MACHINE VISION FOR AUTOMATING VISUAL INSPECTION OF WOODEN RAILWAY SLEEPERS

SAJJAD PASHA MOHAMMAD

2008

Master Thesis
Computer Engineering
Reg.No:E3551D
Abstract

Condition monitoring procedure within rail transportation domain is carried out manually by skilled personnel. Visual inspection, the most common approach used in condition monitoring process by human operators and is applied widely ensuring transport safety. Since the human perception has the capability to acquire, integrate all the visual information around oneself. However, in this thesis work the automatic visual inspection used for monitoring the condition of wooden railway sleepers adapts such intuitive human skills for the development of robust and reliable condition monitoring system.

Railway sleeper inspection is a typical example of condition monitoring procedure which is carried out by hand, and an inspector appointed for its maintenance examines the condition of each sleeper in turn for the presence of cracks on the wooden sleepers along with the metal plates and nails used in rail fastening’s are working with proper condition or not. For visual inspection process the inspector in-charge of sleeper maintenance has to walk along the track to examine each and every sleeper in turn. This visual inspection process is very slow and expensive, since this activity is labour intensive, and operators are prone to fatigue, which could have faulty results due to human errors. Machine vision, however, is now capable of providing valuable assistance by drawing attention to predefined events of interest for the reliability of the visual condition monitoring process.

Today, machine vision is used to ensure the quality of everything from tiny computer chips to massive space vehicles. Machine vision has evolved to become a vital and effective automation tool that enables computers to replace human vision in many high-speed and high-precision manufacturing applications. Machine vision is the solution designed and applied for
the visual inspection of railway sleepers to identify the flaws associated with the sleepers such as cracks on the sleepers, and condition of the rail fastenings in the scene etc. Machine vision has applications in many aspects of transport like traffic monitoring, aerial navigation, traffic safety, etc. As it can be used to collect and understand data useful in transport planning and safety. It can also play an active role in navigation and vehicle guidance due to its robustness and reliability. This machine vision algorithm is experimentally evaluated using 200 real images of the railway sleepers to determine the condition of those wooden railway sleepers. In this process image acquisition, the first stage of any vision system is carried to capture the real images.

Next to the image acquisition process, a machine vision algorithm using various methods of image processing techniques is applied to the image data collected in order to extract its features such as the number of cracks, length of the crack, width of the crack, and length of the metal plate. Finally, these extracted features using machine vision algorithm are then stored in separate feature vectors and were used further for the classification task using pattern recognition techniques. Classification of the features extracted is carried out by using classifiers like Support vector machines (SVM) and Radial Basis function (RBF), though there are several other classifiers available. Both the pattern classifiers, SVM and RBF, have been experimented and compared with each other for the reliability and accuracy of the classification. However, sound analysis is yet another technique useful in detecting the condition of the rail fastenings involved in this condition monitoring process. And sound features could be fused along with the image features during pattern classification.

Results achieved experimentally demonstrate that SVM with Gaussian kernel has good percentage of classification rates. The classification rates of both the training and the test sets are 95.33% and 86.00% respectively using the SVM classifier. But comparatively the RBF classifier with k-means has classification rates of the training set and test sets as 94.67% and 86.00% respectively. Thus, the condition of the wooden railway sleepers whether they are good or bad for its further usage is determined with the system developed. These experiments have been helpful to improve the robustness and reliability of the system in automating the visual inspection process of condition monitoring. However, the results might tend to report some unexpected results when the system fails and it is possibly due to unclear image data collected during image acquisition.
Acknowledgement

I, Sajjad Pasha Mohammad, do acknowledge my thesis supervisor Mr. Siril Yella, a PhD student at the department of computer science, Högskolan Dalarna, for all the support that he has provided with having immense knowledge in the field of Artificial Intelligence, and NDT area which happened to be very helpful in the successful accomplishment of my thesis work.

And also, I am thankful to the department of computer science, Högskolan Dalarna for providing all the resources required to complete my thesis work.
Table of Contents

Introduction..........................................................................................................................1
  1.1 Problem-Railway sleeper inspection ...........................................................................3
  1.2 Proposed Solution .......................................................................................................4
  1.3 Goals and Scope ........................................................................................................6
Chapter 2..............................................................................................................................7
  2 Machine Vision in Transportation ..............................................................................7
    2.1 Machine Vision in Rail transportation ....................................................................8
    2.2 Potential Difficulties ...............................................................................................9
      2.2.1 Image Acquisition ............................................................................................10
      2.2.2 Machine Vision Algorithm ...............................................................................13
Chapter 3............................................................................................................................17
  3 Feature Extraction .......................................................................................................17
    3.1 Number of cracks ....................................................................................................24
    3.2 Length of the crack ...............................................................................................24
    3.3 Width of the crack .................................................................................................25
    3.4 Length of the Metal plate .......................................................................................25
Chapter 4............................................................................................................................31
  4 Pattern Recognition .......................................................................................................31
    4.1 Support Vector Machines .......................................................................................33
    4.2 Radial Basis Function ............................................................................................35
Chapter 5............................................................................................................................37
  5 Results and Discussion ...............................................................................................37
Chapter 6............................................................................................................................41
  6 General Conclusions and Future Work ......................................................................41
References..........................................................................................................................43
List of Figures & Tables

Chapter 2
Figure 2.1: Good Sleeper Image (Top View).
Figure 2.2: Plate Image of a Good Sleeper (Horizontal View).
Figure 2.3: Bad Sleeper Image (Top View).
Figure 2.4: Plate Image of a Bad Sleeper (Horizontal View).

Chapter 3
Figure 3.1: Bad Sleeper Image after image resizing.
Figure 3.2: Good Sleeper Image after image resizing.
Figure 3.3: Bad Sleeper Image after gray-level operation.
Figure 3.4: Good Sleeper Image after gray-level operation
Figure 3.5: Bad Sleeper Image after Gaussian filter operation
Figure 3.6: Good Sleeper Image after Gaussian filter operation.
Figure 3.7: Bad Sleeper Image after binary operation
Figure 3.8: Good Sleeper Image after Binary operation.
Figure 3.9: Bad Sleeper Image after majority operation.
Figure 3.10: Good Sleeper Image after majority operation.
Figure 3.11: Bad Sleeper Image after imclose operation.
Figure 3.12: Good Sleeper Image after imclose operation.
Figure 3.13: Bad Sleeper Image after Skeleton operation.
Figure 3.14: Good Sleeper Image after Skeleton operation.
Figure 3.15: Bad Sleeper Image after Pruning operation.
Figure 3.16: Good Sleeper Image after Pruning operation.
Figure 3.17: Bad Sleeper Image to calculate plate length
Figure 3.18: Good Sleeper Image to calculate plate length
Figure 3.19: Bad Sleeper Image after Gray-level operation
Figure 3.20: Good Sleeper Image after Gray-level operation
Figure 3.21: Bad Sleeper Image after Gaussian filter operation
Figure 3.22: Good Sleeper Image after Gaussian filter operation
Figure 3.23: Bad Sleeper Image after Canny edge-detection operation
Figure 3.24: Good Sleeper Image after Canny edge-detection operation
Figure 3.25: Bad Sleeper Image after morphological Skeleton operation
Figure 3.26: Good Sleeper Image after morphological skeleton operation
Figure 3.27: Bad Sleeper Image after Pruning operation
Figure 3.28: Good Sleeper Image after pruning operation

Chapter 4
Figure 4.1: Linear separating hyperplanes for the separable classes.
Figure 4.2: Linear separating hyperplanes for the non-separable classes.
Figure 4.3: Radial basis network architecture.

Chapter 5
Table 5.1: Confusion matrix.
Table 5.2: Training set Error rate using SVM.
Table 5.3: Confusion matrix.
Table 5.4: Test set Error rate using SVM
Table 5.5: Confusion matrix.
Table 5.6: Training set Error rate using SVM
Table 5.7: Confusion matrix.
Table 5.8: Test set Error rate using RBF
Table 5.9: Pattern classification using two different methods: SVM and RBF.
Chapter 1

Introduction

In today’s world transport or transportation is the movement of people and goods from one place to another. Industries involved with the business of providing equipment, actual transport, transport of people or goods and services used in transport of goods or people make up a large broad and important sector of most national economies, and are collectively referred to as transport industries. The field of transport has several aspects, divided into a triad of infrastructure, vehicles, and operations. Infrastructure includes the transport networks such as roads, railways, airways, waterways, etc. that are used. The vehicles generally ride on the networks, such as automobiles, buses, trains, and aircrafts. The operations deal with the way the vehicles are operated on the network. However, the design of networks are the domain of civil engineering and urban planning, the design of vehicles of mechanical engineering and specialized subfields such as nautical engineering and aerospace engineering, and the operations are usually specialized, though they might appropriately belong to operations research or systems engineering. And, now-a-days the transport industries have widely adopted condition monitoring routines in many of its subject areas with the main purpose to ensure safety and reliability in its day-to-day operations. Due to this heavy duty of transportation and its dynamic nature it is very hard to have absolute control over all the condition monitoring operations in the transportation domain, because the operating environment is always subjected to change. However, the infrastructure being used in transportation is regularly monitored in order to make sure that they continue to work with proper condition. Since any failure of the safety management system in order to control the risk which could be resulting in accidents and other relevant troubling incidents like serious destruction of the infrastructure or property and also could lead to injury or death of passengers or crew members. Forexample, the Hatfield accident of October 2000 in the U.K. (Rail Safety & Standards Board, 2006; Siril Y., et al., 2006) and the Liverpool accident on 26 October 2005; train derailment due to track fault which was reported in the BBC News (BBC News, 2007) were considered as condition monitoring failures. Another incident, the Grayrigg derailment on 23 February 2007 is one of the fatal railway accident in England occurred due to a faulty set of points (Grayrigg derailment, 2007) showing the importance and need for condition monitoring. And condition monitoring applications extensively deploy the usage of non-destructive testing (NDT) procedures (Vary A., 1972) within the transportation domain.

