A New Inspection Method Based on RGB-D Profiling

Affan Ahmed Siddiqui

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Mechanical Engineering

Tomonari Furukawa, Chair
Mehdi Ahmadian, Co-Chair
Alfred Wicks
Shashank Priya

September 18, 2015
Blacksburg, Virginia

Keywords: Rail inspection method, RGB-D profiling , Map-based identification

A New Inspection Method Based on RGB-D Profiling

Affan Ahmed Siddiqui

ABSTRACT

This thesis presents an inspection method based on RGB-D profiling for the rail industry. The proposed approach uses inexpensive RGB-D cameras to generate color and geometrical information of the observations, and stitches each consecutive scan from the sensor to form a map, provided that the two scans contain the information from the same observation. Using a technique known as pairwise registration, the errors between these consecutive scans are minimized using error minimization algorithms such as Iterative Closest Point and Normal Distributions Transform. Once the error between each consecutive scan is minimized, the scans are then converted into a global co-ordinate frame work to form a global map of all the added scans. The proposed approach could be used as a map-based identification technique by comparing the past global map to newly acquired scans while also reducing computation time effectively. The effectiveness of this approach is demonstrated by developing a system that uses multiple RGB-D cameras to detect railway defects such as spikes. The applicability of the proposed approach to other applications is then evaluated by profiling long lengths of road.

Dedication

I would like to dedicate this thesis to my parents, Nadeem Parvaiz Siddiqui and Fauzia Nadeem Siddiqui for their loving support. I would not be in this position if it had not been for their moral support and guidance. This is for you Papa and Mamma.

Acknowledgments

This project has been funded by the Railway Technology Laboratories (RTL) at Virginia Polytechnic Institute and State University.

First of all, I would like to thank my advisor Dr. Tomonari Furukawa for his continued guidance towards my degree. I was able to identify my faults and able to rectify them under his direction. Dr. Furukawa's constant feedback on my performance has helped me build a character that is able perform well in both academic and professional life. I offer my deepest gratitude to him for giving me the opportunity to work under him to purse my interest in robotics. I would also like to thank my co-advisor Dr. Mehdi Ahmadian for the series of technical discussions with him regarding this project. I also appreciate my committee members, Dr. Alfred Wicks and Dr. Shashank Priya for their inputs on my thesis.

Secondly, I would like to give my praise to the members of my research laboratory: Computational Multi-physics Systems (CMS) Lab. Through their help and input, I was able to conduct the experiments required for my project. A hearty thanks to my colleagues: Kuya Takami, Boren Li, Howard Brand, Christopher Mobley, Luan Doan, Rich Fedora, Hangxin Liu, Yazhe Hu, Tian Yi, Wil Blanton, Orson Lin, Jason Doyle, Murat Ambarkutuk, Yoonchang Sung and CMS alumni: Dr. Xianqiao Tong and Sriramvarun Nidamarthy.

Lastly, I would like to thank my parents and my older siblings Adeel Siddiqui and Jibran Siddiqui for their continued moral support and guidance on life decisions. It was through them that I was able to adjust myself through any hardships I faced during my academic life at Virginia Tech.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

From the early years of the railroad industry, rail inspection has been a keen topic of interest with regards to preventing defects. Rail defects emerged as a significant concern when a derailment in Manchester, New York in 1911 resulted in the death of 29 people and caused injuries to 60 others [1]. Early investigation into this accident found that the cause of the accident were caused by defects which is not possible to detect by human inspection. In addition, external defects which are in millimeter scale cannot be quantitatively analyzed by human inspection. Furthermore, cracks can propagate as fatigue failure [2, 3, 4]. Defects in the railroad industry branched into track defects as well, when there were cases when tracks would delineate from their positions under sporadic weather conditions because of expansion and contraction effects [5, 6]. Fastening components such as tie plates and wooden ties would get worn out due to long periods of use. Finally, overgrown vegetation near the tracks caused obstacle issues for the track. As such, non-destructive testing methods have been developed to detect such defects preventing train derailment while also providing ride comfort.

