Improving Underwater Photogrammetric 3D Reconstruction Processing of Shipwreck Sites

Daniel Adams[1,2], Petra Helmholz[3], David McMeekin[1], Andrew Woods[2]

[1]School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth, Australia, david.mcmeekin@curtin.edu.au
[2]Curtin HIVE (Hub for Immersive Visualisation and eResearch), Curtin University, Perth, Australia, daniel.adams1@curtin.edu.au
[3]School for Earth and Planetary Sciences, Curtin University, Perth, Australia, petra.helmholz@curtin.edu.au

Keywords: Underwater Photography, Computer Vision, Image Processing, Digital 3D Models.

Abstract

Applying image processing algorithms to enhance the clarity of underwater images can significantly assist in the visual interpretation of subsea environments. Taking advantage of techniques to maximise the information extracted from the captured input data to produce the highest quality output ensures a higher utilisation of the valuable recorded content which may have been collected at the high cost of a laborious diver survey or an expensive ROV (Remotely Operated Vehicle) expedition. Adopting photogrammetric 3D reconstruction processing in underwater imaging has enabled the creation of high-fidelity digital 3D models which have been used to map subsea infrastructure for maintenance and repair, coral reefs for ecological monitoring, and in the primary context of this article, shipwrecks sites for maritime archaeology and accident investigation. Alongside exploring algorithms to reduce the characteristic haze of underwater photography, this article illustrates the advantages of optimising the processes utilised at distinct stages of the photogrammetric 3D reconstruction workflow to improve the photorealism of the digital 3D models.

1. Introduction

Photogrammetric 3D Reconstruction (P3DR) processing involves taking hundreds or in some cases hundreds of thousands of photographs of a static object or environment from multiple perspectives to create a digital 3D model of the object or location. The process works by comparing mathematically unique 2D features (Olvera et al., 2014) detected in each image to the features detected in every other image of the photographic set (Tu & Dong, 2013). A bundle adjustment (Brown, 1976) process is then used to triangulate the location of matched 2D features to estimate the relative position and orientation of the images at the point of capture. It also allows for solving the intrinsic parameters of the camera. A sparse reconstruction algorithm/spatial resection can then reproject triangulated 2D features into a sparse 3D point cloud (Wu, 2013). Once the images’ intrinsic and extrinsic properties are known a much denser 3D point cloud can be generated and a mesh can be wrapped over those points. The original photographs are then projected back onto the mesh to give it a photorealistic appearance.

P3DR processing of underwater cultural heritage sites (McCarthy, 2016) has grown to prominence in recent years allowing maritime archaeologists, researchers, and the general public to analyse and explore shipwreck sites that were previously inaccessible (Drap, 2012). The resulting 3D reconstructions are not limited to viewing on a flat screen - they can be easily incorporated into interactive visualisations, 3D printed, or even loaded onto a VR (Virtual Reality) headset allowing people to travel back in time to when the photographs were captured to dive on a virtual shipwreck or explore a site in an immersive experience (Figure 1).

For VR and immersive applications, the geometry of the reconstructed object should be spatially accurate and should also look visually accurate, in other words: photo-realistic. Focusing on the P3DR of underwater objects, the physical properties of the underwater medium present several challenges when compared to in-air P3DR processing. Firstly, unlike in-air drone photography, underwater photography is often only semi-structured and so accurate flight planning for optimal geometry can be difficult to achieve. Secondly, light is attenuated with depth and distance to the object of interest, where the faster absorption of its longer wavelengths results in a greenish-blue colour cast, and scattering effects often add a hazy appearance (Treibitz & Schechner, 2009).

Figure 1. a) “Holo-wreck” – a 3D visualisation of the Batavia wrecksite, b) HMAS AE1 wreck as a full-colour 3D print, c) 3D print of HMAS AE1 featured in the “Brickwrecks: Sunken Ships in LEGO® Bricks” exhibition launched at the WA Maritime Museum in Fremantle alongside “Holo-wreck”.

This effect can be observed in a typical example of underwater image data used as the input for 3D reconstruction processing (Figure 2). If a situation arises where the camera utilises an automatic white balance mode or there is the absence of a colour chart in the scene, an accurate colour calibration of the photography is not possible. Alternative image enhancement algorithms in this case become more acceptable to employ to assist in creating photorealistic models.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-2024-7-2024

Figure 2. Original image of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern).

Whether a particular dataset is suitable for reliable P3DR processing will depend upon a variety of factors, some of which can negatively impact the photographic textures and geometric detail of a reconstruction or the amount of coverage in the model. This may include the geometry of the landscape, the turbidity of the water, the quality of the lighting, the amount of backscatter in the images, the amount of sand that might be moving across the site, the sharpness of the photography, the amount of motion blur, the amount of film grain or sensor noise in the images, and how much the environment changes from image to image. Poor visibility in the captured images can often lead to matching errors, image triangulation errors, low visual fidelity, and large gaps in the final model.