Non-destructive testing (NDT) is the branch of engineering concerned with all the methods of detecting and evaluating flaws in materials. And it is not just a crystal ball but it is the scientific use of physical principles to study the fitness for service of a material, component, or assembly of parts without impairing its serviceability. NDT is extremely diverse and multi-disciplinary, involving the various fields related to applied physics, artificial intelligence, biomedical engineering, computer science, electrical engineering, electronics, materials science and engineering, mechanical engineering, and structural engineering (SEALIFT, 2007). And flaws in materials can affect the materials serviceability or structure while NDT testing is important in guaranteeing safe operation as well as in quality control and assessing the life of the infrastructure. The flaws may be cracks or inclusions in welds and castings, or
variations in structural properties which can lead to loss of strength or failure in service. NDT testing is used for in-service inspection of operating infrastructure. It is also used for measurement of components and for the measurement of physical properties such as hardness and internal stress. The essential feature of NDT is the test process itself, produces no deleterious effects on the material or structure under test. The subject of NDT has no clearly defined boundaries; it ranges from simple techniques such as visual examination of surfaces, through the well-established methods of radiography, ultrasonic testing, magnetic particle crack detection, to new and much specialised methods such as the measurement of Barkhausen noise and positron annihilation. NDT methods can be adapted to automated production processes as well as to the inspection of localised problem areas (BINDT, 2007). NDT makes it possible to examine the condition or the quality of an object without damaging it. It is used to improve safety and ensure quality which is indeed very important to many industries and businesses such as energy sector, offshore industry, construction industry, transport sector, machine factories, etc. The key objectives of NDT:

- Improving performance in terms of reliability and punctuality
- Increasing capacity and availability
- Maintaining cost-effectiveness
- Integrating all of these to compete effectively with other transport modes (or complement them as appropriate) and deliver a sustainable future

NDT implementation is very important to describe what shall be found and what to reject. A completely flawless production is almost never possible. For this reason testing specifications are indispensable. And these days there existing a great number of standards and acceptance regulations. They describe the limit between good and bad conditions, but also often which specific NDT method has to be used. The reliability of NDT method is an essential issue. But a comparison of methods is also significant when referring to the similar kind of task. Each NDT method is having its own set of advantages and disadvantages and, therefore, some are better suited than others for a particular application. And by the use of artificial flaws, the threshold of the sensitivity of a testing system can also be determined. If the sensitivity is to low then defective test objects are not always recognized. If the sensitivity is too high parts with smaller flaws are rejected which would have been of no consequence to the serviceability of the component. With statistical methods it is possible to look closer into the field of uncertainly. Methods such as Probability of Detection (POD) or the ROC-method "Relative Operating Characteristics" are examples of the statistical analysis methods. Also the aspect of human errors has to be taken into account when determining the overall reliability (NDT, 2007). But the sensitivity of the testing system is not the major issue in this work.

At the same time a personnel qualification is also an important aspect of non-destructive evaluation. NDT techniques rely heavily on human skills and knowledge for the correct assessment and interpretation of test results. Proper and adequate training and certification of NDT personnel is therefore a must to ensure the capabilities of the techniques being exploited completely. There are a number of published international and regional standards covering the certification of competence of personnel. The EN 473 (Qualification and certification of NDT personnel - General Principles) was developed specifically for the European Union for which the SNT-TC-1A is the American equivalent (NDT, 2007). During recent years the NDT is being used for condition monitoring of various materials and civil engineering structures with a goal of assessing embedded flaws quickly and accurately in a cost effective fashion (Vary...
A., 1972). Samuel and Robert state that all NDT test procedures have five essential features (S A Wenk et al., 1987):

1. Supplying a suitable form and distribution of energy from an external source to the test object.
2. The energy distribution within the test object is modified as a result of discontinuities, flaws or other variations in material properties which correlate to serviceability.
3. Detecting the change in energy properties by a sensitive detector.
4. Indicating or recording the energy measurement from the detector in a form useful for interpretation, and
5. Interpreting the indication and judging the corresponding serviceability of the test object.

NDT has made the task of automatic interpretation of the data for condition monitoring very easy to carry out operations.

In the recent years, the issues like automatic interpretation of NDT data using artificial intelligence (A.I.) had been interesting for the NDT research community. NDT is one such area that has been using AI techniques for the automatic interpretation of NDT methods to reduce the intervention of a human operator. Current non-destructive assessments demand efficient knowledge, storage and retrieval facilities and also involve one or more combinations of methods for the accurate evaluations of the structures. And AI methods promise a fast, reliable classification of flaws and better ways of data interpretation, facilitating efficient information storage and retrieval techniques. Accordingly, many areas of transportation engineering have adapted NDT procedures to make key assessments of infrastructure and other engineering structures. Since several areas of transportation are in demand for reliable, accurate and fast solutions to avoid catastrophic results. For example, catastrophic results could be like derailment of trains, plane crashes, and some other less visible but equally troubling events which are considered to be challenging problems in the transport sector. There are several kinds of techniques available in A.I. for the purpose of automatic interpretation of NDT data such as neural networks, case-based reasoning, machine vision, expert systems, fuzzy logic and genetic algorithms, etc. Because the data yielded by NDT techniques or procedures are usually in the form of signals, images etc. which often do not present direct information of the condition of the structure or product. And the data is interpreted by a skilled human operator but are often unreliable, since human operators are challenged by many factors. However, the automated computer-based techniques, such as AI, offers a significant attraction for the interpretation of NDT data since they offer advantages in improved speed and accuracy of analysis, especially when large-volume automated inspection information is provided, and been applied to a wide range of problems in the area.

1.1 Problem-Railway sleeper inspection

Railway sleeper inspection is a typical example of condition monitoring procedure which is carried out by hand, and an inspector appointed for its maintenance examines the condition of each sleeper in turn for the presence of cracks on the wooden sleepers along with the metal plates and nails used for rail fastenings are in proper condition or not. During visual inspection, the inspector in-charge of sleeper maintenance has to walk along the track to
examine each and every sleeper in turn. This process of visual inspection is very slow and expensive. And also this activity is labour intensive, and operators are prone to fatigue, which could result in the human errors. Machine vision, however, is now capable of providing valuable assistance by drawing attention to predefined events of interest for the reliability of the visual condition monitoring process.

Within this transportation sector, rail inspection is one of the major problems that railroad companies have faced since the earliest days, for the prevention of service failures in track. As is the case with all modes of high-speed travel, failures of an essential component can have serious consequences. The North American railroads have been inspecting their most costly infrastructure asset, the rail, since the late 1920's. With increased traffic at higher speed, and with heavier axle loads in the 1990's, rail inspection is more important today than it has ever been. Although the focus of the inspection seems like a fairly well-defined piece of steel, the testing variables present are significant and make the inspection process challenging (Rail Inspection, 2007). Rail accidents that happen every year are due to its heavy duty of transportation, which result in serious destruction of property and injury or death of passengers and crew members. These accidents are caused by dangerous failures and potential rail defects such as the rail flaw, sleeper failure, fracture and other potential breaks. Various track components are taken into account after the analysis of the rail failure phenomenon (Bogdaniuk et al, 2003):

- Rail type (weight, grade, heat treatment)
- Sleepers (wooden, concrete)
- Axle loads
- Annual tonnage
- Train speed and traffic structure
- Track vertical and horizontal geometry

Assessing flaws is considered crucial and most of the countries usually have their own testing rail flaw detection equipment. For example, in USA an ultrasonic hi-rail vehicle is used in which the vehicle uses the B-scan based test system (Clark et al, 2004). And a company named MER MEC, in Italy, is producing maintenance vehicles for the support of MER MEC Track Surface Defect Detection, is a vision system designed and manufactured for real-time monitoring of track condition and automatic recognition of resulting defects using high-speed line-scan cameras for track images acquisition (Mer Mec, 2007). Another company in Germany, BVSYS, bvSys Bildverarbeitungssysteme GmbH, System is developed for CrackCheck that inspects concrete sleepers and ballastless track surface conditions (BVSYS, 2007). Two cameras provide images for each track with up to 0.4mm x 0.5mm pixel precision, enabling CrackCheck to detect concrete cracks and run-offs early on. And eventually effort is made to build a new system for the purpose of automating visual inspection of wooden railway sleepers through machine vision approach.

**1.2 Proposed Solution**

There is always a need for a system to automatically interpret data to reduce the intervention of a human operator, to gain efficient insight knowledge about the material being inspected and classified. For the purpose of acquiring complete knowledge about the wooden railway...
sleepers, a machine vision approach has been designed and developed in order to automate the process visual inspection in condition monitoring. This is to identify the flaws associated with the sleepers such as cracks on the sleepers, and condition of the rail fastenings in the scene etc. This approach involves with various phases like image acquisition for the purpose of capturing real images, and applying image processing and analysis techniques through the machine vision algorithm to analyse images for its features such as the number of cracks, length of the crack, width of the crack and length of the plate. These features are then considered for the classification using pattern recognition techniques. This machine vision approach has been proposed and found to be an appropriate solution in the process of automating the visual inspection process of wooden railway sleepers.

Machine vision has applications in many aspects of transport like traffic monitoring, aerial navigation, traffic safety, etc. Since it can be used to collect and understand data useful in transport planning and safety. It can also play an active role in navigation and vehicle guidance due to its robustness and reliability. And it provides methods to enhance picture quality for interpretation of events and monitor complex scenes. Example applications include Number-Plate Identification, Tracking People, Face Recognition and Intruder Monitoring. Machine Vision provides the ability to track objects leading to safer and efficient non-obtrusive monitoring (BMVA, 2007). This machine vision algorithm is experimentally evaluated using 400 real images on either side edges of the sleepers making a total of 800 images; 400 images on left side and 400 images on right side including the images of both wooden sleepers and the metal plate of the wooden railway sleepers in order to analyse their features and determine the condition for its future use.

Image acquisition is the first stage of any vision system and is carried out acquiring the real images. After image acquisition the next stage is the image analysis, and the machine vision algorithm uses various methods of image processing techniques to the images captured and extracts its features such as the number of cracks, length of the crack, width of the crack, and length of the metal plate. These extracted features are then stored in feature vectors and were further classified. In this machine learning process pattern recognition technique is used for the classification task.

Because, pattern recognition is the research area that studies the operation and design of systems that recognize patterns in data which encloses sub-disciplines like discriminant analysis, feature extraction, error estimation, cluster analysis (together sometimes called statistical pattern recognition), grammatical inference and parsing (sometimes called syntactical pattern recognition). Whereas its important application areas being image analysis, character recognition, speech analysis, man and machine diagnostics, person identification and industrial inspection. Classification of the features extracted is carried out by using classifiers like Support vector machines (SVM) and Radial Basis function (RBF), though there are several other classifiers available. The pattern classifiers, SVM and RBF have been tested and compared with each other for the reliability of the results obtained by both the classifiers.
1.3 Goals and Scope

Aim of this machine vision approach is to automate the visual inspection process of wooden railway sleepers in order to detect the flaws on them. Because automated visual inspection systems are essential to improve the speed and accuracy of the condition monitoring process and also reduces the human intervention. This approach not only does the operation of condition monitoring in a faster way but also helpful in working cost-effective manner thereby providing a more intelligent, secure, robust and economical feasible solution. This approach analyses the surface condition of the sleepers based on the image features extracted from the images acquired; features such as the number of cracks, length of the crack, width of the crack and length of the metal plate used for the fastening the rails is sunken or not have been considered to be the most important factors in determining the condition of the wooden railway sleepers. Currently, this machine vision approach is capable of analysing these features efficiently.

