# Image processing techniques for the detection and characterisation of features and defects in railway tracks

# Image Processing Techniques for the Detection and Characterisation of Features and Defects in Railway Tracks

CHRISTOPHER IAN JOHNSON

B.Eng. (Hons), Manchester Metropolitan University, UK

A thesis submitted in partial fulfilment of the requirements for the Manchester Metropolitan University for the degree of Doctor of Philosophy

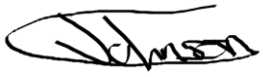
SCHOOL OF ENGINEERING

MAY 2013

**DECLARATION**

This is to certify that the work referred to in this thesis is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). This work has not previously been submitted in any form for another degree or qualification at Manchester Metropolitan University or to any other institute of learning.

Signed:



Christopher Johnson

Date: 01 May 2013

## COPYRIGHT

**ACKNOWLEDGMENTS**

I wish to express my sincere gratitude to my Director of Studies Prof N.J. Bowring and Dr J. Jaiswal of TATA, UK, without whom many things, including this research would not have been possible. Thank you to the whole Sensing and Imaging Group for their continued support and advice. Special mention should be given to Dr Dave Leonard for his continued spirited approach to research.

In a professional role, I would also like to thank Ms. Margaret Fowler and Dr Dave Southall who have supported the development of my professional experience within MMU. Thank you to Dr Peter Twigg for his professionalism and capabilities in dealing with university paperwork; Dr Dalil Benchebra and Shatha Benchebra for their support and friendship, and to all my friends and family, especially my fiancée, thank you for putting up with all the boring conversations about rail.

As a final note, I must thank Dave Rogerson for always finding the perfect solution to a stressful week.

**DEDICATION**

I would like to dedicate the Thesis to Jack, Louise, Manchester Royal Infirmary and The Christie NHS Foundation Trust. Thank you to all those that helped.

**ABSTRACT**

This thesis describes the research that led to the development of a machine vision system in collaboration with TATA, UK and Sheffield Supertram. This was part of a European initiative for Predictive Maintenance employing non-intrusive inspection and data analysis known as PM'n'Idea. The hardware and software design, construction, and evaluation of a prototype for predictive maintenance are presented. The prototype was tested on Sheffield and Warsaw's tram systems.

The prototype has been designed with due account of a specified set of environmental constraints such as a high level of vibrations and space restrictions of the target trams. Special computer vision techniques have been specifically developed to be used with the prototype. Various image processing techniques and algorithms have been evaluated for the purpose of detection and characterisation of a series of rail abnormalities and faults. The system described in this thesis makes use of a number of standard and modified image processing techniques, not only to alleviate the requirements for manual inspections, but also to allow continuous monitoring and tracking of any defects or abnormalities in a rail track. Currently, detecting defects in their earlier stages can only be achieved by using close visual inspection i.e. line walking.

Extensive testing and evaluation of the performance of the prototype inspection system at Sheffield Supertram indicated that the system was able to detect abnormalities with a resolution down to 0.1 mm. Evidence of the classification rates for the standard and modified algorithms that are implemented in the system are presented in this thesis. The algorithms developed show an average success rate of 88.9% in detecting surface bound abnormalities.

**CONTENTS**

## LIST OF FIGURES

## LIST OF EQUATIONS

## LIST OF TABLES

## LIST OF APPENDICES

## NOMENCLATURE

| | |
|---|---|
| PM'n'Idea | Predictive Maintenance employing non-intrusive inspection and data analysis |
| RCF | Rolling Contact Fatigue |
| MVS | Machine Vision System |
| WP1 | Work Package 1 |
| NDT | Non-destructive testing |
| DOF | Depth of Field |
| PCA | Principle Component Analysis |
| LoG | Laplacian of Gaussian |
| FFT | Fast Fourier Transform |
| NCC | Normalised Cross Correlation |
| CAD | Computer Aided Design |
| PCB | Printed Circuit Board |
| IC | Integrated Circuit |
| ROC | Receiver Operating Characteristic |
| FOV | Field of View |
| CCD | Charge-Couple Device |
| MF | Manual Focus |
| AF | Auto Focus |
| FPGA | Field Programmable Gate Array |
| LRT | Light Rail Transit |
| USB | Universal Serial Bus |

| | |
|---|---|
| GPS | Global Positioning System |
| MOSFET | Metal-Oxide-Semiconductor field effect transistor |
| RAM | Random Access Memory |
| SATA | Serial Advanced Technology Attachment |
| JPEG | Joint Photographic Experts Group |
| PNG | Portable Network Graphics |
| AT | Aluminothermic |
| FB | Flash Butt |
| HAZ | Heat Affected Zone |
| S&C | Switches and Crossings |
| NN | Neural Networks |
| ROI | Region of Interest |
| SURF | Speeded up Robust Features |
| IID | In Image Defects |
| MID | Multiple Image Defects |
| TID | Time Image Defects |
| TP | True Positive |
| FP | False Positive |
| TN | True Negative |
| FN | False Negative |
| D2SBM | Dual Data Suppression and Binary Mapping |
| LED | Light Emitting Diode |

# CHAPTER 1 INTRODUCTION

## 1.1 BACKGROUND

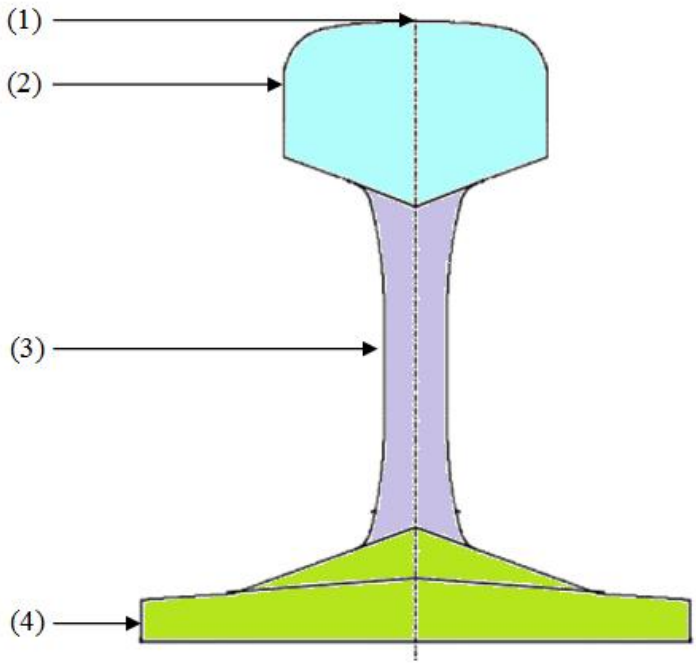
Railway and Tramway track has an expected in-life service of around 45 years. This is dependent on various factors, but mainly: traffic volume i.e. the number of vehicles using the rail; traffic load i.e. the weight of the vehicles using the track and the location i.e. rail on curved parts of track have a lower life expectancy as they are subject to greater stresses. The 45-year mark is the point in which potential fractures of the rail can be expected, although rail has been known outlast this period. Defects of the rail can cut this lifespan to as little as few months.

There are two categories of defects, internal defects and surface bound defects. Internal defects occur during rail production and while internal inspection of the rail does occur after rolling, some internal defects may only start to appear after a period of use. Internal flaws such a failure of steel crystalline structures to bond can cause cracks to form after stress is applied to the rail. Internal defects can also be flaws such as cavities and cracks under just under the rail surface.

External surface rail defects come in many forms, natural corrosion can cause issues, but rail wheel interaction with vehicles causes the majority of defects. Surface bound defects are caused by bending and shear stresses, wheel/rail contact stresses, thermal stresses, residual stresses and dynamic effects that all propagate through the rail head when a wheel of a vehicle passes over its running surface. The common defects that fall within the scope of this thesis are rolling contact fatigue (RCF) cracks, squats, wheel burns, and corrugation. These defects occur on the running surface of the railhead Figure 1.1 shows the relevant parts of a typical rail and the naming conventions of its features.

Bending and shear stresses, and wheel/rail contact stresses can cause micro-fractures on the rail-running surface known as rolling contact fatigue. The same stresses can cause a deformation of the railhead, where an underlying shallow defect can cause two indented lobes on the rail head know as a squat. Wheel slip occurs when the friction coefficient between the rail and the wheel is overcome from the torque applied to the wheel set. This slipping of the

wheel produces a large amount of heat and stress leading to a plastic deformation of the railhead, this type of defect is known as a wheel burn. Finally, corrugation is periodic wear and deformation of the rail surface caused by the wheel set and the bogies' suspension system reacting to a defect such as a squat or an inconsistency in rail placement. An oscillatory reaction occurs as the suspension system reacts to the impact on the wheel set. This causes increased stress in sections of the track between 20 to 200 mm after the source. The stress leads to abnormal wear further propagating the oscillatory action along the rail.




Features:

1. Running Surface
2. Railhead
3. Web
4. Foot

Figure 1.1 Noting the common names of rail track features.

Elmer Ambrose Sperry, who was responsible for inventing the first method of detecting internal rail defects, pioneered non-destructive basic rail inspection (A.B.Allison, 1968). The methods developed have changed little, using electromagnetic induction and Hall Effect sensors, changes in the magnetic field denote the occurrence of defects. E. Sperry used his invention to set up Sperry Rail Service, an American company whose sole purpose became the development and application of non-destructive testing (A.B.Allison, 1968). In 1949, ultrasonic inspection was introduced in an attempt to complement the inductive methods already employed. The advantage of eddy current inspection is its ability to inspect metal nearer the surface while ultrasonic inspection allows a deeper inspection below the surface of the metal (Figure 1.2).

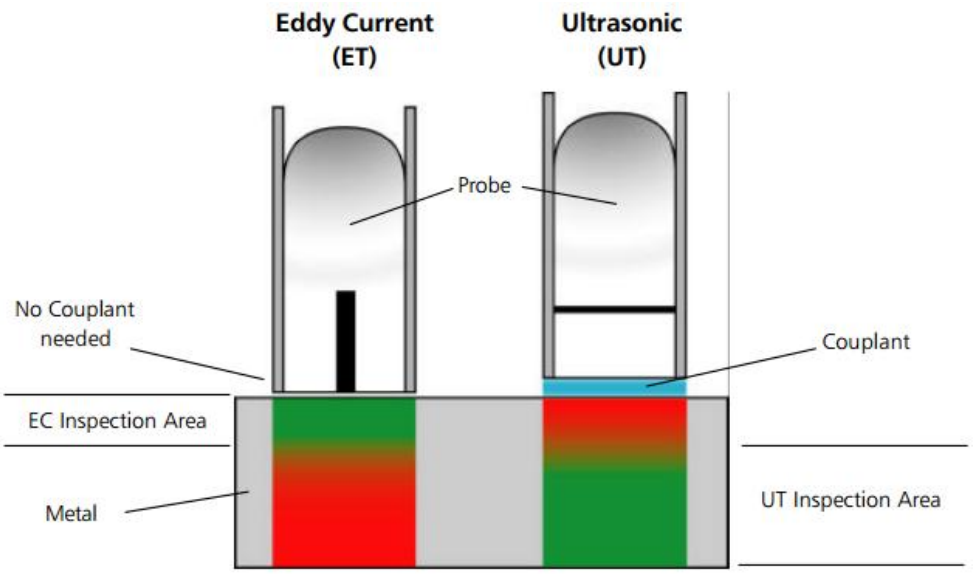



Figure 1.2 Illustrating the typical different inspection areas covered by electromagnetic and ultrasonic inspection methods (Krautkrame, 2012).

In 2000 an assessment financed by the Office of the Rail Regulator and the Health and Safety Executive (Reiff, 2000) recommended that Railtrack, now Network Rail, should pursue the introduction of automatic ultrasonic inspection methods. In the latter half of 2001, Sperry Rail Service was awarded the contract to install and design such a system for the UK rail system (R. Clark, 2004). The system is relatively young in comparison to the rail system; current inspections take place at about 50 mph or 80 Km/h and still require manual inspection to verify the results.

There are several limitations with electromagnetic induction and ultrasonic detecting methods and these limitations are often mistaken for failures of the systems to perform when classifying surface or near surface defects (D.F.Cannon, 1996). These methods were designed to look below the surface of the material, such as the railhead, for internal defects. The limitations of these systems are that they are not capable of distinguishing small surface defects from any noise present during measurements. In some cases, these surface defects have, when left untreated, led to extreme failure of the rail (Office of Rail Regulation, July 2006).

Ultrasonic methods have a strong signal source at the centre of the head of the rail as the signal reaches the surface its intensity dissipates and thus any changes to this signal caused

by surface defects is small and often indeterminable from noise in the system. Electromagnetic inductance has a stronger signal source near the surface of the rail. The standard depth of penetration is 37% of the surface induction value. This is calculated through a test frequency and the magnetic permeability and conductivity of the material.

The permeability and conductivity of rail can change due improvements in the metallurgic combination attempting to improve its lifespan and wear resistance, the foundry in which it was rolled, and even the quality of materials used during production. Larger coils inspect a greater volume of the material since the magnetic field can flow deeper, which is preferred for rail inspection. Smaller coils could be used that are more sensitive to small defects however variations in the permeability of the railhead generates noise that limits the resolution because greater background variations makes their use ineffective.

Electromagnetic induction and ultrasonic are supplemented with the use of visual inspection by rail operators to make up for the limitations. The connection between surface defects and internal defects has yet to be proven conclusively but there are several research cases that provide evidence to link the two (Zili Li, 2008)(D.F.Cannon, 1996)(Corus Rail Technologies, April 2006). It is however, well established that surface defects such as rolling contact fatigue (RCF) cracks can cause major accidents (Office of Rail Regulation, July 2006) and naturally increase in size and occurrence with greater traffic loads on rail systems (S. Bogdanski, 1996). This would suggest that the quicker the defect is located, the sooner it can be repaired, thus limiting the build-up and eventual failure of a rail section.

The adaptation of ultrasonic (Martens George D, 1982) (Jaeggi, 1998)(Pagano, 2013) and electromagnetic induction inspection techniques (Bernard Prasil, 1985)(Ichiro Fujitomi, 2007) by industry and rail operators shows that they are reliable methods of detecting internal rail defects. Nevertheless, few methods can reliably locate surface bound defects such as RCF cracks, early stage squats, or corrugation. As it currently stands, the honoured way of detecting these defects is to have an individual visually inspect the head of the rail for any determining differences from the norm. This is why it remains within the rail assessment policies of rail operators such as Network Rail (Network Rail, 2010)

An RCF crack size can be 0.25 mm or less in thickness and anything above 20 mm in length is considered to be ‘Heavy’ (Corus Rail Technologies, 2000) or severe. It is unlikely that an individual will spot this type of defect every time and defects such as squat and corrugation that are larger and more visible, do not have pre-defined scaling factors to classify severity. The result is that a classification could change between inspections.

Visual inspection carries a severe risk to personal injury or of individuals walking the line due to heavy traffic fatality (Rail Accident Investigation Branch, July 2006). This is more severe on tramways with quicker and quieter electric trams. With a greater demand on tramways and main rail to operate continuously, the timeframe for safe manual inspection is getting smaller and smaller. The same principal applies for track maintenance. Preventing or catching a defect in its earliest stage can allow a faster maintenance procedure to be applied.

Attempts to improve safety have been made with the attachment of cameras on inspection vehicles allowing visual inspection without an individual’s presence on the line (M. Bocciolone, 2007). However human vision techniques are unreliable in detecting and classifying defects, therefore it would be a logical step to move to machine vision system (MVS) as a safer and more productive method. This is a tried and proven step in other areas of industry (C. Demant, 2013).

While MVS are well established, there is an important difference to note between the standard applications within industry and the proposed system. In a majority of cases, machine vision has been introduced to replace individuals inspecting items in a controlled environment (Davies, Jan 2005). In such applications, the object to be inspected is moved within the field of view of the camera at a constant controlled speed. In this application, the inspected object is fixed and camera is mounted on a vehicle moving at various speeds. This introduces several additional factors that have to be addressed such as the variable speed of the vehicle, image blur and distortion, and vibration.

With the introduction of more variables on the acquisition side, the hardware design becomes more significant, requiring a comprehensible specification for each defect’s vision requirements. This will allow acquisition to work in hand with image analysis methods, improving overall system performance.

## 1.2 DEFINING THE REQUIRED OUTCOME

There are two main applications in of MVS in this domain, product accountancy and product maintenance. Product accountancy is the inspection for an objects presence within a frame. It focuses on classification of object type and position, looking for irregular placement and/or non-present items. Product maintenance is a much larger field for an application. The same principles of product accountancy, that of locating objects and their position are required. In addition to this, product maintenance requires the in-depth analysis of features or characteristics that can determine the state within an operating lifetime. This, for example, could be shape, colour, texture and size data, all of which have to be determined for variations over time.

The requirement for both fields varies considerably. The MVS applied to rail inspection is designed for the specific application of product maintenance. The requirements of the end user and the inspection of objects over their life span lead to this application. The classification of an object's position may not be enough to classify the requirement for possible maintenance of the object in question.

The required outcome of this thesis is specifically to develop and test a MVS capable of producing images that allow the classification of rail infrastructure for the purpose of preventative maintenance under the broader terminology of product maintenance. To achieve automated product maintenance of rail, a set of image processing algorithms and techniques have be investigated to examine rail characteristics. Specifically image analysis algorithms have be evaluated and designed to inspect the running surface of the railhead for abnormalities.

## 1.3 RESEARCH ACHIEVEMENTS

This thesis describes the development of the required hardware for successfully acquiring high-resolution images from high-speed line scan cameras in a MVS design. The design processes involved can be applied to the development of specific MVS for various fields. The hardware design uses various technological approaches to associate acquired images with additional information such as the speed of a vehicle during image acquisition, relative GPS co-ordinates, and lighting conditions applied. The hardware design includes the lighting control using high intensity LED lighting technology, the purpose of which is produce images with consistent lighting irrespective of external conditions.

Bespoke algorithms have been developed to work with the hardware to achieve a satisfactory level of detection. This thesis presents an approach for the detection and classification of a range of tramway defects and the classification of some infrastructure. A comparison of algorithms is made with existing techniques, or approaches to similar problems where previous research cannot be obtained. This thesis presents algorithms for the detection of the following features of tramway infrastructure and defects that may occur with this infrastructure:

- Rail edge classification and analysis.
- Running band detection and deviance analysis.
- Squat defects presence with running band data.
- Wheel burn analysis and differentiation from squats.

This thesis additionally contributes to the field of image analysis with developments in:

- Image stabilisation through key point analysis.
- Region growing algorithms and their application methods.

The work of the thesis has also been a key contribution to the PM'n'Idea (Furio, 2012) specifically work package 1 (WP1). PM'n'Idea was a European research project for the development of novel inspection and sensor technologies for rail track infrastructure. The content of this thesis presents significant contributions to the following deliverables:

- D1.1 A laboratory proven image acquisition systems.
- D1.2 Intelligent Image Analysis of selected key prioritised features – Interim.
- D1.2a Extension of the scope of D1.2 to include a methodology of combining the outputs of vehicle mounted sensors with intelligent Image Analysis track inspection.
- D1.2b Intelligent Image Analysis of selected key prioritised features – final.
- D1.6 Additional development of the image acquisition hardware and intelligent image analysis software for the demonstration of the system on Warsaw Tramways.

A key conclusion from this work is that many image-processing techniques cannot readily be applied to solving image analysis problems within industry. A prime example, presented in this thesis, is the application of edge detection, specifically the application of linear edge detection in the classification and analysis of tramway rail edge.

This work will show that standard edge detection methods avert measurement accuracy of imaged objects when applied to higher resolution images. The basis of pre-existing edge detection expects an edge to occur over a smaller range of pixels. While with lower resolution imagery of the past few decades this may be true, with the increase in resolution in digital cameras this edge will start to occur over a greater number of pixels than these algorithms have accounted for. While reducing the resolution of the image to detect edges is easily achieved, this reduction interferes with positioning the correct location of the edge within the original image for measurement purposes. It is for this reason that a bespoke convolution filter that can adjusted to respect the change of edge characteristics achieves better results (Chapter 7 ).

## 1.4 THESIS ORGANISATION

The following thesis is divided into two main sections. The first discusses hardware and the development of the MVS. The importance of the supporting design in image analysis techniques are also addressed (Chapters 3 – 4). The second section describes the image analysis techniques with the detection algorithms and possible defect detection techniques in individual chapters (Chapters 7 - 10). In the case of defects occurring on the head of the rail itself, each technique of defect detection is analysed in its individual chapter to prevent a crossover of descriptions.

Chapter 6 describes specifically the layout of chapters dealing with image analysis techniques. It introduces the data sets involved and the methodology.

# CHAPTER 2 LITERATURE REVIEW

The following literature review divides the scope of research into the following relative fields for investigation.

- Non-destructive testing (NDT) systems that are designed and employed within industry, focusing mainly on those that discuss relevant applications and have similar design features that may be applied or investigated for the purpose of rail inspection.
- A review of systems currently employed for the non-destructive testing of rail, focusing on vision systems that employ video inspection.
- Defect analysis and infrastructure inspection of rail that is not openly linked to the hardware involved in inspection.
- Image analysis techniques that complement the features of defects to be inspected, including object classification for the recognition of infrastructure components.

Image processing concepts relevant to the scope of the thesis are also reviewed, such as:

- Feature extraction i.e.
    - Edge detection
    - Shape analysis
    - Texture analysis
    - Image segmentation
- Object recognition mainly designed for infrastructure classification
- Neural Networks

Programming and streamlining techniques that can improve the efficiency of algorithms are also reviewed focusing on pipelining the image analysis process. This section will examine both the software approaches and modern developments in hardware.

## 2.1 MACHINE VISION SYSTEMS (MVS) IN INDUSTRY

Elias N. Malamas and associates' survey on industrial vision systems (2003) discusses not only the development of a MVS but also the type of image analysis techniques employed within the processing of the resultant image. They demonstrate the varied application the MVS have within industry, from PCB inspection (M. Moganti, 1996) to automatic fruit harvesting (A.R. Jimenez, 1999). After the discussion of some basics of MVS they describe a range of systems, examining what items are inspected, the application of the system, and the depth of field (DOF) required.

Their discussion on the development of a MVS covers the type of hardware required to analyse the images, focusing more on the type of software tools required. They have produced a list of desirable features that are sought after in image analysis software tools. They devote a large amount of discussion to two particular image analysis techniques, namely neural networks and fuzzy logic systems. This paper is particularly useful for use by industry; however, the hardware discussed focuses on area scan cameras rather than line scan cameras.

H. Golnabi and A. Asadpour offer some key design points of a MVS in detail (2007). Separating the design stage into six categories, they discuss the types of application a MVS can have. They note the important design requirements that need to be met in order for a MVS to work effectively. There is significant mention of the importance of lighting within the system to discern image features. This work pays no particular attention to hardware available on the market; however, it does mention 3D systems and the recommendation of laser based lighting.

Yud Ren Chen and associates describe the required components for a MVS focus on Multi-spectral and Hyper-spectral imaging. They list the relevant hardware required and various types available however, not their particular applications (2002). Here, agricultural inspection is discussed, that is a continuous industrial process with no discrete beginning or end. This paper importantly shows that different forms of lighting and light filters can have significant advantages to a MVS.

F. Lahajnar also presents a design of a MVS for continuous industrial inspection (2002). The significant importance of this paper is the MVS design to acquire images for measurement inspection of metal electric plates. To do this they use a multi-camera system, the images are stitched together, and software analyses appropriate dimensions. While current hardware would allow laser measurement to be used, this paper still demonstrates that with appropriate hardware selection, measurements can be made to within ±0.03 mm. Lighting is described as a crucial part of the MVS design.

## 2.2 CURRENT NON DESTRUCTIVE TESTING SYSTEMS USED FOR RAIL INSPECTION

Current research within the field of machine vision for rail/tramway defects is limited to a few specific areas and is usually confined to inspecting for an individual defect. Cannon, Edel, Grassie, and Sawley have produced a comprehensive report of the formation of rail defects (2003). Their work mainly focuses on rolling contact fatigue (RCF); examining both current methods of detecting such a defect and the maintenance methods employed, their scope for application is comparatively large. This paper examines some causes and discusses some modelling techniques, showing why ultrasound does not work in detecting RCF and squat defects (Figure 2.1).

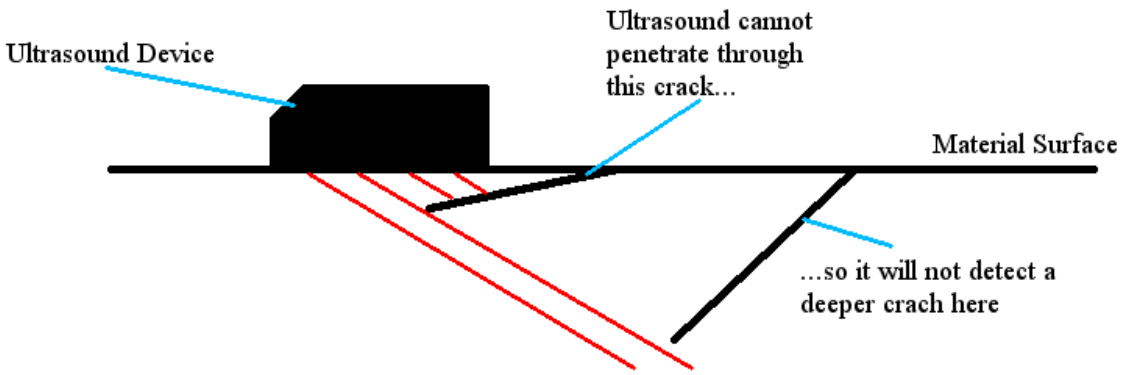



Figure 2.1 Demonstrates how shadowing or masking of an ultrasonic signal can occur. This prevents more severe deeper cracks being detected due to the reflection of the immediate signal (D. F. Cannon, 2003).

Cannon and associates, cover in much detail the control management of defects however, do not go into discussing the possible application of MVS. They focus their investigation to the application of eddy currents and ultrasonic technologies, placing sensors in different patterns to evaluate the potential of the technology.

Clark looks into the needs for future development in rail defect detection (R. Clark, 2004) focusing on the work by Sperry Rail company an American based firm. The paper examines

extensively the work currently put into ultrasonic testing. This work limits itself to how the ultrasonic and electromagnetic systems have evolved. The paper points to how the future will see a move away from specialised vehicles. The paper does highlight the available market and its continual growth rate.

A review of non-destructive testing methods employed in rail inspection is presented by M. Ph Papaelias, C. Roberts and C L Davies (2008). Here they not only discuss the importance of the current types of rail inspection but also the failure of their ability, and the inaccurate maintenance procedure models developed based on rail damage present in the rail network. This paper has various non-destructive testing methods including those involving visual cameras with reference to what speed of analysis can be achieved.

They key conclusion the authors present is that no individual non-destructive method alone can produce a substantial improvement in maintenance and defect detection. Instead, it is suggested that a combination of the methods being employed, to account for the deficiencies of each method, is the only valid path for a solution. This is a fair conclusion since a MVS would have little chance of detecting internal defects just as ultrasonic and eddy current methods struggle to detect surface bound defects.

L. F. Molina and colleagues offer a distinctive insight to current systems under development and those being applied to rail defect detection and infrastructure analysis (2010). Their insight comes from the University of Illinois's contribution to current non-destructive testing for the Federal Railroad Administration. They discuss various MVS's employed throughout the world inspecting different aspects of rail. These include surface defects, rail wear, and infrastructure.

Their summary of inspection technologies is particularly useful, showing that the systems available, do demonstrate the potential for MVS to enhance the inspection of railway infrastructure. They also offer significant data about the causes of derailments from 1998-2009, showing that defects with the rail, joint bar, and rail anchoring caused around 1500 derailments. Track geometry also causes around 1300 derailments.

The demonstration of a Gabor filter is presented by L. F. Molina and colleagues for the detection of switch point bolts, however they fail to ascertain the location of the bolts in respect to the rail. Use of this method could suffer difficulties occurring from noise, alternative infrastructure, and foreign objects without referencing other infrastructure.

S.Sawadisavi et al present a system for detecting and analysing tie plates and ties used with wooden sleepers (2008). This type of fitting is used less and less in favour of more efficient and secure elastic e-clips (more commonly referred to as Pandrol clips). The results presented are of a good quality with the method working in determining location of components by using masks to eliminate ballasted areas. The paper also notes problems caused by ballast interfering with the linear edges from sleepers. Their methods clearly present a positive case for the use of machine vision methods in replacement of current manual inspection methods.

Detection of tie plates and spikes is achieved by using edge detection followed by the Hough transform to classify the tie plate and sleeper. This allows for shape detection through template matching when detecting the spike. In their work, they use area scan cameras in particular; they address the difficulties of lighting this area consistently with use of a strobe light in synchronism with the camera acquisition and the use of a sun shield. Their main concern is the distance at which they operate, making it hard for sun shields to take effect. They show that the use of a line scan camera would solve the illumination issues by referencing a system that inspects joint bars or fishplates (A. Berry, 2008) .

The same research by A. Berry and associates was presented earlier in 2005 to the American Railway Engineering and Maintenance-of-Way Association (B. Nejikovsky, 2005). Using edge detection algorithms to segment the fishplate into regions, template matching is applied to inspect for abnormalities such as cracks. Each crack detected is assessed for length and width characteristics to differentiate from miscellaneous scratches. This report publishes statistics on the methods ability to successfully recognise fishplate cracks. Out of 250 suspected cracks only 98 (40%) were confirmed. A. Berry et al. improves on these results and offers full disclosure about false positives showing the improved algorithm has a failure rate of 17%. While they offer the possibility of 100% true positive classification, they show that the false positive detection

rate will subsequently increase to nearer 100%. It would suggest that the methods alone are not capable detecting a fishplate defect in all cases and additional algorithms are required to improve results.

A. Berry and associates explanation for the false detection is the presence of frayed bond wire, rust, grease spots, weeds or debris. A future development would be to produce algorithms to detect if a possible defect belongs to the fishplate data or could be associated with any of these alternatives being present within the image.

Mandriota et al. have produced a method for the detection of a rail defect (2001), this time focusing on the head of the rail itself and using Gabor filtering to locate and examine corrugation. It is concluded that Gabor filters are successful in corrugation detection with few misclassifications. The paper does not mention if these misclassifications are non-defects being wrongly classified or defects being missed.

Mazzeo and associates produce some interesting research for the detection of fastening bolts and Pandrol clips (2004). The images they use are taken directly over the head of the rail. By pre-processing the image using a discrete wavelet transform and principle component analysis (PCA), and using additional image features such as the position of the sleepers, they demonstrate good results in bolt and Pandrol clip infrastructure detection. Using these methods, they are able to segment the relevant sections of the image so that a neural network can be used to classify the presence or lack thereof fastening bolts and Pandrol clips used to secure the rail.

S.Yella et al. present interesting research in the automated inspection of wooden railway sleepers (2007). As well as using a MVS, they also use an impact acoustic method to support their analysis. This method proved less successful in classifying defective and non-defective sleepers with a typical success rate of 76%. They claim that the MVS they used presented a success rate of 90%. This data is not supported by the original master's thesis presented by S. P. Mohammad (2008). Here the success rate is presented at 86% for both methods of Support Vector Machining and Radial Basis Function. Both publications use 150 training images and only 50 test images. It

is likely that their training set is simply not large enough. Similarly, their test set is restricted in size.

H. Trinh et al. present a more complete system for the detection of infrastructure surrounding the rail (2012). Using a multiple camera-based system the application of tie plate detection is addressed. Tie plates are used to secure rail to wooden sleepers. Assuming that tie plates can be successfully detected, the paper addresses the use of cross correlation for positive detection. It uses the location of the tie plate from each camera looking at either side of the rail head and corrects miss-classification by taking appropriate positioning of tie plates and inferring a corrected position if they do not correlate.

The system demonstrates how data from various sources or camera views increase the robustness and capabilities of machine vision systems. The investigation assumes that a tie plates will be positioned perfectly opposite itself each side of the railhead, and it assumes that the cameras FOVs are also aligned perfectly. Due to vehicle movement and flaws in track installation, this method will require extensive development. On straight sections of track, the method proposed produces favourable results however; upon entering any rail curvature, tie plates are not aligned. Subsequently on these curves the cameras field of view (FOV) will change, as a result the algorithms may suffer from a reduction in accuracy.

Recent work by Q. Li et al. discusses how vision systems can be used to inspect the railhead for abnormalities at speed (2012). A high-speed line scan camera is used to acquire an image of the rail at speeds up to 75 km/h. This speed is limited by the FOV and camera set up. They use multiple stages of contrast enhancement to evaluate for surface bound abnormalities. Their system is simple and robust; it relies on the reflective nature of the rail surface to define the running surface of the rail. Over sampling of the image is applied by manipulating the contrast. By examining the longitudinal and lateral data of the image, a region of interest can be formed around the abnormalities to evaluate the size. The system proposed however, is limited just to the railhead surface and any other features such as the rail edge and supporting infrastructure cannot be evaluated.

J. Espino et al. discuss in depth various application for the detection of rail within an image (2012). They assume a camera is mounted on the front of the vehicle, approaching the problem looking from a driver's perspective. Evaluating current methods, they propose the best approach is to use a simple edge detection method to pick out the key characteristics within the image. This is then followed by a robust but strict set of rules that work with the characteristics of the railway structure by evaluating the position of the rail a fixed point and following its path throughout the image. Although relying on correctly classifying the first point the potential for development is apparent with learning conditions that could be introduced. Such a method could be to use a successful classification to estimate the pose position of the rail in the next occurring frame.

B. T. Nassu and M. Ukai show an application of tracking the rail from a driver's perspective (2012). A series of cameras are used to predict the position that a camera should be focused and directed, i.e. the direction the driver would look. The rail is modelled in the image as a sequence of parabola segments to estimate the pose for a pan and tilt camera control system. A robust detection system is presented by evaluating the edge classification results over sections of the image. The rail is modelled over a short distance in which the rail will be prominent and show less change in direction. It is then modelled over a long distance, in which more noise will be present and curvature of the rail may be visible. Estimation of the direction of the rail compared to some known patterns can then be applied.

The almost template matching type approach allows a scaling system using the more reliable data from the short distance section to evaluate a direction trend. The less confident long distance data can then be used with this to evaluate a pan and tilt direction for a camera. While some good results are presented, it is noted that complexities caused by rail infrastructure and other large objects in the distance can cause significant error in the estimated pose for the camera position. It is worthy to note again that edge detection with some robust rules for matching the rail characteristics has proven successful as with work presented by J. Espino et al (2012).

## 2.3 IMAGE ANALYSIS TECHNIQUES

### 2.3.1 Feature Extraction

Feature Extraction is a pre-requisite in any image analysis algorithm that intends to perform some form of classification. Within image analysis, a large focus is put onto two common features, edges, and corners. Both have several methods of classification and with detection of any of these features, there is the obvious need to classify their position relevant to each other or known positions within the image.

The Hough transform (Hough V, 1962) is a common and useful shape detection technique, particularly for line detection (for example when detecting the rail edge). It has a well-tested history that may not have use for detecting defects directly, however it will aid in feature detection that will assist in infrastructure classification. Linear edge detection is extremely useful for detecting the edges of rail and keeper, if present. This allows for the infrastructure of the program to be organised so that relevant features will be inspected for on relevant sections of track. It also allows for segmentation of the image so that task parallelisation can be performed when looking for features in different regions of the acquired image.

For the Hough transform to work accurately, a well-designed edge detection algorithm is applied to the image first. The theory of edge detection is a well-established research area with machine vision (Davies, Jan 2005) and image processing (M. Sonka, 2007). There are many well-known edge detection techniques. These include Canny (1986) , Sobel(I. Sobel, 1978), Roberts (Feb 1988), Laplacian of Gaussian (LoG) (Milan Sonka, Sep 1998) and the zero-cross method (D. Marr, 1980). These methods use a variety of different statistical techniques to detect edges within an image. All of these methods are very efficient; however, results vary depending on settings and feature types within an image.

A zero-crossing point is the point in which a sine wave intersects the x-axis transferring from positive to negative value or vice versa. Marr and Hildrith use this as the basis of their work, describing that an edge within an image is a point of intensity change developing from a defined

Gaussian operator. Limiting the rate at which intensities can change, LoG (1980) is based and developed on blob detection, and is the same as convolving the image with a smoothing function and taking a Laplacian function of the results. Applying a Laplacian function has the effect of creating a double-edged image from which the edge can be derived with the smoothing function removing any noise (R. C. Gonzalez, 2004).

Guest, et al. demonstrate the use of statistical methods such as the F and T tests to reduce false edges caused by clutter, texture within an image, and cross over from edges (2004). This method shows how each angle has to be examined to give a clear detection of edges. This requirement is further demonstrated by the Prewitt and Sobel methods (I. Sobel, 1978) that allows the orientation of a mask to be changed so that edges along certain angles are detected. The Prewitt is similar to the Sobel method however computationally it is simpler to implement. It is important to note that in practice produces more false classifications that may affect further measurements and statistical calculations (R. C. Gonzalez, 2004).

Konishi, Yuille, and Coughlan extend their own work with Zhu (1999) to demonstrate different forms of statistical inference to classify edge cues within an image (2003). This work shows alternative statistical tests must be applied to multi-scale edge detection to determine which form of statistical analysis produces the best results.

The Canny edge detector is a robust method. The image, in a manner similar to other methods, is smoothed using a Gaussian filter. A gradient magnitude image is then formed to determine the location of the edges (Canny, 1986). The Canny edge detector is used widely for machine vision purposes (Milan Sonka, Sep 1998). There are still possibilities, however, for changing the pre filters used on the image or adapting the method itself to ignore features with certain dimensional differences such as small areas of noise (Lijun Ding, 2001).

There has been several papers published that look into possible ways of resolving the best edge detection methods. Such as work done by Heath et al. who reference and use human vision scoring systems (1997) whereas others, try for a more reliable statistical approach such as

Yitzhaky and Peli that uses receiver operating characteristics (ROC) and the Chi-Square test to produce statistical data for analysis (2003). In this work, a method based on the two approaches has been developed to ensure a more optimal selection of edge detection methods.

Neural networks are an alternative to the proposed edge detection techniques, offering some advantages and disadvantages to using a neural network that are discussed in the Chapter 2, section 3.4.

The methods discussed so far only use greyscale images to determine the location of edge cues within an image. With the advancement of high-speed colour camera, using colour images as the source for analysis is a future possibility. Koschan reviews the use of colour information from images in the use of edge detection (1995). Koschan reviews methods of colour edge detection against adapted existing greyscale methods such as Sobel.

Martin, Fowlkes and Malik have produce an extensive paper that introduces not only a possible way of testing the performance of edge detection within colour images but also a novel set of rules for interpreting edge cues (2004). The importance of this paper is that it introduces the use of several features to determine edge cues such as brightness, colour, and texture cues. This paper clearly demonstrates how the use of more than one analysis technique can be used to improve upon previous methods.

A requirement of this work is to remove unwanted noise within the image and the join the strong edge features to form a complete line (e.g. for rail edge detection). The Hough transform, previously discussed is an established algorithm for this process. However, it struggles with complex shapes with broken lines. Line following algorithms as demonstrated by Chouinard and Plamondon (2007) or Ni and Guansekaran (2004) are good alternatives and can be used to record the XY co-ordinates for subsequent analysis of line position and length.

Many edge detection algorithms may detect an edge cue over multiple pixels. This can cause edge following algorithms to struggle in their classification of edge direction. Line thinning

algorithms as demonstrated by Lee and Wang (C.L. Lee, 1999) and more recently by L.Liu and associates who develop line thinning with complex 3D skeleton descriptors (L. Liu, 2010) can improve line following algorithms. L.Liu et al. have developed an algorithm that removes itself from the iterative process of line thinning algorithms such as those used by Lee and Wang.

Due to many rail defects having distinctive colours and features, and having distinctive colour variations it would be beneficial if this data can be used. Colour segmentation and colour grouping can fall into the same group of literature. Cheng et al. discuss in depth the advances in colour image segmentation techniques (2001); they provide a clear description of how a colour image can be used to greatly improve results in comparison to grey scale methods.

Swain and Ballard show how by using colour histogram analysis comprehensive solutions can be formed by creating stable object representations and analysing the changes in data(1991). They call this method '*color indexing*'. Their application is in the use of robotics and used for the solving of location problems. Funt and Finlayson extend this method in an attempt to account for the changes in varying illumination conditions (1995). This is applicable to some machine vision systems as lighting conditions can vary due to a multiple of environmental influences.

There is an interest within this field to see if keeper wear or the head of the rail could be extracted from the 2D images acquired. L.G. Roberts presents a Ph.D Thesis on the subject (L. G.Roberts, 1963). His works shows that if the depth of perception of a 2D image is known a sufficiently accurate model can be assumed to calculate 3D data within the image. While complex, he provides proof that using a set of invariant shape models the use of mathematical transforms can be utilised to provide depth information. His work focuses on basic shapes with ideal images. While applying such research to a noisy image such as that acquired from a MVS, may be difficult there is a possibility 3D data could be extracted. The accuracy of this data cannot be predicted however. Recent advances in 3 dimensional image acquisitions with the use of lasers have left any such research largely redundant (Stemmer Imaging, 2011).

R. Archibald et al. present interesting research for the extraction and segmentation of a human brain from an MRI scan (2003). They utilise the Gegenbauer reconstruction method (D. Gottlieb, 1992) to successfully improve the segmentation of an image from edge detection results. They demonstrate that noise within the image can be suppressed by the Gegenbauer method. The work suggests that with an advanced mathematical map of the expected key features within an image, segmentation is achievable regardless of noise. A simpler approach of producing a plausibility map detailing where the rail edge can and cannot appear for example will allow for increased success in classification algorithms.

### 2.3.2 Texture analysis

The idea of texture analysis for machine vision is widely researched and there are numerous techniques. Gabor functions and Wavelets are a well-developed methods demonstrated by Manjunath and Ma work that presents a comprehensive paper on using a Gabor filter to provide good pattern retrieval accuracy (1996). This technique however, like many others use adaptive filters to improve quality of any image with various content, while useful it is not ideal. This technique however can be adapted to filter out just the features that are required.

Mondriotaet al. demonstrates the application of this technique for the detection of corrugation (C. Mandriota, 2004). This paper clearly shows how Gabor filters can be applied to rail defect detection for texture analysis. The paper however, fails to take into count the receiver operating characteristics and evaluate the methods poor misclassification of defects.

There are several clear alternatives to a Gabor, wavelet, or combined filter. A mathematical filter can easily be formed using previous research into texture analysis in materials science. Bunge gives a clear description and analysis method of how to examine textures of different materials (Bunge, 1982) however, like most research in materials science it is based on greater resolution of imagery than that which can be obtained from a moving train.

Using a mathematical filter would not be applicable for polymer defect classification were the texture of non-defect areas change due to the use of different polymer compositions. A more

advanced method would be to use a neural network that can be trained up on all textures being a defect or not. Ruan et al. demonstrate the use of colour in texture analysis utilising neural networks by looking for scab infection on wheat (R. Ruan, 1998).

Chen Pau and Wang have reviewed papers that deal with texture analysis techniques (C.H. Chen, 1999). Fukunaga presents a good background using statistical pattern recognition or a mathematical filter for texture analysis (Fukunaga, March 1999). A clear set of statistical approaches for classifying sections of the image are given such as the Bayesian, linear, Quadratic, Piecewise and Sequential classifiers that all use various statistical techniques.

The flaw with a mathematical filter is in its execution time. Unlike a neural network each possible type of classification will be examined to assign the most likely designation of type. Whereas a neural network will examine all the data, and based on its order will assign the most likely designation without having to compare all possible routes. Y. Pao further demonstrates a neural network's capability for texture analysis and discusses the four most common functionalities of texture analysis, clustering, learning functional mappings, classification through associative recall, and optimization (Pao, 1999).

A. Monadjemi presents a paper on the user of Gabor filters for detecting abnormalities in textures (2005). A bank of pre-designed filters is applied to an image the results are then analysed. Rather than applying a blob detection filter to detect anomalies a quantisation algorithm is used. This quantisation algorithm is applied to a normalised image. Analysing this image the number of features or possible abnormalities can be calculated. The use of such a filter bank could be considered for corrugation, squat, and wheel burn detection.

### 2.3.3 Object recognition

The technique of object recognition within image analysis is used extensively like edge detection. Object recognition is deemed a useful scope of research for infrastructure such as fishplates and Pandrol clip detection. While fishplates occur anywhere between every 9 – 120 meters Pandrol clips are present at a maximum of once every meter on standard rail. Loncaric

gives a good analytical review of object recognition methods (1998) reviewing the background techniques and grouping articles of individual solutions. Loncaric lays down a clear understanding of what features are important in object detection and a clear understanding of what each method is capable of doing and its subsequent limitations.

Statistical methods are the most practiced and are extremely useful in detecting complex reoccurring patterns. Schneiderman and Kanade discuss in detail the method of statistical object recognition in face recognition (2000). They discuss in detail the use of normalising the image and the requirement to use differently orientated patterns to locate facial features at different angles. Schneiderman and Kanade use Bayes decision rule that is also used in texture analysis referred to as a Bayesian filter.

Olson and Huttenlocher demonstrate a method for object recognition by using edge-orientated pixels to allow quicker classification of an object (Clark F Olson, 1997). They show that rather than using a whole mask of an object, selecting specific points of interest or defining points i.e. orientations of certain edges, that an object location can be more efficiently located and classified.

A neural network has the potential of superseding a statistical approach however, the use of a Fourier analysis method may be considered as a competitive alternative. The Fourier analysis method offers significant improvements on previous statistical methods. Tsai an Lin show this by using the efficient fast Fourier transform (FFT) with normalised cross correlation (NCC) in the application of defect detection in printed circuit boards (PCB's) and missing integrated circuits (IC's) (2003). Lewis also discusses the use of NCC and its application (1995) paying particular reference to its computational speed he expands to discuss it performance in the application of special effects image processing.

Cootes et al. present work in object classification with the use of Active Shape Models (1995). Thier work uniquely limits the deformation of the contour models to thresholds presented by the training data. The methods presented for object classification are robust and could be used to evaluate deformation of the object due to damage. The method however is limited by the ability

to classify key points of the object. These key points presented by corners and edges need to be sufficient in quantity to allow classification.

In the application of rail classification, there are significant limitations. If a line scan camera is used only two points, the railhead edge, can be used to classify the shape of the model. This would be an insufficient amount and changes between each acquired line of data would prevent a model of the whole rail within an image being used. If an area scan image were used there would be far more reliable key points presented by the edge of the rail. Potentially Cootes et al. method could correctly classify the railhead even when presented with noise from the ballast however; the algorithms robustness relies on an iterative approach to evaluate an object. This iterative cost of computational time might be outweighed by simpler classification methods. There is also the reduced benefit of this system since the rail head will not hold enough detail to allow rail wear to be evaluated. Cootes et al. approach holds much more significance where more sufficient data can be presented such as the data produced by 3D profiling systems.

### 2.3.4 Neural Networks

Neural networks are extremely useful with image analysis problems. They remove the need for a complex function to be applied to the image data in favour of hiding this from the developer and forming its own based on data that it is trained upon. Neural networks allow complex functions to be solved quickly. There are numerous fields that they can be applied to, for instance to classify edge cues which is a recent application. The resultant neural network has produced promising results. Zwaag and Slump demonstrate its capabilities when trained correctly in performing consistently well against current edge detection techniques (2002).

Neural networks are also a solution to image segmentation problem. Pal and Pal have reviewed several segmentation techniques (1993) and concluded that new neural network model based algorithms were an appropriate and efficient replacement for older techniques. It must be noted however, that they base their research on a methods capability of dealing with noise and not in its classification ability.

A neural network works similar to a statistical method by examining every pixel and its surrounding points. The advantage is that a neural network only inspects a given pixel once. Ripley (2008) heavily criticises neural networks based on their lack of a parallelised architecture and there execution upon a single processor. This is now not the case and such a view is already heavily outdated as new computer architecture allows this parallelisation quickly and easily.

Carpenter and Ross (1995) demonstrate the effectiveness of neural networks especially when continuous training is introduced they train a neural network not only on an original set of data but also data of which it struggles to classify. Carpenter and Ross show that only with accurate and comprehensive training can the neural network succeed statistical methods. With a solution in parallelisation, having been developed over the last few years a neural network can now supersede statistical methods.

## 2.4 PROCESSING AND STREAMLINING TECHNIQUES

### 2.4.1 Software

Large software libraries of image processing techniques are available. Some of these are commercialised, closed source projects and have good user support. The alternative are open source projects that are not as focused, but allow better configuration and optimization. These libraries are often a collection of well-known image processing techniques such as Hough transforms, neural networks, common edge detection methods (Sobel, Canny, etc.), and others that are utilised in vision systems. Using such libraries can cut the time required for the development of MVS, as this optimised code can organise and extract key image data points for quicker analysis.

One of the better well-known libraries is the Open Source Computer Vision Library more commonly referred to as OpenCV. Originally developed by Intel the library is now supported by robotics research laboratory Willow Garage. Llorens and associates examine the OpenCV image-processing library in the application gesture recognition (2002). While their system is a basic example of OpenCV's abilities, their conclusion is legitimate. They note the ease of which with the C++ language that the libraries are produced in can be incorporated into many others such as .NET (stated as Visual C++ and Visual Basic), Java and Matlab.

Ji and Haralick propose a method of using the variance of a convolution, on synthetic and real image data sets, to actively calculate edge detection performance of different techniques (Q. Ji, 1999). A comparison plot of the calculated variance and of the Receiver Operating Characteristic (ROC) curves for false alarms against misdetection rate shows, that by examining the output variance of an edge detection algorithm, using different kernel sizes, a balance between performance and accuracy can be estimated. Performance is increased by selecting the smallest possible kernel size without sacrificing accuracy, the smaller the kernel the faster the edge detection method.

Bowyer, Kranenburg, and Dougherty propose a similar method of evaluating edge-detecting algorithms against one another in correct classification of known edge cues (K. Bowyer, 1999). While such a testing method has a limited scope to applications in which data sets are consistent, MVS fall consistently within this scope. Data sets recorded by MVS will always be of a consistent type. Bowyer, Kranenburg, and Dougherty use a set of images in which edge cues are known. Each edge detection algorithm is applied to these images. ROC curves are then used to examine which algorithms produce the highest accuracy. ROC curves of each algorithm are compared to suggest the most compatible for the training set.

Their research fails to account for the use of adjusting the convolution kernel size that while effecting execution performance can increase edge cue classification. Accounting for the kernel size when applying such a test would accommodate for an appropriate edge detection algorithm to be selected for specific MVS. The ROC curve as proven by Ji and Haralick can also be used to evaluate the ideal convolution size with the sacrifice of the minimum amount of accuracy.

It is the sacrifice of a small amount of accuracy for a great improvement in performance that is often used to speed up image analysis. With less data, any algorithm will be quicker, this boils down to a hardware limitation on the speed of which any amount of data can be calculated. With advances in hardware, such superannuated views of reducing data sets should be put to rest in favour of pipelining the processing operation.

The increasing hardware advancements in CCD sensor size and lens quality means key features such as, an edge cue that previously would occur over an individual pixel could now occur over several. Ji and Haralick propose the underlying theory of their work is important for the design of a new edge detector. They suggest there is a requirement for new advanced algorithms that make use of this extra available data.

### 2.4.2 Hardware

M.N. Fesharaki and G.R.Hellestrand present a custom VLSI design for an Integrated Circuit (IC) that performs the Kolmogorov-Smirnov test for edge detection (1993). They utilise MODAL a hardware description language to produce a parallel test for four edge orientations (Hellestrand, 1980). They pay particular attention to the timing constraints of their image acquisition device set at about 40 ms per frame. They successfully produce a design that can allow the parallel examination of an image for edge detection of 512x512 greyscale images. They successfully prove that IC based hardware can improve the efficiency of image analysis by producing a heavily parallel architecture. This allows pixels to be examined under multiple rules in one pass rather than multiple passes.

I.S. Uzun et al. present a similar application producing and field programmable gate array (FPGA) to implement a wide range of Fast Fourier Transform (FFT) algorithms, such as radix-2, radix-4, split-radix and fast Hartley transform (FHT) (2005). FFT algorithms are widely used for signal processing and in image processing for filtering and template matching. They produce a comparative study of the FPGAs execution time verses the execution time of similar general-purpose processors and purpose built digital signal processors (DSP). Comparison of execution time is flawed comparing their quickest implementation (radix-4) against undisclosed FFT algorithms implemented by other processors and DSPs.

I. S. Uzun et al. also compare the use of parallel executions, something that FPGA architectures are designed to implement. The study shows that parallel executions over multiple threads can increase the performance of FFT algorithms. Their work shows how the increase of performance is more evident for the algorithms that have a longer execution time, such as the FHT method. An increase in performance factor of x7 is seen for FHT when eight parallel processors are implemented. The performance gain shown is related to the original execution time of the algorithms, with the least performance increase being shown for the faster radix-4 FFT algorithm. I. S. Uzun et al.'s work demonstrates that by increasing the parallel processors involved the closer the execution time becomes across different implementations of the FFT.

P. Dillinger et al. present a medical imaging problem based on image segmentation (2006). They compare the software implementation of the 3D Gray-Value Structure Code (3-D-GSC) for image segmentation against an FPGA hardware implementation. Their application uses a specialised FPGA co-processor board that allows a large amount of managed random access memory and in-built camera link interface to manage the passing of camera data directly to the FPGA. They do not implement the algorithm on live acquisition but compare performance increase on a simulated 3D MRI brain dataset.

P. Dillinger et al. presents a performance increase factor of x30 – x100 depending on resolution. The greater performance is an evaluation of software performance with no optimisation versus hardware application with full optimisation. This performance increase demonstrates that hardware implementation of algorithms is faster than software however until live acquisition data is presented to the hardware other limiting factors such as data transfer speeds cannot be evaluated. P. Dillinger et al. present their performance increase for estimation purposes only.

# CHAPTER 3 DESIGN OF A MACHINE VISION SYSTEM

## 3.1 INTRODUCTION

There are numerous aspects to consider in the design of a Machine Vision System (MVS), for this application. It is necessary to produce a clear and concise specification, covering aspects such as the Field of View (FOV) and the smallest detail to be examined within the frame i.e. the resolution.

These requirements play an active role in several aspects of the systems design; the type of camera employed the type of lens, and the type of lighting. Generally, there is a compromise between the FOV and the resolution due to the technology available. Consideration must be given to the environment in which the MVS is to be deployed, namely the space for camera and lighting deployment. If a large FOV is required, the mounting height of the camera becomes important. This will affect the obtainable resolution; a good rule of thumb is that an object will appear twice as small as the camera-to-object distance is doubled.

## 3.2 DESIGN CONSIDERATIONS

The Depth of Field (DOF) plays an important role in the design of this particular system. Many MVS inspect a surface at a particular distance. The DOF of an image relates to the nearest object within the FOV and the farthest object. The objects that fit into a set DOF will appear in focus and those that are outside of the DOF appear blurred. The DOF is limited by the distance to the object and relates to a lens's focal ratio (f-number) and focal length.

The DOF is calculated as follows (Larmore, 1965):

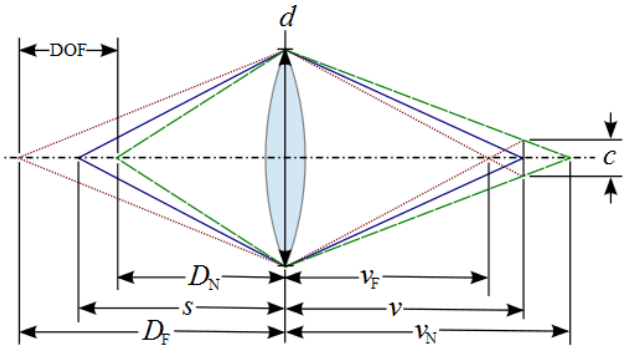



Image: http://en.wikipedia.org/wiki/Depth_of_field

$$D_N = \frac{sf^2}{f^2 + Nc(s-f)} \qquad \& \qquad D_f = \frac{sf^2}{f^2 - Nc(s-f)}$$

Equation 3.1 Method to calculate Depth Of Field (DOF) achievable with lens and environment specifics (Larmore, 1965).

Where:

- $D_N$ = Depth, Near Field
- $D_f$ = Depth, Far Field
- N = f-number
- f = focal length
- s = prime objects distance
- c = circle of confusion (diagonal width of CCD)

An important consideration in machine vision application is the light required to saturate the Charge-Couple Device (CCD). This becomes an important factor when a high focal ratio as used to increase DOF. Increasing the focal ratio reduces the aperture diameter and thus reduces the light received from the object to the CCD. This in turn causes increased diffraction and places a practical limit on the DOF by the amount of light intensity that can be applied to a particular FOV (Equation 3.2). A subsequent problem is that the amount of noise within the image increases with diffraction. In MVS design, where objects are not a consistent distance from the camera, the DOF and lighting must be carefully considered.

MVS differ further from standard photography due to the need for the constant acquisition of images. The application of high-speed digital photography using CCD technology introduces a constraint on the responsivity of the CCD, or its quantum efficiency. The responsivity is the time required for a CCD to be fully saturated to acquire an image from visible light or more generally an accurate measurement of the device's electrical sensitivity to light photons. The responsivity differs between colour cameras and greyscale ones with the latter type of CCD having a far higher responsivity and the possibility of a higher acquisition rate.

The quantum efficiency of a CCD is usually measured over several wavelengths. This is due to a photon's energy being inversely proportional to its wavelength. The frequency at which a colour camera can acquire frames is usually limited by the responsivity of the blue and green light spectrum that has the shortest wavelength of 475 nm to 520 nm (Equation 3.2). Greyscale cameras utilise a broader light spectrum (400 – 900 nm) and subsequently have a higher responsivity allowing a higher frequency of image acquisition.

$$I_o = \frac{16 * N^2 * I_s}{In_p * 1e^5}$$

Equation 3.2 Calculation for the energy required per unit area to saturate a CCD.

Where:
$I_o$ = Intensity Out
N = f-number/focal ratio
$I_s$ = Fluence needed for saturation of a Pixel (wavelength dependant)
$In_p$ = Integration time of a pixel

In some MVS applications, if the fine details within the image are not required for image analysis then the FOV can be quite large and subsequent resolution low. This is often the case with large items that have simple shapes.

In this application, the resolution is important due to the smallest object to be viewed i.e. rolling contact fatigue crack. The FOV must be calculated to ensure the object fits within it and the CCD is of sufficient size to give the required resolution. In such conditions, the DOF can be limited in the achievable ranges of resolution and FOV; however, this can be addressed by increasing the

focal ratio of the lens. The design then relies on providing an intense light source that can still ensure saturation of the CCD. This is often obtainable as the object is closer to the camera and related light sources.

If the object cannot fit into a FOV with a desired resolution then a crossover condition exists that will require greater care in the design stage and the use of multiple cameras.

The FOV is calculated as follows:

$$\frac{o}{d} = \frac{i}{f}$$

Equation 3.3 Method to calculate Field Of View (FOV) with given environment, lens, and camera specifics.

Where:

- f is equal to the focal length of the lens
- d is the object distance from the lens
- ***o*** is the object dimension (or "field of view" perpendicular to and bisected by the optical axis).
- i is the image dimension (or "field stop" perpendicular to and bisected by the optical axis).

In the application of tram and light rail inspection there is a crossover condition. If certain parts of tramway infrastructure are to be analysed, then multiple cameras will be required. In previous research, multiple cameras have been used (L. F. Molina, 2010)(S. Sawadisavi, 2008). The difficulty with this technique becomes evident within the data processing stage where the data is multiplied by the number of cameras used. A large amount of existing research focuses on the analysis of individual parts of infrastructure e.g. Pandrol clips, whereas this thesis addresses the inspection of multiple parts of the infrastructure and the head of the rail.

The camera position is limited by its placement under an existing tram bogie. This has a significant effect on the FOV. It is due to these limitations that the resolution is only dictated by

the image analysis algorithms and the cost of higher resolution cameras. Particular attention has been focused on the mounting position and camera angle for maximising image analysis success.

The DOF is a particular complication within the environment in question. The distance of the head of the rail to the foot is sufficiently large to require a large focal ratio. To provide saturation of a CCD at a high acquisition rate with such a large focal ratio is considerably difficult. This is limited practically by the lighting technology available and considerations for safe use.

## 3.3 MVS HARDWARE SELECTION

### 3.3.1 Camera Types

Railway inspection is a continuous inspection process, for which two types of cameras are suitable; line scan and area scan cameras. Both cameras are similar in their image acquisition abilities, however each is suited for different applications. Area scan cameras capture an area/image similar to that of a standard digital camera, but with generally higher resolution and a greater continuous speed. The main industrial applications are component inspection, where items to be examined can be stopped and inspected by eye, or by automated means, and where the object is contained within an individual frame (E. N. Malamas, 2003)(H. Golnabi, 2007).

**Advantages**

- Shorter Exposure Time: All pixels of the charged–coupled device (CCD) can be saturated simultaneously, whereas a line scan must build up an image.

**Disadvantages**

- Fixed resolution in both width and length.
- Fill factor of 80 per cent or less for the CCD.
- Resolution limited by the number of pixels on a 2D sensor.
- Higher resolution can be achieved, at a greater cost.

- Blurring effect at speeds.
- Uniform lighting of the frame is difficult.

Line scan cameras capture an individual line at a high rate. Successions of lines are acquired to form an image. The main application areas of this type of camera is the continuous industrial process, where the object being inspected moves at a constant speed and where there is no discrete beginning or end to the inspection.

**Advantages**

- Low cost for high resolution.
- Easier to produce a uniformed light source due to smaller area being acquired.
- 100 % fill factor.
- Better suited to applications with no discrete beginning or end.

**Disadvantages**

- Determining line rate and synchronising it into the movement of the product.
- Alignment and focus are more critical.
- Dealing with the data, that forms images can be more complex - e.g. if there is movement between line scans.

There will inevitably be an area of crossover, where the entire item does not fit within a desired FOV. There are two trails of thought for this problem. The first is to employ an area scan camera and take a multitude of images and either stitch them together to form much larger image or alternatively inspect each frame according the part being detected.

There are inherent difficulties with aforementioned methods to negate FOV limitations. A key difficulty will be lighting. A non-uniform lighting source can introduce shadowing or bright spots that can interfere with the image analysis stages. The main issue that limits the capabilities of image stitching algorithms is that they require several points of which to match one image edge with the other. The overlap required for image stitching must also be factored into FOV calculations making the system much harder to design.

A more valid solution would be to employ a line scan camera. The key advantage of these types of camera is that the image size is not theoretically limited. There is an inherent need to store the images in sections, but stitching them together is as simple as adjoining the bottom of one image with the top of another.

Line scan cameras have a significantly larger FOV than area scan cameras; this difference is down to technology limitations, and the increased cost of CCD die size. Where a line scan camera only requires an individual line of semiconductors, an area scan camera requires many more. A higher resolution is therefore available from line scan cameras at a reduced cost. The die size is limited by technology, meaning a limited resolution for both cameras of how many semiconductors can fit onto an individual die.

Crossover conditions often require more than one camera to be employed. There are numerous ways of implementing these, two examples are shown Figure 3.1. In this example, four area scan cameras are required to obtain the dimensions of the coloured box. Only two line-scan cameras are required to obtain the same data. There are further mechanical complexities involved with the area scan camera configuration, where all four FOVs must align or there overlap accounted for in software. With line scan cameras, this complexity is halved and with vertical binning of pixels during acquisition, alignment is considerably easier.

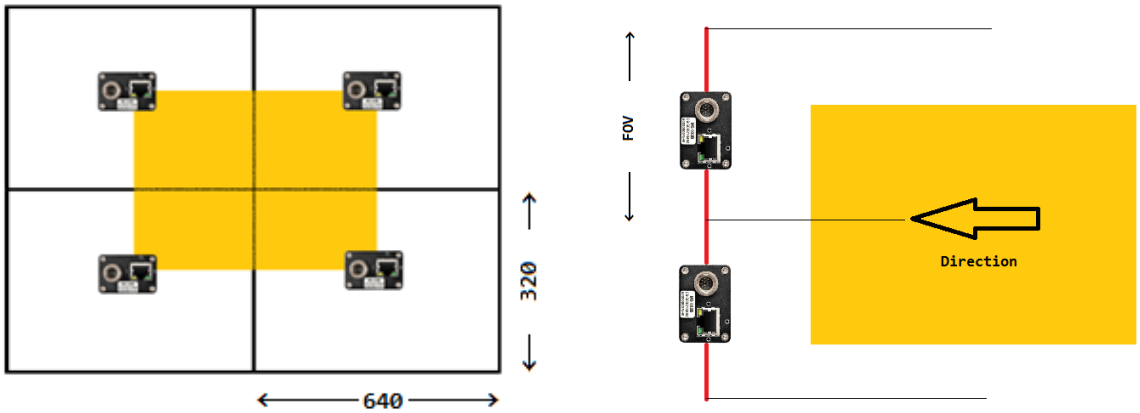



Figure 3.1 An example camera layout configuration for comparison between area scan cameras (left) and line scan cameras (right) when acquiring a large object (yellow) that exceeds a cameras' field of view (FOV).

### 3.3.2 Colour Verses Greyscale

The biggest factor in greyscale or colour camera selection should not be cost or speed but in determining the image processing methods to be involved in feature detection.

The principle selection can be done by determining how an individual sees a feature. If it is a colour difference, this does not mean that a greyscale camera cannot be employed. If the colour features are strong , prime colours then a greyscale camera can still be used. If it is a texture difference then selection is more difficult, although a greyscale camera can still be used by utilising coloured lighting and various filters.

It is generally preferred to use Greyscale cameras as they have a higher acquisition speed and lower cost than colour cameras. However, it must be noted that there is no direct alternative to a colour camera as it produces three different light spectrums of data that can increase the capability of feature detection algorithms.

### 3.3.3 Preliminary Investigations

Preliminary investigation can be made by using a standard digital camera. While these cameras do not have adequate resolution in comparison to area scan cameras, there technology is in effect the same. By examining the results, the mounting position camera angle and type of camera required can be derived. If the objects to be examined fit within the FOV at the required height and their occurrence can be accounted for, then an area scan camera can be opted for. If they do not then a line scan camera can be more appropriate.

By taking images of the feature to be examined at different angles, an appropriate camera angle can be chosen that will aid in the image analysis. An appropriate camera angle will provide unique datasets of a chosen feature to aid in its determination from surrounding data. This could be, for example, a unique colour, bright spot, shadow, or combination of them. Experimentation can be achieved at a far smaller cost with this type of approach as demonstrated in Figure 3.2.

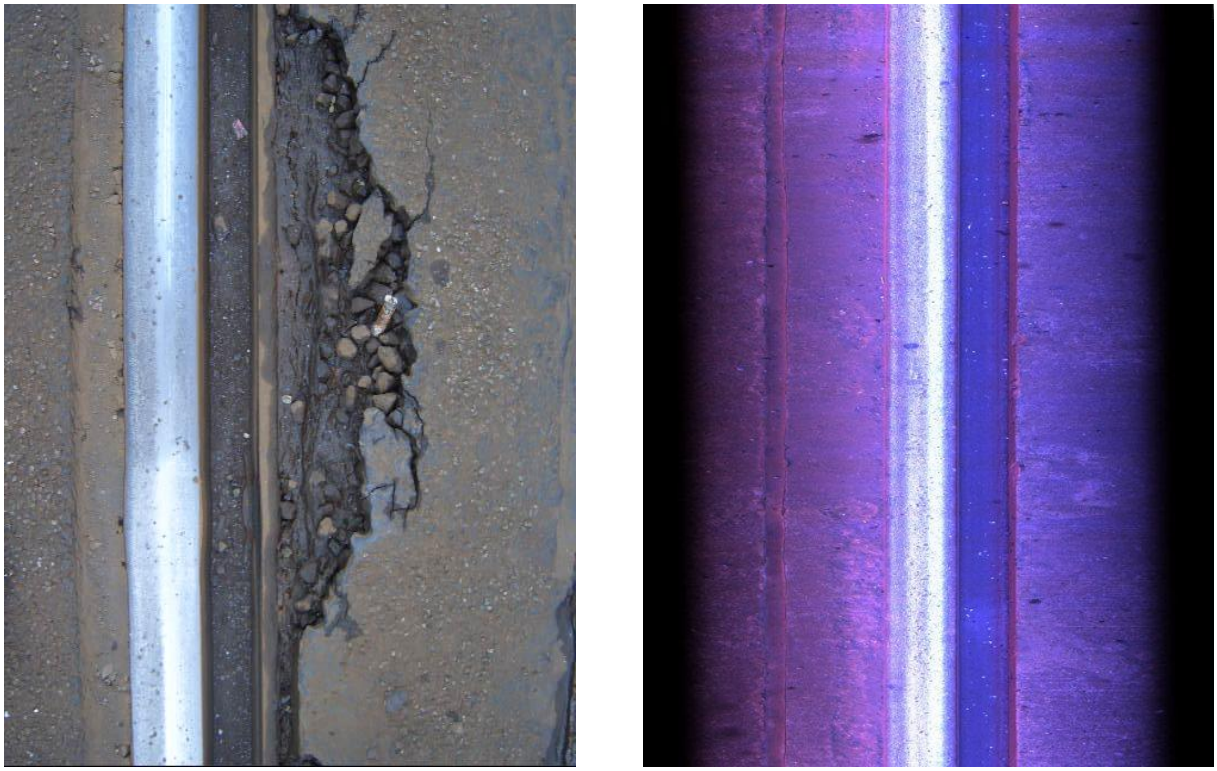


Figure 3.2 Example of an image acquired from standard digital camera (left) compared to that acquired from high speed line scan camera (right).

An alternative to using a standard digital camera for preliminary investigation is the use of simulated models. 3D computer generated models allow scenes to be generated with realistic light sources. While this process can be time consuming, it can have significant importance where images of objects are to be acquired in harsh environments. An output similar to that of a digital camera can be achieved allowing multiple camera positions and angles to be evaluated.

With a selection of sample images, it is important to apply some basic image analysis techniques. The performances of methods such as the Hough transform (Hough V, 1962) to detect linear features within the images will give a good insight into the correct camera position and angle for the feature.

### 3.3.4 Lens Selection

A majority of modern MVS cameras have a lens attachment so that standard lenses can be employed. The quality of the optics employed determines the loss of light as it passes through a lens system. If a MVS is acquiring images at speed this loss of light could affect the saturation of the CCD if the light source does not provide enough energy to account for this loss.

Lens selection is extremely important as it affects the FOV and subsequently the resolution, speed of acquisition, and depth of field. To aid in lens selection, a good specification needs to be developed and a range of lenses around that selection should be tested.

Fish-eye lenses are desirable in applications where a wide angle from a restricted height is required, as they will produce a greater FOV. The resulting effect on the acquired image is curved outer edges that make objects harder to recognise for human and computer algorithms as features are distorted. It is important that Fish-eye lenses are not employed unless image correction problems have been examined. The image distortion can be corrected through calculating the camera intrinsic. This can be done using information about the camera or through images taken of objects with known specifics, fixed patterns with fixed spaces between

detectable points. The sacrifice of this is that data in the outlining field of view is suppressed and lost in the distortion correction.

Prime lenses are often recommended, a prime lens traditionally has a fixed focal length. If the situation the camera is to be employed suffers from severe movement or the camera distance from the object is variable prime lenses cannot be employed. For testing, a zoom lens is adequate and a prime lens can later be employed. It is important for testing purposes that a lens is selected with Manual Focus (MF) as MVS cameras do not regularly employee Auto Focus (AF) capabilities. It is also recommended that lenses should support different focal ratios so the DOF can be adjusted.

When forming a lens specification, the first point of interest is the lens working length (focal length). Equation 3.3 (Page 67) is implemented in different ways according to the camera type employed. The type of camera effects the *image dimension* value *i* which is the diagonal dimension of the CCD.

This in turn affects the result of the *object dimension 'o'* which is the diagonal measurement of the FOV in mm. To calculate the actual height and width of the FOV it is important to realise that the FOV is directly proportional to the size of the CCD. See Appendix 1 FOV Calculation using a Line Scan Camera and Appendix 2 FOV Calculation using an Area Scan Camera for the different calculations employed depending on camera type.

### 3.3.5 Lighting

As shown in research of MVS, lighting plays an important role in the quality of image that can be acquired (H. Golnabi, 2007)(F.i Lahajnar, 2002)(Yud-Ren Chen, 2002). The type of lighting employed depends on the type of camera to be employed. Traditionally with area scan cameras a strobe light is employed that is linked with the acquisition frequency of the camera. This allows similar operation to the flash of a standard digital camera; however, with a line scan camera this approach is not ideal.

A line scan camera forms an image by taking a succession of lines and building a full picture line by line. This means a constant light source is required due to the high rate of acquisition. This is usually achieved by a line light or similar configuration.

Another consideration is the environment in which the light is to be deployed. If there is interference from daylight, the available light within the environment must be considered, and the hardware light dimmed, or the use of extra hardware such as a sunshield should be used.

There are two types of lighting readily available; the traditional halogen style lighting that provides a high intensity light. The heat dissipation from this type of light makes it hard to focus without expensive glass type lenses. These lenses are not suited for harsh environments. The bulbs employed in this lighting are liable to break under vibration. This in turn affects image acquisition and quality, proving costly both in hardware and application costs. The lifespan of a halogen bulb is typically 750 – 1250 hours that limits its usefulness in continuous inspection tasks. Despite this, it has become popular as it produces a wavelength spectrum similar to that of humans visual perception of brightness known as the luminosity function. This allows images to be acquired that visually resemble the object that was photographed.

Alternatively, the use of LED lighting has become more efficient and practical. These are more reliable and more capable of operating in harsh environments. Over the past few years, the technology has evolved to produce LED with a light intensity comparable to halogen light sources at lower power requirements. These LEDs, while producing lower heat dissipation, allows for heat produced to be directed via the use of heat sinks making focussed beams easier to work with.

A key advantage, but also a disadvantage, of LED lighting is the selection of optimum light wavelength dissipation. Greyscale cameras have peak sensitivity to lower and near infrared wavelengths between 600 – 800 nm. LED lighting can be designed to provide more energy at those particular wavelengths. A disadvantage does occur however with colour cameras, as typical LEDs tend not to have a luminosity function similar to a humans visual perception. The

results are images, which do not appear, in colour at least, the same as the objects that were photographed. This can easily be corrected with colour correction algorithms.

### 3.3.6 Frame Grabbers

Frame grabbers are traditionally designed to convert the data received from the camera into a format for computer use. This is usually in the form of an image. With the increased speed of frame grabber interfaces and development of Field Programmable Gate Arrays (FPGA), the ability for the frame grabber to pre-process images acquired in real time has become possible. FPGA processors are integrated circuits that can be designed to operate as streamline processors to perform specific tasks. While such frame grabbers do not allow real time processing of images that require complex algorithm analysis, corrections to the image such as histogram equalisation can easily be achieved. While custom analysis can be implemented, the design is time intensive.

The advantage of the more modern frame grabbers is that the requirement on the computer processor has been alleviated. This allows the processor to concentrate on complex algorithms either for image analysis or for the correct storing and deployment of the data to storage medium.

Such devices, however do pave the way for more complex FPGA designed systems. These frame grabbers demonstrate, the real time processing possibilities for the application of image analysis using external FPGA designed systems. Such systems would not deal with the image acquisition and focus only on image analysis, helping speed up the system. Such devices could be imperative in the application of neural networks. FPGA systems allow a heavily paralleled architecture to be implemented at low relative cost.

The minimum requirements of a frame grabber are laid out by the capabilities of the camera employed. Any frame grabber must be able to compete with the acquisition speed and resolution of the camera. This defines both camera interface type and PC interface type.

## 3.4 SUMMARY

In any potential MVS system, pre-processing of the image should be eliminated by hardware design. Such pre-processing as histogram equalisation for example corrupts data received directly by the camera, potentially allowing software to overwrite key characteristics of features. This may introduce complexities in image analysis algorithms. Before any pre-processing is introduced, images should be inspected in the raw format to grade the ability of the hardware architecture and allow better design of image analysis algorithms.

In designing of the specification of a MVS and its testing the following flow should be taken in determining the hardware. Using educated choices based on the area discussed, and using the appropriate equations to confirm selection MVS design can be optimum for the solution.

1. Environment

- Estimate a set of appropriate constraints, a mounting point of the camera height from the inspection area, camera angle, and available space for adjustment.
- The mounting area for lighting and distance from the source without interfering with a cameras view should be evaluated.

2. Camera Selection:

- Classify the type of inspection task. If it is continuous then select a line scan camera, if it is of an individual object then an area scan camera would be appropriate.
- Appraise if a colour camera is required, this should be evaluated primarily on the presence of colour-based features.

3. Resolution:

- Determine the smallest object or measurement that must be detected.

4. Lens Selection:

- Select a lens according to the focal distance of the object to be imaged.
- Calculate the FOV to see if the resolution required is achieved and if the required object features fall within the FOV.
- If the requirements are not met, reconsider the lens selections, if no alternative lens is available select and evaluate the use of an alternative camera or the use of multiple cameras.

5. Depth of Field (DOF):

- Calculate the DOF and decide upon an appropriate focal ratio (f-stop).
- If the DOF is not achieved, review the Lens selection before reassessing the environment settings.

6. Lighting:

- An appropriate lighting source based on camera type and object features must be selected. In some cases, having a contrasting light source at an angle may complement features. Care should be taken to match the luminosity function to spectral sensitivity of the camera in favour to that of humans. This will improve image quality overall and colour correction can always be made to match the luminosity function displayed in the image to that perceived by human vision.
- If saturation of the CCD cannot be achieved using the light source selected it must either be enhanced to provide greater power either optically or electronically, before reconsidering in order, the DOF, Lens Selection, Camera Selection and then finally reassessing the environment by moving the camera or lighting position.

