TRINITY COLLEGE DUBLIN
COLÁISTE NA TRÍONÓIDE
THE UNIVERSITY OF DUBLIN

Detection and Assessment of Damage in Underwater Marine Structures using Image Processing Based Techniques

Michael O'Byrne

A thesis submitted for the degree of Doctor of Philosophy to the University of Dublin, Trinity College.

September 2014

DECLARATION

I hereby declare that this thesis is entirely my own work and has not been submitted as an exercise for a degree at this or any other university. I hereby agree to permit the library to lend or copy this thesis upon request.

Dublin, September 2014

Michael O'Byrne

SUMMARY

The monitoring of underwater structures is often beset by limited access, exorbitant costs and inherent safety issues. This thesis focuses on developing and implementing automated image processing based Non-Destructive Testing (NDT) techniques to facilitate inspections. Adopting such techniques can vastly improve the condition of monitoring, reduce the operational complexities and partially offset the financial burden of regular inspections. To date, there has been very little work carried out on image based techniques for the purposes of detecting and quantifying the extent of structural damage particularly affecting the submerged part of marine structures. This research endeavours to bridge that knowledge-gap through development and performance evaluation of a series of advanced techniques and strategies. Underwater imaging is exposed to numerous challenges such as luminous complexities, poor visibility, light attenuation and backscatter which diminish the ability of the camera, and subsequent image processing algorithms, to effectively identify and quantify instances of damage. This research addresses the deleterious effects of these environmental conditions and phenomena, and through a repository driven approach, maps the impact that lighting, turbidity, and surface type have on the performance of developed techniques.

The developed techniques include a crack detection algorithm, colour and texture based damage detection algorithms, and a 3D shape recovery algorithm. The crack detection algorithm adopts a percolation based approach to automatically locate and quantify cracks in an efficient manner, removing the need for inspectors to manually undertake this tedious task. The damage detection algorithms employ advanced image segmentation methods to identify and quantify the severity of damage on the surface of infrastructural elements, based on textural information or colour information. Each algorithm is naturally suited to

different applications, depending largely on whether the damage form under consideration is more separable from the background based on colour or on texture. Both of these algorithms are validated on real world instances of infrastructural damage acquired from diving expeditions, as well as on above-water instances of damages. A stereoscopic based approach is employed for recovering 3D shape, which utilises a dual camera set-up to simultaneously photograph a specimen of interest from slightly different viewpoints. As part of this approach, an efficient pyramidal loopy Belief Propagation (BP) Markov Random Field (MRF) stereo correspondence algorithm is developed. This algorithm is applied to a submerged pile covered with artificial marine growth in a large-scale testing facility in Boulogne, France, and is subsequently validated on a real word structure in Cork Harbour, Ireland. Having accurate 3D shape information of submerged structural members is of great practical importance when analysing the forces exercised by the waves, winds and currents.

A protocol for the implementation of image processing techniques is established, which specifies the technical requirements for capturing imagery and outlines a set of best practice guidelines to ensure the acquired imagery is suited for quantitative analysis.

ACKNOWLEDGEMENTS

There are a number of people without whom this thesis could not have been written, and to whom I am greatly indebted.

First and foremost, I would like to express my sincere gratitude to my primary supervisor Dr. Bidisha Ghosh for her kindness, enthusiasm and patience. Dr. Ghosh was always available to me if I had any issues and her wealth of knowledge helped me at every stage of my research. Her contribution to my enjoyment and learning has been immeasurable and I could not have asked for a better supervisor for my PhD study.

I would like to give a very special thanks to my co-supervisor Dr. Vikram Pakrashi, from the Department of Civil and Environmental Engineering at University College Cork, for without his motivation and encouragement during my undergraduate degree I would most likely not have considered a postgraduate career path. It was under his tutelage that I was introduced to the richness of the applications of image processing in engineering problems. I doubt that I will ever be able to convey my appreciation fully, but I owe him my immense gratitude for his hard work, persistence, direction and his seemingly boundless wisdom on every subject.

I would like to thank Professor Franck Schoefs from the Department of Physics at University of Nantes for his expert insight and for always finding time to provide invaluable feedback. His suggestions helped shape this research and have led me to work on some diverse and exciting projects, for which I am extremely grateful.

I wish to thank all of the lab staff in the Department of Civil, Structural and Environmental Engineering in Trinity College Dublin. In particular to Dr. Kevin Ryan, who put in a huge amount of work in helping me set-up experiments and was always on hand to provide technical assistance. I would also like to put on record my appreciation to Mr. Mark Kavanagh, School of Natural Sciences, Trinity College Dublin, for generously lending me equipment.

Thanks to all of the postgraduates who have made my time here so enjoyable, especially to those with whom I have had the privilege to share an office and who have endured the highs and frustrations of PhD life beside me - it has been an absolute pleasure and a truly enriching experience.

I would like to thank my family who have been an unwavering source of support during the course of my PhD. In particular, to my mother Evelyn, for the myriad of ways in which, throughout my life, she has inspired me to find and realise my potential.

Finally, I recognise that this research would not have been possible without the financial assistance of the Irish Research Council for Science, Engineering and Technology (IRCSET) and CAPACITES/IXEAD society under the Enterprise Partnership Scheme (EPS), and I express my gratitude to those agencies.

DEDICATION

This thesis and the personal endeavour it represents is dedicated to my family for their unconditional love, encouragement and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

4DIS	4 Dimensional Input Space
ACFM	Alternating Current Field Measurements
ADA	Atmospheric Diving Suit
AE	Acoustic Emission
ASM	Angular Second Moment
BP	Belief Propagation
CDF	Cumulative Distribution Function
CWI	Custom-Weighted Iterative
DSLR	Digital Single Lens Reflex
FOS	First Order Statistics
FPS	Frames Per Second
GLCM	Gray Level Co-occurrence Matrix
HDR	High Dynamic Range
HE	Histogram Equalization
HSV	Hue-Saturation-Value
IMS	Infrastructure Management System
LED	Light Emitting Diode
MP	Mega-Pixel
MPI	Magnetic Particle Inspection
MRF	Markov Random Field
NCC	Normalised Cross Correlation
NDT	Non-Destructive Testing
NTU	Nephelometric Turbidity Unit
PSD	Power Spectral Density

RBF	Radial Basis Function
REMPS	Regionally Enhanced Multi-Phase Segmentation
RGB	Red-Green-Blue
ROC	Receiver Operating Characteristics
ROV	Remotely Operated Vehicle
SAD	Sum of Absolute Differences
SDR	Standard Dynamic Range
SfM	Structure from Motion
SfS	Structure from Shading
SHM	Structural Health Monitoring
SIFT	Scale Invariant Feature Transform
SSD	Sum of Squared Differences
SVM	Support Vector Machine
TSS	Total Suspended Solids
TTL	Through-The-Lens
ULTIR	Underwater Lighting and Turbidity Image Repository
US	Ultra-Sonic
ZNCC	Zero-mean Normalised Cross Correlation

Chapter 1

Introduction

1.1 Background

The deteriorating condition of infrastructure worldwide and the excessive costs required for reparatory work necessitate the invention of smart inspection strategies. Recent estimates suggest that $1.6 trillion dollars will be needed for the rehabilitation, replacement and maintenance of current infrastructure networks within the next 20 years for the United States alone (Adeli and Jiang, 2009). Structures in marine environment make up an important part of the infrastructure network. The harsh and corrosive conditions make these structures especially susceptible to rapid ageing and deterioration compared to inland structures, and as such, the need for efficient inspections is greatly felt. The optimal inspection strategy would be to inspect at the right location, right time, with the right tool, and at the lowest cost (Rouhan and Schoefs, 2003). In today's economic climate, cost is becoming an unavoidably compelling factor. An increasingly popular low-cost strategy involves the use of image processing based techniques as a means to strengthen conventional visual inspections by introducing a source of quantitative information.

Visual inspection are affected by the ability of inspectors to observe and objectively record details of defects, and are prone to considerations such as operator boredom, lapses in concentration, subjectivity, and fatigue, all of which contribute to increased variability and reduced accuracy (Agin, 1980; Estes and Frangopol, 2003; Komorowski and Forsyth,

2000). Furthermore, visual inspections almost always capture photographs to include in the inspection report to corroborate the inspector's comments; however, these photographs are rarely exploited to their fullest potential in either a qualitative or a quantitative fashion. Adopting effective image based techniques can provide accurate quantitative information with minimal human supervision to supplement visual inspection techniques and increase reliability.

Image processing methods have applications in many of areas of Structural Health Monitoring (SHM). Rytter's (1994) four level damage detection hierarchy is often referred to when discussing SHM objectives. The four levels are: (i) detection of the presence of damage, (ii) detection of the location of damage (iii) quantification of the severity of damage, and (iv) prediction of the remaining service life of the structure. The success at any given level usually depends on having previously achieved success at lower levels. Image based methods developed for crack detection and surface damage detection are usually concerned with the second and third levels, i.e. detecting the location of damage and quantifying shape information such as width or area. Having accurate knowledge of the location of damage is particularly useful when image processing methods are used in conjunction with other specialised NDT tools such as ultra-sonic (US) measurements, where the aim would be to identify appropriate zones on the structure for further in-depth scrutiny. 3D Shape recovery algorithms are restricted to the third level, i.e. quantification of the severity of damage. The fourth level (prediction of service life) is beyond the scope of this thesis, however, the quantitative nature of the data obtained from image analysis is important and naturally lends itself to numerous applications, including for developing new damage models, or strengthening existing ones, which are used to forecast the rate of propagation of damage as the structure continues to operate.

This research is worth tackling as there are considerable advantages of using image based methods as part of the inspection process. The extensive effort and expense associated with undertaking inspections warrants the development of sophisticated image based techniques that can fully exploit the available scene information. This should be accompanied by a thorough exploration of their on-site performance levels so that inspectors are equipped with tried and tested NDT methods. Finally, there is no agreed protocol of image collection and subsequent interpretation within the inspection framework (Phares et al., 2004). This research aims to fill that void by detailing a full and comprehensive set of best practice guidelines.

1.2 Thesis Outline

The thesis is organised in 6 chapters following this chapter. The majority of the work focuses on image based techniques applied in an underwater setting, but cases of above-water applications are also considered

Chapter 2 provides a literature overview summarising NDT techniques that are employed as part of the inspection process, with a focus on the inspection of marine structures. Particular attention is paid to image based methods such as crack detection techniques, surface damage detection techniques, and 3D shape recovery techniques. Gaps in the literature are highlighted, which this thesis endeavours to fill. The novel contributions of this thesis are outlined and it is shown how they constitute an improvement over current methods.

Chapter 3 presents a semi-automatic, enhanced texture segmentation approach to detect and classify surface damage on infrastructural elements and successfully applies them to a

range of images of surface damage. The approach involves statistical analysis of spatially neighbouring pixels in various colour spaces by defining a feature vector that includes measures related to pixel intensity values over a specified colour range and statistics derived from the Grey Level Co-occurrence Matrix. Parameter optimised non-linear Support Vector Machines are used to classify the feature vector. A Custom-Weighted Iterative model and a 4-Dimensional Feature Space model are introduced.

Chapter 4 presents techniques relating to: 1) crack detection, 2) surface damage detection, 3) 3D shape recovery using stereo-vision, and 4) video tracking. All of these techniques are based on colour information as opposed to textural information. The crack detection technique adopts a percolation based approach that offers greater efficiency over other percolation based methods, whilst still achieving good detection results, as the search space for cracks is confined to places where there is a sharp image intensity gradient as found by the Sobel operator. Furthermore, additional criteria are proposed for classifying potential cracks. A new surface damage detection technique, termed REMPS (Regionally Enhanced Multi-Phase Segmentation), is developed that integrates spatial and pixel relationships to identify, classify, and quantify the area of damaged regions to a high degree of accuracy. REMPS consists of identification, classification and enhancement phases. Combining each of these constituent phases in an effective manner creates a powerful and robust detection algorithm. To further improve the detection accuracy of REMPS, High Dynamic Range (HDR) imagery is examined. For the purpose of recovering 3D shape, an efficient pyramidal loopy Belief Propagation (BP) Markov Random Field (MRF) stereo correspondence algorithm, termed PaLPaBEL for short, is described. This algorithm is evaluated on synthetic and real world specimens. Finally, a video tracking technique is outlined, and is demonstrated on a vibrating suspension bridge spanning a river in Cork, Ireland, for the purpose of identifying the natural frequency.

Chapter 5 describes an open-source Underwater Lighting and Turbidity Image Repository (ULTIR) that features images of various damage forms, material types and shapes, which are photographed under controlled lighting and turbidity levels. This rich resource is intended to give inspectors and researchers a platform to efficiently gauge the performance of image based methods under realistic operating conditions and on relevant specimens. It is of great practical importance for inspectors to have access to this information so that they can, first and foremost, rationalise the use of image based methods as part of the inspection regime; secondly, create conditions that are conducive to good performance; and finally, choose an appropriate image analysis algorithm that maximises detection accuracy.

Chapter 6 introduces a protocol that describes a set of best practice guidelines. This protocol addresses each stage of the imaging pipeline from equipment set-up right through to post-processing and quality assurance (performance measurement). Such a protocol is needed to maintain the long-term integrity of an Infrastructure Management System (IMS), which is dependent on the quality and consistency of the input information. The protocol forms the basis for a set of tests where image techniques are applied within the context of a large scale experiment and to a real world structure. As part of a large scale multi-disciplinary experiment to track the extent of marine growth, the stereo based 3D shape recovery technique, PaLPaBEL, is applied to a cylinder covered with artificial marine growth. PaLPaBEL is then performed in a real world setting on the submerged part of a pier located in Cork harbour.

Chapter 7 provides a summary of the research work carried out in the thesis and identifies the primary contributions. The major findings and their impacts are appropriately

discussed. Future research paths leading on from this thesis are suggested and the feasibilities of undertaking such work are discussed.

The structure and flow of this thesis can be summarised as follows: Chapter 2 reviews existing image processing techniques and practices in the domain of SHM, Chapter 3 and Chapter 4 develop image algorithms that improve upon the current state, Chapter 5 collects data for the purpose of calibrating and evaluating these developed algorithms, Chapter 6 completes the validation of all algorithms on real world structures in adherence with a proposed protocol that sets out best practice guidelines for underwater image acquisition. Chapter 7 reflects on the results of this thesis, draws conclusions, and discusses future work.

Chapter 2

Literature Review

This chapter provides an overview of the literature related to the detection and assessment of damage in underwater marine structures using image processing based techniques. The following sections are concerned with: (i) the inspection process, (ii) the main forms of NDT techniques recently developed for topside applications (i.e. NDT techniques intended to be used in dry conditions), (iii) underwater NDT techniques, (iv) image processing techniques with a focus on techniques that have been applied in the domain of NDT, and finally, (v) attempts at characterising the performance of image based methods.

2.1 Inspection of Marine Structures

The deteriorating condition of marine structures worldwide and the excessive costs required for reparatory work places an emphasis on devising comprehensive strategies for the periodic inspection and monitoring of these structures (Gangone et al., 2011; Schoefs et al., 2011). Having a comprehensive strategy is especially relevant nowadays given the increased loads and ever challenging environments that these structures are being installed in (Cusson et al., 2011; Xia et al., 2011).

The information obtained from inspections is useful when analysed in the context of an Infrastructure Management System (IMS), which allows for improved estimates of the remaining service life (Rong-Yau et al., 2010). A reliable IMS can help the decision

makers to optimise intervention plans and enable them to make informed decisions regarding the future course of action that would maximise the potential of their investments. This aspect has attracted a growing interest in recent years as the importance of life cycle optimisation and the related financial benefits continue to be recognised (Sarma and Adeli, 1998; Sirca Jr and Adeli, 2005). For a well calibrated IMS, it is important that the input information from an NDT technique is accurate, which is often hard to achieve when the inspection is being performed in a difficult environment.

2.2 Underwater Inspections

Underwater inspections introduce new challenges for inspectors, which are thoroughly summarised by Busby (1979) and Ramos (1992). These authors detail the legal requirements of undertaking underwater inspections, the modes of inspection, and the available damage diagnostic tools.

The different modes of carrying out inspections include using divers, remotely operated vehicles (ROVs), manned submersibles, and Atmospheric Diving Suits (ADS) (Goldberg, 1996). Diver and ROV based inspections are the dominant approaches. Divers offer significant dexterity, are capable of covering a wide area, descend to depths of up to 50 m, and can facilitate inspecting at a number of scales, from a detailed close-up assessments with the help of specialised equipment if necessary, to broader-scale visual assessments. The use of any specialised equipment will require additional training in its operation.

ROVs may be equipped with numerous sensors. Typically, they carry at least one or more camera/video systems. They can be deployed to much greater depths and for longer periods than divers. They are however considerably more expensive, inflexible, and

vulnerable to failure in comparison to a diver based approach. Consequently, their safe operation requires great attention to detail. Operators must be trained in piloting the ROV, which may require formal qualifications. This research embraces a diver based approach as the emphasis lies in creating an economical and easily deployable system, which only utilises readily available equipment and requires minimal training.

2.3 NDT Techniques

There exists a broad range of NDT techniques available to inspectors capable of detecting various forms of damage. In many situations, NDT techniques offer the only practical way of detecting damage. Before listing available techniques, it may be useful to specify what damage is as the definition varies somewhat in the literature. This thesis adopts the definition of damage as given by Doebling and Farrar (1998), which states that damage is defined as "changes, either intentional or unintentional, to the material and/or geometric properties of structural systems, including changes to the boundary conditions and system connectivity, which adversely affect the current or future performance of that system". A highly attractive feature of NDT techniques is that they can evaluate a material without introducing further damage.

NDT techniques can be partitioned into two categories: non-visual and visual based techniques. There is a wealth of examples from each category in the literature related to topside inspections, i.e. inspections carried out in dry conditions. Amongst the non-visual NDT techniques are ultrasonic scanning (Iyer et al., 2005), surface wave simulation (Kim and Kwak, 2008), acoustic emission techniques (Li et al., 2011; Sohn et al., 2008), ground penetrating radar (Belli et al., 2008), eddy current testing (Yusa et al., 2006), as well as a lot of recent interest in vibration based techniques (Adewuyi and Wu, 2011; Cruz and

Salgado, 2009; Jafarkhani and Masri, 2011; Li et al., 2011; Talebinejad et al., 2011). Visual based techniques offer a good way of detecting anomalies such as corrosion, impact damage and surface-breaking defects. Visual based techniques include image processing techniques, which are outlined in greater detail in Section 2.6. There are several specialist visual techniques such as remote visual inspection (Nugent and Pellegrino, 1991), yet the most common visual based approach is standard visual inspections carried out by trained engineers.

2.4 Underwater NDT techniques

Underwater NDT techniques are adapted from topside NDT techniques. The modifications usually entail waterproofing and tuning the instruments to suit the underwater environment. The most common underwater NDT techniques are: (i) Visual, (ii) Electromagnetic, (iii) Ultrasonic (US), (iv) Radiography, (v) Acoustic Emission (AE), and (vi) Vibration Analysis based methods (Berger et al., 1983). The capabilities and the drawbacks of each of these techniques are explained in this section.

2.4.1 Visual Inspections

Visual inspections are the most popular underwater inspection approach (Watson, 1992). They are the most basic and straight-forward inspection method. The main drawbacks associated with visual inspections are the requirement of good visibility, and the lack of penetration below the surface. Additionally, the presence of bio-fouling can inhibit visual inspections by masking the underlying structure, and must be cleaned beforehand.

The shortcomings of visual inspections when carried out by trained engineers/divers in isolation - without any supplementary tools - have been well observed in the literature. Agin (1980) and Komorowski and Forsyth (2000) found that the quality of the assessment largely depends on the ability of the inspectors to observe and objectively record details of defects. Factors such as operator boredom, lapses in concentration, subjectivity, and fatigue contribute to the variability and reduced accuracy of visual inspections. Gallwey and Drury (1986) found that incorporating image processing techniques into the visual inspection regime offered far greater reproducible and measurable performance over conventional visual inspection techniques.

2.4.2 Electromagnetic Methods

Electromagnetic methods are the second most popular method after visual inspections (Rizzo, 2013). Electromagnetic methods include Magnetic Particle Inspection (MPI) (Groves and Connell, 1985), eddy currents (Goldberg, 1998; Yusa et al., 2006), electrical potential techniques (Dover and Bond, 1986), Alternating Current Field Measurement (ACFM) (Wei et al., 2013) and magnetic flux leakage techniques (Butcher et al., 2013). These techniques provide information about surface and near-surface defects, and about the effectiveness of cathodic protection systems for metallic structures. MPI methods are commonly used underwater to detect surface and near-surface flaws in ferromagnetic materials. They are capable of defining the true length of discontinuities. MPI methods can only be applied to bare metal surfaces and, since the process is time-consuming, they are generally only applied in small areas that have been identified as being susceptible to cracking (Lindgren et al., 2002). ACFM techniques are non-contact electromagnetic techniques that offer the capability to detect and size surface-breaking cracks in a range of different materials and through coatings of varying thickness. ACFM methods can also be

used for the purpose of evaluating marine growth thickness. This task requires a diver to push a probe against the marine growth for a few seconds. The distance between the probe and the face of the underlying metal structure is then determined. While this approach is accurate and straightforward, the output is a series of spot measurements which is not as intuitive to interpret and visualise as image based approaches.

2.4.3 Ultrasonic Methods

Ultrasonic methods can be applied to a wide range of materials and offer the capability to detect both external and internal defects, thickness measurements, and weld examinations (Rose et al., 1983). US methods are based on the propagation of ultrasonic waves generated by one or more probes through the structure. They are essentially local methods, and therefore, would require a lot of time if applied to large structures.

2.4.4 Radiographic Methods

Radiographic methods can detect internal flaws in any material (Correa et al., 2009). Radiographic methods have the advantage of being accurate, the output is easy to interpret and visualise, and they have good resolution. However, the major drawback of radiographic methods is that there are inherent safety concerns when deployed in an underwater environment due to the emitted radiation.

2.4.5 Acoustic Emission

Acoustic Emission tools can be used to monitor the progression of damage, estimate the extent of corrosion in reinforced concrete structures and detect cracks (Blitz and Simpson,

1996; Thaulow and Berge, 1984). Acoustic emission methods are attractive for monitoring large structures such as offshore platforms. The continuous wave and current loads that act on these structures provide an eternal stimulus that causes stress waves to propagate throughout the structure. Detection and analysis of the acoustic emission signals can reveal the location of any defects, as well as offering a qualitative assessment of their severity. However, the defect may remain undetected if there is not sufficient loading to produce an acoustic event.

2.4.6 Vibration Analysis

Finally, there has been some recent interest in vibration based techniques such as (Hillis and Courtney, 2011). Vibration analysis techniques rely on the natural vibration modes, which in a marine environment are typically excited by wind and waves, and are dependent on the characteristics of the structure. Like acoustic emission methods, vibration analysis techniques and are well suited for monitoring large structures, however, a crucial drawback is that they cannot determine the presence of damage until it becomes large enough to affect the natural vibration frequency of the structure. Moreover, environmental factors such as temperature must be accounted for and the precise location of a broken member cannot be determined.

2.5 Image processing based techniques

A convenient way of categorising damage is by dimensionality. Cracks, for instance, are characterised by their fine-structured, almost linear, shape so are therefore considered as a one-dimensional damage form. Visible damages such as corrosion or spalling, which typically appear in larger areas on the surface of infrastructural elements, are treated as a

two-dimensional damage form, while damages such as marine growth, where the objective is to retrieve shape information, is categorised as a three-dimensional problem. This thesis lightly considers the fourth dimension (i.e. three of dimensions of space, and one of time) through the use of video analysis.

Image processing techniques in each dimensionality category are naturally very different, but even techniques within the same category follow vastly different methodologies. The subsequent sub-sections provide an overview of the literature related to image processing techniques in each category.

2.5.1 Cracks

Given the pervasive nature of cracks, attempts to automate the crack detection process have received a lot of attention. This has led to the development of a number of techniques such as wavelets (Khanfar et al., 2003), neural network approaches (Choudhary and Dey, 2012), statistical filters (Sinha and Fieguth, 2006) and percolation based methods (Yamaguchi and Hashimoto, 2008). Zou et al. (2012) have addressed the issue of challenging scene conditions by creating a robust three step algorithm involving shadow removal, tensor voting to create a crack probability map, followed by graph modelling. Wang and Huang (2010) conducted a comparison of crack detection methods and identified percolation-based methods as being particularly suited to detecting ambiguous cracks. However, they also recognised that conventional percolation based methods are computationally expensive given that they operate on a local basis at each pixel location in the image. The crack detection method developed in this thesis is a percolation based method that offers improved efficiency.

2.5.2 Surface Damage

Image processing based damage detection algorithms use either colour information, texture information, or a combination of both, to isolate similar regions in an image. The effectiveness of colour based segmentation algorithms and texture based segmentation algorithms will vary according to the surface and damage type under consideration as certain damages are more separable from the undamaged surface based on either their colour or texture attributes. Texture is an innate property of surfaces (Haralick et al., 1973). There are numerous image processing based techniques which attempt to characterise texture; wavelet analysis (Lu et al., 1997), Laws' texture energy (Choi et al., 2011), First Order Statistics (FOS) (Gill, 1999), GLCM (Gadelmawla, 2004) or hybrid methods (O'Byrne et al., 2013). The main forms of surface damage encountered on ageing infrastructural elements (corrosion, leaching, etc.) are often characterised to a greater extent by the change in colour from the undamaged surface than a change in texture. With this in mind, texture based segmentation methods for isolating damaged surfaces in the field of NDT can be classified as suitable for specific applications where the damaged regions have a noticeably different texture than the surroundings.

Colour based segmentation algorithms may be grouped into four major categories: thresholding, edge detection using gradient information, region growing, and hybrid methods (Abdel-Qader et al., 2008). Existing literature contains a variety of these segmentation methods applied in the domain of NDT. Many of these methods are designed for a particular application such as the detection of weld defects (Anand and Kumar, 2009; Kasban et al., 2011; Vilar et al., 2009; Yazid et al., 2011) or pipe deterioration (Zheng et al., 2012), and/or for particular image sources such as optical (Yazid et al., 2011), thermal (Abdel-Qader et al., 2008; Heriansyah and Abu-Bakar, 2009; Huang and Wu, 2010),

ultrasonic (D'Orazio et al., 2008; Molero et al., 2012) and radiography (Anand and Kumar, 2009; Kasban et al., 2011; Vilar et al., 2009; Yazid et al., 2011). As such, while these techniques may be effective for their designated purposes, they are understandably unlikely to perform well when applied to richly detailed, high-resolution optical images of a broad range of surface types and damage forms in complex natural scenes. There exist very few studies that have developed powerful image processing techniques to cater for the detection of damage in challenging circumstances. Thus, the emphasis lies in the development of a new technique that can characterise features of interest in natural scenes with credibility (Lu et al., 1997; Naccari et al., 2005).

2.5.3 3D Shape Recovery Algorithms

In recent decades, considerable advances have been made in developing underwater 3D shape recovery techniques involving coded structured light (Bruno et al., 2011; Roman et al., 2010), photogrammetry (Cocito et al., 2003) and stereo photography (van Rooij and Videler, 1996). These techniques originate mainly from the fields of marine biology and archaeology. No literature has been found of a shape recovery technique being utilised as an underwater NDT technique. Coded structured light approaches use a light projector to project a light pattern onto an object. A camera then captures how the pattern interacts with the shape. Given the reliance on a light source to encode depth information, the success of this approach is more susceptible to the effects of absorption and scattering that occur in an underwater environment. Additionally, the projected light interferes with the scene and may mask an object's natural colour, which could be an important factor when analysing the imagery for other purposes (e.g. for identifying marine growth species).

Photogrammetry techniques are easily implemented and have the potential to achieve good results; however they are prone to breakdown if the tracked feature points are lost, which can cause the camera to lose its bearing with respect to the subject. Stereo photography, which is also known as stereo vision, is principally based on finding matching pixels in two given input images taken from slightly horizontally separated points. There are a variety of stereo matching algorithms in existing literature. An extensive review of modern stereo correspondence is provided by Olofsson (2010). A stereo based approach is developed in this thesis as it is felt that the importance of inspections, coupled with the time and effort involved, should warrant a more dependable system, which lowers the risk of false matches by enforcing certain constraints and allows useful depth information to be extracted from a single stereo image pair independent of other imagery. Additionally, the stereo vision can be integrated into a photogrammetry framework in order to recover the full 3D profile of the subject.

2.5.4 Video Analysis

The use of video cameras in conjunction with image processing techniques makes it possible to access spatial and temporal dimensions. Video cameras are an attractive and user-friendly option compared to other ways of recording spatial and temporal dimensions such as through a complicated network of accelerometers or similar instruments. One application in SHM where this is of relevance concerns vibrating structures. Wahbah used a video camera and two LEDs to measure displacements of a particular area of a cable-suspension bridge. Gehle and Masri (1998) used video analysis to track the motion of a cable as part of the control and monitoring of smart structures. Olaszek used multiple cameras to measure displacements of a particular location in a highway bridge with the

span length of 33 m, and achieved measured displacements with an accuracy of 0.1–1.0 mm for bridges with a fundamental frequency up to 5 Hz.

2.6 Performance Characterisation

For a well calibrated IMS, it is important that the input information from an NDT technique is accurate and comprehensive. The measure of the onsite performance of a NDT tool remains a pertinent question in the majority of cases (Schoefs et al., 2012). The most suitable NDT for a given application will not only depend on the damage to be detected, but will also depend on the operating environment. Applying NDT techniques in an underwater environment typically results in a decline in performance compared to topside application in ideal conditions. This is similarly true for image based techniques.

While there has been a relatively great deal of attention devoted to developing image based techniques, very little follow-up work has been done to map their performances under varying operating conditions. The limited work includes that done by Bianco et al. (2013), who considered the effects of lighting and total suspended solids (TSS) on a single stereo vision technique, and by O'Byrne et al. (2013), who investigated the effects of turbidity and lighting on a surface damage detection technique. There is thus scope to develop a unified framework that offers the capability to assess the performance of any image processing technique in an underwater setting. This would enable inspectors to decide on the feasibility of adopting an image processing based approach prior to inspection and to isolate conditions that are conducive to good performance.

2.7 Scope for Integrating Image based Techniques into Inspections

Image processing techniques are a mainstay in many other fields such as in bio-medical engineering or traffic engineering. However, there exists significant scope for further development in the domain of SHM, as currently, the full potential of cameras is not being fully exploited. Cameras are a convenient and versatile tool capable of making visual data a part of quantitative assessment for a wide array of applications in SHM. In order for inspectors to properly capitalise on the power of cameras and integrate them into the inspection framework, they must be aware of each camera's limits, know how to acquire photographs effectively, and have access to efficient image processing techniques.

This thesis aims to address each of these issues in a number of ways. Firstly, new techniques are developed to address the unique problems encountered in an underwater infrastructural setting, and existing image processing techniques are customised to suit SHM applications. Secondly, a comprehensive image repository is created and an image acquisition protocol is established so that inspectors are equipped with the necessary knowledge to carry out camera based inspections.

Chapter 3

Texture Analysis based Damage Detection of Ageing Infrastructural Elements

3.1 Introduction

There are a wide range of damage forms on the surface of infrastructural elements that are similar in colour to the non-damaged surface yet are still easily identifiable by human observers as they are markedly different in terms of their textural composition. Such examples include spalling, honeycombing, erosion etc. Detecting these damage forms using computational methods presents new challenges. Conventional image based methods that rely on colour information will not produce good detection results. Instead, texture analysis based methods provide a more appropriate solution. This chapter presents a novel, semi-automatic texture segmentation approach to detect and classify surface damage on infrastructure elements. Physical properties of the identified damage, such as area, may be extracted with knowledge of a real world scale. Possible ways of determining the scale factor include practical approaches such as placing an object of known dimensions alongside the damaged region or by using a stereo system to obtain 3D information. A calibrated stereo-rig, which consists of two cameras viewing the scene from two slightly different vantage points, is capable of providing a fully scaled metric scene reconstruction at the expense of greater algorithmic complexity and the requirement of an additional camera. Placing an object of known dimensions in the scene and reading the corresponding pixel dimensions in the acquired imagery is a straightforward solution.

The texture analysis approach involves generating feature vectors at each point in the image, which are populated with statistics derived directly from the image as well as from a Grey Level Co-occurrence Matrix (GLCM). Parameter optimised non-linear Support Vector Machines (SVMs) are used to classify these feature vectors. This chapter introduces two new SVM classification models; a Custom-Weighted Iterative (CWI) model and a 4-Dimensional Input Space (4DIS) model in which the feature vectors are mapped to a four dimensional input space. These models differ in terms of performance accuracy and computational time. Parameter optimisation is based on a Receiver Operating Characteristic (ROC) framework. The proposed technique is evaluated in various colour spaces (Red-Green-Blue (RGB), Hue-Saturation-Value (HSV) and $L^*a^*b^*$) in order to determine the best segmentation space. The technique is successfully applied to a range of images that feature a variety of damage forms.

3.2 Methodology

The texture analysis based damage detection algorithm that has been proposed in this chapter involves two main stages. The first stage computes a texture characteristics map from an input colour image, and the second stage subsequently classifies pixels as being either damaged or undamaged using SVMs. These stages are explained in Section 3.2.1 and Section 3.2.2 respectively. The overall methodology for the proposed technique is illustrated in the flowchart in Figure 3.1.

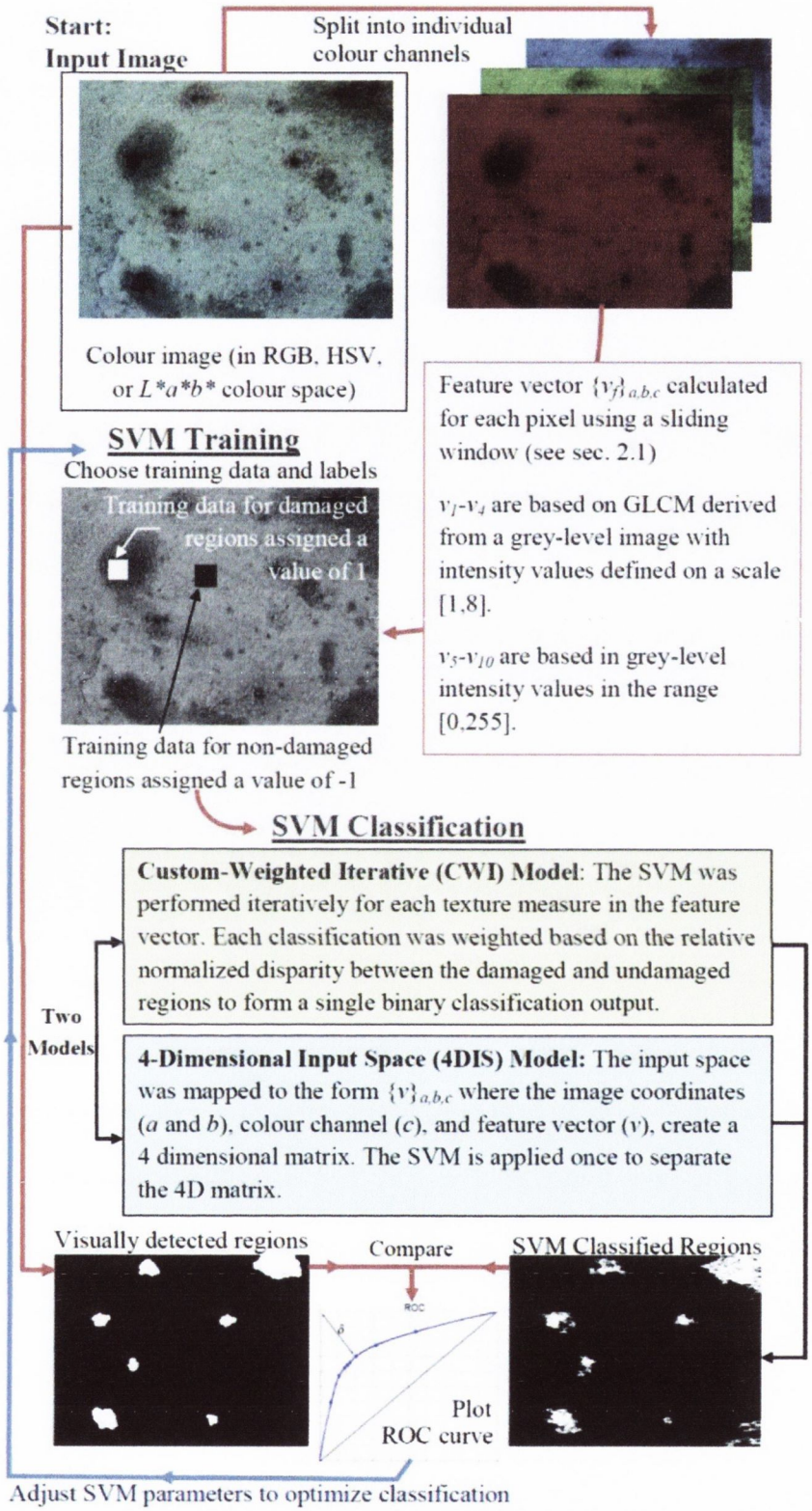


Figure 3.1 Methodology Flow Chart.

3.2.1 Texture Characteristics Map

A texture characteristics feature vector $\{v_f\}_{a,b,c}$ has to be generated for each pixel within the original image, I, for each colour channel, c, where f indicates the index of the vector element and (a,b) indicates the spatial coordinates of the pixel. The first four elements of $\{v_f\}_{a,b,c}$ are obtained by computing statistics derived from a GLCM. These statistics are: 1) Angular Second Moment (ASM), 2) homogeneity, 3) contrast, and 4) correlation. The GLCM is primarily calculated for grey images yet may be readily extended to individual colour channels. The remaining six first order texture features are based on measures calculated from the original pixel values mapped over a range of [0,255]. This range corresponds to 8-bit greyscale images, which are capable of representing 256 (2^8) different intensity levels (or shades of grey). A value of 0 represents black, while a value of 255 represents white. These features are Shannon entropy, mean, variance, range, skewness and kurtosis.

The feature vector for each pixel is calculated separately for each colour channel and can be combined together to form a 4 dimensional array. The feature vector is generated for each pixel using a sliding window, *SW*, that moves throughout the image and provides the basis for the GLCM statistics and the distributions used for calculating descriptive statistics and Shannon entropy. The window starts at the top left-hand corner of the image and moves horizontally in steps of one pixel until it reaches the end of a row, at which point it progresses onto the leftmost point in the next row. The centre is indicated as (a,b) and the size of the window (N-pixel x N-pixel) is optimised for best performance. A trial and error approach is used to determine the optimal size. This optimisation step may be worthwhile if large batches of images featuring similar damaged surfaces are being processed, however, it was experimentally found that the classification accuracy of the technique was

not overly sensitive to the window size. An increase in the size of *SW* is accompanied by a marginal increase in the overall computational time. In this chapter, a nominal window size of 10 x 10 square pixels was used.

3.2.1.1 GLCM Features

The process in which the GLCM is created is illustrated in Figure 3.2.

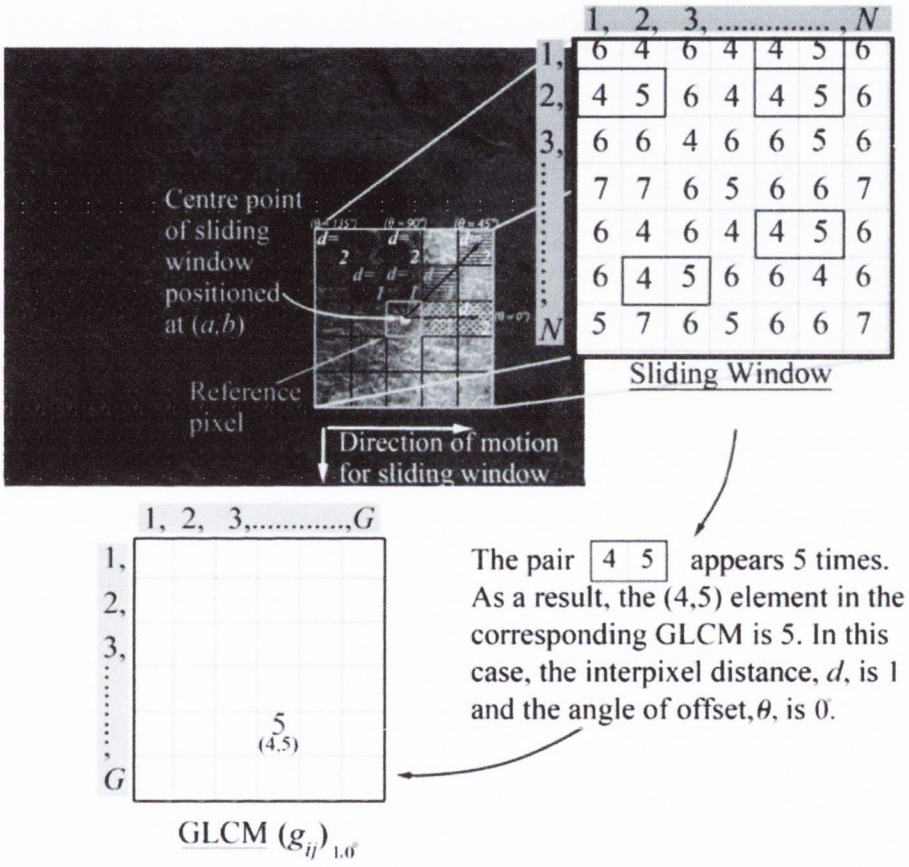


Figure 3.2 Overview of the GLCM process.

The GLCM is a matrix of frequency values of paired combinations of pixel intensities as they appear in certain specific spatial arrangements within an image or sub-image. The GLCM for each pixel is generated through a sub-image that is a sliding window, *SW*, centred on the pixel. Combinations of various pixel pairs within *SW* were counted and the

resulting total was assigned to the g_{ij}, in the GLCM which corresponds to the spatial arrangement of the pixel pairs being summated. The spatial indices i and j of the GLCM match the grey level in the reference pixel and the destination pixel respectively. The spatial arrangement of the reference pixel and destination pixel in relation to each other in SW are governed by two parameters; the interpixel distance, d, and the angle of offset, θ. The grey levels are defined using integer values between 1 and G. In this chapter, the grey levels are defined on a scale of 1 – 8 (G =8) instead of a larger scale such as [0,255]. Quantising in this manner increases computational parsimony at the expense of making the GLCM less sensitive to minute fluctuations in pixel intensity values within the sliding window. Despite this reduced sensitivity, the discriminatory capabilities of the GLCM remain largely unperturbed as perceivable changes in intensity values between neighbouring pixels continue to be taken into account, thus creating condensed yet descriptive matrices.

An illustrated example of the creation process for a GLCM is presented in Figure 3.2. In this example, the number of occurrences of pixels with a quantised grey level of 4 and 5 appearing horizontally alongside each other in the sliding window (d = 1 and θ = 0°) are computed. The number of occurrences of this pair is then assigned to the (4,5) element in the GLCM corresponding to the chosen value of d and θ. It was experimentally found that paired combinations of intensity values of pixels that are spatially neighbouring tend to be more relevant than combinations that involve spatially distant pixels. With this in mind, a value of 1 was chosen for d to ensure a certain level of spatial proximity. The angle along which the interpixel distance is counted is defined as the angles of offset. Four angles for the offset were chosen: θ = 0°, θ = 45°, θ = 90°, θ = 135°. So, this generated a set of 4 GLCMs (d =1; θ = 0°, 45°, 90°, 135°) for each colour channel at each pixel location.

The GLCM for each pixel is populated as:

$$\left(g_{ij}\right)_{d,\theta} = \sum_{u=1}^{N}\sum_{z=1}^{N} A \quad \text{where } A = \begin{cases} 1 & \text{if } s_{uz} = i \text{ and } s_{uz}^{d,\theta} = j \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

where s_{uz} is the pixel intensity expressed in quantised grey levels for the reference pixel located at row u and column z within the sliding window; $s_{uz}^{d,\theta}$, is the pixel intensity expressed in quantised grey levels for the destination pixel located at an interpixel distance d along an angle θ from the reference pixel. The GLCMs are normalised as per:

$$p(i,j)_{d,\theta} = \frac{\left(g_{ij}\right)_{d,\theta}}{N(N-1)} \tag{3.2}$$

Each texture feature determined form the GLCM loosely relates back to some physical textural property of the photographed surface (Baraldi and Parmiggian, 1995). This can be seen in Figure 3.3 and Figure 3.4 which show how the GLCM texture features vary as the textural properties of the illustrated surface also vary. The four texture features, and their physical meaning, as described as follows:

Angular Second Moment (ASM) represents the uniformity of distribution of grey level in the image.

$$(v_{f=1})_{d,\theta} = \sum_{i=1}^{G}\sum_{j=1}^{G}\left(p(i,j)_{d,\theta}\right)^2 \tag{3.3}$$

v_1 ranges from $1/G^2$ to 1. A high value of 1 indicates a uniform image, as may be expected of a smooth painted surface that lacks tonal variation A low ASM value, on the other hand,

may be expected of natural surfaces, which typically have a higher degree of tonal variation.

Homogeneity gives a measure of the similarity of grey levels in the image.

$$(v_{f=2})_{d,\theta} = \sum_{i=1}^{G}\sum_{j=1}^{G} m \cdot p(i,j)_{d,\theta} \quad \text{where } m = |i - j| \tag{3.4}$$

v_2 ranges from 0 to G-1. A value of 0 indicates a strong similarity of grey levels in the image, as may be expected of a surface that has a constant or periodic form.

Contrast is a measure of the local variations present in an image. The contrast will be high if there is a high amount of variation.

$$(v_{f=3})_{d,\theta} = \sum_{m=0}^{G-1} m^2 \left\{ \sum_{i=1}^{G}\sum_{j=1}^{G} p(i,j)_{d,\theta} \right\} \tag{3.5}$$

v_3 ranges from 0 to $(G\text{-}1)^2$. A value of 0 indicates a uniform image, as may be expected of a smooth man-made material or a painted surface.

Correlation is a measure of the grey level linear-dependencies in an image. Correlation will be high if an image contains a considerable amount of linear structure, which is often the case with man-made objects.

$$(v_{f=4})_{d,\theta} = \sum_{i=1}^{G}\sum_{j=1}^{G} \frac{(ij)p(i,j)_{d,\theta} - \bar{\mu}_1\bar{\mu}_2}{\sigma_1\sigma_2} \tag{3.6}$$

where $\overline{\mu}_1$, $\overline{\mu}_2$, σ_1 and σ_2 are the means and standard deviations of the marginal probability matrices, P_1 and P_2, obtained by summing the rows and columns of $p(i,j)_{d,\theta}$ respectively. v_4 ranges from -1 to 1. A value of 1 indicates a perfectly positively correlated image. The denominator in Equation 3.6 will be equal to zero in the case of a perfectly uniform image. This will result in the value of correlation being undefined, however, these undefined values are ignored by the SVM classifier and their influence on the classification accuracy is negligible as these cases tend to appear infrequently, in particular, when working with images of natural scenes.

3.2.1.2 Descriptive Statistics and Shannon Entropy

The feature vector was further populated by considering five descriptive statistics of the pixel intensity values, along with Shannon entropy. These six features were derived for each pixel using the same sliding window approach employed to calculate the GLCM features. Unlike the GLCM approach, the intensity values used in the distribution adopted the scale [0,255] for several reasons. Firstly, the nature of these First Order Statistics (FOS) generated directly from the intensity values differed from that of the GLCM statistics as it was the magnitude of the intensity values that was considered and not their frequency of occurrence. As such, it was more important for the intensity values to contain as much information as possible which required them to be accurately and precisely defined. Having a bigger sample space provided more sensitive information for characterising texture. Conversely, the GLCM statistics produced more meaningful results by having similar intensity values grouped together as separately counting perceptually close values may understate their prominence in the sliding window. Secondly, the number of grey-levels employed in the GLCM generation stage directly affected the size of the GLCM, which in turn affected the computational time of the algorithm. The intermediate

GLCM generation stage already accounted for a significant portion of the algorithm time so for this reason it was desirable to keep the size of the GLCM to a minimum. Employing quantised intensity values for the descriptive statistics and Shannon entropy on the other hand resulted in no perceivable benefits in terms of increased computational efficiency.

A key point to note is that the range of intensity values differed for various colour channels. To ensure equality and compatibility, a standardisation procedure was employed which linearly scaled the original pixel intensity value $I_{a,b,c}$ in each plane to a new pixel intensity, $I'_{a,b,c}$, such that it is always in the range [0, 255] as per:

$$I'_{a,b,c} = 255 \times \frac{I_{a,b,c} - \min(I_{a,b,c})}{\max(I_{a,b,c}) - \min(I_{a,b,c})} \tag{3.7}$$

This standardisation procedure is necessary as the customary range of values in some colour channels would be less conducive to statistical analysis. For instance, the typical range for the a^* and b^* channels in the $L^*a^*b^*$ colour space is [-128, 128]. Proceeding with this range would lead to the Shannon entropy to produce meaningless and undefined results for all distributions that have values in the bottom half of the range, [-128,0]. Using scaled pixel intensity values in the sliding window, denoted by $\tilde{S}_{uz}$, avoids this problem.

As with the GLCM based statistics, each of the FOS in the feature vector describe some aspect of the textural composition in a sliding window. The meaning and contribution of each statistic is discussed. Shannon entropy, v_5, is a statistical measure of the uncertainty associated with a random variable.

$$(v_{f=5}) = -\sum_{u=1}^{N}\sum_{z=1}^{N} \tilde{s}_{uz} \log_2 \tilde{s}_{uz} \tag{3.8}$$

v_5 ranges from -(N^2.max($\tilde{S}_{uz}$).$\log_2$(max($\tilde{S}_{uz}$))) to infinite, which for a pixel intensity range of [0,255] becomes [-2039.N^2,∞].

Mean gives the arithmetic average of the intensity values in a window.

$$(v_{f=6}) = \frac{1}{N^2}\sum_{u=1}^{N}\sum_{z=1}^{N} \tilde{s}_{uz} \tag{3.9}$$

v_6 can range from the minimum value of $\tilde{S}_{uz}$, 0, to maximum value of $\tilde{S}_{uz}$, 255.

Variance is a measure of how far a set of numbers is spread out from the mean.

$$(v_{f=7}) = \frac{1}{N^2}\sum_{u=1}^{N}\sum_{z=1}^{N} (\tilde{s}_{uz} - v_6)^2 \tag{3.10}$$

v_7 ranges from 0 to $\frac{\left(\max(\tilde{s}_{uz}) - \min(\tilde{s}_{uz})\right)^2}{4}$, which equates to 1.625x10^4 for the [0,255] range.

Range gives the difference between the maximum and minimum intensity values in the distribution:

$$(v_{f=8}) = \max(\tilde{s}_{uz}) - \min(\tilde{s}_{uz}) \quad \forall(u,z) \tag{3.11}$$

Skewness is a measure of the asymmetry of the data around the sample mean.

An estimate for the skewness is:

$$(v_{f=9}) = \frac{1}{v_7^{3/2}} \sum_{u=1}^{N} \sum_{z=1}^{N} (\tilde{s}_{uz} - v_6)^3 \tag{3.12}$$

v_9 ranges from -∞ to ∞.

Kurtosis is a measure of the peakedness of a distribution. A positive value for kurtosis indicates that the distribution has a greater peakedness than that predicted by a normal distribution, while a negative value indicates that the distribution is less peaked than predicted by a normal distribution. An estimate for the kurtosis is given by:

$$(v_{f=10}) = \frac{1}{v_7^2} \sum_{u=1}^{N} \sum_{z=1}^{N} (\tilde{s}_{uz} - v_6)^4 - 3 \tag{3.13}$$

v_{10} ranges from -2 to ∞.

As with the GLCM statistics, undefined values, or infinite values, can result for certain descriptive statistics such as skewness and kurtosis when the intensity values in the window are perfectly uniform, i.e. when the standard deviation is equal to zero. The value of entropy may also be undefined in the case of pixel intensities having a value of zero in a given distribution. These undefined values are ignored by the SVM classifier. Their influence on the classification accuracy is negligible however as not only do the undefined values tend to appear infrequently, but by having a large feature vector containing a greater number of correctly defined texture measures, their effect is vastly diminished. Moreover,

since the sensitivity of each texture measure varies according to the surface type and damage form, having a large feature vector is useful as it ensures that the influence of any texture measure that is ineffective at differentiating between damaged and undamaged regions is offset by other texture features that have a higher sensitivity to regions of contrasting texture. Figure 3.4 plots the GLCM and descriptive statistics along the profile line shown in Figure 3.3, which passes over damaged and undamaged regions.

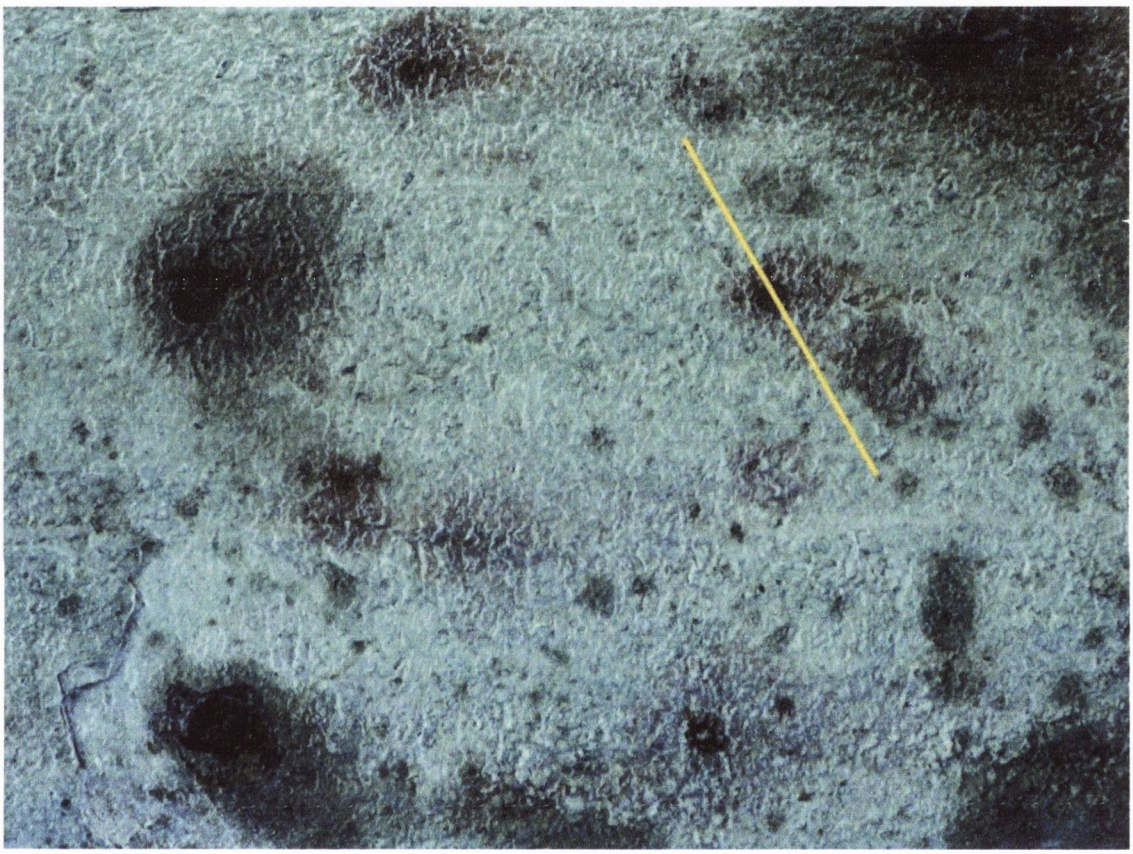


Figure 3.3 Profile line through a corroded region.

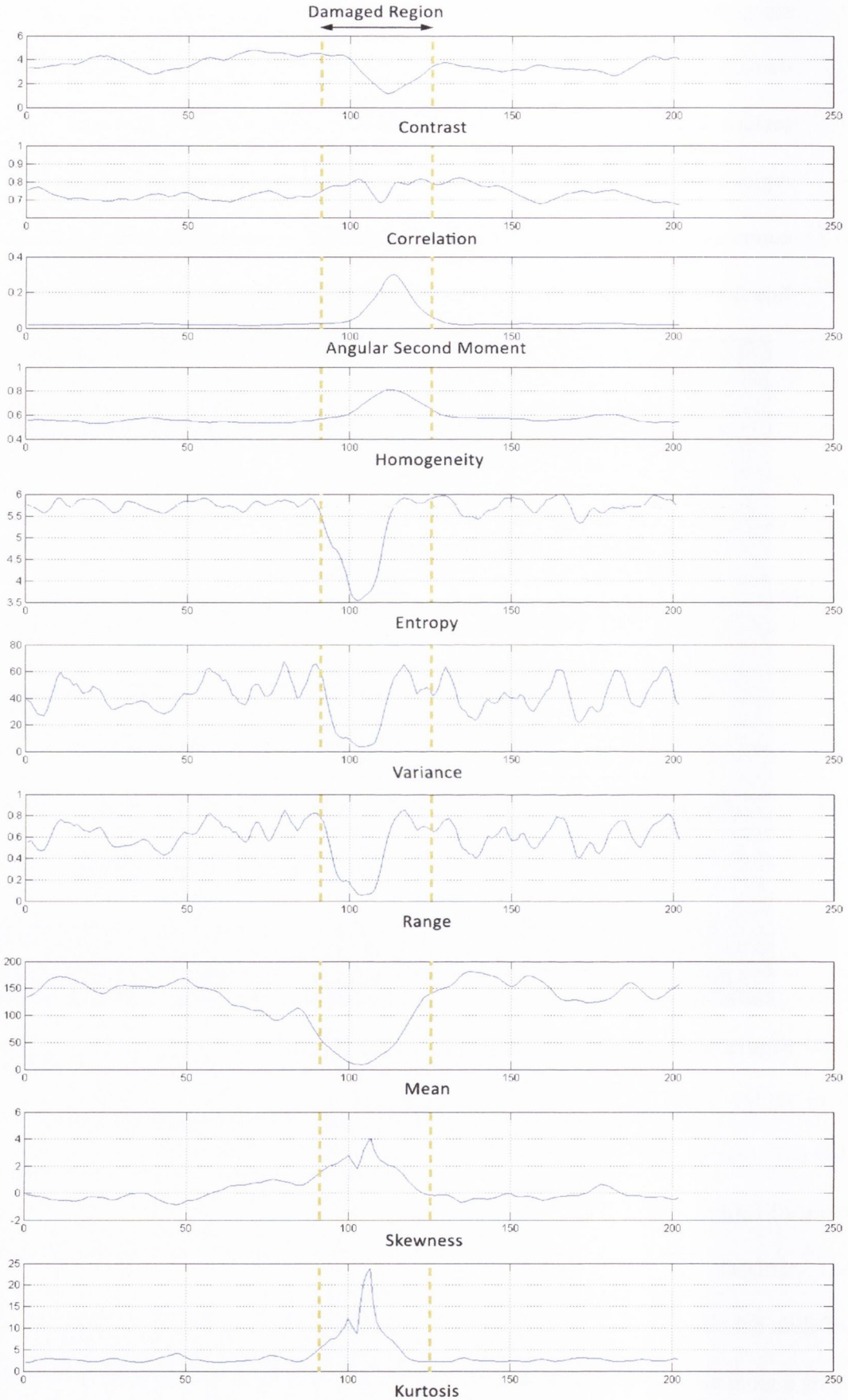


Figure 3.4 Variation of GLCM and descriptive statistics along profile in Figure 3.3

3.2.2 Non-Linear SVM Classification

SVM are used to classify pixels as being either damaged or undamaged, based on the texture feature vector assigned to each pixel. SVM is a supervised learning classifier based on statistical learning theory (Vapnik, 1995). The linear SVM is used for linearly separable data using a (*k-1*) dimensional hyperplane in *k* dimensional feature space (Vapnik, 1996). This hyperplane is called a maximum-margin hyperplane which ensures maximised distance from the hyper-plane to the nearest data points on either side in a transformed space. For linearly non-separable data a non-linear SVM is used which relies on kernel function and maximum-margin hyperplane. The proposed technique adopts a Radial Basis Function (RBF) kernel function for non-linear classification instead of the dot product between the data points and the normal vector to the hyper-plane as used for the linear classification. The kernel function concept is used to simplify the identification of the hyperplane by transforming the feature space into a high dimensional space (Boser et al., 1992; Cortes and Vapnik, 1995; Cristianini and Shawe-Taylor, 2000). The hyperplane found in the high dimensional feature space corresponds to a non-linear decision boundary in the input space.

In SVM the classifier hyperplane is generated based on training datasets. Given a training dataset of *l* points in the form $\{(x_h, y_h)\}_{h=1}^{l}$ where *h* denotes the h^{th} vector in the dataset, x_h is a *k*-dimensional input vector $(x_h \in R^n)$ and y_h is an instance label vector $(y_h \in \{1,-1\}^l)$; for this study, a value of +1 indicates presence of damage and -1 indicated absence of damage. To identify the maximum-margin hyperplane in the feature space, the SVM requires the solution of the following optimisation problem:

$$\{w,e\} = \arg\min_{w,e,\xi}\left(\frac{1}{2}w^T w + C\sum_{h=1}^{l}\xi_h\right);\ \ C>0 \tag{3.14}$$

$$\text{subject to } y_h(w^T\varphi(x_h)+e) \geq 1-\xi_h;\ \ \xi_h \geq 0$$

The function φ maps the training vectors x_h into a higher dimensional space. The vector w is the weight vector which is normal to the hyperplane, e is the bias, ξ is the misclassification error and C is the cost or penalty parameter related to ξ .The solution to the problem is given by:

$$\min_{\alpha}\frac{1}{2}\sum_{h=1}^{l}\sum_{q=1}^{l}\alpha_h\alpha_q y_h y_q K(x_h,x_q) - \sum_{h=1}^{l}\alpha_h \tag{3.15}$$

with constraints:

$$\sum_{h=1}^{l}\alpha_h y_h = 0 \tag{3.16}$$

$$0 \leq \alpha_h \leq C,\quad h=1,...,l$$

where K is the kernel function, α is the Lagrange multiplier, q is the index of the input point x_q. The Radial Basis Function (RBF) kernel has been used here,

$$K(x_h,x_q) = \exp\left(-\gamma\left\|x_h - x_q\right\|^2\right),\ \gamma>0 \tag{3.17}$$

where γ is a kernel parameter. There are two preselected parameter values for the SVM; C and γ. To estimate the optimum parameter values, a novel ROC curve based optimisation framework was employed. In this chapter, the training dataset was obtained using texture features from both damaged and undamaged regions in an analysed image.

3.2.3 SVM Models

Two models were explored in order to determine the most accurate and efficient approach. The first stage was common to both models and involved training of the SVM with a training set of data. However, the dimensions of the input vectors in the training datasets were different for each model. Also, different implementation methods were carried out in the SVM classification stage. The performance and computational time of each approach was noted.

3.2.3.1 Custom-Weighted Iterative (CWI) Model

The input vector of each pixel comprises of 30 elements: the ten features stacked together for 3 colour channels. A single binary output was achieved by introducing a weighting system which gave a greater prominence to texture measures relating to greater difference between damaged and undamaged regions. The damaged and undamaged zones were identified from the training data. The equation for the weight, W, is as follows:

$$W_{v_f} = \left| \frac{\bar{v}_{f,\,damaged} - \bar{v}_{f,\,undamaged}}{\bar{v}_{f,\,total}} \right| \tag{3.18}$$

where $\bar{v}_{f,\,damaged}$ and $\bar{v}_{f,\,undamaged}$ are the averages for the f^{th} texture descriptor in the feature vector for the damaged and undamaged regions in the training data respectively. The average of the overall training data is $\bar{v}_{f,total}$. The normalised weight, ω_{v_f}, is then assigned to each texture feature, v_f.

$$\omega_{v_f} = \frac{W_{v_f}}{\sum_{v_f} W_{v_f}} \tag{3.19}$$

A fundamental issue with the CWI model is that it required 30 separate applications of SVM classifier, one for each of the 10 texture features in each of the three colour channels. This does not represent the most effective approach in terms of computational time. However, the weighting system was found to be quite successful in terms of classification accuracy.

3.2.3.2 Four-Dimensional Input Space (4DIS) Model

The other model considered a 4D input space where the feature vector and the colour channels create two dimensions along with image coordinates for the remaining two dimensions. It is arranged as (*a, b, c, f*), where *a* and *b* are the pixel coordinates, denoting the image row and the image column respectively. *c* denotes the colour channel, which, for an RGB image, indicates either the red, green or blue channel. Finally, the fourth dimension, *f*, identifies the texture related statistic (e.g. mean, contrast, skewness, etc.). The SVM is applied once to separate the 4D input space into damaged and undamaged segments using a cubical space. It was found that this approach offered the fastest classification time with comparable classification accuracy to the CWI model.

3.2.4 Performance Measure

The performance of the texture analysis based detection in conjunction with each of the SVM models is evaluated by plotting performance points as a coordinate in the Receiver Operating Characteristic (ROC) space where the Detection Rate, *DR*, and the Misclassification Rate, *MCR*, are the vertical and horizontal coordinates respectively

(Rouhan and Schoefs, 2003; Schoefs, 2009). The *DR* and *MCR* are represented as a percentage between 0% and 100%. The *DR* and *MCR* are defined as:

$$DR \approx \frac{Card(Q)}{n_c} \text{ with } \quad Q = \{g \in \Im;\, y_k = 1\} \tag{3.20}$$

$$MCR \approx \frac{Card(\partial)}{n} \text{ with } \quad \partial = \{g \in \Im;\, y_k = -1\} \tag{3.21}$$

where *Card(.)* indicates the cardinality of a particular set, $\Im = \{1, \ldots, n\}$, n_c denotes the number of corroded pixels, and ∂ gathers situations of incorrectly detected pixels and undetected corroded pixels while Q gathers the correctly detected ones. The ROC space provides a common and convenient tool for graphically characterising the performance of NDT techniques and its usage has been extended to image detection (Pakrashi et al., 2010). A box counting approach (O'Byrne et al., 2011) was employed to calculate n_c for each image in each colour space. The *DR* and the *MCR* values formed the basis of selecting the performance point in the ROC space employing the α-δ method (Baroth et al., 2011; Schoefs et al., 2012). This method relies on calculating the angle, α, and the Euclidean distance, δ, between the best performance point, defined as an ideal NDT technique with 100% detection and 0% misclassification rates and represented in the ROC space with coordinates (0,1) and the considered point to give a measure of the performance of the considered point. As this chapter does not deal with risk analysis where the shape to the ROC acts as a key factor, only the delta, δ, parameter is required as a measure of performance. A low value for δ is indicative of a strong performing technique.