This paper contributes to the creation of photo-realistic 3D reconstructions in the absence of a colour calibration charts by proposing an image processing pipeline designed to enhance underwater images. It also outlines a modified set of steps in the P3DR workflow which can be utilised to extract as much detail as possible out of the input data. In some cases, this workflow can overcome challenges that would otherwise prevent turnkey P3DR packages from being able to generate complete reconstructions from the full dataset all in one go.

The paper is structured as follows: in Section 2, a literature review is performed investigating existing methods and the challenges using them. The new methodology is introduced in Section 3 as well as illustrating the cascading advantages of these image enhancement operations on the final 3D model, followed by the evaluation (Section 4) and conclusions (Section 5).

2. Background

The impact of underwater imaging spans across many different disciplines of scientific research and its relevance goes beyond just maritime archaeology. It has applications ranging from understanding the effect of climate change in coral reefs and coral growth rates, corrosion detection of subsea infrastructure to the localisation and mapping systems for ROVs (Liu et al., 2020). Underwater operations are often complex, so it is important that we navigate and interpret these subsea environments reliably. With respect to maritime archaeology and cultural heritage, it’s important to remember that these stories of centuries lost merchant vessels and sunken warships, are very much human stories. Given that these stories often take place during significant times in human history, these virtual recreations can serve as a digital time capsule reminding us of where we’ve come from.

While many image processing algorithms have been investigated to enhance image quality (Mangeruga, et al., 2018) such as Contrast Limited Adaptive Histogram Equalisation (CLAHE) (Hitam et al., 2013), Unsharp Masks (Zheng et al., 2016), automatic colour correction (Ancuti et al., 2018) (Akkaynak & Treibitz, 2019) (Berman et al., 2021) and even a weighted fusion (Ancuti et al., 2012) (Ancuti & Ancuti, 2013) of all the above, they are not without their shortfalls. Dehazing methods often assume uniform natural lighting or are calibrated for in-air photography and may not perform well in the underwater environment where wavelengths of light are attenuated at different rates with depth and distance. Problems can arise using weighted fusion-based methods if the weights are either not initialised wisely or are not flexible enough to handle dynamic lighting conditions. Some 2D convolution matrices have the tendency to over amplify noise and other methods such as basic histogram equalisation can create large ugly changes to the global contrast if the image contains clusters of extreme bright and dark spots. Most importantly their effect on complete photogrammetry datasets and the entire P3DR pipeline from feature detection through to the final model texturing is relatively unexplored (Agrafiotis et al., 2017), which leaves room for new discoveries and optimisations.

When it comes to improving the visual fidelity of underwater photogrammetry models for scientific analysis or museum exhibitions, commonplace image processing algorithms in the field of computer science can be used to enhance the initial image clarity as a pre-processing step and offer significant advantages. The general idea is that given a hazy underwater image, how can distinguishable features be made to stand out more, in order to improve the outcome of any process that derives from the use of those images, such as feature detection, feature matching, object classification, anomaly detection etc. Our goal is for these processing tasks to operate more robustly in the presence of outliers, make less mistakes, converge to solutions faster, those solutions containing less noise, and this resulting in fewer gaps in the coverage. This work tries to address these issues by employing techniques to maximise the information extracted from the captured input data to produce the highest quality output.

There are two algorithms that showed the most promise in the literature review for their simplicity and robust performance - Contrast Limited Adaptive Histogram Equalisation (CLAHE) and Grey World colour constancy. Both these algorithms can work effectively in their own right to enhance underwater photography and overcome some of the characteristic haziness of the transmission medium, but when structured in a logical pipeline the cascading effects of extracting as much visual detail from the dataset is noteworthy.

3. Methodology

The proposed image processing pipeline can be broken down into several distinct steps that are used to transform the original input image observed in Figure 2 into the final pre-processed image in Figure 7. Then how these enhanced images are used to create a photorealistic 3D model. We propose a three-stage strategy (Figure 3). Firstly, a local image enhancement is applied to each individual image. Then, using the enhanced images, a photogrammetric 3D reconstruction is performed. After the model is created, a global texture enhancement is performed on the model textures.

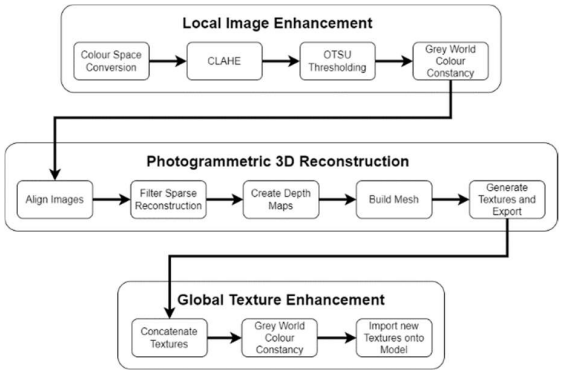

Figure 3. Image enhancement workflow. The Local Image Enhancement stage is performed per image and the Global Texture Enhancement is performed once on the final 3D reconstruction texture.