The need for non-destructive methods arose when visual inspection was not sufficient for preventing future failures. Inspection methods such as Laser Line scanning [7, 8, 9, 10, 11] or Computer Vision [12, 13, 14, 15, 16, 17, 18] help us quantize the geometrical entities of the subject or provide texture information for defect identification. The inclusion of computer algorithms and systems in recent years have made this inspection method quite versatile to many different applications. For example, non-contact sensors could be used for road fault detection. For these reasons and applications, development of non-contact measuring sensors are integral to the structural health community. However, each sensor is developed for one particular function:measuring geometry or providing texture information. There has not been an inexpensive testing method developed for a sensor that can provide both geometry and texture information. For this reason, an inspection method that uses a sensor which uses RGB-D(color and depth) information, must be investigated. Furthermore, the road industry is in constant need for inspection systems as they have to maintain consumer acceptability. This thesis presents an inexpensive inspection method for rail profiling based on RGB-D reconstruction. Although there are many inspection methods which allow you to measure geometry or deduce color information, there is not a single inexpensive inspection method which can provide both functions at the same time.

1.2 Objectives

- Construct a map building method which can be used to measure large and small geometrical defects associatively.
- Develop the system using multiple sensors and implement on tracks to identify raised spikes using a map-based identification approach.
- Expand map-building method to other fields.

1.3 Defect Identification using RGB-D Profiling

This thesis presents an approach for a railroad and road inspection method using RGB-D cameras. The inexpensive RGB-D cameras are able to produce 3-D geometrical profiles while also preserving color information. Unlike a line scanner which uses a single pulse, the RGB-D camera is able to see the same object in two consecutive scans due to its large field of view. The proposed approach produces a 3-D reconstructed map in a global co-ordinate framework. The global map from the proposed approach can be used as a reference for new scans in a map-based identification approach, thus reducing computation time. Multiple RGB-D cameras are used in this approach to reconstruct the rail and road profiles. The cameras are mounted inside an enclosure as these cameras are sensitive to natural light. Uniform LED lighting has been planted on the rim of the enclosure to provide a controlled environment for the matching process.

1.4 Summary of Original Contributions

The original contributions of this thesis are listed as follows.

- Design and development of a scan-matching method and system for rail and road profiling.
- Multiple integration and synchronization of RGB-D cameras.
- Map-based identification of track defects.

1.5 Outline

This thesis is organized as follows.

- Chapter 2 describes the past inspection methods used in the rail and road industry.
- Chapter 3 gives a brief overview of the principles of the the RGB-D sensor, 3-D reconstruction and mapping.
- Chapter 4 presents the map-based identification approach along with the system design and development for rail profiling.
- Chapter 5 presents the experimental results from the rail profiling system.
- Chapter 6 presents the system design and development for road profiling and its results.
- Chapter 7 summarizes conclusions and proposes future work.

1.6 Summary

This chapter provided a background to the structural health monitoring situation that has been presented in the rail and road industry. A need for non-destructive inspection methods was then established. A new inspection method which included detection of geometry and texture information was then introduced along with the objectives. With these objectives in mind, the original contributions of this project were then summarized. Finally, an outline listing the remaining chapters of this thesis was presented.

Chapter 2

Literature Review

This chapter provides an overview of the past non-destructive testing (NDT) methods used in the railway and road industry. This chapter provides a brief overview of the current inspection methods such as Ultrasound, Eddy Current, Magnetic Inspection Method as well as emerging visual inspection methods such as Computer Vision and Laser Line Scanning. With the principles of each technique explained, the advantages and disadvantages are then investigated.

2.1 Railroad defects

Rail defects can be classified into two areas: rail defects and track defects [5, 6, 19]. Rail defects are caused by improper handling of the rail [3]. A consequence of improper handling can be wheel-burn defects. Cracks on the railway head can also be propagated by the continuous loading and unloading load cycles also known as Rolling Contact Fatigue(RCF) [2, 4]. RCF alone contributed to 60% of rail defects in the 1990s [3]. Rail defects are mostly caused due to manufacturing defects or RCF.

Track defects on the other hand are formed over time unlike manufacturing defects. During hot or cold climates, the tracks will expand or contract respectively. Tracks can delineate from their mean positions in a longitudinal or transverse fashion in such weather conditions [5, 6]. In addition this, grown vegetation on the tracks can cause problems as it requires expensive equipment and numerous amount of man power to remove weeds [20]. Lastly tracks can have component failures such as raised spikes which can derail the tracks as mentioned in [14]. An example of a track defect such as buckling can be seen in Figure 2.1.

