

3D RECONSTRUCTION AND MESH OPTIMIZATION OF UNDERWATER SPACES FOR VIRTUAL REALITY

A. El Saer¹, C. Stentoumis¹, I. Kalisperakis¹, L. Grammatikopoulos², P. Nomikou³, O. Vlasopoulos^{1,3}

¹ up2metric P.C., 11521, Athens, Greece – (andreas.saer, christos, ilias)@up2metric.com

² Dept. of Surveying and Geoinformatics Engineering, University of West Attica, 12243, Egaleo, Greece - lazaros@uniwa.gr

³ Dept. of Geology and Geoenvironment, National and Kapodistrian University of Athens, 15784, Athens, Greece -
evinom@geol.uoa.gr

Commission II, WG II/9

KEY WORDS: 3D Reconstruction, VR, Unity, Underwater, Refractive SfM, Image Processing, Mesh Optimization.

ABSTRACT:

In this contribution, we propose a versatile image-based methodology for 3D reconstructing underwater scenes of high fidelity and integrating them into a virtual reality environment. Typically, underwater images suffer from colour degradation (blueish images) due to the propagation of light through water, which is a more absorbing medium than air, as well as the scattering of light on suspended particles. Other factors, such as artificial lights, also, diminish the quality of images and, thus, the quality of the image-based 3D reconstruction. Moreover, degraded images have a direct impact on the user perception of the virtual environment, due to geometric and visual degenerations. Here, it is argued that these can be mitigated by image pre-processing algorithms and specialized filters. The impact of different filtering techniques on images is evaluated, in order to eliminate colour degradation and mismatches in the image sequences. The methodology in this work consists of five sequential pre-processes; *saturation enhancement*, *haze reduction*, and *Rayleigh distribution adaptation*, to de-haze the images, *global histogram matching* to minimize differences among images of the dataset, and *image sharpening* to strengthen the edges of the scene. The 3D reconstruction of the models is based on open-source structure-from-motion software. The models are optimized for virtual reality through *mesh simplification*, *physically based rendering* texture maps baking, and *level-of-details*. The results of the proposed methodology are qualitatively evaluated on image datasets captured in the seabed of Santorini island in Greece, using a ROV platform.

1. INTRODUCTION

In the past few years, there has been a massive adaptation of the Virtual Reality (VR) technology on a variety of application fields, but even more in the entertainment and cultural heritage domains. In particular, the radical hardware advancements made it possible for the VR applications to run efficiently and with high quality graphics. It is, now, widely accepted that VR offers a much better understanding of a represented scene and creates a higher and more immersive user experience than conventional 3D environments. Latest advancements in human-machine interaction have also influenced the way people interact with a virtual world. VR offers to the public, but also to the scientific community, an unprecedented way of accessing environments that are typically inaccessible to common people without specialized equipment and, at some cases, huge budget. *Underwater environments are a significant user scenario for VR technologies to demonstrate the potentials of immersive experiences* since the underwater world hosts an incredible treasure of cultural and marine biodiversity. Besides, seas provide a broad range of economic and marine engineering activities, such as pipe and drilling equipment construction, operation and inspection, telecommunications, and fish farms. Hence, underwater VR can also assist in interactive and immersive education and training of public and experts.

Underwater VR relies on imaging as a passive, non-invasive, non-contact and cheap technique to capture reality. The exploitation of images from the underwater environment poses significant difficulties comparing to the images acquired in typical in-air conditions. Underwater imagery shares a series of common quality issues due to absorption and scattering of light in the water. The water absorbs light as a function of the distance

from the surface and the wavelength of the light spectrum. As an immediate result is the *low visibility* range, which is limited at 20m in clear water and much less in water that contains particles, or it is disturbed. Moreover, as the depth increases, the larger wavelengths are absorbed faster, thus the red colour (780 – 622nm) disappears faster, in contrast to the blue colour (492 – 455nm) that can penetrate at larger depths (Hitam et al., 2013); the difference in light absorption according to the wavelength leads to *colour (spectral) distortion* (the “bluish” effect) noticed in underwater images (Figure 1). Underwater images also present *noise*, *blurriness*, *low contrast*, and *bright spots* owed to sun reflectance. In some cases, artificial light is used in order to increase the diminished visibility, which in return causes a *non-uniform lighting* of the scene at a much greater extent than at grabbing flashed images in the air, and *scattering* on particles in the water. The abovementioned issues degrade the visual quality of underwater images and, thus, the quality of image-based 3D reconstruction and the quality of the VR experience. Processing underwater images is an increasingly active domain, as it is related to ocean exploration by remotely operated underwater vehicles (ROVs) and autonomous underwater vehicles (AUVs). Overall, presenting real underwater spaces through immersive and interactive VR experiences relies on the advances in several research domains.

3D reconstruction of underwater spaces suffers from geometric and visual inaccuracy; hence, this work aims to address these issues by proposing a methodology that improves the overall image quality of a dataset before reconstruction. The image pre-processing algorithm proposed in this work consists of five individual enhancements: *saturation enhancement*, *haze reduction*, *Rayleigh distribution adaptation*, *global histogram matching*, and *image sharpening*. The methodology proposed in this work

was evaluated on custom datasets from Santorini island, acquired by a ROV provided by NKUA. The 1st dataset consists of 352 images depicting a shipwreck, whereas the 2nd one is a collection of 492 images presenting a small part of the dykes of Santorini volcano island -i.e. dykes are a fluid driven (magma driven) extension fracture.

