Environment-Aware Design For Underwater 3D-Scanning Application

Josephine Antoniou and Charalambos Poullis\textsuperscript{1}

Immersive and Creative Technologies Lab, Cyprus University of Technology

November 22, 2012

\textsuperscript{1}\texttt{charalambos@poullis.org}
Abstract

Underwater archaeologists and exploration groups often face a challenge with the documentation and mapping process, which must take place underwater and be able to accurately capture and reconstruct the specific archaeological site. The automation of the scanning and reconstructing process is quite desirable for the underwater archaeologists, however, such automation entails quite a few hurdles from a technological perspective, in terms of data acquisition, processing and final reconstruction of the objects situated underwater. This paper focuses on the design of the 3D scanning application, for the purpose of reconstructing the underwater objects/scenes, such that it is environment-aware. By environment-aware the paper refers to the identification of aspects of an underwater environment that need to be considered in a 3D scanning process and, furthermore, the designing of a system that considers these aspects when scanning objects/scenes found in underwater environments. In this process, several decisions need to be made, with regards to the setup, the method and the analysis, considering issues that may arise in such environments.
0.1 Introduction

The paper proposes a three-dimensional (3D) scanning application, especially designed for scanning objects/scenes in underwater environments. The paper focuses on the design of the 3D scanning system, for the purpose of reconstructing the underwater objects/scenes, such that it is environment-aware. By environment-aware the paper refers to the identification of aspects of an underwater environment that need to be considered in a 3D scanning process and, furthermore, the designing of a system that considers these aspects when scanning objects/scenes found in underwater environments. The design firstly identifies the features of such environments that need to be considered for the successful 3D underwater scanning. Underwater environments are challenging environments for many reasons. Firstly, images taken in such an environment are susceptible to high noise content, because of the medium, i.e. water, as well as due to the motion that is induced in the setup because of the water motion in sea environments. Furthermore, the induced motion in the setup results in no guarantees in terms of system setup parameters, i.e. calibration parameters, which are necessary to achieve the scene reconstruction. Scene reconstruction usually utilizes a sequence of images taken from a statically placed setup, so that information about the scene can be extracted, by using information about the setup as the constant feature. In motion-susceptible environments such task proves to be extremely challenging. Once the environment features, which make the required task more challenging, have been identified and given the selection of setup approaches, scanning techniques, analysis methods, and image processing options that are available, the design decisions for the 3D scanning application are made. Overall, decisions regarding the setup itself, i.e. the placement of the system components, as well as the scanning methodology and image analysis, are guided by the environment limitations, so that they can be overcome by the application design. The paper is organized as follows. Section 2 focuses on background information and related work, organized into three sub-categories: the identification of underwater vision issues (2.1), the 3D scanning techniques (2.2), and the noise handling through filtering techniques (2.3). Section 3 then presents the proposed design for the environment-aware 3D scanning application, by discussing separately the system setup (3.1), the generation of the wrapped phase during the scanning process (3.2), the generation of the unwrapped phase during the scanning process (3.3), and the reconstruction considerations (3.4). Finally, Section 4 offers some conclusions.

0.2 Background and related work

In this section related work is divided into three different subsections, which present necessary background knowledge to be used for the targeted application. The first subsection identifies features of the underwater environment that the design and implementation of a 3D scanning application needs to consider, whereas the second subsection focuses on the presentation of currently existing
3D scanning methodologies, explaining why the structured light methodology is the most appropriate for underwater-enabled scanning, as well as selecting a particular type of structured light scanning. Finally, since underwater environments are more challenging, i.e. may produce very noisy images, the third subsection briefly discusses filtering options for addressing the possible noise in images taken in such environments.

0.2.1 Issues of Underwater Vision

Underwater environments pose undoubtedly many challenges for computer vision applications, especially in terms of data-collection and data-processing tasks, and thus require increasing effort from a technological point of view. Both the fact that the application needs to deal with a topologically structure-less space, as well as the fact that the medium itself imposes additional difficulties by its very nature, limit the options of the application developers. Overall, underwater environments may impose high noise content on images due to the water medium and the induced motion (Hou and Wiedemann, 2007). This is especially true in the case of computer vision applications since such applications rely heavily on the accuracy of the input, closely linked with the predictability and control of the environment, which is not the case in the underwater world. In fact, feature detection is made more difficult by the effect of depth, i.e. as depth increases the contrast in underwater images decreases (Reynolds, 1998). Moreover, sea motion causes long-term accurate position estimation to be problematic (DeAngelis and Whitney, 2000). Processing images that have been captured in underwater environments needs to consider the challenges imposed by such environments with regards to the reliability and accuracy of the obtained results, e.g. on calibration parameters, as compared to similar processes occurring in ideal environments, i.e. air medium environments that allow for the static setup of the system components (Bryant et al, 2000). Given the above-mentioned hurdles, the interest in exploring underwater environments and finding solutions towards implementing computer vision applications is more and more evident in many areas, especially in the area of underwater archaeology. It is aimed through this work to propose an environment-aware application design for underwater 3D scanning in order to assist underwater archaeologists with the automation of the documentation process.

