DESIGN OF FUSION-BASED UNDERWATER IMAGE ENHANCEMENT METHOD IN VIDEO STREAM

Miss Pann Mya Hmue

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Electrical Engineering
Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2018
Copyright of Chulalongkorn University

การออกแบบวิธีการปรับปรุงภาพใต้น้ำด้วยกระบวนการรวมภาพในกระแสวีดีทัศน์

น.ส.ปั้น เมีย มิว

Thesis Title	DESIGN OF FUSION-BASED UNDERWATER IMAGE ENHANCEMENT METHOD IN VIDEO STREAM
By	Miss Pann Mya Hmue
Field of Study	Electrical Engineering
Thesis Advisor	Assistant Professor SUREE PUMRIN, Ph.D.

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Engineering

.... Dean of the Faculty of Engineering
(Associate Professor SUPOT TEACHAVORASINSKUN)

THESIS COMMITTEE

.... Chairman
(Assistant Professor WANCHALERM PORA, Ph.D.)

.... Thesis Advisor
(Assistant Professor SUREE PUMRIN, Ph.D.)

.... External Examiner
(Associate Professor Sanya Mitaim, Ph.D.)

ปืน เมีย มิว : การออกแบบวิธีการปรับปรุงภาพใต้น้ำด้วยกระบวนการรวมภาพในกระแสวีดิทัศน์. (DESIGN OF FUSION-BASED UNDERWATER IMAGE ENHANCEMENT METHOD IN VIDEO STREAM) อ.ที่ปรึกษาหลัก : สุรีย์ พุ่มรินทร์

วิทยานิพนธ์ฉบับนี้เป็นงานวิจัยที่ออกแบบวิธีการปรับปรุงภาพใต้น้ำด้วยกระบวนการรวมภาพสำหรับภาพถ่ายและภาพกระแสวีดิทัศน์ ปัญหาหลักของการถ่ายภาพใต้น้ำคือการผิดเพี้ยนของสีภาพ ความสว่างที่ไม่เพียงพอ และ การปนเปื้อนของอนุภาคในน้ำ ซึ่งวัตถุประสงค์หลักของวิธีการปรับปรุงภาพใต้น้ำนี้จะเป็นการได้มาซึ่งภาพผลลัพธ์ที่เหมาะกับการมองเห็นของมนุษย์ ในงานวิจัยนี้มุ่งเน้นไปที่การประมวลผลภาพหรือภาพกระแสวีดิทัศน์ใต้น้ำจากการการคำนวณการประมวลผลที่ ไม่ซับซ้อน เนื่องจากการสร้างภาพหรือภาพกระแสวีดิทัศน์ผลลัพธ์ได้มาจากการเก็บกู้ภาพบางส่วนจากภาพดั่งเดิมทั้งหมด

ระบบส่วนแรกใช้ชุดส่งสัญญาณและตัวรับสัญญาณ 5.8 GHz สำหรับการส่งข้อมูลจากกล้อง ซึ่งตัวกล้องจะต้องอยู่ใต้น้ำและภาพถ่ายจะถูกส่งไปแสดงผลที่คอมพิวเตอร์ผ่านตัวรับ-ส่งสัญญาณ เมื่อมีสัญญาณตอบรับจากคอมพิวเตอร์แล้ว ภาพผลลัพธ์ที่ได้ตอนนี้จะเป็นภาพที่สลัวไม่ชัดเจน จากนั้นระบบนำภาพไปทำการปรับความสมดุลสีขาว (white-balanced) เพื่อที่จะกำจัดสีที่ผิดเพี้ยนออกจากภาพ จากนั้นระบบจะทำการเพิ่มความคมชัด (contrast-enhanced) เพื่อเป็นการกู้คืนรายละเอียดของภาพ เสร็จแล้วระบบจะทำการชดเชยค่าสีที่ไม่สมดุลย์จากภาพที่ได้จากสถานที่ ที่มีความสว่างไม่เพียงพอ ผลลัพธ์จากกระบวนการนี้จะนำไปประมวลค่าน้ำหนักอยู่สามค่า ได้แก่ น้ำหนักความชัดส่วนกลาง (global contrast weight), น้ำหนักความชัดเฉพาะที่ (local contrast weight) และ น้ำหนักที่โดดเด่น (saliency weight) ค่าน้ำหนักเหล่านี้จะเป็นตัวกำหนดคุณภาพของแต่ละจุดภาพ (pixel) และนำไปสู่ภาพผลลัพธ์สุดท้าย การรวมภาพหลายขนาดนั้นจะถูกทำเป็นชั้นต่อชั้นกับทั้งสองพีระมิด นั่นคือ พีระมิดแบบลาพลาเซียน (Laplacian pyramid) และ พีระมิดแบบเกาส์เซียน (Gaussian pyramid) โดย พีระมิดแบบลาพลาเซียน เป็นส่วนรับเข้า (input) ของความสมดุลสีขาว และ การปรับปรุงความคมชัด (contrast-enhanced) ของรูปภาพ ส่วน พีระมิดแบบเกาส์เซียน เป็นส่วนรับเข้า ของการทำให้เป็นมาตรฐาน (normalization) ของค่าน้ำหนักสามค่าที่ได้กล่าวมาข้างต้น ในส่วนสุดท้ายของกระบวนการนี้จะทำให้ได้ภาพที่ได้มีสีที่ชัดเจนขึ้นและ มีสัญญาณรบกวนของภาพที่ลดลง

ในงานวิจัยนี้ได้สร้างระบบทดสอบการถ่ายภาพใต้น้ำเพื่อทำการปรับปรุงภาพด้วยกระบวนการรวมภาพ ซึ่งการทดสอบประสบความสำเร็จกับภาพใต้น้ำจากสถานที่ ที่มีสัญญาณรบกวน ระบบที่พัฒนาขึ้นเป็นระบบการเฝ้าสังเกตใต้น้ำที่ราคาไม่แพงที่สามารถนำไปใช้กับภาพที่มีความละเอียดของภาพที่หลากหลายและมีแนวโน้มของการการปนเปื้อนจากสิ่งแปลกปน งานวิจัยนี้จะสามารถนำไปประยุกต์ใช้กับระบบการเฝ้าสังเกตใต้น้ำ หรือ สถาปัตยกรรมการทำฟาร์มปลาอัจฉริยะในอนาคตได้

สาขาวิชา	วิศวกรรมไฟฟ้า	ลายมือชื่อนิสิต ....
ปีการศึกษา	2561	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

Pann Mya Hmue : DESIGN OF FUSION-BASED UNDERWATER IMAGE ENHANCEMENT METHOD IN VIDEO STREAM. Advisor: Asst. Prof. SUREE PUMRIN, Ph.D.

This thesis has designed a fusion-based enhancement method for underwater images and video streams which are mainly distorted by color changes, poor illumination, and water particles. The main objective of image fusion is to achieve a comprehensive result which is appropriate for the human visual system. In this thesis, our goal is to use a straight forward and not computationally intensive reconstruction. As a result, the underwater images or video streams are partially restored from the original ones.

Firstly, the 5.8 GHz transmitter and receiver set are used to carry the input data from the camera which is placed inside the underwater scene to the computer for the enhancement stage. The corresponding data which is received by the computer then becomes the input or original hazy image. The input is white-balanced to remove the distorted color and contrast-enhanced to recover the fine details. Color compensation is carried out to remove misbalanced color in a poorly illuminated underwater scene. The image resulted from these stages are then judged by three weights: global contrast weight, local contrast weight, and saliency weight. The purpose of these weights is to decide the quality of each pixel to let them contribute to the final output. The multiscale image fusion is carried out level by level of two pyramids: Laplacian pyramid where the input is the white balance and contrast-enhanced image, and Gaussian pyramid where the input is the normalization of three weights. Finally, the fused output results with more appealing color and less noise.

The thesis has constructed a testbed to implement the proposed fusion-based underwater image enhancement method and successfully processed the noisy underwater scene. The research operates as a mid-range underwater monitoring system with a multi-resolution fashion which is prone to artifacts. This thesis will serve as a basis for future underwater monitoring or smart fish farming architectures.

Field of Study: Electrical Engineering   Student's Signature ....
Academic Year: 2018   Advisor's Signature ....

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Page

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivations and Significance of the Research Problems

The underwater imaging systems are drawing attention during the past few decades as more and more applications such as ocean exploration, underwater vehicles and intelligent fish farming are becoming popular. Underwater observation systems are in need as ecosystem indicators to provide management of sustainable sources in the ocean. Usually, it is quite challenging to collect reliable data of underwater scenes because of their poor visibility and the complications in deploying profitable sensors in a marine environment. The lack of mid-range underwater monitoring system which could provide a desirable result with a reasonable price makes the exploration and monitoring parts difficult.

One significant problem is that images or videos resulting from these systems are mostly polluted by insufficient lighting, water medium, floating particles, and many other conditions. To recover color-balanced and noise-free images, several approaches have been introduced ranging from image restoration (physical model-based) to image enhancement (non-physical model-based). Due to the complexity of physical model-based methods, which need prior knowledge (e.g. water property, light scattering pattern) of a specific underwater scene, enhancement methods are quite popular since they are not based on the physical knowledge like the former one. Moreover, there are two types of methods in tackling the underwater image enhancement: single input and multiple input methods. Multiple input methods need a strong dataset in which the images of the certain underwater scene are taken under various environmental conditions where single input methods can blend in most scenes. In this thesis, the single input approach is considered due to its compactness.

When it comes to imaging devices, latest solutions such as laser range-gated technology can bring accurate and less contaminated results. A variety of acoustic and optical imaging systems are constantly developed for unraveling the foggy condition in water. However, not only the cost and complexness but also the difficulties in utility

make the solutions a failure in real scenarios. To add up, high noise and reduced contrast are also the reasons that sonar images are not popular. On the contrary, the usefulness of most of the low-range products already in the market, aiming for monitoring applications is quite doubtful. A system with decent performance on enhancing and recovering lost details in images and videos is actually in need.

After encountering the issues which are previously stated, another problem arises. Although there are appropriate image enhancement methods regarding the underwater investigation, only a few systems which applied these methods are available. Within my personal knowledge, it is almost impossible to build a real-time handling system for these underwater images or videos as processing time and delays exist in stream processing. Although numerous underwater image improvement systems are proposed, there are only a few which have suggested as real-time applications. One of the main factors is the processing time of each method as it takes several seconds to process each frame or image. This thesis will focus on a soft real-time application with a reduced frame rate to meet the deadlines of a stream. Moreover, the scheme of this research can potentially be applied in underwater monitoring practices.

1.2   Objective

The main objective of this thesis is to design a fusion-based underwater image enhancement method to apply it in a video stream as an employable system. The design criteria include the image and video acquisitions, the frequency band selection for video transmission, and the implementation of enhancement method. Another objective is to make a comparison between the performance of the system and products which are on the market.