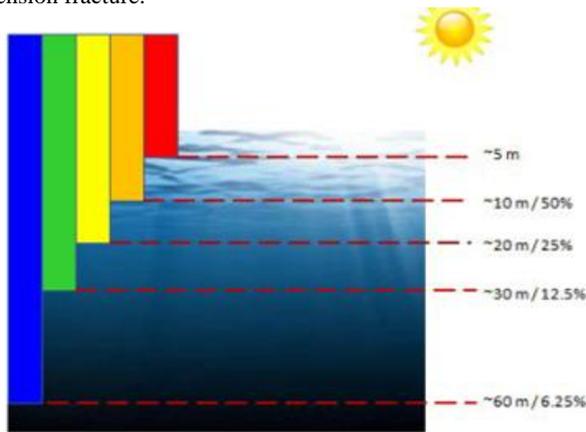


Figure 1. Infiltration of light in the water. [after Hitam et al. 2013]

2. PREVIOUS WORK

2.1 Enhancement of Underwater images

Corchs and Schettini (2010) published a thorough review on methods that aim at enhancing the resolution, contrast and range visibility on underwater images. They categorized the methods that restore underwater images as enhancement and restoration. Image restoration methods are based on complex physical and mathematical *image degradation* models and the scene depth. On the other hand, *image enhancement methods*, aim at the visual amelioration of the underwater images comparing to in-air images. Lu et al. (2017) published an updated review of underwater image processing in which the physics around underwater imaging are described, along with the difficulties of defining a degradation model. Moreover, passive and active methods of image acquisition are discussed, as well as typical methods of image restoration and the categorization of restoration methods in hardware-based and software-based, depending on the acquisition method. Image quality assessment methods are also discussed. Li et al. (2020) propose a newly acquired benchmark dataset (UIEB) under which state-of-the-art algorithms are reviewed.

Hitam et al. (2013) propose an image enhancement method that combines the advantages of RGB and HSV colour spaces; the image in each colour space is processed via adaptive histogram equalization limited by a threshold for contrast enhancement. The results are qualitatively evaluated on their own acquired images. A similar work utilizes the Rayleigh distribution to modify by modifying the histogram in RGB and HSV colour spaces and combine the results (Ghani and Isa, 2014). Qiao et al. (2017) built on the previous work to improve contrast and added wavelet transformation for better denoising of images of underwater cucumbers. The efficiency of this method was qualitatively and quantitatively evaluated on 120 greyscale underwater images. Emberton et al. (2018) presented a method that achieves significant improving (dehazing) via segmenting the water areas in the underwater image. Moreover, some methods explicitly estimate the medium transmission (Drewns-Jr et al., 2013), based on the generic work on Dark Channel Prior for estimating haze in outdoor scenes (He et al., 2011). This approach relies on the

statistics of outdoor images via the pixels of exceptionally low colour values to compute a haze prior that allows the estimation of the thickness of haze in new images and thus image restoration. Concurrently, these approaches estimate an image depth map.

The advance in deep neural networks and data-driven approaches in the latest years has also pushed the advance in image restoration methods. An interesting approach proposed the fusion generative adversarial network to correct colour degradation (Li et al., 2019), while “Water-Net” proposed a generic convolutional neural network to enhance underwater images (Li et al., 2020). The depth estimation can also be important to accurately determine a transmission model for de-hazing an image and can be estimated by a neural network (Ding et al., 2017). (Li et al., 2015) proposed another method for dehazing images including underwater images. Their method exploits an MRF framework to simultaneously optimize for depth and cleared colour values in image sequences. The energy function is built on photo-consistency, fog scattering and smoothness term adapted to underwater conditions.

2.2 3D Reconstruction of Underwater Scenes

In underwater imaging, the path of the light ray through media with different refractive indices (air/glass/water) deviates significantly from the straight line. Especially in the case where an imaging system with a flat refractive interface is utilized (housing with flat glass port), the adoption of the perspective camera model (also known as SVP – Single View Point model), cannot cope with the refraction error, leading to inaccuracies in calibration and 3d reconstruction (Treibitz et al., 2008). This incompatibility arises mainly from the fact that refractive distortion is related to the distance of the point from the imaging system and particularly from the housing interface. In fact, it can be illustrated in Figure 2, where object point P is projected to the image point p_1 following the actual path of the light-ray. The incident ray is refracted according to Snell’s law, at point K due to the different density of the media (air/water) and, afterwards, it defines point p_1 on the image plane. The adoption of the perspective camera model defines a new image point p_2 that deviates substantially from its actual position p_1 . The establishment of a new object point P' at a greater distance from the housing interface along the incident ray PK , further distorts image point location (point p_3) that confirms the inadequacy of the central projection model.

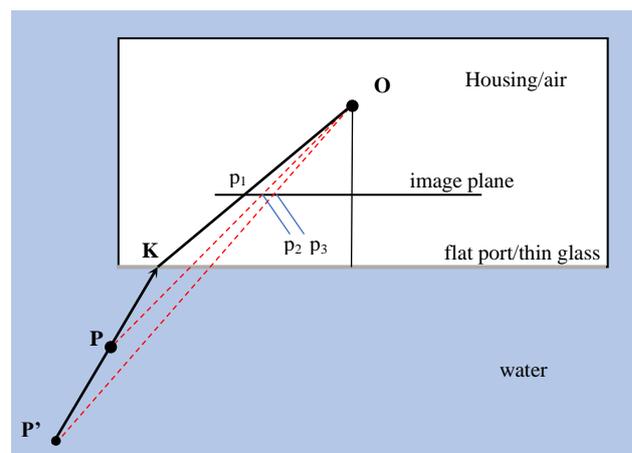


Figure 2. Refracted light-ray path and invalidity of Single View Model.

In this context, Treibitz et al. (2008) have presented the caustic, a geometrical interpretation of the refractive geometry that represents the locus of all viewpoints, which can be derived in closed form as the surface tangential to the bundle of refracted rays. In a similar manner, following a geometrically driven approach, Telem and Filin (2010) proposed a method that can handle light propagation in the different media by explicitly interpreting the ray path. Their approach adapts collinearity model to underwater imagery by the integration of Snell's law of refraction, accommodating two distinct camera configuration cases: the optical axis being strictly perpendicular to the housing interface and the optical axis has a random and unknown orientation. Agrawal et al. (2012) first modelled the geometry of flat refraction systems using an axial camera. By exploiting two additional constraints, they achieved accurate calibration and pose estimation of a monocular camera employing a linear initialization at first, and then a final non-linear optimization. Jordt et al. (2016) used the results of Agrawal et al. (2012) and implemented a complete scalable 3D reconstruction framework for underwater acquisition, which combines an RSfM (Refractive Structure from Motion) and the Refractive Plane Swipe algorithm.