3.1 Local Image Enhancement - Pre-processing

It is important to maximise the information extracted from the data collected during a laborious diver survey or a costly ROV expedition to produce the highest quality output. The Local Image Enhancement pre-processing is structured as follows. First a colour space conversion. Next, CLAHE is applied, followed by an OTSU binarization to create a mask that finally feeds into the Grey World colour constancy algorithm.

3.1.1 Colour Space Conversion

As a preliminary step every standard RGB image is converted to the LAB colour space, which in a 3-channel image lets you separate out the luminance channel from the other two colour channels. The Alpha and Beta channels represent the two opponent colour components related with chrominance ranging from Red to Green (± a) and Yellow to Blue (± b) (Connolly & Fleiss, 1997). The advantage of splitting the LAB image into its three independent channels, is that it allows the application of a histogram equalisation solely on the luminance channel.

3.1.2 Contrast Limited Adaptive Histogram Equalisation (CLAHE)

Contrast Limited Adaptive Histogram Equalisation (CLAHE) works by applying a localised histogram equalisation across the image within a sliding window. The dimension of the sliding window is the first parameter that can be configured in the algorithm, the second parameter is the clip limit also known as the contrast limit. CLAHE redistributes any pixel intensities that exceed this contrast limit and would otherwise be clipped so that they are uniformly distributed across the histogram. This process is intended to boost the contrast in the image.

In Figure 4 a histogram of a sample image is shown for the original image of the SS *Bonnie Dundee* (Figure 2), where the X axis represents the luminance values from 0 – 255 and the Y axis representing the frequency that pixels in the image have those intensities. From the blue line in Figure 4 depicting the intensity distribution in the original image, it can be observed that values are clustered in a narrow range of the full intensity spectrum. In this particular example, the two peaks along the blue line represents the lighter regions and the darker regions in the image. An exception to this includes cases where images contain high levels of gaussian noise. This condition presents as a single peak in a unimodal histogram and can cause issues with the subsequent OTSU thresholding step.

When CLAHE is utilised solely on the luminance channel of the image to stretch and redistribute pixel intensities across the full intensity spectrum (observed along the red line in Figure 4), the three individual channels are merged back together producing a single image (Figure 5). This also has the noticeable effect of smoothing out the luminance histogram as seen in Figure 4 when comparing the blue line before any image processing and the red line after CLAHE is applied.

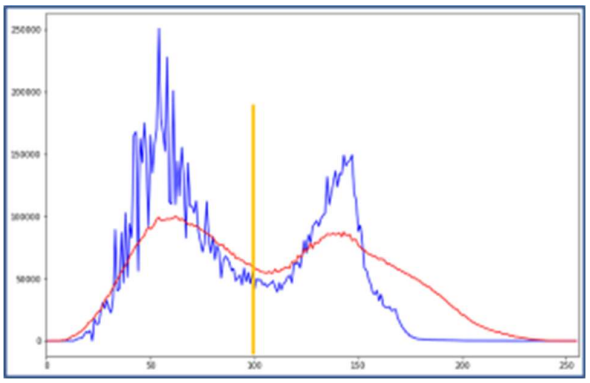

Figure 4. Luminance histogram before (blue) and after (red) CLAHE, with the OTSU threshold line (yellow).

Figure 5. CLAHE image of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern).

Viewing the original image (Figure 2) against the CLAHE processed image (Figure 5), the application of CLAHE on the luminance channel has clearly improved the image contrast. The resultant image appears to have a wider dynamic range, the edges appear sharper, there is better clarity in the fine details and the algorithm preserved all original colour information up to this point by solely operating on the luminance channel.

3.1.3 Colour Effects

As light travels through water, red wavelengths are attenuated at a greater rate than other wavelengths such as blue and green. This effect increases as the path distance increases – either due to increased water depth, or longer path distance between the light source, object of interest, and the image sensor. These processes have the effect of giving underwater images their distinctive blue-green look, which in turn negatively affects the perceived photorealism of the images. The next stages of processing addresses aspects relating to underwater light attenuation and colour.

3.1.4 OTSU Thresholding

If the input image can be classified as a bimodal image, OTSU thresholding (Otsu, 1979) is a technique which can be used to

create a binary mask separating the well-lit and lighter areas from the darker regions. In our experience underwater images are often bimodal in nature. An image can be classified as a bimodal image if the luminance histogram (Figure 4) has two peaks. OTSU binarisation operates by selecting the threshold point in a bimodal image where the variance between the two peaks is close to equal. This point can be observed on the yellow line in Figure 4 and the resultant mask can be observed below in Figure 6.