1.3   Scope of the Thesis

The scope of this research are as follows:
1. Design and implement an actual deployable underwater image enhancement system with video stream transmission

2. Compare the adapted fusion-based method to other contemporary methods to evaluate the performance

3. Build an own underwater test place to verify the proposed architecture.

Chapter 2

Background and Literature Review

2.1 Background

2.1.1 Imaging in Water with Hazy Environment

When a light array acts as a source and an imaging system (e.g. camera) captures an image underwater, there arise three main problems: absorption, refraction, and reflection. Absorption can reduce the volume of light that reaches a specific scene and reflects from the object to image plane inside the camera. This limits the camera viewing range to the object. The intensity of light source as well as the sensitivity of the camera has a direct effect on this viewing range. The loss of contrast in underwater images or videos is mostly caused by refraction. This effect is more obvious in turbid water than the clear one. It diffuses the light not only from the source but also from the object-to-camera part, reducing the light intensity. The critical factor among the three, reflection, is caused by the suspended particles or objects which are within the camera view. The illuminating light reflects from them go inside the camera, causing the backscatter effect [1].

Other than these three problems, the hazy environment caused by dust, water particles and other absorbed particles, is the main reason for poor visibility under water. Haze also contrast details and clearance of underwater scene. The depth of haze model is directly proportional to the appearance of each scene. Among the image degradation models, the model presented by McCartney [2] is quite popular. The model proposed that the observer or the imaging system receives a fraction of light. The rest of the light is diminished or attenuated into the water environment. Regarding the model, the hazy images have linear combination among the air light (A) and direct attenuation (D) [3].

$$I_h = D + A$$
$$I_h = I * t(x) + A_\infty * (1 - t(x))$$

(1)

Where I_h is the hazy image, I is the haze-free image, A_∞ is the air light constant and $t(x)$ is the transmission medium of the amount of light which reaches the imaging device

without scattering. The transmission t can be calculated, assuming a homogeneous environment, as follows:

$$t(x) = \exp(-\beta * d(x)) \tag{2}$$

Where β is the attenuation coefficient of scattering and d is the distance between the scene and observer.

2.1.2 Fusion-based Image Enhancement Method

In order to deal with hazy and noisy images, Ancuti et al. [4] projected a fusion method. The benefit of this method is that the input is a single hazy or color distorted image. This fusion-based method improves the clarity or contrast of such images. Primarily, it is started with the derivation of two images from hazy underwater input. The first image is a white balanced one to make the white objects of the original input white, balancing the color distortion and keeping a more natural color. The second image is the contrast-enhanced image where the lost details due to haze are recovered. Then, weight maps are estimated for each pixel to decide the contribution of two derived images. This results in the final output image with a more appealing state.

Regarding the weight maps, different kinds of weights are used according to one's objective. They aim to compute pixel size manner to enhance details in the spatial domain. For the last stage, multi-scale fusion is applied to fuse the inputs and the normalized weights driven by the Laplacian pyramid and Gaussian pyramid, respectively.

A.   White Balancing Methods

Firstly, it is expected that only one light source illuminates the scene. The sensitivity function of the imaging device or camera and the original color of light source both have an influence on the observed color of an image. It is stated in [5] that three factors: the surface spectral reflectance, the illuminant spectral power distribution, and the sensor spectral sensitivities define the image values.

$$P(x) = \int_{\omega} I(\lambda) S(x, \lambda) C(\lambda) d\lambda \tag{3}$$

Where $P(x)$ is the image values (P_R, P_G, P_B) at the pixel coordinate x, the wavelength of the light λ, visible spectrum ω, $I(\lambda)$ is the spectral power distribution of illuminant, $S(x, \lambda)$ is the spectral reflectance of the corresponding surface, and $C(\lambda)$ indicates the spectral sensitivities of the imaging device.

The white balancing step is performed to eliminate the distorted color in underwater images where the water is in turbid condition. The traditional principle for white balancing step is carried out, using the approximation of the illuminant and estimates the light source's color in the image values (P). Then, by using the estimated illumination, the original values are transformed. The transformation model is defined as:

$$F_t = D \times F_u \tag{4}$$

$$\begin{pmatrix} R_c \\ G_c \\ B_c \end{pmatrix} = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} R_u \\ G_u \\ B_u \end{pmatrix} \tag{5}$$

In here, F_u is the original output without the prior information about the light source, D is denoted as a diagonal mapping matrix between the two images using canonical illuminant c [32], which is shown in Equation, while F_t is the transformed image using D under the condition of c.

However, this basic model fails to define cases like inter-reflections and highlights. Moreover, the illuminant spectral power distribution and sensor sensitivity of the imaging device are generally unknown, causing the problem to be constrained. Most algorithms are usually based on restrictions and assumptions. Recently, a framework with constant color is proposed by Van de Weijer [6] and the model is as follows:

$$\left(\int \left| \frac{\partial^n f^\sigma(x)}{\partial x^n} \right|^p dx \right)^{\frac{1}{p}} = k e^{n,p,\sigma} \tag{6}$$

Where $\|.\|$ is the Frobenius norm, n indicates the derivative order and σ is the scale parameter of the Gaussian filter, where the image is convoluted by the derivative with that filter. There are a few popular white balancing algorithms which are based on the manipulation of (n, p, σ) variables from Van de Weijer's model. By assuming that the

average reflectance of the scene is achromatic, a method called Gray World [34] is constructed, while the variables (n, p, σ) are set to (0, 1, 0), respectively. White Patch method [7] can be generated by setting the variables to (0, ∞, 0) with the supposition of the maximum reflectance being achromatic in a scene. With the statement of *pth* - Minkowsky norm being achromatic and the setting of (0, *p*, 0), Shades of Grey [8] algorithm is obtained. Gray Edge [9] method is gained by assuming the *pth* - Minkowsky norm of the first-order derivative being achromatic in a scene and manipulating the variables to (1, *p*, σ).

B. Contrast Enhancement

Contrast enhancement stage is added to amplify the find details and visibility. However, we need to be careful about the fact that sometimes the details will be probably destroyed due to over enhancement.

C. Weight Maps

Restoration of a hazy cannot be completed only with white balancing and contrast enhancement stages. We need weight maps to define how much each pixel contributes to the final output [10].

The brightness of a light is measured in Luminance and hence, it is used as a weight map to assign small values to bad visibility regions and high values to good visibility regions. This following equation is used to calculate the Luminance weight:

$$W_L^k = \sqrt{1/3[(R^k - L^k)^2 + (G^k - L^k)^2 + (B^k - L^k)^2]} \tag{7}$$

Where L is the Luminance, k indicates the derived inputs and (R, G and B) denote the three-color channels of these inputs. The weakness of this weight is that it reduces the color information and the overall contrast of the image.

Chromaticity weight map is usually applied to control the gain of saturation for colorfulness in the resulting output. It is calculated as follows:

$$W_C^k(x) = \exp\left(-\frac{(S^k(x) - S_{max}^k)^2}{2\sigma^2}\right) \tag{8}$$

Where σ is the standard deviation, S represents the saturation of input, S_{max} is the maximum saturation range (i.e. 1 for high saturated pixels) and C is the Chrominance.

Another parameter that highlights the objects' existence or a pixel with its neighborhood is visual saliency to grasp the significance. The equation for this weight is computed form the Achanta et al.'s [11] algorithm as follows:

$$W_S^k(x) = \left\| I_k^{\omega_{hc}}(x) - I_k^{\mu} \right\|$$ (9)

Here, I_k^{μ} denotes the arithmetic mean value of derived input and $I_k^{\omega_{hc}}$ represents the blurred form of the input, eliminating high-frequency components.

The Local Entropy weight map is denoted with a structuring element of a certain pixel neighborhood and its complexity in a color image. This weight mainly manipulates the changes in the distribution of local gray level. This weight map divides the image into separate regions as different entities of information. The entropy of a window which has a small neighborhood of Ω_k window with $M_k \times N_k$ size can be calculated as:

$$W_E^k = \sum_{j=0}^{G-1} P_j \log_2 \left(\frac{1}{P_j} \right)$$ (10)

Here, an image has gray level G with a neighborhood of Ω_k, then the probability of gray level j can be obtained from the equation, $P_j = \frac{n_j}{M_k \times N_k}$ and n_j is the amount of pixels in that window.

Moreover, the human visual system contributes to this visibility weight map with clarity as a characteristic. The weight with size $M \times N$ can be defined by:

$$W_V^k = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{|F(m,n) - \mu|}{\mu^{\alpha+1}}$$ (11)

Where derived input is indexed by k, the gray value at (m,n) pixel position is denoted by $F(m,n)$, visual constant α and mean intensity value μ of the derived input.

The process of approaching an image to black and white when it has low saturation is decided by saturation weight map where the color of the intensity plays the main role. It [12] is calculated by:

$$W_S^k(x,y) = \sqrt{\frac{(R^k - m)^2 + (G^k - m)^2 + (B^k - m)^2}{3}} \qquad (12)$$

Where m is calculated by the equation, $m = \dfrac{R^k + G^k + B^k}{3}$ and derived input is represented by k.

D. Multi-scale Fusion

To combine the information from several images into a single one, the simple process, image fusion is usually applied. Image fusion levels can be classified into pixel level, feature level, and the decision-making level. The low level, pixel level fusion, can retain original information more since the images are processed pixel by pixel from original image or data. This level can also act as a support or input for other level fusion. It also requires more metric for image matching since it works on the pixel level. Feature level fusion [13] is a method that fuses characteristics like point, edge, angle, and texture which are extracted from input images to obtain desired results. This level also supports as an input for the decision-making level. It does not require a high-level metric like pixel level fusion and data size is also diminished. This makes it easier for information compression and data transmission. The top level, decision-making level fusion [14], receives the data from the pixel or feature levels as an input and makes a decision based on them. Then, objective extraction and classification are performed on source images and process them regarding the criteria and specific objective. This level is mainly used in artificial intelligence application and object recognition.

To focus on the pixel level fusion, the intention of it is to process content details of multiple images and combine them to gather the desired information. The multi-scale fusion is performed in a per-pixel manner to provide a fast operation for image fusion. Initially, the weight maps resulting from derived inputs are normalized to provide the same scale since applying directly the weights can produce halo artifacts due to naive blending. The Gaussian pyramid is then used with the low pass filter to estimate the

normalized weight-map. Another step is performed by applying a Laplacian operator along with a bandpass filter and down-sampling over the derived inputs. This action can boost details specifically around the edges. The final step is to fuse two pyramids level by level, forming a fused pyramid by the equation as follows [10]:

$$F_l(x) = \sum_k G_l\{W^k(x)\} L_l\{I_k(x)\}$$ (13)

Where l denotes the pyramid level, $L\{I\}$ is the input of the Laplacian pyramid and while $G\{W\}$ characterizes the input of the Gaussian pyramid. Finally, the de-hazed output is produced.