On the other hand, the exploitation of a camera housing with a dome port can significantly alleviate the previous geometrical implications. Nevertheless, dome port devices are not as convenient to use as flat ones, while, at the same time, exhibit severe field curvature (Chadebecq et al., 2019). However, considering a limited range of object distances and for moderate accuracy requirements, the approximation of the perspective camera model (with lens distortion) is proved to be sufficient for many underwater 3D reconstruction applications.

2.3 Evaluation and Datasets

Public datasets and evaluation platforms play an essential role in developing algorithms. The ground truth data for underwater images are almost impossible to get, thus, image datasets that contain original and enhanced images are based on subjective relative qualitative assessment (Chen et al., 2014). Li et al. (2019) published the U45 image dataset for developing and evaluating algorithms to restore the degraded images of underwater scenes. The dataset consists of image extracted from other public datasets and it is split to green, blue and haze categories. Moreover, the images are restored via seven algorithms as a baseline. Data-driven approaches can exploit the large dataset of Liu and Chen (2019), which is based on RGB-D captured indoor images transformed via degradation models. Ground truth data are also relying on the assessment of a group of people. The underwater image enhancement benchmark (UIEB) consists of large-scale real-world images with reference images created by top algorithms and assessed by fifty people (Li et al., 2020). Since UIEB includes 890 images, it is adequate for training deep networks for image enhancement. On a different benchmarking approach, Berman et al. (2020) published a newly collected (via SLR cameras) underwater stereo-images dataset to assist the research on image restoration. The images depict colour charts in multiple distances from cameras and ground truth information from a variety of dive sites, 57 sites in total. The stereo-pairs are calibrated, although the refraction is disregarded, and the depth of the scene is estimated to assist with colour restoration. Another different approach is the one proposed by Li et al. (2018) to create synthetic underwater images from real in-air images, so that the ground truth exists.

Typical assessment measures include mean square error, peak signal to noise ratio and entropy, although others have been proposed, such as a linear combination of standard deviation of

chroma, mean of saturation and lightness contrast (Yang and Sowmya, 2015). It is noticeable that quantitative measures can contradict qualitative evaluation of underwater image restoration (Emberton et al., 2018).

Beyond the image datasets for image restoration, Ferrera et al. (2019) published a dataset for simultaneous localization and mapping, which can be exploited for evaluation of 3D reconstruction purposes. The data are captured via ROV and depict industrial, archaeological and physical scenes. Similarly, Mallios et al. (2017) published a dataset of images and sensor data acquired via a ROV under scuba diver guidance, which depicts a complex of caves and can be useful for SLAM. The provided data include raw RGB images, inertial and sonar measurements from a variety of navigation and perception sensors, as well as ground truth points for relative accuracy estimation, and cameras' calibration estimation.

2.4 VR Experiences of Underwater Scenes

Although the inaccessible underwater environments pose an ideal scenario of exploiting VR potentials, the work in the field has not matured, yet. Research in the field of underwater VR includes interactive experiences based on 360° videos, 3D reconstructed real scenes, and even fully immersive experiences via underwater HMI equipment and trackers (Costa et al., 2017). (Osone et al., 2017) evaluate the underwater VR experience of a head-mounted display (HMD). Lately, companies such as ("Ballast VR," n.d.), have offered mass underwater VR experiences, whereas equipment for capturing underwater scenes for VR have been made available to the public ("Vuze," 2020). "TheBlu: An Underwater VR Experience," (2017) expedition invited the public to experience swimming with whales in the Natural History Museum of Los Angeles.

Underwater *cultural heritage* is an important application field for VR, as it enables the public to taste, but also the experts to study, the hidden treasures of the oceans. In McCarthy and Martin (2019), a VR diving experience for maritime archaeology is presented and the writers argue that the implemented 2.5D approach of a pre-set navigation is more appropriate than full 3D interactive experience for the public. Bruno et al. (2016b, 2016a, 2016c) proposed a methodology for surveying underwater archaeological sites and presenting them interactively in VR. VR for *education* is also an emerging topic; Calvi et al. (2018) developed a game played in oceanic environment hoping to raise awareness on the delicate maritime environments. The VR environment was constructed via a mixed approach combining the real 3D model as reconstructed from images and artificial 3D models designed from scratch. Furthermore, VR can support *training* for operating in harsh environments, which pose a threat for human safety, such as diving in the ocean. Jain et al. (2016) developed a VR system for diving that arouses a variety of senses compared to bare HMD and also imitates the neutral buoyancy of the ocean. Hatsushika et al. (2019a, 2019b) developed a tether head mounted display (HMD) for underwater usage in order to train scuba divers in realistic conditions and supported the VR experience with artificial 3D scenes and content. Zhang et al. (2019) developed a VR simulator for training operators to submarine vehicles. In Sinnott et al. (2019), a typical scuba diving mask was converted to a head-mounted display for underwater experiences, and although the goal was to study the effects of neutral buoyancy to VR experiences for NASA, the proposed approach is a step towards more immersive underwater experiences. Fernandez et al. (2019) and Amores et al. (2019) have pioneered on an interactive, immersive VR experience that it is based on videos and 3D of oceanic life and a collection of

physiological, chemical and EEG signals, in order to influence the relaxation state of a user.