Figure 2.1: Buckling of railway tracks [19]

Keeping these defects in mind, non-destructive methods must be implemented beforehand so that failures can be taken care of appropriately. The following NDT methods can be categorized into current and emerging methods. The focus of this thesis is directed towards emerging technologies, however current technologies will also be discussed briefly to demonstrate the historical significance of rail inspection.

2.2 Current Inspection methods

During the early 1900s when train derailments were a detrimental issue, Elmer A. Sperry developed a method known as the "induction method" [21, 22]. In the induction method, a large amount of low current is passed onto the rail head to create a magnetic field according to Faraday's Law. When there was a defect present on the rail, the material would not be able to hold the magnetic flux and as such magnetic flux leakage would occur. When there was a change in magnetic flux, a current was induced and could be recorded on a strip chart. The induction method had a serious limitation as it could only detect transverse fissures in the rail [3]. This limitation led to the development of the ultrasound method which could detect internal flaws in the rail.

2.2.1 Ultrasound

By the 1960s, the induction method soon evolved into the ultrasound technology to introduce flaw detection in other areas such as bolt-holes [23]. The ultrasound technique consists of a pulser which emits ultrasound waves and a receiver which detects changes in the reflected signal in terms of the signal strength. This technique is able to detect subsurface cracks. The Electromagnet acoustic transducer(EMAT) pulser also known as the ultrasound receiver/emitter can be oriented normal to the railhead (pulse-echo) to detect spalling or an angle to the rail (pitch-echo) to detect transverse fissures [24]. The problem with EMATs is the fact that their efficiency decreases drastically with increasing distance between rail and sensor. In addition to this, EMATs are considered better receivers than emitters [23].

By the late 1970s, laser ultrasonics was proposed as a NDT approach [25]. This approach utilizes a high frequency laser to generate elastic waves in a solid medium [26]. When the laser is focused onto the rail head, heating occurs. Some of the energy from the heating

is converted to an elastic stress wave which is then picked up by the receiver. Air and water have been used as a medium to generate ultrasound frequencies. However this technique introduced issues as the air-coupled ultrasound generators were less sensitive than piezoelectric sensors. As such, it was challenging to distinguish the defect from the noise [27]. An improvement to this technique is the laser-air hybrid ultrasonic(LAHUT). Kenderian et al. [28] used a remote pulse laser to generate acoustic waves across the rail. The reflected acoustic wave was then sensed by the air-coupled ultrasonic detector. LAHUT achieved effective results given the fact that it could detect Vertical Split Heads and Transverse Detail Defects. The distinction this method brought was that it could detect rail base cracks while other conventional methods could not.

The ultrasound technique can inspect internal rail defects at high speeds with high accuracy. However, the weight of these sensors poses a problem as it might induce vibrations on the mounted system. In addition to this, external flaws cannot be detected using this method.

2.2.2 Eddy Current

A method which often complements ultrasound is the eddy current testing method [29, 30, 31, 32]. An alternating electric current is passed through a coil to create a magnetic field. When this coil is brought close to the rail, the varying magnetic field induces current flow loops in the material known as Eddy currents. Eddy currents can produce its own magnetic field to characterize defects. If there is a defect, the electrical impedance will change [32]. In the early 1990s, Garnham and Beynon [29] developed a machine that could detect conditions of creepage found in rail-wheel contact. Eddy currents can be used to detect fatigue cracks early in their life. Pohl et al. [30] proposed the approach of detecting head check defects

appearing at the gage corner of the rail. The technique was performed at speeds of 100 km/h and provided promising results by detecting surface defects such as wheel-burns and short-pitch corrugations. Song et al. [31] focused crack identification on railhead using eddy current testing. Defects were identified on an artificial and a real railhead sample by analyzing the amplitude and phase changes of the eddy current.

There are a few limitations with eddy current testing. The largest limitation would be the fact that only conductive materials can be inspected, therefore components such as wooden ties or ballast cannot be inspected. Furthermore, ferromagnetic materials need special treatment for magnetic permeability before eddy currents can be produced.