2.1.3 Quality Assessment Methods

Regarding the verification of the performance of image fusion, a method of evaluation or quantitative assessment is needed. There are two categories in quality assessment methods: subjective and objective evaluation methods. A variety of these methods can also be seen in [15]. The subjective assessment method is constructed by portraying the human visual system. Since this category alone cannot evaluate the performance, the objective evaluation methods are used. The following methods are examples of the objective assessment methods.

A.   Mean Square Error (MSE)

MSE or the average of the errors [16] is a difference between the original image and the enhancement image with the equation:

$$MSE = \frac{1}{A \times B} \sum_{A,B} [I_1(A,B) - I_2(A,B)]^2$$ (14)

The lower the MSE represents the less noise in resulting image.

B.   Peak Signal to Noise Ratio (PSNR)

PSNR [16] is the ratio between the maximum power and power of affecting noise in a signal. It can be defined by an equation as:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$ (15)

PSNR is also basically used for evaluating noisy images. The higher the value of PSNR, the better the image with less noise.

C.   Edge Intensity (EI)

For the evaluating the textural details in an enhanced image, EI [17], [18], [19], is applied. It is designed with Sobel operator as follows:

$$EI = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} Gij \tag{16}$$

where Gij is the calculated value from the enhanced image.

2.1.4 First-person-view (FPV)

First-person-view (FPV) is a video piloting method mainly used in drones to radio-control using a camera mounted on them. Video transmitting and receiving devices are produced to enable this short range controlling FPV method. There are two transmission channels in FPV:

(i)   a control uplink channel that transmits control signal from the pilot to the drone using the remote control (RC), and

(ii)   a video downlink channel that sends video stream from the camera to ground station.

In a basic setup, either for video or remote-control link, three elements are needed. The first one is a transmitter for sending the video signal to a ground station. The second one is a receiver which is placed at the ground station to receive the signal sent by the transmitter. Lastly, a correct type of antenna that suits both the transmitter and receiver is required either for short range or long range. In other word, the data captured by a camera is sent to the video transmitter. It is then further transformed into a radio signal by the transmitter and transferred to the receiver from its attached antenna. The receiver then converted the signal from the transmitter to data and display on the monitor.
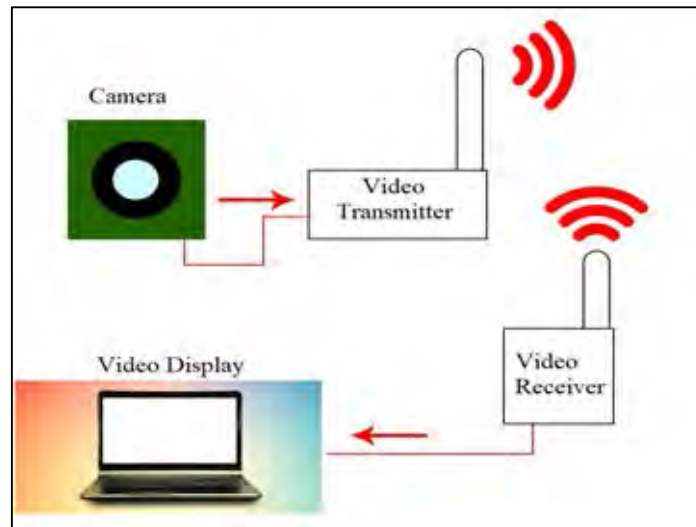
Figure 2.1: Basic FPV setup.

At 2.4 GHz or 5.8 GHz frequencies, the FPV system permits the transmission up to 2 to 3 kilometers (around 1 to 2 miles). Regarding the long-range communication, more equipment is needed. to your system. On the Remote-Control part, an additional ultrahigh-frequency (UHF) 433 MHz module which has a superior function than the 2.4 GHz or 5.8 GHz RC transmitter, a high-power battery (approximately 500 milli-watts) for the UHF 433 MHz module [15]. For the video part, an external video receiver, a battery, and a high gain antenna are required. The following table shows the specification of the popular FPV frequencies.

Figure 2.2: FPV transmitter and receiver.

2.1.4.1 FPV Frequencies: Pros & Cons

There are four types of frequency that can be utilized for video transmission. Each of it has different characteristics and range so that the users can choose according to their preferences in the application.

Table 2.1: FPV frequencies specifications [20].

Frequencies	Range and penetration	Pros	Cons
900 MHz	- medium-range - ideal for low and behind objects flights (Non-Line of sight)	- great penetration	- old technology, bad design equipment, and sensitivity - little interference caused by the cell phone companies - hard to go beyond 5 Km (3 miles) - bad and small antennas, or very big antennas - picture quality is the worst because of the lowest video bandwidth
1.2/ 1.3 GHz	- long-range - ideal for low and behind objects flights (Non-Line of sight) - in need of high gain antenna (around 8dbi) and external receiver with extra battery	- great range and penetration	- degrade GPS performance due to GPS satellite frequency utility - big antennas

2.4 GHz	- long-range	- great range - latest technology, - plenty of design choices with high performance - small and excellent quality antennas	- poor penetration (walls and trees) - Signal interferences in the urban area (Heavy Wi-Fi/Bluetooth area)
5.8 GHz	- short-range - the range can be variable according to antenna usage	- plenty of design choices with high performance - Does not require ground station - tiny antenna - portable setup - excellent video quality	- poor penetration - problem with humidity

Regarding the Relation between frequency and the size of antenna, the antennas are larger in low frequencies (900 MHz, 1.3 GHz) than high frequencies (2.4 GHz, 5.8 GHz) because of their longer wavelength.
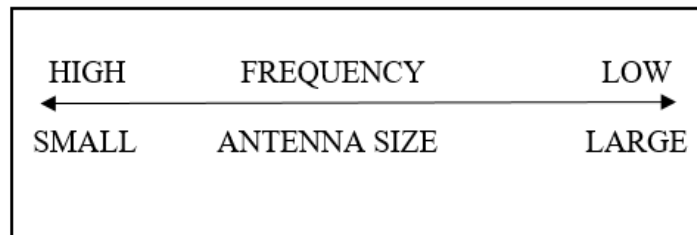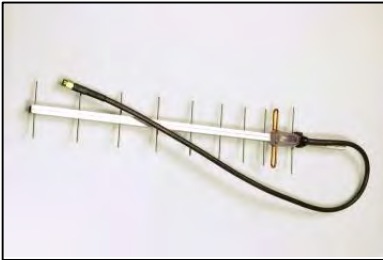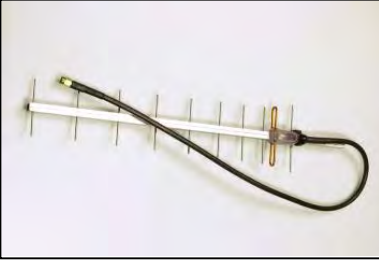
Figure 2. 3: Relation of frequency and antenna size.

2.1.4.2 Antenna Types

There are a variety of antennas available for both 5.8 GHz and 2.4 GHz frequencies ranging from the small antenna to the huge one. The table below presents the types and specifications of each antenna.

Table 2.2: Frequencies and antenna types [20].

Frequencies	Types of Antenna for 2.4 GHz Frequency	Specifications
2.4 GHz (Receiver)		Whip (3dbi) Omnidirectional Length: 11 cm Range: 1-2 km
		Whip (7dbi) Omnidirectional Length: 34 cm Range: 7 km
		Yagi (11dbi) Directional Length: 50 cm Range: 7-15 km

		Patch (14dBi) Directional Length: 21 cm Range: 7-12 km
		Yagi (19dBi) Directional Length: 100 cm Range: 20-30 km
2.4 GHz (Transmitter)		Cloverleaf Circular polarized antenna Omnidirectional
		Whip (7dbi) Linear Polarized antenna Omnidirectional
5.8 GHz (Transmitter/Receiver)		2.4 Spiro NET RP-SMA Omnidirectional

		Whip (3dbi) Omnidirectional

2.2 Literature Review of Image Enhancement Methods

Several researchers have proposed different methods in order to deal with noisy underwater images. As it is mentioned earlier, there are basically two approaches to handle the problem: physical model-based and non-physical model-based methods. Since physical model-methods need prior information about the specific scene, they do not blend in every scenario. This thesis will be focused on non-physical model-based methods where any physical information about the scene like the optical properties and lighting conditions are not required. Under the non-physical model-based approach, a variety of spatial and transform domain enhancement methods are proposed. These methods will be reviewed by categorizing into two parts: multiple inputs and single input methods.

For grayscale images with low exposure value due to underwater dim light condition, Kaur and Verma [16] suggested a method using histogram equalization and Artificial Neural Networks (ANN) to retrieve the information of interest. Underwater images are mostly affected with a hazy appearance and color changes due to light absorption. Chiang and Chen [21] used wavelength compensation and image de-hazing (WCID) algorithm which is inspired from dark-channel prior (DCP) [22]. This method, DCP, is mainly used for evaluating the distance between object and camera. Then images are segmented into foreground and background regions with a depth map. The influence of artificial light source is removed first to prevent the unbalanced light distribution and volume of haze is estimated from DCP. Then, haze is removed from the original image by differencing the interpolated haze model. To remove the dominant color, energy compensation of each color channel according to water depth is carried out. Khan et al. [23] also offered a method for hazy images in which the original image

is duplicated first. Each duplicated image becomes the inputs to color balancing stage and contrast enhancing stage, separately. For color correction, histogram stretching is performed in 'V' component of HSV (Hue-Saturation-Value) color space. The resulted image is then changed back to RGB space and histogram stretching is further implemented in all three RGB components, bringing out a bright and color balanced image. To attain a contrast enriched output, contrast limited adaptive histogram equalization (CLAHE) [24], [25], [26] is carried out. Next step of the system includes accomplishing the wavelet fusion [27], [28] which is composed of three sub-steps: decomposition, fusion and inverse composition. Firstly, the 2-level 2-dimensional image decomposition is implemented. In the first level, the rows of the original image are delivered to the low pass filter and high pass filter, each of which has a down-sampling function (a factor of 2). This results in horizontal approximations and details. Then, the columns are processed as the same way and this would produce approximate (LL), vertical detail (LH), horizontal detail (HL) and diagonal detail (HH), as four sub-images. Secondly, the approximate (LL) becomes the input of the second level and the scale-down process is reiterated. Finally, wavelet coefficients are fused into one image according to maximum value coefficients, which is presented in Figure 2.4.

$LL_{level-2}$	$HL_{level-2}$	$HL_{level-1}$	
$LH_{level-2}$	$HH_{level-2}$		
$LH_{level-1}$		$HH_{level-1}$	

Figure 2.4: Fusion stage for decomposed wavelet coefficients.