This machine vision approach could be more interesting if the possibility to find the condition of the fastenings used. That is, not just to find the presence of the nails and bolts used in fastening but also to identify the strongness in holding the metal plate to the rails laying on the wooden railway sleepers. However, the condition of the fastenings could be easily identified by using sound analysis techniques like impact acoustics (Siril, 2006), where the sleeper is hit by using an axe to emit sounds which are then recorded and analysed based on the rattling sound produced from the sleepers. Fusion method, combination of both image analysis technique and sound recognition technique using impact acoustics together classify the image data features and sound data features. This kind of fusion is expected to give more reliable and robust results of machine vision in condition monitoring process.
Chapter 2

2 Machine Vision in Transportation

Today, Machine vision has evolved to become a vital and effective automation tool that enables computers to replace human vision in many high-speed and high-precision manufacturing applications. Machine vision is used to ensure the quality of everything from tiny computer chips to massive space vehicles. However, choosing the right image acquisition technology in this machine vision process is quite difficult and relative work is being carried by the automated imaging association, a global machine vision industry’s trade group (AIA, 2007). Machine vision has application in many aspects of transport. For example, it can be used to collect and understand data useful in transport planning and safety. It can also play an active role in navigation and vehicle guidance (BMVA, 2007). And traffic monitoring, aerial navigation, transport safety, etc are said to be some of the areas where machine vision application is used in its transportation. Monitoring traffic (BMVA, 2007), is a vital part of transport planning. Human observers are currently used, but this approach is time consuming and the data captured is also limited. An automated system to study traffic movements is being developed which aims to identify vehicles entering an area under observation, and monitor their movements. Using images from a single camera, vehicles can, thereby, be located precisely and tracked as it moves in the scene. To obtain information on the actual position of the vehicles, the system is given a model of the scene and the exact position of the camera. Using this geometrical knowledge, a full understanding of vehicle movements can be obtained for subsequent analysis. And currently, the navigation of airborne vehicles relies on the use of active remote sensing equipment. These techniques are very expensive and an automated system to interpret aerial images, using passive image sensing, combined with machine vision techniques, is being developed. The system operates by identifying road networks, and matching them with stored maps of the roads in the area. The roads are identified by processing the image with computational filters that pick out road-like features. And further processing modifies the roads into a series of straight line segments. The characteristics and relationships between these segments can be compared with the stored map which has been similarly processed. This system can work over a range of altitudes and has potential use in the autonomous guidance of aircraft (BMVA, 2007). Video monitoring by human operators is applied widely to transport safety. This activity is labour intensive, and operators are prone to fatigue.

Machine vision, however, is now capable of providing valuable assistance by drawing attention to predefined events of interest. Systems using neural nets possess the flexibility to ignore acceptable events, but to generate an alarm in unacceptable situations. To do this, the neural net must be trained to accept the normal range of lighting and visibility likely in an outdoor situation. Events such as intruders, or obstructions in the scene, can then be identified by the system. To make the system more robust, training continues whilst it is in place. This allows it to adapt to slowly changing situations, while retaining its sensitivity to unacceptable occurrences (BMVA, 2007).

In the transportation sector, the Intelligent Transportation System (ITS) is a novel paradigm to improve conventional road-vehicle systems by using microelectronics, artificial intelligence,
and other related technologies. ITS, is based on three key functions: communication, control, and sensing. Machine-vision systems play important roles by detecting lane markings, vehicles, pedestrians, road signs, traffic conditions, traffic incidents, and even driver drowsiness. And an approach towards real-time vision-based road markings detection; working on flat roads with painted road markings has been demonstrated to be robust with respect to extremely critical shadow conditions and global illumination changes. This approach allows detecting the road markings through an extremely simple and fast morphological processing, to overcome completely the annoying problems caused by a non uniform illumination (shadows), and to implement efficiently the detection step on massively parallel SIMD architectures in order to obtain real-time performances. It has been implemented on the special-purpose and low-cost massively parallel system PAPRICA and integrated onto the MOB-LAB land vehicle, reaching a processing rate of about 17 Hz (A. Broggi, 1995). And also it has been reviewed that cameras are clearly having a much greater role in transport; driverless cars at the Darpa Grand Challenge 2005, has been guided solely by machine vision and travelled a 212.4km course across rough terrain. And BMW has designed a system that guides cars in a straight line on the motorway, and another system wakes up the driver when it sees his or her eyelids are drooping. The images are used in inventory management systems of roads, cataloguing every sign maintaining accuracy which is very important in these applications. And the cameras are supposed to have a large optic dynamic range, to cope with the significant variance in intensity caused by bright reflections. Clearly effective illumination is an important feature of these systems that can be helpful to capture and analyse the images to complete the process of feature extraction such as licence plate recognition and traffic analysis (Imaging & Machine Vision Europe, 2007).

Within transportation domain several methods have been developed for solving different condition monitoring problems and machine vision approach is being used extensively. As all the transportation areas demand reliable, accurate and fast solutions in order to avoid catastrophic results. Machine vision approach is applied in various fields of transport; such as road, aerial, water and rail transport areas. However, the current work is involved in the condition monitoring process of wooden railway sleepers (Siril Y, 2007).

2.1 Machine Vision in Rail transportation

Machine vision is being used in rail transportation domain. And also a serious attention is paid towards condition monitoring of wooden railway sleepers. Because many of the rail incidents have been occurring due to the poor condition of the railway sleepers and also the rail fastenings used on it. And over years the visual inspection as a condition monitoring process of wooden railway sleepers is being carried by human operators and which are prone to fatigue resulting in the serious destruction of property and infrastructure through derailment, etc (Siril Y, 2007). This human intervention of visual inspection is very slow, time consuming, laborious and hence it is essential to automate the process of visual inspection for condition monitoring; to identify the flaws associated with the sleepers such as cracks on the sleepers, and condition of the rail fastenings in the scene etc within the proper time avoiding such hazardous situations in the rail industry.

In order to decrease human intervention and to increase the speed, quality in the performance of condition monitoring process a company named MER MEC in Italy, has been developing...
and deploying the equipment like vehicles, sensors, softwares etc for the purpose of railway maintenance. MER MEC Track Surface Defect Detection, is a vision system designed and manufactured for real-time monitoring of track condition and automatic recognition of resulting defects. This monitoring system uses high-speed line-scan cameras for track images acquisition. Enhanced vision algorithms carry out measurements and classify defects according to their properties and/or their position onto the the rail structure. A special illuminating system suitably designed by Mer Mec allows the system to operate properly in every light-conditions. System basic modular architecture consists 3 subsystems: Track Surface Inspection, Joint Gap & Head Checks Inspection, Track Surface Measurement. System acquires, stores, analyzes images up to 160 km/h. It is provided with a software application for data analysis that allows fast processing and real-time visualization of measurements/acquisitions. MER MEC offers a wide product range for enhanced data processing by providing software applications for intelligent optimization of maintenance (further more information could be found on Mer Mec webpage) (MER MEC, 2007).

A vision-based system (Gadelmawla, 2004) has been developed for surface roughness characterisation. Software named GLCMSurf has been developed to analyse the surface roughness using a grey level co-occurrence matrix. Schutter has proposed an automated image analysis system for monitoring cracks on concrete objects. But this image analysis system is not specific about the concrete objects used in condition monitoring process (Siril Yella, 2006).

Another company named BV SYS in Germany, bvSys Bildverarbeitungssysteme GmbH (BV SYS, 2007), has developed a system for Crack Check that inspects the concrete sleepers and ballastless track surface conditions. Two cameras provide images for each track with up to 0.4mm x 0.5mm pixel precision, enabling Crack Check to detect concrete cracks and run-offs early on.

2.2 Potential Difficulties

Visual inspection, an NDT method used extensively to evaluate the condition or the quality of a component or object (Krstulovic et al., 1996). It is easy to perform and this method requires good vision, good lighting and operating knowledge. However, in this process of condition monitoring two major difficulties have been identified; (a) image acquisition and (b) the machine vision algorithm development. Image acquisition is the process involved to collect the image data for visually analysing the images for flaws with the help of image processing tools. Machine vision algorithm is the algorithm developed using these image processing tools to analyze and identify the flaws from the image data collected through the image acquisition procedure in this condition monitoring process. For the image acquisition procedure the camera is the key element and its positioning and lighting setup used for capturing the real images for processing are considered as the real problems. Since the position and aperture of the camera focusing the object should be capable of capturing complete information necessary for further processing. And also the machine vision algorithm development is another challenging task since it should be efficient enough to analyse even the finer details available in the images acquired. Thus these two factors are the real problems and the solution towards automating the visual inspection in condition monitoring of the
wooden railway sleepers in order to identify the flaws associated with the sleepers, such as cracks on the sleepers, and condition of the rail fastenings in the scene etc.

Machine vision algorithm pre-processes the images acquired using image processing tools and extracts the features such as the number of cracks, length of the crack, width of the crack, and length of the metal plate. However, image acquisition is considered to be the initial and very essential phase that needs great attention to capture the images from which all the necessary information can be attained through the feature extraction procedure and then those extracted features are classified using the pattern classification technique in the machine vision process.

2.2.1 Image Acquisition

Image acquisition, a process of gathering image data, is essential for automating the visual inspection routine and is quite complex as well (Siril Y, 2007). This is the first step towards designing an image analysis system i.e., machine vision system. Acquiring real-time or still images could be quite a challenge. A two-dimensional image that is recorded by the sensors like optics is the mapping of the three-dimensional visual world (Acharya and Ray, 2005). Sometimes noisy images or degraded image is recorded because of the optical lens system in a digital camera used to acquire the visual information. And if the camera is not appropriately focused then the image acquisition process results in blurred images and the reason is the defocused camera. And there are several factors making the image acquisition process a real challenging task in achieving the desired output. For example, factors like equipping the vehicle running along the track with a camera that is calibrated to acquire the required images can itself be a very difficult task. Operational and structural constraints may pose definite threats, i.e., the system or vehicle being considered for the process should be properly built so as to keep it within the load gauge limit, and focus of the camera should be able to provide complete details about the object from its image, and also lighting conditions, etc in order to proceed further with the necessary experiments.