0.2.2 Techniques for 3D Scanning

The main purpose of a 3D scanner is to collect information about the shape of a real-world object or scene so that it may be represented in three dimensions. A successful such process for the required application, would enable the digitized documentation of underwater scenes, in terms of shape, height measurements and colour, which would be particularly helpful for underwater archaeologists and exploration groups in terms of efficiency, since they often resort to time-consuming manual documentation of their findings. There exist many types of 3D scanning technologies, but the three most commonly used technologies
are: contact scanning, laser triangulation and structured light scanning (Ikeuchi, 2001). Contact Scanning technologies give quite accurate results, however, they require quite a bit of time to complete the scan. The proposed application design aims for a more rapid solution of underwater 3D scanning. Laser triangulation technologies use the motion of a laser line that scans over the object/scene. It is required that for the time the laser line scans the object/scene the system setup should be as static as possible, however, such guarantee is not possible in an underwater environment, especially in sea environments, where water motion exists. Therefore, the accuracy of using laser triangulation in underwater 3D scanning applications is not expected to reach satisfactory levels. The third type of scanning technology, structured light scanning can be, under certain conditions, a much more promising scanning solution in terms of accuracy and reliability. Structured light scanning systems illuminate the object surface with a light pattern and collect the images of the modified pattern (because of the shape of the scanned object), which are in turn analyzed to give back the real-world object features such as its shape and height. One of the most commonly used patterns projected is what is known as a sinusoidal fringe pattern, examples of three different frequencies of such a pattern are given in Figure 1.

\[ I(x, y) = I_b I_m \ast \cos(\theta(x, y) + a) \]  

where \( I(x, y) \) is the intensity of the fringe pattern at a generic point \((x,y)\), \( I_b \) is the background intensity, \( I_m \) is the fringe modulation intensity, \( \theta(x,y) \) is the phase difference, and \( a \) is a known added phase. This equation gives the distribution for an ideal fringe pattern; the actual distribution may be different depending on the set-up parameters and optical components. In the 3D scanning process, in addition to deciding to scan the object/scene with a sinusoidal
fringe pattern, one needs to be aware of the analysis methodologies and use the one most appropriate for the environment-aware application. Generally, there exist two broad categories of analysis approaches of a sinusoidal pattern. By analysis we refer to the way the captured pattern on the object/scene is processed to extract the required shape information. The two types are: the spatial analysis, and the temporal analysis. In the spatial analysis, information for the object shape is gathered by studying the modified pattern over the frame as a whole or a collection of pixel neighbourhoods, i.e. sub-frame windows. On the other hand, in the temporal analysis, information for the object shape is gathered by studying the same pixel(s) in different images. In an underwater scanning application, however, water motion prohibits any guarantees regarding the preservation of information between frames, and thus for the purpose of designing an underwater-enabled application, the processing approach selected is the spatial one so that there are no dependencies between subsequent frames.

0.2.3 Addressing Noise through image filtering

Although issues arising due to the existence of motion are mostly addressed with the choice of scanning and analysis techniques, there might still exist noise effects in the captured image. Noise may be the result of other factors, in addition to the motion factor. Noise is an important issue to overcome in challenging environment conditions, such as those imposed by an underwater sea setting. Noise can be addressed as a pre-processing or a post-processing issue, i.e. prior to applying the spatial analysis onto the captured images or once the captured image has been spatially processed. Image filtering is considered as a tool to either remove noise prior to fringe analysis or smooth out noise effects from an already analysed image. Digital images are prone to noise due to a variety of reasons. In underwater environments, noise is the result of the various features of the environment causing errors in the acquisition of the image data, such as motion or depth, resulting in pixel values that do not reflect their real-world intensities. Common filtering tools include: linear, median and adaptive filtering. Linear filtering includes filters such as averaging or Gaussian filters, which are useful for removing grain type of noise and overall smoothing out the image. The averaging filter sets each pixel to the average value of the pixels in its neighbourhood, thus reducing local variations caused by grain type of noise. Similarly, the Gaussian filter removes noise according to the Gaussian distribution. Median filtering is similar to averaging filter; however, median filtering determines the value of a pixel by the median of the neighborhood pixels, rather than the average. Median filtering is better able to remove noise without reducing the sharpness of the image since the median is much less sensitive to outliers than the average. Adaptive filtering applies any of the above-mentioned filters to an image in an adaptive way, tailoring itself to the local image variance, i.e. where the variance is large, little smoothing is applied, whereas where the variance is small, more smoothing is applied. Often the adaptive approach produces better results than simply using linear or median filters on their own.
0.3 Design of Underwater 3D Scanning Application

In this section the 3D scanning application is designed step by step according to the issues discussed in Section 2. The steps are outlined in Figure 2 and further elaborated with appropriate examples in the following subsections.