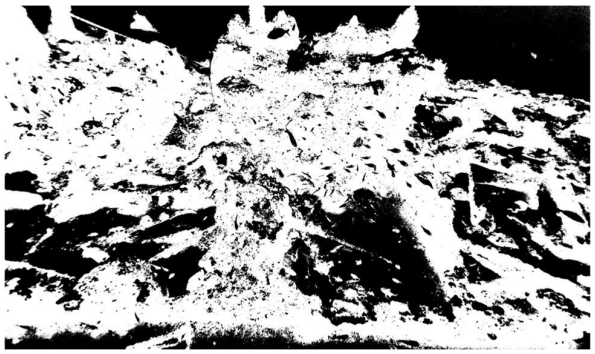

Figure 6. OTSU binary mask segmenting image

3.1.5 Grey World Colour Constancy

The OTSU binary mask generated is fed into the Grey World colour constancy algorithm which attempts to estimate the average colour cast of an image relative to a neutral grey and then applies a colour shift to remove it. Using the original image in Figure 2 as an example, averaging the pixel values along the Alpha channel in the LAB colour space from Green to Red, one might deduce that the average pixel value is leaning more towards the Green side. To account for this, all pixels values in the Alpha channel are shifted back by the difference from this mean value to the midpoint in the colour range. This operation is also applied independently on the Beta channel from Yellow to Blue. The benefit of utilizing the OTSU binary mask is that it makes the algorithm more robust. For example, if the scene is not uniformly lit and there are harsh spotlights in use or there are areas in the image that are over exposed, this can sometimes bias the average colour estimate and the colour shift will drastically overshoot in the opposite direction. Limiting the application of the colour processing based on the OTSU mask results in the final pre-processed image in Figure 7 which offers much better clarity when compared over the original (Figure 2).

Figure 7. Final Pre-Processed image of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern).

3.2 Photogrammetric 3D Reconstruction

The next stage is P3DR processing – creating 3D models from a series of still images of an object. The workflow is structured as follows: First the alignment of the pre-processed images. Next, 3D points from the sparse reconstruction are filtered, then image depth maps are generated from the filtered point cloud, followed by the meshing of the depth maps. Finally, the 3D mesh is textured with the local image enhancement pre-processed images to create the first fully textured 3D model. The consumer photogrammetry software used for this article is Agisoft Metashape Professional 2, however we are also developing our own software pipeline that will run on a supercomputer.

3.2.1 Align Images

The challenges of processing underwater imagery can present in several ways, the most impactful on the quality of a reconstruction or the amount of coverage in the model being the initial image alignment. For this reason, the most reliable way to reduce the likelihood of problems with the imaging system or harshness of the real-world scene impeding a successful alignment, is to maximise the number of features detected and matched from image to image. These features are heavily boosted by algorithms like CLAHE, but this sort of preprocessing works best when the settings are configured optimally in the photogrammetry software. Images were processed at their native resolution and not downscaled, the maximum number of keypoints allowed for detection and matching was limited to 2,000,000 with no limit on the number of tie points reprojected into the sparse reconstruction and guided matching is a setting always used if available.

Even if the situation arises where lower quality settings are still able to yield a valid image alignment, the general philosophy followed is to maximise the detail one can squeeze out of the dataset early in the processing which can be easily filtered in a later stage. This is in opposition to the practice of using lower quality settings early on, which just leaves less data to work with later. Less detail in the sparse reconstruction can result in lower quality depth maps, which then in turn creates less detailed mesh geometry, and the cascading effects of leaving data on the table becomes self-evident by this point.

3.2.2 Filter Sparse Reconstruction

The approach to outlier filtering follows two workflows, the first being a manual selection of points that can be visually identified as noise for deletion. Having the cleanest sparse point cloud conforming to the real shape of the object or environment will yield the cleanest depth maps and therefore less problematic mesh geometry. The second approach being an iterative selection of all points falling outside a specified RMS (Root Mean Square) reprojection error, and then deleting them before rerunning the bundle adjustment. The process is iterative in the sense that the new bundle adjustment can reposition points outside of the chosen RMS threshold and the process can be repeated until the remaining points are of a negligible quantity and realignment no longer has a noticeable effect.

3.2.3 Create Depth Maps

The generation of the image depth maps is a step that can require a small amount of finetuning based on the quality of the sparse reconstruction. Often the optimal quality level or smoothing factor for the depth map generation won’t reveal itself until the subsequent step of meshing the depth maps is completed. Thus, this process can be repeated at multiple levels starting at the medium setting then creating a mesh, inspecting if the mesh has holes, gaps or artefacts resembling dissolving structures apparent in Figure 8.

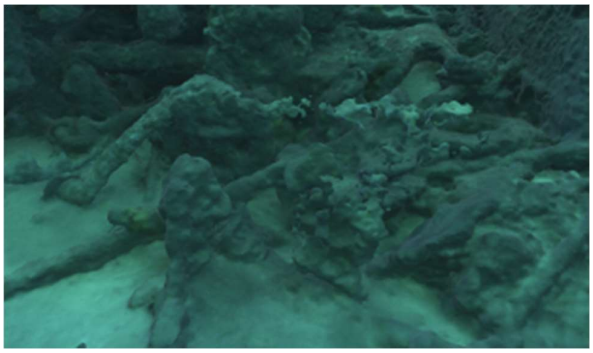

Figure 8. Dissolving thin structures on the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern).

If there are problems with the mesh, the depth map quality can be reduced, introducing a higher smoothing parameter. Alternatively, if the mesh looks correct, the depth map quality can potentially be increased, reducing the amount of smoothing which will bring out more of the finer detail.

