CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Underwater Image Enhancement By Fusion

by

Shabbar Abbas

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Computing
Department of Computer Science

2019

Dedicated to

My Parents and my Wife

CERTIFICATE OF APPROVAL

Underwater Image Enhancement By Fusion

by

Shabbar Abbas

(MCS153015)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Ayyaz Hussain	IIUI, Islamabad
(b)	Internal Examiner	Dr. Nadeem Anjum	CUST, Islamabad
(c)	Supervisor	Dr. Mohammad Masroor Ahmed	CUST, Islamabad

Dr. Mohammad Masroor Ahmed
Thesis Supervisor
November, 2019

Dr. Nayyer Masood
Head
Dept. of Computer Science
November, 2019

Dr. Muhammad Abdul Qadir
Dean
Faculty of Computing
November, 2019

Author's Declaration

I, **Shabbar Abbas** hereby state that my MS thesis titled "**Underwater Image Enhancement By Fusion**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Shabbar Abbas)

Registration No: MCS153015

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Underwater Image Enhancement By Fusion**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been dully acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Shabbar Abbas)

Registration No: MCS153015

Acknowledgements

I would like to render my heartfelt regards to my supervisor, **Dr. Mohammad Masroor Ahmed**, his continuous guidance made this research possible. His advice and friendly demeanour not only made him accessible but also encourage me to discuss all sort of difficulties in a rather warm environment. I highly appreciate his patience and expertise.

I am in great indebtedness to my colleagues and friends Shuja Uddin Khawar and Zeeshan Ur Rehman for helping me find the research material and with the computational resources.

Finally, I would like to thank my beloved family for standing by me during the whole period and their constant encouragement and advice. I will always be grateful and pray for them.

(Shabbar Abbas)

Registration MCS153015

Abstract

Underwater images usually lack contrast and suffer from color distortion due to light beam scattering and attenuation. Light scattering is due to the presence of suspended particles in water in form of both organic and inorganic material which reflects and deflects the light in an unpredictable manner before it reaches the sensor and results in an image which is low in contrast. Water as a medium readily absorbs light, and moreover different wavelengths of light as absorbs at different rates. Furthermore, the longer wavelength is absorbed first and it results in the underwater environment with a dominant green-bluish tone.

Our research has resulted in improvement upon the well-established underwater image fusion method. Our underwater image enhancement system can broadly be divided in three major components; splitting of the original image into illuminance and reflectance slices, applying a linear piecewise color correction algorithm on the reflectance slice and contract enhancing techniques on the illuminance part. Finally, to reconstruct the output image we apply the multi stage fusion technique which is based upon Gaussian and Laplacian pyramids.

Experimental results show an improvement on the already available fusion based techniques. Also we provide a comparison with two other methods.

Contents

List of Figures

List of Tables

Abbreviations

ACE	Automatic Color Enhancement
AG	Average Gradient
ARC	Automatic Red Channel restoration
CLAHE	Contrast Limited Adaptive Histogram Equalization
DCP	Dark Channel Prior
EME	Enhancement Measure Estimation
EMEE	Enhancement Measure Estimation by Entropy
HE	Histogram Equalization
log**AMEE**	Logarithmic Assessment by EMEE
MSE	Mean Square Error
QTS	Quad Tree Subdivision
RGB	Red Green Blue
SNR	Signal to Noise Ratio
STD	Standard Deviation
UICM	Underwater Image Colorfulness Metric
UIConM	Underwater Image Contrast Metric
UIQM	Underwater Image Quality Metric
UISM	Underwater Image Sharpness Metric
WB	White Balance

Chapter 1

Introduction

Similar to light traveling in the air, the underwater light propagation suffers from scattering and absorption. Nonetheless, the magnitude of absorption and scattering is enormous. While the light attenuation coefficients in the air are measured in inverse kilometers for an underwater environment it is in inverse meters. Such severe degradation of light poses serious challenges for imaging sensors to capture the information of the underwater area of interest. Unlike air, water is only transparent to the visible part of the electromagnetic spectrum and opaque to all other wavelengths. Furthermore, the constituent wavelengths of the visible spectrum are absorbed in different rates with longer wavelengths are absorbed more rapidly. The decay of light energy in water is truly remarkable. In the crystal clears waters of the middle oceans less than 1% of light energy remain by the depth of 150m. Hence the visibility degradation is such that the object is harder to see beyond the 20m range and in turbid coastal waters the visibility falls below the 5m mark. Also no natural light from sun reaches below 1km of sea.

Hence, the amount of light with in water is always less than the amount of light over the surface of water. Therefore images obtained under water generally have low visual quality. The scarcity of light under water is usually because of two unavoidable facts. One, the light under water is loses its true intensity, and second the chances for scattering of light within water is quite high. The immediate impact of this insuffcient amount of light is the color distortion and illumination

Table 1.1: Loss of light energy at 1m depth in ocean

Light Energy decay at 1m depth of ocean								
	Violet		**Blue-green**			**Yellow**	**Orange**	**Red**
λ (μm)	0.3	0.45	0.45	0.51	0.53	0.59	0.62	0.7
open seas, clearest	17%	3%	2%	3%	5%	8.5%	30%	41%
open sea, turbid waters	57%	16%	11%	10%	13%	19%	36%	55%
near coastal sea, average		64%	38%	30%	29%	31%	46%	75%

of the under water scene visibility. Two of the most deteriorating effect on underwater image quality are the absorption of light energy and random path change of the light beams and at travel in water medium filled with suspended particles. The part of light energy which enters the water is rapidly absorbed and converted into other forms of energy like heat which in returns makes the water molecules get energized and become warmer and tend to evaporate. Also some of the light energy is used up by the tiny plants based organism which use it for photosynthesis. This absorption degrades the true color intensity of the underwater objects. As explained in the previous paragraph part of light which is not absorbed by the water molecules may not travel in straight line but follows a random Brownian motion. Presence of suspended matter in water is responsible for this. Water, particularly the sea water contains dissolved salts and both organic and inorganic matter which reflects and deflects the light beam in new directions and it may also leave the water surface and decay back in air. Light scattering of light in water medium is further divided into two categories. The forward scattering; which is the light beam deflected after hitting the object of interest and reaching the image sensor. This type of scattering usually makes an image appear blur. The other type of scattering is the back-scattering which is the light beam hitting the image sensor without first reflecting back from the object. In a sense this the light energy which is not carrying any information about the object. It just adds to degraded contrast of an image.

One way of improving the visibility underwater is by introducing the artificial light source, although it adds its own flavor to the problem. Apart from the problem mentioned above like scattering and attenuation of light, artificial light tends to illuminate the scene of interest in a non-uniform manner and usually produce a bright spotting at the center of and image with darker shades around it. In addition to this the lighting equipment is heavy and costly. Also they required constant supply of electricity either in form of batteries or wired from the surface ship.

Henceforth, the underwater imaging system not only affected with the low light conditions, severely degraded visibility, diminishing contrast, burliness, light artifacts but also restrict color range and random noise. Consequently, the standard image processing techniques which serve us well for terrestrial imaging enhancement need to be modified or abundant completely and we need to come up with new solutions. And by improving the quality of underwater images can lead to better image segmentation, an improved feature extraction and better underwater navigation algorithms to steer autonomous underwater vehicles. Also offshore drilling platforms may also benefit from the clearer imagery for assessing the structural strength of the underwater part of their rigs.

1.1 Characteristics of Underwater Images

Unlike conventional imaging taken above sea in open air, underwater photography shows a strong dominance of bluish and greenish colors. On the other hand, the strong attenuation of light in the water with respect to the air and a greater diffusion of the incident light have the consequence of considerably reducing the visibility. Thus, objects at distant distances from the acquisition system or the observer but also at medium distances, or even relatively short in some cases, are hardly visible and poorly contrasted with respect to their environment [1]. In addition, in the presence of particles suspended in water (sand, plankton, algae, etc.), the incident light is reflected by these particles and forms a kind of *inhomogeneous* mist that adds to the scene observed. This turbidity of the water, most often white, also affects the visibility but also the colour dynamics of the objects contained in the image by tarnishing or veiling them. On the other hand, the formation of an underwater image is highly dependent on the nature of the water in which it was acquired. Natural waters can have very varied constitutions in terms of plants or minerals dissolved or suspended in water. The behaviour of the propagation of light in such a medium is strongly governed by this factor [2].

1.1.1 Bio-optical Properties of Natural Waters

Natural waters and their IOPs depend on the various elements that go into their compositions. While clear waters will mostly diffuse blue light, organic-rich waters will emit a greener, sometimes even yellow. Numerous measurements have been made which have made it possible to establish a link between the optical properties of absorption and diffusion and the chemical nature of the main components of the water [3].

Bukata, et al. [4] have established a first approximation in a classification according to the total concentration of chlorophyll-based pigments including phytoplankton. A second component corresponding to Dissolved Organic Matter, or yellow substance, was added later improving the model. From optical measurements and chemical measurements, spectral attenuation curves for different concentrations were obtained. An attenuation model was then established by regression of these data, making it possible to define a function expressing the attenuation coefficient directly from the concentrations.

1.1.2 Components and Mitigation by Water

The first component of natural waters is the water molecule itself. If we consider the total attenuation coefficient, it is possible to decompose this term into several components corresponding respectively to attenuation by pure water and to different types of particles. In fact, the components that most affect attenuation in natural waters are respectively the chlorophyll-based pigments present in some living cell organisms in suspension and the dissolved organic matter. These relatively

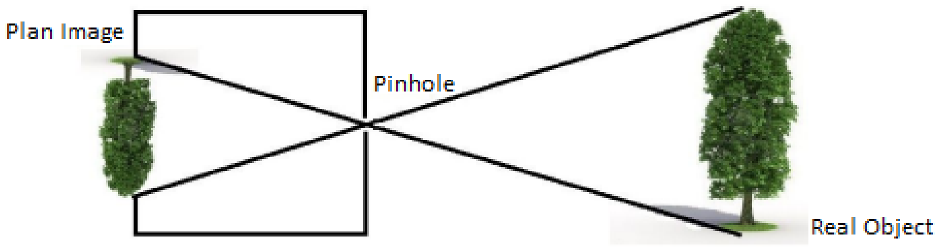


FIGURE 1.1: Pinhole principle

large particles scatter more light at higher wavelengths than water, thus affecting the colour perceived by the observer (Gibson, 2015). According to Chiang and Chen, (2011) it is possible to restrict to three coefficients corresponding to water, pigments and dissolved matter.

In conventional cameras the stenop'e hole is replaced by a lens which, thanks to a larger surface area and its light-focusing property, allows the lens to be illuminated. Get a brighter image and have a better clean. The elements that make up a basic camera are the following: one or more lenses, the image plane (film, CCD, CMOS), a focusing device that makes it possible to move the lens with respect to the image plane, the diaphragm that controls the amount of light received by the lens and the shutter that controls the aperture time of the diaphragm [5]. The quality of the image observed will depend on sensor optics, electronics (amplification, quantification and sampling) as well as the environment in which the light will propagate before reaching the objective and lighting conditions [6].

1.2 Mathematical Model for Underwater Images

There are several mathematical models to describe the formation of an image underwater. For this thesis the mathemcatical model we used is described below. It constitute the summation of three parts:

$$E_T = E_d + E_f + E_b \tag{1.1}$$

E_T is the total light energy falls on the image sensor which forms the image. The first variable on the right side of the equation is the direct transmission E_d of the energy of the surface of an object E_o. The E_f is the forward scattered light which although reflected of the object but suffer from deflections before entering the sensor. E_b is back-scattered light which falls on the sensor but carries no image information. The light beam after reflecting of the object is further attenuated and as explained before the absorption is depended upon the wavelength of light. and therefore creates a false color map for the object under observation. Mathematically the reflective component can be described as:

$$E_d = E_o e^{-c_\lambda r} \tag{1.2}$$

Here the variable $\boldsymbol{r}$ is a length of space between the sensor and the object and $\boldsymbol{c}$ is a attenuation coefficient constant. Subscript denotes that c depends on the wavelength. This equation also explains why white balance may only be used to correct registered colors if all registered points are approximately equidistant to the camera. In this case the shift in color is similar for all the points and therefore white balance may succeed. However, it is important to remember, that applying white balance to the image will not restore true colors - it will make the image look more naturally, but it is based on some general assumption (e.g. so called grey world assumption) which may not be true for a given picture a therefore one should not expect to get an accurate color reconstruction. Back-scattering, also called the veiling light, happens when the ambient light present around the sence of interest is reflected and deflected without touching the object of interest and enters the camera. It is directly responsible for making the image appear foggy or low contrast. In a scene the back-scattering can be compared to making an image in a foggy or hazy environment. In case of total absence of natural light and object is being observed with the help of artificial light only it is important to keep the light equipment away from the camera so the light cone does not creates bright light spots for the image. Back-scattering created by virtue of natural light

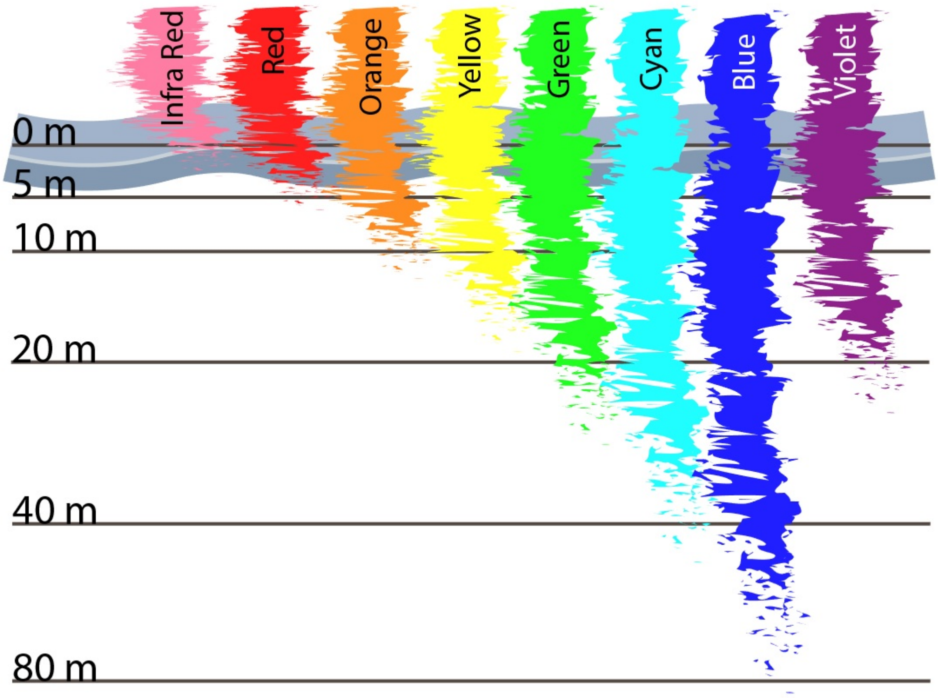


FIGURE 1.2: Attenuation of light underwater

can be mathematically descried as:

$$E_d = B_\infty(1 - e^{-c_\lambda r}) \tag{1.3}$$

Here B_∞ describes the veiling light signal. we can think it of as the amount of light entering the sensor if there was no object or scene to observe just the open waters extending to the infinity. Furthermore, this signal carries zero information. Although it dependent upon the natural lighting present in the scene; greater the distance between the object and the sensor bigger the amount of back-scattered light is available to fall on the sensor and lower the value for the contrast would be.