7. Frame grabber:

- A majority of high-end cameras used in the market will come with specific frame grabber support however if certain aspects of image processing are speed dependant a frame grabber with programmable FPGA support may be required. A frame grabber should support all available functions of the camera selected.

8. Image Quality

- Finally image quality must be assessed by applying some know algorithms to determine an objects features. If desired results cannot be obtained or if the feature is not visible within the image then lighting should be re-examined.

# CHAPTER 4 DESIGN OF A MVS FOR THE ACQUISITION OF TRAMWAY RAIL IMAGES

A precise specification was developed before the design of a MVS was undertaken. A specification was developed between MMU, Tata, and Stagecoach Supertram; this ensured all parties understood what should be expected from the MVS. Once formed, a laboratory system was designed to test and help aid the design of additional hardware, particularly the lighting. Appropriate cameras were chosen, based upon FOV and DOF calculations.

The system required the development of a bespoke lighting system. Two such systems were developed to produce a uniform very high intensity lighting source. The lighting sources make use of modern LED technology to produce uniform lighting for an image with a varying DOF. Associated fittings for the hardware and lighting systems were developed in partnership with Stagecoach Supertram, allowing the hardware to be mounted onto existing tramway bogies.

## 4.1 DESIGN SPECIFICATION FOR A TRAMWAY MACHINE VISION SYSTEM

These design specification points were provided by Tata and Stagecoach Supertram, based upon the details previously presented and initial investigation within the field. The following specification describes the requirements of the MVS that acquired high-speed images of tramway track and infrastructure.

- Must fit on an existing vehicle within fleet
    - Current Non-Destructive testing methods such as eddy currents and ultrasonic are often implemented on an individual inspection train that has the sole purpose of rail inspection. The cost to Stagecoach Supertram in acquiring new vehicles is currently around £3.5 million. A system on this scale is not seen as an acceptable cost for the company to incur; therefore, the designed system must adaptable to mount on current vehicles within the fleet.

- Non Vehicle Specific
    - The design must be of a size that it can be mounted on any vehicle within the fleet. Due to the limited amount of 25 vehicles at Stagecoach Supertram, only one vehicle is usually not in service. This vehicle is situated inside the depot for maintenance. The design must be of such that it can be mounted upon a vehicle with minimal effort, as to prevent interruptions to operations outside the depot preventing any delay to customers. As the Sheffield Supertram vehicles were made, and purchase at different times their layout has changed slightly through new devices being installed to old ones being removed or reduced in size. The design must account for the smallest room on a vehicle within the fleet.

With these specifications and by consultation with design engineers at Stagecoach Supertram, it was suggested that mounting the device onto a bogie would give a consistent view of the head of the rail. If the camera were to be mounted to the main body of the tram itself, the swing of the body would cause the position of the head of the rail to move unacceptably within the acquired images when on a bend in the track.

- Limited space for deployment
    - Retrofitting a design onto current vehicles within the fleet and developing the design so it is non-vehicle specific reduces the space allowed for deployment. Figure 4.1 shows where on the tram the largest available space is underneath the vehicle is located. The space available denotes a design size must fit in the space of 550 mm height by 650 mm width, taking into account the swing of the bogie. See Figure 4.2 an image of the suggested mounting position for acquisition device deployment for reference to the space available. The Mounting bracket denotes where the system can be fixed to the vehicle bogie.
    - The Stagecoach Supertram system has one of the tightest turning circles currently employed by LRT (Light Rail Transit) systems in the UK. This turn is situated within the depot itself; however, on circuit the tram body can still experience a swing of up to a meter. This means that any protrusion by the design into the

650 mm width will risk interaction with the side skirts of the tram body and risk damage to the tram itself and/or the equipment employed.

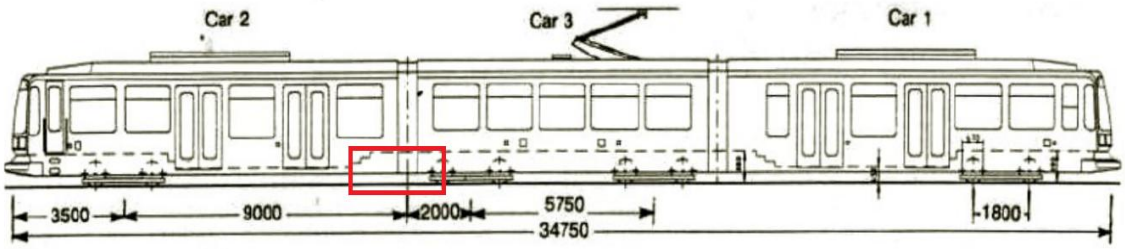



Figure 4.1 Location of the suggested mounting position for acquisition device deployment.

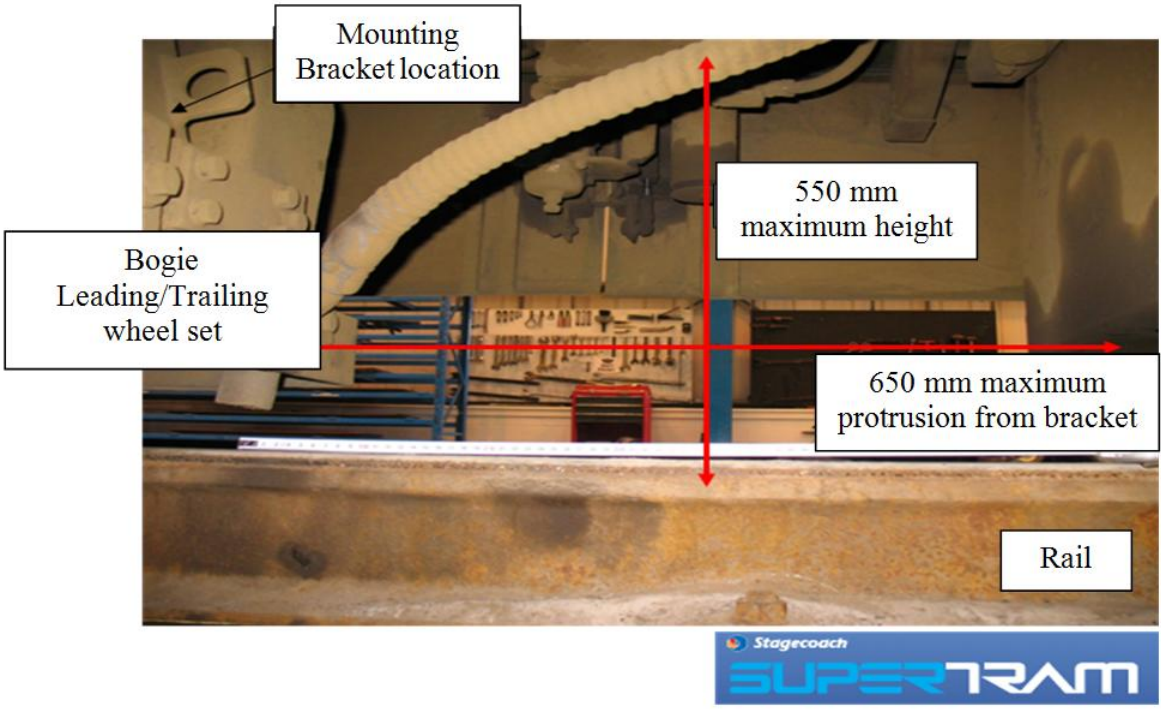



Figure 4.2 Image from the suggested mounting location for the acquisition equipment.

- Record Location of Image Acquisition
    - It is essential that any image acquired has associated GPS data so that it can be visually inspected and examined. There is a two part agenda to using GPS positional data. The first is to test and establish the systems capabilities in comparison to visual inspection techniques. The second is to establish the use of GPS in directing maintenance crews to a position in which they can locate the defect. Both agendas must be met in order to make the systems deployable.

- The visible infrastructure within the images acquired
    - It was agreed that Tata and Sheffield would generate a documentation that would list infrastructure currently installed on Stagecoach's Sheffield tramway network. While visible infrastructure is limited not only by the hardware employed but also the limited height for deployment it was agreed that at the very least the system must be able to visually see the head of the rail and keeper where applicable, and fixtures such as Pandrol clips, Fish plates and relevant bolts employed. All other visible infrastructure would be of a beneficial nature to the system.
    - The documentation of the infrastructure will also be used in setting out design of algorithms by organising it by infrastructure that requires visual inspection on the most frequent basis.

### 4.1.1 Additional

- Target Speed of Acquisition agreed at 30 Kph
    - After initial testing of a system, it was agreed that an acceptable target acquisition speed should be set at 30 Kph.

## 4.2 ADDITIONAL SPECIFICATIONS TO IMPROVE DESIGN FEATURES

- Easily upgradable
    - A pre-requisite for many real world designs that utilise modern technology is to allow for developments in hardware to be easily integrated into the system. Camera speed and resolution is constantly improving, allowing both image quality and image acquisition speed to be increased as the technology advances. This upgradability also applies to the lighting design with LED technology improving, and rapidly replacing halogen light sources. It would imperative that advances in this hardware are utilised along with the camera upgrades.
    - The computer architecture will also change to cope with the increase in data from the cameras, therefore it is also required that the acquisition recording and hardware is modular and upgradable meaning an integrated processor unit is less desirable than standard PC components.

- Adaptable
    - With the computer architecture, camera and lighting likely to change as technology advances, it is useful if the mounting units they are attached to are adaptable. They should allow for additional lighting, cameras, and acquisition hardware to be attached, exclusive of the hardware already attached, thus preventing the requirement for new mounting hardware to be made for each unit. This does not apply to additional mounting fittings but just to the main base plate that will interface with the bogie.

- Camera Angle
    - For effective image analysis algorithm design and repeatable performance, a fixed camera angle must be used. After careful examination of the environment and additional factors, a fixed camera angle has been decided at 90˚, perpendicular to the head of the rail. This will provided an image of the head of the rail from directly above. While this will limit the field of view, it has several advantages. It is easier to set up and change between vehicles. Calibration is not as complex

since fixing the camera at a specific angle to the head of the rail is not a requirement. This angle does not have to be calibrated with the lighting as well. It also carries significant advantage to lighting design as all items to be inspected fall within a specific length from the camera, allowing for a simpler lighting design. It also results in a less intensive light source being required.

- With the camera fixed to the bogie, a large amount of movement is removed from the infrastructure in the imagery. The head of the rail itself will swing less within the image making its location to be easily detected. The angle also removes a majority of vertical vibration from the camera as it only changes focus slightly, this will not affect imagery at such a small focal length.

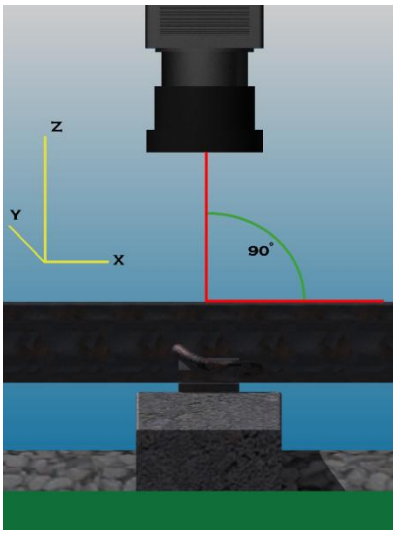



Figure 4.3 CAD model of specified camera angle in relationship to the rail.

- Vibration resistant
  - While the cameras angle removes a majority of the vertical vibration along the Z-axis, there are still vibration issues that can occur along the X and Y-axis. There should be relative design features that can counteract this. Such as the ability to add vibration dampening materials/connections and rigid support bars to the main mounting interface. Inevitably, there will be vibration that cannot be filtered out

through hardware adaptation and therefore computational algorithms should be examined.

- Waterproof
    - While not a pre-requisite for initial testing of designs, the ability to waterproof the hardware is essential. The hardware is expected to run at various points throughout the year and while more extreme weather conditions are not require to be catered for, the minimum requirement is a waterproof ability of the final design for full deployment. This includes all acquisition hardware and relevant cabling and mounting points.

## 4.3 LABORATORY PROVEN IMAGE ACQUISITION SYSTEM

### 4.3.1 Camera Choice

The initial camera chosen for testing was a Dalsa Piranha Colour 4K line scan camera (PC-30-04K80). This provides a colour image for testing and a method of interfacing the acquisition speed with the speed of the tram by controlling the line rate.

A test rig was constructed for evaluating the camera and lighting system. This assisted with calculations of the field of view, development of analysis algorithms, suitability testing for the type of camera and mounting specifications.

Whilst suitable for feasibility tests and development work, this particular camera is incapable of providing the acquisition speeds required for routine tram mounted inspections (e.g. at > 30 kph) as laid out in the amendment to the specification. Potentially a colour image gives more options for image processing algorithms, however there are currently no colour cameras with a high enough acquisition speed to support the project needs. Therefore a greyscale option was decided upon, the Dalsa Piranha 2 Monochrome (P2-42-02K40).

This camera is a mid-range speed, monochrome alternative that will allow a higher acquisition speed and allow further proof of the systems capabilities at useful inspection speeds. Table 4.1 examines the possible acquisition speeds of three cameras a Dalsa ES or HS alternative with a 4K resolution and pixel size of 14 µm and a line rate of 110 kHz. A Basler Sprint series with a 4K or 2K resolution and pixel size of 10 µm and a line rate of 140 kHz. Dalsa Piranha 2 with a 2K resolution and pixel size of 10 µm and a line rate of 68 kHz. This is the current camera to be employed as it has a shorter lead-time.

| Feature \ Camera | Dalsa HS/ES | Balser Sprint | Dalsa Piranha 2 2K |
|---|---|---|---|
| CCD Horizontal ($i_h$) | 57.344mm | 40.960mm | 20.48mm |
| CCD Vertical ($i_v$) | 0.014mm | 0.01mm | 0.01 |
| CCD Diagonal (i) | ≈ 57.344mm | ≈ 40.96 mm | ≈ 20.48 mm |
| 28mm Lens | | | |
| FOV width ($O_v$) | $716.8\,mm$ | 512 mm | 256 mm |
| FOV height ($O_h$) | $0.175mm$ | 0.125 mm | 0.0.125 mm |
| Speed | $69.3Km/h$ | $63Km/h$ | $30.6Km/h$ |
| 18mm Lens | | | |
| FOV width ($O_v$) | $1115\,mm$ | 796.4 mm | 398.2 mm |
| FOV height ($O_h$) | $0.2722mm$ | 0.19 mm | 0.19 mm |
| Speed | $107.79km/h$ | $95.76km/h$ | $47.6\,km/h$ |

Table 4.1 Chart comparing the capable acquisition speed of the selected mid-range camera and alternative high-end cameras. The chart shows how the mid range Dalsa Piranha 2 2K performs against the faster Dalsa and Basler line scan cameras, showing that its selection meets the acquisition speed laid out in the specification for the MVS.

### 4.3.2 Low Speed Trolley and Laboratory Testing

To allow testing and design of a plausible system, a simple lab based trolley was designed. The design was easily changeable to allow different camera models to be fully tested. It allowed the camera to be mounted up to 600 mm away from a rail section being analysed. The resultant FOV from the trolley was representative of that which could be obtained from live acquisition from the tram bogie. Investigation of the camera angle was also achievable and conformation of the perpendicular FOV to the head of the rail was ensured to be the most productive for image analysis purposes.

The simple adaptability of the design allowed different lighting types and positions to be investigated. Appropriate lens selection and associated investigation was also instigated to show that the overall quality of lens employed had small effect on the quality of image acquired. The focal length and focal ratio of available lenses was investigated. This confirmed that while lens selection was overall controlled be appropriate focal length, a smaller focal ratio was required. While a larger focal ratio improved DOF, the required light intensity to achieve CCD saturation was difficult to achieve.

Figure 4.4 CAD model and photograph of the produced low speed trolley for lab based image acquisition of rail.

Using the designed testing rig (Figure 4.4), calculations were confirmed for the FOV. The lab testing systems also allowed completion of software for acquisition that tied in the vehicles speed to the acquisition rate of the camera and GPS positioning data. Testing allowed the required hardware for the control of the lighting to be independent of the computer acquisition. This allowed for design adjustments and for control refinement for high image quality.

Figure 4.5 Photographs of light testing on low speed trolley for lab based image acquisition of rail.

## 4.4 DESIGN OF BESPOKE HIGH INTENSITY LED LINE LIGHT

As discussed in Section 3.2.5 (Lighting of MVS Hardware Selection) the lighting employed in a MVS system is essential to its overall success, and the ability of image analysis techniques employed. LED technology is utilised to produce a lighting system that provides better stability for the type of environment the MVS will be deployed in. Standard halogen lighting is too fragile for the system and can easily malfunction and break due to the vibration of the tram bogie.

LED's are not as fragile, but do require an efficient cooling system to keep them within operable limits. To achieve this water-cooling must be used. Standard PC water cooling components that are similar to those used in air compression systems are ideal for this purpose. Their small form factors and resistance to corrosion is ideal.

Current LED line light designs in production have an operating focal length of 50 – 100 mm. While this is acceptable in a majority of machine vision applications, as the environment is controlled, it will not be acceptable for this application. To overcome this bespoke LED line lights were design with optical calculations that allow intense collimated beams of light to be projected at a focal length of 500 mm. This is the maximum operating distance applicable.

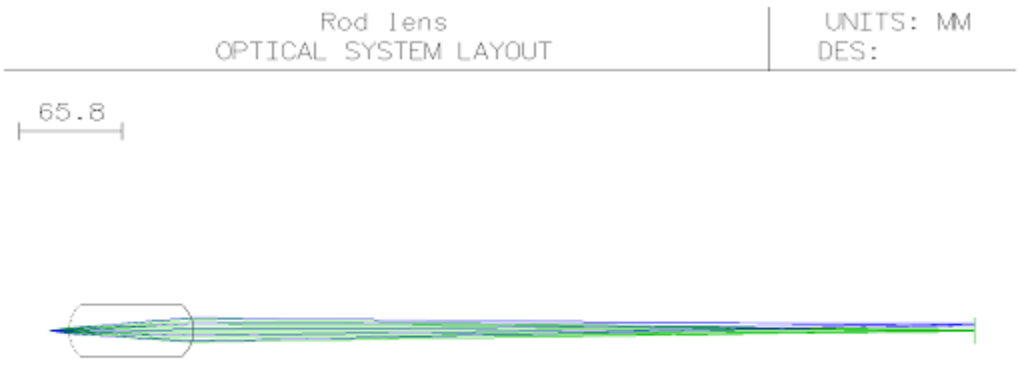



Figure 4.6 Optical designs of rod lens employed in line light design. Focal distance 500 mm.

Due to the reflective nature of the running band of the railhead and the type of camera selected, a bespoke LED line light was designed with controllable intensity. While such lights are available on the market, a novel approach was taken by dividing the light into individual blocks. Each of these blocks are independently controllable allowing a non-uniform lighting pattern to be used. This allows the light being directed towards the highly reflective surface of the railhead to be controlled and allows a greater amount to be applied to image area of a more light absorbent nature. The result is an image with uniform lighting and controlled contrast ranges i.e. no bright or dark spots within the image. Figure 4.7 demonstrates how the light intensity will mirror the reflectivity of the object within the FOV.

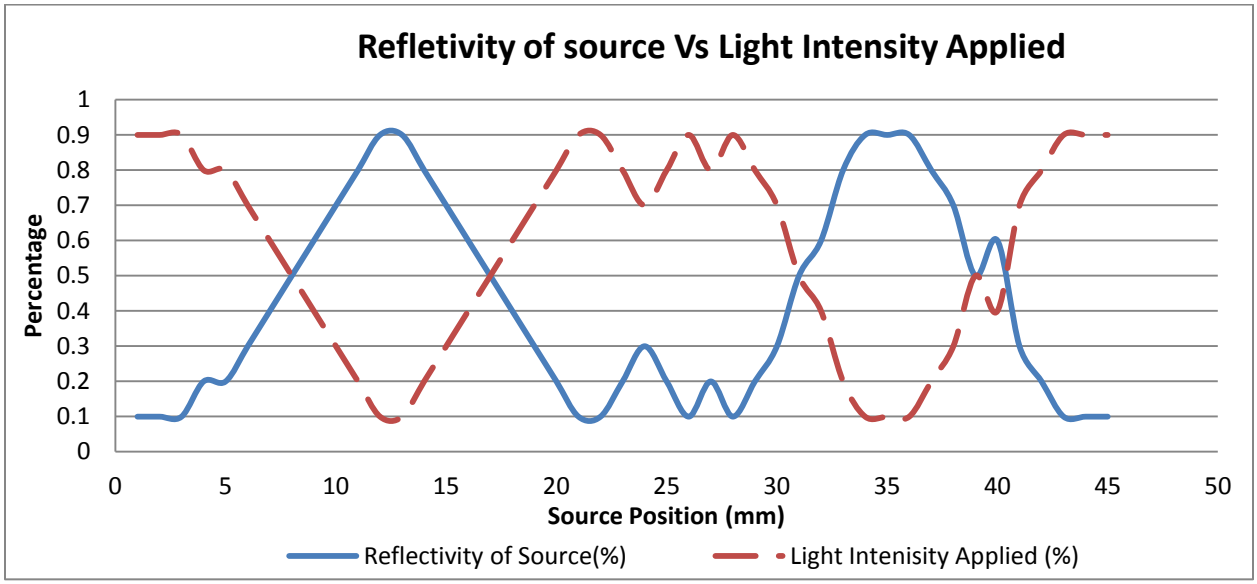



Figure 4.7 Plot to show the Percentage of power intensity applied to a source verses its reflectivity of the object within the scene.

Voltage control is a common method of dimming light sources; this is done by switching on and off the voltage supplied to the light source at speed. Varying the duty cycle i.e. the time the voltage is on increases the average voltage seen by the light source and makes it appear brighter. However, due to the high rate in which a line scan cameras acquire their data, voltage controlled lighting is not acceptable. The high rate of image acquisition would allow the camera to see when the light is switched on and off making the image appear brighter and darker between frames. Therefore, each block of LED lighting must have a constant controllable current source.

This allows the LEDs to remain on as an uninterrupted voltage is supplied but by starving the light of current, the brightness of the light is reduced. The constant current source is an adapted version of a Microchip design (2007).

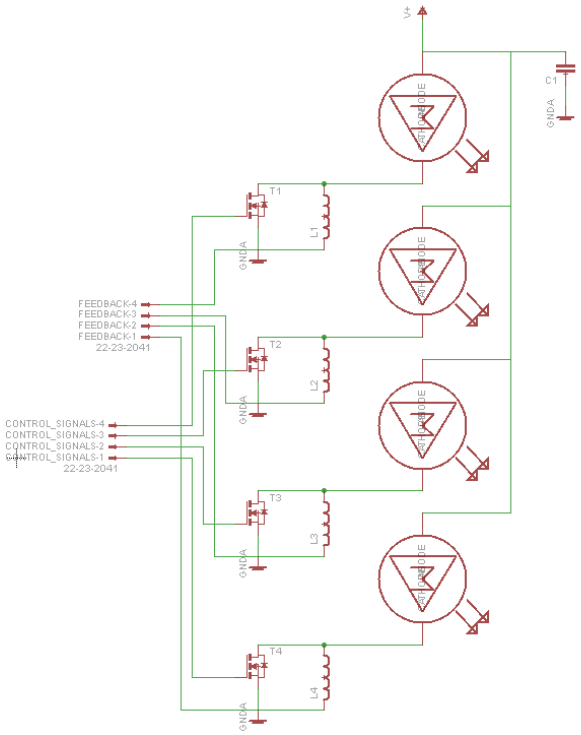



Figure 4.8 Circuit design for a constant current control of high power LED's with feedback. Based on Microchip design (S. Bowling - MICROCHIP, 2007).

### 4.4.1 LED Line Light Version 1: Single

The initial design was based on current line light designs. A long light capable of illuminating the whole FOV. Due to the limited space available, the size of the light was limited and therefore illuminating the full FOV was not achievable.

Line light design ‘Version 1’ uses high intensity LED lights and an Acrylic lens that has a focal lens of 500 mm. It is designed using nine individual blocks of lights that have their individual intensities controlled. This designed mounted the required MOSFET’s for the current control directly on the light. This allows heat dissipation to be managed with the cooling of the light. Feedback circuitry was not implemented in the design however it was adaptable to allow such circuitry to be implemented. Control signals were sent to each block using individual cables. This design was not produced to be waterproof and was an initial testing design for the lighting control, intensity, and focal length of the designed lens.

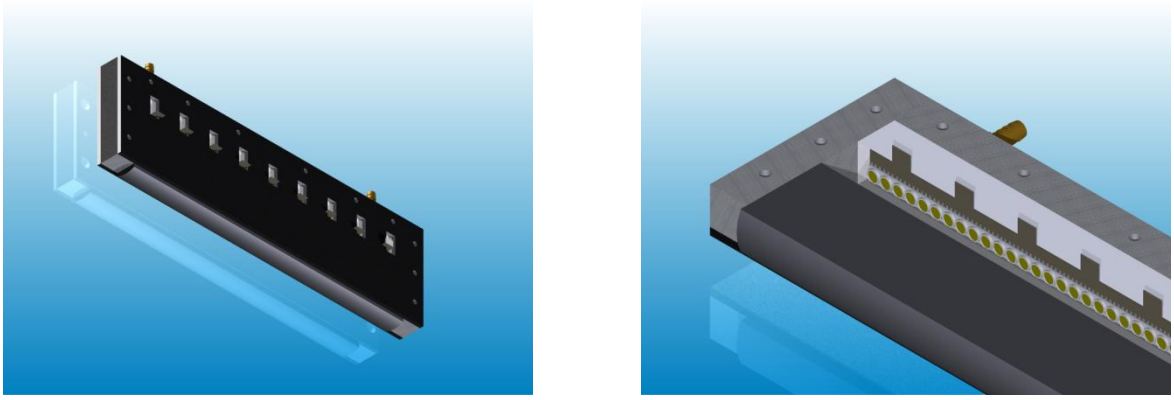


Figure 4.9 CAD model of the LED line light version 1.

The design developed demonstrated the lens calculations were accurate and showed that the required intensity of light could be provided by LED technology at the required distance. The design also showed that control of the light intensity could be achieved however, a re-design of this circuitry was necessary to waterproof the unit. The designed water-cooling was capable of providing adequate heat dissipation allowing the light to be used for long periods.

In lab testing and deployment, it was clear that the intensity of the light upon the foot of the rail was not sufficient. This was a result from the mounting position of the light. As the light was mounted in front of the camera, it was angled toward the FOV at the head of the rail. This resulted in a focused light at the head of the rail; however, the lighting angle resulted in the focus beam missing the FOV at the rail foot Figure 4.10 shows this.

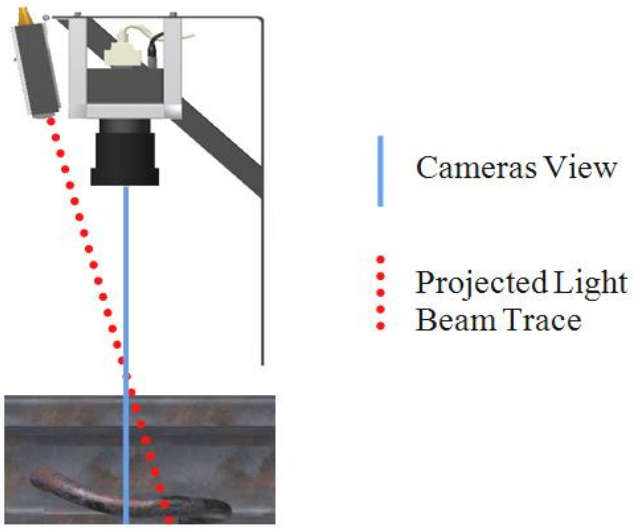



Figure 4.10 CAD model showing the camera view and LED line light projection problem.

This is a rare occurrence within MVS; it is caused by the DOF and the need to inspect objects on the head of the rail and its foot. A solution within a controlled environment would be to use two lights either side of the camera with lenses that would allow adequate distribution to the centre of the FOV. The space available under the tram prevents this as a solution.

### 4.4.2 LED Line Light Version 2: Tri-linear

A unique approach has been taken to the solution of the problem encountered in the first design, where the foot of the rail was outside the focal beam of the light. The image has been divided into 3 independent areas to be lit by individual line lights. The railhead and each side of the foot will have an independent light source allowing for better control of the lighting conditions within the image. It allows correct focus points of the light for each area of the FOV.

The second-generation design was developed to allow equal illumination of the railhead and foot. It was designed to be splash proof with all electrical components held securely inside the aluminium and steel chassis. The intensity of the high brightness LED's is individually controlled with appropriate logic signals.

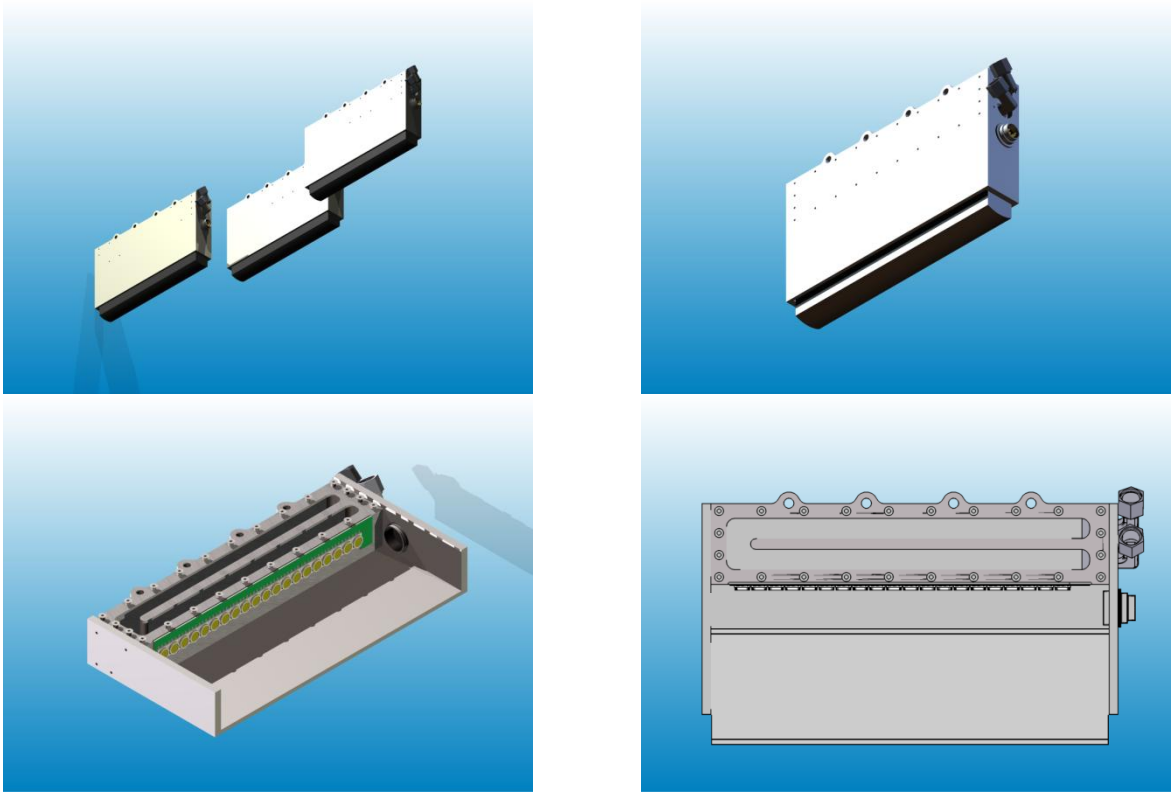


Figure 4.11 CAD models of the LED line light version 2 'tri-linear' design.

The current control circuitry was situated apart from the lights and the required power was transmitted to the lights through an external interface. The power drain is controlled by the external circuitry. Each light has four individual blocks of LEDs, the intensity of each block is controlled independently. The lens employed is of the same acrylic design as 'Version 1' of the line light. The design is mechanically easy to produce. While control and design are of a substantial improvement to the initial design, further development needs to be made to produce a watertight unit.

## 4.5 VEHICLE MOUNTING DESIGN

The mounting hardware was designed in association with Stagecoach Supertram to ensure that it reached required safety specifications. In the initial testing stage, an L-shaped bracket was decided upon. This allows for adaptability and upgrades to be easily mounted. A supporting bracket was later added to prevent vibration. The run was completed in April 2010 that provided excellent images using line light ‘Version 1’ included in Appendix 3.

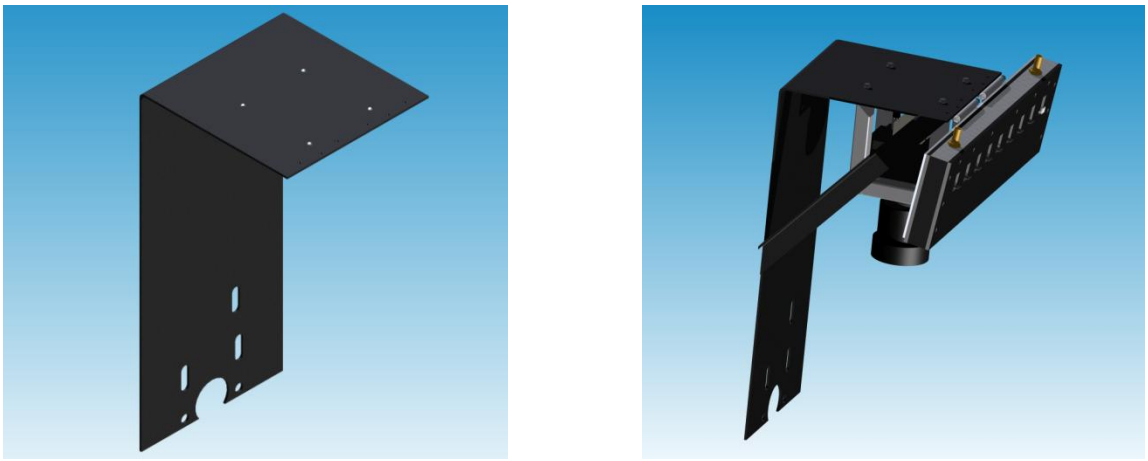


Figure 4.12 CAD model of the 'L' plate mounting unit and full view of additional fixings and supports.

Figure 4.12 shows the CAD design of the mounting plate and Figure 4.13 shows the device in deployment underneath one of Stagecoach Supertrams’ vehicle. In this stage of testing the additional acquisition, hardware was deployed inside the tram. This allowed closer monitoring of the acquisition in real time. Deployed under the tram was the line scan camera and LED lighting system.

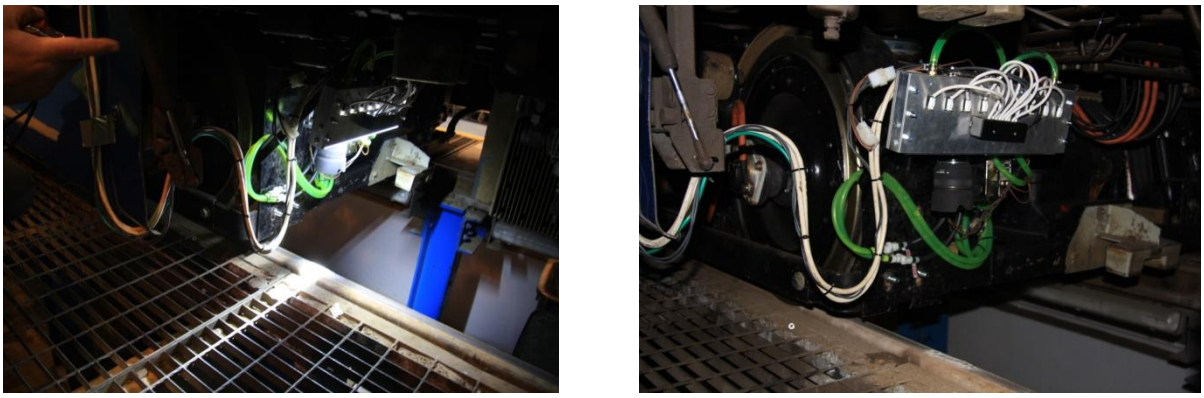


Figure 4.13 Photographs showing the deployment of the vehicle mount design for testing in April 2010.

With testing of the initial hardware, a proposed design to move all associated hardware underneath the tram and attach it to the tram bogie was developed. This stage of development extended the initial design further to implement inspection of each side of the rail simultaneously. This required the development of self-contained acquisition computers.

The design of the acquisition computers is based on standard PC architecture, as this is more appropriate for the type of frame grabbers being used. They each have an individual PCI-Express slot for the frame grabber. Large quantity of USB ports for extra peripherals such as the GPS devices and external interface, the possibility of on-board graphics (processor dependant) and 8 GB of RAM. The operating systems are stored on solid-state hard drives to cope with vibration and acquisition images are stored on 3.5” SATA hard drive(s).

The amount of acquisition images that can be stored is dependant of the number of hard drives installed. Factors such as image type and compression also play a large role in the amount of data that can be stored. A 2048 x 480 pixel image in JPEG (Joint Photographic Experts Group) lossless format takes an average 1.10 MB of physical disk space. This results in 867’153 JPEG images per Terabyte that can be stored resulting in 91.57 Kilometres of track. A 2048*480 pixel image in PNG (Portable Network Graphics) format takes an average of 2.82 MB of physical disk space. This results in 338’251 PNG images per Terabyte that can be stored resulting in 35.72 Kilometres of track.

The computers are stored in waterproof housing with adequate size and heat dissipation to meet the processing requirements. There are two external sealable USB ports for peripherals such as keyboard, mouse, and a USB to VGA adapter. No direct video interface was supplied as this threatened the waterproof integrity of the case.

The computers have an external interface for control of the lights, via USB. This allowed them to be altered from the PC hardware and maintain their values independently. They were not directly integrated into the PC unit due to the large current they were dealing with and a different heat dissipation requirement.

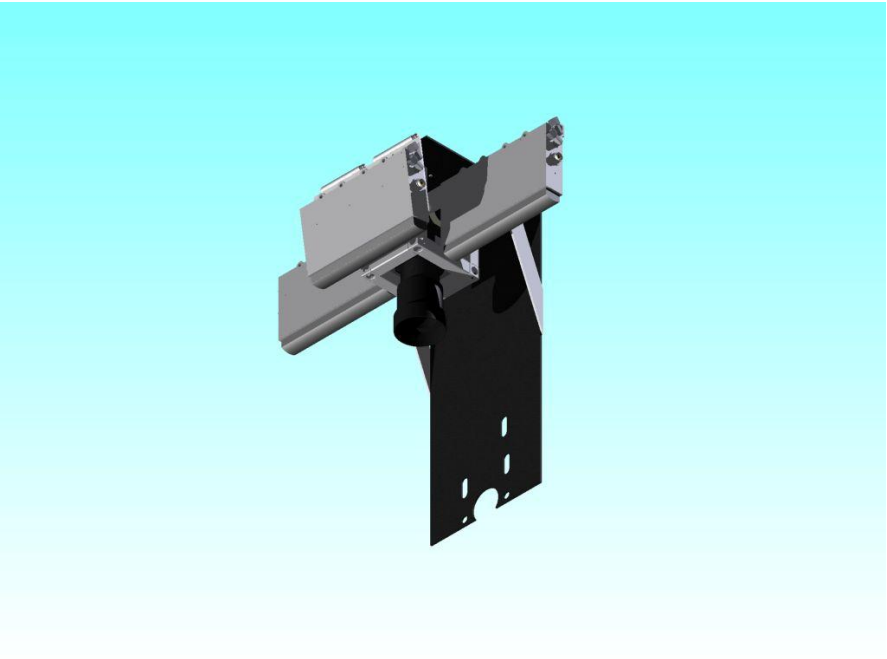


Figure 4.14 CAD modelling of 'L-plate' mounting with tri-linear LED design implemented.

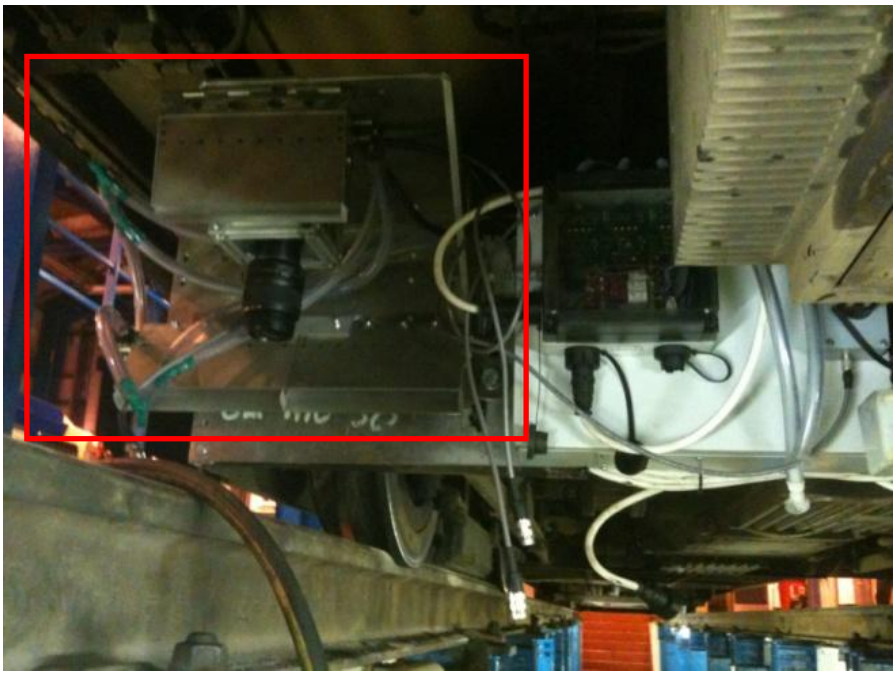


Figure 4.15 Photograph of the Tri-linear LED design implemented and mounted on Sheffield Supertram vehicle.

Figure 4.16 Photograph of the Tri-linear LED design implemented and mounted on Warsaw Tramway (Tramwaje Warszaws) vehicle.

## 4.6 FURTHER WORK: FUTURE DESIGN OF THE MVS SYSTEM FOR TRAMWAY RAIL INSPECTION

### 4.6.1 Industrialisation of the Image Acquisition unit

Ruggedisation of the image acquisition system is still required. The deliverables required to achieve ruggedisation of an industrialised system are:

- Mounting of the camera within a self-contained waterproof unit, a glass screen using a non-stick ClearShield$^{TM}$ or LiquiGlide$^{TM}$ screen to prevent water or dirt from affecting image acquisition ability. In addition or alternatively the use of a liquid repellent nano-coating material could be considered.
- Active dampening suspension to aid in removing horizontal vibration generally observed in sections of corrugated track.
- Redesign of the lighting system in an attempt to remove water-cooling completely and use new developments of LED technology. Alternatively, to remove cooling of each unit individually through copper pipe transmission to a radiator and single cooling based area.
- Reduction in the size of the computer acquisition hardware into an adapted chassis allowing better use of space for storage.
- A caddy based system for hard drive storage, allowing quick removal of data and replacement of fresh storage medium.
- Self-calibrating hardware that will inform correct calibration of lens focal lengths.
- Self-adjustment of lighting through live image analysis.
- Investigation into the ability to share relevant information from active measurement systems such as speed and GPS data using a wireless based connectivity system.

### 4.6.2 Upgrading of GPS positional hardware

A recommended improvement for combining of data from the image acquisition system to additional sensors is the increased accuracy of GPS data. Access to military grade GPS hardware

can make use of an extra transmission signal to correct for ionosphere signal degradation; this can increase accuracy to within 0.5 meters. This is beneficial, however securing accurate time basses from individual GPS data units is essential to relate the readings from the images acquisition runs completed on different dates. Alternative methods of tying in accelerometer readings with the GPS data can allow similar accuracies to that of military grade GPS.

# CHAPTER 5 DEFECTS AND INFRASTRUCTURE TO BE DETECTED AND ANALYSED

## 5.1 INTRODUCTION

Tramways and railways alike have a large amount of infrastructure associated with them. The entire infrastructure requires regular attended inspection and associated maintenance. While ideally individual hardware systems could be employed to eliminate the requirement for attended inspection, the cost and the greater complexity of their implementation and associated networking often prevents this. A system such as a MVS has a greater appeal as it can be employed to inspect a larger amount of infrastructure than individually designed hardware.

The following chapter discusses the infrastructure that could be potentially inspected using the MVS described in Chapter 4. While there is potential for a variety of infrastructure to be inspected, the scope of the following thesis chapters are narrowed for appropriate research objectives to be met. A full catalogue of potential objects for MVS inspection was developed as part of the PM'n'Idea (Tata Steel (formerly Corus), 2010).

## 5.2 TRAMWAY AND RAIL INFRASTRUCTURE FOR POTENTIAL MVS INSPECTION

This thesis covers the application of a MVS mounted underneath the bogie of a tramway-based vehicle. Due to height limitations with mounting the camera, and associated hardware in such an inhospitable environment, the FOV provides a limitation to what infrastructure can be examined. Only those objects of infrastructure that are visible to the observable FOV (Table 4.1:- 'Owidth' 398.2 mm) will be examined for this work.

All images are provided through the PM'n'Idea report unless mentioned otherwise (Tata Steel (formerly Corus), 2010). The following railhead features and infrastructure items fall within the scope of the research; methods for their detection and classification are discussed within this thesis, together with algorithms for the detection of all of the following features.

- Rail edge classification and analysis
- Running band detection and deviation analysis
- Squat defects detection within the running band data
- Wheel burn analysis and differentiation from squats

### 5.2.1 Types of Rail on a Tram Network

There are two typical types of rail present on a tram network. There are numerous grades and formats for the types used; grades referring to the type of material; and properties of the material that the rail is made out of and the formats, being the dimensional changes and size of the rail features. The two types are Flat bottom rail (Figure 5.1) and Grooved/Embedded rail (Figure 5.2). The more traditional Flat bottom rail is the one seen on cross-country networks and is the one most familiar. Grooved rail is used in order to make street running feasible, allowing a polymer to edge the rail and concrete or similar solid structure to be poured to the level of the railhead.


Figure 5.1 Graphical Model and photograph of Flat Bottom Rail.

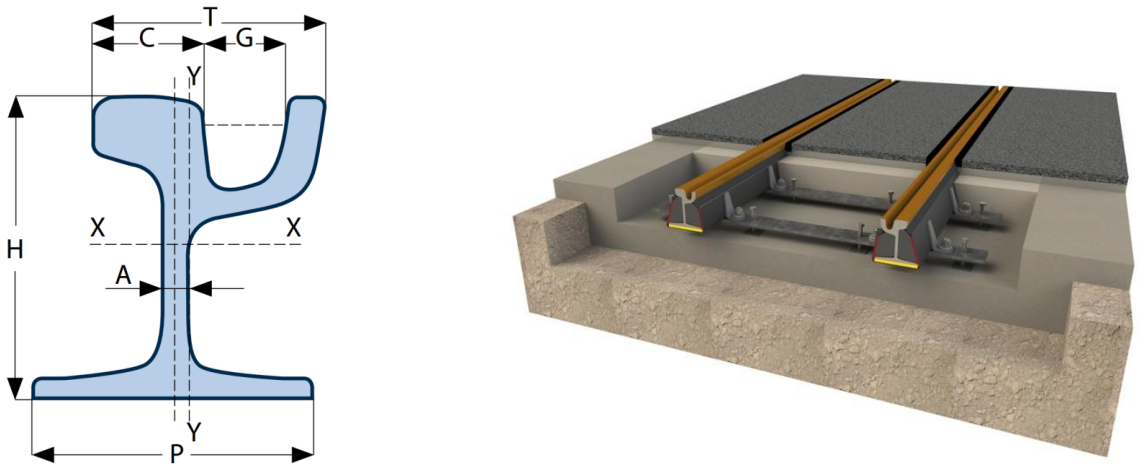



Figure 5.2 Graphical Model and CAD model of Grooved Rail.

## 5.2.2 Rail Joints

### 5.2.2.1 Standard/Heavy Duty Fishplated Rail Joints

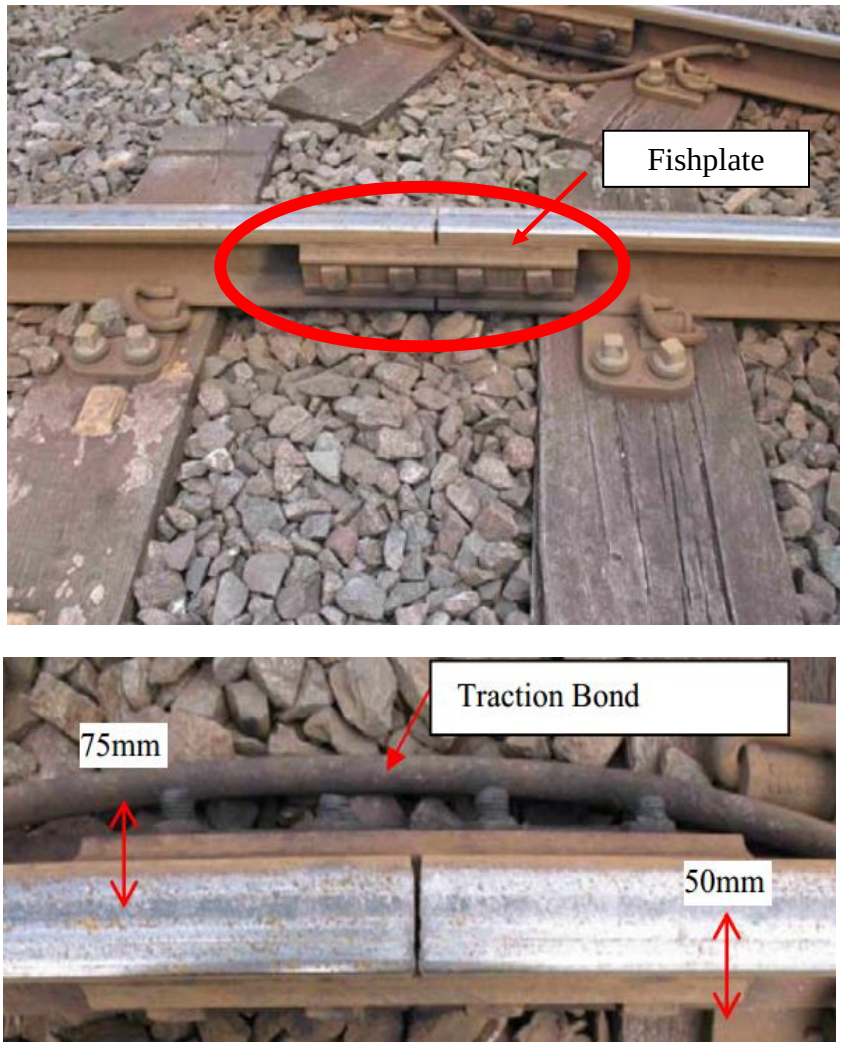



Figure 5.3 Photographs of a fishplate joint circled in red (above) with approximate dimensions and in addition the traction bond sharing the earthing point across the rails (bellow). A fishplate allows to sections of track to be joined.

Fishplates are used for joining two rail sections together. Their occurrence throughout the tramway system is sporadic. Their key features are the square shaped top view that interferes with the foot of the rail. Defects are noted as either a degradation e.g. splitting of the main plate or a loosening / loss of bolts. Loosening of bolts is a minor level defect that can lead to more severe defects. A missing bolt is a severe defect as is cracking or splitting of the main plate. Often a Traction Bond is associated with these fishplated joints allowing rails to share an earthing point. Detecting loose bolts is potentially very difficult.

#### 5.2.2.2 Insulated Rail Joints

Figure 5.4 A photograph of pre-prepared insulated fishplate joint. An insulated fishplate joint allows two sections of track to be joined but remain electrically isolated.

Figure 5.5 A photograph of a production sample insulated joint. An insulated joint allows two sections of track to be joined but remain electrically isolated.

Insulated fishplate joints contain the same features as discussed with standard fishplates. The key differences are the size of the fishplate and the number and the types of bolts used. Many of the same image analysis techniques will be required with the potential complexity of evaluating the insulating material for degradation. The requirement may be surplus to standard maintenance needs since fishplates require greasing annually.

### 5.2.2.3 Aluminothermic (AT) / Flash Butt (FB) Welded Rail

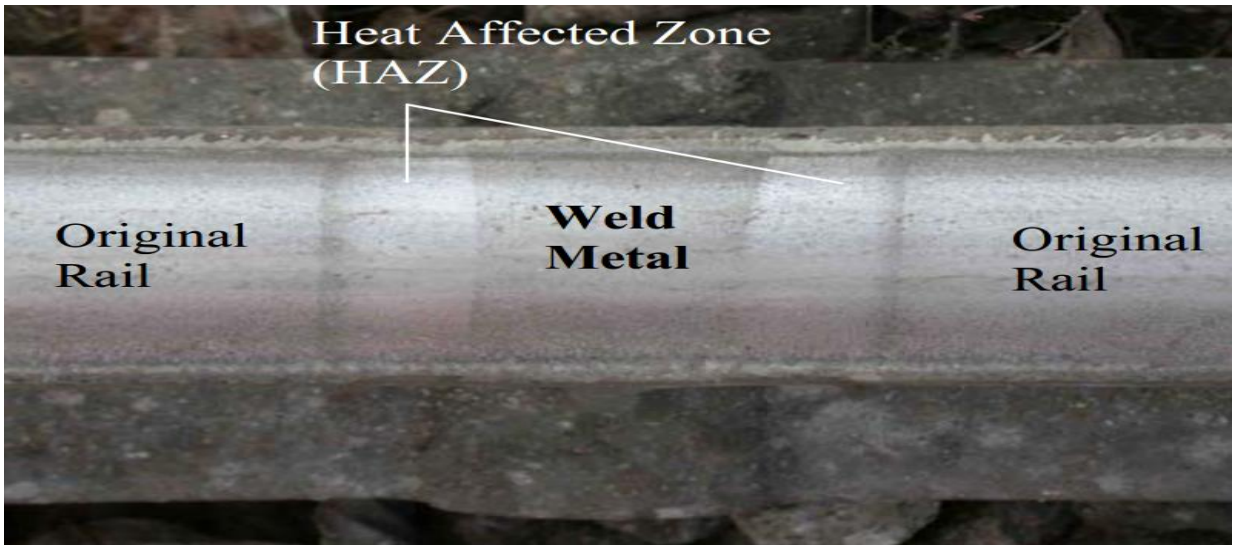



Figure 5.6 A photograph of Aluminothermic (AT) welded joint in situ. An aluminothermic weld joins two section of rail without the need for additional fixings.

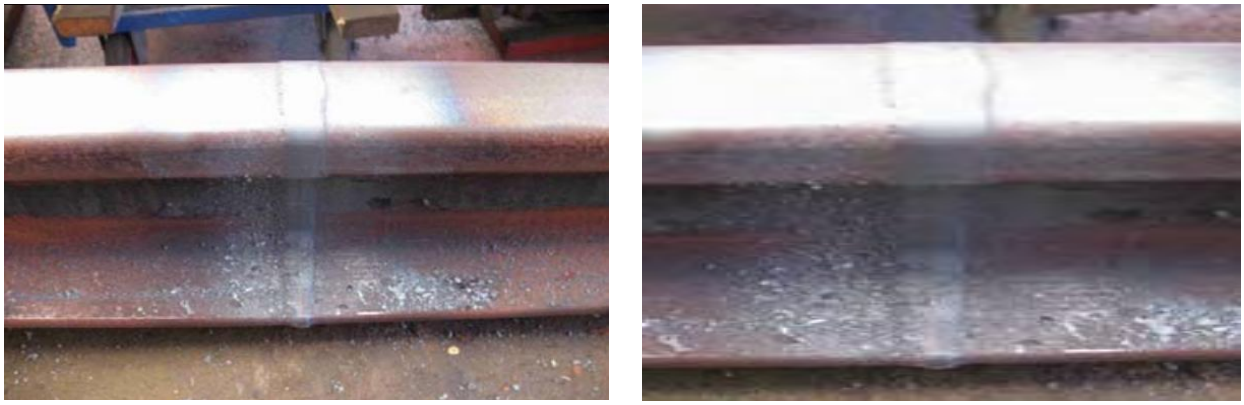


Figure 5.7 A photograph of Flash Butt (FB) welded joint in situ. An flash butt weld joins two section of rail without the need for additional fixings.

AT and FB are two methods used to join rail ends together, the advantage being a stronger, defect resistant join, better than the use of fishplates. The biggest problem with this type of joint is incorrect welding procedure that can lead to failure. AT welding, uses a filler material and provides a wider affected area that can have a different texture than that of the normal running band. FB welding does not require a filler and results in a much smaller affected area. Both methods produce a small deformation on the foot of the rail.

The occurrence of welded joints, like fishplates is sporadic, and with the methods involving key infrastructure i.e. rail itself; any associated features will fall within the designated FOV. A change in texture when compared to the railhead can allow classification however; changes are subtle and make accurate classification difficult. Incorrect smoothing of the head of the rail can cause dipping or cupping leading to potential failure and defects such as corrugation. A large data set of images would be required before attempting to classify welded joints. Evaluating dimensions for dipping or cupping would be difficult for the MVS and a depth imaging system would be more practical.

#### 5.2.2.4 Linear / Lateral Expansion Joints

Figure 5.8 A photograph of a rail expansion joint (breather or adjustment switch) (left) and a lateral expansion joints (right) in situ. Rail expansion joints allow track to expand and contract in a linear direction. Lateral expansion joints allow track to expand and contract in a lateral direction to the track.

The method of joining rail using a rail expansion joint or a lateral expansion joint is complex and its occurrence within a network is less than that of welded or fishplated joints. It allows expansion and contraction of rail due to weather. Linear expansion joints (pictured left) have infrastructure that falls within the expected FOV and would allow analysis for any degradation or defects. Lateral expansion joints however (pictured right), have a much larger amount of infrastructure associated with them, some of which cannot be seen within the set FOV. Expansion joints maintain a smooth running surface for the vehicle wheel to travel.

The associated bolts and clips can potentially aid recognition of both types of expansion joints. In the case of linear expansion joints, locating these can instigate custom algorithms for inspecting the head of the rail for “batter” of the rail ends. Certain infrastructure associated with lateral expansion joints, such as bolts and clips, can be evaluated for degradation or defects. Associated infrastructure, such as Tie bars, cover the expanse of the rail-to-rail track bed. This leaves a majority of this infrastructure outside the FOV for analysis although, the parts of the Tie bars that interact with the rail could be evaluated.

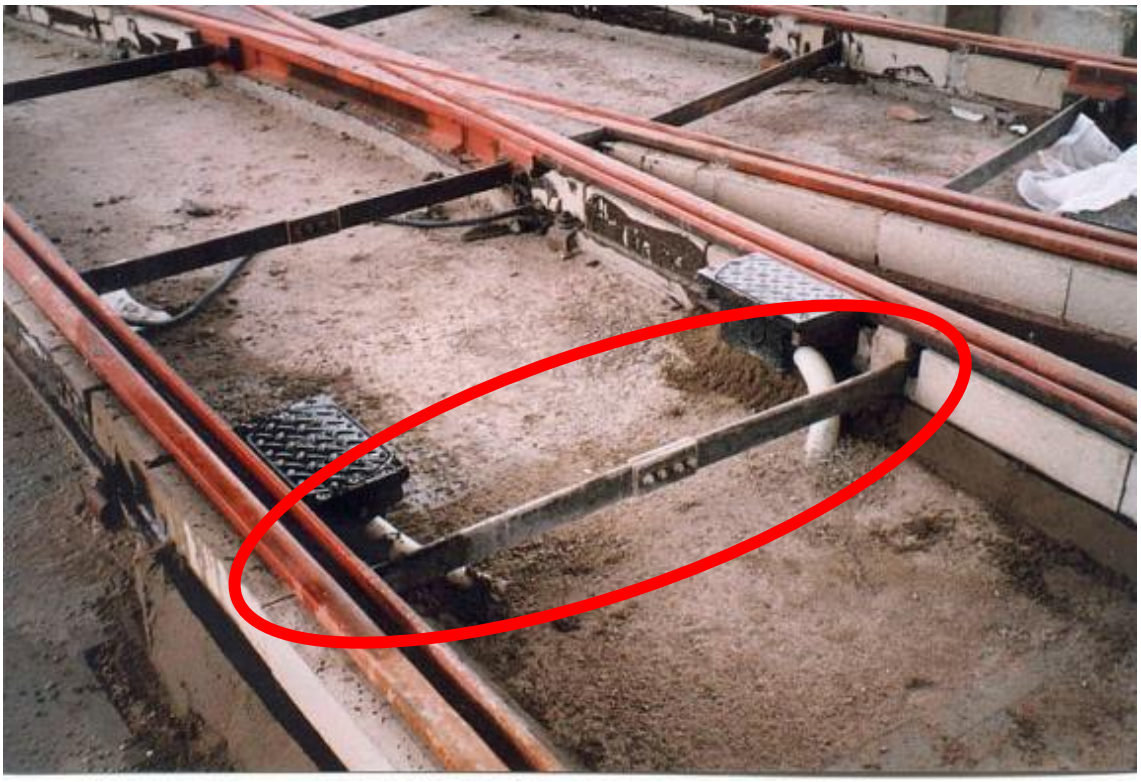


Figure 5.9 A photograph of Tie bars with no ballast present. Tie bars maintain the distance between rails.

Further research would be required to see what points of the lateral expansion joints fail and if the parts of the infrastructure visible would assist in the maintenance of the equipment.

### 5.2.2.5 Transition Rail

Figure 5.10 A photograph of transition rail in situ. This aids in the transition of wheelsets from flat bottom track to grooved track and vice versa.

Transitional rail occurs when Flat bottom rail joins Grooved rail. The features are simple and clear, falling with the FOV of the MVS. Accurately locating the transition point without producing errors could be difficult. By evaluating the post and proceeding infrastructure, custom algorithms could be run for a more accurate evaluation. Cracks can occur near the weld however, wear is similar to that of standard rail. Deformation of the angle between the head of the rail and flange should be evaluated.

### 5.2.3 Rail Fastenings

#### 5.2.3.1 Pandrol ‘e’ type clip Cast Base Plate on a Wood Sleeper

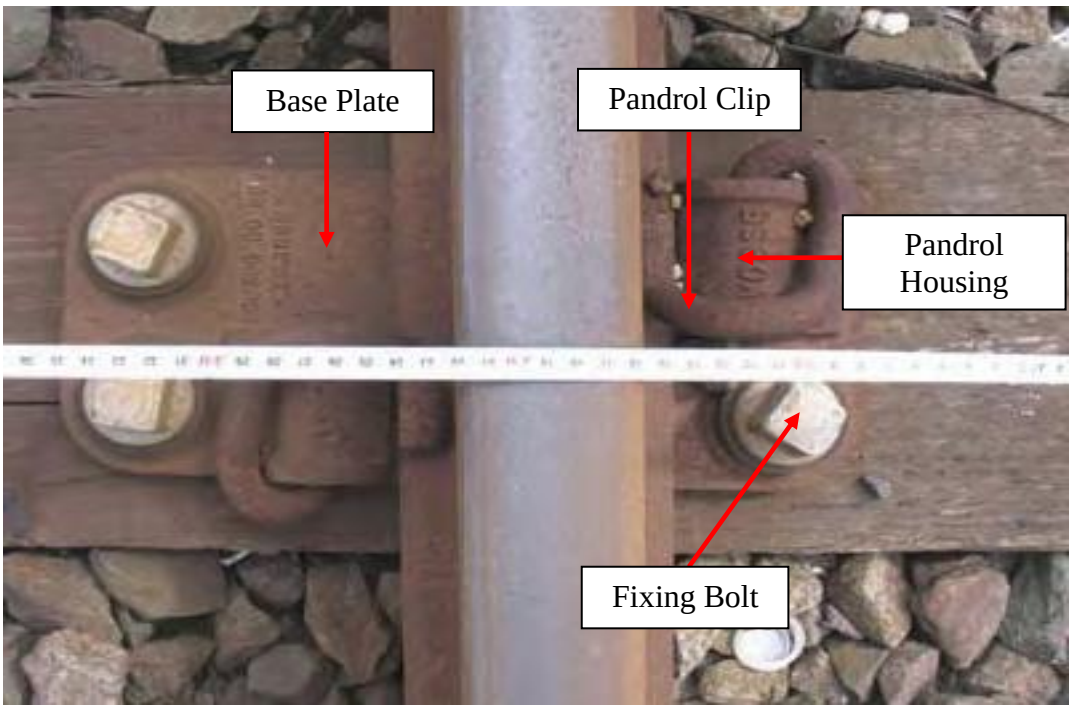



Figure 5.11 A photograph and measurement of a Pandrol 'e' type clip in situ with associated base plate and fixing bolts. A Pandrol clip secures rail in position.

As a method of securing the rail to wooden sleepers, the occurrence of Pandrol clips in particular ‘e’ clips is extremely common. While general corrosion can affect these types of clips, movement in the housing leading to complete loss of the clip can occur. It is unknown how the rate of movement relates to the frequency of traffic and its type. Loosening of bolts can also occur; however evaluating the movement of bolts could be complex, whereas simply inspecting their presence is a simpler task. There is the potential for fracture and breaking of the base plate that must be considered.

### 5.2.3.2 Pandrol Clip ‘e’ Type on Lug Base Plate with Check Rail

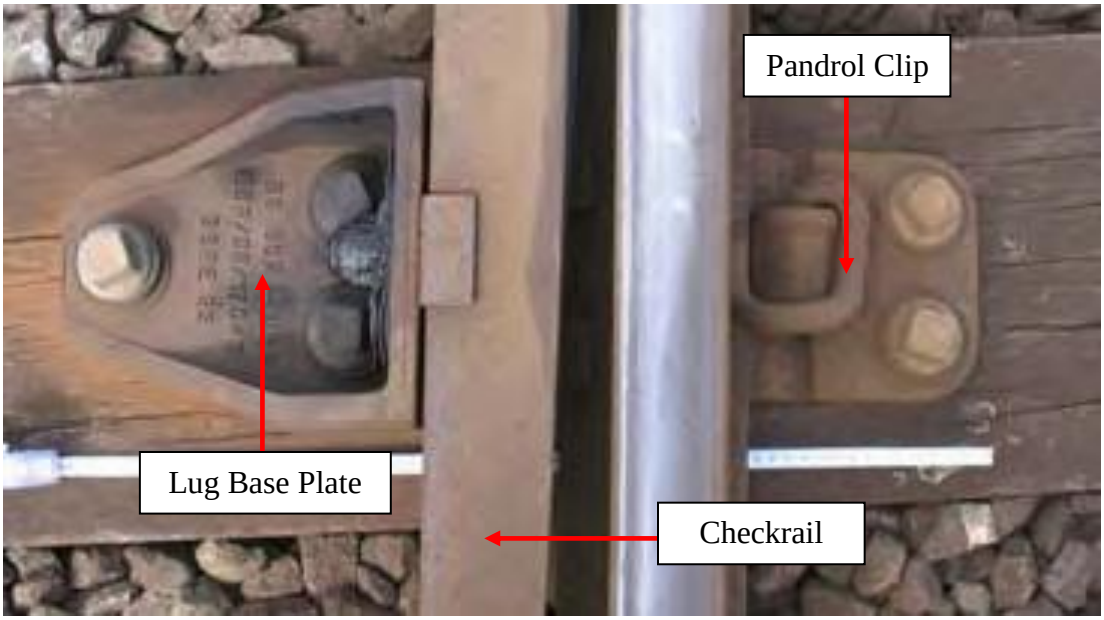



Figure 5.12 A photograph of a Pandrol 'e' type clip on Lug base plate in situ. A Pandrol clip secures rail in position while a Lug base plate provides mounting for a checkrail on wooden sleepers in ballasted track areas.

Check rail is a key part of infrastructure that prevents derailment on tight curves. The flange way clearance needs to be monitored; this is the distance between the head of the rail edge and the checkrail. The farthest bolt associated with the check rail fastening falls outside the FOV that may lead to potential difficulties in evaluating their condition. The inner fixing’s however fall within the FOV and can be evaluated once the check rail is located. The detection of loose bolts is desirable but potentially difficult.

The occurrence of check rail and the associated lug base plate and ‘e’ clip fixings is infrequent. Methods employed in detecting the rail and its edge cues can be employed in evaluating the checkrail as well. Check Rails are not always employed on both sides of the rail.

### 5.2.3.3 Pandrol ‘e’ clip with Twin Flange-planned Base Plate with Guardrail

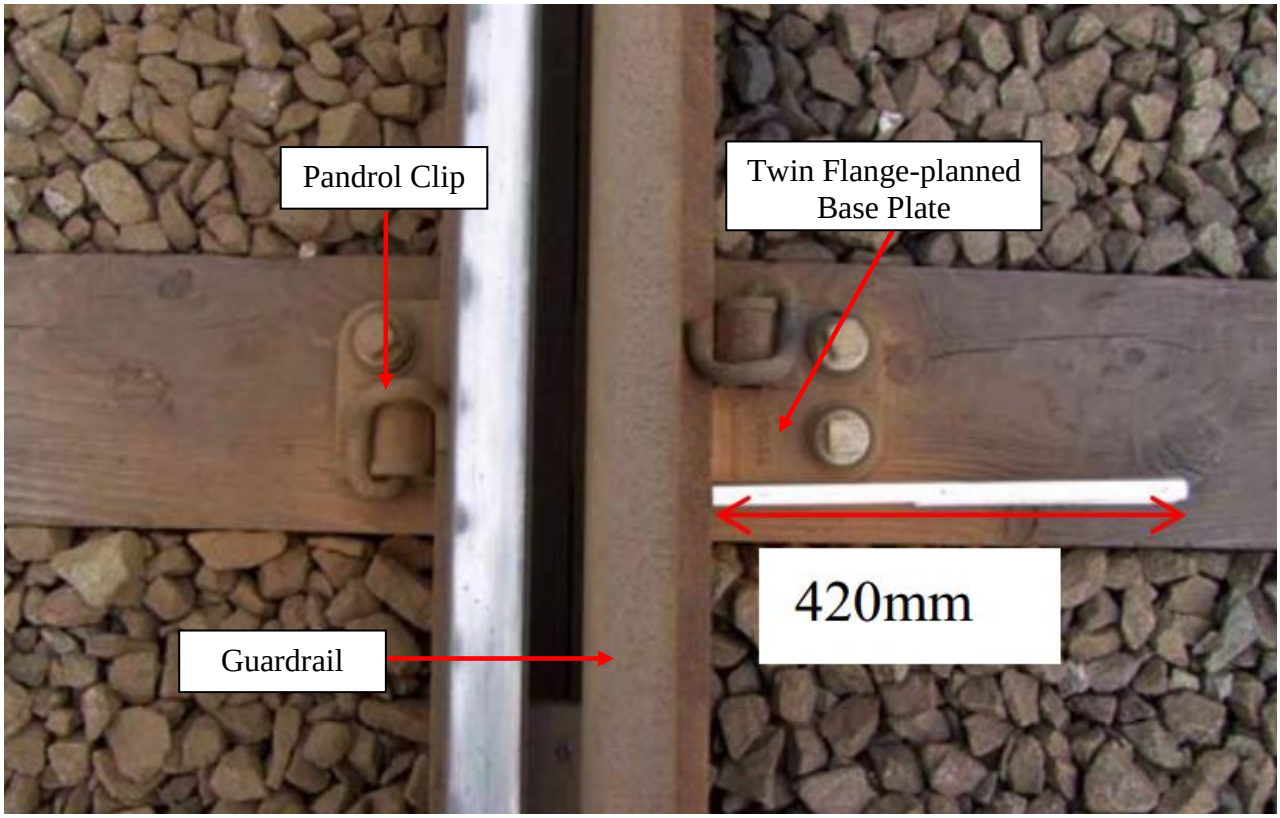



Figure 5.13 A photograph and measurement of a Pandrol 'e' type clip on twin flange-planned base plate in situ. A Pandrol clip secures rail in position while a flange-planned base plate provides mounting for rail and a guardrail on wooden sleepers in ballasted track areas.

.

The Pandrol ‘e’ clips associated with this infrastructure have the same pattern as seen with those on a cast base plate on a wood sleeper. The potential difficulty with this type of infrastructure is that the guardrail pushes the associated Pandrol clip base plate outside the FOV. This may cause the associated bolts not to be classified, making any evaluation of loosening impossible. The Pandrol clips can still be evaluated and fall within the accepted FOV. The bolt associated with the base plate on the rail side also falls within the FOV and can be evaluated. Unlike with check rail the flange way clearance does not require evaluation, as guardrail should not suffer from wear. Guardrails are fitted to both sides of the rail.

### 5.2.3.4 Twin Flange-planned Fabricated Steel Base plate

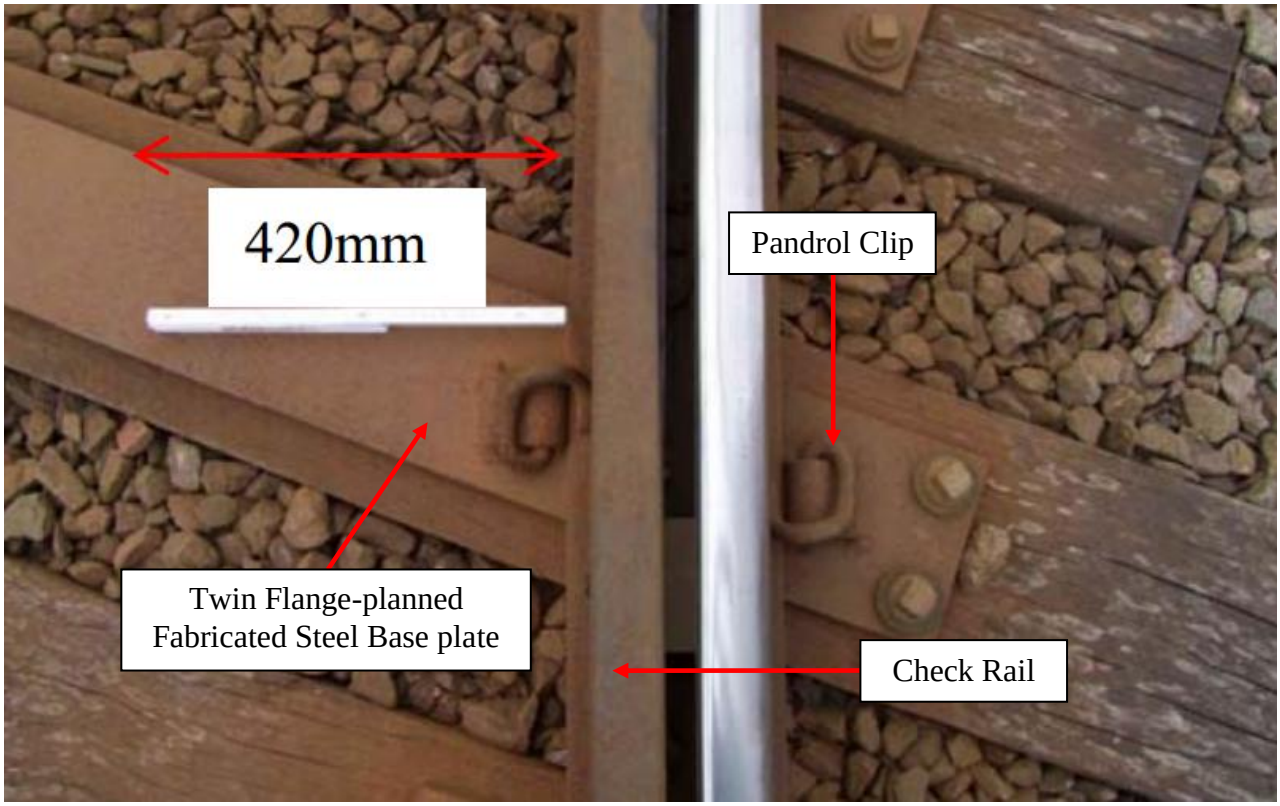



Figure 5.14 A photograph of a Twin flange-planned fabricated steel base plate in situ. This base plate allows secure mounting for both rail and a checkrail on wooden sleepers in ballasted track areas.

The advantages of a fabricated steel base plate is that the bolts that were potentially outside the FOV with a non fabricated base plate, are simply not there. The angular positioning of the bolts present on the rail side can be used for classification. Inspection for deterioration of the fabricated base plate can only be made for the part of the infrastructure that falls within the FOV. A key issue, likely to affect such analysis, is ballast interfering with the base plate edges.

Evaluation of this infrastructure is similar to that of a standard base plate. The occurrence of this feature is irregular and algorithms to analyse the fabricated base plate would rely on the detection of the guard rail, Pandrol clips, and the angle between the bolt installations.