2.2.3 Magnetic Particle Inspection

Another approach to detect top-level rail head cracks and subsurface cracks is the magnetic particle inspection method [33]. Firstly, a magnetic field is generated across a ferromagnetic surface. In theory, magnetic field lines will flow from the North pole to the South pole. If a crack is present, magnetic flux leakage occurs and flows outside of the material. Magnetic particles are laid over the material parallel to the magnetic field. If a defect is present, the magnetic particles will accumulate around that point to indicate that defect. Heyder and Girsch [34] investigated the application of Heat treated rails (HSH) using magnetic particle inspection. Li et al. [35] performed an experiment and simulation study on using 3-D magnetic fields. The simulation study showed that 3-D magnetic field measurements could give comprehensive positional information about defects in terms of shape and orientation.

Similar to eddy current testing, a limitation of this technique is that only ferromagnetic materials can be inspected. In addition, precise alignment is needed when establishing the magnetic field to detect cracks. Lastly, the material will have to be demagnetized after

inspection.

2.3 Emerging Inspection methods

Taking the current inspection methods into mind, one must consider the fact that these methods could only give us a quantitative measure of the flaws inside the rail, and describe the flaws using sensor characteristics such as magnetic flux leakage, ultrasound waves or current. However, there comes a point where we can use the rail's profile or characteristics for visual inspection methods. Such emerging methods can help us give a visual representation of the fault. The scope of this thesis is focused on these emerging visual inspection methods.

2.3.1 Computer Vision

A rather modern class of testing approaches is using automated vision techniques to identify and classify flaws such as raised spikes, worn tie-plate and missing bolts [12, 13, 14, 15, 16, 17, 18]. Mair and Fararooy [12] describe the need for computer vision technologies as a track inspection method. The paper describes that most of the objects on the rail track are made of simple geometric shapes such as hexagonal nut bolts and spikes, hence algorithms could be developed in computer vision to detect such shapes. For instance, Sawadisavi [36] proposed the use of machine vision technology(a subset of computer vision) to detect defects in wooden-ties, rail anchors, crib ballast and turnout components. Machine vision is based on the principle that certain features of the track can be used as a template and subsequently compared with the new image. Stella et al. [16] proposed the idea of using machine vision and neural networks to detect hexagonal bots while Marino et al. [18] performed a quantitative analysis on hexagonal bots detection using Field Programmable

Gate Array(FGPA). Edwards [17] investigated the detection of track component such as ties and spikes using edge detection [37] and texture information. Gaussian mixture models have also been used in the paper to refine defect classification.

Such a technique is able to pick out the location of the fault, however it would be not be able to extract geometrical properties of the fault. Machine learning is also computationally inefficient as it requires a large amount of scans to produce a template.

2.3.2 Laser line scanning

High speed laser in conjunction with line-scan cameras(Laser Line Scanning) have also been used to create high resolution 3-D datasets [38, 39, 40, 41]. Laser Line Scanning is based on the principle of triangulation [42]. This approach gives more of a visual representation to the rail as opposed to the laser ultrasound method. Lorente et al. [38] used a high speed laser in conjunction with a scan-line camera to create a 3-D cloud of points with visual information coming from the scan-line camera. An odometer and an inertial measurement unit (IMU) were used to store the trajectory information. The system could detect and identify various parts of the rail using edge detection. The approach could also detect missing fasteners by manually comparing the 2-D images of the fasteners with the expected result. Once the location has been found, the 3-D point cloud of the location is extracted and subtracted from the expected 3-D point cloud. Li et al. [14] used a structured light approach to detect and measure rail wear. This method used a single line structure light vision sensor to compute the registered measured profile with the reference profile. 3-D coordinates of rail profile can be computed by integrating the light stripe center points from the laser with the camera's coordinate frame. Due to the inclusion of high speed laser as the sensor, the system could reach a limit of 500 frames per second (FPS) measurement frequency.

Laser Line Scanning when used exclusively with a laser is not able to produce color information. An inclusion of a laser and camera can produce geometrical properties, however, synchronization will need to be performed for the camera and laser to work together.