Since the reconstruction of the image is the main goal and empirically the light absorption, back-scattering and forward scattering affect the image formation in there own manner. The light absorption reduces the overall amount of light available and reduces the color intensity of an image but does not effects the other qualities of an image. Haziness is primarily caused by the forward scattering of light. And lower contrast is result of back-scattering. It is important to note that the major reasons for underwater image degradation are light absorption and back-scattering and forward scattering is not a whole lot responsible for an image with lower quality.

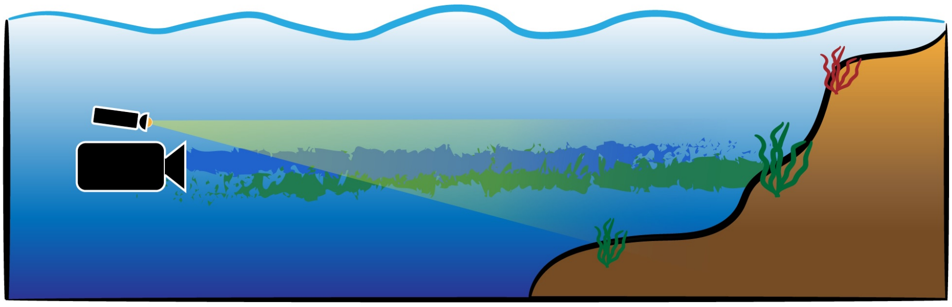


Figure 1.3: Light Propagation from object to camera. Green describes the light travelling from object to the camera and attenuating exponentially on it way. Blue is the back-scattered light which is responsible for the lower contrast and gets stronge with the increase in distance.

1.3 Motivation

Software based underwater image enhancement techniques are usually work by controlling some aspect of the mathematical model of underwater to compensate for the degrading effects introduced by the water's light absorption and the presence of organic and inorganic particles in water.Current state of the art method for underwater image restoration are typically designed for a single image input as using multiple images for the processing usually require more computational resources and may not be suitable for the real-time applications. Amongst the single image enhancements method, those based on the image fusion methodologies shows the most promising results [3]. The general principle followed by fusion based techniques is splitting the input image into two, processing the split images separately calculating the weights based on the features of each image and finally using the calculated weights to fuse the images into a final result. Although this algorithm does produce good results for most of the cases, it does suffer from over compensation of the color and sometime distort the contrast in a negative way. Different variations of this techniques exists and there is room for improvement. Specifically, improving the color degradation part of the method can make use of the linear function proposed by [7]. Thus combining the color improvement algorithm with the image fusion principle could improve the overall quality of the enhanced image even more.

1.4 Objectives

The objectives of this research is to develop an improved single image enhancement technique with is able to work under a variety of conditions and using a standard formulation for assessment of underwater image quality. Specifically, the goals of this research are as follows.

- Analyze and study the effects on light propagation under water and and its impact on the degradation on image quality.
- Studying the decomposition of a single degraded underwater image into its constituent parts for independent processing to improve color and enhance contrast and finally fusing the processed images into a single output image with better visual qualities.
- Comparing the results of the proposed solution using established techniques with other state of the art solutions to the problem of enhancing underwater images.

1.5 Problem Statement

This is a general observation that the quality of an image taken with in water is always degraded. It loses the actual tonal quality and the contrast necessary for distinguishing the object of interest present in the image. The situation becomes more challenging when the neighboring objects have very minor differences in pixel intensity values. This situation poses a serious challenge to extract finer details from the data and reduces the performance of the algorithms used to extract information from the images. Therefore, there is a pressing demand for images taken underwater much be processed in such a way that they should represent their true tonal details. Underwater imagery has wide range of applications for example investigation of aquatic life, quality of water, defense and security purposes etc. therefore images or videos obtained for meeting these objectives much carry the exact details.

Chapter 2

Literature Review

2.1 Introduction

This chapter presents a comprehensive review of literature focusing on various techniques that are used to enhance under water image and restore visibility and quality of the image. There is a wide variety of technologies as well as techniques used for underwater image restoration and therefore the following discussion categorises and groups them in relation to their distinguishing features.

2.2 The Processing of Underwater Images

A large number of underwater image processing methods have been developed. Different approaches have been considered: inversion of the light propagation phenomenon, colour filtering, frequency filtering, etc. In this study, we propose to expose approaches based on colour filtering methods based on the physical model on the one hand, and colour consistency methods allowing unsupervised processing. These methods rely on an adjustment of the luminance of the pixels of the image on each of the colour channels to improve the contrast and colours of the image [8]. Existing methods of processing by the physical model require knowing one or more references allowing the establishment of the processing chain, as well

as some knowledge on the conditions of acquisition. Chromatic adjustment methods are unsupervised but provide a chromatic correction that is not necessarily in keeping with reality. The goal is therefore to reconcile these two aspects: unsupervised and chromatic adjustment based on the physical model and there four methods that have particularly caught our attention because they provide good correction results [9]
Several methods based on the propagation model exist. Based on an estimate of the attenuation, they use controlled acquisition conditions. These systems make it possible to estimate the spectral attenuation and diffusion coefficients necessary for the model either by having reference in the image or by using more elaborate specific systems than the simple camera. This section presents two methods corresponding respectively to each of these two categories [2].

2.2.1 Underwater Image Correction based on Spectral Data

This correction method, proposed by [10], is based on an estimate of the attenuation coefficient. The estimation is performed using known reflectance values of a grey reference target present in the images considered. This object, named Spectralon, consists of a plastic surface with the property of reflecting light very strongly by following a Lambertian reflection of almost perfect light [10]. On the other hand, the camera used, its properties and its position are known. Thanks to a priori knowledge about the content of the scene, the photometric behaviour of the objects and the acquisition system, the luminance reaching the spectral is known as well as the value of the luminance reaching the camera for each of the three chromatic channels. In this method, the camera is placed vertically at the ocean floor. Thus, the size of the water column corresponds to the depth. The luminance can therefore be written as a function of the depth represents the luminance received by the camera. The author makes the following three hypotheses allowing correction [11]:

- The photographic seabed has a reflection following a Lambertian distribution;

- The spectral receives as much light as the surrounding environment;
- The camera has stable sensitivity curves with respect to illumination variations.

Having all these elements, it is possible to express the attenuation coefficient as a function of depth and luminance. The reflectance at a given point of an object corresponds to the ratio of the incident luminance (incoming) Lin received by the spectral and of the diffused luminance (outgoing) Lout received by the camera [12] Having measured the depth and the attenuation coefficient, Peng, Zhao, and Cosman, (2015) applied the Beer's law, for all the pixels of the image and to find the value corrected pixels. We could not reproduce the experiment, which involves an acquisition campaign with the reference object. As far as one can judge by the results given in the correction is of good quality, from a visual point of view. However, the study is limited to clear water and photographs taken in shallow waters. It is likely that high turbidity of the water causing white scattering of the incident light would distort the estimation of the attenuation coefficient and background luminance. Moreover, this method requires mastered acquisition conditions (illumination, depth, position of the camera, etc.).

2.2.2 Underwater Image Correction with Polarization Filter

This second method, described in [13], is based on the use of a light polarizing filter. One of the problems of inversion of the physical model is to estimate the distance between the objects and the camera. Due to attenuation and diffusion, the visibility in the images is very small. The purpose of this method is to increase visibility. The authors therefore propose the acquisition of two images with different polarizations in order to recover additional information. The authors first describe the mechanism of image formation at the camera level in order to then implement a treatment to compensate for the effect of diffusion to increase the visibility in the image. This method does not rely solely on the inversion of

propagation but also on the nature of the acquisition system and its properties[1]. By relying on the light propagation model, the signal is considered as the sum of the direct transmission D, i.e. the attenuated light coming from the objects, and the forward scatter, in other words the ambient light scattered in a direction forming an angle close to the line of sight of the camera. Thus the signal corresponding to the image is measured and the direct transmission is attained. Forward scatter makes the image blurry. The authors use a special function called Point Spread Function (PSF) to compensate for this phenomenon. A PSF that can be used in the underwater setting accounts for the distance, the inverse Fourier transform and the spatial frequency in the plane image [14].
Generally, a degraded image may negatively impact its interpretation by the human eye as well as the performance of a computer vision system. In this section, this study shows the phenomena that come into play in the different stages of the scene acquisition chain and that can alter the quality of the captured image. The chapter then moves on to describe the methods used to improve the quality of the image [15].

2.2.3 Polarization

Underwater diffusion involves polarization effects. The method exploits these effects to compensate for the degradation of visibility. Considering a light source illuminating the particles of the line of sight, an incidence plane is formed by a ray coming from the source and the line of sight. The backscattered light is partially polarized perpendicularly to this plane. For this reason, typical natural backscattering in the underwater environment is partially horizontally polarized [16]. In order to measure the different polarization components, the scene is acquired through a polarizing filter. Since backscattering is polarized, its intensity depends on the orientation of the filter around the optical axis. There are two orthogonal orientations for which the transmittance of the backscattered light reaches maximum values Bmax and Bmin. Thus, there are two linear polarization components [17].

When the polarizer is mounted, the intensity of each pixel in the image depends on a cosine function of the orientation angle. Similar to backscattering, there are two intensity extremes, and the visibility enhancement algorithm compensates for the haze effect caused by the broadcast [18].

2.2.4 Total Variation Model

From the local characteristics of the image a global regularity is measured. This regularity is called "total variation" and is calculated as the sum of the local gradients of the image. In the case of an additive noise, the image observed. The image noise is made by looking for the image. TV is used as a term of regularization which allows to penalize the big variations and to allow the discontinuities along the sufficiently regular outlines. The denoising force is controlled and the larger it is, the smaller the total variation of the resulting image [19]. The disadvantage of this type of smoothing is that the textures can be considered as noise and be erased. Yang, [20] also observed the creation of staircase effects. This method, in addition to noise suppression can also be used to restore images degraded by fuzziness. Note that the "Total Variation Model" method can be interpreted as a special case of the Bayesian approaches.

2.2.5 Bayesian Approaches

The Bayes formula expresses the posterior distribution of having an image in there is a random noise. Denoising is maximizing this probability. The probability is known and plays the role of a normalization constant, and is the likelihood and is determined from the model of data formation. The maximum likelihood (ML) estimate consists of looking for the value of that maximizes the likelihood. For reasons of simplicity, it is preferable to estimate the log of this product of probabilities. In general, the maximization is done by looking for image which satisfies the probability density in the case of a Gaussian noise. The Maximum a posteriori (MAP) estimation has the advantage of being able to take into account

the prior. The MAP consists of finding the image which maximizes, which gives, by applying the logarithmic function to the Bayes formula [21].
Markov fields are often used to determine the prior probability. For this reason, the image is considered as an arrangement of atoms found on several energy states, the grey levels. The state of each atom depends only on its neighbors. The Hammersley-Clifford theorem gives the expression of this probability in which there is a normalization constant, the sum of potential functions computed on cliques (neighborhoods). A priori corresponds to the choice of cliques and potential functions (differential operator for example) [22]. To determine the MAP, there are two families of methods [13]:

- Deterministic algorithms that are fast, but may converge to a local minimum far from the global minimum. We mention for example the gradient descent algorithm and its variants.
- Stochastic algorithms that are slow but provide convergence towards a global minimum.

2.2.6 Transformed into Wavelets

Wavelets were introduced in the early 1980s, in order to overcome a problem related to the Fourier transform which does not allow to locate the frequencies of the signal in the time. Denoising consists in keeping the coefficients cmn and jo mn having a significant value, considering that the low values correspond to the noise, then there is a reversing equation in order to recover the image without noise [23]. In order to recover the coefficients which interest analysts, a threshold must be found to detect the coefficients corresponding to the noise, for which a great diversity of methods exists (strong, soft thresholding). Note that there are other methods belonging to the same family as wavelets (time-frequency representation) such as curvelet or contourlet. In addition to noise-related problems, the image may suffer from loss of contrast [24].

2.2.7 Contrast Enhancement

Incomplete or very weak illumination may induce loss of contrast in the image. In the case where the information of the contrast is always available, there are several methods allowing its enhancement of global or local way. Contrast enhancement is very useful for many computer vision applications such as image segmentation or pattern recognition.