3. PRE-PROCESSING UNDERWATER IMAGES

Typically, the water as a medium for travelling light causes to images the following problems: *greenish* and *blueish* effect, *haze* effect, *low contrast*, *noise* and *lack of vividness*. Additionally, 3D reconstruction from underwater filtered images adds certain limitations to the image filtering techniques and enforces global image adjustment. Moreover, the water reduces the visibility range and the attempt to increase it with artificial lights causes bright spots due to intense reflections on particles and uneven lighting. The algorithm proposed in this work consists of five individual enhancements: *saturation enhancement*, *haze reduction*, *Rayleigh distribution adaptation*, *global histogram matching*, and *image sharpening* (Figure 3).

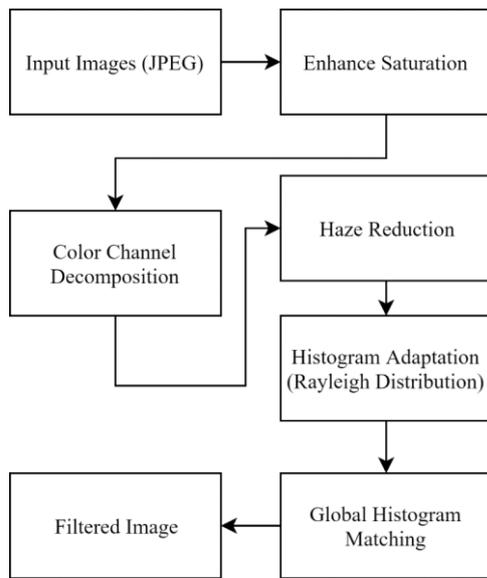


Figure 3. Flow chart of pre-processing underwater images for 3D reconstruction purposes.

3.1 Saturation Enhancement

There are three main colourmap representations of the RGB domain, the HSL, HSV, HSB. In this work, we used the HSV colourmap to isolate the saturation parameter (Hitam et al., 2013). Saturation is the intensity of a colour and how dominant can be, which is particularly important on underwater images. The selected colourmap transformation between the HSV and the RGB domain was proposed by Smith (1999). Initially, the normalized $[-0,1]$ - image is converted via the RGB to the HSV hex cone colourmap model to enhance the saturation value S . The reason for this is that underwater images also lack in vividness. Experiments showed that scaling the S value by 1.5 is adequate; it is important not to over-saturate the images. In some cases, this scaling might need tuning after visual assessment of the results on a sample of the dataset. After changing S , the image is remapped to the RGB domain.

3.2 Haze Reduction

After saturation enhancement, the image is “de-hazed” via the dark channel prior, after He et al. (2011) to remove the image blurriness. As observed in Figure 5, the enhancement in underwater images is exceptional. The resampled image J is

$$J(p) = \frac{I(p) - A}{r(p)} + A \quad (1)$$

where I is the observed intensity, A the atmospheric light, and r the transmission. If at least one colour channel of an RGB image has some pixels, whose intensity values are close to zero, then the dark channel J^{dark} is

$$J^{\text{dark}}(p) = \min_{q \in \Omega(p)} \left(\min_{c \in \{r,g,b\}} J^c(q) \right) \quad (2)$$

where J is the scene’s radiance, J^c is a color channel and $\Omega(p)$ is the local neighbourhood of a pixel p .

After computing the dark channel prior, the colour imbalances are minimized by white balancing the image following Park et al. (2014). To estimate the atmospheric light, it is assumed that it diffuses to a larger part of an image and its intensity peaks are in a smaller part. Initially, the RGB corresponding grayscale image is subdivided into blocks of 30×30 size and the pixel values of each block are replaced with their minimum. Then, a quad-tree subdivision of pre-specified repetitions on the grayscale image is enforced. The atmospheric light is the vector which minimizes the Euclidean norm,

$$\|(r_p, g_p, b_p) - (1,1,1)\| \quad (3)$$

where p is the selected pixel. Afterwards, the transmission $r(p)$, which offers a nice contrast to the dehazed image, is computed from an objective function $f_{\text{objective}}$ which relies on, f_{entropy} and f_{fidelity} .

$$\begin{cases} f_{\text{entropy}}(r) = - \sum_{i=0}^{255} \frac{h_i(r)}{N} \log \frac{h_i(r)}{N} \\ f_{\text{fidelity}}(r) = \min_{c \in \{r,g,b\}} s^c(r) \\ s^c(r) = \frac{1}{N} \sum_{p=1}^N \delta(p), \delta(p) = \begin{cases} 1, & 0 \leq J^c(p) \leq 255 \\ 0, & \text{otherwise} \end{cases} \end{cases} \quad (4)$$

where N is the number of the pixels in the image, $h_i(r)$ is the number of pixels that have intensity i in the greyscale image J . The transmission value r is computed, which acts as local optimum for each subdivided block r_k^{block} ,

$$r_k^{\text{block}} = \arg \max_{r \in \{0.01 \leq r \leq 1\}} f_{\text{objective}}(r) \quad (5)$$

At the end, a refinement of the transmission variable is needed for the dehazing to perform seamlessly along the blocks.

3.3 Rayleigh distribution adaptation

Rayleigh distribution is a continuous probability distribution for positive values. In image processing aggregates pixel values to the middle range of the intensity level and it is basically, a bell-shaped distribution. After haze removal, some pixel values are either too dark, or too bright. Although haze reduction and high contrast typically result to a nice visualized image can lead to information loss mainly during the 3D reconstruction process. The probability distribution function (PDF) of the Rayleigh distribution is

$$\text{PDF}_{\text{Rayleigh}} = \left(\frac{x}{a^2} \right) e^{-\frac{x^2}{2a^2}} \quad (6)$$

where a is the distribution parameter and x is the input pixel value.

3.4 Global histogram matching

While filtering techniques are typically treating the images individually, the 3D reconstruction process is favoured by a consistent appearance of the scene across the image sequence used in structure-from-motion and dense matching. Underwater image sequences suffer from abrupt light changes. Global adjustments are very sensitive and only small changes can be compensated without totally distorting the histogram but for 3D reconstruction purposes, even minor corrections can add value. Initially, the coefficient variation CV for each image is estimated

$$CV(I) = \frac{\text{std}(H)}{m(H)} \quad (7)$$

where I is the image and H is the computed histogram of I . The histogram with the lowest coefficient values is selected and all the other images are transformed so that their histograms approximately match to the histogram of the lowest CV. The transformed image dataset contains minimized colour differences among all images. This step facilitates the texture generation of the model during the 3D reconstruction process.