Another common method for interpreting the performance of a technique using ROC analysis is by calculating the area under the ROC curve corresponding to that technique. In order for this method to be accurate, there should be a lot of performance points that should be relatively evenly distributed along the ROC curve. Acquiring a lot of performance points is not always feasible or possible. For instance, in the case of mapping the effect that an environmental condition has on a particular technique, and where there are only a limited number of levels available (e.g. low, medium, high), the resulting ROC curve would be restricted to that same number of performance points (plus the two end points (0,0) and (1,1)). This is not sufficient to reliably construct a curve.

In cases where a lot of performance points can be generated that would allow a good curve to be constructed, there is often an issue with inefficiencies. For instance, to investigate the effect that a controllable parameter has on a technique (e.g. an adjustable detection threshold), a number of values for the parameter can be selected and the resulting ROC curve is created from the plotted performance points. However, in practice, we often have a good idea of what the value should be within some reasonable range. Selecting values outside this relevant range is time consuming and does not have any practical significance besides getting performance points near the extremities of the ROC curve (i.e. near the points (DR = 0, MCR = 0) and (DR = 1, MCR = 1)).

Image processing based detection techniques feature a trade-off between the DR and the MCR. A low alarm threshold setting of the technique, which favours high damage detection rates, will result in a higher DR but will also produce a higher MCR. Conversely, a high alarm threshold will result in a low DR and a low MCR. The best DR and MCR balance corresponds to performance point in the ROC space with the lowest value of δ.

3.3 Evaluation of the Technique

The proposed damage detection technique was applied on six images of various forms of damage on the surface of infrastructural elements. In order to assess the robustness of the technique, the six images were chosen to reflect a broad range of surfaces, damage forms, viewing angles, lighting conditions and image resolutions as shown in Figure 3.5. The sample images in the figure depict, (a) pitting corrosion on metal sheet piling in marine conditions, (b) corroded metal sheeting in coastal regions, (c) corrosion at a half joint on bridge span, (d) staining through bridge deck shown from underneath, (e) marine growth on the surface of underwater steel pile wharf and (f) exposed concrete bridge deck through wear of pavement surfacing; all in RGB colour space. The sample images are shown in HSV and $L^*a^*b^*$in Figure 3.6 and Figure 3.7 respectively. The technique was performed on the images in all three colour spaces (RGB, HSV and $L^*a^*b^*$) so as to determine whether a particular colour space offered a superior level of performance.

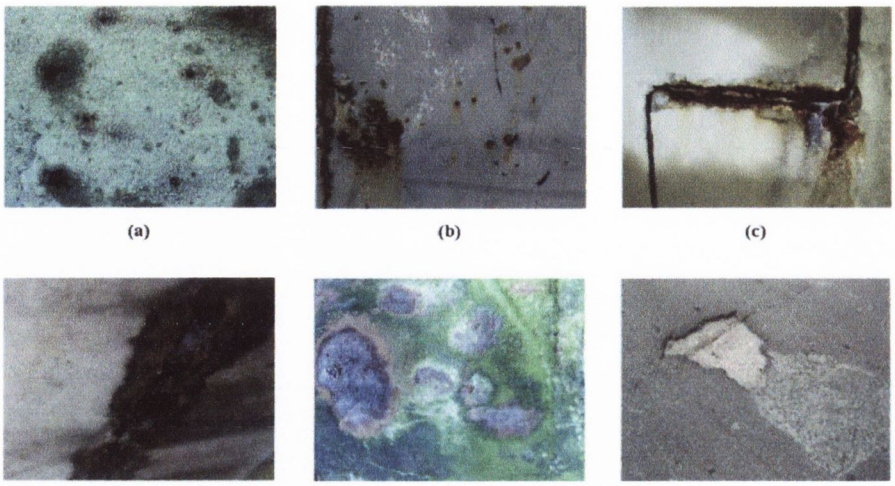


Figure 3.5 Sample images in the RGB colour space: (a) pitting corrosion on metal sheet piling, (b) corroded metal, (c) corrosion at a half joint, (d) staining through bridge deck, (e) marine growth on steel surface, (f) exposed bridge deck.

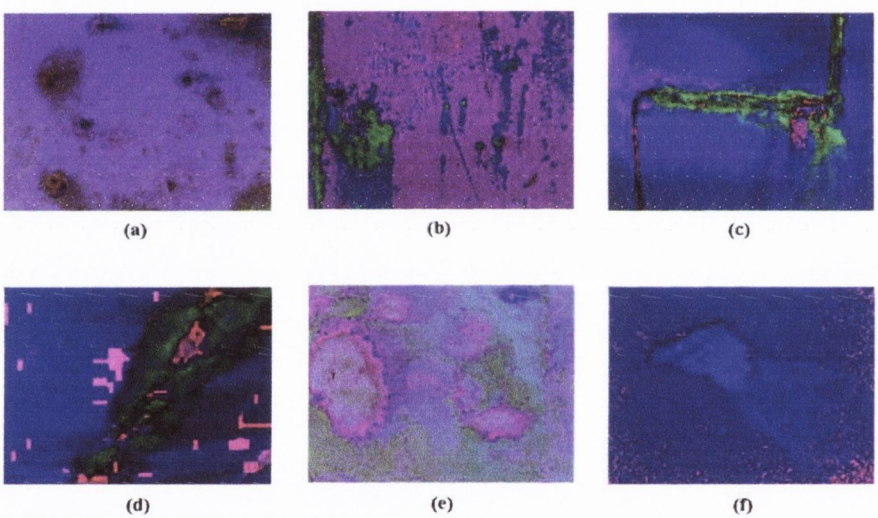


Figure 3.6 Sample images in the HSV colour space.

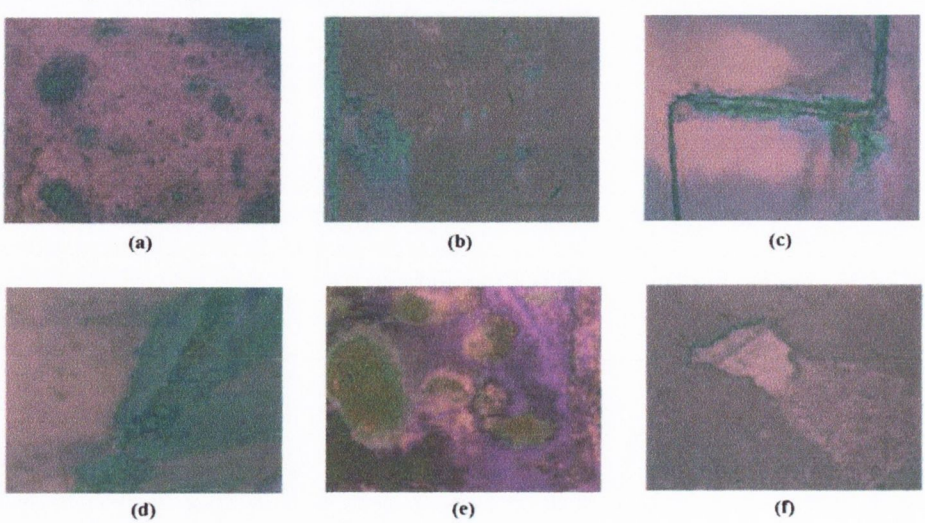


Figure 3.7 Sample images in the *L*a*b colour space.**

3.3.1 Results

The following subsections present the results obtained from the CWI and 4DIS models for each colour space. The final subsection details the procedure for selecting the SVM parameters: the penalty parameter of the error term, C, and the kernel parameter, γ so as to optimise the classification accuracy.

3.3.1.1 The Custom-Weighted Iterative (CWI) Model

The detected regions using the CWI model are shown for each colour space in Figures 3.8-3.10. The detection and misclassification rates are summarised in Table 3.1. The CWI model performed well in terms of identifying the locations of the damaged regions in the sample images. However, these regions were often poorly defined in many instances, resulting in reduced DR. An example of this is Figure 3.8(c), where the damaged regions in the RGB space have been located, but the identified damage is not homogeneous and the outer boundaries of the damaged areas are inadequately identified. This problem is also observed in the other colour spaces.

The classification accuracy was found to be dependent on colour space. The δ values in Table 3.1 provide a quantitative measure of the variation in classification accuracy among colour spaces. HSV colour space achieved a high level of performance on a consistent basis while the RGB and $L^*a^*b^*$ colour spaces were prone to more varied performance levels. The *DR* values for the sample images in $L^*a^*b^*$ space were generally high but were accompanied with a high *MCR* as well. The images in RGB colour space on the other hand showed moderate *DR* and high *MCR* values.

While the HSV colour space generally outperformed the RGB and $L^*a^*b^*$ colour spaces, a notable exception to this was in the case of the image in Figure 3.5(f), which produced the best results in the $L^*a^*b^*$ colour space. The L^*, or lightness, plane in $L^*a^*b^*$ is essentially the original image with the colour data reduced to shades of grey only. The L^* plane typically responded well to feature extraction by means of statistical analysis. Conversely, extracting textural features through statistical analysis in the a^* and b^* planes in $L^*a^*b^*$ generally yielded quite poor results as it was found that these planes were relatively non-descript and offered little distinction between damaged and undamaged regions in terms of texture. As a consequence, the CWI model had to rely disproportionately on the texture descriptors in the L^* plane in order to obtain a reasonable result. However, this issue was not a major factor in the case of Figure 3.5(f) as the dominant colours in the image were varying shades of grey resulting in the L^* plane containing a high proportion of the original image data. As a result, this image was largely unaffected by the poor performances of the statistical analysis in the a^* and b^* planes.

The explanation for the poor performance in the RGB colour space may be attributed to the high correlation between its Red, Green, and Blue components (Cheng et al., 2001). The pixel intensities from the Red, Green, and Blue colour channels are all correlated as they contain the same light and contrast information as received by the scene. Hence, the image descriptions in terms of these components make discriminating damaged and undamaged regions difficult. Descriptions in terms of hue-saturation-brightness are often more distinct and therefore more relevant for detection purposes, a point reinforced by the good results attained from the six sample images in the HSV colour space.

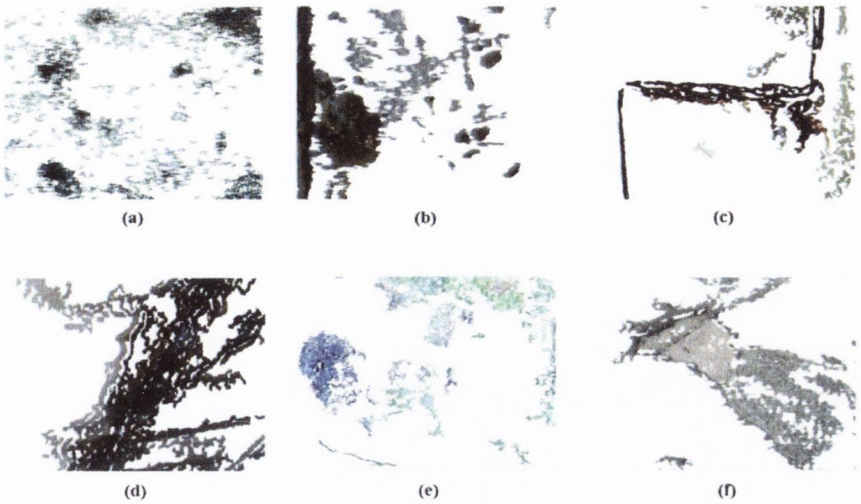


Figure 3.8 Detected regions using the CWI model for the RGB colour space.

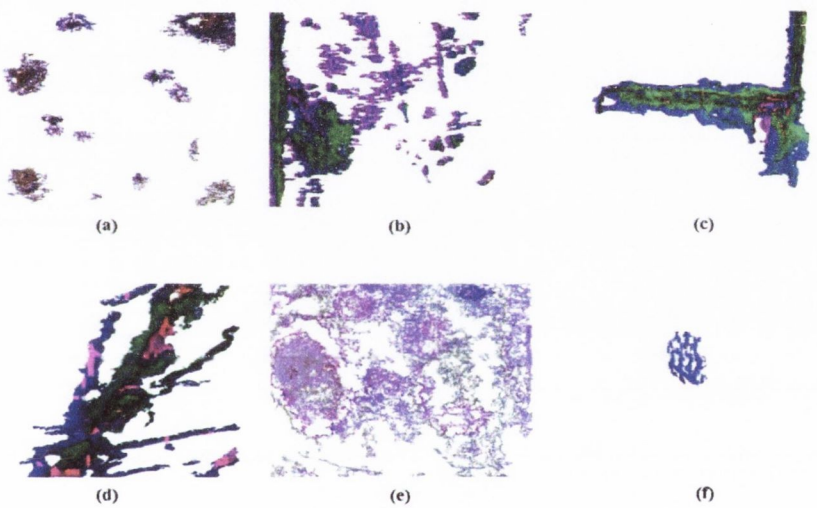


Figure 3.9 Detected regions using the CWI model for the HSV colour space.

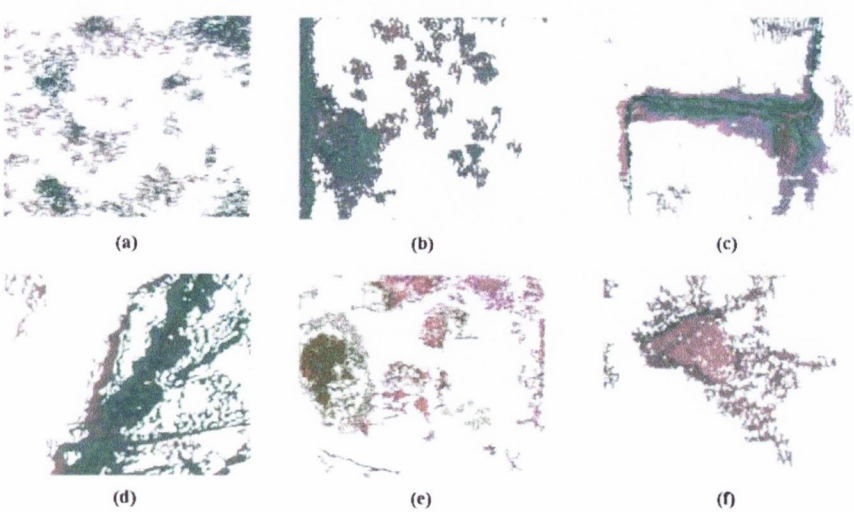


Figure 3.10 Detected regions using the CWI model for the *L*a*b colour space.**

Table 3-1 Performance for the six images in each colour space using the CWI model

Sample Image	*Colour Space*								
	RGB			***HSV***			***L*a*b****		
	DR	*MCR*	*δ*	*DR*	*MCR*	*δ*	*DR*	*MCR*	*δ*
(a) Pitting Corrosion	84.0%	30.9%	0.35	84.3%	7.2%	0.17	85.5%	34.6%	0.38
(b) Corroded Metal	96.3%	19.7%	0.2	98.0%	20.1%	0.2	94.1%	26.0%	0.27
(c) Half-Joint Damage	70.8%	10.8%	0.31	94.7%	12.0%	0.13	100.0%	23.2%	0.23
(d) Stained Deck	83.6%	36.4%	0.4	79.5%	23.5%	0.31	80.6%	26.1%	0.33
(e) Marine Growth	70.2%	40.5%	0.5	43.8%	20.9%	0.6	43.3%	20.1%	0.6
(f) Exposed Deck	66.6%	29.1%	0.44	54.4%	6.1%	0.46	73.9%	9.1%	0.28

3.3.1.2 The Four Dimensional Input Space (4DIS) Models

The detected regions for the 4DIS are shown for each colour space in Figures 3.11-3.13. The detection and misclassification rates are summarised in Table 3.2. The 4DIS model succeeded at defining the damaged regions to a better extent than the CWI model. In the majority of the cases, there were only a few spurious regions that were misclassified as being damaged. However there were occasional cases where comparatively large portions of undamaged regions in the images were misclassified such as in Figure 3.12(f) and Figure 3.13(a). As with the CWI model, the performance levels varied significantly between the colour spaces. The HSV colour space was deemed as the best option. The slightly worse performances in the RGB and *L*a*b** colour spaces may be attributed to the same reasons as outlined for the CWI model.

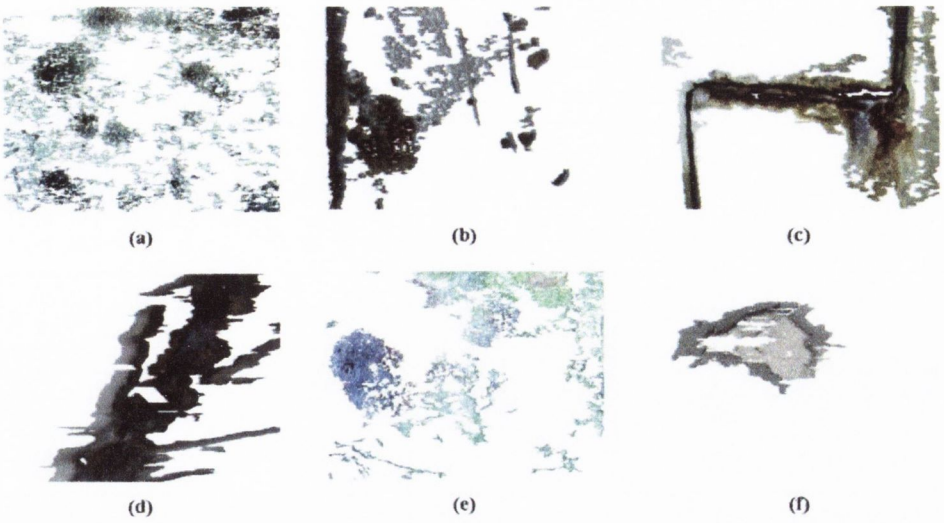


Figure 3.11 Detected regions using the 4DIS model for the RGB colour space.

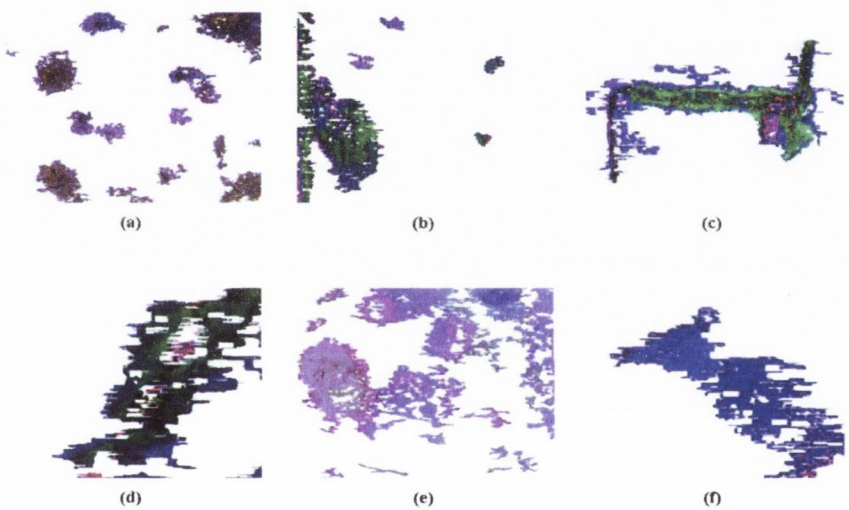


Figure 3.12 Detected regions using the 4DIS model for the HSV colour space.

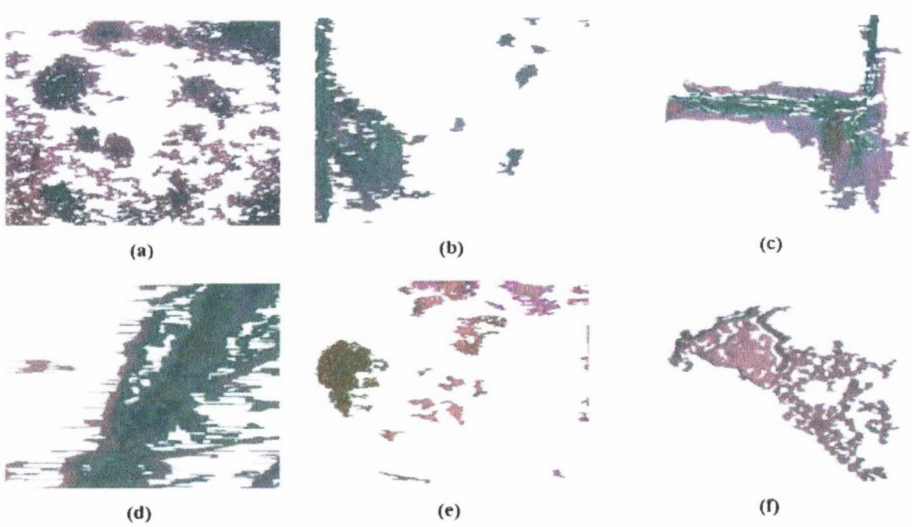


Figure 3.13 Detected regions using the 4DIS model for the *L*a*b colour space.**

Table 3-2 Performance for the six images in each colour space using the 4DIS model.

Sample Image	*Colour Space*								
	RGB			***HSV***			***L*a*b****		
	DR	*MCR*	δ	*DR*	*MCR*	δ	*DR*	*MCR*	δ
(a) Pitting Corrosion	77.8%	32.0%	0.39	88.7%	10.5%	0.15	89.9%	30.2%	0.32
(b) Corroded Metal	95.9%	24.1%	0.24	80.3%	10.0%	0.22	86.1%	14.7%	0.2
(c) Half Joint Damage	94.7%	25.9%	0.26	92.5%	8.0%	0.11	83.2%	11.0%	0.2
(d) Stained Deck	71.4%	19.4%	0.35	67.2%	23.0%	0.4	53.3%	16.1%	0.49
(e) Marine Growth	64.1%	29.0%	0.46	67.9%	22.5%	0.39	43.5%	16.0%	0.59
(f) Exposed Deck	52.1%	10.1%	0.49	86.5%	23.7%	0.27	96.7%	36.2%	0.36

3.3.2 Comparison of Model Performances

A graphical comparison of the models for each image and colour space is provided in Figure 3.14, in which the performance points corresponding to each model and colour space are plotted in the ROC space. Performance points that are near the Best Performance Point represent cases that produced high detection and low misclassification. The Line of Chance is the 45° diagonal line that connects the point (0,0) with the point (1,1) in the ROC space. It corresponds to random chance.

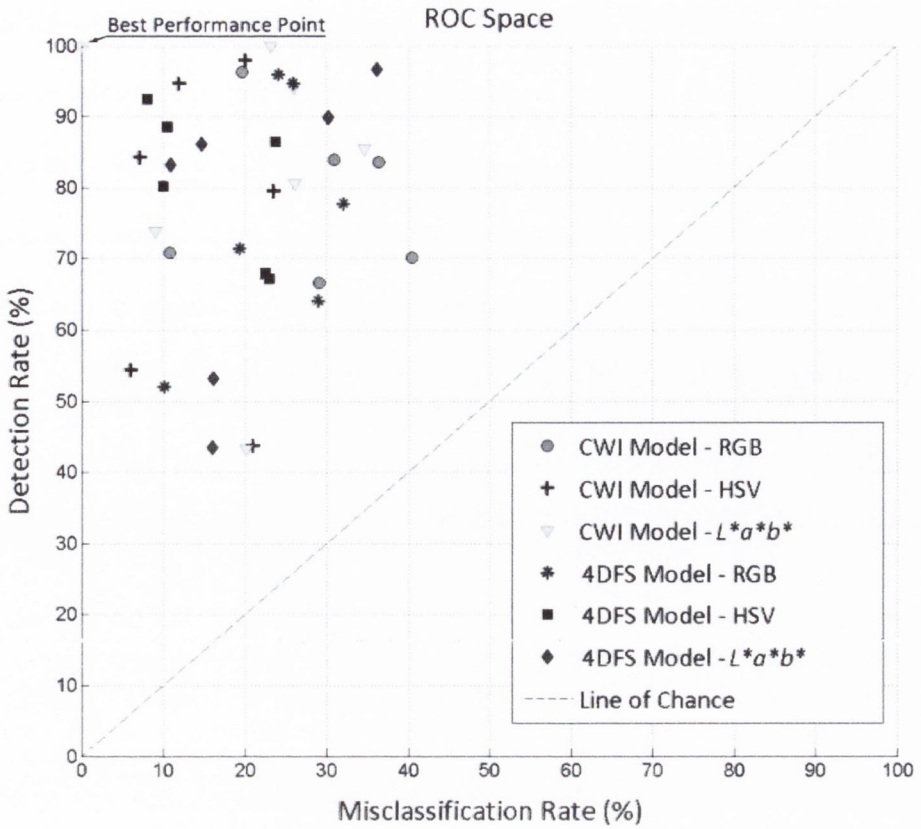


Figure 3.14 Performance points in the ROC space showing the performance of the classification models in each colour space for images in Figure 3.9.

To ensure comparability between the CWI and 4DIS model, the same training data were used for each model and colour space. Of the six images tested in each colour space, the 4DIS model outperformed the CWI model in 61% of cases. However, the performance of the models varied from image to image, with some images responding better to classification via CWI while the other images attained comparatively better results with 4DIS model (Figure 3.14).

Both the models performed consistently in different colour spaces and the HSV colour space typically provided the best results. It is evident in Figure 3.14, where the performance points corresponding to the HSV colour space for both the CWI and 4DIS models are far closer to the best performance point as compared to the other points in the

ROC space. The RGB colour space used in conjunction with the 4DIS model and the $L^*a^*b^*$ colour space used with both the CWI and 4DIS models achieved similar performance levels, reflected by the δ values for a given image in these colour space – model combinations. The images in the RGB colour space analysed using the CWI model showed the poorest performance accuracy while the images in the HSV colour space analysed using 4DIS model achieved the best performance accuracy.

3.3.3 Computation Times of Models

Whilst both models produced apparently comparable classification accuracy, their respective computational times provide a conclusive source of differentiation, with the 4DIS model being the superior option. The computation times for all sample images in RGB space for both models are presented in Table 3.3 for illustrative purposes. The other colour spaces demonstrate similar results. The different computation times for the sample images can be attributed to the image size.

A significant portion of the time in the CWI model may be attributed to its weighting system, which is required to calculate the dissimilarity between texture features in the damaged and undamaged zones, along with the inefficient application of the SVM classifier which is required to be iteratively performed for 30 times.

Table 3-3 Classification times for the 4DIS and CWI models.

Image	*Image Size (sq pixels)*	*Time Taken (seconds)*	
		4DIS Model	*CWI Model*
(a) Pitting Corrosion	1056x1408	46.5	893.6
(b) Corroded Metal	635x846	12.6	322.3
(c) Half Joint Damage	436x648	6.5	130.6
(d) Stained Deck	441x427	3.0	97.6
(e) Marine Growth	1056x1408	51.5	875.2
(f) Exposed Bridge Deck	255x391	2.5	47.0

3.3.4 Parameters of the SVM classifier

SVM classification requires a penalty parameter of the error term, C, and the kernel parameter, γ, which define the decision boundary. The parameters should be chosen carefully in order to produce an effective classifier. An ROC based optimisation framework was adopted through which the two components were independently optimised (Schoefs et al., 2012). Whilst theoretically it would be preferable to search for the optimum (C,γ) pairing, it was found experimentally that this exhaustive and computationally intensive approach was largely unnecessary as the C and γ values were largely independent of each other. For illustrative purposes, the performances of various C and γ values for Figure 3.5(a), analysed through the 4DIS model, are presented in Tables 3.4 and 3.5 respectively. The corresponding ROC curves are displayed in Figures 3.15 and 3.16 respectively.

Table 3-4 Performance of SVM for various C values (γ kept constant at 1).

C-value	*DR*	*MCR*	δ
0.001	57.00%	11.88%	0.45
0.25	67.90%	17.65%	0.37
0.5	70.20%	20.20%	0.36
0.75	74.90%	26.92%	0.37
1	77.80%	31.90%	0.39
10000	77.61%	33.13%	0.40

Table 3-5 Performance of SVM for various γ values (C kept constant at 0.5)

γ-value	*DR*	*MCR*	δ
0.25	39.72%	6.24%	0.61
0.5	57.24%	10.65%	0.44
0.66	62.45%	13.72%	0.40
0.75	64.44%	15.44%	0.39
0.8	66.14%	16.74%	0.38
1	70.28%	20.22%	0.36
1.33	77.74%	31.45%	0.39
2	86.79%	53.83%	0.55

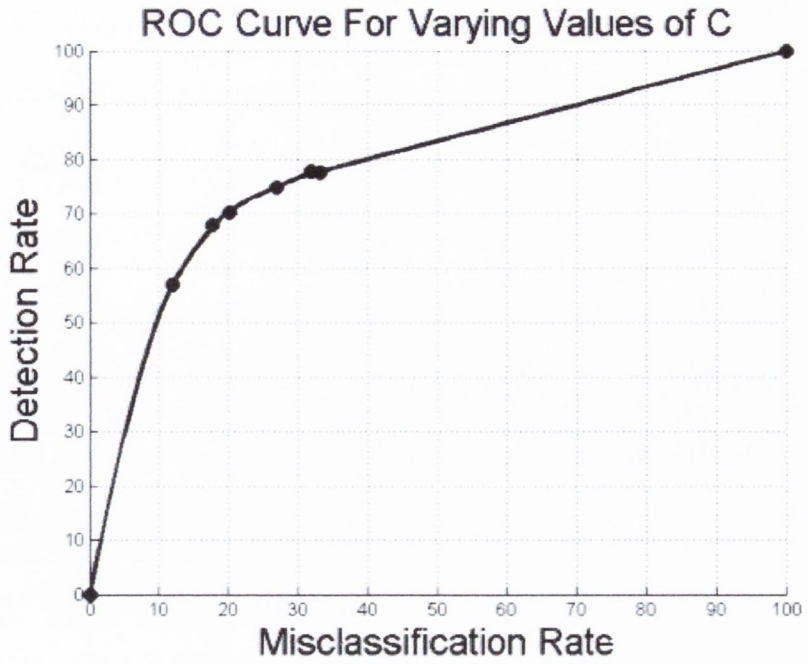


Figure 3.15 ROC curve for varying values of C (γ kept constant at 1).

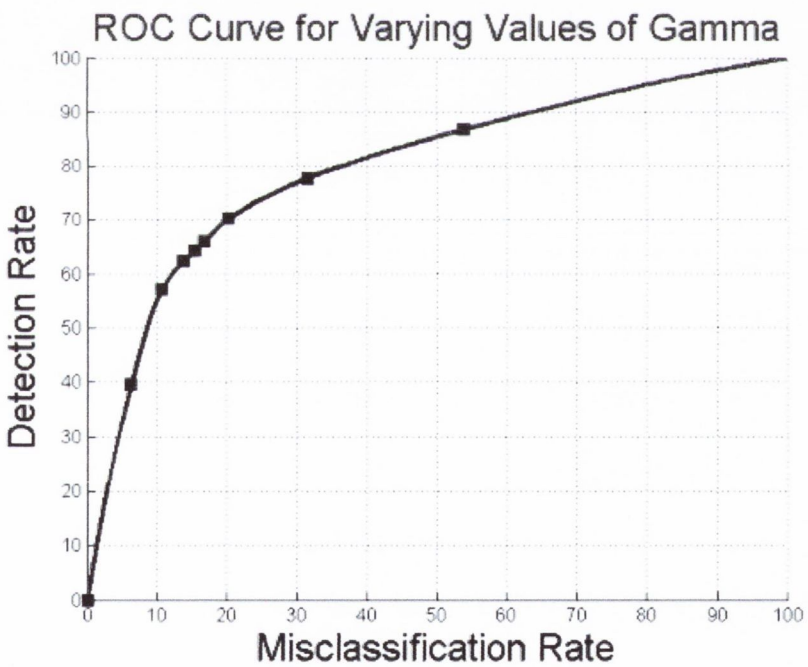


Figure 3.16 ROC curve for varying values of Gamma (C kept constant at 0.5).

The δ values attained for the set of parameter values trialled indicate that the optimum values for C and γ were 0.5 and 1 respectively. It was found that combining these independently optimised parameters provided satisfactory results; with negligible

differences from that of the jointly optimised (C, γ) pair. Moreover, the classifier demonstrated a low sensitivity to deviations from the optimal pairing suggesting that a highly optimised pairing was not integral to the classifier performance. This was especially true for the penalty parameter, C, which returned similar performance levels across a range spanning multiple orders of magnitude.

3.4 Conclusion

This chapter presents a semi-automatic texture analysis based technique for the detection and classification of damaged regions on the surface of infrastructural elements. The technique involves generating a texture feature vector for each pixel in the image including information derived from GLCM matrix based on a quantised grey-level scale along with statistical and energy information from the pixel intensity values. The pixels are consequently classified through non-linear Support Vector Machines (SVM) models.

The proposed technique has a number of favourable aspects:

- Each pixel is qualified through a large feature vector containing ten texture related measures representing both grey-levels and pixel intensities in appropriate scales providing a well-rounded description of the image in terms of the textural characterising. This aspect also increases the robustness of the technique as some measures may be good at differentiating regions in one image and may not necessarily be particularly useful in another image. This robustness is showcased by the ability of the technique to perform effectively when applied to images featuring a broad range of surfaces and damage forms, exposed to various lighting conditions, viewing angles and resolutions.

- The technique is more immune to variations in lighting conditions than colour based techniques, where only the pixel intensity values are considered as opposed to texture based segmentation techniques in which the relationship between adjacent pixel intensity values are considered. This relationship is often maintained to a significant extent even when inherent chromatic and luminous complexities are introduced to the scene.
- The technique requires only three parameters to be optimised: the size of the sliding window and two SVM parameters. A ROC curve based optimisation framework has been presented which shows a simple means of attaining suitable values for the SVM parameters. The size of window can be independently chosen through trial-and-error.

Two SVM classification models have been explored; a CWI model and a 4DIS model. The CWI model employed a weighting scheme based on the relative differences of textural descriptors in the damaged and undamaged training data. The SVM was applied iteratively to each texture measure in each colour channel. The 4DIS model offered a more efficient approach, requiring only one application of the SVM. The 4DIS model had the fastest computational time and, overall, achieved slightly better classification accuracy over the CWI model.

The proposed technique was performed in RGB, HSV and $L^*a^*b^*$colour spaces. The HSV colour space, in conjunction with the 4DIS model, offered a consistently high level of performance in a time efficient manner, and is thus concluded to be the best combination. These positive aspects underline the potential of this technique in other very critical fields as marine growth classification from picture collection from Remote Operated underwater Vehicle (Boukinda Mbadinga et al., 2007).

Chapter 4

Colour Based Image Analysis Techniques

4.1 Introduction

In chapter 3, the use of texture analysis was explored for detecting damage on the surface of infrastructural elements. Texture analysis approaches are especially useful for detecting damaged regions that can be distinguished from the non-damaged surface mainly based on their textural composition, however, common damage forms found in marine environment (e.g. corrosion, cracks, leaching, etc.) are often accompanied by a perceivable change in colour. In such cases, colour analysis techniques can be effectively employed for detecting damage. Additionally, image based methods for recovering the 3D shape of damage rely heavily on colour information. This chapter presents three colour information based techniques relating to; 1) crack detection, 2) surface damage detection, and 3) 3D shape recovery using stereo-vision. These techniques have been specifically designed to handle the challenging conditions that are encountered during inspections of marine structures, in particular, for the underwater portion of the inspection which suffers from reduced visibility. A video tracking technique is also presented for the purpose of analysing bridge vibrations. Each of the techniques are introduced and discussed in the following sections.

4.2 Crack Detection

The development of an image based crack detection algorithm is motivated by the widespread occurrence of cracks, which affect a range of civil structures such as bridges, pavements, pipes, columns, etc. The presence of cracks in a structure can be unsightly, may cause a loss in serviceability, or can lead to structural failure in more serious cases. Visible cracks provide an indication of the structural degradation and are an important factor when diagnosing the condition of a concrete structure. However, the identification and quantification of cracks is often a costly, tedious, subjective, and error prone task for visual inspectors, yet is a crucial task for establishing a safe infrastructure network. Traditional monitoring methods often rely on regular visual inspections which require inspectors to travel to the location of a structure in order to determine its current state. As part of the inspection, observed cracks are often mapped, counted, quantitatively measured and photographed. Even with great diligence, measuring the true extent of cracks remains a difficult task for inspectors as their assessment is often subjective in nature and prone to error.

Adopting an image processing based approach to automatically count and quantify the length and width of cracks can enhance inspections, and in turn, lead to significant monetary savings or more frequent inspection cycles. This chapter presents an image processing approach to efficiently and objectively detect cracks. The approach employs a percolation based method which is applied to locations in the image where there is a large colour change, or pixel intensity gradient, suggestive of edges or surface irregularities such as cracks. A new classification criterion is introduced which considers the pixel intensity values on either side of a crack. Detected regions that match the expected shape of a crack

yet have disparate sides are rejected based on the assumption that they represent non-harmful edge boundaries.

The study has also been extended to consider above surface cracks in scenes suffering from a variety of challenges, such as bright spots, difficult viewing angles, road markings/ tyre marks, and low-contrast monochromatic scenes, all of which may mislead the crack detection algorithm. The Receiver Operating Characteristic (ROC) analysis, as described in Section 3.2.4, is used to assess the performance of the technique in the presence of such variability. The methodology of the proposed technique is discussed in the following section.

4.2.1 Percolation based Crack Detection Method

Cracks are generally characterised by their narrow shape and their lower brightness in comparison to the surroundings. Percolation based methods take both of these characteristics into account. The percolation method is based on tracing out dark pixels in a sliding window, or sub-image, starting at the centre point of the window, and spreading out until the boundary of the window. The resulting pattern of dark pixels is then analysed. Cases where a narrow or linear pattern is traced out is indicative of a crack, while irregular or radial diffusion patterns typically correspond to the non-cracked background. The methodology of the proposed percolation based technique is illustrated in Figure 4.1.

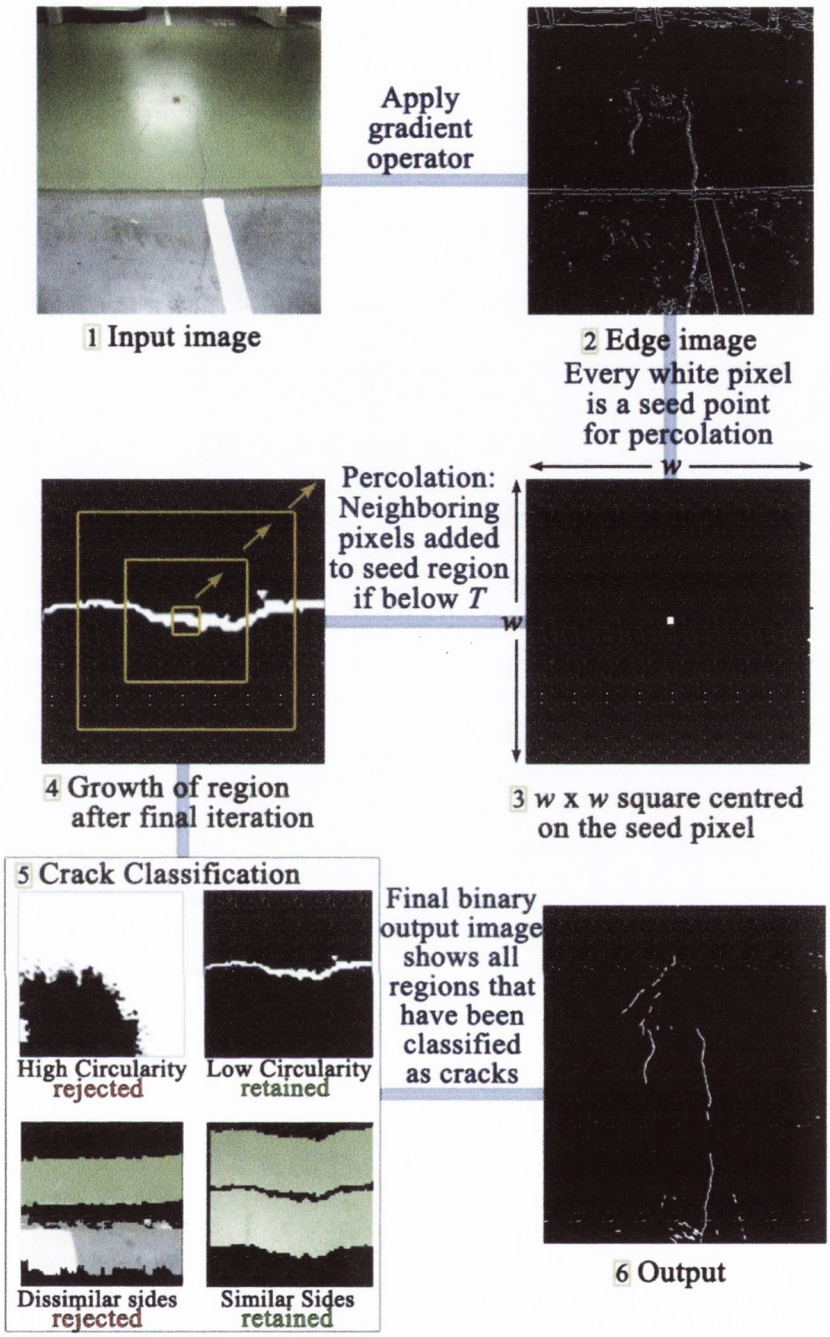


Figure 4.1 Crack detection flowchart.

The proposed method is described as follows with reference to Figure 4.1:

Step 1: The input is a colour or greyscale image, *A*. The Sobel operator is applied to this image which gives an approximation of the image gradient. The output from this operation is a binary image, where white pixels represent locations that have a large gradient. In the

case of colour images, the average gradient magnitude from the three colour channels is used.

Step 2: This binary image identifies points in the image where performing percolation would be relevant whilst overlooking points in the image where cracks are unlikely to be present. This cuts down on the computation time considerably. It is not essential that the entire crack is detected with the Sobel operator. Partial detection should be sufficient as the window will likely overlap with non-detected portions of the crack. The detected pixels from this step are considered for percolation.

Step 3: A window of size ω by ω is centred on each detected pixel from the binary edge image. The pixel at the centre of the window, p_c, is the seed point of the percolated region, R_p. It has an intensity value of $A(p_c)$. Neighbouring pixels are added to the percolated region if they are equal to, or below, the Threshold, T, which is defined as:

$$T = \max(\text{mean}(R_p), \min(R_n)) \tag{4.1}$$

where R_n denotes pixels neighbouring the percolated region R_p. Essentially, T assumes the maximum value between either the mean intensity value of pixels already in the percolated region, or the minimum intensity value of the pixels neighbouring the percolated region. This ensures that there will always be at least one pixel added to the region following each iteration. Taking the mean intensity of the pixels in the percolated region is a meaningful action as it reduces the effect of outliers which could set the threshold to be unreasonably high, and it also allows the seed points obtained from the Sobel operator stage to be only loosely located near the dark crack for percolation to succeed. If the seed point is located

just outside the crack, the percolated region should gravitate towards the interior of the crack and continue to proceed along the crack line.

Step 4: The process outlined in step 3 is repeated until the percolated region, R_p, reaches the boundary of the window. Percolation ceases and the resulting percolated shape is fed into the classification stage.

Step 5: The first stage of the classification phase involves evaluating the shape of the percolated region. This is done by calculating the circularity parameter $\aleph$. The circularity is defined as:

$$\aleph = \frac{4.P_{count}}{(\pi.P_{width}{}^2)} \tag{4.2}$$

where P_{count} is the total number of pixels in the percolated region and P_{width} is the maximum width of the percolated region. $\aleph$ can take values in the range 0 to 1. A value close to 1 corresponds to a circular shape, while a value close to 0 corresponds to a line segment that could represent a crack. With this in mind, percolated regions having circularity lower than an arbitrary threshold are retained.

The second part of the classification phase involves checking the pixel intensity values on either side of the percolated region in order to discount cases where the perceived crack has an increased likelihood of being a false alarm, perhaps arising from some innocuous line boundary in the scene. This is achieved by dilating the percolated region in a perpendicular direction on both sides by a distance, ϖ, and comparing the mean of the intensity values from the original image, A, on both sides, Z_1 and Z_2 respectively. A value

of 20 pixels was taken as ϖ throughout this chapter. If the difference between the mean of the pixel values from both side, denoted as ι, is greater than a preselected threshold than the percolated region is rejected. ι is expressed as:

$$\iota = \left| \frac{1}{n_{Z_1}} \sum A(\hbar) - \frac{1}{n_{Z_2}} \sum A(\wp) \right| \qquad \hbar \in Z_1, \ \wp \in Z_2 \tag{4.3}$$

where n_{Z_1} is the total number of pixels on one side of the crack while n_{Z_2} is the number of pixel on the other side.

Step 6: Retained percolated regions from the classification stage are added to the binary output image.

4.2.2 Error Analysis

The performance of the proposed crack detection technique is evaluated through the use of performance points in the ROC space, as described in Section 3.2.4. The Detection Rate, *DR*, and the corresponding Misclassification Rate, *MCR*, are determined by comparing the detected cracks with a visually segmented image. The visually segmented image is created by a human operator who must manually identify the cracks in an image. This visually segmented image acts as the control as it is assumed it shows the true extent of cracking. The visually segmented image only needs to be created when it is wished to gauge the performance levels of the technique under scrutiny.

4.2.3 Data Analysis

The proposed crack detection technique is demonstrated on a real cracked concrete specimen in an underwater setting (Figure 4.2) for varying controlled levels of turbidity and lighting. The quality of these photographs is assumed to be chiefly affected by luminosity, sharpness (focus accuracy), contrast and noise. These quality factors are directly related to the on-site operating conditions, of which lighting and turbidity are the most influential.

Turbidity is defined as the cloudiness in a liquid caused by the presence of suspended solids (Kirk, 1985). These suspended solids scatter and absorb light and therefore reduce visibility. The level of turbidity may be exacerbated by interference during the inspection process, such as from a mechanical source like a boat or human contact with the river bed, which may aggravate and disturb sediment.

Lighting also plays a pivotal role for achieving good visibility. Ambient light may be sufficient for near-surface inspections; however, it is unlikely to be sufficient at greater depths at which point artificial light sources become necessary. These artificial light sources may introduce luminous complexities such as 'bright-spots', whereby areas of high light intensity may be focused at a spot which may fool the crack detection algorithm. In this section, three levels for turbidity were chosen (0 NTU (clear water), 6 NTU, and 12 NTU) and three lighting levels were used (100 Lux, 1000 Lux, 10000 Lux).

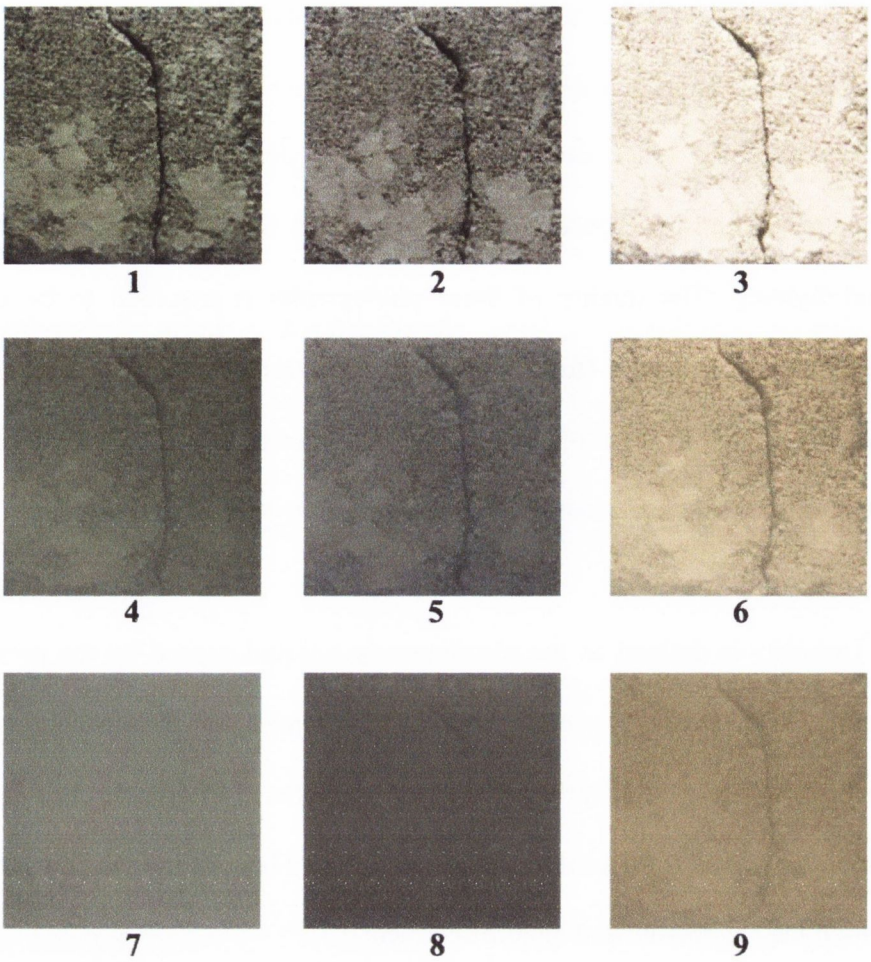


Figure 4.2 Cracked concrete specimen under vary lighing and turbidity conditions. Columns: Low, Medium, High Light. Rows: Low, Medium, High Turbidity.

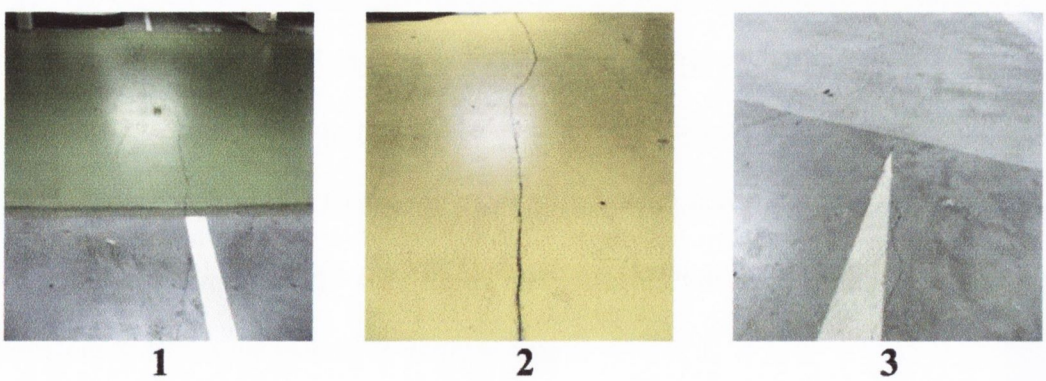


Figure 4.3 Cracks in a conventional above-water setting.

Although underwater scenes generally provide some of the poorest visibility levels, interpreting images from above-water scenes can also be severely affected by some adverse conditions. It is thus important that the influence of some of the more common challenges for crack detection should be investigated. Figure 4.3 shows three image of cracks, each presenting one or more challenges, namely, (1) Road markings, (2) illumination complexities (i.e. bright-spots), (3) monochromatic scene with vague cracks.

4.2.4 Results

This section presents and discusses the results obtained following application of the crack detection technique on the underwater and above water images. The results for the underwater concrete specimen for changing lighting and turbidity levels are shown in Figure 4.4.

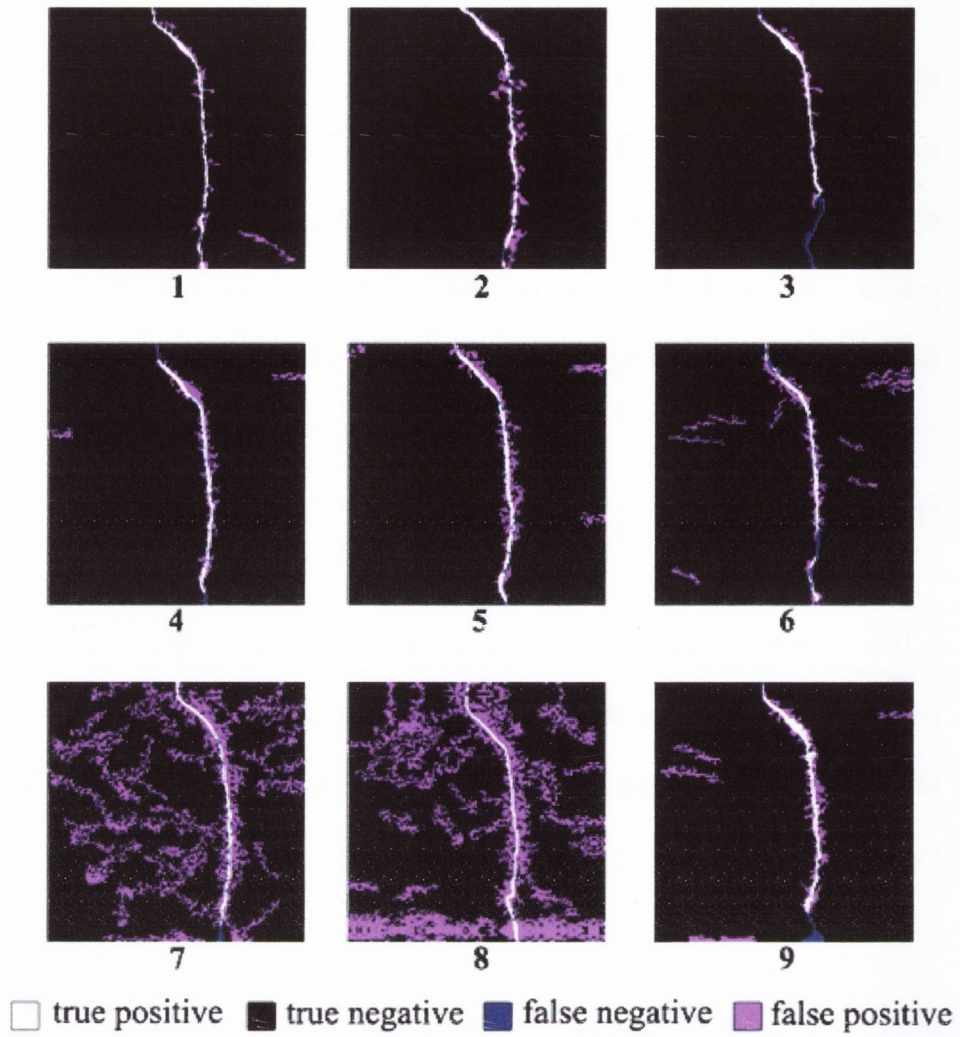


Figure 4.4 Detected cracks corresponding to the images in Figure 4.2.

The results in the case of the above-water images are shown in Figure 4.5.

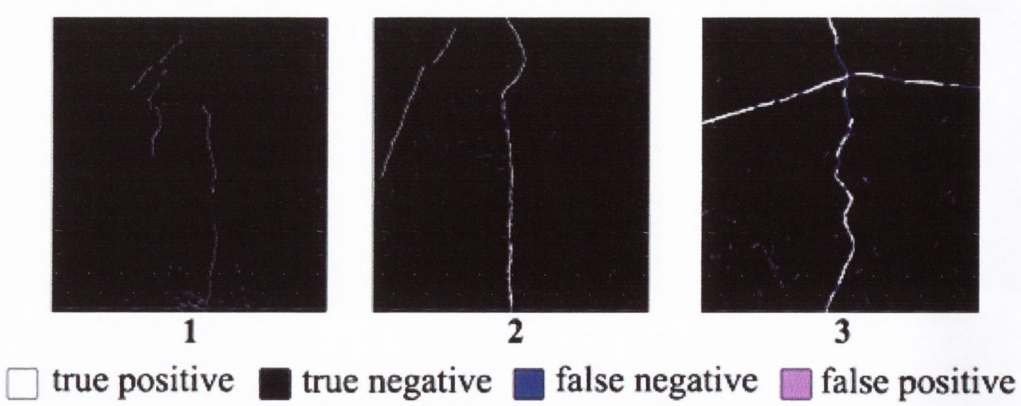


Figure 4.5 Detected cracks for the above-water images in Figure 4.3.

The performance of the crack detection technique is quantified in Table 4.1 and the associated performance points are plotted in the ROC space in Figure 4.6.

Table 4-1 Performance of the crack detection technique.

Image	Condition	(DR)	(MCR)	δ
Underwater cracks				
1	Low Light, Low Turbidity	89.8%	1.7%	0.10
2	Medium Light, Low Turbidity	87.4%	2.5%	0.13
3	High Light, Low Turbidity	62.6%	1.5%	0.37
4	Low Light, Medium Turbidity	76.3%	2.5%	0.24
5	Medium Light, Medium Turbidity	86.6%	3.7%	0.14
6	High Light, Medium Turbidity	68.6%	4.2%	0.32
7	Low Light, High Turbidity	80.0%	21.3%	0.29
8	Medium Light, High Turbidity	94.0%	26.8%	0.27
9	High Light, High Turbidity	78.1%	4.2%	0.22
Above-water cracks				
1	Road markings	73.0%	0.4%	0.27
2	Illumination complexities	69.2%	0.9%	0.31
3	Monochromatic scene	78.1%	0.6%	0.22

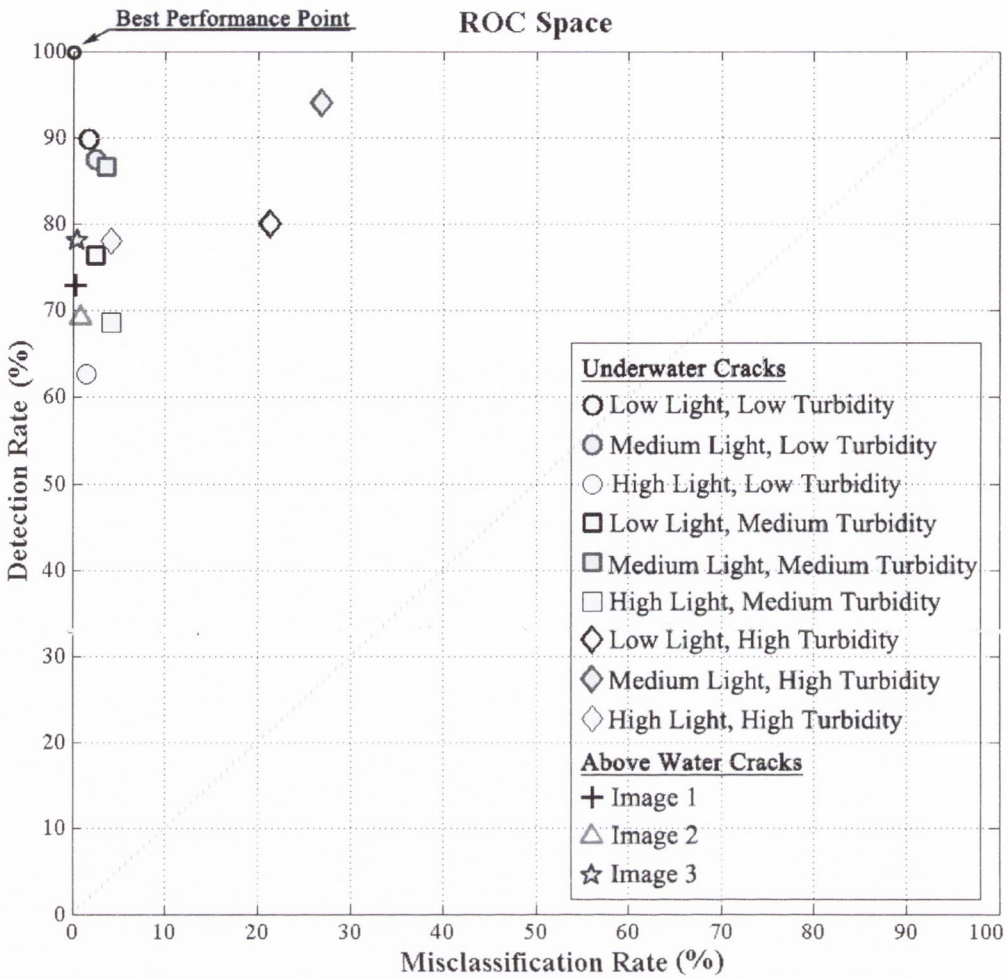


Figure 4.6 Evaluation of crack detection technique through the use of performance points in the ROC space.

It may be observed from the detected regions in Figure 4.4 that the ability of the detection technique is strongly related to the degree of clarity in the original images (Figure 4.2). Unsurprisingly, images that feature poor visibility conditions produced poor performance levels; while on the other hand, images featuring clear and sufficiently lit scenes produced quite good detection results.

At the same time however, the results suggest that an overly bright light source can be harmful to the detection accuracy. The nature of underwater lighting means that an overly bright light source will induce a bright-spot at the centre of the light beam which will

gradually fall-off, resulting in a non-uniform scene. It is this non-uniformity in lighting which the camera cannot effectively compensate for. An example of this can be seen in Figure 4.4(3) where the strong lighting creates a bright area on the surface of the specimen. This washes out some of the detail, thereby preventing a portion of the crack from being detected.

A clear trend that emerges from analysis of Table 4.1 is that the misclassification rate increases with increasing turbidity. However, the relationship does not appear to be linear. Instead, there is a gradual increase in *MCR* values from the low to the medium turbidity levels, while there is a pronounced increase between the medium and high turbidity levels. This suggests that as the turbidity approaches the operating limits there is a rapid deterioration in performance. This is especially evident for the high turbidity images in Figure 4.7 (7 and 8) where there is a high degree of false positives contributing to high *MCR* values in comparison to the lower turbidity levels.

An exception to the decline in performance associated with increasing turbidity and excessive lighting is in Figure 4.4(9). In this case, the high turbidity partially mitigates the high lighting through absorption and diffusion which limits the formation of a bright spot. Overall, the best results are obtained for the low and medium light levels at the lowest turbidity, and the medium lighting level at the medium turbidity level (Figure 4.4, images 1, 2, and 5).

Analysis of the results in Table 4.1 and the corresponding performance points in the ROC space (Figure 4.6) reveal that each image performs quite consistently in spite of the varying challenges faced in each image. The first image (Figure 4.3(1)) demonstrates the usefulness of employing the classification criteria which rejects 'cracks' having dissimilar

sides. It may be observed from the image of the detected cracks (Figure 4.5(1)) that there are very few false positives, even though the boundary between some road markings bears a resemblance to a crack (fine structured and darker than the surroundings).

The bright spot in Figure 4.3(2) has a deleterious effect on the detection accuracy as the crack remains partially undetected coinciding with this spot. This suggests that attempts to minimise the glare through careful choice of camera viewing angle/position during the on-site image acquisition stage would be worthwhile.

The cracks in Figure 4.3(3) appear quite vague at certain places. There are also numerous instances of staining on the floor which may mislead the crack detection algorithm. As such, it may be expected that the *DR* would suffer and there would be a high *MCR*, however the *DR* remains reasonably high and the *MCR* remains low. This may be attributed to percolation based methods which have been recognised as being well suited for detecting ambiguous cracks (Wang and Huang, 2010).

4.2.5 Crack Detection Method Discussion

A percolation based method for detecting cracks is proposed in this section and applied to series of images under a host of varying conditions. The proposed method offers greater efficiency than other percolation based methods, whilst still achieving good results, as the search space for cracks is confined to places where there is a sharp image intensity gradient as found by the Sobel operator. Furthermore, a new set of criteria is developed for the percolation stage and for the subsequent classification of the percolated regions.

Understanding how the operating conditions affect the performance of an image based crack detection technique is of great practical importance. This section examines the effect of two parameters; turbidity and lighting, for an image processing based NDT tool for underwater structural health monitoring purposes. These parameters affect the underwater visibility, which is crucial for any image based NDT to succeed. The influence and relative importance of these two environmental conditions are investigated in greater detail in Chapter 5.

Although environmental conditions such as turbidity are generally uncontrollable in practice, lighting can be easily adjusted. Even in conventional above-water settings, consideration should be given to the lighting conditions in order to avoid bright spots. It is readily attainable to have appropriate lighting conditions which are conducive to structural health monitoring and damage detection. The results from this study may facilitate an inspector when deciding on the appropriateness and implementation of an image processing based crack detection technique, under a given set of environmental circumstances.

4.3 Surface Damage Detection

The motivation behind the development an improved image based damage detection technique is to provide owners/operators of infrastructure with an efficient source of quantitative information. The aim of the image analysis stage is to locate and quantify the area occupied by visible mechanical damage (typically larger than 10^{-6} m^2) on the surface of infrastructural elements with minimal human supervision. There exists scope for significant development of damage detection algorithms that can characterise features of interest in challenging scenes with credibility. This section presents a new Regionally

Enhanced Multi-Phase Segmentation (REMPS) technique that is designed to detect a broad range of damage forms on the surface of civil infrastructure. The technique is successfully applied to a corroding infrastructure component in a harbour facility. REMPS integrates spatial and pixel relationships to identify, classify, and quantify the area of damaged regions to a high degree of accuracy. The image of interest is pre-processed through a contrast enhancement and colour reduction scheme. Features in the image are then identified using a Sobel edge detector, followed by subsequent classification using a clustering based filtering technique. Finally, Support Vector Machines (SVM) are used to classify pixels which are locally supplemented onto damaged regions to improve their size and shape characteristics. Combining each of these constituent phases in an effective manner creates a powerful and robust detection algorithm. To further improve the detection accuracy of REMPS, High Dynamic Range (HDR) imagery is examined.

The superiority of REMPS over existing segmentation approaches is demonstrated, in particular when considering High Dynamic Range (HDR) imagery. It is shown that REMPS easily extends beyond the application presented and may be considered an effective and versatile standalone segmentation technique. Situations where adopting a colour analysis based approach over a texture analysis based approach are highlighted.

4.3.1 High Dynamic Range (HDR)

HDR imagery is a set of techniques that are used to allow a greater dynamic range of luminance values between the brightest and darkest regions of an image than standard digital images. SDR images can typically only accommodate a very limited range bracket of the full tonal spectrum in a real world scene. Therefore, a dynamic range bracket would have to be chosen in the knowledge that all luminance values outside the range would not

be represented correctly. The broad principle behind HDR imagery is that multiple SDR images of the same scene, each taken at a different exposure, and thus capturing a different range bracket of the tonal spectrum, may be merged to form one HDR image that has a wider dynamic range (Reinhard et al., 2008). Combining SDR images can be done using various merging algorithms (Debevec and Malik, 2008; Naccari et al., 2005).

The benefits of adopting HDR imagery as an imaging protocol may be observed in Figure 4.7 which shows three SDR images (an underexposed, a normally exposed and an overexposed image) and the corresponding HDR image. These images depict a 30-year old corroded steel pile in the tidal area in a wharf situated off the French Atlantic Ocean. The pixel dimensions of the images are 2816 pixels by 2112 pixels and the dimensions of the corroding metallic surface are approximately 0.3 m by 0.3 m. It may be observed that HDR imagery is particularly useful here since the shiny metallic surface gives rise to naturally high dynamic ranges. Generally, scenes which have a wide dynamic range due are likely to especially benefit from the adoption of HDR as a protocol.