2.2.7.1 Histogram Equalization

The purpose of this type of approach is to modify the histogram of the image by assigning new values to the pixels of the input image. The histogram of images with low contrast occupies a small portion of the intensity range. The goal of equalization is to spread the histogram over a larger beach. For this, from the histogram of the image, the approach calculates the cumulative histogram and applies it (after normalization) to the image in order to spread its histogram uniformly over the entire range of dynamics [1]. There are also other functions such as logarithmic, exponential, power and others to obtain a histogram with a certain shape. Histogram equalization often gives better results when applied locally [25].

2.2.7.2 Local Histogram Equalization

In many cases, the histogram of the image covers a broad dynamic. In this case a local histogram equalization is necessary to bring out the contrasts of the different parts of the image. For this, the image is scanned with a small window and the equalization principle described above is applied to each window separately. Then, in order to eliminate the generated block effects, due to the difference of the histograms between neighboring blocks, a bi-linear interpolation is used [1]. This method is called Contrast Limited Adaptive Histogram Equalization (CLAHE). The defect of this type of method is the over-amelioration of contrasts: it brings out false details. Because of the local character of the method, it requires more processing time than a global equation [26].

2.2.7.3 Retinex

Retinex is the combination of the words R'etine and cortex. The method is based on the observation that the human visual system perceives the contrast and colour of an object relatively in the same way under different illumination conditions. This is not the case for camera sensors because the intensity value of a pixel depends strongly on the photon flux. The objective is to build, from a given image, a new image illuminated by a constant white light [27]. Retinex has a scale that applies a nonlinear operation to the logarithmic input image. There is also Multi Scale Retinex (MSR) which as the name suggests, a combination of several retinex (usually 3) made at different scales (different sizes). Experimentally, it has been shown that a uniform weighting gives good results. The last step of the algorithm is normalization which brings the result back to the definition interval of the image using an affine operation. The retinex algorithm is simple and automatic but requires a large signal-to-noise ratio to obtain a satisfactory result [28]. On the other hand, in order to improve the processing time and to be able to process large images more quickly, it is customary to replace the convolution in the spatial domain by a multiplication in the frequency domain. In the following, we tackle another problem which is the white balance [29].

2.2.8 White Balance

In some cases, the white areas in the scene appear coloured in the pictures. Generally invisible to the human eye, which is able to do the correction automatically, it causes extremely saturated colours after the application of a restoration method. This problem is particularly sensitive in underwater conditions during the restoration [12]. Following are some methods that are interesting to improve the images before the application of restoration methods.

2.2.8.1 Gray World

The Gray World is one of the oldest methods used to perform a white balance. It considers that the average intensities of the three RGB channels must be equal. The principle is to keep the green channel unchanged and to multiply the red and blue channels, respectively by the gains. Some authors such as Wong, [30] has proposed a combination of this method with Retinex.

2.2.8.2 Max White

This approach assumes that the maximum intensity on each channel is white. The gain of each channel in the case of 8 bit image is calculated in a simple way. In the case of underwater images, a white balance can be used to obtain better restoration results. There are other ways to improve the colours of the image such as histogram matching which uses two images, one of which is not degraded, which serves as a reference for improving the colours of the other. Methods using a combination of the methods presented in this section are also used [12].

2.2.9 3D Reconstruction: Systems and Methods

The first step in 3D structural reconstruction is the acquisition of shapes that can be processed in different ways. There are various types of systems as explained below:

2.2.9.1 Non-optical Systems

The robustness of the non-optical sensors (Radar, Sonar, etc.) and more particularly their range and their resistance to environmental variations (climate, fog, brightness, etc.) makes them well adapted to operate in difficult atmospheric conditions as in the case robotic applications in terrestrial outdoor environments [31]. For example, the Massot-Campos, and Oliver-Codina [32] introduced a geometric method for 3D reconstruction of the external environment using a camera and a panoramic microwave RADAR. They relied on the complementarity of these two sensors, given the robustness of the radar and its ability to generate a depth map and the high spatial resolution of the camera. In the same context, several works exist in the literature exploiting the use of SONARs in combination, in particular, with the "Time-of-Flight" methods, especially for long distances. These methods are based on the analysis of the propagation time of the signal which, knowing its speed in the medium in which it moves, allows us to estimate the distance of the object that generated an echo. These sensors are commonly referred to as profiling sonars because they are mainly used to collect bathymetric data. Nevertheless, the use of these devices in marine environments is less advantageous because of changes in water temperature, salinity and pressure [33].

2.2.9.2 Active Optical Systems

The extraction of primitives from structured light images has been used in several works such as [22]. These systems consist mainly of a video projector that projects a sequence of specific patterns of light (black and white or grayscale stripes, coloured line patterns, specific targets, etc.), while one or more Digital cameras record the distortion of projected patterns on objects. Different techniques are proposed for triangulation 3D reconstruction [22]. In Xi, Rauschenbach, and Daoliang [34] the combination of image-based techniques and information from the laser system provides a very reliable reconstruction. In front of the laser, a diffractive optical element is inserted allowing to modify the shape of the beam in a set of parallel lines. These lines are projected onto the stage and recovered by the camera. Once the relationship between a pixel and the laser plane is known, the 3D information can be calculated by triangulation. The major advantage of these devices is the high resolution and accuracy obtained leading to the high quality 3D surface reconstruction. Nevertheless, the main disadvantage is that they are limited to static objects and a large number of images must be acquired [34].

2.2.9.3 Passive Optical Systems

The reconstruction of the 3D form of objects from multiple images is one of the oldest issues in the field of computer vision. The literature offers few real-time methods, which estimate precisely the 3D form of objects filmed by one or more cameras. Reconstruction approaches that use a single camera rely on the extraction of certain characteristics of the image to determine the depth information [35]. In the field of archaeology, the use of structured light or 3D scanning methods have been recommended for better reconstruction accuracy provided that the object can be transported and oriented so that it can capture all its faces. In the opposite case, it is the stereo-vision based methods that are used the most. Stereoscopy is based on the extraction of characteristic points from the same scene from images acquired simultaneously by two cameras. These points are matched to calculate their corresponding 3D. Indeed, by knowing the relative position of the camera and the location of the same characteristic point in the two images, the 3D coordinates of this point in the frame of the scene can be calculated by triangulation. These methods, however, have the disadvantage that the accuracy is often limited because many sites or archaeological objects are rather without texture and the problem of correspondence becomes difficult (Johnson-Roberson, et al., 2017).

In order to overcome the problem of correspondence, Sarafraz and Haus [36] proposed to combine stereovision methods with "Shape-from-X" and photometry techniques. The word "Shape-from-X" is a generic name for image extraction techniques only. Shape-from-Shading (SfS) is a reconstruction method that improves the shape of an object from shading information from a single image [7]. The actual number of applications of this method is rather limited and is mainly due to the fact that the directions of the light sources and the reflection properties must be perfectly known to achieve good results. In this context, refinement algorithms have been proposed to improve the quality of reconstruction. Shape-from-Motion (SfM) is a method of triangulation that involves taking multiple images of an object or scene using a single camera. From the camera orientations, the image characteristics are detected and compared between consecutive shots to estimate

the relative movement of the camera and, therefore, determine its trajectory in 3D. This approach is the cheapest in terms of material and the easiest to achieve. The only restriction is to have enough memory on the camera or video recorder to save a full view of the scene [37].
Photometry is a reconstruction technique that requires a reduced number of images captured under different lighting conditions. A sphere is placed in the scene for calibration of the light source when capturing data. By changing the location of the light source, the 3D information is obtained while keeping the camera and the object in a fixed position. Since the depth of a scene in an image is variable, the volume of the illuminated particles is also variable. This implies that, under scattered light, backscatter is directly related to the depth of the scene which allows us to estimate a depth map [38].

2.2.10 3D Reconstruction

Although there is a large body of work, in the literature, on omnidirectional camera calibration, location, structure estimation from motion or SLAM(*Simultaneous Localization and Mapping*) applications, less numerous are the research that has been conducted in 3D reconstruction with catadioptric cameras. In submarine applications, the nature of the aquatic environment has refractive effects of the light passing through it. Several studies have been carried out in the same context, for example by Nichols [39], showing the relationship between the good estimation of navigation conditions and the 3D modeling accuracy of the structure of the environment. In addition, 3D scene rendering algorithms must take into account refraction effects either by simulating an estimate of the vision system parameters from underwater images, or by inserting these effects into the tracking system rays crossing the various media from the 3D point to its 2D projection on the image plane. Obtaining the depth information from the 2D data can be done by various methods [40].

2.2.10.1 Epipolar Geometry

Epipolar geometry describes the relationship between 3D points and the positions of corresponding points in a pair of images. It can be established from a few image matches and is used to simplify the search for more matches, calculate the movement between cameras, and reconstruct the scene [41], the researchers studied the epipolar geometry of central catadioptric systems. They show that the epipolar lines correspond to conics in the omnidirectional image that are reduced to radial lines in the case of vertically aligned catadioptric cameras. Several methods take the advantage of this configuration either by a combination of two catadioptric cameras one above the other or by using a double mirror with only one camera (stereo pairing on the same image).
The panoramic system, is composed of a catadioptric camera and a light projector. The 3D reconstruction method used is based on epipolar geometry to define relations between the corresponding points in the stereo images of the same scene (Remijan, 2018). The 3D coordinates of a point of the real world are calculated using the distance between two corresponding points which depends, in fact intrinsic and extrinsic parameters of the vision system. A 3D representation is performed using various software [42].

2.2.10.2 Triangulation

The 3D position of a world point can be calculated using classical triangulation as traditionally done in a stereo-vision system. Triangulation is a technique that makes it possible to determine the depth of a point of space, from its pixel projections in the two stereoscopic images. In its simplest form, a 3D point is associated with two unique matched pixels, each located on an image of the pair of stereo images [43]. In [44] a triangulation is applied on the pairs of points extracted from the stereo images in order to obtain an estimate of the 3D structure, by using the geometry of the stereo system and the internal parameters of the camera. In order to improve the 3D structure of the objects, a rectification of the stereo images is performed and a dense depth estimation algorithm is applied. The depth map is

transformed into a surface mesh using the Delaunay triangulation algorithm. The 3D representation is then visualized in a more realistic aspect by providing a 3D model with texture mapping. The reconstruction process is performed offline from a set of underwater stereo images since it requires a long computing time [44].
Onmek in [45] presented a navigation system is developed to create 3D models of real objects photographed in an underwater environment. The algorithm is a succession of several processing processes. It uses the stereoscopic information to create a disparity map and generate the 3D positions of the 2D feature points. Between two successive acquisitions, an algorithm based on the RANSAC approach is applied to estimate the movement of the camera. The rotation and translation of the system are optimized using the Levenberg-Marquardt minimization algorithm. The resulting point cloud model is then converted to a volumetric representation and a textured polygonal mesh is finally extracted.

2.2.10.3 Geometry of the Ray Trace

Several studies on underwater imaging are developed incorporate the properties of light refraction in water. Three-dimensional shape / structure reconstruction can be achieved using specific imaging models and numerical techniques. The underwater environment, not being a homogeneous medium, makes that ray of light does not follow a linear trajectory during its passage from one medium to another. Indeed, the refractive index depends on the celerity of the light ray which varies from one medium to another. The latter undergoes, at each separating surface, a refraction which depends on several factors such as temperature, salinity and water pressure. In addition to these characteristics, the refractive index also varies considerably with the colour, the wavelength of the radiation considered [46].
When the calibration of the vision system is performed in the air, the refraction effects are integrated into the 3D reconstruction algorithm. In this context, the ray tracing algorithm is the most recommended in literature. Indeed, this process allows the precise reconstruction of the scene by the modeling and the follow-up of the refraction of the light. It is well suited to complex refraction across multiple interfaces and requires a complete knowledge of the 3D projection process [47]. In

Łuczyński, Pfingsthorn, and Birk [48], a system of catadioptric cameras immersed in an aquatic environment was used for the acquisition of omnidirectional images. The 3D coordinates of the vector ray crossing the air and refracted on the surface of the mirror is determined from the characteristic parameters of the camera (calibrated in the air), the intrinsic parameters of the mirror and the image coordinates. The equations giving the relationship between the refracted vector rays on the different surfaces (air / glass and glass / water) can be determined by the Ray-tracing method.
The 3D rendering algorithms, presented in [49] are inspired by the geometry relationships of the ray trace. In general, they consist of using the geometry of the vision system to switch from 2D to 3D information. This is done by following the trajectory of the light beam carrier vectors and using Snell laws to determine the refractive angles corresponding to each change of medium. In Shortis, [50] a new calibration algorithm in both water and underwater was presented. The study of the effects of distortions is based on the ray tracing approach as well as the refraction expressions of the light beam crossing the different media. The angles of refraction are determined by applying Snell's law.

2.2.11 Markov Random Field (MRF)

Torres-Méndez and Dudek [51] applied the Markov Random Field (MRF) for colour restoration of underwater images and concluded that Markov fields can be used in unobservable random field estimation problems. In the context of image processing, this modeling makes it possible, in particular, to perform statistical segmentations of textured and noisy images by correlated noise. Unlike the hierarchical model MRF based model makes it possible to apply the Bayesian methods MPM and MAP without resorting to any approximation. The main difference with Hidden Markov Field Modeling lies in the fact that the a priori law of the random field of classes is not necessarily a Markov law (so the model is not a hidden Markov field because the field hidden is not necessarily Markov). Anderson and Smith, (2018) concluded that it is possible to synthesis textured class images on the one hand, and their segmentation by the Bayesian MPM method, on the other hand. The simulations show that one can obtain, in a framework of relatively simple models, as well different homogeneities of the images of classes, as different correlations of the noise, giving visually different textures. In spite of the visual importance of the noise (it is difficult to distinguish the classes in the noisy images) the segmentation by the MPM gives encouraging results. The application of the proposed model in segmentation of textured images and, possibly, noisy by correlated noise, real would require a method of estimating its parameters. The search for such methods is a natural perspective for the continuation of our work.