Another approach in underwater image enhancement is to use reference images (in-air images) and the color information of these images is applied in manipulating the color of underwater images. A system is projected by Li et al. [29] in which it transfers semantic color values from the reference image and these values are employed on input underwater images while keeping their original contents and structure. This method is not data sensitive unlike WaterGAN [30] which is based on the data information and

pre-acquired images of the environment. The underlying structure is to perform a mapping between the source (underwater image) and target (air image) images by calculating three loss functions: Adversarial loss, Cycle Consistency loss, and SSIM loss. He et al. [5] also came up with a system which adapted Reinhard-based method for correcting color inequality. It is then combined with Pulse-Coupled Neural Networks (PCNNs) for enhancing contrast. To compensate the underwater images which suffer from greenish-blue appearance, the input image is transformed to $l\alpha\beta$ space and its distribution is balanced regarding the natural scene (in-air) image. After balancing the green channel, I component from HIS (Hue-Intensity-Saturation) color space is separately enhanced with PCNN network.

Chapter 3

Research Methodology

Unlike air images, underwater scenes are mostly suffered from the haze, which is a phenomenon triggered by fine dust, smoke, dispersed particles, light vapor and makes an opalescent appearance. Air transparency inside the underwater scene is reduced by scattering and creating a less visible environment. The images taken under this influence are automatically affected in clarity and contrast details. In this thesis, a method to recover noisy images from a single input image is presented. Regarding the system constraints, the limitations of the research are as follows:

- The system aims to focus on single input image processing since there will be no prior data of a specific environment or dataset.
- The programs must have low computational time because the usage is intended to be nearly real-time.

The overall system involves three separate tasks where the underwater video is acquired by the camera and it is sent to the computer through 5.8GHz Transmitter/Receiver set with frame by frame architecture. The frames are then received by the receiver module which is plugged into the computer. The MATLAB program in the computer accepts the data and runs the underwater image enhancement code to produce de-hazed output.

Figure 3.1: Data acquisition tools.

3.1   Architecture of Proposed Fusion-based Image Enhancement Method

The success of a noble fusion method depends on the input choices and decent weight map definitions. One advantage of this method is that it does not need multiple inputs as multi-scale fusion and enhancing can be carried out with one degraded input. Firstly, the color cast of original underwater image can be corrected with several balancing methods (e.g. white balancing) as underwater images are dominant by greenish or bluish color tones. Secondly, noise reduction and preprocessing techniques are applied. The multi-scale fusion aims to remove the artifacts caused by illumination and enrich the image quality [32], [33]. The framework of the fusion method is presented below:
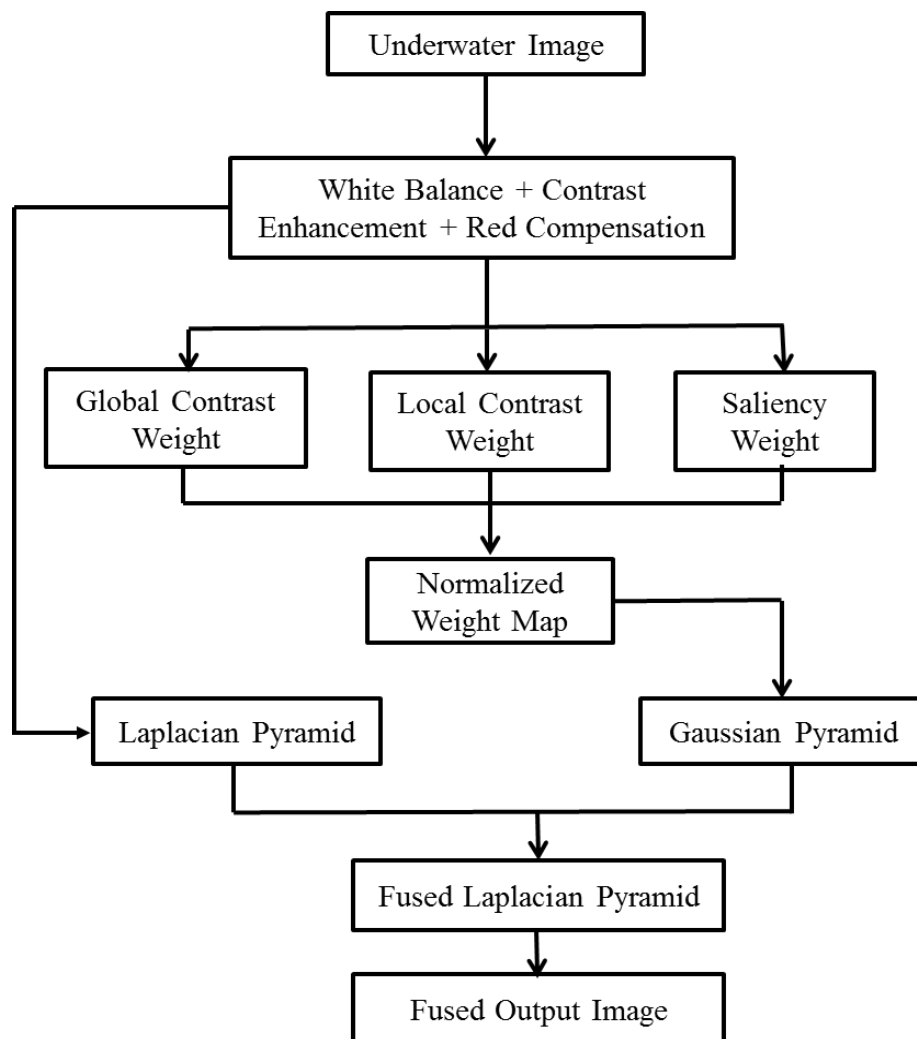

Figure 3.2: Proposed fusion-based method.

A.   White Balancing

Color histogram stretching method is used in this thesis to perform the white balancing step on color-distorted underwater images. The general aim of this algorithm is to stretch the entire histogram range so that the appearance frequencies of pixels are evenly distributed. In this case, the value zero is on the lowest intensity side and the highest value would be on the highest, providing a good distribution with a lighter appearance for light pixels and black for dark pixels. Since a color image has three channels (R, G, B), three histograms for each channel will be acquired. When a scene is captured, and we see a peak in all three channels, that's the point where the image is neutral with correct color temperature. To give an example, when the green channel has many highlights the greenish image is gained.

The idea of color histogram stretching algorithm in order to get an image with no color cast is that two thresholds are calculated first and the color histogram for each channel is stretched according to these thresholds. This algorithm can be designed to an equation as follows:

$$C_{out} = \frac{(C_{in} - L)}{(H - L) \times range + c\_min} \tag{17}$$

Here, C_{in} and C_{out} are the input and output tonal value of a pixel respectively, the parameter H is the highest threshold while the parameter L is the lowest in acting as limits to the color histogram, c_min is the lowest tonal value of pixels (default value, 0) and $range$ is the output tonal range (default value, 255). H and L are defined as the value which is higher than the other 99 percent and lower than other 99 percent of pixels respectively. And then, every pixel from RGB channels can be processed according to them.

The following images are the comparison of color histogram stretching methods used in this thesis and other contemporary white balancing methods. For the objective assessment, the results can be seen in Wang et al.'s [34] with the saturation value comparison between the proposed method, Gray World (GW) [35], White Patch (WP) [7], Shades of Gray (SG) [8], and 2order-Grey Edge (GE2) [6].

(a)   (b)

(c)   (d)
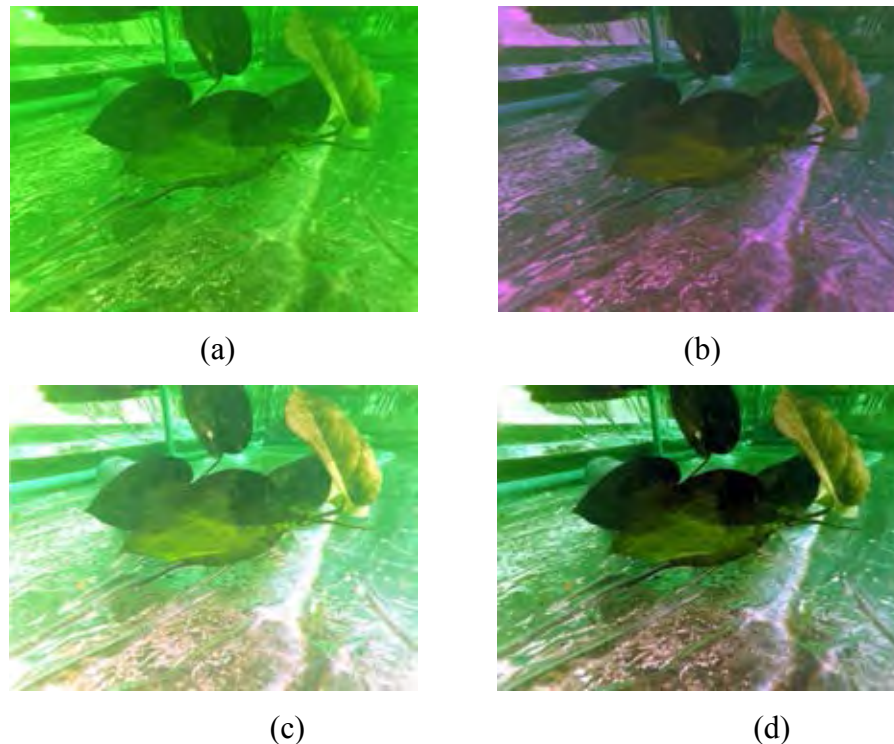
Figure 3.3: (a) original image (b) Gray World [35] (c) White Patch [7] (d) Color Histogram Stretching [34]

From Figure 3.3, the performance of three white balancing methods can be seen. Gray World algorithm produces a reddish blue image while the White Patch algorithm introduces an over-enhanced result. Color histogram stretching method used in this thesis shows a better performance with stable enhancement result. Moreover, according to Wang et al.'s [34], it has the lowest computational cost than other methods. It would be suitable to use in real-time systems.

B.   Contrast Enhancement

Contrast Limited Adaptive Histogram Equalization (CLAHE) [36] is used in order to improve the local contrast of the white-balanced input in this thesis. Firstly, the histogram of a pixel with its neighboring block is equalized. Then, the regions where the contrast is lighter or darker will be enhanced by using the transformation function or cumulative distribution function (CDF) resulted from its neighboring pixels. CLAHE uses the clipping limits at predefined value in the histogram before computing the transformation function. The limit depends on the histogram normalization. This stage reduced the noise amplification more than the original adaptive histogram equalization

(AHE) method. This stage also restricts the slope of the function. The mathematical equation for linear histogram stretching is as follows:

$$Y(n_1, n_2) = \frac{255}{x_{max} - x_{min}}(x(n_1, n_2) - x_{min}).$$ (18)

The following images display the effect of CLAHE on the underwater images.