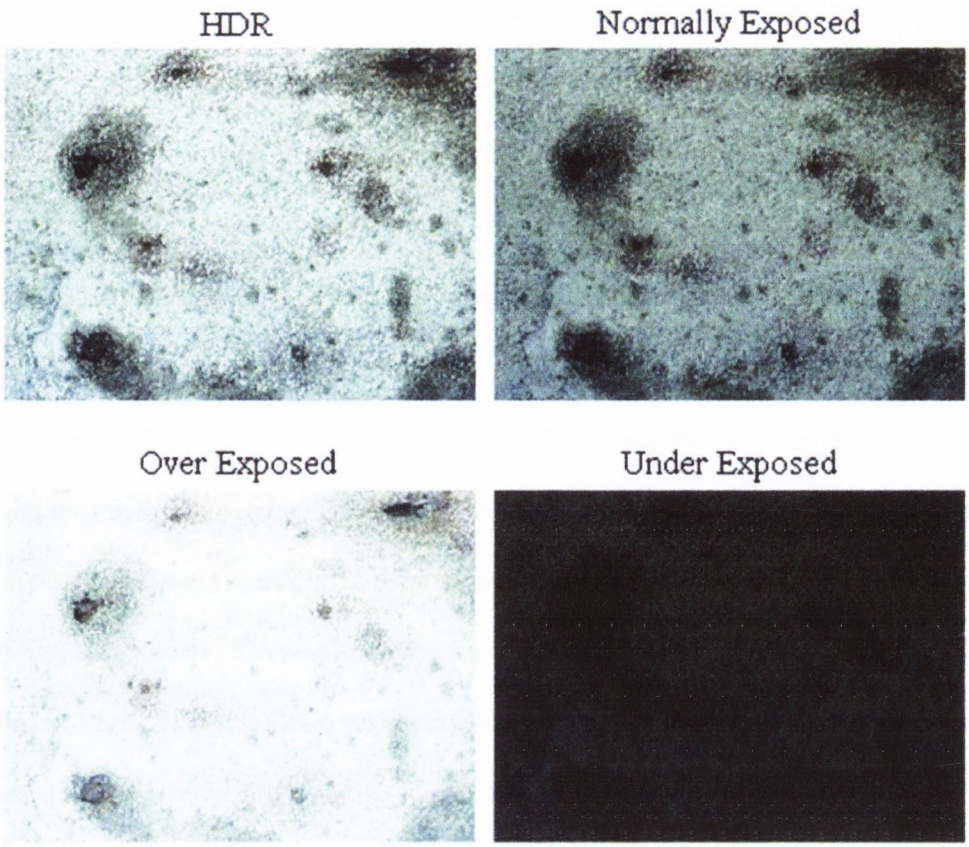


Figure 4.7 A High Dynamic Range (HDR) image of corroding steel formed by merging the normally, over and under exposed images together.

4.3.2 Regionally Enhanced Multi-Phase Segmentation (REMPS) Methodology

The REMPS technique integrates three feature detection methods. A flowchart illustrating the order of the feature detection methods is presented in Figure 4.8. The first method involves the application of the Sobel edge detector on a modified image in order to form closed geometries corresponding to objects in a scene. Statistical properties are calculated for each of the closed geometries. Statistical based approaches are popular owing to their computationally inexpensive nature and their robustness (Giralt et al., 2013; Li et al., 2013). These statistical properties serve as input to a clustering based filtering phase which retains closed geometries that have statistical properties characteristic of damaged regions whilst discarding closed geometries that have statistical properties

characteristic of non-damaged regions. SVMs are then used to identify potentially damaged pixels adjacent to these filtered closed geometries in order to improve the definition of the damaged regions.

REMPS attempts to utilise the advantages of these three techniques most effectively. The low complexity of the Sobel edge detector and the clustering based filtering techniques are complimented by the strategic application of the high complexity SVMs. For instance, the robustness and generality of the Sobel edge detector serves as a natural precursor to the closed geometry clustering stage. This clustering stage performs well at classifying the presence of damage, however, it is only after the pixel supplementation stage that the shape and size characteristics of the retained closed geometries are sufficiently realised. Finally, a Receiver Operating Characteristic (ROC) based optimisation framework may be employed to determine the best input parameters. Each stage is discussed in the following sub-sections.

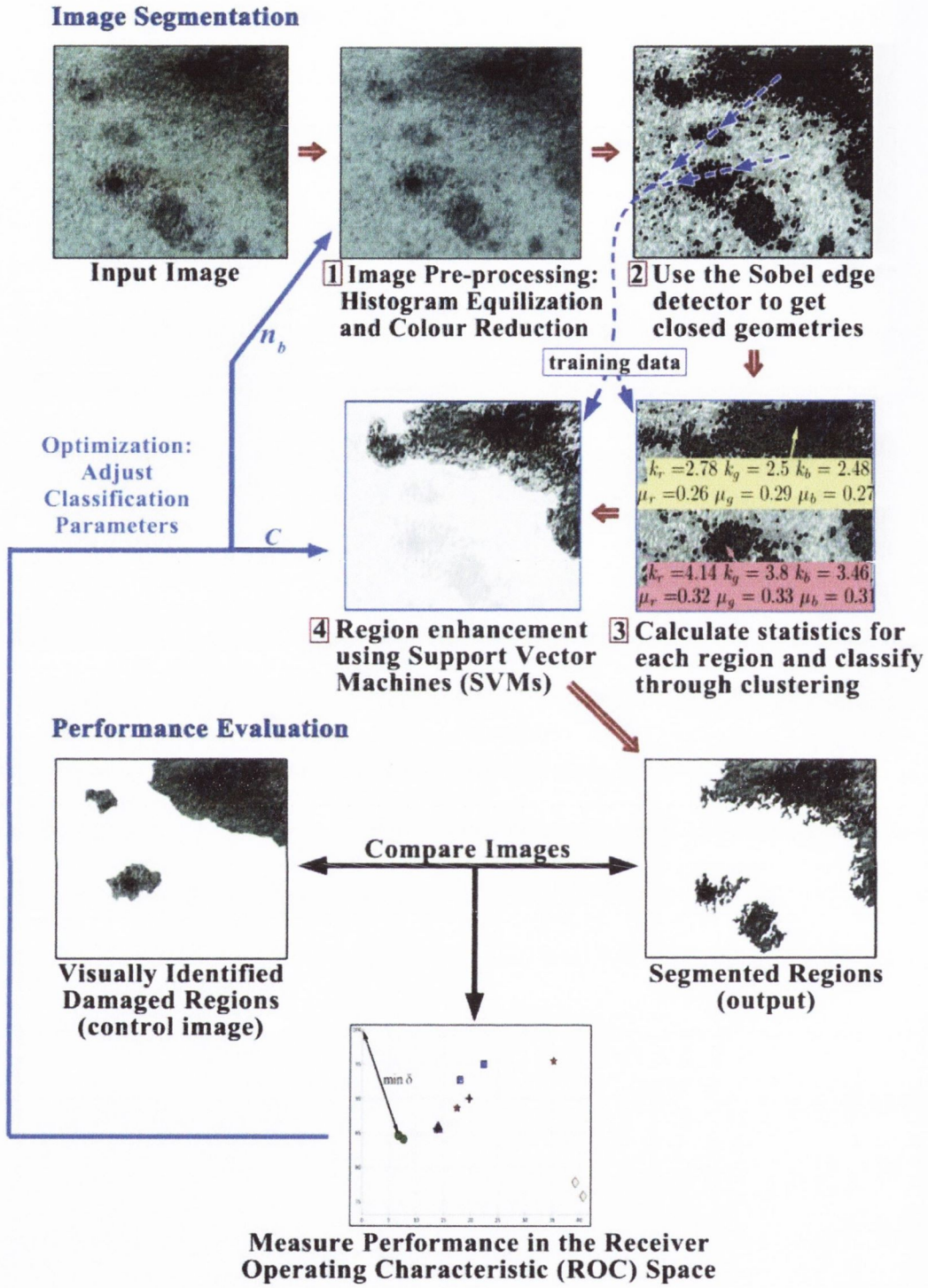


Figure 4.8 REMPS Flowchart applied to a zoomed-in portion of the normally exposed image from Figure 4.7.

4.3.2.1 Identification

The first stage of the damage identification process involves the creation of a temporary image which undergoes contrast enhancement and colour reduction operations. These operations help make the boundaries of features of interest in a scene more apparent and to focus the following analysis on key regions. This is an important step before the application of the Sobel operator, as often in natural scenes, the transition from damaged to undamaged zones is ambiguous, resulting in an increased likelihood that an edge boundary may be undetected. Contrast is amplified through a process known as Histogram Equalisation (HE) (O'Gorman et al., 2008; Pizer et al., 1987) whereby the intensity values are uniformly spread over the full range of each colour channel in the image [0, 255]. Let A denote an image represented as an M_x x M_y x CC matrix of pixel intensity values, where M_x and M_y are the row and column lengths and CC is the number of colour channels. For colour images, there are three colour channels (e.g. RGB images are comprised of a Red, Green, and Blue plane). For HE, each colour channel is operated on separately. The discrete frequency and the fraction $p_c(v)$ of a pixel having intensity v in the c^{th} colour channel are assimilated and the fraction is defined as

$$p_c(v) = \frac{n_{v,c}}{n} \quad \text{for } 0 \le v < L;\ n_{v,c} \le n \tag{4.4}$$

where n is the total number of pixels in the image and $n_{v,c}$ is the number of occurrences of pixels with intensity value v in the c^{th} colour channel; L is the total number of intensity levels in the image A (for images defined on a scale of [0, 255], $L = 256$). L assumes the same value for all colour channels. The Cumulative Distribution Function (CDF), P_c,

which provides the accumulated normalised histogram for the c^{th} colour plane can be computed as:

$$P_c(v) = \sum_{\tau \leq v} p_c(\tau) \tag{4.5}$$

With knowledge of the CDF, the general equation for generating the HE image may be written as:

$$heq_c(v) = round\left(\frac{P_c(v) - P_c(v=0)}{(M_x \times M_y) - P_c(v=0)} (n_b - 1) \right) n_b \leq L \tag{4.6}$$

where $heq_c(v)$ is the histogram equalised intensity value in the c^{th} colour plane. This equation incorporates colour reduction which quantises the intensity values in each colour channel into n_b discrete bins. Finally for this preliminary modification stage, a 2D greyscale image, B, of size M_x x M_y, is formed by averaging the intensity values from each colour channel as per Equation 4. This equation also includes a scaling term, $\frac{255}{n_b - 1}$, for restoring the image's range from [0, n_b-1] to the original range of [0, 255].

$$\bar{b}(v) = round\left(\frac{255}{n_b - 1} \frac{1}{CC} \sum_{c=1}^{CC} heq_c(v) \right) \tag{4.7}$$

where $\bar{b}(v)$ is the corrected intensity value in image B. The function of the rounding operator is to ensure that the intensity values in image B remain discrete integers. A value of 14 was used for n_b for each application of REMPS on the sample images shown throughout this section. This value was chosen as it was experimentally found to offer a

sufficient number of distinct bins and provide suitable grouping of perceptually similar pixels within each bin. The n_b parameter can be optimised methodically through a Receiver Operating Characteristic (ROC) based optimisation framework as discussed in Section 3.2.4, or simply using a trial and error approach, which involves trying various n_b values until satisfactory segmentation results are obtained. Initial detection of features of interest may now be carried by applying an edge detector to the modified image.

The Sobel operator edge detector works by calculating approximations of the first derivatives of an image in horizontal and vertical directions respectively (Abdou and Pratt, 1979). Denoting G_{horz} and G_{vert} as the two masks which give the horizontal and vertical derivative approximations at each point as

$$G_{horz} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * B \qquad \text{and} \qquad G_{vert} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * B$$

(4.8)

where the asterisk denotes the 2-dimensional convolution operation. A padding with a thickness of one pixel may be applied around the border of image B during the convolution process thereby enabling the computation to be performed at the image extremities. The intensity values in the padding assume the value of the neighbouring pixel in the original image. At each point in the image, the resulting gradient approximations can be combined to give the gradient magnitude, using

$$G_{total} = \sqrt{G_{horz}^2 + G_{vert}^2}$$

(4.9)

A large value of G_{total} represents a sharp change in image intensity which in turn is indicative of an edge boundary. Since the preliminary histogram equalisation and colour reduction steps prevent the occurrence of weak edges, all non-zero values of G_{total} may be taken as being representative of an edge. The detected boundaries for the HDR image are shown in Figure 4.9. The region enclosed by the boundary is denoted by R_j where j is the index of the region (j=1,…,J).

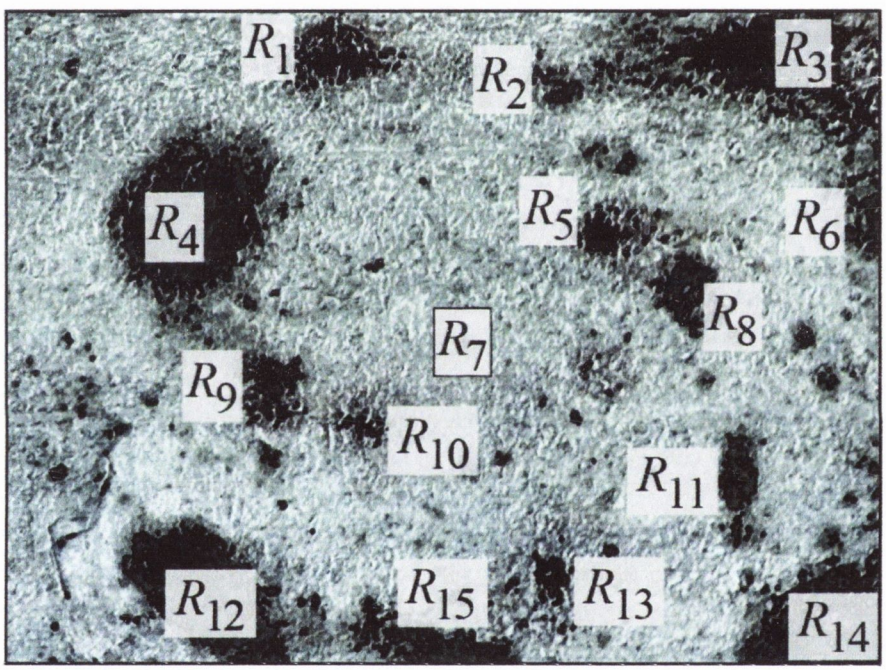


Figure 4.9 Detected closed geometries following pre-processing and application of the Sobel operator.

It may be observed from Figure 4.9 that many closed geometries detected by the Sobel operator are of a negligible size which tend to represent spurious regions rather than damaged regions. For computational parsimony and classification accuracy purposes, closed geometries below a certain size are not considered for future analysis. The chosen threshold area can be viewed as the minimum defect size, below which regions are

considered to present an insignificant degree of damage. It may be convenient to represent the threshold area as a function of the overall image size. For instance, it could be specified that closed geometries less than 1% of the total image area should be discarded. A priori knowledge of the damage type and its relationship to the decision process (repair, detailed inspection etc.) may be used as a factor in choosing the threshold area. The remaining closed geometries are classified by means of a clustering technique.

4.3.2.2 Clustering based Filtering

Given a set of closed geometries ($R_o = R_1 , R_2,..., R_J$), the clustering algorithm aims to partition the J observations into two sets $S = \{S_1, S_2\}$ such that the Euclidean distance between the centroid of R_o and the centroid of the set which it is assigned to is minimised. S_1 corresponds to the cluster representing damaged regions while S_2 represents the undamaged cluster.

The centroid of each closed geometry is given by the point $(\mu_1,\mu_2,\mu_3,k_1,k_2,k_3)_o$ where $\mu_{1\text{-}3}$ is the mean, and $k_{1\text{-}3}$ is the kurtosis, of the pixel distribution for each of the three colour channels within the o^{th} closed geometry. The mean of each colour channel for each region is computed by:

$$\mu_{c,o} = \frac{1}{n_{R_o}} \left(\sum_{t=1}^{n_{R_o}} v_{t,c,o} \right) \qquad t \in R_o \tag{4.10}$$

while the kurtosis is given by:

$$k_{c,o} = \frac{m_4}{m_2^2}$$

$$\text{where } m_4 = \frac{\sum_{t=1}^{n_{R_o}} (v_{t,c,o} - \mu_{c,o})^4}{n_{R_o}} \tag{4.11}$$

$$\text{and } m_2^2 = \left(\frac{\sum_{t=1}^{n_{R_o}} (v_{t,c,o} - \mu_{c,o})^2}{n_{R_o}} \right)^2 \qquad t \in R_o$$

where $v_{t,c,o}$ denotes the intensity value for the pixel with index t within the o^{th} region for the c^{th} colour channel, while n_{R_o} denotes the total number of pixels in the o^{th} region. Kurtosis provides a measure of the peakedness of a probability distribution. It is helpful for assessing the homogeneity in a set of distributed values by computing the shape of the probabilistic mode. While kurtosis gives a measure of shape, it fails to provide information about the centre of a distribution. With this in mind, using kurtosis in conjunction with the mean is an effective way to describe the pixel distribution within each closed geometry. Representing a closed geometry solely based on the mean is susceptible to error as closed geometries with disparate pixel distributions may yield similar values. Introducing kurtosis offsets this issue and creates a more well-rounded description of each closed geometry. Its amplitude independent nature means that it is less affected by variations in contrast levels between images within a batch, ensuring that training data selected in one (or more) image(s) remains relevant to other images in the batch. Furthermore, it was experimentally found to provide a good description of damaged regions. A scatter plot of the mean and kurtosis values for the numbered regions in Figure 4.10 is illustrated in Figure 4.10(b). The centroid of the o^{th} closed geometry, R_o, is given by $(\mu_1,\mu_2,\mu_3,k_1,k_2,k_3)_o$. The centroids of the damaged and undamaged clusters are obtained from the training data. The training data is comprised of two closed geometries which are representative of a damaged and undamaged zone. These regions must be manually selected. In the illustrated example, the

region labelled "R_1" in Figure 4.9 was used as the damaged training data while the background was chosen as the undamaged region (labelled "R_7"). The cluster centroid for the damaged cluster, S_1, is thus given by the vector $(\mu_1,\mu_2,\mu_3,k_1,k_2,k_3)_1$, while the centroid of the undamaged cluster, S_2, is given by $(\mu_1,\mu_2,\mu_3,k_1,k_2,k_3)_7$.

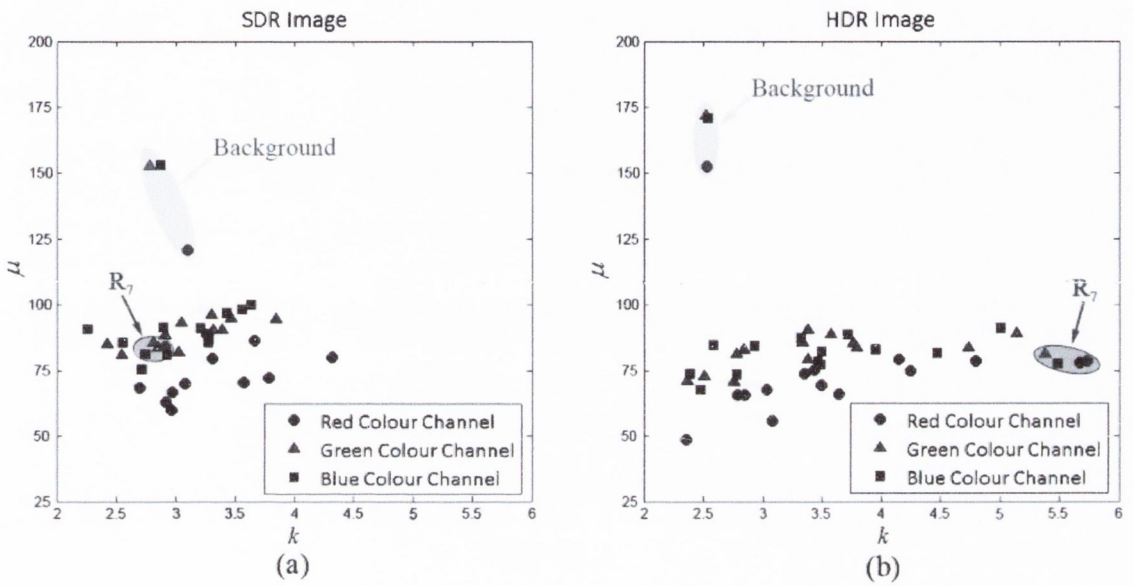


Figure 4.10 Scatter plot of mean and kurtosis values for each closed geometry in the (a) SDR and (b) HDR images.

Figure 4.10 illustrates the difference between the range of μ and k values in the SDR and HDR images. It may be observed from the HDR image scatter plot (Figure 4.10(b)) that there is a greater degree of separation, according to μ, between the background and the other closed geometries in comparison to the SDR image scatter plot (Figure 4.10(a)). Additionally, the scatter points for the HDR image are more dispersed, according to k, which should, in theory, facilitate clustering by reducing the likelihood of ambiguous closed geometries that lie on the decision boundary.

R_o is assigned to the set which minimises the Euclidean distance between the observation centroid and cluster centroid as:

$$R_O \in \begin{cases} S_1, \text{ if } \sum_{\kappa} \left\| R_{O,\kappa} - R_{damaged,\kappa} \right\|^2 \leq \sum_{\kappa} \left\| R_{O,\kappa} - R_{undamaged,\kappa} \right\|^2 \\ S_2, \text{otherwise} \end{cases}$$

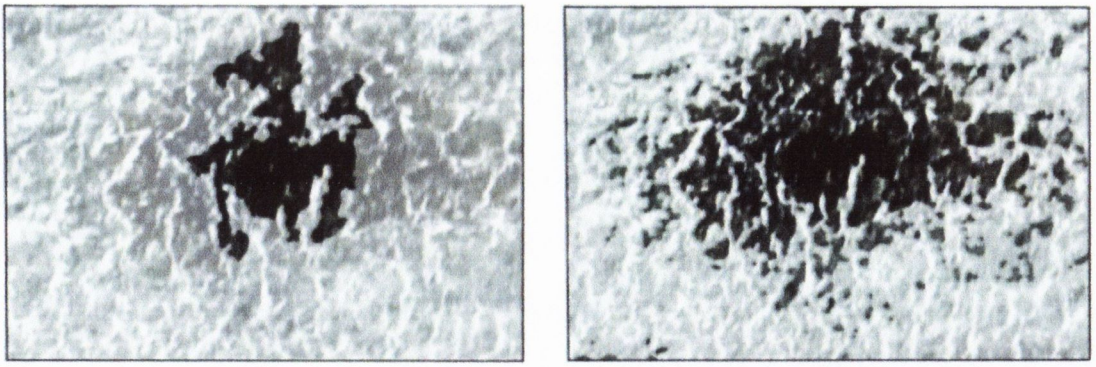
(4.12)

where $\kappa = 1,2,...,6$ denotes the index of the elements in R_o.

Once the closed geometries have been grouped into their respective clusters, it is necessary to enhance their size and shape characteristics.

4.3.2.3 Enhancement

Following the region based clustering stage, there still exists many damaged pixels around the periphery of the region that remain undetected Performing SVM in the neighbourhood of these regions and then locally supplementing the closed geometries with the SVM pixels produces better defined features of interest. This is conveyed by comparing the closed geometry R_1 before and after the local application of SVM classified pixels (Figure 4.11(a) and 4.11(b) respectively).



(a) (b)

Figure 4.11 Close-up of R1, (a) Before local enhancement, and (b) Detected pixels from SVM classification.

SVMs are used to classify pixels as being either damaged or undamaged based on the intensity values for each colour channel. SVM is a supervised learning classifier based on statistical learning theory. The linear SVM is used for linearly separable data using a $(\mathfrak{d}-1)$ dimensional hyperplane in $\mathfrak{d}$ dimensional feature space (Boser et al.; Cortes and Vapnik, 1995; Cristianini and Shawe-Taylor, 2000; Vapnik, 1995). This hyperplane is called a maximum-margin hyperplane which ensures maximised distance from the hyperplane to the nearest data points on either side in a transformed space. The linear kernel function is the dot product between the data points and the normal vector to the hyper-plane. The kernel function concept is used to simplify the identification of the hyperplane by transforming the feature space into a high dimensional space. The hyperplane found in the high dimensional feature space corresponds to a decision boundary in the input space.

In SVM the classifier hyperplane is generated based on training datasets. The same damaged and undamaged regions used in the clustering stage are used as the training data. Given a training dataset of l points in the form $\{(u_h, v_h)\}_{h=1}^{l}$ where h denotes the h^{th} vector in the dataset, u_h is a real $\mathfrak{d}$-dimensional input vector containing the mean and kurtosis values

associated with each region $u_h \in \Re^3$ and v_h is an instance label vector $(v_h \in \{1,-1\}^l)$; for this study, a value of +1 indicates presence of damage and -1 indicated absence of damage. To identify the maximum-margin hyperplane in the feature space, the SVM requires the solution of the following optimisation problem:

$$\{w,e\} = \arg\min_{w,b,\xi}\left(\frac{1}{2}w^T w + C\sum_{h=1}^{l}\xi_h\right);\quad C>0$$
$$\text{subject to } v_h(w^T\varphi(u_h)+e) \geq 1-\xi_h;\quad \xi_h \geq 0 \tag{4.13}$$

The function φ maps the training vectors u_h into a higher dimensional space. The vector w is the weight vector which is normal to the hyperplane, e is the bias, ξ is the misclassification error and C is the cost or penalty parameter related to ξ. The solution to this constrained problem is obtained by introducing Lagrange multipliers as expressed in the following Quadratic Programming (QP) optimisation problem:

$$\min_{w,b,\xi}\max_{w,b,\xi}\left(\frac{1}{2}w^T w + C\sum_{h=1}^{l}\xi_h - \sum_{h=1}^{l}\alpha_h[v_h(w^T\varphi(u_h)+e)-1+\xi_h] - \sum_{h=1}^{l}\beta_h\xi_h\right)$$
$$\text{with } \alpha_h,\beta_h \geq 0 \tag{4.14}$$

where α and β are the Lagrange multipliers. By substituting $w = \sum_{h=1}^{l}\alpha_h v_h u_h$, it can show that the dual form of this problem reduces to the following optimization problem where the aim is to maximise the Lagrange multipliers (Vapnik, 1996):

$$\tilde{L}(\alpha) = \left(\sum_{h=1}^{l}\alpha_h - \frac{1}{2}\sum_{h=1}^{l}\sum_{q=1}^{l}\alpha_h\alpha_q v_h v_{x,y} K(u_h, u_{x,y,c})\right) \tag{4.15}$$

This can be framed as a minimisation problem:

$$\min_{\alpha} \left(\frac{1}{2} \sum_{h=1}^{l} \sum_{q=1}^{l} \alpha_h \alpha_q v_h v_{x,y} K(u_h, u_{x,y,c}) - \sum_{h=1}^{l} \alpha_h \right) \tag{4.16}$$

With Constraints:

$$\sum_{h=1}^{l} \alpha_h v_h = 0$$
$$0 \le \alpha_h \le C, h = 1, \ldots, l \tag{4.17}$$

where K is the kernel function α_h and α_q are the Lagrange multipliers, $v_{x,y}$ is a label vector ($v_{x,y} \in \{1,-1\}$) for the input point $u_{x,y,c}$. This formulation has the advantage of not having to use w, or the mapping function φ, as well as having a unique optimal solution. This problem may be solved using QP techniques. The linear kernel has been used here, which is given by:

$$K(u_h, u_{x,y,c}) = u_h^T u_{x,y,c} \tag{4.18}$$

There is one preselected parameter value for the SVM, namely the cost parameter C, which may be optimised in a similar fashion to the n_b parameter from the colour reduction stage using an ROC based optimisation framework/trial and error approach.

The enhancement process firstly examines pixels that are immediately adjacent to each retained region, R_j. A pixel is considered to be adjacent to a region if it shares an edge or corner with any pixel on the periphery of that region. SVM classification is applied to these adjacent pixels utilising their original intensity values ($a_{x,y,c}$) to classify each of these pixels as representing damaged surface or not. Pixels which are classified using SVMs as representing damage become a member of the region, R_j. Pixels in immediate vicinity of

the newly identified member pixels of R_j are further subjected to classification using SVMs. This process is repeated until there are no more adjacent damaged pixels that can be added to a region. For computational parsimony, no individual pixel is subjected to classification using SVMs more than once in the entire region enhancement step.

4.3.3 Evaluation of REMPS

This section presents the results obtained by REMPS when applied to the SDR (normally-exposed) and HDR images of pitting corrosion (Figure 4.7). The performance is investigated for several colour spaces. These colour spaces are introduced in the first subsection. A comparison of the results for each colour space is provided in the second subsection, followed by a subsection which compares the performance of REMPS against established segmentation techniques. The final subsection presents the performance of REMPS alongside a texture analysis based damage detection technique when applied to a variety of damage forms.

4.3.3.1 HSV and *L*a*b* Colour Spaces

Two additional colour spaces were considered in order to determine whether this could improve the accuracy of detection, namely the HSV and *L*a*b** spaces. HSV (Hue, Saturation, and Value) is one of several variations of colour spaces characterised by the factors in the parenthesis. It is often used in computer vision and image analysis for feature detection or image segmentation as the hue component is believed to be especially useful for separating objects with different colours (Vapnik, 1995). Often, detection algorithms applied to colour images are extensions to algorithms designed for greyscale images whereby each of the three colour channels is separately passed through the same algorithm.

It is important, therefore, that the features of interest can be distinguished by the colour dimensions used. Because the red, green and blue components in an RGB image are all correlated with the same amount of light hitting the object, and therefore with each other, image descriptions in terms of these components can make object discrimination difficult. Descriptions in terms of hue-saturation-brightness are often more relevant due to this separation of chromatic and achromatic information. The HDR image of pitting corrosion is shown in the HSV colour space in Figure 4.12(b).

The $L^*a^*b^*$ colour space also offers some interesting benefits over the RGB space, especially in cases where the colour of damaged zones is perceptually close to the colour of the undamaged surface. The $L^*a^*b^*$ space consists of a luminosity layer L^*, and chromaticity layers a^* and b^*. The L^* component is similar to the V component HSV space. It closely matches the human perception of lightness. Being able to isolate the lightness layer is helpful for making accurate colour balance corrections which is useful when lighting levels cannot be controlled (O'Byrne et al., 2013). The colour information is stored in the a^* and b^* layers. The a^* component indicates where the colour lies on the red-green axis, while the b^* component indicates where the colour lies on the blue-yellow axis. The $L^*a^*b^*$ space attempts to reflect a uniform change in perceived colour with a corresponding uniform change in the L^*, a^*, and b^* components. The HDR image in the $L^*a^*b^*$ colour space is shown in Figure 4.12(c).

4.3.3.2 Comparison of Colour Spaces

Different colour spaces encode and numerically represent colour in various ways. Consequently, some colour spaces are more receptive to certain segmentation techniques than others. REMPS is applied to the SDR and HDR images in the RGB, HSV and $L^*a^*b^*$

colour spaces to explore whether a particular colour space responds well to the proposed technique. The detected regions for the HDR image in each colour space are shown in Figure 4.12. The performance of REMPS for the SDR and HDR images in each colour space are quantified in Table 4.2 and the associated performance points are plotted in the ROC space in Figure 4.13.

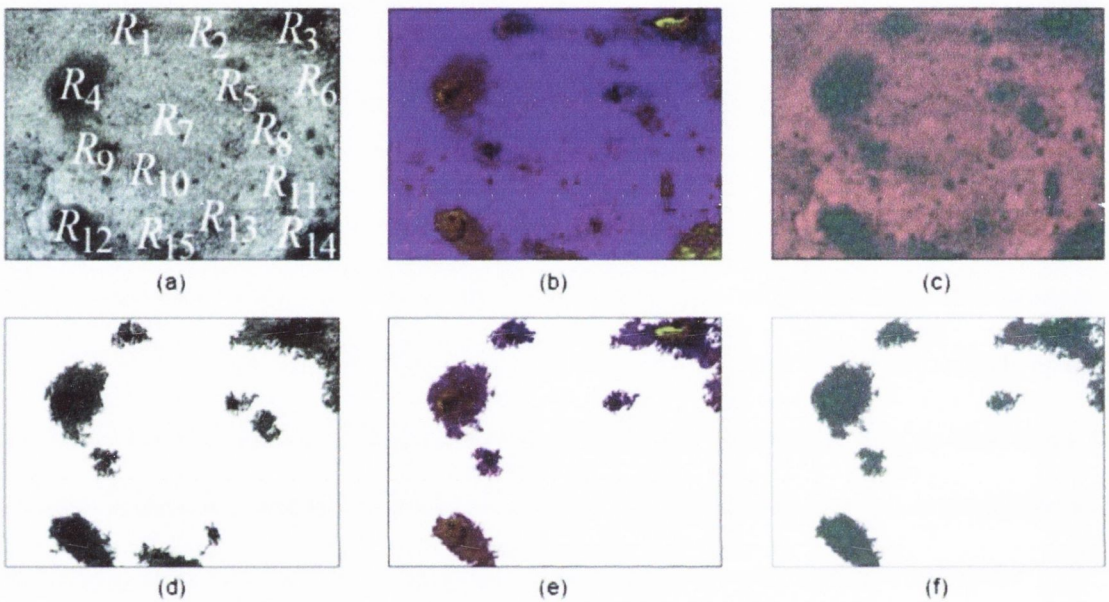


Figure 4.12 HDR Image of Pitting Corrosion. (a) Image in RGB space, (b) Image in HSV space, (c) Image in L*a*b* space, (d) Detected Regions in RGB, (e) Detected Regions in HSV, (f) Detected Regions in L*a*b*.

Table 4-2 Detection accuracy for the SDR and HDR image of pitting corrosion for each colour space.

Colour Space	SDR Image			HDR Image		
	DR	***MCR***	***δ***	***DR***	***MCR***	***δ***
RGB	84%	8%	0.18	85%	7%	0.17
HSV	85%	7%	0.17	84%	5%	0.17
*L*a*b**	76%	5%	0.24	85%	5%	0.16

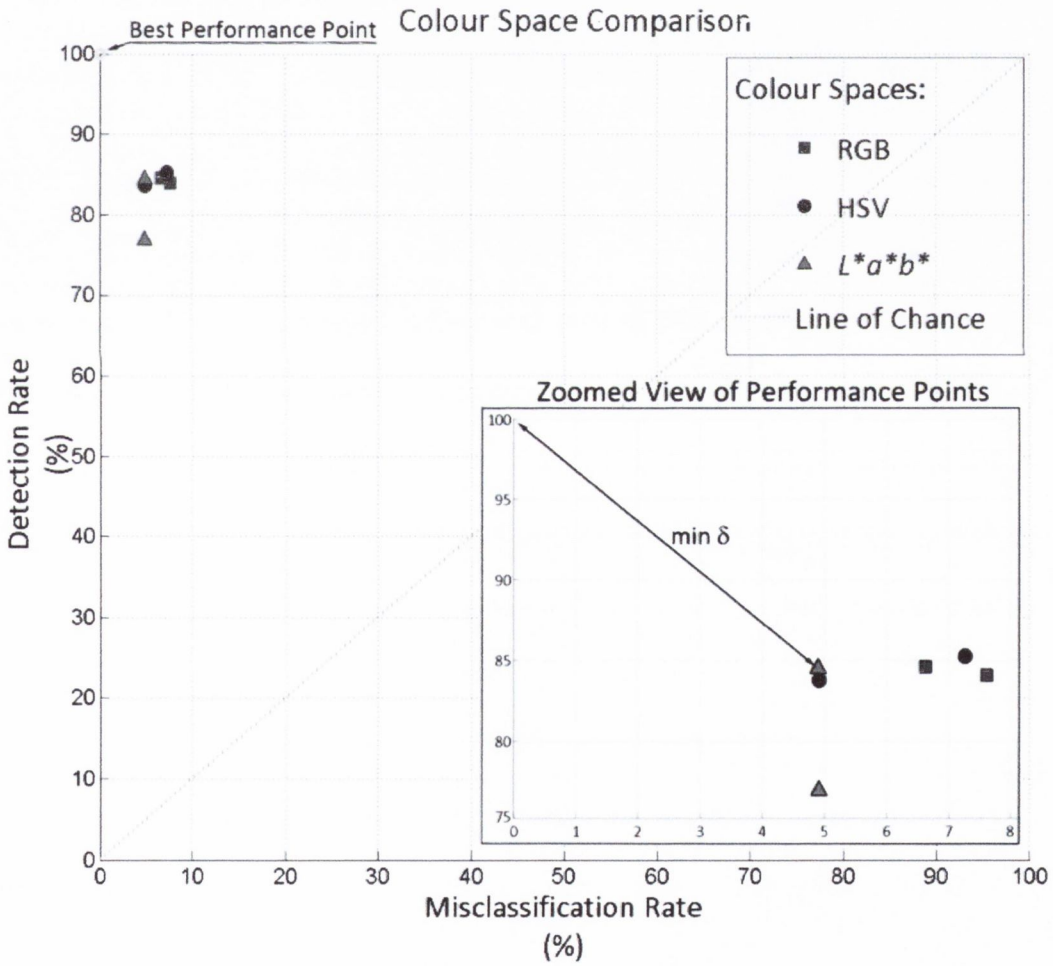


Figure 4.13 ROC curves depicting the performance of REMPS algorithm in various colour spaces.

It may be observed from the relatively compact nature of performance points in Figure 4.13 that the accuracy of REMPS is not heavily reliant on the colour space. Despite this, some interesting findings emerged. Analysis of the δ values in Table 4.2, which represent the best balance between DR and MCR for each image, reveals that HDR imagery typically offers a superior performance over the SDR image. Overall, it was the HDR image in the $L^*a^*b^*$ space that achieved the best performance while, conversely, the SDR image in $L^*a^*b^*$ was the worst performer by noticeable margin. This suggests that adopting a HDR protocol is especially relevant when operating on images in the $L^*a^*b^*$ space.

The performance order of the colour spaces might be somewhat expected given the visual appearance of the HDR image in each colour space (Figure 4.12(a)-4.12(c)). It may be noted that damaged regions in $L^*a^*b^*$ appear relatively homogenous and are readily discernible against the background. The damaged regions in the HSV colour space on the other hand are composed of several colours making object detection more difficult. The RGB space is slightly more effective than the other spaces at locating the presence of damage while the HSV and $L^*a^*b^*$ spaces perform well at defining the shape and size of damaged regions.

The success of REMPS is influenced to varying extents by the performance of each phase in a given colour space. While there is a heavy reliance on the ability of the Sobel edge detection phase to isolate damaged regions, it is the clustering stage which has the greatest influence. This stage determines whether clusters should be retained or discarded so it can have a significant impact on both the misclassification and detection rate. Conversely, the SVM stage has a relatively minor effect on the detection accuracy as it is largely confined to a role as a supplementary tool to enhance already detected regions.

Thus, colour spaces that do not respond well to the edge detection stage and particularly the clustering stage are greatly handicapped.

4.3.4 Comparison with Traditional Colour based Techniques

In this section, REMPS is compared with established detection techniques, such as Otsu's Method (Otsu, 1979), Chan-Vese Method (Chan and Vese, 2001), Delaunay Triangulation (Cheddad et al., 2008), Region Growing (Adams and Bischof, 1994), and Graph-Based Segmentation (Felzenszwalb and Huttenlocher, 2004). The comparison serves to highlight the effectiveness of REMPS in relation to the other segmentation techniques. The regions detected using these techniques on the HDR image are shown in Figure 4.14, and their respective performances are quantified in Table 4.3 for both the SDR and HDR images, as well as being graphically illustrated by means of performance points in the ROC space in Figure 4.15.

The performance of colour based segmentation techniques is affected by whether the technique is contextual or non-contextual. Non-contextual techniques (e.g. thresholding) do not take into account any spatial relationships between pixels in an image, but rather segment pixels at a global level on the basis of some attribute, e.g. colour intensity. Contextual techniques (e.g. REMPS or region growing techniques) on the other hand do consider spatial relationships. If a contextual relationship is an important factor for segmenting a particular image, than non-contextual techniques will have limited success compared to techniques which exploit the contextual relationship. Adopting the α-δ method allowed for a clear comparison between various (DR,MCR) pairs. Analysis of the δ parameter in Table 4.3 reveals that all of the established techniques performed noticeably better when applied to the SDR image rather than the HDR image. In some cases, such as

for the Region Growing technique (Figure 4.14(d)), the SDR image offered an appreciably improved performance suggesting that the increased local contrast associated with HDR has an adverse effect. This may be observed from the ROC space in Figure 4.15 which illustrates the relatively separate nature of the two performance points associated with the technique. However, an exception to this trend emerged in the case of REMPS, whereby the performance was slightly enhanced when HDR imagery was considered as an imaging protocol.

The performance levels obtained from each technique varied markedly. The Chan-Vese method (Figure 4.14(b)) and Delaunay Triangulation (Figure 4.14(c)) performed quite well when applied to the SDR image, whilst Otsu's Method (Figure 4.14(e)) performed reasonably well on both the SDR and HDR images despite its simple and non-contextual nature. The Graph Cutting technique (Figure 4.14(f)) on the other hand produced poor results, as demonstrated by the performance points in the ROC space lying closer to the line of chance than the best performance point (Figure 4.15). Overall, the REMPS achieved the best detection results, especially when performed on the HDR image.

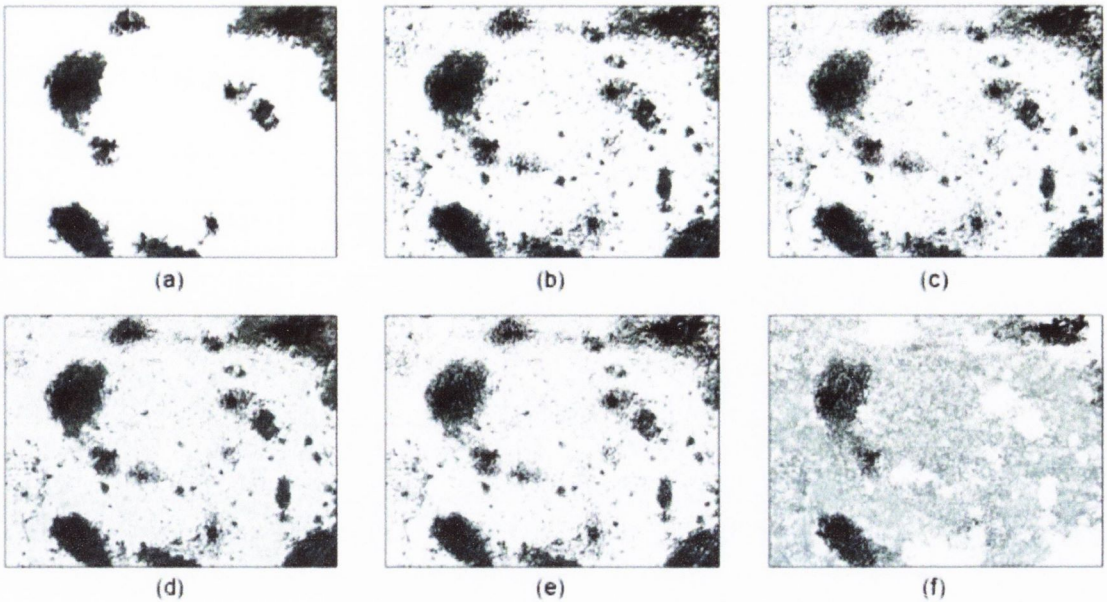


Figure 4.14 Detected regions from: (a) REMPS, (b) Chan-Vese Method, (c) Delaunay Triangulation, (d) Region Growing, (e) Otsu's Method, (f) Graph-Based Segmentation.

Table 4-3 Comparison of techniques.

Segmentation Technique	Normally Exposed Image			HDR Image		
	DR	***MCR***	***δ***	***DR***	***MCR***	***δ***
REMPS	84%	8%	0.18	85%	7%	0.17
Chen-Vese	93%	18%	0.19	95%	22%	0.23
Delaunay Triangulation	86%	14%	0.20	85%	14%	0.20
Region Growing	89%	17%	0.21	95%	35%	0.36
Otsu's Method	86%	14%	0.20	90%	20%	0.22
Graph Cutting	78%	39%	0.45	76%	41%	0.47

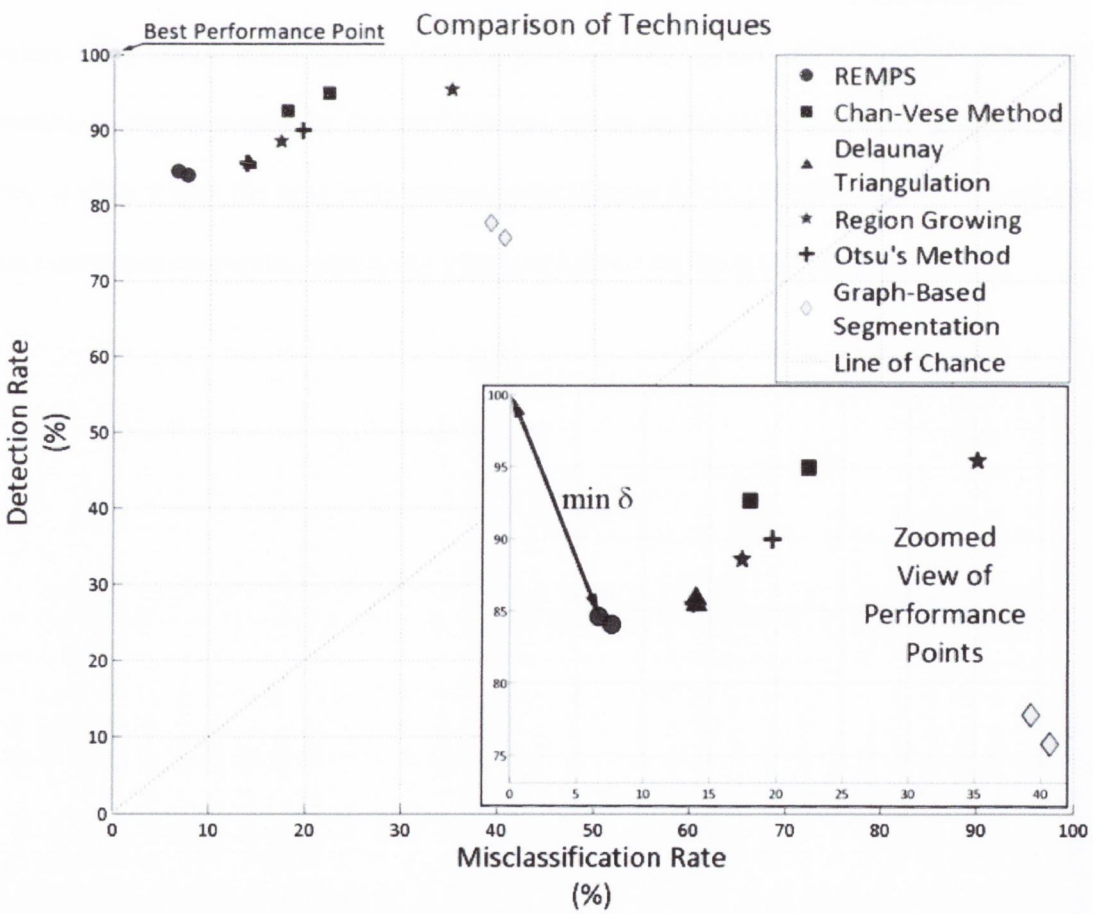


Figure 4.15 Comparison of detection techniques in the ROC space.

REMPS was also applied to a standard image in a non-structural scene, significantly disparate from the corroded example presented in this section, to showcase its credentials as a standalone segmentation technique. A visual comparison with some of the detection techniques previously mentioned is presented in Figure 4.16, which further illustrates the potential of REMPS and underlines its credentials as a high performing standalone technique beyond damage identification applications.

Figure 4.16 (a) Original Image, Detected Regions from: (b) REMPS, (c) Chan-Vese Method, (d) Delaunay Triangulation, (e) Region Growing, and (f) Otsu's Method.

It may be observed from Figure 4.14 and Figure 4.16 that the detected regions from REMPS produce a much 'cleaner' image of detected regions that is not contaminated by speckles of spurious regions which is a feature of all the other techniques. Having a 'cleaner' image is important for many post-processing applications such as calculating the propagation rate for damaged regions. For such an application, labelling and numbering of damaged regions may be a necessary prerequisite which would be inhibited by the presence of the many small and insignificant spurious regions.

4.3.5 Comparison with a Texture Analysis based Technique for Damage Detection

Image processing based techniques include colour intensity based methods and texture analysis based methods. Naturally, the techniques in each group are suited to different applications, depending largely on whether the damaged regions are more separable from the background based on colour or on texture. This section assesses the performance of REMPS alongside a texture analysis based technique (O'Byrne et al., 2012) previously proposed in the domain of NDT, in order to give an indication of the performance levels that can be expected when a range of damage forms and surfaces are under consideration. This should enable the end user to better decide on which approach is most appropriate to their needs.

Both methods are applied to four different images shown which feature various damage forms, lighting conditions, viewing angles, resolutions etc. These images are shown in Figure 4.17, along with the regions detected using REMPS and texture analysis approach.

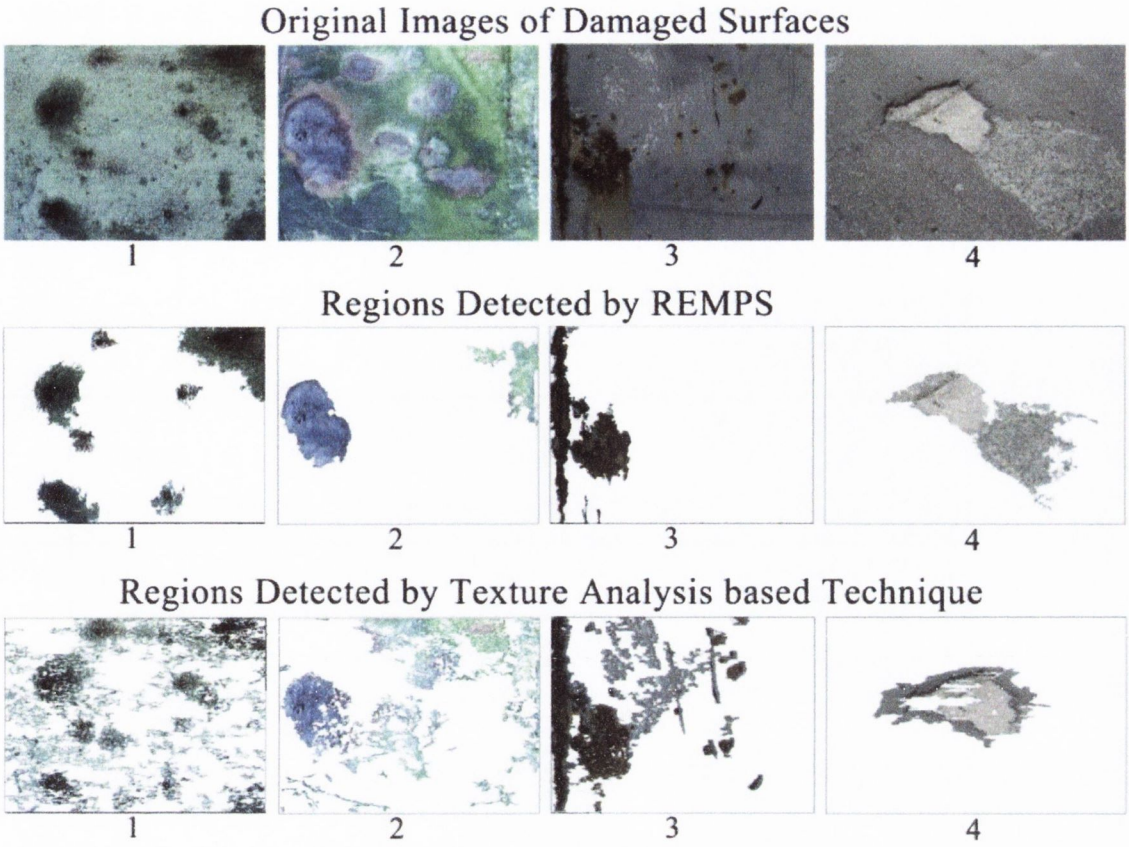


Figure 4.17 Top row: original images featuring damage: (1) pitting corrosion, (2) marine growth, (2) metal sheet with various corrosion damage, and (4) road delamination. Middle row: regions detected using REMPS. Bottom row: regions detected using texture analysis.

The performances for each damage detection technique are presented in Table 4.4.

Table 4-4 Comparison of REMPS technique and a texture analysis technique.

Sample Image	REMPS			Texture Analysis		
	DR	*MCR*	δ	*DR*	*MCR*	δ
(1) Pitting Corrosion	84%	8%	0.18	78%	32%	0.39
(2) Marine Growth	64%	8%	0.37	64%	29%	0.46
(3) Metal Sheet with various Corrosion Damage	90%	3%	0.11	96%	24%	0.24
(4) Road Delamination	93%	10%	0.12	52%	10%	0.49

It may be noted from these results that REMPS was quite successful for the majority of cases with the exception of the marine growth image. The poorer detection results for this image may be explained by the fact that the damaged regions throughout the image were not characterised by one single colour. Instead they took on numerous contrasting shades which often overlapped with the non-damaged background. Generally however, REMPS proved effective at locating the presence of damage as well as accurately defining the shape and size of damaged regions.

The texture based method was effective at locating the presence of damage as may be observed from Figure 4.17, however it did not perform as well as REMPS at defining the extent of damage which resulted in poor *DR*, *MCR* and δ values in Table 4.4. Many small spurious regions were detected unlike REMPS which produced a 'cleaner' and more homogenous detection.

These results demonstrate the applicability of REMPS for a wide range of damage forms, and show that it offers an improvement over the texture analysis based damage detection approach for the presented scenarios.

4.3.6 REMPS Discussion and Conclusion

Various forms of NDT techniques have been employed to assess civil infrastructure since the advent of SHM, however it is only with the relatively recent introduction of computer based systems that quantitative information on the health status of structural components can be obtained on-site. There is thus an emphasis on devising sophisticated damage detection techniques that can effectively capitalise on the ever increasing level of computational efficiency. This section has presented an image analysis based damage detection technique, REMPS, which is intended to supplement and strengthen existing visual inspection methods by providing a quick and convenient source of quantitative information. The development of REMPS was necessitated by a lack of sophisticated image based damage detection techniques that can be applied to a broad range of surface types, damage forms, and lighting conditions that are typically encountered in infrastructures. The specific application presented in this section demonstrates the immediate success of the method as an NDT tool to assist visual inspections where an improved detection directly influences the owner of infrastructure systems during a decision-making process.

REMPS adopts a multi-phase segmentation methodology which incorporates features from three standard image processing and data analysis techniques. Since these techniques are well-known and described in the literature, REMPS may be easily replicated and implemented. A key benefit of REMPS is its ability to produce better defined and more

homogenous regions of interest without being affected by isolated extraneous pixels. REMPS achieves this cleaner segmentation by efficiently integrating pixel and spatial relationships. The presented results indicate that improvements can be made to the detection accuracy by segmenting in the $L^*a^*b^*$ colour space and adopting a HDR protocol. Furthermore, the credentials of REMPS as a standalone segmentation technique are underlined as it is shown that REMPS outperforms several established detection techniques for various scenes.

4.4 Stereo Matching Algorithm for 3D Shape Recovery

Recovering shape information is a challenging but useful task, and has wide applicability in many areas of Structural Health Monitoring (SHM). For underwater inspections, one such example concerns the tracking of marine growth thickness on offshore structures, which is undesirable as it increases the hydrodynamic forces acting on the host structure (Boukinda Mbadinga et al., 2007). The shape recovery technique discussed in this section is based on stereo-vision, or stereo for short, which employs a dual camera set-up to simultaneously photograph a specimen of interest from slightly different viewpoints. 3D information can then be extracted by examining of the relative positions of objects captured by each camera. An overview of the stereo process is shown in Figure 4.18.

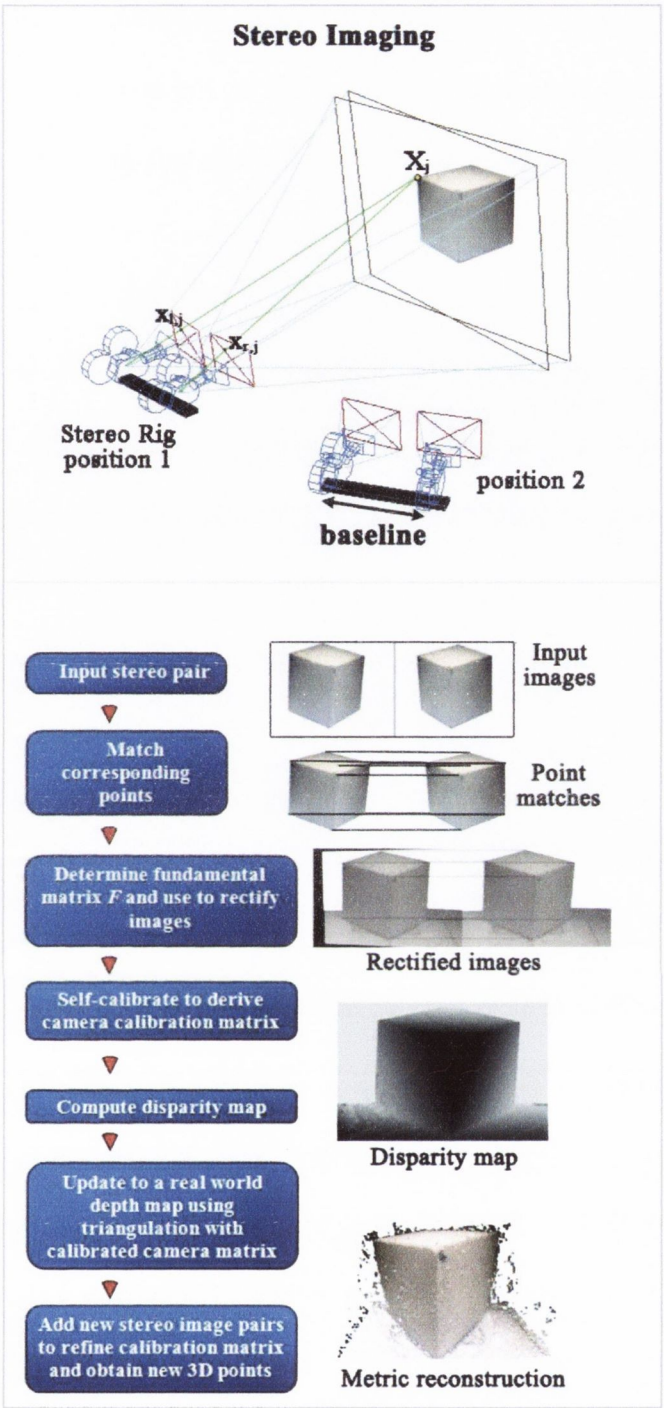


Figure 4.18 Shape recovery using stereo imaging flowchart.

A key part of the stereo based shape recovery process is the matching phase, which seeks to locate the same object in the left and right camera images. Although a wide range of stereo matching algorithms have been proposed to solve this problem, these are often designed with specific applications and operating requirements in mind such as real time

processing. The development of a new stereo matching algorithm is motivated by the need for a technique that can effectively deal with the poor visibility conditions encountered in an underwater setting and efficiently exploit the high resolution of images to provide good depth sensitivity.

The challenging underwater conditions attenuate the contrast of the subject and can introduce artefacts such as floating particulate, both of which increase the likelihood of unreliable or false matches. To combat this, the PaLPaBEL (Pyramidal Loopy Propagated BELief) stereo correspondence algorithm technique was developed. PaLPaBEL is based on a Markov Random Field (MRF). The MRF provides smoothness - an important factor in a noisy underwater environment, where matching exclusively based on colour information cannot be relied upon. Adopting a pyramidal scheme is not only efficient, but it also enables MRFs to be used that can consider a wide disparity range without having to downsize the image dimensions so as to avoid exceeding the computer's memory limitations. This in turn enables accurate and precise depth information that can fully utilise the original image resolution.

A brief overview of the relevant terminology is provided in the section. The methodology of the proposed stereo approach is outlined in Section 4.4.2. The technique is demonstrated on a synthetic scene as well as a real world specimen photographed underwater.

4.4.1 Stereo Terminology

This section summarises the main terminology used in the context of stereo imaging.

Disparity - The distance in pixels between two corresponding points in the left and right image of a stereo pair.

Depth Map/Disparity Map - By matching every pixel in the left image with its corresponding pixel in the right image and computing the relative distances between matched pixels (i.e. the disparity), we end up with an image, or map, which shows the disparity at every point in the image.

Rectification - Rectification of a stereo pair involves applying a transformation (warping) to each image such that corresponding points are separated by a horizontal offset only and not by a vertical one. Rectification makes the process of matching pixels in the left and right image considerably faster as the search will be confined to the horizontal direction only.

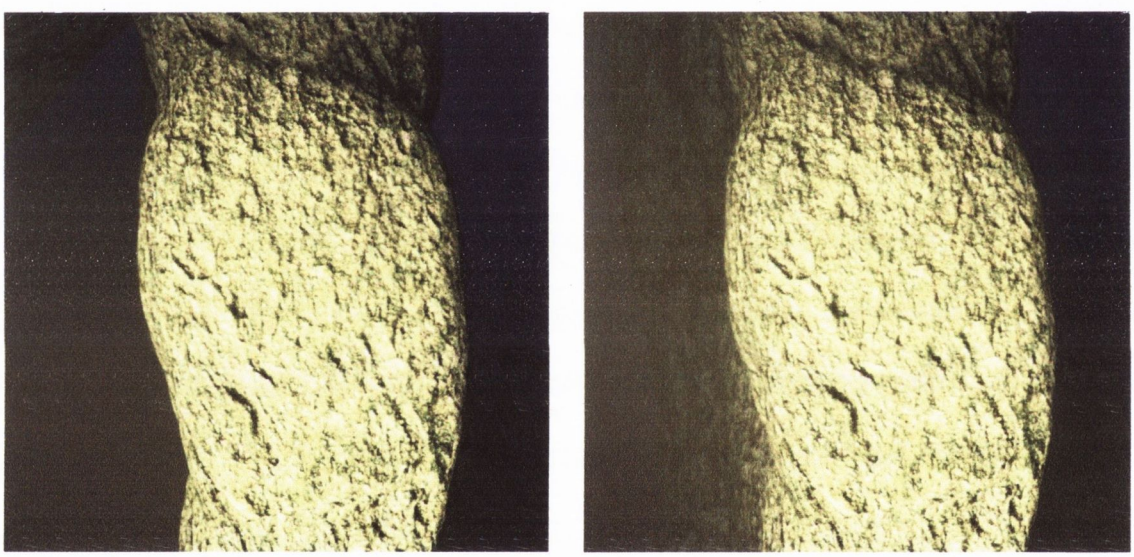
Occlusion - Occlusion occurs if a 3D point in the real world space is only depicted in one image and is hidden in the other image.

Fundamental Matrix - The fundamental matrix is a 3×3 matrix which relates corresponding points in stereo images. It may be estimated from at least seven point correspondences. Its seven parameters represent the only geometric information about cameras that can be obtained through point correspondences alone.

Baseline - The real world distance between the two camera centres (measured in millimetres).

4.4.2 Methodology

The methodology is outlined in five steps. These are: 1) Image pre-processing using histogram stretching, 2) rectification, 3) matching, 4) occlusion handling and false match elimination, and 5) Reconstruction. The synthetic stereo pair shown in Figure 4.19, which depicts a marine growth affected pile, is used to illustrate each of these steps. Synthetic data was used to imitate the textures and irregular shapes typical of marine growth.



Left Image Right Image

Figure 4.19 Original input images featuring synthetic marine growth.

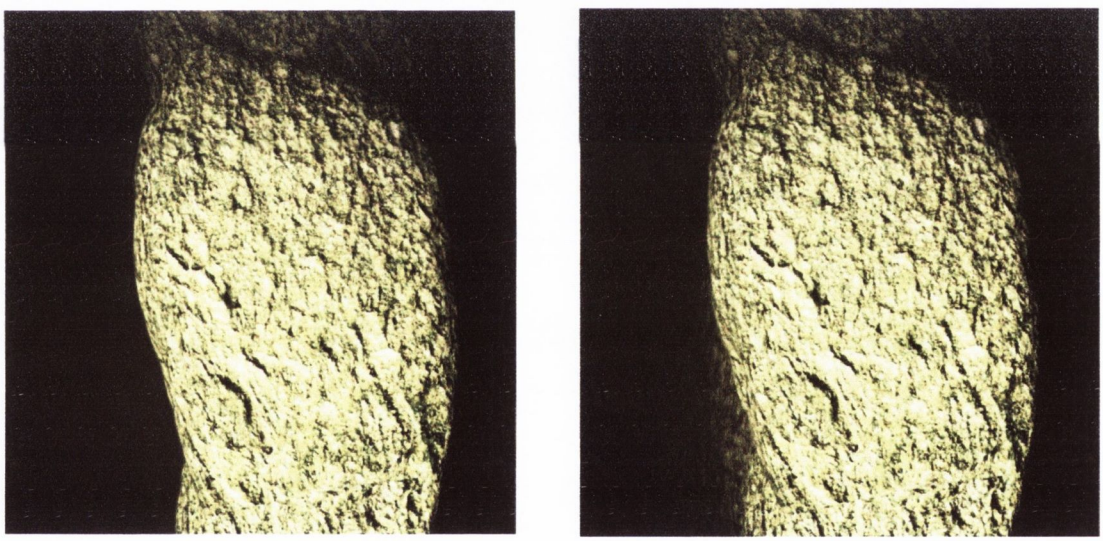
4.4.2.1 Image Pre-processing: Histogram Stretching in the $L^*a^*b^*$ Colour Space

This step is intended to improve the matching accuracy. The images are converted from the RGB colour space to the $L^*a^*b^*$ colour space in order to isolate the 'Lightness' component, L^*. Adjusting the L^* component re-maps the brightness producing greater contrast in a way that does not significantly affect the colour balance. This procedure is

done by stretching the values in the L^* channel at each pixel location so that they occupy the full range - from 0 to 1, as per:

$$L^*_{stretched} = \frac{L^*_{input} - \min(L^*_{input})}{\max(L^*_{input}) - \min(L^*_{input})} \quad \textbf{(4.17)}$$

where $L^*_{stretched}$ is the output lightness component, which replaces the original input lightness component, L^*_{input}, in the updated contrast enhanced image. The updated $L^*a^*b^*$ image is converted back into the RGB colour space once again. Histogram stretching is highly applicable to underwater images so as to restore some of the lost contrast. The resulting enhanced images are shown in Figure 4.20.