2.2.12 The Dark Channel Prior Algorithm

Chao and Wang [52] used the dark channel prior to recover the original clarity in objects in underwater images. The technique was used to solve the scattering of particles. The study showed that the existence of purplish pixels is mainly due to the scattering of incident light. In fact, when the fog sensation is strong at one point, the luminance of the pixel is strong due to the scattering of the incident light. When the inversion of the attenuation is performed, the values of the red channel are greatly increased while the additional luminance due to the incoming scattering is not subtracted [53]. In order to remove this purplish appearance from the corrected images, the value of all the pixels of is adjusted by adding to the correction the effect of the luminance and the red diffusion coefficient in the propagation. This procedure is applied through an iterative process to successively reduce and minimize the number of defective pixels detected. The algorithm employed is based on a least squares method seeking to minimize the cardinal [54]. Schettini and Corchs [55] conducted a review of various image enhancement techniques and concluded that there are significant challenges in obtaining objects' visibility at short and long distance in underwater settings. The study concluded that numerous approaches are available for image enhancement, but majority of them are limited to ordinary images and only a small number of approaches are particularly focused on enhancement of underwater images [56].

He, Sun, and Tang [27] used the dark channel prior algorithm to remove haze from single image and concluded that this method improves the quality of colour images. The study proposed a simpler implementation to transfer colour information to neighbouring pixels. For this it applied a simple morphological closure to the transmittance map obtained. The results showed that in places that do not contain colours there is a zero intensity that corresponds to the "Dark Channel". Note that no smoothing was performed, hence the artifacts generated on the contours. The defect of this method is that it applies only to images containing everywhere colors [53].

Chapter 3

Underwater Image Enhancement by Fusion

3.1 Introduction

To handle two of the major issues of degraded underwater images many software based techniques have presented and evolved over the years. Techniques based on decomposing the images into its constituent parts and dealing with the decomposed parts independently and finally fuse them to obtain a single images shows the great prospects.

In this chapter an improved underwater image enhancement technique is presented which is an improvement upon the image fusioning technique [57] [12] [3]. Starting from just a single image as input, it is first decomposed into two components the *Reflective* and *Illuminance*. At this stage a *Linear piece-wise* function is applied to the *Reflective* component of the original image. This application results an image with improved colors. For the *Illuminance* part a dehazing algorithm is used. Weights of each component are calculated separately and a multi-stage fusion process is followed to output a single improved image is produced with better colors and less haziness.

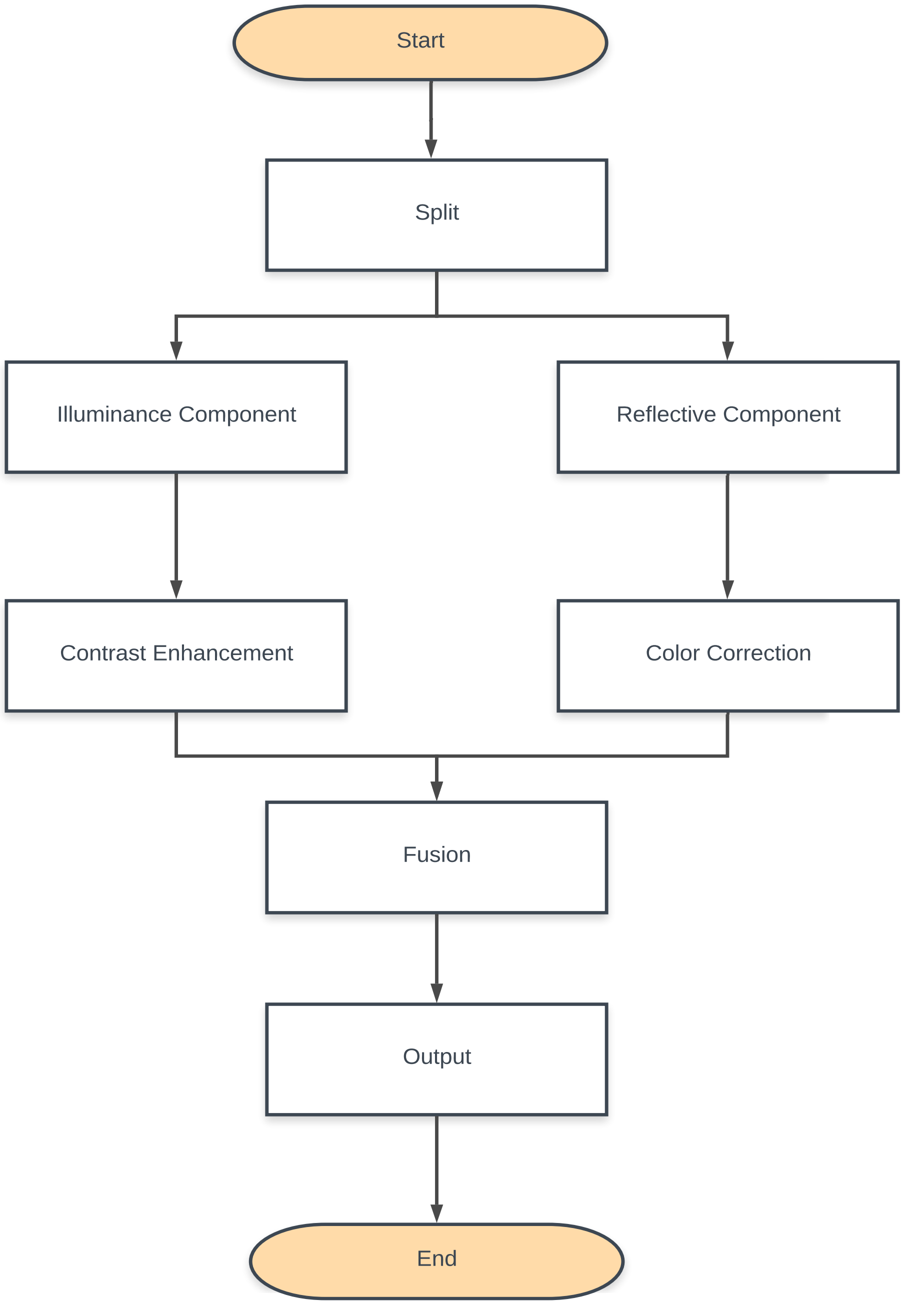


FIGURE 3.1: Block diagram of the Enhancement Process

3.1.1 Image Decomposition

As an important pre-process method, image decomposition is widely used technique to split images and videos into several parts and then execute different processing methods according to their features. Image decomposition methods like wavelet transform and weighted decomposition are most popular methods in this domain. For the case of underwater image processing, we want to separate images into several parts and each part contains a specific features obtained from original image, such as illumination and reflectance. Wavelet transform method is widely used in image denoising and enhancement, however, for this case, it is not suitable because it is time consuming and can not extract the features we want from the original image. Thus, an image decomposition method is used. This technique splits the input image into two of its constituent parts. one part describes the illumination part and other contains the reflectance part of the input image.

The image formation by the sensor is primarily rests on the intensity of light reflected back from the object under observation which is described by the reflective component. For the images taken in the clean air medium the intensity of white light is evenly presented and the average value of the image tends to get closer to the color gray as it should be according to the Gray world assumption. Nonetheless in case of hazy or blurred images the amount of light coming in the camera after reflecting off the object is attenuated. Hence for such situations (underwater imaging is one such situation) its important to decompose the input image into to reflectance and luminance parts and work on then separately. Input underwater image can be described as:

$$I_c = I_c^R(x) + I_c^I(x) \tag{3.1}$$

Here $c \in \{red, green, blue\}$ describes the RGB color channels, $I_c^R(x)$ represents the reflectance part of the image and $I_c^I(x)$ is the illumination part of the original image.

$$I_c^R(x) = \gamma \cdot I_c(x) \tag{3.2}$$

$$I_c^I(x) = (1 - \gamma) \cdot I_c(x) \tag{3.3}$$

Where γ is parameterized wieghted constant to keep the bright areas of the image bright and improve the contrast in the refectance componenet by minimizing the backsacttering light coming in the image. Mathematically γ can be derived as

$$\gamma = \xi \cdot \frac{I_c(x)}{I_c^{max}} \tag{3.4}$$

In the above equation, I_c^{max} represents the max channel value in a pixel while ξ is the parameterized constant to calculate the weight of the reflectance slice of the input image. The illuminance component and the reflectance component selection value depends upon the value of the control parameter ξ. If $\xi = 0$, we can treat the image as illuminance component, if $\xi = 1$, we can treat the image as a reflectance component. For further image processing, the backscatter effect can be considered for the illuminance component and color distortion can be considered for the reflectance component. suffers from color distortion.

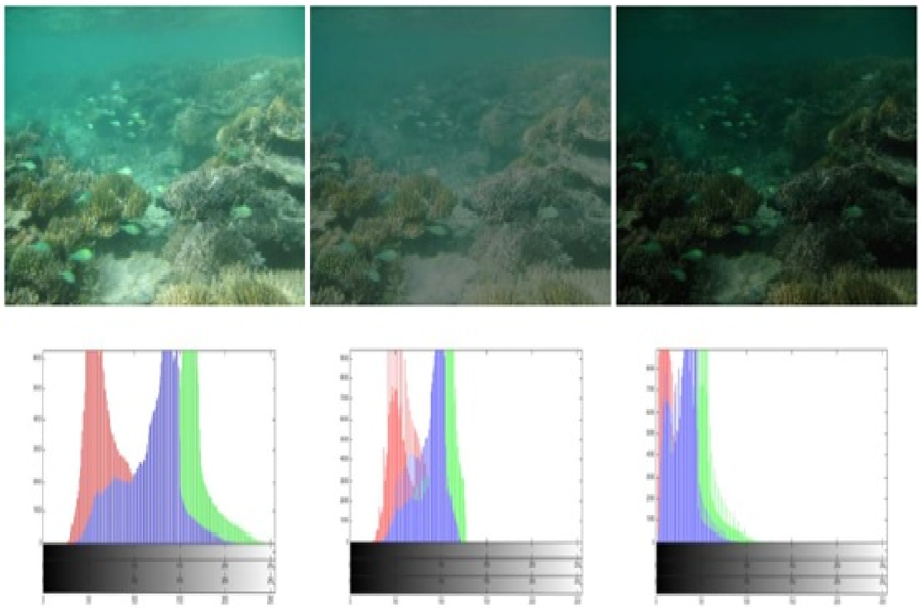


FIGURE 3.2: Original Image, Reflectance and Illuminance Component

3.1.2 Color Correction

Since the energy of light is exponentially attenuated the true colors of an object cannot be restored. The current conventional converters are failed to recover the lost color components. Due to poor intensity and dimness of light underwater, images cannot be clear. We prefer the hypothesis of gray-world for color correction which uses the linear transformation function. Color correction techniques which are built upon the gray-world assumption requires that the average of an image is achromatic or gray, hence the name Gray-World assumption. In the 8 bit image, the average turns out to be equal to 128.

A piece-wise linear transformation function is applied for stretching the mean of the image towards 128. S variable is the input image. For each color channel $C \in r, g, b$ the max, the min and the mean is computed. The mathematical expression of it as follows,

$$S_{CR}^{C} = \begin{cases} \left(S^{C} - S_{mean}^{C}\right) \frac{S_{min-128}^{C}}{S_{min}^{C} - S_{mean}^{C}} + 128, & S_{mean}^{C} \leq 128, \\ \left(S^{C} - S_{mean}^{C}\right) \frac{S_{max-128}^{C}}{S_{max}^{C} - S_{mean}^{C}} + 128, & S_{mean}^{C} > 128 \end{cases}$$

S_{CR}^{C} Shows the color corrected image while $S^{C}, S_{min}^{C}, S_{max}^{C}$ are the mean, maximum and minimum of the channel. The mean decides the direction to stretch. In some cases, due to the red color wavelength, the systems show an over correction by excessive stretching. An improved version of the system in present in [7] by adding one more equation like:-

$$S_{CR}^{C} = \begin{cases} S^{C} - \lambda\left(S_{mean}^{C} - 128\right), P^{c} > 0.7, \\ \left(S^{C} - S_{mean}^{C}\right) \frac{S_{min-128}^{C}}{S_{min}^{C} - S_{mean}^{C}} + 128, & S_{mean}^{C} \leq 128, \\ \left(S^{C} - S_{mean}^{C}\right) \frac{S_{max-128}^{C}}{S_{max}^{C} - S_{mean}^{C}} + 128, & S_{mean}^{C} > 128 \end{cases}$$

here λ behaves as a controlling parameter, its a positive number and manipulates the shifting range and P^{c} describes the probability of a pixel values which are equal to or less than 40. It handles the over-correction rather effectively. The final step is to compute $S_{CR} = min(max(S_{CR}, 0), 255)$, and it is used to steer clear of exceeding the range of a pixel. It is noteworthy that all operations are component-wise.

FIGURE 3.3: Original Image, output of gray-world, proposed by Fu [7]

3.2 Processing the Illuminance Component

3.2.1 Global Underwater Background Light Estimation

Underwater images can be distorted by the dust particles present under the water with different sort of wavelengths. The light absorption causes the color fading, due to which the objects become bluish and greenish as these colors have shorter wavelengths than the red light. As dust particles cause fading color of the captured objects, it also causes the contrast degradation by the scattering and reflection of the light phenomenon into arbitrary directions. Due to the absence of the black body radiation, multi scattering process shows the objects and scenes to be covered by a slight haze. It becomes the background light in the hazy environment. Usually, the background light can be estimated as the brightest pixel value or the average value of the brightest area in the images due to the presence of a large amount of haze. For underwater environments, this estimation method provides the intrinsic information of light absorption at different wavelengths.