Moreover, such experiments might also be challenged by various factors such as different light and weather conditions etc. The advances in technology has made the present day cameras a more capable, compact and amazingly powerful than they were a few years ago. Image acquisition is becoming simple and able to maintain quality. A point to be noted is that the difficulty in image acquisition is strictly dependent on the kind of examination that is being focused. For example, images concerning railway track profile, rail surface, railway sleepers, examining aircraft composites etc, demand different camera positions in each case. So, it should be realised that determining any standard approach of acquiring images is highly impossible.

No surprise, camera being the key element or component of the machine-vision system performing the operations of image acquisition. The camera sets the level of detail, or resolution, that the system can distinguish. It also sets an upper limit on the system's frame rate, or the speed at which the system can generate images, and the shutter speed or image-capture time. And digital cameras do not need any external digitizers and have the additional advantage of being free from the resolution and frame rate of video standadards. The explosion of digital cameras in consumer applications, from photography to cell phones, has stimulated
significant advances in the technology during the last decade. Digital cameras are now available in a range of resolutions with an equally broad selection of achievable frame rates, at ever-decreasing cost. The machine-vision systems use digital cameras as their "eyes", because camera is the first major component of any image processing system that captures the image of a three-dimensional object (Acharya and Ray, 2005).

There are a variety of trade-offs to consider in selecting a digital camera. One of the most significant is resolution versus frame rate. In general, the higher the camera's resolution, the lower the frame rate it can achieve. The core of a digital camera is a CCD (charge-coupled device) sensor that captures the images. And basic digital camera has only one charge sensor, so reading an image out of the camera occurs one pixel at a time. This one-at-a-time action ties the camera's frame rate to its resolution. This CCD-process technology is having an upper limit to the speed at which the transport array can move the pixel charges. So, the larger the array, the longer it takes to read out an entire frame. But this trade-off is not absolute. Camera CCDs with multiple charge sensors are available. They break the image into non-overlapping blocks that you can read out in parallel, increasing the achievable frame rate. However, balancing the multiple conversions from charge to digital value so that the independent blocks will produce matching images complicates the design effort. And it is also possible to increase the frame rate of a high-resolution camera by using only a subset of the total image. Cameras that offer this feature allow the control system to specify an image area of interest for readout rather than shift out the entire array. This approach results in a smaller image at a faster frame rate (Richard, 2007), which is important to consider during the camera selection process for the image acquisition.

In this current approach, a Nikon D70 camera is mounted on the vehicle which can roll over the rails at a fixed length and proper angle is maintained through the entire process of image acquisition which appeared to be the best possible angle that exploits complete details of the wooden railway sleepers visually from the images. A digital camera can capture images in various resolutions. For example, 320 x 240, or 352 x 288, or 640 x 480 pixels on the low to medium resolution range to 1216 x 912 or 1600 x 1200 pixels on the high resolution size. And this Nikon D70 (Phil Askey, 2007) has image resolutions of pixels 3008 x 2000 [Large] (6.01 million), 2240 x 1488 [Medium], 1504 x 1000 [Small] that captures large, medium and small sized images respectively. And at each pixel one can normally have 16 million colours. But, colour is not an important factor in the present work as the images gets converted into gray level images during image analysis. For maintaining the quality of the images during image acquisition the aperture, focal length, flash etc needs proper adjustments according to the circumstances of the setup. The SB-800, announced on 22 July 2003, flash light is used in this approach and it is Nikon's best flash (Ken Rockwell, 2007). Since it works with every Nikon camera made for the past 40 years and works properly with the Nikon D70 and the new i-TTL exposure system. The SB-800 adds a reverse flash guide number mode for flash where you set the aperture and the flash sets the level automatically based on distance which reflects the flashmatic modes of the 1950s, backwards. The SB-800 has a rated GN of 125 at ISO 100. At full manual power it has a recycle time of 6 seconds (alkaline) or 4 seconds (Ni-MH) and provides 130 shots (alkaline) or 150 shots (Ni-MH).

The camera, Nikon D70, is fixed to the vehicle running over the rails and the image of each sleeper is captured in turn with the images being stored temporarily within the camera itself. Currently in this work, a vehicle to run over rails was built with supporting arms which were
used to fix the camera. And the camera fixed is properly positioned and flash light, focal length of the aperture is maintained constant focusing towards the surface of the sleepers. And then the setup is run over the rails and image of each and every sleeper is captured with great attention. Camera is positioned in two different views i.e., top and horizontal view of the sleepers. This is in order to capture the wooden railway sleepers from the rail to its outer edge since most of the necessary information which really affects the condition of the sleeper is available in this region of the sleeper. Using this image acquisition procedure a total of 800 real images of the wooden railway sleepers of either side edges have been acquired for its process of image analysis in order to extract the required features available from those images. After image acquisition, those images have been transferred to the computer and image analysis is carried out applying the algorithm developed for analysing and calculating those image features.

During image acquisition process the images of the wooden railway sleepers were acquired by considering only the image part from the rail to its outer edges on either side of the sleepers (Siril Y, 2007). Since, this part of the wooden sleeper is mostly concentrated during the visual inspection process by the human inspectors, and this region of the sleeper is giving more details about its condition in most of the cases. Totally, the number of images collected were 800; with 400 images on the left side where 200 images are showing the wooden part of the sleeper (see Figure 2.1) and 200 images are showing the metal plate and nail(s) used for fastening the rails (see Figure 2.2), and 400 images on the right side with respective to the left side i.e., same sleeper with images on its right side; out of which 200 images are showing the wooden part of the sleeper (see Figure 2.3) and 200 images are showing the metal plate and nail(s) used for fastening the rails (see Figure 2.4). Since this is the part of the sleeper which could mainly effect the rail transportation and also helpful in exhibiting the real condition of the wooden railway sleepers in decision making for its future use. Even many human inspectors have been found knocking the edges with an axe during the visual inspection. And in this process two different image views have been considered, one from top view to capture the images of wooden part of the railway sleepers (see Figures 2.1 and 2.3) and one horizontal view capturing the images of metal plates and nail(s) or bolt(s) used for fastening the rails (see Figures 2.2 and 2.4). After image acquisition, the image features are extracted by developing an efficient algorithm. The figures below depict the camera view used in capturing the wooden railway sleepers in both good and bad conditions respectively.

![Wooden Railway Sleeper](image_url)
2.2.2 Machine Vision Algorithm

Machine vision (Bruce, 2007), a subset of systems engineering, is concerned with the engineering of integrated mechanical-optical-electronic-software systems for examining natural objects and materials, human artifacts and manufacturing processes, in order to detect defects and improve quality, operating efficiency and the safety of both products and processes. It is also used to control machines used in manufacturing. Machine Vision necessarily involves the harmonious integration of elements of the following areas of study; mechanical handling, lighting, optics (conventional, fibre optics, lasers, diffractive optics), sensors (video cameras, UV, IR and X-ray sensors, laser scanners), electronics (digital,
analogue and video), signal processing, image processing, digital systems architecture, software, industrial engineering, human-computer interfacing, control systems, manufacturing, existing work practices and QA methods.

Machine vision is the application of computer vision to industry and manufacturing. However the computer vision is mainly focused on machine-based image processing, machine vision most often requires also digital input/output devices and computer networks to control other manufacturing equipment such as robotic arms. Machine vision is also known as a subfield of engineering that encompasses computer science, optics, mechanical engineering, and industrial automation. One of the most common applications of machine vision is the inspection of manufactured goods such as semiconductor chips, automobiles, food and pharmaceuticals. Just as human inspectors visually inspect the objects or its parts to judge the quality of workmanship, so machine vision systems use digital cameras, smart cameras and image processing software to perform similar inspections. Machine vision systems are programmed to perform narrowly defined tasks such as counting objects on a conveyor, reading serial numbers, and searching for surface defects on the objects which is our main focus in the current work. However the machine vision systems for such visual inspections requires high-speed, high-magnification, 24-hour operation, and repeatability of measurements. Frequently these tasks extend roles traditionally occupied by human beings whose degree of failure is classically high through distraction, illness and circumstance. However, humans may display finer perception over the short period and greater flexibility in classification and adaptation to new defects and quality assurance policies.

The goal of a machine vision is to create a model of the real world from the images i.e., machine vision system retrieves useful information about a scene from the images with its two-dimensional projections of the three-dimensional real world. And the information or its features are extracted from these images with having the knowledge about the objects in the images and also about the projection geometry involved in the image acquisition process. However, many fields are related to machine vision because techniques developed from many areas are used for extracting information from the images. And image processing is a well-developed field. The image processing techniques usually transform images into other images and the task of extracting information will be according to the desired objective of the work. And techniques from pattern recognition play an important role in the machine vision for recognizing and classifying the object patterns (Machine vision, 1995).

The design of robust machine vision algorithm is one of the most difficult parts of developing and integrating automated systems. An efficient algorithm development technique that utilizes modes of the human visual system is equally important task in comparison to the process of image acquisition in this work. Once the image acquisition phase is completed, the digital images having acquired from both the views; top and horizontal views, then the image data is very much in need to be processed for image analysis using the image processing methods in order to extract the necessary image features. Hence, development or design of robust algorithm is considered as a challenging task. However, a machine vision algorithm is developed with the desired objective of analysing a large set of real images, 800 real images currently used in this work, in order to produce more accurate and appropriate results as a part of condition monitoring process.
The machine vision algorithms majorly have certain qualities like: 1. Machine vision is an engineering discipline, not a mathematical or philosophical exercise, 2. There is no unique way to perform any given image processing function; every algorithm can be implemented in many different ways, and 3. A given calculation may be very difficult using one implementation technology but very easy with another.

Image acquisition process is used in acquiring the image (raw) data through a digital camera and then the acquired images are stored in a computer to be used for further processing. And then image processing techniques using the algorithm is applied on the acquired raw image data for the process of feature extraction which will be further used in decision making during the condition monitoring process. The algorithm developed is described briefly in the next section (see Section 3) of feature extraction, i.e., the implementation of various image processing tools used in pre-processing the image data to analyse and extract the required image features. Hence, all the necessary image features have been extracted and stored in the system separately as feature vectors for its pattern recognition task of machine vision process which is helpful in making decisions related to the future use of the wooden railway sleepers in the condition monitoring process.