(a)   (b)   (c)
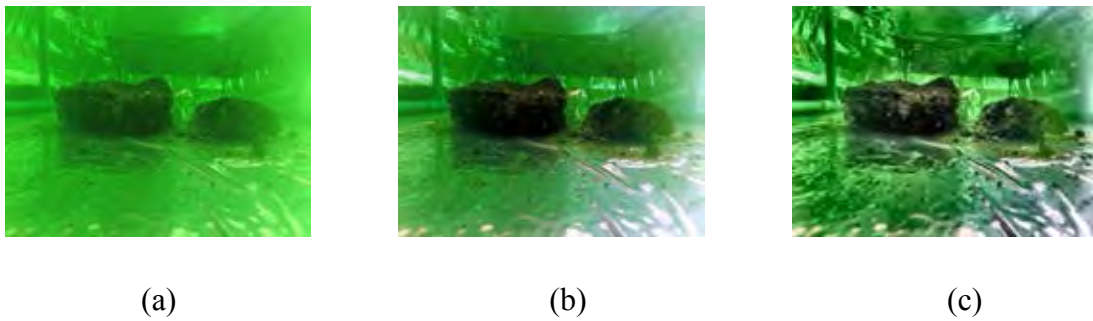
Figure 3.4: (a) Original image (b) Before applying CLAHE (c) After applying CLAHE

C.   Red Compensation

Although most of the color cast are removed by the usage of white balancing method and contrast Enhancement method, in some scenarios such as highly turbid water, light attenuation can cause severe red artifacts. A specific amount of mean value from the red channel results in over-compensation in some neighborhood which has a presence of red color. This can be manipulated by adding a portion of green channel to red channel [32]. The mathematical expression for this observation is as follows:

$$I_{rc}(x) = I_r(x) + \alpha.(\bar{I}_g - \bar{I}_r).(1 - I_r(x)).I_g(x)$$ (19)

Here, $I_{rc}(x)$ is the compensated red channel, $I_r(x)$ and $I_g(x)$ represent the original red and green channel, $\bar{I}_g$ and $\bar{I}_r$ are the mean value of each channel, and α is the constant parameter ($\alpha = 1$ for various illumination and image acquisition conditions). Figure 3.5 illustrates the effects of red compensation in images.
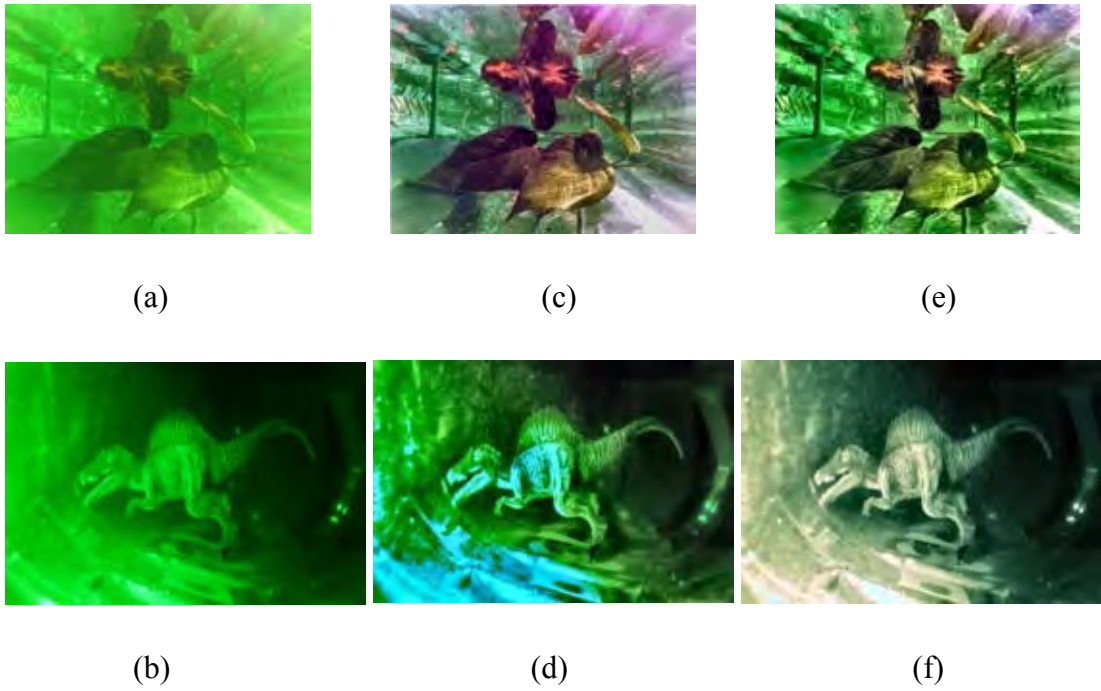
Figure 3.5: (a)-(b) Original noisy images, (c)-(d) Image enhancement without red compensation, (e)-(f) Image enhancement with red compensation

D.   Weight Maps

The first weight map in the fusion-based system is the global contrast weight [37]. The luminance value of the input is filtered by Laplacian filter and produced the absolute value. It is mainly used to extend the depth of edge and texture regions. Since the first weight map only is not sufficient and it cannot differentiate the flat and ramp regions well, the local contrast weight map is used [33]. This weight mainly strengthens the highlighted regions of the contrast enhanced image. The mathematical expression of the local contrast weight is as follows:

$$W_{Lc}(x, y) = \left\| I^k - I^k_{\omega_{hc}} \right\| \tag{20}$$

Where, I^k and $I^k_{\omega_{hc}}$ represent the luminance channel of original input and the low pass version of it, respectively. The cutoff frequency value ω_{hc} is $\pi/2.75$.

The third weight map is the visual saliency one which is known as the perceptual quality, making the object or pixel to stand out among the neighboring environment

and capturing attention. The saliency map for an image can be formulated using Achanta's approach [11] as:

$$S(x,y) = \left| I_\mu - I_{\omega_{hc}}(x,y) \right| \tag{21}$$

Where I is the image with width W and height H pixels, I_μ is the arithmetic mean of the image and $I_{\omega_{hc}}$ is the blurred output of Gaussian filter for elimination of texture, noise, details, and artifacts. By using the norm of difference, the magnitude is obtained, and it is efficient in computational time. Finally, each pixel position is gained as a $[L,a,b]^T$ vector using *Lab* color space while calculating the Euclidean distance and fulfilling salient region detection.

| (a) | (b) | (c) |
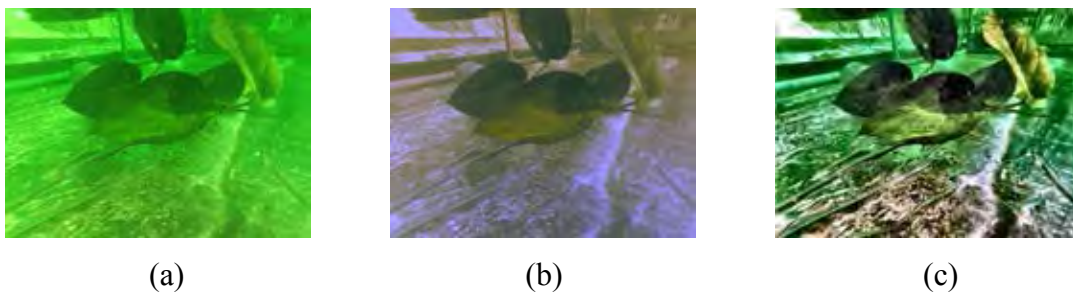
Figure 3.6: (a) Original image (b) Ancuti et al.'s result [32] (c) Proposed method's result

Figure 3.6 makes a comparison with subjective evaluation on the outputs of Ancuti et al.'s method and the proposed method with different weight maps. It can be seen that Ancuti's method can get rid of the green color cast, however, global contrast is reduced. The proposed method provides a good effect in white balancing part with brighter output.

E.   Quality Assessment Methods

Among the various methods for performance evaluation, the methods used for underwater image enhancement are quite a few because it is problematic in acquiring the ground truth. Since a reference image which is captured with standard illumination for comparison cannot be easily obtained, methods which need ground truth like PSNR [16] and MSE [16] are not an option. In this thesis, two objective metrices which do

not require information like the property of water body or reference image are applied to test the performance.

The first metric is the measure of enhancement (EME) [38] which is mainly used for evaluating the quality of improvement. Assume that an image $X\ (n,\ m)$ is divided into blocks $k_1 k_2$ of size ($l_1 \times l_2$), we can express the EME as:

$$EME = \max_{\Phi \in \{\Phi\}} \chi \left(\frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} 20 \log \frac{I_{max;k,l}^{\omega}}{I_{min;k,l}^{\omega}} \right) \tag{22}$$

Where $I_{max;k,l}^{\omega}$ is the maximum value and $I_{min;k,l}^{\omega}$ is the minimum value of image $X\ (n,\ m)$ inside the block $\omega_{k,l}$, and the sign function $\chi(x) = x(or) - x$.

The second metric is the underwater color image quality evaluation (UICQE) [39] which is based on three parameters: CIELab chroma, contrast, and saturation. This metric is mainly used to measure the sharpness and colorfulness of an underwater image. In this metric, CIELab space [40] is used since it is uniform and device independent. For degraded underwater images and survey images, it is assumed that linear combination of these three metrics can represent the colorfulness of images. The variance of chroma is correlated to human perception for color of underwater images while providing a good reason to use. Another metric is a good parameter to use in turbid water environment and it is also one of the most perceived factors. For measuring local contrast of the region of interest from a uniform background, contrast metric is added. It can be defined with an equation as follows:

$$UCIQE = c_1 \times \sigma_c + c_2 \times con_1 + c_3 \times \mu_s \tag{23}$$

Here, with pixel values Ip of an image, σ_c denotes the standard deviation or variance of chroma, con_1 represents the luminance contrast, μ_s is the average value of saturation, and (c1, c2, c3) are weighted coefficients. In luminance channel, the parameter, con_1, can be computed from the difference from the bottom 1 percent and top 1 percent of all pixel values. For images with reduced contrast and color cast, the coefficients are set as c1=0.4680, c2=0.2745, c3=0.2576.

3.2 Frequency Band Selection for Video Transmission

As it is already mentioned in Chapter 2, there are several choices ranging in different frequency spectrums. Among them, 2.4 GHz and 5.8 GHz transmitter/receiver sets are the most popular ones and they are seen in many applications due to their modern technology and availability of a wide variety of choices. The specifications of both frequencies are summarized as follows to highlight what each frequency offers [20]:

[1] Although 2.4GHz frequency also offers small antenna set, it needs a proper ground station while 5.8GHz frequency only needs a tiny antenna and no ground station. It can be used with just a simple monitoring device.

[2] As for penetration ability, both frequencies are almost at the same level with poor penetration. Only lower frequencies like 900MHz have an advantage here as it can pass through obstacles more easily.

[3] 2.4 GHz frequency is more suitable for long-range applications while 5.8 GHz only provides around 2-kilometer.

[4] 2.4 GHz has a great problem of interference in urban areas with heavy Wi-Fi applications since all the remote-control equipment run on this frequency. 5.8 GHz frequency outrange the former frequency in this area.

In this research, the 5.8 GHz transmitter and receiver will be used due to its flexibility, low power consumption capability, great picture quality, and compactness. It also does not have an issue when it comes to portability because with only a small antenna and a monitoring device, it will do its work.