### 5.2.3.5 Pandrol ‘e’ Clip with Cast Base Plate, Fixed by Spring Loaded Bolt Fastenings on a Concrete Slab

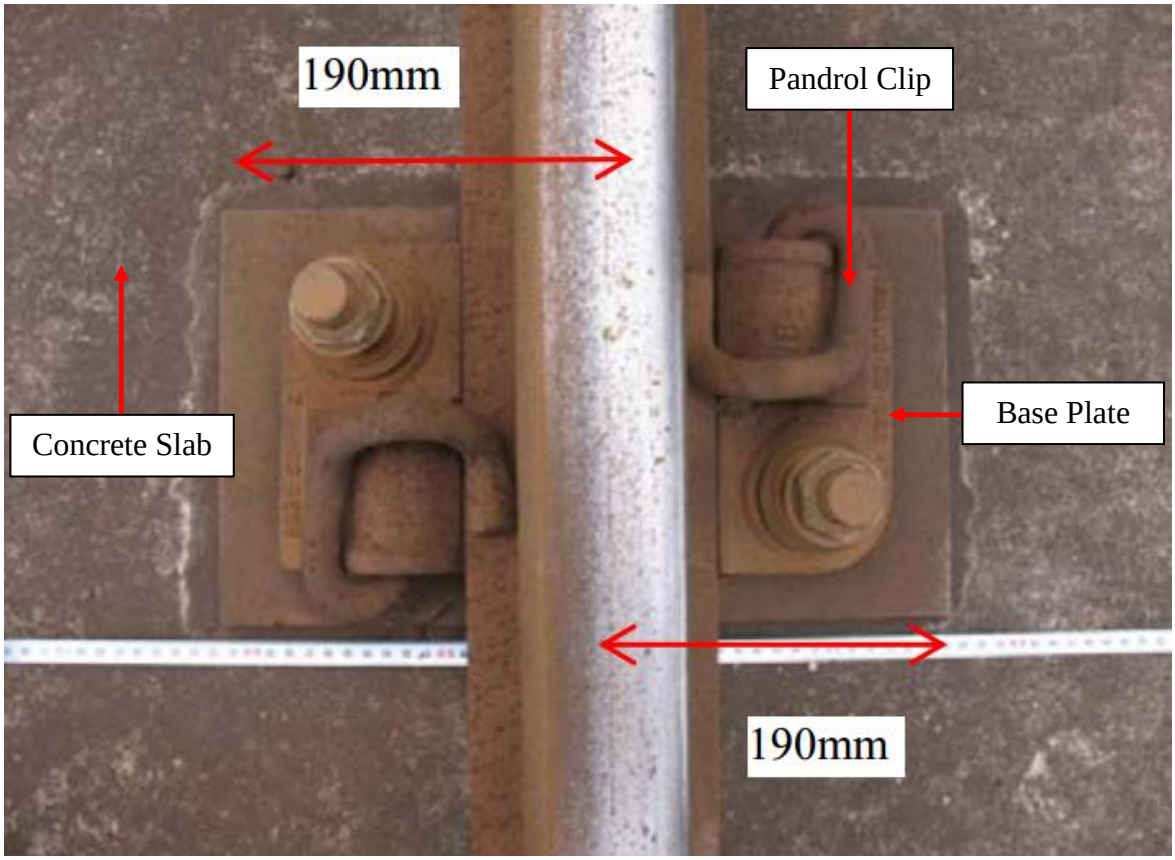



Figure 5.15 A photograph of a Pandrol ‘e’ clip with a cast base plate in situ. A Pandrol clip secures rail in position while the cast base plate provides a secure mounting for rail in concreted areas.

The lack of ballast when using concrete slabs to mount rail benefits potential detection algorithms, as foreign objects interfering with infrastructure will be minimised. Movement, or loss of the Pandrol clips, along with bolt loosening are features that need to be evaluated. As with all bolt fixtures, the detection of their presence can be performed, but loosening is hard to evaluate with matching rail section over a set periods and examining for rotation.

Of added interest is the cracking of the concrete slab. Concrete within the FOV could be examined, but accuracy would depend on defining its position in relationship to ballast and would only be reliable on cracks greater than any other linear features caused by dirt etc.

### 5.2.3.6 Pandrol ‘e’ Clip on Staton Bonna Bi-Block Concrete Sleeper

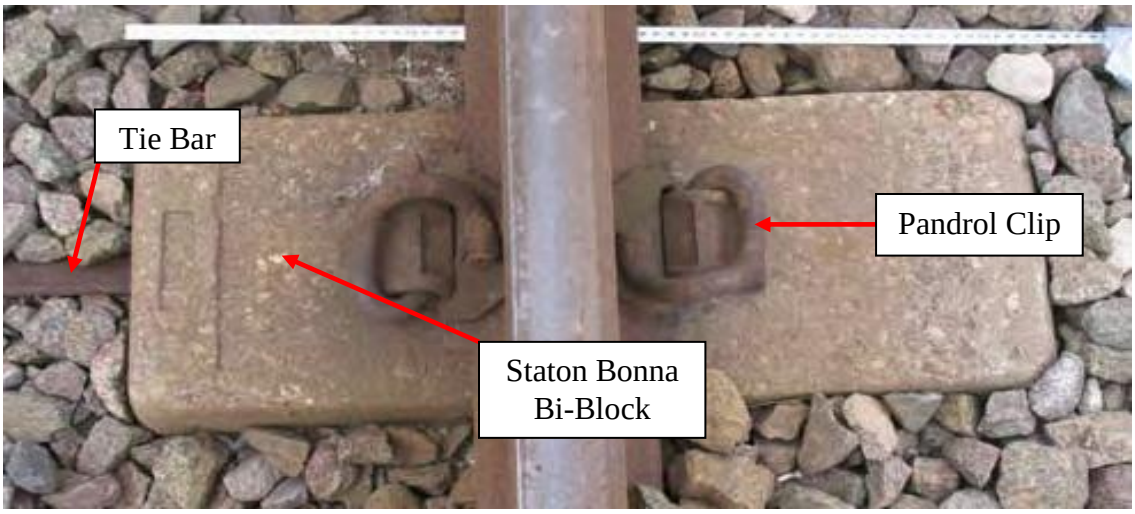



Figure 5.16 A photograph of a Pandrol 'e' clip on Staton Bonna bi-block in situ. A Pandrol clip secures rail in position while the Staton Bonna bi-block provides a resilient base plate for mounting rail on ballasted track. A Tie bar is used to maintain distance between Staton Bonna bi-blocks.

This infrastructure is common within a tram network such Sheffield. There is often interference with ballast that could prevent classification of Pandrol clips. Dirt and grime effect forms of texture analysis. If Pandrol clips can be identified, movement or loss can be evaluated. If clear from ballast interference, what can be seen of the Stanton Bonna bi-block could potentially be evaluated for degradation such as cracks. The associated Tie Bar falls outside of the FOV and cannot be evaluated with the proposed MVS. It is often covered with ballast that could make identification with a MVS mounted in a different location difficult.

### 5.2.3.7 Pandrol ‘e’ Clip on Tiflex Resilient Base Plate on Concrete Slab or Plinth Track and Lords Resilient Base Plate on Concrete Slab or Plinth


Figure 5.17 Photographs of Pandrol ‘e’ clips on a Tiflex resilient base plate in situ. A Pandrol clip secures rail in position while a Tiflex resilient base plate reduces ground born vibration in non-ballasted track areas.


Figure 5.18 Photographs of Pandrol ‘e’ clips on Lords resilient base plates in situ. A Pandrol clip secures rail in position while a Lords resilient base plate provides noise dampening in urban environments.

Similar to Pandrol ‘e’ clips with a cast base plate on concrete, the lack of ballast may improve the performance of classification algorithms for features such as the Pandrol clips. The difference in bolt fittings, if identified, would allow bespoke classification algorithms to inspect the associated infrastructure. While previously discussed, movement in infrastructure requires evaluation; the resilient layer can de-bond from the base plate. While this movement of the rail

can be evaluated further data would be required to see if this de-bonding can be visually assessed within the FOV and camera angle.

#### 5.2.3.8 BS80A Rail on Lord Base Plate with Billet Guardrail on Fabricated Mounting

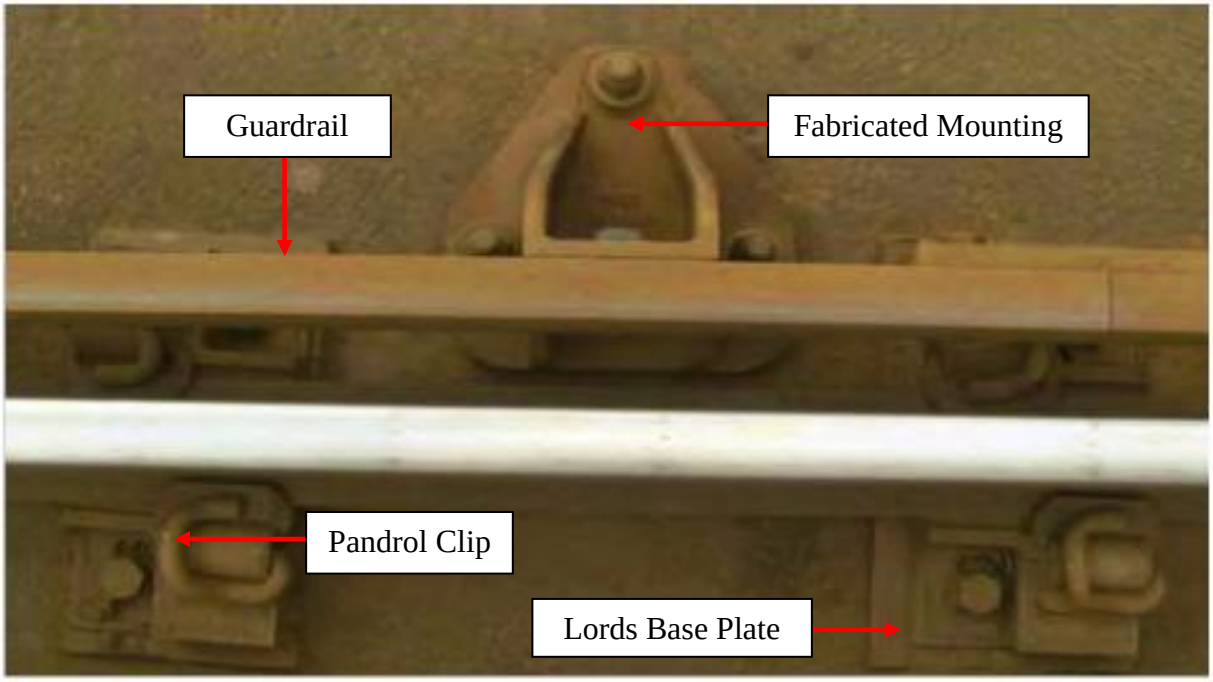



Figure 5.19 A photograph of a BS80A rail on lord base plate with billet guardrail on fabricated mounting in situ. Guardrail is used to prevent potential derailment of a wheelset.

The additional infrastructure introduced with a Billet guardrail makes evaluation of the infrastructure bellow impossible. With the added fittings holding the guard rail, the outer most bolt is potentially outside the FOV. Evaluation of the entire infrastructure is not possible although the lack of ballast potentially allows infrastructure that is within the FOV to be evaluated with a greater accuracy. As above, the resilient layer can deteriorate and further data is required before specific algorithms can be developed.

### 5.2.4 Switches and Crossings (S&C)

#### 5.2.4.1 A Switch Blade on a Variable Slide Chair Mounting


Figure 5.20 A switch blade on a variable slide chair mounting in situ, used to allow traffic to change lines.

A switch blade is a vital point of infrastructure; visual inspection is rated as very important. A switch blade enables railway vehicles to be guided from one track to another, and failure can cause derailment. The ability to inspect the points on a regular basis will greatly improve network safety. The movement and forces that are placed on such infrastructure cause irregular wear. Deformation of the tip or breakage can be classified as a severe defect. Due to this, the tip of the switch blade requires careful evaluation. There is no scale for severity and as such, one will need to be developed.

There are two sides to each Switch Blade, the tip of which should be visible. While changes in the rail edge and running band will show the presence of a Switch Blade, the surrounding infrastructure, such as the fixings and slide chair can improve their detection. Once a Switch Blade is confirmed within an image bespoke algorithms can be developed to asses in detail the measurements taken from image analysis algorithms. Assessing all fixings is unlikely to be possible as bolts that secure the slide chair are outside the FOV.

Associated infrastructure such as tie bars can only be inspected from what part is visible within the FOV. The drive mechanism is also part of the infrastructure for which visual assessment is required. As a majority of this lies outside the FOV the effectiveness of evaluating the drive mechanism by evaluating the visible infrastructure requires further research.

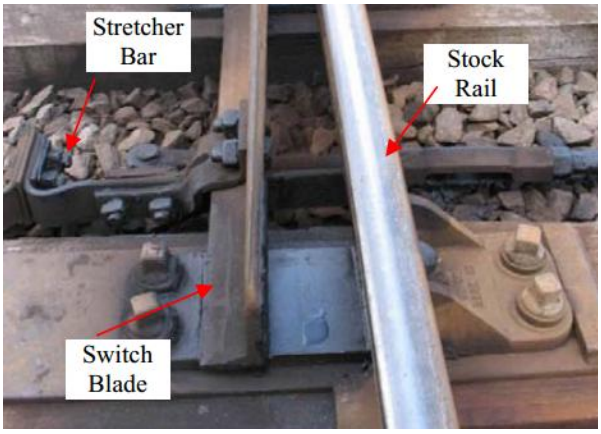



Figure 5.21 A photograph of a switch blade on a variable slide chair mounting in situ, used to allow traffic to change lines.

#### 5.2.4.2 PR7 Diamond Mono Block Crossing

Figure 5.22 A photograph of a PR7 diamond mono block crossing in situ, used to allow traffic to cross lines.

The size of Diamond Block crossings makes detection reasonably easy. The break in the running band needs to be addressed when assessing for defects. The key interest location is the nose of the Diamond Block where deformation can cause a severe defect. Classification for severity needs to be developed. The associated fastenings are also of interest. As the Pandrol clips are partially obscured, the evaluation of their position will require bespoke algorithms. Not all of the associated fixings are visible within the FOV, and the effectiveness of evaluating just the visible infrastructure will require investigation.

### 5.2.4.3 Fabricated ("Frog") Crossing

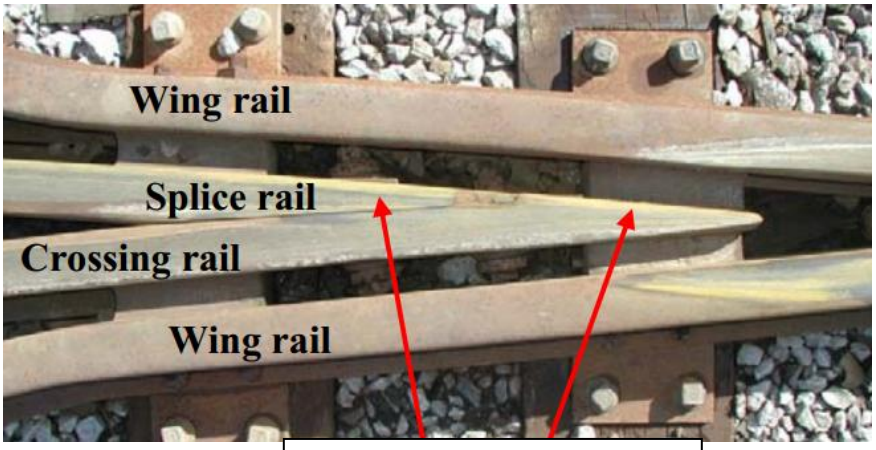



Figure 5.23 A photograph of a fabricated "frog" crossing in situ, used to allow traffic to cross lines.

Fabricated crossings differ from a Diamond Mono Blocks and are machined from existing rail, but serve the same purpose. The nose requires precise evaluation for degradation. The associated fixings are visible with this infrastructure, making analysis easier; however classifying the position due to the similarities with standard rail can become complicated.

### 5.2.4.4 Heel Block

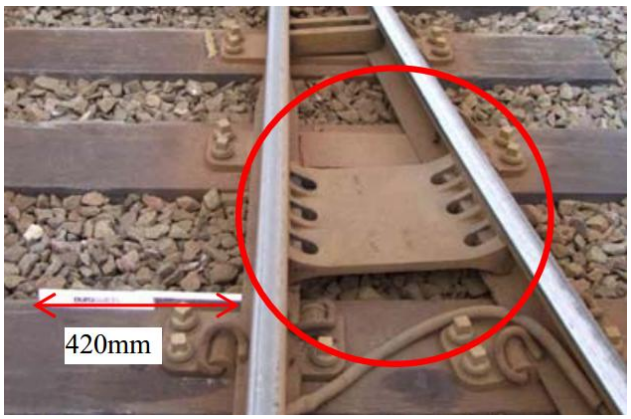



Figure 5.24 A photograph of a Heel block in situ, used to maintain rail spacing at crossings.

A heel block maintains spacing of a switch blade and rail, the features are unique which would allow an single algorithm to detect and classify them. The whole heel block cannot be seen within the FOV of the proposed system; however, the associated bolts and the infrastructure directly related to the railhead are visible. This should allow inspection for fracturing of the item.

### 5.2.4.5 Jumbo Lubricators

Figure 5.25 A photograph of Jumbo lubricators in situ. Lubrication is used to prevent excessive wear on curved track and prevent rolling contact fatigue occurring.

Jumbo lubricators are present on tight curves within a rail or tram network. The occurrence of these items may warrant the use of more sensitive algorithm thresholds when evaluating the rail edges to look for wear. The supply of lubricant cannot really be evaluated with the proposed system without more data examining the effect on the rail when the lubricant runs out. Cleaning may be addressed, however this would still require more data to evaluate the appearance of the infrastructure when cleaning is required compared to a cleaned component.

### 5.2.4.6 Cable Attachments to Rail

Figure 5.26 A photograph of a Broken cable attachment that provides a traction current return.

Cable attachments of various sizes and for different functions are common throughout a rail or tramway network. They can be drilled bolted or studded to the rail. For those fixings that are associated with traction current return and for signalling systems, the loosening of fixings, theft, or breakage of cable can be a significant problem. Within the FOV, the fixing holding the cable can be evaluated providing that some key feature protrudes outside of the head of the rail. Locating this fixing will allow algorithms to detect the presence of cable directly associated with the fixing type. Loosening can be evaluated providing the fixing bolt is visible.

### 5.2.4.7 VIS Loop and Signalling apparatus

Figure 5.27 Photographs of VIS loop (left) and signalling apparatus (right) in situ. VIS loop indicates to signals and a control room the relative location of a vehicle on the network. Signalling apparatus is used to control and direct traffic around the network.

As the signalling apparatus is situated within the centre of the rail, it cannot be evaluated with the proposed system. The limited height means an appropriate FOV cannot be introduced to inspect this infrastructure and the head of the rail. An independent MVS could be introduced to inspect these components and other tracked components such as stretcher bars and drive mechanisms of points.

### 5.2.4.8 Distance "Chainage" Marker

A distance marker shows a distance from a known datum. The recognition of such markings can allow, together with GPS data, the matching of image sets acquired over time. This would allow an effective data library that can be used to examine changes of specific infrastructure points over time. There are significant problems with these items being covered with dirt and while in a new state they are clear against background image information, after a period of time colouration from dirt and foreign objects can obstruct detail.


Figure 5.28 A photograpgh of a distance marker on Staton Bonna bi-block. A distance marker shows the distance from a known datum

Inspection for a distance marker is not required throughout the image set. GPS data can be used to find the relative location of one of these points before bespoke algorithms inspect the reduced image set around this point to

locate and classify its position. This data can then be used to position images and infrastructure accordingly.

#### 5.2.4.9 Pedestrian Crossing

Figure 5.29 A photograph of a typical pedestrian crossing.

A majority of a pedestrian crossing is outside the FOV of the proposed system. The relative edge associated with the rail can be inspected for damage. The distance between rail and crossing platform can also be evaluated, looking for shifts in movement that may eventually cause a collision with a wheel set.

## 5.2.5 Components – Embedded Track

### 5.2.5.1 Grooved Rail Embedded in Polymer


Figure 5.30 A photograph and graphical model of grooved rail. This is used in urban environments where vehicle and pedestrian traffic intersects with the track.

Figure 5.31 Photographs of damaged polymer. Polymer protects the rail from degradation such as rust and helps dampen noise and ground vibration.

Grooved rail is present on much of a tramways network and is visible within the FOV. They are mounted inside a surrounding polymer and laid on concrete, monitoring the degradation is important. There are numerous polymer types in use that require examination for wear and damage.

The surrounding concrete can also suffer from cracking and degradation. This type of damage is difficult to assess since cobbles and tarmac may be placed with, or over, the concrete. While it would be reasonable to assume that polymer damage could occur together with any damage to the concrete this cannot be ensured.

The flange on this type of rail is also of importance, any wear or fracture could show incorrect laying of the track or running of wheel sets. While associative wear of the railhead may cause this kind of degradation to the flange, evaluating this wear is difficult without depth data or evaluating the rail from a different angle.

### 5.2.5.2 Twin Flange-Planned for Checkrail and Attached Bulb Flat Polymer Support

Figure 5.32 Photographs of Twin flange-planned for checkrail. This is used in small sections of embedded track where grooved rail would not be suited.

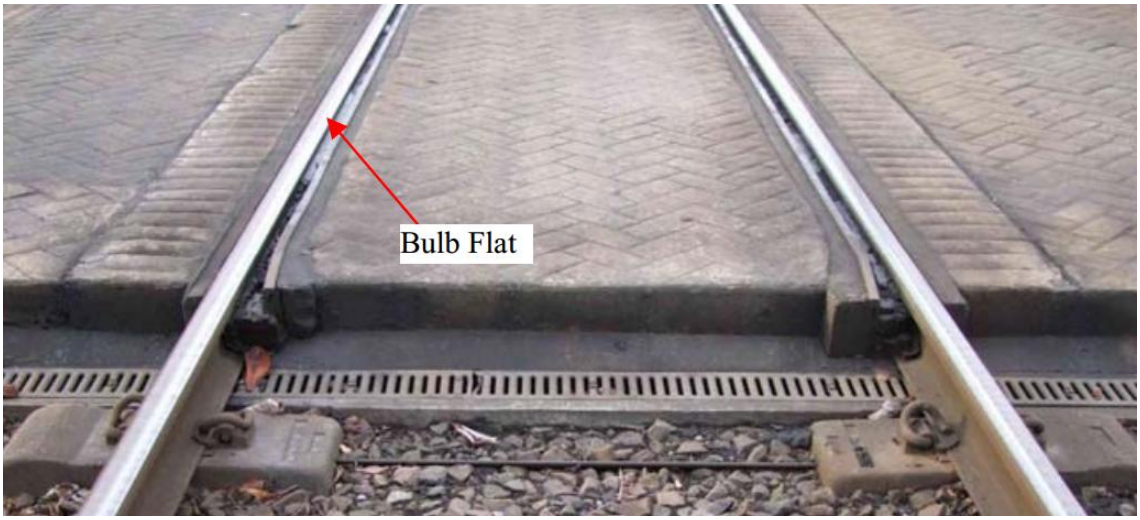



Figure 5.33 A photograph of an attached bulb flat polymer support. This is used in transition between flat bottom and twin flange-planned for checkrail rail and vice versa.

Infrastructure associated with a twin flange-planned is similar to embedded grooved rail, except a second section of standard rail is used to make up the flange. The same evaluation is required as with embedded grooved rail, however evaluation of the flange is different. Both methods are preferred over introducing small sections of grooved rail.

### 5.2.5.3 Adjacent Flange-Path Aco Derailment Channel

Figure 5.34 Photographs of adjacent flange-path Aco derailment channel. These are designed to aid in drainage and provide a path for a potential derailed wheelset to follow

A derailment channel serves two purposes. The first is that it provides drainage on the structure that the infrastructure is built upon. Secondly and more importantly, it provides a channel to contain the wheel if derailment occurs. The plastic drain covers used to cover the channel can become smashed by vandalism leading to blockage by debris such as leaves and soil etc.

As the drains cover the channel, blockages cannot be inspected however smashed covers can be examined and when replaced a clean out could be instigated. As the channel replaces the flange of the embedded track, its analysis needs to be reliable. All of the cover is visible within the FOV.

#### 5.2.5.4 Embedded Expansion Joint

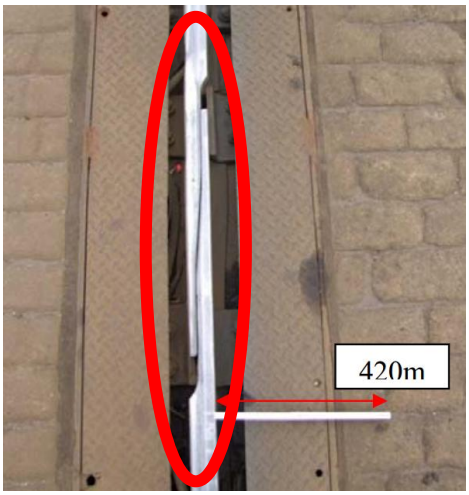



Figure 5.35 A photograph of an Embedded expansion joint. This allows linear expansion of embedded rail due to weather.

An expansion joint itself can be evaluated in the same fashion as standard ballasted track, with the obvious exception of associated fixing infrastructure. The fixings are obscured by a metal plate that in itself may be evaluated for degradation however; the use of such evaluation should be addressed before implementation for usefulness.

### 5.2.6 Switches and Crossings (S&C) Embedded Track

#### 5.2.6.1 Switch Blades

Figure 5.36 A photograph of an Embedded expansion joint. This allows lateral expansion of embedded rail due to weather.

The same requirements for evaluating the tip of an embedded switch blade should be maintained as with ballasted track, however infrastructure such as tie bars are hidden. While the expansion of the keeper will show the presence of a switch blade, the evaluation of some of key components is not possible.

#### 5.2.6.2 Crossings Embedded Track

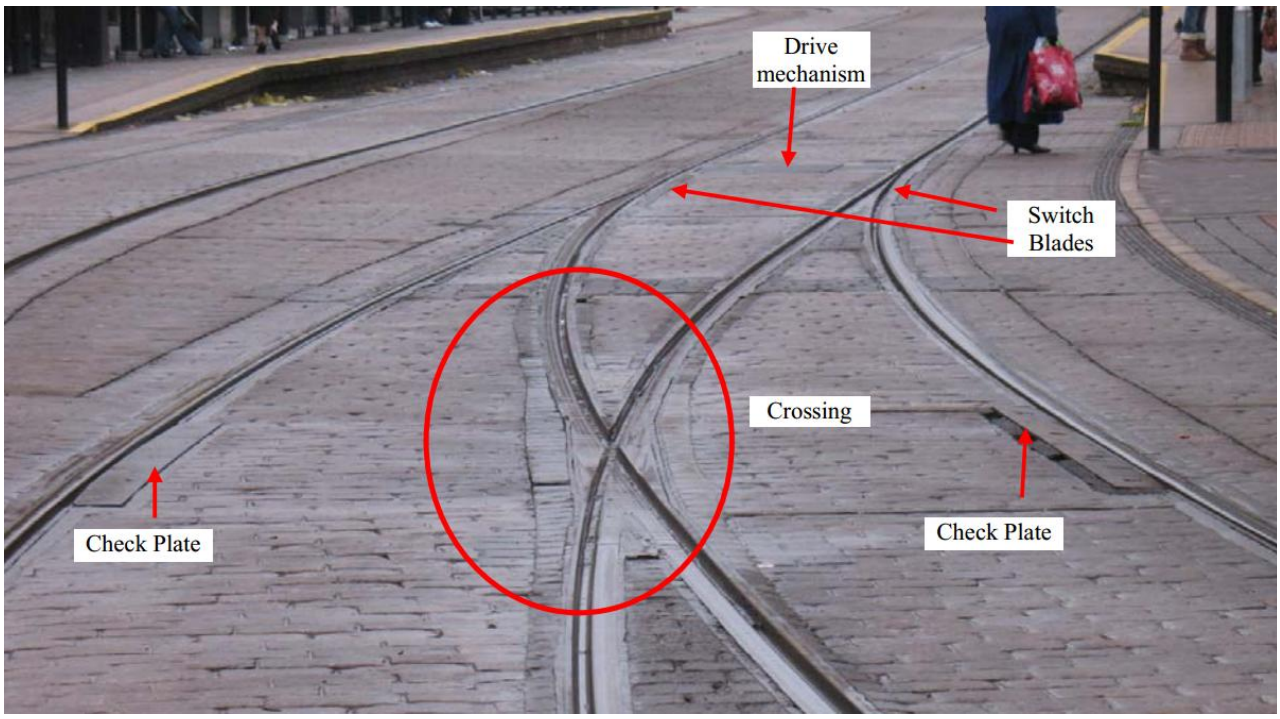



Figure 5.37 A photograph indicating an embedded diamond crossing with supporting infrastructure. A diamond crossing allows vehicles to cross tracks.

Crossing in embedded track varies from ballasted track and as such will need its own independent algorithm developing. The added complexity of evaluating the polymer and surrounding concrete will add complexity to any such algorithm. The lack of associative infrastructure will also make pre-recognition by fixing evaluation more difficult.

#### 5.2.6.3 Replaceable Check Plates

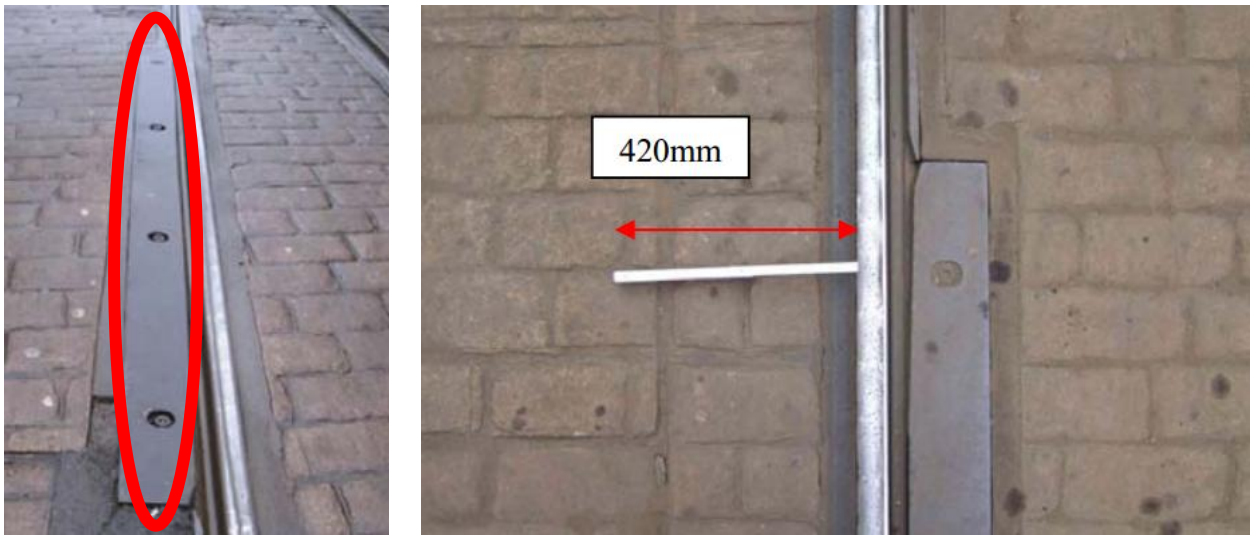



Figure 5.38 Photographs of replaceable check plates in situ. Check plates maintain and align the tracking of wheelsets preventing potential derailment.

These are used to prevent the derailment of a wheel sets at some embedded switch points and crossings. This infrastructure falls fully within the FOV. The space between the check plate and rail edge will need monitoring for wear and adjustment. Surrounding polymer will need examining for integrity; however, the polymer on the check plate side may fall outside the FOV limiting the evaluation of this feature.

### 5.2.7 Other Features

#### 5.2.7.1 Grooved Rail Drain/Drain Gulley Catch Pots

These provide drainage for grooved rail and are clearly visible, falling within the FOV. Other than the square shape, a break in the flange or keeper also denotes its position. As there is a clear view down to the foot, there is the potential to examine for blockages. The plates coving the catch pot can be evaluated for degradation, similarly the leading and trailing edges of the flange or keeper that leads to the infrastructure.

Figure 5.39 A Photograph of a drain gully catch pot for drainage of water.

### 5.2.8 Rail Track and Degradation

#### 5.2.8.1 Wear

The wear rate of the railhead is essential when determining the lifetime of perspective rail; inspection can prevent damage to wheel sets and prevent a potential derailment. Flat bottom rails can suffer vertical and side wear to the gauge face. There is a potential to measure this if the running band changes significantly, or the measured width of the rail is affected disproportionately to measurements of ideal rail width (Chapter 7.4, Chapter 9.4).

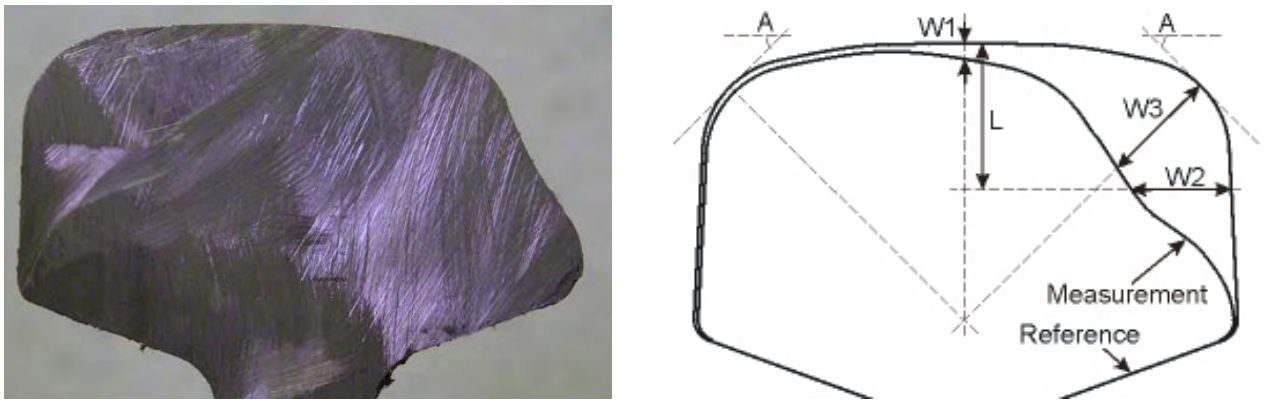



Figure 5.40 An example of rail wear and the measurements used to grade the severity.

Grooved rail can suffer vertical and side wear as with flat bottom rail; however, keeper wear is also a potential problem. The same evaluation can be approached with the head of the rail as with flat bottom rail. Keeper wear can only be evaluated if the width of the keeper is affected disproportionately to measurements of ideal keepers. In both cases, a 3D profiling system would provide more useable and accurate statistics.

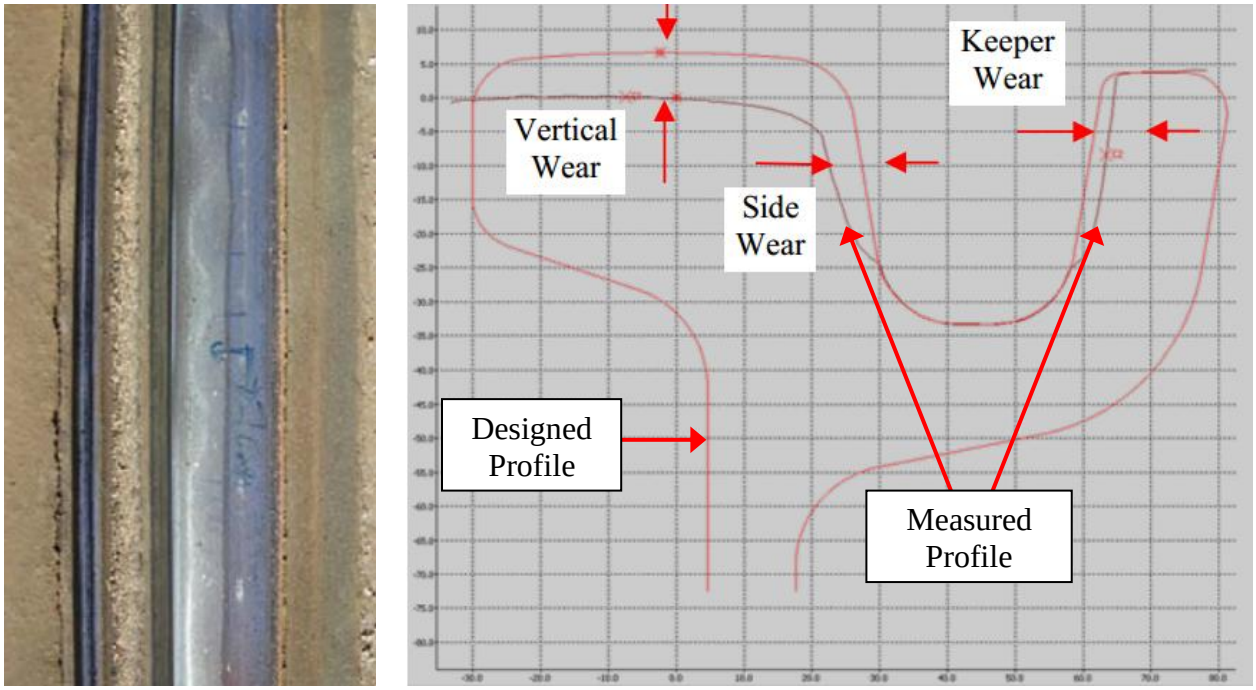



Figure 5.41 An example of rail and keeper wear with the wear measurements in relation to the designed profile.

### 5.2.9 Other Rail Defects

#### 5.2.9.1 Corrugation

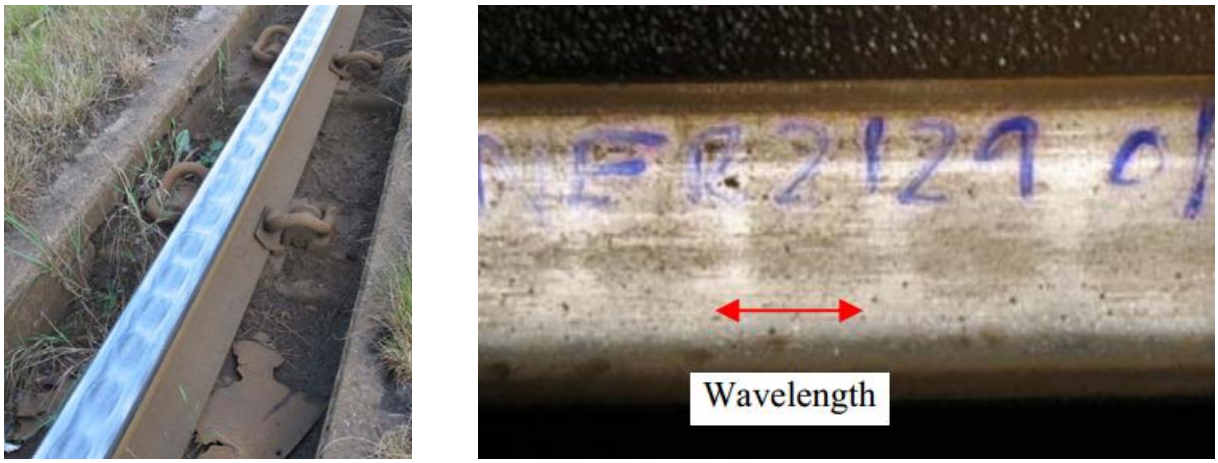



Figure 5.42 Photographic examples of corrugation on rail.

Corrugation is a serious and expensive problem that leads to rail replacement and damage to wheel sets. The periodic wear and deformation of the surface can occur over several tens of meters and often has a fixed wavelength usually between 20 to 200 mm. A discolouration of the rail surface and a textural difference are the key features to be examined for. Once detected, the length, amplitude, and wavelength, need to be evaluated. As this is a defect on the railhead, it is seen as extremely important.

### 5.2.9.2 Plastic Deformation

This defect can be harder to determine and analyse especially in its earlier stages. Plastic deformation is characterised by a “mushrooming” of the rail profile. The deformation will affect measurements of the thickness of the head of the rail. In earlier stages, the deformation will likely be bellow measurable thresholds, resulting in classification being difficult without introducing numerous false classifications (Chapter 7.4).

Figure 5.43 A photograph of a cut piece of rail with plastic deformation

More data will be required to compare different measurements of rail thickness without wear across a rail or tramway network. Data from such research will allow a set threshold of which plastic deformation can be confirmed above any possible false measurements. Once detected, location and magnitude needs to be assessed. There is no standard for severity and thus one will need to be developed to set priority of maintenance.

### 5.2.9.3 Rolling Contact Fatigue (RCF)

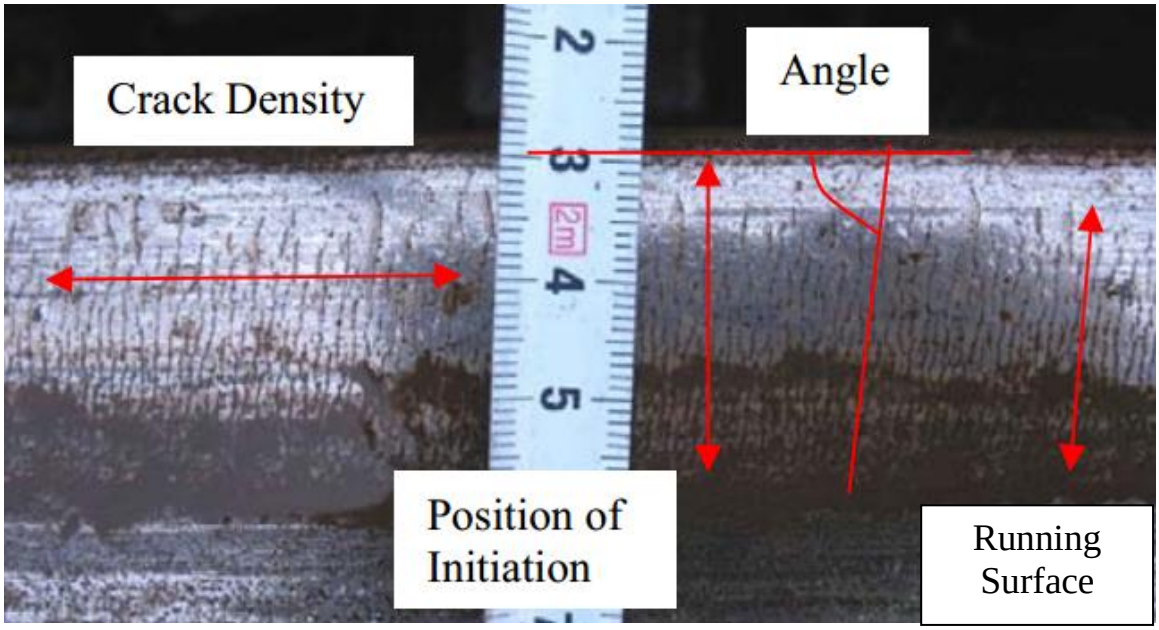



Figure 5.44 A photograph of Rolling Contact Fatigue (RCF) and the density of its occurrence on a rail surface.

Rolling contact fatigue is, while in its early stages, a repairable defect on ballasted track and grooved rail. Grinding, an intrusive rail repair method, allows RCF to be removed, and is preferred over rail replacement. RCF cracks are a precursive effect of underlying defects or stress occurrence of the rail through traffic load or wheel slip. Currently underlying defects can be detected by current sensing methods, such as eddy current inspection or ultrasonic testing, although these suffer limitations in detecting the early onset of RCF (Innotrack, 2008)(M. Papaelias, 2008).

Detection of such cracks only occurs when they reach a certain size, which propagates vertically throughout the railhead, making them more visible. The proposed MVS should be able to detect these defects at an earlier stage given a great enough resolution within the images acquired. The required resolution to detect these defects limits the top speed of the MVS systems by limiting the maximum FOV.

Once RCF cracks reach a certain size, they have the potential to become severe defects, causing spallation and rail cracking. The length of the crack is an indicator of the severity (Corus Rail Technologies, 2000). Other useful statistics will be the angle of crack from the gauge distance, the position of initiation with reference to the gauge side and the crack density i.e. the number of cracks per 100mm.


Figure 5.45 A Photograph of rail cracking caused by RCF development over time.

#### 5.2.9.4 Squats and Wheel Burns

Figure 5.46 An example of a severe squat on the running surface of a flat bottom rail.

Squats and wheel burns are caused by different means, but their characteristics are very similar. Both defects lead to a plastic deformation of the railhead, causing a shift in running band and/or a discoloration of the affected area. Squats, caused by underlying defects of the railhead, can develop rapidly over time. Detection of squats requires immediate action, usually resulting in rail replacement.

Figure 5.47 A photograph of wheel burns on the running surface of flat bottom rail.

Wheel burn has the unique characteristics of appearing on both sides of the rail. These are caused by wheel slip. When traction is lost between wheels and the railhead both sets of wheels generate massive amounts of friction heat that causes a deformation of the railhead. Wheel burn actions depend on the size and severity of the defect and usually occur around stations and signals, where vehicles stop and start regularly. Both squats and wheel burn requires measurements of their length, however classification by their area is possible using a MVS system (Chapter 10.4). Any deviation in running band or spalling of the rail surface overtime will also need recording.

#### 5.2.9.5 Broken Rail

Figure 5.48 Photographic examples of broken grooved rail (left) and broken flat bottom rail (right).

This severe defect needs to be addressed immediately. While a rail break is not generally clean, careful examination should be made to prevent false classification of rail joints and isolated rail joints. The use of GPS data may be required to eliminate some false classifications and if image sets can be effectively matched, known rail joints can be eliminated from classification. If

detected immediate action is required. The severity should be based upon measuring the distance between rail ends.

## 5.3 SUMMARY OF INFRASTRUCTURE AND ASSOCIATED DEFECTS TO BE ANALYSED IN THE FOLLOWING CHAPTERS

### 5.3.1 The Head of the Rail and Associated Rail Defects

Detecting the head of the rail is vital, not only is it an important part of infrastructure that requires analysis, but it also allows relevant infrastructure to be evaluated with its position in relationship to the railhead.

As the railhead carries several possible defects, the majority of research is applied to this field. Initial investigations focus on locating the rail within the image, before performing an analysis of the edge cues to evaluate deformation of the railhead. This analysis also allows classification of rail joints and potential breaks. Once correctly classified, further algorithms are described that look for surface defects.

Before a defect such as a squat or wheel burn can be detected, the running band needs to be successfully segmented from the image. Analysis of the edge of this running band is also useful to evaluate potential defects such as corrugation, squats, wheel burns and incorrectly laid rail. When classified, the running band data can be specifically analysed for irregularities. RCF cracks are also present within the running band; some work has been undertaken that can be applied to their detection.

### 5.3.2 Rail Joints and Fastenings

With the large quantity of infrastructure available for inspection, the scope of research needs to be focused sufficiently, without sacrificing purpose of the thesis. Infrastructure such as fish plates interact directly with the rail and carry a high potential for defects. Pandrol clips also have a close relationship with the rail and their detection is important. While this thesis does not attempt to cover the application of algorithms to detect such infrastructure, its detection is seen as an important aspect of a MVS for rail inspection and will be investigated within future iterations of the described MVS.

# CHAPTER 6 IMAGE ANALYSIS CHAPTER LAYOUT AND METHODOLOGY

## 6.1 INTRODUCTION

The following chapter describe the layout and data sets used in the image analysis chapters. The design of the image analysis algorithms focus on robustness and execution time. These factors are important due to the high volume of images that the algorithms are to be applied. Taking the Sheffield Supertram network as an example, their network consists of 29 Km of track operating mainly in both directions roughly totalling 58 Km of track. In total assuming a proposed resolution of 0.1 mm per pixel and a fixed image height of 480 pixels, a rough total of 1.2 million images per rail (left and right) will be acquired.

The robustness of algorithms is vital, not only to deal with potential changes in weather and lighting but also to deal with hardware alteration. As the hardware was developed alongside image processing algorithms small changes to the Field of View (FOV), camera angle, camera type and lenses employed can affect imagery. Similarly, the results from the initial investigation process can change the deployment of the hardware stipulating factors such as lighting conditions and camera angle.

## 6.2 DATA SETS

There are four data sets involved in the development of the image analysis algorithms described in the following chapters 7, 9, and 10. Table 6.1 shows the number of images per data set and describes the source of the images. Chapter 8 only makes use of 1000 images from the 4$^{th}$ Data set. Data set 1, used within the 'Investigative Research' section of the chapters 7, 9, and 10, is a selection of stock imagery. The contents varies from images of specific defects such as corrugation, squats, wheel burns and rail without defects, Figure 6.1 shows some example images. These images were used to evaluate common image processing algorithms in detecting characteristics of features and defects of the rail under differing conditions. Changes in lighting, camera hardware, field of view, and camera angle were made to aid in the development of the

machine vision system (MVS) as described in Chapter 4 Design of a MVS for the Acquisition of Tramway Rail.

Data set 2 is a set of 20 lower resolution images taken from a machine vision system. Data set 2 was used in conjunction with data set 3 in developing the image processing algorithms. The FOV of the data sets is similar to that described in Chapter 4 Design of a MVS for the Acquisition of Tramway Rail. These data sets allowed improvement in feature detection algorithms to be made while carefully evaluating how changes in image resolution affected results. Figure 6.2 shows a selection of images from data set 2 and Figure 6.3 data set 3.

| Data Set Number | Total Images | Chapter Use | Description |
|---|---|---|---|
| 1 | 969 | 'Investigative Research' | A selection of stock images from various sources. |
| 2 | 20 | Initial Evaluation | Low-resolution images taken from a machine vision system. |
| 3 | 5 | Initial Evaluation | High-resolution images taken from a machine vision system. |
| 4 | 5000 | Final Evaluation | High-resolution images taken from machine vision system on the Sheffield Supertram network. |

Table 6.1 Reference table for data set contents used for the development of image analysis algorithms.

Data set 4 is used as a final evaluation data set it was used to test the final design of the image analysis algorithm. The images were acquired from the machine vision system described in Chapter 4 after all alterations had been made to hardware. The larger data set of 5000 images covers 240 meters of track and is sub divided into 48-meter sections from random positions around the network. Figure 6.4 shows some of images from data set 4. All images presented from the 'Final Evaluation' data set 4 will have been corrected for vibration using the algorithm described in Chapter 8 allowing for consistency in visual assessment.

Figure 6.1 Example images from data set 1 used within the investigative research stages (8 of 969). Data set 1 consists of a selection of stock images from various sources. Images from data set 1 feature various characteristics and are taken using a digital camera.

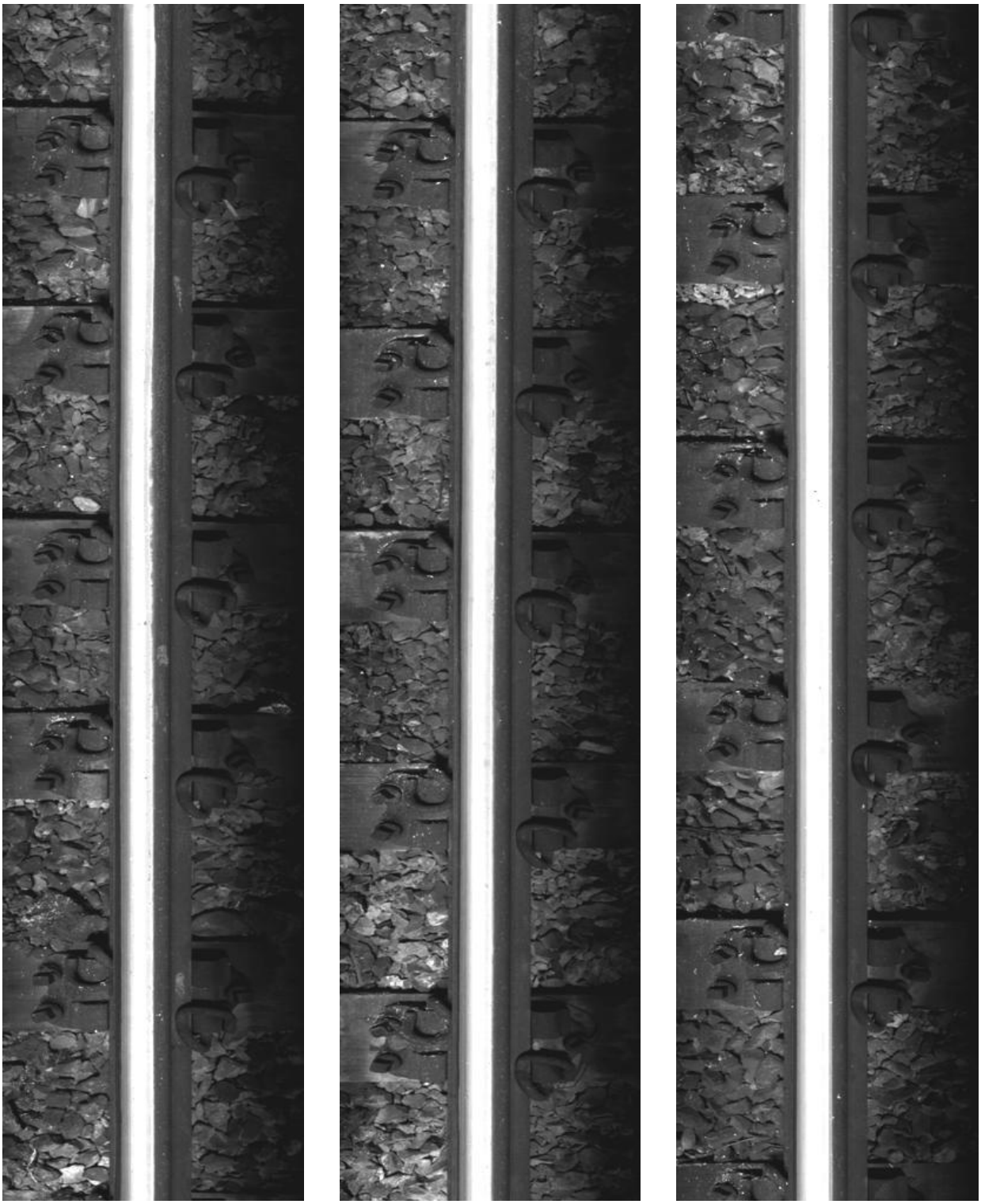


Figure 6.2 Example images from data set 2 (3 of 20) used within the initial investigation stages. Data set 2 consists of low-resolution images taken from a machine vision system.

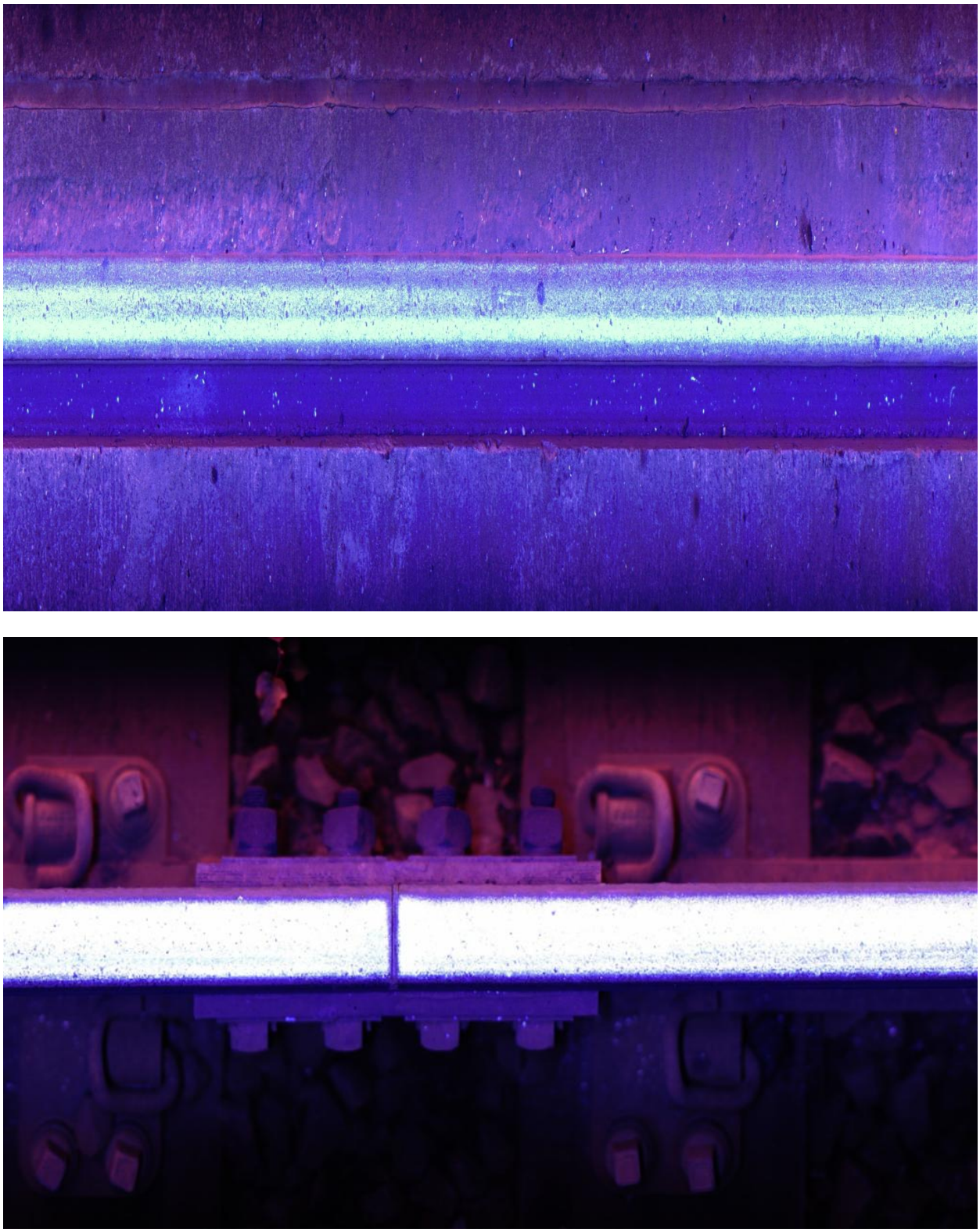


Figure 6.3 Example images from data set 3 (2 of 5) used within the initial investigation stages. Data set 3 consists of a selection high-resolution images taken from a machine vision system.

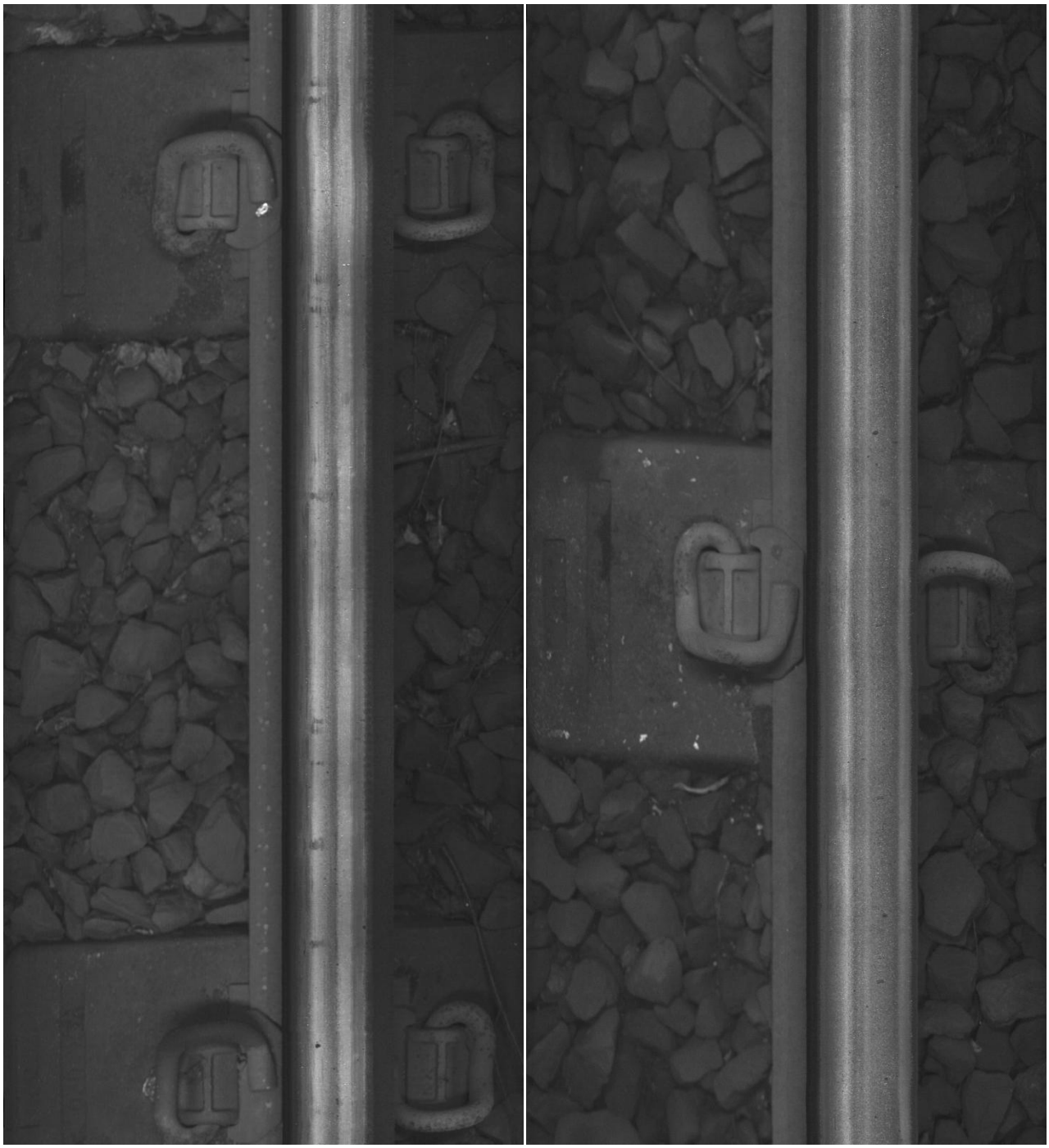


Figure 6.4 Example images from data set 4 (2 images of a set of 10 joined together, 20 of 5000 total images) used for the final evaluation of image processing algorithms. Data set 4 consists of 5000 High-resolution images taken from machine vision system on the Sheffield Supertram network.

## 6.3 METHODOLOGY

### 6.3.1 Chapter 7 , Chapter 9 , and Chapter 10 layout and Methodology

The first stage of development for chapters 7, 9, and 10 was to apply some common image processing algorithms used for the particular feature that were present. As described in Chapter 3.3.3 page 72, images acquired for preliminary investigation of the FOV and camera angle were used to evaluate image-processing methods i.e. Data set 1(examples within Figure 6.1) . These images were used with the 'Investigative Research' sub heading of each chapter. The source images for these test where varied to display the best part of the features under examination.

After 'Investigative Research', results were observed under the section heading of 'Method Evaluation and Observations'. This allowed for an educated decision in the development of algorithms that could then be made upon the most successful and robust methods. The objective was to increase the capabilities of basic methods and adapt them to complement the hardware and image resolution. The results were algorithms that were robust, capable with dealing with small changes in lighting and hardware, but also highly adaptable. Since algorithms were designed around the feature, adaptations could be made for changes in FOV or camera angle.

Suitable methods that showed some successful detection of the features being examined were expanded upon with an 'Initial' and 'Secondary' evaluation data sets 2 and 3 (examples within Figure 6.2 and Figure 6.3). The initial data set 2 was taken from an MVS with a low quality output. The secondary data set 3 was acquired after hardware adjustments to satisfy finding from the 'Investigative Research' at a much higher resolution. Adjustments to algorithms were made until satisfactory results were seen from these two smaller data sets. This increased algorithm design efficiency as; the data sets were considerably different as to allow effective evaluation of robustness, an evaluation of its adaptability when hardware changes are introduced, and the algorithms efficiency in execution time could be inspected.

Once an algorithm performed well on the smaller data sets, the algorithm was applied to the larger data set 4 of 5000 images (examples within Figure 6.4). Assessment of the algorithms

performance must be visually assessed meaning the 5000 images, although limited in size, was felt to be of a significant enough sample to evaluate an algorithms performance for future development.

### 6.3.2 Chapter 8 layout and Methodology

Chapter 8 differs from the other image analysis chapters, as its application is only relevant to the larger data set 4 of 5000 Images. Chapter 8 describes methods to correct for vibration within the image caused by camera movement. The ability of image correction is dependent on the performance of the algorithm presented in Chapter 7 and therefore methodology is split up into successfully classifying key points within the image and then applying stabilisation algorithms to correct for camera movement. A selection of 1000 images was taken from the data set 4 to evaluate the performance of the algorithm.

All images presented from the 'Final Evaluation' data set 4 will have been corrected for vibration using the algorithm described in Chapter 8 allowing for consistency in visual assessment.

# CHAPTER 7 LINEAR EDGE DETECTION IN THE CLASSIFICATION AND ANALYSIS OF TRAMWAY RAIL EDGE

## 7.1 INTRODUCTION

The rail edge itself is the most important re-occurring feature within rail image analysis. Detection and classification of the rail edge within the image plays an important role in locating features and defects. More significantly, successful classification aids in the parallelisation of the image analysis software architecture. Locating defects is important but also the ability to segmenting the image into its relevant areas allows the data quantity, during specific analysis steps, to be minimised, improving the efficiency and accuracy of bespoke defect analysis methods.

The rail edge is defined, in this scope of research, as the left and right hand side of the head of the rail. The foot of the rail is not considered to be the rail edge as this feature is not present in embedded track and is interfered with by infrastructure such as Pandrol clips and ballast.

Detecting the rail edge allows for severe defects such as breaks and gaps between rail joints to be measured and classified according to severity. A novel method has been developed that detects the relative location of the head of the rail using FFT template matching algorithms. Using results from this method the location of the head of the rail edge cues can be estimated. This produces a reduced data set that allows bespoke convolution filters to be applied in classifying the defined rail edge.

The results from the convolution filter can be processed to determine features such as exact location within the image, railhead deformation, and width along the length of rail. Breaks within the rail can also be determined; however, additional algorithms are required to classify a difference between known rail joints and fracture of the rail.

The application of detecting the rail edge also allows for vibration from camera movement to be corrected. Vibration from vehicle movement is passed through the bogie, causing a linear movement of the camera, perpendicular to the direction of acquisition. Chapter 8 , 'Novel Feature Matching Algorithm for the Suppression of Tramway MVS Vibration for Image Stabilisation' discusses the application.

The vibration causes a wave type effect throughout the image demonstrated in Figure 7.1. By detecting the rail edges using a bespoke convolution filter the key points detected can be used for re-alignment of the image. The method of vibration correction is applied after successfully classifying the rail edge; as such, vibration does not directly affect the classification of the rail edge, and does not affect any measurement methods applied. All images presented from the 'Final Evaluation' data set 4 will have been corrected for vibration using the algorithm described in Chapter 8 allowing for consistency in visual assessment.

Original Image

Corrected Image

Figure 7.1 Example application of the vibration correction algorithm developed in Chapter 8 applied to series of 8 consecutive images joined to form 1 larger image of a rail section.

## 7.2 INVESTIGATIVE RESEARCH

Edge detection is a well-documented field within image analysis and is discussed in detail under Chapter heading 2.3.1: Feature Extraction. Well-tested edge detection methods such as the Canny and Sobel algorithms have been examined in comparison with more advanced methods such as back propagation neural networks.

As the rail edge is consistent in both shape and size, a range of filters and templates have been examined with a view to classifying the rail edge location within the image. The additional benefits of these methods are the generation of a minimised location for the likelihood of the rail edge to occur. Bespoke linear convolution filters have been shown to improve upon 'standard' edge detection methods in comparison to statistical analysis as an alternative to a convolved filter. Neural networks are investigated as a robust method of rail edge classification that can produce a more bespoke filter.

Results are assessed visually and statistically for accuracy, complexity, robustness, and execution time. Algorithms were tested on several images before the most successful were taken on to further development and application. All images that the algorithms were tested on were part of data set 1 unless stated otherwise, see chapter 6.2 Data Sets (page 142) for reference. They are from a selection of stock imagery totalling 969 images the content varies in image quality, camera angle, and lighting conditions. Those image presented show typical results from the data set and where possible have a similar FOV and camera angle to that of the final data set.

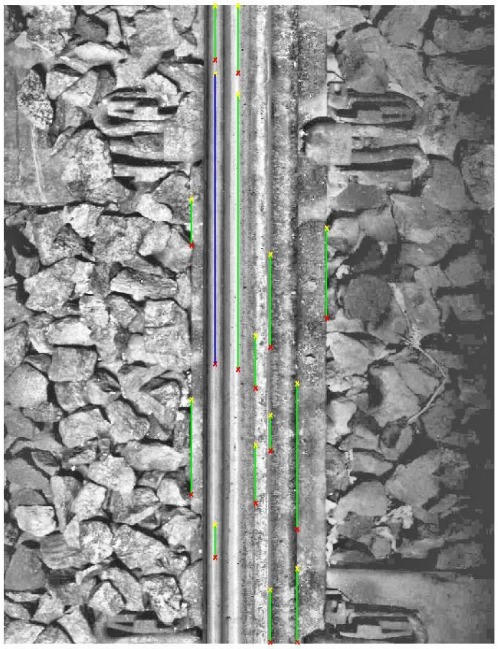
### 7.2.1 Standard Edge detection Methods for Locating the Rail Edge

When using standard edge detection methods the benefits are often clearer than the disabling factors. Using well-tested methods such as Sobel (I. Sobel, 1978) and Canny (1986) with larger Gaussian filters (C. Demant, 2013) can produce a binary image in which the rail edge feature can be determined easily. The filter however is an intermediate stage designed to remove noise from a cluttered image. Noise in this application describes the false classification of edge type features from the edge detection algorithm. Such noise can be caused by linear edge cues associated with

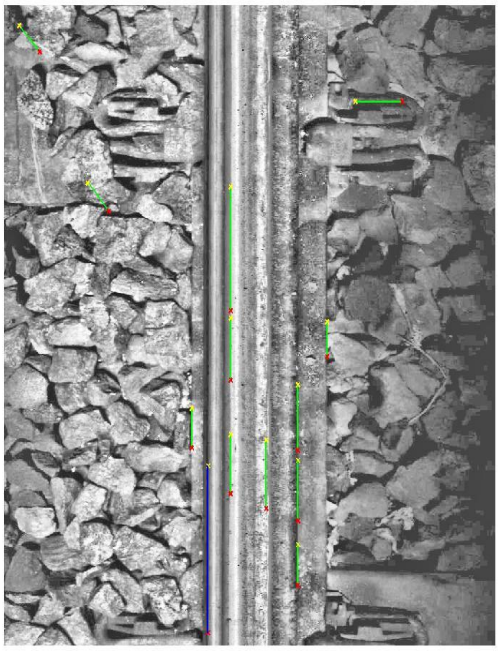
ballast or other features present within the image. This differs from noise caused during image acquisition that is to be considered negligible, due to acquisition hardware being resistant to environmental effects.

There are two distinct methods for dealing with the noise in the binary image caused by false classification. The first is to use the Hough transform (Hough V, 1962) this method uses a voting system that can be used to favour linear features occurring in the right direction. This method then allows the interim step of a Gaussian filter to be removed, however a distinct improvement is visible when using a Gaussian filters, as shown in Figure 7.2.

Hough Results/Canny Edge detection

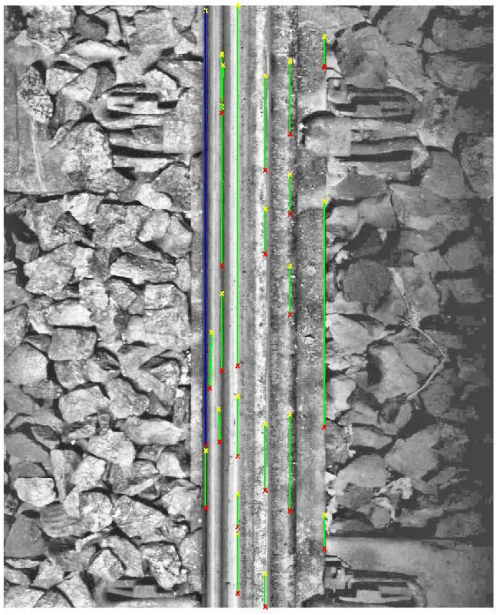


With Gaussian Filter

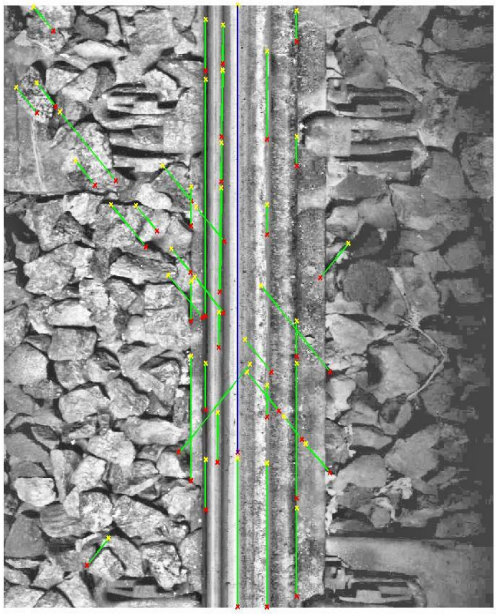


Without Gaussian Filter

Hough Results/Sobel Edge detection



With Gaussian Filter



Without Gaussian Filter

Figure 7.2 Demonstration of the improved results from Hough transform analysis when a Gaussian filter is used in the application of rail edge classification.

The unwanted classifications found by the Hough transform when no Gaussian filter is used can be eliminated by only utilising the co-ordinates in the vertical plane that are over a certain size. The Hough transform allows the classification of the rail edge yet also allows features such as the foot and running band to be classified as additional linear features. To ensure correct classification, the Hough transform cannot be adjusted to take a set number of primary edges; if this is applied misclassification of the rail edge can be caused by the running band. The running band is often the primary linear feature and if multiple or split running bands are present this can prevent the rail edge being located.

To distinguish between the linear features, data set constants from features such as accepted widths for the railhead can be used. As the rail thickness remains a stable constant, there will be two points generated by the Hough transform that fall within this range. Sections along this line could be further analysed if required to ensure that the rail edge occurs at that point.

As discussed, calibrating the Hough transform to classify the total rail edge is difficult Figure 7.2 shows results achieved when applying the Hough transform. Even if post processing of the data could allow correct classification addition methods would be required that will construct the edge data required for analysis.

### 7.2.2 Filters and Templates

A fast method of rail edge detection is to produce a filter for each edge of the rail that can be applied to the whole image fast and efficiently. This method can improve the results presented by standard edge detection techniques, as edges that only conform to those found on the rail will be found. This eliminates a majority of noise caused surrounding features. If a single edge filter is applied, linear features such as the foot of the rail can still interfere with results. The use of a bespoke filter or template allows multiple edges to be located in one sweep and allows edges of a certain distance apart to be located.

A convolution filter can be used that will show all edges cues over a certain plane much like in a standard edge detection method, a threshold can then be used to determine positive matches. The benefits of such a filter is that the mask size can be made larger and more linear (Equation 7.2) than that of which can be applied by standard edge detection algorithms (Equation 7.1). While noise can still affect the filter, larger lines are easier to determine programmatically. Figure 7.3 shows that applying convolution filter produces better results that that produced by the Sobel edge detection method.

Sobel's edge detection algorithm was visually determined to be a better algorithm in comparison to the Canny alternative. Using the Sobel method there is a distinct correlation between the linear features detected, and noise within the image that can be clearly seen from Figure 7.3. Noise can be filtered out from the image by the use of non-maximum suppression and/or hysteresis thresholding. This is due to the image being transformed into an evaluation of intensity gradients.

$$Sobel\ (vertical) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \& \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

Equation 7.1 Standard Sobel vertical edge detection filters

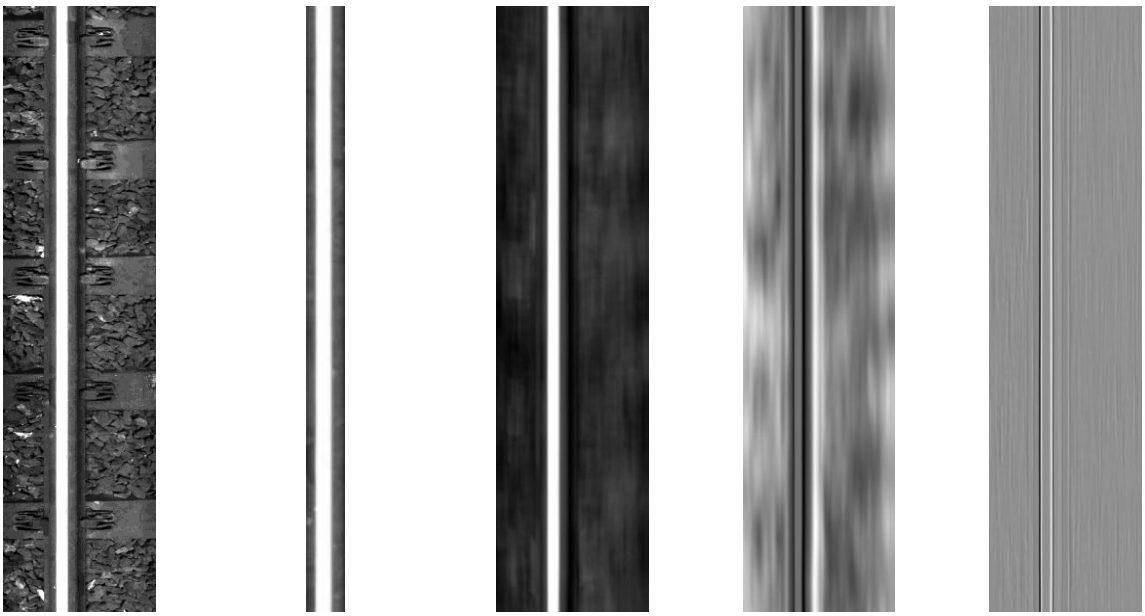
$$Convolution\ Filter_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -3 & 0 & 3 \\ -4 & 0 & 4 \\ -3 & 0 & 3 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \& \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 3 & 0 & -3 \\ 4 & 0 & -4 \\ 3 & 0 & -3 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

Equation 7.2 Linear convolution filter based on Sobel algorithm

Original Image | Standard Sobel Vertical Edge Detection | Linear Convolution Filter based on Sobel Algorithm

Figure 7.3 The difference in results from standard Sobel edge detection results and extended bespoke convolution filters.