For background light estimation, and efficient algorithm is select such pixel which lies furthest away from the image sensor since the degradation of color and contrast depends on the distance between the object and the sensor.As the distance increases between the object and sensor the amount of backscattered light also increases causing bright spots on the image. For the correct result, this scheme must not be applied as the variance and scene pixel values are lower in the presence of denser haze. To pick the correct pixel we used a search algorithm built upon the quad-tree subdivision. In this technique the input image is decomposed into four equal rectangular parts and for each part, the mean value is calculated and afterwards the standard deviation is subtracted from the mean value.

$$Score_l = \frac{1}{3N} \sum_{\lambda \in r,g,b} \sum_{x=1}^{N} I_l^{\lambda}(x) - \frac{1}{3} \sum_{\lambda \in r,g,b} \sqrt{\frac{\sum_{x=1}^{N} \left(I_l^{\lambda}(x) - \bar{I}_l^{\lambda}\right)^2}{N}} \tag{3.5}$$

In this equation, $l = 1, 2, 3$ shows the four areas of the input images and N is the pixel number. The selection of region is based on the region with smallest variance. These steps are repeated till the pre-determined threshold value is achieved. Within this lowest variance region, we calculate the mean vector. The mean vector and obtained light are considered the brightest value with the image.

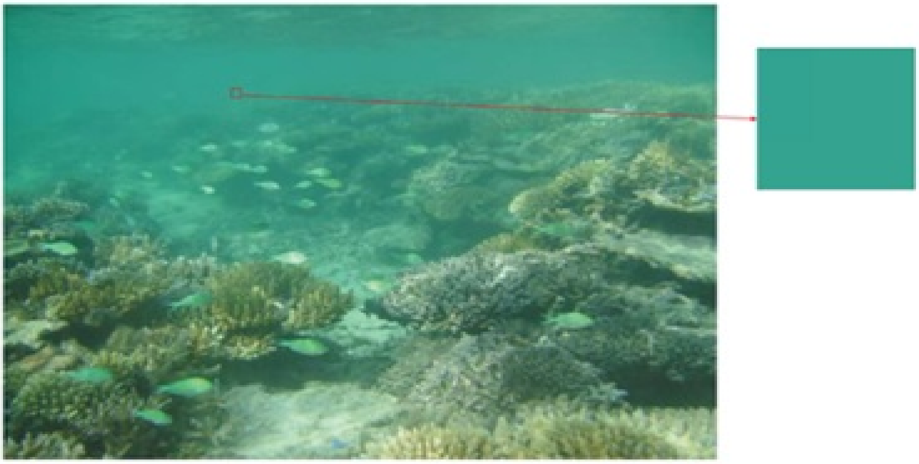


FIGURE 3.4: Background light from the correct image block

3.2.2 Transmission Map Estimation

The scene depths within the local area varies spatially. Further, the block-based local constant transmission value yields artifact and weakens the contrast. To address this issue many solution have been devised. for example Wong used the Gaussian-Markov random field model [30] which was guided by the original image for the refining of the transmission map. He also used the soft matting method. By solving the sparse linear system,

$$(L + \lambda \cdot U) \cdot t = \lambda \cdot \tilde{t} \tag{3.6}$$

L Is the matting Laplacian matrix and U is the identity matrix. Both matrices have the similar dimensions and λ is the regularization parameter. Its elements can be represented as,

$$L_{ij} = |\omega|\big(\delta_{ij} - W_{ij}\big) \tag{3.7}$$

W is the number of pixels and δ is the kronecker delta used in the neural network which needs to be updated after each iteration. In order words W is the weight update equation of the neural network. Thus the transmission map is refined due to the weight update equation of the neural network.

3.2.3 Enhancement of Transmission Map

A rather smooth transmission map is obtained after the refinement in which weights are updated regularly on each iteration. Textures and details are improved by the use of the refined transmission map. Hence the image can be separated into the smoothing component and the detailed component. The smoothing component can be obtained by applying the blurring filtering and a low-pass Gaussian filter is also an effective smooth filtering technique.

$$t = t_{smooth} + t_{detail} \tag{3.8}$$

Mathematically the Gaussian convolution kernel is described as,

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{\frac{(x^2+y^2)}{2\sigma^2}} \tag{3.9}$$

The smoothing part can be attained by smoothing of the refined transmission map with the help of Gaussian filter

$$t_{smooth} = t * G \tag{3.10}$$

And the detailed component can be obtained as

$$t_{detail} = t - t_{smooth} \tag{3.11}$$

Hence the enhanced transmission map can be calculated as

$$t_{enhanced} = t_{smooth} + \alpha \cdot t_{detail} \tag{3.12}$$

3.2.4 Enhanced Transmission Map for each Color Component

Since each wavelength of light attenuated at different rates in water each constituent wavelengths have their own transmission map. The image is usually dominated by one color only in most cases as the other colors are absorbed by the water. Even though there is one dominate color but still average of pixel value can be used to calculate the attenuation coefficient. So the average pixel value is used to calculate the enhanced transmission map for all color channels.

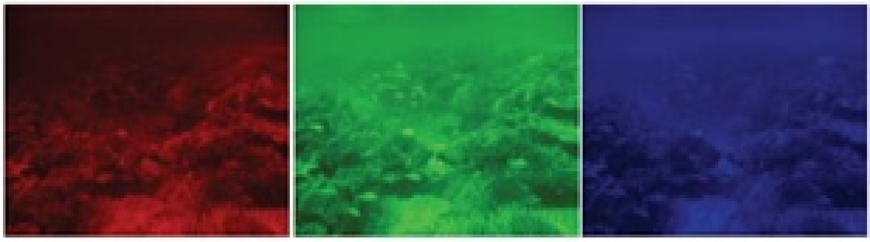


FIGURE 3.5: Enhanced transmission map for each channel

3.2.5 Dehazing

By using the simplified physical formation model, we can restore underwater image.

$$J_\lambda^T = \frac{1}{t_\lambda(x)} \cdot (I_\lambda(x) - A(\lambda)) + A(\lambda) \tag{3.13}$$

The aforementioned mathematical model cannot bring back the actual colors of the underwater image. A normalization of all intensity values is used for the mapping it to the display interval. For the color correction the piece-wise linear function is used which is described in the previous section.

3.3 Processing the Reflectance Part

Since reflective element of the decomposed image only contains the texture information of the underwater image thus, it is considered to be free of backscattered light. Hence the reflective component has only the distorted colors and which as previously stated is primarily caused by the absorption of light energy by the water molecules. To deal with the color distortion aspect of the problem an efficient color correction algorithm is needed which as explained in previous section. This linear piece-wise function is shown to produce good results.The simplest color balance (SCB) is a also employed to further improve the quality of the color distortion. This is efficient way of dealing with the color degradation.

3.4 Fusion Process

3.4.1 Weights Calculation for the Fusion

Once the improved reflectance and illuminance parts have been derived, the focus of attention is on the fusion process which combines these two images and provides a final result which has reduced backscattering and improved color range. The main step of the fusion process is that the weights need to be updated at each iteration. We studied comprehensive methods on weight update of the network so we divided the weights into three categories.

1. Luminance weight map.
2. Saliency weight map.
3. Exposedness weight map.

Luminance map describes the luminance parameter W_L. The standard deviation amongs the color channels $red, green, blue$ is computed for generating the weight maps. The luminance value is defined by l and is derived by the following equation:

$$l = \alpha \cdot r + \beta \cdot g + \gamma \cdot b \tag{3.14}$$

Where $\alpha + \beta + \gamma = 1$. α, β, γ are weighted parameter of each color channel. We set the values of $\alpha = 0.299, \beta = 0.587, \gamma = 0.114$. This weight map provides the corrected degree of reflected and luminance component of the image and shows improved enhancement of the image that was degraded previously. This weight map also ensures that the effects on contrast and colorfulness are negative. The following two weight maps were introduced in the algorithm.

Saliency weight map; A lot of research has been done by researchers for the generation of saliency map. Achanta et al.[3] which was built around biological concept of centered surround contrast and is applied in our case because it is simple computationally, efficient and time saving algorithm. A major discrepancy of this algorithm is that it overestimates the highlighted areas. So an exposedness weight map is used to ensure the accuracy and precision of the image.

Exposedness weight map, W_E evaluate the details of the exposed pixels. It keeps

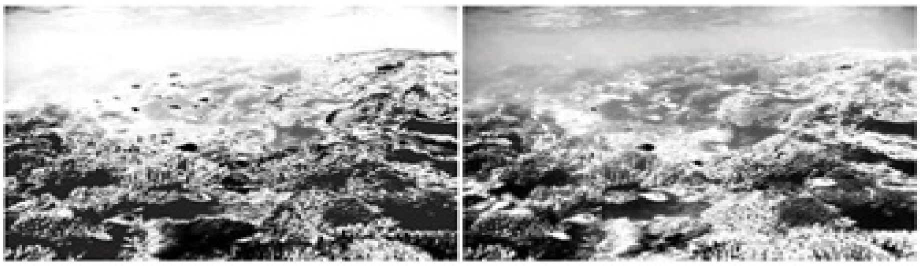


FIGURE 3.6: Exposedness weight map of reflectance and illuminace components

the local contrast closer to its original value and not be understated in the final output image. Mathematically, the map can be written as:

$$W_E = e^{-\frac{(I(x)-\bar{I})^2}{2\sigma^2}} \tag{3.15}$$

In the above equation, σ^2 is the standard deviation, $I(x)$ indicated the pixel value and $\bar{I}$ represents the mean value. Those pixels that lay nearer to the mean for the channel have higher weight as compare to the pixel lie closer to periphery. Also pixel further away from the mean means underexposed regions. So these three maps provide the well appearance and better accuracy in the fused images.

3.5 Multi-scale Fusion Process

For more than a single scale fusion process, these three weight maps can express the features of input images in a better manner. For the normalization of weight map from these three featured weight maps, we can calculate the normalized weight map as,

$$W_{norm} = \sum_{k=1}^{K} \cdot W_k \tag{3.16}$$

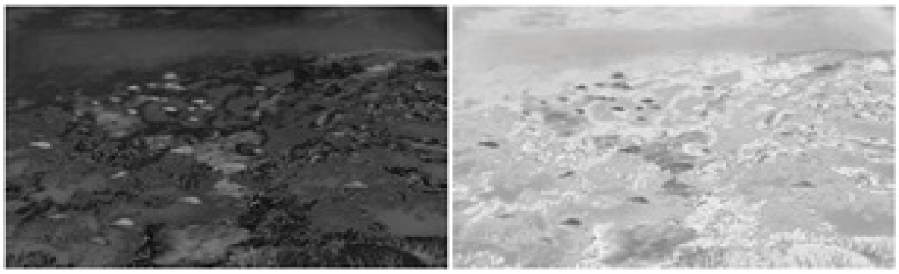


FIGURE 3.7: Normalized weight map of reflectance and illuminace components

The final back-scatter free and color enhanced can be derived as

$$R_\lambda = \sum_{n=1}^{N} W_{norm}^n \cdot J_\lambda^n \tag{3.17}$$

Where R is the color component of the final output image.

FIGURE 3.8: Final result via naive fusion

In order to further improve the performance, the Gaussian and Laplacian pyramid scheme is employed for the execution of the fusion process. Using the Laplacian operator and Gaussian kernel, every input is decomposed into several layers with different scales. Differentiation of original image yields in higher layers' generation and Gaussian pyramid provides us the filtered image. The Laplacian pyramid that we use is a quasi-bandpass version. The entire Fusion process can be written in terms of Gaussian and Laplacian pyramids as:

$$R_{\lambda}^{l} = \sum_{n=1}^{N} G^{l}\left\{W_{norm}^{n}\right\} \cdot L^{l}\left\{J_{\lambda}^{n}\right\} \tag{3.18}$$

By using the above Gaussian and Laplacian pyramid techniques, the final output image results as:

FIGURE 3.9: Final result via multi stage fusion

Chapter 4

Results and Analysis

4.1 Evaluation Underwater Image Quality

In most of the case the end users of images and other multimedia are humans. Hence, among all the image quality and accuracy assessments the subjective analysis by a human is often more accurate. Althought the subject analysis of images does produce much higher quality results but it often lacks speed, its expensive and may produce different results on different times dependent upon the mood of the user. Thus for many real world applications the subjective analysis is not the way to proceed. Since human observers are the ultimate users in most of the multimedia applications, the most accurate and also reliable way of assessing the quality of images is through subjective evaluation. However, subjective evaluations are expensive and time consuming, which makes them impractical in real-world applications. Besides the subjective technique may also be affected by the equipment, lighting distance from the observer and off course the observers eye sight. Nonetheless for the sack of completeness some of the more popular categories of subjective analysis techniques are described below.

4.1.1 Single Stimulus Categorical Rating

For this method the observer is shown the test images randomly on a screen for a preset amount of time and afterwards the screen goes blank and the observer is requested to rate the images. The rating by the observers falls in five categories **excellent, good, fair, poor and bad**

4.1.2 Double Stimulus Categorical Rating

In this technique the test image and the reference image is shown to the observer simultaneously and like the previous method for a fixed amount of time. Afterwords the user is expected to rate the test image according to the rating categories mentioned above.
In addition to above described categories many international standards have also been proposed [58] which provide reliable results.

1. ITU-R BT.500-11 subject quality rating for television pictures. Its a well adaptive standard and it also details the conditions for viewing, how to assess the subjective data etc.

2. ITU-T P.910 This standard details with the subject analysis of the digital videos quality classification specifically for the bit rate under 1.5 Mb/s.