This algorithm developed is found to be working in an efficient way relatively even on a large set of data extracting the required image features and then classifying those patterns through pattern recognition in its machine vision process of condition monitoring of the wooden railway sleepers. In fact, the automation of visual inspection of wooden railway sleepers is the desired objective and also a part of condition monitoring techniques used in the NDT domain now-a-days. And here the speed of the machine is also given great importance to automate the whole process and it has complications because of two components (Bruce, 2007): 1. Throughput rate (Number of images processed or objects inspected per second), and 2. Latency (Time delay from digitization of an image and appearance of result). However, an on-line inspection system, working in conjunction with a continuously-moving vehicle system, can usually tolerate a high latency.

In this current machine vision approach the system involves with various stages like image acquisition and processing those images with the help of algorithm developed and then classifying those image patterns to make decisions regarding the condition of the wooden railway sleepers. Here in this process the algorithm uses various image processing tools like Image resize, gray-level image conversion to avoid unnecessary details like rust particles, oil-stains, etc., and Gaussian filters to deblurr the images to have only information of the major cracks effecting the wooden sleeper condition, entropy filtering, and canny edge-detector for the plate analysis and also image adjustment is used in plate analysis, and some morphological operations like thinning, skeletons and pruning along with gray-scale morphological operation like ‘imclose’ has been used to analyse the images. These tools have been used in a proper order gradually extracting the image shapes required for feature extraction. And then the algorithm is followed by the pixel operations to calculate the actual length and width of each crack. Because from the gray-level images the pixels either have black or white pixel values with either 0 or 1 respectively where the pixel values having 1 as its value represent the image feature. And a group of pixels connected make a component or a feature of the image being analysed and whose values are calculated. During these calculations, when a similar kind of feature is being calculated more than once actually which are more in number, that is if length of the crack is being calculated and if those cracks are
happened to be more than one then the crack with maximum value is given importance to be used as the important calculated feature for further processing of classification task. Thus when such kind of multiple number of values result from the calculation then the crack having maximum length is reported and stored in a feature vector and similarly width with maximum width is extracted and stored separately in a feature vector. The length of the metal plate is also calculated with the similar procedure and only this feature is extracted from the images captured for metal plate analysis. Since the nails and other information available can be calculated but is of no use as they appear to be present but the fastening cannot be measured.
Chapter 3

3 Feature Extraction

In general the feature extraction involves simplifying the amount of resources required to describe a large set of data accurately that is when performing analysis of complex data one of the major problems arises from the number of variables involved. And analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample during the process of pattern recognition and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

Feature extraction is carried by the algorithm developed in order to extract the required essential image features so as to analyse the condition of the wooden railway sleepers, which is being inspected in the current work and is considered to be an important step towards the automation process of the visual inspection of wooden railway sleepers involved in condition monitoring. During this feature extraction the extracted features are stored as separate feature vectors called number of cracks, length of the crack, width of the crack and length of the (metal) plate and these feature vectors are classified into two different classes called ‘good’ and ‘bad’ i.e., 1, 0 respectively. This classification of feature vectors into good and bad classes is very useful during the further process of pattern classification (Siril Y, 2007).

Image acquisition, the first step towards the feature extraction process, is carried with all the necessary measures to capture the images provided with having a good quality and clarity for the ease of image analysis (see Section 2.2.1). The most important part of the wooden railway sleeper images is from the rail towards its outer edges where all the necessary information required for the condition monitoring process is obtained. And hence, the operation of feature extraction using the machine vision algorithm is experimented in this region with the images acquired in this process. Major cracks which are responsible in degrading the condition of the wooden railway sleepers for its future use are found in this part of the sleepers in vertical direction relative to the images captured. However, the images from the top view (see Figure 2.1 and 2.3) and the horizontal (or side) view (see Figure 2.1 and 2.3) relative to the rail. Both the views have been considered for the crack detection and metal plate, nail(s) or bolt(s) detection respectively. But only the metal plate is considered during image processing to calculate its length because the condition of the fastening is analysed through the sound analysis techniques (Siril Y, 2007).

Wooden railway sleeper image with the top view, i.e., image part from the rail towards its outer edges avoiding the part of the metal plate, is considered for the crack analysis in extracting its crack features since these cracks on the wooden sleepers depending on length and width of its crack shows the condition for its future use. At same time these cracks not only effects the sleepers but also they are responsible in worsening the condition of the fastenings used to hold the rail with the sleepers, and through the horizontal (or side) view of the camera the image part having the metal plate is captured and considered for its feature extraction which is used in analysing the length of the metal plate. This approach with two
different views of image acquisition has greatly reduced the computation burden in having a large set of data for processing. Since, only a part of the sleeper is considered with having all the necessary features required for the condition monitoring process of the wooden railway sleepers the computation time decreases and the processing speed increases which results in development of a faster and effective algorithm (Siril Y, 2007).

Initially, this approach has got problems and importantly related to visual information of the images captured and what would be the image size, etc. This is because the images captured through the digital cameras are of larger size and the computation time required for relatively a large set of images using the algorithm in MATLAB environment increases and the performance of the algorithm decreases. However, the digital images having represented with its pixel values provides valuable information for image processing in the process of feature extraction. However, this particular work aims to extract necessary features such as the number of cracks, length of the crack, crack width and metal plate length using the image processing methods is implemented in this algorithm. These four features have been considered as key elements in determining the condition of the wooden railway sleepers.

In this algorithm initially the acquired images were resized to an image size of 200*300 pixels, since it is very important because the MATLAB environment cannot support processing of the images with relatively large size as reading and loading the image files with larger size requires high amount of memory and also increases the computation time in executing the operations to do so. And the figures following shows the results after each operation involved in process of feature extraction and the figures of both good and bad sleepers are shown to make a clear distinction between them assuring proper functioning of the algorithm in condition monitoring process. Actually, in the case of a good sleeper when the cracks are not present or relatively small (very thin hair line cracks) then the system ignores such cracks and report as if those thin hair like cracks do not exist on the sleeper (See Figures 3.12, 3.14, and 3.16). The images are to be processed further only to remove the unnecessary stones, etc which are not truly required in further processing of feature extraction and pattern recognition. And below observed the wooden railway sleepers with both the good and bad conditions after the image resize operation using the algorithm.

![Figure 3.1: Bad Sleeper Image after image resizing.](image-url)
The image resize is followed by the gray-scale image conversion operation. This conversion of coloured images to gray-level images makes the further processing task very easy, since the pixel values are toned to gray-levels between black and white pixels. This function is a basic image processing tool used for the intensity transformations of grayscale images. As the colour information available in these images is not of much use. The reason is that these images might be having any kind of unnecessary details like rust particles, diesel or oil stains, etc on the wooden sleepers and possibly could cause problem during the process of feature extraction. And importantly, the features that are being extracted mainly depend on the number of pixels available and connected together making a component in the image extracted during processing. And below follows the results of gray-level images obtained after the gray-scale image operation.
Next, Gaussian filter is used as a part of smoothing operation for the process of deblurring the images and reducing the noise. Normally smoothing filters are used for blurring and noise reduction. And blurring is used in pre-processing tasks, such as removal of small details from an image prior to (large) object extraction and bridging of small gaps in lines or curves (Gonzalez and Woods, 1993). Gaussian filter with a standard deviation of 1.4 and threshold of 7 is used in this particular image operation. Since this operation has made it very easy to identify the cracks which were happening to be the real cause for the condition of the wooden sleepers.

Also, followed the structuring elements operation connecting its neighbouring elements to retain the shape and then entropy filtering operation is followed for the removal of unnecessary stones from the image. These basic operations in the pre-processing stage increased the level of details in the image to a greater extent respectively.

Later, the image was converted into binary image (pixels values are either black or white) using a gray threshold of 0.65. A binary image is a digital image that has only two possible values for each pixel i.e., black and white. Binary images are also called bi-level or two-level. Binary images often arise in digital image processing as masks or as the result of certain operations such as segmentation, and thresholding, etc. The interpretation of the pixel's binary value is very useful in pixel operations having pixel value of 0 as black and 1 as white. A
binary image is usually stored in memory as a packed array of pixels. Binary images can be interpreted as subsets of the two-dimensional integer lattice $\mathbb{Z}^2$; the field of morphological image processing is inspired and dependent on its pixel values for further processing in the current work.

Figure 3.7: Bad Sleeper Image after binary operation.

Figure 3.8: Good Sleeper Image after Binary operation.

And, the image is still resized now to some more extent (number of pixels) required for the major removal of unnecessary regions from the images. Later a series of morphological operators were applied. Morphological image processing techniques are useful for extracting image components that are useful in representing and describing region shapes (Gonzalez and Woods, 1993). As first stage of morphological processing, erosion operation is used and further the image was resized by using image cropping and then majority operation was carried out in the second stage of morphological processing. Also followed the structuring elements operation connecting its neighbouring elements to retain the shape determined. These basic operations in the pre-processing stage increased the level of details in the image to a greater extent respectively.
The erosion operation of morphological processing used in this process changes any black pixel adjacent to a white pixel to white. And also undergoes the operation to remove all connected components (objects) having value of pixels less than or equal to 50 from the binary image. This erosion operation of morphological processing removes all small scale bright features and darkens the image improving the contrast between the large scale features (Gonzalez and Woods, 1993). And the majority operation of morphological processing is to set the pixel values to 1 if the pixels following the neighbouring pixels are 1’s (Haralick, 1992; Pratt 1991); otherwise, sets the pixel value to 0. Now structuring element operation is applied to retain the shape. However, even after carrying out these operations, some unfilled spaces on the crack were identified. These spaces were formed as a result of shadows or due to the presence of stones in between the cracks. As a result, one single crack was separated into two if a stone is lying on the crack. Due to this, features such as number of the cracks, crack length etc, were reported incorrectly. Hence, it was desired to handle this problem by joining pixels which are relatively close. After certain experimentation, a distance of 15 pixels or less was determined as a standard for joining pixels. If, the distance between the two crack is less than or equal to 15 pixels, the two cracks are joined by a straight line. Such an approach was considered reasonable so as not to loose the integrity of the crack because of the presence of the stones etc.
Labeling operation is used for the identification and having the count for the image shapes. And then the skeleton operation of morphological processing is used from where the skeleton of the image shape is obtained. Because the object recognition is achieved using shape skeletons and has made the calculation of the image features much easy since the length calculation is carried by the pixel operation. Skeleton of the crack image is like a line and is exactly one pixel wide (Gonzalez, 1998; Ramesh, 1995). Skeleton can be described as a line representation of an object, one-pixel thick, through the "middle" of the object preserving the topology of the object (I.T. Young, 2007). In the case of good sleeper the skeleton of the resultant image is shown below.