An alternative to producing a convolution filter is the use of a template generated from image data. Larger and potentially more accurate filters can be produced using this method. Figure 7.4 shows these results from template matching data from the head of the rail on an Image.



Original Image | Template | Cross Correlation* | Sum of the Squared Difference* | Correlation Coefficient*

*Matches of the upper left point of the template are marked creating a shift in alignment. Normalised versions are the same, however values are adjusted to fall within an upper and lower limit threshold usually -1 0 and 1.

Figure 7.4 Application of FFT template matching in classifying the railhead using a template formed from image data.

The Fast Fourier Transform (FFT) template matching method (Lewis, 1995) is a fast and effective alternative way for detecting the head of the rail, and will usually locate the edges based on the higher intensity running band. This can cause a shift in the matched edges as the running band moves across the head of the rail. Future work could be to use this method in conjunction with standard edge detection algorithms to minimise the region of interest of the image, so that the foot of the rail is the only linear interference within the area being analysed.

FFT template matching is fast and effective in classifying the head of the rail. It is limited when measuring features of the rail like the heads width. An advantage of this method is that it can work for all types of rail if a running band is present. It does not aid however in the classification of the type of rail i.e. if there is a keeper present or not. This will required an extra stage of analysis such as the presence of Pandrol clips, ballast, or sleepers.

### 7.2.3 Bespoke Linear Convolution Filters

To further examine the use of bespoke convolution filters for edge cue classification, data set 2 of 20 lower resolution images was tested with three pairs of convolution filters, see chapter 6.2 Data Sets (page 142) for reference. The adaptability of the algorithm was then tested in its performance on five higher definition images from data set 3 acquired from the more advanced MVS. These data sets were used as they provided a more consistence image FOV from which evaluate the performance of the filters. The three pairs of convolution filters are 6, 12, and 18 pixels in height and conform to the mathematical rule presented in Table 7.1.

$$\begin{bmatrix} -1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ -\left\|m'/2\right\| = m' & 0 & \|m/2\| = m \\ -\|n\ /2\| = m' & 0 & \|n\ /2\| = m \\ -\left\|m'/2\right\| = m' & 0 & \|m/2\| = m \\ \vdots & \vdots & \vdots \\ -1 & 0 & 1 \end{bmatrix}$$

Form for Odd Filter Height

$$\begin{bmatrix} -1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ -\|m'/2\| = m' & 0 & \|m/2\| = m \\ -\|n\ /2\| = m' & 0 & \|n\ /2\| = m \\ -\|n\ /2\| = m' & 0 & \|n\ /2\| = m \\ -\|m'/2\| = m' & 0 & \|m/2\| = m \\ \vdots & \vdots & \vdots \\ -1 & 0 & 1 \end{bmatrix}$$

Form for Even Filter Height

*Where n is equal to the filter height in pixels

Table 7.1 Mathematical representation of bespoke linear convolution filters.

The filter is applied by using a convolution neighbourhood operation, in which the pixel being examined has a value calculated by mathematical operations on itself and its surrounding data. The filters shown in Table 7.1 are applied twice, once in the form presented and again with a 180-degree rotation about their centre. The surrounding data of each pixel is calculated and the

summation of the filters forms a final pixel value. This allows both positive and negative edge gradients to be located within an image.

After the filter is applied, a noise removal algorithm is run to remove smaller edges detected within the image. The requirements of this data suppression are that the edge is under a certain height and the area it occupies within the image is under a set threshold. The Hough line detection (Hough V, 1962) algorithm is then run on the remaining data to find the strong edges within the image.

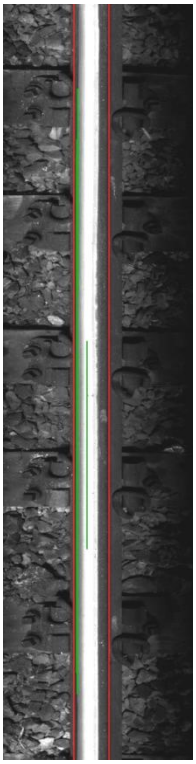
To allow comparison between the advantages of bespoke convolution filters and standard edge detection methods discussed in Chapter 7.2.1 ‘traditional’ edge detectors have been be applied to the images. Both Canny (Canny, 1986) and Sobel (I. Sobel, 1978) methods are implemented under the same comparable constraints of the bespoke convolution filter.

Examples of the results from image evaluation are presented in Figure 7.5. All images were visually assessed and the bespoke convolution algorithms ability to detect the right and left edge of the railhead was scored according to Table 7.2. With the set of 20 Images from data set 2, the maximum possible score is 100. With the higher quality set of 5 images from data set 3 the maximum possible score is 20 this multiplied by 5 to give a maximum of 100. The method with the highest overall score is the most accurate method. Receiver operating characteristics data is not specifically useful in the evaluation of this data as the true negative rate will always be zero.

Application of left edge convolution filter 3x6 example

Application of right edge convolution filter 3x6 example

Output image after edge suppression and small object removal



Results after Hough line classification algorithm is applied

Figure 7.5 Application of convolution filter in rail edge cue classification.

| Description | Weight Given |
|---|---|
| Edge Found. (True Positive) | 2 |
| False edge found, however will not affect rail defect analysis. (False Positive) | 1 |
| No edge located. (False Negative) | -1 |
| False edge found that would affect defect analysis. (False Positive) | -2 |

Table 7.2 Scoring method in the visual assessment of the bespoke convolution algorithms ability to detect the right and left railhead edge.

Each method was evaluated with 4 sub-methods to test for performance of the algorithms. Only the strongest linear feature was taken from the Hough transform. The first column in Table 7.3 and Table 7.4 allowed only 1 pixel either side of a manually assed edge location to be classified by the Hough transform. This evaluated the results from the Hough transform in its ability to

classify an edge at the expected location. The same algorithmic methods were applied with an increased error rate of 5 and 10 pixels either side of the manually classified edge location. The results allow the robustness of the algorithms to be tested in classifying the same rail edge that was classified through manual assessment of the image.

| | Calculated Success Results | | | |
|---|---|---|---|---|
| Method \ Allowed Error (pixels) | 1 | 5 | 5* | 10 |
| 3x6 Convolution | 46% | 49% | 42% | 9% |
| 3x12 Convolution | 45% | 30% | 47% | -32% |
| 3x18 Convolution | 12% | 1% | 38% | -42% |
| Canny | 100% | 53% | N/A | -6% |
| Sobel | 66% | 55% | N/A | 6% |

*No small object suppression is applied

Table 7.3 Performance table of Canny, Sobel and bespoke convolution rail edge detection methods on 20 lower quality image test set.

| | Calculated Success Results | | | |
|---|---|---|---|---|
| Method \ Allowed Error (pixels) | 1 | 5 | 5* | 10% |
| 3x6 Convolution | 55% | 15% | 65% | 15% |
| 3x12 Convolution | 100% | 65% | 55% | 60% |
| 3x18 Convolution | 10% | -45% | 75% | -45% |
| Canny | 40% | 20% | N/A | 5% |
| Sobel | 100% | 40% | N/A | 0% |

*No small object suppression is applied

Table 7.4 Performance table of Canny, Sobel and bespoke convolution rail edge detection methods on 5 higher quality image test set.

The Canny and Sobel methods of edge detection provide better results than the proposed convolution filter in the initial stages of investigation. The robustness of these algorithms however, is less desirable. For both Canny and Sobel methods, the most accurate results are displayed where only 1% error in position is allowed, when this error is increased to 5 and 10% the accuracy soon drops off and the Hough transform sees a stronger false classified edge. This

would lead to the assumption that the Hough line evaluation method detects several lines around the rail edge area preventing accurate classification.

Visually assessing the Sobel edge detection results show, the strongest linear features often cover a significant height of the image. These edges are often located inside the running band. This would suggest that the number of potential false positives is high due to the noise from the running band. The Canny method is more randomised with its strongest linear features. While many still consume the height of the image, their position is sporadic meaning that the potential of a false negative occurrence is unpredictable.

Considering this and occluding the Canny and Sobel results, the smaller 3 x 6 bespoke convolution filter provides the best results, as they are far more robust. In Table 7.4 the 3x12 bespoke convolution filter provides a perfect results with 100% positive classification failing to a minimal error of 55%. The results from applying the bespoke linear filter suggest that the rail edge within the image is not a true linear feature. As the larger mask depreciates in p, it would suggest a slight rotational transform must be considered. The rail within the image is likely to be at a slight angle that would require more accurate modelling and the use of larger bespoke convolution filters.

Evaluating the fall in performance when increasing the allowed error of the bespoke convolution filters also supports the conclusion that the performance increase presented by Canny and Sobel methods are not true. The ability of the bespoke convolution filters between 1 and 5% error shows a robust classification method. The 3x6 filter shows an increase in performance when a larger deviation from the manually assed edge location is allowed in the detection of the rail edges.

The execution time for each method was applied to 5 separate images from the set of 20 lower quality images (data set 2). The average execution time was the taken for comparison, the results are presented in Table 7.5. Canny and Sobel methods are faster in their execution than applying larger convolution filters. This can be expected as the number of mathematical operations

required increase with the larger filters. When applying a smaller 3x6 convolution filter the execution time required is reduced, this difference is related to the application of the Hough line detection. Execution time is significantly affected by the total number of possible lines classified. Through further investigation, it was found that the number of lines classified and that required suppressing when applying Canny or Sobel methods was greater that those generated when applying the 3x6 convolution filter. This further supports the conclusion that the Canny and Sobel methods are not as accurate as initially presented.

| | Method Execution Time(Error @ 1%) (ms) | | | | |
|---|---|---|---|---|---|
| Test number | 3x6 Convolution | 3x12 Convolution | 3x18 Convolution | Canny | Sobel |
| 1 | 1991 | 4098 | 7234 | 2918 | 4804 |
| 2 | 2836 | 4710 | 7481 | 3573 | 4737 |
| 3 | 2764 | 4639 | 7877 | 3905 | 4360 |
| 4 | 2908 | 4946 | 5280 | 3859 | 4852 |
| 5 | 2872 | 5065 | 7972 | 3738 | 4579 |
| Average Time | 2674.2 | 4691.6 | 7168.8 | 3598.6 | 4666.4 |

Table 7.5 Comparison of execution time of edge detection algorithms and different convolution filters.

### 7.2.4 Data Analysis

There is the further analytical approach to locating the rail data within the image. It assumes that the rail edge data is isolated from any noise present within the image. Analysis of individual lines from the image provides information that supports this claim. The running band data within the image is unique, due to its high reflective nature the running band data is formed by high intensity values far greater than the surrounding ballast.

The rail edge will always be within a certain distance of the running band. There should be relevant edge cues that show this rail edge. If the edge cues are large enough then the rail edge can be determined statistically for each line within the image. Figure 7.6 shows the position of the head of the rail edge cues. These are determinable via troughs within the plotted data. There are inherent difficulties in filtering out the alternative troughs that occur within the data.

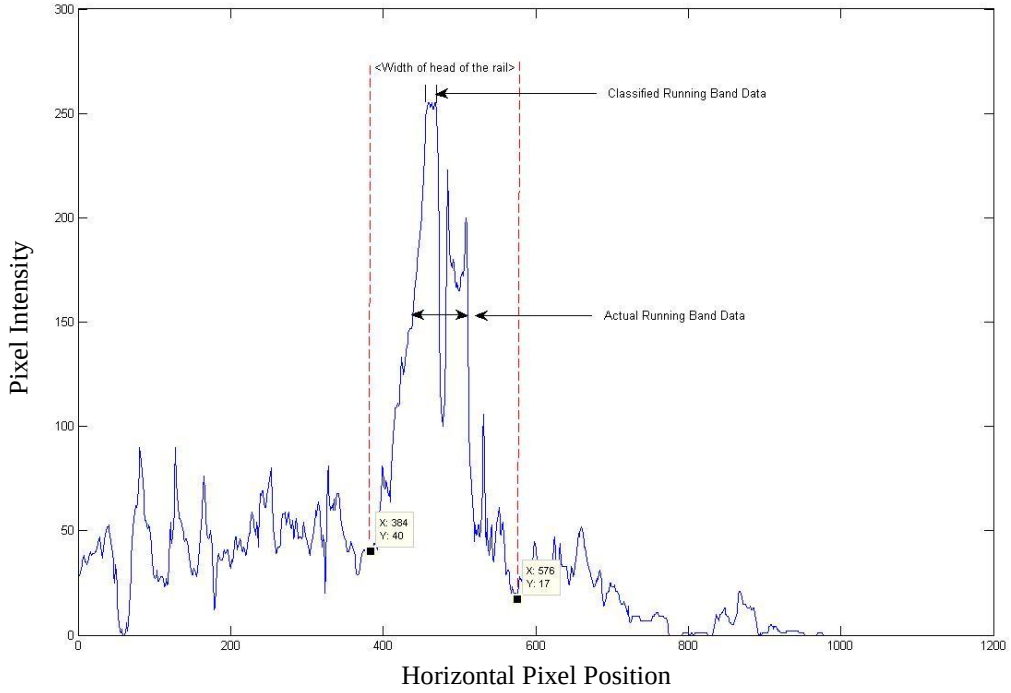



Figure 7.6 Plot to show the position of the head of the rail edge cues and running band within the 1st line of image data.

Figure 7.6 shows that the edge cues from the head of the rail are not directly determinable from alternative edge cues or noise. The only constant is the distance between the left edge cue and right edge cue that will always fall within a certain range. Figure 7.7 shows the number of edge cues that are of an equal distance apart to the width of the railhead. This is representative of the number of false rail edge positives that could occur within a single image. Table 7.6 (*page* 170) shows the tabulated results from the detection of head of the rail edge cues. It shows that on average five possible classifications for the head of the rail is possible.

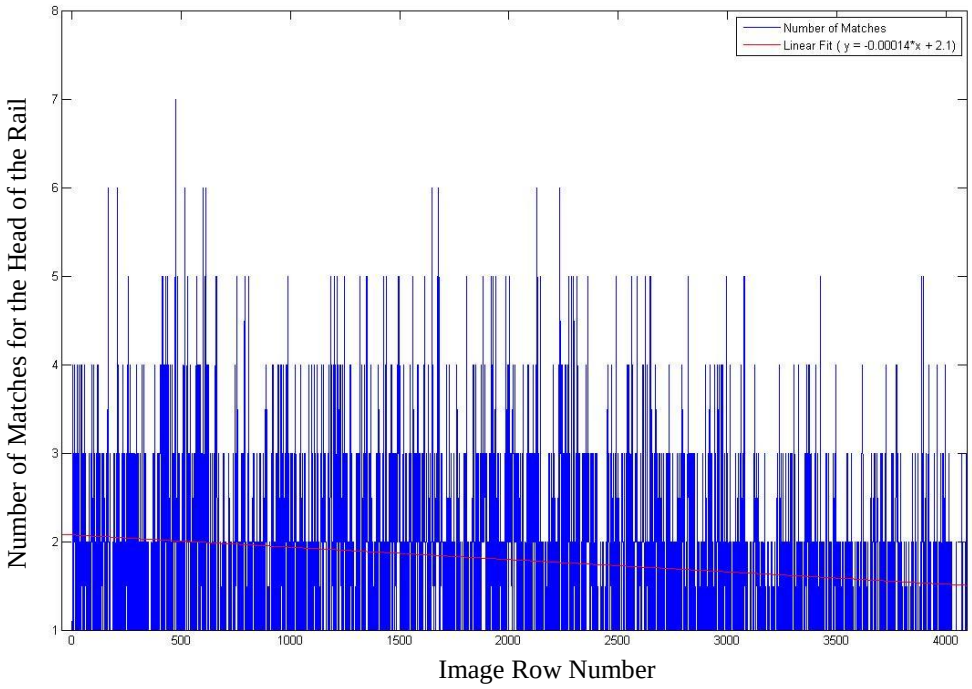



Figure 7.7 Plot to show number of possible classifications for the head of the rail over the height of an image (1024 * 4096).

Smoothing of the image data could be considered to remove some of the false classifications or rail edge cues. This would eliminate edge cues of smaller gradients; however smoothing has its limitations. The risk of eliminating the rail edge cues that are to be located is also high. In practice due to the head of the rail edge cues being of a non-distinct nature this does occur, but the algorithm can be re-enforced. Using previous reference points of the head of the rail edge, any failures to classify the correct rail edge could be overwritten with a statistical average of its most likely point within the image.

This is not overly desirable, however if the edge cues of the rail head can be determined for the overall sum of each column of the image, then errors caused by false positives can be reduced. Figure 7.8 shows the plot of the sum of each column, and the position of the railhead edge cues. As can be seen the location of the railhead is distinct, the right hand side edge is a prominent trough while, the left hand side average produces a peak. While looking for a peak and trough of

a certain distance apart would allow classification of these points, the same approach cannot be applied to each individual line. Previous results have shown that in some cases a trough represents both sides of the railhead edge.

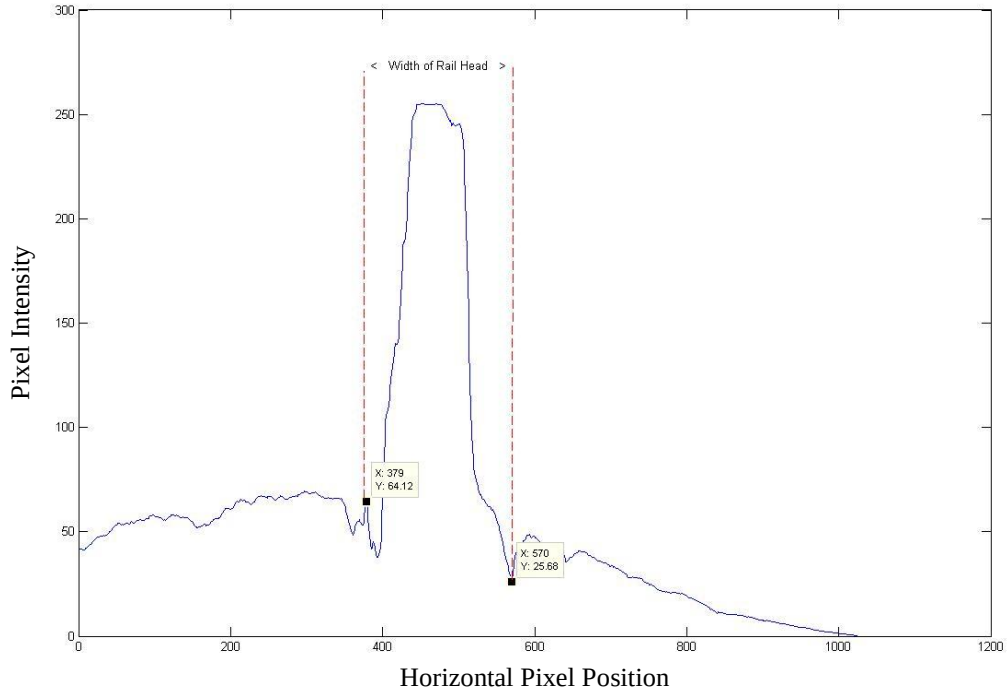



Figure 7.8 Plot to show the position of the railhead edge cues on the average of an image's rows.

Figure 7.8 also shows us that the foot of the rail's edge cue is distinct however, as with determining the railhead edge cues there are difficulties. If this edge could be detected, then the head of the rail would occur a fixed distance away from this point. Looking for this particular feature is not plausible as infrastructure such as Pandrol clips and ballast hide this feature at regular intervals. From observation the infrastructure and associated ballast covering this feature will not allow accurate classification (Figure 7.2 [*page 155*] Figure 7.3[ *page 158*]). A distinct problem in the distribution of the histogram is presented in Figure 7.8 with the left hand side of the image showing a larger average intensity from the surrounding infrastructure.

The image data also shows distinct characteristics related to the mounting position of the camera, while over the head of the rail the foot is far more visible on the right hand side of the image. By improving lighting and camera position, to provide an equal view of the left and right hand sides of the foot of the rail the quality of results may improve. Figure 7.9 is taken from a higher resolution image, with improved lighting, and a camera position direct over the head of the rail. The plot is based on the same principle as presented in Figure 7.8. It can be observed that the edge cues of the head of the rail are clear troughs.

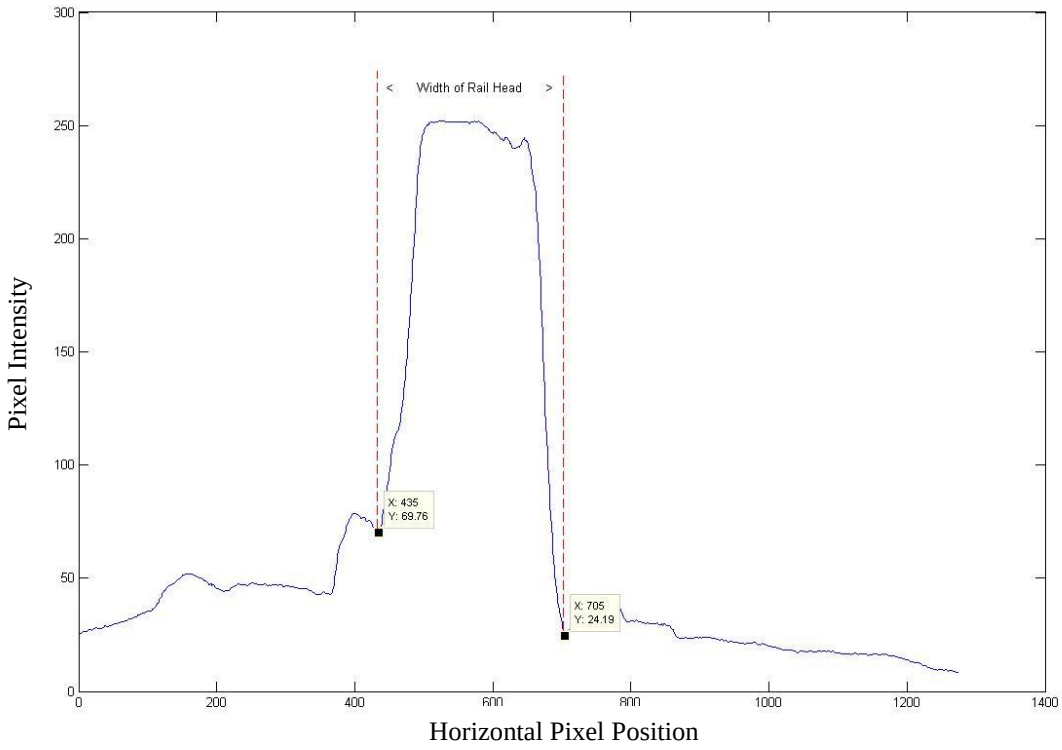



Figure 7.9 Plot to show the position of the head of the rail edge cues on the average of an higher resolution image's rows.

Despite the apparent improvements, the number of false classifications has increased with resolution. This is likely due to the added noise causing false edge cues within the images. Again, smoothing could be applied to improve results and statistical rules could be developed to discard the small edge cues. Table 7.6 compares these results to those achieved by applying the same data analysis on a lower resolution image.

| | Number of Possible Heads of Rail per Image Row |
|---|---|
| Min | 1 |
| Max | 14 |
| Mean | 5.55 |

Results from testing on lower resolution image

| | Number of Possible Heads of Rail per Image Row |
|---|---|
| Min | 2 |
| Max | 30 |
| Mean | 12.08 |

Results from testing on higher resolution image

Table 7.6 Comparison of railhead edge detected in low resolution images and high resolution images over the height of the image.

While the higher resolution image results in more false classifications, it is plausible that by introducing further statistical analysis, of edge cues, that this error could be reduced. The same is achievable with the lower quality image. Further data analysis would be responsible for examining each edge cue detected and eliminating the weaker cues. This could be strengthened by referencing the sum of the image's rows for an ideal position. This analysis could also reduce the amount of data examined for edge cues and improve accuracy to a desirable amount.

### 7.2.5 The use of Back-propagation Neural Networks for Railhead Edge Detection

Back- propagation Neural Networks (NN) provide the most adaptable results from all systems discussed. NNs are robust and accurate given a large enough data set. NN have been used for edge detection purposes and challenge all current methods of edge detection (Berend Jan van der Zwaag, 2002).

A mask of data e.g. 3x3 pixels is taken from the image, fed into the NN and a chosen output is given to indicate a required edge or a region with no edge, the more details used within the training set data the better the results will be. The settings of the NN are as important as the organisation of the training data. This is covered extensively by Lecun, Bottou, Orr and Muller (1998) who show that by following a few simple steps the effectiveness of edge detection from a NN can be improved. Appendix 4 shows a simple example of NN capabilities in edge detection.

The only inherent drawback of NN systems is the cost in execution time and memory commitments. To train a NN on a large set of data a machine must have enough memory to store the NN as well as the training data set involved.

The execution time of analysing a larger image can also be costly in implementing a NN. Each pixel needs to be analysed and data surrounding it, if applicable, needs to be fed into the NN. The NN then has to calculate the output, the more complex the NN and the larger the data sets involved the longer the execution time.

A NN can be applied to this problem in two ways. The first to examine for edge cues by sliding a mask over the image. This is similar to the standard edge detection where the surrounding data of each pixel is considered to examine for a set edge cue. As discussed with edge filters the size of this mask can be adjusted to favour the longer linear features of the rail edge reducing the effects of noise. Edges such as the foot of the rail and running band will still be detected. The benefits of using a NN is that the edge of the head of the rail appears to be detected as reliably as the

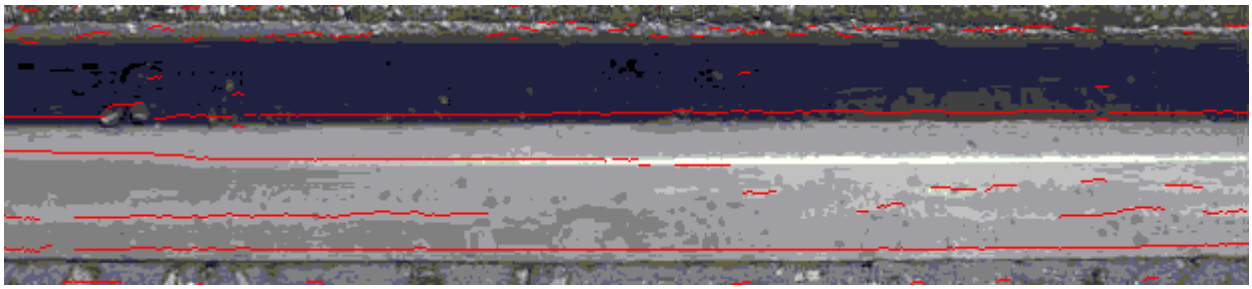
running band and a mathematical rule could be developed to determine the foot and head of the rail edge by the related distance between each point.

The alternative method is applying a NN to solve the mathematical problem forming a statistical rule. Training a NN on selected set points of data relating to the location of the running band will allow a NN to determine the position of the head of the rail. A NN will be able associate points such as the edge cues from the foot of the rail and running band in helping determine the correct location of the head of the rail. A NN will solve the mathematical problems involved while adapting for noise and slight variations in the rail edge location.

A NN applied to solving a mathematical rule can have a significant performance advantage over the application of a sliding mask alternative. There are significant performance advantages in using a linear set of data, rather than a matrix. Less rearrangement of data is required to achieve a data format acceptable for a NN input. The size of the linear data set is often smaller in comparison. This is dependent on the quantity of data used for classification of the edge location in comparison to the size of a sliding mask required for similar classification. In such an application, the number or processing iterations is often reduced.

Figure 7.10 shows the output from a NN trained to detect the rail edge of embedded track. A mask size of 15x15 pixels was used and the NN was trained with 262 positive edges 222 negative edges. The larger mask size was required to reduce noise from smaller edge cues. The results are comparable to standard edge detection methods. There is an offset of 15 pixels from the detected edge. As can be observed false edges such as small rocks and parts of the polymer interfere with the edge detection, further post processing can potentially reduce the effects these false classifications may have on analysis of the rail edge.

Original Image



Results from NN

Figure 7.10 Results from a neural network in the application of rail edge detection

## 7.3 METHOD EVALUATION AND OBSERVATIONS

Standard edge detection and statistical analysis are both methods that have potential to be developed into successful algorithms. Their robustness however is limited with simple changes in lighting conditions, noise, or camera types affecting their ability to correctly classify the edge of the head of the rail. It is this lack of robustness that prevents their sole use without some form of image pre-processing.

Back- propagation Neural Networks are incredibly robust to both lighting conditions and noise. They can potentially overcome changes in camera type assuming FOV remains relatively unchanged. Their design and training cost however outweigh the usefulness in this particular application. The cost of re-designing the training set if the FOV changes considerably, with the potential of recreating a NN that cannot perform as well, leave a NNs reliability much lower than required for this application. While NNs are ideal for complex situations where specific algorithms cannot be mathematically solved, research has shown that filters and templates have an equivalent potential to correctly classify rail edge cues with a smaller design time.

The chosen route for the investigation in the application of, linear edge detection in the classification and analysis of tramway rail edge was the use of bespoke convolution filters with data analysis, in conjunction with template matching methods. The rationale for this choice is that while the performance of a bespoke convolution filter is an improvement on edge detection methods, it can be greatly improved by reducing the region of interest within the image. To allow automated cropping of the area in which the railhead is likely to appear the use template matching is ideal (Appendix 5)(Lewis, 1995). Edge cues can be tailored for detection within the acquired images by using results from template matching to minimise the region of interest, substantially reducing potential noise and false classification. The data analysis will be based on using previously classified edge cues to predict the potential location of following edge cue, within the next row of the image.

### 7.3.1 Method Evaluation

Evaluation of the chosen algorithm is completed in two stages, one using smaller data sets, the 'initial evaluation', and another with a broader unknown data set, the 'final evaluation'. This approach allows the algorithms robustness to be evaluated on a smaller scale. This method of evaluation is chosen to aid algorithm development during testing.

#### 7.3.1.1 Initial Evaluation

A set of 20 lower quality images forms the first test set Data set 2, while 5 higher quality images from a distinctly different MVS forms the second test set Data set 3, see chapter 6.2 Data Sets (page 142) for reference. The initial investigation focused on the template matching of the head of the rail. With both test sets, the first image within the set is used to form the template data. This template was then applied to the test sets using two template-matching algorithms. The first used the whole head of the rail data to form one template to match the railhead in each image. The second used data forming the left and right individual edges.

The most successful algorithm for matching the edge cues for the head of the rail was then used to minimise the image data. The ideal resulting image contained the railhead and its relevant edge cues. A bespoke convolution filter was then applied to classify these edge cues. A strong line following algorithm then traces the strongest matched edge through the image.

The results of the template-matching algorithm were visually evaluated. Images were overlaid with the classified image edge cues and the accuracy of the results, once visually assessed, were assessed to examine for variations in the width of the rail. A width threshold of ±10 pixels, or 1.6mm, was used to determine classification ability. If the width of the railhead varied more than this a detailed visual inspection was carried out looking for a failure in edge tracking. The total variation across the height of the image was accumulated to give an insight into the accuracy of the edge cue classification and line following algorithm.

#### 7.3.1.2 Final Evaluation

Once confirmation of the technique was achieved, the process was applied to data set 4 of 5000 images. These images were acquired from the Sheffield Supertram network using the finalised specification for the acquisition system described in Chapter 4. Variations in the rail head width and the tracked edge cues starting and finishing position within the image are used to determine images that require visual inspection. A further random 100 images were chosen as a control set to offset results, looking for possible false negative failures.

## 7.4 RESULTS: TEMPLATE MATCHING WITH BESPOKE CONVOLUTION EDGE CUE CLASSIFICATION

### 7.4.1 Template Matching

The use of template matching methods often correlate with that of convolution filters as features such as edges and corners can be extracted to improve detection. Convolution filters are often required when applying template matching methods such as cross correlation. The difference between this FFT template matching methods is that a filter forming the whole object can be used rather than a filter designed around suppressing non-matching features.

Execution speed is of particular importance even though images can be post processed after image acquisition. Acquiring images from the Sheffield Supertram network for example would produce around 1.2 million images per rail (left and right). A processing time of 1second per image would require over 333 hours of processing. If the whole of the UK rail network was processed it would require 188900 hours or 21.5 years.

The linear convolution filter is designed to show all edge cues within the image according to a certain direction and over a certain size. A template matching method will remove the added steps of finding the location of the correct edge within the image. Instead, the majority of the edge cues present within the image should be those associated with railhead. Using algorithms such as cross-correlation calculations on a large image matrix can be computationally expensive; therefore, the Fast Fourier analysis method discussed in the literature review will be investigated. (Du-Ming Tsai, 2003)(Lewis, 1995).

Using cross correlation for template matching is costly, as the translation and rotation of an object within an image must be accounted for. For each possible translation or rotation of an object, a filter must be applied to each supporting pixel of the image where the object may occur. Equation 7.3 shows the calculation of computational executions for any given filter. A number of filters can be combined, trading a small amount of accuracy for a quicker execution of the

analysis. For the application of rail edge detection the number of transformations and degrees of freedom can be considered to be 1.

$$\left(\left(\frac{T * M}{C}\right) * N\right) * A = E_{tc}$$

Where:

T = Number of possible transformations
M = Number of rotational degrees of freedom (Max 3)
C = Number of filters that can be combined
N = Number of origin pixels for examination
A = Size of filter (Height * Width)
$E_{tc}$ = Total mathematical computational executions

Equation 7.3 Computational loops for convolution/cross correlation template matching algorithm.

To reduce the number of computational loops, the number of origin pixels can be greatly reduced in this application, this is achieved by extracting key features using convolution filters. Since the cross-correlation calculation must be performed on each pixel, reducing the image size and filter accordingly can speed up application. This reduces accuracy and has the potential for a knock on effect in other algorithm calculations.

Using a Fourier analysis method does not reduce the need to perform operations on each pixel however; the process is faster due to hardware implementation on computer processors. A Fourier method removes the number of computational loops required (Equation 7.4, Equation 7.5). It has less of an impact for this particular application as the transformations and rotational degrees of freedom are minimal. It does however remove the factor of origin pixels by retaining this value to 1.

$$\gamma_{(u,v)} = \sum_{x,y} f'(x,y)t'(x-u,y-v)$$

Equation 7.4 Mathematical operation for Cross Correlation* for template matching (R. C. Gonzalez, 2004)

* Assumes mean value has been subtracted.

$$\gamma_{(u,v)} = F^{-1}\{F(f') * F(t')\}$$

Equation 7.5 Mathematical Operation for Fourier transformation for template matching (R. Brunelli, 1993)

The resultant theory concludes that the execution time for the cross-correlation of each pixel within the image to a template will inherently require greater execution time than that of the Fourier Transform alternative (Equation 7.6). Therefore, for template matching applications the Fourier Transform methods should be applied in favour of the alternative methods.

$$E_{tc} * t\left\{\sum_{x,y} f'(x,)t'(x-u,y-v)\right\} > A * t\{F^{-1}(F(f') * F(t')\}$$

Where:

$E_{tc}$ = Total Mathematical Computational Executions (Equation 7.3)
t = time
A = Size of Image (Height * Width)

Equation 7.6 Theory in the conclusion for the greater execution time of cross-correlation in comparison to the Fourier transform application.

Table 7.7 shows the results from timing both methods. The cross-correlation was not applied to every pixel within the image and only to initial row, as this would provide and accurate matching of the railhead location within the image. The output results are therefore comparable to those presented by the FFT template-matching algorithm. A performance increase of x11.7 is observed, showing that the conclusion was justified, as there is significant decrease in execution time when applying an FFT template-matching algorithm in comparison to cross-correlation template matching algorithm.

| | Method Execution Time (ms) | |
|---|---|---|
| | FFT | Cross-Correlation |
| Average | 1112.2 | 13043.22222 |

Table 7.7 Execution time comparison of FFT VS cross-correlation template matching methods.

#### 7.4.1.1 Fast Fourier Transform (FFT) Template Matching Results

To produce results that are can be compared two test were applied using the FFT template matching method. The first uses a template of the complete head of the rail matching both the left and right edge features. The second individually matches the left and right edges. The second test is more comparable to the methods applied in the bespoke convolution filter. The Hough transform applied in the bespoke convolution filter investigation only takes into account the relative position of the left or right rail edge. It does not consider the distance between these; this potentially prevents false positives that occur due to running bands.

The results presented in Table 7.8 show that applying a FFT template matching method for the detection of the head of the rail within the image is highly adaptable to different image types. The limitation of this method is that the high success rate only occurs when matching all features of the railhead such as the edges and the running band(s) present. In the second investigation in which each edge was examined for individually, the success rate drops with several false positives occurring due to running band interference. As with a bespoke convolution filter, using the distance between the classified edges has the potential to reduce the number of false positives.

| | Success Rate (%) | |
|---|---|---|
| Method/Test Set | Lower Quality | Higher Quality |
| Whole Rail | 100 | 100 |
| Each Edge | 35 | 40 |

Table 7.8 Performance table of FFT template matching methods on lower quality and higher quality image test sets

Results presented in Table 7.9 show that the most successful method of matching the whole head of the rail within the image incurs the largest execution time. This is to be expected due to the scale of the template data being matched. If the template size is reduced, the execution time is also reduced. This is demonstrated with the method of matching each rail edge individually that

requires a pass of the data for both templates. As each template is smaller in size than the data forming the whole of the railhead, the execution time is reduced.

| Test number | Whole Rail (ms) | Each Edge (ms) |
|---|---|---|
| 1 | 10930 | 8581 |
| 2 | 10698 | 8510 |
| 3 | 10885 | 8599 |
| 4 | 10689 | 8618 |
| 5 | 10883 | 8778 |
| Average Time | 10817 | 8617.2 |

Table 7.9 Table of execution time for FFT template matching methods for the head of the rail and each edge templates on the lower quality image set.

### 7.4.2 Bespoke Convolution Filter

Using FFT template matching, the position of the head of the rail can be estimated. The FFT template-matching algorithm does not classify rotation of the head of the rail within the image. Therefore, the true position of the edge cues along the length of the image must be evaluated as the edge cues are not always situated at a true vertical angle.

When using a convolution filter, features such as the foot or running band produced interference, causing false classification of rail edge cues. Using the estimated position, provided by template matching, the suppression of unwanted linear edge cues can be achieved. These local regions, containing the estimated data for the left and right rail edge cues can be evaluated for linear edge cues with the use a bespoke convolution filter. The convolution filters applied were the same as discussed in Chapter 7.2.3.

Figure 7.11 shows the results of applying the convolution filter, after suppressing data not associated with the estimated position of the head of the rail. As can be observed, the edge cues are detectable; however, methods for the classification of the correct edge cues are still highly susceptible to noise. Two methods were evaluated; the first method examined the localised regions for the highest potential edge. This method did not have any line following logic built into it a followed the rule presented in Equation 7.7. The equation was applied on the localised edge cue regions, left and right, individually.

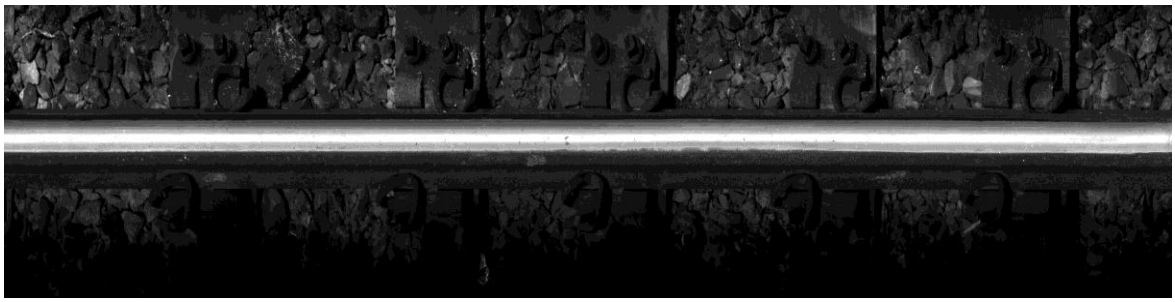
$$EC = \sum_{x=0,y=0}^{m,n} I_{(x,y)} > T$$

Where:

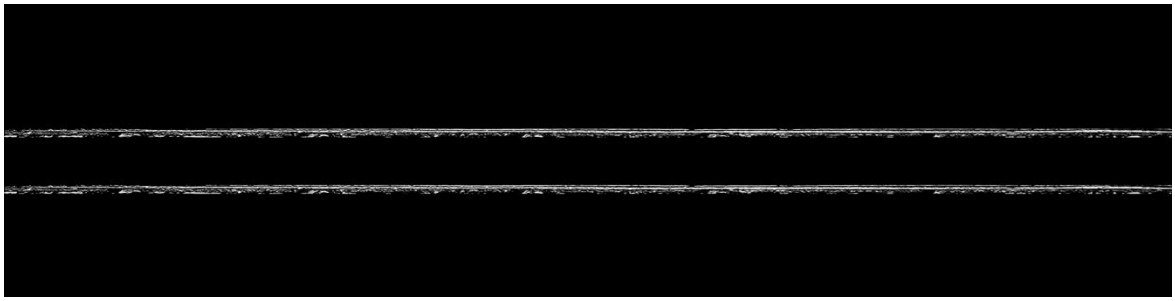
$EC$ = Edge Cue
$I$ = Image Pixel Being Examined
$T$ = Threshold

Equation 7.7 Equation to classify head of the rail edge cue

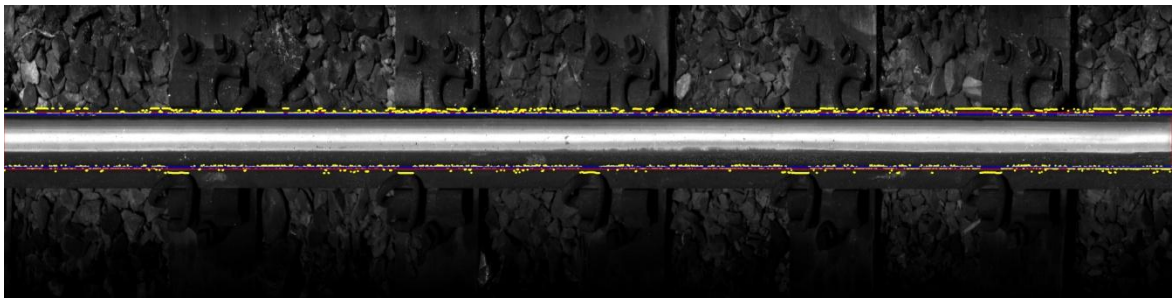
The second method employed the Hough transform for the detection of linear features contained within the localised edge cue regions (Hough V, 1962). Similarly, the method was applied to the left and right regions individually. Only the highest line feature was accepted and was assumed to be the classified railhead edge.



Original image



Data suppression and convolution filter results



Tracked convolution results applied to image

Figure 7.11 Example of the results from data suppression and bespoke convolution filter in detection of rail edge cues (evaluated to be the best approach).

### 7.4.3 Initial Evaluation

The results from classifying the rail edge are presented in detail in Appendix 5 where the classification accuracy for determining the rail width is evaluated. Both applied methods are evaluated on their classification of the head of the rails width within the image. Both the average calculated values and the variance of the results are inspected. While the average shows the potential error in pixels, the variance relates to how much noise in evaluating this feature measurement can be expected. Details of the evaluation method are discussed in chapter 7.3 Method Evaluation and Observations (*page 174*).

Results from both test sets show an improvement in performance when a line following algorithm is applied over that of the Hough transform. Table 7.10 shows the overall results from Appendix 5, the tables examine the calculated width of the railhead. It looks at the width calculated from the line following algorithm and the Hough transform method. The results are compared to a visually assessed width. The 1st and 3rd column present the average measurements, a large difference can be seen from the classified Hough transform edges and those from the line following algorithm. The 2nd and 3rd column show the variance of the measurements and which permits the relative performance of the algorithms to be evaluated. Table 7.10 shows that the line following algorithm is more accurate and robust than applying the Hough transform method.

| Variable | Test Set 1: Average | Test Set 1: Variance | Test Set 2: Average | Test Set 2: Variance |
|---|---|---|---|---|
| Line following Average Width (pixels) | 195.2 | 2.56 | 505.8 | 122.24 |
| Hough Width (pixels) | 182.7 | 5.64 | 591.67 | 373.89 |
| Manually Assessed Width (pixels) | 195 | 0 | 500.8 | 116.68 |
| Diff of Manually Assessed VS Hough | 12.3 | 5.64 | -31 | 274.83 |
| Diff of Manually Assessed VS Average | -0.2 | 2.56 | -5 | 12.58 |

Table 7.10 Comparison of bespoke convolution performance against Hough transform performance in the application of railhead edge cue classification.

The Hough transform works more proficiently after non-maximal suppression is applied to images in which convolution has been applied. This relates to the successful application of the method with edge detectors such as Sobel in other applications. There is potential to improve the results from the Hough transform by altering the pre-processing of the image. The results show however, that the line following algorithm employed has a higher performance without the added stage and potential complexity of evaluating image variants, such as light changes within frames and infrastructure present.

The Hough transform presents advantages in test set 1 by classifying the inside edge cue. This results in a smaller value being recorded for the rail width. This allows only head of the rail data to be successfully segmented for further analysis in all images from this test set. The Hough transform however, struggles with the higher quality test set, failing to classify either right or left edge cues on two images. This is due to the higher resolution of the image, edge detection techniques classify an edge other greater width of pixels. This prevents the Hough transform from evaluating the edge cues detected as a continuous line feature. Image size could be reduced to account for this, subsequently effecting accuracy of measurement that can be achieved from the image, the loss of accuracy is not seen as acceptable.

While the smaller measurement prevents potential false classification of running band data it prevents accurate measurement of the head of the rails features. The results from the convolution filter are more desirable however determining the head of the rail edge cues from the foot of the rail can potentially cause false classifications. The results show that by adding an offset of 6 pixels to the classified width greater effectiveness can be obtained in the classification of the running band data. This offset will not affect railhead measurement, as it will only be applied when segmenting the running band data.

### 7.4.4 Final Evaluation

Having evaluated the algorithm's robustness and, finding it to be acceptable in its ability to classify the rail edge, the algorithm was applied to a larger 5000 image data set 4. This data set is split into 5 sets of 1000 images, each taken from separate section of the Sheffield Supertram network. Results are presented in Table 7.11 showing 8.4% of the images require visual assessment through the potential of having a defect present. Details of the evaluation method are discussed in chapter 7.3 Method Evaluation and Observations (*page 174*).

| Sub Data Set Number | Number of Images | Classified Errors | Percentage Error (%) |
|---|---|---|---|
| 1 | 1000 | 35 | 3.5 |
| 2 | 1000 | 27 | 2.7 |
| 3 | 1000 | 187 | 18.7 |
| 4 | 1000 | 71 | 7.1 |
| 5 | 1000 | 100 | 10 |
| Total | 5000 | 420 | 8.4 |

Table 7.11 Classified edge cue results from final evaluation data set.

Table 7.12 shows the results from visual assessment of 420 classified errors. There were no true errors found. The causes of the false positives, through visual assessment, were found to be:

- Running Band Interference: Edge cues from the running band being classified falsely as the rail edge cue. This was caused due to the edge cues being a stronger feature within the reduced image were assessment for the rail edge cues failed. Occurrence was on the RHS of the railhead, were the running band approached the railhead edge.

- Loss of Rail Edge Tracking: Caused by interference from debris surrounding the railhead. Often there was no statistical difference between the expected location of the edge cue and the surrounding debris. This makes rail edge classification impossible.

- Infrastructure Interference: Within sub data set 3, the presence of a fishplate produced several false positives. This could be eliminated by post-processing, when fishplates are examined for.

| Set Type | Number of Images | True Positive | False Positive | True Negative | False Negative |
|---|---|---|---|---|---|
| Classified Errors | 420 | 0 | 420 | N/A | N/A |
| Control Set | 100 | N/A | N/A | 100 | 0 |

Table 7.12 Results from visual assessment of errors and control set.

The control set results show that in all 100 images examined, the railhead was successfully classified and width evaluated correctly.

## 7.5 SUMMARY

The results presented show that the use of edge cue analysis in determining the railhead position within the image is not effective. Reduction of the region of interest (ROI) removes noise caused by ballast and rail infrastructure improving performance. Applying the Fourier template matching method allows the ROI to be calculated for each frame. Results confirm that with the reduced ROI edge cue detection can allow the successful classification of the railhead edge.

The results demonstrate that the 'traditional' edge detectors (Canny (Canny, 1986) and Sobel (I. Sobel, 1978)) do not perform as productively when applied to high-resolution images. The high resolution affects the 'traditional' edge detectors ability to classify edge cues; this is due to edges being present over larger areas than what 3, 5 or even 7 pixels mask can classify. While it is accepted that image size can be reduced allowing for their successful application this significantly reduces the accuracy of measurements taken within the image.

Alternatively, these methods can be built upon to produce bespoke convolution filters that work for the larger edge resolution available. The performance in classification of the railhead edge cues and the investigated execution time difference between Sobel and Canny edge detection methods in comparison to that of a bespoke convolution filter show, that bespoke filters for object features should be applied and that the 'traditional' edge detection method should only be used for testing purposes.

For specific object detection in which features such as edge cues and corners need to be located, bespoke convolution filters should be applied. This is targeted to MVS in which large execution time is considered detrimental to overall performance. In the application of large object detection with strong and consistent feature characteristics and minimal transformation rotation of objects then Fourier Transform methods can be applied. Reducing the image size to decrease execution time and performance of 'traditional' edge detectors can be detrimental to accuracy of measurements and should be avoided.

In the classification of the railhead edges, a bespoke convolution filter provides the best performance in classification of railhead edge cues. In an evaluation of the results, using simple line following algorithm to classify the best potential edge cue has proven more efficacious and robust than that of applying the Hough transform method. It is accepted that further pre-processing before applying the Hough transform can improve the results for this method however, as discussed variation of lighting and infrastructure within the image can affect the ability of the pre-processing and whereas the effectiveness of applying a line following algorithm has already been established.

# CHAPTER 8 NOVEL FEATURE MATCHING ALGORITHM FOR THE SUPPRESSION OF TRAMWAY MVS VIBRATION FOR IMAGE STABILISATION

## 8.1 INTRODUCTION

The proposed MVS for tramway track inspection mounts a camera on the bogie of a moving vehicle. This allows some vibration of vehicle to be transmitted through the camera system. Vibration in the vertical direction is not visible due the use of a single line CCD sensor. This is also minimised due to the movement of the vehicle and mounting hardware. Vibration perpendicular to the rail direction is apparent, and requires image stabilisation algorithms to be developed and applied. Figure 8.1 shows the effect of vibration during the image acquisition.

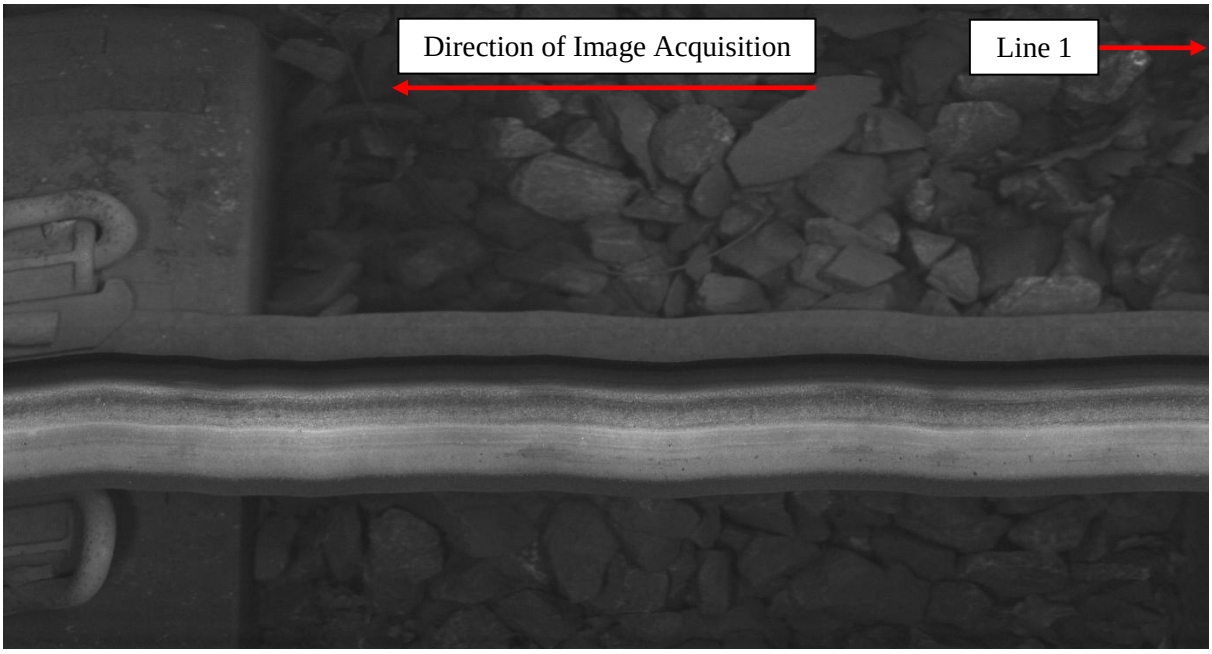



Figure 8.1 Example image acquired with camera vibration

A novel approach has been developed to suppress the vibration within the image. The algorithm segments each individual line of the image in the same fashion that the data is acquired. This data is then examined for repetitive defining features labelled as key points. Such defined key points are produced by the railhead edge and are present throughout each line of the image.

These key points are examined for movement such as that caused by vibration. Each line can then be aligned to the preceded one by calculating the movement of the key points, successfully suppressing the vibration.

The key points classification is achieved by the process described in Chapter 7 “Linear Edge Detection in the Classification and Analysis of Tramway Rail Edge”. In particular, correction of the image is done after the rail edge has been successfully classified. The application of correction allows for the detection of infrastructure such as Pandrol clips to be achieved in future work. As template matching is the most prominent method to achieve the classification of such objects the image correction ensures they have an expected shape for an algorithm to detect.

## 8.2 INVESTIGATIVE RESEARCH

The methodology behind the image stabilisation is similar to the way SURF (Speeded up Robust Features) description template matching is applied (H. Bay, 2006). Locating key points that correlate with a templates key points allows classification of objects within an image. With line scan cameras, the same theoretical method can be applied, where the acquired data has consistency between visual components. In this application, mathematical key points can be calculated for the railhead, a consistent visual component that should be seen through a majority of images.

There are significant problems in the application of image stabilisation in the location and classification of tramway rail defects. Irregular occurring features such as ballast and other rail infrastructure cause differences in key points that will be hard to match between two lines. The railhead is the only remaining feature that is consistent throughout a majority of the acquired image datasets. This allows the data set in each line to be reduced for matching around the railhead; this is demonstrated in Equation 8.1.

$$f\left(\sum_{j=x}^{n} l_1\right) \approx \sum_{i=2}^{m} f\left(\sum_{j=x-y}^{n+y} l_i\right) \Rightarrow O_x$$

Where:

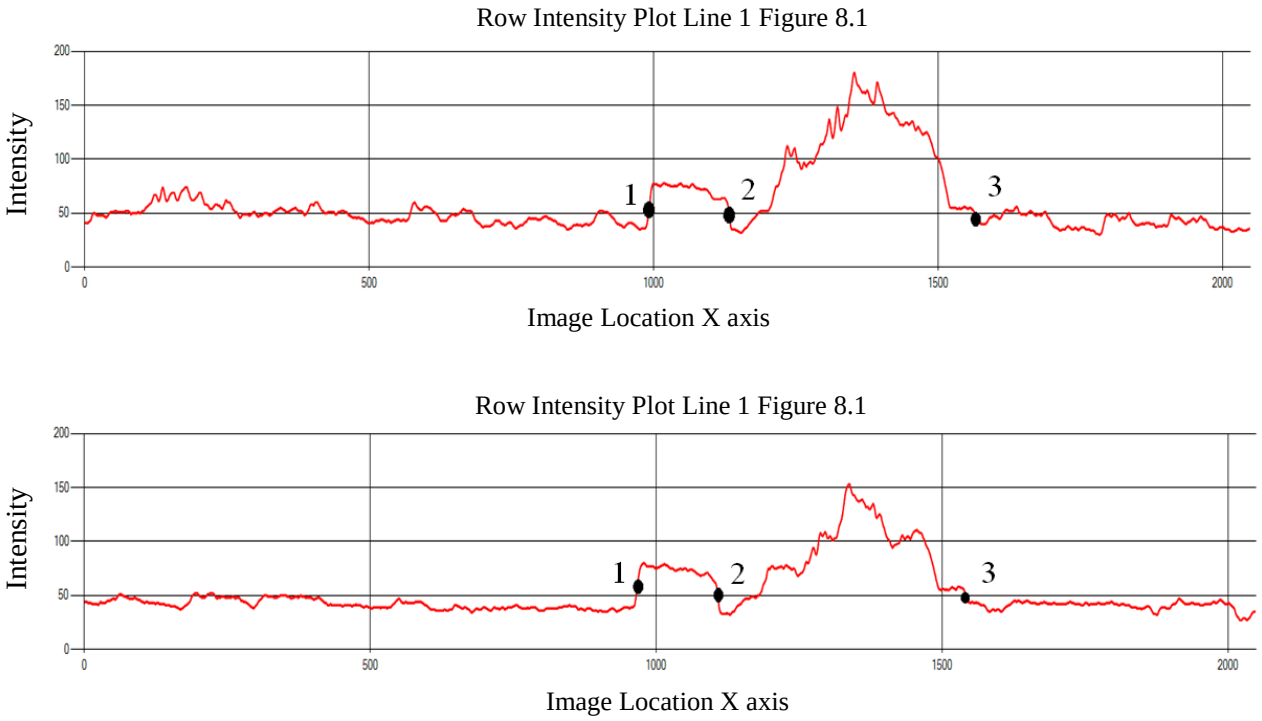
$f$ = Function applied to data
$x$ = Start position within key point data set to begin examination
$y$ = Region of interest offset
$n =$ End position of key point data set
$m =$ Image Height
$l$ = Line data set under examination
$O_x$ = Offset to align features

Equation 8.1 Image stabilisation theory applied to a minimised data set.

The key to good image stabilisation is the function applied in extracting the key points from each line data set. An ideal function will employee noise suppression in extracting appropriate key points, however bespoke pre-processing of the data may be required.

Classification of key points relies on determining a prominent feature that is consistent through a majority of image sets acquired. Key points can be decided upon visually before methods are developed to determine key points throughout the data set. Simple visual assessment shows that the key points are associated with railhead features (Figure 8.2). Since these are edge cues, classifying key points requires the use of edge detection theory.

### 8.2.1 Classifying Key Points



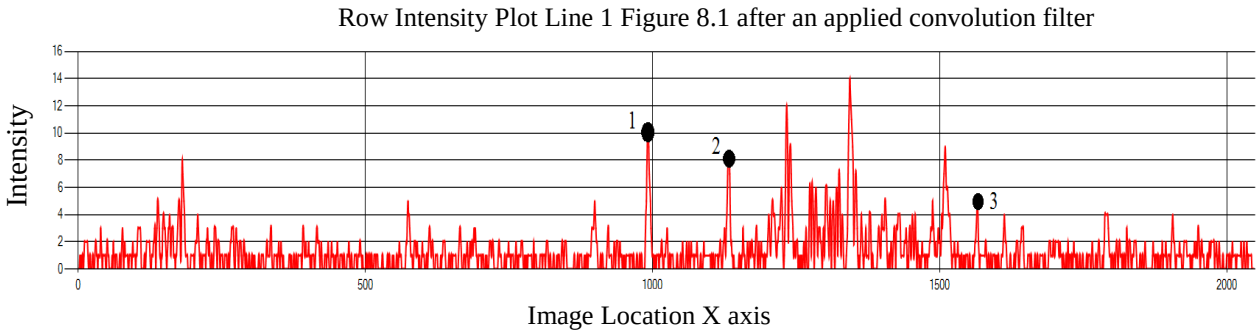

[1] Rail foot edge cue LHS

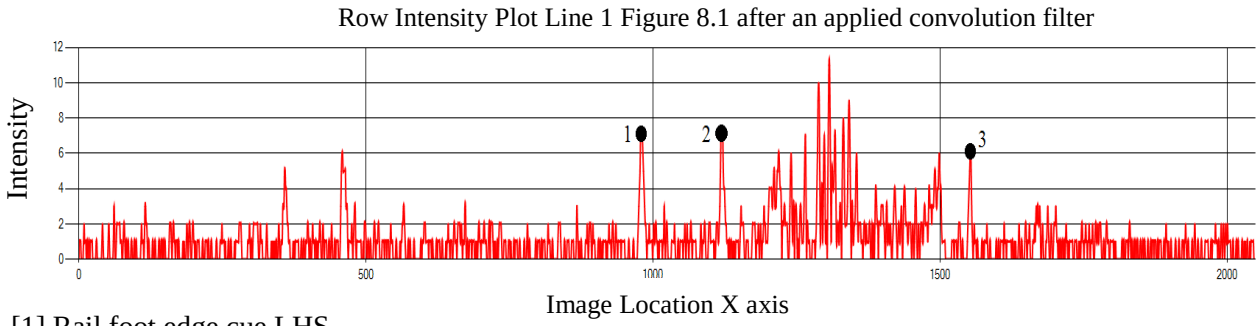
[2] Railhead edge cue LHS

[3] Railhead edge cue RHS

Figure 8.2 Plots to show visually assessed edge cue key points from two image rows.

In edge detection, it is often the application of convolution filters that classify edge cues before further noise suppression is applied. In this case, a bespoke linear convolution filter is applied to the data as described in Chapter 7.2.3 page 160. Figure 8.3 shows how the prominent edge cues of the railhead are visible by eye. However, noise and edge cues from ballast and the running band can interfere.




[1] Rail foot edge cue LHS

[2] Railhead edge cue LHS

[3] Railhead edge cue RHS

Figure 8.3 Plot to show detected edge cue key points through visual assessment.

To eliminate noise, week edge cues must be removed from the data by setting a simple threshold. This will ensure that only strong edge cues remain. These are logically the most likely to continue onto the next image line. Figure 8.4 shows the application of this methodology. A majority of significant edge cues can be observed continuing onto the next line of the image data. It can be assumed that since these edge cues are linear in nature, that if they are aligned to each other by shifting the data, then the vibration can be eliminated in this axis.

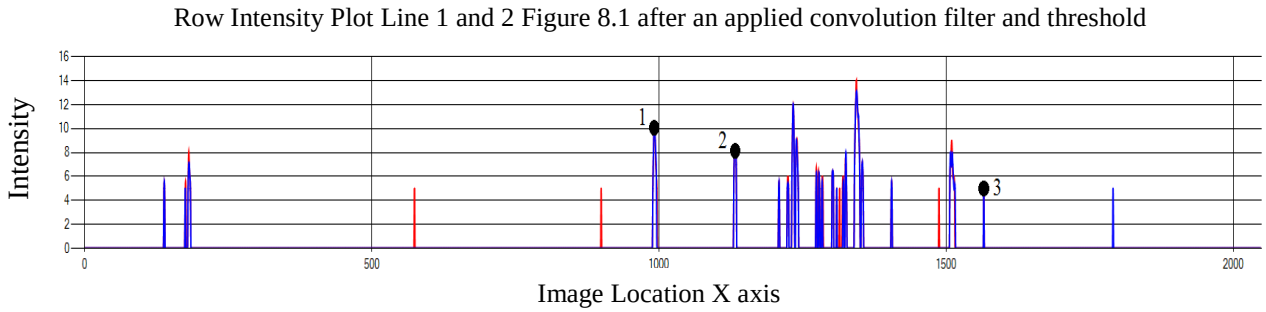



Figure 8.4 Plot comparing detected edge cue key points of line 1 and line 2 of image Figure 8.1 after applied convolution filter and an intensity threshold value of 4 is applied.

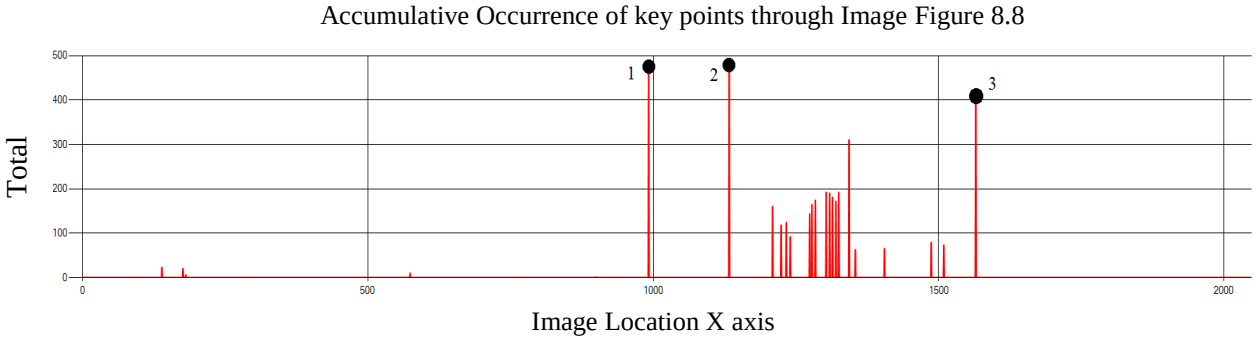
### 8.2.2 Edge Cue Key Point Comparison

Assuming that edge cue key points can be determined successfully from the data, the application of key point matching will eliminate vibration from the image. A comparison between key point data has been developed so that only certain key points are used throughout the data set. As smaller edge cues, caused by noise, have been eliminated by setting a threshold, eliminating key points that are not consistent is achieved by examining data for edge cues that does not continue onto the following lines.

To classify the significant edge cues non-maximal suppression is applied, this minimises the co-ordinates in which key points are most prominent. Each key point is cycled through and compared to the preceding lines results. If the key point's x co-ordinate location corresponds to a previous key point, no vibration between lines is apparent. If they are off by a set amount, they could be demonstrating a direction of vibration through key point movement.

Interference cause by noise and running band edge cues now needs to be suppressed. A key point list of each line is formed and compared to the initial lines key points. Classification of the most prominent features is then achieved by examining how often this key point iterates throughout the data set. Figure 8.5 shows the correlation of the key points through each line of the image.

The three edge cues visually assessed as key points are apparent, being matched throughout most of the image lines.




[1] Rail foot edge cue LHS

[2] Railhead edge cue LHS

[3] Railhead edge cue RHS

Figure 8.5 Accumulative Key point correlation through each line of image data.

By using the minimised list of key points, the location of classified vertical edge cues within each line of the image can be determined. Tracking these key points and their movement can suppress the apparent vibration within the image. The average change in position of the key points is taken and each line of the image is corrected so that the average position lines up through the image. Figure 8.6 shows the tracking of the three key points throughout the image height and, Figure 8.7 the average movement of the edge cues.

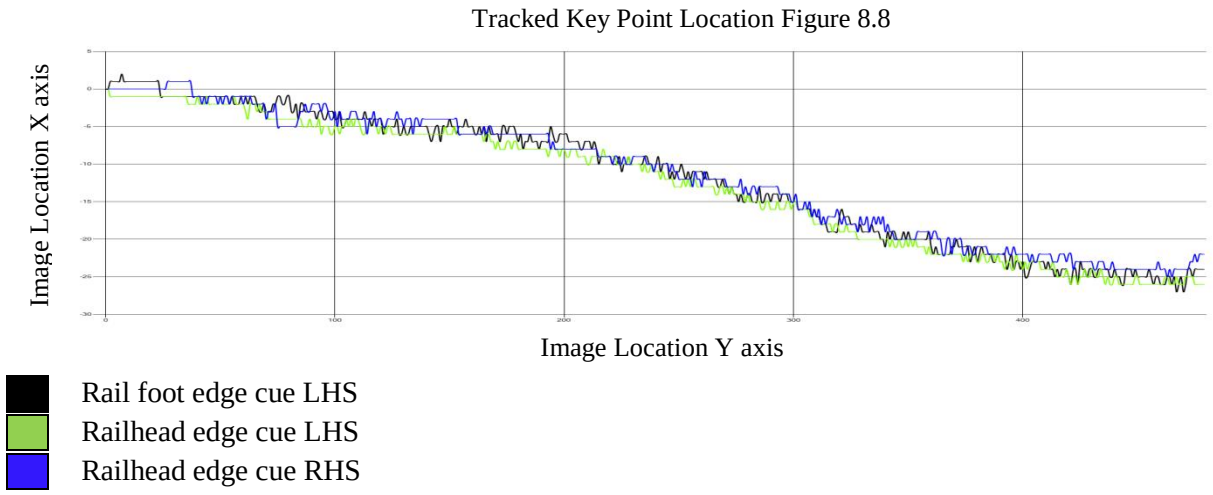



Figure 8.6 Tracking of all three edge cue key points within an image (Figure 8.8).

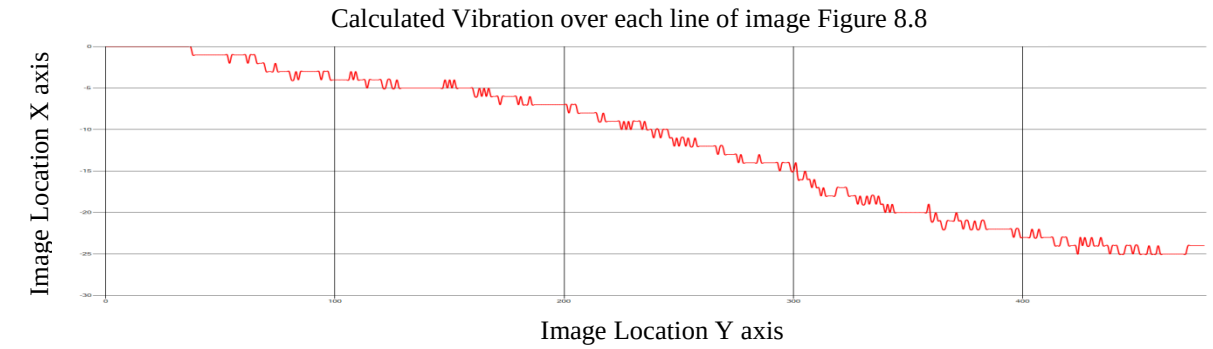



Figure 8.7 Average movement of all three edge cue key points
within a image an image (Figure 8.8).

Further filtering of the average movement of the edge cue key points is still required. A simple averaging filter smoothes the noise still present and allows for a more improved correction of vibration within the image. Figure 8.8 show the tracked key points throughout each line of the image and Figure 8.9, the stabilised image.

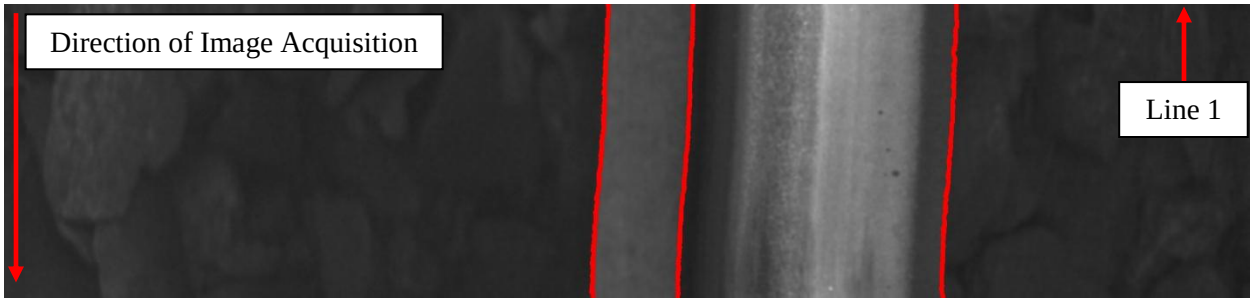
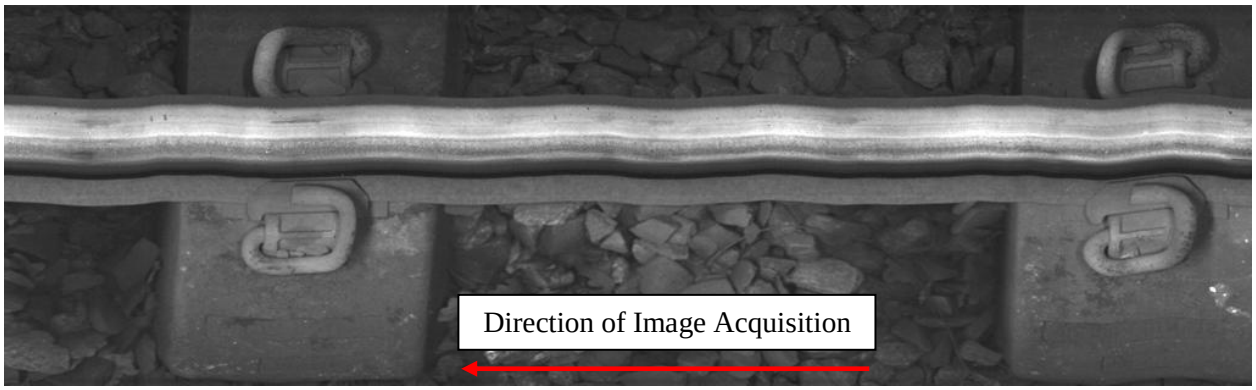



Figure 8.8 Tracked key points displayed on the rail image

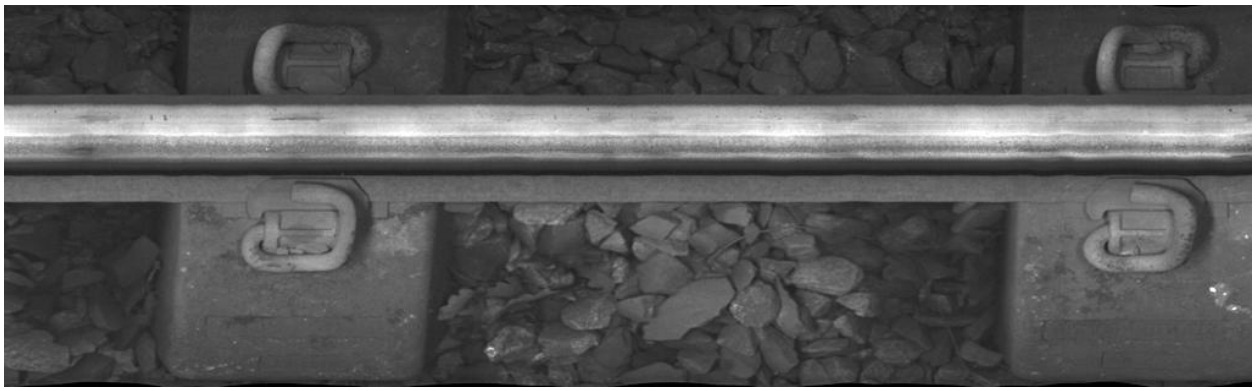
Figure 8.9 Stabilised image

## 8.3 RESULTS

The stabilisation algorithm was applied to a test set of 1000 images acquired from the Sheffield Supertram network forming data set 4. The sub data set has various amounts of vibration apparent as well as different kinds of infrastructure. Images were visually assessed for the performance of vibration correction. Figure 8.10 shows the results of applying the stabilisation algorithm over a set of 10 images. After stabilisation, the data is used to align the rail in each image to the preceding image. Table 8.1 shows the success rate of the alignment for the data set.



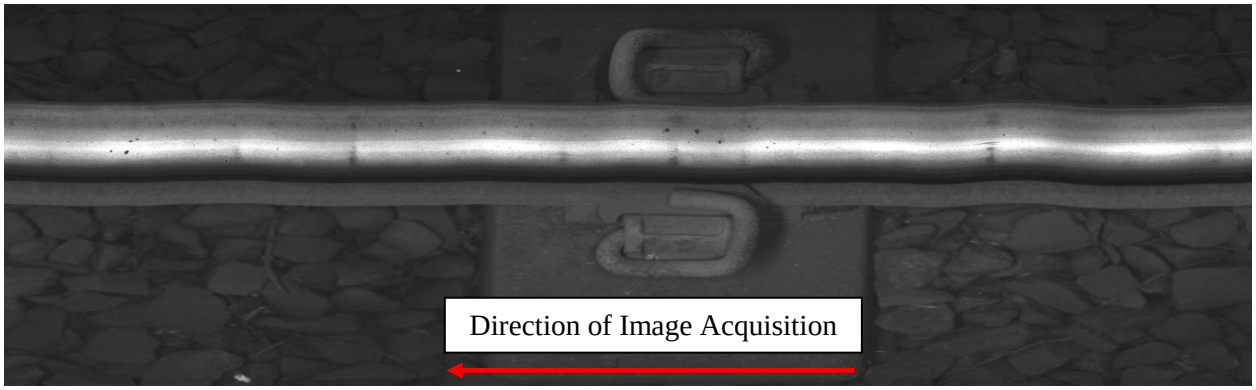

Vibration present over 10 images



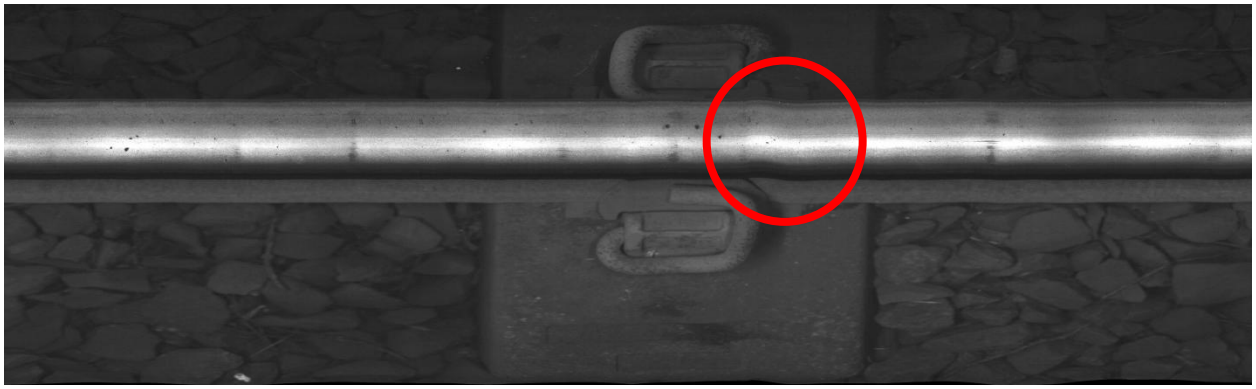
Corrected vibration over 10 images

Figure 8.10 Example of vibration correction over a series of 10 images.