Chapter 4

Experiments and Results

4.1 Implementation of Proposed Fusion-based Enhancement Method on Color-distorted Water

Before we implement the final testbed, the performance of the enhancement method is tested using a container filled with green color-distorted water in a mild condition. The illuminating source is the daylight and the imaging device is Yi action camera. Moreover, the results of several existing methods (i.e. DCP [22], NLD [41], SP [42], Singh's method [43] and Fu's method [44]) can be seen along with the results from proposed method to make a comparison. The MATLAB R2019a simulation tool is used on Intel® Core™ i5-3337U CPU @ 1.80GHz 8.00 GB RAM computer. The images extracted from the camera are (4608x3456) resolution and the images are resized to (640x480) resolution for consistency.

| (a) | (b) | (c) |

| (d) | (e) | (f) |

(g)   (h)   (i)

(j)   (k)   (l)

(m)   (n)   (o)

(p)   (q)   (r)

(s)   (t)   (u)

Figure 4.1: (a)-(c) Original image, (d)-(f) DCP method [22], (g)-(i) NLD method [41], (j)-(l) SP method [42], (m)-(o) Singh's method [43], (p)-(r) Fu's method [44], (s)-(u) Proposed method.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)

(m)

(n)

(o)

(p)   (q)   (r)

(s)   (t)   (u)

Figure 4.2: (a)-(c) Original image, (d)-(f) DCP method [22], (g)-(i) NLD method [41], (j)-(l) SP method [42], (m)-(o) Singh's method [43], (p)-(r) Fu's method [44], (s)-(u) Proposed method.

Regarding Figure 4.1 and 4.2, images resulting from DCP method and NLD method do not show much difference from the original images with the domination of green color even after the enhancement. SP method can remove green color and provides clarity more than the two previous methods. Singh's method presents a non-uniform color cast, making the resulting images more distorted. Fu's method offers more desirable outputs with green color elimination, but some parts of the images are covered with redness. The proposed method can exclude the green color cast to a greater extent. The overall brightness of the image is enhanced, and haziness is also removed.

The second setup is made by adding brown color and powder to the green-colored water, portraying the moderate level turbidity. The room light is acting as the illuminating source in this setup. The conduction of the test can be witnessed as follows.

Figure 4.3: Test container.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)
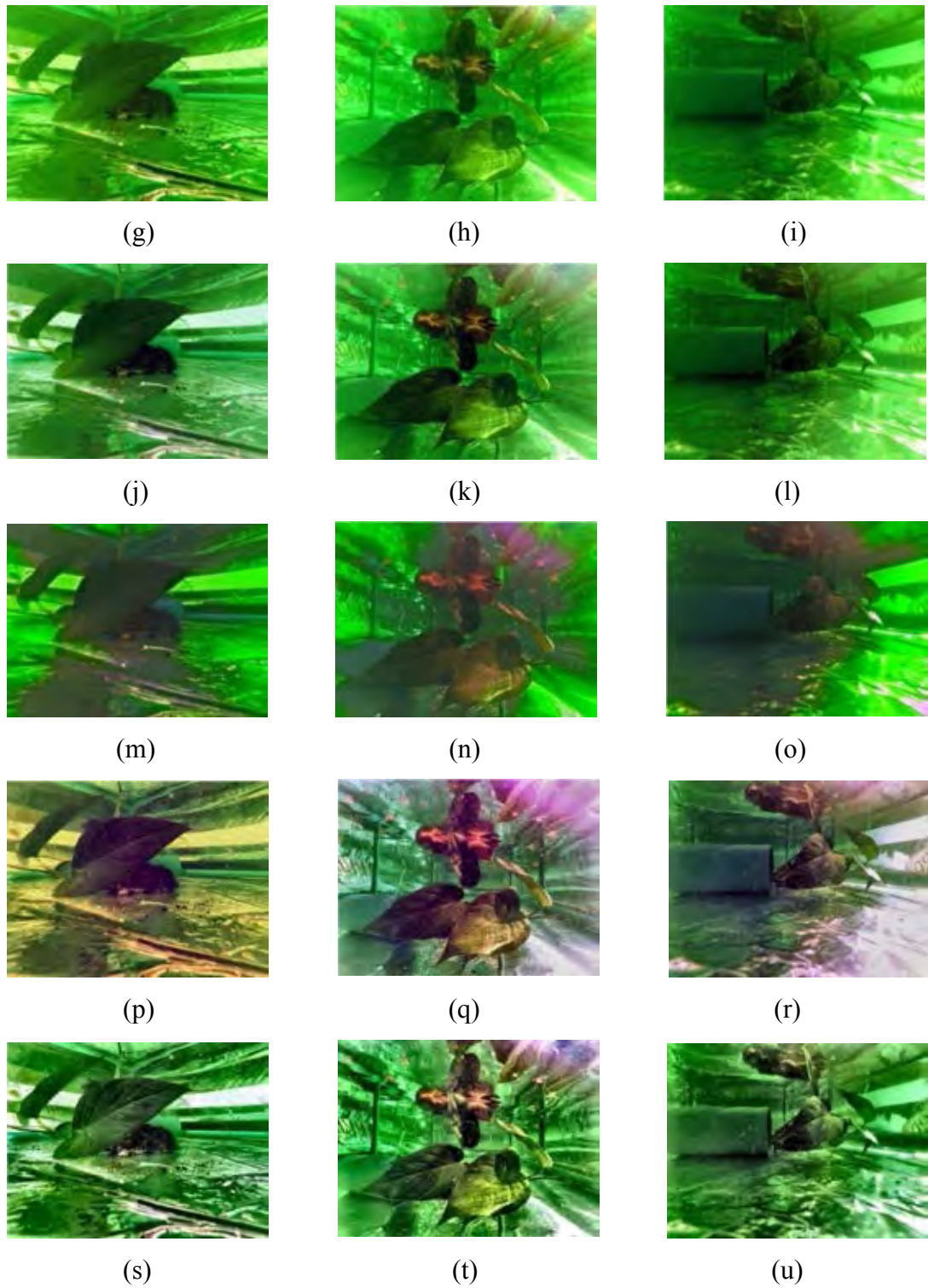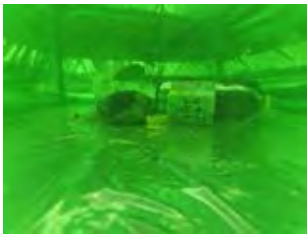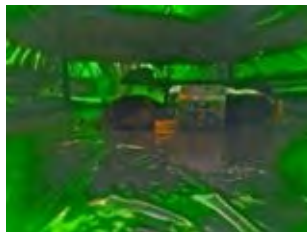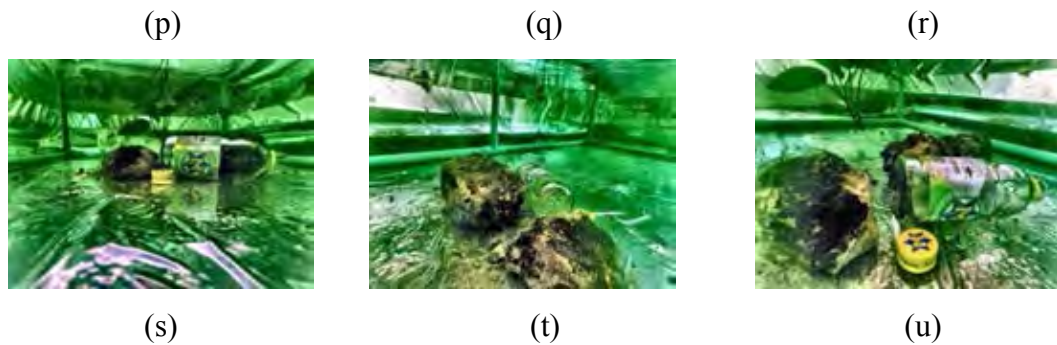
(j)

(k)

(l)

(m)

(n)

(o)

(p)

(q)

(r)

(s)

(t)

(u)

Figure 4.4: (a)-(c) Original image, (d)-(f) DCP method [22], (g)-(i) NLD method [41], (j)-(l) SP method [42], (m)-(o) Singh's method [43], (p)-(r) Fu's method [44], (s)-(u) Proposed method.

Mostly, the same pattern is witnessed with the DCP method failing to remove the color cast and the NLD method showing small improvements. The SP method provides output images with less color cast and Singh's method again introduced the non-uniformity with green colors seen in most part of the images. Fu's method can remove the green color evenly, however, it makes the resulting images yellowish. The proposed method gets rid of the color cast and hazy regions while the original colors of the objects are maintained.

The performance evaluation is done on Singh's method, Fu's method and proposed method because the other methods perform poorly, and they do not differ much from the original image. Table 4.1 shows the quality assessment done on images from Figure 4.1. Table 4.2 displays the result of the evaluation from Figure 4.2. Finally, Table 4.3 represents the images from Figure 4.4.

Table 4.1: Underwater Image Enhancement Evaluation based on EME and UCIQE Metrics. The corresponding images are from Figure 4.1.

	EME [38]	UCIQE [39]
Original image,	4.45800	27.9578
DCP method [25]	5.70701	31.3848
NLD method [26]	7.79280	34.4946
SP method [27]	9.58375	35.5071
Singh's method [28]	6.46529	**38.2668**
Fu's method [21]	12.71191	33.5304
Proposed method	**18.01400**	35.9022
Original image,	4.31294	26.7083
DCP method [25]	5.16367	30.7738
NLD method [26]	6.99734	36.0654
SP method [27]	10.32516	**37.9634**
Singh's method [28]	6.20887	37.1944

Fu's method [21]	13.71588	33.4510
Proposed method	**18.22439**	37.1258
Original image,	4.05109	30.8123
DCP method [25]	4.91939	34.8417
NLD method [26]	6.01293	36.8160
SP method [27]	7.28054	37.4565
Singh's method [28]	5.32939	**41.7490**
Fu's method [21]	9.52816	32.1456
Proposed method	**14.50748**	35.4006

Table 4.2: Underwater Image Enhancement Evaluation based on EME and UCIQE Metrics. The corresponding images are from Figure 4.2.

	EME [38]	UCIQE [39]
Original image,	3.85917	22.9743
DCP method [25]	5.63928	27.0868
NLD method [26]	6.83353	31.1486
SP method [27]	9.47194	34.7039
Singh's method [28]	6.18961	34.1547
Fu's method [21]	11.56502	30.4663
Proposed method	**18.74669**	**35.2500**
Original image,	4.46346	26.9183
DCP method [25]	6.92036	32.5044
NLD method [26]	9.02858	35.5788
SP method [27]	10.65350	35.8055
Singh's method [28]	6.27664	**38.7770**
Fu's method [21]	13.35131	32.6064
Proposed method	**19.31541**	35.8966
Original image,	4.38565	27.4405
DCP method [25]	6.57879	31.9265
NLD method [26]	8.75858	35.7142
SP method [27]	10.28729	36.5428
Singh's method [28]	6.42321	**40.1696**

Fu's method [21]	14.98564	31.7912
Proposed method	**19.03156**	36.7561

Table 4.3: Underwater Image Enhancement Evaluation based on EME and UCIQE Metrics. The corresponding images are from Figure 4.4.

