submerged structures and underwater archaeological finds are widely used in various and different applications, such as monitoring, analysis, dissemination, and inspection. Underwater environments are characterised by poor visibility conditions and the presence of marine flora and fauna. Consequently, the adoption of passive optical techniques for the 3D reconstruction of underwater scenarios is a highly challenging task.

This article presents a performance analysis conducted on a multi-view technique that is commonly used in air in order to highlight its limits in the underwater environment and then provide guidelines for the accurate modelling of a submerged site in poor visibility conditions. A performance analysis has been performed by comparing different image enhancement algorithms, and the results have been adopted to reconstruct an area of 40 m² at a depth of about 5 m at the underwater archaeological site of Baiae (Italy).

Section: RESEARCH PAPER

Keywords: 3D reconstruction; Underwater Cultural Heritage; Image enhancement; Underwater imaging

Citation: Alessandro Gallo, Fabio Bruno, Loris Barbieri, Antonio Lagudi, Maurizio Muzzupappa, Performance evaluation of underwater image pre-processing algorithms for the improvement of multi-view 3D reconstruction, Acta IMEKO, vol. 8, no. 3, article 11, September 2019, identifier: IMEKO-ACTA-08 (2019)-03-11

Section Editor: Egidio De Benedetto, University of Salento, Italy

Received November 12, 2018; **In final form** May 28, 2019; **Published** September 2019

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Corresponding author: Loris Barbieri, e-mail: loris.barbieri@unical.it

1. INTRODUCTION

The 3D reconstruction of submerged structures or archaeological finds has achieved notable popularity in Underwater Cultural Heritage (UCH) preservation, as the method may allow for the exploration of sites located in inaccessible and hostile environments.

In the last decade, techniques and tools for 3D reconstructions have been widely employed in the underwater archaeology field according to the guidelines of UNESCO, which suggest the in-situ preservation of underwater heritage [1]. Among the different 3D imaging techniques that are suitable for underwater applications, photogrammetry represents a valid method of reconstructing 3D scenes from a set of images taken from different viewpoints [2]. The popularity of this technique is also due to the acquisition devices (still or a movie camera with appropriate waterproof casings) that are affordable and easy to use compared to dedicated devices like LIDAR, multi-beam, etc. [3]. Furthermore, these devices can be handled by scuba divers or mounted on underwater robots [4]. Unfortunately, image-

based acquisition suffers due to the poor environmental conditions. The depth of the water, flora, fauna, weather conditions, and sea currents are all factors that affect visibility, refraction, and lighting conditions. Consequently, these factors limit underwater photogrammetry's scope to close-range applications, and further efforts are required to improve the radiometric quality of the images. For these reasons, the enhancement of underwater images is still a necessary step in improving the accuracy of 3D reconstruction and creating realistic textures.

This article presents a performance evaluation of underwater image pre-processing algorithms for the improvement of multi-view 3D reconstruction. Two existing colour enhancement models, i.e. ACE (Automatic Colour Equalisation) and PCA (Principal Component Analysis) algorithms, have been tested to compare their results with those provided by a new method based on histogram stretching and manual retouching (HIST). To this end, an experimental campaign has been planned using Design of Experiment (DOE) [5] criteria to investigate the factors that affect reconstruction accuracy. The experimental

campaign has been carried out in the underwater archaeological site of Baiae (Naples, Italy), where the seafloor, with a water depth ranging between 2.5 and 20 m, offers a particularly interesting environment, as it encompasses a submerged area of many hectares and presents a wide range of different architectural structures with a number of decorations that are still preserved. The research has been conducted in the context of the iMARECulture project [6], [7], [8], which aims to develop new tools and technologies for improving public awareness of UCH.

The article is organised as follows. Section 2 presents related works about underwater image pre-processing methods. Section 3 describes the image acquisition and pre-processing stage. In section 4, the results of the statistical analysis are detailed. Section 5 presents the 3D reconstruction obtained by means of the enhanced images, and finally, conclusions are presented in section 6.

2. UNDERWATER IMAGE PRE-PROCESSING ALGORITHMS

Underwater pictures generally suffer from light absorption, which causes some defects mostly on the red channel (the first component of the light spectrum that is absorbed), and this effect is already noticeable at only a few metres of depth. The pre-processing of underwater images can be conducted with two different approaches: image restoration techniques or image enhancement methods [9], [10]. Image restoration techniques need some environmental parameters to be entered, such as scattering and attenuation coefficients, while image enhancement methods do not require a priori knowledge of the underwater environment.

The physical effects of visibility degradation have been analysed in [11], showing that the degradation effects can be associated mainly with the partial polarisation of light. The developed algorithm is based on a couple of images taken through a polariser at different orientation, improving contrast and colour and doubling the underwater visibility range. The work of [12] presents an image restoration filter based on a simplified version of the Jaffe-McGlamery underwater image formation model, which can be used for images with limited backscatter in diffuse lighting.

The ACE algorithm [13] is inspired by the human vision, which is able to adapt to highly variable lighting conditions, extracting visual information from the underwater environment [14]. The algorithm combines the Patch White algorithm with the Gray World algorithm, taking into account the spatial distribution of colour information. In the first stages of the ACE method, chromatic data and pixels are processed and adjusted according to the information contained in the image. Subsequently, colours in the output image are restored and enhanced [15]. Different to the ACE algorithm, which can adapt to widely varying lighting conditions and can extract visual information, in order to reduce the number of variables considerably while still retaining much of the information in the original dataset, it is possible to adopt the Principal Component Analysis (PCA) algorithm. PCA is one of the most popular multivariate statistical techniques that analyses a data table representing observations described by several dependent variables and extracts the important information in the form of a set of new orthogonal variables called principal components. In this specific application, the PCA algorithm allows us to extract a dominant colour of the image. Hence, in most cases, the water colour also provides good results in term of colour enhancement.