2.4 Road inspection

In addition to rail inspection, there is also a need for road inspection(or road profiling). Roads or rough terrains can contain defects such as crack, potholes and gutting [43]. As such, there becomes a need to characterize these defects geometrically to prevent road accidents. Tire companies are in need of a system to profile roads to maintain consumer acceptability. The parameter of interest when it comes to roads is roughness. By roughness, one can tell the ride comfort level. One such parameter for measuring roughness is the International Roughness Index(IRI)[43]. The IRI is measured in "inch/mile" to describe the height deviations per longitudinal length. Measurement of the IRI contributed to the development of road profilers.

2.4.1 Static profilers

By definition, an apparatus could only be considered a road profiler if it could measure the reference elevation, the height relative to the reference and the longitudinal distance. The first road profiler was a "rod and level" technique [43]. The apparatus included a level for reference elevation while the readings from the rod provided the height relative to the reference. This apparatus could only be used in a static case and was considered a true profiler. By late 20th Century, a new road profiler "Dipstick" by Face Company was introduced for static cases. The Dipstick involved an inclinometer that measured the relative difference between the two base supports of the apparatus. The apparatus is moved

by pivoting it 180 degrees about its leading foot, positioning the former foot from behind to the front. The reference elevation is the height value of previous point. The Dipstick is used as a ground truth measurement for profiling approaches as well [44, 45].

2.4.2 Inertial profilers

The final type of road profilers are the inertial road profilers for dynamic cases. These profiling systems use laser [7, 8, 9, 10, 11], ultrasound or optical sensors as the measuring device. Accelerometers are attached to these systems to measure the vertical acceleration produced by the sensor. A double integration of this acceleration gives us the effect of the vehicle dynamics which can then be removed from the height measurements [46]. Most of the inertial road profilers use laser line scanning as the sensor of choice rather than ultrasound or optical sensors. For instance, Cigada et al. [8] used two laser transducers to estimate the vehicle velocity that is needed to acquire the road texture from the sensor time histories. Laurent et al. [10] used an automatic 3-D profiling method using the Laser Crack Measurement System(LCMS). The system was able to detect ruts and cracks using intensity data. Li [11] proposed the idea of using 3-D lasers to detect pavement defects such as rutting. The approach used a statistical outliers removal technique to distinguish rut features from other defects. In another approach, González et al. [47] solely used the accelerometers of the inertial system to determine the Power Spectral Densities of the profile(PSD). When using profilers systems, the key topic of interest is the instrument's accuracy. For this reason, correlation analysis is performed for the profiler to quantify the measurement error. Sayers et al. [48] and Gillespie et al. [49] talk about the need for calibration and correlation for frequency response profiling systems. With these two procedures in mind, one can determine the repeatability and reproducibility of the apparatus.

As mentioned in the rail profiling literature review, although lasers produce high accuracy results, it is not able to produce color information without the integration of other sensors such as color cameras. In addition, the integration of other sensors such as accelerometers or IMUs require further synchronization. Hence the need for a new inexpensive inspection method which utilizes both color and geometry comes into play. This need can be accomplished using a 3-D reconstructed map building approach along with an appropriate sensor as opposed to inspecting single scans at a time using computer vision or laser approaches. The fundamentals of the 3-D reconstruction for this map building approach are explained in the next chapter.

2.5 Summary

This chapter provided an overview of the defects found in the railway and road industry. Past non-destructive testing methods such as Ultrasound, Computer vision, Eddy current, Magnetic Induction and Laser Line Scanning were discussed in this section. It was found that Ultrasonic, Eddy Current and Magnetic induction could be used to detect subsurface cracks but could not be used to create any geometrical information or color information without the use of other sensors. Emerging methods such as Computer Vision methods allow us to detect and classify defects based on image processing. However it cannot reproduce geometrical information. Laser Line scanning incorporates a triangulation method which makes it possible to see 3-D datasets but is not able to see color information for defect classification. As such, a new approach is needed which can reproduce color and depth information without the inclusion of other sensors. The following chapter explains the 3-D reconstruction process for this new approach.