And more importantly the pruning of images was carried to have clear images of cracks. So far the image of a bad sleeper is experimented using the machine vision algorithm necessarily for the crack detection and the figures illustrated below are the results obtained. These detected cracks are further processed with the same algorithm for its calculations. And if the detected number of cracks detected is more in number (greater than 1) then the crack with maximum length is considered to be the output, which is actually useful to determine the condition of the sleeper.
3.1 Number of cracks

Number of cracks, the first feature, is an important feature depending on which the length of each crack is calculated. First when the wooden railway sleeper images are considered for feature extraction it is found that the number of cracks appearing on each sleeper is giving clear information related to the condition of the wooden railway sleeper. From whose visual information further analysis of the condition of the sleeper like length of each crack and also its width is carried further. In order to calculate the number of cracks the images have already been operated using various image processing methods like image resizing, gray-level conversion, smoothing operation is applied using Gaussian filters to the images and then those images are converted into binary images from where the number of cracks is easily calculated. And the number of the cracks in an image is given by labels, that is, labels having the pixels connected to each other with a similar value in the binary image. From the labelling it is very clear that the pixels labelled ‘1’ makes up one object (or crack), and the pixels labelled ‘2’ make up second object, and so on (Matlab, 2007). And these labelling numbers represents the number of cracks available and identified in the images during condition monitoring process using the machine vision algorithm.

3.2 Length of the crack

Now, after calculating the number of cracks in the image the length of all those cracks detected is calculated, and is considered to be the second important feature in this approach. And the crack with maximum length is reported for further processing through pattern recognition. Crack length is nothing but the area of white pixels in the binary image. And length of the crack has been calculated rather easily using the skeleton image of the crack. Skeleton can be described as a line representation of an object which is of one-pixel thick,
through the "middle" of the object preserving the topology of the object. Since, skeleton of the crack image is like a line which is of exactly one pixel wide (Gonzalez, 1998; Ramesh, 1995). Two end points of the crack from the skeleton image whose length is to be determined is first fixed. The number of white pixels encountered in the path is kept track and is reported as the crack length. In case of multiple cracks, the crack with the largest area is only reported though the length of each crack is calculated. Justification for using such an approach is to be able to report the uniform number of features to the pattern classifier for all images. Moreover, the crack with the largest area is often the worst case and hence attention has been focussed.

3.3 Width of the crack

Width of the crack, the third important feature in the current approach and is calculated through simple means. In this particular work, width of the crack is identified and calculated by dividing the image area of the largest crack following the operation of finding the length of the cracks. And this image area is nothing but the white pixels on a black background since the images have already been converted into gray level images where the images have only white and black pixel values. However, width of the cracks can be easily observed after applying the ‘imclose’ operation to the images as part of feature extraction using the image processing method. Figures 3.11 and 3.12, shows the images after the imclose operation from where calculation of the width of the crack is found beneficial in order to have the exact and real information regarding the width of the crack. The image obtained after imclose operation is used to calculate the width that is calculating the number of white pixels representing the width and in vertical direction with respect to the crack since possibly most of the cracks are in horizontal direction and calculated accordingly which have been considered to be effecting the condition of the sleepers. Similarly, along with the length of the cracks, the cracks having much greater width is considered to be big problem which affects the condition of the sleeper so badly. And width of the crack is considered and given equal importance as the length of the crack in condition monitoring process of the wooden railway sleepers. Hence, the crack with much greater width is reported for further processing of pattern recognition.

3.4 Length of the Metal plate

Metal plate is the fourth important feature in this work and it is used in rail fastenings. How far the metal plate is sunken determines the condition of the wooden railway sleepers. And its calculation determines the presence or absence of the metal plate but the condition of its fastening is not determined, like how strong or how good is the condition of the fastening. This fastening condition is considered as a limitation, using this method. Length of the metal plate is calculated with the similar kind of approach applied for calculating the length of the crack in section 3.2. But with necessary minor modifications like Gaussian filter’s having threshold value as 7 and its standard deviation value as 2, using canny edge-detection operation, etc.

Edge detection is the approach used most frequently for segmenting images based on abrupt changes in intensity. And canny edge detector is used in this process as a part of image segmentation technique. And generally canny’s approach is based on three basic objectives: 1.
low error rate, 2. edge points should be well localized, and 3. Single edge point response. And the canny edge detection algorithm consists of the following basic steps:

1. Smooth the image with a Gaussian filter.
2. Compute the gradient magnitude and angle images.
3. Apply nonmaxima suppression to the gradient magnitude image.
4. Use double thresholding and connectivity analysis to detect and link edges.

Because of these reasons and experiments conducted on the images canny edge detection was used to obtain the edges with 1 pixel thick which follows step 4 in its edge detection algorithm while smoothing is accomplished by convolving the input with a Gaussian mask (Gonzalez and woods, 1993).

As explained in section 3.2, similar to the length of the crack, the length of the metal plate is also nothing but the area of white pixels in the binary image. And the length of the metal plate is also calculated using its skeleton of the image. Between the two points the length of the metal plate can be easily calculated by constantly checking for the next white pixel in the neighbourhood with similar pixel value in a single clockwise direction until the end of the skeleton is encountered. The number of white pixels encountered in the path is kept track and is reported as the length of the plate. In case of multiple skeletons, the skeleton with the maximum area is reported as the length of the metal plate and is considered for further processing of pattern recognition. Moreover, attention towards the length of the metal plate is focussed to determine whether the metal plate is sunken or not in order to determine the condition of the wooden railway sleeper.

In this approach objects like nails, bolts etc used for fastening also appears in the images but ignored in this image analysis approach, because we can only find its presence but not the actual condition of the fastening (as described in the previous section). And a careful observation is made to cut-down the images to extract only the required part of the image where the plate edges can be detected for its length calculation. And figures illustrated below are showing the difference between the good and bad sleepers during their length calculation process.

![Figure 3.17: Bad Sleeper Image to calculate plate length.](image-url)
Figure 3.18: Good Sleeper Image to calculate plate length

Figure 3.19: Bad Sleeper Image after Gray-level operation

Figure 3.20: Good Sleeper Image after Gray-level operation

Figure 3.21: Bad Sleeper Image after Gaussian filter operation
Figure 3.22: Good Sleeper Image after Gaussian filter operation

Figure 3.23: Bad Sleeper Image after Canny edge-detection operation

Figure 3.24: Good Sleeper Image after canny edge-detection operation
Figure 3.25: Bad Sleeper Image after morphological Skeleton operation

Figure 3.26: Good Sleeper Image after morphological skeleton operation

Figure 3.27: Bad Sleeper Image after Pruning operation
But in these images probably we may find two objects related to nail and plate. Since it is difficult to identify the exact condition of the fastening, the image part with metal plate is segmented and has been considered to identify the condition whether the plate is sunken or not. If it is sunken then the length of the metal plate is possibly less than the actual length but will be very less and also completely disappears like the one in the bad case of sleeper image. And the length of the metal plate of a good sleeper is possibly more or approximately of actual size, but actual size of the metal plate is not determined and not compulsion that it may vary from sleeper to sleeper.

After extracting all the possible features; number of cracks, crack length, crack width, and plate length, through the image processing using the machine vision algorithm then those features are stored in separate feature vectors for its classification using pattern recognition (Siril Y, 2007).
Chapter 4

4 Pattern Recognition

Pattern recognition, is comprised of many different methods supporting the development of numerous applications in different areas. Pattern recognition (B.D. Ripley, 1996), has a long and respectable history within engineering but the cost of the system hardware both to acquire the data (images and signals) and its computation capabilities, which has increased its importance. And the system hardware advances has made it widely applicable and becoming feasible to design automated systems to replace the human experts, and either performing better or cloning the expert. However, the primary assessment of a system will be by its performance.

Pattern recognition (Christopher, 2007) is the research area that studies the operation and design of systems that recognize patterns in data. It encloses sub-disciplines like discriminant analysis, feature extraction, error estimation, cluster analysis (together sometimes called statistical pattern recognition), grammatical inference and parsing (sometimes called syntactical pattern recognition). However, the important application areas being image analysis, character recognition, speech analysis, man and machine diagnostics, person identification and industrial inspection.

Pattern recognition, is a sub-topic of machine learning, can be defined as "the act of taking in raw data and taking an action based on the category of the data". Most research in pattern recognition is about methods for supervised learning and unsupervised learning. Pattern recognition aims to classify data (patterns) based on either a priori knowledge or on statistical information extracted from the patterns. The patterns to be classified are usually groups of measurements or observations, defining points in an appropriate multidimensional space. A complete pattern recognition system consists of a sensor that gathers the observations to be classified or described; a feature extraction mechanism that computes numeric or symbolic information from the observations; and a classification or description scheme that does the actual job of classifying or describing observations, relying on the extracted features.

The classification or description scheme is usually based on the availability of a set of patterns that have already been classified or described. This set of patterns is termed the training set and the resulting learning strategy is characterized as supervised learning. Learning can also be unsupervised, in the sense that the system is not given an a priori labeling of patterns, instead it establishes the classes itself based on the statistical regularities of the patterns. The classification or description scheme usually uses one of the following approaches: statistical (or decision theoretic), syntactic (or structural). Statistical pattern recognition is based on statistical characterizations of patterns, assuming that the patterns are generated by a probabilistic system. Syntactical (or structural) pattern recognition is based on the structural interrelationships of features.

Structural pattern recognition is often used in combination with statistical classification and/or neural network methods in hybrid approaches in order to achieve complex pattern recognition tasks. For example, in applications like two-dimensional or three-dimensional object
recognition. A wide range of algorithms from very simple Bayesian classifiers to much more powerful neural networks have been applied for pattern recognition. Typical applications of pattern recognition are automatic speech recognition, classification of text into several categories (e.g. spam/non-spam email messages), the automatic recognition of handwritten postal codes on postal envelopes, or the automatic recognition of images of human faces. The last two examples form the subtopic image analysis of pattern recognition that deals with digital images as input to pattern recognition systems. However, in the current application machine vision algorithm is used on two-dimensional objects or images.

From classification trees to neural networks, there are many possible choices for the classifier to be used. Multi-Linear Perceptron (MLP) is a neural network classifier and most commonly used in pattern recognition. However, the pattern recognition in this work is carried using classifiers like support vector machines (SVM) and radial basis function (RBF). SVM differs radically from MLP like SVM is a rule based classifier, and the SVM training always finds a global minimum. The main difference between MLP and SVM is the principle of risk minimization. In case of SVM, structural risk minimization principle is applied by minimizing an upper bound on the expected risk whereas in MLP, traditional empirical risk minimization is used minimizing the error on the training data. The difference in risk minimization is to improve the generalization performance of SVM compared to MLP (Samanta et al., 2003).