Left Image Right Image

Figure 4.20 Histogram Stretched Images.

4.4.2.2 Rectification using SIFT (Scale Invariant Feature Transform) Features

Rectification is the process of transforming the images in a stereo pair, such that corresponding points in each image are separated by a horizontal distance and not a vertical distance.

Rectification is achieved by locating and matching distinctive points in the left and right images. SIFT features are invariant to scaling and rotation of the image, and are resilient to illumination changes, the presence of noise, as well as minor changes in perspective. Additionally, the features are efficiently described, which leads to a low probability of false matches. The process of rectifying two images based on SIFT Features is described by Lowe (1999):

All points in the rectified images should satisfy the epipolar geometry of a rectified image pair (i.e. that the images are aligned horizontally). This may be expressed as: if a point *P1*, in the left image, corresponds to *P2*, a point in the right image, then:

$$\begin{bmatrix} P2 \\ 1 \end{bmatrix} F \begin{bmatrix} P1 \\ 1 \end{bmatrix} = 0 \qquad \textbf{(4.18)}$$

where F is the fundamental matrix. Figure 4.21(a and b) show the unrectified and rectified images respectively,

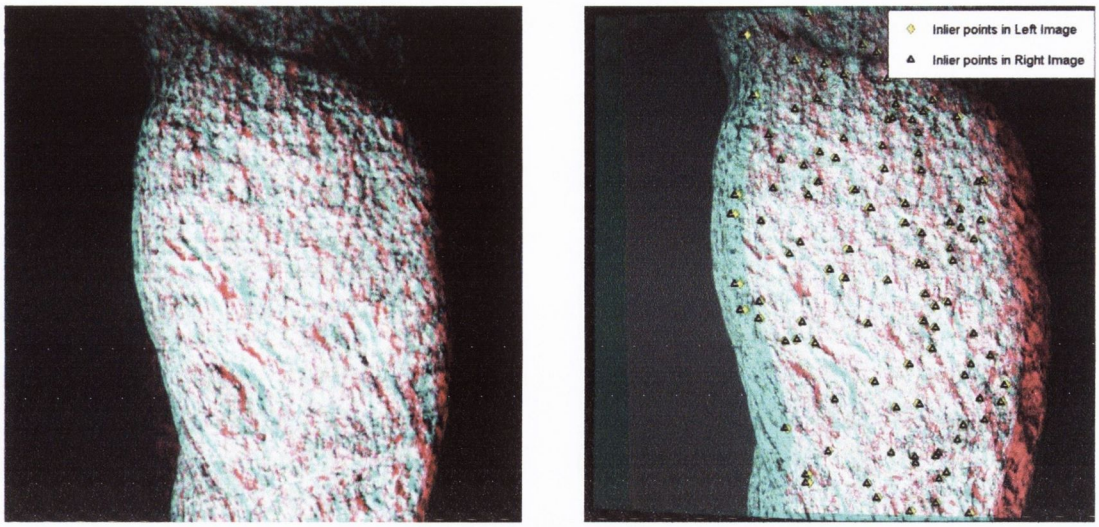


Figure 4.21 (a) Unrectified Images (blue = left image, red = right image), (b) Rectified image where the keypoints are now horizontally aligned.

4.4.2.3 Matching

This section details the PaLPaBEL technique. Using an approach such as loopy Belief Propagation (BP) applied to Markov Random Fields, as described by Ho (2012), typically offers an improvement over naive block matching as it promotes a smooth and continuous depth map, which is often more reflective of the scene under consideration. Attempting to recover the depth map through basic block matching, which selects the best match based only on colour information whilst ignoring context/spatial information, is prone to producing inconsistent and inaccurate depth maps. In reality, it is typically the case that pixels in close proximity to each other share similar disparity values, unless the pixels represent an edge boundary in the scene.

The first step is to measure the similarity of points in the left and right images of the stereo pair. There are numerous ways of calculating a measure of similarity, such as Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD), Normalised Cross

Correlation (NCC) and Zero-mean Normalised Cross Correlation (ZNCC). The described technique employs ZNCC as a measure of similarity as it provides better robustness than NCC since it tolerates uniform brightness variations thanks to the subtraction of the local mean, albeit at the expense of increased computation time. Calculating the ZNCC involves a sliding window, which moves horizontally over the right image within the predefined disparity range. The contents of the window at every disparity value are compared with the reference window in the left image. The zero mean normal cross correlation coefficient, ς, has a high value when there is a high degree of similarity between two sets. It is defined as:

$$\varsigma = \frac{\sum_{u,v \in SW} (DS(u,z)_{\Delta} - \bar{DS}_{\Delta})(CG\ (u+\Delta,z) - \bar{CG}\ )}{\sqrt{\sum_{u,v \in SW} (DS(u,z)_{\Delta} - \bar{DS}_{\Delta})^2 (CG\ (u+\Delta,z) - \bar{CG}\ )^2}} \tag{4.20}$$

where $CG(u,z)$ denotes the intensity value in the reference window (left image), $DS(u,z)$ denotes the intensity value in the sliding window, SW, in the right image, u and z are the horizontal and vertical spatial indices of the of SW, Δ is the disparity distance at which the template has been shifted, and $\bar{DS}$ and $\bar{CG}$ denote the mean value of the destination image within the area of the template, and in the mean value in the reference window respectively.

(a) MRF Formulation

MRFs are undirected graphical models that have application in many image processing tasks (Kinderman and Snell, 1980). For stereo matching purposes, MRFs enable spatial dependencies between nearby pixels to be encoded. MRFs are comprised of nodes and

links and may be cyclic, meaning there is no guarantee of convergence regardless of the number of iterations. An example of how MRFs can be used to model the stereo problem is shown in Figure 4.22 for a 3×3 image.

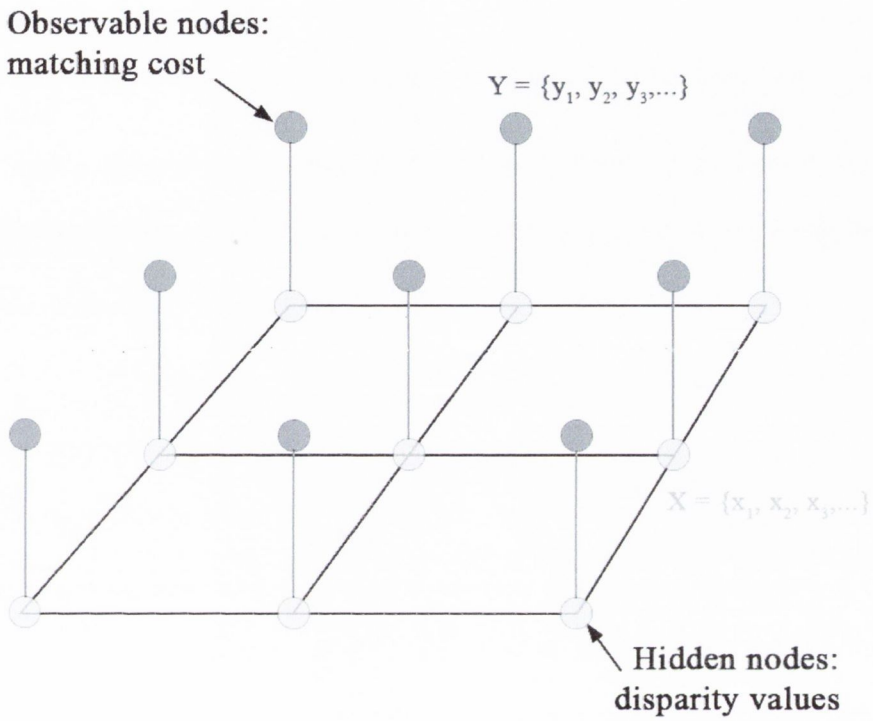


Figure 4.22 MRF for a 3 x 3 image.

A dark node represents an observed variable, which is the matching cost associated with a particular disparity value. A light node represents a hidden variable, which directly refers to the disparity value. Disparity values are widely referred to as 'labels' in this context.

The links between each node in Figure 4.22 denote a dependency. The value of a hidden node will depend on the four immediately adjacent hidden nodes, along with the observable node. The traditional formulation for the stereo problem in terms of an MRF is given by the following energy function:

$$E(Y,X) = \sum_{f} MC(y_f, x_f) + \sum_{\Upsilon=\text{neighbours of } f} SC(x_f, x_\Upsilon) \quad \textbf{(4.19)}$$

where the variables Y and X are the observed and hidden node respectively, E is the energy or total cost, f is the pixel index, Υ are the neighbouring nodes of node x_f. MC is the matching cost associated with assigning a label value of x_f to data y_f, and SC is the smoothness cost, which promotes good agreement between neighbouring pixels. The energy function essentially sums up the cost at each point given a matching cost Y and some labelling X. The objective is to find a labelling for X that produces the lowest energy.

The smoothness cost promotes consistent and steady labelling across neighbouring hidden nodes. It penalises cases where adjacent labels are markedly different from each other. The smoothness cost increases as the label difference increases, as per:

$$SC = \min(\rho - |f - \Upsilon|, |f - \Upsilon|) \quad \textbf{(4.21)}$$

where ρ is the number of labels, i.e. the number of possible integer disparity values that can be assigned to a node. This function means that the smoothness cost increases as the difference between adjacent labels also increases, up to a point where the difference between adjacent nodes becomes half the entire label range, at which point the smoothness cost begins to decrease once again. This decrease may seem counter-intuitive at first, but it is necessary for the pyramidal stage in order to facilitate transitioning from nodes that have the same absolute disparity value. This issue arises when the course disparities computed at a previous pyramid level are refined in the subsequent level. The finer disparities in the new level begin with a disparity value based on the coarser disparity value computed at the previous pyramid level. This smoothness cost function allows seamless interaction and

smooth overlapping of disparity values for neighbouring pixels, even if these pixels were represented by different disparities at lower pyramid levels. This is illustrated on a small sample stereo image pair in Figure 4.23.

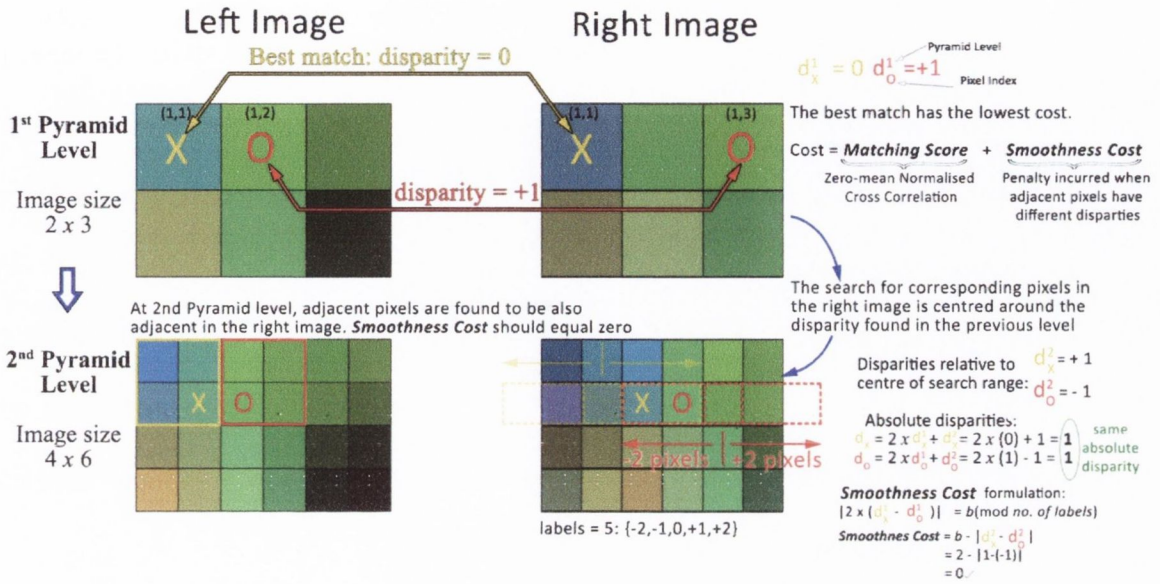


Figure 4.23 Example of smoothness cost formulation.

(b) Loopy Belief Propagation (BP)

The loopy BP algorithm is a message-passing algorithm that operates on graphical models such as MRFs. The algorithm passes a message from one node to an adjacent node after it has received messages from all of the other connected adjacent nodes. It is worth pointing out that there is no return message from the receiving nodes back to the message dispensing nodes. Formally, this message passing is defined as:

$$msg_{f \longrightarrow \Upsilon}(p) = \min_{p' \in \text{all labels}} \left[\sum_{k=\text{neighbours of } f \text{ except } \Upsilon} MC(y_f, p) + SC(p, p') + msg_{k \longrightarrow f}(p') \right] \tag{4.22}$$

Essentially, this states that a node f sends a message to node Y about what it believes the most suitable label b should be. It is node f's belief about node Y regarding label b. The messages are stored in a matrix of size M_x *x* M_y *x 5 x No. of labels*, where M_x and M_y are the image dimensions and 5 corresponds to the 4 directions from which the message is incoming plus the outgoing message (from x_1 to x_2). For high resolution imagery and large disparities, the size of this matrix would become prohibitively large for most computers to handle. For example, an image of size 2250 x 2250, and having a representative disparity range of 400 pixels, would create an array of 2250 x 2250 x 5 x 400, which would require more than 10 billion elements. This demonstrates just why it would be infeasible to employ the traditional MRF approach without adopting a pyramiding scheme.

(c) Incorporation into a pyramidal scheme

The primary challenge of using MRFs as part of a pyramidal scheme is accommodating high disparity values alongside small disparity values with access to only a limited number of labels. This proposed approach achieves this by disregarding the absolute disparity values, and only concentrating on the disparity relative to the depth computed at the preceding pyramid level. The pyramidal structure is illustrated in Figure 4.24, which shows the image dimensions at each level, starting at a coarse, poor resolution, and finishing at the top with a fine, large resolution.

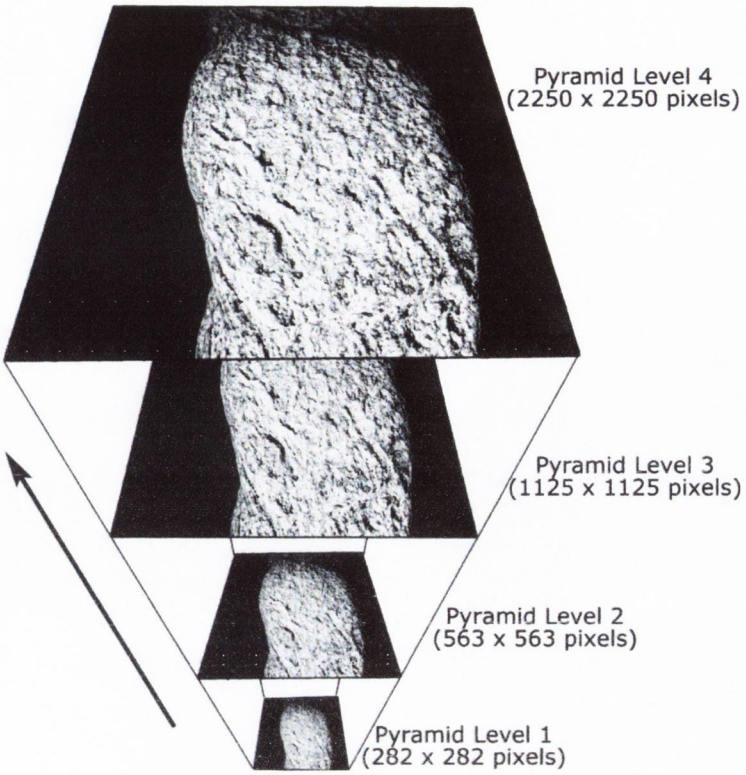


Figure 4.24 Pyramid levels.

The process is described as follows:

1) The process starts at the first pyramid level (i.e. the most down-sampled image, in the case of illustrated image it is of size 282 x 282 pixels). The disparity range at this level is determined by scaling down the disparity range for the full size image by the appropriate factor, in this case, by a factor of 8 (= $2^{\text{No of Pyramid Levels - 1}}$). The loopy BP algorithm is applied to this downsized image, which results in the output smooth depth map image as shown in Figure 4.25.

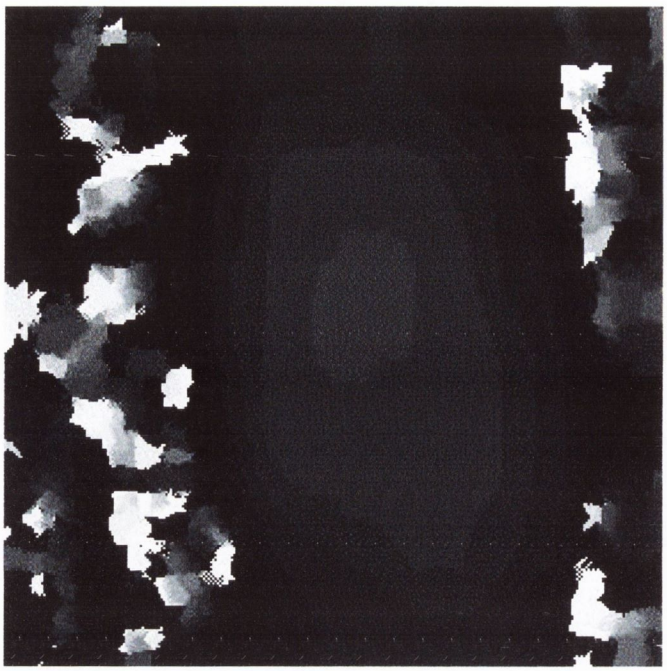


Figure 4.25 Depth map output from the first pyramid level.

It may be observed that this depth map is quite coarse and could significantly benefit from further refinement. Working towards this objective, the depth map is resized to the size of the image at the next pyramid level (pyramid level 2, image size = 563 x 563 pixels). This depth map is then used as input for next level.

2) Pyramid level 2: The disparity search commences at the second pyramid level. The starting point for the refined disparity search begins at the depth computed from the previous pyramid level at each pixel location. So for example, if a given pixel in the image was assigned a disparity of -4 from the 1st pyramid level than the correspondence search at the 2nd level is centred at -8 (bearing in mind that when the increase in the size of the image is taken into account, the disparity values scale accordingly, so -4 at the first pyramid level translates to -4*2 = -8 at the 2nd pyramid level). A search then takes place over the range -8 ± (No. of labels)/2. The progressive refinement of the depth map is

shown with the aid of a diagram in Figure 4.26, where it may be noted that more depth detail is captured in later levels.

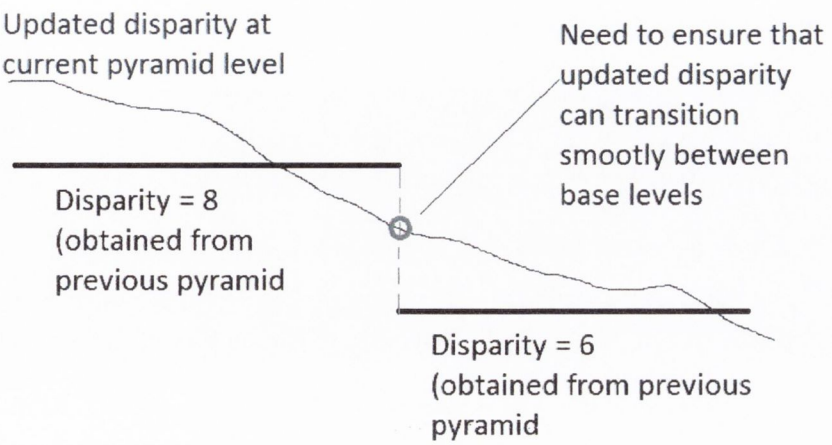


Figure 4.26 Ensuring continuity in the updated depth map.

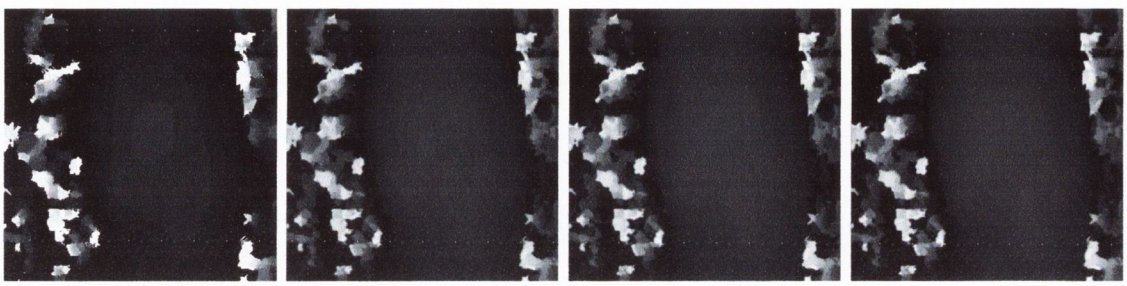
Continuous overlapping is achieved using a modulus based indexing structure. The array that stores the messages is updated at each iteration of the belief propagation as follows:

$$msg_{f\to\Upsilon}(\dot{p}) = \min_{\dot{p}' \in \text{all labels}} \left[\sum_{\phi=\text{neighbours of } f \text{ except } \Upsilon} MC(y_f, \dot{p}) + SC(\dot{p}, \dot{p}') + msg_{\phi\to f}(\dot{p}') \right] \tag{4.23}$$

In this chapter, a value of 20 iterations was chosen. It was found that going beyond this value was unnecessary as there is no further significant change in energy or cost. After 20 iterations, the final cost is available, which is calculated from:

$$Cost(x_f = \dot{p}) = MC(y_f, \dot{p}) + \sum_{\phi=\text{neighbours of } x_f} msg_{\phi\to f}(\dot{p}) \tag{4.24}$$

The disparity associated with the lowest cost is assigned to the pixel under consideration. The BP procedure is repeated at each pyramid level. The output at each pyramid level is shown in Figure 4.27.



Level 1 Level 2 Level 3 Level 4

Figure 4.27 Depth map output from PaLPaBEL at each pyramid level.

It may be noted that the depth map gets progressively refined at each successive level. This is also conveyed in Figure 4.28 which plots the disparities along a profile line for each pyramid level.

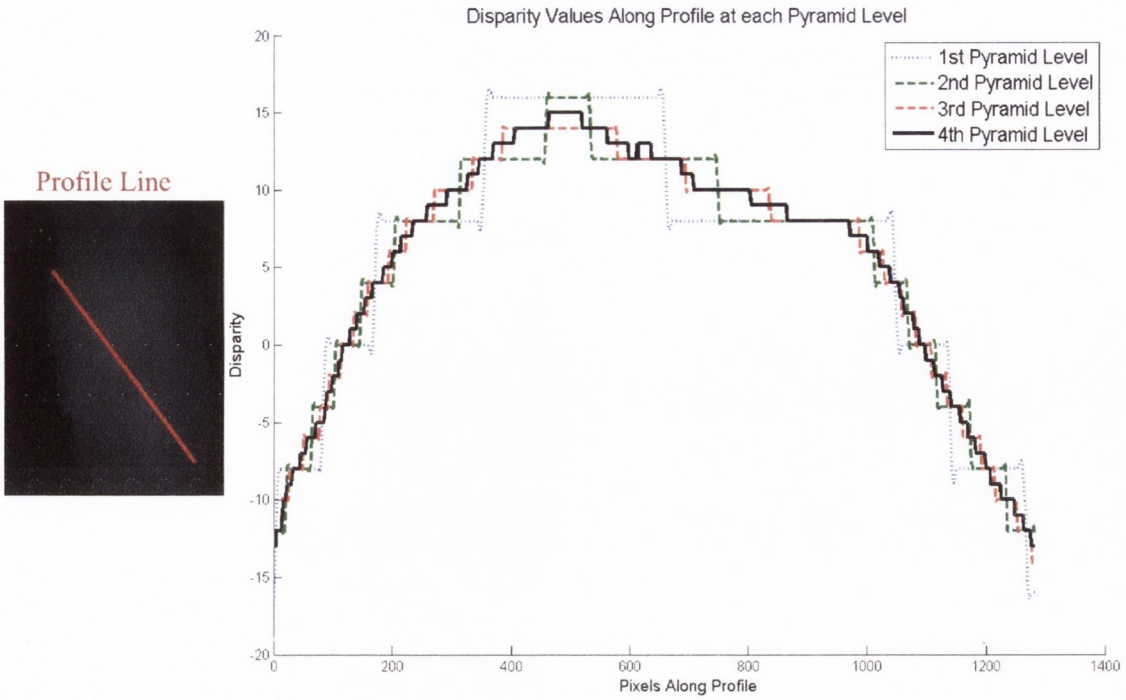


Figure 4.28 Disparity profile for varying pyramid levels.

It may be observed that the first pyramidal level produces a coarse profile in which the disparity values can assume only one of five discrete values. At each successive level, the disparities are refined, so that by the 4^{th} pyramidal level, the disparity can assume integer values in the whole range; from the minimum disparity to the maximum disparity.

4.4.2.4 Occlusion Handling and False Match Elimination

A notable feature of the resulting depth maps is that occluded/poorly matched regions are characterised by high fluctuating disparity values, i.e. there are a high frequency of small and large (black and white) values in close proximity. For this reason, a standard deviation filter is applied to the depth map to identify and discard these regions. This involves convoluting the depth map image with a window of the same size as the window used at the final pyramid level during the correspondence stage and then calculating the standard deviation of all values within that window for each pixel location. The standard deviation computed from each window is then assigned to the pixel at which the window was

centred upon. Areas of high standard deviation, represented as white in Figure 4.29, are regarded as the background or poorly matched areas and are discarded.

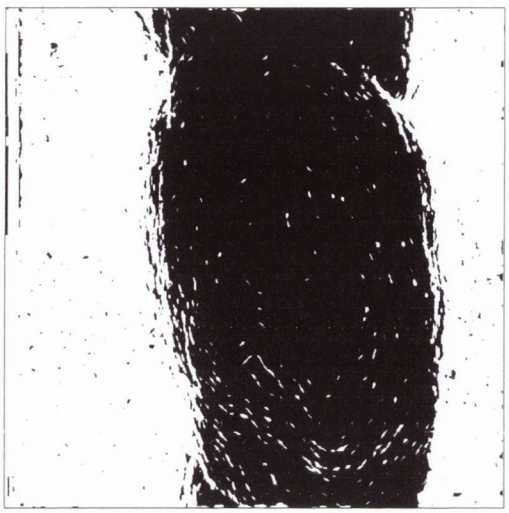


Figure 4.29 Standard Deviation Filter (Threshold = 1).

4.4.2.5 Reconstruction

With knowledge of the camera intrinsic parameters and the fundamental matrix we could reconstruct the scene points up to a scale factor (a Euclidean reconstruction). For this, the cameras must be calibrated. The calibration must be done under the same conditions as that of the image acquisition stage. Self-calibration, which is also referred to as auto-calibration, is an attractive way of determining the intrinsic and extrinsic camera parameters. Self-calibration refers to the process of obtaining a calibrated camera matrix using the constraints in the scene. The stereo based shape recovery technique in this section adopts a self-calibration procedure as outlined in the work of Zhang et al. (1994). Views of the reconstructed synthetic specimen are shown in Figure 4.30.

Figure 4.30 Projective reconstruction.

One of the primary motivations behind adopting image processing techniques is to provide a high degree of automation. The process of image based 3D shape recovery should strive towards being a 'click-and-go' operation with minimal human supervision. With this in mind, there should be as few input parameters as possible and these should not require significant attention or drastically affect performance. The input parameters for the proposed technique are: the number of pyramid levels, the number of iterations for the loopy belief propagation algorithm, the disparity search range and the window size for matching. The default number of pyramid levels is taken as 3. This provides a good balance between accuracy and memory management/computational efficiency.

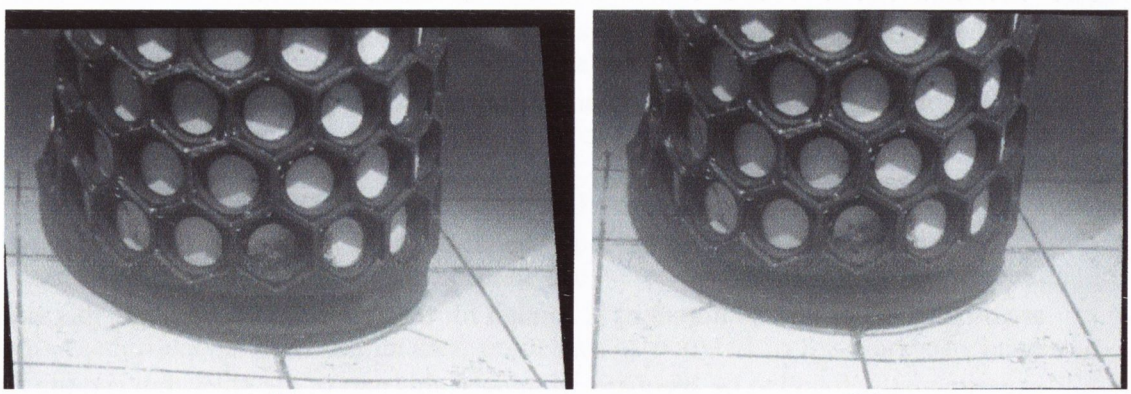
The number of iterations for the loopy belief propagation is taken as 20. Although there are no guarantees that 20 iterations is enough for the algorithm to converge, this value was experimentally found to be adequate. Moreover, this part of the algorithm is quite fast so there is not much computational expense incurred by extending beyond 20 iterations.

The size of the window used for matching should be carefully chosen. A balance must be struck between being large enough to include enough intensity variation to be sufficiently distinctive - crucial for reliable matching, but small enough to remain unique to the point in the image at which the window is centred upon. Generally, a window size of 7 pixels provides good results.

Finally, the stereo matching algorithm requires a disparity range to be specified. This range is generally over-estimated so the search considers a wider range of candidate disparities than what is really needed. This has the effect of using more computation and memory; however, there is little adverse effect on the matching accuracy.

4.4.3 Demonstration on Real World Specimen

The technique is applied to a real world specimen in an underwater environment. The specimen is a rubber mat wrapped around a cylinder, which has intricate depth variations - a feature that is reflective of what would be encountered in real-world scenes. This specimen is illustrated in Figure 4.31.



(a) **(b)**

Figure 4.31 (a) Left, and (b) right, rectified stereo images of a real world specimen in an underwater setting.

The depth map produced from PaLPaBEL is shown in Figure 4.32.

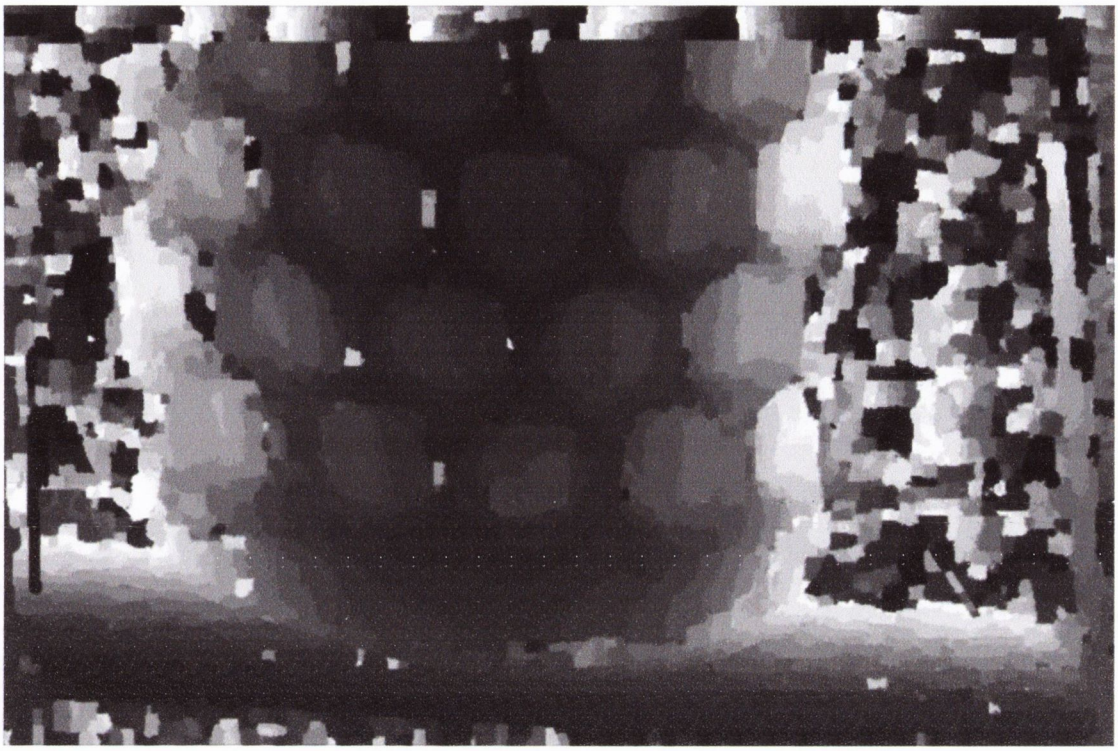


Figure 4.32 Depth map produced from the PaLPaBEL stereo correspondence algorithm.

Two views of the resulting reconstructed shape are shown in Figure 4.32.

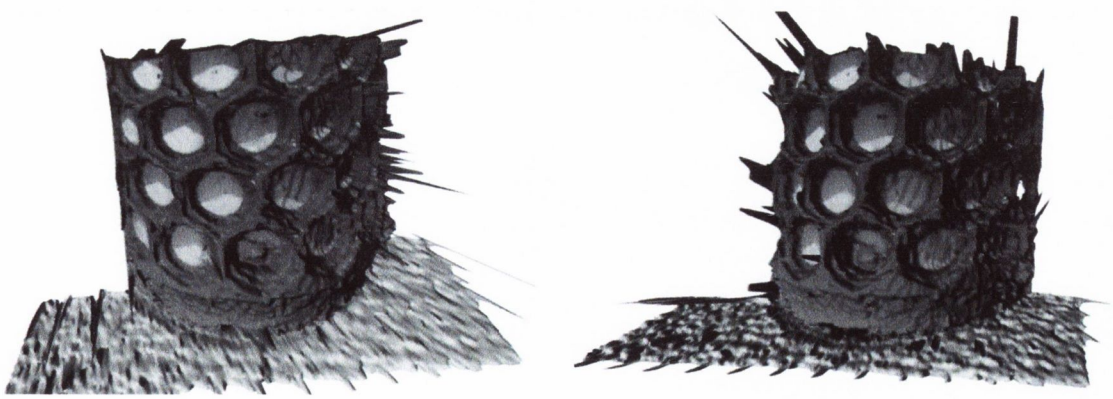


Figure 4.33 Two views of the reconstructed shape.

The results indicate that stereo can be a useful approach for capturing depth information from only a single stereo pair, and this acquired depth information can be easily visualised in the reconstructions.

4.4.4 Stereo Method Discussion

The development of a new stereo matching algorithm is motivated by the need for a technique that can effectively deal with the poor visibility conditions encountered in an underwater setting. To address this issue, a pyramidal Belief Propagation (BP) Markov Random Field (MRF) stereo correspondence algorithm, termed PaLPaBEL, has been described in this section. Adopting a pyramidal scheme is not only efficient, but it also enables MRFs to be used without exceeding the computer's memory limitations. Pyramiding also has an inherent smoothing effect as well as the smoothing derived from the MRF. Such smoothing is useful in noisy underwater conditions. This approach also enables accurate and precise depth information that can fully utilise the original image resolution.

4.5 Video Based Analysis of Vibration Measurements for a Suspension Bridge

4.5.1 Introduction

Image processing techniques can be extended for video analysis, making it possible to access spatial and temporal dimensions. This section describes a video tracking technique, which is demonstrated on a vibrating suspension bridge spanning a river in Cork, Ireland, for the purpose of identifying the natural frequency. The only equipment required is a

conventional digital video camera. The following sub-sections discuss: the capabilities and challenges of video tracking, the suspension bridge under consideration and the test set-up, and the methodology of the adopted video tracking technique. The results are presented and interpreted in the final sub-section.

4.5.2 Capabilities and Challenges

Video based tracking is capable of analysing bridge vibrations that are characterised by large amplitudes and low frequencies (i.e. typically below 20 Hz). The maximum frequency is controlled by the sampling rate, or the video frame rate. The digital video camera used in this case study was a Canon 600D, which records video at a rate of 60 Frames per Second (FPS) at a resolution of 1280 pixels x 720 pixels, or at 35 FPS at a resolution of 1920 pixels x 1080 pixels. The 35 FPS option was chosen as the need for the highest resolution was greatly felt in order to compensate for the extensive distance between the video device and the mid-span of the bridge. This extensive distance was due to the fact that the video camera had to be positioned relatively far away on the bank of the river as shown in Figure 4.34. By Nyquist's criterion, the highest frequency that can be coded is half of the sampling rate, which for a frame rate of 35 FPS is 17.5 Hz.

Figure 4.34 Daly's Suspension Bridge in Cork, Ireland, and the location of a mounted video camera on the river bank.

There are a number of challenges associated with tracking a moving point in a video sequence. Most notably, the tracked point may drift or become completely lost as a result of temporary occlusion or luminous complexities in the scene such as shadows or light reflections. For best results, a prominent and distinct point (i.e. a patch in the image with high local contrast) should be selected for tracking. If needed, multiple trials can be carried out by selecting various high-contrast regions near the mid-span of the bridge and assessing how well they are tracked over the duration of the video.

Another challenge relates to locating the position of the tracked region as precisely as possible. Even at the highest video resolution, the motion of the bridge translates to minute changes of only a few pixels in the video frames. In order to address this, a quadratic is fitted to surrounding values for sub-pixel and sub-scale interpolation as described by Brown and Lowe (2002).

4.5.3 Data acquisition

The bridge used as a case study is Daly's Bridge, known locally as 'The Shaky bridge'. Vibrations were induced in the bridge by a male and female pedestrian, along with a cyclist, all travelling at a typical pace. Their crossings were captured by the video camera, which was focused at the mid-span of the bridge. In the video, three points were chosen to be tracked as these represented distinct points (e.g. a corroded spot that was easily discernible from the background, region of discolouration etc.) as shown in Figure 4.35. This allowed for more robust and effective tracking from frame to frame.

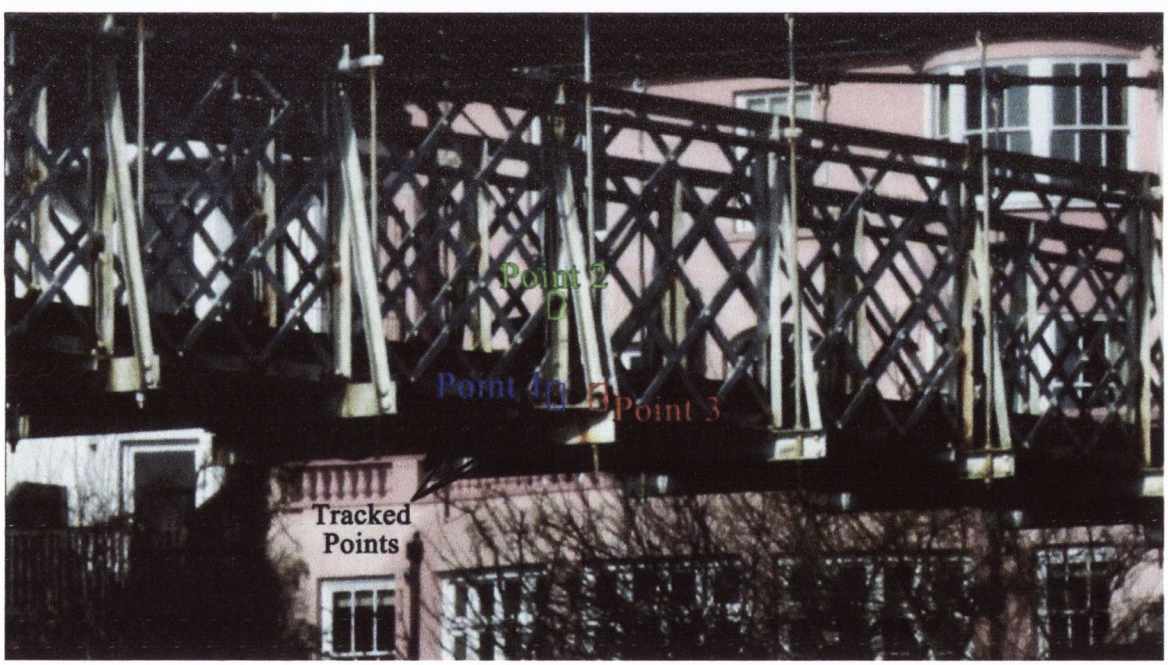


Figure 4.35 Tracked points at mid-span of the bridge.

Pixels units are related to metric units by a scale factor which was computed by measuring the length of an object in a video frame in pixels and then comparing it with the known real world length of the corresponding object in the scene. In this case study, the hanger length at the mid-span was used as the object. This approach is valid as long as the

bridge movement is roughly in the same plane as the measured object and the length of the object in the image in pixels is linearly related to its real length (i.e. there is negligible distortion such as perspective change between the top and bottom of the object). This study only considers bridge deformations in the more prominent vertical plane.

4.5.4 Video Tracking Technique

The video analysis operates by tracking the movement of points on the bridge whilst in an excited state. Tracking is done by picking a small patch, or window, *SW*, containing the chosen point in the first frame of a video sequence and following it throughout the duration of the video clip. For every successive frame, the point is located by finding the patch which best correlates with the patch in the previous frame. The best correlation was determined using the Zero-mean Normalised Cross Correlation (ZNCC) metric, which was introduced in Section 4.4 for the purpose of matching corresponding points in the left and right images of a stereo pair. For video tracking, the window centred on the tracked point in a frame Γ_n, is matched to the corresponding window in the next frame, Γ_{n+1}. The search space for locating the corresponding window is confined to a stationary region in the video, *SS*, which encloses the tracked point throughout the entire video. Confining the search space to the local neighbourhood allows for greater computational parsimony and minimises the risk of false matches. For this task, the ZNCC is defined as:

$$\psi = \frac{\sum_{u,z \in SW} (\Gamma_{\ell+1}(u,z)_{ab} - \bar{\Gamma}_{\ell+1,ab})(\Gamma_{\ell}(u,z) - \bar{\Gamma}_{\ell})}{\sqrt{\sum_{u,z \in SW} (\Gamma_{\ell+1}(u,z)_{ab} - \bar{\Gamma}_{\ell+1,ab})^2 (\Gamma_{\ell}(u,z) - \bar{\Gamma}_{\ell})^2}} \quad a,b \in SS \tag{4.25}$$

where $f_n(u,z)$ and $f_{n+1}(u,z)$ denote the intensity values in the window, *SW*, centred on the point being tracked in the n^{th} and $n+1^{th}$ frame respectively , u and z are the horizontal and vertical spatial indices of the of *SW*, a and b are the horizontal and vertical indices of the stationary region *SS*, and $\bar{\Gamma}_n$ and $\bar{\Gamma}_{n+1}$ are the mean pixel intensity values within the area of the *SW* for the n^{th} frame, and within *SS* for the $n+1^{th}$ frame. The zero mean normal cross correlation coefficient, ψ, has a high value when there is a high degree of similarity between two patches in each frame. The new location of the tracked point in frame $n+1$ is taken as the one which corresponds to the highest value of ς. This procedure is repeated until the final frame in the video sequence. The pixel locations for the tracked point are recorded at each frame, which are then plotted as displacement versus time. Spectral analysis on the signal is then performed in order to identify the major frequency components.

4.5.5 Results

This section presents a representative sample of the results obtained from video tracking. Three excitation sources are considered for the case study: i) pedestrian induced vibration - male, ii) pedestrian induced vibration - female (less weight than male), and iii) cyclist induced vibrations. Displacement versus time and Power Spectral Density (PSD) graphs corresponding to each scenario are shown in Figures 3.36-3.38. Each of the graphs relate to a time period of ~22 seconds, the centre point of the displacement versus time graph (i.e. 11 seconds) represents the point at which the pedestrian/cyclist is adjudged to be crossing the mid-span of the bridge.

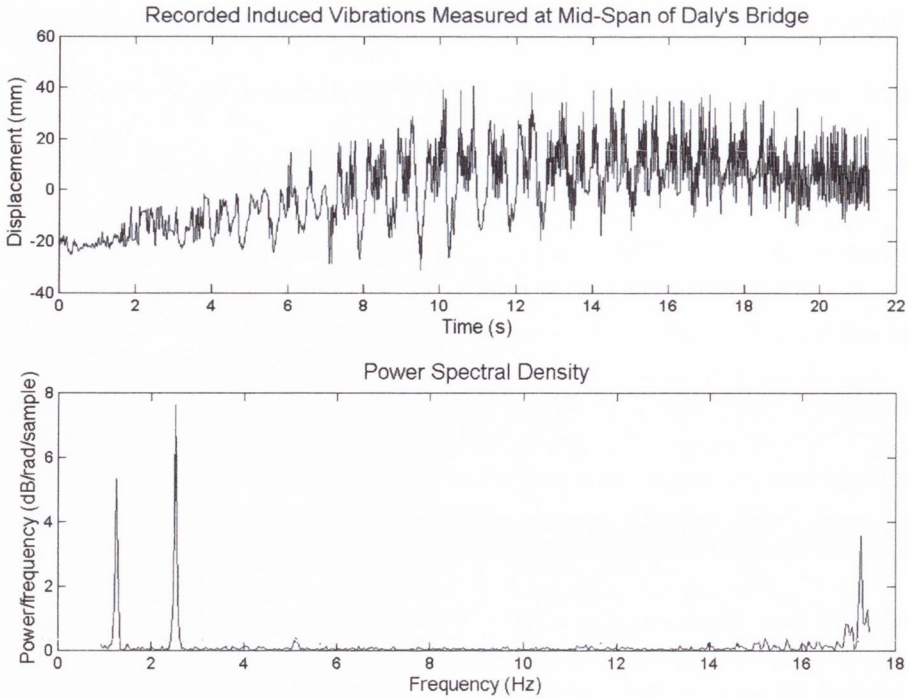


Figure 4.36 Pedestrian induced vibrations - male.

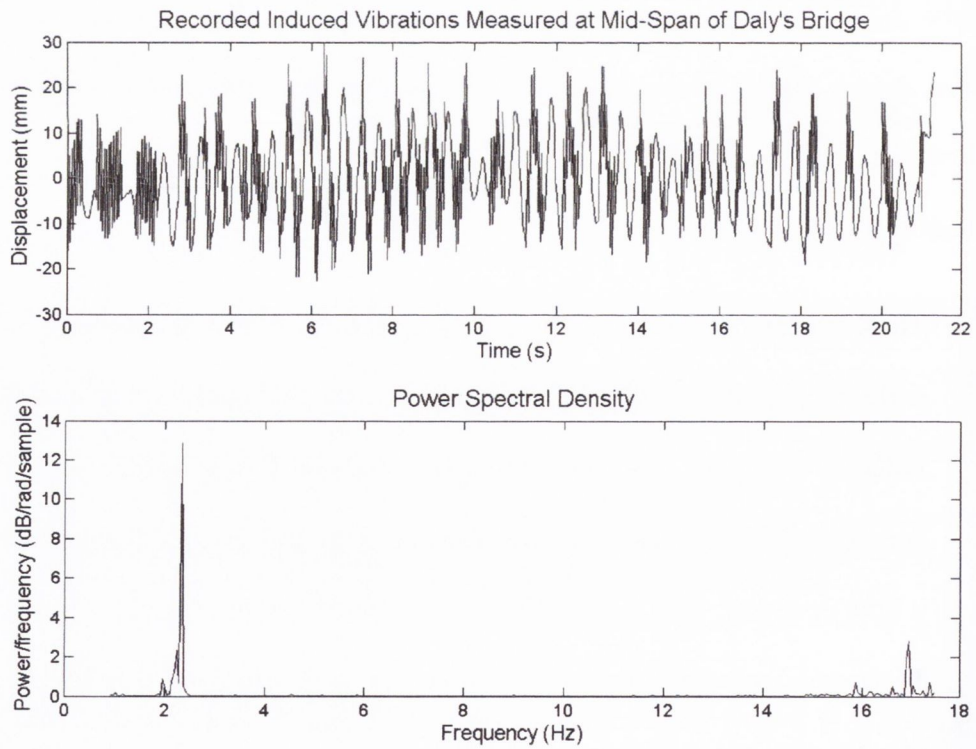


Figure 4.37 Pedestrian induced vibrations - female.

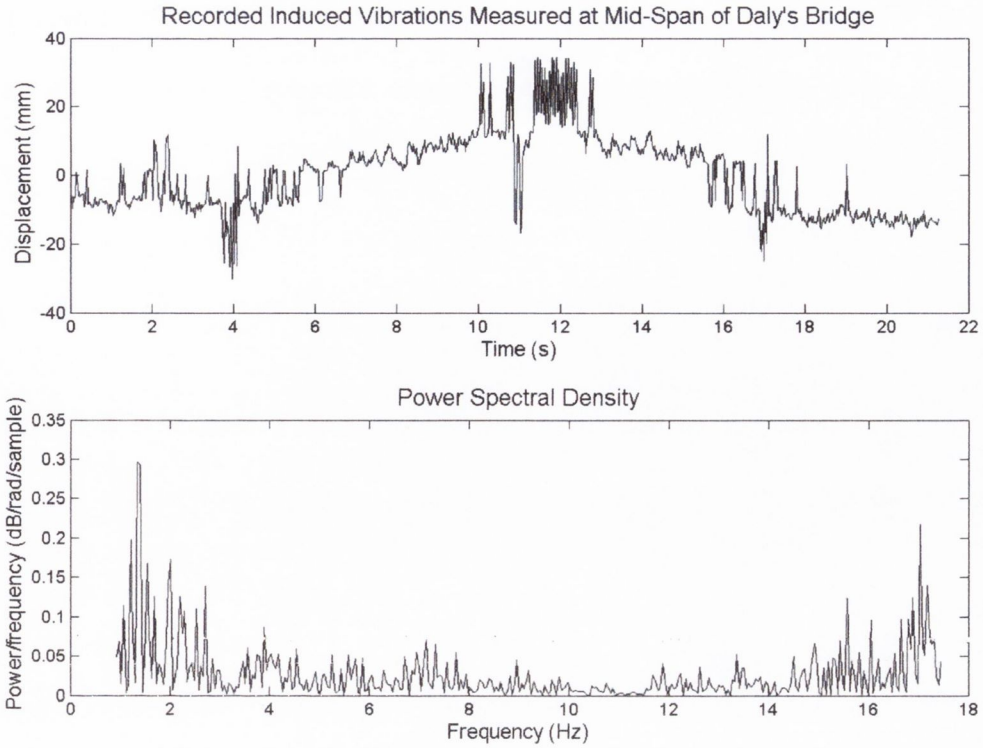


Figure 4.38 Cyclist induced vibrations.

It may be observed from each of the PSD plots that there is a peak at around 2 Hz, which corresponds to the natural frequency of the bridge. Accelerometers placed at the mid-span of the bridge indicate similar values for the bridge's natural frequency. A summary of the results for both sensor types is presented in Table 4.5.

Table 4-5 Natural Frequency Results from Image Analysis Method and from Accelerometers on the Bridge.

	Mean Natural Frequency	Standard Deviation from Mean, %
Accelerometers	2.265 Hz	4.41
Video Analysis	2.11 Hz	1.56

From these results, it may be concluded that video tracking offers a convenient and accurate way of identifying the natural frequency of bridges characterised by large amplitudes and low frequencies.

4.6 Conclusion

This chapter draws together a range of colour analysis based techniques for detecting and evaluating the extent of various damage forms, as well as for identifying the modal parameters such as natural frequency for a bridge. The image analysis techniques have been specifically designed to cope with the challenging underwater conditions and to provide an efficient source of quantitative information for inspectors. The developed image analysis techniques relate to; 1) crack detection, 2) surface damage detection, and 3) 3D shape recovery using stereo-vision.

The crack detection method proposed in this chapter adopts a percolation based approach. It is demonstrated on a series of images, which present a variety of challenges such as bright spots, reduced clarity due to a turbid medium, the presence of potential false detections arising from line markings or other artefacts. The proposed method offers greater efficiency over other percolation based methods, whilst still achieving good results, as the search space for cracks is confined to places only where there is a sharp image intensity gradient as determined from an edge detector, namely, the Sobel operator. Furthermore, a new criterion was introduced for classifying potential cracks, which sought to exclude the false detection of lines that separated two dissimilarly coloured regions.

This chapter also presents an image analysis based damage detection technique, REMPS, which is intended to supplement and strengthen existing visual inspection methods by providing a quick and convenient source of quantitative information. The

development of REMPS was necessitated by a lack of sophisticated image based damage detection techniques that can be applied to a broad range of surface types, damage forms, and lighting conditions that are typically encountered in infrastructures. REMPS adopts a multi-phase segmentation methodology which incorporates features from three standard image processing and data analysis techniques. Since these techniques are well-known and described in the literature, REMPS may be easily replicated and implemented. A key benefit of REMPS is its ability to produce better defined and more homogenous regions of interest without being affected by isolated extraneous pixels. REMPS achieves this cleaner segmentation by efficiently integrating pixel and spatial relationships. The presented results indicate that improvements can be made to the detection accuracy by adopting a HDR protocol.

A stereo based 3D shape recovery technique is also presented in this chapter. Having an accurate 3D shape reconstruction of underwater infrastructural elements is of great practical importance when analysing the forces exercised by the waves, winds and currents. The development a pyramidal Belief Propagation (BP) Markov Random Field (MRF) stereo correspondence algorithm in this chapter is motivated by the need for a technique that can effectively deal with the poor visibility conditions encountered in an underwater setting and efficiently exploit the high resolution of images to provide good depth sensitivity.

Finally, this section implements a video tracking technique, which is demonstrated on a pedestrian suspension bridge spanning a river in Cork, Ireland, for the purpose of identifying the natural frequency. Video analysis is an attractive option as the only equipment required is a conventional digital video camera and the output is easy to interpret in comparison to other methods for identifying the natural frequency such as a

network of accelerometers. Furthermore, the results indicate that video tracking can be an accurate solution.

Chapter 5

An Online Repository for Evaluating the Effects of Onsite Conditions on Image Processing based NDT techniques

5.1 Introduction

In chapters 3 and 4, a range of image based techniques were developed for detecting and quantifying numerous damage forms on marine structures such as cracks and corrosion, as well as for recovering 3D shape information. The next step is to assess the performance of these techniques when applied in an underwater setting, which is generally characterised by reduced visibility. There are three important and practical aspects for inspectors to have an understanding of how the onsite conditions affect the performance of image based methods. Firstly, a better understanding of underwater conditions for image based inspection can be used by inspectors to rationalise the use of image based methods as part of the inspection regime. Secondly, such understanding creates conditions conducive to good performance during inspection. Finally, it also allows choosing an appropriate image analysis algorithm that maximises detection accuracy. This chapter addresses these issues by developing an open-source Underwater Lighting and Turbidity Image Repository (ULTIR) that contains images of various damage forms, material types and shapes, which are photographed under controlled lighting and turbidity levels.

The need for this repository is warranted by the extensive effort and expense associated with undertaking underwater inspections, where inspectors can ill-afford to rely on

untested NDT methods. For the most part, algorithms intended for underwater application are developed and presented without assessing their performance levels for various conditions. This may be partly attributed to the lack of good quality imagery or controlled imagery available when evaluating image based methods. Researchers working on damage detection and 3D shape recovery algorithms would benefit from having large and diverse datasets of images for developing and refining their algorithms, as well as for training and validation purposes. Thus, there are clear benefits of having a large, standardised, well-annotated and freely-available database of images and associated metadata.

ULTIR is the first repository of this kind and may be accessed through a user-friendly web-interface (available at: http://www.ultir.net), which allows users to browse and download images of controlled and partially controlled specimens in an underwater setting along with the associated ground-truth data. ULTIR consists of three categories relating to; 1) 1D crack detection, 2) 2D surface damage detection, and 3) 3D shape recovery using stereo-vision. The imagery contained within each category was captured under three lighting levels and three turbidity levels, resulting in nine images for each specimen. The specimens thus cover a wide range of geometric and photometric properties. This resource affords researchers the opportunity to efficiently gauge the performance of damage detection methods under realistic operating conditions and on relevant specimens. Furthermore, researchers can compare their methods against existing methods and an upload facility allows users to contribute their own imagery. This chapter develops the comparison facet by applying some common image methods, as well as newly established techniques in the domain of NDT, to the images in the repository. Results show that the choice of image processing method and the environmental conditions are important factors and these should be given careful consideration to establish the domain of operation and efficiency of different methods and sites.

5.2 Design and Content

ULTIR consists of images generated from experiments conducted in an underwater setting that are accessible via the internet. This section describes the contents of the repository, the experimental set-up used to generate the imagery, and the web-interface. A schematic of the design and contents of the repository is illustrated in Figure 5.1.

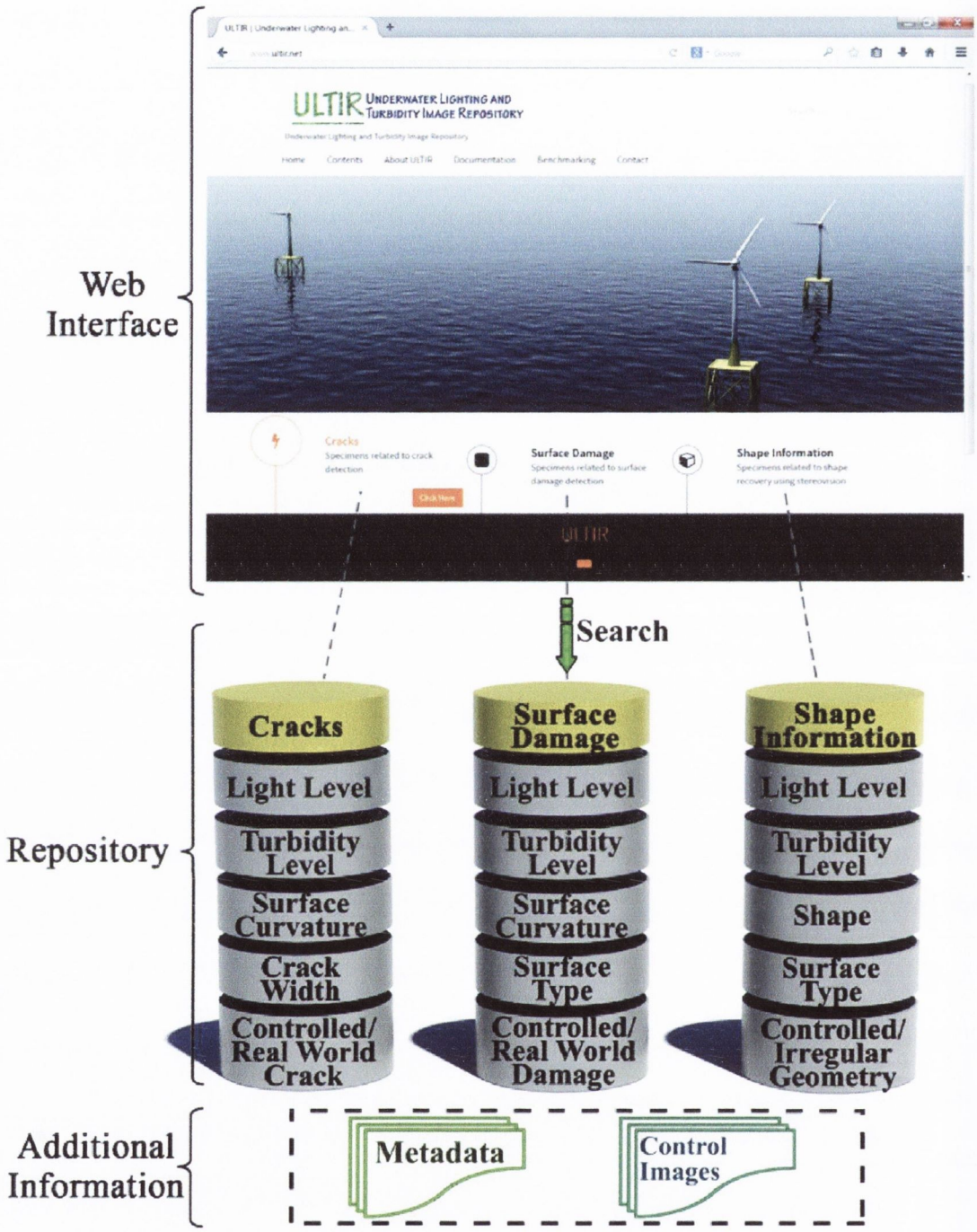


Figure 5.1 Contents of the repository.

5.2.1 Contents of the Repository

ULTIR contains a large and diverse array of images which cover a wide range of conditions and situations encountered underwater. This diversity is necessary for proper

and comprehensive validation of algorithms. A full breakdown of the contents of the repository are summarised in Table 5.1 and the imagery is shown in Appendix I. The images in the repository are characterised by several attributes - most notably the turbidity and lighting conditions under which the image was captured and the type of damage under consideration - but also by other attributes such as the object curvature and surface type. The significance of each of these attributes is discussed below.

Table 5-1 Breakdown of the contents of the repository.

Section	No. of Specimens	Level of Control	Surface Type	Shape/ Curvature	Light Levels	Turbidity Levels	No. of images
Cracks	9	7 controlled, 2 real cracks	8 concrete, 1 textured concrete	Surface curvature: 4 flat, 5 curved	3 levels: 100 lux, 1000 lux, 10000 lux	3 levels: 0 NTU, 6 NTU, 12 NTU	81
Surface damage	10	9 controlled, 1 real damage	4 concrete, 3 textured concrete, 3 metallic	Surface curvature: 4 flat, 3 cylindrical, 3 spherical	3 levels: 100 lux, 1000 lux, 10000 lux	3 levels: 0 NTU, 6 NTU, 12 NTU	90
3D shapes	12	9 controlled, 3 irregular shapes	4 concrete, 4 metallic, 3 plastic, 1 rubber	3 cubes, 3 cylinders, 3 spheres, 3 irregular shapes	3 levels: 100 lux, 1000 lux, 10000 lux	3 levels: 0 NTU, 6 NTU, 12 NTU	(108 x 8) 864

5.2.2 Controlled and Partially Controlled Images

The image specimens are classified as either controlled or partially controlled. For the controlled specimens, the 'damage' is artificially created such that the true dimensions of the specimen and the damage are precisely known. The partially controlled specimens feature real world instances of damage or are irregular shaped objects. In this case, precise knowledge of the size of the damage is not known beforehand and must be visually identified by a human observer from the images. The visually segmented image is created by a human operator who manually identifies damage in an image. It acts as the control as it is assumed it shows the true extent of damage. The visually segmented images then act as the control. Approximately 80% of the repository is made up of controlled specimens while the remainder are partially controlled specimens. The controlled specimens have the primary purpose of algorithm validation, while the partially controlled are intended for testing.

5.2.3 Damage Type

The repository is partitioned according to the type of damage. There are three general damage forms considered: 1D cracks, 2D surface damage, and damages quantified by 3D shape. Although the primary purpose of ULTIR is to benefit inspectors and researchers directly working with damage detection techniques, it is also well suited for testing other image algorithms such as image de-noising and contrast enhancement. However, these additional applications are beyond the scope of this chapter. The background and nature of each of the aforementioned damage forms are discussed below:

5.2.3.1 Cracks

Cracks provide an indication of the structural degradation and are an important factor when diagnosing the condition of concrete structures. Marine structures are especially susceptible to cracks due to the high dynamic loading and the presence of corrosive salts which are absorbed into the concrete and corrode the reinforcing steel, causing volume expansion and consequent cracking. Traditional visual monitoring methods typically require an inspector to be at on-site in order to map, count, quantitatively measure and photograph the observed cracks. This tedious process can be significantly improved through image processing based methods which can automatically count, classify and quantify the length and width of cracks, thereby enhancing inspections, and in turn, leading to significant monetary savings or leading to more efficient inspection cycles.

Image based crack detection algorithms work by identifying features of cracks such as their narrow shape and their lower brightness in comparison to the surroundings. For the controlled part of the repository, cracks were simulated by offsetting two halves of a split concrete specimen by a fixed distance. The controlled cracks range in width from 1 mm to 5 mm. These simulated cracks share all the features of real cracks such as the fine structure and lower brightness. Each of the cracked specimens in the repository is shown in Figure 5.2. The images shown were captured under medium lighting (1,000 lux) and in clear water (0 Nephelometric Turbidity Units (NTU)) conditions. Full details of illuminance and turbidity values are discussed in Section 5.2.3.4. ULTIR contains two cases of real-world cracks on different surfaces (Figure 5.2(h-i)).