3.5 Sharpening

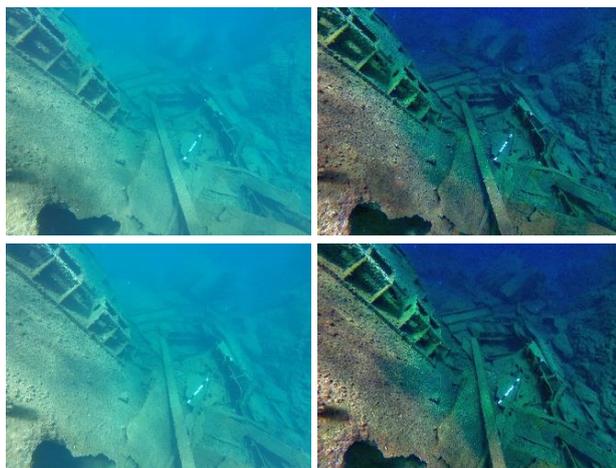
As a final enhancement, unsharp masking is applied, since, based on these experiments, the global histogram matching equalization could result in blurring some features of the image I , such as object edges.

$$I_{\text{sharpened}} = I_{\text{original}} + (I_{\text{original}} - I_{\text{blurred}})a \quad (8)$$

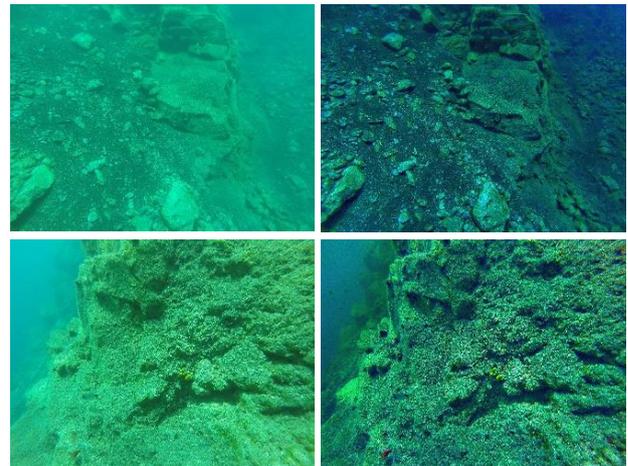
where a is the strength value of the sharpening effect.

3.6 Image Pre-processing Results

The five image enhancements proposed in this work were evaluated on two custom datasets before the 3D reconstruction and their results are presented in Figure 4. The left columns present the initial images and the right columns



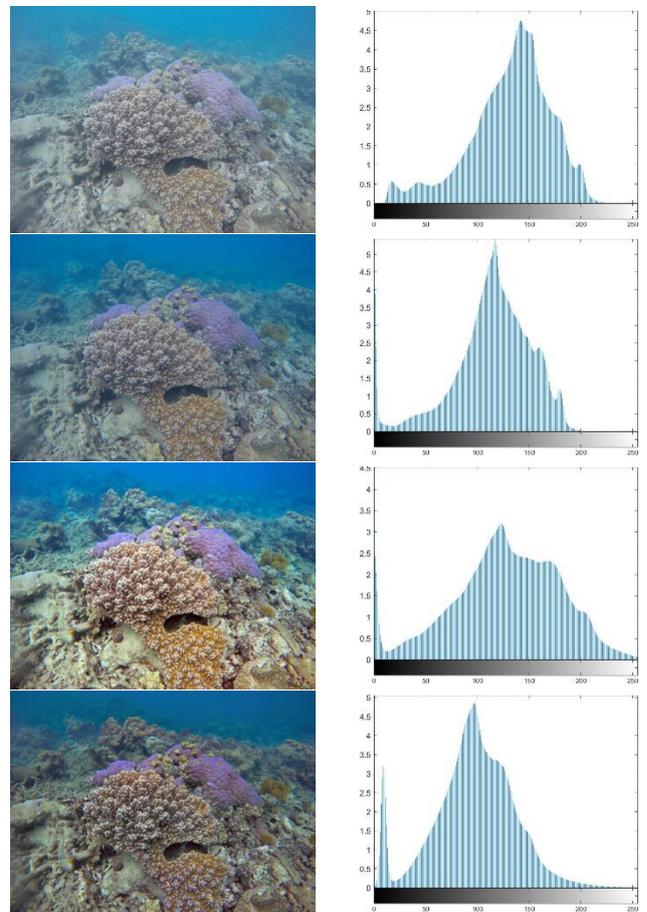
1st dataset



2nd dataset

Figure 4. Image pre-processing results. *Left*: initial images; *right*: enhanced images.

In Figure 5, the image pre-processing steps were also evaluated via the corresponding histograms. The results are cumulative, as each algorithm is applied on the result of the previous step. The image depicts a healthy coral, including both hard and soft coral, in the northern region of Australia's Great Barrier Reef [Credit: NASA/JPL-Caltech/BIOS] and it is chosen for its vivid colours to present the results of the enhancement algorithms.