3.2.4 Build Mesh

After the image depth maps have been created, the mesh generation is performed directly from the depth maps. The distinction should be made that the common process of building the dense reconstruction is skipped entirely in preference to building the mesh straight from the image depth maps. This avoids introducing common problems with using the dense reconstruction to generate a mesh whereby even the smallest amount of noise can negatively impact the quality of the mesh in terms of total coverage, holes, poor resolution in the finer details and frequent issues reconstructing thin structures. Building the mesh directly from the depth maps also removes the need to perform another stage of point filtering and cleaning to account for this sort of noise. Meshing the depth maps are run with no depth filtering to get the closest representation to the original data as possible, and with extrapolation settings enabled if available to maximise the extent of the final 3D model coverage. Any excess mesh geometry created and often cocooning the 3D model due to the extrapolation setting as seen in Figure 9 can be manually removed.

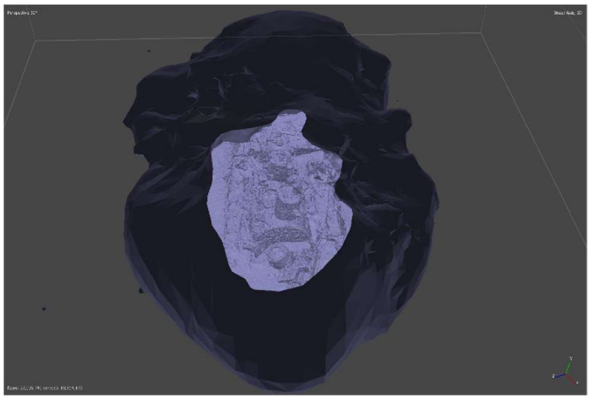

Figure 9. Extrapolated mesh cocooning the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern).

3.2.5 Generate Textures and Export

The final stage of generating the photogrammetric 3D reconstruction is to project the original images onto the mesh to create a near photorealistic 3D model. In order to maximise the visual fidelity of the 3D model, a total of four 8K texture files are generated. This maximises visual texture detail. These textures are finally exported to multiple JPG or PNG files to feed into the next Global Texture Enhancement stage.

3.3 Global Texture Enhancement - Post-processing

It has been our experience that the texture files from the previous step can still benefit from some additional image enhancement. First, all the texture files generated from the reconstruction are concatenated into a single global texture image. Next, the Grey World colour constancy algorithm is used to estimate the global colour cast of the reconstruction and remove it, then finally the global texture image is split back into the four texture files and imported back onto the reconstruction in place of the original texture files.

3.3.1 Concatenate Textures

To achieve a global texture enhancement of the entire 3D reconstruction, the data must first be structured in a way that enables global colour cast estimate to be calculated over the whole dataset. If each texture was processed independently from each other, the resultant textures when loaded back onto the model would produce a patchwork of inconsistent colours with visible seams following the shape of the UV map. Thus, it is necessary to concatenate all exported texture files into a single texture image (Figure 10) which can be then fed into the Grey World colour constancy algorithm.

Figure 10. Concatenated textures of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern).

3.3.2 Grey World Colour Constancy

As described previously the concatenated texture image is again converted from the RGB colour space into the LAB colour space. This enables the estimation of the average global colour cast to be performed on the isolated Alpha (Red to Green) and Beta (Yellow to Blue) channels and for pixels to be shifted back by the difference from this average to the midpoint in the colour range.

3.3.3 Import New Textures onto Model

The final step of the processing pipeline is to split the global texture image that has just been colour enhanced back into the individual texture files. This process must maintain the original resolution and order of concatenation. Order is important as the names of the textures will need to match the original UV maps. The output of this final step in the processing pipeline can be observed below in Figure 11 alongside the Sketchfab link to the explorable 3D reconstruction.

Figure 11. 3D Reconstruction of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern) https://skfb.ly/oEWYn

4. Evaluation

In this section we perform some evaluation of the results and methods presented above in the context of two underwater cultural heritage (UCH) datasets – the SS *Bonnie Dundee* and the American B17 Bomber "*Black Jack*".

The distinction should be made that the approach to evaluation outlined in this section does not refer to the validation of an accurate radiometric colour correction. There were no physical colour charts used in the survey to compare in-air and underwater colour correction values. Thus, in the context of this research paper, the evaluation process seeks to demonstrate how the parameters will give consistent results from the proposed method across multiple datasets.

4.1 The Datasets

The dataset which has been the primary focus of this research paper so far is the SS *Bonnie Dundee,* a 40-metre-long steamship built in Scotland in 1877 for its Australian owners. This dataset shows the Stern of the historic wreck, which is located in 25 meters water depth off the coast near Swansea, New South Wales (NSW), Australia where it was lost on 10 March 1879 following a collision with another vessel, the SS *Barrabool*. 742 photos were captured with a NIKON D90 DSLR camera by recreational diver Grant Thomas and the digital 3D model was generated by Daniel Adams at the Curtin University HIVE.