Figure 8.11 shows a falsely aligned image where correction has occurred without vibration (circled in red). Vibration occurring has been significantly corrected in other parts of the image. The Pandrol clip in this occurrence has interfered with classification of the rail edge key point.




Vibration present over 10 images



Falsely aligned vibration over 10 images

Figure 8.11 Example of vibration falsely corrected over a series of 10 images.

| Correctly Aligned Images | Falsely Aligned Images |
|---|---|
| 944 | 56 |

Table 8.1 Image vibration correction results

## 8.4 SUMMARY

A 94.4% successful classification of key point data (Table 8.1) shows that successful suppression of vibration present within an image can be achieved using edge cue key point analysis. While this algorithm is not a final solution in its current state, the techniques discussed show that significant vibration within the image from camera movement can be corrected for.

In the 5.6% falsely aligned images, correct key points were still identified. The false classification of key points caused by linear edge cues of Pandrol clips and debris caused interference that prevented alignment of the rail edge within the image. Linear effects of the running band data also introduced noticeable noise. The running band edge cues differ from the railhead edge cues and could be eliminated by examining the image.

# CHAPTER 9 LINEAR FEATURE DETECTION IN THE APPLICATION OF RUNNING BAND ANALYSIS

## 9.1 INTRODUCTION

The running band is the area where vehicle's wheel sets interacts with the head of the rail. Analysing the space that the running band occupies can show numerous features such as surface defects (Squats, Wheel burns corrugation and RCF cracks). Along with segmenting the data, having the ability to analyse its edge and its width over distances allows for the analysis of rail wear; movement in track beading; incorrect installation of track; missing fixtures and fittings, including defects just under the surface of the head of the rail and early the detection of surface defects.

Assuming that the edges of the rail can be determined, any method developed will have to work with data forming the head of the rail only. This data will therefore consist primarily of two constructs; the running band data, silver and reflective in nature and the non-running band data with a brown rust like appearance. In greyscale images, the running band will appear brighter with a white/grey colouration. The non-running band data will consist of data with a dull to dark gray colouration.

There are difficulties encountered when tracking the running band, these are due to maintenance procedures for early repairs of defects. A method of early repair is to grind the area of track where a defect has been located, on the approach to and the exit of the affected section. This produces an unusual problem, an edge that can be seen by the human eye but can be very difficult to detect by existing machine vision techniques. Shadowing caused by different light sources and environmental effects such as rain can also interfere with classification.

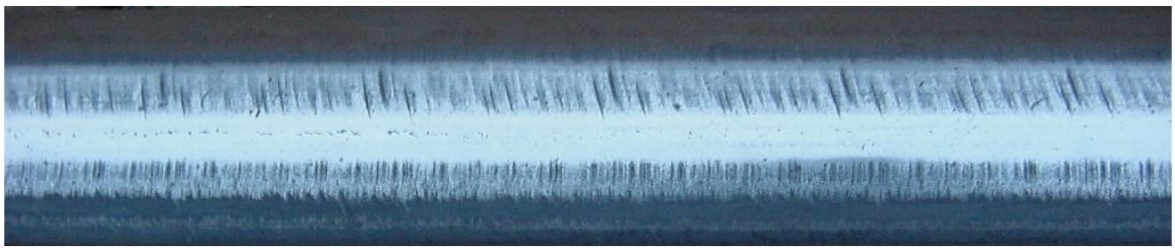


Figure 9.1 An example image of grinding marks from the removal of a track surface layer for defect repair.

A method has been developed using histogram analysis that classifies running bands. This is compared to standard edge detection methods, region classification methods, and colour matching/thresholding techniques. Certain edge classification algorithms such as Canny and Sobel, can successfully be applied if certain aspects of the image are constant. This is improbable due to different MVS types and varying environmental effects discussed such as lighting and the effects of weather such as rain.

If constant image aspects are not available due to environmental effects such as changing light conditions it needs to be determined if changes in the data forming image features can be evaluated from the image data and subsequently used to adapt the application of feature analysis algorithms and increase their performance. A possible method is histogram analysis. The histogram of an image shows the spread of pixel values over the range of 0-255. It can also show the maximum and minimum pixel intensities, and can be used to calculate the overall mean value.

Figure 9.2 demonstrates how useful histogram analysis can be in determining values for use in thresholding techniques and region growing methods. If image content remains constant, pre-calculated means can be applied. Histogram analysis allows an intensity value to be evaluated for each image.

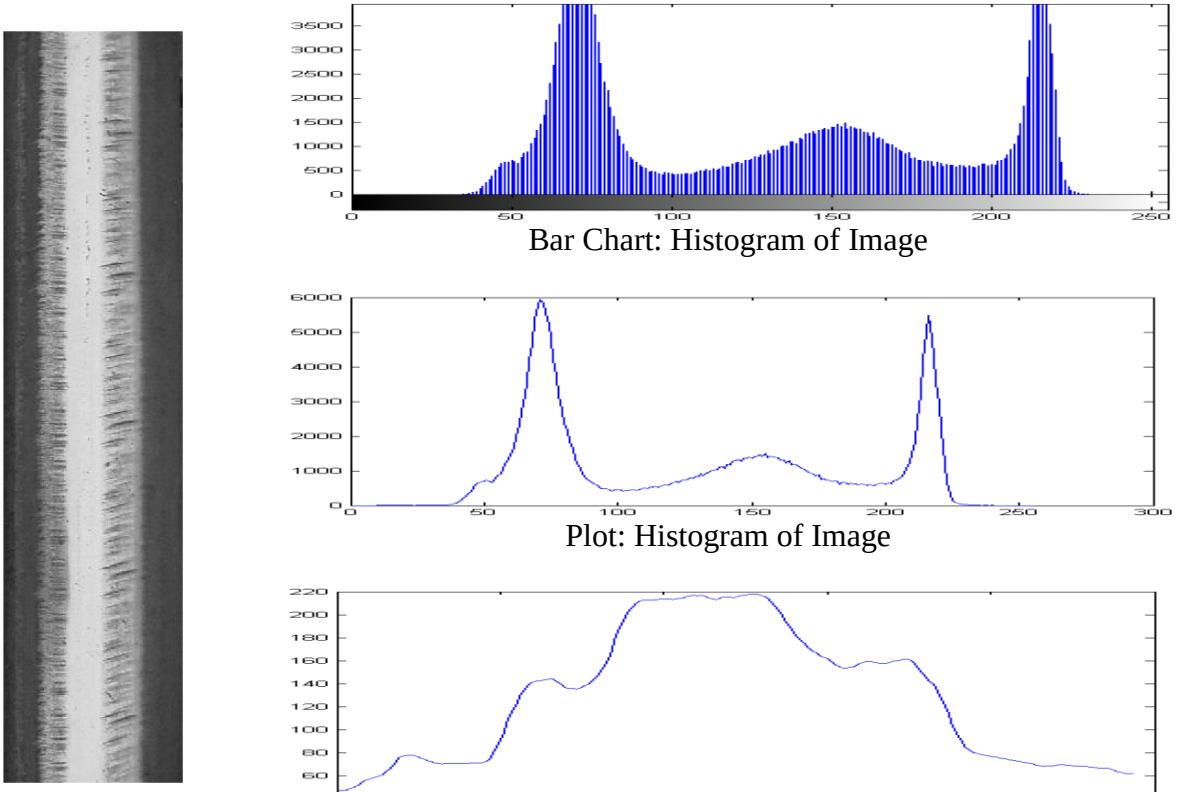



Figure 9.2 Plots of the data provided from image histogram and statistical analysis.

In Figure 9.2, *'Bar Chart: Histogram of Image'* and *'Plot: Histogram of Image'* show two distinct peaks and a third lesser peak. These peaks correspond to regions within the image. The left peak corresponds to the pixel values that make up the non-running band data. The centre peak corresponds to the old running band and the right peak corresponds to the current running band.

In Figure 9.2, *'Plot: Mean of Image Columns'* shows an alternative method of viewing this data, it uses the mean value of each column, and plots that accordingly. In this plot, each plateau corresponds similarly to the histogram peaks. Non-running band data is represented by the initial lower plateau's location. The middle plateau corresponds to the old running band and the higher plateau corresponds to the current running band.

Using this data, a bespoke histogram analysis algorithm has been developed with novel noise suppression algorithms. The algorithm allows a precise classification of the running band and its associated edge cues. Algorithms were then developed to analyse these edge cues for characteristics matching potential defects.

A key discovery from this work is the ability to classify multiple running bands upon the railhead. Each running band can be associated with the path particular wheel sets take from different vehicles. Using classified edges from all running bands defects, such as corrugation can be seen in there earlier stages. The classification of defects from this running band data is addressed by separating them into three areas, In-Image Defects, Multiple Image Defects, and Time Image Defects.

The running band path along the railhead surface is expected to be parallel to the direction of the rail, changes or variations in this path display characteristics of a defect. In-Image Defects (IID) are generally associated with singular defects such as a severe squat or wheel burns that effect the running band path sporadically. Multiple Image Defects (MID) are associated with corrugation where the running band is effected over several images in an oscillatory pattern. Finally Time Image Defects (TID) are those that cannot be directly associated with a surface bound abnormality however, show a significant change in running band position over time.

## 9.2 INVESTIGATIVE RESEARCH

Many techniques readily employed are discussed in the Literature Review Chapter heading 2.3.1 Feature Extraction. To evaluate the comparative performance of the developed algorithms current standards in edge feature classification were applied to the images. The results are statistically and visually assessed to evaluate their performance and accuracy. Methods were then expanded upon for application in running band analysis.

All images that the algorithms were tested on were part of data set 1 unless stated otherwise, see chapter 6.2 Data Sets (page 142) for reference. They are from a selection of stock imagery

totalling 969 images the content varies in image quality, camera angle, and lighting conditions. Those image presented show typical results from the data set and where possible have a similar FOV and camera angle to that of the final data set.

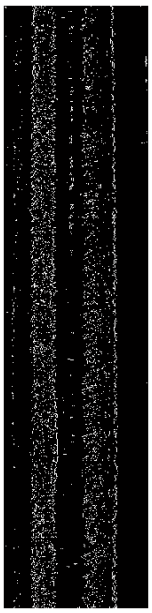
Region Classification methods are evaluated under different sections. Statistical filters, using mean and standard deviation algorithms, are examined alongside colour matching and thresholding techniques. A simplistic region-growing algorithm is also investigated. To evaluate robustness different images were evaluated where the colour/shading of the running band differed.

### 9.2.1 Standard Edge detection Methods in Classifying the Running Band

Standard edge detection methods for the detection of linear features have been previously discussed in Chapter 7.2.1. A majority of the details discussed remain true for this application; however, there are some differences in the nature of the running band. While a majority will have a linear nature defects will alter this. Methods for detecting the scale of this variation must also be considered during investigation. Figure 9.3 shows the difference in methods, with Sobel edge detection performing poorly and Canny more proficiently. Also shown are the results from applying the Hough transform after Canny edge detection has been applied. As can be observed, the results are mixed. While the Canny edge detection method can detect the grinding marks, to enable this sensitivity interference from marks not linked with the running band itself is inevitable even using the shown optimal settings.

In Figure 9.3, the Canny edge detector and Hough transform are applied to 6 separate images of running band from a range of sources. The success of the method is based on its ability to deal with varying greyscale effects within running bands. Varying changes in light or shadowing within the image inhibits the Hough transforms ability to distinguish between a true edge and false edges as the Canny edge detector classifies these as an edge cues regardless of length. This noise can be generated by defects themselves, however determining the difference between a defect or the alternative shadowing from the generated results is difficult.

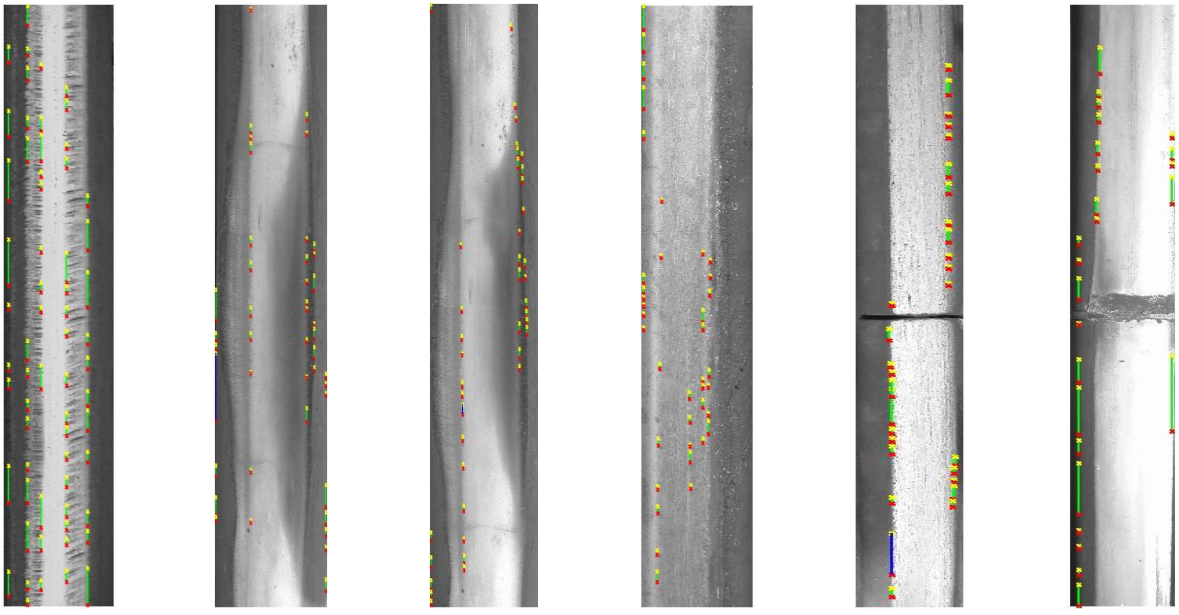
Original Image



Sobel Edge detection
Threshold 0.3, Vertical

Canny Edge detection
Threshold 1 0.2, Threshold 2 7

Hough Results from Canny Edge Detection Applied to Various Images



(1) (2) (3) (4) (5) (6)

Figure 9.3 Application of Hough line algorithm in classifying running band edges using Canny edge detection.

### 9.2.2 Region Classification

As the running band and non-running band data form large regions within the image, which cannot be reliably classified by locating their associated edge cues using standard techniques, an alternative is to use region classification techniques. Using pre-determined data or constants formed from image data the regions that make up the image can be determined. The edge cues can then be calculated by mapping the edge of the classified regions. This in principle is similar to edge detection methods however, all edge cue orientations are analysed uniformly in an individual pass.

The ability of a region classification technique can be evaluated more efficiently by comparing the total region it classifies against an expected region. This expected region is produced through manual inspection using pre-set constraints, such as no noise suppression, that the region classification algorithm is not limited by. This method allows computational evaluation that is more efficient then visually inspecting the edge cues of the regions detected.

Region classification is achieved by using a variety of methods; the most common is the use of colour matching and thresholding techniques. While these are in application the same process, colour matching will be used descriptively for work with colour imagery where three threshold values are used. Greyscale images have only one threshold value, and thus will be referred to as thresholding because of the differences in data.

Other such methods include the use of mathematical analysis such as statistical filters. These evaluate each pixel and its surrounding data to evaluate its associative contents. Such methods allow noise suppressive algorithms to be implemented within the filtering method. There are two distinct application methods of applying a statistical filter. The first is linear, in which each pixel within the image is examined regardless of position. The second is region growing methods in which a specific point is located that belongs to the data set; only pixels directly linked to this point are considered for analysis. This potentially reduces the execution time. Both methods are discussed and investigated.

As discussed, each region classification algorithm is applied to detect the whole region and evaluated on its output. An algorithm that is successful in classifying the calculated region will be more likely to be successful in the classification of the edge cues associated with the running band data. This is discussed further in the Method Evaluation and Observations section (*page 228*).

#### 9.2.2.1 Statistical Filters

If the information that forms the running band and non-running band data is consistent then it is plausible that the data can be segmented by forming statistical filters. Using information about each pixel and its associated neighbours, a statistical filter should be able to associate each pixel as belonging to the running band or non-running band.

To assess the performance and noise suppressive abilities of differing filter sizes, two tests were employed. The designed filters ability to classify running band data is assessed by comparing the resultant binary map of classified running band data points with that of manually produced binary map of the same image. Receiver operating characteristics (ROC) data is calculated from this comparison.

ROC data allows the performance of binary classifiers, such as statistical filters, to be assessed as a discrimination threshold is varied. ROC data is represented by plotting the fraction of true positives against the fraction of false positives at the varying threshold settings. By assessing these graphs, appropriate thresholds can be decided upon on visual basis. As the area under the graphs represent how a classifier will rank, a visual representation of how a sacrifice in true positives can affect the false positives is achieved.

Two plots are presented for the ROC analysis an ROC curve and a decision threshold graph. An ROC curve shows the representative relationship of the true positives and false positives as a percentage allowing reliability to be evaluated Figure 9.4 shows how reliability can be seen. The closer to a right angle the curve approaches the more reliable the method as 100% true positives

can be seen with 0% false positives. A Decision Threshold Graph allows a threshold value to be evaluated from plotting know data. Figure 9.7 shows how the threshold value can be evaluated to balance the true positive false negative rate and the true negative false positive rate. In an ideal situation, the two peaks will be separated and not overlap.

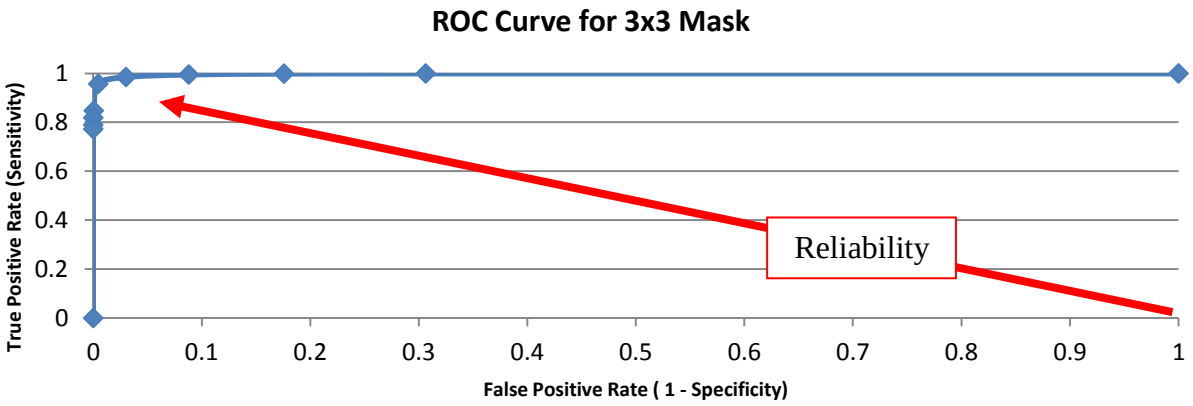



Figure 9.4 Demonstration of evaluating reliability of algorithm from ROC curve data.

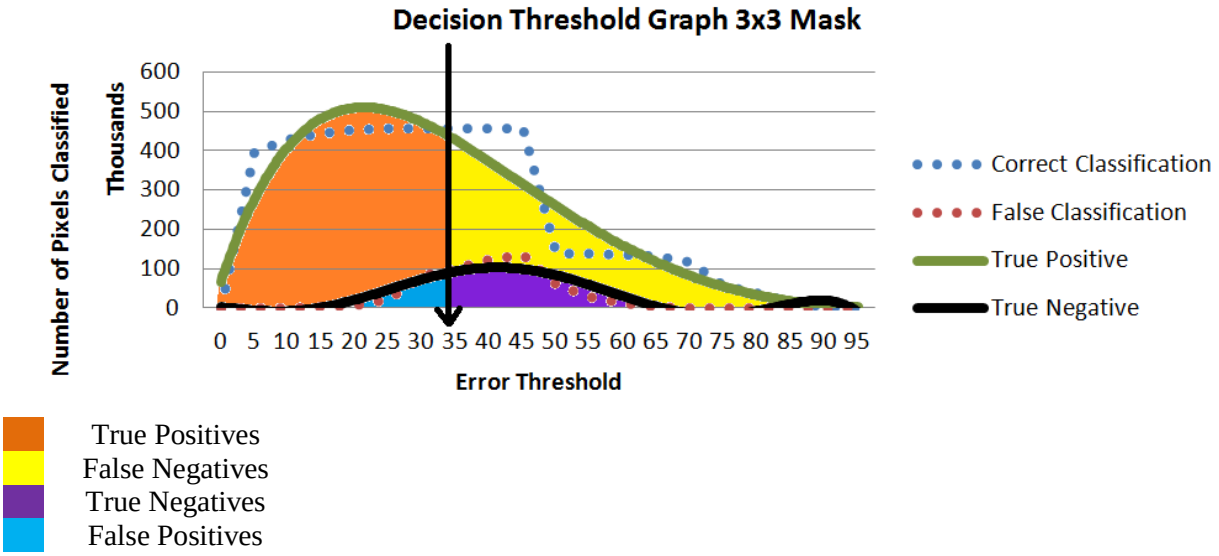



Figure 9.5 Demonstration of evaluating ideal error threshold of algorithm from ROC data.

The first test of the designed filters ability to classify running band data examines a series of increasing mask sizes calculating only the mean of each individual pixel and surrounding data and then comparing it to the calculated mean of the mask. Several error rates are analysed to produce ROC data for comparison. The second test will examine the standard deviation of the sum of the rows and the sum of the columns.

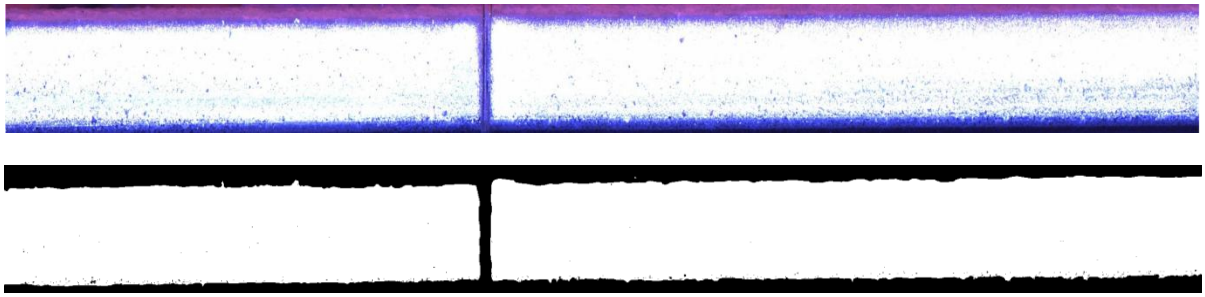


Figure 9.6 Manual binary map of running band data using visual assessment

Due to the manual creation of the binary map, it is unlikely that any method employed will reach 100% accuracy. To examine a technique's ability to classify running band data, the best filter size and thresholds can only be estimated from the results. These methods must then be employed on a larger data set and the results manually inspected for success. By employing two filters, the results from both methods can be examined side by side for any extreme variations. A subsequent decision on the relative merits of the method can then determined based on its performance.

For a comparative test of both methods, the mask was produced from the image data and had no filtering employed. The mask was formed by data around a single X, Y pixel location (100, 150).

#### 9.2.2.1.1 Mean Testing Results

Mean testing examines a pixel and its surrounding data or neighbours. Calculating the mean of all the pixels produces a value that can be compared to a threshold. A decision based on this threshold can then be made in order to classify or bin the pixel data. The overall success of the mean testing method is likely to be below that of other methods. The mean testing method although quick, is affected by noise. This effect can be suppressed by the use of Gabor filters and larger mask sizes.

Table 9.1 shows the overall percentage error of utilising an averaging mask filter of different size to detect and classify running band data. Evidence suggests that a 10x10 mask allowing a 40% error will produce the best available results. The supporting ROC decision threshold plot Figure 9.7 supports this finding but suggest a lower threshold of 35% would produce a better True positive to false positive rate. Results from mask sizes 3x3, 5x5, 20x20 and 30x30 are presented in Appendix 10 for comparison.

| | Overall Percentage Error (%) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mask size | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| 3x3 | 10.14 | 4.95 | 2.52 | 1.81 | 3.48 | 13.68 | 16.85 | 20.29 | 22.40 |
| 5x5 | 10.08 | 4.73 | 2.25 | 1.67 | 3.56 | 13.49 | 16.52 | 19.50 | 21.22 |
| 10x10 | 9.75 | 4.81 | 2.25 | 1.61 | 4.02 | 12.76 | 15.25 | 17.10 | 17.82 |
| 20x20 | 11.22 | 5.67 | 2.51 | 2.01 | 5.07 | 9.29 | 10.03 | 10.21 | 10.21 |
| 30x30 | 11.52 | 5.46 | 1.88 | 2.04 | 3.07 | 4.03 | 4.03 | 4.03 | 4.03 |

Table 9.1 Overall percentage error of average filtering mask for running band detection.

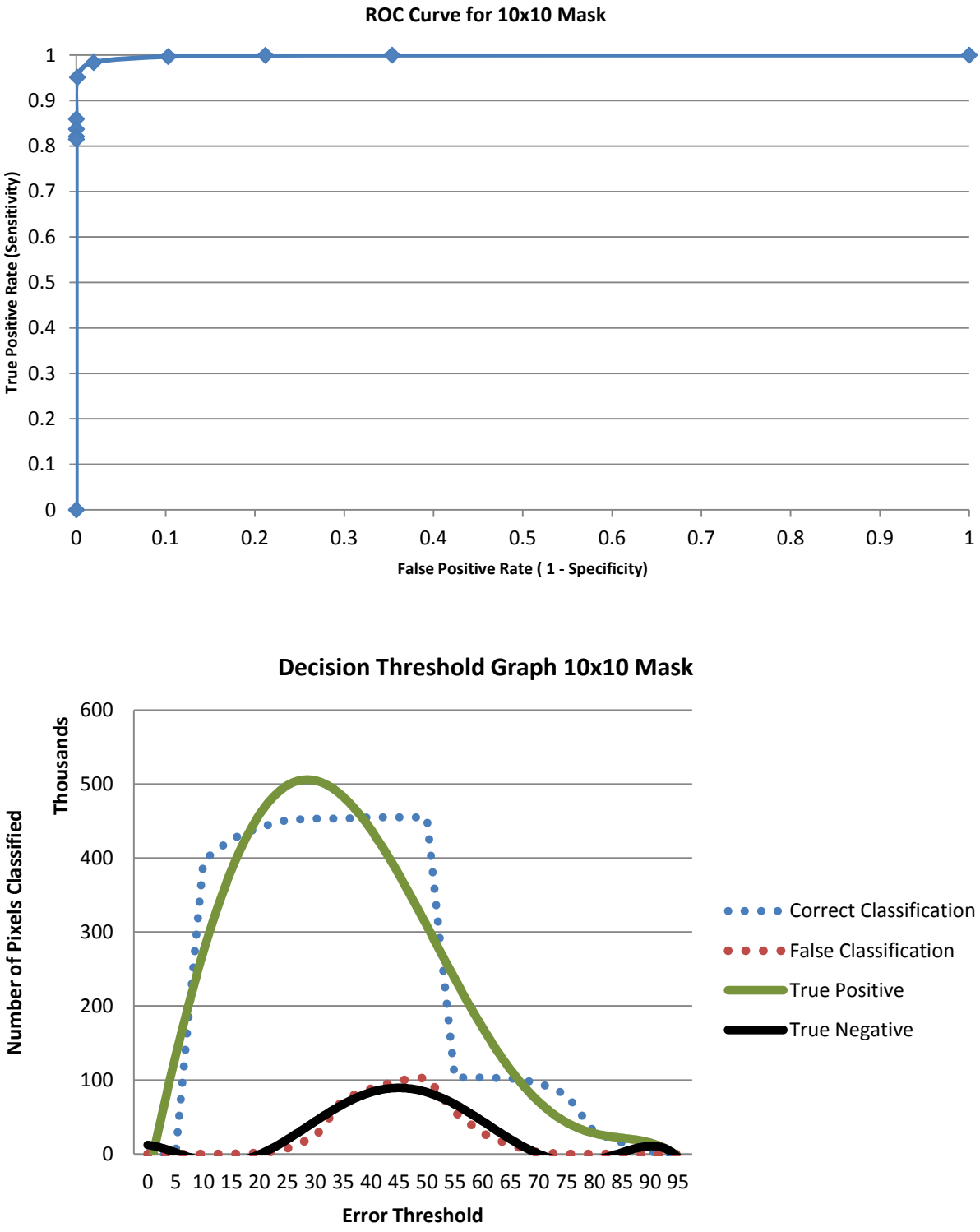



Figure 9.7 ROC curve and decision threshold graph for averaging 10x10 mask.

#### 9.2.2.1.2 Stand Deviation of the sum of Columns and Rows Testing Results

The standard deviation was calculated utilising data from the area to be examined and the selected mask data. The sum of the rows from the selected pixel area and the sum of the rows from the mask were concatenated with the sum of the columns of the selected area and the sum of the columns of the mask before standard deviation was calculated. This method allowed noise to be accounted for in a similar fashion to the mean testing filter, before the standard deviation between the mask and the pixel data was examined.

Summation of Columns and Rows for Mask and Pixel

| | | $f_1$ $\sum_{x=1}^{n}(x_n-1,n)$ | $f_2$ $\sum_{x=1}^{n}(x_n,n)$ | $f_3$ $\sum_{x=1}^{n}(x_n+1,n)$ |
|---|---|---|---|---|
| $f_4$ | $\sum_{y=1}^{n}(n,y_n-1)$ | (x-1, y-1) | (x, y-1) | (x+1, y-1) |
| $f_5$ | $\sum_{y=1}^{n}(n,y_n)$ | (x-1, y) | (x, y) | (x+1, y) |
| $f_6$ | $\sum_{y=1}^{n}(n,y_n+1)$ | (x-1, y+1) | (x, y+1) | (x+1, y+1) |

$$S_n = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_n - \bar{f}\right)^2}$$

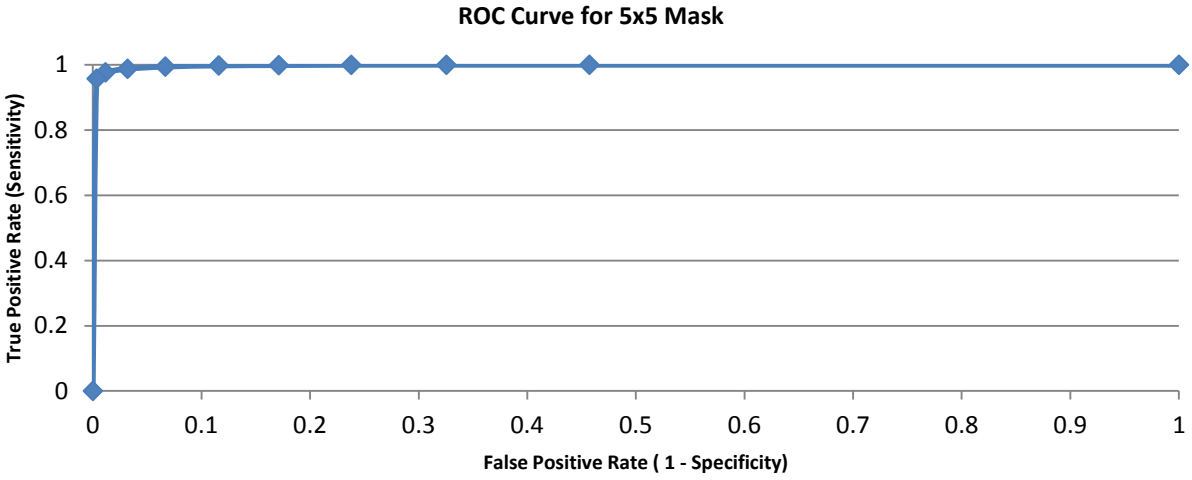
Where:

- $S_n$ = Standard Deviation
- $f_n$ = Sum of Row or Column
- $\bar{f}$ = Mean of $\sum f_n$

Equation 9.1 Standard deviation for concatenated sum of rows and columns.

| | Standard Deviation Constant (K) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mask size | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| 3x3 | 16.68 | 10.28 | 6.89 | 4.68 | 3.15 | 2.18 | 1.81 | 2.08 | 3.37 |
| 5x5 | 18.12 | 10.38 | 6.72 | 4.80 | 2.95 | 1.93 | 1.63 | 2.05 | 3.46 |
| 10x10 | 19.19 | 10.33 | 6.75 | 4.64 | 3.01 | 1.89 | 1.50 | 2.17 | 4.05 |
| 20x20 | 20.70 | 10.95 | 7.87 | 5.45 | 3.55 | 2.38 | 2.18 | 3.97 | 5.04 |
| 30x30 | 21.09 | 11.81 | 8.52 | 5.81 | 3.59 | 1.91 | 1.95 | 2.78 | 3.90 |

Table 9.2 Overall percentage error of standard deviation filtering mask for running band detection.



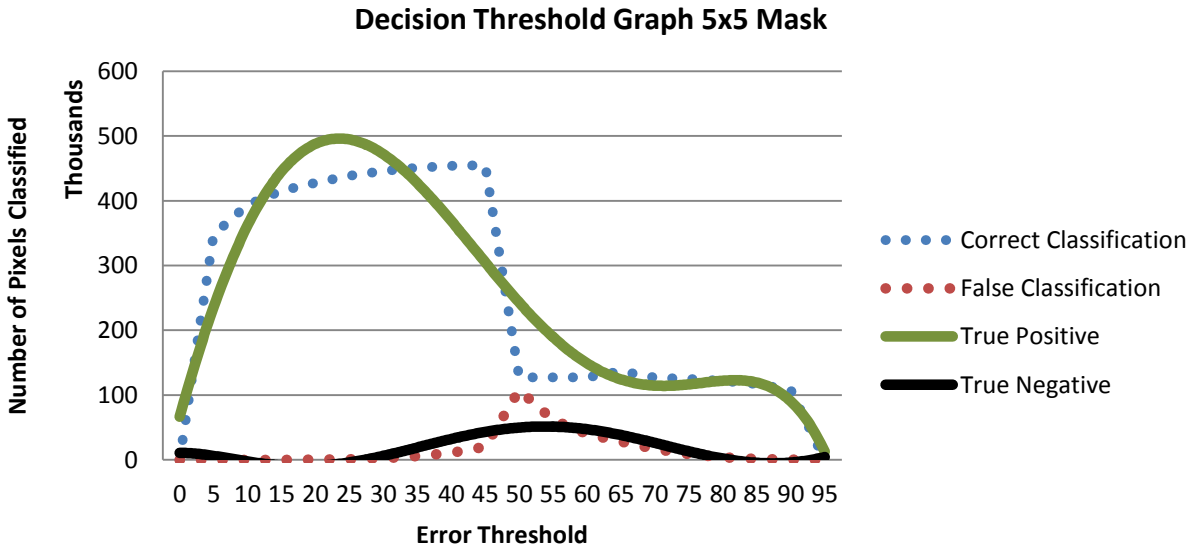




Figure 9.8 ROC curve and decision threshold graph for standard deviation 10x10 mask.

Results show that a 10x10 mask size produces the best results from the overall percentage error data. This is not supported with the ROC data (Appendix 11), that suggests a 5x5 mask with a standard deviation threshold of less than 30 would produce the best classification accuracy.

Both calculation methods show that a statistical filter can be used in the classification of running band and non-running band data. As with any segmentation, method additional steps would be required to trace the edge of the running band data. These could be included into the algorithm to improve efficiency.

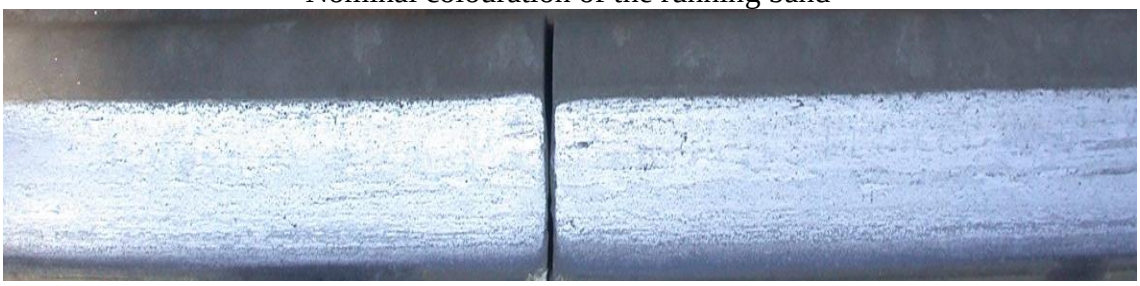
By comparing the ROC data produced by the mean testing filter and the standard deviation test shows that the standard deviation algorithm has a much higher success rate. This complexity is evident in the algorithms execution time. Any future investigation should include the use of multiple statistical examinations. More than one statistical filter would be utilised to improve accuracy when certain conditions are met. Such conditions may be small areas of classified data i.e. running band data or after a long set of classified data results in non-running band classification.

#### 9.2.2.2 Colour Matching

Colour matching has a distinct advantage over gray level thresholding methods, as it allows the analysis of 3 sets of data that contain the boundary information whereas thresholding only uses the one set of this data collection.

Colour matching has some merit of appeal for this application, as the distinguishing features of the non-running band data is its orange/red rust colour, compared to the polished steel look of the running band itself. However, many different factors can change the colour of the running band, most common of which is the weather condition i.e. rain. Interference from oil and lubricants make this type of technique more complex. Environmental interference is presented in Figure 9.9.

Nominal colouration of the running band



Environmental discolouration of the running band

Figure 9.9 Changes in non-running band colouration due to environmental factors.

The non-running band colour is subject to change, however the running band itself is more resilient to colour change as shown in Figure 9.9. It is usually shinny in comparison and the brighter part of the image. A reasonable solution would be to segment the running band data itself and ignore all other data that does not fall within this boundary. Figure 9.10 shows the binary results when segmenting the data that is consistent with the rail head its self. A comparison between colour matching and greyscale thresholding techniques is shown. The results indicate that a greyscale thresholding technique produces comparable results.

#### 9.2.2.3 Thresholding

In the top three images of Figure 9.10, the constant subtracted from the image was 150 for all three-colour spectrums, red, green, and blue. As can be seen in the bottom images of Figure 9.10 applying the same methods to a greyscale image has the same desirable results. While the colour matching method is a plausible solution to segmenting the running band data, the results shown suggest that the same operation can be applied on a greyscale image.

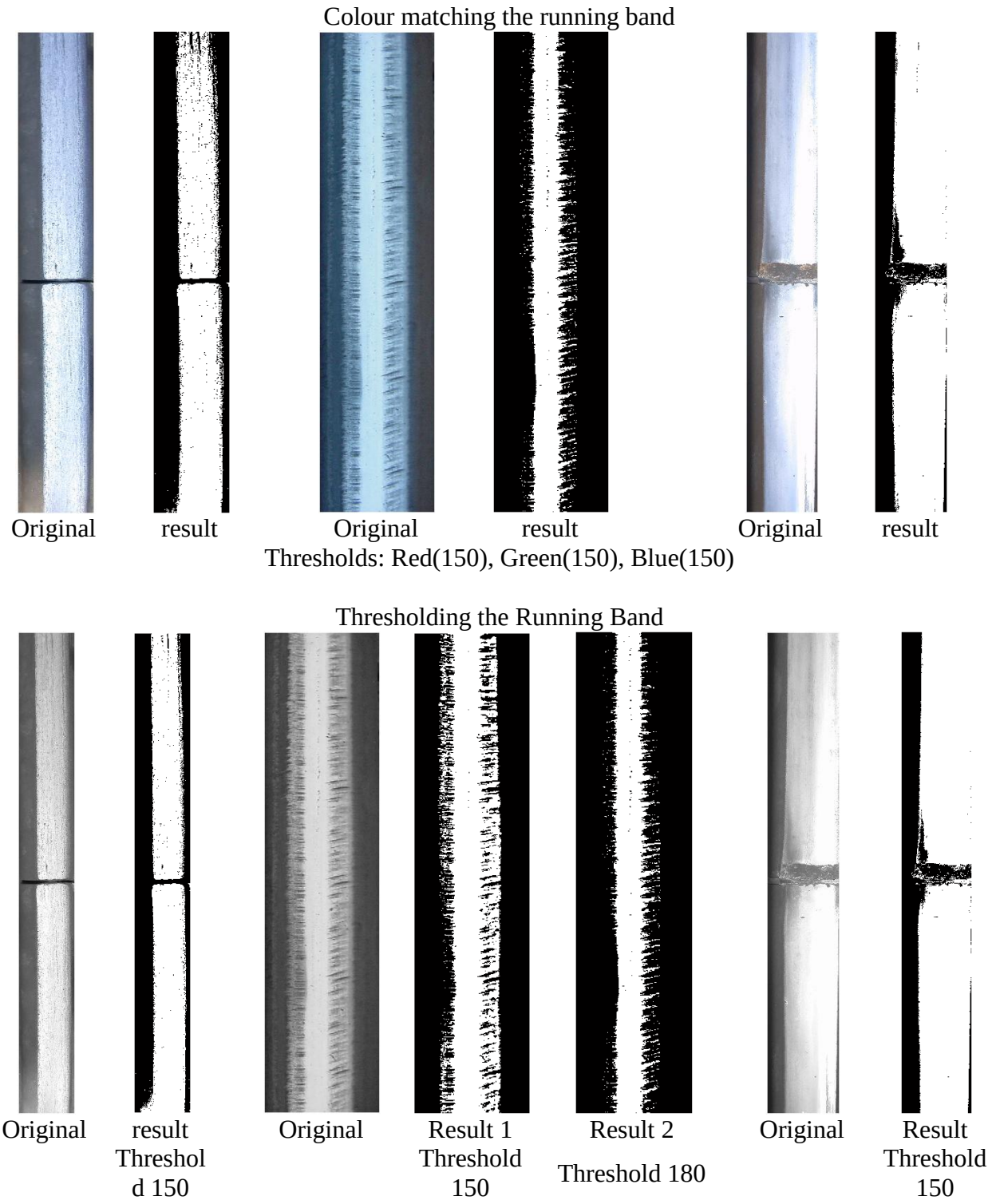



Figure 9.10 Comparison of colour matching and thresholding techniques in the classification of running band data.

Thresholding does have inherent problems when it comes to grinding marks, as the colour matching method does. Since thresholding only uses a third of the data set, limitations are to be expected. The difference between thresholding on a greyscale level and colour matching is that with grinding marks, the old running band is detected and the current running band is lost. It is possible to detect both running bands by subtracting two different constants.

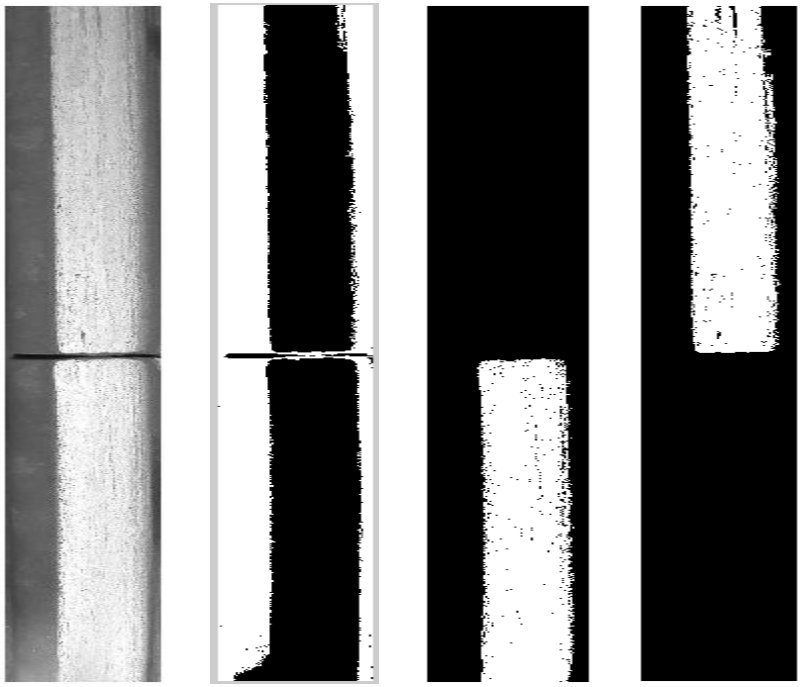
If data within the image is used to determine the threshold limit to be applied during data suppression, results can be improved. This ensures that the current running band can be segmented. There are varying unique data sets that can facilitate this, for example mean pixel value, maximum pixel value, minimum pixel value and average pixel value. These are only a few possibilities.

#### 9.2.2.4 The Detection of the Running Band by Region Growing

Similar to the statistical filters previously discussed, region-growing algorithms reduce the potential data set required for examination. The same mean comparison algorithm used for the statistical filter was used for the classification rule of the region-growing algorithm. The method applied is genetic region growing algorithm with random growth patterns discussed in Chapter 10.2.4 (*page 279*) and Equation 10.1.

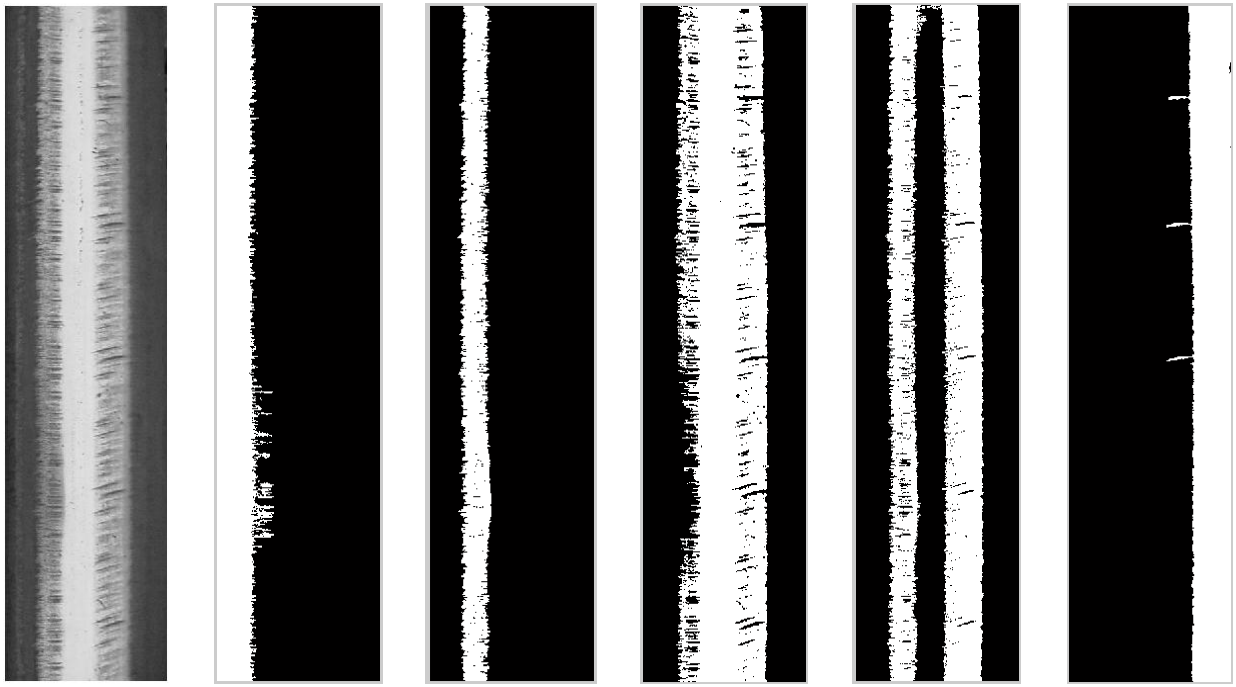
As with the alternative region classification methods discussed, a constant was used for initial analysis. Figure 9.11 shows the results from adapting a basic region-growing algorithm employed to locate a region with a pre-calculated mean. It is only designed to locate a single region at a time. However is set to locate all possible regions until 90% or more of the image data is classified. The start point is manually calculated, as the binary map of pixels associated or not associated with regions is not fully utilised.

Image 1



Original Region 1 Region 2 Region 3

Image 2



Original Region 1 Region 2 Region 3 Region 4 Region 5

Image 3

Original Region 1 Region 2 Region 3

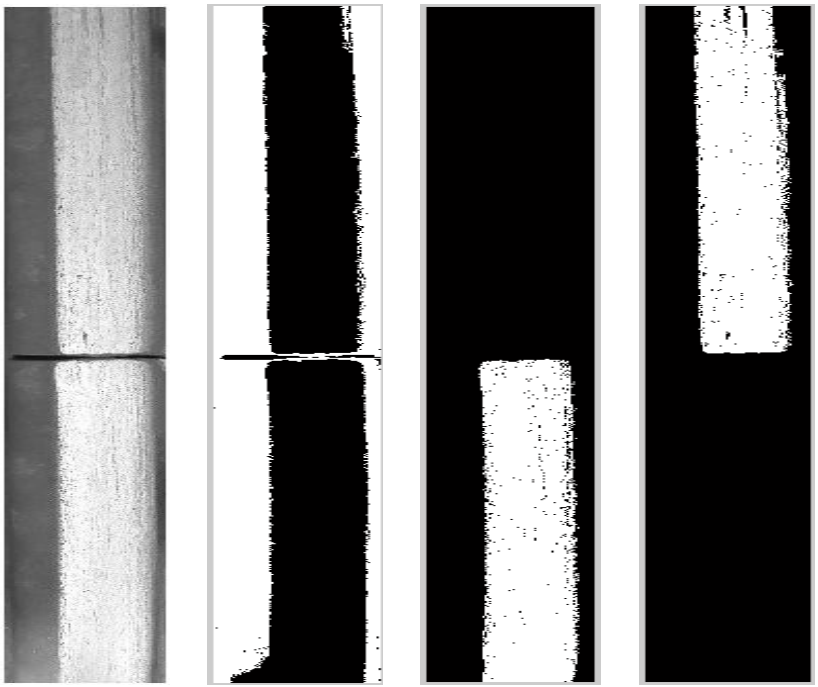
Figure 9.11 Region Growing Mean Algorithm: Image segmentation results.

The investigation was extended to utilise the mean values required for region classification through analysis of the accumulation of the image columns. The region growing method was then applied to the image in order to associate as many pixels as possible to an appropriate region using these mean values. The starting points for the region-growing algorithm were pixels in which this mean value occurred that had not already been classified within a region. The results are comparable to the self-mean calculating region-growing algorithm shown in Figure 9.12, Figure 9.13, and Figure 9.14.

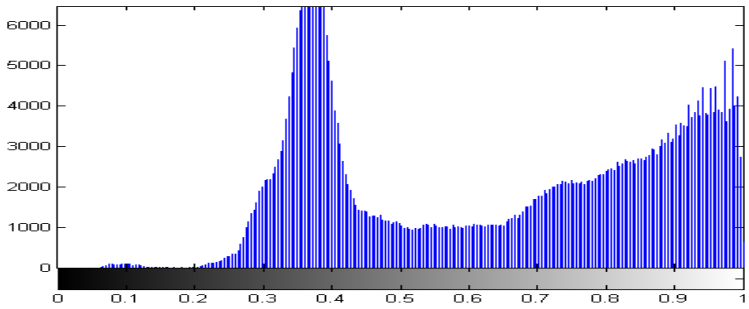
Figure 9.12 shows identical results while Figure 9.13 shows complications of cross over regions that would be eliminated with the use of a universal binary map. A binary map allows the algorithm to see what areas have been classified as belong to the region being inspected. The use of a universal binary allows the algorithm to see what pixels within the image have been classified as belonging to any region. Figure 9.14 shows good results with only 3 main regions

being detected, however some misclassified data is visible in the non-running band region due to a more severe change in shadowing. There is echoed non-running band data within region 4, which again could be eliminated with a universal binary map. The results show that data corresponding to that of the running band can be utilised in detecting defects on the head of the rail.

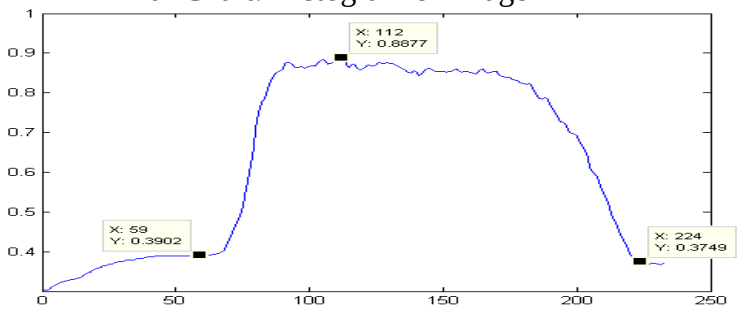
Image 1



Original Region 1 Region 2 Region 3



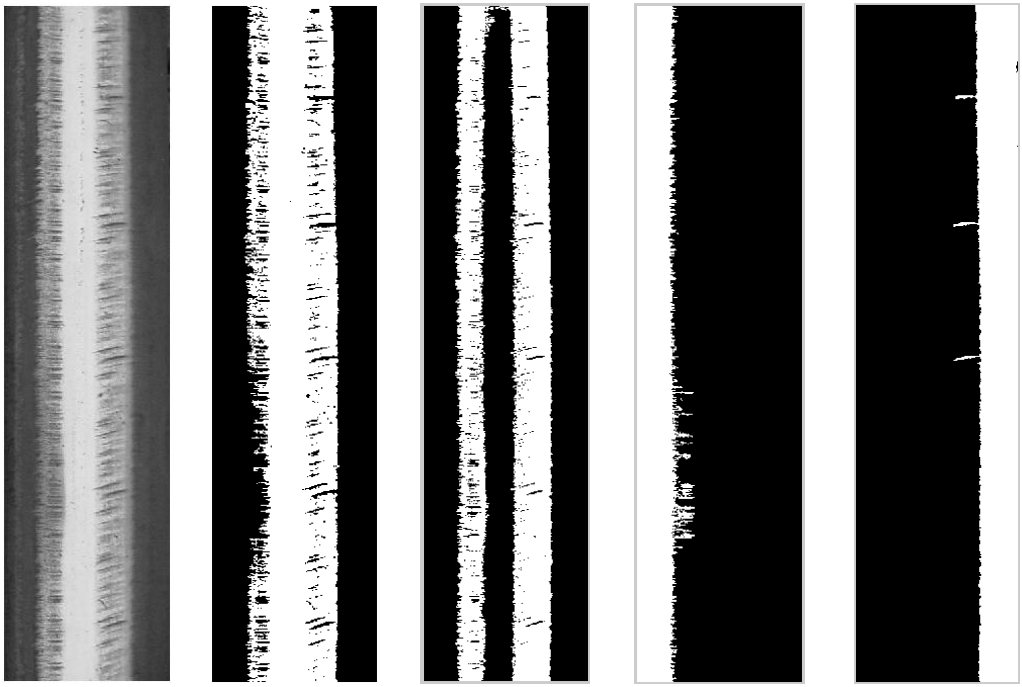

Bar Chart: Histogram of Image



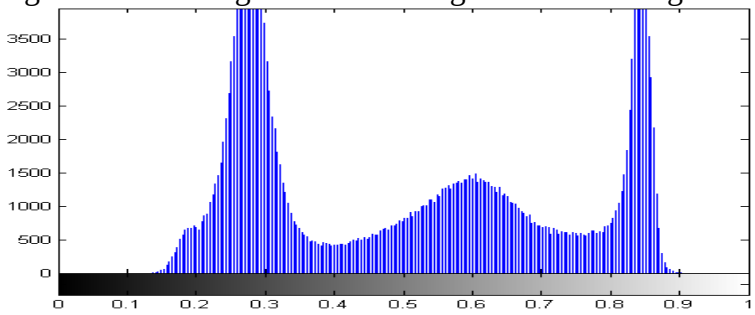

Plot: Mean of Image Columns

Figure 9.12 Histogram analyses and results from applying the mean test region-growing algorithm (Image 1).

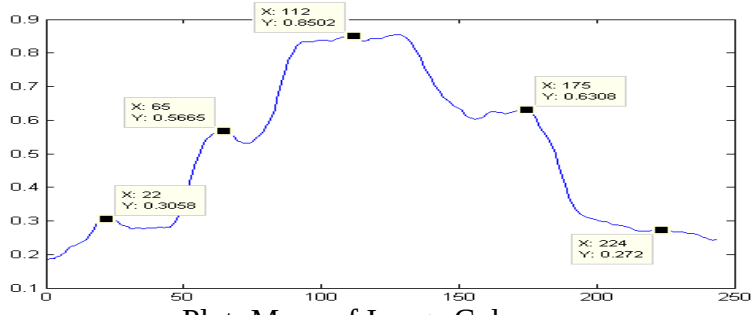
Image 2



Original Region 1 Region 2 Region 3 Region 4



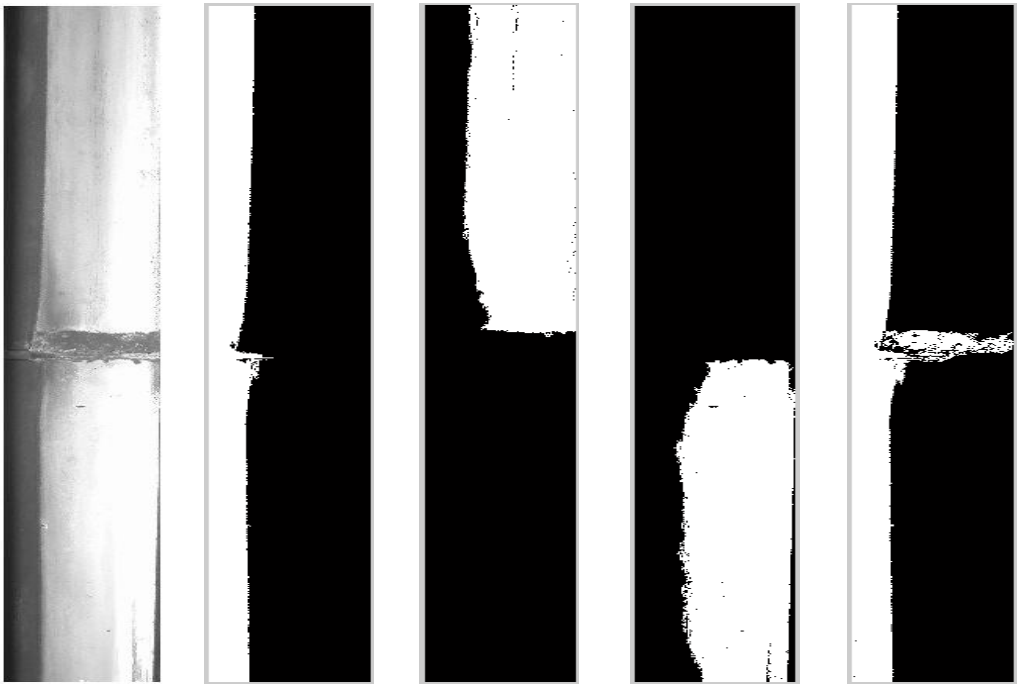

Bar Chart: Histogram of Image



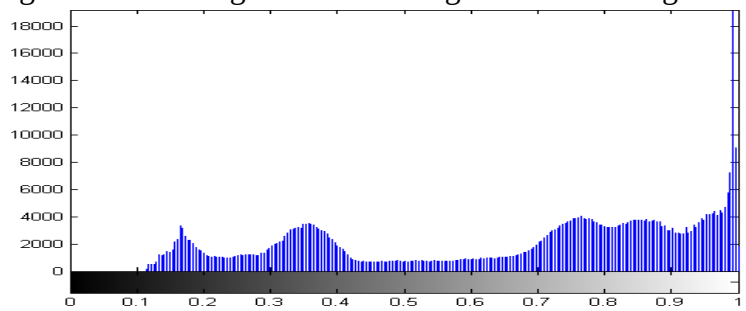

Plot: Mean of Image Columns

Figure 9.13 Histogram analyses and results from applying the mean test region-growing algorithm (Image 2).

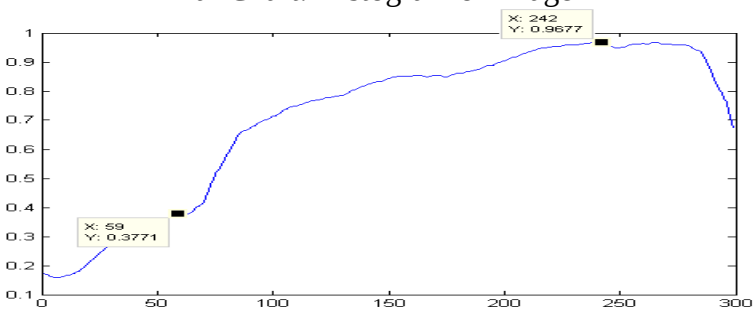
Image 3



Original Region 1 Region 2 Region 3 Region 4




Bar Chart: Histogram of Image




Plot: Mean of Image Columns

Figure 9.14 Histogram analyses and results from applying the mean test region-growing algorithm (Image 3).

## 9.3 METHOD EVALUATION AND OBSERVATIONS

The varying data contents forming the running band produces an unusual problem, an edge that can be seen by the human eye but can be very difficult to see through machine vision techniques. Figure 9.15 is a prime example of when this is true. Being an edge based problem, standard edge detection methods have been examined showing that their ability of detecting the running band under varying circumstance is poor. Region growing methods and statistical filters proved successful in classifying running band data. With the ability to classify running band data, the boundary data of these regions are also the edge cues required for defect analysis.

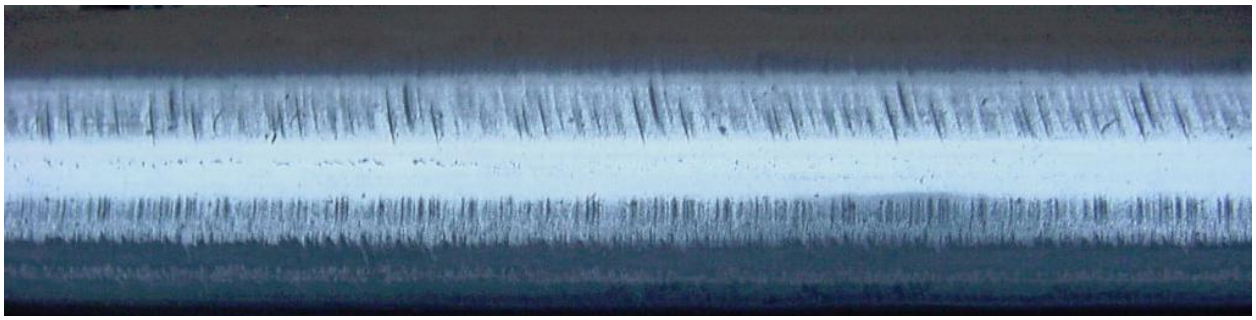


Figure 9.15 An example of a running band that a computer algorithm will struggle to classifiy

As discussed in the application of a region-growing algorithm a possible method of increasing the performance of region growing and statistical algorithms is to use calculated means from image contents. With simple thresholding and colour matching proving, equally successful in the classification of running band data the complexity of statistical based methods proves a disadvantage. The simpler threshold suppression techniques are easier to implement across various image types and in terms of execution times are more favourable.

Similarly, to statistical methods, calculated intensity values from image contents can be used to supplement the thresholding technique algorithms more robust to environmental effects, such as rain. Appendix 6 shows the results from using histogram analysis to calculate a constant for thresholding. This method shows inherent advantages to a region-growing algorithm. Foremost is the increase in performance, with each region segmentation taking 0.1% of time to complete

compared to a region-growing algorithm. This together with the relevant simplicity in programming terms compared to the more complex region-growing algorithm makes thresholding a desirable method.

Region growing methods should still not be ignored as they can produce comparable results to the thresholding method. Whereas thresholding techniques will produce results similar to those expected, even if the data is inconsistent. Data suppression through quantitative analysis is the suggested method for linear feature detection in the application of running band analysis.

## 9.4 RESULTS: IMAGE DATA SUPPRESSION THROUGH QUANTITATIVE ANALYSIS IN RUNNING BAND CLASSIFICATION

### 9.4.1 Initial Evaluation

Methods for detecting the correct value for image data suppression through examination of the image have been investigated. The methods designed have been tested on two sets of images; a set of 20 lower quality images (data set 2) and a set of 5 higher quality images (data set 3), see chapter 6.2 Data Sets (page 142) for reference. These data sets are acquired from two different MVS to test the adaptability as well as the validity of the algorithms. The image set data consist of only data forming the head of the rail.

Evaluating the results from the algorithm is complicated. As with a majority of applications visual assessment must be employed. Visually assessing each edge pixel of the running band data is unreliable. To allow for a faster more reliable testing of the algorithms success rate a set suppression point was evaluated through visual image assessment. The set point chosen for each image was used to suppress non-running band data and form a binary map. This process is discussed in Chapter 8.2.2.1 Statistical Filters. Due to irregularities in this technique, a positive classification is considered true when an algorithm classifies at least 90% of the binary map successfully as running band data.

The data suppression algorithm is designed around one of the lower quality images and the applied to the remaining 19 images forming the first test set (data set 2). Once the algorithm is deemed successful, it was applied to a second test set of 5 higher quality images (data set 3). These images still only consist of data forming the head of the rail. Since the second higher quality image set does not relate to the first set in hardware standards but only in data form, then the algorithms ability to adapt to various environmental and hardware factors is tested. As the second test set is smaller, a higher success rate of classification is required for an algorithm to be considered robust enough for final evaluation.

### 9.4.2 Quantitative Analysis

A value for data suppression can be applied to the image in a variety of methods. A method could directly apply the value to individual pixels. This can be as simple as subtract that value from all pixels and count any remaining data that is not suppressed to zero as belonging to the running band data. More advanced methods may suppress the data that is bellow the calculated threshold but leave any data that is not, to ensure original data characteristics remain intact. Although in principle, since all data will have a known constant applied to it, the only real purpose for this would be to serve as an aid for visual assessment.

Computationally it quicker to implement the same operation on all pixels then inspect a pixel to see if a value should be applied. It is important to consider this when optimisation is implemented however, the tracking of operations applied to the image to reconstitute the original pixel values can add unnecessary complexity to algorithms. It is often simpler to keep a copy of the original data for algorithms to reference.

This method of *Linear Feature Detection in the Application of Running Band Analysis* is unique. The data that forms the image is not adjusted or interfered with. The value calculated for suppression is used as a threshold to evaluate edge cues associated with running band data. This method of application can allow for greater efficiency when compared to applying a single operation to all pixels. This is only true if Equation 9.2 is satisfied, where the following must be true, $C_i = nC_m$ and $A_{NR} = \frac{A}{n+k}$.

$$C_m A > (C_m + C_i) A_{NR}$$

If

$$C_i = nC_m$$

$$A_{NR} = \frac{A}{+k}$$

∴

$$C_m A > (C_m + nCm) \frac{A}{n+k}$$

$$C_m A > \frac{nC_m A}{n+k}$$

$$C_m A > \frac{C_m A}{k}$$

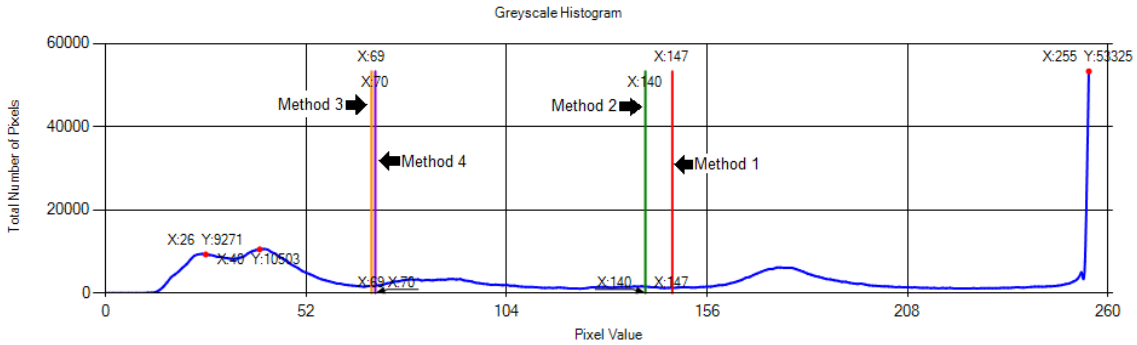
Where:
$A$ = Image Area
$A_{NR}$ = Non-Running Band Area
$C_m$ = Clock Cycles for mathematical operation
$C_i$ = Clock Cycles for Value Comparison
$n$ = Constant
$k > 0$

Equation 9.2 Equation to evaluate pixel inspection efficiency against applying mathematical operation to all pixels.

In initial investigation, the histogram and mean of the image columns were investigated for algorithms that would allow the calculation of the suppression threshold. Both sets of data show the distinct difference between running band and non-running band data. Quantitative analysis is performed on both sets of data to determine which set provides results that are more accurate. While there are other methods of representing the scale of the data within the image these two methods are seen to be more resilient to noise. This data is also readily available and easily manipulated to examine its potential uses.

### 9.4.2.1 Histogram Analysis

Examining the histogram of typical data forming the head of the rail shows two distinct peaks. Each peak is the collection of running band and non-running band data. The area underneath the graph between these peaks is the potential noise. Finding a point between these peaks can produce the values required for suppression of non-running band data. Figure 9.16 displays the histogram for the first image of the first test set. Four methods of calculating a suppression rate have been developed and applied to the image. The histogram analysis carries the potential for a noise-omitted design however; there are potential difficulties in producing reliable results.




Histogram for first image of the first test set

Figure 9.16 Graph showing the greyscale histogram of the head of the rail, with calculated suppression points shown

Dealing with multiple running bands is a complication with histogram analysis. Within the histogram data, the occurrence of multiple running bands will appear as multiple peaks according to the number of running bands present. To reduce effects such as wheel burns, marks, or breaks within the rail, a natural threshold must be developed to suppress peaks that are not considered a running band. Due to size of the railhead within an image in comparison to that of a defect upon the railhead, the dominant values of each peak can be used to analyse its potential area.

Figure 9.17 shows three potential points that have been classified as running bands. These are marked with red dot markers and have the X and Y values presented with them. An averaging

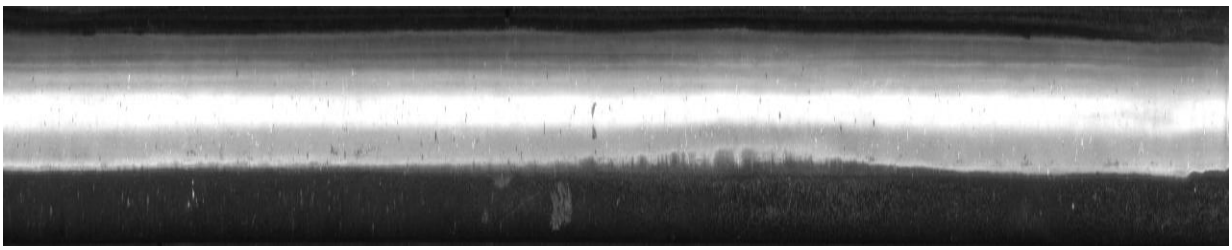
smoothing filter and maximum suppression is used to find all of the peaks within the histogram data. For each peak found, Equation 9.3 is used with a set threshold to determine if the data occupies a significant part of the image data. The threshold value varies depending upon the scale of the image of the image. With the lower quality image set (smaller image size) the threshold value is set at 0.01 (10% of the data) and the higher quality (larger image size) 0.003 (3% of the data).

$$P > A * T$$

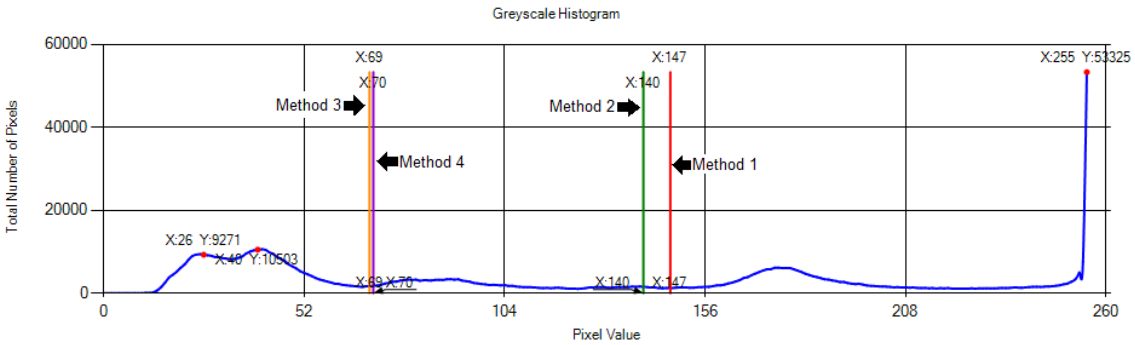
Where:
$P$ = Peak Value
$A$ = Image Area
$T$ = Threshold < 1

Equation 9.3 Equation to evaluate if a peak value belongs a running band or non-running band data set.

Analysis of the histogram data representing the first image displayed in Figure 9.17 shows that the far right peak represents the running band data. The peaks value of 53325 shows that pixels with the value of 255 consume 6.61% of the data forming the head of the rail image. Similarly, the two far left peaks representing non-running band data (pixel values of 26 and 40) have peak values of 9271 and 10503 forming 1.15% and 1.3% of the image data respectively. Using the lower threshold of 1% in the case of the lower quality images allows multiple running bands to be accepted by the classification algorithm.



First image of the first test set




Histogram for first image of the first test set

Figure 9.17 Graph showing the greyscale histogram of the head of the rail, with calculated suppression points shown.

The Four Methods that have been evaluated for calculating the suppression rate are described below:

Method 1 takes the centre point between the furthest right peaks. This method takes into account that on an ideal image with one running band and no noise there should be two prominent peaks. The first of a higher nature being the running band that is reflective in nature. The second would be the non-running band data something towards the lower values of the greyscale spectrum. A value somewhere within the middle should suppress the unwanted non-running band data.

The results show that in practice Method 1 falls short in producing the required threshold value. This mainly falls down to the properties of the head of the rail image. Rather than seeing a definitive edge that divides running band and non-running band data, there is a shadowing effect

that makes up the boundary data. This data is has a significantly low enough value to fall below the calculated threshold value and thus be falsely suppressed.

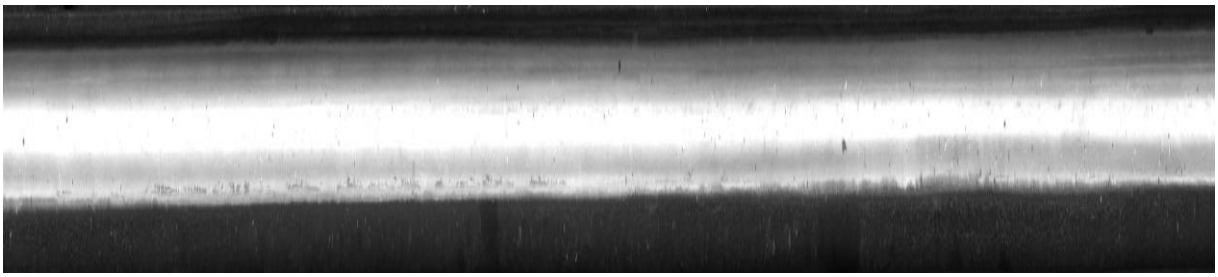
Method 2 purely relies on their only being one running band present within the image. This is an acceptable assumption considering that in a majority of images this is true. Method 2 finds the midway point between the lowest peak and the highest peak and ignores any other peaks detected. While this is only a small shift in suppression value it produces an increase of 85% accuracy. Despite improvements over the result presented in Method 1 the method still fails to classify a nominal suppression value.

Method 3 is designed to allow more calibration, allowing the method to be tailored towards changing hardware within MVS. Accepting that the suppression values are two high when the midpoint between peaks is taken, at a point ratio of 1/5.3 or around a 19% between the two furthest left and right peaks is calculated. A value closer to a nominal suppression value can be obtained consistently. As in Method 2, this method takes the assumption that only one running band is present within the image. The ratio is decided by comparing the first image of the first test set with the expected binary map and adjusting the value so that a minimum number of false negatives are recorded.