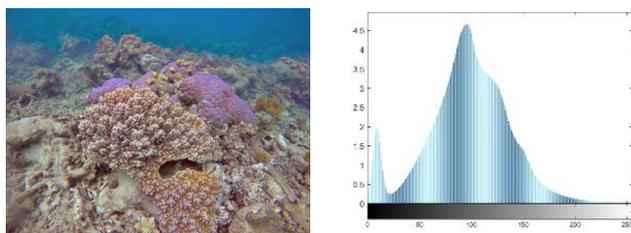
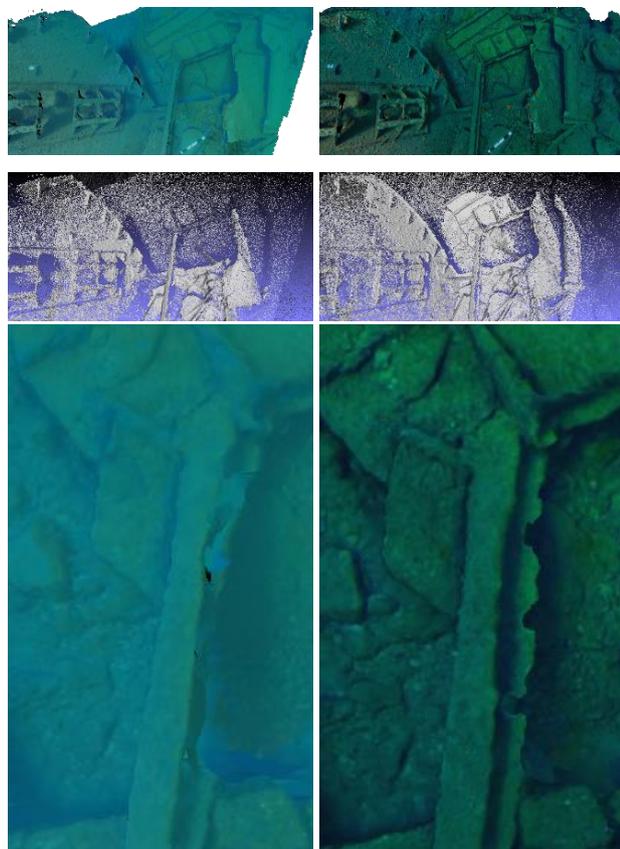


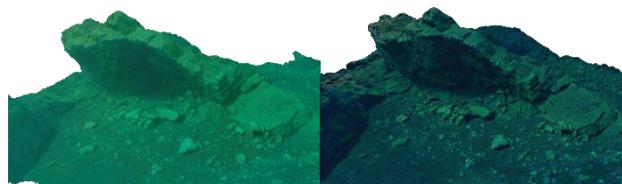
Figure 5. Results of pre-processing on the images and the histograms, i.e. saturation enhancement (S), haze reduction (RH), Rayleigh distribution adaptation (RD), global image equalization (GE), and sharpening (SH). 1st row: initial image; 2nd row: S; 3rd row: S – RH; 4th row: S – RH – RD; 5th row: S – RH – RD – GE; 6th row: S – RH – RD – GE – SH.

4. 3D RECONSTRUCTION

The 3D models were reconstructed using “Meshroom”, a prominent open source photogrammetry software. The geometric and visual quality of the reconstructed scene heavily depend on the quality of the images. Hence, the image pre-processing can significantly assist the reconstruction process, as shown in our evaluation. Holes, ambiguities and degenerated geometry that result from the degradation of the impact of the water medium on the unfiltered images were corrected, while the overall 3D model became more vivid and realistic. As presented in Figure 6, the reconstruction results from feature extraction and dense matching algorithms on pre-processed images, outperform the results produced by the initial images.



1st dataset



2nd dataset

Figure 6. 3D reconstruction results, from initial (*left*) and enhanced images (*right*). 1st dataset. 1st row: textured mesh; 2nd row: triangular mesh; 3rd row: reconstruction detail from the shipwreck, where the improved geometry and texture are presented. 2nd dataset. Details from the dykes.

Even though the SfM error was slightly higher when calibrating the filtered images, 10% more images of the dataset were included in the solution, comparing to the initial image dataset. Also, the geometry is better recovered from the processed images, as one can notice from the details and complex structures in the scenes. Nevertheless, in most VR games, cultural and tourism applications, geometric accuracy does not play a significant role. Hence, a more important effect of the pre-processing proposed in this work is that the image enhancements add realism and vividness to the reconstructed scene and significantly improved the VR experience, as presented in Figure 6.

5. 3D MODEL OPTIMIZATION FOR VR

The obtained mesh from the 3D reconstruction approach is often too complicated and complex to be handled directly in real-time Virtual Reality applications. To overcome this, a typical workflow from the computer graphics domain is adopted. This includes: i) *mesh simplification*; ii) *physically based rendering (PBR) texture maps baking*; and iii) *level-of-details (LODs) generation*. Moreover, a de-lighting effect is applied to remove the natural lights and shadows of the environment, where possible. To simplify the geometry of the reconstructed 3D mesh the “Instant Meshes” tool (Jakob et al., 2015) is employed. The original 3D surface is remeshed into an isotropic triangular or quad mesh via a local based approach that optimizes both the edge orientations and vertex positions in the output mesh (Figure 7). The first step computes an orientation field, i.e., a set of directions that the edges of the simplified mesh should align with. The second step computes a local *uv* parameterization, whose gradient is aligned with the orientation field and which is discontinuous over edges. Finally, a 3D triangular, or quad, mesh is extracted from the two fields.

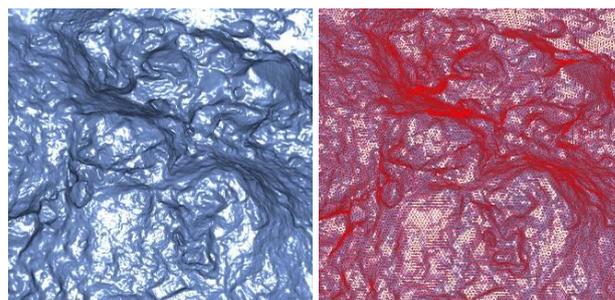


Figure 7. Mesh simplification. *Left*: original mesh; *right*: simplified mesh.

The photo-texture of the original 3D mesh model is then applied to the simplified mesh by typical *uv* unwrapping and

interpolation techniques. Alternatively, photo-texture can be estimated from the original oriented images. For better visualization in VR, PBR is adopted. Normal maps, height maps and ambient occlusion maps for the simplified 3D mesh are estimated from the more complex geometry of the original mesh using “xNormal” tool (Figure 8).