An automatic enhancement algorithm that does not require any correction parameter has been proposed in [16], where each source of errors is corrected sequentially. The first step removes the moiré effect, then a homomorphic or frequency filter is applied to equalise brightness and to enhance the contrast. Regarding the acquisition noise, a wavelet denoising filter followed by an anisotropic filtering has been applied. Finally, dynamic expansion is applied to increase contrast, followed by colour equalisation. The process is performed on one channel, specifically the YCbCr colour space, in order to optimise the computation time. Even if this last step speeds up all the following processes avoiding the need to process each RGB channel each time, it is important to point out that the use of a homomorphic filter affects the geometry and could generate errors on the reconstructed scene. The effectiveness of the use of different colour spaces for the enhancement of underwater images has been demonstrated in [17], where a slide stretching algorithm has been used both on RGB and HSI colour spaces. After a contrast stretching on RGB colour space has been performed, the resulting images have been converted to HSI colour space and processed through saturation and intensity stretching in order to increase the true colour and solve the problem of lighting. The aim of underwater colour correction is not only to obtain better quality images, but also to improve the performance of feature extraction algorithms in terms of the detection of feature points. The effects of different image pre-processing methods on the performance of the SURF (Speeded Up Robust Features) detector [18] have been investigated in [20], and the IACE (Image Adaptive Contrast Enhancement) method has been proposed. In particular, the IACE method enhances the intrinsic features in images, like corners, edges, and blobs, along with maintaining the relative contrast between the pixels. Thanks to this capability, the IACE method has proven better than other techniques, like Histogram Equalisation and Multiscale Retinex algorithm for Color Enhancement, in terms of the repeatability of their SURF detector and the robustness and distinctiveness of their SURF descriptor.

Different to previous works [20] that focus on the comparison of different image enhancement algorithms, with all other conditions being equal, this article presents a performance analysis based on a DOE approach, which takes into account the main influential factors that affect 3D reconstruction quality.

3. EXPERIMENTATION

The experiment was undertaken in the underwater archaeological site of Baiae, which is located few kilometres north of Naples (Italy). The submerged environment of Baiae is characterised by highly critical visibility conditions due to water turbidity and the heavy presence of flora and fauna. The area selected for the experimentation is the thermal room of 'Villa Protiro', with a size of 5 x 8 m at an average depth from the sea level of 5 m. The choice of this area is due to the presence of different building materials (bricks, mortar, tile floors, etc.) and a strong colonisation of various bio-fouling agents. Because of the critical visibility conditions, 3D reconstruction techniques based on the multi-view stereo method are not sufficient for performing an accurate 3D reconstruction of the submerged archaeological area. The underwater images require, then, a pre-processing stage involving the adoption of image enhancement algorithms that could have a relevant or mediocre impact on the quality of the final 3D reconstructed model.

3.1. Experimental setup

The experimental setup consists of a camera, its underwater housing, and two underwater strobes. The camera is a Nikon D7000 reflex device equipped with a CMOS (Complementary Metal-Oxide Semiconductor) sensor size of 23.6 x 15.8 mm and a resolution of 4928 × 3264 pixels (16.2 effective megapixels), as well as a AF-Nikkor 20 mm lens. The underwater housing, manufactured by Ikelite, is equipped with a spherical port. The flashguns are connected to the camera housing with a pair of articulated arms at a distance of 45 cm. The two strobes have been fixed at a distance of 5 cm behind the dome pointing outwards to illuminate the object with the 'edge' of the light beam. A calibration panel, produced by Lastolite, has been used in the beginning of the survey to acquire a colour calibration image to perform in-situ white balance correction, while a digital depth gauge has been used to maintain a constant depth from the seabed.

3.2. Image acquisition

The photogrammetric survey of the submerged area has been carried out in two different dive sessions, in the north and south parts of the site. The survey has been carried out according to a standard aerial photography layout: The diver swims at a distance from the submerged structures of about 2.5 m, taking overlapping pictures along straight lines that cover the whole area in the north-south direction. Another set of images has been acquired in the east-west direction. The occluded areas have been acquired using oblique photographs. At the end of the survey activity, the dataset included a total of around 700 images.

3.3. Colour enhancement of underwater images

The original images (OR) have been then enhanced by means of three algorithms (ACE, PCA, HIST) and corrected through an in-situ white balance correction procedure.

The ACE (Automatic Colour Enhancement) algorithm proposed in [9] and the PCA algorithm proposed in [21] have been adopted.

The HIST (Histogram Stretching and Manual Retouching) algorithm is a semiautomatic enhancement methodology that has been developed for this study. It is based on histogram stretching and a manual colour retouching procedure and has been implemented using batch actions in a graphics editor to rescue the maximum amount of information from a set of defective and noisy pictures. In particular, the HIST method consists of the following three-step procedure: preliminary histogram stretching to improve the contrast; mixing of the colour channels to balance the missing information on the red channel; creation of a set of adjustment layers, including saturation enhancement for some missing hues, contrast masks, colour balancing, and equalising.

In addition to the enhanced algorithms, the images have been processed by performing an in-situ white balance correction procedure (WB) performed by means of a Lastolite waterproof panel.

The following figure shows an original uncorrected image (Figure 1a) and those enhanced with the WB procedure (Figure 1b), ACE (Figure 1c), HIST (Figure 1d) and PCA (Figure 1e) algorithms.