Chapter 3

Three Dimensional Reconstruction and Offline Mapping

When it comes to railroad profiling using inertial/optical sensors, we are looking at three aspects: area of the rail to be captured by multiple sensors, the location of each point using co-ordinate frames and the specific features of the profile. The optical sensor that can accomplish all these aspects is the proposed RGB-D camera. The RGB-D camera's 3-D reconstruction principle allows an environment to be modeled as a virtual three dimensional object with the use of algorithms. However 3-D reconstruction is not enough as color information would also have to be synchronized with the depth values to achieve a meaningful model. This chapter will cover the how an RGB-D camera re-constructs a depth model of the object and synchronize color information to get a color(RGB) point cloud of the model. Then, certain features from this point cloud can be extracted using feature detectors and extractors. Finally, we can then use the extracted points for local and global point cloud matching.

3.1 RGB-D Reconstruction

3.1.1 RGB-D camera

An RGB-D camera consists of three elements: An infrared(IR) illuminator or projector, an IR camera and a color(RGB) camera. The RGB camera records color information, the IR projector emits an infrared laser while the IR camera works as a sensor for the infrared laser. The RGB-D camera discussed in this thesis is a Time of Flight(ToF) camera which is the Kinect v2 sensor as shown in Figure 3.1a. The Kinect v2's field of view is 70.6° horizontal and 60° vertical as shown in Figure 3.1b. The values in the Z-direction, or depth values, are calculated using the Time of Flight (ToF) principle [50] through Equation 3.1 whereas the x and y values are determined by using homogeneous image coordinates u and v, explained extensively in [51]. The calculation of the x and y values are shown in Equation 3.2 and 3.3 respectively. The depth measurements are explained in detail in Section 3.1.2. The collection of these points is known as a 3-D point cloud. Each scan of the Kinect v2 has an RGB resolution of 1920x1080 pixels and a IR camera resolution of 512x424 pixels. Keeping these resolutions in mind, the Kinect v2 outputs 217,088 colored 3-D points every frame. These acquisitions can be performed up to 30 frames per second (FPS). The output stream from the Kinect v2 is shown in Figure 3.2 [52].

$$z = \frac{\Delta\varphi C}{4\pi F} \tag{3.1}$$

where z is the depth in meters, $\Delta\varphi$ is the phase shift, C is the speed of light and F is the modulation frequency.

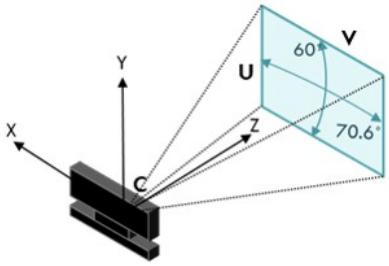
$$x = \frac{u - c_x}{f_x} \tag{3.2}$$

where x is the horizontal position, u is the vertical image coordinate, c_x is optical center in the X-direction and f_x is the focal length in the X-direction

$$y = \frac{v - c_y}{f_y} \tag{3.3}$$

where y is the vertical position, v is the horizontal image coordinate, c_y is optical center in the Y-direction and f_y is the focal length in the Y-direction.

(a) Kinect v2's components



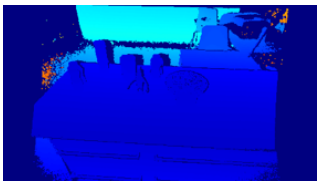
(b) Kinect v2's field of view

Figure 3.1: Kinect v2: a)Components b)Field of View.

3.1.2 ToF sensor depth measurement principle

The Kinect v2 is the successor of the Xtion Pro Live RGB-D camera. For the Xtion Pro Live RGB-D camera, the IR camera analyzed the fixed speckle pattern projected by the IR projector and computed depth values by triangulation. The analysis of this pattern is commonly referred to as the Structured Light(SL) approach [53]. In a SL approach, a memorized IR pattern stored in the RGB-D camera's computer architecture is projected

(a) RGB image



(b) Colored depth image

(c) 3-D colored point cloud

Figure 3.2: Output stream from Kinect v2: a)RGB image b)Colored depth image. c)3-D colored point cloud [52]

onto the screen and is compared with the current pattern on the screen. If there are any obstacles in the way, the IR pattern changes shape from which the depth values can be deciphered. The Xtion Pro Live's successor Kinect v2 however, uses ToF technique [50] to acquire depth values The sensor measures the time it takes for the modulated laser pulses from the IR projector to reach the object and then back to the IR camera. As such, a depth value can be measured for each pixel instead of comparing the distances between each pixel in the structured light approach. The overview of both SL and ToF approaches is shown in Figure 3.3.