3. ITU-R BT.814-1 This deals with the contrast and brightness of display devices.

4. ITU-R BT.1129-2 A standard for the SD videos.[*Standard Definition*].

Subjective image quality assessment could be biased, and it requires costly resources and time.But more important is the fact the subjective assessment techniques cannot be automated. To overcome the slowness of the subjective image quality assessment a lot of research is being conducted to create mathematical models to make the whole process of the image quality assessment automatic and to have a reliable objective evaluation measure. One such method is the mean value and standard deviation which are the basic and widely used evaluations to measure the quality of images; are able to reflect the intensity and contrast information of images well. Generally, gray average value of an image reflects the integral intensity, and higher the gray average is, higher the intensity is, and the standard deviation computation of a image implies a higher frequency part, which is connect to the contrast level of the image. Higher values of standard deviation not only reflects high contrast but also greater color information. Meanwhile, Jobson et al. [59] indicate that an image shows good integral quality performance when its mean lays between 100 and 200 and the standard deviation is between 35 and 80 after they analyzed and statistic a large amount of images. Still for underwater image quality assessment this rather simple statistics model does not suffice. A more comprehensive method is proposed by Penata which specifically deals with the image quality assessment of underwater images.

4.2 Objective Image quality Assessment

There are two main classification of objective assessment techniques; non-reference assessment and fully-reference assessment. Fully reference assessment requires availability of *"True Color"* and *"ideal Contrast"* image set. Since the availability of ground-truth images of underwater imaging is rather challenging, hence, the non-reference techniques are rather more desirable. This task get more complicated due the fact it needs to comprehend the Human Visual System (HVS) and how it perceive the quality in it entirety. A few image assessment techniques like curve analysis, receiver operating characteristic (ROC) and the Histogram measurement have been devised to automate the task of image quality evaluation but are not specifically target the underwater images. Aforementioned techniques does not take in account the light properties in water and how it attenuate dramatically with depth.

4.2.1 Underwater Image Quality Metric

The Underwater Image Quality Metric (UIQM) [60] is different from more commonly used Human Visual System (HVS)-based measures, like peak signal-to-noise ratio (PSNR) and mean square error (MSE). MSE and PSNR usually measures the differences between the distorted images and the reference images. UIQM also differ from such measures like gradient structural similarity measure (GSSIM) and structural similarity measure (SSIM) which tries to quantify the differences visually. The UIQM does not need any reference images and it utilizes the Human Visual System which tallies well with the perceived underwater image quality. UIQM, is made up of three measurements, which are independent of each other. These three measurements are:

1. Underwater image colorfulness measure (UICM).
2. Underwater image sharpness measure (UISM).
3. Underwater image sharpness measure (UISM).
4. Underwater image contrast measure (UIConM).

4.2.1.1 Underwater Image Colorfulness Measure

Most of under-water images are degraded by color casting, and growth with depth increases, while different colors show various attenuating ratio. Generally, red component is disappeared firstly because of the shortest wavelength, while blue and green components attenuate slowly, so underwater scenes are often demonstrated to have a green or blue tint. Moreover, as previously mentioned the light attenuation severely degrades the colors of an image. So to assess the performance of color correction algorithms, the UICM evaluates the Red-Green (RG) and Yellow- Blue (YB) color components:

$$RG = R - G \tag{4.1}$$

$$YB = \frac{R+G}{2} - B \tag{4.2}$$

Considering that underwater image often suffers from heavy noise, the alpha trimmed statistical values are employed to measure underwater images colorfulness,

$$\mu_{RG} = \frac{1}{N - T_L - T_R} \sum_{x=T_L}^{N-T_R} Intensity_{RG}(x) \tag{4.3}$$

here N is the total pixels in the RG component and and all pixels of the image are sorted such that $x_1 < x_2 < ... < x_N, T_L = \alpha_L \cdot N$ and $T_R = \alpha_R \cdot N$ are the number of smallest and greatest pixel values to be truncated or discarded from the sorted sequence $x_1 < x_2 < ... < x_N$. The first-order statistic mean value μ_{RG} represents chrominance intensity, and the average value that is closer to zero in the $RG - YB$ opponent color component implies a better white balance, which means none of the colors are dominant. Further, the second-order statistic variance is defined by:

$$\sigma^2_{RG} = \frac{1}{N} \cdot \sum_{x=1}^{N} \left(Intensity_{RG}(x) - \mu_{RG}\right)^2 \tag{4.4}$$

σ^2_{RG} represents the pixel activity and a greater variance corresponds to a higher dynamic range. What's more, the first and second order statistic information σ_{YB} and σ^2_{YB} of the yellow-blue component can be computing in the similar way.
The overall colorfulness coefficient metric which is used for measuring underwater image colorfulness is able to demonstrated in

$$UICM = -0.0268 \cdot \sqrt{\mu^2_{RG} + \mu^2_{YB}} + 0.1586 \cdot \sqrt{\sigma^2_{RG} + \sigma^2_{YB}} \tag{4.5}$$

TABLE 4.1: Statistical intermediate values of UICM

	μ_{RG}	μ_{YB}	σ_{RG}	σ_{YB}	***UICM***
Left Image	-1.067	61.696	13.194	11.045	-0.575
Right Image	-1.432	5.03	21.560	25.770	4.843

The ***UICM*** value of restored image is much greater than original image.

UICM = -0.575 UICM = 4.843

FIGURE 4.3: UICM example of original image and restored image

4.2.1.2 Underwater Image Sharpness Measure

Sharpness reflects the details and edges of an image, and fine captured images are likely to show better sharpness. However, for images captured under the water, severe blurring and distortion occur due to backscatter and absorption. In order to measure the sharpness, the Sobel operator is first applied on each color component to generate the edge maps. Then the obtained edge maps are multiplied to original color component to calculate the gray-scale edge maps. By doing this more efficient, the enhancement measure estimation (EME) measure [27] is utilized to measure the sharpness

$$EME = \frac{2}{m \cdot n} \cdot \sum_{k=1}^{m} \cdot \sum_{l=1}^{n} \log\left(\frac{I_{max}, k, l}{I_{min}, k, l}\right) \tag{4.6}$$

where the image is divided into $m \cdot n$ blocks, and obtain the maximal and minimal pixel values in each block, $\frac{I_{max},k,l}{I_{min},k,l}$ indicates the relative contrast ratio within each block. Then the underwater image sharpness measure (UISM) can be written as:

$$UISM = \sum_{c=1}^{3} \lambda_c \cdot EME(grayscale\ edge_c) \tag{4.7}$$

where λ_c, is the weight coefficient of each color component, normally, $\lambda_R = 0.299, \lambda_G = 0.587, \lambda_B = 0.114$ for red, green and blue color channels.

UISM = 6.683 UISM = 7.654

FIGURE 4.6: UISM example of original image and restored image. Left: underwater backscatter image, right: descattered image by proposed method, it seems that original image suffers heavier blurring effect.

4.2.1.3 Underwater Image Contrast Measure

Contrast is the attribute related to underwater visual performance. For underwater images, contrast degradation is usually caused by backscattering. The contrast performance can be measured by the logAMEE measurement, and it is defined by:

$$\log AMEE = \frac{1}{m \cdot n} \cdot \sum_{k=1}^{m} \sum_{l=1}^{n} \frac{I_{max,k,l} - I_{min,k,l}}{I_{max,k,l} + I_{min,k,l}} \cdot \log \left(\frac{I_{max,k,l} - I_{min,k,l}}{I_{max,k,l} + I_{min,k,l}} \right) \tag{4.8}$$

and the underwater image contrast measure can be written as:

$$UIConM = \log AMEE(Intensity) \tag{4.9}$$

4.2.1.4 Underwater Image Quality Measure

It has been demonstrated that underwater images can be modeled as linear superposition of absorbed and scattered components [27]. Meanwhile, the water absorption and backscatter by dusk-like particles are able to cause color casting, sharpness attenuation and contrast degradation. Therefore, it is reasonable to use the linear model for generating the overall underwater image quality measure, thus

UIConM = 0.651 UIConM = 2.234

FIGURE 4.9: UIConM example of original image and restored image. Left: original underwater image, right: contrast enhanced image by proposed method, which shows greater improvement of contrast.

the underwater image quality measure (UIQM) is given by:

$$UIQM = \alpha \cdot UICM + \beta \cdot UISM + \gamma \cdot UIConM \tag{4.10}$$

where the colorfulness, sharpness and contrast measure are combined together through the linear function designed above, and α, β and γ are the weight coefficients to control the importance each measure and balance their values. Generally, these parameters are set to be $\alpha = 0.0282, \beta = 0.2953$ and $\gamma = 3.5753$.

4.3 Experiment Results and Evaluations

In this subsection, we use the measurements mentioned above to evaluate the performance of proposed underwater objects visibility enhancing method as well as comparison to the state-of-art methods. For this case, we compare the performance of our method to several state-of-art methods in recent years, where most of them also achieves excellent performance for the underwater imaging issue. Generally, we introduce the methods proposed by Ancuti [12], Histogram equalization (HE), zhang [57], Ancuti [3], Fu [7] and method proposed by this thesis.

Diver Image Open Scene

FIGURE 4.12: Test images for evaluation and comparison

4.3.1 Evaluation of Different Methods

Images captured in the different underwater environments often show various attenuation and degradation degree, which causes that an underwater imaging method may performs well for several certain underwater environment conditions, but weak for other conditions. We choose a set of images for evaluation.

4.3.1.1 Processing and Evaluation of Images

We first process first image shown in the Figure 4-13 by several state-of-art methods and the result is shown below. We see that on the diver image, the histogram equalization (HE) method fails to correct and enhance the visibility of degraded image, while Ancuti [3] method shows good color correction as well as proper contrast enhancement, the restored results generated by other three methods also achieve color correction and contrast enhancement, their mean value and standard deviation is greatly improved, as shown in the table 4.2. Expect the highest mean value, standard deviation and average gradient, the objective visual performance of the result obtained by our method are better than other methods. Moreover, we can see from the RGB color space of the original images and restored results that all the pixels of original image maps into the RGB color space are gathered in the left corner, where is large green value and small blue and red value, and other

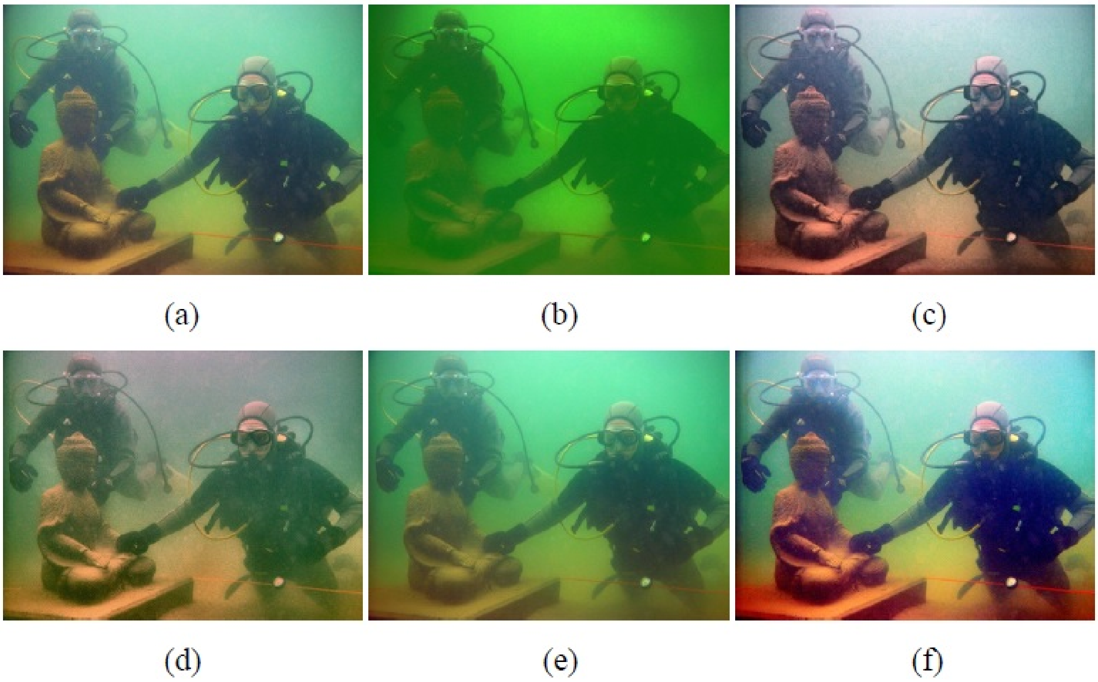


FIGURE 4.13: Results on diver image via different methods. (a) Ancuti [12] (b) Histogram equalization (HE). (c) zhang [57]. (d) Ancuti [3] (e) Fu [7] (f) Proposed method.

TABLE 4.2: Statistical Analysis

	Orig	Ancuti[12]	HE	zhang [57]	Ancuti [3]	Fu [7]	Proposed
Mean	88.7	117.3	70.3	114.0	103.9	104.2	**124.4**
σ	21.0	45.63	25.94	61.76	38.81	42.87	**80.33**

methods achieve more or less stretch of the pixel mapping results. It is obvious that the mapping result of our method is best and it maps the largest area in the RGB color space, which means it has the best integral dynamic compression range. However, many mapping pixels of our method is mapped onto the boundary, i.e., dark region (pixel value is 0) and brightest region (pixel value is 255), it indicates that there is information loss of our method. For this diver image, our method achieves the integrally best performance, although it seems that some information loss exists. Actually, our method not only achieves excellent performance for heavy color casting images, like diver image above, but also performs well on images captured in different underwater environments. For the open scene the results is presented below.