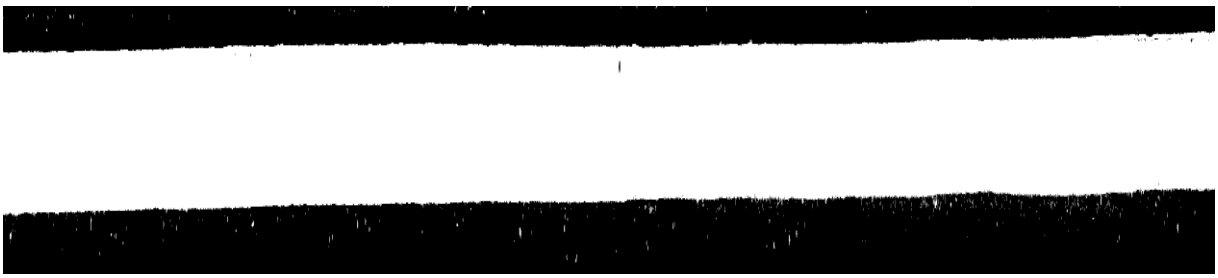
Method 4 combines the techniques applied in Method 1 with the techniques presented in Method 3. Accepting that an image may have more than one running band this method takes the far right peak value (the running band data) and the next occurring peak to its left (the expected point of the data belonging to other running band or non-running band data) and calculates a midpoint position $1/7^{th}$ of the total width (14%). As displayed in Figure 9.17 this calculated value is similar to the one calculated when applying Method 3 however, allows the potential to detect multiple running bands.

Each value calculated from the methods is applied to the image in suppressing non-running band data. From this, a binary map was formed and compared to a manually produced binary map of the running band. An example of a binary map used to evaluate the suppression value is

presented in Figure 9.18. Table 9.3 displays the tabulated results of the histogram analysis algorithms performance when the method was used with data set 2. Table 9.4 displays the tabulated results of the histogram analysis algorithms performance when the method was used with data set 3. Appendix 7 presents an extract from the data set for Method 1 on which the tables were formed



Original image



Binary map used for comparison

Figure 9.18 Example of a manually produce binary map used to evaluate suppression value performance.

.

| Method | False Positive Error (%) | False Negative Error(%) |
|---|---|---|
| 1 | 0 | 100 |
| 2 | 0 | 45.96151438 |
| 3 | 0 | 8.889557192 |
| 4 | 0.532661266 | 5.868958292 |

Table 9.3 Tabulated results from using histogram analysis to calculate a data suppression value in data set 2.

**Greyscale Spectrum**

| Method | False Positive Error (%) | False Negative Error(%) |
|---|---|---|
| 1 | 0.668282924 | 59.91527775 |
| 2 | 14.12302609 | 34.4991778 |
| 3 | 32.98266921 | 7.849535888 |
| 4 | 12.99740717 | 47.81389691 |

**Blue Spectrum**

| Method | False Positive Error (%) | False Negative Error(%) |
|---|---|---|
| 1 | 10.25325617 | 26.24725084 |
| 2 | 10.25325617 | 26.24725084 |
| 3 | 10.25325617 | 26.24725084 |
| 4 | 10.25325617 | 26.24725084 |

**Green Spectrum**

| Method | False Positive Error (%) | False Negative Error(%) |
|---|---|---|
| 1 | 0 | 81.56828539 |
| 2 | 11.04449331 | 33.16351972 |
| 3 | 33.93640654 | 0.218163783 |
| 4 | 10.36825131 | 59.23194137 |

**Red Spectrum**

| Method | False Positive Error (%) | False Negative Error(%) |
|---|---|---|
| 1 | 0 | 93.20159085 |
| 2 | 11.28148874 | 40.06458818 |
| 3 | 35.81590599 | 18.59311672 |
| 4 | 3.66117329 | 57.36274545 |

Table 9.4 Tabulated results from using histogram analysis to calculate a data suppression value in data set 2 in different colour spectrums.

### 9.4.2.2 Accumulation of Column Data

Accumulation of the images column data shows a clear separation consisting of the running band and non-running band data. With this representation of the image data, the required threshold for each running band is displayed as the points with the steepest gradient. Figure 9.19 shows the accumulation of the columns with the addition of the differential plotted. The differential shows that the despite some noise the points with the highest gradient are similar to the expected value for suppression of the non-running band data.

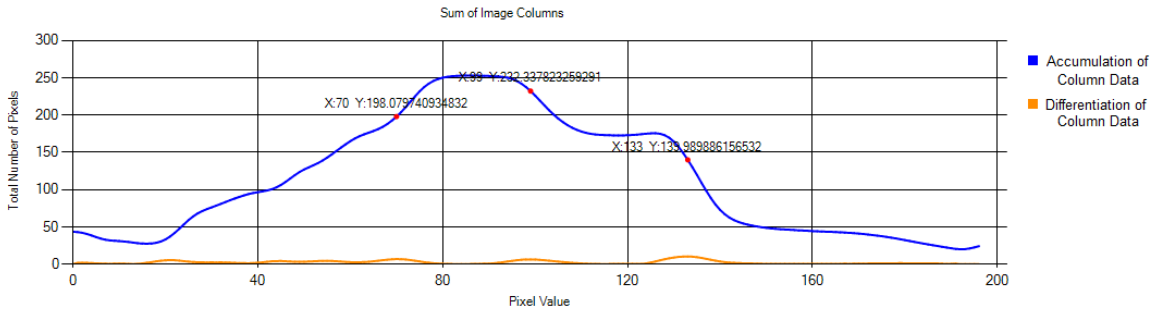



Figure 9.19 Graph showing the greyscale accumulation of columns of the head of the rail and the differentiation of the column data. Markers show the calculated suppression points for suppression of non-running band data.

Multiple running bands are easy to detect by locating the peaks of the gradient, after smoothing to remove noise. This method is more susceptible to noise, when a running band is not linear and varies along the height of the image. In such cases, the gradient point can become more difficult to differentiate from noise. Despite this potential flaw, utilising this method is far more accurate when multiple running bands are present within the image. It could also be argued that if the running band does vary considerably preventing accurate classification of the running band data than a defect must be present requiring further inspection.

As both methods of histogram analysis and accumulation of column data use similar steps there are no real execution time differences to compare, therefore each method can only be evaluated by its application on each test set. Suppression values calculated from the accumulation of the

column data is calculated using the rule set presented in Equation 9.4. This rule set allows a maximum of 3 running bands to be classified within the image.

$$SV = \Delta P[0 + i] + \Delta P[n - i]$$

Where:
$\Delta P$ = Differential of the Sum of the Columns > Threshold
$SV$ = Suppression Value
$n$ = total number of DiffPeaks
$i$ = Running Bands Present = ($\Delta P\ Count$ / 2)

Equation 9.4 Equation to evaluate if a differential peak value belongs a running band or non-running band data set.

### 9.4.3 Initial Results

Table 9.3 and Table 9.4 (*page 237*) display the results from the four methods employed to evaluate the calculation of a data suppression value from the available histogram data. They show that by using the available histogram data, an appropriate value can be calculated to successfully suppress non-running band data. The more adaptable Method 3 and Method 4 algorithms are more successful on the first test set (data set 2). With Method 4 providing the best false negative error with a small sacrifice in the false positive rate.

The second test set results (data set 3) are less successful, with a much higher false positive rate achieved as the false negative rate decreases. To expand on the results, all threes spectrums were analysed alongside the greyscale image, to inspect closer on the potential cause of the error. These results show that the data forming the head of the rail is not ideal. The blue spectrum constantly falls through to a default state, where, only one peak is classified. In such cases, the non-running band takes up a minor part of the image data and falls below the calculated threshold. Both Green and Red spectrum analysis shows a large error rate for all methods.

Table 9.5 and Table 9.6 present the results from using the accumulation of the column data to calculate a data suppression value. Similar to the histogram analysis methods, when applying this method to the first lower quality test set it is more successful than when applying the same method to the higher quality test set. The results how that while, this method can calculate an approximate value, the algorithm is not as accurate as histogram analysis.

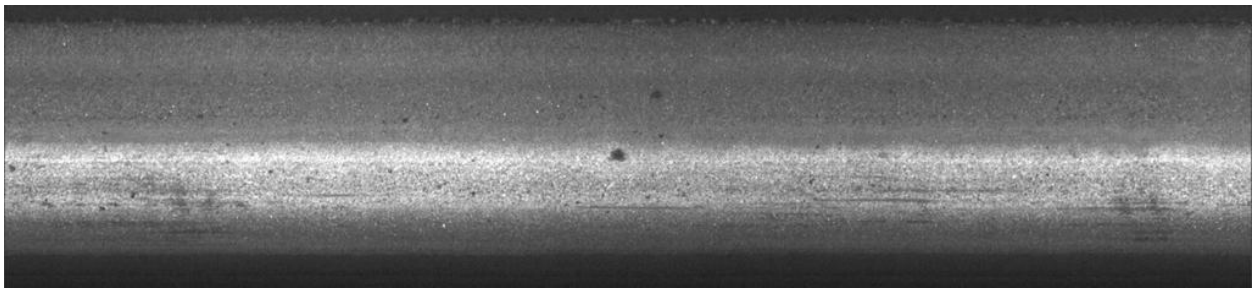
| Image | True Positive | False Positive | True Negative | False Negative |
|---|---|---|---|---|
| 1 | 425426 | 0 | 338414 | 43072 |
| 2 | 456150 | 0 | 320232 | 5954 |
| 3 | 451307 | 0 | 325484 | 5545 |
| 4* | 0 | 0 | 0 | 0 |
| 5 | 370603 | 0 | 299470 | 104071 |
| 6 | 465557 | 8177 | 310353 | 2345 |
| 7 | 484079 | 0 | 238253 | 35428 |
| 8 | 504293 | 27674 | 249697 | 672 |
| 9 | 422476 | 0 | 293995 | 45385 |
| 10 | 456807 | 0 | 273022 | 60699 |
| 11 | 462494 | 0 | 269475 | 54463 |
| 12 | 388969 | 0 | 279052 | 126603 |
| 13 | 410782 | 0 | 296662 | 87180 |
| 14 | 479563 | 44876 | 269831 | 354 |
| 15 | 484956 | 96039 | 229915 | 98 |
| 16 | 448113 | 0 | 320257 | 38542 |
| 17 | 397615 | 0 | 300196 | 105005 |
| 18 | 500982 | 80015 | 221334 | 485 |
| 19 | 488116 | 0 | 309974 | 12918 |
| 20 | 496909 | 69709 | 243821 | 569 |
| Average | 429759.9 | 16324.5 | 269471.9 | 36469.4 |
| Error | | 3.798517 | | 13.53366 |

*Cells marked Red are where the algorithm was unable to detect a running band

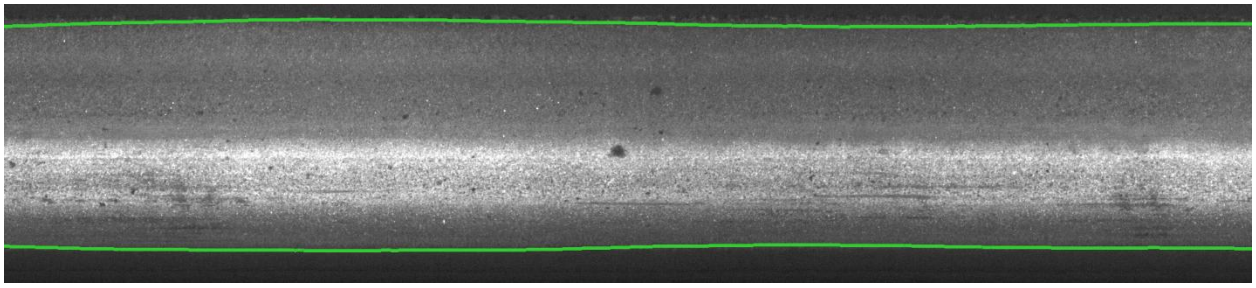
Table 9.5 Results from using the accumulation of the column data to calculate a data suppression value for data set 2.

| Image | True Positive | False Positive | True Negative | False Negative |
|---|---|---|---|---|
| 1 | 983423 | 675173 | 583640 | 324 |
| 2 | 1710216 | 7142 | 299613 | 2869 |
| 3 | 678446 | 1284839 | 71914 | 1 |
| 4 | 1581172 | 383457 | 51338 | 33 |
| 5 | 468310 | 67185 | 68928 | 377 |
| Average | 247662.9 | 117530.6 | 50325.25 | 161.35 |
| Error | | 47.45587 | | 0.320614 |

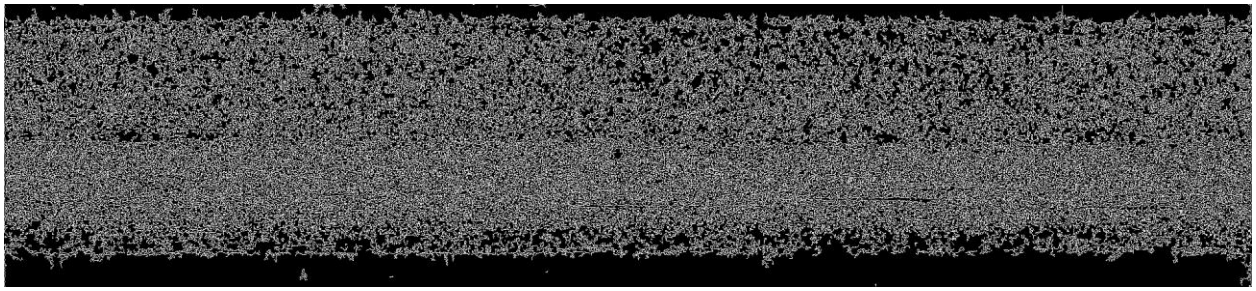
Table 9.6 Results from using the accumulation of the column data to calculate a data suppression value for data set 3,



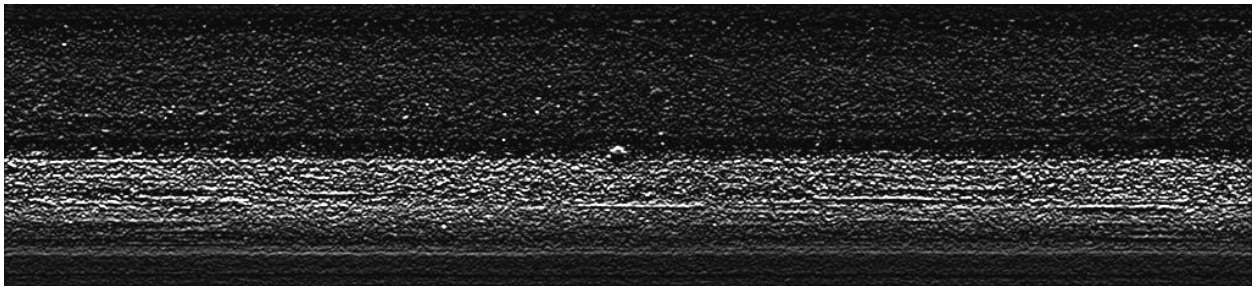
Original image



Data suppression through quantitative analysis (Running band can be determined)



Canny edge detection (Threshold = 101, linking = 39) (Running band indeterminable)



Sobel edge detection (X-Order = 0, Y-Order = 1, Aperture Size = 7) (Running band indeterminable)

Figure 9.20 Comparison of Canny, Sobel, and the data suppression through quantitative analysis algorithms in running band classification.

The poorer performance in classifying the running band in the higher quality second test set (data set 3) is echoed in both algorithms. Despite this the algorithm employing histogram analysis has a higher success rate, therefore it is the most appropriate algorithm to be employed in the suppression of the non-running band data. Figure 9.20 demonstrates the ability of data suppression through quantitative analysis in comparison to Canny and Sobel algorithms in running band classification.

As histogram quantitative analysis allows for the suppression of non-running band data it allows classification of running band edge cues. The accuracy at this stage it variable as it is susceptible to noise of running band type data being present in non-running band regions. Suppression of the noise could be applied at the thresholding stage however; complexities of this algorithm being implemented at this level will affect performance. This noise, as its location is distinct from the running band region, can be suppressed when evaluating the edge cues for the classified running band data.

### 9.4.4 Analysing the Running Band Segmented Data

With successful classification of running band and non-running band data, the algorithms employed must be adapted for the classification of running band edge cues. Both the histogram and accumulative algorithms employed detect a threshold value to suppress non-running band data. This means that by scanning across each row of the image a left and right edge cue of the running band region can be located. The edge location of the running band data can be considered true when the data of the row rises and falls below this value.

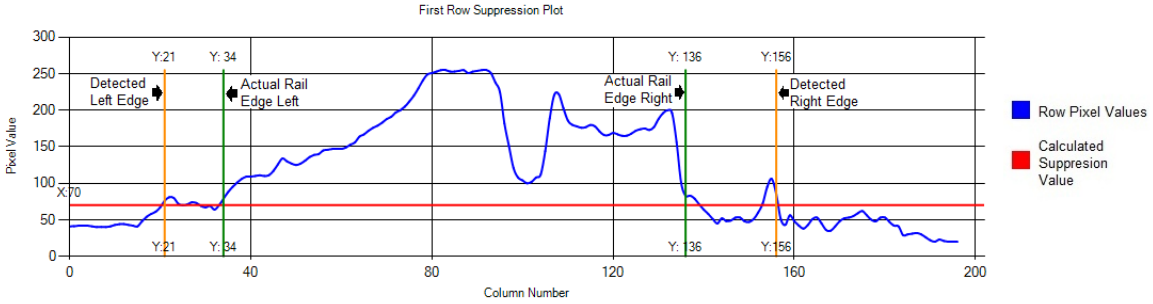



Figure 9.21 Graph showing the first row of test set 1 image 1 with the suppression value to be applied

This is shown in Figure 9.21, as can be seen from the plot; noise is present that prevents the accurate classification of the rail edge without using a filter. Figure 9.22 shows the classification of the running band edge with no filtering applied.

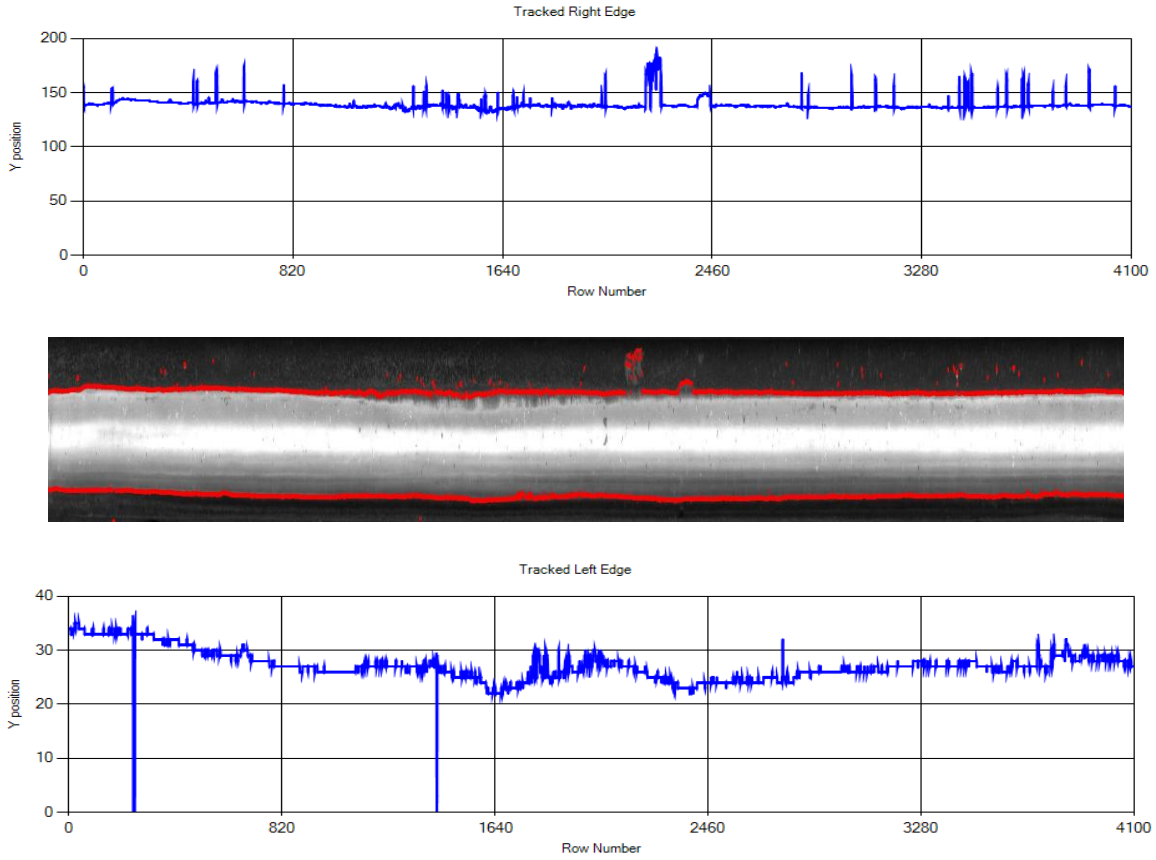



Figure 9.22 Initial detection of the running band edge with noise present.

A simple averaging filter will enable some of the noise to be suppressed. An averaging filter, while efficient in suppressing smaller values of noise, struggles in cases of greater amounts of noise. To suppress the noise present in an image a more advanced noise filter is required. Figure 9.23 shows the initial detection of the running band edge with intermittent noise and an n-th order averaging filter applied as shown in Equation 9.5 where $n = 7$ & $m = 5$.

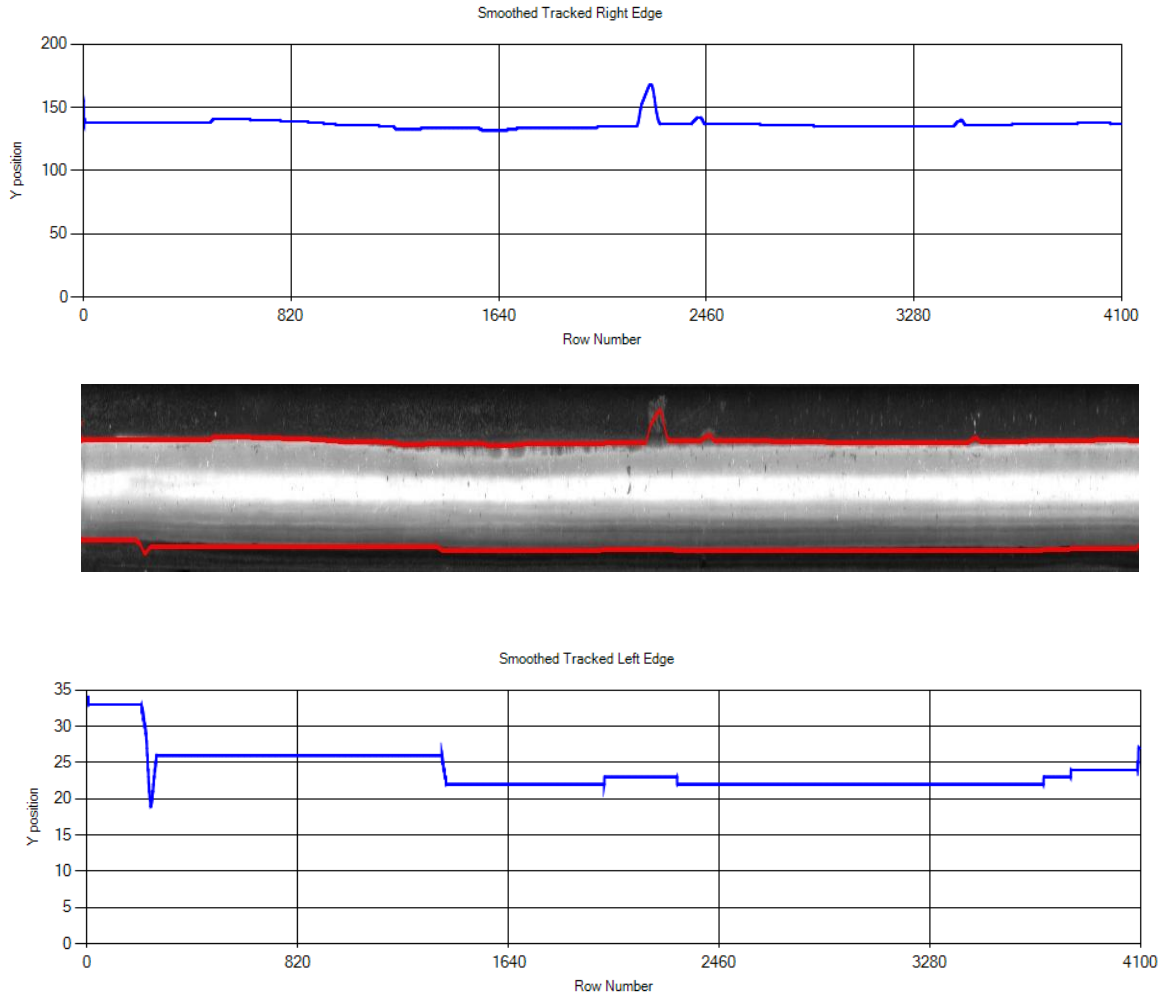



Figure 9.23 Initial detection of the running band edge with intermittent noise and averaging filtering applied.

$$SD_{(x_i)} = \sum_{j=1,i=0+n}^{m,n} \frac{E[x_{i-n}] + E[x_{i-(n-1)}] \ldots + E[x_i] + \ldots E[x_{i+(n-1)}] + E[x_{i+(n)}]}{n}$$

Where:
$E$ = Edge X co-ordinate
$SD$ = Smoothed Data
$m$ = number of averaging passes
$n$ = distance to average data over
$i$ = Array position of Edge X co-ordinate

Equation 9.5 N-th Order averaging filter for suppression of non-associated running band data.

As can be observed while an averaging filter can successfully eliminate noise close to the running band it fails where noise is present at the outer edges of the head of the rail. There is a resultant loss of accuracy as the increased number of passes required to eliminate this noise starts to eliminate any accurate classification of curvature in the running band edge data. Two bespoke noise filters have been designed to suppress noise picked up from miscellaneous white spots upon the head of the rail. Each algorithm can be implemented according to what should be considered a defect of the running band.

The first algorithm assumes that all lighter areas that carry the same characteristics as the running band data, and are over a certain size, can still be considered a possible contact patch with the wheel of a tram. This will allow the suppression of smaller items of noise but allow for larger areas of running band type data to remain for defect analysis. The algorithm shown in Equation 9.6 can successfully eliminate noise present at the edge of the image however; the averaging filter presented in Equation 9.5 is still required to eliminate noise close to the running band edge. The advantage being that the intensity of smoothing using the averaging filter is reduced resulting in classification of the running band data with minimum loss of accuracy.

$$\sum_{j=1,i=1}^{m,n} I\left[x_{i+1}, y_{j+1}\right].X - I\left[x_i, y_j\right].X > DT \Rightarrow \xi_{(x_{i+1}, y_{j+1})} = I\left[x_{i+1}, y_{j+1}\right]$$

$$\sum_{j=\xi.Y+1,i=\xi.X+1}^{m,n} \left\{\left(I\left[x_{i+n}, y_{j+m}\right].X > \xi.X * 0.9\right) \,\&\, \left(I\left[x_{i+n}, y_{j+m}\right].X < \xi.X * 1.1\right)\right\} \Rightarrow$$

$$\xi_{(x_{i+n}, y_{j+m})} = I\left[x_{i+n}, y_{j+m}\right]$$

$$N = \xi_{(x_{i+n})} - \xi_{(x_{i+1})} > ST$$

Where:
$I$ = Running Band Edge Location within the Image
$DT$ = Distance threshold
$SD$ = Data Suppression Threshold
$m$ = Image Height
$n$ = Image Width
$j$ = Image Row number
$i$ = Image Column number
$N$ = Noise

Equation 9.6 First bespoke noise suppression of non-running band data with allowance for non-associated running band data.

Figure 9.24 shows the results of applying the first bespoke noise suppression algorithm. The setting applied for noise suppression where $DT = 10, ST = 30$ and with the averaging filter $n = 5$ & $m = 3$ were used. As can be observed the correctly classified curve of the running band present within the image has not been suppressed. The two larger features of noise remain, allowing for the classification as potential defects for visual inspection.

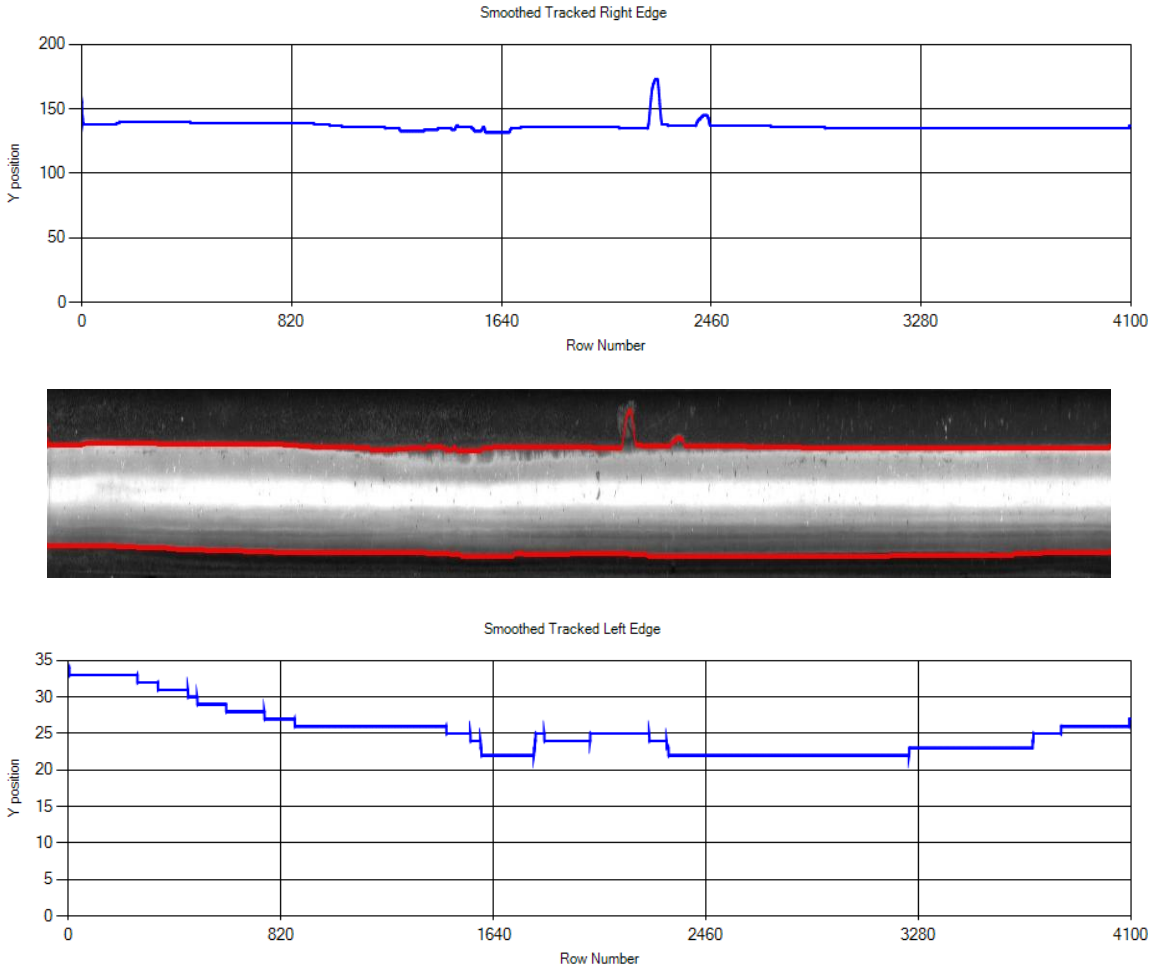



Figure 9.24 Detection of the running band edge with the first bespoke noise suppression algorithm employed

The second bespoke noise filter assumes that the only point of contact with the rail is the largest area of running band data. Any running band data that is not associated with the main running band area is eliminated. While a region-growing algorithm could be employed to achieve this kind of suppression a far more efficient algorithm has been designed. This algorithm is presented in equation form in Equation 9.7; it inspects each point of the running band detected edge looking for large differences in the position from previously identified running band edge cues. If this difference is considered too high, the algorithm identifies if the point is the only occurring edge cue before the opposite side edge cue of the running band. If not, the data before the newly determined edge cue is suppressed.

$$\sum_{j=1,i=1}^{m,n} I\left[x_{i+1}, y_{j+1}\right].X - I\left[x_i, y_j\right].X > DT \Rightarrow \xi_{(x_{i+1}, y_{j+1})} = I\left[x_{i+1}, y_{j+1}\right]$$

$$\sum_{i=\xi.X+1}^{n} (I[x_{i+n}, \xi.Y] < T) \Rightarrow \xi.X = I\left[x_i, y_j\right].X$$

Where:
$I$ = Running Band Edge Location within the Image
$DT$ = Distance threshold
$n$ = Image Width
$j$ = Image Row number
$i$ = Image Column number
$T$ = Running Band Threshold

Equation 9.7 Second bespoke noise suppression of non-running band data with suppression of non-associated running band data.

The second bespoke noise filter requires the pre-condition that the initial edge cue is compared against the final edge cue from the previous image within the set. In such circumstances where the image being analysed is the first, then initial data points should be neglected when examining for potential defects. If this is deemed unacceptable then the average edge cue location should be used as an initial value for $I\left[x_i, y_j\right].X$ where $i = 0$ & $j = 0$ for $I\left[x_{i+1}, y_{j+1}\right].X$.

The results are presented in Figure 9.25 show the effectiveness of the second bespoke noise suppression filter in accurately determining running band edge cues. This method eliminates all noise except that directly associated with the main area of running band data. The algorithm was run with Distance threshold $(DT) = 5$, the averaging filter shown in Equation 9.5 is applied to normalise the data and remove noise that fails bellow this threshold. The averaging filter is run with $n = 5$ & $m = 3$.

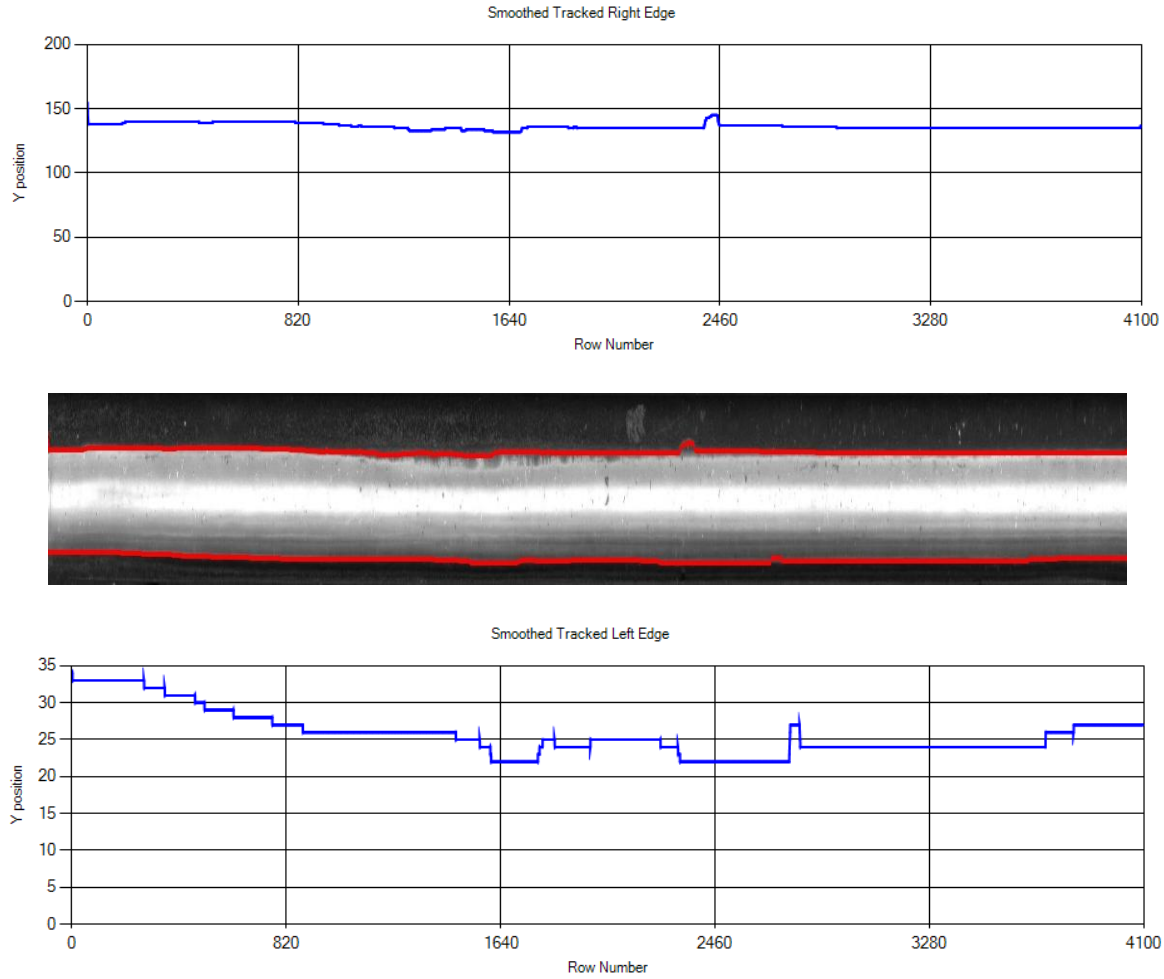



Figure 9.25 Detection of the running band edge with the second bespoke noise suppression algorithm employed.

### 9.4.5 Final Evaluation

The method of classifying multiple running bands was applied to a larger test set of 5000 images (data set 4), see chapter 6.2 Data Sets (page 142) for reference. These images were acquired from the Sheffield Supertram network using the acquisition system described in Chapter 4. All 5000 images were processed using the techniques described in Chapter 7 for the detection of the railhead. The possible false classification of the rail head were compared to errors detected within the running band analysis to exclude repeated false classifications this resulted in 4638 images for running band analysis.

Furthermore, in this evaluation, three different characterised running bands were detected. The first being the whole running band as evaluated in the initial investigation. The second and third were running bands that occur within this wider area. These sub-running bands are hard to visually segment from the images however become visible when computational suppression of the data is applied. Figure 9.26 demonstrates these three running bands detected. Appendix 8 shows the thresholding results from the running band detection.

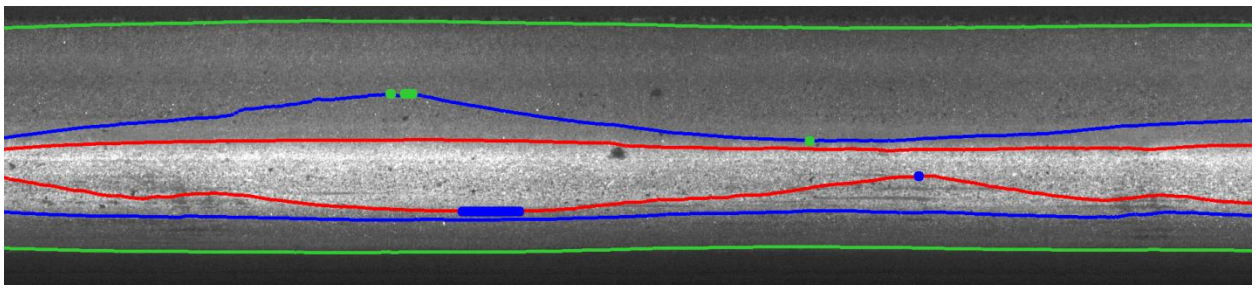


Figure 9.26 Classification of multiple running bands

The inner phase 1 blue running band is the path that only some wheel sets take; the inner phase 2 red running band is the path that a majority of wheel sets take across the railhead. While the inner phase 2 is visually determinable from the image, the inner phase 1 is not. As can be seen in Figure 9.26, corrugation is displayed in the inner phase 1 running band and confirmed in the inner phase 2 running band.

The classified running band data was evaluated using two rules to examine for potential defects causing a variation in the running bands location. The first rule presented in Equation 9.8 uses the minimum and maximum-recorded location of the running band in the x-axis, showing a defect if this distance between these two points breached a set value. The second rule presented in Equation 9.10 examined the gradient of the running band as it was processed. If there was a change of pixels values that caused the gradient to exceed a set value, a defect was considered.

In such cases that Equation 9.8 and Equation 9.10 are true, it can be considered that a severe defect is present. In cases in which Equation 9.8 was true the likelihood of corrugation being the potential classified defect was high. In cases in which Equation 9.10 was true the early development of a smaller surface bound defect such as a squat or wheel burn would be considered to be interfering with running band classification.

During this work the running band width was not examined, which remains a potential analysis method to further examine the severity of corrugation.

### 9.4.6 Final Results

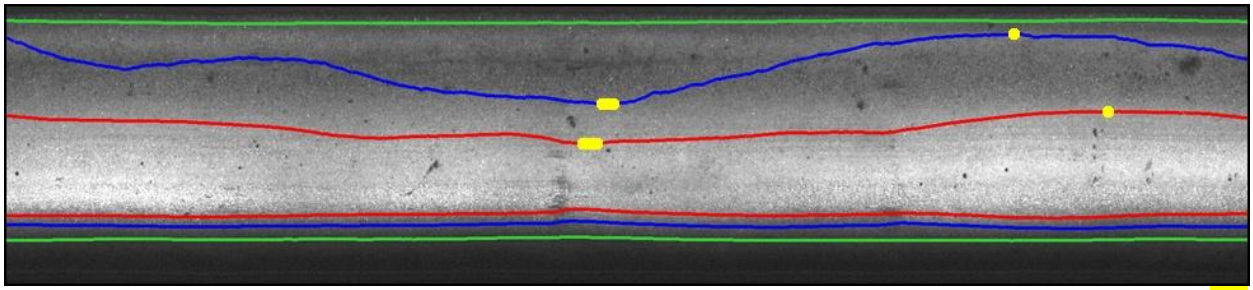
Table 9.7 shows the classified defects and non-classified defects from running band classification using data suppression through quantitative analysis. Classified defects occur when Equation 9.8 (*page 259*) is satisfied and the characteristics of corrugation are shown. Non-classified defects occur when Equation 9.10 (*page 260*) is satisfied and Equation 9.8 is not satisfied. Due to the high volume of classified corrugation images 1 in every 10 images was visually assed for classified defects, and for non-classified defects each image was inspected.

| | | | | Outer Running Band | | Inner Phase 1 Running Band | | Inner Phase 2 Running Band | | |
|---|---|---|---|---|---|---|---|---|---|---|
| File | Images | REE | AI | CD | ND | CD | ND | CD | ND | Error Rate |
| 12 | 1000 | 16 | 984 | 86 | 0 | 270 | 0 | 327 | 0 | 0 |
| 14 | 1000 | 25 | 975 | 46 | 1 | 553 | 90 | 612 | 55 | 14.97 |
| 17 | 1000 | 157 | 843 | 147 | 0 | 454 | 0 | 523 | 0 | 0 |
| 53 | 1000 | 69 | 931 | 185 | 1 | 525 | 1 | 771 | 4 | 0.64 |
| 56 | 1000 | 95 | 905 | 150 | 1 | 785 | 92 | 682 | 2 | 10.49 |
| Total | 5000 | | 4638 | 614 | 3 | 2587 | 183 | 2915 | 61 | 26.11 |

REE = Rail Edge Detection Errors
AI = Total Images after Admissible Errors
CD = Classified Defects
ND = Non-Classified Defects

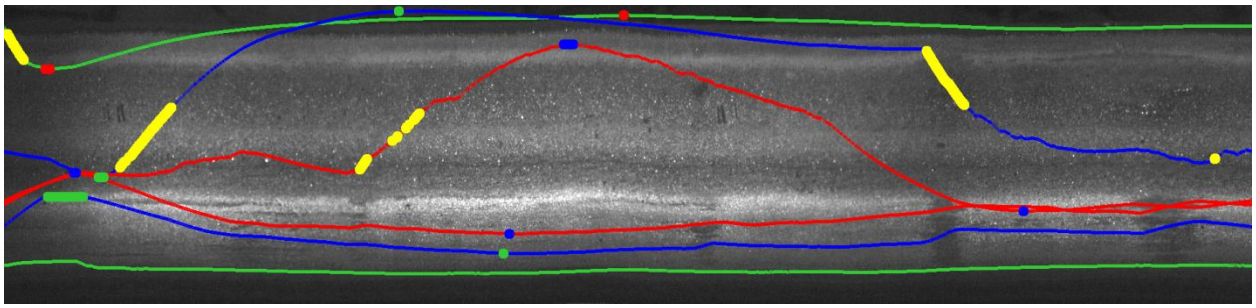
Table 9.7 Results of running band analysis using data suppression through quantitative analysis.

Visual assessment of the results showed that the high volume of corrugation detected was correct and that the corrugation was present throughout the image sets. All the non-classified images were a direct result of corrugation however in such cases Equation 9.8 was not satisfied, as the change in the running band was not significant to breech the set threshold. Figure 9.27 demonstrates a classified defect i.e. corrugation and Figure 9.28 shows a non-classified defect as a direct result of corrugation.



Classified Defect=

Figure 9.27 Example of a classified defect i.e. corrugation.



Non-Classified Defect=

Figure 9.28 Example of a Non-classified defect as a result of corrugation.

The large variation in running bands shown in Figure 9.28 are caused by the development of data with similar characteristics of the inner sub-running band (marked red) in the out sub-running band (marked blue). With noise filtering and smoothing a large ballooning in the tracked edge can be observed. While visually the same path of running band is not seen assessing the image using a sliding threshold the tracked edges can be seen, a similar effect can be observed in Appendix 8.

## 9.5 CLASSIFICATION AND GRADING OF DEFECTS

The pre-processing algorithms presented allow the noise to be filtered accurately before the running band edge is inspected for classification of defects. Defects are classified by a set of rules based on previous methods of defect categorisation. Rules are also designed to make use of data catalogues and examine change over time using standard deviation calculations.

The defect classification algorithms are separated into 3 main fields. In-Image Defects (IID), which can be classified using analysis from individual images. Multiple Image Defects (MID) where defects can be classified only by the examination of previous image data. Time Image Defects (TID), that required the catalogue of multiple passes over the same point of tramway network and statistical analysis comparing data acquired from multiple runs. The techniques employed for detecting defects depends on MVS setup and image details.

If an individual image contains rail data over a small distance, then evaluating its properties individually may not result in accurate classification of defects. Combining multiple smaller images will allow IID classification to be applied. While in principle this is the same as MID defect analysis MID is designed to work over a greater range.

### 9.5.1 In-Image Defects (IID)

IID detection examines three characteristics or the running band edge cues located within an image. The first is the overall change within the image looking at the maximum and minimum positional change. This would allow quick classification of a defect present upon the head of the rail. The second examines the starting point of the running band as it enters the image and the finishing point as the running band leaves the image. A large enough variation in positions would denote a shift in running band possibly relating to a shift in ballast or incorrect placement of support for the rail. The third examines the differential plot of the position of the edge cues, while this may produce similar results to the first method it allows accurate placement within the image and better classification from potential size of the defect. Examining the differential plot also eliminates potential false defect classification produce by some types of infrastructure.

These equations presented in Equation 9.8, Equation 9.9, and Equation 9.10, all require thresholds based upon the type of MVS systems employed. It is implausible to give a set of fixed constants for all MVS due to the varying factors of image quality and image size. While a railhead may be 100 pixels wide in one image it can quite easily be 200 pixels in another due to change in the field of view due to lens, camera variations and mounting positions.

$$\sum_{j=1,i=1}^{m,n} E_j \geq E_i \Rightarrow \xi_{max}$$

$$\sum_{j=1,i=1}^{m,n} E_j \leq E_i \Rightarrow \xi_{min}$$

$$\xi_{max} - \xi_{min} > T_D \Rightarrow D$$

Where:

$D$ = Defect Present
$T_D$ = Threshold for Defect Classification
$E$ = Edge Cue Location
$i = j$ = Edge Cue Location Under Examination
$m = n$ = Data set Size
$\xi$ = Unique Edge Location

Equation 9.8 IID formulas for defect classification examining running band shift.

$$\xi_{Start} = E_1$$

$$\xi_{Exit} = E_n$$

$$|\xi_{Start} - \xi_{Exit}| > T_D \Rightarrow D$$

Where:
$D$ = Defect Present
$T_D$ = Threshold for Defect Classification
$E$ = Edge Cue Location
$n$ = Image Height
$\xi$ = Unique Edge Location

Equation 9.9 IID formulas for defect classification examining running band entrance and exit change.

$$\sum_{i=1}^{n-x} \Delta(E_i - E_{i+x}) > T_D \Rightarrow D$$

Where:
$D$ = Defect Present
$T_D$ = Threshold for Defect Classification
$E$ = Edge Cue Location
$i =$ Edge Cue Location Under Examination
$n$ = Image Height
x = Distance of which to perform Differentiation over

Equation 9.10 IID formulas for defect classification examining running band differential change throughout the image.

### 9.5.2 Multiple Image Defects (MID)

MID detection examines for potential degradation or shifts in ballast that would not be visible by examining data within an individual image. The range of the running band shift examined in Equation 9.8 is expanded to reflect this in Equation 9.11. The data examined is taken over a pre-determined range. For example if a shift of 20 pixels was considered a defect over a range of 3 meters then $T_D$ = 20 and the $m$ would be calculated based upon the FOV measurements from the MVS and images involved.

$$v = \frac{X}{I_h}$$

$$\sum_{k=1,j=1,i=1}^{v,m,n} E_j \geq E_i \Rightarrow \xi_{max}$$

$$\sum_{k=1,j=1,i=1}^{v,m,n} E_j \leq E_i \Rightarrow \xi_{min}$$

$$\xi_{max} - \xi_{min} > T_D \Rightarrow D$$

Where:
$I_h$ = Image Height in mm
X = Distance for Inspection
$D$ = Defect Present
$T_D$ = Threshold for Defect Classification
$E$ = Edge Cue Location
$i = j$ = Edge Cue Location Under Examination
$u$ = Number of Images to examine across
$n$ = m = Data Set Size
$\xi$ = Unique Edge Location

Equation 9.11 IID formulas for defect classification examining running band shift over multiple images.

### 9.5.3 Time Image Defects (TID)

A unique ability of employing a MVS in linear feature detection in the application of running band analysis is the ability to track changes of data over time. With a sufficient data library, the increasing size of the running band can be correlated with running traffic on the system. This will allow a much better life profile to be generated for track based on real world data.

With this type of analysis, the key issue is the matching of location to the each set of data. Despite control of running location during inspection routines, identical image sets will be acquired. The running band data cannot be used in itself to match location, as there should be a significant change in key features, such as defects, as maintenance is carried out to remove them.

Using key features within an image set, such as start and end of embedded track sections, crossings, and associating the images with GPS data, good matching between acquired data sets can be assumed. Together with data location matched simple statistical analysis such as standard deviation, F-test (comparing variance (Lomax, 2007)), and T tests (Hurst, 1995) over set distances can be applied. The results from these various methods will be representative of changes in running band due to wear and can be directly correlated over time.

It is not possible at this time to determine what change could be classified as a defect. It has yet to be determined if any form of matching can be obtained by using GPS data and whether feature-matching algorithms can successfully be used to compare data.

## 9.6 SUMMARY

A novel set of algorithms have been presented that can classify the edge cues of running band data. The algorithm uses image data to correlate its settings and allow accurate classification. This suppresses noise from environmental factors such as grinding marks, changes in light intensity, and weather interference. Bespoke noise suppression algorithms have been developed to work alongside the edge cue detection algorithm and reduce false classification from noise present on the railhead. The algorithms have been demonstrated to outperform standard edge detection methods in running band classification.

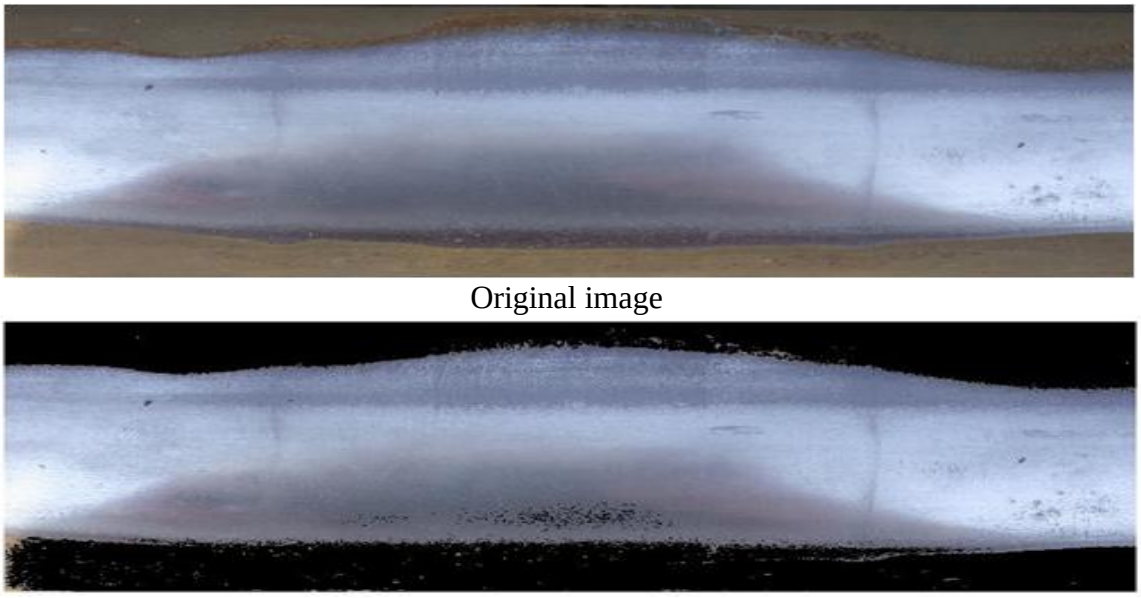
This algorithm is robust for different MVS and has been optimised to reduce computation time. Algorithms to analyse the edge cue data from running band classification have been developed and presented. The analysis algorithms present has been shown to successfully identify corrugation. These algorithms allow scaling to real world measurements based upon MVS statistics.

# CHAPTER 10 DETECTION OF SURFACE BOUND ABNORMALITIES WITHIN THE RUNNING BAND – SQUAT AND WHEEL BURNS

## 10.1 INTRODUCTION

Surface bound abnormalities such as squats and wheel burns have a similar appearance on the railhead and cause discolouration of the running band. Image analysis techniques that can be employed for their detection are therefore shared. Due to them having a similar statistical nature, they are treated as a single problem for a designed algorithm. The most common feature allowing an algorithm to distinguish between the two types is the occurrence of wheel burns on both rails at the same location. Squat and wheel burn analysis impacts running bad analysis and segmentation.

It is assumed that for successful squat and wheel burn analysis, that the running band data has been successfully distinguished from the non-running band data. Figure 10.1 demonstrates the expected segmentation of running band and non-running band data that will be achieved before squat and wheel burn detection is applied.



Original image

Segmented running band data

Figure 10.1 Results from successful image segmentation of running band data

A bespoke approach has been developed to detect abnormalities within the running band data. This novel approach uses previous research into running band classification (see Chapter 8 ). Running band analysis produces a threshold value that suppresses non-running band data. As squat and a wheel burns have a similar threshold characteristic to non-running band data, the threshold value can be used to classify their presence within the running band. By using the classified running band each line is scanned until a datum point is located that carries non-running band data characteristics. Once detected, a region-growing algorithm is used to classify the size of the abnormality. This size can then be used to classify severity.

Differentiation between squats and wheel burns still requires the occurrence of the same abnormality of the railhead to occur on both sides of the rail in a similar location.

## 10.2 INVESTIGATIVE RESEARCH

Two avenues of research were followed in the classification of squat and wheel burn data. The first examines texture analysis. Using neural networks investigation was implemented to examine if the texture of the squat or wheel burn could aid in the determination of its characteristics against that of the running band data. Both greyscale and colour images were used.

Statistical filters were included due to their ability to classify running band data. Mean testing and standard deviation algorithms were implemented, with two purposes. The first, to see if squat or wheel burn data could be determined from running band data given a statistical norm for the abnormal data. The second role investigated if the same methods, given a statistical norm for the running band, would classify the squat or wheel burn data as not belonging to the running band data.

Building on the running band classification using quantitative analysis (Chapter 8 ), the investigation into thresholding values was extended to see if the classified running band edge could be used to implement a dual data elimination method in which, known running band and non-running band data either side of this could be suppressed . Similarly, as with the initial investigative research within Chapter 8, a region-growing algorithm was implemented for region classification. The region-growing algorithm was developed to produce a novel genetic growth algorithm for region classification.

Results were visually and statistically assessed over a series of images to evaluate the most robust and suitable algorithm for further development. All images that the algorithms were tested on were part of data set 1 unless stated otherwise, see chapter 6.2 Data Sets (page 142) for reference. They are from a selection of stock imagery totalling 969 images the content varies in image quality, camera angle, and lighting conditions. Those image presented show typical results from the data set and where possible have a similar FOV and camera angle to that of the final data set.

### 10.2.1 Back-propagation Neural Networks

A neural network (NN) is ideally suited for texture recognition due to its ability to self adapt and become more robust through training on varied image sets making the method more appealing than that of more complex statistical method. Although as discussed, its execution time can be large for analysing the greater quantity of data involved within the running band.

Further advantages of applying NN's are not only their performance but also their capability of analysing the size and scale of the defect using the same algorithm as used for its data classification, reducing the steps involved. There is however, a disadvantage in that this technique carries, it only provides the ability to calculate the total amount of data that corresponds to squat defects resulting in the possibility of false scale being reported by two or more small squats or wheel burns being summed into a larger more severe defect.

As discussed by LeCun et al. (1998) the data selected and its formation in training a NN can have a severe effect on its performance. In an ideal world, the more data provided to a NN the more reliable its ability to correctly classify positive from negative inputs. In practice however, this is not always true as it depended on the quality of the training data. This can be clearly demonstrated with the application of a NN to texture analysis.

#### 10.2.1.1 Back-Propagation Neural Network for Greyscale Image Texture Recognition

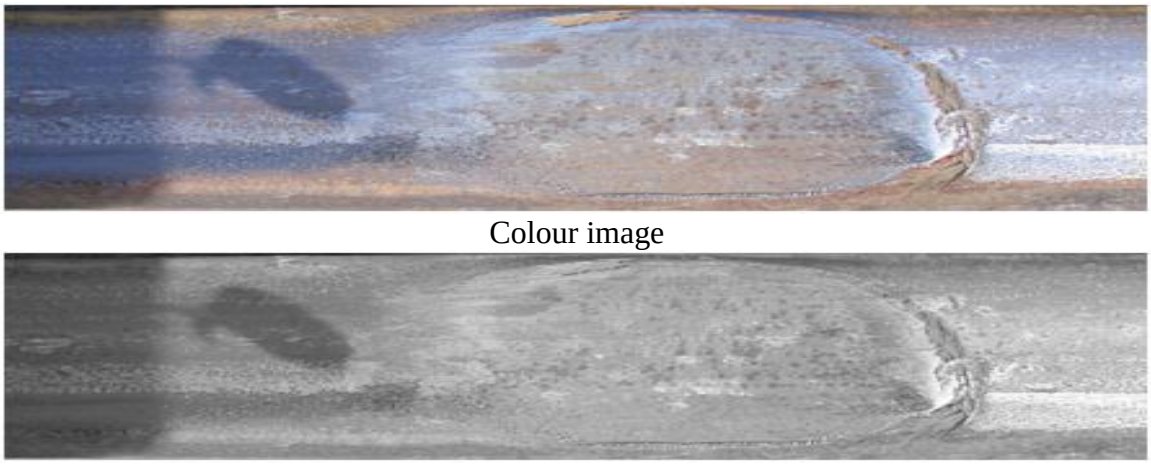
In this application a, NN has been provided a training set of data formed from the image. It consists of two sets. The first data set represents squat and wheel burn texture and the second data set represents running band data with no defects present. As previously stated as the data to be analyses will only consist of running band data any false classification of non-running band data may be ignored.

The NN task was to classify any data that does not conform directly to that of running band data as a potential for a defect. Any areas that show large a large cluster of classifications should be considered a defect such as a squat or wheel burn. A mask size of 16x16 pixels was used to train the NN. This mask was applied to each pixel of the images and the NN then asked to classify the results.

Appendix 9 Shows the results from the NN, with each image is the resultant binary map in which, white areas are the classified squat data. A NN was trained for each image to ensure that the data for classification was consistent with the data to be classified. While the results show some correct classification of the squat data, the results are neither satisfactory nor consistent.

There are factors that could have affected the results of the NN. It can be reasonably assumed that the data presented to the NN is not sufficient for accurate classification.

Figure 10.2 shows a comparison of a colour squat image and a greyscale squat image. It is clear from the greyscale image that the data that forms the centre of the squat is not easily determinable from its surrounding data. Improved image quality could address this problem. Should the data become clearly determinable the use of Gaussian filters to pre-process the image should be considered.



Colour image

Greyscale image

Figure 10.2 Comparison of colour squat image to greyscale squat image.

While the output from a desired Gaussian filter will have the effect of blurring the image, what this means is that the data that forms the running band will become more consistent. Similarly, with the squat data this will allow a NN to have greater potential in converging and allow classification of the data correctly.

#### 10.2.1.2 Back-Propagation Neural Network in Colour Imagery Texture Recognition

To address the possibility of image quality and data quantity playing a negative role on a NN performance further investigation was made into the use of colour images. This triples the amount of data for classification, with the potential of increasing NN performance. The same training sets were used however, each section of image data had the red, green, and blue spectrum analysed.

There are several ways of implementing a NN with the training data. Two methods were selected, the first involved a set of three input groups for a NN where the three colour spectrums of data were fed into separate grouped input nodes (Layer 1 and 2) and then the results of these where analysed again in a third layer. The second approach was to remain with a single grouped input NN and join each 3 spectrums of data in a specific order and feed the full data set into the NN.

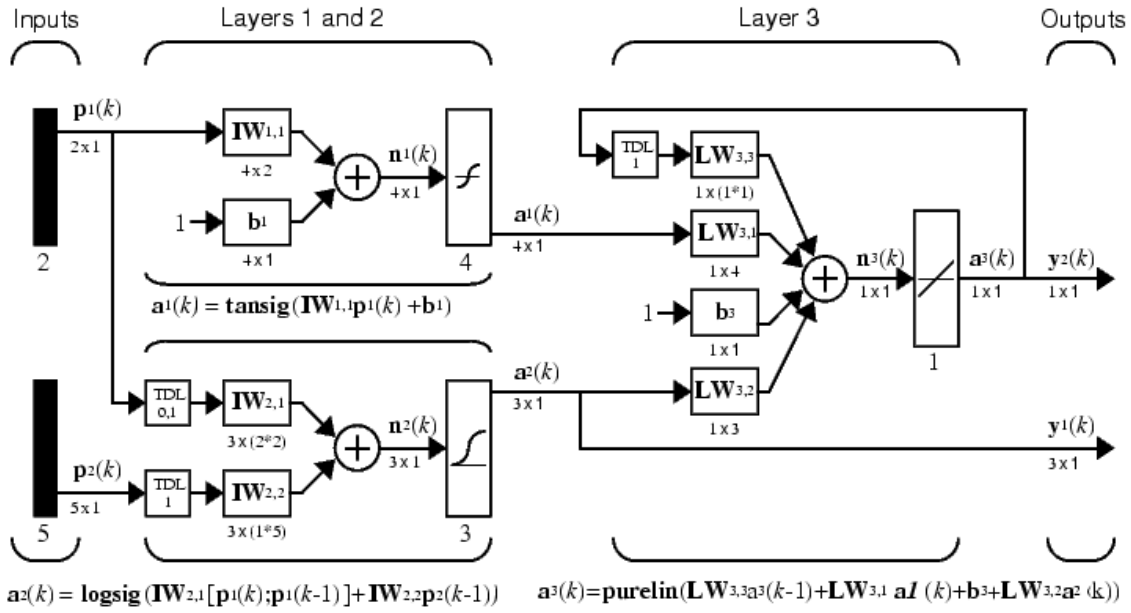



Figure 10.3 Multiple input neural network design example

The results from the 3 input NN demonstrated a failure to determine the centre squat data from the running band data. The poor classification ability was echoed through all three-colour spectrums producing poorer performance to that of the NN greyscale image application. Figure 10.4 shows these results and how, with the added data, the NNs ability to correctly classify the squat and running band data has decreased.

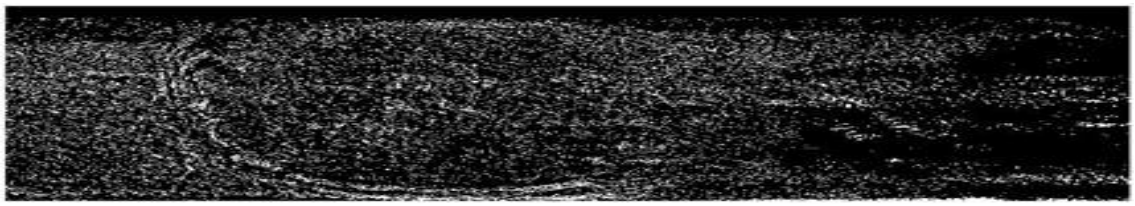


Figure 10.4 Results from three input neural network examining all three colour spectrums.

Significant improvements are shown in Figure 10.5 where all three spectrums of colour are concatenated into an individual data set. This has the effect of providing the NN with more information about the context of data that forms a squat. The NN now has data sets that show a distinguishing difference from the squat data (positive/white) and the running band data (negative/black) Figure 10.5.

Original Image

Neural Network Results

Figure 10.5 Results from single input neural network examining all three colour spectrums.

The results suggest that if the image quality was improved to show a distinct difference between squat data and running band data in a greyscale image, then a NN should be able to classify the two textures.

### 10.2.2 Statistical Filters

The use of statistical filters in the classification of running band and non-running band data has already been examined. The ability of such methods to classify running band data has been shown. It is therefore logical that the same techniques can be applied to the detection of squat and wheel burn detection. A similar testing approach was employed, however, since it is assumed that running band data can be successfully determined this data is suppressed in the results.

A key factor in the use of a statistical filter is its inability to work with images without a uniform lighting source. Histogram analysis and correction could be employed to account for this; however, histogram correction can potentially cause feature corruption altering data that allows features to be easily classified.

A mean testing algorithm will be employed in comparison to a more advanced standard deviation algorithm using the concatenation of the summation of rows and columns of the mask and filter. Each algorithm will employ two tests. Test 1 will attempt to classify data belonging only to squats present. The second will examine for running band data. While neither test is more efficient as each examines all pixels within the running band, the comparison of methods will show if squat data is identical over multiple defects and if the running band data is consistent.

The tests employ two binary maps for evaluating the statistical filters performance in classifying the required data. A suppression map is used to prevent non-running band data either side of the running band being analysed (this non-running band data does not include the squat data). A manually produced binary map of the running band data and separate squat data is then utilised to evaluate the statistical filters performance.

The mask sizes to be compared are ‘3x3, 5x5’ and ‘10x10’. These mask sizes are deemed the most effective from the running band classification test carried out in Chapter 8. It is useful to note that the use of different shaped masks could increase an algorithms ability. This should be

investigated if the method proves to be the most successful. Receiver operating characteristics curves are formed from comparison of the results to the manually produced binary map.

#### 10.2.2.1 Squat Classification Filter

##### 10.2.2.1.1 Mean Testing Results

Analysis of the results in Figure 10.7 show that the mean testing algorithm is successful at determining squat based data from running band data. In this particular test, the smaller 3x3 mask size was the most proficient. Analysis of the Receiver operating characteristics (ROC) data suggest that a setting above 65% error threshold would provide the best ratio of true and false, positives and negatives, Figure 10.6. While, the percentage error data suggest a lower setting of 30% allowed deviation in threshold settings would produce results that are more favourable in term of less false negatives.

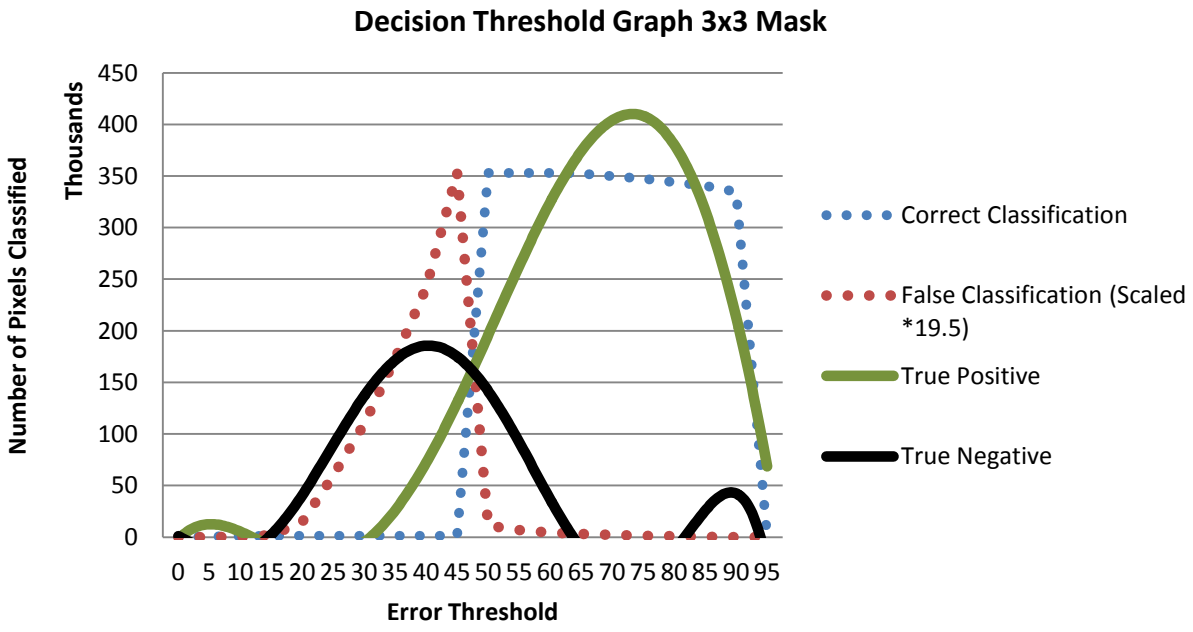
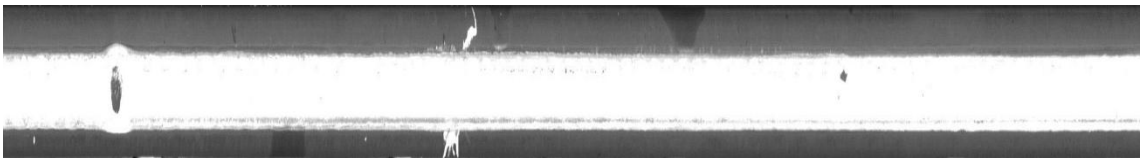



Figure 10.6 ROC decision threshold graph for averaging 3x3 mask.

As can be observed from Figure 10.7, the results are comparable with the lower 30% error threshold providing the better results. However when using a higher 45% error threshold, similar to that suggested by the ROC data, areas of the running band are falsely classified as squat data. There is a significant difference in the methods effectiveness to pick up the smaller squat on the right hand side of the image. This would suggest that the squat data is not uniform across defects and each defect will have a portion of unique or semi-unique data that does not occur with all such defects.



Original image

Results from 30% error of squat filter

Results from 45% error of squat filter

Figure 10.7 Results from applying mean comparison statistical filter using squat data.

#### 10.2.2.1.2 Standard Deviation Testing Results

As demonstrated in previous testing (8.2.2.1 Statistical Filters) comparable results should be expected from a standard deviation filter. Improvements are down to the more complex algorithm comparing pixel to mask data more accurately. In this application however, this is not true. This more complex algorithm has the inverse effect of classifying the pixel data too well meaning more false negatives occur for each true positive than the mean testing algorithm. This can be observed by examining the results of which similar true positive matches were made in Table 10.1.

| Mean Algorithm 3x3 Mask | | | | |
|---|---|---|---|---|
| Error (%) | TP | FP | FN | TN |
| 10 | 752 | 580 | 3 | 352927 |
| 30 | 1113 | 216 | 100 | 352830 |
| 60 | 1261 | 71 | 5630 | 347300 |
| 90 | 1325 | 7 | 18127 | 334803 |

| Standard Deviation Algorithm 3x3 Mask | | | | |
|---|---|---|---|---|
| Threshold (K) | TP | FP | FN | TN |
| 10 | 759 | 573 | 13 | 352917 |
| 30 | 1141 | 191 | 664 | 352266 |
| 50 | 1262 | 70 | 7204 | 345726 |
| 70 | 1324 | 8 | 20475 | 332455 |

Where:

- TP = True Positive
- FP = False Positive
- FN = False Negative
- TN = True Negative

Table 10.1 A comparison of mean algorithm and standard deviation algorithm false negatives rate using a squat statistical filter.

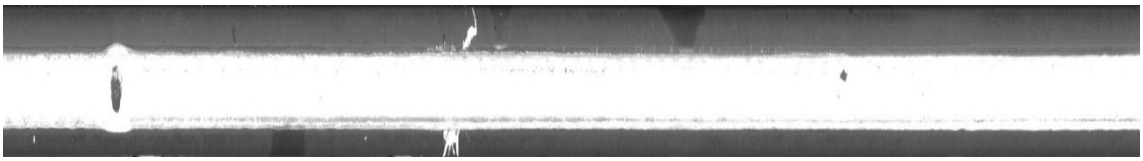
The preferred use of this algorithm is also warranted by the reduced execution time. This would be further suggestive proof that the data forming squats within an image are not consistent enough to allow such an accurate evaluation with a statistical filter examining for such data.

### 10.2.2.2 Running Band Classification Filter

#### 10.2.2.2.1 Mean Testing Results

With the use of a filter for determining running band data, some improvements could be expected over the squat filter results. This is due to the running band having a more uniform data construction than squat data. Analysis of the results supports this theory. There is a clear contradiction between analysis of the statistical threshold data and that produced by ROC analysis.

ROC analysis suggests that using 30% error threshold would produce the best ratio of true and false, positives and negatives. This is true for both the 3x3 and the 5x5 mask sizes. The 5x5 mask size ROC curve shows a small increase in accuracy over the 3x3 mask curve. Statistical results suggest that a higher 40% and 50% error threshold would be better suited for the 3x3 mask and 5x5 mask respectively. While some credit can be given to the statistical results, the false negative rate is unnecessarily high.



Original image

Results from 30% error using 3x3 mask

Results from 45% error using 5x5 mask

Figure 10.8 Results from applying mean comparison statistical filter using running band data.

The differences between the 3x3 mask and 5x5 mask are minor. Visual inspection of the results from Figure 10.8 shows that there is a slight increase in the detail that a 3x3 mask picks up, which may help in accurately evaluating the size of a squat defect.

#### 10.2.2.2.2 Standard Deviation Testing Results

When applying the statistical filters as a squat filter the mean testing algorithms had a greater ability than that of the standard deviation algorithm. This is not the case with the application of the statistical filter to running band classification. As when applying a statistical filter for the classification of the running band data, the standard deviation algorithm shows improvement. This is evident from Table 10.2 where there are a reduced number of false negative classifications.

| Mean Algorithm 3x3 Mask | | | | |
|---|---|---|---|---|
| Error (%) | TP | FP | FN | TN |
| 10 | 312845 | 40085 | 0 | 1332 |
| 30 | 344035 | 8895 | 43 | 1289 |
| 40 | 351042 | 1888 | 134 | 1198 |
| 50 | 352909 | 21 | 1004 | 328 |

| Standard Deviation Algorithm 3x3 Mask | | | | |
|---|---|---|---|---|
| Threshold (K) | TP | FP | FN | TN |
| 20 | 313179 | 39751 | 0 | 1332 |
| 60 | 345924 | 7006 | 44 | 1288 |
| 70 | 349924 | 3006 | 102 | 1230 |
| 90 | 352908 | 22 | 324 | 1008 |

Where:

- TP = True Positive
- FP = False Positive
- FN = False Negative
- TN = True Negative

Table 10.2 Comparison of the mean testing algorithm and the standard deviation testing algorithm for false negative rate when using a running band statistical filter.

These results show that not only does the more complex algorithm have a better reliability of correct classification of the running band but also that the data forming the running band is consistent. This suggests that while squat data cannot easily be statistically classified, the running band data can. By applying this form of methodology, any data that does not conform to running band data can therefore be considered abnormal and have the potential for belonging to a defect.

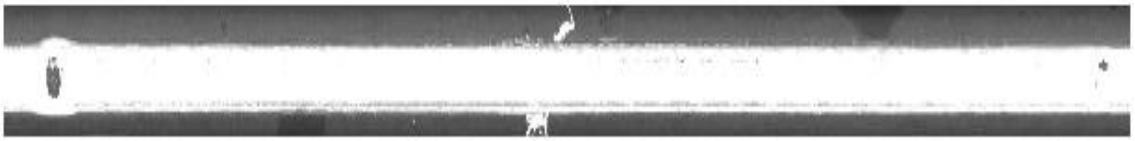
### 10.2.3 Novel Histogram Analysis, Dual Data Suppression and Binary Mapping (D2SBM) for Detection of Surface Bound Abnormalities within the Running Band

Histogram analysis and data suppression have proven successful in the segmentation and analysis of running band data. The same techniques may also be employed for squat and wheel burn detection. The most common details observed from these defects is the discoloration of the area. It is often similar to that associated to the colouration of the non-running band area. Reversing the methods of data suppression and utilising binary mapping, that records the classification of regions, allows for errors within the running band to be segmented.

If the process of segmenting the running band data is inverted, all data associated with the running band would be purged and non-running band data would now become dominant. This would include any squat or wheel data. By utilising data achieved by tracking the running band edge, all data points that fall within the non-running band can be assigned and suppressed within this image by utilising a binary mapping technique.

The resultant data consists of all points that are not associated with running band data or that of the non-running band data. To improve the results further analysis is required; this includes the suppression of noise, that is data groups that are smaller than a set constant in size and do not correspond to a wheel burn or squat.