The second dataset is the American B17 Bomber "*Black Jack*" (lost 1943). The “*Black Jack*” was on a bombing mission to Rabaul from Port Moresby, but bad weather and engine failure forced it to ditch on the return journey. It was 23 metres long and had a wingspan of 32 metres. It ditched in waters 50 meters deep on 11 July 1943 off the coast of Papua New Guinea near Boga Boga (~300km East of Port Moresby) and all 10 crew survived the ditching. 1,983 photos of the wreck were captured with a Canon EOS 5D Mark III by recreational divers Grant Thomas and Andrew Hamilton in June 2023 and the digital 3D model was generated by Daniel Adams at the Curtin University HIVE.

4.2 Impact of CLAHE parameters

The parameters requiring finetuning in the Local Image Enhancement stage are those required to configure the CLAHE algorithm. The first being the size of the sliding window and the second being the clip limit. The size of the sliding window represents how many pixels are considered in the local histogram. The default used in OpenCV – the library this image processing code has been implement in - is a window size of 8x8 pixels. With high resolution imagery this setting may create a little too harsh an effect and result in a bright halo around edges of widely varying contrast. For this reason, the window size used in this pipeline was 16x16 for a milder enhancement. The chosen sliding window works alongside the clip limit parameter which acts as a contrast limit on pixel intensity values after histogram equalisation stretches the image. Any pixel intensities that exceed this contrast limit and would otherwise be clipped are uniformly distributed across the histogram. The default value used in OpenCV is 2.0, which for datasets where images are already of an already decent quality may be too high and may result in an over-enhancement where edge detail appears too bright and unnatural. This is not to say that if a dataset is particularly hazy and unclear, that a value of 2.0 may be suitable, but for the two datasets analysed in this article a value of 1.0 was found to be sufficient. The effect of the CLAHE parameters can be observed below in Figures 12 and 13. From left to right, the first image being the original, the second image being the recommended 16x16 sliding window and clip limit 1.0, and the third image being the default 8x8 sliding window and clip limit 2.0.

Figure 12. An illustration of the effect of CLAHE processing parameters on a single image of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern). a) original image. b) recommended parameters. c) default parameters.

Figure 13. An illustration of the effect of CLAHE parameters on a single image of the B17 “*Black Jack*” Aircraft wreck (1943). a) original image. b) recommended parameters. c) default parameters.

4.2.1 Analysis

The advantages of applying CLAHE to the input images extends beyond revealing more detail in the final 3D model textures. The reduction in the underwater haze achieved by increasing the contrast and sharpness of the images also has the benefit of boosting the number of SIFT keypoints extracted from the dataset – as shown in Figures 14 and 15.

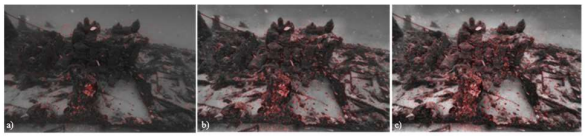

Figure 14. Comparison of SIFT keypoints vs CLAHE parameters on a single image from the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern). a) original image. b) recommended parameters. c) default parameters. Red markers indicate extracted keypoints.

Figure 15. Comparison of SIFT keypoints vs CLAHE parameters on a single image from the B17 “*Black Jack*” Aircraft wreck (1943). (a) original image, (b) recommended parameters, and (c) default parameters.

A comparison of the detected SIFT keypoints for the SS *Bonnie Dundee* and the B17 “*Black Jack*” is summarised in Table 1 below. In both Figure 14 and Figure 15, the first image (a) being the original image, the second image (b) being the recommended 16x16 sliding window and clip limit 1.0, and the third image being the default 8x8 sliding window and clip limit 2.0.

SIFT keypoints	Figure 14	Figure 15
a)	4,801	799
b)	17,669	8,299
c)	28,254	21,322

Table 1. Comparison of the detected SIFT keypoints

It could be argued that CLAHE processing may amplify noise in the image and by extension increase the number of false positive feature matches in the subsequent image alignment step, so there is a trade-off to be considered. But in the case of the datasets analysed in this article, the noise can be minimised by using the recommended settings over the default sliding window and clip limit. Counter to this argument, there are some cases whereby the underwater haze has seemingly prevented the alignment of certain images and as a result can leave gaps in the 3D model coverage. So, in this case an algorithm like CLAHE that boosts the number of features available to match between images becomes advantageous to experiment with. All things considered, given that the parameters are trained by eye and can be finetuned to each dataset, the benefits of using this technique will often outweigh the costs. Ultimately the noisy features reprojected into the subsequent sparse reconstruction point cloud can always be filtered through the same method previously described in section 3.2.2.

4.2.2 Visual Inspection

When we evaluate the cascading effects of the local and global (pre- and post-processing) enhancement procedures on the two datasets, it can be observed in Figure 16 that the most noticeable improvement comes from the local enhancement. This is not to say that the global enhancement hasn’t produced more subtle improvements that still add to the clarity and ease of interpretation. Even from a distant viewpoint, the global enhancement paints less a murky picture of the SS *Bonnie Dundee* (3D model visible here: https://skfb.ly/oEWYn) wreck over the original unprocessed images. The dynamic range in the scene has improved considerably with more detail visible between both the lighter and darker areas, the sand has a more visually accurate white point as the characteristic bluish-green haze is measurably reduced and the two Wobbegong sharks resting on top of the steam engine are clearly visible.