3.4. Design of the experimental campaign

The experimental campaign has been planned according to the DOE criteria with the purpose of identifying the most influential factors affecting the results of the 3D reconstruction in the underwater environment. Particular attention has been

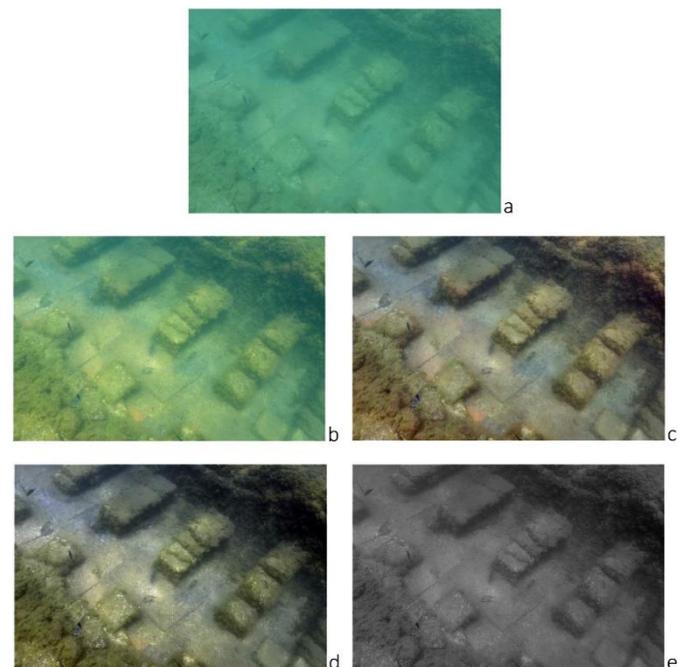


Figure 1. Sample original image (a) corrected with the in situ white balance measurement (b), enhanced with the ACE method (c), HIST method (d), and PCA method (e).

given to the effect of the image enhancement methods on the camera orientation and the self-calibration bundle adjustment process. The measured data has been compared and analysed by means of standard statistical tools to verify if a particular factor (or a combination of factors) has an impact on a parameter with a certain confidence level. On the basis of the results of this analysis, it is possible to find the best combination of factors that should be used for an accurate and dense 3D reconstruction by using a multi-view stereo technique.

3.5. Influencing factors

The first step is the selection and identification of the influencing factors that could have an influence on the quality of a 3D reconstruction performed with a multi-view technique in the underwater environment. The influencing factors have been chosen among those that cannot be controlled in situ, such as the presence of marine organisms in motion and the level of turbidity of the water. Furthermore, factors that can be considered as a direct consequence of others have been taken into account. For instance, focus settings depend on the distance from the subject, and the focal length can be set according to the required field of view and to the working distance. The factors selected for the experiment are reported in Table 1.

The first factor (EN) is related to the original images (OR) and to the colour enhancement algorithm (WB, HIST, ACE and PCA) adopted to improve underwater images.

The second factor refers to image resolution (PYR). The full resolution images have not been used in order to save computational time. The raw images (4928 x 3264 pixels) have been resized by means of the Mitchell-Netravali Cubic Filter [22] in order to create the following levels: level 1 for images of 2464 x 1632 pixels; level 2 for images of 1232 x 816 pixels; and level 3 for images of 616 x 408 pixels.

The third factor is represented by the composite RGB image and its three R (Red), G (Green), and B (Blue) components. This factor has been taken into consideration in order to investigate the influence of a single-colour channel on the reconstruction

Table 1. Influential factors and related symbol and levels.

Influential factor	Symbol	Level
Colour enhancement method	EN	OR, HIST, ACE, PCA, WB
Image pyramid level	PYR	1, 2, 3
Colour channel	CH	RGB, R, G, B
Image set	SET	SET 1, SET 2, ..., SET 7

quality with respect to the grayscale image obtained by combining the RGB components.

In order to perform a quantitative analysis of the impact of camera layout on the processing results, the last influencing factor is referred to the type of image set (SET). Seven subsets have been created, which differ among each other according to the type of shot (aerial vs. oblique), working distance, and overlapping pictures. In particular, the first two sets include photos that have been taken with a standard aerial layout. The third set includes pictures characterised by high overlap and good visibility due to the reduced distance from the submerged structures. The fourth and fifth sets cover the outside masonry structures of the outer walls, while the sixth and seventh sets group oblique pictures with variable working distances.

3.6. Measured parameters

The 3D reconstruction quality has been evaluated by means of four different parameters: the mean number of extracted features; the percentage of matched features; the percentage of oriented cameras; and bundle adjustment mean re-projection error. The mean number of extracted features ($num\bar{F}$) has been calculated according to the SIFT (Scale Invariant Feature Transform) operator [23] that consists of the following relationship:

$$num\bar{F}_{EN,PYR,CH,SET} = \frac{numF_{EN,PYR,CH,SET}}{numIm_{SET}} \quad (1)$$

where $numF$ is the total number of extracted features for each configuration, and $numIm$ is the number of images included in each set.

The percentage of matched features ($matchedF\%$) and percentage of oriented cameras ($cam\%$) parameters have been evaluated using Bundler [24], through the following relationships:

$$matchedF\%_{EN,PYR,CH,SET} = \frac{matchedF_{EN,PYR,CH,SET}}{numF_{EN,PYR,CH,SET}} \quad (2)$$

$$cam\%_{EN,PYR,CH,SET} = \frac{cam_{EN,PYR,CH,SET}}{numIm_{SET}} \quad (3)$$

where $matchedF$ represents the number of matched 3D points in the sparse scene reconstruction resulting at the end of the bundle adjustment process, and cam is the number of oriented images for each configuration.