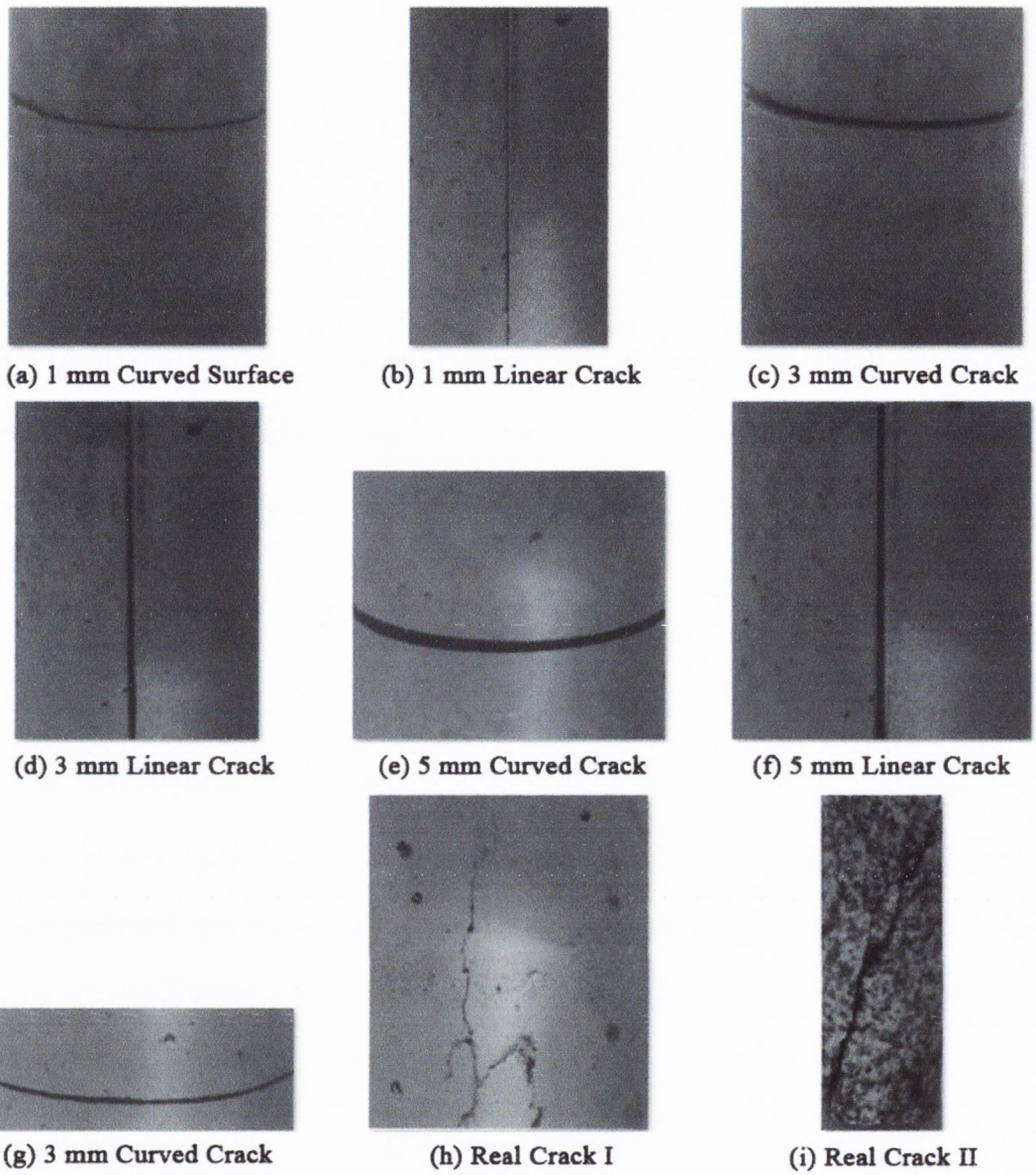
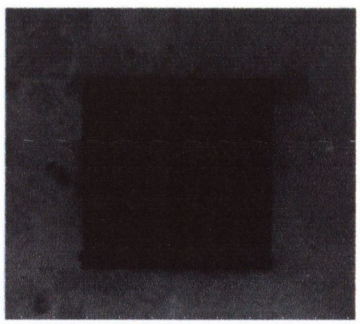


Figure 5.2 Specimens in the crack section of the repository. These illustrated images were captured under medium light (1,000 lux) and low turbidity (0 NTU) conditions.

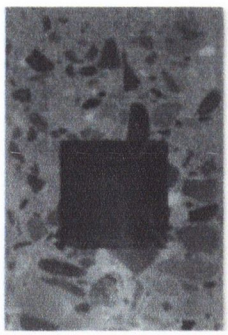
5.2.3.2 Surface Damage

The aim of the image based 2D surface damage algorithms is to locate and quantify the area occupied by visible damage (typically larger than 10^{-6} m^2) on the surface of infrastructural elements with minimal human supervision. Marine structures are affected by a wide range of visible damage forms such as corrosion, scour, erosion, leaching,

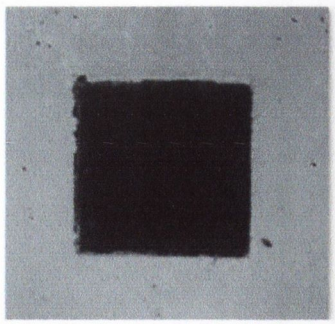
spalling, impact damage etc. These damages come in an array of shapes, sizes and pose varying levels of significance to the health of a structure. While some damage forms such as pitting corrosion are generally detected using colour information, other damages such as spalling and erosion may be more differentiable based on their textural properties in comparison to the undamaged surface. The contents of the repository reflect this variety by including specimens with different surface textures, as shown in Figure 5.3. For the controlled part of the repository, damage is simulated by applying a typically rust-coloured paint to standard geometric shapes such as squares and rectangles of known dimensions on different surfaces and surface curvatures. The repository also contains a real-world case of rust spotting (Figure 5.3(j)).



(a) Flat Concrete Surface

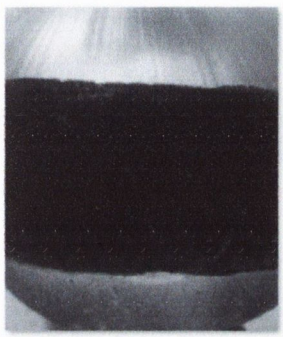


(b) Flat Textured Concrete

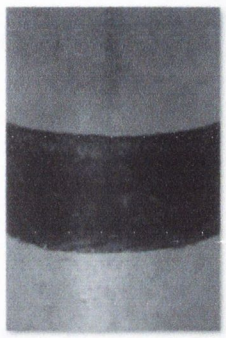


(c) Flat Metallic Surface

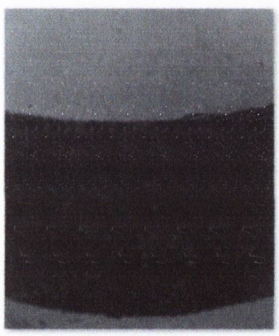
(d) Cylindrical Metallic Surface



(e) Spherical Metallic Surface



(f) Cylindrical Concrete Surface



(g) Spherical Concrete Surface

(h) Spherical Textured Concrete Surface

(i) Cylindrical Textured Concrete Surface

(j) Real Rust Spotting

Figure 5.3 Specimens in the surface damage section of the repository. These images were captured under medium light (1,000 lux) and low turbidity (0 NTU) conditions.

5.2.3.3 Shape Information

Recovering shape information is a challenging but useful task, and has wide applicability in many areas of Structural Health Monitoring (SHM) such as defect detection (LiMei et al., 2005) and accurate crack reconstruction (Yatchev et al., 2012) in topside inspections. For underwater inspections, one such application concerns the tracking of marine growth thickness on offshore structures. Marine growth is undesirable as it increases the hydrodynamic forces acting on the host structure by increasing the diameter of structural members and the roughness coefficients (Boukinda Mbadinga et al., 2007). For fixed offshore structures such as Jacket platforms, the hydrodynamic loading caused by marine growth is considerable in terms of the percentage on each component, and consequently on the total load acting on the entire structure.

There are a number of practical image based approaches capable of recovering 3D shape. These include structured lighting, Structure from Motion (SfM), and stereo photography. Other techniques such as Structure from Shading (SfS) (Shaomin and Negahdaripour, 1997), depth from defocus (Nayar et al., 1995) and the use monocular cues (Barrois and Wohler, 2007) are not discussed further as they are not well suited for accurate quantitative shape recovery as part of underwater inspections. Structured lighting techniques utilise a light projector and a camera to project a light pattern onto an object and capture how it interacts with the shape (Bruno et al., 2011). Given the reliance on a light source to encode depth information, the success of this approach is particularly susceptible to absorption and scattering.

Structure from Motion relies on an image sequence/video acquired from a single moving camera to obtain 3D shape information. Features in the scene are tracked across successive

images. The trajectories of each feature over time provide information on the camera motion as well as the 3D position of the feature in space. SfM algorithms require a scene to remain relatively rigid which is not always possible in an underwater setting as deforming surfaces, floating particulate, and illumination changes are all sources of non-rigidity. Additionally, SfM has been shown to be less reliable than stereo vision based methods when applied underwater (O'Byrne et al., 2014).

A stereo system consists of two synchronised cameras viewing the scene from two slightly different vantage points. The captured images from both cameras are collectively known as a stereo pair. By examining the relative positions of objects in each image, 3D information can be extracted. Stereo systems are capable of providing a fully scaled metric scene reconstruction once the system is calibrated, using either standard techniques (Zhang, 2000) or by auto-calibration (Faugeras et al., 1992). A fully scaled metric reconstruction is an accurate representation of the scene; true angles are preserved and object dimensions are in real world units. ULTIR is populated by stereo pairs of numerous specimens. The controlled specimens are standard geometric shapes in the form of spheres, cubes and cylinders, as shown in Figure 5.4. The controlled specimens were chosen as these primitive shapes are the building blocks for more complex shapes. Therefore, it is important to get a fundamental understanding of how stereo-matching algorithms handle the various curvatures. The uncontrolled specimens are irregular shapes with more intricate depth variations which is more reflective of what would be encountered during real-world inspections.

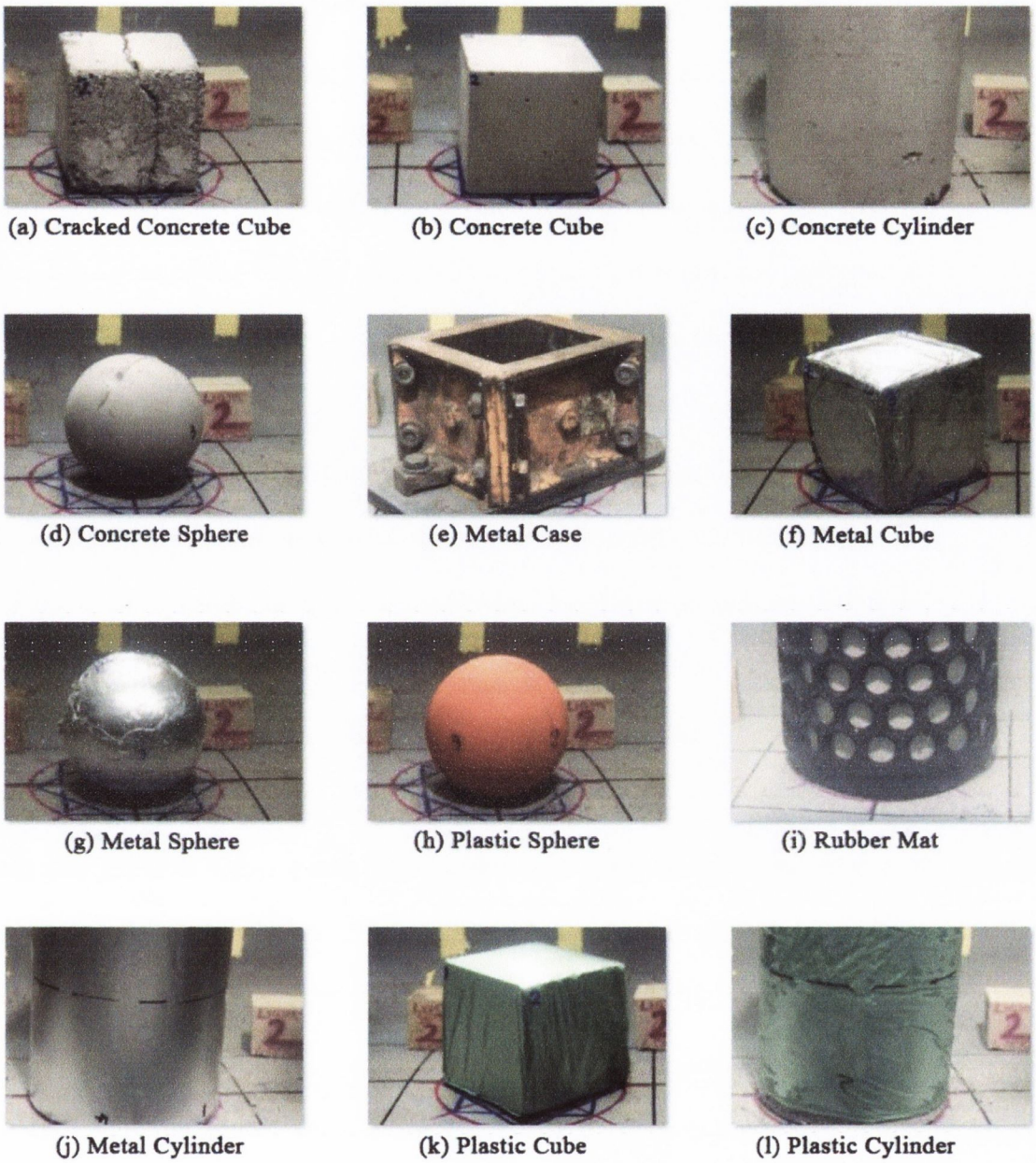


Figure 5.4 Specimens in the surface damage section of the repository. These images were captured under medium light (1,000 lux) and low turbidity (0 NTU) conditions.

5.2.3.4 Turbidity and lighting

Image quality is assumed to be chiefly affected by luminosity, sharpness (focus accuracy), contrast and noise. These quality factors are directly related to the on-site operating conditions, for which turbidity and lighting are the most influential (Mahiddine

et al., 2012). Lighting is crucial for all photographing applications, while turbidity may be regarded as an integrative parameter, which takes into account other parameters that affect visibility such as water colour, total suspended solids and dissolved solids. The other notable factors that affect image quality are the quality of the camera sensor, the choice of appropriate camera settings, and the adherence to good photographing practices. These issues are discussed in greater detail in Chapter 6.

Turbidity is defined as the cloudiness in a liquid caused by the presence of suspended solids (Kirk, 1985). These suspended solids scatter and absorb light and therefore reduce visibility. As can be expected, underwater imaging is severely hampered by turbidity, which results in reduced contrast, loss of details and colour alteration. Turbidity can be caused by organic particles, such as decomposed plant and animal matter, and algae; or by inorganic particles such as silt and clay. In rivers and lakes, the level of turbidity can fluctuate due to a number of factors such as; heavy rains, flooding and spring runoff, landslides and bank erosion, algae blooms, interference during the inspection process (e.g. a boat may aggravate and disturb sediment on the river bed), human activities such as construction, and storm water pollution from urban areas. Given these factors, significant variations in turbidity can be expected. For example, the measured turbidity in the lower Waitaki river, New Zealand, ranges from between 1.2 and 23 NTU (Graham, 1990). In the Minnesota river basin, United States, it was found that high levels of turbidity occurred during periods of excessive stream flow as a result of increased suspended inorganic sediment from watershed runoff, stream banks, and channel contributions (Lenhart et al., 2010). In the open ocean, turbidity is affected mostly by seasonal phytoplankton blooms, however it is generally low.

For ULTIR, three levels for turbidity were chosen: 0 NTU, 6 NTU, and 12 NTU. Clear water has a turbidity of 0 NTU, water that is visibly cloudy has a turbidity of 6 NTU, while water that is murky has a turbidity of 25 NTU. A cut-off point of 12 NTU was chosen based on observations that it becomes increasingly difficult to interpret and extract useful information from images captured in more turbid water. Turbidity can be measured on site using a digital nephelometer or by using a Secchi disk which is a black and white disk lowered into the water until it is no longer visible. The depth of the disk is a measure of the transparency of the water which is inversely related to the turbidity. This has the advantage of being a quick, inexpensive and simple approach for measuring turbidity. Additionally, many water bodies, such as the River Lee in Ireland, are increasingly being monitored by arrays of sensor networks, which provide near real-time data on water quality parameters including turbidity (Lawlor et al., 2012).

The effects of turbidity can be overcome by moving the camera closer to the subject, however, it must be noted that if the camera is too close to the subject then the resulting imagery loses context and can only depict small areas. The imagery in ULTIR is generated with a camera(s) and specimens kept at a fixed distance of 80 cm apart. This distance is regarded as being a practical balance between capturing enough of the scene whilst mitigating the effects of turbidity.

Lighting also plays a pivotal role for achieving good visibility. Ambient light may be sufficient for near-surface inspections; however, it is unlikely to be sufficient at greater depths at which point artificial light sources become necessary. Three light levels were used: 100 lux, 1000 lux, 10000 lux. To put this in perspective; the approximate level of light, or illuminance, on a very dark overcast day is 100 lux, a moderately overcast day is 1000 lux, and full daylight (not in direct sunlight) is 10,000 - 25,000 lux (Schlyter, 2009).

A specimen from each section of the repository is shown under the varying levels of lighting and turbidity in Figures 5.5 - 5.7.

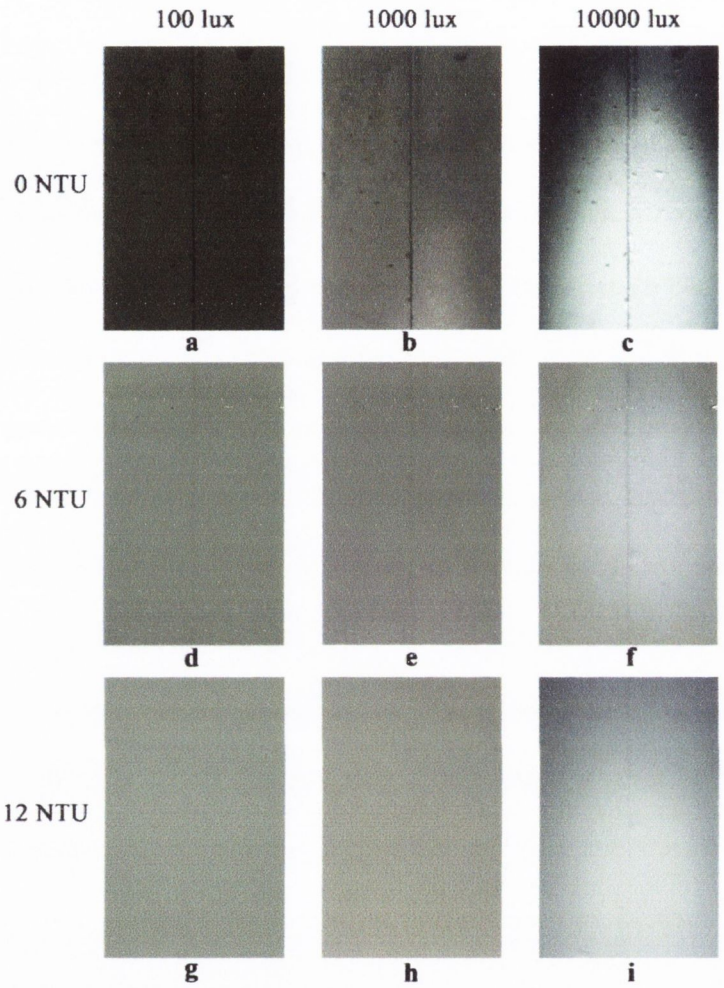


Figure 5.5 Controlled crack specimen from ULTIR shown under varying lighting and turbidity conditions. Columns: Low (100 lux), Medium (1000 lux), High (10000 lux) Light. Rows: Low (0 NTU), Medium (6 NTU), High (12 NTU) Turbidity.

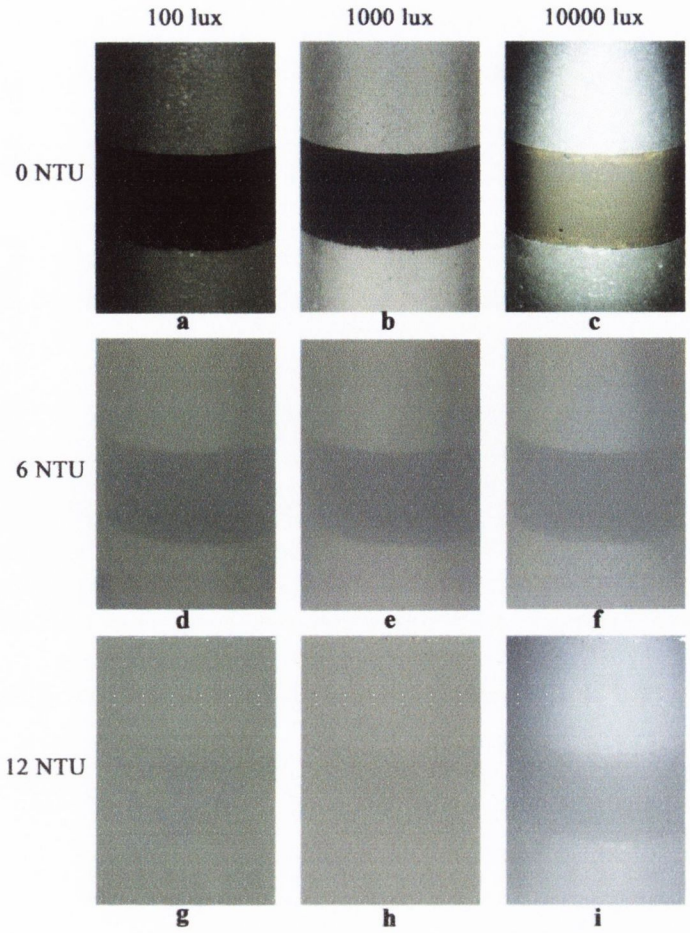


Figure 5.6 Controlled surface damage specimen from ULTIR shown under varying lighting and turbidity conditions. Columns: Low (100 lux), Medium (1000 lux), High (10000 lux) Light. Rows: Low (0 NTU), Medium (6 NTU), High (12 NTU) Turbidity.

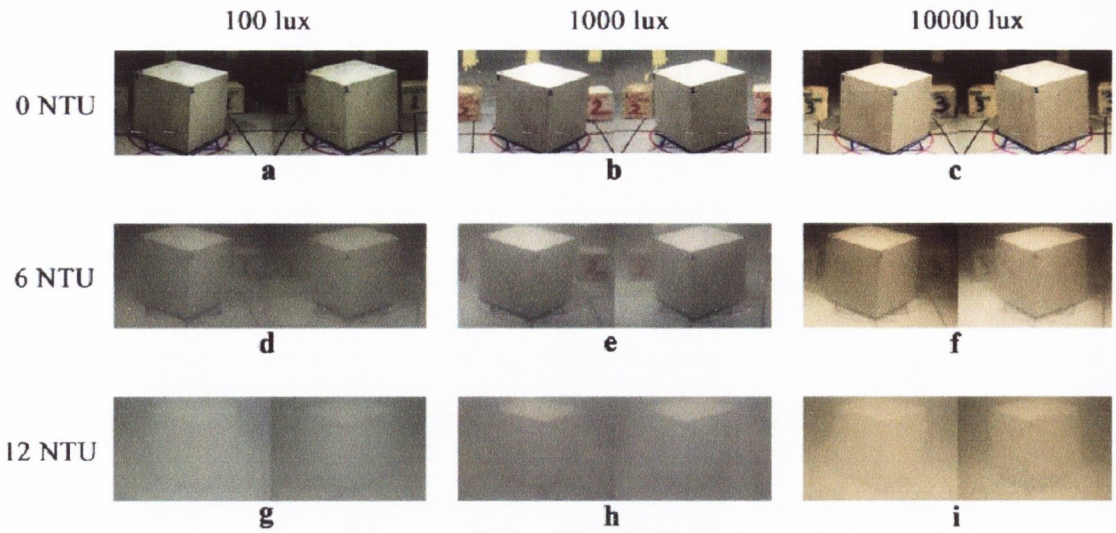


Figure 5.7 Stereo pairs featuring a concrete cube from ULTIR shown under varying lighting and turbidity levels. Columns: Low (100 lux), Medium (1000 lux), High (10000 lux) Light. Rows: Low (0 NTU), Medium (6 NTU), High (12 NTU) Turbidity.

5.2.3.5 Surface type

The main construction materials for marine structures are concrete and metal. These materials have different photometric properties. Metallic surfaces often appear shiny to the eye due to a large amount of light being reflected in a specular fashion whilst concrete surfaces reflect very little specular light, instead reflecting light in a scattered fashion which produces a dull appearance (Dana et al., 1997). High specular reflections can present some problems from an image analysis perspective as the shine hides details and creates artefacts that could mislead algorithms. This is especially problematic when strong light sources are being used.

The texture of a surface is another important property that has an effect on image analysis. Texture may be qualified by terms such as fine, coarse, smooth, rippled, molled,

irregular, or lineated (Haralick et al., 1973). The influence of texture can be seen during the stereo-matching process in which corresponding points in two images must be matched. Matching points on smooth or uniform surfaces is more ambiguous than matching points on coarsely textured surfaces. Additionally, the surface texture may be a consideration when deciding on what type of surface damage detection algorithm to use. If damage is characterised more so by textural composition than by colour from the undamaged surface then it may be worthwhile segmenting based on texture. Finally, crack detection algorithms applied to surfaces with a rippled texture can produce a lot of false alarms. To cater for these issues, ULTIR contains metallic surfaces with high specular reflectivity as well as diffuse concrete surfaces with various textural finishes.

5.2.4 Experimental Set-up

The images in ULTIR were generated from experiments that were conducted in a water basin. These experiments were run in two phases. The set-up for the first phase entailed having a single underwater camera focused on damaged specimens. This phase produced the imagery for the crack and surface damage sections of the repository. The second phase employed on a dual-camera set-up as shown in Figure 5.8. This phase produced the stereo imagery for the 3D shape information section of the repository.

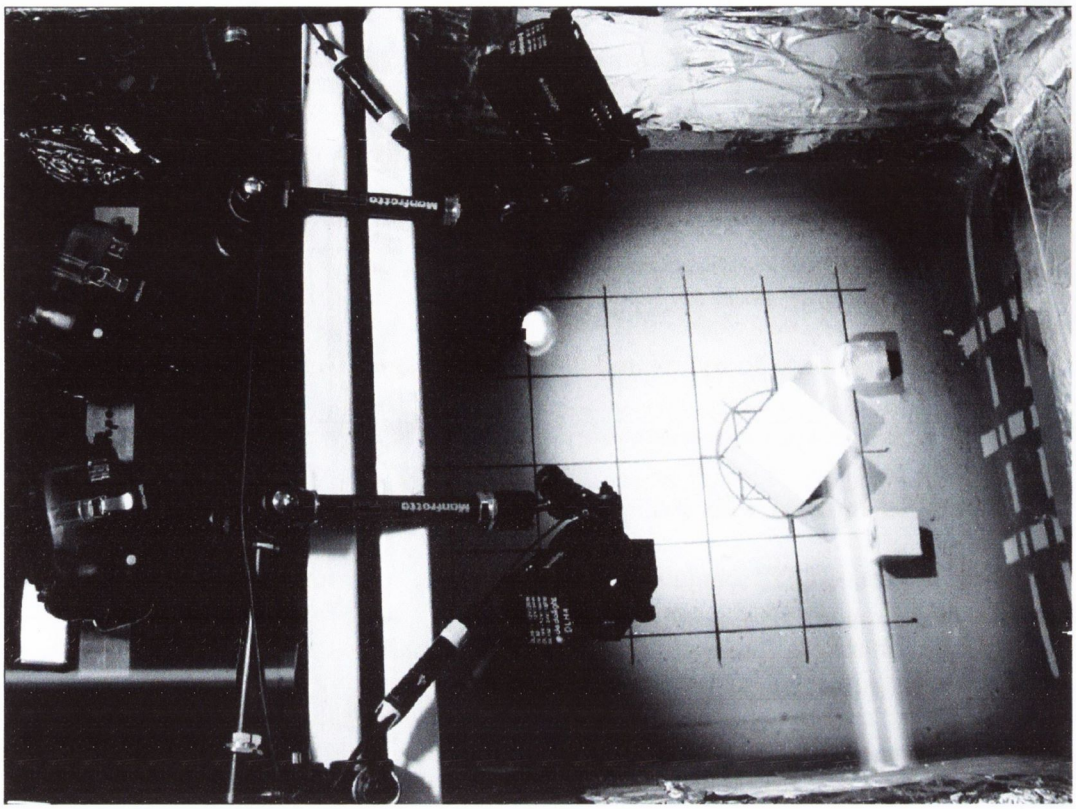


Figure 5.8 Plan view of experiment set-up.

5.2.4.1 Controlling the turbidity and light levels

Both phases of the experiment followed the same methodology in terms of controlling turbidity and light levels. Turbidity was measured using a digital turbidimeter (model: Hach 2100P), which operates by measuring the loss of intensity of transmitted light through a water sample. Initially, the basin was filled with clear water. After all the photographs were captured at this turbidity level, finely sieved kaolin was to added in order to bring the turbidity up to 6 NTU, and later, to 12 NTU. Kaolin is a soft white clay consisting principally of the mineral kaolinite which goes into suspension when mixed with water. Regular stirring was carried out to ensure kaolin remained in suspension and was distributed uniformly in the basin.

The light level was measured using a lux meter. Ambient lighting and a configuration of light sources were used to produce light levels of 100, 1,000, and 10,000 lux. Light was measured at the same position just above the specimen. While the illumination over the surface of a specimen will vary somewhat as a result of the slightly different distances from the light sources, the lux values span multiple orders of magnitude so slight variations within a given light level are negligible compared to the differences between levels.

5.2.4.2 Camera parameters

The imagery was captured using two DSLR (Digital Single Lens Reflex) cameras, namely the Canon EOS 600D (Canon Rebel T3i), with standard kit lens (Canon EF-S 18–55mm lens f/3.5–5.6), which were enclosed in underwater housings. There are a number of camera settings which must be configured in order to get the desired quality of imagery. These settings include the shutter speed (exposure time), aperture, and ISO. The aperture, along with focal length, determine the depth of field and the shutter speed determines the amount of motion blur. Together, the shutter speed and aperture control how much light arrives at the camera sensor. In low light conditions such as an underwater environment, a slow shutter coupled with a large aperture is necessary to ensure sufficient light exposure.

The other important setting is the ISO which is the degree of sensitivity of a camera to the available light. The lower the ISO number, the less sensitive the camera sensor is to incoming light, while a higher ISO number increases the sensitivity. Higher ISO settings are generally used in darker situations to allow for faster shutter speeds, however, raising the ISO means a similar decrease in quality, with an increase in noise. While early digital cameras had objectionable levels of noise at ISOs as low as 800, most modern DSLRs produce good quality images at ISOs up to 1600 and above. A balance must be struck

between all of these camera settings. The shutter speed was chosen as the critical parameter as diver-held cameras employed in underwater inspections are prone to shaking which leads to unacceptable levels of motion blur if long shutter speeds are being used. For this reason, shutter priority mode was used, which is a semi-automatic shooting mode that allows the user to specify the shutter speed. The camera then automatically decides the best aperture and ISO sensitivity for the specified shutter speed to get the correct exposure. For ULTIR, the shutter speed was set to be faster than 1/15 seconds.

The second phase of the experiment that involved collecting stereo pairs for the purpose of 3D shape recovery uses the two cameras at once. This phase required both cameras to operate simultaneously and to each have the same camera parameter configuration. In order to synchronise the cameras, a remote control that was linked to both cameras was used to initially trigger the photographing and from then on, successive images were captured at a pre-set time interval.

The cameras were configured to save images in RAW and JPEG format with pixel dimensions of 5184 x 3456. The JPEG format is significantly smaller in terms of file size compared to the RAW format with little perceptible loss in image quality. For this reason, the JPEG format is used to disseminate the images via the web-interface for ease of transmission.

5.2.5 Web-Interface

Users of the repository will first encounter the web-interface, which facilitates navigation through the repository. There are a range of links on the homepage that direct users to the three sections of the repository , along with links to supplementary material such as a user

guide and a description of the repository. Some of the key web pages are shown in Figure 5.9.

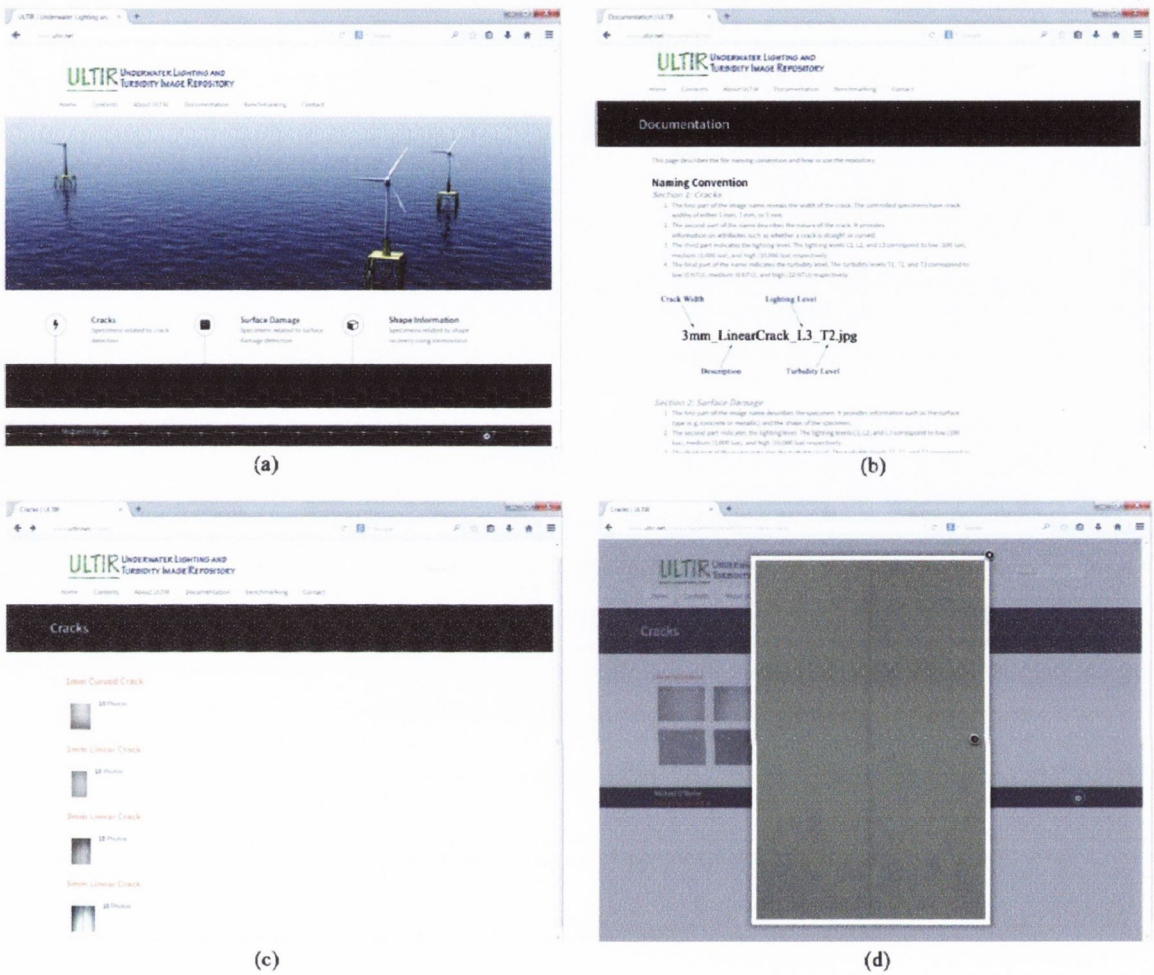


Figure 5.9 (a) ULTIR home page, (b) ULTIR documentation and user guide, (c) Crack page which lists the specimens and nature of the crack, (d) sample image from the crack part of the repository captured in low light and high turbidity.

The home page, as shown in Figure 5.9(a), allows quick access to the three sections of the repository pertaining to crack detection, surface damage detection, and the recovery of 3D shape information. For example, clicking on the 'Cracks' section will bring up a list of all the specimens and details about the nature of the crack such as the crack width, as

shown in Figure 5.9(c). Clicking on any thumbnail in this list will bring up the full assortment of associated images for each turbidity and lighting level, along with a binary control image which shows the damaged region in white and the undamaged in black, as shown in Figure 5.9(d). Metadata such as the aperture, focal length and ISO remain embedded in the images.

A menu bar on the home page provides links to several relevant pages. These pages outline the need and purpose of the repository, describe the contents, present a user guide, and showcase the performance of various techniques. The user guide, shown in Figure 5.9(b), explains the naming convention and advices how to navigate through the repository. The naming convention was adopted to efficiently convey important information such as the specimen description, turbidity level and lighting level.

5.3 Application of Image based Techniques

This section discusses some of the main algorithmic approaches related to each category of ULTIR, namely crack detection, surface damage detection, and 3D shape recovery. A number of algorithms are applied to the representative datasets as shown in Figure 5.5-5.7, the results from which are illustrated and the performance levels are quantified for ranking purposes. The performance of submitted techniques are evaluated and ranked with the help of performance points in the Receiver Operating Characteristic (ROC) space.

5.3.1 Crack Detection

5.3.1.1 Crack Detection Algorithms

The pervasive nature of cracks and the tedious task of manually counting and measuring them has led to a growing interest in utilising image processing based techniques to automate the detection process. A number of techniques have been devised which are capable of identifying crack-like features which are characterised by their narrow shape and lower brightness in comparison to the surroundings. These include the percolation based method proposed in Chapter 4 (O'Byrne et al., 2014), eigenvalue analysis of the Hessian (Frangi et al., 1998), Kirsch templates (Kirsch, 1971), neural networks (Choudhary and Dey, 2012) and statistical filters (Sinha and Fieguth, 2006). This chapter applies the first three of these methods to the 1 mm controlled crack data set (as shown in Figure 5.5) from the repository to investigate the effects of changing turbidity and lighting levels on the detection accuracy for each technique. Other lesser factors that could be investigated include the orientation of the crack. The orientation matters as square filters will give a slightly different response when considering diagonal cracks versus horizontal or vertical cracks.

The percolation method is based on tracing out dark pixels in a sliding window, or sub-image, starting at the centre point of the window, and spreading out until the boundary of the window. The resulting pattern of dark pixels is analysed. Cases where a narrow or linear pattern is traced out is indicative of a crack, while irregular or radial diffusion patterns typically correspond to the non-cracked background. The eigenvalue analysis of the Hessian method detects narrow crack-like paths by calculating the direction of smallest curvature where there is minimum change in intensity, which is usually along the crack

path. The Kirsch templates method detects line-like objects using spatial filtering involving templates orientated in eight different directions followed by thresholding.

The results for each of these techniques are shown in Figure 5.10, the performance levels are summarised in Table 5.2 and the associated performance points are plotted in the ROC space in Figure 5.11.

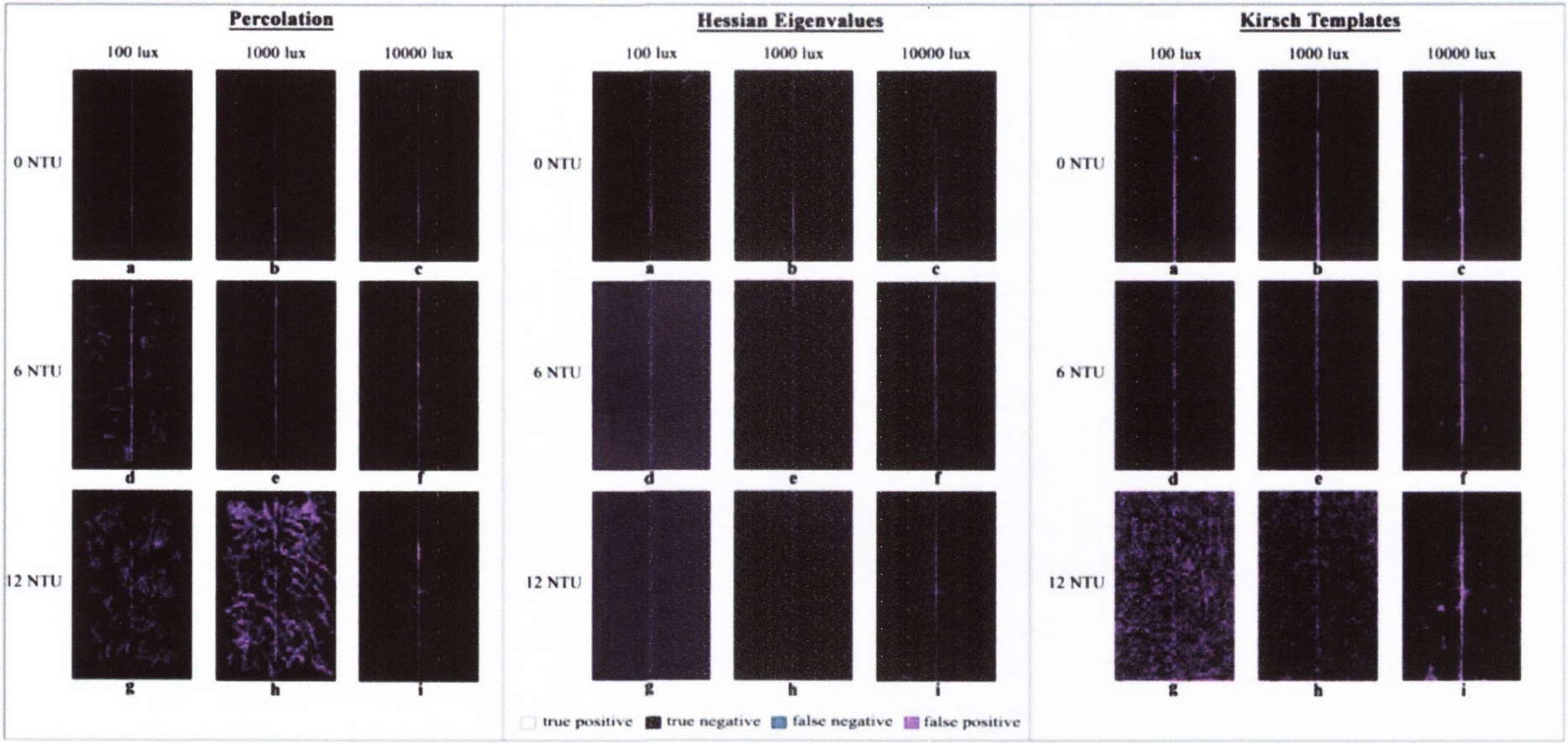


Figure 5.10 Detected cracks corresponding to the images in Figure 5.5.

Table 5-2 Performance of the crack detection techniques.

Image	Condition	Percolation			Hessian Eigenvalues			Kirsch Templates		
		(DR)	(MCR)	δ	(DR)	(MCR)	δ	(DR)	(MCR)	δ
a	Low Light, Low Turbidity	94.9%	0.7%	0.05	96.7%	2.3%	0.04	92.3%	5.3%	0.09
b	Medium Light, Low Turbidity	91.5%	0.8%	0.09	91.6%	1.2%	0.09	94.1%	4.0%	0.07
c	High Light, Low Turbidity	81.7%	0.7%	0.18	90.7%	3.9%	0.10	96.8%	5.3%	0.06
d	Low Light, Medium Turbidity	75.0%	3.9%	0.25	58.6%	50.2%	0.65	67.9%	13.2%	0.35
e	Medium Light, Medium Turbidity	88.0%	1.0%	0.12	54.6%	34.1%	0.57	79.4%	4.5%	0.21
f	High Light, Medium Turbidity	94.1%	1.2%	0.06	95.4%	7.7%	0.09	92.7%	4.8%	0.09
g	Low Light, High Turbidity	21.5%	8.4%	0.79	44.6%	48.7%	0.74	56.4%	54.2%	0.70
h	Medium Light, High Turbidity	54.8%	21.7%	0.50	25.2%	29.8%	0.81	55.3%	36.1%	0.57
i	High Light, High Turbidity	91.7%	1.7%	0.08	83.7%	24.6%	0.30	92.5%	11.7%	0.14

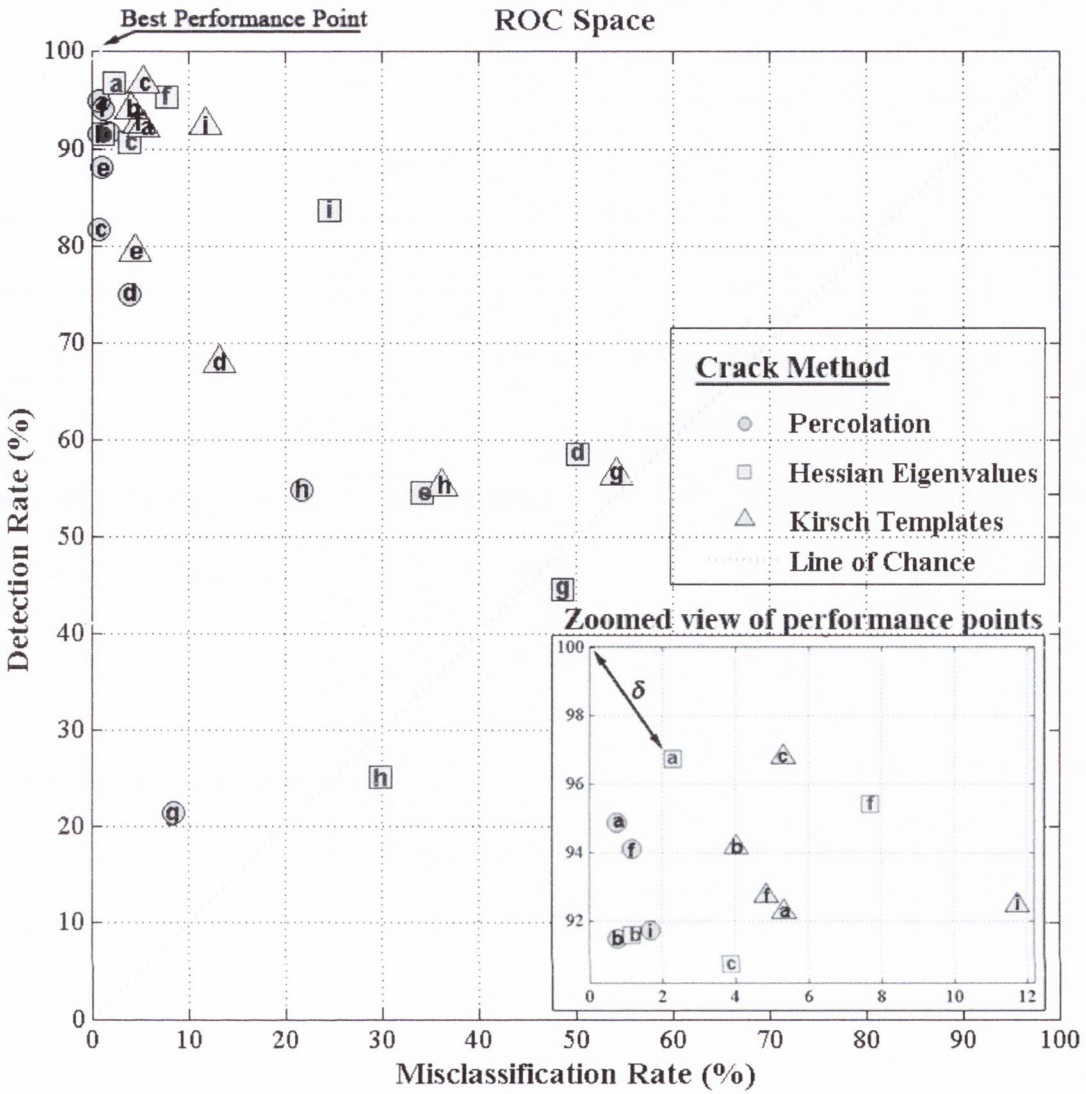


Figure 5.11 Evaluation of crack detection techniques through the use of performance points in the ROC space.

It may be observed from the detected cracks in Figure 5.10 that each technique performs quite well for images that feature clear and sufficiently lit scenes. Unsurprisingly, the performance deteriorates when the turbidity levels increase. In the worst visibility conditions - low light and high turbidity (Figure 5.5(g)) - all of the techniques produce poor results suggesting the adoption of image based crack detection approaches under these conditions is not practical. However, the results show that having high lighting can mitigate the effects of high turbidity. In these situations, the increased absorption and

diffusion in turbid water limits the formation of a bright spot which would otherwise impair detection.

Adopting the α-δ method allowed for a clear comparison between various detection rate and misclassification rate pairs, i.e. (*DR*, *MCR*). Analysis of the δ parameter in Table 5.2 reveals that the percolation based method performs the best for the medium turbidity level. The cracks are well delineated and there are relatively few misclassified pixels. The eigenvalue analysis of the Hessian method produces a high misclassification rate at the lower light levels, while the Kirsch templates method also produces a lot of small spurious regions.

The shape of the specimen and the crack width had a minor effect on the success of the crack detection techniques. Cracks with a larger width could be more easily detected at higher turbidities, and on curved surfaces, slightly more of the crack could be observed. For the real crack on the textured concrete surface, the textural pattern gave rise to a lot of false detections, indicating that the surface type is a key factor that controls the detection performance.

Overall, there is relatively little difference in terms of detection accuracy for the techniques considered when applied to the same samples, indicating that the choice of technique is not the most crucial factor. Instead, the visibility conditions have by far the greatest influence on the output.

5.3.2 Surface Damage

Most image processing based damage detection algorithms consist of segmentation followed by subsequent classification of the segmented regions. Ideally, the segmentation methodology should identify and accurately define all regions of interest in an image whilst minimising the inclusion of extraneous regions. In reality, perfect segmentation is difficult to achieve given the inherent chromatic and luminous complexities encountered in natural scenes. Image processing based techniques include colour intensity based methods and texture analysis based methods. Naturally, the techniques in each group are suited to different applications. The effectiveness of colour based segmentation algorithms and texture based segmentation algorithms will vary according to the surface and damage type under consideration as certain damages are more separable from the undamaged surface based on either their colour or texture attributes. This section assesses the performance of two colour based methods, the REMPS technique proposed in Chapter 4 (O'Byrne et al., 2014) and Otsu's thresholding (Otsu, 1979); along with the texture analysis based technique proposed in Chapter 3 (O'Byrne et al., 2013).

The performance of colour based segmentation techniques is affected by whether the technique is contextual or non-contextual. Non-contextual techniques such as Otsu's method do not take into account any spatial relationships between pixels in an image. REMPS, which operates by classifying whole regions, is an example of a contextual technique which does consider spatial relationships. If a contextual relationship is an important factor for segmenting a particular image, than non-contextual techniques will have limited success compared to techniques which exploit the contextual relationship.

These techniques are applied to the imagery in Figure 5.6 and the results are shown in Figure 5.12, the performance levels are quantified in Table 5.3 and the associated performance points are plotted in the ROC space in Figure 5.13.

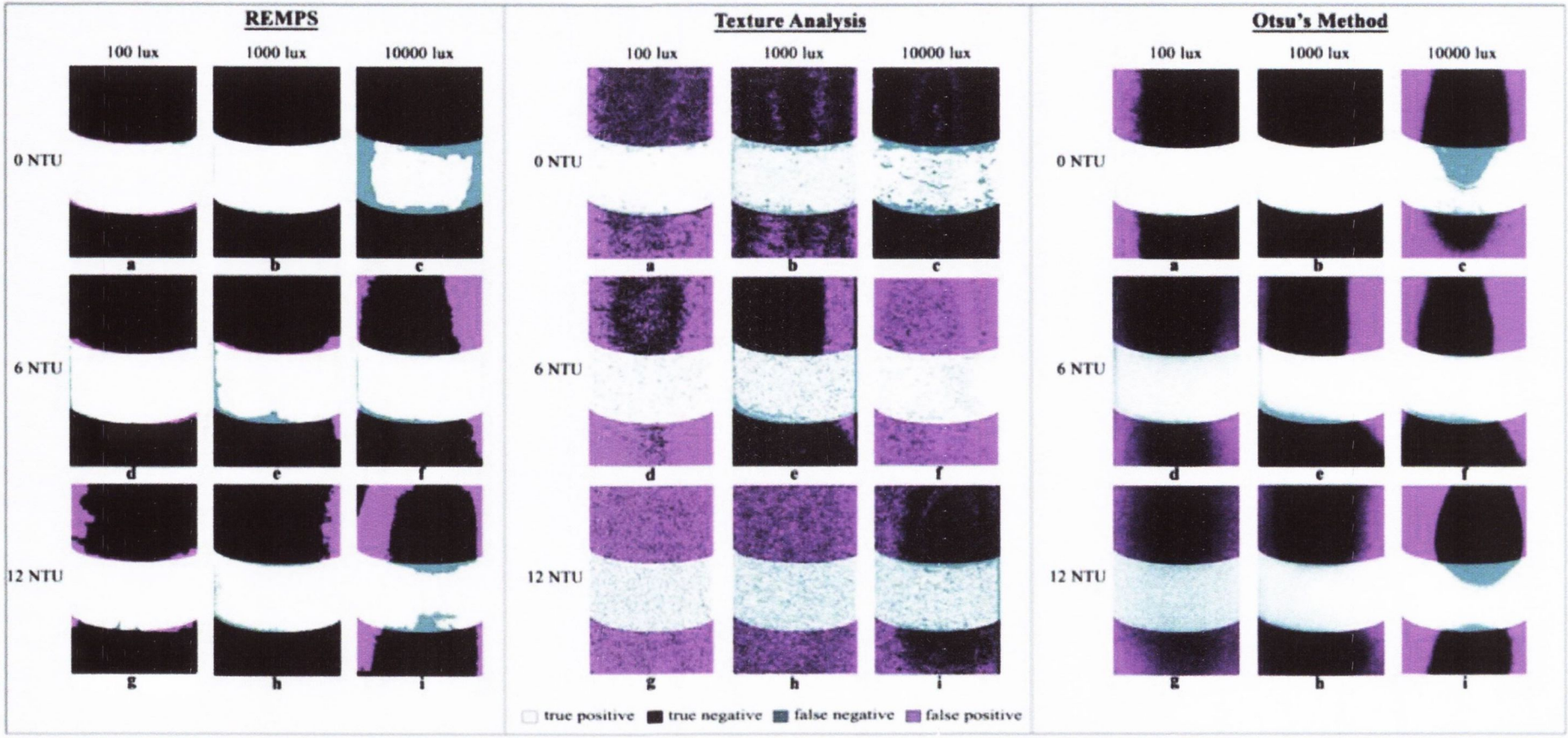


Figure 5.12 Detected damage corresponding to the images in Figure 5.6.

Table 5-3 Performance of the damage detection techniques.

Image	Condition	REMPS			Texture Analysis			Otsu's Method		
		(DR)	(MCR)	δ	(DR)	(MCR)	δ	(DR)	(MCR)	δ
a	Low Light, Low Turbidity	99.5%	2.1%	0.02	97.2%	49.3%	0.49	98.1%	14.8%	0.15
b	Medium Light, Low Turbidity	97.5%	1.2%	0.03	89.9%	33.1%	0.35	97.5%	1.2%	0.03
c	High Light, Low Turbidity	60.3%	14.2%	0.42	75.5%	15.8%	0.29	76.5%	39.8%	0.46
d	Low Light, Medium Turbidity	98.1%	2.5%	0.03	98.0%	51.3%	0.51	95.7%	22.2%	0.23
e	Medium Light, Medium Turbidity	91.9%	5.7%	0.10	82.6%	20.6%	0.27	92.6%	18.9%	0.20
f	High Light, Medium Turbidity	93.3%	16.7%	0.18	98.8%	61.5%	0.61	94.1%	24.6%	0.25
g	Low Light, High Turbidity	99.2%	7.0%	0.07	90.1%	61.8%	0.63	83.5%	37.5%	0.41
h	Medium Light, High Turbidity	93.4%	6.9%	0.10	83.8%	58.0%	0.60	93.9%	20.9%	0.22
i	High Light, High Turbidity	85.4%	24.0%	0.28	76.8%	34.0%	0.41	82.9%	30.7%	0.35

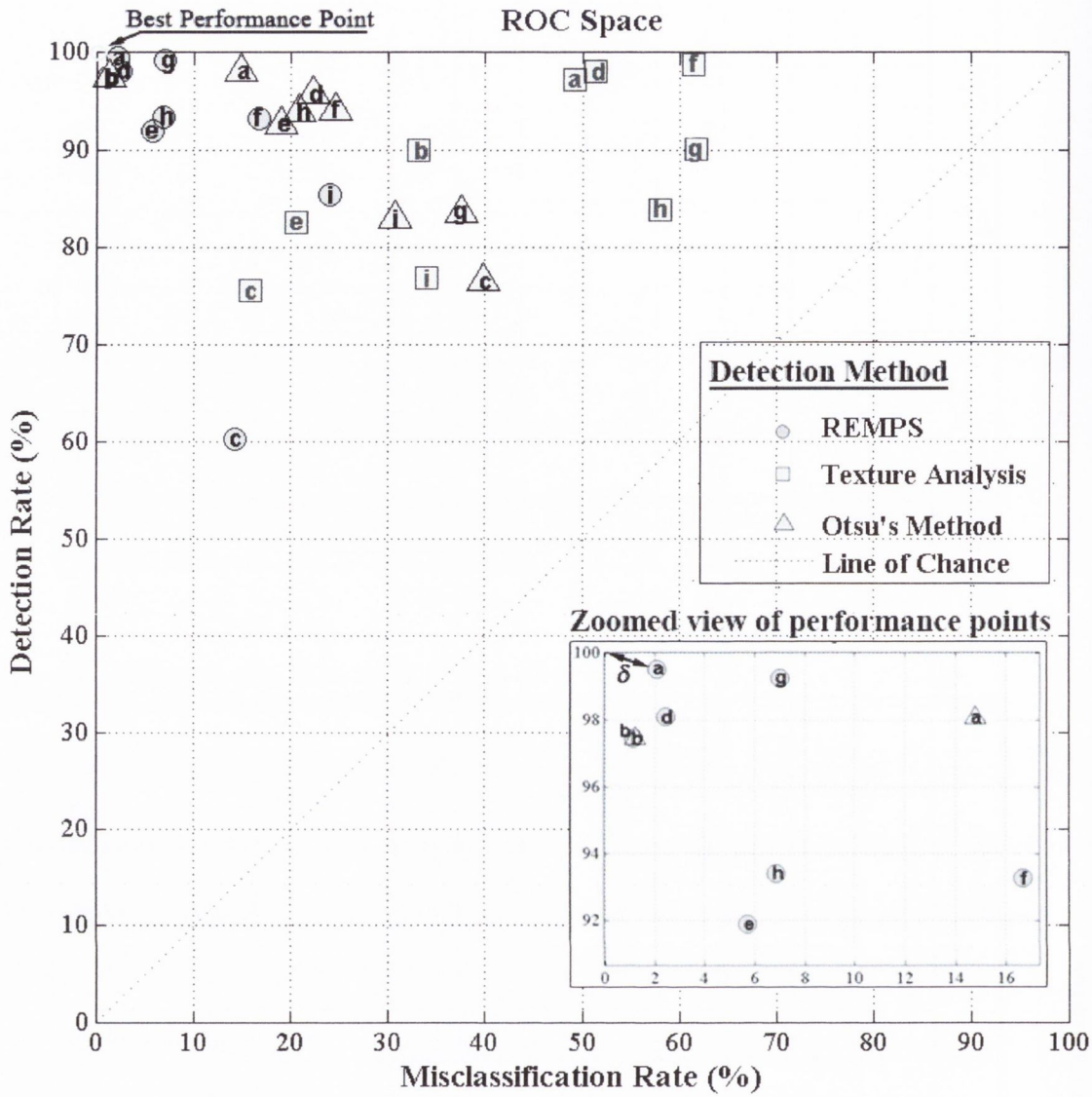


Figure 5.13 Evaluation of damage detection techniques through the use of performance points in the ROC space.

It may be noted from these results that REMPS was quite successful for the majority of cases with the exception of the overly bright image (Figure 5.6c). It proved effective at locating the presence of damage as well as accurately defining the shape and size of damaged regions. The success of Otsu's method may be explained by the fact that the damaged region for the specimen is characterised by one single colour which is sufficiently distinct from the background. The relatively apparent distinction between damaged and

non-damaged regions in this case diminishes the importance of exploiting any spatial relationships as part of the image segmentation technique, meaning that Otsu's method is not adversely affected. As such, the limitations of Otsu's method are not fully exposed. Both Otsu's method and texture analysis had many small spurious regions unlike REMPS which produced a 'cleaner' and more homogenous detection.

The texture based method was effective at locating the presence of damage as may be observed from Figure 5.12, however it did not perform as well as the colour based methods at defining the extent of damage which resulted in poor *DR*, *MCR* and δ values in Table 5.3. The exception to this was for the low turbidity and high light image (Figure 5.6c) where the texture analysis outperformed the colour based methods. While the high light and the shiny metallic surface created luminous complexities that misled the colour based methods, the high light illuminated and brought out some of textural properties of the surface which benefitted the texture analysis technique.

The performance of the damage detection techniques were affected by the shape and surface type of the specimens. Specimens with a flat surface performed quite well in comparison to the cylindrical and spherical surfaces as there was less severe luminous complexities such as bright-spots and light fall-off. Light fall-off occurs when less light reaches parts of the curved surface, which are noticeably darker as a result. The underside of the spherical specimens were particularly dark due to self-shadowing. These effects misled and hampered the detection algorithms. The textured concrete surface (e.g. Figure 5.6(i)) had a lot of false detections as the coarse aggregate was similar in colour to the damaged area. However, the texture analysis algorithm performed well for this case.

While the visibility conditions had a major impact on the detection of cracks, the performance of surface damage techniques do not rapidly decline with deteriorating visibility condition. Instead, the results indicate that choice of technique is more critical.

5.3.3 3D Shape Recovery using Stereo Vision

Stereo vision is the process of recovering depth from images, typically taken by two parallel cameras that are separated by a horizontal distance. The recovery of 3D structure using stereo vision requires two sub problems to be solved: the correspondence problem, in which image points corresponding to the same real world object are matched, and the reconstruction problem, in which the matched image points are reconstructed into 3D information.

The correspondence problem is often difficult because of ambiguous correspondences between points in the two images. This is especially problematic for uniform surface types where the lack of distinct features causes a high number of false matches. In other instances, correspondence cannot be established because a region is occluded in one of the images. The correspondence problem has been extensively researched and a wide range of stereo matching algorithms have been proposed. Scharstein and Szeliski (2002) provide a thorough taxonomy of the most notable stereo correspondence algorithms.

It is during this correspondence phase where the effects of lighting and turbidity will be noticed. For this reason, this section focuses on the output from the correspondence stage which is a disparity map (translatable to a range image) that tells how far each point in the physical scene is from the camera.

This section compares the performance of three types of stereo correspondence algorithms. The first is PaLPaBEL (a Pyramidal Loopy Propagated BELief method using a Markov Random Field (MRF)), which is introduced in Section 4.4 of this thesis. It is a hierarchical (coarse-to-fine) algorithm that operates on an image pyramid, where results

from coarser levels are used to obtain increasing finer results as successive levels. The MRF model takes into account the differences between pixel intensity values between corresponding points and the spatial relationship between the horizontal disparities. The goal is to find a piecewise smooth horizontal disparity map consistent with the observed data which minimises the total energy. The second method is based on the well-known (Birchfield and Tomasi, 1998) matching cost that is insensitive to image sampling. Rather than just comparing pixel values shifted by integral amounts (which may miss a valid match), this method compares each pixel in the reference image against a linearly interpolated function of the other image. It does not rely on any smoothness constraints, but rather, the disparity is computed by selecting the minimal (winning) aggregated value at each pixel. The final method is a global optimisation technique called scanline optimisation (Scharstein and Szeliski, 2002) that operates on individual horizontal slices and optimises one scanline at a time. The overall matching cost along the scanline is minimised.

These techniques are applied to the stereo imagery featuring a concrete cube in Figure 5.7 and the resulting disparity maps, or depth maps, are shown in Figure 5.14. The performance levels are quantified in Table 5.4. The performance is based on comparing the disparity maps with a ground truth, which is build by visually matching corresponding vertices of the cube. The performance for each algorithm is measured by the extent of the deviations from the resulting ground truth disparity map. Deviations greater than 5 pixels are regarded as a badly detected pixel while pixels within 5 pixels of the ground truth disparity are counted as good pixels. The percentage of good pixels and the percentage of bad pixels will always add up to 100%.

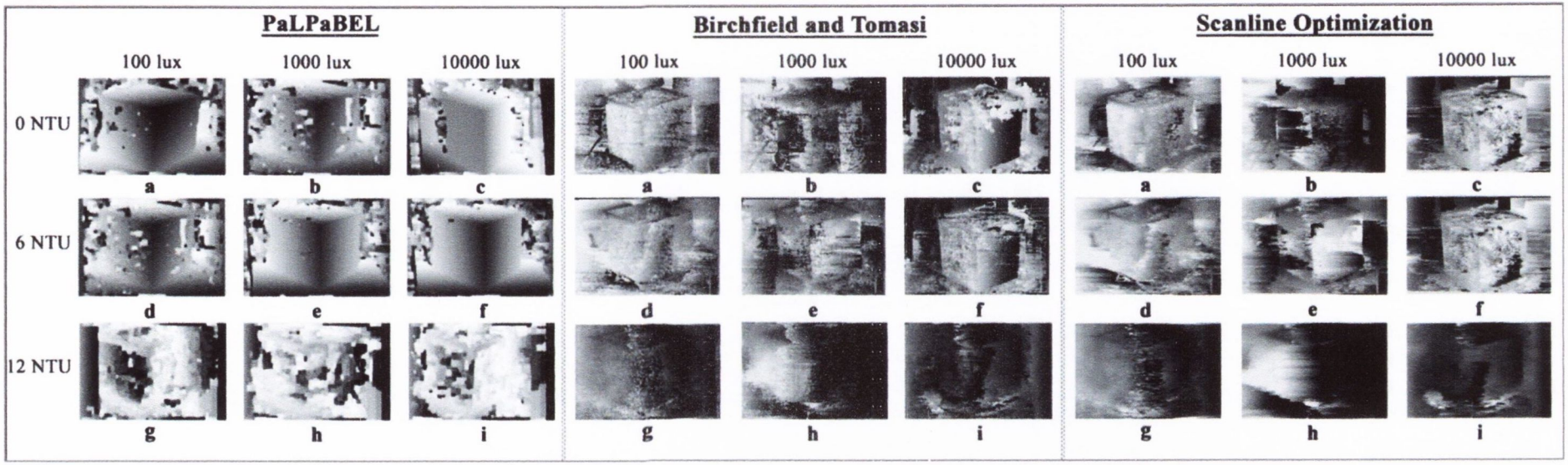


Figure 5.14 Disparity maps for the stereo imagery in Figure 5.7.

Table 5-4 Performance of the stereo correspondence algorithms.

Image	Condition	PaLPaBEL		Birchfield and Tomasi		Scanline Optimisation	
		Good Pixels	Bad Pixels	Good Pixels	Bad Pixels	Good Pixels	Bad Pixels
a	Low Light, Low Turbidity	91.0%	9.0%	46.1%	53.9%	52.9%	47.1%
b	Medium Light, Low Turbidity	89.4%	10.6%	26.7%	73.3%	49.1%	50.9%
c	High Light, Low Turbidity	91.2%	8.8%	45.2%	54.8%	66.4%	33.6%
d	Low Light, Medium Turbidity	78.7%	21.3%	60.4%	39.6%	69.9%	30.1%
e	Medium Light, Medium Turbidity	91.7%	8.3%	65.2%	34.8%	62.1%	37.9%
f	High Light, Medium Turbidity	90.5%	9.5%	67.3%	32.7%	68.6%	31.4%
g	Low Light, High Turbidity	44.0%	56.0%	66.6%	33.4%	72.1%	27.9%
h	Medium Light, High Turbidity	41.6%	58.4%	59.4%	40.6%	62.6%	37.4%
i	High Light, High Turbidity	44.9%	55.1%	75.5%	24.5%	70.9%	29.1%

It is apparent from the disparity maps in Figure 5.14 and analysis of the results in Table 5.4 that PaLPaBEL is far superior to the other techniques in this case. This may be attributed to a number of factors. Firstly, the concrete surface is quite uniform which creates ambiguous matches. This issue is compounded at higher turbidity and light levels when some of the surface detail is masked or washed out. In these cases, the inclusion of a smoothness term in the algorithm is particularly valued. Unlike the PaLPaBEL algorithm, the Birchfield and Tomasi approach does not take any smoothness term into account, while the scanline optimisation method only considers horizontal smoothness without considering vertical inter-scanline smoothness, which produces a streaky appearance in the depth maps. PaLPaBEL performs quite well up until the highest turbidity level at which point it breaks down due to the complete lack of contrast in much of the scene. This indicates that the operating limit has been reached.