	EME [38]	UCIQE [39]
Original image,	3.23862	28.6997
DCP method [25]	5.78111	38.0535
NLD method [26]	7.53201	**38.8596**
SP method [27]	5.97126	35.6210
Singh's method [28]	3.72576	34.0343
Fu's method [21]	6.68795	28.7219
Proposed method	**13.44797**	35.9374
Original image,	2.60483	22.8688
DCP method [25]	4.00899	32.7713
NLD method [26]	5.51089	36.4465
SP method [27]	5.46081	37.4429
Singh's method [28]	3.17500	25.5360
Fu's method [21]	6.54554	31.7088
Proposed method	**11.86958**	**39.5250**
Original image,	3.84410	33.3546
DCP method [25]	5.04090	37.7402
NLD method [26]	6.15382	**40.7895**
SP method [27]	6.64520	37.5080
Singh's method [28]	4.47096	36.4142
Fu's method [21]	6.68383	32.0790
Proposed method	**10.26205**	36.3837

According to the three tables, it is concluded that the proposed method has the highest value in every image than other five methods in EME metric, hence, a success in improvement factor. For the sharpness and colorfulness factors (UCIQE metric), it is ranked at the top three level, showing an appropriate result in the regarding field. The

proposed method is suitable for middle range turbidity while the highly turbid water condition becomes a limitation for it. Also, the temporal changes between the frames of the video stream should be taken into consideration so that the processed video frames will be smooth.

4.2   Testing FPV Transmitter/Receiver Set's Data Transmission

TS832 5.8 GHz 48 channel transmitter is placed in the transmission part, connecting to the Yi action camera to get the images or video stream and changing it to signal for transmission. The 5.8GHz UVC receiver is used at the receiving part which is connected to the computer with USB 2.0 port. It received the signal from the transmitter and change it to corresponding images or video stream. Figure 4.5 displays the hardware used in this thesis.
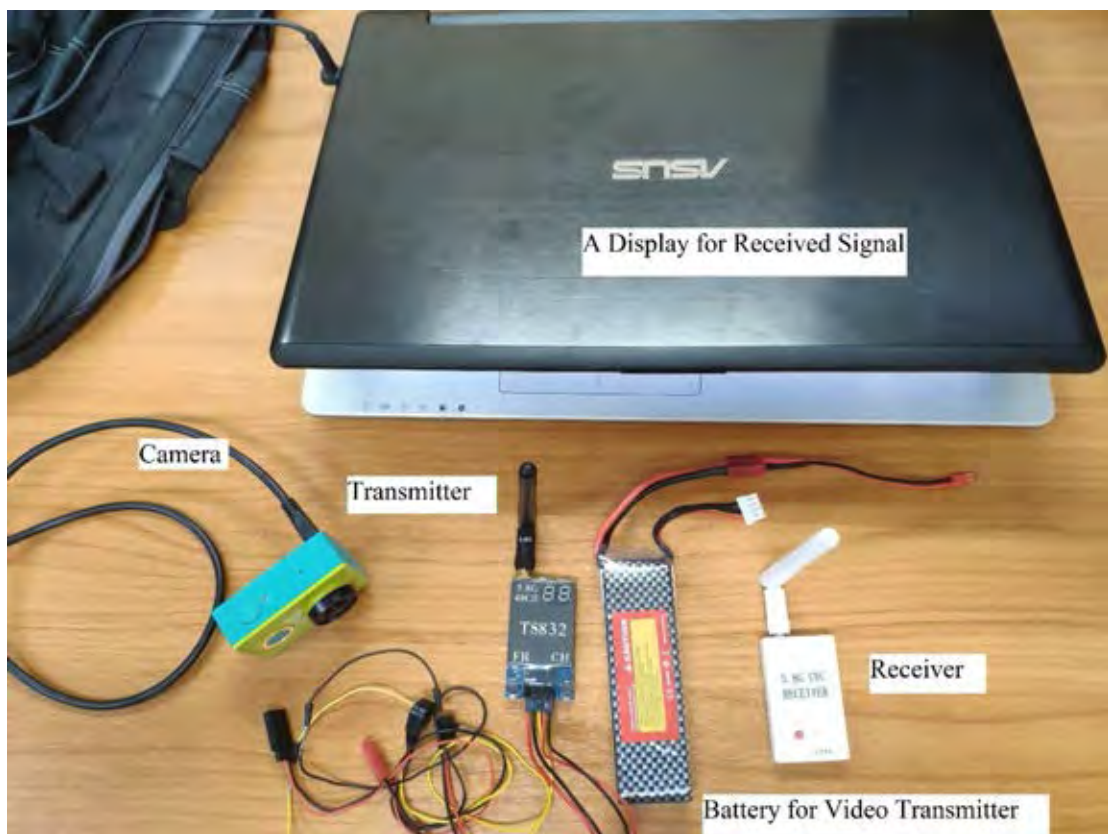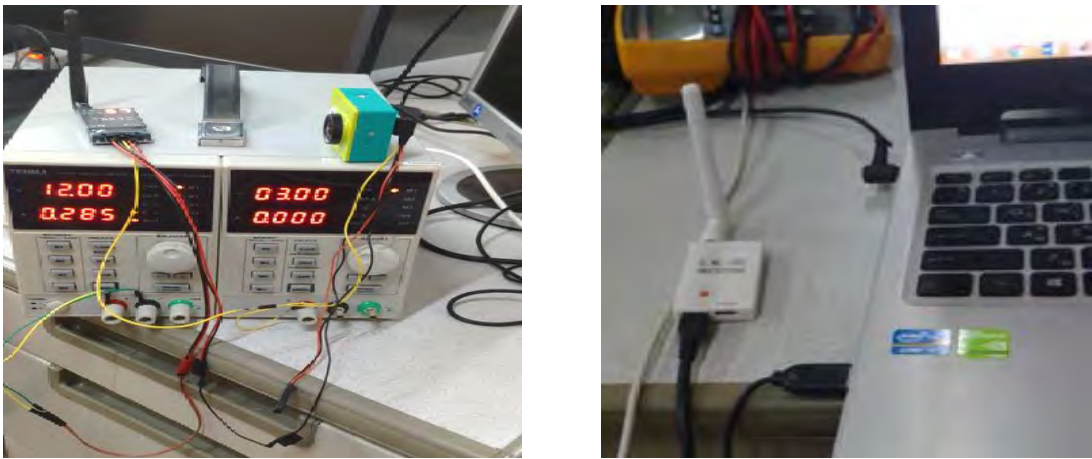

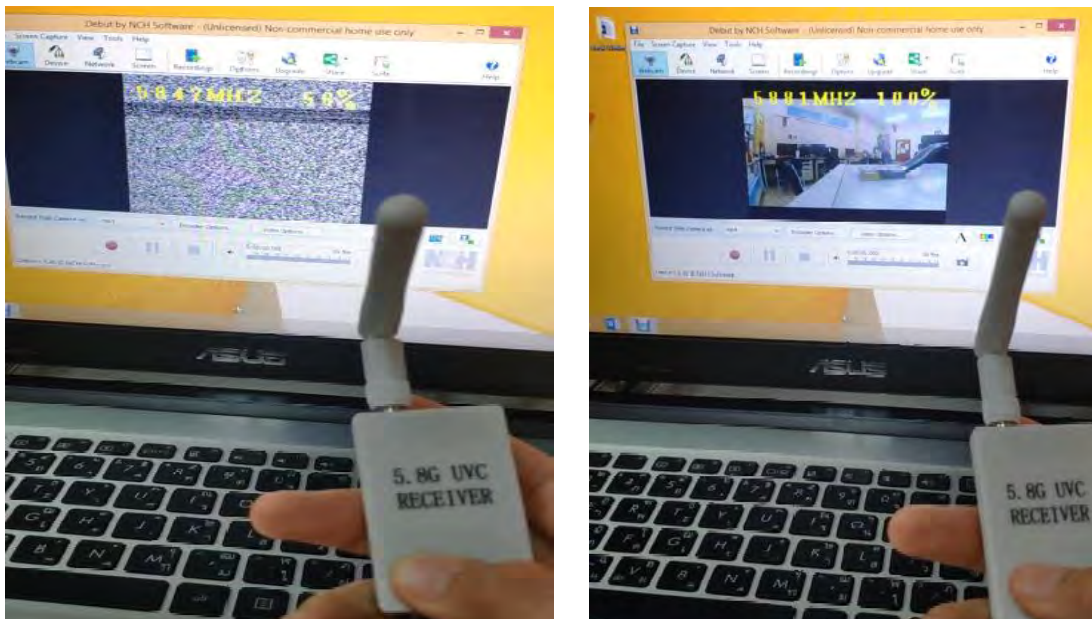
Figure 4.5: Hardware of the system.

The transmitter works at the 12V input voltage and the antenna must be connected first before powering for product safety. The transmitter is linked to the video output of the

camera using a micro USB port from the camera. The receiver is linked with the computer using the USB 2.0 port not only for input power but also for data display. Figure 4.6 shows the hardware connections of transmitter/receiver set. First, the transmitter and receiver must be on the same channel to make the transmission and receiving parts successfully. Figure 4.7 demonstrates that the receiver is tuned to reach the same channel as the transmitter. As regards the receiver used in the system, the set fits at the 5881 MHz and the data is displayed at the computer. The data from the camera come into the computer and MATLAB as 'USB 2.0 PC camera' as shown in Figure 4.8.

Figure 4.6: FPV transmitter/ receiver set in setup.

Figure 4.7: Tuning receiver.

(a)
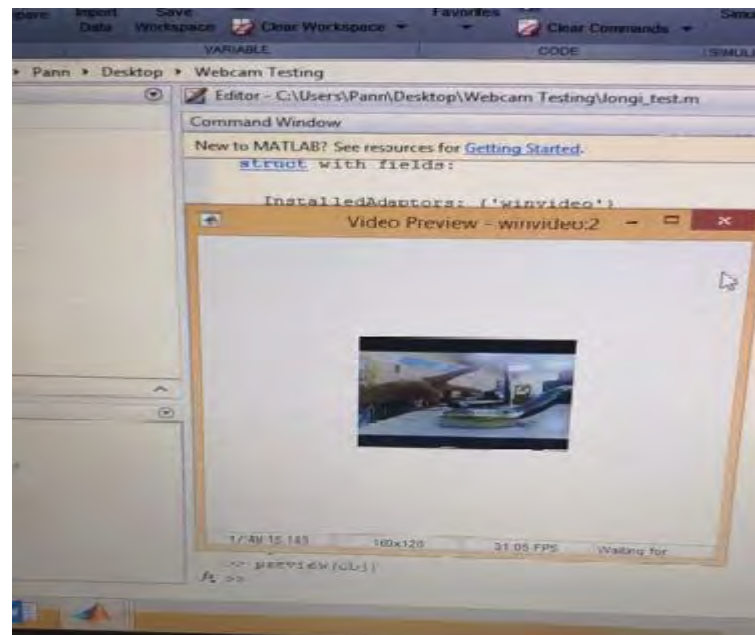
(b)

Figure 4.8: (a) Video input object into MATLAB (b) Video preview from the camera.