The bundle adjustment mean re-projection error (available as output in the Bundler log file and measured in pixels) is the result of a minimisation problem applied to the sum of distances

between the projections of each track (a connected set of matching key points across multiple images) and its corresponding image features.

3.7. Dataset generation

As mentioned above in section 3.5, the whole dataset has been grouped into seven subsets according to: camera orientation; distance from the subject; pictures taken with flash; and the heavy presence of dark and bright areas. The grouping procedure has meant a selection and reduction in the number of images to 196 pictures.

A Matlab script has been programmed in order to manage the selected images and apply the different image enhancement algorithms. Firstly, the enhancement methods (ACE, PCA, WB, HIST) and the WB correction technique have been applied to the original full resolution images. Secondly, the red, green, and blue colour components have been extracted only from images enhanced with WB correction, ACE, and HIST methods. No action has been taken on images enhanced with the PCA method because this method produces a single-channel output. Lastly, all the images have been resized according to the pyramid levels.

4. STATISTICAL ANALYSIS

Table 2 shows the mean values of the measured parameters, described in section 3.6, computed for each influential factor.

The data has been analysed by means of statistical instruments and have been compared by performing an ANOVA test with a 95 % confidence level. The summary of the results is presented in Table 3, in which the measured parameters have been computed from the main source only. The Tukey post hoc test has been performed in order to find out significant differences between the groups of each influential factor.

4.1. Mean extracted features

Figure 2 reports the mean values of the extracted features for all the factors of influence (summarised in Table 2). The results show that the number of extracted points strongly depends on image resolution: A greater number of features is obtained from images with a higher resolution. In this regard, the data summarised in Table 3 shows a statistically significant difference

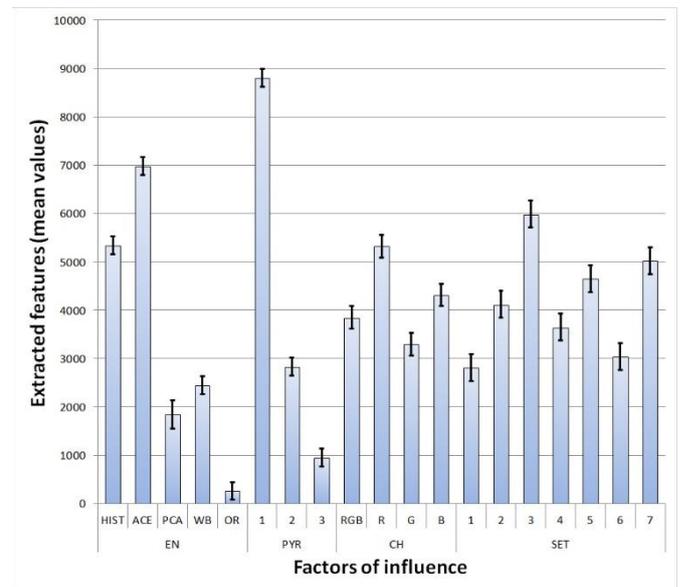


Figure 2. The mean values of extracted features for all the influential factors.

Table 2. Mean values of the measured parameters computed for each influential factor.

Factors	Mean extracted features (1)		% matched features (2)		% oriented cameras (3)		Bundle adjustment mean re-projection error (pixels)	
	F	p-value	F	p-value	F	p-value	F	p-value
EN	HIST	5344.9	2.41		43.24		0.19	
	ACE	6984.8	1.45		37.72		0.22	
	PCA	1841.2	2.07		33.38		0.18	
	WB	2452.1	1.21		21.23		0.22	
	OR	260.1	1.14		9.26		0.10	
PYR	1	8806.0	1.61		37.22		0.24	
	2	2829.6	2.04		32.58		0.16	
	3	954.2	1.35		20.42		0.13	
CH	RGB	3848.9	2.20		35.98		0.18	
	R	4315.0	1.36		30.51		0.17	
	G	3300.1	2.19		34.82		0.18	
	B	5322.4	0.91		18.97		0.16	
SET	1	2813.7	1.71		28.44		0.14	
	2	4120.1	1.81		25.97		0.15	
	3	5986.0	2.39		48.46		0.13	
	4	3647.2	2.56		47.02		0.22	
	5	4654.6	0.96		22.62		0.24	
	6	3040.7	1.43		26.59		0.24	
	7	5022.6	0.98		20.90		0.22	

among the three levels of the PYR factor, and this result shows that the number of features has a strong relationship with image resolution. Nevertheless, the images at level 1 (four times more pixels than level 2), produces only three times more features than the images at level 2 and nine times more features than the images at level 3 (containing 16 times less pixels). The image enhancement methods ACE and HIST have returned the best results. On the contrary, the colour channel appears to be less influential than enhancement algorithms, because these last ones operate a mixing between the various colour channels in different ways. The results related to the factor ‘image set’ show a difference in behaviour within the seven datasets. In particular, the highest number of features is extracted from the images belonging to set 3, which includes pictures taken at a reduced distance from the subject. The first two sets are related to the same area: The second set has been acquired after having removed the sand that covered the tiled floor in order to improve the reliability of point detection. Set 6 shows a lower number of features, as the oblique pictures have been taken from a greater distance, and the presence of the blue background is more evident.