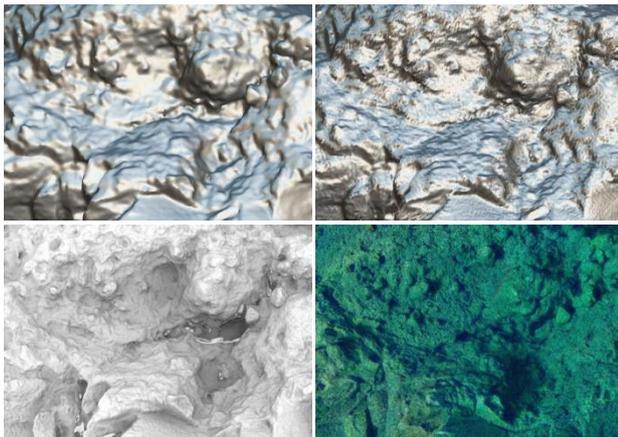


Figure 8. PBR texture maps. *Upper-left*: visualizations of the simplified mesh geometry without normal map; *upper-right*: with normal map; *lower-left*: ambient occlusion texture; and *lower-right*: final PBR render.

Finally, “simplygon” software was used for generating LODs to ease the rendering process in the VR software (Unity 3D) (Figure 9). The latter is essential to achieve real-time performance for the exceptionally large meshes that are typical from images in the cases where high fidelity scenes needs to be modelled.

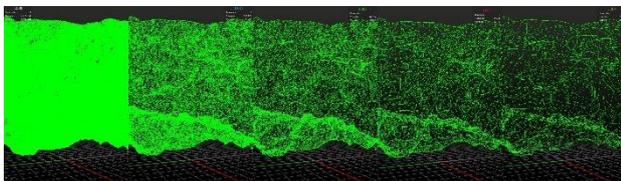


Figure 9. LODs visualization. Triangular meshes per LOD. From left to right: the full-resolution model to the simplified model.

6. CONCLUSIONS

This work proposes a methodology to 3D reconstruct underwater scenes and efficiently present it in an immersive and interactive VR experience. The difficulties that water causes as a medium for light-rays, in contrast to in-air images, are discussed. Here, an image pre-processing pipeline is adopted to assist the reconstruction process and facilitate the VR result. The methodology proposed in this work was evaluated on custom datasets. As a next step, this approach will be evaluated on public datasets and new algorithms on image de-hazing will be tested.

ACKNOWLEDGEMENTS

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code T1EDK-02210).

REFERENCES

- Agrawal, A., Ramalingam, S., Taguchi, Y., Chari, V., 2012. A theory of multi-layer flat refractive geometry, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
- Amores, J., Fuste, A., Richer, R., 2019. Deep Reality: Towards Increasing Relaxation in VR by Subtly Changing Light, Sound and Movement Based on HR, EDA, and EEG, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, NY, USA, 1–2.
- Ballast VR [WWW Document], n.d. URL <https://ballastvr.com/about> (accessed 4.30.20).
- Berman, D., Levy, D., Avidan, S., Treibitz, T., 2020. Underwater Single Image Color Restoration Using Haze-Lines and a New Quantitative Dataset. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1.
- Bruno, F., Lagudi, A., Barbieri, L., Muzzupappa, M., Cozza, M., Cozza, A., Peluso, R., 2016a. A VR system for the exploitation of underwater archaeological sites, in: 2016 International Workshop on Computational Intelligence for Multimedia Understanding, IWCIM 2016. IEEE.
- Bruno, F., Lagudi, A., Barbieri, L., Muzzupappa, M., Ritacco, G., Cozza, A., Cozza, M., Peluso, R., Lupia, M., Cario, G., 2016b. Virtual and augmented reality tools to improve the exploitation of underwater archaeological sites by diver and non-diver tourists, in: *Lecture Notes in Computer Science*. Springer Verlag, 269–280.
- Bruno, F., Lagudi, A., Muzzupappa, M., Lupia, M., Cario, G., Barbieri, L., Passaro, S., Saggiomo, R., 2016c. Project VISAS: Virtual and Augmented Exploitation of Submerged Archaeological Sites-Overview and First Results. *Mar. Technol. Soc. J.* 50, 119–129.
- Calvi, L., Santos, C.P., Relouw, J., Endrovski, B., Rothwell, C., Sara, A., Lucrezi, S., Palma, M., Pantaleo, U., 2018. A vr game to teach underwater sustainability while diving, in: 5th IFIP Conference on Sustainable Internet and ICT for Sustainability, 2017. IEEE, 1–4.
- Chadebecq, F., Vasconcelos, F., Lacher, R., Maneas, E., Desjardins, A., Ourselin, S., Vercauteren, T., Stoyanov, D., 2019. Refractive Two-View Reconstruction for Underwater 3D Vision. *Int. J. Comput. Vis.*
- Chen, Z., Jiang, T., Tian, Y., 2014. Quality assessment for comparing image enhancement algorithms, in: of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition., 3003–3010.
- Corchs, S., Schettini, R., 2010. Underwater image processing: State of the art of restoration and image enhancement methods. *EURASIP J. Adv. Signal Process.*
- Costa, R., Guo, R., Quarles, J., 2017. Towards usable underwater virtual reality systems, in: Virtual Reality. IEEE Computer Society, 271–272.
- Ding, X., Wang, Y., Zhang, J., Fu, X., 2017. Underwater image dehaze using scene depth estimation with adaptive color correction, in: OCEANS 2017. IEEE, 1–5.