Figure 3.4 shows the depth measurement principle of a ToF RGB-D camera. A timing generator in the sensor chip is used to create a clock signal for modulating the 860 nanometers laser light source (transmitter) and the return signal (receiver) [54]. The reflected or the return signal reaches the sensor lens which lights up a 512x424 pixel array(also known as a depth image). All signals are measured in voltage. Due to multiple captures, a 10 bit 2GS/s Analog to Digital Converter(ADC) is employed as well [55]. For each depth measurement, light is illuminated on the A side (left half) and B side (right half) of each pixel depending

on the clock signal. The A side of the pixel is illuminated when the light signal(synchronized with the clock) is on and B side of the pixel is on when the light signal is off. When the clock is high, the A side of the pixel is illuminated which indicates the fact that the object is very close to the sensor. As the object is moved farther away, the ratio of the photons on B increases compared to A. The ratio of the photos on the A and B side are then compared using the A-B out signal shown in Figure 3.4 whose output voltage depends on the returning light level and the time it takes to arrive with respect to the clock [54]. This is how the phase shift is measured in Equation 3.1.

As shown from Equation 3.1 that depth resolution is a function of the modulation frequency. Increasing modulation frequency can increase resolution, however it will also create ambiguities due to phase aliasing . To counteract this problem, the Kinect v2 uses multiple modulation frequencies ranging from 16MHz to 120MHz [54, 55, 56]. Kinect v2 takes two measurements where the first measurement is done using low modulation frequency with no room for uncertainties. The second measurement is taken with high modulation frequency using the low modulation frequency as an estimate to remove any ambiguities. Keeping this in mind, we can determine the depth resolution using the ADC amplitude resolution of 10 bits. Firstly, the range resolution is taken into account using Equation 3.4 which in turn is used for calculating the depth resolution in meters(m) using Equation 3.5.

$$Range = \frac{C}{2F} \tag{3.4}$$

$$\Delta z = \frac{Range}{2^{bits}} \tag{3.5}$$

For example, a 16MHz modulation frequency will have a depth resolution of $9.16 * 10^{-3}$ m whereas a 120 MHz will have a resolution of $1.22 * 10^{-3}$ m. The Kinect v2 was designed to have a depth resolution within 1 percent of the distance as explained in [54]. More information on the accuracy and standard deviation of this sensor chip can be found in [55].