4.2. Percentage of matched features

The most influential factor on the parameter ‘percentage of matched features’ (Table 2) is the image enhancement algorithm used. As depicted in Figure 3, it is noticeable that the HIST algorithm allows for matching a higher number of features. The second factor in terms of influence is the colour channel: RGB images and the green channel allow for matching the maximum number of features.

The results related to the factor ‘image pyramid level’ reflect the low influence deduced from the ANOVA analysis (Table 3), but it must be pointed out that resized images lead to a higher percentage of matched points. For all the three enhancement algorithms used in the experimentation, PYR levels 1 and 2 leads to a higher performance of the feature-matching algorithm. One of the reasons for this behaviour is the fact that by reducing the resolution, it is possible to find more robust features, as they are extracted from the more evident details. The images at level 3 have not shown good results due to the lack of reliable details.

Regarding the ‘image set’ factor, the matching of the images included in sets 5, 6, and 7 leads to poor results, since these are composed by oblique photographs only. The best results have

Table 3. Summary of the results of the ANOVA analysis.

Factors	Mean extracted features (1)		% matched features (2)		% oriented cameras (3)		Bundle adjustment mean re-projection error	
	F	p-value	F	p-value	F	p-value	F	p-value
EN	1065.41	0	21.07	0	91.08	0	9.17	0.0002
PYR	1457.10	0	5.93	0.005	20.61	0	13.82	0
CH	47.80	0	14.54	0	12.39	0	0.18	0.9129
SET	92.79	0	7.76	0.0001	18.64	0	6.11	0.0002

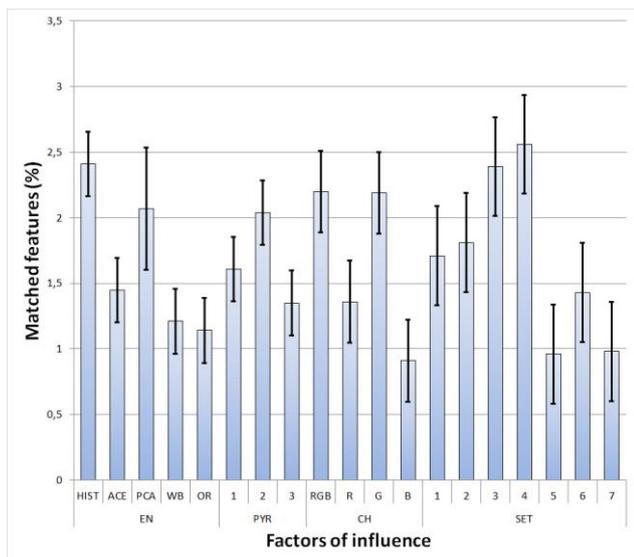


Figure 3. The mean values of the percentage of the matched features for all the influential factors.

been obtained for sets 3 and 4, which include images taken with a reduced working distance and a greater picture overlap.

4.3. Percentage of oriented cameras

The image enhancement algorithm is the most influential factor on the parameter ‘percentage of oriented cameras’ (Table 2). The performances obtained using each method are clearly better than the results obtained with the original images. For the latter, only 9 % of the cameras have been oriented, while about 43 %, 37 %, and 33 % of cameras have been successfully oriented while using the HIST, ACE, and PCA enhancement methods, respectively.

The second most influential factor is image resolution. By analysing the results in Figure 4 and the outcomes of the Tukey post hoc test, it is noticeable that there are no statistically relevant differences between the values for the first and second levels of the image pyramid. This means that it is possible to obtain the maximum number of oriented cameras with a lower resolution, also saving computational time.

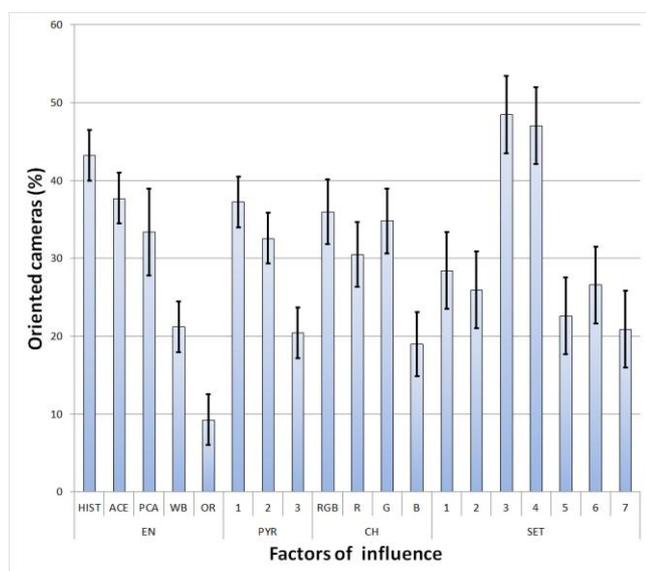


Figure 4. Mean values of the percentage of oriented cameras for all the influential factors.

4.4. Bundle adjustment mean re-projection error

The ANOVA results (Table 3) show that the most influential factor for the parameter ‘mean re-projection error’ is image resolution. If we consider the pixel size of the subsampled images and the mean distance from the subject of 2.5 m, the first and second pyramid levels have demonstrated errors measured on the ground of 0.29 and 0.38 mm, respectively.

By analysing the results presented in Table 2 and depicted in Figure 5, it is noticeable that a higher error has been measured on the image sets containing oblique photographs only. The presence of the blue background in almost all the pictures reduces the accuracy of the bundle adjustment process.

Furthermore, the ANOVA analysis results (Table 3) reveal that there is not a statistically significant difference among the different colour channels (CH).