The performance of the stereo matching techniques had varying degrees of success when applied to the other shapes and surface types in the repository. For the cylindrical shapes, only the centre-front regions of the specimens could be properly matched as the perspective difference between the left and right cameras became more pronounced for points on the cylinder that were further away from the centre. The spherical shaped objects also had this problem, as well as suffering from poor lighting on the underside of the sphere due to self-shadowing. Like the concrete surface, the relatively uniform plastic and metallic surfaces benefitted greatly from algorithms that incorporated a smoothness term. The uncontrolled specimens in the repository are all richly textured, which explains why they were successfully matched in most cases.

For SHM applications, the ultimate objective is generally to obtain a fully metric reconstruction. The disparity map found from the correspondence phase can be converted

to a 3D shape once the stereo geometry and camera parameters are known. The stereo system can be pre-calibrated using a well-known checkerboard procedure (Zhang, 2000) or through self-calibration using the static scene as a constraint on the camera parameters (Zhang et al., 1994). Reconstruction is demonstrated in Figure 5.15 which shows what can be achieved under good conditions, i.e. acceptable turbidity levels and adequate lighting, and using an efficient technique, such as PaLPaBEL.

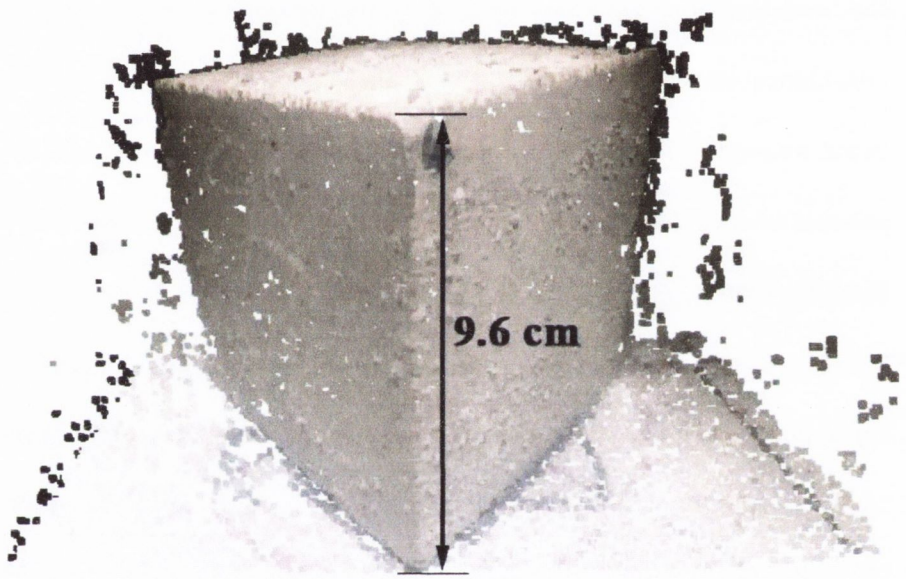


Figure 5.15 Reconstructed shape using PaLPaBEL.

In this example, the known dimensions of the cube - 10 cm for each side, correspond quite well with the dimensions from that of the reconstructed cube which had a side length measured at 9.6 cm. This example demonstrates that a good understanding of the operating conditions and an informed choice of stereo correspondence algorithm are important for obtaining good results.

5.4 Conclusions

While image based methods are beneficial in terms of efficiency and convenience, knowledge of their performance levels for a host of operating conditions is crucial. The need for an online image repository that provides inspectors and researchers with a platform to efficiently gauge the performance of image processing based damage detection methods under realistic underwater operating conditions and on relevant specimens is thus greatly felt. ULTIR allows inspectors to make informed decisions when assessing the feasibility of adopting image based approaches under a given set of environmental conditions and the knowledge acquired from using the repository can be used in real world inspections to by creating conditions that are conducive to good performance. Although the turbidity levels are generally uncontrollable in practice, appropriate lighting can be easily obtained. The resource also assists researchers when developing and evaluating new or existing image algorithms intended for application underwater.

Image processing methods have applications in many of areas of SHM including crack detection, surface damage detection, and 3D shape recovery using stereo-vision; where a wide and diverse array of algorithms have been devised for each category. However, there exists no standardised approach for choosing a technique that can best deal with unfavourable environmental conditions. ULTIR helps inspectors to identify techniques that are likely to perform well in the more challenging conditions. Various algorithms are compared in this chapter and specific features that favour good detection are singled out and discussed.

The results show that the choice of image processing method is an important factor for some categories, especially for 3D shape reconstruction, while environmental conditions

are the most critical factor for detecting cracks. Specific attributes of strong performing image algorithms are identified such as the incorporation of smoothness constraints which are helpful when handling the noise-contaminated and fuzzy appearance of images in high turbidity conditions. The results reveal that in many cases high turbidity can be partially mitigated by having bright light source. It is expected that ULTIR will continue to evolve as new imagery is added and more techniques are evaluated in the near future.

The creation of this resource serves as one step towards improving the quality of inspections by characterising the on-site conditions. Building upon this, the quality and consistency of image based inspections can be further improved through the development of a protocol that deals with the technical aspects of image acquisition in an underwater infrastructural setting.

Chapter 6

Protocol Establishment and Full Scale Testing

6.1 Introduction

In the previous chapter, an image repository was created that addresses how the on-site operating conditions affect the performance of image algorithms. The other notable aspect that affects the performance of image algorithms is the ability of the inspector or diver to effectively and consistently acquire good quality imagery. This chapter introduces a protocol that describes a set of best practice guidelines for obtaining imagery suited for quantitative image processing applications. Such a protocol is needed to maintain the long-term integrity of an Infrastructure Management System (IMS), which is dependent on the quality and consistency of the input information.

The protocol is put into practice as part of multi-disciplinary experiment that is conducted in a large scale testing facility in Boulogne, France, for the purpose of recovering the shape of a submerged cylinder covered with artificial marine growth using the PaLPaBEL technique. Following this, PaLPaBEL is validated by applying it to the submerged part of a real world structure located in Cork harbour.

6.2 Underwater Stereo Imaging Protocol

This protocol aims to describe the procedural method for acquiring underwater images that are to be used for quantitative analysis. The following protocols are established:

- Image storage
- Imaging protocol
- Diving protocol
- Combined protocol for underwater inspections.

Special attention is given to stereo based 3D shape recovery systems. Stereo systems involve increased operational complexity due to the fact that two cameras must be configured and also synchronised.

6.2.1 Image Storage

The creation of an image library for image storage requires a set of predefined guidelines to ensure that all contributions are consistent. This, in turn, ensures an organised and manageable library. The proposed image storage protocol addresses the following aspects:

- File format
- Metadata and information to be recorded
- File naming convention and cataloguing.

There are a wide variety of image formats available, each offering different advantages and drawbacks. Typically there is a trade-off between the image file size and the amount of information retained in it. As storage space is inexpensive nowadays, this protocol recommends capturing images in both JPEG and RAW format, if applicable.

JPEG images, which have a ".jpg" extension are one of the most popular image formats. JPEG is naturally suited for image analysis and processing tasks as the compacted file size allows images to be accessed quickly and put little strain on the computer's memory reserves during processing tasks. JPEG uses a method of lossy compression which means that the image quality degrades slightly after the image has been saved. The degree of compression can be controlled; enabling the user to choose how much detail should be sacrificed for an associated reduction in file size. Minimal loss of image quality can generally be achieved with a compression ratio of 10:1 from RAW to JPEG.

The RAW image format should be retained for archival purposes. RAW images contain minimally processed data from the image sensor. They are capable of storing a greater level of information from a scene (i.e. wider dynamic range and colour gamut) than other image formats. The RAW format differs depending on the camera model and camera manufacturer, although each format contains essentially the same data and metadata. For example, the RAW format for Canon cameras have file extensions *.crw* or *.cr2,* while Nikon cameras have file extensions *.nef* or *.nrw*. An additional step, if desired, would be to convert the original RAW formats to an open standard and well supported format, namely the Digital Negative (DNG) format, which is a popular and freely available format developed by Adobe Systems™. The DNG format has some notable benefits. Its popularity ensures that it well be well supported in the future which is important as some of the images captured today may be of interest over the course of a structures lifespan,

possibly several decades. Since there is a wide range of proprietary RAW formats, it is hard for applications and programs to guarantee future compatibility with them all, especially for some of the lesser known and lesser used RAW formats. Thus, the additional step of converting to the DNG format from the original RAW format would be of particular value to users that have, or expect to have, imagery acquired from a number of devices, and would like to unify the RAW formats into a common format that retains all of the original information. However, RAW images have quite a limited role overall as they are intended only for archival/back-up purposes and for use in rare cases where the extra detail offered by RAW images over JPEG images is needed. Everyday tasks, such as image processing and analysis are performed on the more manageable and light-weight JPEG images.

A comparison between JPEG and RAW images is shown in Figure 6.1. It may be observed that the RAW image exhibits greater tonal definition, however, the difference is barely perceivable.

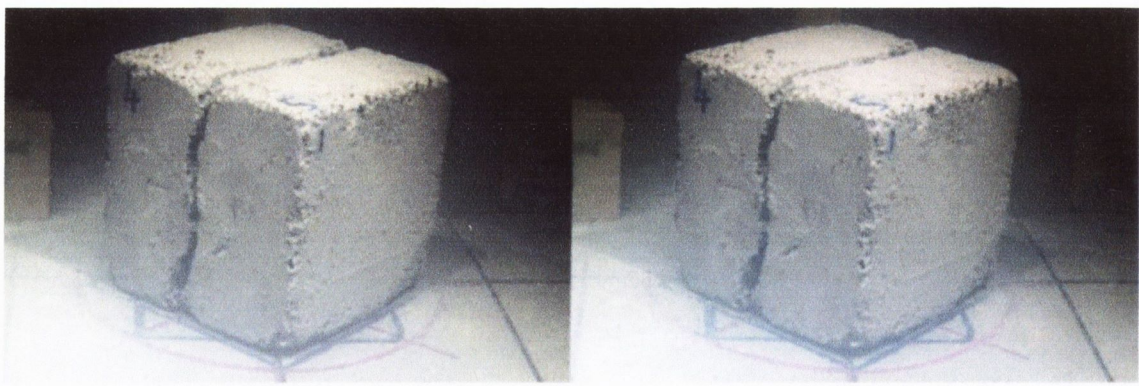


Figure 6.1 JPEG image vs. RAW image.

As there may be a variety of camera/video recording models contributing to the database, some devices may not support outputting to the RAW format. In this case, the native

format should be stored for archival purposes and a JPEG duplicate should be created for processing purposes. In the case of video, still frames should be extracted at relevant intervals and saved as JPEG files as analysing ever frame is unnecessary. The intervals will be primarily determined by the speed of the camera relative to the subject. If a recording device is moving quite quickly then more frames should be extracted. Generally, extracting three frames per second should be sufficient for most cases. As before, the original video should be stored in its native format. Both imagery and video have accompanying metadata which contains useful information about the content and context of the file.

Metadata is automatically embedded into each digital photo file. It provides information such as the time and date of capture, camera model, exposure information etc. To ensure that the integrity of the database is maintained, it is vital that the time and date of all contributing cameras are precisely set as this is provides a convenient way for identifying the synchronised stereo image pairs.

Some metadata must be manually added such as the baseline distance for stereo imaging, or the camera calibration data in a separate text file that is clearly linked to associated imagery. The specific nature of this metadata will vary according to the task. In the case of stereo imaging, the imagery should be clearly identified as coming from the left or right cameras, or simply be placed in separate folders.

6.2.2 Imaging Protocol

The first objective of establishing imaging protocols will involve specifying the technical requirements of capturing images of underwater marine structures. These technical requirements deal with the following aspects:

- Overcoming challenging environmental effects
- Underwater housings
- Choice of camera settings; exposure, aperture, ISO etc.

Underwater imaging must try to overcome the challenging environmental conditions (Kay, 2003). Firstly, there are optical issues with regards using cameras enclosed an underwater housing. There will be refraction as the light rays change mediums from water to the glass port of the underwater housing, and from the glass port to the air within the housing. As a result of this, the images will be slightly distorted. This issue can be accounted for in the case of stereo imaging by calibrating the cameras as the calibration stage takes any distortion into account.

Secondly, there is a loss of colour and contrast when the subject is separated from the camera by any significant distance, even in perfectly clear water. Water absorbs the red component of light to a greater extent which results in underwater subjects having a blue-green tinge. Colour diminishes with distance so subjects further away will appear indistinct and devoid of colour detail. This issue may be addressed by using artificial lighting in the form of strobes to restore some lost colour, or by photographing the damaged region from as close up as possible. An added complication of using artificial lights is the phenomenon of backscatter, where the light reflects off particles in the water. The best method for

limiting backscatter is positioning the lights away from the axis of the camera lens. This avoids the situation whereby the water immediately in front of the camera is exposed to a high intensity light source, while still allowing the subject to be suitably illuminated. Photographing the subject from relatively a close distance minimises any loss of colour.

The minimum distance between the camera(s) and the subject is constraint by the minimum focusing distance of the lens. A camera will not be able to focus on the subject properly if it is closer than this minimum focusing distance. The minimum focusing distance is typically specified on the lens; however in practice, good camera-subject distances typically exceed the minimum focusing distances. A recommended choice of working distance is in the region of 1 m – 1.5 m from the subject as this enables close-up, detailed imagery to be obtained. It is worth noting at this point that image stitching algorithms are widely available that can combine individual images with other partially overlapping images to create a high resolution photo mosaics that provide both a high level of information and greater context.

The distance between the camera and the subject will have an effect on the depth of field. A large camera-subject distance will lead to a greater depth of field, while closer distances will result in a shallower depth of field. Additionally, lenses with shorter focal lengths will produce greater depths of field. It is recommended that normal lenses are used (typically lenses having focal lengths of around 50 mm, although this varies slightly based on factors such as the crop factor and the camera's sensor size), as the resulting imagery does not suffer from significant distortion. This is in contrast to wide angle and telephoto lens, which have shorter and longer focal lengths than normal lenses respectively, both of which introduce perspective distortion in the imagery. For a wide angle lens, objects that are near to the lens appear abnormally large, while distant objects appear abnormally small.

Conversely, telephoto lens compress the relative distances between objects in the scene as distant objects appear abnormally large in comparison to closer objects. Extracting quantitative information from photographs in light of these aberrations is thus problematic. Although post-processing software exists that can fix these distortions, it is best avoided at source by choosing a normal lens.

There are cases where a wide angle lens or a telephoto lens (long lens) may be more appropriate. A wide angle lens provides a wide field of view which is useful in reduced visibility conditions where a short distance must be kept between camera and subject as it allows a sufficient portion of the scene to be captured and with a large depth of field, while a telephoto lens is useful when limited accessibility means that inspectors cannot get close to a structure but still wish to capture detailed, zoomed-in imagery (Stroebel and Zakia, 1993).

Lenses may also be partitioned into two general categories: zoom lenses and prime (fixed focal length) lenses. Zoom lenses have variable focal lengths (e.g. 18 mm - 70 mm). They offer greater versatility than prime lenses which have fixed focal lengths (e.g. 50 mm), however, comparably priced prime lenses generally have better optical quality and faster lens speeds. The lens speed refers to the maximum aperture of a lens which controls how much light reaches the sensor. Fast lenses have a large aperture which allows more light to reach the sensor at the expense of a shallower depth of field, meaning only objects within a confined range will be in focus and will appear acceptably sharp. A small aperture will produce a greater depth of field, however the resulting image will be dark/under-exposed unless the shutter stays open for an extended period of time enabling enough light to reach the sensor. The length of time that the shutter remains open for light to impinge on the camera's sensor is known as the shutter speed. A high degree of blur will likely result if the

shutter is allowed to remain open for too long a period (i.e. a slow shutter speed), especially for handheld cameras. As a rule of thumb, the minimum shutter speed should be no slower than the inverse of the focal length (i.e 1/50 second for a 50 mm lens).

One other important camera setting is the ISO, which measures the sensitivity of the image sensor, with common settings being: 100, 200, 400, 800, 1600, 3200. Higher ISO settings tend to be used in darker situations to amplify the available light; however this comes at a cost of increased noise in the imagery. This protocol recommends the user to use the ISO that is the native sensitivity of the sensor, which is usually indicated by the camera manufacturer or can be found experimentally by comparing photographs from a certain camera captured at different ISOs. Choosing an ISO that is much higher than the recommended ISO will result in too much grain in the imagery. As an example, for the main camera model used throughout this thesis - the Canon 600D, it is recommended to keep the ISO below 400, although ISOs up to 800 produce images that are useable.

Many modern cameras have simplified the process of choosing good combinations of the aperture, ISO, and shutter speed for a given scene through various automatic exposure modes and the use of through-the-lens (TTL) metering. However, it is important to impose certain limits on these settings in order to achieve a sharp image and to avoid adverse effects such as motion blur and excessive noise. The best combination of camera settings will depend on the task at hand. For the purpose of photographing cracks and 2D surface damages, where the damage is typically is on a single plane, the emphasis lies on ensuring the image is properly exposed and there is no motion blur. As there is there is little depth variation in the scene, the depth of field does not need to be that large and therefore does not need to be prioritised. The damaged region of interest should remain in sharp focus as long as the focal plane approximately coincides with the plane of the damaged surface.

With this in mind, the largest available aperture should be used in order to allow for a faster shutter speed that reduces the risk of motion blur.

For stereo photography of 3D damages, it is necessary that the depth of field encompass the full range of points on the 3D shape of interest. As a result, the cameras must be configured to provide a larger depth of field than in the case of the crack and 2D damage forms. This will involve selecting a smaller aperture, and balancing it against a slower shutter speed to maintain proper exposure. Recommendations of the limits are summarised in Table 1. Figures 6.2-6.3 show images with varying values of aperture and shutter speed. These figures showcase the deleterious effects on image quality by exceeding the stated limits. If it is found that these limits are insufficient, increasing the scene illumination, having a smaller focal length and/or increasing the distance between cameras and subject are measures that can be undertaken to increase the depth of field.

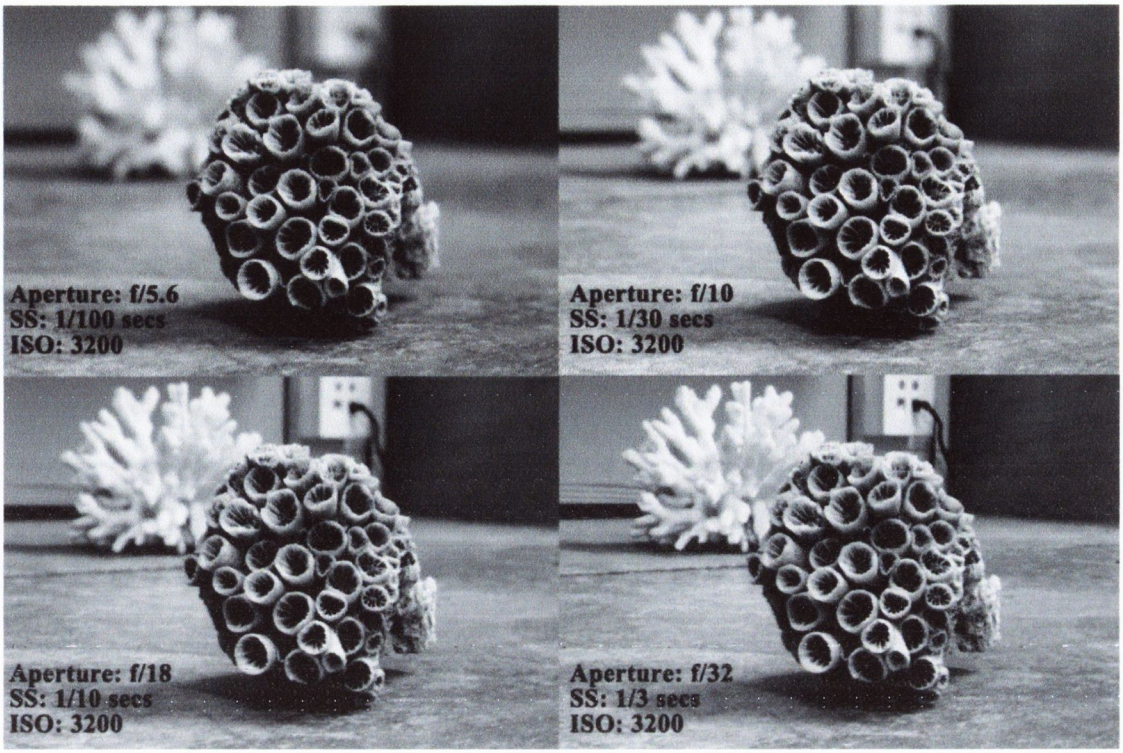


Figure 6.2 Varying aperture captured with a 55 mm lens; large apertures (e.g. f/5.6) result in blurry background objects whilst small apertures (e.g. f/32) have a greater depth of field.

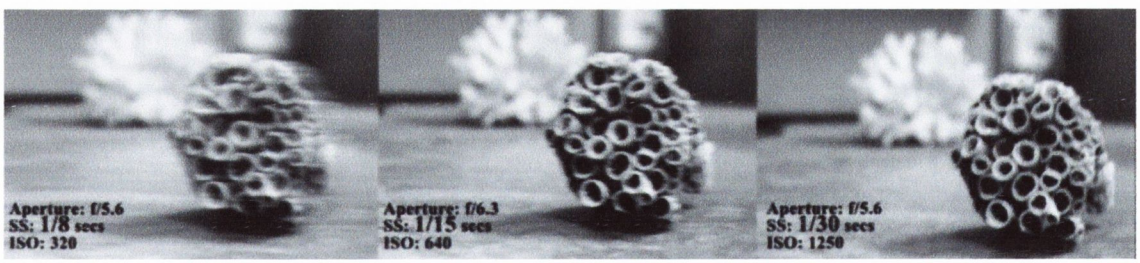


Figure 6.3 Varying shutter speed whilst undergoing gentle shaking with a 55 mm lens *f*/5.6; images having slower shutter speeds are more affected by motion blur.

Table 6-1 Recommended ranges of camera settings for image analysis purposes.

Camera Setting	Recommended Range
Aperture	The largest available aperture should be used that keeps the region of interest in focus. This could be a large aperture for 1D cracks and 2D planar damages (e.g. f/1.8). For stereo photography, a smaller aperture will be required (e.g. f/8 or smaller)
Shutter Speed	As a rule of thumb, the minimum shutter speed for handheld cameras is the inverse of the focal length, i.e. minimum Shutter Speed (secs) = 1/Focal Length (mm). So for a 50mm lens it is 1/50 sec. The imagery obtained from the cameras should be reviewed at regular intervals. If it is apparent there is too much motion blur present in the images, the shutter speed should be adjusted to a faster setting.
ISO	The recommended ISO setting is the native sensitivity of the sensor as specified by the camera manufacturer. An ISO value of 400 is a good compromise for photographing in dimly-lit conditions whilst still controlling noise/graininess. In the case of stereo imaging, the ISO should be fixed at the same value in both cameras.
Focal Length	It is recommended that a normal lens is used as the resulting imagery does not suffer significant distortion.

For all damage types, a good strategy would be to first fix the ISO to the camera's native sensitivity, followed by choosing a shutter speed that is the inverse of the focal length (aforementioned rule of thumb), and finally, choosing an aperture that provides the correct exposure. If the depth of field is too shallow or the images are still too dark, it is a sign that additional artificial light sources are required.

6.2.3 Diving Protocol

It is vital that any unnecessary time spent underwater by the diver is kept to a minimum. With this in mind, the diver should be presented with a clear and concise brief outlining the task at hand and specifying the target sites and certain things to look out for. The proposed protocol addresses the following aspects which should be included in the brief:

- Lighting and turbidity considerations.
- Choosing the appropriate distance between cameras and subject
- Logistical considerations (testing equipment prior to usage, route planning etc).

Both lighting and turbidity are crucial factors which affect the underwater visibility and consequently the image quality (Mahiddine et al., 2012). Artificial lighting in the form of underwater strobe lights are required in dim lighting conditions, especially if the subject is bumpy/rugged in which case the lights would assist with revealing regions that would otherwise remain in shadow.

Water is seldom optimally clear, and the dissolved and suspended matter can reduce visibility by both absorption and scattering of light. While turbidity may not be easily reduced, there are some precautions which can be taken to offset the deleterious effect in relatively high turbid waters. Firstly, caution should be taken in shallow waters to avoid disturbance with the sea/river bed which may unsettle fine sediments through either direct contact or from turbulence created from the ship. Anchoring the vessel downstream away from the inspection site can also prevent additional sediments that would reduce the water clarity. Secondly, the distance between the cameras and the subject under consideration should be reduced. In seas and oceans, the water is generally clear so measures to counteract poor visibility need not be considered.

Choosing the distance between the subject and the diver is a trade-off between a number of factors. The ideal distance should be kept in the range 50 - 150 cm; however this will vary depending on the following circumstances:

- Visibility
- The size of the subject in the scene
- Acceptable error tolerance
- Baseline shift between the cameras in the case of stereo imagery

When the underwater visibility is poor, the diver can photograph in close proximity to the subject (up to 30cm). Going any closer than 30cm leads to large perspective differences between the stereo image pairs which can hamper the matching process and result in a myriad of occluded regions (Matthies and Shafer, 1987). In clear underwater conditions, the diver can photograph from a distance up to 2.5m before the stereoscopy breaks down and the error tolerance becomes unacceptably high (Olofsson, 2010).

If the object under inspection is quite large (e.g. a wide diameter pile) and it is wished to include the whole structure within an image than photographing from further back will provide better context bearing in mind that the 2.5m limit should not be exceeded. If, however, a greater accuracy of the structure's macro geometry is required than the diver should keep within 1.2 metres of the subject.

Finally, the baseline distance for stereo imagery will influence the choice of subject-camera distance. While the baseline shift will normally be fixed at a certain width according the constraints of the available equipment, it is important to note the effect it has on the accuracy. Theoretically, a wider separation between the cameras results in a lower percentage error in the depth estimation. However, the advantages of having a wide separation are mitigated by the creation of large perspective differences as previously eluded to. The baseline shift should be in the range 10cm to 30cm. Additionally, the cameras should be aimed inwards at an angle θ (known as the vergence angle) such that their centrelines intersect approximately at the face of the subject as shown in Figure 6.4. This is to ensure that the cameras capture as many of the same points in both images as possible. The vergence angle has previously been used to model the error in depth by (Sahabi and Basu, 1996). For real world scenes, they found that vergence angles in the range 5° - 10° provided the lowest errors.

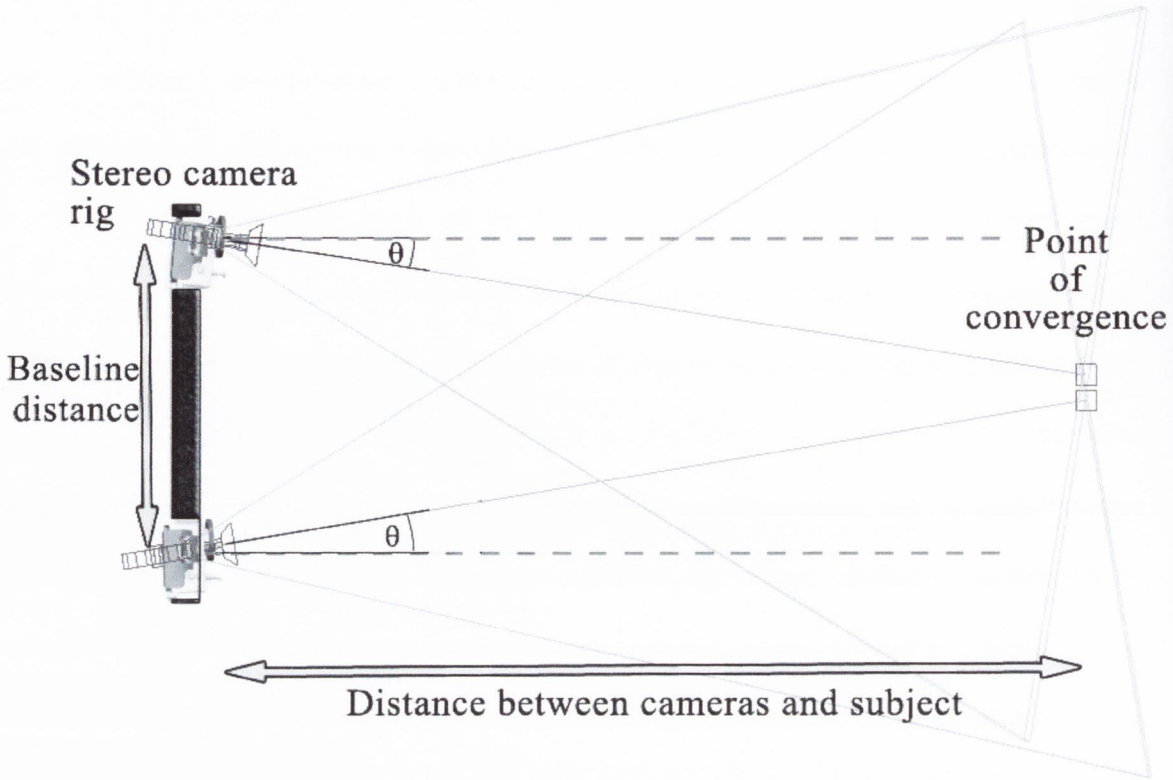


Figure 6.4 Stereo rig set-up.

All the equipment should be checked above water prior to each inspection. The underwater housings should be checked for any signs of leakage by firstly submerging them without their contents. If necessary, the o-rings should be relubed/replaced as per manual instructions. Care should be given to ensure that the time and date of the cameras are precisely set, that there is enough storage capacity in the SD cards and the battery is sufficiently charged. The cameras, ideally, should be focused at a point on the subject as illustrated in Figure 6.4. Appropriate settings should be configured for each camera/video recorder ensuring that the shooting modes in both are identical. It is advised to initiate filming immediately before the diver submerges as it is easier to control the simultaneous triggering of both cameras when above water. The captured imagery should be reviewed at regular intervals during the inspection and any adjustments should be made accordingly. Dives that produce sub-standard imagery should be repeated. Necessary props should be

prepared such as lighting equipment, an object of known dimensions to attach onto the structure or a checkerboard for calibration if applicable. Moreover, this would be a good opportunity to record the baseline distance and the dimensions of the object of known size such as a checkerboard/wand. Finally, the diver should familiarise himself/herself with the blueprint of the structure and identify any components that are of particular interest. A suitable route should then be planned based on this groundwork. In cases where the diver cannot photograph a particular component from all sides due to restricted access, blockage by other obstacles etc., he/she should endeavour to photograph as much of it as possible.

6.2.4 Combined Underwater Protocol

Greater care and attention is required for stereo imaging as two cameras must act in unison. A flowchart showing the summarised methodology for the combined underwater protocol for stereo imaging is shown in Figure 6.5.

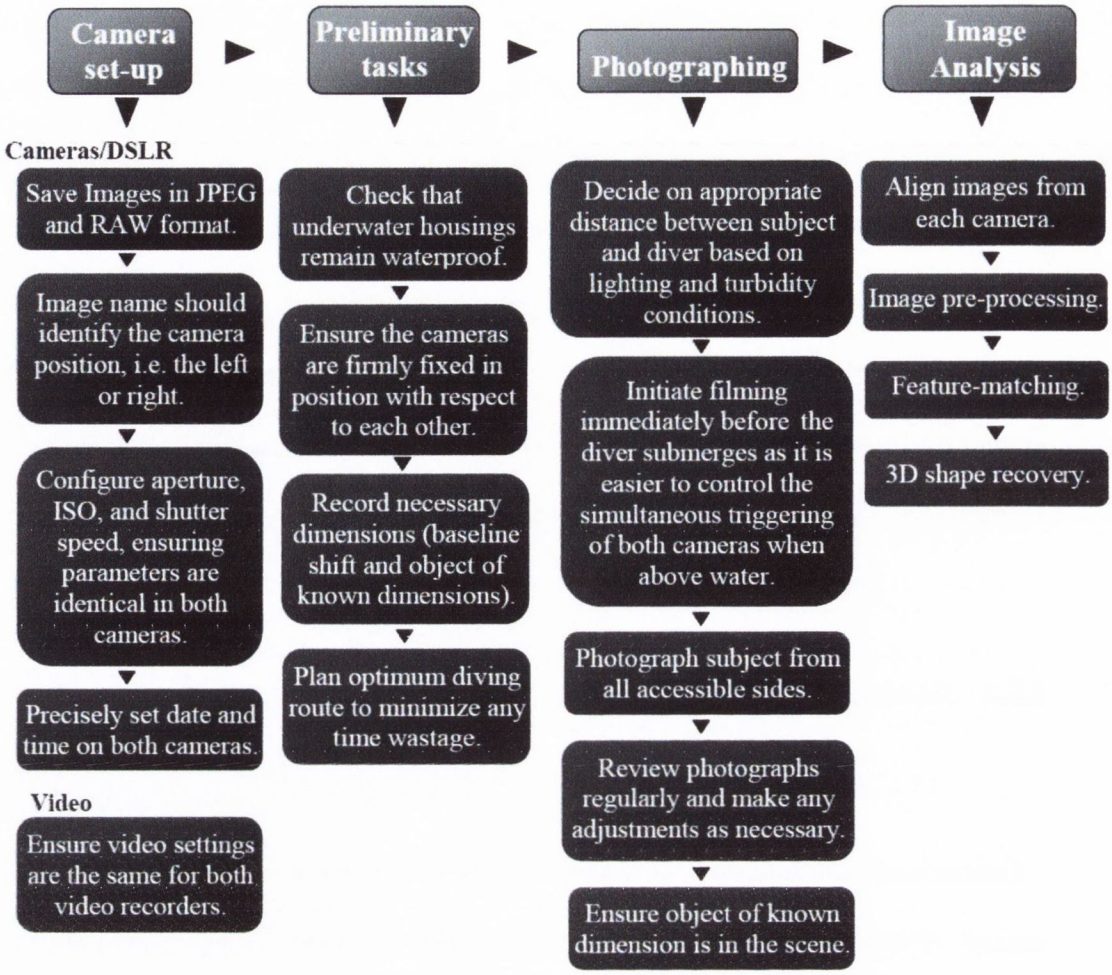


Figure 6.5 Key steps involved in the stereo imaging pipeline.

6.2.5 Conclusion

Problems in the acquired imagery, such as blurring or under/over exposure, are hard to fully appraise on-site as, in a lot of cases, the only way of reviewing imagery is by looking at small screens on the imaging devices. The true quality of the imagery can only be fully appreciated back at the office. Having to repeat inspections due to the sub-par performance of any NDT technique is a daunting and frustrating prospect. In particular, the effort and expense associated with undertaking underwater inspections warrants significant forethought and planning to ensure the every technical aspect of data acquisition is covered. To address this issue, a comprehensive protocol is developed for the first time that sets out best practices for underwater image acquisition. The protocol is developed based on personal experience, logical rationale and information from the literature. This developed protocol is used as a guide for the large scale experiment and the real world testing, which are discussed in the following sections.

6.3 Large Scale Experiment: Testing of 3D Shape Recovery Method in Water Basin

The developed protocol was utilised as part of multi-disciplinary experiment that was conducted in a large scale testing facility in Boulogne, France. The protocol was provided to a diver who was charged with capturing stereo imagery of a submerged cylinder covered with artificial marine growth. The protocol served as an important tool for providing guidance and for conveying the intricacies of stereo imaging to a diver who may not have been unfamiliar with the technique. PaLPaBEL was evaluated by applying it to the obtained stereo imagery in order to recover the shape of the roughened cylinder. The following sub-sections present the background of the experiment, the experimental procedure, and the results and conclusions.

6.3.1 Background of the experiment

Marine growth quickly forms on offshore structures soon after they have been installed. The thickness of marine growth increases at a rate of about 1 cm every year, although this varies depending on a number of factors such as the location of the structure and how much sunlight reaches the marine growth. For structures like jacket platforms, the increased thickness of structural members due to marine growth colonisation is considerable relative to the member's original diameter. Increasing the diameter of structural members, as well as the roughness coefficients, leads to increased hydrodynamic forces acting on the structure (Boukinda Mbadinga et al., 2007), as given by Morison's Equation:

$$F_{Morison} = \frac{1}{2}\rho C_d \phi \left((U_c + \chi)|U_c + \chi| \right) + C_m \frac{\rho \pi \phi^2}{4} \dot{\chi} \tag{6.1}$$

where $F_{Morison}$ is the force per unit length of the member, C_d is the drag coefficient, C_m is the inertia coefficient, ρ is the fluid density, ø is the member diameter, χ is the wave velocity, and U_c is the velocity of current. As such, the shape information is essential for determining the forces acting on a structure.

The aim of this experiment is to analyse the forces acting on a cylinder that is covered with artificial marine growth in a flowing water basin. This roughened cylinder has known dimensions: the inner cylinder is 200 mm in diameter, which is surrounded by rows of six 120 mm diameter spheres that are staggered and inserted to 1/3 of their diameter in the cylinder. The spheres therefore protrude beyond the face of the cylinder by 80 mm, giving a maximum outside diameter of 360 mm. The roughened cylinder is shown in Figure 6.6In the context of this thesis, the experiment will serve as an opportunity to test and evaluate the proposed protocol and gauge the level of attainable accuracy of the 3D shape recovery algorithms against the known geometry of the roughened cylinder. The roughened cylinder and the testing facility are shown in Figure 6.6.

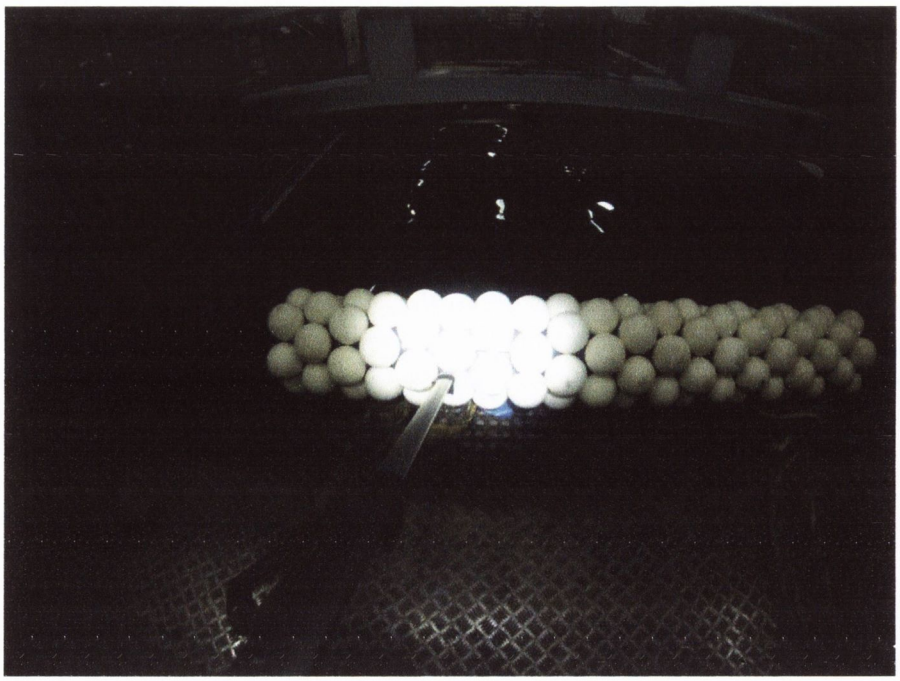


Figure 6.6 The roughened cylinder before submersion at facility in Boulogne, France.

6.3.2 Experiment procedure

The stereo imagery is acquired by a diver in accordance with the protocol outlined in Section 6.2.

6.3.2.1 Equipment

The stereo system is shown in Figure 6.7. This system consists of two GoPro HERO3+ Black Edition cameras, enclosed in underwater cases. These cameras are capable of capturing 12 MP photos at up to 30 frames per second. Two underwater flashlights provide additional light, which was necessary as the water basin is situated indoors. Finally, a stick, or a wand as it is commonly referred to as, is attached to the stereo system, which is always visible in both cameras. The wand length should be recorded. This is needed for calibration purposes using standard procedures such as (Mitchelson and Hilton, 2003; Pribanic et al, 2009), and for getting a scale factor that relates pixel dimensions to real world units.

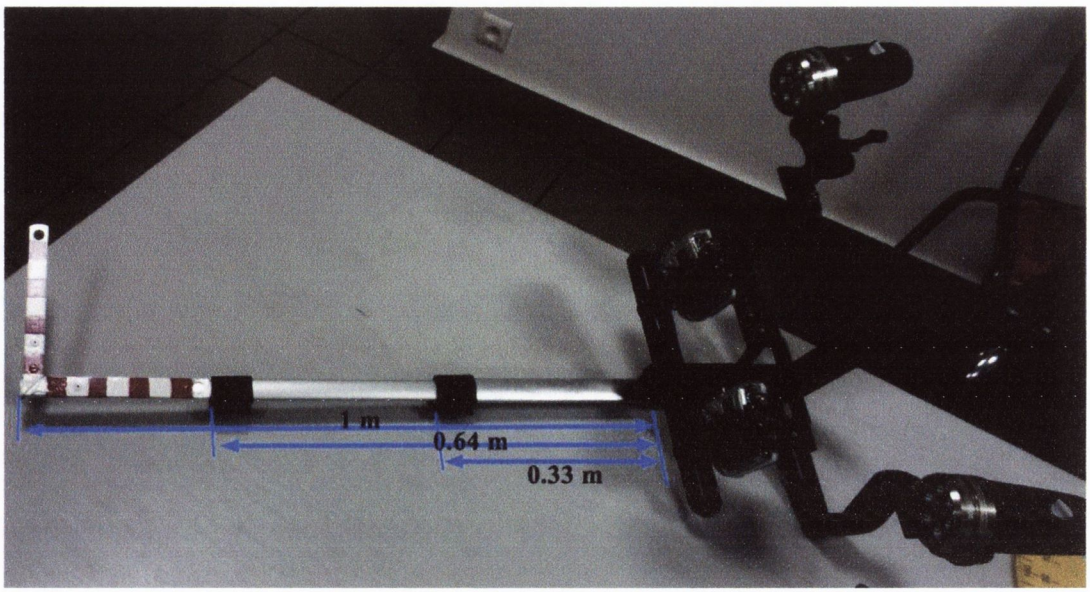
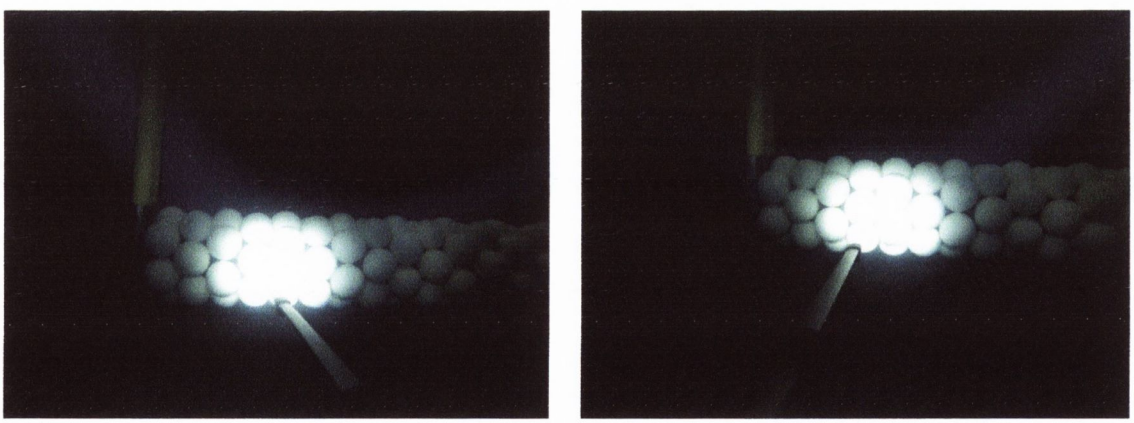


Figure 6.7 GoPro underwater stereo system.

6.3.2.2 Data Acquisition

The cameras were configured to capture 12 MP still-frame images and were carefully synchronised above water. The cameras were set to automatically capture images at a predefined time interval. Some of the acquired imagery is shown in Figure 6.8.

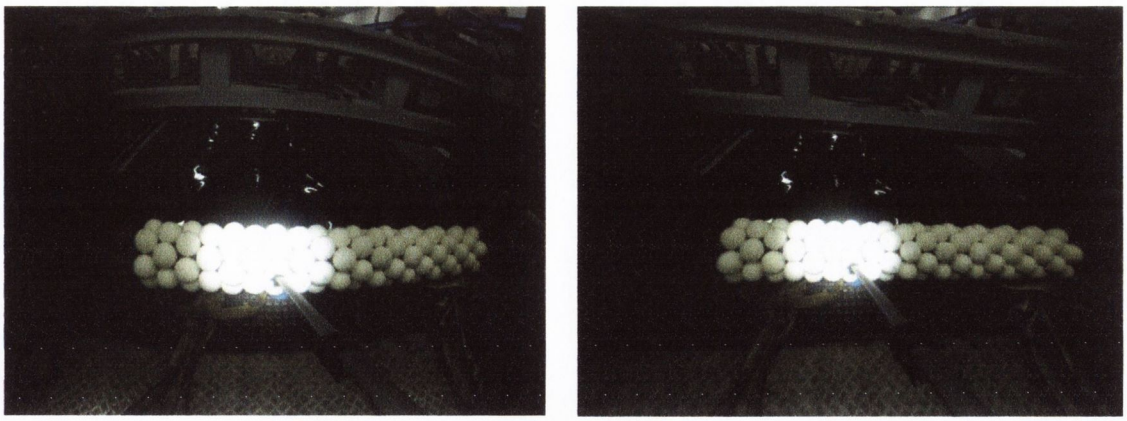


(a) **(b)**

Figure 6.8 Artificial marine growth pile in the (a) left, and (b) right camera.

6.3.2.3 Shape Recovery

There are a number of stages in the shape recovery process. The first stage involves correcting the images. Unprocessed images directly from the GoPro cameras have a distinctive fisheye effect. This is beneficial in some respects as it gives a large field of view (FOV), allowing more of the scene to be captured. However, this distortion impedes the stereo matching stage. The fisheye effect can be removed by applying a "lens adjustment", which is a non-destructive process that uses the lens profile to correct the imagery. There are many lens profiles available online, especially for popular makes of cameras and lenses. The result of distortion is best illustrated by using an image with a lot of structure, such as in Figure 6.9(a), where it is immediately apparent that the lines which should be straight are curved. It can be seen that the undistorted version of the image in Figure 6.9(b) fixes this problem.



(a) (b)

Figure 6.9 Undistortion process: (a) Original input image, (b) undistorted image.

The next stage is to rectify the images from the left and right cameras so that points in the left image lie on the same horizontal row as the same point in the right image. This is followed by the application of the PaLPaBEL stereo matching technique. These processes are discussed in Chapter 4. The rectified images are shown in Figure 6.10.

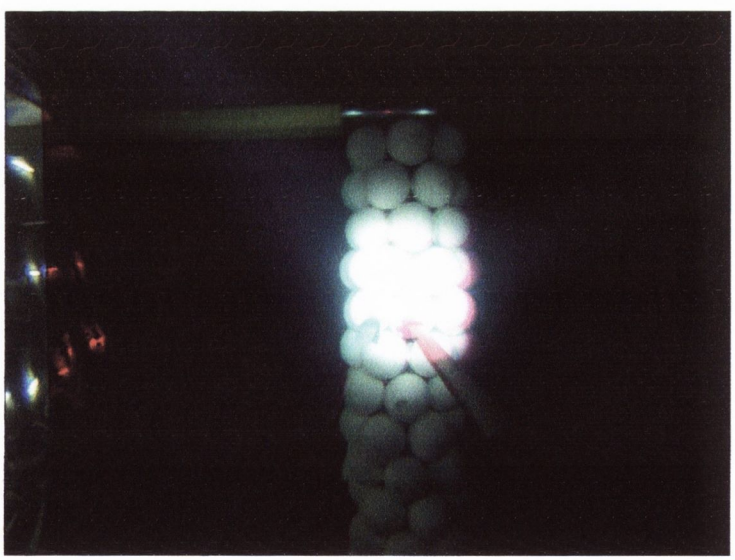


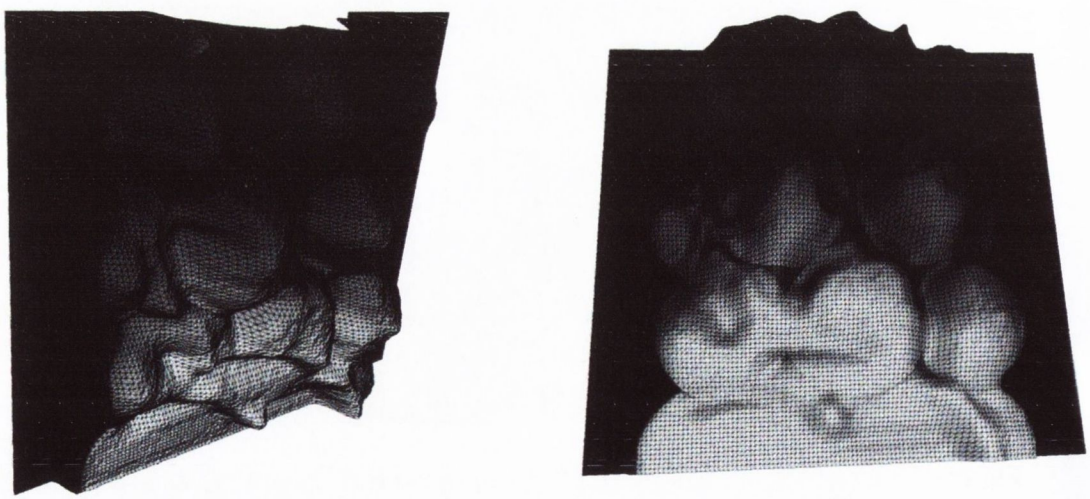
Figure 6.10 The left (cyan) and right (red) input images are automatically aligned onto a common image plane to simplify the problem of finding matching points.

6.3.3 Results and Discussion

The depth map obtained from applying PaLPaBEL to the stereo image pair in Figure 6.10 is shown in Figure 6.11. This depth map relates to a segment of the artificial marine growth pile, above the bright spot, where matching was successful.

Figure 6.11 Depth Map (Segmented Foreground).

The associated reconstruction is shown in Figure 6.12, which is based on calibration procedure outlined by Mitchelson and Hilton (2003).

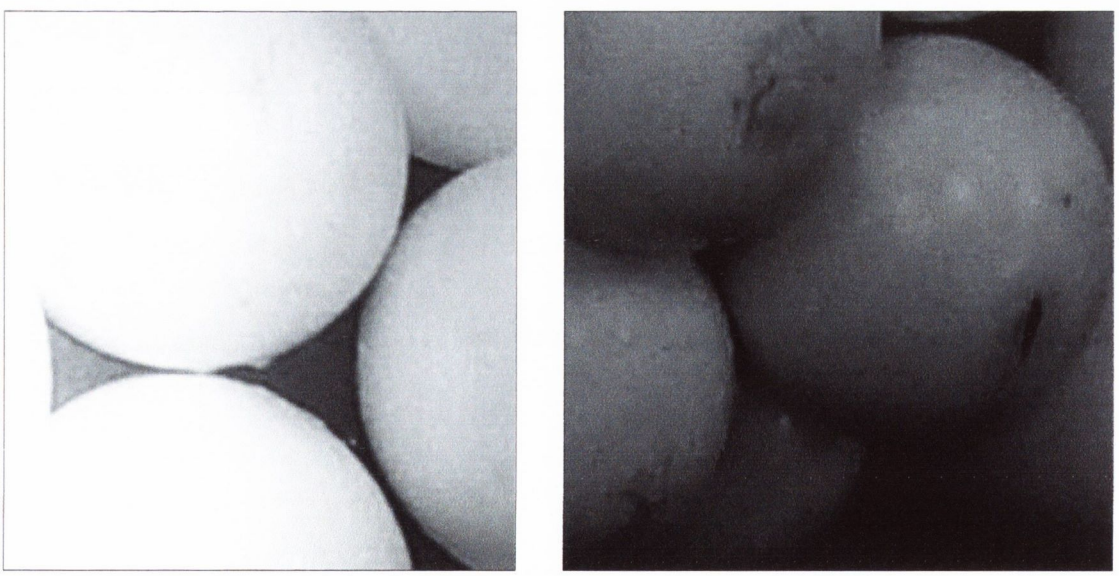


(a) **(b)**

Figure 6.12 Two views of the reconstructed surface.

With reference to Figure 6.12, the reconstructed surface reveals that the general shape is represented quite well near the top of the surface, although, the spheres do appear somewhat deformed. Reconstruction is unsuccessful towards the bottom of the surface as a consequence of the bright spot. The bright spot washes out details on the surface of the roughened cylinder making the task of stereo matching more challenging.

This problem is compounded by the relatively textureless surface of the roughened cylinder, which is conveyed in Figure 6.13(a). The surface does not have any prominent or distinctive points, which makes it hard to find the correct correspondence during the matching process and means that the matching algorithm must rely heavily on smoothness assumptions. It would help matching accuracy and reliability if there were more features on the surface of the spheres. By comparison, the region shown in Figure 6.13(b), which is slightly dirtier and has some distinctive markings performed significantly better at matching.



(a) **(b)**

Figure 6.13 Close-up view of the (a) textureless surface, and (b) the surface which gave better matches.

A complete 3D profile was created by rotating and applying the reconstructed surface displayed in Figure 6.12 in six positions around the cylinder to create a closed form shape as shown in Figure 6.14. In practice, the subject would have to be photographed from all sides to develop a full 3D shape profile, however, that was not possible in the case of this experiment due to accessibility issues. The volume of the reconstructed shape was compared with that of the known volume of the roughened pile for the same length of pile, which covered three rows of spheres.

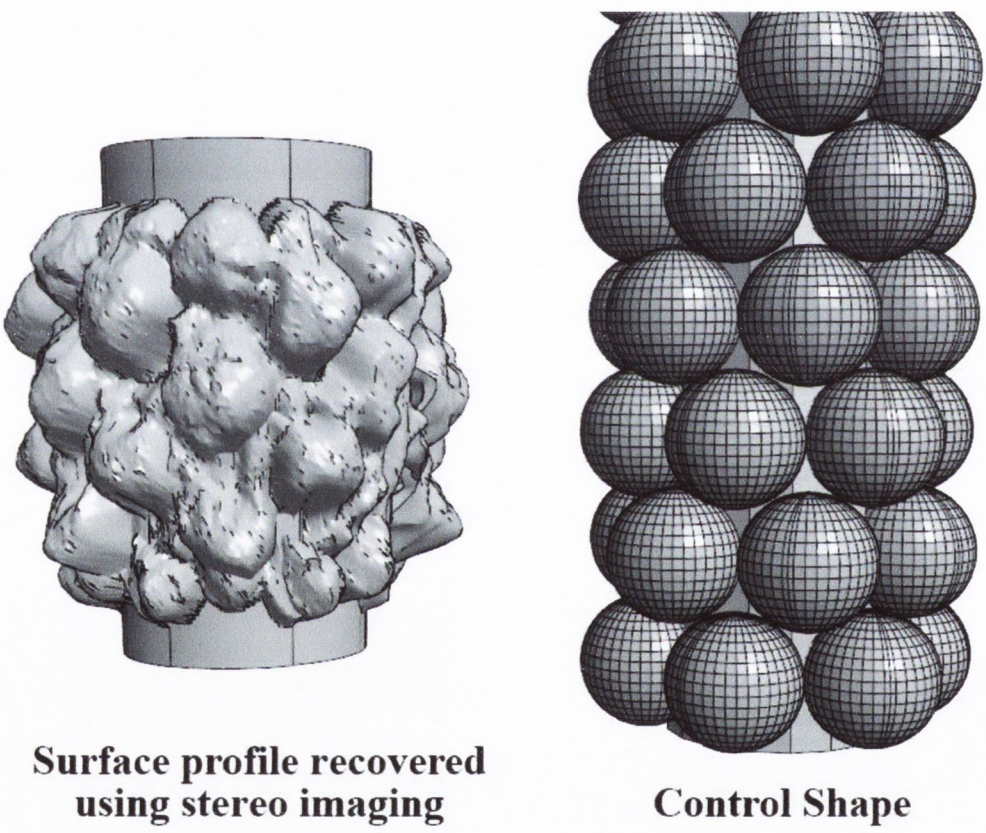


Figure 6.14 Reconstructed 3D shape of the roughened cylinder and actual shape.

The volume of the roughened cylinder for 3 rows of spheres is 25,899 cm^3 (control volume), while the volume of the reconstructed object is 21,072 cm^3. There are a few things that affected this accuracy.

6.3.4 Conclusion

This experiment provided an opportunity to test the performance of the developed stereo matching algorithm PaLPaBEL on a large scale object under realistic conditions, as well as implementing the protocol devised in Section 6.2. The purpose of the experiment was to analyse the hydrodynamic forces acting on a cylindrical structural component affected by marine growth. A stereo system comprised of two lights and two GoPro cameras was employed to characterise the shape of the roughened cylinder, with a view to evaluating its performance before being used on part of real world structures. The GoPro cameras offered great convenience given that they are a low-cost, pre-built system. The lights created some challenges as bright spots were a feature of the acquired imagery. However, this may be partially attributed to the relatively glossy surface of the roughened cylinder, which also suffered from being poorly textured. Both of these aspects had a deleterious effect on the performance of PaLPaBEL. Natural instances of marine growth are typically richly textured so this is not envisaged as being a problem for real world structures. In future, steps will be taken to diffuse the light so that it spreads more evenly throughout the scene and is not focused on a small area. Overall, the performance of the shape recovery method was good at identifying the general shape of the roughened cylinder excluding areas of high light intensity.

6.4 Real World Structure: Shape Recovery of Marine Growth

Testing PaLPaBEL in a laboratory setting allowed for significant control, however, this control came at the expense of sacrificing realism in certain ways. The poorly textured surface of the artificial marine growth pile and the clear water conditions are generally not reflective of reality. Increased validity is achieved by conducting a field experiment where turbidity is a factor and the surfaces are natural. This section applies the PaLPaBEL technique to a marine growth affected pier located in Cork Harbour, Ireland. This study is conducted in order to validate the developed stereo based shape recovery technique under real word conditions and on a real world structure. The following sub-sections i) provide an overview of the structure under consideration and the underwater conditions at the testing site, ii) outline the image acquisition procedure, and iii) present and discuss the results.

6.4.1 Structure and Conditions at the Test Site.

Cork harbour is a natural harbour at the mouth of the River Lee. The structure under consideration is a long-serving pier that is primarily used for recreational and leisure purposes. An aerial view of the testing site is shown in Figure 6.15, and view from the shore is shown in Figure 6.16.

Figure 6.15 Aerial view of testing zone in Currabinny, Cork Harbour. Source: Google Maps (retrieved 15/09/2014).

Figure 6.16 View of the structure at ground level.

It was a clear and sunny day when the testing was carried out, with the luxmeter recording measurements in excess of 10,000 lux above the water. Thus, there was no requirement for additional artificial light sources as ambient light was sufficient in the shallow waters. The turbidity varied somewhat around the structure. The turbidity was higher in the shallower water near the shoreline as a lot of particles were being kept in suspension thanks to the breaking action of the waves. The water was a bit clearer at the main testing location on the structure which was at a depth of about 1 m; however, it was still above 12 NTU. As a result, the imagery had to be acquired at a close range from the subject, which was not ideal as less of the scene could be captured.

6.4.2 Experiment Procedure

The stereo imagery was obtained by photographing the submerged part of a pile from a number of perspectives. The equipment and the process are detailed in this section.

6.4.2.1 Equipment

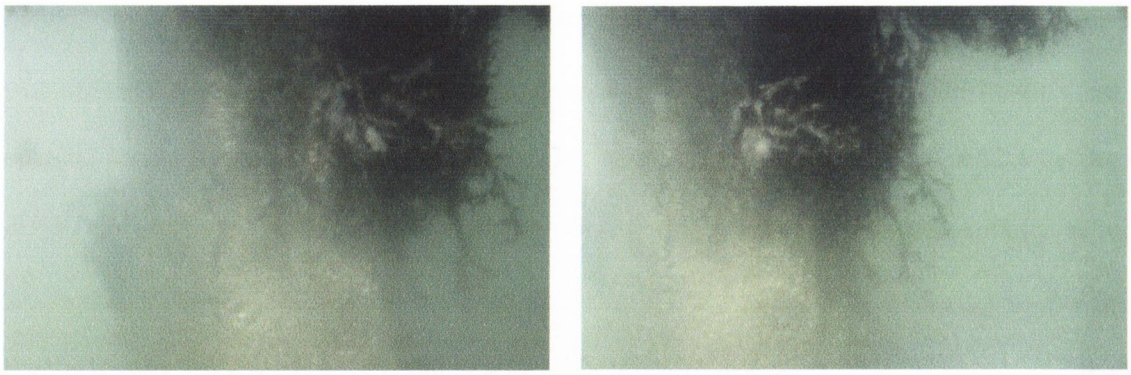
The stereo system is shown in Figure 6.17. This system consist of two Canon 600D DSLR cameras, enclosed in underwater cases, which are securely attached to a graduated stereo bar. The graduated stereo bar allows the baseline distance to be easily recorded. In this case, the centres of the cameras were separated by 15 cm. Knowledge of the baseline is necessary for obtaining the scale factor that converts the dimensions of the reconstructed 3D shape to real world units. The Canon 600D cameras were both configured to simultaneously capture high resolution 18 MP images in RAW and JPEG format at a time interval of two seconds. The high resolution enables good depth sensitivity as each pixel in the image may be viewed as representing a 3D point in space. Any resulting 3D

reconstruction will therefore be able to offer more precise depth information. The cameras were in shutter priority mode in accordance with the protocol and had a shutter speed of 1/20 seconds. The minimum focal length was selected, which was 18 mm, so that as much of the scene could be captured as possible.

Figure 6.17 The stereo system used to capture the imagery.

6.4.2.2 Data analysis

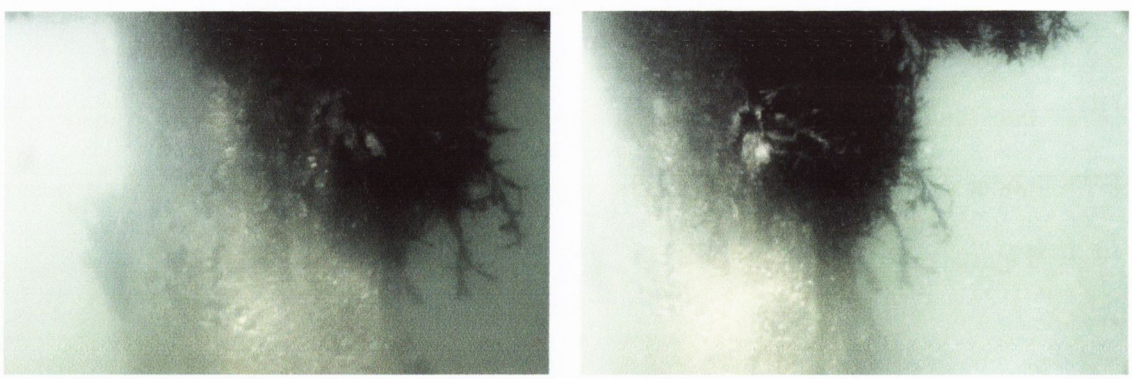
Marine growth can be classified as either hard or soft. Hard fouling types include barnacles, mussels, corals etc. Soft types include seaweed, algae, hydroids etc. The pile under consideration is predominantly affected by soft fouling organisms as shown in Figure 6.18.



(a) (b)

Figure 6.18 Marine growth affected pile from the (a) Left, and (b) Right cameras.

It may be observed that the high turbidity creates a fuzzy image lacking in contrast. In this case, the imagery greatly benefits from the image enhancement stage of the PaLPaBEL technique. The image enhancement stage employs histogram stretching to restore some of the lost contrast. The resulting enhanced images are shown in Figure 6.19 for the left and right images.



(a) (b)

Figure 6.19 Stereo images following enhancement.

6.4.2.3 Shape Recovery

The resulting depth map following application of PaLPaBEL on this stereo pair is shown in Figure 6.20. The successfully matched area is quite small relative to the size of the original images. This stems from the fact that the images had to be acquired at a close distance from the subject, which means that a smaller area of the pile is overlapped in both the left and right images.

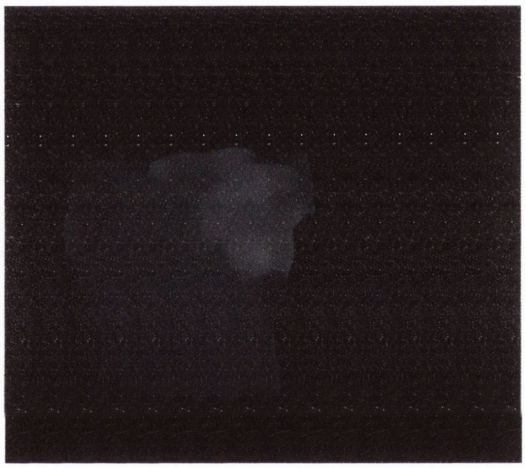


Figure 6.20 Depth Map (Segmented Foreground).