Chapter 5

Conclusion

We designed a fusion-based image enhancement system and focus on the underwater noisy images as the dataset. The purpose of this thesis is to implement the image enhancement system to the underwater video stream. By using the 5.8GHz FPV transmitter/receiver set, the desired underwater images or video stream can be manipulated nearly real-time through the antenna. Our own test place is constructed using the green and brown coloring liquids to portray the turbid water condition. The hardware and imaging device are put in a color-distorted water container to acquire the data and it is sent to the receiver part which is connected to the computer from the transmitter antenna. After we received the data from the receiver, it is accepted in the MATLAB R2019a software. Firstly, the white balanced step and contrast enhanced steps are carried out to remove the green color cast. Then, three weight maps (global contrast weight, local contrast weight and saliency weight) are put to decide whether a specific pixel is good enough to contribute in the final output. The results from the weight maps are normalized to set the consistency. Finally, multi-scale fusion is applied on the normalized results to produce a final fused output. The video steam is manipulated in this manner frame by frame although there is a small computational time for the software. We made a quality evaluation of the output with five contemporary methods. The assessment methods used for evaluation are the measure of enhancement (EME) and the underwater color image quality evaluation (UICQE). According to the results of these metrics, our proposed method shows a strong performance with the highest values in EME metric and top three level results in UICQE metric. This work confirms that it shows a good approach in using it in real application. However, it has a limitation on severely turbid water which definitely need a strong artificial illumination system. For a future work, the effects of using artificial light sources and their contribution to the image acquisition should be considered.

REFERENCES

[1]   "TV Camera Lighting in Turbid Water Condition", Kongsberg Simrad Ltd, April 1998.

[2]   E. J. McCartney, Optics of the Atmosphere: Scattering by Molecules and Particles. New York, NY, USA: Wiley, 1976.

[3]   Nitish Gundawar, V. B. Baru, "Improved Single Image Dehazing by Fusion", May. 2014, vol.3, Issue.5.

[4]   C. O. Ancuti and C. Ancuti, "Single image dehazing by multiscale fusion," IEEE Trans. Image Process, vol. 22, no.8, pp. 3271-3282, 2013.

[5]   G. D. Finlayson, S. D. Hordley, and P.M. Hubel, "Color by correlation: a simple, unifying framework for color constancy", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 11, pp. 1209-1221, Nov. 2001.

[6]   J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy", IEEE Trans. Image Processing, vol. 16, no.9, pp. 2207-2214, Sept. 2007.

[7]   E.H. Land, "The retinex theory of color vision", Scientific Am., vol. 237, no. 6, pp. 108-128, Dec. 1977.

[8]   G.D. Finlayson and E. Trezzi, "Shades of gray and color constancy", Proc. IS&T/SID Color Imaging Conf., pp. 37-41, 2004.

[9]   J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy", IEEE Trans. Image Processing, vol. 16, no.9, pp. 2207-2214, Sept. 2007.

[10]   M. Dalvi, D. Shetty, K. Shah and A. Vanmali, "Improved Weight Map Guided Single Image Dehazing", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 5 Issue 03, March-2016.

[11]   R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 1597-1604.

[12]   T.Mertens, J. Kautz and F.V. Reeth. "Exposure fusion", Computer Graphics and Applications. Pacific Conference, 0:382-390,2007.

[13]   G. Piella, "A general framework for multiresolution image fusion: From pixels to regions [J]. Information Fusion", 4(4):259-280, 2003.

[14]   B. Jeon, D. A. Landgrebe, "Decision fusion approach for multi-temporal classification [J]", IEEE Trans. on Geoscience and Remote Sensing, 37(3):1227-1233, 1999.

[15] Pann Mya Hmue, Suree Pumrin, "Image Enhancement and Quality Assessment Methods in Turbid Water: A Review Article", 4th International Conference on Consumer Electronics Asia 2019 (ICCE-Asia 2019), Bangkok, June 2019.

[16] M. Kaur and K. Verma, "A Novel Hybrid Technique for Low Exposure Image Enhancement using Sub-Image Histogram equalization and Artificial Neural network", 2016 International Conference on Inventive Computation Technologies (ICICT), vol. 2, pp.1-5, Aug. 2016.

[17] K. He , R. Wang , D. Tao , J. Cheng and W. Liu, "Color Transfer Pulse-Coupled Neural Networks for Underwater Robotic Visual Systems", IEEE Access, vol.6, June 2018.

[18] S. Yuan, D. Xiang, X. Liu, X. Zhou, and P. Bing, "Edge detection based on computational ghost imaging with structured illuminations", Opt. Commun., vol. 410, pp. 350-355, Mar. 2018.

[19] Y. Yao, G. Yang, X. Sun, and L. Li, "Detecting video frame-rate up-conversion based on periodic properties of edge-intensity", J. Inf. Secur. Appl., vol. 26, pp. 39-50, Feb. 2016.

[20] Oliver A., "Introduction to FPV and TBS Discovery", September 2013.

[21] Y. Chiang and Y. Chen, "Underwater Image Enhancement by Wavelength Compensation and Dehazing", IEEE Transactions on Image Processing, vol. 21, Issue 4, April 2012.

[22] He K.; Sun J., Tang, X., "Single image haze removal using dark channel prior", IEEE Trans. Pattern Anal. Mach. Intell., 2011, 33, 2341–2353.

[23] A. Khan, S. S. A. Ali, A. S. Malik, A. Anwer and F. Meriaudeau, "Underwater Image Enhancement By Wavelet Based Fusion", IEEE International Conference on Underwater System Technology: Theory and Applications (USYS), Dec. 2016.

[24] A. Khunteta, D. Ghosh, and Ribhu, "Fuzzy rule-based image exposure level estimation and adaptive gamma correction for contrast enhancement in dark images", IEEE 11th International Conference on Signal Processing, vol. 1, pp. 667–672, 2012.

[25] S. Pizer, "Adaptive Histogram Equalization and Its Variations.pdf", Computer Vision, Graphics and image processing. pp. 355–368, 1987.

[26] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," Graphics G., P. S. Heckbert, Ed. San Diego, CA, USA: Academic Press Professional, Inc., 1994, pp. 474–485.

[27] A. Anoop Suraj, M. Francis, T. S. Kavya, and T. M. Nirmal, "Discrete wavelet transform based image fusion and de-noising in FPGA," J. Electr. Syst. Inf. Technol., vol. 1, no. 1, pp. 72–81, 2014.

[28] H. Zhang and X. Cao, "A Way of Image Fusion Based on Wavelet Transform", IEEE 9th International Conference on Mobile Ad-hoc and Sensor Networks, pp. 498–501, 2013.

[29] C. Li, J. Guo and C. Guo, "Emerging from Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer", IEEE Signal Processing Letters, vol. 25, issue 3, pp. 323-3, March 2018.

[30] J. Li, K. A. Skinner, R. M. Eustice and M. J. Roberson, "WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images", IEEE Robotics and Automation Letters, vol. 3, Issue 1, Jan. 2018.

[31] K. He, R. Wang, D. Tao , J. Cheng and W. Liu, "Color Transfer Pulse-Coupled Neural Networks for Underwater Robotic Visual Systems", IEEE Access, vol.6, June 2018.

[32] Codruta, Ancuti & Ancuti, Cosmin & Vleeschouwer, Christophe & Bekaert, Philippe, "Color Balance and Fusion for Underwater Image Enhancement", IEEE Transactions on Image Processing, pp. 1-1, 2017.

[33] Ancuti, C & Codruta, Ancuti & Haber, Tom & Bekaert, Philippe, "Enhancing Underwater Images and Videos by Fusion", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2012.

[34] Su Wang, Yewei Zhang, Peng Deng, Fuqiang Zhou, "Fast Automatic White Balancing Method by Color Histogram Stretching", 4th International Congress on Image and Signal Processing, Oct. 2011.

[35] G. Buchsbaum, "A spatial processor model for object color perception", Franklin Inst., vol. 310, no. 1, pp. 1-26, July 1980.

[36] Zuiderveld, Karel, "Contrast Limited Adaptive Histograph Equalization.", Graphic Gems IV. San Diego: Academic Press Professional, 1994. 474–485.

[37] T. Mertens, J. Kautz, and F. V. Reeth. Exposure fusion. Comp. Graph. Forum, 2009.

[38] Sos S. Agaian, Karen P. Lentz, and Artyom M. Grigoryan, "A New Measure of Image Enhancement", IASTED Int. Conf. Signal Processing Communication, 2000-Sep.-1922.

[39] M. Yang, A. Sowmya, "New Image Quality Evaluation Metric for Underwater Video", IEEE Transactions on Image Processing, vol. 24, No. 12, December 2015.

[40] D. Hasler and S. E. Suesstrunk, "Measuring colorfulness in natural images," Proc. SPIE, vol. 5007, pp. 87–95, Jun. 2003.

[41] Dana B., Avidan S., "Non-local image dehazing", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1674–1682.

[42] Morel J.M., Petro A., Sbert C., "Screened Poisson equation for image contrast enhancement", Image Process. Line 2014, 4, 16–29.

[43] Singh R., Biswas M., "Contrast and Color Improvement based Haze Removal of Underwater Images using Fusion Technique", 4th IEEE International Conference on Signal Processing, Computing and Control (ISPCC 2017), Solana, India, Sep 21-23, 2017.

[44] Xueyang Fu, Zhiwen Fan, Mei Ling, Yue Huang, Xinghao Ding, "Two-Step Approach for Single Underwater Image Enhancement", 2017 International Symposium on Intelligent Signal Processing and Communication Systems, Xiamen, China, November 6-9, 2017.

VITA

NAME	Pann Mya Hmue
DATE OF BIRTH	08 June 1995
PLACE OF BIRTH	Mandalay, Myanmar
INSTITUTIONS ATTENDED	University of Technology (Yatanarpon Cyber City), Myanmar. (2011-2016) Chulalongkorn University, Bangkok, Thailand. (2017-2019)
HOME ADDRESS	Mandalay, Myanmar.
PUBLICATION	[1] K. R. G. Operiano, P. M. Hmue, W. Pora, S. Pumrin, "American Alphabet Hand Sign Language Detection and Recognition Using Haar Cascades and Convolutional Neural Networks", 11th Regional Conference on Electrical and Electronic Engineering (RCEEE2018), Penan, Malaysia, 14-15 Aug 2018. [2] P. M. Hmue, S. Pumrin, "Image Enhancement and Quality Assessment Methods in Turbid Water: A Review Article", 4th International Conference on Consumer Electronics Asia 2019 (ICCE-Asia 2019), Bangkok, Thailand, 12-14 June 2019.