![Figure 2: Steps to consider in designing the 3D scanning application](image)

0.3.1 System Setup

In order to scan any object/scene it is required that we have the following components: A sinusoidal fringe pattern to be projected A projector to project the pattern and a camera to capture the modified pattern. The sinusoidal fringe pattern can be easily generated with the following expression:

\[
\frac{1 + \sin(a \cdot Y \cdot p)}{2}
\]  

where \(a\) is the angle, \(p\) is the period and \(Y\) is the fringe orientation. It is recommended that the fringes generated are as thin as possible, i.e. that the frequency of the generated fringe pattern is high. This will result in more information about the object shape being collected from the subsequent analysis.
Figure 3 presents such a sinusoidal pattern projected onto a selected object in the laboratory. Note that the brighter white spots are due to differing texturing on the head itself.

Figure 3: Sinusoidal pattern captured as it is projected onto an object from a parallel optical axes system setup

The camera and the projector need to be strategically placed in relation to one another and in relation to the object itself so that processing the subsequent image does not depend on the their static placement, especially between the object and the rest of the setup as the camera and projector need to be placed in appropriate waterproof casing, thus their topological relation with regards to each others relative position can be fixed to remain constant. Given the subsequent analysis of the captured frame, the most convenient setup is one where the optical axis of the camera and the optical axis of the projector are parallel (Takeda and Mutoh, 1983), as it slightly simplifies the subsequent analysis process, in terms of extracting height information from the unwrapped phase map (see subsection 3.3), since the fringe pattern projected onto the object/scene remains in its regular form regardless of the position of the projector, e.g. whether it is placed next to the camera or on top of the camera. In order to simulate an effect closer to an image captured in an underwater environment, and in particular, the effect of motion onto the captured fringe pattern, a motion filter has been applied to the image to observe such effects as expected in the targeted environment (Figure 4). Edge sharpness is clearly affected by the motion filter.

Pre-filtering of the image can be applied to alleviate the effects of motion and consequently, increase the contrast between dark and light fringes. Pre-processing of the image with a simple filter that enhances the contrast in the image is presented in Figure 5.
Figure 4: Motion filter applied to captured image of a sinusoidal fringe pattern projected onto an object.

Figure 5: Pre-processing the motion-filtered image in order to restore the contrast between dark and light fringes.

The purpose is to restore the contrast between the fringes in order to be able to analyse them at correct phase values during the scanning process, i.e.
the recognition of the shape of the object/scene at its real world measurements. Figure 5 does not achieve the sharp contrast existing in the preliminary image in Figure 3; however, it manages to alleviate the effects of motion by slightly increasing the contrast between the fringes.

0.3.2 3D Scanning Obtaining Wrapped Phase Map

Once the image of the modified fringe pattern is captured, it needs to be processed in order to extract the necessary shape information. The analysis technique that has been selected for the proposed design is the Windowed Fourier Transform (WFT) analysis (Kemao et al, 2010). Fourier transform analysis, with the use of the fast Fourier transform algorithm, is a well-known spatial analysis method, which is based on a time-frequency analysis that can demodulate a given fringe pattern in order to yield the required shape information. WFT uses this concept and operates on pixel sub-neighbourhoods in order to better capture the shape of the object, especially for shapes that are not so uniform. Given the kind of images that we are expecting in an underwater environment, like the image in Figure 4, we need an analysis method that can give accurate results and a windowing technique is acting locally in each sub-pixel neighbourhood, prior to giving the global wrapped phase map. The WFT analysis technique has been achieved with the use of Matlab Software (Kemao, 2009). A useful option of the tool is that it allows at this stage of the processing to be able to view, in addition the wrapped phase map or better the filtered phase map (windowed Fourier filtering algorithm is used, which makes use of the fast Fourier transform algorithm), the quality map, which basically present the absolute value of the filtered phase value for each pixel, which basically represents the magnitude of the complex image (image without the fringe pattern). An example of a quality/magnitude map for the selected object with the use of this software is presented in Figure 6. This bi-product is not a height map; it simply gives the recognized shape stripped from the pattern without being the case that each pixel value corresponds to a height value.