The associated reconstruction is shown in Figure 6.21, which is based on the auto-calibration technique described by Faugeras et al. (1992). Auto-calibration is a convenient way of obtaining the intrinsic camera parameters of a stereo system as it does not require the use of any props, which can be awkward and cumbersome for a diver to carry out. Auto-calibration relies on static scene to act as a constraint for the five degree-of-freedom pinhole camera model. The five intrinsic parameters are the focal length in pixel dimensions, aspect ratio, skew, and the two principle points. A minimum of three views are needed for complete calibration assuming fixed intrinsic parameters between views (i.e.

the focal length, or any other parameter, does not change between views). However, in reality, the principle points can usually be estimated to be at the image centre and quality modern imaging sensors and optics also provide further prior constraints such as zero skew and unity aspect ratio. Integrating these priors will reduce the minimum number of views required to two. The only additional information needed is the baseline distance which enables a properly scaled reconstruction. The accuracy of the camera parameters obtained through auto-calibration is usually lower than that of the conventional checkerboard based pre-calibration procedures; however, the practical advantages often outweigh this reduction in accuracy. Moreover, errant camera parameters do not always translate to an appreciably errant reconstruction.

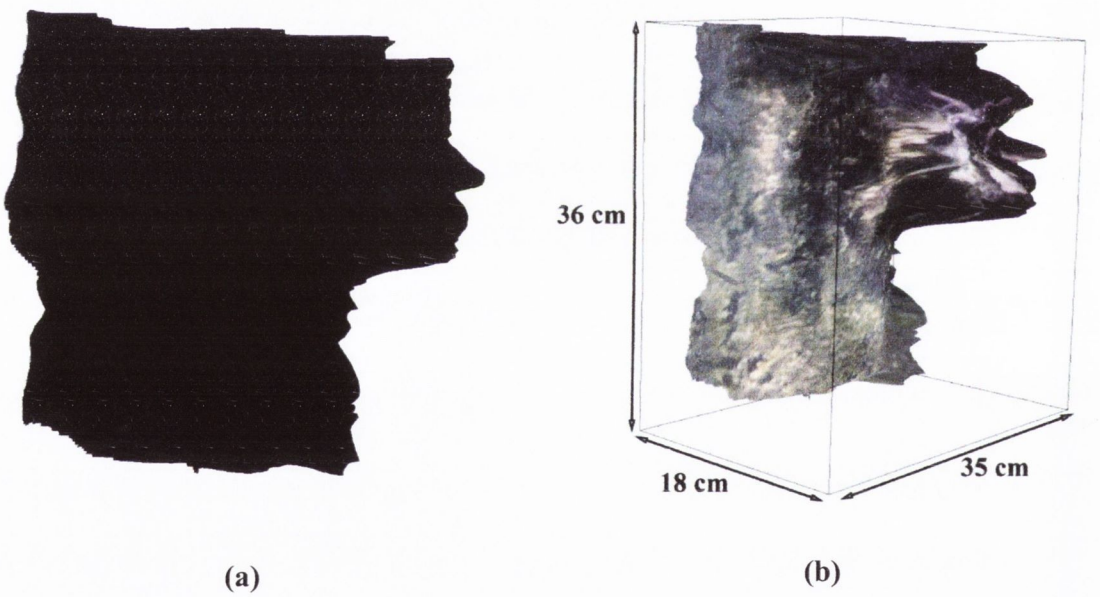


Figure 6.21. Two views of the reconstructed surface, with (b) showing the dimensions.

The reconstructed surface is a good reflection the true pile shape and the shape of the protruding marine growth. As the testing was done on an uncontrolled surface (i.e. where the shape was not known), a quantitative measure of the performance could not be determined, however, the recovered dimensions showed good visual agreement with

reality. Closer examination of the reconstructed surface shows that it is quite 'bumpy' in places which should be quite flat, however this effect is mild. PaLPaBEL performed decently overall despite the high turbidity. It benefitted from the richly textures surface of the pile and the natural marine growth.

6.5 Conclusions

Selecting appropriate camera settings can be an overwhelming task for inspectors given the wide array of options. Selecting the most suitable combination of these settings for a given environment involves a trade-off between minimising negative image quality factors such as blur and noise, whilst retaining good brightness and ensuring enough of the subject is in focus. Ineffective image acquisition practices directly affect the performance of image algorithms. The protocol introduced in this chapter provides guidance for obtaining imagery suited for quantitative image processing applications. Such a protocol is needed for maintaining the quality and consistency of the input information from inspections.

The protocol is put into practice as part of multi-disciplinary experiment that is conducted in a large scale testing facility in Boulogne, France, for the purpose of recovering the shape of a submerged cylinder covered with artificial marine growth using the PaLPaBEL technique. The results demonstrate that an image based approach is suitable for recovering 3D shape as long as the lighting is not excessive and the surface is sufficiently textured to facilitate good stereo matching. Following this large scale experiment in a laboratory setting, PaLPaBEL was applied to a real world marine structure. The underwater conditions here differed from the laboratory setting as turbidity became a critical factor. Despite the high turbidity however, a reasonable shape reconstruction was nevertheless obtained.

Chapter 7

Conclusions

7.1 Summary of Research

This research aims to enhance the quality of image processing based techniques for inspection of infrastructure assets. The thesis, to a significant extent, addresses underwater inspection using image processing, which is a very important but not very developed field of study as of yet. This research advances the field in three ways. Firstly, new robust damage detection and assessment algorithms are developed that take into account the challenging conditions of an underwater environment. Secondly, the challenging underwater conditions are characterised through a repository driven approach, which maps the effect of lighting, turbidity, and surface type on the performance of techniques. Finally, a protocol has been developed, which outlines a set of best practice guidelines for obtaining imagery suitable for quantitative analysis in relation to underwater inspections. The work carried out in this thesis can be expected to be used by inspectors, owners and managers of marine engineering structures looking to rationalise and optimise the use of image methods as part of their inspection regimes. Although the developed techniques are focussed on underwater application, they are quite general and can be readily used in a much broader context.

7.2 Detailed Results

The contributions of the thesis and are listed in detail in this section.

The first part of the research deals with devising sophisticated methods for inspecting and assessing infrastructure elements. The developed image based techniques include a crack detection algorithm, colour and texture based damage detection algorithms, and a 3D shape recovery algorithm. The crack detection algorithm adopts a percolation based approach to automatically locate and quantify cracks in an efficient manner, removing the need for inspectors to manually undertake this tedious task. The developed algorithm offers improved efficiency over traditional percolation based crack detection methods by searching in a concise space, whilst still obtaining good results thanks to novel classification criteria. The performance of the technique is investigated in the presence of challenges such as luminous complexities, ambiguous cracks, spurious markings and poor visibility. High turbidity is particularly harmful to detection accuracy. In low light and low turbidity conditions (100 lux, 0 NTU), a detection rate of 89.8% can be achieved for a cracked concrete specimen, however this drops to a detection rate of 62.6% in low light and high turbidity conditions (100 lux, 12 NTU), for similar values of the misclassification rate.

Two algorithms were developed for locating and identifying damaged regions based on colour and texture respectively. The texture analysis method involves generating a texture feature vector for each pixel in the image based on information derived from a Gray Level Co-occurrence Matrix (GLCM) matrix as well as directly from the pixel intensity values. The pixels are consequently classified through non-linear Support Vector Machines (SVM) models. The colour based method, known as REMPS (Regionally Enhanced Multi-Phase

Segmentation) relies on a multi-phase segmentation methodology that incorporates features from three standard image processing and data analysis techniques. REMPS and the texture analysis technique were both validated on images featuring a variety of damage forms obtained from inspection campaigns. Each image presented a challenge of some sort, such as difficult viewing angles or shinny metallic surfaces, which created luminous complexities. It was found that REMPS consistently achieved good performance and was capable of producing better defined and more homogenous regions of interest. Slight improvements can be made to the detection accuracy of REMPS by segmenting in the $L^*a^*b^*$ colour space and by adopting High Dynamic Range (HDR) imaging. A δ value of 0.16 was obtained for $L^*a^*b^*$ colour space applied to HDR image of pitting corrosion, versus the slightly worse value of 0.18 for a Standard Dynamic Range (SDR) version of the image. Texture analysis fared a lot worse for the same image, recording a δ value of 0.39. Visually, it was apparent that the output from the texture analysis method was often affected by isolated extraneous pixels. Texture analysis only performed well in specific cases where the damage was distinguishable from the undamaged background based on colour or texture characteristics.

A stereo matching algorithm with a novel pyramidal formulation was developed for the purpose of recovering 3D shape. The developed technique, PaLPaBEL, is a pyramidal loopy Belief Propagation (BP) method that operates on a Markov Random Field (MRF). Moreover, the pyramidal scheme allowed full image resolution to be used efficiently, which enabled accurate and precise depth information that can fully utilise the original image resolution. PaLPaBEL was evaluated as part of a large scale experiment that was conducted in a testing facility in Boulogne, France, for the purpose of recovering the shape of a poorly textured submerged cylinder covered with artificial marine growth. The volume of a reconstructed section of the artificial marine growth cylinder was calculated to be

21,072 cm^3, which compared against a value of 25,899 cm^3 the control volume over same length of cylinder. There are a few things that affected this accuracy. In particular, the high brightness, which induced bright spots that masked details on the surface of the cylinder, as well as a poor texture. Following this, PaLPaBEL was validated on a real word structure in Cork Harbour, Ireland, where there was high levels of turbidity. Nevertheless, it was found that a shape reconstruction that showed good visual agreement could be obtained.

A video tracking method was customised and demonstrated on a vibrating suspension bridge spanning a river in Cork, Ireland, for the purpose of identifying the natural frequency. The video tracking method found that the natural frequency was ~2 Hz, which was corroborated by accelerometers placed at the mid-span of the bridge. From these results, it may be concluded that video tracking offers a convenient and accurate way of identifying the natural frequency of bridges characterised by large amplitudes and low frequencies.

A large Underwater Lighting and Turbidity Image Repository (ULTIR) was generated from a set of controlled experiments conducted in a water tank. This is the first repository of its kind for underwater inspection. ULTIR provides a means of characterising the effects of turbidity and lighting on an image analysis method. ULTIR was populated with images relating to three categories: 1) 1D crack detection, 2) 2D surface damage detection, and 3) 3D shape recovery using stereo-vision. The imagery contained within each category was captured under three lighting levels and three turbidity levels, resulting in nine images for each specimen. The controlled specimens were specifically chosen to cover a wide range of geometric and photometric properties. Specimens featuring real instances of damage were also included. The performance of various algorithms in each category were examined under the different lighting and turbidity levels and specific features that

favoured good detection were singled out and discussed. Algorithms that incorporated a smoothness term were found to benefit greatly. Additionally, the results show that the choice of image analysis method is an important factor, especially for 3D shape reconstruction where it is reaffirmed that the incorporation of smoothness constraints is helpful when handling the noise-contaminated and fuzzy appearance of images in high turbidity conditions. For example, PaLPaBEL achieved 91.7% rate of accurate pixels compared to 62.1% for Scanline Optimisation for the same turbidity and lighting levels. It is also found that high turbidity can be partially mitigated by having a bright light source. For instance, PaLPaBEL produced a 90.5% rate of accurate pixels at high light (10,000 lux) and medium turbidity (6 NTU) compared to 78.7% at low light (100 lux) and the same turbidity.

Finally, a protocol was developed that promotes effective image acquisition practices. Special attention was given to stereo imagery as this requires two cameras to be configured and synchronised as oppose to just one. The protocol was put into practice as part of the large scale experiment in Boulogne and the real word experiment in Cork Harbour. The protocol also served an instrumental role during the laboratory testing as it conveyed important details and the intricacies of stereo imaging to the diver who may have been unfamiliar with the technique.

7.3 Critical Assessment of the Developed Work

The role of light is seen to play a crucial role in the success of image processing algorithms. The repository considers light of three different magnitudes: 100, 1000, and 10,000 lux, however, in reality the magnitude of the light is only one aspect. Another issue is the shape of the light beam. A diffuse light will affect the performance of image algorithms differently than a light beam that is focussed on a spot. Given this, a possible recommendation would be to also assess the effect that the uniformity of light has on.

Accurate detection of damage poses many practical problems from an imaging perspective. A variety of reasons can cause a technique to perform badly, including the poor environmental conditions, the nature of the damage itself, poor image acquisition practices etc. This is especially problematic for methods like 3D stereo imaging which solve a number of sub-problems like rectification, matching, and calibration. If failure occurs at any of these stages, then the entire technique is jeopardised. However, taking steps during the image acquisition phase to recognise potential problems can help alleviate this problem.

7.4 Future Work

The research work carried out in this thesis could be extended in a number of ways.

First, there is scope to expand ULTIR to include more specimens, especially specimens that feature real instances of damage. Adding more damaged specimens in this manner would serve to make the repository more relevant and comprehensive. Additionally, assessing the effect that the uniformity of light has on performance may be a worthwhile endeavour.

Secondly, an avenue of future research that offers significant potential involves extending texture analysis to the critical field of marine growth classification. The two key parameters associated with marine growth colonization of offshore structures are the thickness and the roughness, which are both needed for structural reliability computation. These parameters vary around and among structural components. Currently, stereo imaging is a validated means of obtaining the thickness. It may be worth exploring whether image analysis could be used to find a correlation between the perceived texture of the marine growth and the actual roughness coefficients. This would create a powerful tool that could improve marine growth time variant probabilistic models and allow for more accurate hydrodynamic loading assessments.

As for applications, a third potential avenue of research relates to the continued growth of the offshore renewable energy sector. In response this trend, it would be relevant to consider the role that imaging systems could play for the close observation and monitoring of these structures. As these structures often operate in large offshore farms, where the

water is comparatively clear, it would appear that image methods can prove an attractive way for monitoring the multiple targets efficiently and economically.

REFERENCES

Abdel-Qader, I., Yohali, S., Abudayyeh, O., and Yehia, S. (2008). "Segmentation Of Thermal Images for Non-Destructive Evaluation of Bridge Decks." *NDT & E International*, 41(5), 395-405.

Abdou, I. E., and Pratt, W. (1979). "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors." *Proceedings of the IEEE*, 67(5), 753-763.

Adams, R., and Bischof, L. (1994). "Seeded Region Growing." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6), 641-647.

Adeli, H., and Jiang, X. (2009). *Intelligent Infrastructure: Neural Networks, Wavelets, and Chaos Theory for Intelligent Transportation Systems and Smart Structures*, CRC Press, Taylor & Francis, Boca Raton , FL, USA.

Adewuyi, A. P., and Wu, Z. (2011). "Vibration-Based Damage Localization in Flexural Structures using Normalized Modal Macrostrain Techniques from Limited Measurements." *Computer-Aided Civil and Infrastructure Engineering*, 26(3), 154-172.

Agin, G. J. (1980). "Computer Vision Systems for Industrial Inspection and Assembly." *Computer*, 13(5), 11-20.

Anand, R., and Kumar, P. (2009). "Flaw Detection in Radiographic Weldment Images using Morphological Watershed Segmentation Technique." *NDT & E International*, 42(1), 2-8.

Baraldi, A., and Parmiggiani, F. (1995). "An Iinvestigation of the Textural Characteristics Associated with Gray Level Cooccurrence Matrix Statistical Parameters." *Geoscience and Remote Sensing, IEEE Transactions on*, 33(2), 293-304.

Baroth, J., Breysse, D., and Schoefs, F. (2011). *Construction Reliability*, Wiley, Hoboken, NJ, USA.

Barrois, B., and Wohler, C. (2007). "3D Pose Estimation Based On Multiple Monocular Cues." *Proc., CVPR '07. IEEE Conference on Computer Vision and Pattern Recognition, 18-23 June 2007*, IEEE, 3018-3025.

Belli, K., Wadia-Fascetti, S., and Rappaport, C. (2008). "Model based Evaluation of Bridge Decks using Ground Penetrating Radar." *Computer-Aided Civil and Infrastructure Engineering*, 23(1), 3-16.

Berger, H., Brackett, R., and Mittleman, J. (1983). "Underwater Inspection of Naval Structures." *Proc., OCEANS '83, Proceedings*, 29 Aug.-Sept. 1 1983, San Francisco, CA, USA, 555-559.

Bianco, G., Gallo, A., Bruno, F., and Muzzupappa, M. (2013). "A Comparative Analysis Between Active and Passive Techniques for Underwater 3D Reconstruction of Close-Range Objects." *Sensors*, 13(8), 11007-11031.

Birchfield, S., and Tomasi, C. (1998). "Pixel Dissimilarity Measure that is Insensitive to Image Sampling." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4), 401-406.

Blitz, J., and Simpson, G. (1996). *Ultrasonic methods of non-destructive testing*, Chapmann and Hall, London, UK.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). "A Training Algorithm for Optimal Margin Classifiers." *Proceedings of the fifth annual workshop on Computational learning theory*, ACM, Pittsburgh, Pennsylvania, July 1992, United States, 144-152.

Boukinda Mbadinga, M. L., Schoefs, F., Quiniou-Ramus, V., Birades, M., and Garretta, R. (2007). "Marine Growth Colonization Process in Guinea Gulf: Data Analysis." *Journal of Offshore Mechanics and Arctic Engineering*, 129(2), 97-106.

Brown, M., and Lowe, D. (2002). "Invariant Features from Interest Point Groups." *Proc., BMVC2002: British Machine Vision Conference 2002, 2-5 Sept. 2002*, British Machine Vision Assoc., 253-262.

Bruno, F., Bianco, G., Muzzupappa, M., Barone, S., and Razionale, A. (2011). "Experimentation of Structured Light and Stereo Vision for Underwater 3D Reconstruction." *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), 508-518.

Bruno, F., Bianco, G., Muzzupappa, M., Barone, S., and Razionale, A. V. (2011). "Experimentation of Structured Light and Stereo Vision for Underwater 3D Reconstruction." *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), 508-518.

Busby, F. R. (1979). "Underwater Inspection/Testing/ Monitoring of Offshore Structures." *Ocean Engineering*, 6(4), 355-491.

Butcher, J. B., Day, C. R., Austin, J. C., Haycock, P. W., Verstraeten, D., and Schrauwen, B. (2013). "Defect Detection in Reinforced Concrete using Random Neural Architectures." *Computer-Aided Civil and Infrastructure Engineering*, 29(3), 191-207.

Chan, T. F., and Vese, L. A. (2001). "Active Contours Without Edges." *Image processing, IEEE transactions on*, 10(2), 266-277.

Cheddad, A., Mohamad, D., and Manaf, A. A. (2008). "Exploiting Voronoi Diagram Properties in Face Segmentation And Feature Extraction." *Pattern Recognition*, 41(12), 3842-3859.

Cheng, H. D., Jiang, X. H., Sun, Y., and Wang, J. (2001). "Color Image Segmentation: Advances and Prospects." *Pattern Recognition*, 34(12), 2259-2281.

Choi, B., Han, S., Chung, B., and Ryou, J. (2011). "Human Body Parts Candidate Segmentation using Laws Texture Energy Measures with Skin Color." *Proc., 13th International Conference on Advanced Communication Technology: Smart Service Innovation through Mobile Interactivity, ICACT 2011, February 13, 2011 - February 16, 2011*, Phoenix Park, Korea, 556-560.

Choudhary, G. K., and Dey, S. (2012). "Crack Detection in Concrete Surfaces using Image Processing, Fuzzy Logic, and Neural Networks." *Proc., 2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI), 18-20 Oct. 2012*, IEEE, 404-411.

Cocito, S., Sgorbini, S., Peirano, A., and Valle, M. (2003). "3-D Reconstruction of Biological Objects using Underwater Video Technique and Image Processing." *Journal of Experimental Marine Biology and Ecology*, 297(1), 57-70.

Correa, S., Souza, E., Oliveira, D., Silva, A., Lopes, R., Marinho, C., and Camerini, C. (2009). "Assessment of Weld Thickness Loss in Offshore Pipelines using Computed Radiography and Computational Modeling." *Applied Radiation and Isotopes*, 67(10), 1824-1828.

Cortes, C., and Vapnik, V. (1995). "Support-Vector Networks." *Machine Learning*, 20(3), 273-297.

Cristianini, N., and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and other Kernel-based Learning Methods*, Cambridge University Press.

Cruz, P. J. S., and Salgado, R. (2009). "Performance of Vibration-Based Damage Detection Methods in Bridges." *Computer-Aided Civil and Infrastructure Engineering*, 24(1), 62-79.

Cusson, D., Lounis, Z., and Daigle, L. (2011). "Durability Monitoring For Improved Service Life Predictions Of Concrete Bridge Decks In Corrosive Environments." *Computer-Aided Civil and Infrastructure Engineering*, 26(7), 524-541.

D'Orazio, T., Leo, M., Distante, A., Guaragnella, C., Pianese, V., and Cavaccini, G. (2008). "Automatic Ultrasonic Inspection for Internal Defect Detection in Composite Materials." *NDT and E International*, 41(2), 145-154.

Dana, K. J., Nayar, S. K., van Ginneken, B., and Koenderink, J. J. (1997). "Reflectance and Texture of Real-World Surfaces." *Proc., Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 17-19 June 1997*, IEEE Comput. Soc., 151-157.

Debevec, P. E., and Malik, J. (2008). "Recovering High Dynamic Range Radiance Maps from Photographs." *ACM SIGGRAPH 2008 classes*, ACM, Los Angeles, California, 1-10.

Doebling, S. W., and Farrar, C. R. (2011). "Statistical Damage Identification Techniques Applied to the I-40 Bridge over the Rio Grande River." *Proc., Proceedings of the 1998 16th International Modal Analysis Conference, IMAC. Part 1 (of 2), February 2, 1998 - February 5, 1998*, SEM, 1717-1724.

Dover, W., and Bond, L. (1986). "Weld Crack Characterization on Offshore Structures using AC Potential Difference and Ultrasonics." *NDT international*, 19(4), 243-247.

Estes, A., and Frangopol, D. (2003). "Updating Bridge Reliability Based on Bridge Management Systems Visual Inspection Results." *Journal of Bridge Engineering*, 8(6), 374-382.

Faugeras, O. D., Luong, Q. T., and Maybank, S. J. (1992). "Camera Self-Calibration: Theory and Experiments." *Computer Vision — ECCV'92*, G. Sandini, ed., Springer Berlin Heidelberg, 321-334.

Felzenszwalb, P. F., and Huttenlocher, D. P. (2004). "Efficient Graph-Based Image Segmentation." *International Journal of Computer Vision*, 59(2), 167-181.

Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A. (1998). "Multiscale Vessel Enhancement Filtering." *Proc., Medical Image Computing and Computer-Assisted Intervention - MICCAI'98. First International Conference. Proceedings, 11-13 Oct. 1998*, Springer-Verlag, Berlin, Germany, 130-137.

Gadelmawla, E. S. (2004). "A Vision System for Surface Roughness Characterization using the Gray Level Co-Occurrence Matrix." *NDT and E International*, 37(7), 577-588.

Gallwey, T. J., and Drury, C. G. (1986). "Task Complexity in Visual Inspection." *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 28(5), 595-606.

Gangone, M. V., Whelan, M. J., and Janoyan, K. D. (2011). "Wireless Monitoring of a Multispan Bridge Superstructure for Diagnostic Load Testing and System Identification." *Computer-Aided Civil and Infrastructure Engineering*, 26(7), 560-579.

Gehle, R., and Masri, S. (1998). "Tracking the Multi-Component Motion of a Cable Using a Television Camera." *Smart materials and structures*, 7(1), 43.

Gill, R. S. (1999). "Ice Cover Discrimination in the Greenland Waters using First-Order Texture Parameters of ERS SAR Images." *International Journal of Remote Sensing*, 20(2), 373-385.

Giralt, J., Rodriguez-Benitez, L., Moreno-Garcia, J., Solana-Cipres, C., and Jimenez, L. (2013). "Lane Mark Segmentation and Identification using Statistical Criteria on Compressed Video." *Integrated Computer-Aided Engineering*, 20(2), 143-155.

Goldberg, L. (1996). "Diversity in Underwater Inspection." *Materials Evaluation*, 54(3), 401-403.

Goldberg, L. O. (1998). "Eddy Current Testing, an Emerging NDT Method for Ferritic Weld Inspection." *Materials evaluation*, 56(2).

Graham, A. A. (1990). "Siltation of Stone-Surface Periphyton in Rivers by Clay-Sized Particles from Low Concentrations in Suspention." *Hydrobiologia*, 199(2), 107-115.

Groves, D., and Connell, D. (1985). "Offshore Structure Fabrication Experience with Magnetic Particle Inspection." *NDT international*, 18(2), 85-88.

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). "Textural Features for Image Classification." *Systems, Man and Cybernetics, IEEE Transactions on*, 3(6), 610-621.

Heriansyah, R., and Abu-Bakar, S. (2009). "Defect Detection in Thermal Image for Nondestructive Evaluation of Petrochemical Equipments." *NDT & E International*, 42(8), 729-740.

Hillis, A. J., and Courtney, C. (2011). "Structural Health Monitoring of Fixed Offshore Structures using the Bicoherence Function of Ambient Vibration Measurements." *Journal of Sound and Vibration*, 330(6), 1141-1152.

Ho, N. (2012). "*Loopy Belief Propagation, Markov Random Field, Stereo Vision*". Available: http://nghiaho.com/?page_id=1366. Last accessed 25th Nov 2014.

Huang, Y., and Wu, J.-W. (2010). "Infrared Thermal Image Segmentations Employing the Multilayer Level Set Method for Non-Destructive Evaluation of Layered Structures." *NDT & E International*, 43(1), 34-44.

Iyer, S. R., Sinha, S. K., and Schokker, A. J. (2005). "- Ultrasonic C-Scan Imaging of Post-Tensioned Concrete Bridge Structures for Detection of Corrosion and Voids." *Computer-Aided Civil and Infrastructure Engineering*, - 20(- 2), - 94.

Jafarkhani, R., and Masri, S. F. (2011). "Finite Element Model Updating using Evolutionary Strategy for Damage Detection." *Computer-Aided Civil and Infrastructure Engineering*, 26(3), 207-224.

Kasban, H., Zahran, O., Arafa, H., El-Kordy, M., Elaraby, S. M. S., and Abd El-Samie, F. E. (2011). "Welding Defect Detection from Radiography Images with a Cepstral Approach." *NDT&E International*, 44(2), 226-231.

Kay, P. (2004). *Underwater Photography*, Guild of Master Craftsman. East Sussex, United Kingdom.

Khanfar, A., Abu-Khousa, M., and Qaddoumi, N. (2003). "Microwave Near-Field Nondestructive Detection and Characterization of Disbonds in Concrete Structures using Fuzzy Logic Techniques." *Composite structures*, 62(3), 335-339.

Kim, J. H., and Kwak, H.-G. (2008). "Nondestructive Evaluation of Elastic Properties of Concrete using Simulation of Surface Waves." *Computer-Aided Civil and Infrastructure Engineering*, 23(8), 611-624.

Kindermann, R., and Snell, J.L. (1980). "Markov Random Fields and Their Applications." American Mathematical Society, Providence, United States.

Kirk, J. O. (1985). "Effects Of Suspensoids (Turbidity) on Penetration of Solar Radiation in Aquatic Ecosystems." *Perspectives in Southern Hemisphere Limnology*, B. R. Davies, and R. D. Walmsley, eds., Springer Netherlands, 195-208.

Kirsch, R. A. (1971). "Computer Determination of the Constituent Structure of Biological Images." *Computers and biomedical research, an international journal*, 4(3), 315-328.

Komorowski, J. P., and Forsyth, D. S. (2000). "Role of Enhanced Visual Inspections in the New Strategy for Corrosion Management." *Aircraft Engineering and Aerospace Technology*, 72(1), 5-13.

Lawlor, A., Torres, J., O'Flynn, B., Wallace, J., and Regan, F. (2012). "DEPLOY: A Long Term Deployment of a Water Quality Sensor Monitoring System." *Sensor Review*, 32(1), 29-38.

Lenhart, C., Brooks, K., Heneley, D., and Magner, J. (2010). "Spatial and Temporal Variation in Suspended Sediment, Organic Matter, and Turbidity in a Minnesota Prairie River: Implications for Tmdls." *Environ Monit Assess*, 165(1-4), 435-447.

Li, D., Xu, L., Goodman, E. D., Xu, Y., and Wu, Y. (2013). "Integrating a Statistical Background-Foreground Extraction Algorithm and SVM Classifier for Pedestrian Detection and Tracking." *Integrated Computer-Aided Engineering*, 20(3), 201-216.

Li, H., Huang, Y., Chen, W. L., Ma, M. L., Tao, D. W., and Ou, J. P. (2011). "Estimation and Warning of Fatigue Damage of FRP Stay Cables Based on Acoustic Emission Techniques and Fractal Theory." *Computer-Aided Civil and Infrastructure Engineering*, 26(7), 500-512.

Li, H., Huang, Y., Ou, J., and Bao, Y. (2011). "Fractal Dimension-Based Damage Detection Method for Beams with a Uniform Cross-Section." *Computer-Aided Civil and Infrastructure Engineering*, 26(3), 190-206.

LiMei, S., XingLin, Z., XingHua, Q., KeXin, X., and LiNa, L. (2005). "Novel 3D Sphericity Evaluation Based on SFS-NDT." *NDT& E International*, 38(6), 442-447.

Lindgren, A., Shull, P., Joseph, K., and Hagemaier, D. (2002). "Magnetic particle." *Nondestructive Evaluation: Theory, Techniques, and Applications. Marcel Dekker, Inc, 270 Madison Avenue, New York, NY 10016, USA, 2002.*, 193-259.

Lowe, D. G. (1999). "Object Recognition from Local Scale-Invariant Features." *Proc., Proceedings of the Seventh IEEE International Conference on Computer Vision, 20-27 Sept. 1999*, IEEE Comput. Soc, Kerkyra, Greece, 1150-1157.

Lu, C. S., Chung, P. C., and Chen, C. F. (1997). "Unsupervised Texture Segmentation via Wavelet Transform." *Pattern Recognition*, 30(5), 729-742.

Mahiddine, A., Seinturier, J., Boi, D. P. J., Drap, P., Merad, D., and Luc, L. (2012). "Underwater Image Preprocessing for Automated Photogrammetry in High Turbidity Water: An Application on the Arles-Rhone XIII Roman Wreck in the Rhodano River, France." *Proc., Virtual Systems and Multimedia (VSMM), 2012 18th International Conference on*, Milan, Italy, *2-5 Sept. 2012*, 189-194.

Matthies, L., and Shafer, S. (1987). "Error Modeling in Stereo Navigation." *Robotics and Automation, IEEE Journal of*, 3(3), 239-248.

Mitchelson, J., and Hilton, A. (2003). "Wand-Based Multiple Camera Studio Calibration." *Center Vision, Speech and Signal Process*. Surrey, UK.

Molero, M., Aparicio, S., Al-Assadi, G., Casati, M., Hernández, M., and Anaya, J. (2012). "Evaluation of Freeze–Thaw Damage in Concrete by Ultrasonic Imaging." *NDT & E International*, 52, 86-94.

Naccari, F., Battiato, S., Bruna, A., Capra, A., and Castorina, A. (2005). "Natural Scenes Classification for Color Enhancement." *Consumer Electronics, IEEE Transactions on*, 51(1), 234-239.

Nayar, S. K., Watanabe, M., and Noguchi, M. (2005). "Real-Time Focus Range Sensor." *Proc., Proceedings of IEEE International Conference on Computer Vision,* Cambridge, MA, USA, *20-23 June 1995*, IEEE Comput. Soc. Press, 995-1001.

Nugent, M. J., and Pellegrino, B. A. (1991). "Remote Visual Testing (RVT) for the Diagnostic Inspection of Feedwater Heaters." *Proc., 1991 International Joint Power Generation Conference,* October 6, 1991 - October 10, 1991, Publ by ASME, 55-62.

O'Byrne, M., Ghosh, B., Pakrashi, V., and Schoefs, F. (2013). "Effects of Turbidity and Lighting on the Performance of an Image Processing Based Damage Detection Technique." *Proc., 11th International Conference on Structural Safety and Reliability, ICOSSAR* 16-20 June, 2013, New York, NY, USA, 2645-2650.

O'Byrne, M., Ghosh, B., Schoefs, F., and Pakrashi, V. (2014). "Regionally Enhanced Multiphase Segmentation Technique for Damaged Surfaces." *Computer-Aided Civil and Infrastructure Engineering*, 29(9), 644-658.

O'Byrne, M., Pakrashi, V., Ghosh, B., and F., S. (2014). "Effects of Turbidity and Lighting on an Image Processing based Crack Detection Technique." *Proceedings of Civil Engineering Research in Ireland,* Belfast, Northern Ireland, 28-29 August, 2014.

O'Byrne, M., Pakrashi, V., Ghosh, B., and Schoefs, F. (2011). "Receiver Operating Characteristics of a Modified Edge Detection for Corrosion Classification." *Forum Bauinformatik 23rd European Conference,* Cork, Ireland.

O'Byrne, M., Pakrashi, V., Schoefs, F., and Ghosh, B. (2014). "A Comparison of Image Based 3D Recovery Methods for Underwater Inspections." 7th European Workshop on Structural Health Monitoring EWSHM 2014, 8-11 July, 2014, Nantes, France.

O'Byrne, M., Schoefs, F., Ghosh, B., and Pakrashi, V. (2013). "Texture Analysis Based Damage Detection of Ageing Infrastructural Elements." *Computer-Aided Civil and Infrastructure Engineering*, 28(3), 162-177.

O'Gorman, L., Sammon, M. J., and Seul, M. (2008). *Practical Algorithms for Image Analysis with CD-ROM*, Cambridge University Press.

Olaszek, P. (1999). "Investigation of the Dynamic Characteristic of Bridge Structures using a Computer Vision Method." *Measurement*, 25(3), 227-236.

Olofsson, A. (2010). "Modern Stereo Correspondence Algorithms: Investigation and Evaluation." Thesis from Department of Electrical Engineering, Linköping University, Linköping, Sweden.

Otsu, N. (1979). "A Threshold Selection Method from Gray-Level Histograms." *IEEE Transactions on Systems, Man and Cybernetics*, SMC-9(1), 62-66.

Pakrashi, V., Schoefs, F., Memet, J. B., and O'Connor, A. (2010). "ROC Dependent Event Isolation Method for Image Processing Based Assessment of Corroded Harbour Structures." *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and performance (NSIE)*, 6(3), 365-378.

Phares, B., Washer, G., Rolander, D., Graybeal, B., and Moore, M. (2004). "Routine Highway Bridge Inspection Condition Documentation Accuracy and Reliability." *Journal of Bridge Engineering*, 9(4), 403-413.

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., and Zuiderveld, K. (1987). "Adaptive Histogram Equalization and its Variations." *Computer Vision, Graphics, and Image Processing*, 39(3), 355-368.

Pribanic, T., Sturm, P., and Peharec, S. (2009). "Wand-based calibration of 3D kinematic system," *Computer Vision, IET* , 3(3), 124-129.

Ramos, E. J. (1992). "Underwater Inspection - The State Of Art." *Non-destructive Testing 92*, C. Hallai, and P. Kulcsar, eds., Elsevier, Oxford, 517-519.

Reinhard, E., Khan, E. A., Akyz, A. O., and Johnson, G. M. (2008). *Color Imaging: Fundamentals and Applications*, Natick, MA, USA.

Rizzo, P. (2013). "NDE/SHM of Underwater Structures: a Review." *Advances in Science and Technology*, 83, 208-216.

Roman, C., Inglis, G., and Rutter, J. (2010). "Application of Structured Light Imaging for High Resolution Mapping of Underwater Archaeological Sites." *Proc., OCEANS 2010 IEEE-Sydney*, IEEE, 1-9.

Rong-Yau, H., Mao, I. S., and Hao-Kang, L. (2010). "Exploring the Deterioration Factors of RC Bridge Decks: A Rough Set Approach." *Computer-Aided Civil and Infrastructure Engineering*, 25(7), 517-529.

Rose, J. L., Fuller, M., Nestleroth, J., and Jeong, Y. (1983). "An ultrasonic global inspection technique for an offshore K-Joint." *Old SPE Journal*, 23(2), 358-364.

Rouhan, A., and Schoefs, F. (2003). "Probabilistic Modeling of Inspection Results for Offshore Structures." *Structural Safety*, 25(4), 379-399.

Rytter, A., and Kirkegaard, P. H. (1994). *Vibration based Inspection of Civil Engineering Structures*, PhD Thesis, Aalborg Universitetsforlag, Denmark.

Sahabi, H., and Basu, A. (1996). "Analysis of Error in Depth Perception with Vergence and Spatially Varying Sensing." *Computer Vision and Image Understanding*, 63(3), 447-461.

Sarma, K. C., and Adeli, H. (1998). "Cost Optimization of Concrete Structures." *Journal of Structural Engineering* 124(5), 570-578.

Scharstein, D., and Szeliski, R. (2002). "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms." *International Journal of Computer Vision*, 47(1-3), 7-42.

Schlyter, P. (2009). "Radiometry and photometry in astronomy." Available at www.stjarnhimlen.se/comp/radfaq.html [Accessed 8 November 2013].

Schoefs, F., Abraham, O., and Popovics, J. (2012). "Quantitative Evaluation of NDT Method Performance: Application Example Based on Contactless Impact Echo Measurements for Void Detection in Tendon Duct." *Construction and Building Materials (CBM)*(Special Issue 'NDT').

Schoefs, F., Bernard, O., Capra, B. & Aduriz X. (2009). "Comparison of Additional Costs for Several Replacement Strategies of Randomly Ageing Reinforced Concrete Pipes" *Computer Aided Civil and Infrastructure Engineering (CACAIE),* Wiley ed, 24(7), 492-508.

Schoefs, F., Boéro, J., Clément, A., and Capra, B. (2012). "The αδ Method for Modelling Expert Judgment and Combination of NDT Tools in RBI Context: Application to Marine Structures, Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and performance (NSIE)." *SMonitoring, Modeling and Assessment of Structural Deterioration in Marine Environments*, 8(Special Issue), 531-543.

Schoefs, F., Yanez-Godoy, H., and Lanata, F. (2011). "Polynomial Chaos Representation for Identification of Mechanical Characteristics of Instrumented Structures." *Computer-Aided Civil and Infrastructure Engineering*, 26(3), 173-189.

Shaomin, Z., and Negahdaripour, S. (1997). "Recovery of 3D Depth Map from Image Shading for Underwater Applications." *Proc., Oceans '97. MTS/IEEE Conference Proceedings, 6-9 Oct. 1997*, IEEE, 618-625.

Sinha, S. K., and Fieguth, P. W. (2006). "Automated Detection of Cracks in Buried Concrete Pipe Images." *Automation in Construction*, 15(1), 58-72.

Sirca Jr, G. F., and Adeli, H. (2005). "Cost Optimization of Prestressed Concrete Bridges." *Journal of Structural Engineering*, 131(3), 380-388.

Sohn, H., Kim, S. D., and Harries, K. (2008). "Reference-Free Damage Classification Based on Cluster Analysis." *Computer-Aided Civil and Infrastructure Engineering*, 23(5), 324-338.

Stroebel, L. D., and Zakia, R. D. (1993). *The Focal encyclopedia of photography (3rd ed.)*. Focal Press, Boston, USA.

Talebinejad, I., Fischer, C., and Ansari, F. (2011). "Numerical Evaluation of Vibration-Based Methods for Damage Assessment of Cable-Stayed Bridges." *Computer-Aided Civil and Infrastructure Engineering*, 26(3), 239-251.

Thaulow, C., and Berge, T. (1984). "Acoustic Emission Monitoring of Corrosion Fatigue Crack Growth in Offshore Steel." *NDT international*, 17(3), 147-153.

Van Rooij, J. M., and Videler, J. J. (1996). "A Simple Field Method for Stereo-Photographic Length Measurement of Free-Swimming Fish: Merits And Constraints." *Journal of Experimental Marine Biology and Ecology*, 195(2), 237-249.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*, Springer-Verlag New York, Inc. New York, NY, USA.

Vapnik, V. N. (1996). "Book Review: The Nature of Statistical Learning Theory." *Technometrics*, 38(4), 400-400.

Vilar, R., Zapata, J., and Ruiz, R. (2009). "An Automatic System of Classification of Weld Defects in Radiographic Images." *NDT & E International*, 42(5), 467-476.

Wahbeh, A. M., Caffrey, J. P., and Masri, S. F. (2003). "A Vision-Based Approach for the Direct Measurement of Displacements in Vibrating Systems." *Smart materials and structures*, 12(5), 785.

Wang, P., and Huang, H. (2010). "Comparison Analysis on Present Image-Based Crack Detection Methods in Concrete Structures." *Proc., Image and Signal Processing (CISP), 2010 3rd International Congress on*, IEEE, 16-18 Oct. 2010, Yantai, China, 2530-2533.

Watson, J. (1992). "Underwater Visual Inspection and Measurement using Optical Holography." *Optics and Lasers in Engineering*, 16(4–5), 375-390.

Wei, L., Guoming, C., Xiaokang, Y., Chuanrong, Z., and Tao, L. (2012). "Modeling and Simulation of Crack Detection for Underwater Structures using an ACFM Method." *Proc., Review of Progress in Quantitative Nondestructive Evaluation, 15-20 July 2012*, American Institute of Physics, Denver, CO, USA, 436-440.

Xia, Y., Ni, Y.Q., Zhang, P., Liao, W.Y., and Ko, J.M. (2011). "Stress Development of a Supertall Structure during Construction: Field Monitoring and Numerical Analysis." *Computer-Aided Civil and Infrastructure Engineering*, 26(7), 542-559.

Yamaguchi, T., and Hashimoto, S. (2008). "Improved Percolation-Based Method for Crack Detection in Concrete Surface Images." *Proc., Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, IEEE, 1-4.

Yatchev, I., Stancheva, R., Marinova, I., Li, M., and Lowther, D. (2012). "NDT Sensor Design Optimization For Accurate Crack Reconstruction." *COMPEL - The international*

journal for computation and mathematics in electrical and electronic engineering, 31(3), 792-802.

Yazid, H., Arof, H., Yazid, H., Ahmad, S., Mohamed, A., and Ahmad, F. (2011). "Discontinuities Detection in Welded Joints Based on Inverse Surface Thresholding." *NDT & E International*, 44(7), 563-570.

Yusa, N., Janousek, L., Rebican, M., Chen, Z., Miya, K., Chigusa, N., and Ito, H. (2006). "Detection of Embedded Fatigue Cracks in Inconel Weld Overlay and the Evaluation of the Minimum Thickness of the Weld Overlay using Eddy Current Testing." *Nuclear Engineering and Design*, 236(18), 1852-1859.

Zhang, Z. (2000). "A Flexible New Technique for Camera Calibration." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330-1334.

Zhang, Z., Luong, Q. T., and Faugeras, O. (1994). "Motion of an Uncalibrated Stereo Rig: Self-Calibration and Metric Reconstruction." *Proc., Proceedings of 12th International Conference on Pattern Recognition, 9-13 Oct. 1994*, IEEE Comput. Soc. Press, 695-697.

Zheng, L., Genest, M., and Krys, D. (2012). "Processing Thermography Images for Pitting Corrosion Quantification on Small Diameter Ductile Iron Pipe." *NDT&E International*, 47, 105-115.

Zou, Q., Cao, Y., Li, Q., Mao, Q., and Wang, S. (2012). "CrackTree: Automatic crack Detection from Pavement Images." *Pattern Recognition Letters*, 33(3), 227-238.

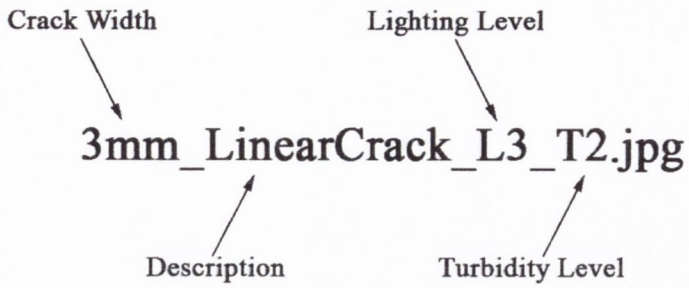
APPENDIX

Contents of the Underwater Lighting and Turbidity Images Repository

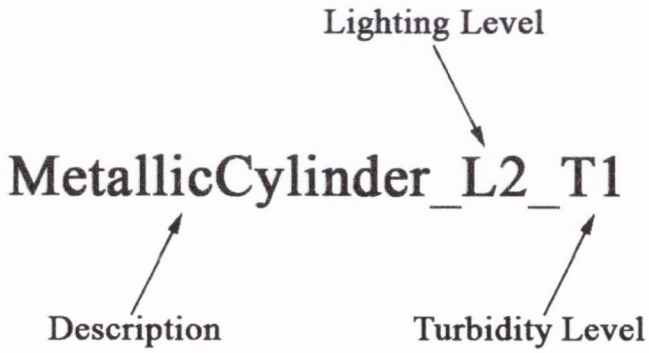
Naming Convention

1. The first part of the image name reveals the dimensions if the damage (if applicable), e.g. the crack width.
2. The second part of the name describes the nature of the specimen, such as the surface type or shape.
3. The third part indicates the lighting level. The lighting levels L1, L2, and L3 correspond to low (100 lux), medium (1,000 lux), and high (10,000 lux) respectively.
4. The final part of the name indicates the turbidity level. The turbidity levels T1, T2, and T3 correspond to low (0 NTU), medium (6 NTU), and high (12 NTU) respectively.
5. For the stereo images, the specimens are photographed from 8 sides. Each time the specimen is rotated by 45° from the previous position in order to enable users obtain a full 3D reconstruction of the specimen (excluding the underside). The number at the end of the file name indicates how many times it has been rotated. This number ranges from 1 to 8.

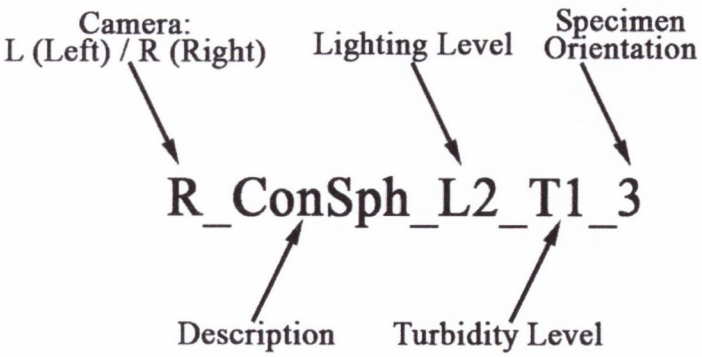
Cracks



Surface Damage



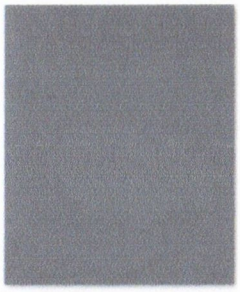
Shape Information (Stereo Imagery)