4.5. Discussion

The statistical analysis allowed for choosing the best combination of factors that should be used to perform the 3D reconstruction of the site. The accuracy of the SfM (Structure from Motion) procedure (namely the mean re-projection error) is mainly related to the camera network orientation. Sets 1, 2 and 3 are characterised by convergent images with a high overlap, forming a more robust network. Moreover, as reported in the previous section, both levels 1 and 2 led to errors below the acceptable value of 0.5 mm. For these reasons, it is possible to save computational time using subsampled images, which also result in a higher percentage of matched features. In fact, in the same conditions and varying only the image resolution, the average reconstruction time for the datasets taken into account in the study presents a saving time of 81 % for PYR2 compared to PYR1, and 92 % for PYR3 compared to PYR1.

HIST and ACE methods considerably increase the performance of image matching. The first method returns better results in terms of the percentage of oriented cameras and matched features, increasing the performance by about 150 % and 50 %, respectively. The analysis of the effects of white balance correction has shown good results for the parameters ‘mean number of extracted features’ and ‘percentage of oriented cameras’. In particular, white balance corrected images have shown a higher number of extracted features and better performances compared with PCA method. The images

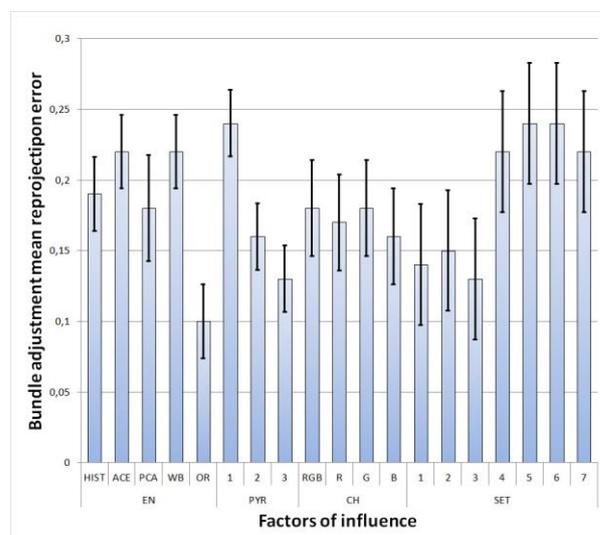


Figure 5. Bundle adjustment mean re-projection error for all the influential factors.

obtained using the custom white balance adjustment lead to less stable results, since the correction is performed at the beginning of the survey.

The results in terms of the number of oriented cameras have shown similar values for the first and second image pyramid levels. Concerning the re-projection error, as described in section 4.4, in this case the best choice is also to use a lower resolution in order to save computational time without affecting reconstruction accuracy. Regarding the colour channel, the data has demonstrated the better performance of the RGB images, particularly the green channel, which outperforms the results for other channels in terms of matched features and oriented images.

Considering the results related to the factor 'image set', it is evident that the highest number of features is extracted from the sequences of images that have been taken using a standard aerial photography layout, with a reduced distance to the subject. On the contrary, oblique pictures, where the presence of the blue background is more evident, returned poor results in terms of percentage of matched features and percentage of oriented cameras.

5. RESULTS

The 3D reconstruction pipeline starts by performing the orientation of the whole dataset of 722 pictures by means of the Bundler software [25]. In the first instance, the 3D reconstruction has been performed on the dataset composed by the original images. The image orientation process failed to orient all the pictures in a single block: the dataset has been divided into two non-overlapping groups, the north and south parts, which have been reconstructed separately. In particular, 384 images have been oriented for the north block and 116 for the south block. This failure is mainly due to the sandy seabed present in the central part of the room, which makes the extraction and matching of features difficult, as a consequence of the low contrast. Furthermore, the lack of overlapping areas in the reconstructed model prevented the alignment of the two blocks.

The results of the statistical analysis allow for choosing the best combination of factors to be used in order to improve the reconstruction process, represented by RGB images resized to

25 % (second pyramid level) enhanced with the HIST method. The enhanced dataset has been processed with Bundler, and a subset of 533 images related to the whole area has been aligned, allowing for the generation of a complete 3D point cloud without the need to register the different meshes (Figure 6). This result shows that colour correction considerably improves the matching process.

The data returned by Bundler (camera positions and camera parameters computed by a self-calibration procedure) and the undistorted images have been processed with PMVS2 (Patch Based Multi-View Stereo) [26] in order to create a dense cloud of about 10 million points related to the whole site.

This algorithm estimates the surface orientation while enforcing the local photometric consistency, which is important for obtaining accurate models for low-textured objects or for images affected by blur due to the turbidity in the underwater environment. Furthermore, PMVS2 automatically rejects moving objects such as fishes and algae. The dense stereo matching algorithm implemented in PMVS2 receives, as inputs, an undistorted set of images and the 3×4 camera projection matrix computed by Bundler. The output is a coloured dense 3D point cloud. The PMVS2 parameters used to fine-tune the 3D reconstruction are the size of the correlation window and the level in the internal image pyramid used for the computation. In our experiment, a fixed correlation window with a size of 7×7 pixels was adopted, while the image resolution (image pyramid level) was chosen according to the results obtained through the variational analysis. Moreover, image triplets instead of pairs were used to increase the robustness of the reconstruction.

The 3D point cloud has been elaborated with Meshlab tools. The first operation was the manual selection and deletion of unwanted areas and outliers caused by the presence of underwater flora and fauna and bad visibility conditions. Then, a watertight surface with about 25 millions of triangles (Figure 7) was obtained through the Poisson Surface Reconstruction algorithm.