Figure 10.9 demonstrates the ability of combining dual data suppression techniques by utilising binary mapping in the segmentation of squat data. While the output shown is a binary map of the squat data this map can be easily use to segment the squat data from the image.



Original Image

Binary Map of Non-Running Band Data

Detected Squats Present Within Running Band Data

Figure 10.9 Results of bespoke data suppression of squat and wheel burn detection.

### 10.2.4 Region Growing Techniques in the Detection of Surface Bound Abnormalities

As the data within the image is formed into specific regions, namely non-running band data, running band data and unknown defect data, then region growing techniques can be applied. The advantages of a region-growing algorithm are that the area of region can be calculated easily, with the boundary (edge) positions recorded, and statistical data such as mean and variance about the region as a whole can be formed.

Such data can allow classification of what the region contains and, given multiple regions, if any of these regions are similar. For example if just the head of the rail were to be examined, as in Figure 10.9, a region-growing algorithm, in a best-case scenario, would detect the non-running band data as two independent regions. Rather than rescan every pixel and calculate statistical values based on this data a well design region-growing algorithm could gather this data during its region classification. Data such as the overall mean or variance could then be used to associate the regions as being similar in characteristics.

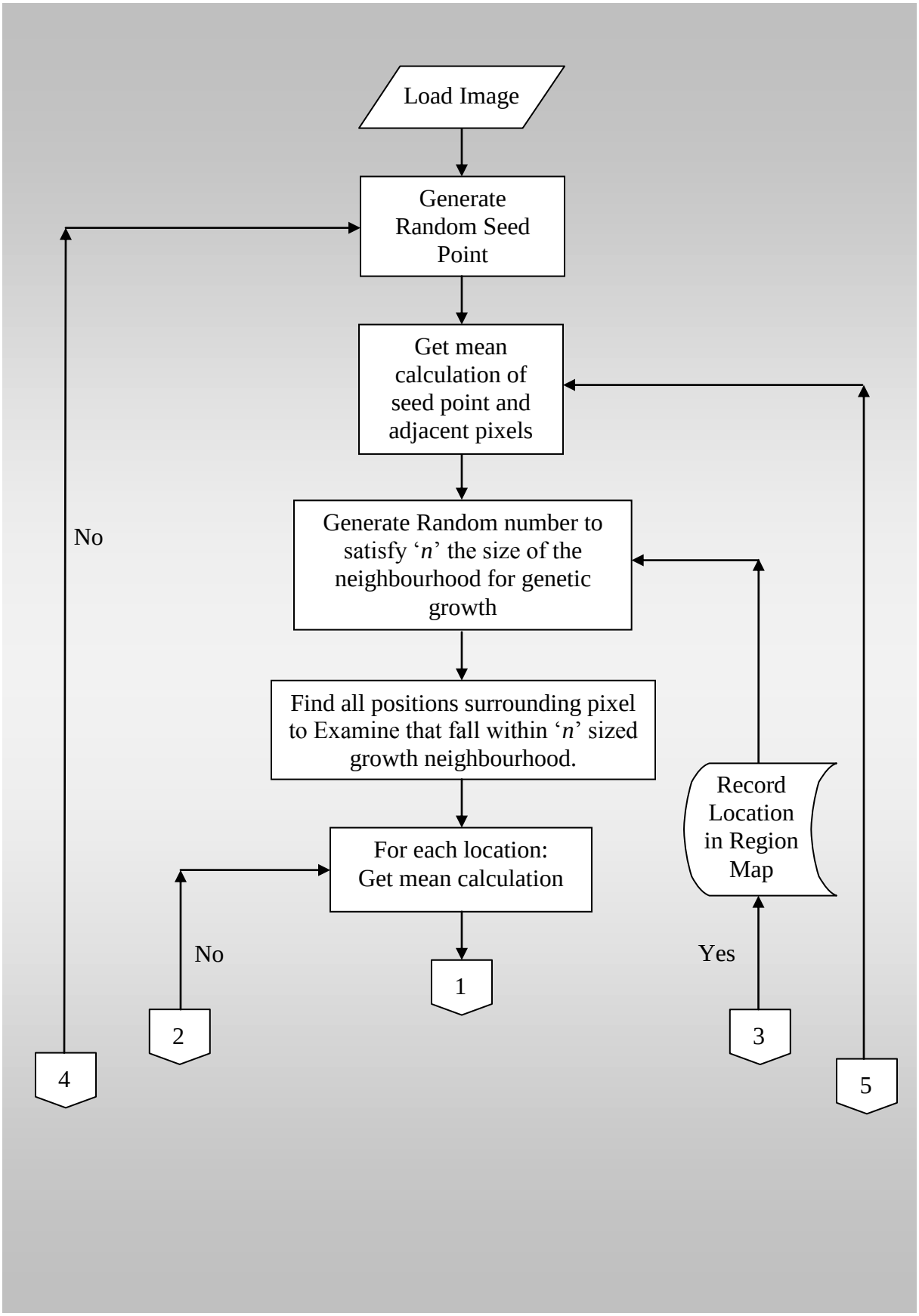
#### 10.2.4.1 Genetic Region Growing Algorithm with Random Growth Pattern

To achieve region-growing analysis a novel program was written. This program allows different analysis methods to be used in classifying region pixels. This program also incorporates the ability to look beyond a pixels immediate boundary. The region growing method applied is based upon a genetic variation in which the pixels examined are not limited to the adjacent pixels. However, their ability to look beyond this boundary is controlled randomly, and the distance that may be examined is randomly generated. The algorithms process is described and an associated work flow diagram is available in Figure 10.10. An algebraic representation is presented in Equation 10.1.

$$\sum_{x=s(lb,rb),y=(H,W)}^{m,n} f\left(I_{(x,y)}\right) \geq T \Rightarrow \sum_{x'=x+s(0,u),y'=y+s(0,v)}^{m,n} f\left(I_{(x\prime,y\prime)}\right) \approx f\left(I_{(x,y)}\right) \Rightarrow R$$

Where:
$lb$ = Left Running Band Edge Cue
$rb$ = Right Running Band Edge Cue
$f$ = Function to be applied in comparing regions
$I$ = Pixel being examined
$s$ = Random Number
$T$ = Threshold Determining Non-Running band Data
$R$ = Classified as belong to Region
$H$ = Image height
$W$ = Image Width

Equation 10.1 Algebraic representation of the genetic region growing algorithm.


Load Image
Generate Random Seed Point
Get mean calculation of seed point and adjacent pixels
Generate Random number to satisfy '*n*' the size of the neighbourhood for genetic growth
Find all positions surrounding pixel to Examine that fall within '*n*' sized growth neighbourhood.
For each location: Get mean calculation
Record Location in Region Map
No
No
Yes
1
2
3
4
5

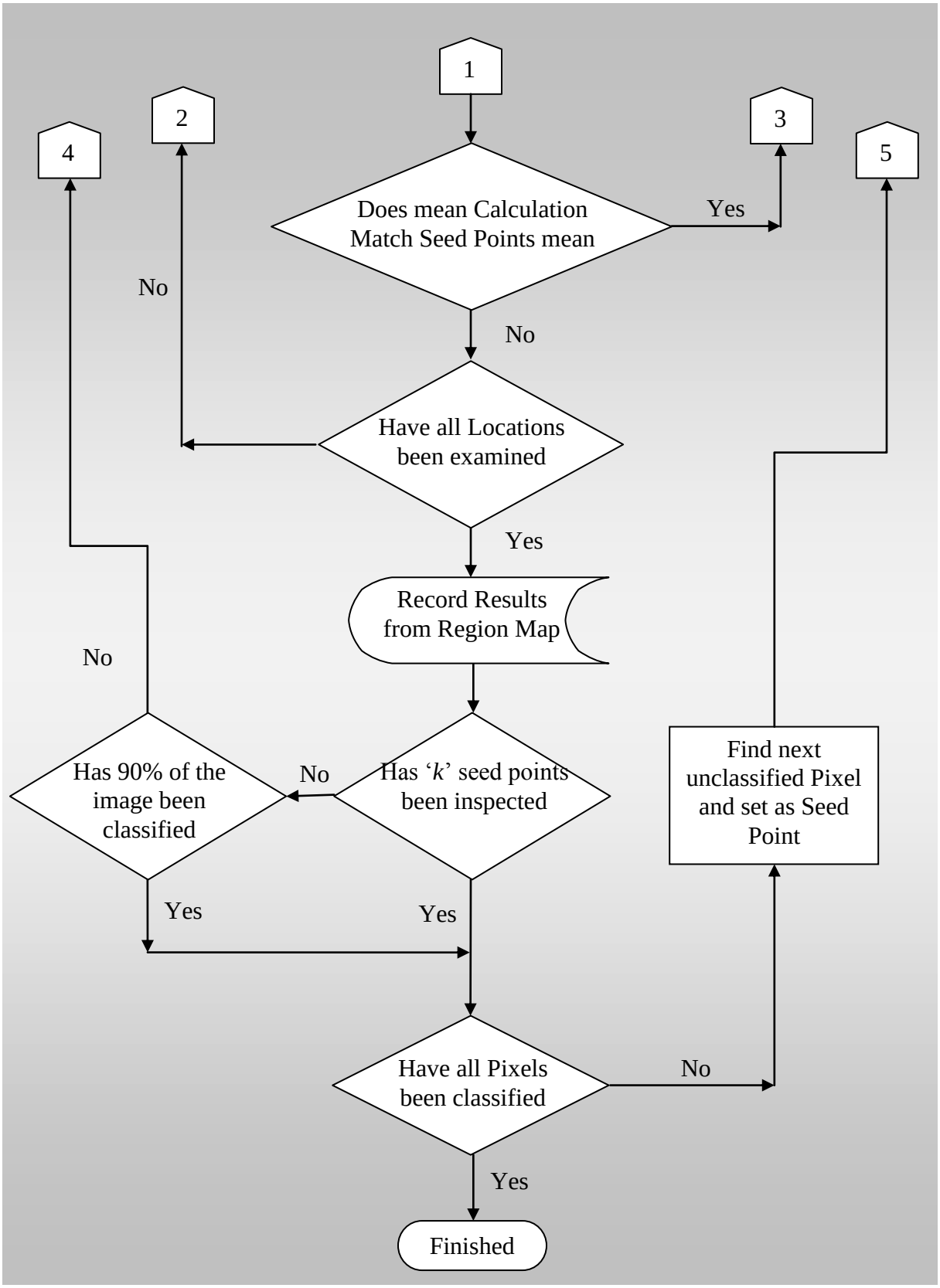



Figure 10.10 Genetic Region Growing Flow Diagram.

**STAGE 1:** A random seed point is chosen within the image and its characteristics are examined. Different tests can be performed, the simplest of which is an average value. In this particular application, an average value of the seed point and its nearest neighbours is calculated. Every pixel examined will have an average value calculated using the same process. If the examined pixel has a value that is within ±10%, for example, of the seed's average value it will be classified as part of the region.

**STAGE 2:** The boundary around the seed point is examined to see if those pixels have the same characteristics as the seed point. The size of the boundary is randomly generated. If any of these pixels are associated with the seed point then the pixel positions are logged and added to the region. Figure 10.11 shows what boundary region is examined depending on the boundary limitation defined by '*n*'. '*n'* is the constant at which the random growth pattern can operate within

For practical reasons the random number '*n'* is limited. The maximum value will be the largest of either the image height or width minus one. Extra checks are performed to ensure that any value calculated falls within the image boundaries. If the maximum value for *'n'* is used then for each seed point generated, all pixel within the image will be tested to see if it's characteristic are similar to the seed point characteristics. Checking of every pixel inhibits the real purpose of applying the region-growing algorithm, i.e. to minimise the number of pixels inspected to a specific region of interest. In such cases, a statistical filter will have identical results and be far more efficient.

| | | | | |
|---|---|---|---|---|
| (x-*n*, y-*n*) | (x-*n*, y-$n_{-1}$) | (x-*n*, y) | (x-*n*, y+$n_{-1}$) | (x-*n*, y+*n*) |
| (x-$n_{-1}$, y-*n*) | (x-1, y-1) | (x-1, y) | (x-1, y+1) | (x-$n_{-1}$, y+*n*) |
| (x, y-*n*) | (x, y-1) | (x, y) | (x, y+1) | (x, y+*n*) |
| (x+$n_{-1}$, y-*n*) | (x+1, y-1) | (x+1, y) | (x+1, y+1) | (x+$n_{-1}$, y+*n*) |
| (x+*n*, y-*n*) | (x+*n*, y-$n_{-1}$) | (x+*n*, y) | (x+*n*, y+$n_{-1}$) | (x+*n*, y+*n*) |

Where:

- *n* = Random Number
- = Seed Point
- = Seed Points Nearest Neighbours

Figure 10.11 Seed point and neighbour pixel examinations.

**STAGE 3:** For all those pixels that are associated with the seed point, the same analysis is repeated. In this case, however, a new random number is generated for '*n*'. This randomly adjusts the jump ability of each pixel examined to look for association with the regions pixels.

'STAGE 3' repeats until all avenues of the region are exhausted. The algorithm uses a map of pre-classified pixels to ensure that pixels that have been examined are not checked again. This is necessary for increasing the inspection speed of the algorithm. The results are a '**random region growth pattern**'.

**STAGE 4:** On completion of a region, a new seed point is randomly chosen from within the image bounds and the algorithm continues from 'STAGE 2' providing the point selected has not been classified as belonging to a region. If the seed point has been classified previously, a new seed point is selected.

To prevent the application of a continuous loop, where the chance of obtaining an unclassified seed point becomes infinitesimally small, the loop will complete when 90% of the image is classified within regions. The remaining unclassified pixels will then be examined sequentially until all pixels are classified within regions.

To prevent memory exhaustion, additional parameters such as the minimum size at which a region should be recorded are also applied. To increase speed on small and large images a parameter was introduced to overwrite the random seed point loop when a set number of random seed points have been successfully created. This assumes that after the set number of random seed points a majority of the image has already been classified and chances of obtaining an unclassified seed point is approaching zero.

#### 10.2.4.2 Testing the Region Growing Algorithm

To evaluate the genetic region-growing algorithm, the results were compared to manually assessed and created binary maps, two for each of the squats and one for the running band. There are unique differences in examining the output results from a region growing method in comparison to the similar results generated by statistical filters. Where previously ROC data was gathered in order to compare the ability of classification, this is no longer useful. ROC analysis presumes that at a set point, an algorithms output, will be entirely accurate and all pixels will be correctly classified. With an algorithm such as the applied genetic region-growing algorithm, this will not happen. The point of which ideal classification happens, falls at a position where the boundary examined is large enough to account for noise however, is sufficiently small enough not to jump boundary lines.

To examine the output accurately statistics based on the correct number of classification of positive and negative pixels will be gathered and the percentage error will be recorded (Table 10.3, Table 10.4). The tests involved increasing the constant at which random growth pattern can fluctuate. The results from this test can be compared to linear seed point analysis, in which each pixel is examined sequentially and if it is not classified within a region, it becomes the seed point

for a new region. The comparison of these results will demonstrate any benefits to the random seed point and growth analysis method.

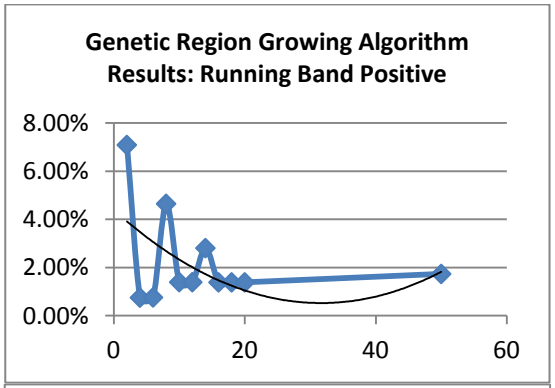
To increase efficiency the maximum number of random seed points was limited to 30 before the remaining pixels were evaluated linearly. As with the statistical filter, the non-running band data was suppressed and not used in analysing the results from the region-growing algorithm. The mean testing error was set at ±30%. Table 10.3 and Table 10.4 show the results from the Genetic Region Growing algorithm. It is important to note that the significant difference in percentage error between running band classification and squat classification is simply down to the size of the region within the image. The squat data makes up a small amount of the image data whereas the running band makes up the majority of the data.

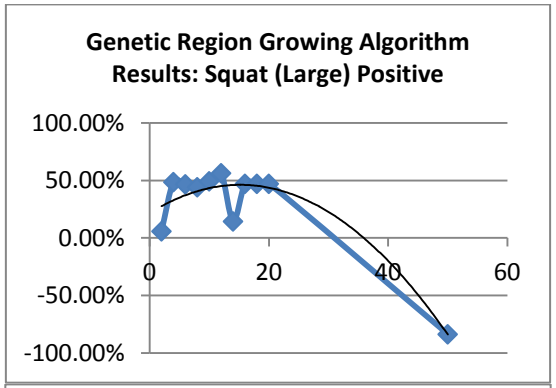
| | | Neighbour Threshold '*n*' (Pixels) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 50 |
| Running Band | Results (P) | | 328493 | 350897 | 350863 | 337108 | 348391 | 348631 | 343629 | 348654 | 348654 | 348654 | 347278 |
| | Compare Map (P) | | 353567 | 353567 | 353567 | 353567 | 353567 | 353567 | 353567 | 353567 | 353567 | 353567 | 353567 |
| | Difference (P) | 0 | 25074 | 2670 | 2704 | 16459 | 5176 | 4936 | 9938 | 4913 | 4913 | 4913 | 6289 |
| | Results (N) | | 26406 | 4002 | 4036 | 17791 | 6476 | 6268 | 11270 | 6245 | 6245 | 6245 | 7481 |
| | Compare Map (N) | | 1332 | 1332 | 1332 | 1332 | 1332 | 1332 | 1332 | 1332 | 1332 | 1332 | 1332 |
| | Difference (N) | 0 | -25074 | -2670 | -2704 | -16459 | -5144 | -4936 | -9938 | -4913 | -4913 | -4913 | -6149 |
| Squat 1 (Large) | Results (P) | | 1080 | 589 | 611 | 640 | 510 | 581 | 981 | 606 | 606 | 606 | 7039 |
| | Compare Map (P) | | 1147 | 1147 | 1147 | 1147 | 1147 | 1147 | 1147 | 1147 | 1147 | 1147 | 1147 |
| | Difference (P) | 0 | 67 | 558 | 536 | 507 | 637 | 566 | 166 | 541 | 541 | 541 | -5892 |
| | Results (N) | | 353819 | 354310 | 354288 | 354259 | 354318 | 354318 | 353918 | 354293 | 354293 | 354293 | 347860 |
| | Compare Map (N) | | 353752 | 353752 | 353752 | 353752 | 353752 | 353752 | 353752 | 353752 | 353752 | 353752 | 353752 |
| | Difference (N) | 0 | -67 | -558 | -536 | -507 | -566 | -566 | -166 | -541 | -541 | -541 | 5892 |
| Squat 2 (small) | Results (P) | | 118 | 142 | 142 | 151 | 142 | 142 | 134 | 142 | 142 | 142 | 7093 |
| | Compare Map (P) | | 170 | 170 | 170 | 170 | 170 | 170 | 170 | 170 | 170 | 170 | 170 |
| | Difference (P) | 0 | 52 | 28 | 28 | 19 | 28 | 28 | 36 | 28 | 28 | 28 | -6923 |
| | Results (N) | | 354781 | 354757 | 354757 | 354748 | 354757 | 354757 | 354730 | 354757 | 354757 | 354757 | 347860 |
| | Compare Map (N) | | 354729 | 354729 | 354729 | 354729 | 354729 | 354729 | 354729 | 354729 | 354729 | 354729 | 354729 |
| | Difference (N) | 0 | -52 | -28 | -28 | -19 | -28 | -28 | -1 | -28 | -28 | -28 | 6869 |

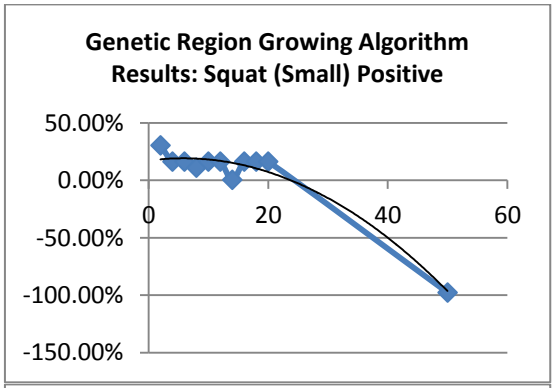
P = Positive, N = Negative

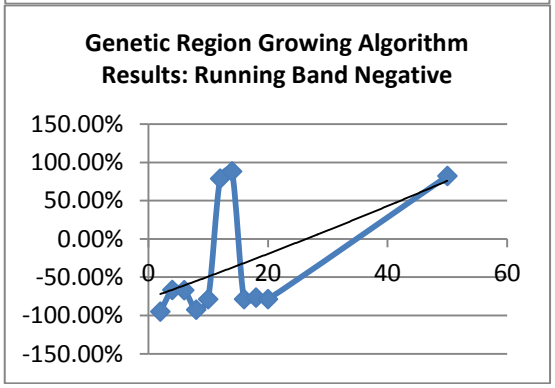
Table 10.3 Results from genetic growth region analysis algorithm.

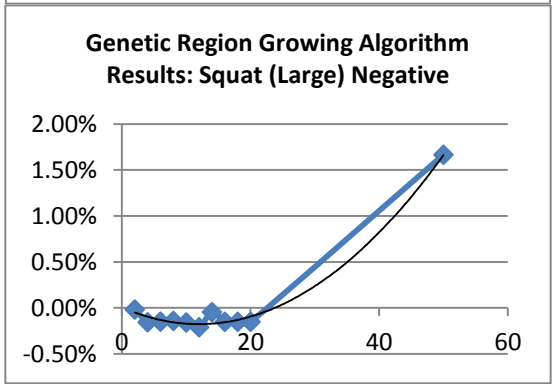
| | | Neighbour Threshold '*n*' | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 50 |
| Running Band | Positive | 7.09% | 0.76% | 0.76% | 4.65% | 1.40% | 1.40% | 2.81% | 1.39% | 1.39% | 1.39% | 1.74% |
| | Negative | -94.96% | -66.72% | -66.99% | -92.51% | -78.78% | 78.75% | 88.18% | -78.67% | -76.67% | -78.67% | 82.19% |
| Squat 1 (Large) | Positive | 5.84% | 48.64% | 46.73% | 44.20% | 49.35% | 56.38% | 14.47% | 47.17% | 47.17% | 47.17% | -83.71% |
| | Negative | -0.02% | -0.16% | -0.15% | -0.14% | -0.16% | -0.21% | -0.05% | -0.15% | -0.15% | -0.15% | 1.67% |
| Squat 2 (small) | Positive | 30.59% | 16.47% | 16.47% | 11.18% | 16.47% | 16.47% | 0.59% | 16.47% | 16.47% | 16.47% | -97.58% |
| | Negative | -0.01% | -0.01% | -0.01% | -0.01% | -0.01% | -0.01% | 0.00% | -0.01% | -0.01% | -0.01% | 1.94% |



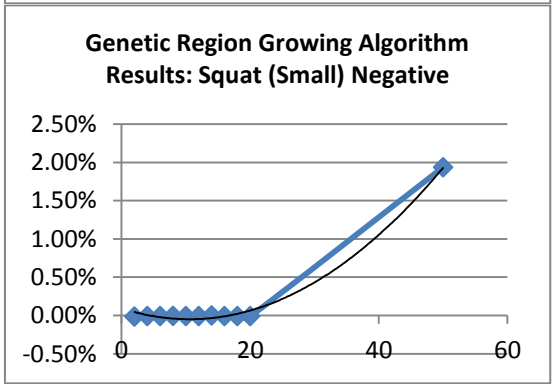




Table 10.4 Tabulated results from genetic growth region analysis algorithm.

Table 10.5 shows the results from a linear region-growing algorithm where only the immediate boundaries of the pixels are examined. These results show that at small values of '*n*' the genetic region growth algorithm holds no significant advantage however at between values of 20 to 50 boundary pixels the ability to offset true positive classification against false positive classification comes apparent. By sacrificing a 0.5% error in true negative classification of the running band the true positive rate of the squat detection approaches ideal values with error rate approaching 0%.

| | | Linear Analysis | Neighbour Threshold (2) | Neighbour Threshold (20) |
|---|---|---|---|---|
| Running Band | Positive | 1.02% | 7.09% | 1.39% |
| | Negative | -73.04% | -94.96% | -78.67% |
| Squat 1 (Large) | Positive | 48.30% | 5.84% | 47.17% |
| | Negative | -0.16% | -0.02% | -0.15% |
| Squat 2 (small) | Positive | 16.47% | 30.59% | 16.47% |
| | Negative | -0.01% | -0.01% | -0.01% |

Table 10.5 Tabulated results from linear region analysis.

The method of the binary map comparison plays an important role in determining an algorithms ability to correctly classify regions. Using a more advanced method of pixel comparison could improve the accuracy of the region growing classification ability, this was demonstrated in chapter 9.2.2.2 Running Band Filter (*page 276*). Due to the random effects caused by the genetic region growth algorithm there are unique issues introduced in deploying such an algorithm. Each calculation from the genetic region-growing algorithm is different. The margin of which is dependent on the constant 'n'. The results presented are the average calculation over 5 tests. This accounts for changes in the random seed points and classification errors.

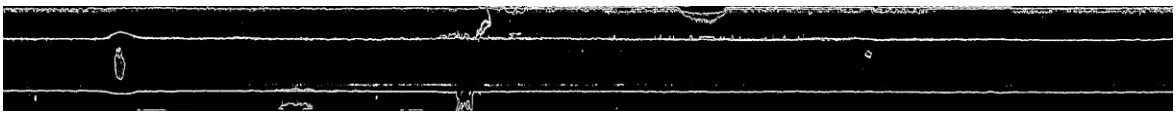
These slight differences mean that to ensure that the region-growing algorithm does not produce errors in defect classification it must be run successive times. This has a significant effect on performance, as the algorithm is processor intensive and cannot be parallelised. However, multiple instances of the same algorithm could be parallelised and each individual output compared. Figure 10.12 shows the relevant outputs achieved from the genetic region-growing

algorithm. Figure 10.12 also demonstrates the ability of extending the region-growing algorithm as an edge detector. The algorithm correctly classifies the running band edge-cues and the squat boundaries.

Squat 1 Region detected

Squat 2 Region detected

Running Band Region detected



Edge detection from algorithm

Figure 10.12 Application of a region growing algorithm in the segmentation of squat defects detected.

## 10.3 METHOD EVALUATION AND OBSERVATIONS

Region Classification algorithms and NN texture recognition algorithms both produce good results. The region classification algorithms produce slightly favourable robustness due to the ease of which they can be adapted to work with changes in image contents. The genetic region-growing algorithm discussed produces good results successfully classifying running band and abnormal squat or wheel burn type data. Performance limitations in executing the algorithm multiple times to average results means such an algorithm will be computationally expensive to implement.

Histogram analysis with dual data suppression and binary mapping (D2SBM) techniques produce comparable results to the other investigated methods. As research into running band classification has found a similar technique to be robust (Chapter 9.4), D2SBM is a preferred implementation. The genetic region growing method is more robust in its classification of data and should be considered for post-processing after the D2SBM has detected possible defects. As a majority of steps would be implemented into running band classification, minimal operations would require executing to evaluate an image that has successfully segmented the surface abnormality data such as a squat.

### 10.3.1 Method Evaluation

With D2SMB being a more efficient technique, the analysis of the squats left within the image still needs to be performed. This allows for classification of the defect features such as size and location upon the railhead. A region-growing algorithm is ideal for this situation it allows the size of the squat to be easily classified and has been proven robust in the classification of squat type regions. This technique has the significant advantage of being able to classify multiple squat and wheel burn defects on one image allowing each defect to be individually classified and subsequently reported.

Large areas of noise would also be suppressed in favour of addressing squat and wheel burn defects detected on a size based severity scale. The application of this method will allow a squat or wheel burn area to be measured to the nearest $mm^2$. A unique measurement from this technique is the ability to scale the railhead abnormality according to the area it consumes. Currently squats are classified by severity according to the height and width, this additional data allows conformation of severity according to this method.

The D2SBM algorithm is applied to data set 4 of 5000 images, see chapter 6.2 Data Sets (page 142) for reference. These images were acquired from the Sheffield Supertram network using the finalised specification for the acquisition system described in Chapter 4. The dataset has been visually inspected and contains various instances of surface abnormalities. These range from severe corrugation, to surface impacts from ballast and sanding. The desired outcome is the detection of these abnormalities. Defect type classification is not applied, as D2SBM would require additional classifiers to determine between squat, wheel burns, or other types of defects.

Only Surface bound abnormalities that occupy more than 1000 pixels are visually assessed. A comparison of the region classification method and current practices in surface bound abnormality measurements are evaluated on 10 randomly selected successfully classified abnormalities from the first 1000 images. The region classification is evaluated by visually assessing the defects characteristics within the image and applying the area of on oval rule.

## 10.4 RESULTS

D2SBM is applied after the rail edge classification algorithm is applied as discussed in Chapter 7 and running band classification is applied as discussed in Chapter 8 .

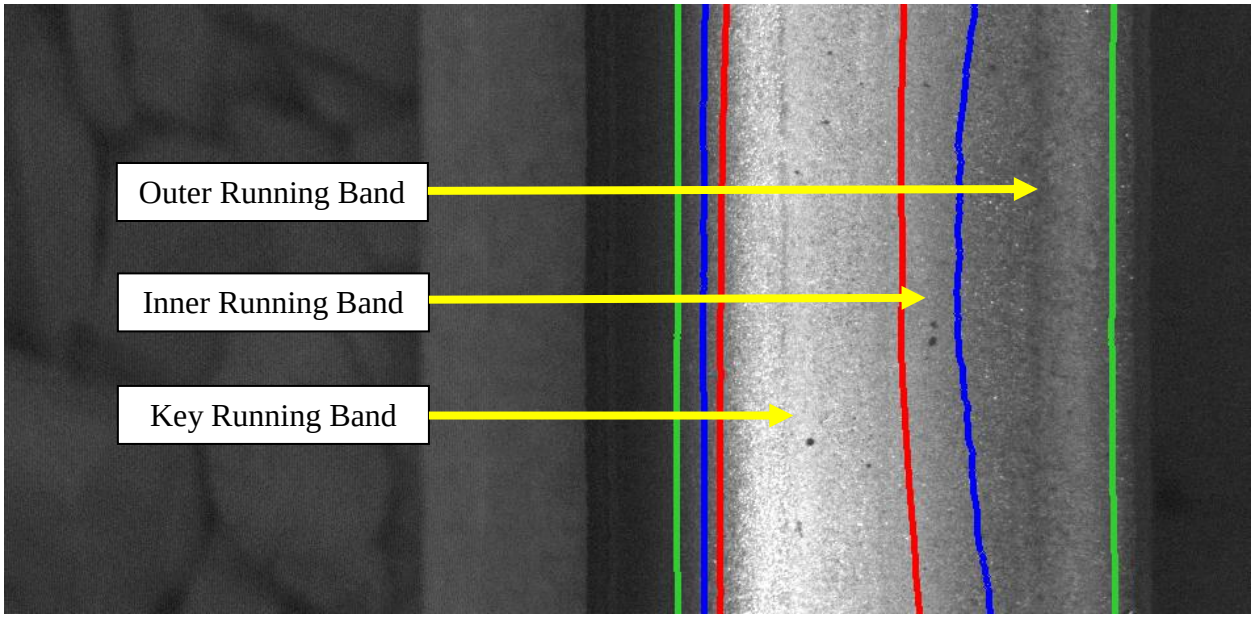



Figure 10.13 Different running band locations on the railhead

| | | Outer Running Band | | | Inner Running Band | | | Key Running band | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sub Set | No. Images | Total Images | T | F | Total Images | T | F | Total Images | T | F | CE(%) | OE(%) |
| 1 | 1000 | 80 | 70 | 10 | 114 | 114 | 0 | 83 | 78 | 5 | 5.42 | 1.5 |
| 2 | 1000 | 18 | 16 | 2 | 975 | 845 | 130 | 753 | 705 | 48 | 10.31 | 18 |
| 3 | 1000 | 109 | 97 | 15 | 183 | 169 | 14 | 195 | 176 | 19 | 9.86 | 4.5 |
| 4 | 1000 | 579 | 515 | 64 | 956 | 838 | 118 | 972 | 904 | 68 | 9.97 | 25 |
| 5 | 1000 | 715 | 675 | 40 | 987 | 987 | 0 | 653 | 633 | 20 | 2.55 | 6 |
| Total | 5000 | 1501 | 1370 | 131 | 3215 | 2953 | 262 | 2656 | 2496 | 160 | 7.5 | 11.06 |

T = True, correctly classified defects
F = False, falsely classified defects
CE = Classification Error
OE = Overall Error

Table 10.6 Results from surface bound abnormality detection showing the true and false classification rate.

Table 10.6 shows the results from the D2SBM algorithm being applied to data set 4 of 5000 images. The table shows the individual results for all 5 sub sets of 1000 images, containing 240 meters of continuous image acquisition. Within the outer running band, 1501 images displayed surface bound abnormalities with 91.27% true positives and visually confirmed as potential defects. The inner running band presented 3215 surface bound abnormalities, with 91.85% true positives. Finally, the key running band presented 2656 surface bound abnormalities with 93.98% true positives.

Table 10.7 shows the difference between the current practice of height and width evaluation to scale surface bound defect severity versus that of the region-growing algorithm. It allows the difference between a bounding box versus measuring the area to be evaluate. To calculate the area from the height and width it is assumed that squat has an oval shape and an area of an oval rule is used to calculate the area.

| Image Number | File | Region Growing Area (Pixels) | Calculated Area (Pixels) | Error of calculated Area (%) |
|---|---|---|---|---|
| 1 | 12203 | 5270 | 8704 | 39.45 |
| 2 | 12519 | 1029 | 720 | 42.91 |
| 3 | 12943 | 2763 | 3801.6 | 27.32 |
| 4 | 12030 | 1794 | 2576 | 30.35 |
| 5 | 12655 | 1436 | 2420 | 40.66 |
| 6 | 12902 | 3366 | 3511.2 | 4.13 |
| 7 | 12104 | 1639 | 3728.8 | 56.04 |
| 8 | 12149 | 1173 | 4932 | 76.21 |
| 9 | 12780 | 1119 | 1971.2 | 43.23 |
| 10 | 12844 | 1139 | 6908 | 83.51 |

Table 10.7 Performance of region growing algorithm versus current practice in squat measurement.

## 10.5 SUMMARY

The results presented within Table 10.6 show that the dual data suppression and binary mapping (D2SBM) algorithm is capable of detecting surface bound abnormalities. The results shown in Figure 10.14 demonstrate how the algorithm is capable of detecting abnormalities caused by corrugation which is often less visible than that of squats or wheel burns. While the algorithm has an average success rate of 88.9% in detecting the abnormalities, there is no method of classifying the defect type. A deviation of up to 23.5% false classifications between 1000 image sub data sets shows that there is still a requirement to improve performance; this could be achieved in improving the algorithms capabilities or in post processing of the results.

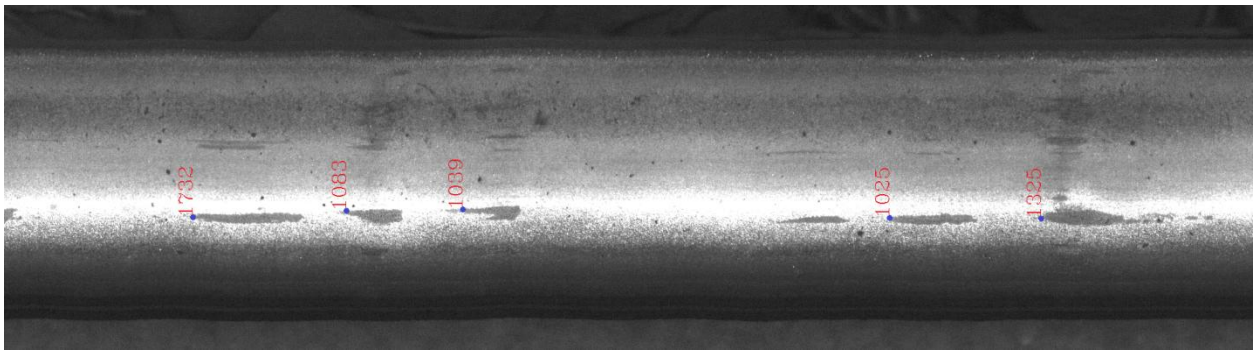



Figure 10.14 Results of surface abnormality detection (D2SBM) algorithm showing measured pixel area of surface bound abnormality.

Further investigation is required into the results obtained to verify if defects consist of different data types, or if determinable characteristics can be obtained. Such characteristics may be shape, location on the railhead and/or pixel intensity difference in comparison to the rail surface. Several false positives occurred due running band filtering, where changes within the running band were sharp. A smoothed occurrence of this feature caused false detection of abnormalities. As a majority of these sharp changes in running band location were caused by corrugation, it is yet to be determined if the surface bound abnormality detection results can be omitted when corrugation is classified.

It can be shown, however, that surface bound abnormality detection can support the detection of corrugation through running band analysis. Using results from both methods can increase

classification performance of this type of defect. Defects such as squats and wheel burns can also be detected, Figure 10.15 show the detection of a smaller squat type features, cause by impact, correctly classified by the D2SBM algorithm. Differentiation between a surface bound defects and clutter, such as paint, rubbish, or impact features upon the railhead still cannot be determined.

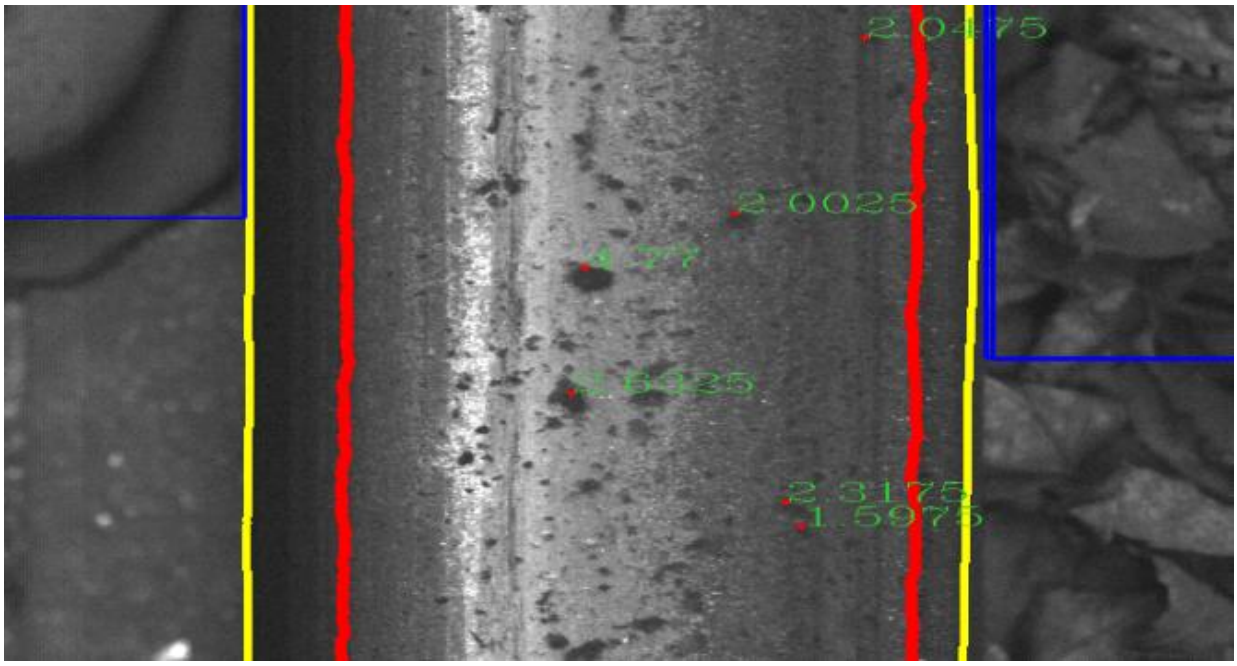



Figure 10.15 Results of surface abnormality detection (D2SBM) algorithm on small squat type defects showing measured pixel area.

Table 10.7 shows that current the manual evaluation method of surface bound defects requires re-evaluation. While expected error rates in the comparison of defect scale classification methods could be put down to region classification inaccuracies, the difference shown is too great to account just for this. The large error rate falls down to the complex shapes and forms of the surface bound defects. A simple mathematical approach will not correctly show the severity of a defect and will leave a possible deficit in the timeliness of the repair actions taken against it.

More significantly, the current method of measuring the height and width of the surface bound abnormality will not show accurately the growth rate over time. The bounding box of the defect can remain the same even though the defect is growing in size.

# CHAPTER 11 CONCLUSION & FUTURE WORK

## 11.1 CONCLUSION

This thesis has discussed the application of a machine vision system in the field of rail and tramway track analysis. A prototype machine vision system (MVS) capable of producing high quality images of tramway track has been developed. In support of this MVS, novel and bespoke algorithms to detect the rail edge, running band, and surface bound abnormalities have been evaluated. The algorithms presented are tested on a dataset 4 of 5000 images acquired by the prototype to evaluate their robustness and performance. While there is still further calibration and design to be applied, the algorithms presented show good results in each of their applications, and exhibit considerable advances against standard machine vision techniques.

A novel algorithm in suppressing vibration caused during image acquisition through means of software correction has been established. The proposed MVS requires some further industrialisation before the system can be applied for rail inspection to minimise this vibration. While this process of industrialisation will introduce hardware to dampen the vibration effect, it is expected that the mechanical assistance will not in itself be sufficient and the novel algorithm presented will aid in suppressing any residual effects.

The algorithms discussed in this thesis demonstrate the advantages of designing image-processing methods around particular characteristics of the image sets. In the application to rail edge and running band classification, this is demonstrated with the poor performance of standard edge detection methods and the superior performance of bespoke algorithms.

Edge detection of railway track shows a high degree of correct classification with results of up to 97.3% accuracy. A drop in performance is seen when severe corrugation is present and further pre-processing may be required to classify such cases. An average classification rate of 91.6% over a 5000 image data set 4 has demonstrated the robust nature of applying a bespoke convolution filter.

A novel method of image data suppression through a quantitative analysis is evaluated in the classification of running band data. Here, an algorithm is presented that is robust to environmental features, such as rain, dirt/grease, and changing light qualities affecting image data. A high level of performance is achieved with an average ability to classify 94.78% of corrugation type defects over 4638 image dataset. The remaining errors are still classified as potential defects however not as the corrugation that caused the abnormality. Further analysis of the width of the running band can still be introduced to further increase the ability of this algorithm if required.

Classification and Grading of Defects has been discussed with potential methods proposed to classify defects over sets of images, as defects occur within individual images, over several images, and the tracking of the defects over time to evaluate their growth. These rules are essential when introducing much larger data sets.

The detection of surface bound defects discusses a novel method of extending the algorithm for running band detection by using a double data suppression method. Surface bound abnormalities can be detected using this method to an accuracy of 0.1mm. Determining the type of defect however, would require the use of further post-processing. It has been shown that surface bound defects such as squats and wheel burns, that have differing characteristics from that of the running band, can be located. A success rate of 89.9% is achieved over a 5000 images of data set 4.

While three types of feature classification algorithms have been developed, the associative defect classification algorithms require refinement. While a defect or presence of a potential defect can be obtained, accurate classification cannot be achieved in the current iteration of the algorithms design. Data set 4 of 5000 images is used to evaluate the performance and robustness of the algorithms discussed. This is shown through the development of the corrugation algorithm that can correctly classify the presence of the defect however, defects such as squats and wheel burns a few and far between and require a large data library to increase classification algorithms.

Finally this thesis has produced a specification for the implementation of a prototype MVS that fits on existing tramway vehicle bogies. The specific purpose of this MVS is for rail inspection in which, 3 algorithms have been discussed that form the basis of automated inspection of the imagery acquired. The aim of the system is to reduce the requirement of manual inspection and while in its current form it will not replace the requirement for manual inspection, the subject matter discussed within this thesis lays the groundwork for its potential reduction.

## 11.2 FUTURE WORK

Refinement of the rail analysis algorithms presented within this thesis is required but what has been presented is a strong groundwork from which to progress. A majority of the future work deals with using the results provided by the rail analysis methods while, improving the robustness of the algorithms by expanding their capabilities to deal with false positive classification of defects.

In detecting the rail edge, future work, will focus on analysing the results for classification of rail joints and determining the characterisations of types of joints i.e. standard and insulated joints. Understanding these will aid in the positive classification of potential fractures in rail. Work with this data will aid in developing the running band analysis as the techniques, such as, changes in gradient with respect to plotted co-ordinates within the image, or variations in distances between mirrored sides of the running band or rail edge can contribute as their application are similar. Rules and thresholds for classifying these changes need to be established and tested to evaluate their capabilities along with evaluating potential methods proposed to classify defects over sets of images, such as defects occurring within individual images, over several images, and the tracking of the defects over time to evaluate their growth.

Classification and Grading of Defects with surface bound abnormalities requires the use of further post-processing. Future work would be to develop the techniques to classify differences between squats and wheel burns evaluating occurrence through, location on the railhead, location on the network and characterisation by the data forming the defect within the image itself.

Future development of new algorithms will deal with the classification and analysis of rail infrastructure. While infrastructure such as switches and crossing will need to be addressed with the refinements of the developed techniques in rail edge and running band classification, detection of such infrastructure as Pandrol clips will need to be instigated. With this, a specification for the data management dealing with the mass of data from image acquisition will need to produce.

# WORKS CITED

A. Berry, B. N. G. A., 2008. *High Speed Video Inspection of Joint Bars Using Advanced Image Collection and Processing Techniques.* Seoul, Korea, s.n.

A.B.Allison, 1968. Sperry Rail Service. *Hull National Railway Historical Soc 33,* p. 6.

A.R. Jimenez, A. J. R. C. a. J. P., 1999. Automatic fruit recognition: a survey and new results using range/attenuation images. *Pattern Recognition,* Volume 32, pp. 1719-1736.

B. Nejikovsky, G. C. C. D. W. J. X. G.-S., 2005. *Automated Joint Bar Inpsection using High Speed Cameras,* Washington: American Railway Engineering and Maintenance-of-Way Association.

B. S. Manjunath, W. Y. M., 1996. Texture Features for Browsing and Retrival of Image Data. *IEEE Transactions on Pattern Anaylsis and Machine Vision,* August, 18(8), pp. 837-842.

Berend Jan van der Zwaag, K. S., 2002. Analysis of Neural Networks for Edge Detection. *Proceedings of the ProRISC workshop on Circuits, Systems and Signal Processing,* pp. 580-586.

Bernard Prasil, B. T. D. S. A. L. H. D. P. S., 1985. *Location of faults in railway rails by means of an eddy current device capable of discriminating between the faults and certain constructional discontinuities of the rail.* Europe, Patent No. EP0159933 A2, DE159933T1, EP0159933A3.

Brian V. Funt, G. D. F., 1995. Color Constant Color Indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* May, 17(5), pp. 522-529.

Bunge, H. J., 1982. *Texture Analysis in Materials Science.* 2nd Edition ed. s.l.:Butterworth-Heinemann Ltd.

C. Chouinard, R. P., 2007. Thinning and Segmenting handwritten characters by line following. *Machine Vision ans Applications,* May, 5(3), pp. 185-197.

C. Demant, C. G. B. S.-A., 2013. *Industrial Image Processing: Visual Quality Control in Manufacturing.* 2nd revised ed. s.l.:Springer.

C. Madoriota, E. S. M. N. N. A. A. D., 2001. *Rail corrugation detection by Gabor Filtering.* s.l., s.n., pp. 626-628.

C. Mandriota, M. N. N. A. E. S. a. A. D., 2004. Filter-based feature selection for rail defect detection. *Machine Vision and Applications,* July, 15(4), pp. 179-185.

C.H. Chen, L. P. P. W. ed., 1999. *Handbook of Pattern Recognition and Computer Vision.* 2nd Edition ed. s.l.:World Scientific Publishing Co.

C.L. Lee, P. W., 1999. A Simple and Robust Thinning Algorithm. *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI),* pp. 357-366.

Canny, J., 1986. A Computational Approach to Edge Detection. *Pattern Analysis and Machine Intelligence,* November, Volume 8, pp. 679-698.

Clark F Olson, D. P. H., 1997. Automatic target recognition by matching oriented edge pixels. *IEEE Transactionson Image Procesing,* 6(1).

Corus Rail Technologies, 2000. *Rolling Contact Fatigue Awareness Training for Visual Inspectors.* s.l.:s.n.

Corus Rail Technologies, 2009. [Online]
Available at: http://www.corusgroup.com

Corus Rail Technologies, April 2006. *DTI Image Analysis Project, Defect catalogue and Running Surface Features,* s.l.: s.n.

D. F. Cannon, K.-O. E. S. L. G. K. S., 2003. Rail defects: an overview. *Fatigue & Fracture of Engineering Materials & Structures,* June, 26(10), pp. 865-886.
D. Gottlieb, C. S. A. S. H. V., 1992. On the Gibbs phenomenon I: recovering exponential accuracy from the Fourier partial sum of a nonperiodic analytic functionstar,. *Journal of Computational and Applied Mathematics,* pp. 81-98.
D. Marr, E. H., 1980. Theory of edge detection. *Proceedings of the Royal Society London. Series B, Biological Sciences,* 29 Febuary, 207(1167), pp. 187-217.
D.F.Cannon, H., 1996. Rail rolling contact fatigue Research by the European Rail Research Institute. *4th International Conference on Contact Mechanics and Wear of Rail-Wheel Systems,* January, 191(1-2), pp. 1 -13.
D.R. Martin, C. F. J. M., 2004. Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* May, 26(5), pp. 530-549.
Davies, E., Jan 2005. *Machine Vision: Theory, Algorithms, Practicalities.* 3rd Edition ed. s.l.:Morgan Kaufmann.
Dillinger, P., Vogelbruch, J., Leinen, J. & Suslov, S., 2006. FPGA-Based Real-Time Image Segmentation for Medical Systems and Data Processing. *Nuclear Science, IEEE Transactions on* , 53(4), pp. 2097 - 2101.
Du-Ming Tsai, C.-T. L., 2003. Fast Normalized Cross Correlation for Defect Detection. *Pattern Recognition Letters,* November, 24(15), pp. 2625-2631.
E. N. Malamas, E. G. P. M. Z. L. P. J. L., 2003. A survey on industrial vision systems, applications and tools. *Image and Vision Computing,* 21(2), pp. 171-188.
Edwards, S. R. H. B. a. A., 2009. *Advancements in Railroad Track Inspection Using Machine-Vision Technology.* Chicago, IL, s.n.
Espino, J. S. B., 2012. Rail extraction technique using gradient information and a priori shape model. *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on,* 7(1), pp. 1132-1136.
F. Llorens, F. J. M. M. P. R. R. C., 2002. *Working with OpenCV and Intel Image Processing Libraries Processing image Data Tools.* [Online]
Available at: http://rua.ua.es/dspace/handle/10045/758
[Accessed 27 June 2011].
F.i Lahajnar, R. B. F. P. S. K., 2002. Machine vision system for inspecting electric plates. *Computers in Industry,* January, 47(1), pp. 113-122.
Fukunaga, K., March 1999. *Handbook of Pattern Recognition and Computer Vision: Statistical Pattern Recognition pp. 33-60.* 2nd Edition ed. s.l.:World Scientific Publishing Co.
Furio, N., 2012. *PM'n'Idea.* [Online]
Available at: http://www.pmnidea.eu/
[Accessed 1 January 2012].
Gail A. Carpenter, W. D. R., 1995. ART-EMAP: A Neural Network Architecture for Object Recognition by Evidence Accumulation. *IEEE Transactions on Neural Networks,* July, 6(4), pp. 805-818.
H. Bay, T. T. L. V. G., 2006. *SURF: Speeded Up Robust Features.* Graz, Austria, Computer Vision – ECCV 2006, pp. 404-417.
H. D. Cheng, X. H. J. Y. J. W., 2001. Color Image Segmentation: Advances and Prospects. *Pattern Recognition,* December, 34(12), pp. 2259-2281.

H. Golnabi, A. A., 2007. Design and application of industrial machine vision systems. *Robotics and Computer-Integrated Manufacturing,* December, 23(6), pp. 630-637.
H. Golnabia, A. A., 2007. Design and application of industrial machine vision systems. *Robotics and Computer-Integrated Manufacturing,* 23(6), pp. 630-637.
H. Ni, S. G., 2004. Image processing algorithm for cheese shred evaluation. *Journal of Food Engineering,* Volume 61, pp. 37-45.
H. Trinh, N. H. S. P., 2012. *Multisensor evidence integration and optimization in rail inspection.* Tsukuba, IEEE, pp. 886-889.
Hellestrand, G., 1980. MODAL: A System for digital hardware description ans simulation. *Journal of Digital Systems,* pp. 241-303.
Henry Schneiderman, T. K., 2000. Probabilistic Modeling of Local Appearance and Spatial Relationships for Object. *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition,* pp. 746-751.
Hough V, P. C., 1962. *Method and means for recognizing complex patterns.* United States, Patent No. 3069654.
Huges, T. P., October 1, 1993. *Elmer Sperry: Inventor and Engineer*. s.l.:The Johns Hopkins University Press.
Hurst, S., 1995. *The Characteristic Function of the Student-t Distribution - Report No. FMRR006-95,* s.l.: Financial Mathematics Research.
I. Sobel, 1978. Neighbourhood coding of binary images fast contour following and general array binary processing. *Computer Graphics and Image Processing,* Volume 8, pp. 127-135.
Ichiro Fujitomi, M. I. H. K. T. Y. K. Y., 2007. *Electromagnetic induction type inspection device and method.* Europe, USA, Patent No. EP1783487 A1, US7696747, US20100052667.
Innotrack, 2008. *D4.4.1 – Rail Inspection Technologies, INNOTRACK Project Number TIP5-CT-2006-031415.* s.l., Innotrack.
J. Canny, 1986. A Computational Approach to Edge Detection. *Pattern Analysis and Machine Intelligence,* November, Volume 8, pp. 679-698.
Jaeggi, J.-P., 1998. *Ultrasonic measuring device of defaults of a railway rail.* Europe, Patent No. EP0676322 B1.
K. Bowyer, C. K. S. D., 1999. *Edge detector evaluation using empirical ROC curves.* Fort Collins, CO , USA , IEEE, pp. 354-359.
Koschan, A., 1995. A Comparative Study On Color Edge Detection. *Proceedings 2nd Asian Conference on Computer Vision, Singapore,* December, Volume 3, pp. 574-578.
Krautkrame, 2012. *Krautkramer.com.au.* [Online]
Available at: http://www.krautkramer.com.au/Eddy%20Current%20Rail%20Inspection.pdf
[Accessed 01 10 2012].
L. F. Molina, E. R. J. R. E. J. H. C. P. L. B. N., 2010. *Condition Monitoring of Railway Turnouts and Other Track Components Using Machine Vision.* Illinois, Department of Civil and Environmental Engineering University of Illinois.
L. G.Roberts, 1963. *Machine Perception Of Three Dimensional Solids, Ph.D Thesis,* s.l.: Massachusetts Institue of Technology, Electrical Engineering.
L. Liu, E. C. D. L. a. T. J., 2010. *A Simple and Robust Thinning Algorithm on Cell Complexes.* Hangzhou, China, Computer Graphics Forum.
Larmore, L., 1965. Introduction to Photographic Principles. pp. 161-166.
Lewis, J. P., 1995. Fast Template Matching. *Vision Interface,* May, Volume 95, pp. 120-123.

Lijun Ding, A. G., 2001. On the Canny edge Detctor. *The Pattern Recognition Society,* Volume 34, pp. 721-725.
Lomax, R. G., 2007. *Statistical Concepts.* 3 PAP/COM ed. s.l.:s.n.
Loncaric, S., 1998. A Survey of Shape Anaylsis Techniques. *Pattern Recognition,* 31(8), pp. 983-1001.
M Ph Papaelias, C. R. a. C. L. D., 2008. *M Ph Papaelias, C Roberts and C L Davis.* s.l., s.n., pp. 367-384.
M. Bocciolone, A. C. A. C. A. C., 2007. A measurement system for quick rail inspection and effective track maintenance strategy. *Mechanical Systems and Signal Processing,* Issue 21, pp. 1242-1254.
M. Farenzena, A. F. a. R. G., 2009. *Structure-and motion pipeline on a hierarchical cluster tree..* Kyoto Japan, IEEE, p. 1489–1496.
M. Moganti, F. E. C. D. a. S. T., 1996. Automatic PCB inspection algorithms: a survey. *Computer Vision and Image Understanding,* Volume 63, pp. 287-313.
M. N. Fesharaki, G. R., 1993. *Designing a video rate edge detection ASIC,* s.l.: [Kensington, N.S.W.] : School of Computer Science and Engineering, University of New South Wales.
M. Papaelias, C. R. C. L. D., 2008. *A review on non-destructive evaluation of rails: state-of-the-art and future development.* s.l., s.n., pp. 367-384.
M. Sonka, V. H. R. B., 2007. *ISE Image Processing Analysis and Machine Vision.* Student ED edition ed. s.l.:Nelson Engineering.
M.Heath, S. S. T. S. K. B., 1997. A Robust Visual Method for Assessing the Relative Performance of Edge-Detection Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* December, 19(12), pp. 1338-1359.
Martens George D, 1982. *An ultrasonic rail testing method and system.* International, Patent No. WO1982003920 A1.
Michael J. Swain, D. H. B., 1991. Color Indexing. *International Journal of Computer Vision,* November, 7(1), pp. 11-32.
Milan Sonka, V. H. R. B., Sep 1998. *Image Processing: Analysis and Machine Vision.* 2nd Edition ed. s.l.:CL-Engineering.
Monadjemi, A., 2005. *A New Multi-Scale/Multi-Directional Method for Detection of Abnormalities in Random Textures.* Benidorm, Spain , IASTED, pp. 480-085.
N. Bowring, E. G. P. Y. F. D. G., 2004. A New Statistical Method for Edge Detection on Textured and Cluttered Images. *Visualization, Imaging, And Image Processing: Forth IASTED International Conference Proceedings,* September.Issue 453.
N. R. Pal, S. K. P., 1993. A Review on Image Segmentation Techniques. *Pattern Recognition Society,* September, 26(9), pp. 1277-1294.
Nassu, B. U. M., 2012. A Vision-Based Approach for Rail Extraction and its Application in a Camera Pan–Tilt Control System. *Intelligent Transportation Systems, IEEE Transactions on ,* 13(4), pp. 1763-1771.
Network Rail, 2010. *Track asset policy part 4 update March 2010. - Network Rail,* s.l.: Network Rail.
NVIDIA Corporation, 2008. *Tesla GPU's Enable Cluster Class Performanceon the Desktop at 1/10th the Power.* [Online]
Available at: http://www.nividia.com/object/io_1227008280995.html [Accessed 29 November 2008]

Office of Rail Regulation, July 2006. *Train Derailment at Hatfield: A Final Report by the Independent Investigation Board,* s.l.: s.n.
P.L. Mazzeo, M. N. E. S. A. D., 2004. Visual recognition of fastening bolts for railroad maintenance. *Pattern Recognition Letters,* pp. 669-667.
Pagano, D. A., 2013. *Method and apparatus for detecting internal rail defects.* USA, Patent No. US20130111997 A1.
Pao, Y.-H., 1999. *Handbook of Pattern Recognition and Computer Vision: Neural Net Computing for Pattern Recognition pp. 125-162.* 2nd Edition ed. s.l.:World Scientific Publishing Co.
Q. Ji, R. M. H., 1999. *Quantitative Evaluation of Edge Detectors Using Minimum Kernel Variance Criterion.* Kobe , Japan , IEEE, pp. 705-709.
Q. Li, S. R., 2012. A Real-Time Visual Inspection System for Discrete Surface Defects of Rail Heads. *Instrumentation and Measurement, IEEE Transactions on ,* 61(8), pp. 2189-2199.
R. Archibald, K. C. A. G. R. R., 2003. Improving tissue segmentation of human brain MRI through preprocessing by the Gegenbauer reconstruction method. *NeuroImage,* pp. 489-502.
R. Brunelli, T. P., 1993. Face Recognition: Fetures versus Templates. *IEEE Trans. Pattern Analysis and Machine Intelligence,* 15(10), pp. 1042-1052.
R. C. Gonzalez, R. E. W. S. L. E., 2004. *Digital Image Processing Using Matlab.* 2nd Edition ed. s.l.:Prentice Hall.
R. Clark, 2004. Rail Flaw Detection: Overview and needs for future developments. *NDT & E International,* March, 37(2), pp. 111-118.
R. Ruan, S. N. A. S. N. R. J. P. C., 1998. Estimation of Fusarium Scab in Wheat Using Machine Vision and a Neural Network. *Cereal Chemistry,* August , 75(4), pp. 455-459.
R.C.Thomas, R. B., Feb 1988. *Computer Vision: A First Course.* s.l.:Alfred Waller Ltd.
Rail Accident Investigation Branch, July 2006. *Trackworker Fatality at Trafford Park 26th October 2005,* s.l.: s.n.
Reiff, K. S. a. R., 2000. An assessment of Railtrack's methods for managing broken and defective rails. *Rail Failure Assessment for the Office of the Rail Regulator,* October 35.
Ripley, B. D., 2008. *Pattern Recognition and Neural Networks.* 1st Edition ed. s.l.:Cambridge University Press.
S. Bogdanski, M. O. J. S., 1996. Numerical Stress analysis of rail rolling contact fatigue cracks. *4th International Conference on Contact Mechanics and Wear of Rail-Wheel Systems,* January, 191(1-2), pp. 14-24.
S. Bowling - MICROCHIP, 2007. *AN1138 A Digital Constant Current Power LED Driver.* [Online]
Available at:
http://www.microchip.com/stellent/idcplg?IdcService=SS_GET_PAGE&nodeId=1824&appnote=en532853
[Accessed 06 07 2011].
S. Konishi, A. Y. J. C. S. C. Z., 1999. Fundamental Bounds on Edge Detection: An Information Theoretic Evaluation of Different Edge Cues. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Volume 1, pp. 1573-1999.
S. Sawadisavi, J. E. E. R. ,. H. C. B. N. A., 2008. *Machine-Vision Inspection of Railroad Track.* Landover, Maryland, s.n.
S. Yella, D. N. K. G. M. D., 2007. *Machine vision for automating visual condition monitoring of railway sleepers.* CA, USA , ACTA Press Anaheim, pp. 289-295.

S.P. Mohammad, 2008. *Machine Vision for Automating Visual Inspection of Wooden Railway Sleepers, Master Thesis, Computer Engineering, Reg.No:E3551D,* Borlänge/Falun: s.n.
Scott Konishia, E.-m. T. C. A. A. Y. A. K. A. Y. J. C., 2003. A Statistical Approach to Multi-Scale Edge Detection. *Image and Vision Computing,* January, 21(1), pp. 37-48.
Stagecoach , 2009. [Online]
Available at: http://www.supertram.com/
Stemmer Imaging, 2011. *http://www.stemmer-imaging.co.uk/.* [Online]
Available at: http://www.stemmer-imaging.co.uk/en/articles/4395-Smart-Approach-to-3D-Imaging
[Accessed 25 11 2013].
Swets, J. A., 1996. *Signal Detection Theory and ROC analysis in psychology and diagnostics: Collected papers.*. New Jersey: Lawrence Erlbaum Associates.
T.F. Cootes, C. T. D. C. J. G., 1995. Active Shape Models-Their Training and Application. *Computer Vision and Image Understanding,* 61(1), pp. 38-59.
Tata Steel (formerly Corus), 2010. *WP1 Image Analysis - Component Catalogue (FP7-234299),* Brussels: PM'n'Idea.
Uzun, I., Sch. of Comput. Sci., Q. U. o. B. U., Amira, A. & Bouridane, A., 2005. *FPGA implementations of fast Fourier transforms for real-time signal and image processing.* s.l., IEEE.
Welch, R. Y. a. G., n.d. Fast Image Segmentation and Smoothing Using Commodity Graphics Hardware.
Y. LeCun, L. B. G. B. O. K. M., 1998. Efficient Backprop. *LNCS: Neural Networks: Tricks of the Trade,* 1524(546), pp. 9-50.
Y. Yitzhaky, E. P., 2003. A Method for Objective Edge Detection Evaluation and Detector Parameter Selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* August, 25(8), pp. 1027-1033.
Yud-Ren Chen, K. C. M. S. K., 2002. Machine vision technology for agricultural applications. *Computers and Electronics in Agriculture,* November, 36(2-3), pp. 173-191.
Zili Li, X. Z. C. E. R. D. M. M., 2008. An investigation into the causes of squats - Correlation analysis and numerical modelling. *Contact Mechanics and Wear of Rail/Wheel Systems,* October, 265(9-10), pp. 1349-1355.

## APPENDIX 1 FOV CALCULATION USING A LINE SCAN CAMERA

The following Appendix contains an example calculation for the Field of View (FOV) when considering a line scan camera. This calculation is essential in determining the objects that can be viewed by a camera given mounting constraints. It also allows real world measurements of specific objects given a known distance from the CCD using image data lone.

Line scan camera details: 4K pixel resolution with 10$um$ pixel size.

Working out the CCD:

i horizontal dimension

$$4096\ Pixels\ @\ 10um = 40960um$$

∴

$$\frac{40960um}{1000} = 40.960mm$$

i vertical dimension (Only 1 line of Pixels)

$$\frac{10um}{1000} = 0.01mm$$

∴

i diagonal

$$\sqrt{0.01^2 + 40.96^2} \approx 40.96$$

Calculating Basic Image Dimension

Therefore the FOV with 50mm Lens:

$$o = \frac{i * d}{f}$$

∴

$$o_{width} = \frac{40.96 * 500}{50} = 409.6\, mm$$

$$o_{height} = \frac{409.6}{4096} = 0.1mm$$

## APPENDIX 2 FOV CALCULATION USING AN AREA SCAN CAMERA

The following Appendix contains an example calculation for the Field of View (FOV) when considering an area scan camera. This calculation is essential in determining the objects that can be viewed by a camera given mounting constraints. It also allows real world measurements of specific objects given a known distance from the CCD using image data lone.

Area Scan Camera Details: 1600 x 1200 pixel resolution with 10 $um$ pixel size

Working out the CCD:

i horizontal dimension

$$1600\ Pixels\ @\ 7.4um = 11840um$$

∴

$$\frac{11840um}{1000} = 11.84mm$$

i vertical dimension

$$1200\ Pixels\ @\ 7.4um = 8880um$$

∴

$$\frac{8880um}{1000} = 8.88mm$$

∴

i diagonal

$$\sqrt{11.84^2 + 8.88^2} = 14.8\ mm$$

Calculating Basic Image Dimension

Therefore, 50mm lens:

$$o = \frac{i * d}{f}$$

∴

$$o_{diagonal} = \frac{14.8 * 500}{50} = 148\ mm$$

FOV α CCD

$$o_{diagonal}\ \alpha\ I_{\text{diagonal}}$$

∴

$$o_C = \frac{14.8}{148} = 0.1$$

*o* horizontal dimension

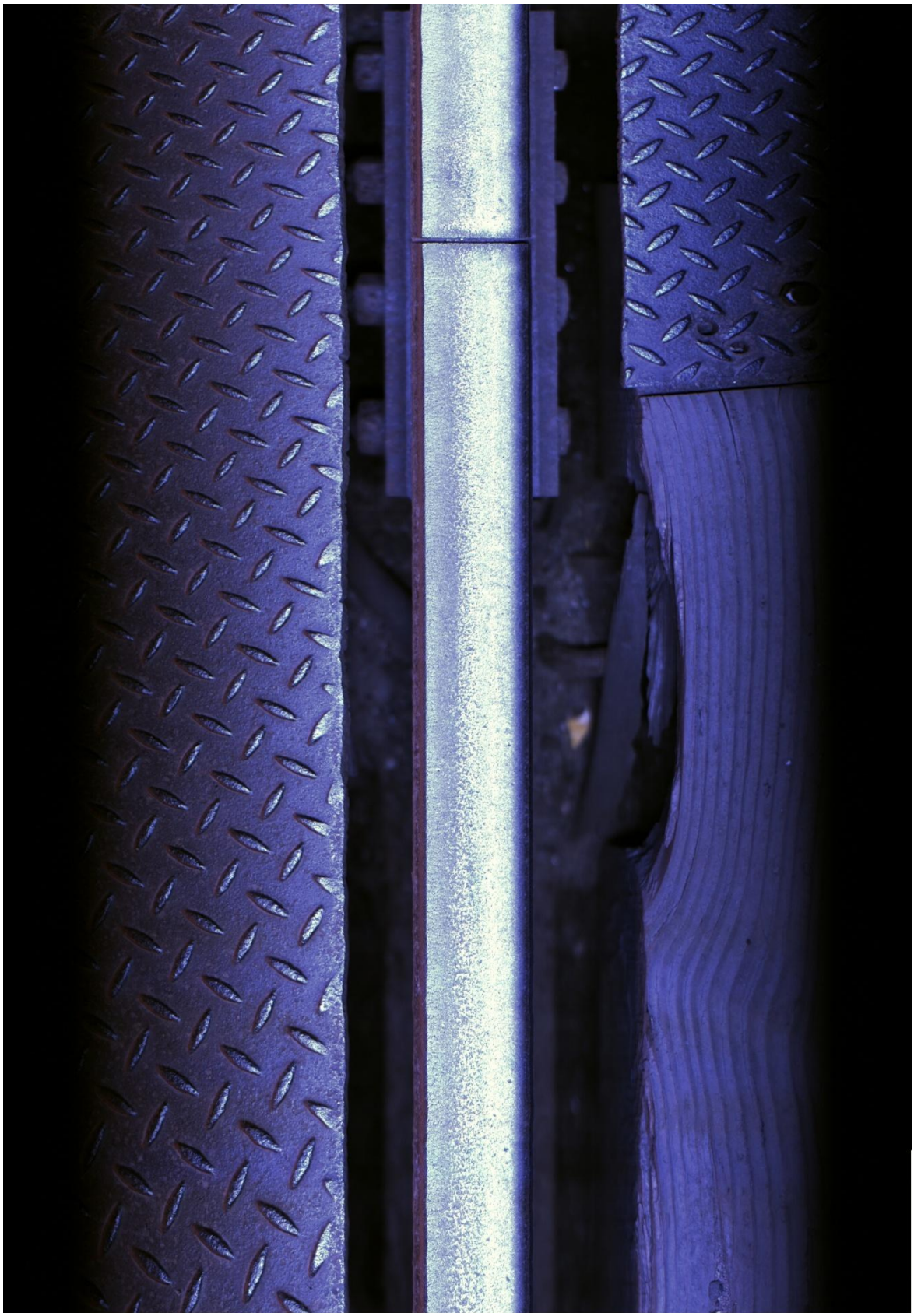
$$11.84/0.1 = 118.40mm$$

*o* vertical dimension

$$8.88/0.1 = 88.8mm$$

## APPENDIX 3 IMAGES ACQUIRED FROM TEST RUN APRIL 2010

The following images were acquired using a Dalsa PC-30-04K80 colour line scan camera. The images were acquired from initial testing of the prototype on the Sheffield Supertram network the run was completed in April 2010.

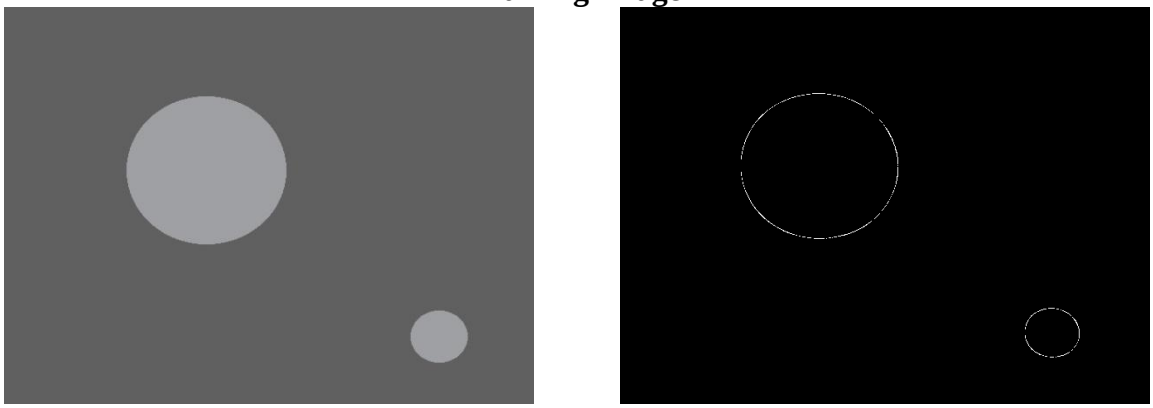


## APPENDIX 4 RESULTS FROM AN NEURAL NETWORK (NN) IN APPLIED EDGE DETECTION

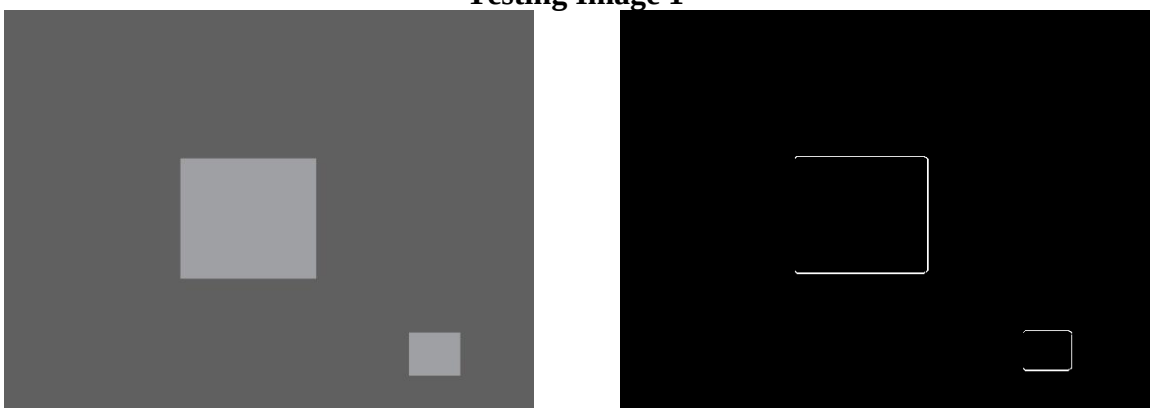
The following Appendix demonstrates the results of a NN capability when applied to edge detection. A mask of data is taken from the image fed into the NN and a chosen result is given to indicate a required edge or a region with no edge. Two types of images are presented. The 'Training Image' is a simple two-tone greyscale images. The 'Training Image' was used to acquired 8 Positives and 8 Negatives 10x10 masks of edges to feed into the NN. The testing images show the results of the NN when applied to an identical two-tone image (Testing Image 1) with different edge features and a computer generated graphical design (Testing Image 2). Testing Image 2 shows the NN capabilities to deal with new data while the two-tone testing image, Testing Image 1 shows how one edge orientation has not been successfully learnt by the NN.

**Original Image** **NN result**

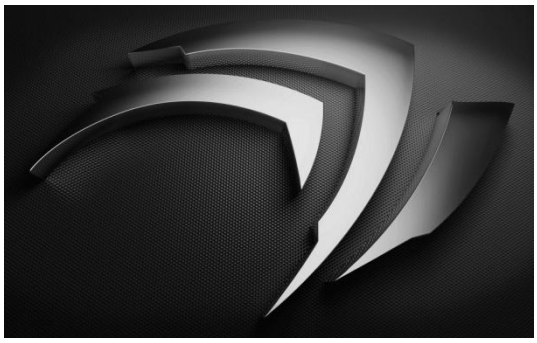
**Training Image**

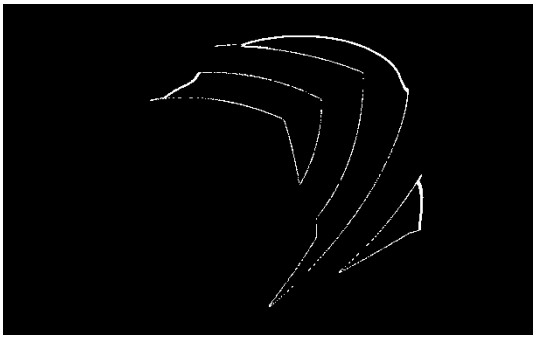


**Testing Image 1**



**Testing Image 2**





Neural Network Settings:

Training Set: 8 Positives 8 Negatives 10*10 Size (100 values)
Neurons: 100
Epochs: 274419

Learning Rate: 0.1
Momentum: 0
Sigmoid Alpha: 2 (Bi-Polar)

# APPENDIX 5 BESPOKE LINEAR CONVOLUTION VS. HOUGH TRANSFORM RESULTS IN RAIL EDGE CUE CLASSIFICATION

This Appendix presents the results from classifying the rail edge, comparing the use of a line following algorithm and the Hough transform. Using FFT template matching, the position of the head of the rail within an image is estimated. The region in which the railhead occurs is then processed for edge cues. A bespoke convolution filter is used to determine likely edge cues that belong to the railhead. The results of the applied edge cue detection are then evaluated to determine the location of the railhead using the line following and Hough methods. The results of these methods are evaluated against a manually determined position of the railhead within the image.

Two test sets are used for the evaluation, both test sets show an improvement in performance when a line following algorithm is applied over that of the Hough transform. A large difference can be seen from the classified Hough transform edges and those from the line following algorithm.