In the case of the B17 “*Black Jack*” Aircraft (3D Model visible here: https://skfb.ly/oJpZr), the wreck rests in deeper water with lower levels of natural light penetration which can add to the challenge of recovering detail. Despite this, the global enhancement in Figure 16 can be observed bringing out much more edge detail and visual resolution over the unprocessed version. The overall effect the global enhancement pipeline has on the final 3D reconstructions is to almost remove the visual characteristics of the water itself.

The argument could be made that producing a rendition of a wrecksite free of the underwater haze and representing it in a way that is more visually familiar to people, exploring these 3D reconstructions can feel more realistic and immersive. If artefacts on a site are less obscured and more easily identifiable, these algorithms have the potential to increase people’s connection to the data and the history.

4.3 3D Reconstruction of Original vs Global Texture Enhanced

Without the presence of physical ground control points in the datasets, it’s difficult to compare the accuracy of the final mesh geometry derived from the original images against one derived from the enhanced images. Instead, the percentage of aligned images and coverage in the final 3D reconstructions can be compared to validate the original hypothesis. That being if techniques to maximise the information extracted from the captured input data to produce the highest quality output are employed, it may result in fewer gaps in the coverage.

When reprocessing the SS *Bonnie Dundee* dataset from start to finish there is less observable difference between the total 3D model coverage comparing the original vs the image enhanced models. All 742 images are correctly aligned in both reconstructions, and the sparse point cloud which the geometry is built on top of has a difference of 58,213 points (1,093,524 in original and 1,151,737 in enhanced). Although this difference in the density of the sparse reconstruction is not insignificant, the actual geometry remains fairly consistent when the same parameters are used. This is a testament to the quality of the original images and clarity of the water column. Natural light in this dataset has also been able to penetrate further at 35 meters water depth, which has translated to less underwater haze that might otherwise hinder image alignment.

Contrary to the first case and resting at 50 meters water depth, reprocessing the B17 “*Black Jack*” dataset from start to finish tells a slightly different story. 1,978 out of the total 1,983 images in the dataset are aligned when the original images are used, and 1,981 out of 1,983 images in the dataset are aligned when the enhanced images are used. This has resulted in the sparse point cloud available to build the geometry from having a difference of 786,298 points (9,377,354 in original and 10,163,652 in enhanced) when processing with the same parameters. In this case, the increase in the density of the points and number of images aligned in the reconstruction from the enhanced dataset has improved the resolution and resolved some errors in the final 3D model, most notably around the 2 thin barrels of the upper gun turret. Although this improvement has come as a result of aligning only a few extra images, these sorts of problems become more likely to be resolvable when encountering hazier images or datasets with less overlap. In these cases, every extra image aligned can translate to a greater difference in the final result.

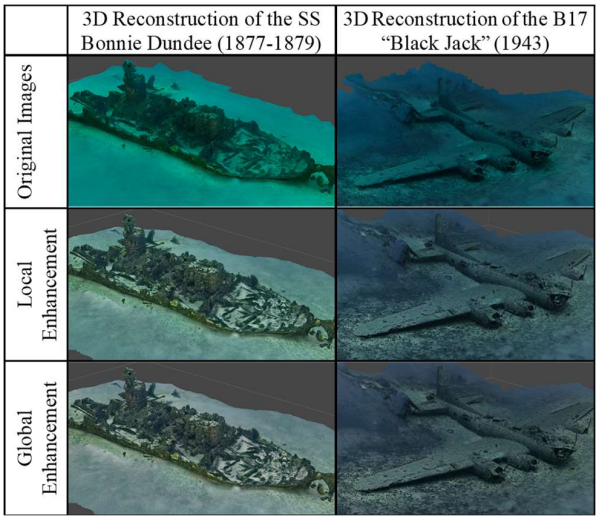

Figure 16. 3D Reconstruction of the SS *Bonnie Dundee* (1877-1879) Shipwreck (Stern) (left, https://skfb.ly/oEWYn) and the B17 “*Black Jack*” Aircraft wreck (1943) (right, https://skfb.ly/oJpZr), using original images (top), using local enhancement only (middle) and using local and global enhancement (bottom).

5. Conclusion

This article has illustrated how image processing and photogrammetric 3D reconstruction processing techniques can be applied and optimised for generating high quality photo-realistic 3D models of underwater cultural heritage sites

(including wrecksites). The cascading effects of these algorithms is remarkable when structured in a logical pipeline where extracting as much visual detail from the dataset is the objective. When it comes to reducing the bluish-green haze often characterising underwater photography, the techniques discussed in this paper enhance the clarity of the original images. Enhancing the clarity of underwater photography and postprocessing the final 3D model textures has the potential to provide a much richer understanding of underwater sites. When it’s not as challenging to differentiate objects from one another, and finer details are less obscured, these virtual recreations can be an excellent source of new information allowing archaeologists and the broader public revisit and explore underwater heritage locations as if they’re frozen in a time capsule.