The resulting surface has been subsequently decimated in a mesh of 6.5 million triangles and 3 million points in order to be handled more efficiently without losing details. Since the camera orientation procedure has been carried out with an unknown scale factor, it is necessary to scale the model by selecting two points with a known distance. In this experimentation, a scale bar has been placed in the scene and reconstructed in order to evaluate the scale factor.

The last step consists in the application of the texture on the 3D surface. Colour information can be extracted directly from the coloured point cloud, but this method does not allow for the

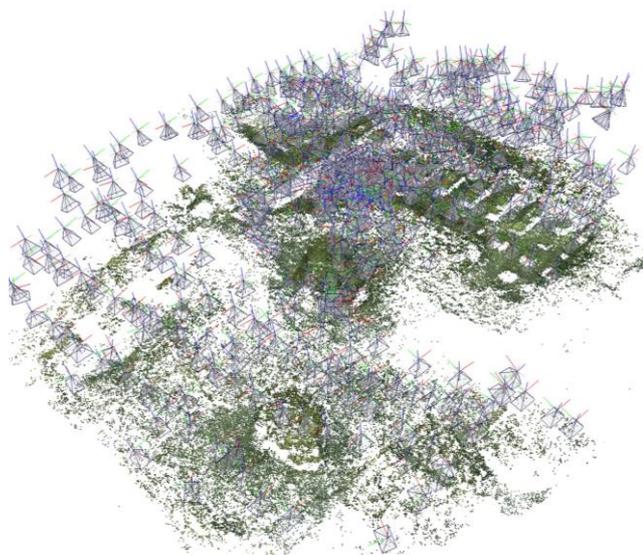


Figure 6. Results of the camera orientation process (enhanced pictures with HIST method): sparse point cloud and 533 oriented cameras.

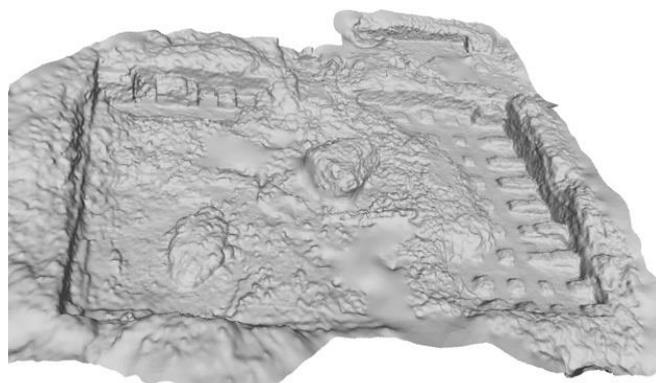


Figure 7. Reconstructed surface.

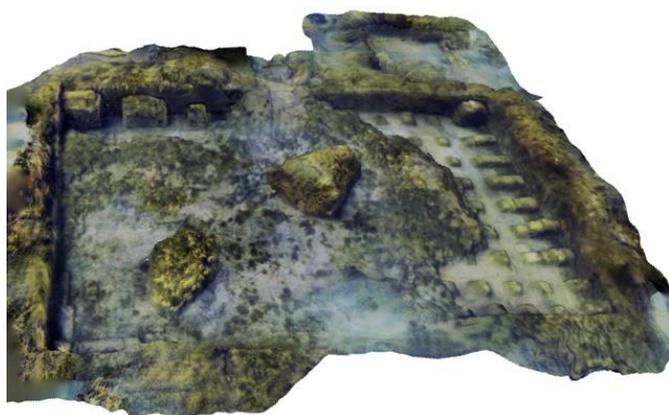


Figure 8. Final textured 3D model.

creation of a high-quality texture, because its resolution depends on the point cloud density. Moreover, since the enhancement procedure is often performed to improve the feature extraction process (by increasing the contrast without taking into account the fidelity of the colours – usually single-component or greyscale images are used), the colour information stored in the pixels cannot be used. Since the camera positions are known, the texture mapping has been carried out by means the projection and blending of high-resolution images directly on the 3D surface. In particular, an image subset has been selected because the averaging among neighbourhood values during the blending on the images works better if a small overlapping area is present.

This subset of images has been extracted from the images enhanced with the HIST method, which also gave the best results in terms of texture quality. This is mainly due to the manual retouching step performed on a sample image and then exported to the whole dataset.

The result of this procedure is a texture with a resolution comparable to the original images (Figure 8).

6. CONCLUSIONS

This paper has presented a performance analysis, based on a DOE approach, of the main influential factors that affect 3D reconstruction quality. The performance of three different colour enhancement algorithms, ACE, PCA, and HIST, have been evaluated by using a variance analysis, including the effects of image resolution and colour channels. The results of the ANOVA analysis show that the factors EN (image enhancement method), PYR (image pyramid level), and SET (image set) are influential, with a confidence level of 95 %, for all the parameters, while the results related to the factor CH (colour channel) have shown a limited influence, since each enhancement method performs mixing operations among the channels.

The ANOVA data allowed for choosing the best combination of factors to optimise the SfM bundle adjustment mean re-projection error, the number of extracted features, oriented cameras, and matched features, also taking the processing time into account. More precisely, the best combination is characterised by RGB images resized to 25 % and enhanced with the HIST method, which returns more stable results.

By using the results of the statistical analysis to correct and process the underwater images, it has been possible to align an unordered sequence of more than 500 images belonging to the entire site. On the contrary, the original images could not be used to align all the cameras. Moreover, the corrected images allowed for creating a model mapped with a high-quality texture,

comparable with original images in terms of resolution and with a fair colour balance, since the whole dataset shares the same colour statistics.