SVM is a rule based classifier which follows the principle of structural risk minimization. And RBF is a neural network classifier and neural network classifiers use the output values of internal calculation nodes to estimate the posterior class probabilities of input patterns (Linda & Richard, 1999). SVM also provides an effective mechanism for incorporating the results of pair wise classification as a means of reducing training times when aiming to achieve an all-together result (Simon and Arnaud, 2005).

The SVM is a novel type of learning machine, based on statistical learning theory, which contains polynomial classifiers, neural networks, and radial basis function (RBF) networks as special cases. In the RBF case, the support vector (SV) algorithm automatically determines centers, weights, and threshold that minimize an upper bound on the expected test error. The SV algorithm provides a principled way of choosing the number and the locations of RBF centers. While from the real-world pattern recognition problem in contrast to a corresponding number of centers chosen by K-means, the centers chosen by the SV algorithm allowed a training error of zero, even if the weights were trained by classical RBF methods. The interpretation of this finding is that the SV centers are specifically chosen for the classification task at hand, whereas -means does not care about picking those centers that will make a problem separable (Bernhard, 1997).

For this classification task to classify good and bad class of railway sleepers from the images in the current work the image data is initially pre-processed by the machine vision algorithm to extract the features required in the condition monitoring process and these features are taken as inputs to the classifiers being used in the current classification task of pattern recognition.
4.1 Support Vector Machines

Support Vector Machines (SVM) (J.P. Marques, 2001), is a distinctive approach to pattern classification, with the principle of structural risk minimization. And support vector machines can provide a good generalization performance independent of the distributions of the patterns. SVM has high generalization performance without the need of adding a priori knowledge, even when the dimension of the input space is very high (Yanning Zhang, 2001). SVM is a new approach of pattern recognition that has arisen with statistical learning theory, introduced by Vapnik (Vapnik, 1998), combining statistical mathematics and learning theory to make a learning algorithm which is practically reliable and theoretically robust. The SVMs are robust even in the presence of noise (Mazzeo, 2003). The central idea of SVM is the adjustment of a discriminating function using the information of the boundary patterns which is considered to be optimally separable (J.P. Marques, 2001). The SVM approach, in its simplest linear version, consists of determining the hyperplane that maximizes the margin of separation, called optimal hyperplane.

SVM performs classification by constructing an $N$-dimensional hyperplane that optimally separates the data into two categories. SVM models are closely related to neural networks and SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron neural network. SVM models are closely related to classical multilayer perceptron neural networks. Using a kernel function, SVM’s are an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training.

The feature selection is an important factor for the pattern classification. And a predictor variable is called an attribute, and a transformed attribute that is used to define the hyperplane is called a feature. Hence, the task of choosing the most suitable representation is known as feature selection. A set of features that describes one case (i.e., a row of predictor values) is called a vector. So the goal of SVM modelling is to find the optimal hyperplane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyperplane are the support vectors. And for example, figure 4.1 shows linearly separated hyperplanes (H1 and H2) for two classes which are separable and the support vectors are circled for clear distinction between those vectors (J.C. Burges, 1998).

![Figure 4.1: Linear separating hyperplanes for the separable classes.](image)
Figure 4.2 shows linearly separated hyperplanes for two classes which are non-separable and the support vectors are circled for clear distinction between those vectors (J.C. Burges, 1998).

The SVM is a rule based-classifier, and rule based classifiers partition the input space into decision regions using threshold logic nodes or rules (Linda & Richard, 1999), and is derived from statistical learning theory first presented in (Boser et al., 1992). The main advantages of SVM when used for image classification problems are: (1) ability to work with high-dimensional data and (2) high generalization performance without the need to add a-priori knowledge, even when the dimension of the input space is very high. Excellent topics relative to the SVM can be found elsewhere (Cristianini and Shawe-Taylor, 2007; Vapnik, 1995). The SVMs is a new generation learning system based on recent advances in statistical learning theory and are considered to deliver state-of-the-art performance in real-world applications such as text categorisation, hand-written character recognition, image classification, bio-sequences analysis, etc., and are now established as one of the standard tools for machine learning and data mining (Cristianini and Shawe-Taylor, 2007). However, SVM with a Gaussian kernel is used in this approach classifying the features extracted from the images of the wooden railway sleepers.

Ideally, SVM analysis should produce a hyperplane that completely separates the feature vectors into two non-overlapping groups. However, perfect separation may not be possible, or it may result in a model with so many feature vector dimensions that the model does not generalize well to other data; known as overfitting.

The idea of using a hyperplane to separate the feature vectors into two groups works good when there are only two target categories, but SVM also works well when more than two target variable having more than two categories is to be classified. Several approaches have been suggested, but two of them are most popular: (1) “one against many” where each category is split out and all of the other categories are merged; and, (2) “one against one” where $k(k-1)/2$ models are constructed where $k$ is the number of categories. However, in this approach there are only two target categories either good or bad depending on which the decision is considered relative to the condition monitoring process.

The accuracy of an SVM model is largely dependent on the selection of the kernel parameters such as C, Gamma, P, etc. There are two methods for finding optimal parameter values, a grid
search and a pattern search. A pattern search generally requires far fewer evaluations of the model than a grid search. A grid search tries values of each parameter across the specified search range using geometric steps. A pattern search (also known as a “compass search” or a “line search”) starts at the center of the search range and makes trial steps in each direction for each parameter. If the fit of the model improves, the search center moves to the new point and the process is repeated. If no improvement is found, the step size is reduced and the search is tried again. The pattern search terminates when the search step size is reduced to a specified tolerance. The weakness of a pattern search is that it may find a local rather than global optimal point for the parameters. To avoid overfitting, cross-validation is used to evaluate the fitting provided by each parameter value set tried during the grid or pattern search process.

Grid searches are computationally expensive because the model must be evaluated at many points within the grid for each parameter. For example, if a grid search is used with 10 search intervals and an RBF kernel function is used with two parameters (C and Gamma), then the model must be evaluated at 10\*10 = 100 grid points. If cross-validation is used for each model evaluation, the number of actual SVM calculations would be further multiplied by the number of cross-validation folds (typically 4 to 10). For large models, this approach may be computationally infeasible.

Many kernel mapping functions can be used – probably an infinite number. But a few kernel functions have been found to work well in for a wide variety of applications. Kernel methods are helpful in providing a powerful and unified framework for pattern discovery, motivating algorithms that can act on general types of data (e.g. strings, vectors or text) and look for general types of relations (e.g. rankings, classifications, regressions, clusters). The application areas range from neural networks and pattern recognition to machine learning and data mining (John, 2007). The concept of a kernel mapping function is very powerful and allows SVM models to perform separations even with very complex boundaries.

4.2 Radial Basis Function

The radial basis function (RBF) network is a feed-forward neural network with a single layer of hidden units. Radial basis function networks are different from back propagation networks because they have only one layer of hidden units, and do not use the sigmoid activation function in the hidden layer unit. Instead, the radial basis function network has fixed-feature detectors in the hidden layer which use a specified basis function to detect and respond to localized portions of the input vector space. One advantage of radial basis networks over back propagation is that, if the input signal is non-stationary, the localized nature of the hidden layer response makes the networks less susceptible to "memory loss," or, as some would say, "weight loss". The radial basis function as implemented has a single hidden layer of units which can use one of three basis functions: Gaussian, thin plate spline, or multi-quadratic. The basis centers or weight vectors are learned during an initial stage of self-organized learning, currently fixed at 15 epochs. Once the basis vectors are set, the output layer weights are adjusted using back propagation (RBF, 2007). And much more information regarding the functionality of RBF network could be found elsewhere (Niranjan and Ranjan, 2005).
There are several parameters to set during the initial and final training phases. To train the basis weights, you must set the autocenter parameter to TRUE. When this parameter is set, only the hidden layer weights (the basis vectors) are adjusted. Training parameters which are used during this phase include the learn rate and tolerance for errors. After the basis weights have been set or trained, a second training phase is required. This is quite similar to training a back propagation network with a single hidden layer. The learning-rate and momentum settings are used; modifying them can have a large effect on the training performance of the network. These values are commonly set from 0.5 to 0.7 for learning rate and 0.0 to 0.9 for momentum. The error tolerance setting controls the final training process. If the data set contains binary targets (0,1), then the tolerance parameter is usually set to 0.1. This means that the output is considered "good" when it is within 0.1 of the desired output (that is, 0.9 for a 1, 0.1 for a 0). When every output is within the tolerance range of the desired output value, the network status is changed to LOCKED and weight updates are stopped. You can also set the epoch update flag. If set to TRUE, the weights are changed only after every complete cycle through the training set. This is true gradient descent. When the epoch update flag is set to FALSE, weights are updated after each pattern is presented. For most problems, the network converges more quickly with the epoch update flag set to FALSE.

An additional factor in training a radial basis network is the order in which training patterns are presented. By setting the randomize flag on the Import object feeding the Network object, you can ensure that the network is presented with a random ordering of the training patterns. This is often useful in avoiding local minima and aids in training speed.

The run-time performance of a radial basis network is relatively fast. The input vector is propagated through the network by first multiplying it by basis vectors and then passing it through the basis function. This vector is then multiplied by the output weight matrix. Next it is passed through the selected basis function to produce the network outputs array which is returned to the application program.
Chapter 5

5 Results and Discussion

Earlier when the experiments were conducted by considering one feature at a time and combining with number of cracks then the SVM classifier reported good classification rates whereas combining with crack length and crack width separately delivered poor rates, thereby indicating that number of cracks and metal plate length as important features in classification. While observing various possible combinations, the best classification rates were achieved by combining all the four features of both left side and right side i.e., four features from each side making altogether 8 features. Thus the need for feature fusion is strongly indicated. And it is clear enough that the number of cracks, crack length, crack width and plate length altogether serve as the most important features for classification in this particular work. As crack width determining the true condition of the crack and from whose skeleton image the true length of the crack is calculated. Also, length of the metal plate is an important feature since it determines the condition of the fastenings of rail to the sleeper. Because the rail accidents recorded till date found that these fastenings condition as one of the reasons in many incidents. Hence, it is suggested that a combination of all these features from both the sides to be used for classification.

In this classification task, a total of 200 patterns were used and segregated into training and test files having 150 and 50 patterns respectively with having 75% of good class patterns and 25% of bad class patterns in each set of file used in classification. The total number of input features used were 8, the number of output classes were 2, and the input feature labels i.e., class labels were 0,1 where normally 0 indicates bad class and 1 indicates good class. But has difference in the resulted classification if the experiment action with the test file selected as train or test. And major difference occurs if the false negative is high that is bad class being classified as good class. From the above figure, Table 5.1, it is clear that SVM has better results compared to RBF. This section explores the results obtained during SVM classification to have an idea about the description of the pattern classification task, and there are several classification methods that could be adopted by the users according to their application keeping in view of advantages and disadvantages of the methods.