**Results for Test Set 1:**

| Image Num. | Left Edge Avg. | Right Edge Avg. | Left Edge Hough | Right Edge Hough | Average Width (pixels) | Hough Width (pixels) | Actual Width (pixels) | Diff Actual / Hough | Diff Actual / Average |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 372 | 568 | 381 | 567 | 196 | 186 | 195 | 9 | -1 |
| 2 | 382 | 576 | 388 | 567 | 194 | 179 | 195 | 16 | 1 |
| 3 | 392 | 588 | 398 | 581 | 196 | 183 | 195 | 12 | -1 |
| 4 | 386 | 581 | 398 | 574 | 195 | 176 | 195 | 19 | 0 |
| 5 | 370 | 567 | 377 | 562 | 197 | 185 | 195 | 10 | -2 |
| 6 | 371 | 572 | 390 | 562 | 201 | 172 | 195 | 23 | -6 |
| 7 | 379 | 579 | 387 | 571 | 200 | 184 | 195 | 11 | -5 |
| 8 | 384 | 582 | 385 | 576 | 198 | 191 | 195 | 4 | -3 |
| 9 | 379 | 577 | 383 | 567 | 198 | 184 | 195 | 11 | -3 |
| 10 | 378 | 570 | 389 | 566 | 192 | 177 | 195 | 18 | 3 |
| 11 | 376 | 572 | 384 | 564 | 196 | 180 | 195 | 15 | -1 |
| 12 | 376 | 571 | 381 | 565 | 195 | 184 | 195 | 11 | 0 |
| 13 | 378 | 571 | 383 | 567 | 193 | 184 | 195 | 11 | 2 |
| 14 | 379 | 571 | 383 | 570 | 192 | 187 | 195 | 8 | 3 |
| 15 | 381 | 573 | 392 | 574 | 192 | 182 | 195 | 13 | 3 |
| 16 | 378 | 572 | 383 | 573 | 194 | 190 | 195 | 5 | 1 |
| 17 | 381 | 577 | 398 | 570 | 196 | 172 | 195 | 23 | -1 |
| 18 | 385 | 578 | 382 | 575 | 193 | 193 | 195 | 2 | 2 |
| 19 | 382 | 575 | 382 | 569 | 193 | 187 | 195 | 8 | 2 |
| 20 | 380 | 573 | 391 | 569 | 193 | 178 | 195 | 17 | 2 |
| Average | | | | | 195.2 | 182.7 | 195 | 12.3 | -0.2 |
| Variance | | | | | 2.56 | 5.64 | 0 | 5.64 | 2.56 |

**Results for Test Set 2:**

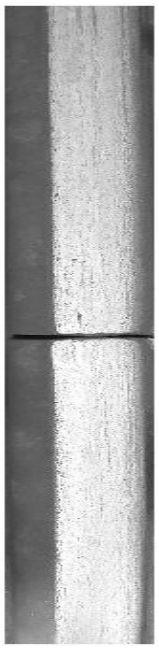
| Image Num. | Left Edge Avg. | Right Edge Avg. | Left Edge Hough | Right Edge Hough | Average Width (pixels) | Hough Width (pixels) | Actual Width (pixels) | Diff Actual / Hough | Diff Actual / Average |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1742 | 2331 | 1742 | 2348 | 589 | 606 | 584 | -22 | -5 |
| 2 | 1683 | 2238 | 1681 | 2245 | 555 | 564 | 554 | -10 | -1 |
| 3 | 1778 | 2360 | 0 | 0 | 582 | 0 | 553 | 553 | -29 |
| 4 | 1838 | 2377 | 1846 | 2451 | 539 | 605 | 544 | -61 | 5 |
| 5 | 887 | 1151 | 0 | 0 | 264 | 0 | 269 | 269 | 5 |
| Average | | | | | 505.8 | 591.6666667 | 500.8 | -31 | -5 |
| Variance | | | | | 122.24 | 373.89 | 116.68 | 274.83 | 12.58 |

= Unsuccessful Classification by Hough Algorithm

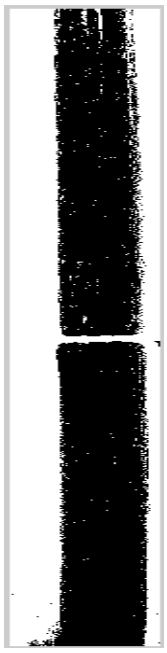
## APPENDIX 6 HISTOGRAM ANALYSES AND THRESHOLDING

This Appendix shows the results from using histogram analysis to calculate a constant for thresholding. An image of the railhead has its histogram statistically evaluate for the most likely value corresponding to that of the railheads reflective surface or more commonly known running band. This method shows inherent advantages to a region-growing algorithm. The relevant simplicity algorithm and thresholding results provide desirable results.

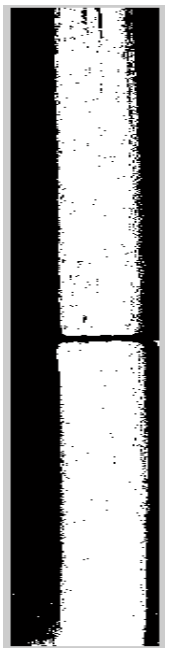
Image 1



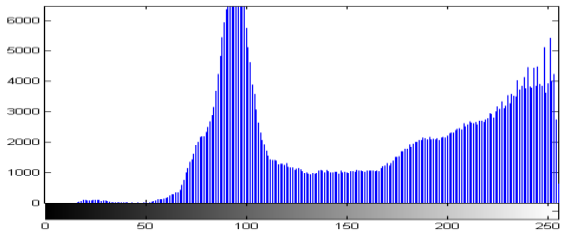
Original



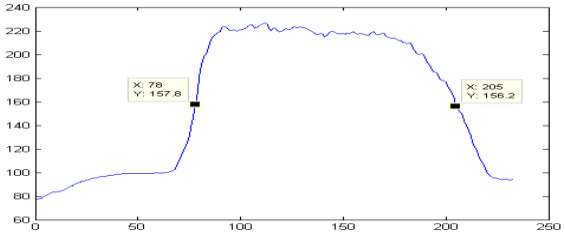
Region 1
Threshold <157



Region 2
Threshold>157

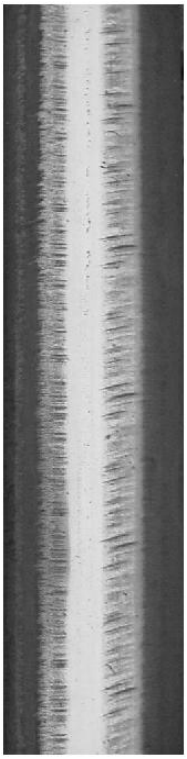



Bar Chart: Histogram of Image



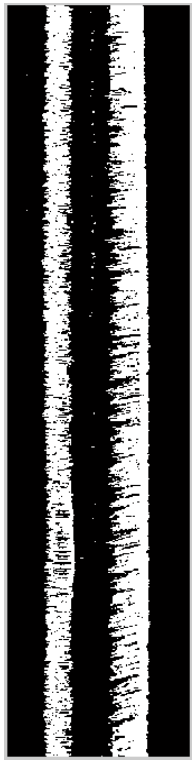

Plot: Mean of Image Columns

Image 2

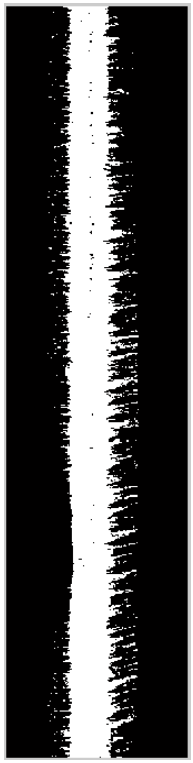


Original

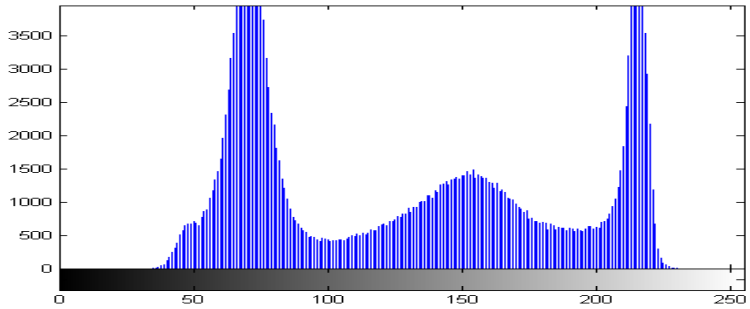
Region 1
Threshold<110



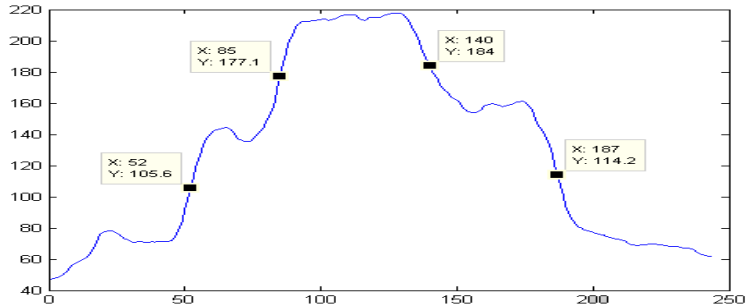
Region 2
110<Threshold<180



Region 3
Threshold>180

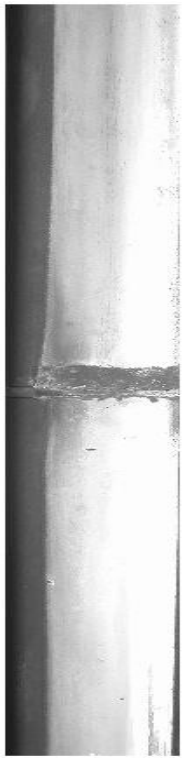



Bar Chart: Histogram of Image

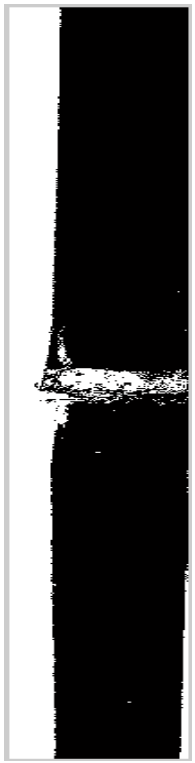



Plot: Mean of Image Columns

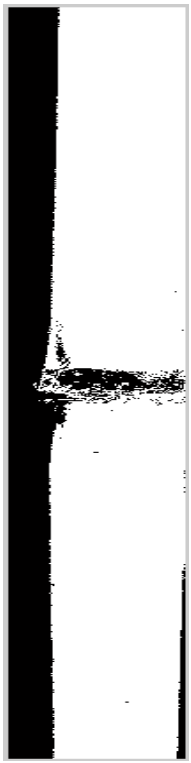
Image 3



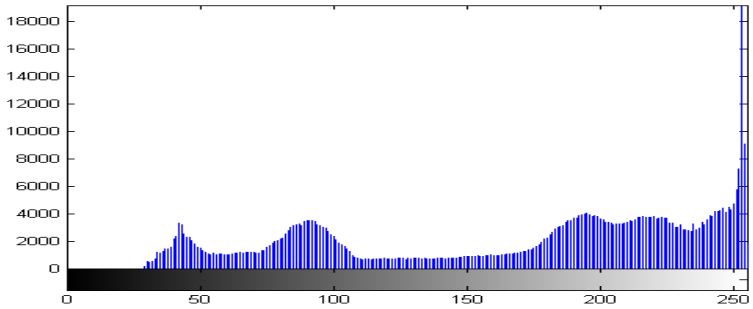
Original



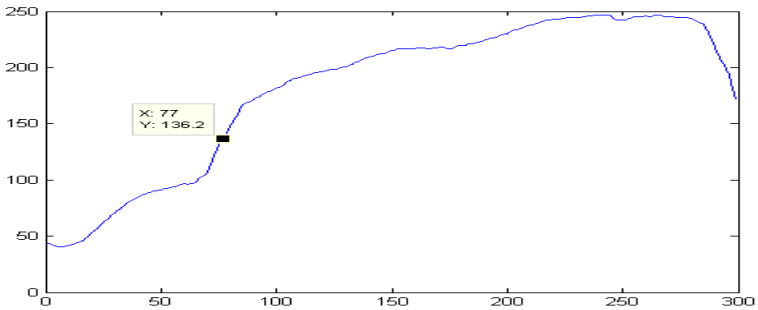
Region 1
Threshold<136



Region 2
Threshold >136




Bar Chart: Histogram of Image




Plot: Mean of Image Columns

## APPENDIX 7 RESULTS TAKEN FROM METHOD 1: TEST SET 1 FROM APPLYING HISTOGRAM ANALYSIS IN THE SUPPRESSION OF NON-RUNNING BAND DATA

This Appendix presents the results from evaluating the classification of the running band data when thresholding is applied. A threshold value is evaluated through histogram analysis. The value determined is designed so that data not belong to the running band is suppressed. Only results for Method 1 are presented as other methods are tabulated within the main section of the thesis.

| Image | TP | FP | TN | FN |
|---|---|---|---|---|
| 1 | 119374 | 0 | 338414 | 349124 |
| 2 | 137142 | 0 | 320232 | 324962 |
| 3 | 132124 | 0 | 325484 | 324728 |
| 4 | 40350 | 0 | 305959 | 415547 |
| 5 | 72997 | 0 | 299470 | 401677 |
| 6 | 135226 | 0 | 318530 | 332676 |
| 7 | 131146 | 0 | 238253 | 388361 |
| 8 | 22646 | 0 | 277371 | 482319 |
| 9 | 87859 | 0 | 293995 | 380002 |
| 10 | 112975 | 0 | 273022 | 404531 |
| 11 | 93295 | 0 | 269475 | 423662 |
| 12 | 126030 | 0 | 279052 | 389542 |
| 13 | 73544 | 0 | 296662 | 424418 |
| 14 | 120478 | 0 | 314707 | 359439 |
| 15 | 80058 | 0 | 325954 | 404996 |
| 16 | 62678 | 0 | 320257 | 423977 |
| 17 | 95832 | 0 | 300196 | 406788 |
| 18 | 65992 | 0 | 301349 | 435475 |
| 19 | 69645 | 0 | 309974 | 431389 |
| 20 | 111127 | 0 | 313530 | 386351 |
| Average | 94525.9 | 0 | 301094.3 | 394498.2 |
| Error | | 0 | | 131.0215 |

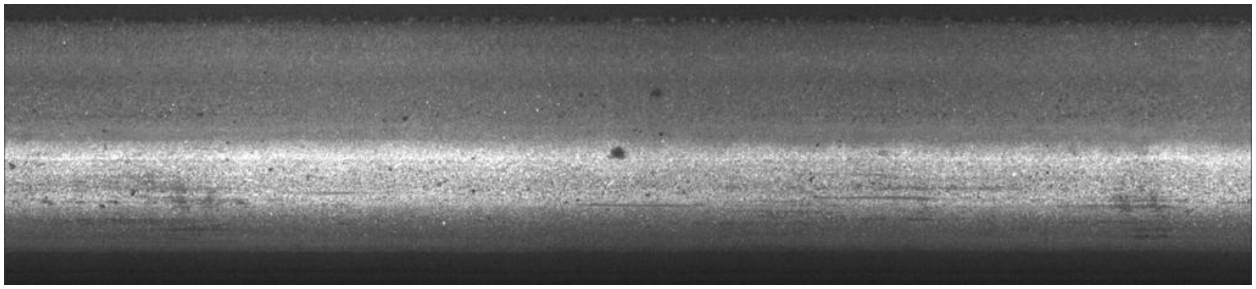
Method 1 takes the centre point between the furthest right peaks of the plotted histogram. This method takes into account that on an ideal image with one running band and no noise there

should be two prominent peaks. The first of a higher nature being the running band that is reflective in nature. The second would be the non-running band data something towards the lower values of the greyscale spectrum. A value somewhere within the middle should suppress the unwanted non-running band data.

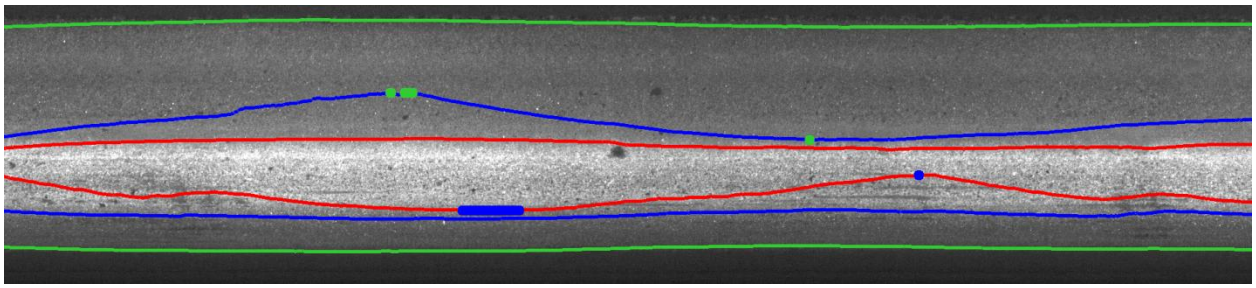
## APPENDIX 8 SEGMENTATION OF RUNNING BANDS USING COMPUTATIONAL SUPPRESSION

This appendix demonstrates the ability to classify 3 phases of running band. The Inner Phase 2 running band represents the path taken by most wheel sets on a vehicle. The Inner Phase 1 shows the lesser travelled path were corrugation can be determined in its early stages. Finally, the wider running band can be classified to demonstrate the total reflective running band area of the railhead. The wider running band area is the only running band that can be evaluated through human visual inspection methods.

Presented with the original image is the classified running band edge cue data from all three running bands.

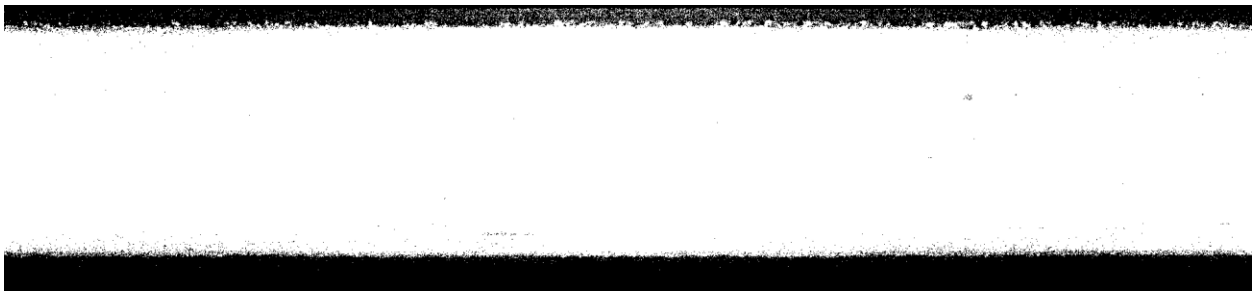


Original Image

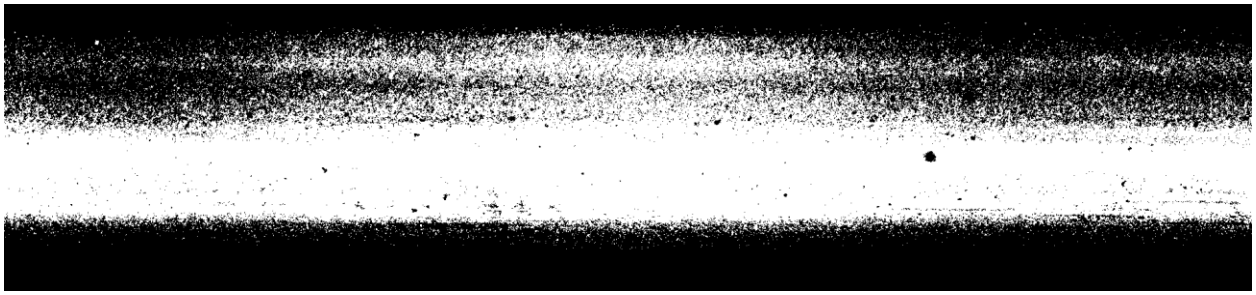


Classification results

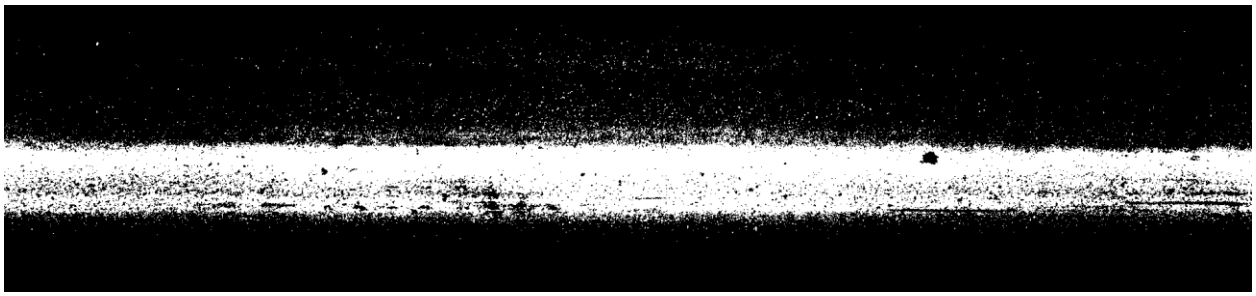
Wider Running Band
Inner Phase 1 Running Band abnormalities in green
Inner Phase 2 Running Band abnormalities in blue



Wider Running Band Threshold Results



Inner Phase 1 Running Band Threshold Results

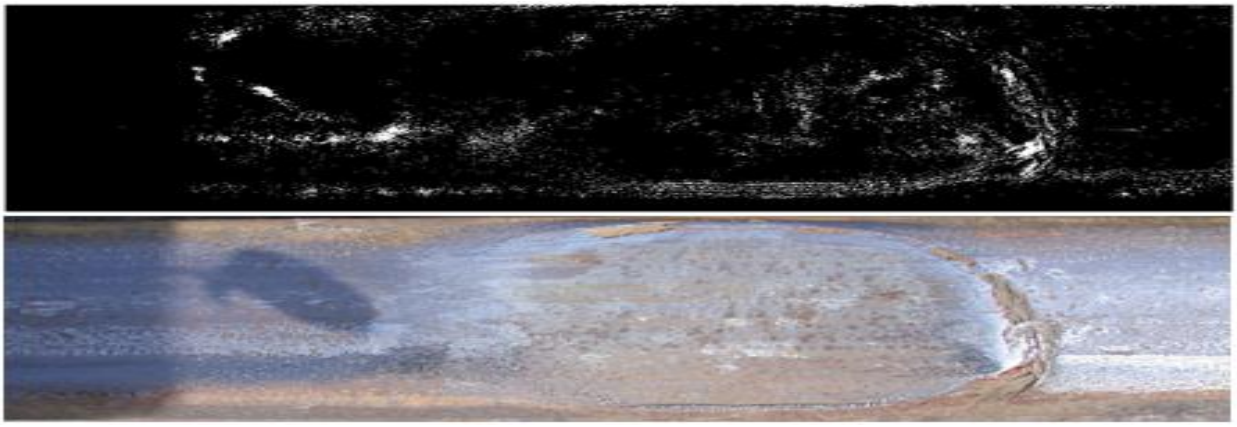


Inner Phase 2 Running Band Threshold Results

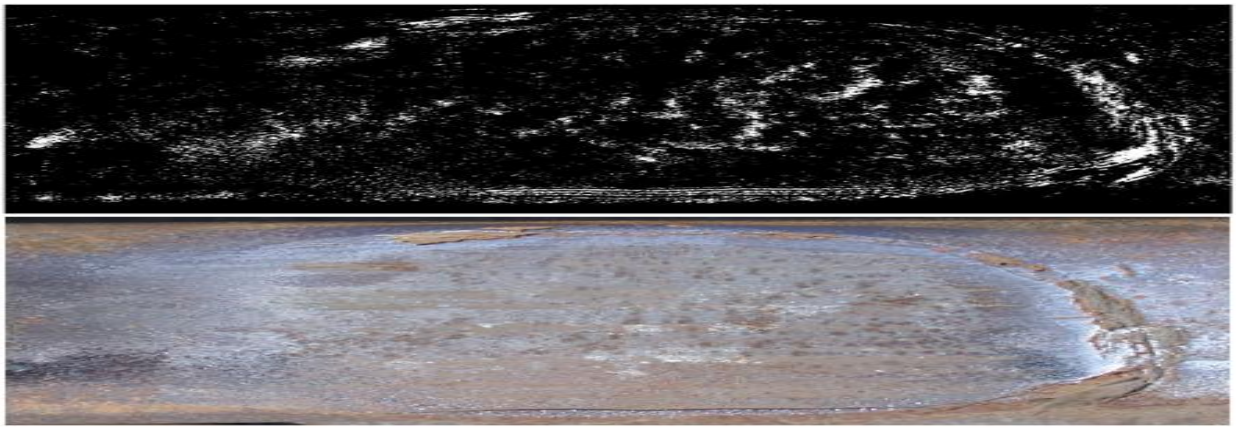
## APPENDIX 9 RESULTS FROM NEURAL NETWORK (NN) FROM TEXTURE ANALYSIS OF A GREYSCALE IMAGE

This Appendix shows the results from a NN, with each image is the resultant binary map in which, white areas are the classified squat data. A NN was trained for each image to ensure that the data for classification was consistent with the data to be classified. While the results show some correct classification of the squat data, the results are neither satisfactory nor consistent.

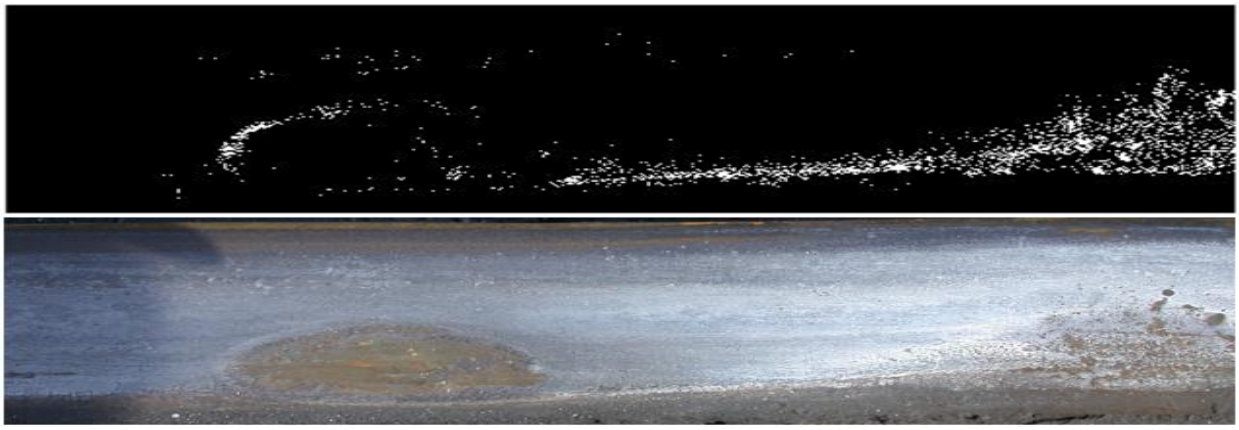
A mask size of 16x16 pixels was used to train the NN. This mask was applied to each pixel with the NN then asked to classify the results. A 100 pixels of squat data and its surrounding mask were used for training with 100 pixels of non-squat running band data.



Results from Squat Detection Image 1



Results from Squat Detection Image 2

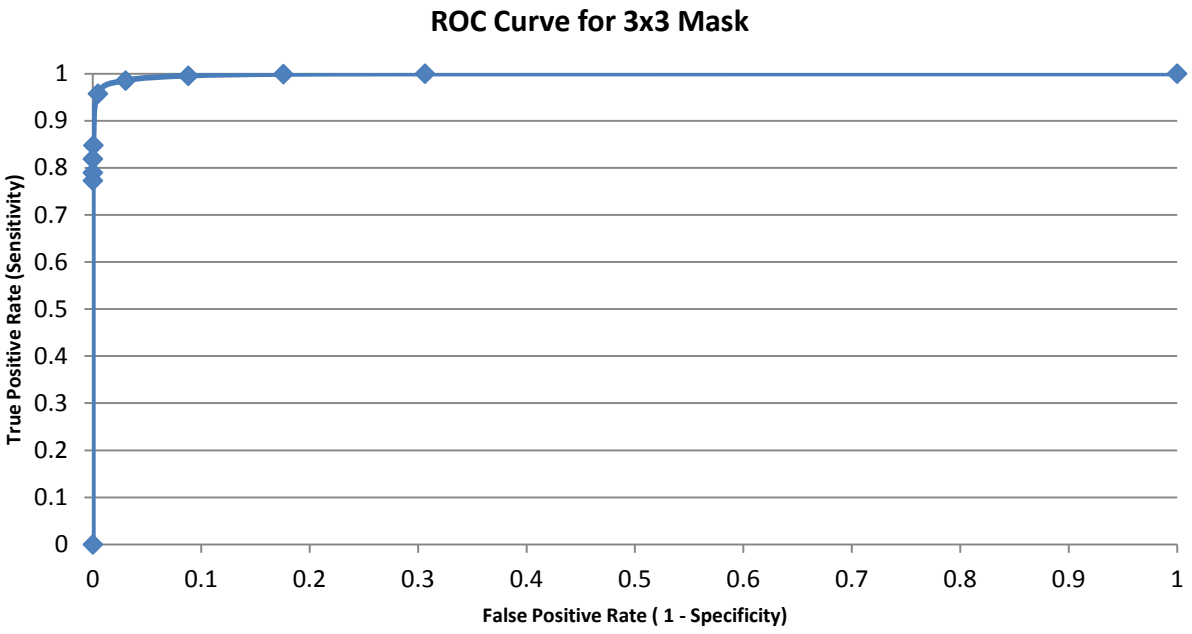


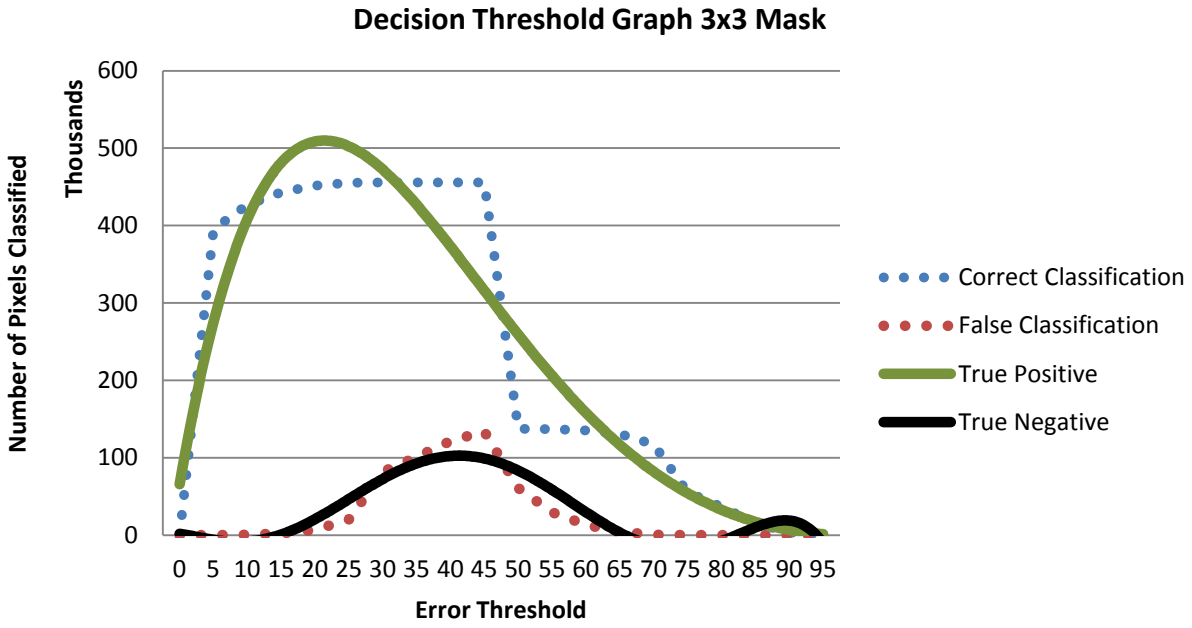
Results from Squat Detection Image 3

## APPENDIX 10 ROC CURVES AND DECISION THRESHOLD GRAPHS FOR AVERAGING MASKS 3X3 - 30X30

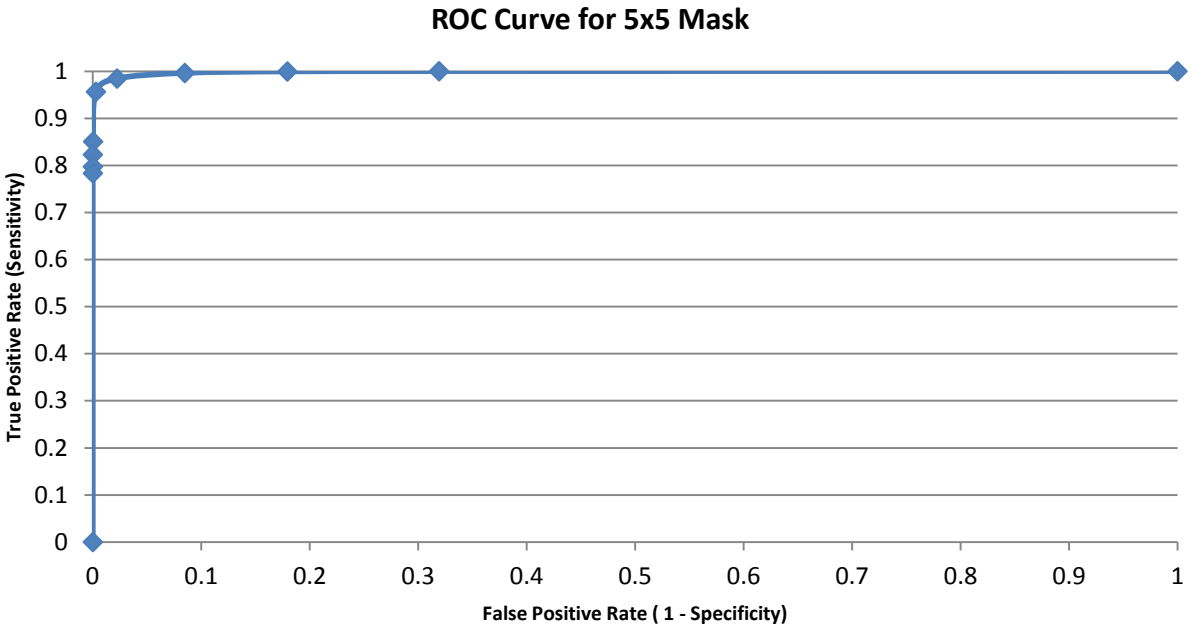
This Appendix presents the ROC decision threshold plots for evaluation of running band data using a Mean testing algorithm. Calculating the mean of all the pixels produces a value that can be compared to a threshold. A decision based on this threshold was then be made in order to classify the pixels as belonging to a surface bound abnormality or to the expected running band.

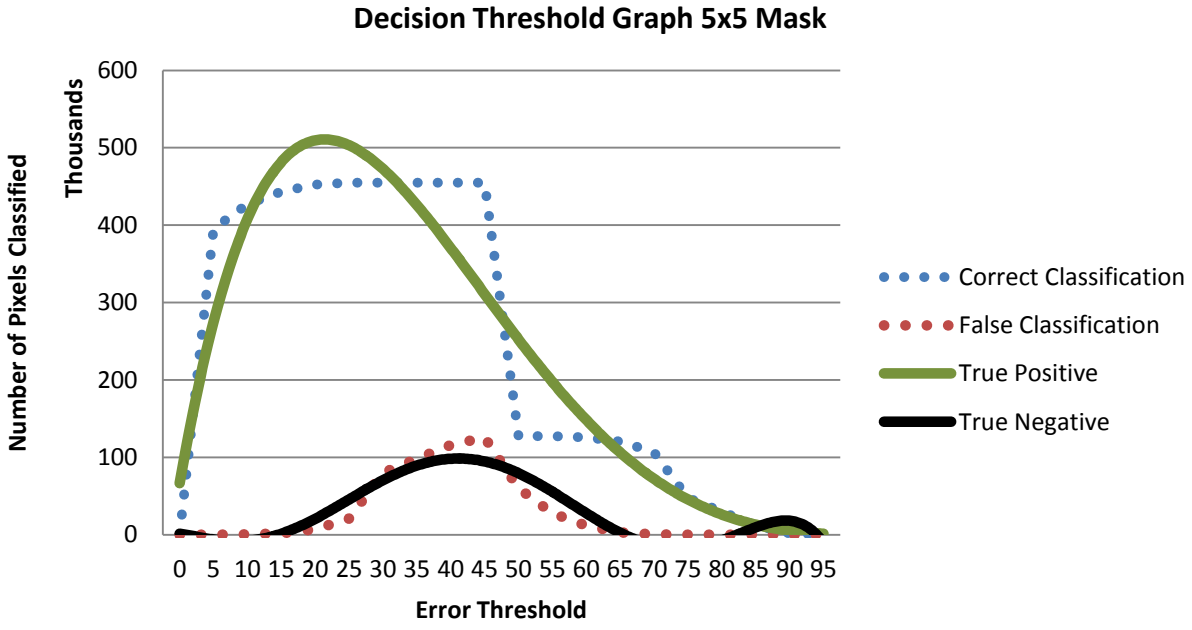
Evidence suggests that a 10x10 mask allowing a 40% error will produce the best available results. The supporting ROC decision threshold plot Figure 9.7 supports this finding but suggest a lower threshold of 35% would produce a better True positive to false positive rate.



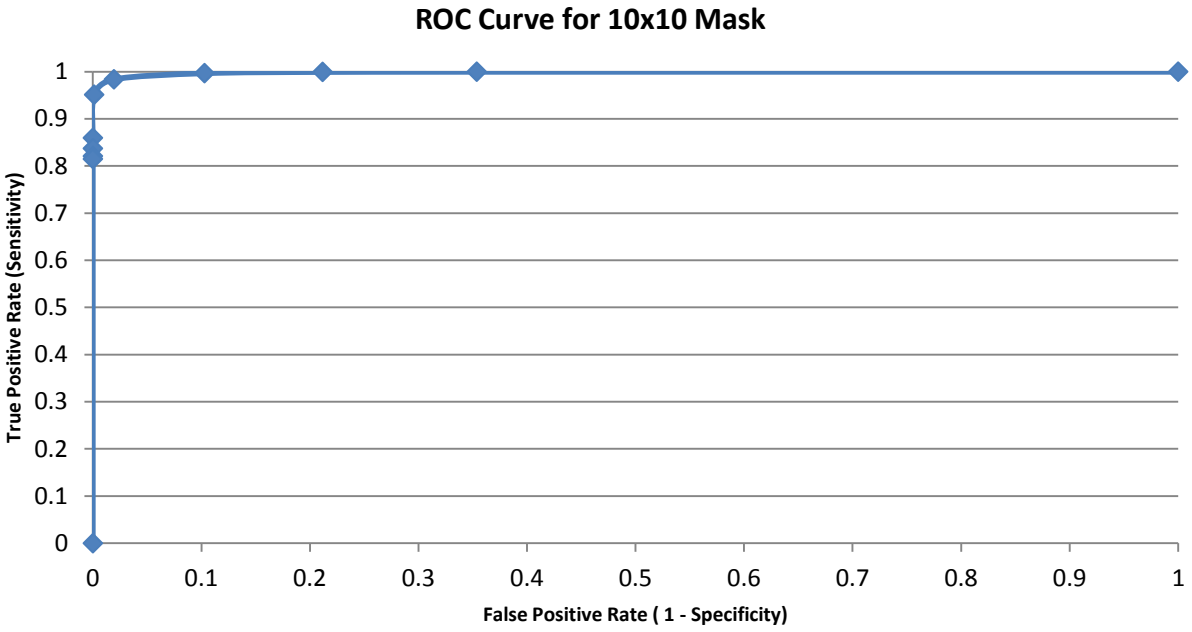



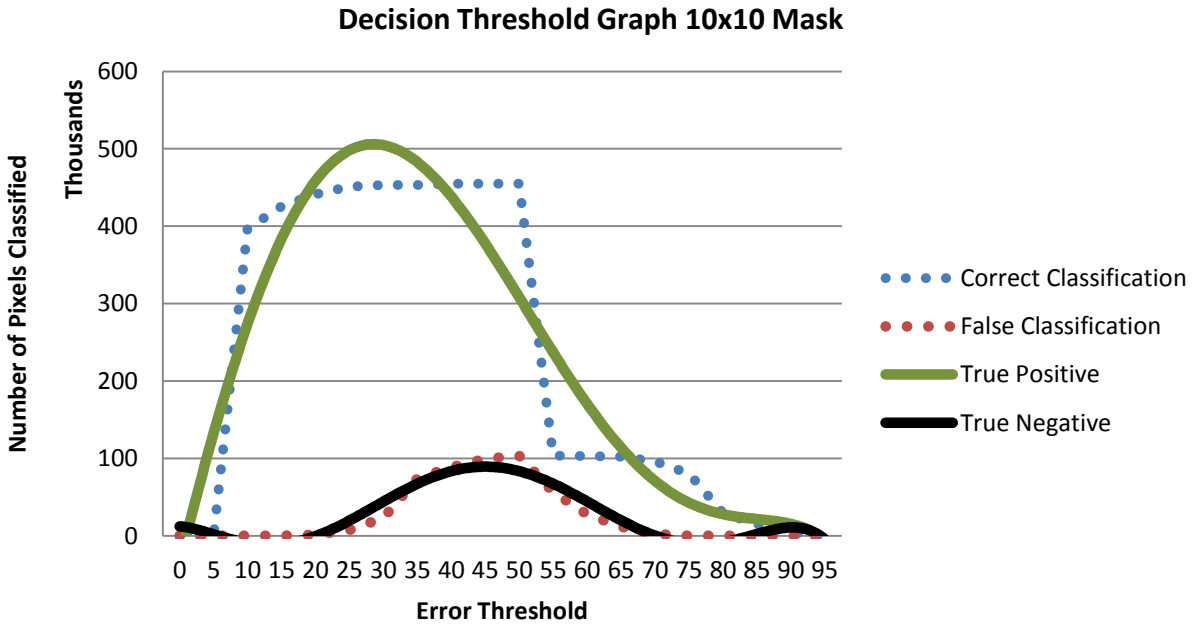

**ROC Curve and Decision Threshold Graph for 3x3 Mask**



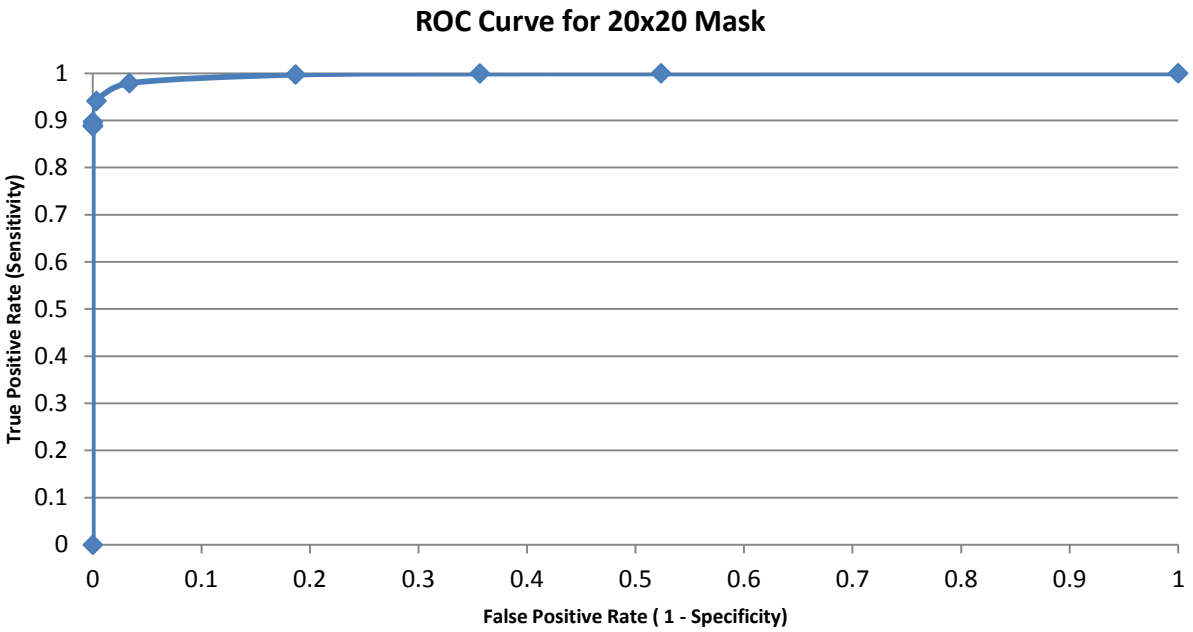



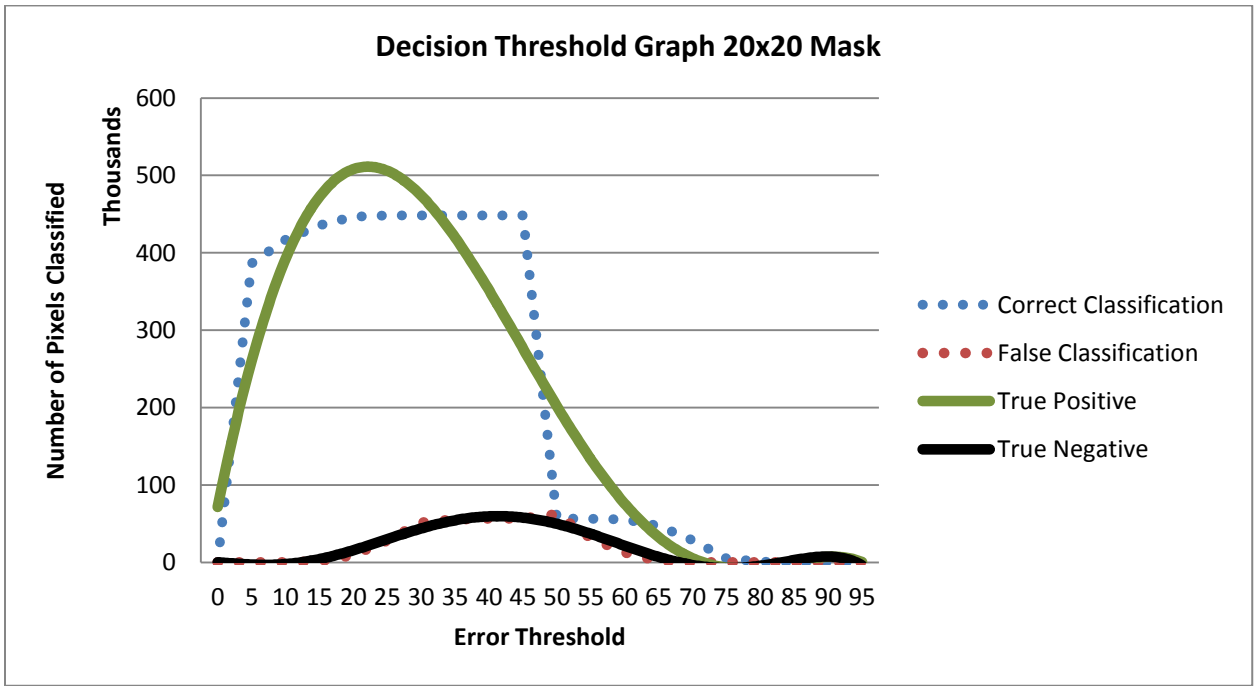

**ROC Curve and Decision Threshold Graph for 5x5 Mask**



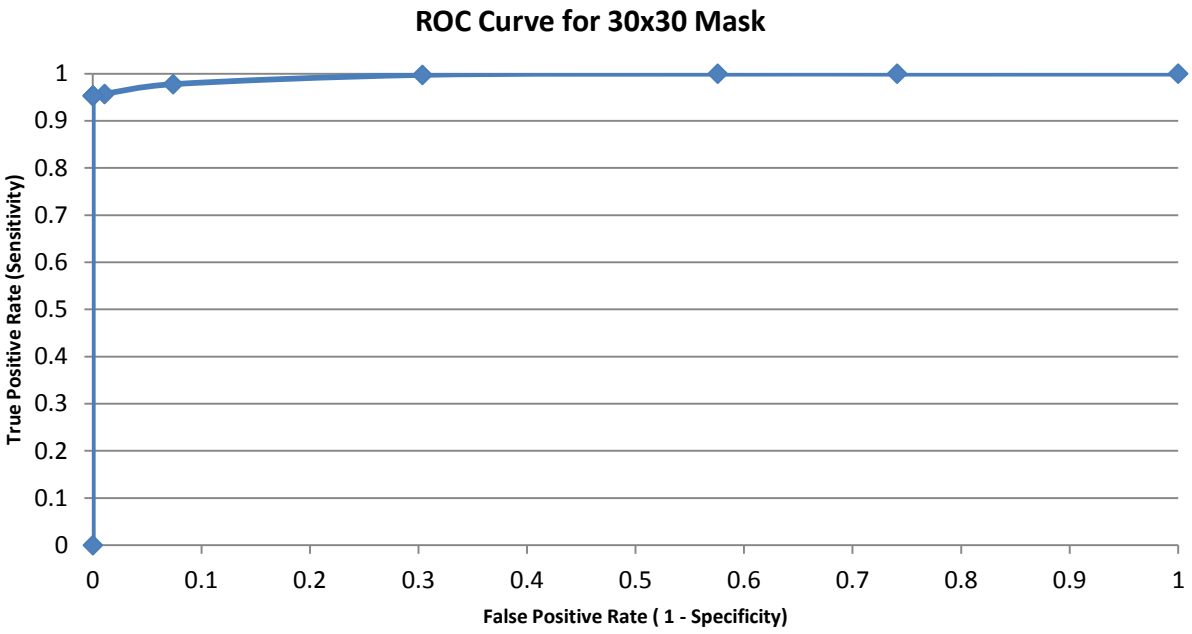



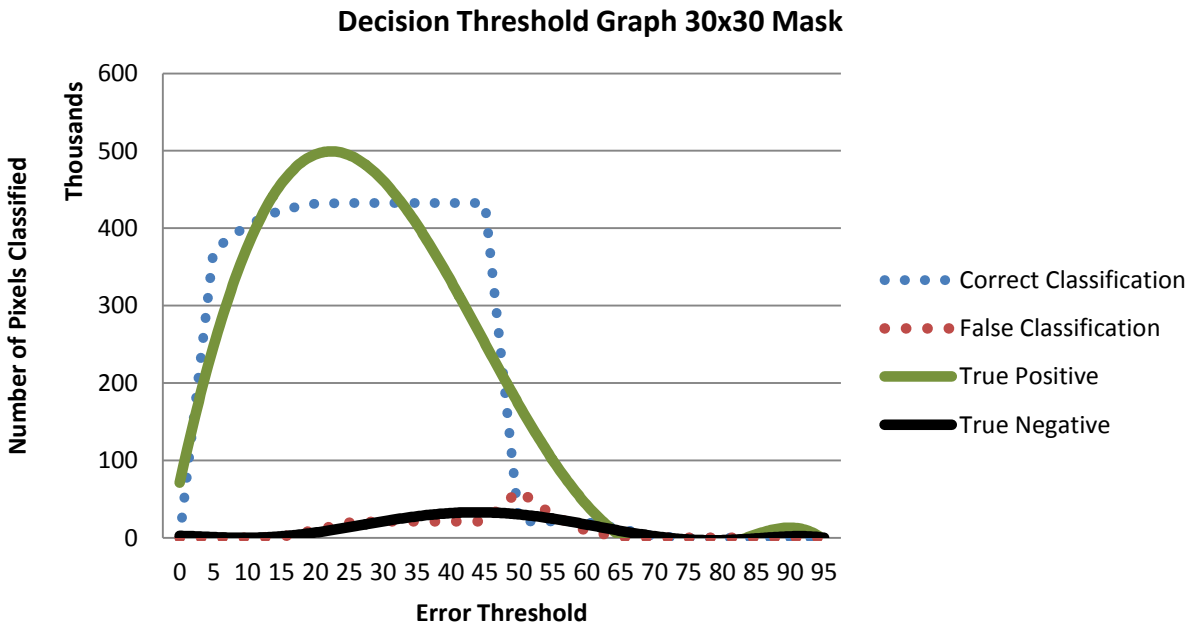

**ROC Curve and Decision Threshold Graph for 10x10 Mask**




**ROC Curve and Decision Threshold Graph for 20x20 Mask**



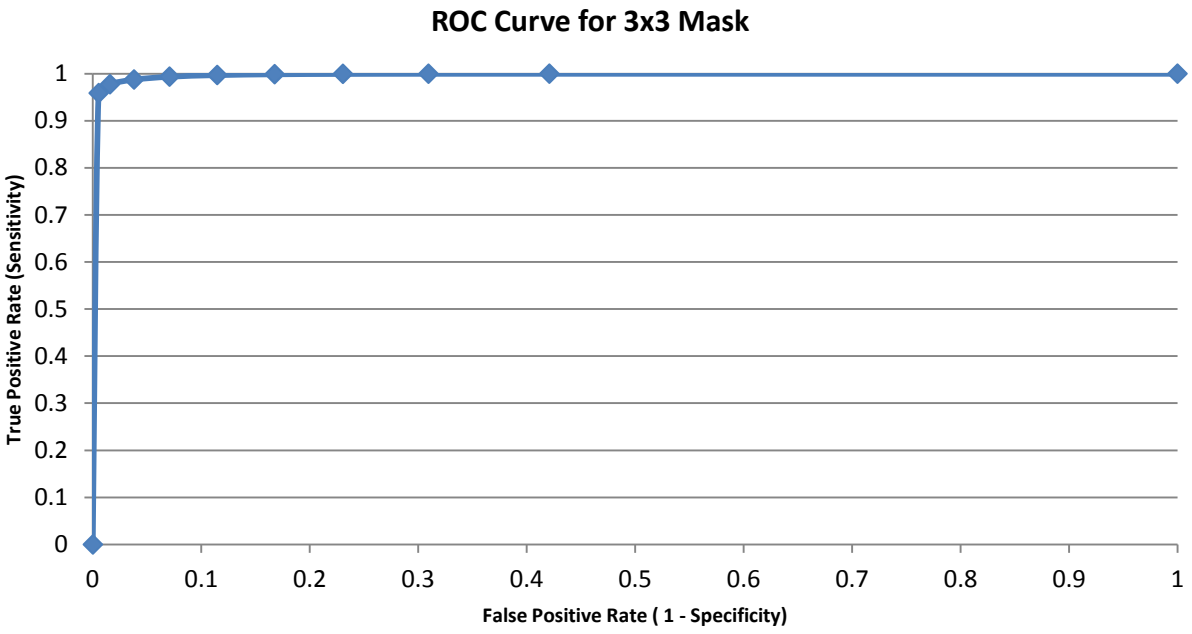



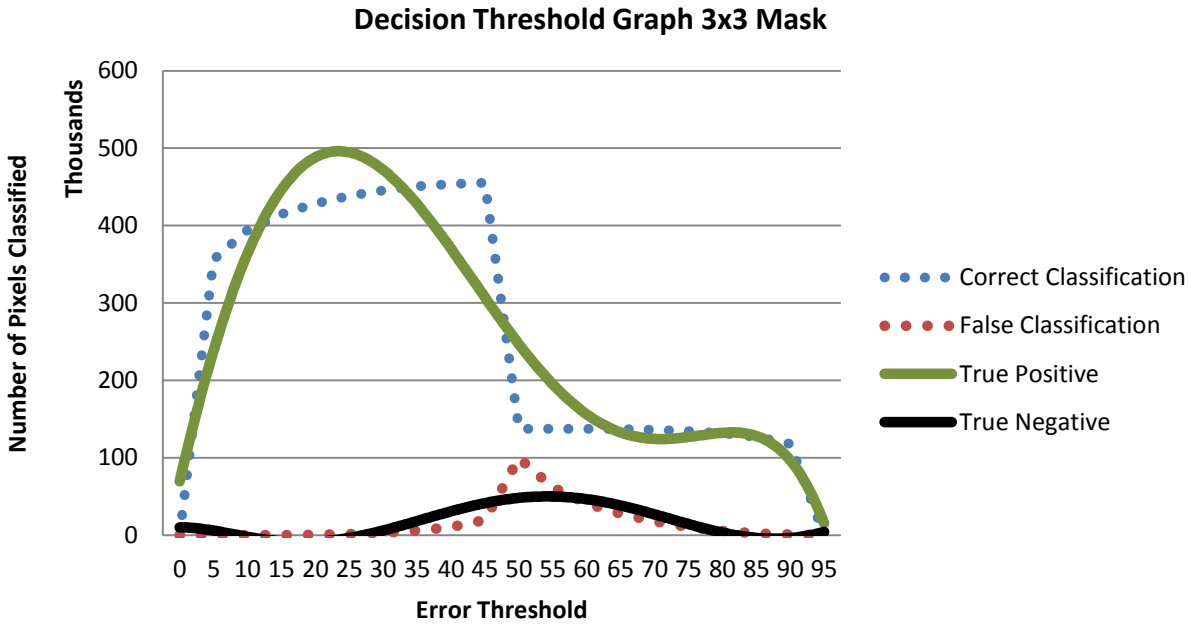

**ROC Curve and Decision Threshold Graph for 30x30 Mask**

# APPENDIX 11 ROC CURVES AND DECISION THRESHOLD GRAPHS FOR STANDARD DEVIATION MASKS 3X3 - 30X30

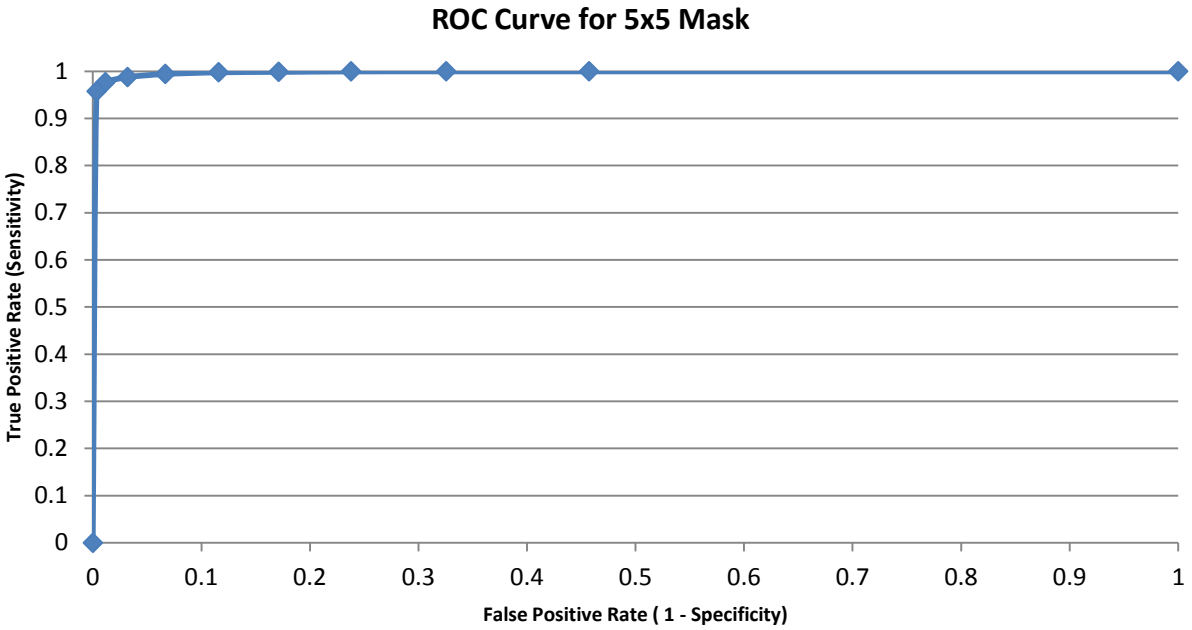
The Appendix presents the ROC decision threshold plots for evaluation of running band data using a standard deviation testing algorithm. Calculating the standard deviation of all the pixels produces a value that can be compared to a threshold. A decision based on this threshold was then be made in order to classify the pixels as belonging to a surface bound abnormality or to the expected running band.

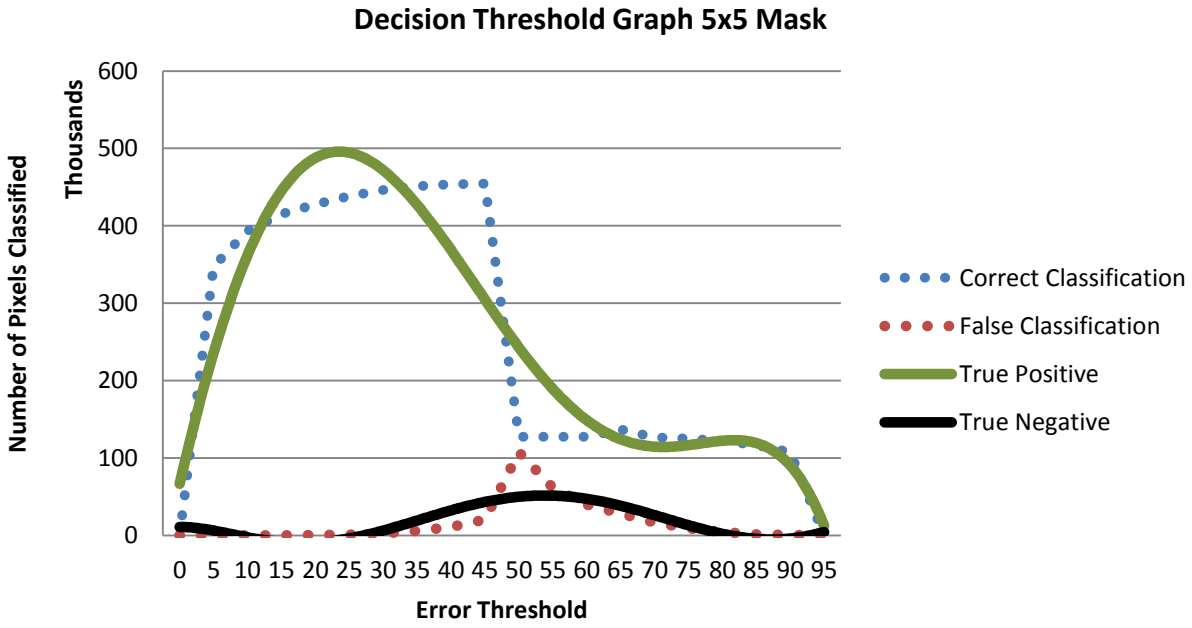
Evidence suggests that a 10x10 mask size produces the best results from the overall percentage error data. This is not supported with the following ROC data that suggests a 5x5 mask with a standard deviation threshold of less than 30 would produce the best classification accuracy.



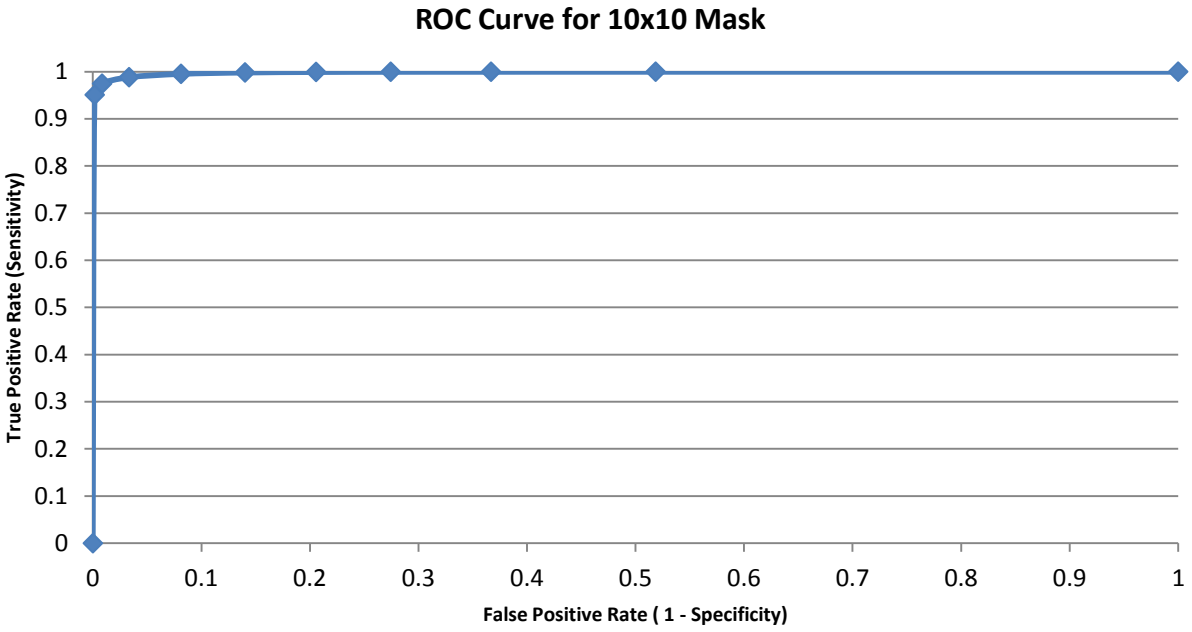



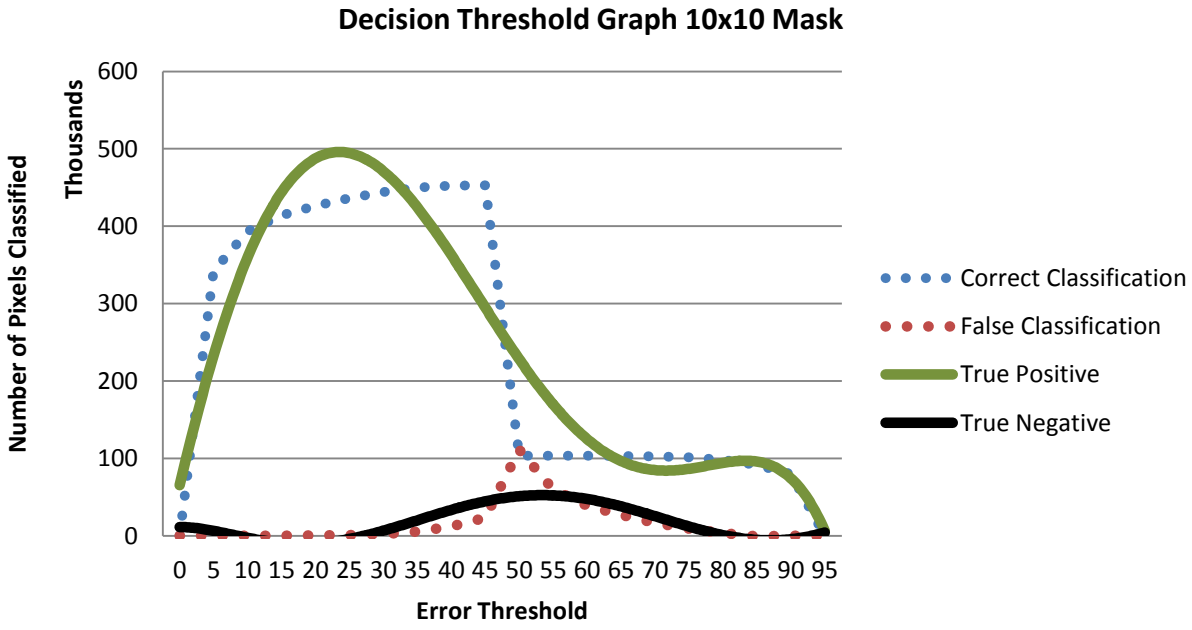

**ROC Curve and Decision Threshold Graph for 3x3 Mask**



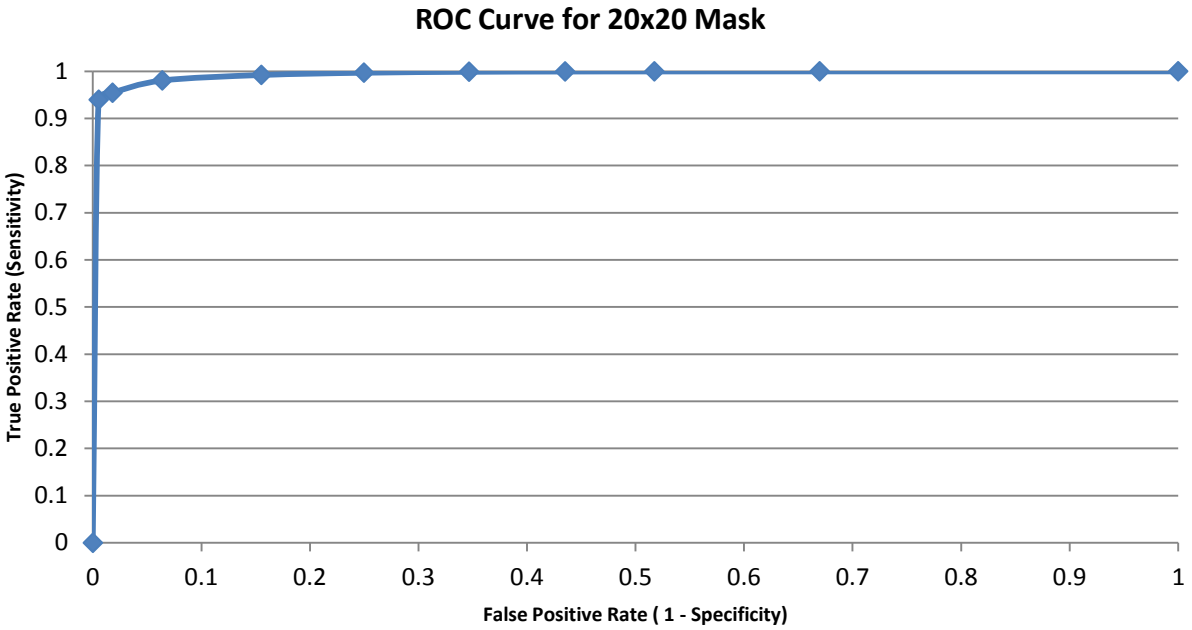



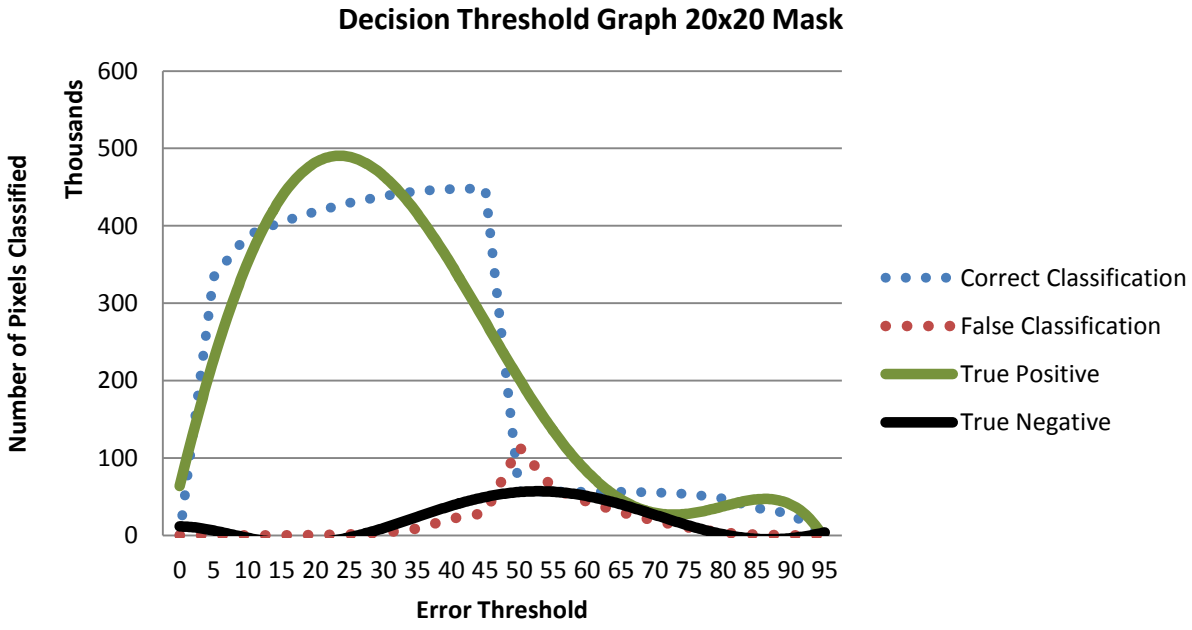

**ROC Curve and Decision Threshold Graph for 5x5 Mask**



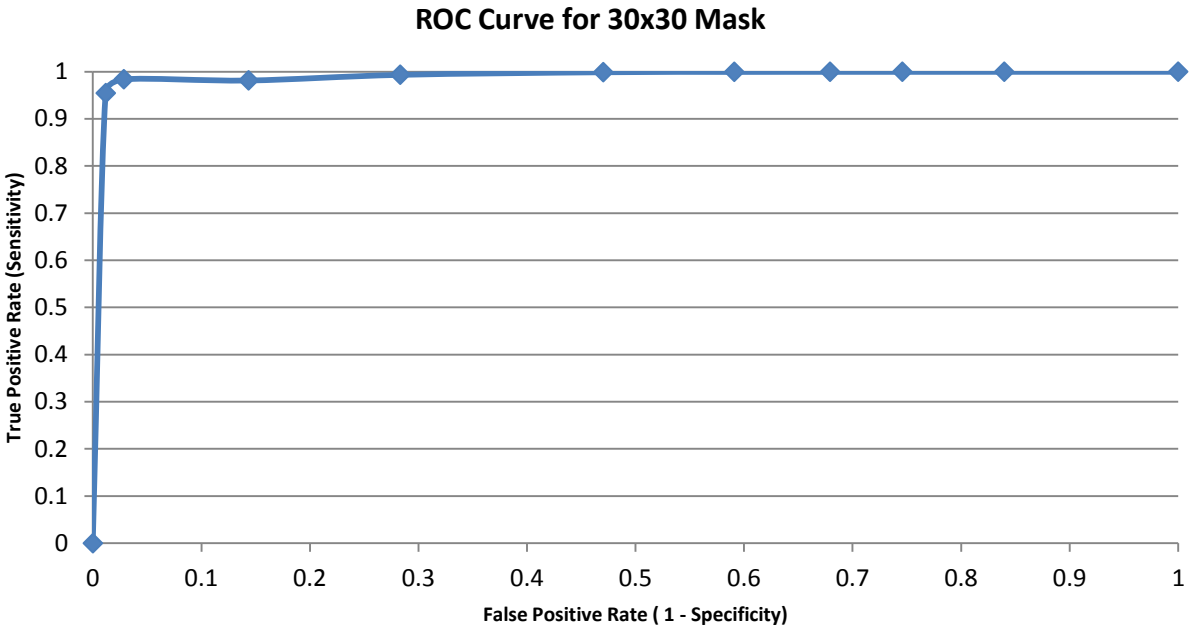



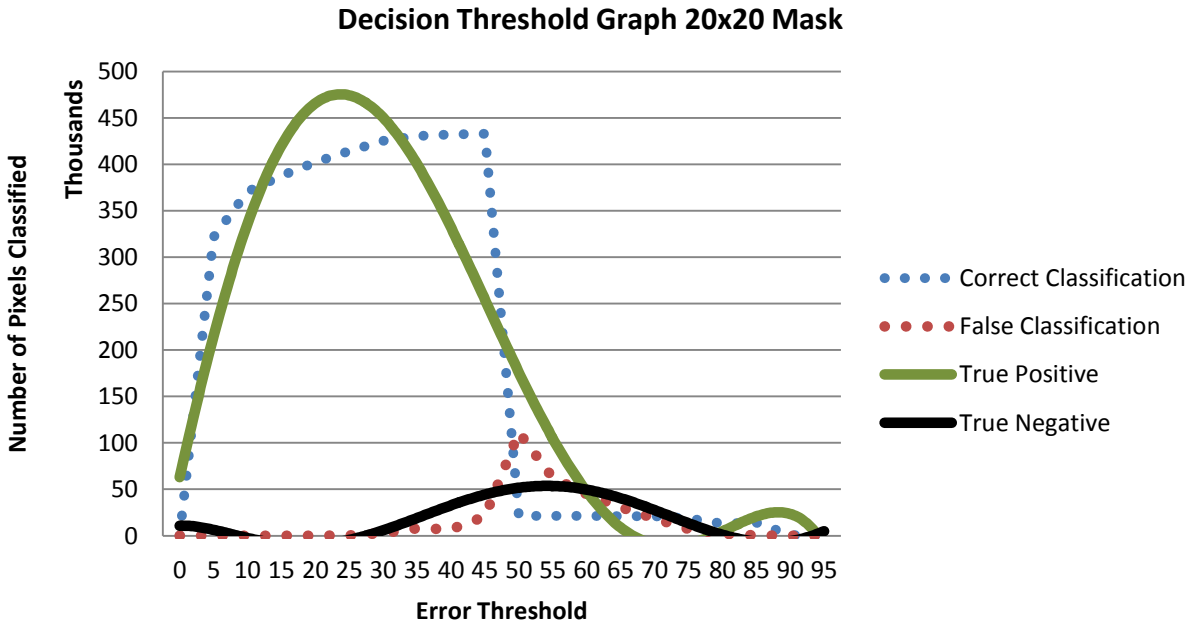

**ROC Curve and Decision Threshold Graph for 10x10 Mask**




**ROC Curve and Decision Threshold Graph for 20x20 Mask**




**ROC Curve and Decision Threshold Graph for 30x30 Mask**