- Drewns-Jr, P., Do Nascimento, E., Moraes, F., Botelho, S., Campos, M., 2013. Transmission estimation in underwater single images, in: International Conference on Computer Vision. IEEE, 825–830.
- Emberston, S., Chittka, L., Cavallaro, A., 2018. Underwater image and video dehazing with pure haze region segmentation. *Comput. Vis. Image Underst.* 168, 145–156.
- Fernandez, J.A., Maes, P., Richer, A.F.R., 2019. Deep reality: an underwater VR experience to promote relaxation by unconscious HR, EDA, and brain activity biofeedback, in: Virtual, Augmented, and Mixed Reality on - SIGGRAPH '19. ACM Press, New York, USA, 1–1.
- Ferrera, M., Creuze, V., Moras, J., Trouvé-Peloux, P., 2019. AQUALOC: An underwater dataset for visual-inertial-pressure localization. *Int. J. Rob. Res.* 38, 1549–1559.
- Ghani, A.S.A., Isa, N.A.M., 2014. Underwater image quality enhancement through Rayleigh-stretching and averaging image planes. *Int. J. Nav. Archit. Ocean Eng.* 6, 840–866.
- Jordt, A., Köser, K., Koch, R., 2016. Refractive 3D reconstruction on underwater images. *Methods in Oceanography* 15, 90–113.
- Hatsushika, D., Nagata, K., Hashimoto, Y., 2019a. SCUBA VR: Submersible-type virtual underwater experience system, in: 26th Conference on Virtual Reality and 3D User Interfaces, VR 2019. IEEE, 962–963.
- Hatsushika, D., Nagata, K., Hashimoto, Y., 2019b. Underwater VR Experience System for Scuba Training Using Underwater Wired HMD, in: OCEANS 2018. IEEE.
- He, K., Sun, J., Tang, X., 2011. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 2341–2353.
- Hitam, M.S., Awalludin, E.A., Jawahir Hj Wan Yussof, W.N., Bachok, Z., 2013. Mixture contrast limited adaptive histogram equalization for underwater image enhancement, in: International Conference on Computer Applications Technology, ICCAT 2013.
- Jain, D., Sra, M., Gro, J., Marques, R., Wu, R., Chiu, J., Schmandt, C., 2016. Immersive scuba diving simulator using virtual reality, in: 29th Annual Symposium on User Interface Software and Technology. ACM, 729–739.
- Jakob, W., Tarini, M., Panozzo, D., Sorkine-Hornung, O., 2015. Instant Field-Aligned Meshes, in: ACM Transactions on Graphics.
- Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D., 2020. An Underwater Image Enhancement Benchmark Dataset and beyond. *IEEE Trans. Image Process.* 29, 4376–4389.
- Li, H., Li, J., Wang, W., 2019. A Fusion Adversarial Underwater Image Enhancement Network with a Public Test Dataset 1–8.
- Li, N., Zheng, Z., Zhang, S., Yu, Z., Zheng, H., Zheng, B., 2018. The synthesis of unpaired underwater images using a multistyle generative adversarial network. *IEEE Access* 6, 54241–54257.
- Li, Z., Tan, P., Tan, R.T., Zou, D., Zhou, S.Z., Cheong, L.F., 2015. Simultaneous video defogging and stereo reconstruction, in: Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 4988–4997.
- Liu, X., Chen, B.M., 2019. A Systematic Approach to Synthesize Underwater Images Benchmark Dataset and Beyond, in: 15th International Conference on Control and Automation (ICCA). IEEE, 1517–1522.
- Lu, H., Li, Y., Zhang, Y., Chen, M., Serikawa, S., Kim, H., 2017. Underwater Optical Image Processing: a Comprehensive Review. *Mob. Networks Appl.* 22, 1204–1211.
- Mallios, A., Vidal, E., Campos, R., Carreras, M., 2017. Underwater caves sonar data set. *Int. J. Rob. Res.* 36, 1247–1251.
- McCarthy, J., Martin, K., 2019. Virtual reality for maritime archaeology in 2.5D: A virtual dive on a fluke wreck of 1659 in Iceland, in: 2019 23rd International Conference in Information Visualization - Part II, IV-2 2019. IEEE, 104–109.
- Osonoe, H., Yoshida, T., Ochiai, Y., 2017. Optimized HMD system for underwater VR experience, in: SIGGRAPH 2017 Posters. ACM.
- Park, D., Park, H., Han, D.K., Ko, H., 2014. Single image dehazing with image entropy and information fidelity, in: 2014 International Conference on Image Processing (ICIP). IEEE, 4037–4041.
- Qiao, X., Bao, J., Zhang, H., Zeng, L., Li, D., 2017. Underwater image quality enhancement of sea cucumbers based on improved histogram equalization and wavelet transform. *Inf. Process. Agric.*
- Sinnott, C.B., Liu, J., Matera, C., Halow, S., Jones, A.E., Moroz, M., Mulligan, J., Crognale, M.A., Folmer, E., MacNeilage, P.R., 2019. Underwater virtual reality system for neutral buoyancy training: Development and evaluation, in: Symposium on Virtual Reality Software and Technology, VRST. ACM, 1–9.
- Smith, A.R., 1978. Color gamut transform pairs. *ACM SIGGRAPH Comput. Graph.* 12, 12–19.
- Telem, G., Filin, S., 2010. Photogrammetric modeling of underwater environments. *ISPRS J. Photogramm. Remote Sens.* 65, 433–444.
- TheBlu: An Underwater VR Experience [WWW Document], 2017. URL <https://www.livescience.com/58279-the-blu-underwater-vr-exhibit.html> (accessed 4.30.20).
- Treibitz, T., Schechner, Y.Y., Kunz, C., Singh, H., 2012. Flat refractive geometry. *IEEE Trans. Pattern Anal. Mach. Intell.*
- Vuze [WWW Document], 2020. URL <https://store.vuze.camera/store/> (accessed 4.30.20).
- Yang, M., Sowmya, A., 2015. An Underwater Color Image Quality Evaluation Metric. *IEEE Trans. Image Process.* 24, 6062–6071.
- Zhang, X., Yin, Y., Wan, F., 2019. Underwater manipulation training simulation system for manned deepsubmarine vehicle, in: 26th Conference on Virtual Reality and 3D User Interfaces, 2019. IEEE, 1271–1272.