0.3.3 3D Scanning Obtaining Unwrapped Phase Map

Most of the existing analysis techniques, both temporal and spatial ones, generate a wrapped phase distribution of this information, i.e. for each pixel a phase value limited to the interval \([-\pi, \pi]\), in which the recovered phase contains artificial discontinuities since the true phase may range over an interval greater than 2. The process of recovering true phase is referred to as unwrapping. The unwrapping of the wrapped phase gives a continuous phase map that can be in turn translated into object height information. It is worth noting that in general, the continuous phase distribution that is obtained by the unwrapping stage contains the sum of the objects shape related phase and the fringe related phase. Therefore, the fringe-related phase must be removed so that the height of the object can be estimated. This effect is negligible in parallel axis setups (Takeda and Mutoh, 1983). In principle, the parallel axes system setup does not require
for the initial phase, i.e. the projected fringe pattern phase at zero height, to be subtracted from the current phase obtained. Therefore, it is important that during setup, the relative position between the camera and the projector in the underwater casing is fixed so that the optical axes of the two components are constantly parallel. The unwrapping algorithm used is provided by the WFT software (Kemao, 2009) and it is based on a quality guided flood fill algorithm (Kemao et al, 2010). The successful unwrapping is based on the quality or magnitude obtained during the WFT analysis of the image as demonstrated in Figure 5. The unwrapped phase map of the selected object is presented in Figure 7. In this image each pixel corresponds to a height value. Subsection 3.4 discusses the extraction of height map from the unwrapped phase map.

0.3.4 Post-processing and Reconstruction

Finally calibration information should be used to map the unwrapped phase distribution to real-world 3D coordinates. Prior to discussing how to extract the height information, we need to perform some post-processing on the unwrapped phase map in order to smooth out the colour grading. The sharp colour changes are a result of some inaccuracies in the pixel values from pre-filtering stage. Post-processing works well with objects and scenes that have a continuous surface, like the selected object for example. Since what is wanted is a smoother colour grading in our final image, a linear filtering technique has been used on the unwrapped image. Figure 8 demonstrates the obtained result, to be used for the phase to height calculation stage. Next, we discuss the phase to height calculation phase as it is a part of the scanning process that has been recently revisited in literature, due to an overlook in the initial calculation presented in (Takeda and Mutoh, 1983). The appropriate calculation for a parallel optical
axes setup is discussed next based on (Maurel et al, 2009). Once the unwrapped

phase map, or post-processed version of it is generated, the correct measurement of the shape of the object needs to be calculated. That requires a relationship between the phase at a pixel and the corresponding height measurement. The original equation presented in (Takeda and Mutoh, 1983) is:

\[ h(x, y) = \frac{L \Delta \phi(x, y)}{(\Delta \phi(x, y) - \omega D)} \]
where \( h(x,y) \) is the height at a point, \( \Delta \phi(x,y) \) is the phase difference between the fringe pattern phase captured on the object and the fringe pattern phase captured on a reference plane of zero height, \( L \) is the distance between the reference plane and either the projector or camera (since they are assumed to be placed at the same distance from the reference plane), \( D \) is the distance between the camera and the projector and \( \omega \) refers to the frequency of the projected fringe pattern. The authors in (Maurel et al, 2009) clarify that equation (3) returns not in fact \( h(x,y) \) but \( h(x,y) \) where:

\[
x = x\left(\frac{h}{L}\right)x
\]

and,

\[
y = y\left(\frac{h}{L}\right)y
\]

This results in a new equation for phase to height calculation, which for a parallel axes optical system is the following:

\[
\Delta \phi(x,y) = \omega\left[\frac{h(x,y)}{L}\right]y
\]

Although, 3D reconstruction of an object/scene is out of the scope of this paper, we present in Figure 9(a-b), snapshots of a generated 3D depiction of the scanned surface of the selected object. In reality multiple frames capturing different perspectives of the object/scene would be necessary to get a detailed and complete reconstruction. Future work plans to consider the integration of multiple frames of different perspectives into the proposed design, so that 3D reconstruction of an object/scene is not limited to only one surface.

Figure 9: A 3D depiction of the scanned object surface.
0.4 Conclusion

The paper takes a pipeline approach toward 3D scanning for underwater environments. The discrete steps, mandatory, e.g. generating the wrapped and unwrapped phase maps, and optional, e.g. pre-processing and post-processing of captured and generated image of the object/scene, are identified and discussed, both as important procedures on their own, as well as in relation to the environment that the application aims to exploit. The environment-awareness is discussed for each pipeline step and a continuous example of an implementation of the proposed design is given throughout Section 3, in which the proposed design is elaborated. In particular, the environment-aware design is given as four consecutive individual processes, where the outputs of a process feed the next process as inputs, resulting in the requested 3D scanned product being the output of the fourth step. The four steps include the system setup, the generation of the wrapped phase, the generation of the unwrapped phase and the reconstruction considerations. The steps incorporate image pre-processing and post-processing options. Furthermore, it is worth mentioning that future work plans to implement the proposed design in an actual underwater set-up, as it is the on-going work of the authors as part of an EU-funded project.

Bibliography

fourier-transform-for-fringe-pattern-analysis-with-gui