In SVM classifier, the parameters used were like; Lagrange Multiplier Upper Bound = 15.000000, while the kernel function used is Gaussian Kernel with its Gaussian Standard Deviation = 2.500000. And the tables below illustrate the results achieved by SVM classifier:

**Training set classification:**

Error rate = 4.67%

Confusion matrix:

<table>
<thead>
<tr>
<th>Desired Class</th>
<th>0</th>
<th>1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>108</td>
<td>0</td>
<td>108</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>115</td>
<td>35</td>
<td>150</td>
</tr>
</tbody>
</table>

Table 5.1: Confusion matrix.
### Table 5.2: Training set Error rate using SVM.

<table>
<thead>
<tr>
<th>Class</th>
<th>Patterns</th>
<th># Errors</th>
<th>% Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>108</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>42</td>
<td>7</td>
<td>16.67</td>
</tr>
<tr>
<td>Overall</td>
<td>150</td>
<td>7</td>
<td>4.67</td>
</tr>
</tbody>
</table>

### Table 5.3: Confusion matrix.

<table>
<thead>
<tr>
<th>Desired Class</th>
<th>0</th>
<th>1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>41</td>
<td>50</td>
</tr>
</tbody>
</table>

### Table 5.4: Test set Error rate using SVM.

<table>
<thead>
<tr>
<th>Class</th>
<th>Patterns</th>
<th># Errors</th>
<th>% Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>14</td>
<td>6</td>
<td>42.86</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>1</td>
<td>2.78</td>
</tr>
<tr>
<td>Overall</td>
<td>50</td>
<td>7</td>
<td>14.00</td>
</tr>
</tbody>
</table>

From the above figures; Table 5.1, Table 5.2, Table 5.3, and Table 5.4 the number of patterns that have been classified correctly and the percentage of the overall patterns classified is clear to make decisions regarding the condition of the wooden railway sleepers. The classification confusion matrix is very much useful to know exactly the acceptable patterns classified. However, the major problem is the condition of false positive of the patterns classified which increases the risk factor in condition monitoring of the wooden railway sleepers.

When the similar data sets; training set and test set, are used with the RBF classifier then there observed relatively no big difference compared to the SVM classifier. But, difference in the patterns classified and a small change in false positive values of the patterns classified is identified.

In RBF classifier, the clustering algorithm used is K-means with its clustering parameters considered were like; Cluster by class - K equal for all classes, Number of Centers (K) =16, and Number of Centers to remove at a time for non-binary k= 2 while the Training Variance Multiplier (h)= 6. And below illustrated the results achieved by RBF classifier:

### Training set classification:
- Error rate: 5.33%
- Confusion Matrix:
Table 5.5: Confusion matrix.

<table>
<thead>
<tr>
<th>Desired Class</th>
<th>0</th>
<th>1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>34</td>
<td>8</td>
<td>42</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>108</td>
<td>108</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>34</strong></td>
<td><strong>116</strong></td>
<td><strong>150</strong></td>
</tr>
</tbody>
</table>

Table 5.6: Training set Error rate using RBF.

<table>
<thead>
<tr>
<th>Class</th>
<th>Patterns</th>
<th># Errors</th>
<th>% Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>42</td>
<td>8</td>
<td>19.05</td>
</tr>
<tr>
<td>1</td>
<td>108</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>150</strong></td>
<td><strong>8</strong></td>
<td><strong>5.33</strong></td>
</tr>
</tbody>
</table>

Testing set Classification:
Error rate: 14.00%
Confusion Matrix:

<table>
<thead>
<tr>
<th>Desired Class</th>
<th>0</th>
<th>1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>7</strong></td>
<td><strong>43</strong></td>
<td><strong>50</strong></td>
</tr>
</tbody>
</table>

Table 5.7: Confusion matrix.

<table>
<thead>
<tr>
<th>Class</th>
<th>Patterns</th>
<th># Errors</th>
<th>% Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>14</td>
<td>7</td>
<td>50.00</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>50</strong></td>
<td><strong>7</strong></td>
<td><strong>14.00</strong></td>
</tr>
</tbody>
</table>

Table 5.8: Test set Error rate using RBF.

From the tables illustrated above in this section; Table 5.5, Table 5.6, Table 5.7, and Table 5.8 it is clear that RBF has classified correctly and results were almost matching with the SVM classifier. Thus, both the classifiers SVM and RBF have performed well and relatively the both results happened to be almost similar to each other in both the training and test sets of data. But, a clear difference could be identified in the misclassified patterns among the two training and test sets of patterns used in the classification task. However, the major problem arises when the misclassified patterns increases the risk factors, with the condition of false positive, leading to the damage of the property like derailment, etc. This indicates for the high attention to be paid to avoid such hazardous conditions by improving the efficiency in image acquisition method and updating the algorithm used for feature extraction leading to the correct classification. Image acquisition approach needs high attention in this process since further processing is completely dependent on the image data acquired through this process in having proper images allowing complete information available through its visual appearance. And the algorithm developed should be capable of analysing even the finer details available in the image and extracting the required features used to determine the condition of the wooden...
railway sleepers which is the main aim of this work. This improvement in image acquisition approach and the algorithm extracting the features will definitely yield better classification increasing the robustness of the machine vision used to automate the visual inspection of wooden railway sleepers and thus making this condition monitoring process very easy, fast and reliable task.

The overall results (see Table 5.9) recorded by the classifiers SVM and RBF with a combination of all the four features; number of cracks, length of the crack, width of the crack and length of the (metal) plate. These four features were extracted from the images through the feature extraction process and combined especially for the classification task making the machine vision process simple and easy task. And the results obtained through classification task using classifiers SVM and RBF have been tabulated for its simplicity and clarity. Initially, each feature was tested individually against the classifier and then the results obtained were used to have a single result by taking average, etc. But the overall result obtained following this procedure was not so realistic. Actually, this approach following the combination of features finding the best results could be applicable and possibly identifying the most important feature of all those features explaining the details of the condition of the sleepers.

<table>
<thead>
<tr>
<th>Classification Method</th>
<th>Training set classification</th>
<th>Test set classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector Machine</td>
<td>95.33%</td>
<td>86.00%</td>
</tr>
<tr>
<td>Radial Basis Function</td>
<td>94.67%</td>
<td>86.00%</td>
</tr>
</tbody>
</table>

Table 5.9: Pattern classification using two different methods: SVM and RBF.
Chapter 6

6 General Conclusions and Future Work

In this work, a machine vision system automating the visual inspection process of wooden railway sleepers is being developed. This system detects the condition of the wooden railway sleepers through the images acquired and processed using the algorithm that has been developed for this purpose ensuring the security of the rail property and safety to the users. Well in advance it is considered that the periodical inspection of railway infrastructure is very important to prevent dangerous situations like accidents causing mass destruction of the railway property and injury or even deaths of the crew or passengers, etc.

For over several years there has been condition monitoring process detecting the flaws associated with the rail infrastructure, but it was heavily involved with the human intervention which is prone to fatigue due to human errors. To avoid these human errors and also to improve the speed of the process with robustness and reliability ensuring safety and security within time a reliable machine is required. And has lead to the development of a machine vision system.

And in this work, wooden railway sleepers have been considered for the condition monitoring since it is observed that most of the rail incidents are due to bad condition of the wooden sleepers which have been used since many years and also due to the bad condition of rail fastenings used. The machine vision system being developed is an NDT approach causing no damage to the rail infrastructure, saving time, increasing processing speed, reliable and robust with its nature.

This machine vision system involves with various processing stages like image acquisition, algorithm development extracting the image features, and pattern classification. And image acquisition is the first and initial stage to capture the real images with two different angles or views; one image having the wooden sleeper part for identifying the cracks and another image having the metal plate part for identifying the rail fastenings. From the rail fastening image part it is easy to find the availability of the fastenings but it is very hard to find the condition of the fastenings used. And in this work image part of the rail fastening is used to find whether the metal plate used for fastening is sunken into the wooden sleeper or not which determines its condition as good when it is not sunken else viceversa.

And the machine vision algorithm is used with various image processing tools in analysing and extracting the image features. This algorithm is well designed with an intention to extract only those features which have been considered to be responsible for degrading the condition of the wooden railway sleepers. And the features extracted were number of cracks, length of the crack, width of the crack, and length of the metal plate. And more particularly the lengths or widths with maximum values are extracted since it is believed that they are the real cause and visually appeal its condition.

And then those extracted features were stored in feature vectors for their classification using pattern recognition classification task. Classifiers like SVM and RBF have been reported to
show the robustness and reliability of the system. However, SVM and RBF classifiers reported almost similar results with a slight difference in its classification of training set data. And this classified information could be found unreliable if there happened to be increasing in its false positive values of the patterns classified. And the system performance is also dependent on its inspection velocity along with reliability ensuring safety and security of the condition monitoring in a non-destructive fashion in this current approach.

Since, this approach was not capable enough to detect the condition of the rail fastenings it is always reasonable to solve such problems with different technique in order to identify its condition. And finally, using fusion methodology with all the necessary features extracted will give proper verdict regarding the condition monitoring process and obviously there will every possibility of developing a robust system. And this current machine vision approach could be further extended to fusion analysis by using sound analysis techniques that could easily detect the condition of the fastenings, that is, with the rattling sounds produced from the wooden railway sleepers when they are struck hard with some material or object. Thus sound analysis will detect the condition of the rail fastenings and helps in the automation of machine vision process of condition monitoring.

Moreover, resistivity analysis yet another technique could also be used in this condition monitoring process, determining the condition of the wooden railway sleepers. Because of the availability of moisture content in the wood which is responsible for the behaviour of wood. And different kinds of wood have been used to have wooden sleepers behaving different in accordance to its environmental conditions. And the wooden railway sleeper on being used for a long time always has the tendency to degrade physically and loose its moisture content, etc. However this approach cannot be considered as a visual inspection process but still has the advantage of using it as a condition monitoring process. And will be a different method in the area of NDT (non-destructive testing). In order to accomplish this task, the current system cannot be used and a different experiment setup is required with a different algorithm to extract features from the raw information.
References


[29] Linda Kukolich, Richard Lippmann, LNKnet user’s guide, MIT Lincoln Laboratory, August 31, 1999.