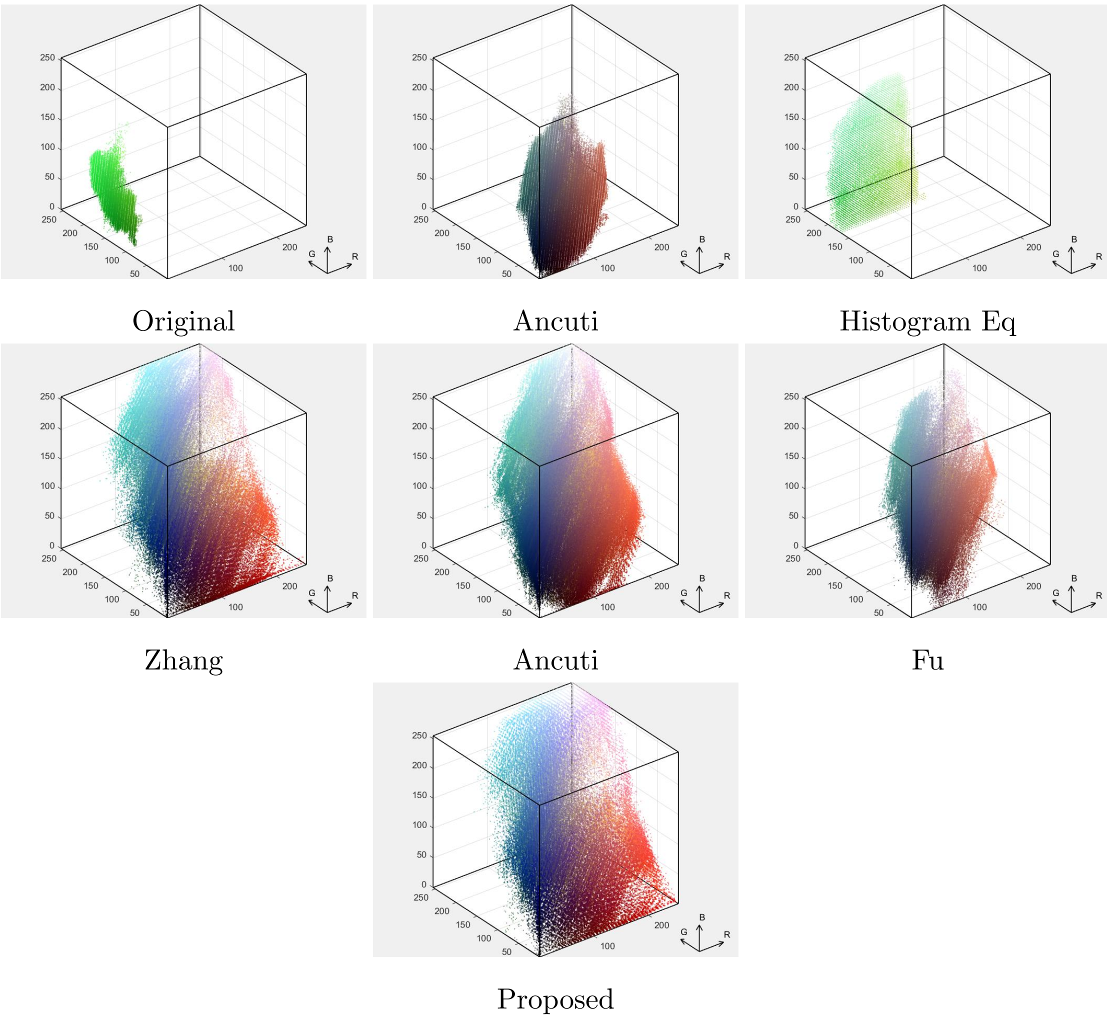


FIGURE 4.21: RGB spaces of divers image

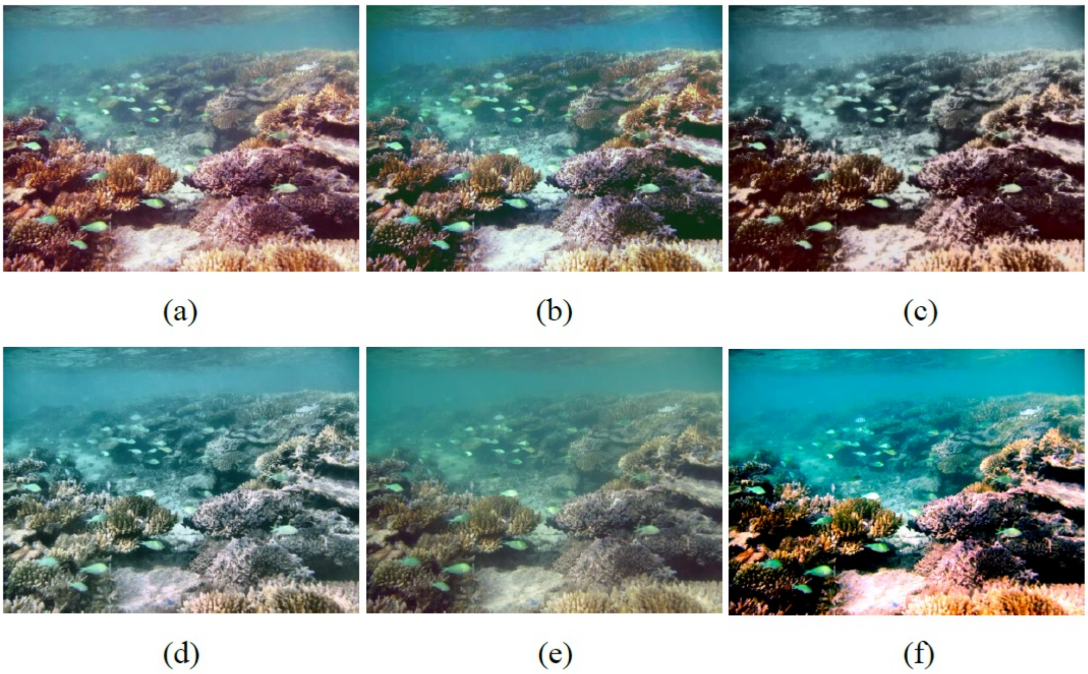
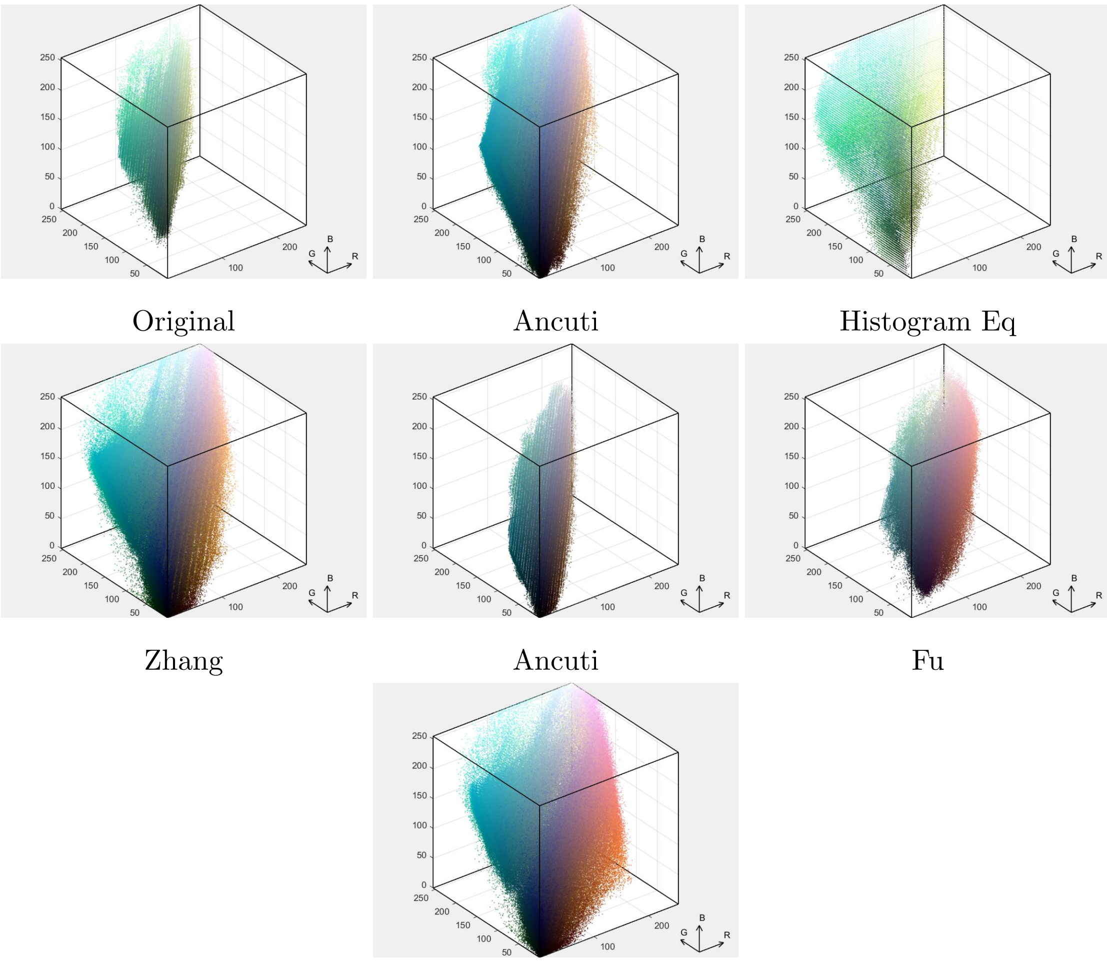


FIGURE 4.22: Results on the open scene image via different methods. (a) Ancuti [12] (b) Histogram equalization (HE). (c) zhang [57]. (d) Ancuti [3] (e) Fu [7] (f) Proposed method.

Table 4.3: Statistical Analysis

	Orig	Ancuti[12]	HE	zhang [57]	Ancuti [3]	Fu [7]	Proposed
Mean	116.1	**125.2**	111.1	101.3	121.1	110.3	108.9
σ	5.64	10.35	10.70	9.64	11.56	6.65	**12.79**



Original Ancuti Histogram Eq

Zhang Ancuti Fu

Proposed

Figure 4.30: RGB spaces of open Scene

Although the Ancuti [12] method outperforms our method in mean value, our method still performs best among all these methods, which has excellent visual performance as well as good RGB color space mapping and higher contrast degrees, but its edge information does not achieve well because of the information loss. Among these methods, we can see that Fu et al. over-enhancing the distorted image, which causes severe color distortion and noise amplification, thus the integral performance of enhanced image is worse. Galdran et al. solves the effect of color casting and contrast degradation to some degree, but it is not enough, the results still tends to be greenish, while Ancuti [3]. gets the good result, not only correcting the color casting but also enhancing the contrast. For histogram equalization method, aim to normalize and regularize the histogram of an image and achieve the contrast enhancement and color correction. By comparison, although the histogram equalization method fails to solve the issue of diver image, it achieves good result for the open scene image, which indicates that this method is not suitable for dense color casting and heavy hazy situation. But for Ancuti [12] method, it performs well in both of two test images.

4.3.2 UIQM Measurement of Tested Methods

Although the measurements used in subsection 4.2.1 gives us some performance information of restoration performance of different methods, these measurements are not used to evaluate the underwater imaging issue, so they can only give some certain parts of the comparing results. In order to overcome the drawbacks of traditional image measurements and evaluate the performance of underwater imaging well, we introduce the underwater image quality metrics (UIQM) [60] measurement to give a comprehensive evaluation. Similarly, we also choose a set of distorted underwater images to be corrected and enhanced by different methods and then utilize the underwater image quality metrics measurement to evaluate their performance. These original images are shown below.
Despite the subjective visual evaluation method, we also introduce the underwater image quality metrics to evaluate the performance of underwater imaging methods well. The underwater image quality metrics is composed by colorfulness metric, sharpness metric and contrast metric, where colorfulness metric is utilized to measure the color correction performance, sharpness plays a role of reflecting textures and edges of the restored image and contrast metric is used to measure the contrast of restored image.

Since human eyes are more sensitive to low frequency component of an image, where low frequency component reflects the intensity and color of the image, and can not show the small change of high frequency component, which reflects the details and texture information of the image. Based on this case, we find that the visual performance of images has the most influence for determining the image quality, and color casting contributes the most effects to visual performance. So we first analyze the color correction performance of these methods by underwater image colorfulness metric (UICM). As shown in the table 4-4, we compare the mean value and standard deviation of Red-Green (RG) channel and Yellow-Blue (YB) channel, followed by the UICM values. It is mentioned before that an image with good color performance when the μ_{RG} and μ_{YB} are both near to zero, while the contrast is measured by σ_{RG} and σ_{YB}, who should be as large as possible. The μ_{RG} and μ_{YB} of original image is large (-106.74, 61.70) due to the color casting caused by light absorption, where red component attenuates fastest and blue component is absorbed and backscattered by dusk-like particles in this specific water environment, while σ_{RG} and σ_{YB} stay at a low level because of haze.

Table 4.4: UICM Analysis scores of diver image

	Orig	Ancuti[12]	HE	zhang [57]	Ancuti [3]	Fu [7]	Proposed
μ_{RG}	-106.74	-20.89	-66.48	**-0.99**	-21.90	-45.83	-20.35
μ_{YB}	61.70	5.59	44.56	**-1.09**	15.21	25.78	9.14
σ_{RG}	13.19	15.38	14.09	10.05	13.62	**20.98**	18.76
σ_{YB}	11.04	17.48	11.27	12.23	12.52	15.29	**24.73**
UICM	-0.58	3.11	0.72	2.47	2.22	2.71	**4.33**

Then the underwater image sharpness metric (UISM) and underwater image contrast metric (UIConM) are calculated. The Ancuti [12] method highest score, 7.06, in UISM measurement, which is closely followed by our method, 7.05. For UIConM measurement, the result of our method stands at the highest level, 0.58, and finally for the comprehensive measurement, underwater image quality metric

(UIQM), our method also achieves the highest score. Thus, under this water environment, our method performs best and successfully realizes the color correction and contrast enhancement.