Even if these techniques have been used in other works related to underwater archaeology, this experiment represents a significant case study for verifying their robustness in the presence of strong turbidity and poor environmental conditions, providing useful guidelines for an accurate modelling of a submerged site.

ACKNOWLEDGEMENT

This work has been supported by the iMARECulture project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 727153.

REFERENCES

- [1] UNESCO, Convention on the Protection of the Underwater Cultural Heritage, 2 November 2001, <http://www.unesco.org>.
- [2] G. Telem, S. Filin, Photogrammetric modelling of underwater environments, *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 5 (2010), pp. 433-444.
- [3] F. Menna, P. Agrafiotis, A. Georgopoulos, State of the art and applications in archaeological underwater 3D recording and mapping, *Journal of Cultural Heritage* 33 (2018) pp. 231-248.
- [4] F. Bruno, A. Lagudi, L. Barbieri, D. Rizzo, M. Muzzupappa, L. De Napoli, Augmented Reality visualization of scene depth for aiding ROV pilots in underwater manipulation, *Ocean Engineering* 168C (2018) pp. 140-154.
- [5] D. C. Montgomery, *Design and Analysis of Experiments*, John Wiley & Sons, New York, 2017, ISBN: 978-1-119-11347-8.
- [6] iMARECulture, <http://www.iMARECulture.eu>
- [7] F. Bruno, A. Lagudi, G. Ritacco, J. Cejka, P. Kouril, F. Liarokapis, P. Agrafiotis, D. Skarlatos, O. Philpin-Briscoe, E.C. Poullis, Development and integration of digital technologies addressed to raise awareness and access to European underwater cultural heritage, An Overview of the H2020 iMARECulture Project, Proc. of the MTS/IEEE Conference Oceans'17, Aberdeen, UK, 19-22 June, 2017.
- [8] D. Skarlatos, P. Agrafiotis, T. Balogh, F. Bruno, F. Castro, B.D. Petriaggi, S. Demesticha, A. Doulamis, P. Drap, A. Georgopoulos, 'Project iMARECulture: Advanced VR, iMmersive serious games and augmented reality as tools to raise awareness and access to European underwater cultural heritage', Proc. of the International Conference on Cultural Heritage, Nicosia, Cyprus, 1-5 November, 2016.
- [9] A. Mahiddine, J. Seinturier, D. Peloso, J. M. Boi, P. Drap, D. Merad, Underwater image pre-processing for automated photogrammetry in high turbidity water, *VSM2012*, 2012, pp. 189-194.
- [10] R. Schettini, S. Corchs, Imaging for underwater archaeology, *American Journal of Field Archaeology* 27, 3 (2000), pp. 319-328.
- [11] Y. Y. Schechner, N. Karpel, Recovery of underwater visibility and structure by polarization analysis, *IEEE Journal of Oceanic Engineering*, 2005, 30(3), pp. 570-587.
- [12] E. Trucco, A.T. Olmos-Antillon, Self-tuning underwater image restoration, *IEEE Journal of Oceanic Engineering* 31, 2 (2006) pp. 511-519.
- [13] A. Rizzi, C. Gatta, From Retinex to Automatic Color Equalization: issues in developing a new algorithm for unsupervised color equalization, *Journal of Electronic Imaging* 13 (2004) pp.75-84.
- [14] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, A. Rizzi, Underwater color constancy: enhancement of automatic live fish recognition, Proc. of the 16th Annual Symposium on Electronic Imaging, 2003, United States, 5293, pp. 157-169.

- [15] F. Petit, Traitement et analyse d'images couleur sous-marines: modèles physiques et représentation quaternionique, Doctorat, Sciences et Ingénierie pour l'Information, Poitier, 2010.
- [16] S. Bazeille, I. Quidu, L. Jaulin, J. P. Malkasse, Automatic underwater image pre-processing, CMM'06 - Caracterisation Du Milieu Marin, 2006.
- [17] K. Iqbal, R. Abdul Salam, A. Osman, A. Z. Talib, Underwater image enhancement using an integrated colour model, IAENG International Journal of Computer Science 32, 2 (2007) pp.239-244.
- [18] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (SURF), Comput. Vis. Image Underst. 110 (2008) pp.346-359.
- [19] R. Kalia, K.-D. Lee, B.V.R. Samir, S.-K. Je, W.-G. Oh, An analysis of the effect of different image pre-processing techniques on the performance of SURF: Speeded Up Robust Feature, Proc. of the 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), Ulsan, South Korea, 9-11 February 2011, pp.1-6.
- [20] M. Mangeruga, F. Bruno, M. Cozza, P. Agrafiotis, D. Skarlatos, Guidelines for underwater image enhancement based on benchmarking of different methods, Remote Sensing 10, 10 (2018) 1652, pp.1-27.
- [21] A. Tonazzini, E. Salerno, M. Mochi, L. Bedini, Blind source separation techniques for detecting hidden texts and textures in document images, Image Analysis and Recognition Lecture Notes in Computer Science 3212 (2004) pp. 241-248.
- [22] D. P. Mitchell, A. N. Netravali, Reconstruction filters in computer graphics, Computer Graphics 22, 4 (1988) pp. 221-228.
- [23] D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60, 2 (2004) pp. 91-110.
- [24] Z. Zhang, A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 11 (2000), pp.1330-1334.
- [25] Bundler software, <http://www.cs.cornell.edu/~snavely/bundler>
- [26] Y. Furukawa, J. Ponce, Accurate, dense, and robust multi-view stereopsis, IEEE Transactions on Pattern Analysis and Machine Intelligence 32, 8 (2010), pp. 1362-1376.