Cracks

1mm_CurvedCrack_L1_T1

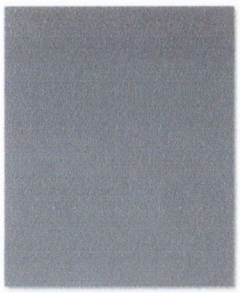
1mm_CurvedCrack_L1_T2



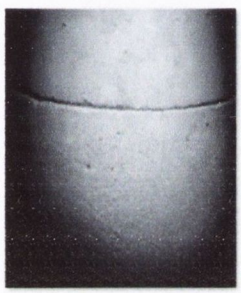
1mm_CurvedCrack_L1_T3

1mm_CurvedCrack_L2_T1

1mm_CurvedCrack_L2_T2



1mm_CurvedCrack_L2_T3



1mm CurvedCrack L3 T1

1mm CurvedCrack L3 T2

1mm CurvedCrack L3 T3

1mm_LinearCrack_L1_T1

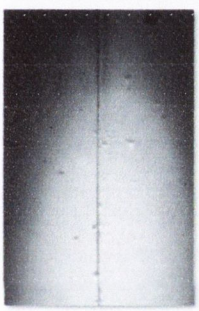
1mm_LinearCrack_L1_T2

1mm_LinearCrack_L1_T3

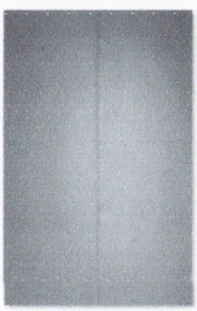
1mm_LinearCrack_L2_T1

1mm_LinearCrack_L2_T2

1mm_LinearCrack_L2_T3

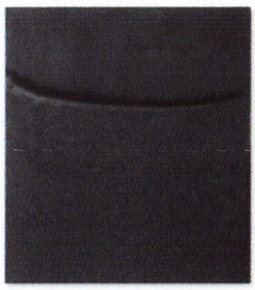


1mm_LinearCrack_L3_T1

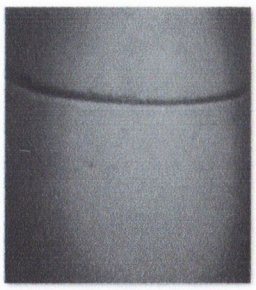


1mm_LinearCrack_L3_T2

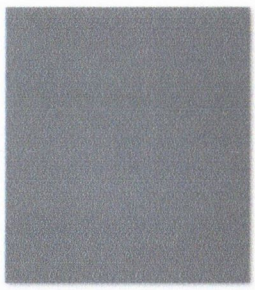
1mm_LinearCrack_L3_T3



3mm_CurvedCrack_L1_T1

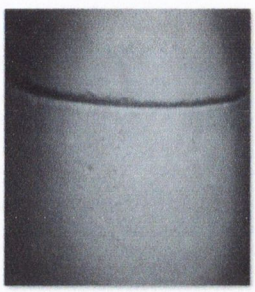


3mm_CurvedCrack_L1_T2



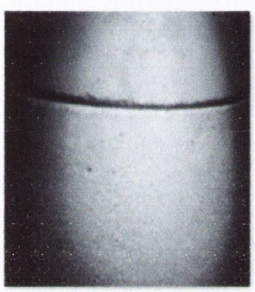
3mm_CurvedCrack_L1_T3

3mm_CurvedCrack_L2_T1



3mm_CurvedCrack_L2_T2

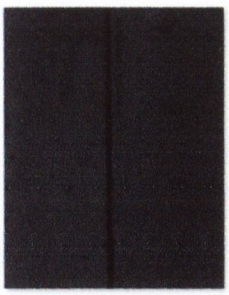
3mm_CurvedCrack_L2_T3



3mm_CurvedCrack_L3_T1

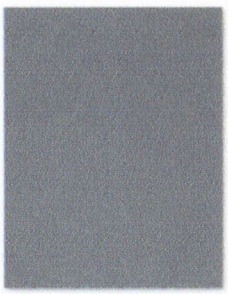
3mm_CurvedCrack_L3_T2

3mm_CurvedCrack_L3_T3

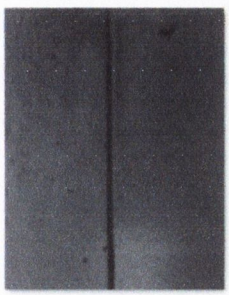


3mm_LinearCrack_L1_T1

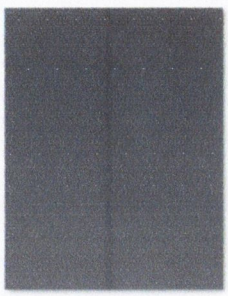
3mm_LinearCrack_L1_T2



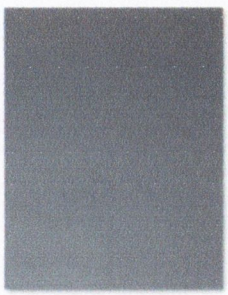
3mm_LinearCrack_L1_T3



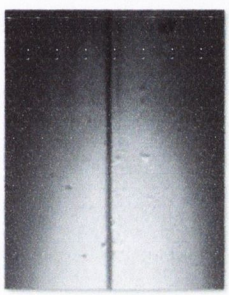
3mm_LinearCrack_L2_T1



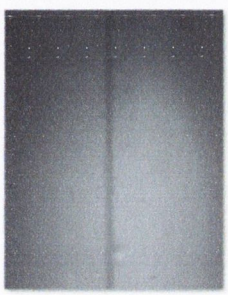
3mm_LinearCrack_L2_T2



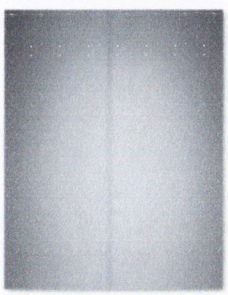
3mm_LinearCrack_L2_T3



3mm_LinearCrack_L3_T1



3mm_LinearCrack_L3_T2

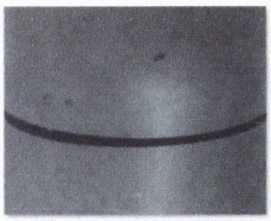


3mm_LinearCrack_L3_T3

5mm_MarkerCurvedCrack_L1_T1

5mm_MarkerCurvedCrack_L1_T2

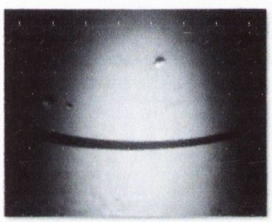
5mm_MarkerCurvedCrack_L1_T3



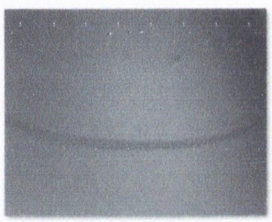
5mm_MarkerCurvedCrack_L2_T1

5mm_MarkerCurvedCrack_L2_T2

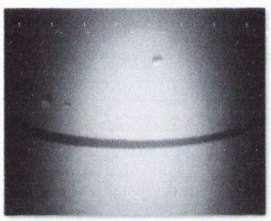
5mm_MarkerCurvedCrack_L2_T3



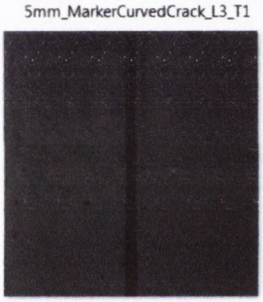
5mm_MarkerCurvedCrack_L3_T1



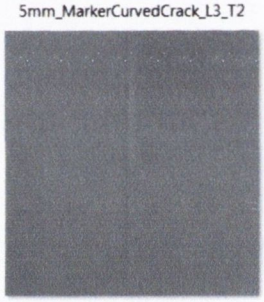
5mm_MarkerCurvedCrack_L3_T2



5mm_MarkerCurvedCrack_L3_T3

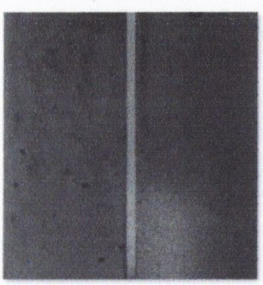


5mm_LinearCrack_L1_T1



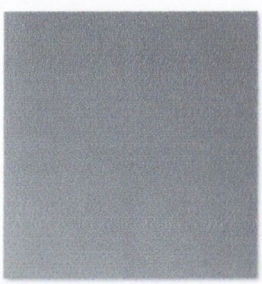
5mm_LinearCrack_L1_T2

5mm_LinearCrack_L1_T3

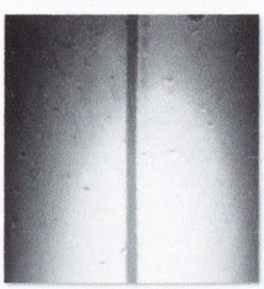


5mm_LinearCrack_L2_T1

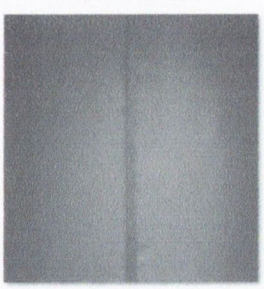
5mm_LinearCrack_L2_T2



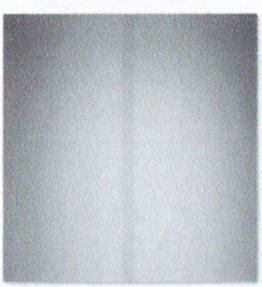
5mm_LinearCrack_L2_T3



5mm_LinearCrack_L3_T1



5mm_LinearCrack_L3_T2

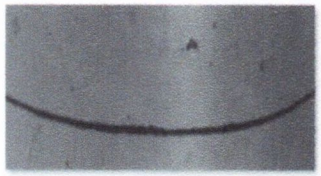


5mm_LinearCrack_L3_T3

3mm_MarkerCurvedCrack_L1_T1

3mm_MarkerCurvedCrack_L1_T2

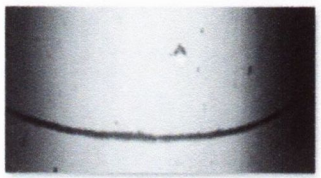
3mm_MarkerCurvedCrack_L1_T3



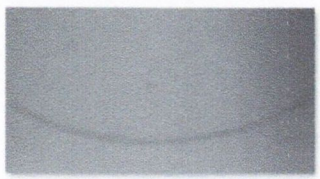
3mm_MarkerCurvedCrack_L2_T1

3mm_MarkerCurvedCrack_L2_T2

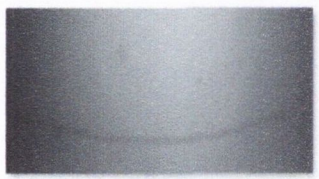
3mm_MarkerCurvedCrack_L2_T3



3mm_MarkerCurvedCrack_L3_T1

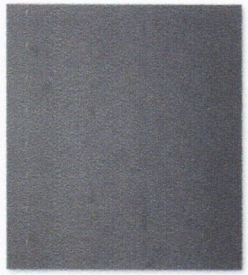


3mm_MarkerCurvedCrack_L3_T2

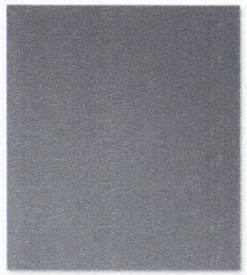


3mm_MarkerCurvedCrack_L3_T3

RealCrackPartiallyControled_L1_T1

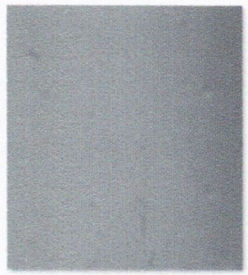


RealCrackPartiallyControled_L1_T2

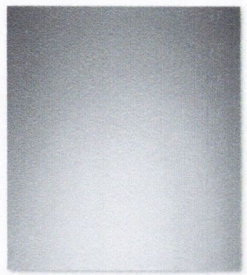


RealCrackPartiallyControled_L1_T3

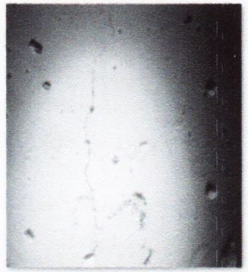
RealCrackPartiallyControled_L2_T1



RealCrackPartiallyControled_L2_T2

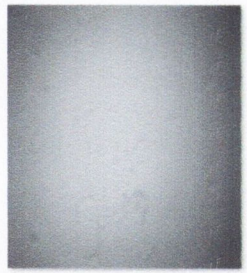


RealCrackPartiallyControled_L2_T3



RealCrackPartiallyControled_L3_T1

RealCrackPartiallyControled_L3_T2



RealCrackPartiallyControled_L3_T3

RealCrackTexturedCylinder_L1_T1

RealCrackTexturedCylinder_L1_T2

RealCrackTexturedCylinder_L1_T3

RealCrackTexturedCylinder_L2_T1

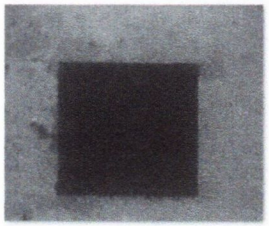
RealCrackTexturedCylinder_L2_T2

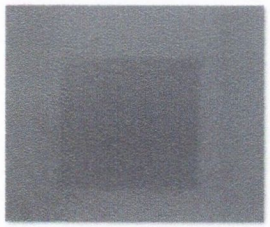
RealCrackTexturedCylinder_L2_T3

RealCrackTexturedCylinder_L3_T1

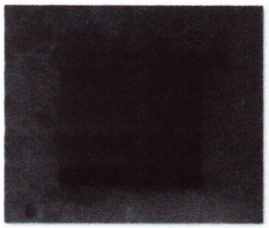
RealCrackTexturedCylinderP_L3_T2

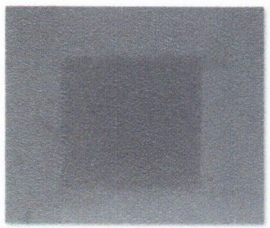
RealCrackTexturedCylinder_L3_T3


50mm_FlatBrownPlainConcrete_L1_T1

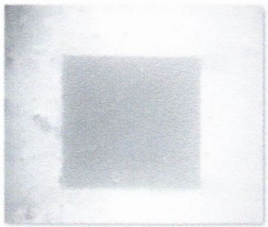

50mm_FlatBrownPlainConcrete_L1_T2

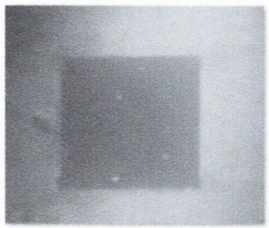

50mm_FlatBrownPlainConcrete_L1_T3

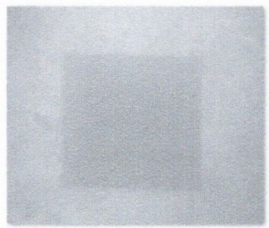

50mm_FlatBrownPlainConcrete_L2_T1


50mm_FlatBrownPlainConcrete_L2_T2

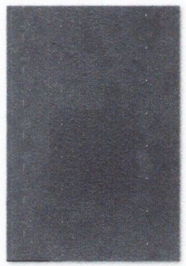

50mm_FlatBrownPlainConcrete_L2_T3


50mm_FlatBrownPlainConcrete_L3_T1


50mm_FlatBrownPlainConcrete_L3_T2


50mm_FlatBrownPlainConcrete_L3_T3


50mm_FlatBrownTexturedConcrete_L1_T1

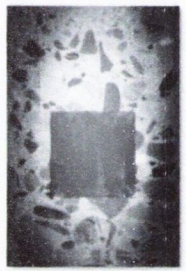

50mm_FlatBrownTexturedConcrete_L1_T2

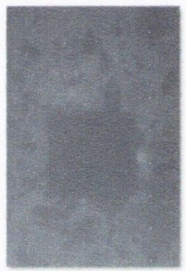

50mm_FlatBrownTexturedConcrete_L1_T3

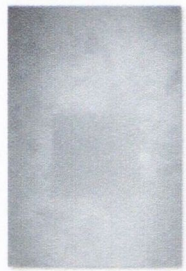

50mm_FlatBrownTexturedConcrete_L2_T1

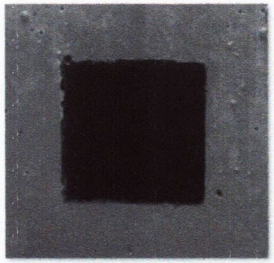

50mm_FlatBrownTexturedConcrete_L2_T2


50mm_FlatBrownTexturedConcrete_L2_T3

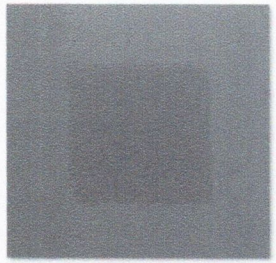

50mm_FlatBrownTexturedConcrete_L3_T1


50mm_FlatBrownTexturedConcrete_L3_T2


50mm_FlatBrownTexturedConcrete_L3_T3

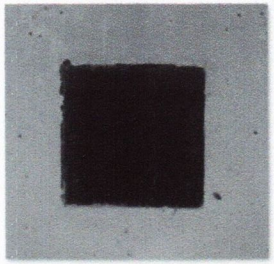


50mm_FlatMetallic_L1_T1

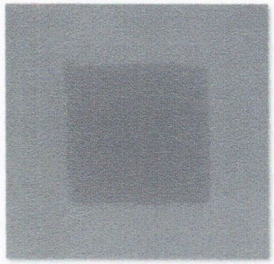


50mm_FlatMetallic_L1_T2

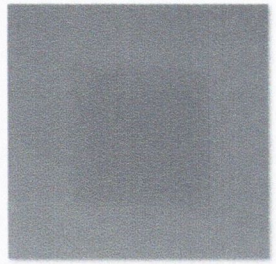
50mm_FlatMetallic_L1_T3



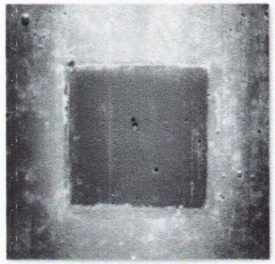
50mm_FlatMetallic_L2_T1



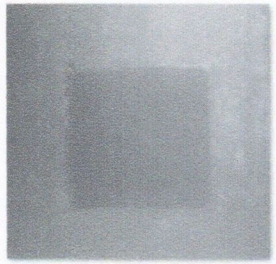
50mm_FlatMetallic_L2_T2



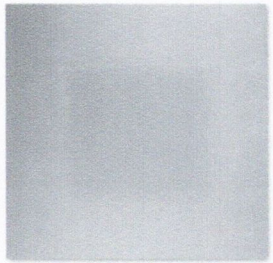
50mm_FlatMetallic_L2_T3



50mm_FlatMetallic_L3_T1



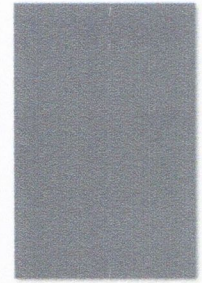
50mm_FlatMetallic_L3_T2



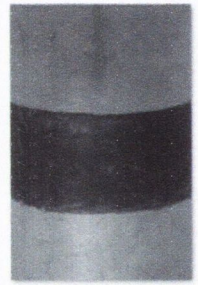
50mm_FlatMetallic_L3_T3

PlainConcreteCylinder_L1_T1

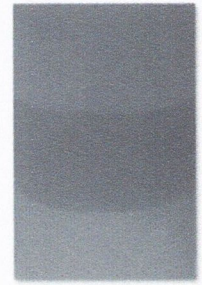
PlainConcreteCylinder_L1_T2



PlainConcreteCylinder_L1_T3



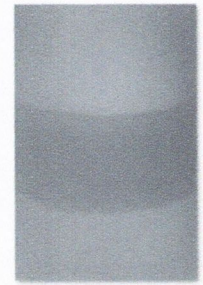
PlainConcreteCylinder_L2_T1



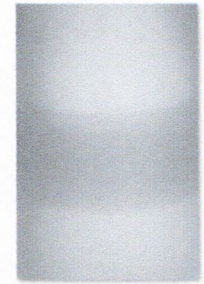
PlainConcreteCylinder_L2_T2

PlainConcreteCylinder_L2_T3

PlainConcreteCylinder_L3_T1

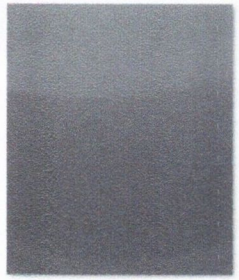


PlainConcreteCylinder_L3_T2

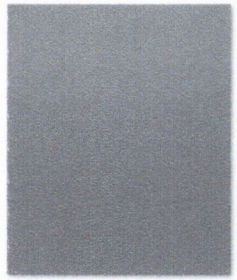


PlainConcreteCylinder_L3_T3

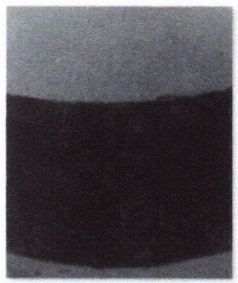
PlainConcreteSphere_L1_T1



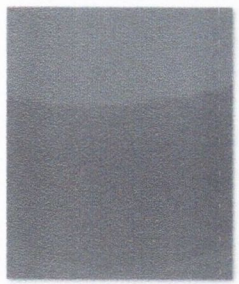
PlainConcreteSphere_L1_T2



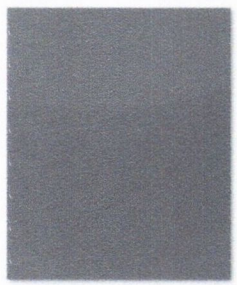
PlainConcreteSphere_L1_T3



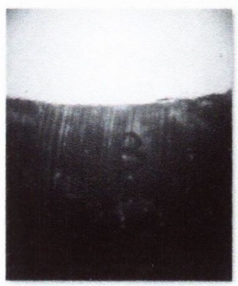
PlainConcreteSphere_L2_T1



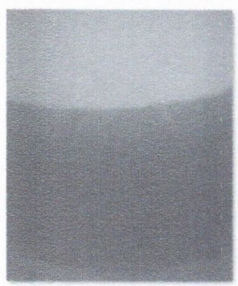
PlainConcreteSphere_L2_T2



PlainConcreteSphere_L2_T3



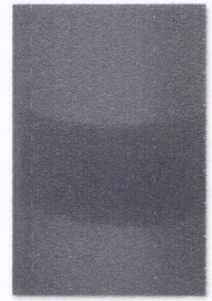
PlainConcreteSphere_L3_T1



PlainConcreteSphere_L3_T2

PlainConcreteSphere_L3_T3

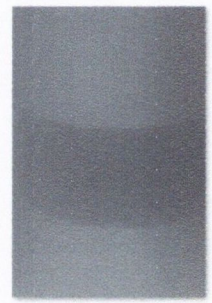
MetallicCylinder_L1_T1



MetallicCylinder_L1_T2

MetallicCylinder_L1_T3

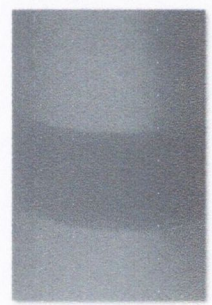
MetallicCylinder_L2_T1



MetallicCylinder_L2_T2

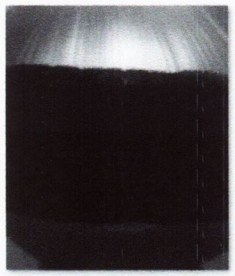
MetallicCylinder_L2_T3

MetallicCylinder_L3_T1

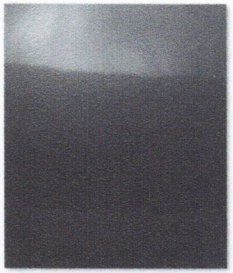


MetallicCylinder_L3_T2

MetallicCylinder_L3_T3

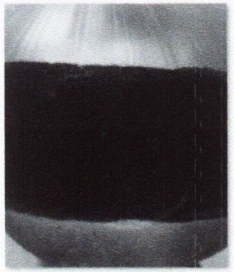


MetallicSphere_L1_T1

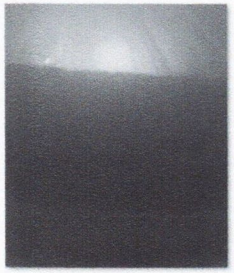


MetallicSphere_L1_T2

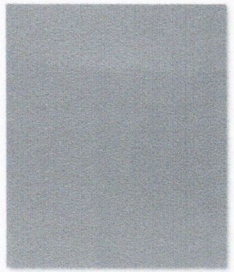
MetallicSphere_L1_T3



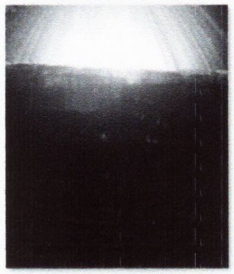
MetallicSphere_L2_T1



MetallicSphere_L2_T2



MetallicSphere_L2_T3

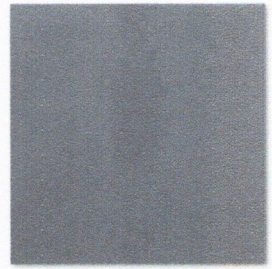


MetallicSphere_L3_T1

MetallicSphere_L3_T2

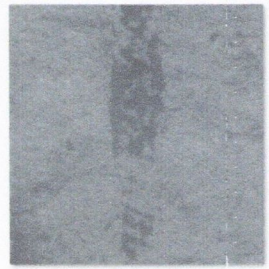
MetallicSphere_L3_T3

Real2dDamage_L1_T1

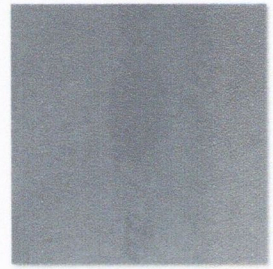


Real2dDamage_L1_T2

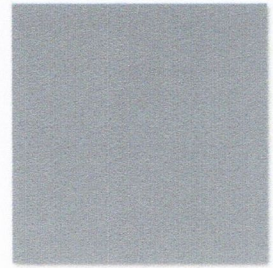
Real2dDamage_L1_T3



Real2dDamage_L2_T1

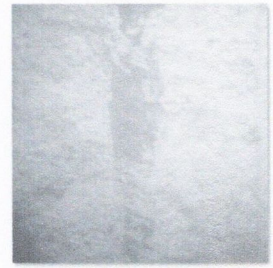


Real2dDamage_L2_T2

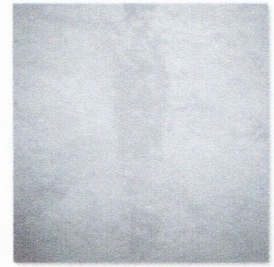


Real2dDamage_L2_T3

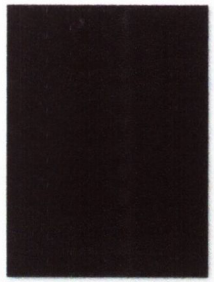
Real2dDamage_L3_T1



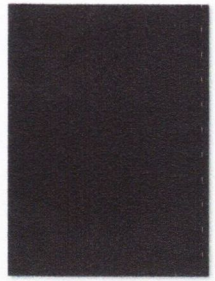
Real2dDamage_L3_T2



Real2dDamage_L3_T3



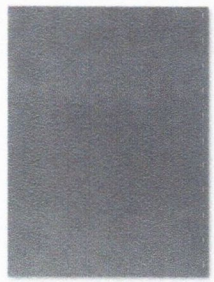
TexturedConcreteCylinder_L1_T1



TexturedConcreteCylinder_L1_T2

TexturedConcreteCylinder_L1_T3

TexturedConcreteCylinder_L2_T1



TexturedConcreteCylinder_L2_T2

TexturedConcreteCylinder_L2_T3

TexturedConcreteCylinder_L3_T1

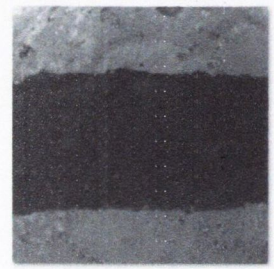
TexturedConcreteCylinder_L3_T2

TexturedConcreteCylinder_L3_T3

TexturedConcreteSphere_L1_T1

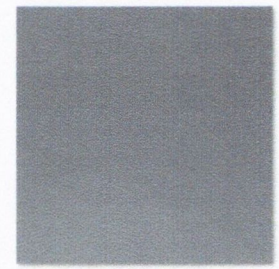
TexturedConcreteSphere_L1_T2

TexturedConcreteSphere_L1_T3



TexturedConcreteSphere_L2_T1

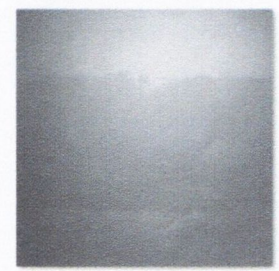
TexturedConcreteSphere_L2_T2



TexturedConcreteSphere_L2_T3

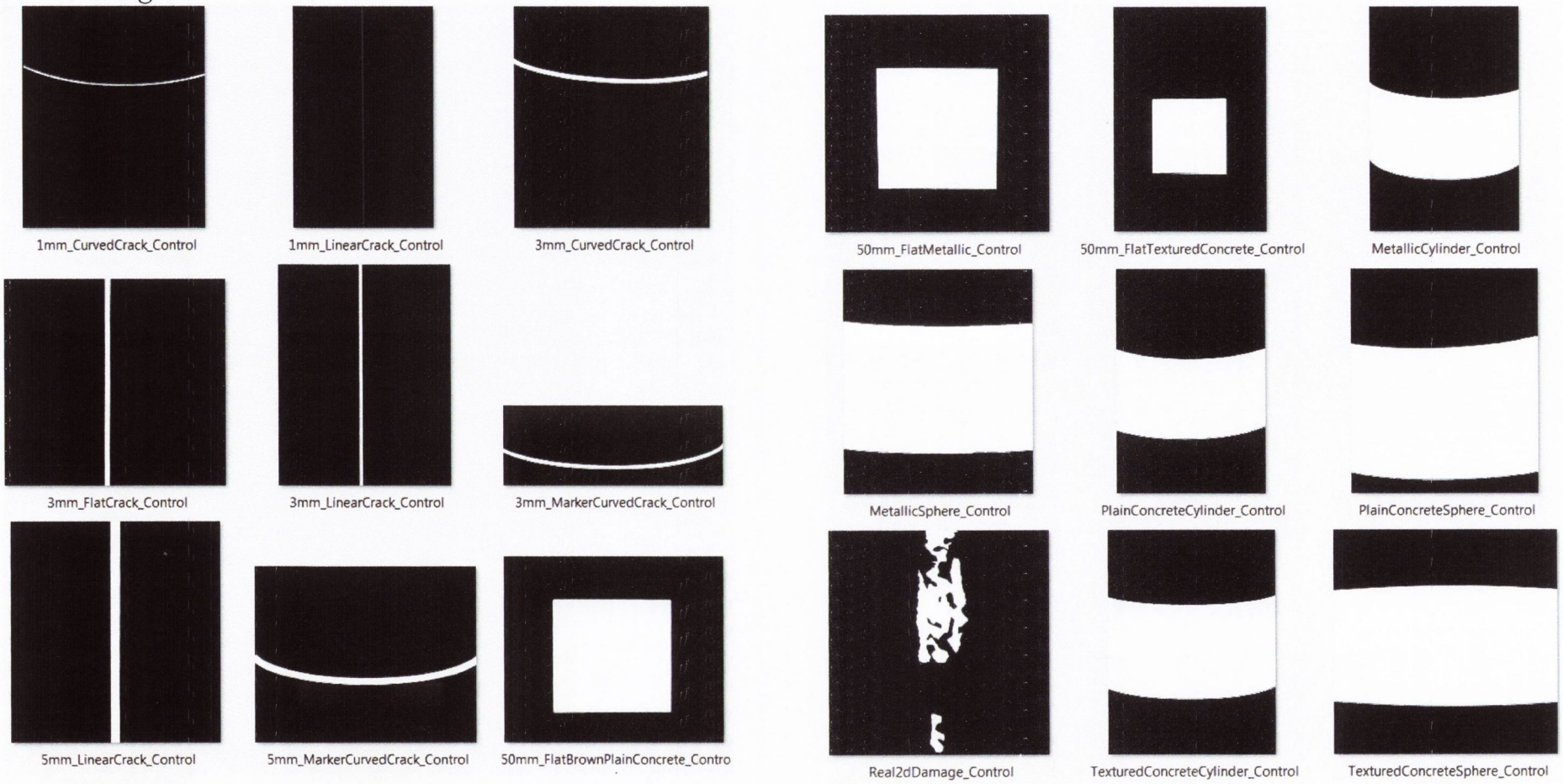
TexturedConcreteSphere_L3_T1

TexturedConcreteSphere_L3_T2

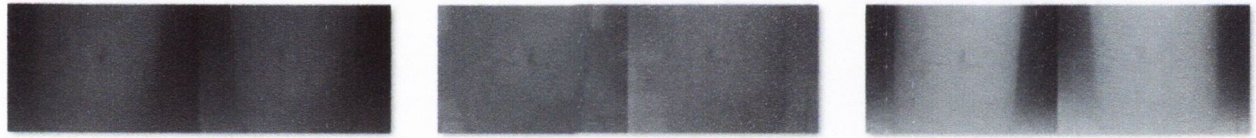


TexturedConcreteSphere_L3_T3

Control Images

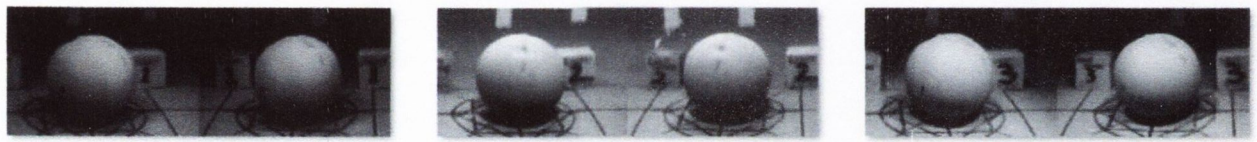


ConCyl_L1_T1_1 ConCyl_L2_T1_1 ConCyl_L3_T1_1

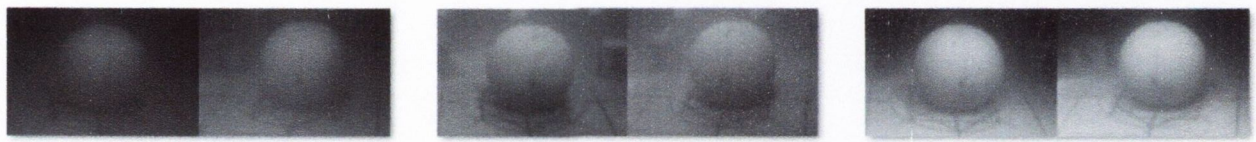


ConCyl_L1_T2_1 ConCyl_L2_T2_1 ConCyl_L3_T2_1

ConCyl_L1_T3_1 ConCyl_L2_T3_1 ConCyl_L3_T3_1



ConSph_L1_T1_1 ConSph_L2_T1_1 ConSph_L3_T1_1



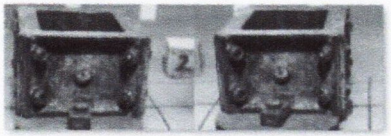
ConSph_L1_T2_1 ConSph_L2_T2_1 ConSph_L3_T2_1

ConSph_L1_T3_1 ConSph_L2_T3_1 ConSph_L3_T3_1

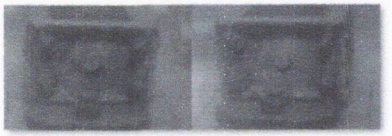
MetCas_L1_T1_1

MetCas_L1_T2_1

MetCas_L1_T3_1



MetCas_L2_T1_1



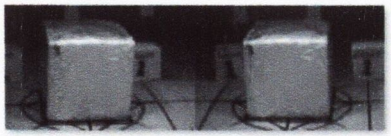
MetCas_L2_T2_1

MetCas_L2_T3_1

MetCas_L3_T1_1

MetCas_L3_T2_1

MetCas_L3_T3_1

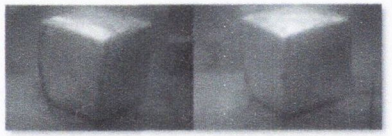


MetCub_L1_T1_1

MetCub_L1_T2_1

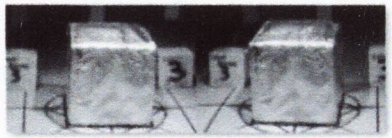
MetCub_L1_T3_1

MetCub_L2_T1_1



MetCub_L2_T2_1

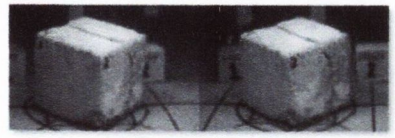
MetCub_L2_T3_1



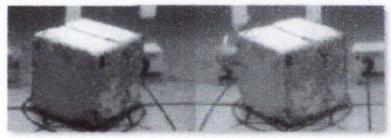
MetCub_L3_T1_1

MetCub_L3_T2_1

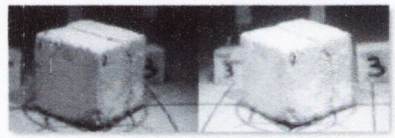
MetCub_L3_T3_1



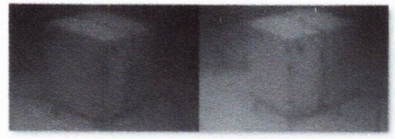
ConCra_L1_T1_1



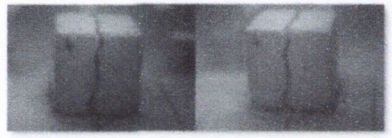
ConCra_L2_T1_1



ConCra_L3_T1_1



ConCra_L1_T2_1



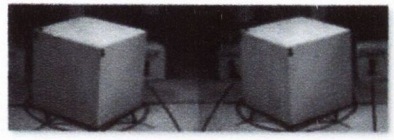
ConCra_L2_T2_1

ConCra_L3_T2_1

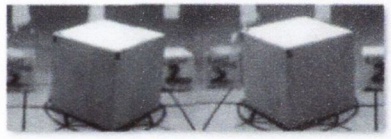
ConCra_L1_T3_1

ConCra_L2_T3_1

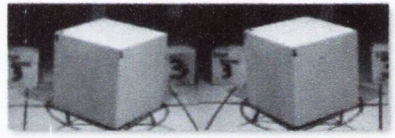
ConCra_L3_T3_1



ConCub_L1_T1_1



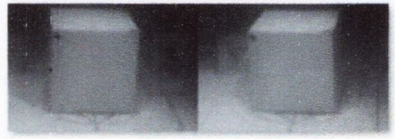
ConCub_L2_T1_1



ConCub_L3_T1_1

ConCub_L1_T2_1

ConCub_L2_T2_1



ConCub_L3_T2_1

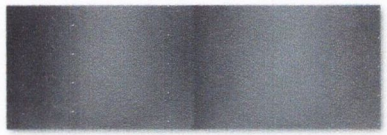
ConCub_L1_T3_1

ConCub_L2_T3_1

ConCub_L3_T3_1

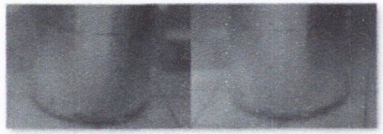
StereoAppendixMetCyl_L1_T1_1

StereoAppendixMetCyl_L1_T2_1



StereoAppendixMetCyl_L1_T3_1

StereoAppendixMetCyl_L2_T1_1

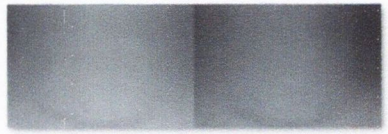


StereoAppendixMetCyl_L2_T2_1

StereoAppendixMetCyl_L2_T3_1

StereoAppendixMetCyl_L3_T1_1

StereoAppendixMetCyl_L3_T2_1

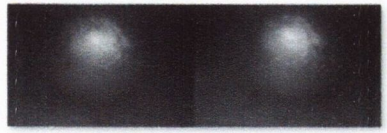


StereoAppendixMetCyl_L3_T3_1

MetSph_L1_T1_1

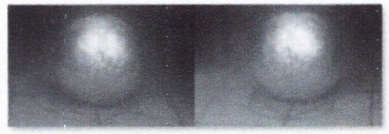
MetSph_L2_T1_1

MetSph_L3_T1_1



MetSph_L1_T2_1

MetSph_L2_T2_1



MetSph_L3_T2_1

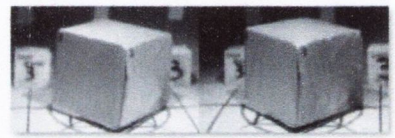
MetSph_L1_T3_1

MetSph_L2_T3_1

MetSph_L3_T3_1

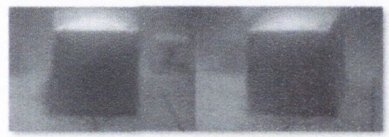
PlaCub_L1_T1_1

PlaCub_L2_T1_1



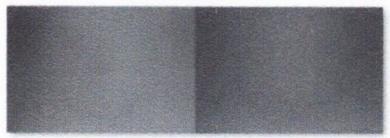
PlaCub_L3_T1_1

PlaCub_L1_T2_1



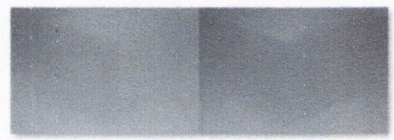
PlaCub_L2_T2_1

PlaCub_L3_T2_1



PlaCub_L1_T3_1

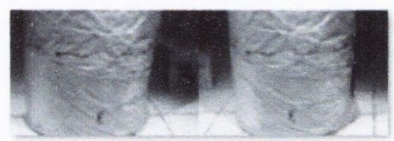
PlaCub_L2_T3_1



PlaCub_L3_T3_1

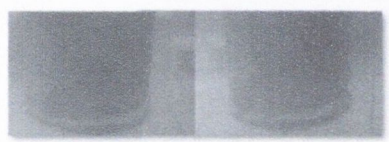
PlaCyl_L1_T1_1

PlaCyl_L2_T1_1

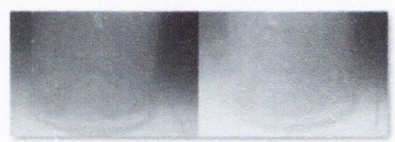


PlaCyl_L3_T1_1

PlaCyl_L1_T2_1



PlaCyl_L2_T2_1



PlaCyl_L3_T2_1

PlaCyl_L1_T3_1

PlaCyl_L2_T3_1

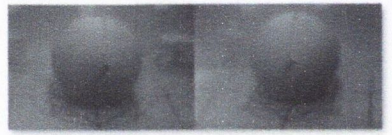
PlaCyl_L3_T3_1

PlaSph_L1_T1_1

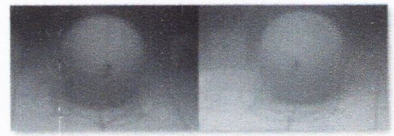
PlaSph_L2_T1_1

PlaSph_L3_T1_1

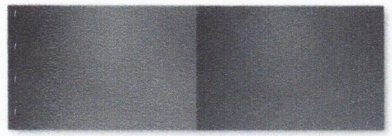
PlaSph_L1_T2_1



PlaSph_L2_T2_1



PlaSph_L3_T2_1

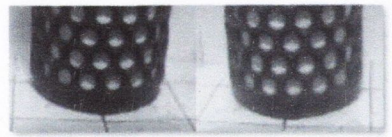


PlaSph_L1_T3_1

PlaSph_L2_T3_1

PlaSph_L3_T3_1

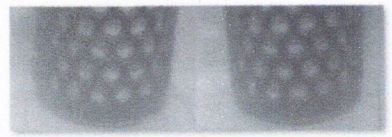
RubMat_L1_T1_1



RubMat_L2_T1_1

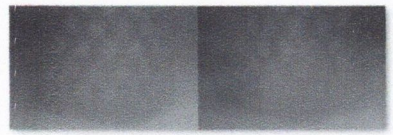
RubMat_L3_T1_1

RubMat_L1_T2_1



RubMat_L2_T2_1

RubMat_L3_T2_1



RubMat_L1_T3_1

RubMat_L2_T3_1

RubMat_L3_T3_1

Matlab™ Code

Contents:

- Crack Detection
- Texture Analysis
- REMPS
- PaLPaBEL Stereo Matching

Crack Detection

```
function [DetectedCracks] = CrackDetection(Image,WindowHalfSize)

%Detection of cracks in images.
%   CrackDetection applies an edge detector to an input colour or gray
%   image, Image. A percolation process is then applied to the detected
%   edges. Image can be of any class and dimension.  The result,
%   DetectedCracks, has the same size and class as Image.
%
%   Notes
%   -----
%
%
%   See also Section 4.2, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%   Example
%   -------------
%       [DetectedCracks] = CrackDetection(Image,35);
%       figure, imshow(DetectedCracks)
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%
% Image:                Colour or grayscale image. Required.
%
% WindowHalfSize:       Integer. Required. Specifies the max window size of the
%                       percolated region. Window size = 2*WindowHalfSize+1


%Convert to grayscale and detect edges
gray = rgb2gray(Image);
edges = edge(gray,'sobel');

    % % Optional: Can sample fewer edge detected points for improved computation↙
efficiency
    % se = strel('disk',2);
    % edges = imdilate(edges,se);
    % se = strel('disk',3);
    % edges = imerode(edges,se);


[M N e] = size(Image); %Get image dimensions

%Initiate variables
DetectedCracks = zeros(M,N);
se = strel('square',3);


[x y] = find(edges == 1); %Gather all edge points

for i = 1:2:length(x)
    cp(1:2,i) = [x(i),y(i)];
end

for i = 1:2:length(x) %Visit every detected edge point

    %Start percolating on each detected each point
    if cp(1,i) > WindowHalfSize+1 && cp(1,i) < M - WindowHalfSize && cp(2,i) >↙
WindowHalfSize+1 && cp(2,i) < N - WindowHalfSize
```

```
        ptx = cp(1,i); pty = cp(2,i);

        Block = (gray(ptx-WindowHalfSize:ptx+WindowHalfSize,pty-WindowHalfSize:↙
pty+WindowHalfSize));


        square1 = zeros(WindowHalfSize*2+1);
        Rjected = zeros(WindowHalfSize*2+1);
        square1(WindowHalfSize+1,WindowHalfSize+1) = 1;



        for s = 1:WindowHalfSize*2+1

            New_square = imdilate(square1,se)-square1; %Grow the percolated region↙
outwardly by one pixel

            CurrentPixels = square1.* Block;
            newPixels = New_square.* Block;

            Tresh = max([(mean(CurrentPixels(CurrentPixels~=0)))   , min(newPixels↙
(newPixels~=0))]);


            if  isempty(Tresh) == 1
                break
            end


            %Update Square
            New_square = ~im2bw(Block,Tresh).*New_square;
            Rjected = Rjected+ im2bw(Block,Tresh).*New_square;
            square1 = New_square + square1;

        end


        %Classification
        cc = bwconncomp(square1);
        stats = regionprops(cc, 'Eccentricity','Orientation'); %Compute the↙
Eccentricity which is related to circularity

        ang = stats.Orientation;

        if   stats.Eccentricity > 0.9

            C=1;
            if sum(sum(square1)) > 0

                SEL1 = strel('line', 20, ang+90); %Extend 20 pixels in both directions
                SEL2 = strel('line', 2, ang+90);
                TwoSides = imdilate(square1,SEL1)-imdilate(square1,SEL2);
                ccCor = bwconncomp(TwoSides,4);
                stats = regionprops(ccCor, 'PixelIdxList');

                try
                    S1r = BlockOrigR(stats(1).PixelIdxList);
                    S1g = BlockOrigG(stats(1).PixelIdxList);
                    S1b = BlockOrigB(stats(1).PixelIdxList);
```

```
                    S2r = BlockOrigR(stats(2).PixelIdxList);
                    S2g = BlockOrigG(stats(2).PixelIdxList);
                    S2b = BlockOrigB(stats(2).PixelIdxList);

                    C=mean(abs([mean(S1r) - mean(S2r),mean(S1g) - mean(S2g),mean(S1b) -↙
mean(S2b) ]));


                catch err
                    C = 1;
                end

                if  C < 0.9

                    DetectedCracks(ptx-WindowHalfSize:ptx+WindowHalfSize,pty-↙
WindowHalfSize:pty+WindowHalfSize) =DetectedCracks(ptx-WindowHalfSize:↙
ptx+WindowHalfSize,pty-WindowHalfSize:pty+WindowHalfSize)+ squarel;

                end

            end



        end
    end

end

figure,imshow(DetectedCracks)
figure,imshow(DetectedCracks,[])


end
```

Texture Analysis

```
function [FeatureArray] = StatisticsComputation(Gray,WindowSize)

%Computation of GLCM and descriptove statistics.
%   StatisticsComputation computes a texture characteristics map. The texture
%   characteristics map comprises 4 Gray  Level Co-Occurance Matrix and 6
%   Descriptive statistics.
%
%   Notes
%   -----
%   Each channel of a colour image may be analysed seperately
%
%
%   Convert to HSV or Lab colour spaces
%
%   %HSV
%   hsv = rgb2hsv(OriginalImage);
%   Gray = hsv(:,:,1); clear hsv;
%
%
%   %Lab
%   cform = makecform('srgb2lab');
%   lab = applycform(OriginalImage,cform);
%   Gray = lab(:,:,1);
%
%   The output, FeatureArray, is subsequently classified as being either
%   damaged or undamaged using SVMs
%
%   See also Section 3.2, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%   Example
%   -------------
%       [FeatureArray] = StatisticsComputation(Gray,WindowSize);
%
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%
% Gray:                 Grayscale image or single channel of a colour image. Required.
%
% WindowSize:       Integer. Required. Specifies the window size for GLCM


[M N] = (size(Gray));
FeatureArray = zeros(M,N,10);


%Calculate half the window size
hbs = (round(WindowSize./2)-1);


%Comput GLCM Statistics
for i = hbs+1:step:M-hbs
    for j = 1+hbs:step:N-hbs
        GLCM2 = graycomatrix(Gray(i-hbs:i+hbs,j-hbs:j+hbs),'Offset',[0 1; -1 1; -1 0;↙
-1 -1],'GrayLimits',[],'NumLevels',12); % -1 1; -1 0; -1 -1
        stats = graycoprops(GLCM2);
        c = cell2mat(struct2cell(stats));

        AvContrast = mean(c(1,:),2);
        AvCorrelation = mean(c(2,:),2);
```

```
        AvEnergy = mean(c(3,:),2);
        AvHomogeneity = mean(c(4,:),2);


        FeatureArray(((i - hbs - 1)/step + 1), ((j - hbs - 1)/step + 1),1:4) =↙
[AvContrast, AvCorrelation, AvEnergy, AvHomogeneity];

    end


end


[B Q e] = size(FeatureArray);

E = im2double(imresize(entropyfilt(Gray),[B Q]));

S = im2double(imresize(stdfilt(Gray,ones(7)),[B Q]));
R = im2double(imresize(rangefilt(Gray,ones(7)),[B Q]));
[M Sk] = meanfilt(imresize(Gray, [B Q]), 5);
K = kurtosisfilt(imresize(Gray, [B Q]), 5);



FeatureArray(:,:,5) = E;
FeatureArray(:,:,6) = S;
FeatureArray(:,:,7) = R;
FeatureArray(:,:,8) = M;
FeatureArray(:,:,9) = Sk;
FeatureArray(:,:,10) = K;
[FeatureArray] = FeatureArray;

end



%% Compute the Kurtosis Statistic

function K = kurtosisfilt(Gray, WindowSize)

Gray = im2double(Gray);

hbs = (round(WindowSize./2)-1);

step = hbs;
[M N] = size(Gray);
K = zeros(M,N);

for i = 1+step:M-step
    for j = 1+step:N-step
        I = (Gray(i-step:i+step,j-step:j+step));

        B = reshape(I,1,[]);


        K(i,j) = kurtosis(B);

    end
end
```

```
end


%% Compute the Mean and Skewness Statistics

function [M Sk] = meanfilt(Gray, WindowSize)

Gray = im2double(Gray);

hbs = (round(WindowSize./2)-1);

step = hbs;
[H V] = size(Gray);

Sk = zeros(H,V);
for i = 1+step:H-step
    for j = 1+step:V-step
        I = (Gray(i-step:i+step,j-step:j+step));

        B = reshape(I,1,[]);

        M(i,j) = mean(B);

        Sk(i,j) = skewness(B);

    end
end

end
```

```
function [DetectedRegions, a] = SVM_Classifier(FeatureArray, Gamma, C,  WindowSize,Dx,Dy,NDx,NDy)

%SVM classification of FeatureArray.
%   SVM_Classifier performs SVM classification on the output from the FeatureArray
%   function which is a texture characteristics map.
%
%   Notes
%   -----
%   Training Data
%
%   The output, FeatureArray, is subsequently classified as being either
%   damaged or undamaged using SVMs
%
%   See also Section 3.2, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%   Example
%   -------------
%       [FeatureArray] = StatisticsComputation(Gray,WindowSize);
%       [DetectedRegions, a] = SVM_Classifier(FeatureArray, Gamma, C, WindowSize,Dx,Dy,NDx,NDy);
%
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%
% FeatureArray:     Output from StatisticsComputation function. Required.
% Gamma:            SVM Parameter. Required.
% C:                SVM Cost Parameter.
% WindowSize:       Integer. Required. Specifies the window size for GLCM.
%                   Same as for the StatisticsComputation function.
% Dx,Dy,NDx,NDy     Specify the locations of the damaged and nondamaged
%                   regions for SVM training




%  Dx = x(1); Dy = y(1); NDx = x(2); NDy = y(2); %For selecting the damaged
%  and no-damaged training regions

% Dx = 241.4211; Dy = 364.1333; NDx = 610.6834; NDy = 344.6194; %For script purposes only

Input = FeatureArray;
halfBlocksize = round(WindowSize./2)-1;



%% Custom-Weighted Iterative Classification

% Training damaged regions
dAvContrast = mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:Dx+halfBlocksize,1) );
dAvCorrelation =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:Dx+halfBlocksize,2) );
dAvEnergy =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:Dx+halfBlocksize,3) );
dAvHomogeneity =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-
```

```
halfBlocksize:Dx+halfBlocksize,4) );
dE =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:↙
Dx+halfBlocksize,5) );
dS =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:↙
Dx+halfBlocksize,6) );
dR =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:↙
Dx+halfBlocksize,7) );
dM =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:↙
Dx+halfBlocksize,8) );
dSk =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:↙
Dx+halfBlocksize,9) );
dK =  mean2( FeatureArray(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:↙
Dx+halfBlocksize,10) );


% Non-damaged region
ndAvContrast =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-↙
halfBlocksize:NDx+halfBlocksize,1) );
ndAvCorrelation =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-↙
halfBlocksize:NDx+halfBlocksize,2) );
ndAvEnergy =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-↙
halfBlocksize:NDx+halfBlocksize,3) );
ndAvHomogeneity =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-↙
halfBlocksize:NDx+halfBlocksize,4) );
ndE =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:↙
NDx+halfBlocksize,5) );
ndS =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:↙
NDx+halfBlocksize,6) );
ndR =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:↙
NDx+halfBlocksize,7) );
ndM =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:↙
NDx+halfBlocksize,8) );
ndSk =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:↙
NDx+halfBlocksize,9) );
ndK =  mean2( FeatureArray(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:↙
NDx+halfBlocksize,10) );


% Overall Image
AvContrast = mean2( FeatureArray(:,:,1) );
AvCorrelation = mean2( FeatureArray(:,:,2) );
AvEnergy = mean2( FeatureArray(:,:,3) );
AvHomogeneity = mean2( FeatureArray(:,:,4) );
E = mean2( FeatureArray(:,:,5) );
S = mean2( FeatureArray(:,:,6) );
R = mean2( FeatureArray(:,:,7) );
M = mean2( FeatureArray(:,:,8) );
Sk = mean2( FeatureArray(:,:,9) );
K = mean2( FeatureArray(:,:,10) );


% Calculate Weights

w = ones(1,10);
%%Get rid of abs and put them on the matrixw later
w(1) = (dAvContrast-ndAvContrast)/AvContrast
w(2) = (dAvCorrelation-ndAvCorrelation)/AvCorrelation
w(3) = (dAvEnergy- ndAvEnergy)/AvEnergy
w(4) = (dAvHomogeneity-ndAvHomogeneity)/AvHomogeneity
```

```
w(5) = (dE-ndE)/E
w(6) = (dS-ndS)/S
w(7) = (dR-ndR)/R
w(8) = (dM-ndM)/M
w(9) = (dSk-ndSk)/Sk
w(10) =(dK-ndK)/abs(K)

for ii = 1:10
    if isnan(w(ii)) == 1
        w(ii) = 0;
    end
end
sum(w)
a = w(isnan(w)==0)./sum(w)

%    createBarPlotfigure1(a)


% Automate Procedure
[M N e] = size(Input);
Input(isnan(Input) == 1) = 0;

Array = zeros(M,N);

for t = 1:e

    % Trainig Set
    group = ([ones(1,(2*halfBlocksize+1)) zeros(1,(2*halfBlocksize+1))])';


    Block = [Input(Dy-halfBlocksize:Dy+halfBlocksize,Dx-halfBlocksize:Dx+halfBlocksize,t); Input(NDy-halfBlocksize:NDy+halfBlocksize,NDx-halfBlocksize:NDx+halfBlocksize,t)];
    %Block(isnan(Block)) = 0;
    svmStruct = svmtrain(Block,group, 'kernel_function', 'rbf', 'rbf_sigma', Gamma, 'boxconstraint', C);

    [B Q f] = size(Block);


    %Add an additional Q columns onto input

    PaddedInput = zeros(M, N+Q, e);
    PaddedInput(:,1:N,:) = Input;

    for i = 1:N%-Q+1
        DamagedPixels = svmclassify(svmStruct,PaddedInput(:,i:i+Q-1,t));
        %size(species)
        Array(:,i) = Array(:,i)+DamagedPixels.*a(t);

    end

end

DetectedRegions = double(Array);

figure, imshow(DetectedRegions)
```

REMPS

```
%Regionally Enhanced Multi-Phase Segmentation.
%   The image of interest is pre-processed through a contrast enhancement
%   and colour reduction scheme. Features in the image are then identified
%   using a Sobel edge detector, followed by subsequent classification using
%   a clustering based filtering technique. Finally, Support Vector Machines
%   (SVM) are used to classify pixels which are locally supplemented onto
%   damaged regions to improve their size and shape characteristics.
%
%   Notes
%   -----
%
%   Convert to HSV or Lab colour spaces
%
%   %HSV
%   hsv = rgb2hsv(OriginalImage);
%   Gray = hsv(:,:,1); clear hsv;
%
%
%   %Lab
%   cform = makecform('srgb2lab');
%   lab = applycform(OriginalImage,cform);
%   Gray = lab(:,:,1);
%
%
%
%   See also Section 4.3, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%   Example
%   -------------
%
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%
% Image:            Input colour image. Standard Dynamic Range or High Dynamic Range↙
Image. Required.



[bw] = ImagePreprocessingAndEdgeDetect(Image); %Involves contrast enhancement procedure
bw=~bw; %Invert image
[labeled,numObjects] = bwlabel(bw,8);
Corrosiondata = regionprops(labeled,'all');


%% GUI for selecting the training data for SVM
I(:,:,1) = bw.*im2double(Image(:,:,1));
I(:,:,2) = bw.*im2double(Image(:,:,2));
I(:,:,3) = bw.*im2double(Image(:,:,3));
imshow(I);impixelinfo;
set(handles.text42, 'Visible','on');
[x,y] = ginput(1);
Dx = x(1); Dy = y(1);
set(handles.text42, 'Visible','off');

set(handles.text43, 'Visible','on');
[x,y] = ginput(1);
NDx = x(1); NDy = y(1);
set(handles.text43, 'Visible','off');
```

```
function [output11, Dam_Training, NonDam_Training] = RegionClustering(Image,numObjects,
Corrosiondata,Dx,Dy,NDx,NDy)

%RegionClustering of FeatureArray.
%   RegionClustering computers the mean and kurtosis for regions

%
%   Notes
%   -----
%
%   See also Section 4.2, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%


Image = im2double(Image);

[M,N,e] = size(Image);

MinArea = 0.01*(M*N) %Minimum defect size. Optional. Can be set to 0 if not to be used

% Iniate Values
SmallestDamDist = Inf;
SmallestNonDamDist = Inf;
DamTrainingData = -1;
NonDamTrainingData = -1;


for i = 1:numObjects
    if Corrosiondata(i).Area > MinArea
        t_a = Corrosiondata(i).PixelList;
        t_b = [round(Dx) round(Dy)];
        t_Nb = [round(NDx) round(NDy)];
        if ismember(t_b,t_a,'rows') == 1
            DamTrainingData = i;
        end
        if ismember(t_Nb,t_a,'rows') == 1
            NonDamTrainingData = i;
        end

    end
end

if DamTrainingData == -1
    for i = 1:numObjects
        if Corrosiondata(i).Area > MinArea
            DamDist = sqrt((Corrosiondata(i).Centroid(1) - Dx).^2 + (Corrosiondata(i).
Centroid(2) - Dy).^2);
            if DamDist < SmallestDamDist
                SmallestDamDist = DamDist;
                DamTrainingData = i; % DamTrainingData is the index of the training
data for the damagd region
            end

        end
    end
```

```
end

if NonDamTrainingData == -1
    for i = 1:numObjects
        if Corrosiondata(i).Area > MinArea
            NonDamDist = sqrt((Corrosiondata(i).Centroid(1) - NDx).^2 + (Corrosiondata↙
(i).Centroid(2) - NDy).^2);
            if NonDamDist < SmallestNonDamDist
                SmallestNonDamDist = NonDamDist;
                NonDamTrainingData = i;

            end
        end
    end
end


output11 = ones(numObjects,6);

for i = 1:numObjects
    if Corrosiondata(i).Area > MinArea

        Bin = Corrosiondata(i).Image;

        [ul_corner] =  round(Corrosiondata(i).BoundingBox);

        Im(:,:,1) = Image(ul_corner(2):(ul_corner(4)+ul_corner(2)-1),ul_corner(1):↙
(ul_corner(3)+ul_corner(1)-1),1).*Bin;
        Im(:,:,2) = Image(ul_corner(2):(ul_corner(4)+ul_corner(2)-1),ul_corner(1):↙
(ul_corner(3)+ul_corner(1)-1),2).*Bin;
        Im(:,:,3) = Image(ul_corner(2):(ul_corner(4)+ul_corner(2)-1),ul_corner(1):↙
(ul_corner(3)+ul_corner(1)-1),3).*Bin;

        kur1 = reshape(Im(:,:,1),[],1);
        kur2 = reshape(Im(:,:,2),[],1);
        kur3 = reshape(Im(:,:,3),[],1);


        if i == DamTrainingData
            Dam_Training = Im;
        elseif i == NonDamTrainingData
            NonDam_Training = Im;
        end

        % Computation of the mean and kurtosis
        output11(i,1) = kurtosis(kur1(kur1~=0));
        output11(i,2) = kurtosis(kur2(kur1~=0));
        output11(i,3) = kurtosis(kur3(kur1~=0));
        output11(i,4) = mean(kur1(kur1~=0));
        output11(i,5) = mean(kur2(kur1~=0));
        output11(i,6) = mean(kur3(kur1~=0));


        %Display statistics overlayed on image
        CentroidX = round(Corrosiondata(i).BoundingBox(1) + 0.5.*Corrosiondata(i).↙
BoundingBox(3)) -1;
        CentroidY = round(Corrosiondata(i).BoundingBox(2) + 0.5.*Corrosiondata(i).↙
BoundingBox(4)) ;
        text(CentroidX,(CentroidY + 33) ,['Kred = ',round(num2str(output11(i,1),2)), '↙
Kred = ',round(num2str(output11(i,2),2)), ' Kblue = ',round(num2str(output11(i,3),↙
```

```
2))],'HorizontalAlignment','center','BackgroundColor',[rand(1) rand(1) rand(1)]);
        text(CentroidX,(CentroidY + 3) ,['Ured = ',round(num2str(output11(i,4),2)), '↙
Ured = ',round(num2str(output11(i,5),2)), ' Ublue = ',round(num2str(output11(i,6),↙
2))],'HorizontalAlignment','center','BackgroundColor',[rand(1) rand(1) rand(1)]);
        clear kur Im p
    end
end

%% Add histogrammed clusters to image
[M N e] = size(Image);
Blank = zeros(M,N,3);



for i = 1:numObjects
    if Corrosiondata(i).Area > MinArea
        sum(((output11(i,:) - output11(DamTrainingData,:))./(output11(i,:))).^2)
        sum(((output11(i,:) - output11(NonDamTrainingData,:))./(output11(i,:))).^2);

        DD =  sum(((output11(i,:) - output11(DamTrainingData,:))./(output11(i,:))).^2);
        NDD = sum(((output11(i,:) - output11(NonDamTrainingData,:))./(output11(i,:))).↙
^2);
        if  DD < NDD
            p = Corrosiondata(i).PixelList;
            [v h] = size(p);
            for k = 1:v
                Blank(p(k,2),p(k,1),:) =  Blank(p(k,2),p(k,1),:)+Image(p(k,2),p(k,↙
1),:);
            end

        end

    end
end


output11 = Blank;



end



function [Output model TimeTakenForAlgorithm] = SVM_RegionGrowth(RGB_Image,output11,↙
Dam_Training, NonDam_Training,p)

%Enhancement stage of REMPS: SVM classification of pixels adjacent to identified
%   damaged regions.
%
%   Notes
%   -----
%
%   See also Section 4.2, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%
```

```
%Seperate Colour Channels
OutputR= RGB_Image(:,:,1);
OutputG= RGB_Image(:,:,2);
OutputB= RGB_Image(:,:,3);

[B Q f] = size(OutputR);


%Prepare SVM Training Data for Damaged Regions
DTR = Dam_Training(:,:,1);
DTG = Dam_Training(:,:,2);
DTB = Dam_Training(:,:,3);
A = DTR + DTG + DTB;
DTR = DTR(A ~=0 );
DTG = DTG(A ~=0 );
DTB = DTB(A ~=0 );

clear A

DTR = DTR(1:(2*halfBlocksize+1));
DTG = DTR(1:(2*halfBlocksize+1));
DTB = DTR(1:(2*halfBlocksize+1));

DTR = reshape(DTR,1,[]);
DTG = reshape(DTG,1,[]);
DTB = reshape(DTB,1,[]);

%NonDamaged
nDTR = NonDam_Training(:,:,1);
nDTG = NonDam_Training(:,:,2);
nDTB = NonDam_Training(:,:,3);
AA=nDTR + nDTG + nDTB;
nDTR = nDTR(AA ~=0 );
nDTG = nDTG(AA ~=0 );
nDTB = nDTB(AA ~=0 );
clear AA

nDTR = nDTR(1:(2*halfBlocksize+1));
nDTG = nDTG(1:(2*halfBlocksize+1));
nDTB = nDTB(1:(2*halfBlocksize+1));

nDTR = reshape(nDTR,1,[]);
nDTG = reshape(nDTG,1,[]);
nDTB = reshape(nDTB,1,[]);

BlockR = [DTR; nDTR];

BlockG = [DTG; nDTG];

BlockB = [DTB; nDTB];
[HH,GG] = size(BlockR);
Block = zeros(HH,GG,3);


Block(:,:,1) = BlockR;
Block(:,:,2) = BlockG;
Block(:,:,3) = BlockB;
l = (halfBlocksize*2+1)./1;
```

```
%Initiate array
FourDimArray = zeros(B,Q+2*halfBlocksize,3);

%Populate the array
FourDimArray(:,1+halfBlocksize:Q+halfBlocksize,1) = OutputR;
FourDimArray(:,1+halfBlocksize:Q+halfBlocksize,2) = OutputG;
FourDimArray(:,1+halfBlocksize:Q+halfBlocksize,3) = OutputB;

FourDimArray(:,Q+halfBlocksize:Q+2*halfBlocksize,:) = FourDimArray(:,Q+halfBlocksize:↙
-1:Q,:);
FourDimArray(:,1:halfBlocksize+1,:) = FourDimArray(:,2*halfBlocksize:-1:↙
halfBlocksize,:);


%Trainig Set


BlockN = reshape(Block,[],1,3);
size(BlockN)


group = ([ones(1,ld*1) zeros(1,lnd*1)])';
size(group)
tic


LibSVM_inputs = ['-s 0 -t 2 -c ', num2str(p)];

LibSVM_inputs = ['-s 0 -t 2 -c 1 -h 0 -g ', num2str(p)];
model = svmtrain2(group, BlockN, LibSVM_inputs);%SVM training



bwoutput = im2bw(output11,0.0001);
bwoutputNew = ones(size(bwoutput));
predicted_label = bwoutput;
counterr = 1;


SE = strel('square', 3);
Bounderies = bwoutput;

sum(sum(bwoutputNew-bwoutput))
[Me Ne] = size(bwoutput);
Vis = zeros(Me, Ne,3);
OldPredlabel = ones(size(bwoutput)); NewPredlabel =  zeros(size(bwoutput));

while isequal(OldPredlabel,NewPredlabel) == 0 && counterr <20


    [B Q f] = size(bwoutput);


    Bounderies = imdilate((predicted_label), SE)-(bwoutput + predicted_label); %Grow↙
the damaged region outwardly by 1 pixel
    [indx]=find(Bounderies == 1);

    OldPredlabel = predicted_label;
```

```
    %SVM Classification
    [predicted_label(indx), ~, decision_values] = svmpredict2(ones(length(indx),1), ↙
[OutputR(indx),OutputG(indx),OutputB(indx)], model);

    NewPredlabel = predicted_label;


    bwoutputNew = im2bw(bwoutput + predicted_label);


    counterr = counterr+1;
end



Output = bwoutputNew;

end
```

PaLPaBel Stereo Matching

```
%PaLPaBel Pyramidal Stereo Correspondance Problem Script

%Read in Left and Right images of the stereo pair
Left = imread('C:\Users\Desktop\Left.jpg');
Right = imread('C:\Users\Desktop\Right.jpg');


base_name = 'Image0001';

%%Light Equalisation and Rectification

[LeftCorr, RightCorr] = EqualizeLight(Left,Right); %See funtion EqualizeLight

%Computes the fundamental matrix, See funcion Rectifier below
[LeftRectified1, RightRectified1, t1,t2, inlierPoints1, inlierPoints2,fMatrix] =↙
Rectifier(imresize(LeftCorr,0.5),  imresize(RightCorr,0.5));


inlierPoints_lr2_1 = inlierPoints1;
inlierPoints_lr2_2 = inlierPoints2;
LeftRectified_lr2 = LeftRectified1;
F_lr2 = fMatrix;


[Y, X] = size(inlierPoints1);


%% Uncomment to display Left and Right Images Overlapping
% figure(1), clf;
% imshow(cat(3,rgb2gray(RightRectified1),rgb2gray(LeftRectified1),rgb2gray↙
(LeftRectified1))), axis image;
% title('Color composite (right=red, left=cyan)');


l_min = -150;% lowest disparity value

l_max =  150; % highest disparity value

label = l_max - l_min; %numbr of labels


Pyramid = 3; %No of pyramidal levels


label = ceil(label./(2.^(Pyramid-1))); %Number of labels at the coursest pyramidal↙
level

l_min = ceil(l_min./(2.^(Pyramid-1))); %Lowest disparity value at the coursest↙
pyramidal level


[depthMap Cost] = P_LBP_MRF(LeftRectified1, RightRectified1, label, l_min);


MinValue = min(min(depthMap))
MaxValue = max(max(depthMap))

%%Write Information To Text File

% open a file for writing
```

```
fid = fopen([ 'C:\Users\Desktop\' base_name '.txt'], 'w');
% print values in column order
% two values appear on each row of the file
texxt = ['Stereo Matching Details \n\n Minimum Disparity = %f  \n Maximum Disparity = %↙
f \n No. of pyramids = %f  \n Labels = %f  \n Lower Disparity Limit  = %f \n'];
values = [MinValue, MaxValue,  Pyramid, label, MaxValue ];
fprintf(fid, texxt,values);
fclose(fid);

%Save the depth map
imwrite((depthMap-min(min(depthMap)))./(max(max(depthMap))-min(min(depthMap))), [ 'C:↙
\Users\Desktop\' base_name '.jpg'],'jpg')
imwrite((Cost-min(min(Cost)))./(max(max(Cost))-min(min(Cost))), [ 'C:\Users\Desktop\'↙
base_name '.jpg'],'jpg')



function [depthMap Cost] = P_LBP_MRF(LeftRectified1, RightRectified1, label, l_min,↙
PyrmaidLevels)
%Pyramidal Loopy Belief Propagation based on a Markov Ranodm Field .
%   Accepts the Left and Right image as input, along with the minimum disparity value,↙
l_min, the
%   number of labels, labels, and the number of Pyramid Levels, PyrmaidLevels.
%
%   Notes
%   -----
%
%   See also Section 4.3, Detection and Assessment of Damage in
%   Underwater Marine Structures using Image Processing Based Techniques
%
%
%
%   $Revision: 1.0.0.0 $  $Date: 2015/02/09 04:16:39 $
%
% LeftRectified1:   Left rectified image. Required.
%
% RightRectified1:  Right rectified image. Required.
%
% label:            Integer. No of disparity levels.
%
% l_min:            Integer. Required. Specifies the minimum disparity level

tic % Start timer

lambda = 1;%Belief propagation parameter

BP_iterations =15; %No of iterations of the belief propagation algorithm
Direction = [1,2,3,4,5]; % 1 = Data,2=Left,3=Right,4=Up,5=Down

%Initiate depth map
depthMap = zeros(size( imresize(RightRectified1(:,:,1),0.5.^(PyrmaidLevels))));

%Cost map - pixels matched with high confidence have low cost
CostMap = zeros(size( imresize(RightRectified1:,:,1),0.5.^(PyrmaidLevels))));

% window_radius = 2.^(PyrmaidLevels-pyr);
```

```
for pyr = PyrmaidLevels:-1:1

    if pyr ~= PyrmaidLevels
        l_min = -floor(label/2);
        %modval= 4;
    end

    %label2 = ceil(label/(PyrmaidLevels-pyr+1));%%

    l_max = l_min + label-1;

    %  window_radius = 2.^(PyrmaidLevels-pyr+1)+3; %window size can increase
    %  with each pyramidal level

    window_radius = (3);

    hwait = waitbar(0,'Please wait...');

    Left = imresize(LeftRectified1,0.5.^(pyr-1));

    Right = imresize(RightRectified1,0.5.^(pyr-1));

    depthMap = round(imfilter(imresize(depthMap,size(Right(:,:,1))).*2,hav));

    CostMap = imresize(CostMap,size(Right(:,:,1)));


    pmax = max(abs(l_max),abs(l_min))+max(max(abs(depthMap)));

    tmpLeft = im2double(padarray(Left,[window_radius window_radius+pmax])); %Pad the↙
image
    tmpRight = im2double(padarray(Right,[window_radius window_radius+pmax]));
    depthMap = im2double(padarray(depthMap,[window_radius window_radius+pmax]));



    [M,N,e] = size(tmpLeft);

    %CAche data costs to avoid recomputation

    mrf = double(zeros(M,N,length(Direction),label));


    %% SAD
    % for y = 1+window_radius:M-window_radius;
    %     for x = 1+l_max+window_radius:N-window_radius+l_min;
    %         for l = l_min:l_max
    %             mrf(y,x,1,l-l_min+1) = sum(sum(sum(abs(tmpLeft(y-window_radius:↙
y+window_radius,x-window_radius:x+window_radius,:)-tmpRight(y-window_radius:↙
y+window_radius,x-window_radius-l:x+window_radius-l,:))))./((double(window_radius).↙
*2+1).^2)./3;
    %         end
    %     end
    %
    % end


    %% Normalised Cross Correlation
```

```
    % %Check if a patch has a standard deviation = 0;
    % J = stdfilt(tmpLeft, ones(window_radius.*2-1));
    % bad = find(J == 0);
    %
    % bb = (window_radius.*2+1).^2;
    %
    % while isempty(bad) == 0;
    %
    % tmpLeft(bad) = tmpLeft(bad) + rand(1)./100;
    % J = stdfilt(tmpLeft, ones((window_radius-1).*2-1));
    % [bad] = find(J == 0);
    % 'hi'
    % end

    if pyr < -4 % In case we wish to use different descriptors for different pyramids
        bb = (window_radius.*2+1).^2;
        for y = 1+window_radius:M-window_radius;
            for x = 1+pmax+window_radius:N-window_radius-pmax;

                for ii = l_min:l_max

                    mrf(y,x,1,mod(ii-l_min+depthMap(y,x),label)+1) = sum(sum(sum(abs↙
(tmpLeft(y-window_radius:y+window_radius,x-window_radius:x+window_radius,:)-tmpRight(y-↙
window_radius:y+window_radius,x-window_radius+ii+depthMap(y,x):↙
x+window_radius+ii+depthMap(y,x),:))))./bb ;

                end
            end
            waitbar(y /(M-window_radius))
        end

    else

        %-min(min(depthMap))

        for y = 1+window_radius:M-window_radius;
            for x = 1+pmax+window_radius:N-window_radius-pmax;

                % %          Descriptors: Can use Sum of Squared Distances or Sum of↙
Absolute Distance as a measure of siimilarity. Default: Zero-mean Normalised cross↙
correlation
                % %          for ii = l_min:l_max
                % %
                % %              %%SSD
                % %              % mrf(y,x,1,ii-l_min+1) = sum(sum(sum((tmpLeft(y-↙
window_radius:y+window_radius,x-window_radius:x+window_radius,:)-tmpRight(y-↙
window_radius:y+window_radius,x-window_radius+ii+depthMap(y,x):↙
x+window_radius+ii+depthMap(y,x),:)).^2)))./bb ;%+ ...
                % %              %  sum(sum(sum(abs(diff(tmpLeft(y-window_radius:↙
y+window_radius,x-window_radius:x+window_radius,:))-diff(tmpRight(y-window_radius:↙
y+window_radius,x-window_radius+ii+depthMap(y,x):x+window_radius+ii+depthMap(y,↙
x),:))))))./bb;
                % %
                % %
                % %              %SAD
                % %              mrf(y,x,1,mod(ii-l_min+depthMap(y,x),label)+1) = sum↙
(sum(sum(abs(tmpLeft(y-window_radius:y+window_radius,x-window_radius:x+window_radius,:)↙
-tmpRight(y-window_radius:y+window_radius,x-window_radius+ii+depthMap(y,x):↙
x+window_radius+ii+depthMap(y,x),:))))./bb ;%+ ...
                % %
```

```
                % %
                % %
                % %         end

                %%Norm Corr

                try

                    L = tmpLeft(y-window_radius:y+window_radius,x-window_radius:x+window_radius,:);
                    R = tmpRight(y-window_radius:y+window_radius,x-window_radius+l_min+depthMap(y,x):x+window_radius+l_max+depthMap(y,x),:);

                    Br = normxcorr2_mex(L(:,:,1), R(:,:,1));
                    Bg = normxcorr2_mex(L(:,:,2), R(:,:,2));
                    Bb = normxcorr2_mex(L(:,:,3), R(:,:,3));

                    Bm = (Br+Bg+Bb)./3;

                    mrf(y,x,1,mod((1:label)-1+depthMap(y,x),label)+1) = Bm(window_radius.*2+1, window_radius*2+1: end - window_radius.*2).*-1+1;



                catch exception


                end


            end

            waitbar(y /(M-window_radius))
        end
    end

    close(hwait)

    size(mrf)

    % %Optional: Considering edge boundaries as a prior to belief propagation.
    % %--> No smoothness assumption at edge boundary
    % h = fspecial('sobel');
    % %edges=conv2(im2double(edges),h');
    %  H = fspecial('gaussian',[7 7],3);
    %
    % edgesH = imadjust(abs(imfilter(rgb2gray(imfilter(tmpLeft(1+window_radius:M-window_radius, 1+window_radius+pmax:N-window_radius-pmax,:),H)),h)));
    % edgesV = imadjust(abs(imfilter(rgb2gray(imfilter(tmpLeft(1+window_radius:M-window_radius, 1+window_radius+pmax:N-window_radius-pmax,:),H)),h')));
    %
    % edges = (edgesV + edgesH).^2;

    mrf = mrf(1+window_radius:M-window_radius, 1+window_radius+pmax:N-window_radius-pmax,:,:);

    mrf = min(mrf,1.0);
```

```
size(mrf)

%View naive disparity by block matching
% [MM,NN,ee] = size(mrf);
% for ii = 1:MM
%     for jj = 1:NN
%         Arr = mrf(ii,jj,1,:);
%         disp_min = mean(find(Arr == min(Arr)));
%         depth_map(ii,jj) = disp_min;
%     end
% end

clear  ii jj Arr disp_min Lambdas mv_av
% imshow(depth_map,[]), colormap gray, title('naive block matching')

%mrf = min(mrf,1);

mrf1 = mrf; %store a copy of mrf

clear mrf
mrf(:,:,1,:) = imresize(mrf1(:,:,1,:),1);
mrf(:,:,2,:) = imresize(mrf1(:,:,2,:),1);
mrf(:,:,3,:) = imresize(mrf1(:,:,3,:),1);
mrf(:,:,4,:) = imresize(mrf1(:,:,4,:),1);
mrf(:,:,5,:) = imresize(mrf1(:,:,5,:),1);

[MM,NN,ee] = size(mrf);

%% mrf = mrf1;

msg = zeros(MM,NN,5,label);

%BW = edge(rgb2gray(tmpLeft),'sobel','vertical');

clear tmpLeft tmpRight Tmp new_msg

hwait = waitbar(0,'Please wait...');

Tmp = zeros(MM,NN,label,label); %Initiate MRF

for tt = 1:BP_iterations
    y = 2:MM-1;
    x = 2:NN-1;

    for Dir = 1:5
        msg(:,:,Dir,:) = sum(mrf(:,:,Direction(Direction~=Dir),:),3);
    end

    for Dir = 1:5


        for i = 1:label

            for j = 1:label

                %Matching cost + smoothness cost
                Tmp(:,:,i,j) = msg(:,:,Dir,j) +  lambda.*min([label - abs(i-j),abs↙
(i-j)]);

            end
```

```
                new_msg(:,:,Dir,i)=min(Tmp(:,:,i,:),[],4);
            end


        end


        for Dir = 2:5
            for i=1:label

                if Dir == 2
                    mrf(y,x-1,3,i) = new_msg(y,x,Dir,i);%.*(1-edgesV(y,x));

                elseif Dir == 3
                    mrf(y,x+1,2,i) = new_msg(y,x,Dir,i);%.*(1-edgesV(y,x));

                elseif Dir == 4
                    mrf(y-1,x,5,i) = new_msg(y,x,Dir,i);%.*(1-edgesH(y,x));

                elseif Dir == 5
                    mrf(y+1,x,4,i) = new_msg(y,x,Dir,i);%.*(1-edgesH(y,x));

                end
            end
        end


        waitbar(tt/BP_iterations)

    end


    depthMap = depthMap(1+window_radius:M-window_radius, 1+window_radius+pmax:N-↙
window_radius-pmax,:);

    [depthMapNew CostMap Energy] = MaxAP_Pyr_mod(mrf,label,depthMap);
    CostMap = (CostMap - min(min(CostMap)))./(max(max(CostMap))-min(min(CostMap)));

    depthMap = depthMap + depthMapNew + l_min-1;

    figure,imshow(depthMap,[]), colormap gray, title(['Level ', num2str(pyr)])

    label =floor(label/2)-mod(label,2) + 1; %reset and update label

    close(hwait)


end


figure, imshow(depthMap(:,:),[]), colormap gray, title(['LBP MRF'])

toc


end


%Image preprocessing: Equalize light
```

```
function [LeftCorr, RightCorr] = EqualizeLight(L1, R1)
%Take rgb image and convert to L*a*b* space and then normalize the light component, L,
and re-convert to rgb

%LeftImage
C = makecform('srgb2lab');

lab = applycform(L1,C);
d = stretchlim(lab(:,:,1));
lab(:,:,1) = imadjust(lab(:,:,1),d,[]);

C = makecform('lab2srgb');
LeftCorr = applycform(lab,C);
figure,imshow(LeftCorr)

%RightImage
C = makecform('srgb2lab');

labR = applycform(R1,C);
%d = stretchlim(lab(:,:,1));
labR(:,:,1) = imadjust(labR(:,:,1),d,[]);
hgram = imhist(lab(:,:,1));

labR(:,:,1) = histeq(labR(:,:,1), hgram);


C = makecform('lab2srgb');
RightCorr = applycform(labR,C);
figure,imshow(RightCorr)
```

Michael O'Byrne Trinity College Dublin

Detection and Assessment of Damage in Underwater Marine Structures using Image Processing Based Techniques

Abstract: The monitoring of underwater structures is often beset by limited access, exorbitant costs and inherent safety issues. This thesis focuses on developing and implementing automated image processing based Non-Destructive Testing (NDT) techniques to facilitate inspections. Adopting such techniques can vastly improve the condition of monitoring, reduce the operational complexities and partially offset the financial burden of regular inspections. To date, there has been very little work carried out on image based techniques for the purposes of detecting and quantifying the extent of structural damage, particularly affecting the submerged part of marine structures. This research endeavours to bridge that knowledge-gap through development and performance evaluation of a series of advanced techniques and strategies. Underwater imaging is exposed to numerous challenges such as luminous complexities, poor visibility, light attenuation and backscatter which diminish the ability of the camera, and subsequent image processing algorithms, to effectively identify and quantify instances of damage. This research addresses the deleterious effects of these environmental conditions and phenomena, and through a repository driven approach, maps the impact that lighting, turbidity, and surface type have on the performance of developed techniques.

The developed techniques include a crack detection algorithm, colour and texture based damage detection algorithms, and a 3D shape recovery algorithm. The crack detection algorithm adopts a percolation based approach to automatically locate and quantify cracks in an efficient manner, removing the need for inspectors to manually undertake this tedious task.

The damage detection algorithms employ advanced image segmentation methods to identify and quantify the severity of damage on the surface of infrastructural elements, based on textural information or colour information. Each algorithm is naturally suited to different applications, depending largely on whether the damage form under consideration is more separable from the background based on colour or on texture. Both of these algorithms are validated on real world instances of infrastructural damage acquired from diving expeditions, as well as on above-water instances of damages. A stereoscopic based approach is employed for recovering 3D shape, which utilises a dual camera set-up to simultaneously photograph a specimen of interest from slightly different viewpoints. As part of this approach, an efficient pyramidal loopy Belief Propagation (BP) Markov Random Field (MRF) stereo correspondence algorithm is developed. This algorithm is applied to a submerged pile covered with artificial marine growth in a large-scale testing facility in Boulogne, France, and is subsequently validated on a real word structure in Cork Harbour, Ireland. Having accurate 3D shape information of submerged structural members is of great practical importance when analysing the forces exercised by the waves, winds and currents.

A protocol for the implementation of image processing techniques is established, which specifies the technical requirements for capturing imagery and outlines a set of best practice guidelines to ensure the acquired imagery is suited for quantitative analysis.