TABLE 4.5: UIQM Analysis scores of diver image

	Orig	Ancuti[12]	HE	zhang [57]	Ancuti [3]	Li [2]	Proposed
UICM	-0.58	3.11	0.72	2.47	2.22	2.71	**4.33**
UISM	6.96	**7.06**	6.94	6.84	7.04	6.97	7.05
UIConM	0.34	0.56	0.47	0.51	0.50	0.45	**0.58**
UIQM	0.26	3.52	1.86	3.91	3.24	2.38	**4.28**

Chapter 5

Conclusion and Future Work

An improved fusion based technique for enhancement of underwater images is presented here. Comparing to other fusion based techniques it is revealed that our techniques improved the underwater images. The technique is simple yet robust. It requires only one image as input. Also the mathematical model of underwater image formation is studied in great detail and it is revealed that the light attenuation in water and veiling light has the most effect on the under water image quality. For fusion inputs the original image is decomposed into two constituent components namely the reflective and illuminance parts. Both components are processed independently and after calculating the weights for the fusion, both images are merged back into a single image as output.

5.1 Limitations

On the contrary there is one major restriction has also been observed by using fusion based technique is of noise or disturbance level has also amplified significantly while using the depth glorious unseen show for distant areas. This will be the work for future to find optimal techniques and filters to reduce the ambiguity of noise or disturbance in the underwater images, while enhancement and restoration has been of better viewing than the techniques being used previously.

Emergent upside down image resolution approaches along with system pass are important to adapt and expand in the techniques devised, which can greatly help in evaluating the data coming from 3 dimensional landscape data. However a close look assures to have more information revealing underwater images as compared to the techniques previously used involving sub-aqueous life.

5.2 Future work

One aspect, if worked on can dramatically improve the execution time is the parallelization of the whole process. After decomposition of the input image into two parts. Both parts have zero dependency on each other. Hence, a good opportunity to make the algorithm massively parallel.
In Conclusion the techniques proposed in the thesis improves upon the existing fusion based under water image enhancement solutions. It performs quite good on many different images taken under different lighting conditions. Although it did show a propensity of over correcting some of the test images, which results in information lost and over exposed patches which are in some cases quite visible to naked eye. Nonetheless in comparison to other fusion techniques our method stood well and in future a more improved version will be presented.

References

[1] Yujie Li, Huimin Lu, Jianru Li, Xin Li, Yun Li, and Seiichi Serikawa. Underwater image de-scattering and classification by deep neural network. *Computers & Electrical Engineering*, 54:68–77, 2016.

[2] Jingjing Li and Hanyu Li. A fusion adversarial network for underwater image enhancement. *arXiv preprint arXiv:1906.06819*, 2019.

[3] Cosmin Ancuti, Codruta Orniana Ancuti, Tom Haber, and Philippe Bekaert. Enhancing underwater images and videos by fusion. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 81–88. IEEE, 2012.

[4] Robert P Bukata, John H Jerome, Alexander S Kondratyev, and Dimitry V Pozdnyakov. *Optical properties and remote sensing of inland and coastal waters*. CRC press, 2018.

[5] Daily Daleno de O Rodrigues, Wagner F de Barros, José P de Queiroz-Neto, Anderson G Fontoura, and José Reginaldo H Carvalho. Enhancement of underwater images in low-to-high turbidity rivers. In *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 233–240. IEEE, 2016.

[6] Sofia Ahlberg Pilfold. *Managing knowledge for through life capability*. PhD thesis, © Sofia Ahlberg Pilfold, 2016.

[7] Xueyang Fu, Zhiwen Fan, Mei Ling, Yue Huang, and Xinghao Ding. Two-step approach for single underwater image enhancement. In *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 789–794. IEEE, 2017.

[8] Jin-Hwan Kim, Won-Dong Jang, Jae-Young Sim, and Chang-Su Kim. Optimized contrast enhancement for real-time image and video dehazing. *Journal of Visual Communication and Image Representation*, 24(3):410–425, 2013.

[9] Chongyi Li, Jichang Guo, and Chunle Guo. Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal processing letters*, 25(3):323–327, 2018.

[10] Paulo LJ Drews, Erickson R Nascimento, Silvia SC Botelho, and Mario Fernando Montenegro Campos. Underwater depth estimation and image restoration based on single images. *IEEE computer graphics and applications*, 36(2):24–35, 2016.

[11] Wenhao Zhang, Ge Li, and Zhenqiang Ying. Underwater image enhancement by the combination of dehazing and color correction. In *Pacific Rim Conference on Multimedia*, pages 145–155. Springer, 2018.

[12] Codruta O Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, Laszlo Neumann, and Rafael Garcia. Color transfer for underwater dehazing and depth estimation. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 695–699. IEEE, 2017.

[13] Huimin Lu, Yujie Li, Xing Xu, Li He, Yun Li, Donald Dansereau, and Seiichi Serikawa. Underwater image descattering and quality assessment. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 1998–2002. IEEE, 2016.

[14] Huimin Lu, Yujie Li, and Seiichi Serikawa. Single underwater image descattering and color correction. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1623–1627. IEEE, 2015.

[15] Khadidja Ould Amer, Marwa Elbouz, Ayman Alfalou, Christian Brosseau, and Jaouad Hajjami. Enhancing underwater optical imaging by using a low-pass polarization filter. *Optics express*, 27(2):621–643, 2019.

[16] Fei Liu, Yi Wei, Pingli Han, Kui Yang, Lu Bai, and Xiaoepeng Shao. Polarization-based exploration for clear underwater vision in natural illumination. *Optics express*, 27(3):3629–3641, 2019.

[17] Bingjing Huang, Tiegen Liu, Haofeng Hu, Jiahui Han, and Mingxuan Yu. Underwater image recovery considering polarization effects of objects. *Optics express*, 24(9):9826–9838, 2016.

[18] Chongyi Li and Jichang Guo. Underwater image enhancement by dehazing and color correction. *Journal of Electronic Imaging*, 24(3):033023, 2015.

[19] Dana Berman, Tali Treibitz, and Shai Avidan. Diving into haze-lines: Color restoration of underwater images. In *Proc. British Machine Vision Conference (BMVC)*, volume 1, 2017.

[20] Miao Yang, Arcot Sowmya, ZhiQiang Wei, and Bing Zheng. Offshore underwater image restoration using reflection-decomposition-based transmission map estimation. *IEEE Journal of Oceanic Engineering*, 2019.

[21] Laurent Guillon, Stan E Dosso, N Ross Chapman, and Achraf Drira. Bayesian geoacoustic inversion with the image source method. *IEEE Journal of Oceanic Engineering*, 41(4):1035–1044, 2016.

[22] Tengyue Li, Bo He, Shizhe Tan, Chen Feng, Shuai Guo, Hanmin Liu, and Tianhong Yan. Optical sources optimization for 3d reconstruction based on underwater vision system. In *2019 IEEE Underwater Technology (UT)*, pages 1–4. IEEE, 2019.

[23] Adarsh Jamadandi and Uma Mudenagudi. Exemplar-based underwater image enhancement augmented by wavelet corrected transforms. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 11–17, 2019.

[24] Xi Qiao, Jianhua Bao, Hang Zhang, Lihua Zeng, and Daoliang Li. Underwater image quality enhancement of sea cucumbers based on improved histogram

equalization and wavelet transform. *Information processing in agriculture*, 4(3):206–213, 2017.

[25] WNJHW Yussof, Muhammad Suzuri Hitam, Ezmahamrul Afreen Awalludin, and Zainuddin Bachok. Performing contrast limited adaptive histogram equalization technique on combined color models for underwater image enhancement. *International Journal of Interactive Digital Media*, 1(1):1–6, 2013.

[26] Haocheng Wen, Yonghong Tian, Tiejun Huang, and Wen Gao. Single underwater image enhancement with a new optical model. In *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, pages 753–756. IEEE, 2013.

[27] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010.

[28] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018.

[29] Xueyang Fu, Peixian Zhuang, Yue Huang, Yinghao Liao, Xiao-Ping Zhang, and Xinghao Ding. A retinex-based enhancing approach for single underwater image. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 4572–4576. IEEE, 2014.

[30] Siaw-Lang Wong, Raveendran Paramesran, and Akira Taguchi. Underwater image enhancement by adaptive gray world and differential gray-levels histogram equalization. *Advances in Electrical and Computer Engineering*, 18(2):109–117, 2018.

[31] Rafał Protasiuk, Adel Bibi, and Bernard Ghanem. Local color mapping combined with color transfer for underwater image enhancement. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1433–1439. IEEE, 2019.

[32] Miquel Massot-Campos and Gabriel Oliver-Codina. Optical sensors and methods for underwater 3d reconstruction. *Sensors*, 15(12):31525–31557, 2015.

[33] Adrian Driewer, Igor Abrosimov, Jonathan Alexander, Marc Benger, Marion O'Farrell, Karl Henrik Haugholt, Chris Softley, Jens T Thielemann, Jostein Thorstensen, and Chris Yates. Utofia: an underwater time-of-flight image acquisition system. In *Electro-Optical Remote Sensing XI*, volume 10434, page 1043404. International Society for Optics and Photonics, 2017.

[34] Qiao Xi, Thomas Rauschenbach, and Li Daoliang. Review of underwater machine vision technology and its applications. *Marine Technology Society Journal*, 51(1):75–97, 2017.

[35] Saeed Anwar, Chongyi Li, and Fatih Porikli. Deep underwater image enhancement. *arXiv preprint arXiv:1807.03528*, 2018.

[36] Amin Sarafraz and Brian K Haus. A structured light method for underwater surface reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114:40–52, 2016.

[37] CJ Prabhakar and PU Kumar. 3d surface reconstruction of underwater objects. *arXiv preprint arXiv:1211.2082*, 2012.

[38] Zak Murez, Tali Treibitz, Ravi Ramamoorthi, and David Kriegman. Photometric stereo in a scattering medium. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3415–3423, 2015.

[39] JM Nichols, JR Waterman, Raghu Menon, and John Devitt. Performance characteristics of a submarine panoramic infrared imaging sensor. In *Infrared Technology and Applications XXXVI*, volume 7660, page 766005. International Society for Optics and Photonics, 2010.

[40] Huimin Lu, Yujie Li, Hyoungseop Kim, and Seiichi Serikawa. Underwater light field depth map restoration using deep convolutional neural fields. In *Artificial Intelligence and Robotics*, pages 305–312. Springer, 2018.

[41] Sudharshan Suresh, Eric Westman, and Michael Kaess. Through-water stereo slam with refraction correction for auv localization. *IEEE Robotics and Automation Letters*, 4(2):692–699, 2019.

[42] Oleari Fabio, Kallasi Fabjan, Lodi Rizzini Dario, Aleotti Jacopo, and Caselli Stefano. Performance evaluation of a low-cost stereo vision system for underwater object detection. *IFAC Proceedings Volumes*, 47(3):3388–3394, 2014.

[43] Pedro DV Buschinelli, Gabriel Matos, Tiago Pinto, and Armando Albertazzi. Underwater 3d shape measurement using inverse triangulation through two flat refractive surfaces. In *OCEANS 2016 MTS/IEEE Monterey*, pages 1–7. IEEE, 2016.

[44] Gianfranco Bianco, M Muzzupappa, F Bruno, R Garcia, and L Neumann. A new color correction method for underwater imaging. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(5):25, 2015.

[45] Yadpiroon Onmek, Jean Triboulet, Sébastien Druon, Arnaud Meline, Silvain Louis, and Bruno Jouvencel. 3d underwater reconstruction of archeology object with a mono camera camcorder: The case study at gourneyras lake. In *OCEANS 2017-Aberdeen*, pages 1–6. IEEE, 2017.

[46] Kazuto Ichimaru and Hiroshi Kawasaki. Underwater stereo using refraction-free image synthesized from light field camera. *arXiv preprint arXiv:1905.09588*, 2019.

[47] HG Maas. A modular geometric model for underwater photogrammetry. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(5):139, 2015.

[48] Tomasz Łuczyński, Max Pfingsthorn, and Andreas Birk. The pinax-model for accurate and efficient refraction correction of underwater cameras in flat-pane housings. *Ocean Engineering*, 133:9–22, 2017.

[49] Yujie Li, Yin Zhang, Xing Xu, Li He, Seiichi Serikawa, and Hyoungseop Kim. Dust removal from high turbid underwater images using convolutional neural networks. *Optics & Laser Technology*, 110:2–6, 2019.

[50] Mark Shortis. Calibration techniques for accurate measurements by underwater camera systems. *Sensors*, 15(12):30810–30826, 2015.

[51] Luz A Torres-Méndez and Gregory Dudek. Color correction of underwater images for aquatic robot inspection. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 60–73. Springer, 2005.

[52] Liu Chao and Meng Wang. Removal of water scattering. In *2010 2nd International Conference on Computer Engineering and Technology*, volume 2, pages V2–35. IEEE, 2010.

[53] Marco Block, Benjamin Gehmlich, and Damian Hettmanczyk. Automatic underwater image enhancement using improved dark channel prior. *Studies in Digital Heritage*, 1(2):566–589, 2017.

[54] Adrian Galdran, David Pardo, Artzai Picón, and Aitor Alvarez-Gila. Automatic red-channel underwater image restoration. *Journal of Visual Communication and Image Representation*, 26:132–145, 2015.

[55] Raimondo Schettini and Silvia Corchs. Underwater image processing: state of the art of restoration and image enhancement methods. *EURASIP Journal on Advances in Signal Processing*, 2010(1):746052, 2010.

[56] Tomasz Łuczyński and Andreas Birk. Underwater image haze removal and color correction with an underwater-ready dark channel prior. *arXiv preprint arXiv:1807.04169*, 2018.

[57] Hao Zhang. Removing backscatter to enhance the visibility of underwater object. *M.Sc. Thesis of NTU*, 2018.

[58] Pedram Mohammadi, Abbas Ebrahimi-Moghadam, and Shahram Shirani. Subjective and objective quality assessment of image: A survey. *arXiv preprint arXiv:1406.7799*, 2014.

[59] Bo Jiang, Glenn A. Woodell, and Daniel J. Jobson. Novel multi-scale retinex with color restoration on graphics processing unit. *Journal of Real-Time Image Processing*, 10, 06 2014.

[60] Karen Panetta, Chen Gao, and Sos Agaian. Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 41(3):541–551, 2015.