This paper has described one stage of our P3DR processing pipeline. We are continuing to explore a range of options to improve the performance of P3DR processing. One of those improvements is using supercomputer and cloud computing to increase the processing speed of large-scale datasets, and hence increasing the size of dataset that can be processed in a reasonable time. That development is ongoing. We are also working on other improvements which will be reported in future papers.

Acknowledgements

The authors would like to thank recreational divers Grant Thomas and Andrew Hamilton who provided the original underwater image datasets used in these reconstructions. This project was made possible by support from the Curtin HIVE.

References

Agrafiotis, P., Drakonakis, G. I., Georgopoulos, A., & Skarlatos, D. (2017). THE EFFECT OF UNDERWATER IMAGERY RADIOMETRY ON 3D RECONSTRUCTION AND ORTHOIMAGERY. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W3*, 25–31. doi:10.5194/isprs-archives-XLII-2-W3-25-2017

Akkaynak, D., & Treibitz, T. (2019). Sea-Thru: A Method for Removing Water From Underwater Images. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 1682-1691). doi:10.1109/CVPR.2019.00178

Ancuti, C. O., & Ancuti, C. (2013). Single Image Dehazing by Multi-Scale Fusion. *IEEE Transactions on Image Processing, 22*, 3271-3282. doi:10.1109/TIP.2013.2262284

Ancuti, C. O., Ancuti, C., De Vleeschouwer, C., & Bekaert, P. (2018). Color Balance and Fusion for Underwater Image Enhancement. *IEEE Transactions on Image Processing, 27*, 379-393. doi:10.1109/TIP.2017.2759252

Ancuti, C., Ancuti, C. O., Haber, T., & Bekaert, P. (2012). Enhancing underwater images and videos by fusion. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 81-88). doi:10.1109/CVPR.2012.6247661

Berman, D., Levy, D., Avidan, S., & Treibitz, T. (2021, August). Underwater Single Image Color Restoration Using Haze-Lines and a New Quantitative Dataset. *IEEE transactions on pattern analysis and machine intelligence, 43*(8), 2822-2837.

Brown, D. (1976). The bundle adjustment-progress and prospect. *XIII Congress of the ISPRS, Helsinki, 1976.*

Connolly, C., & Fleiss, T. (1997). A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *IEEE Transactions on Image Processing, 6*, 1046-1048. doi:10.1109/83.597279

Drap, P. (2012). Underwater Photogrammetry for Archaeology. In D. C. da Silva (Ed.), *Special Applications of Photogrammetry.* Rijeka: IntechOpen. doi:10.5772/33999

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. *2013 International Conference on Computer Applications Technology (ICCAT)*, (pp. 1-5). doi:10.1109/ICCAT.2013.6522017

Liu, R., Fan, X., Zhu, M., Hou, M., & Luo, Z. (2020). Real-World Underwater Enhancement: Challenges, Benchmarks, and Solutions Under Natural Light. *IEEE Transactions on Circuits and Systems for Video Technology, 30*, 4861-4875. doi:10.1109/TCSVT.2019.2963772

Mangeruga, M., Cozza, M., & Bruno, F. (2018). Evaluation of Underwater Image Enhancement Algorithms under Different Environmental Conditions. *Journal of Marine Science and Engineering, 6*. doi:10.3390/jmse6010010

McCarthy, J., Benjamin, J., Winton, T., & van Duivenvoorde, W. (2019). The Rise of 3D in Maritime Archaeology. In J. K. McCarthy, J. Benjamin, T. Winton, & W. van Duivenvoorde (Eds.), *3D Recording and Interpretation for Maritime Archaeology* (pp. 1–10). Cham: Springer International Publishing. doi:10.1007/978-3-030-03635-5_1

McCarthy, M. (ed.), 2016, *From great depths: The wrecks of HMAS Sydney II and HSK Kormoran*. WA Museum and UWA Publishing, Perth.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics, 9*, 62-66. doi:10.1109/TSMC.1979.4310076

Treibitz, T., & Schechner, Y. Y. (2009). Active Polarization Descattering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31*, 385-399. doi:10.1109/TPAMI.2008.85

Tu, L., & Dong, C. (2013). Histogram equalization and image feature matching. *2013 6th International Congress on Image and Signal Processing (CISP)*, *01*, pp. 443-447. doi:10.1109/CISP.2013.6744035

Wu, C. (2013). Towards Linear-Time Incremental Structure from Motion. *2013 International Conference on 3D Vision - 3DV 2013*, (pp. 127-134). doi:10.1109/3DV.2013.25

Zheng, L., Shi, H., & Sun, S. (2016). Underwater image enhancement algorithm based on CLAHE and USM. *2016 IEEE International Conference on Information and Automation (ICIA)*, (pp. 585-590). doi:10.1109/ICInfA.2016.78318