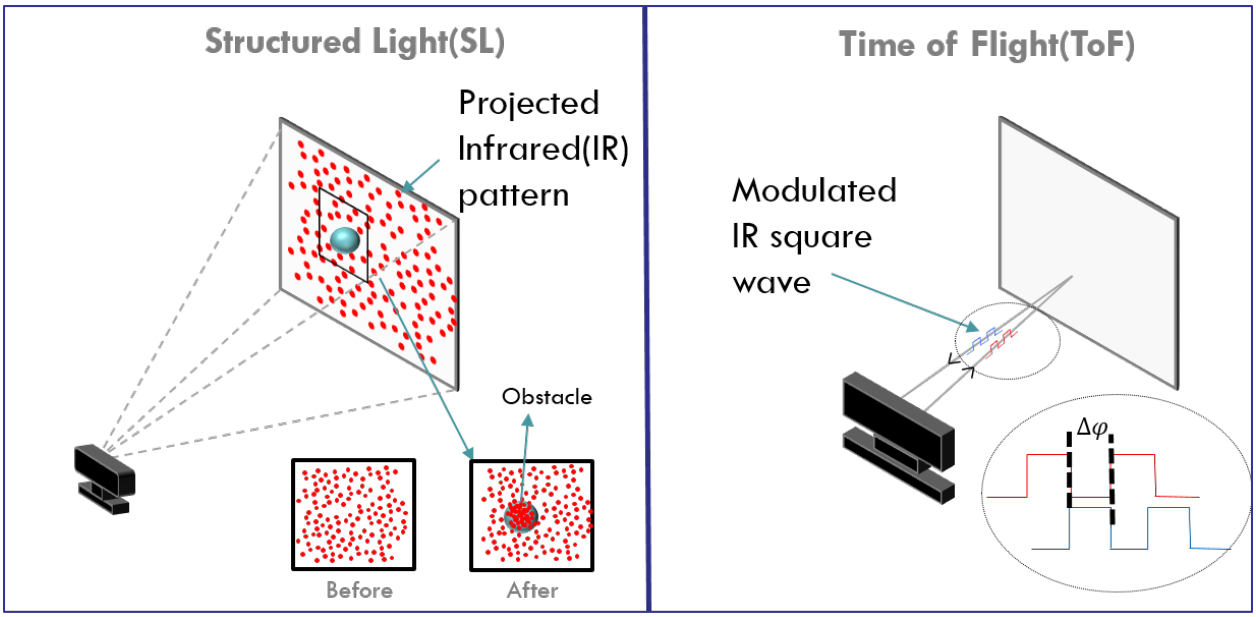


Figure 3.3: Comparision of SL and ToF sensors

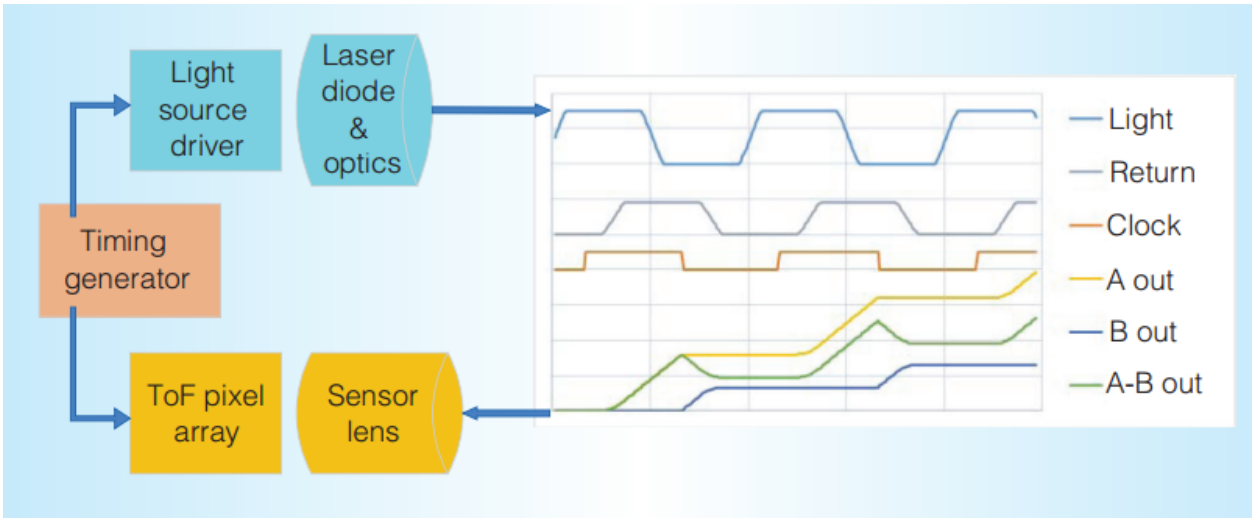


Figure 3.4: Depth measurement from ToF sensor [54]

3.1.3 Multiple RGB-D cameras

Kinect v2's predecessor had a problem acquiring valid data when there were multiple RGB-D cameras whose field of view intersected with each other. This is because of the fact that the projected IR patterns from each camera interfered with each other and as such the camera

was not able to determine depth values [57]. This problem becomes highly evident in outdoor environments where there is a rich source of infrared waves from sunlight. The advantage Kinect v2 holds over its predecessor is the fact that it uses ToF instead of relying on projected IR patterns for computing depth. As such, the interference problem is greatly reduced as the sensor does not have to compute distances between neighboring points on the pattern. Another advantage the Kinect v2 holds over the Xtion, is that the camera has a built in ambient-light rejection method, which makes it possible to use in an outdoor environment with near infrared sources without the fear of interference. The Kinect v2 however, has a limitation in regards to the fact that multiple cameras cannot operate on one computer as each camera consumes over half of the computer's Universal Serial Bus(USB) 3.0 controller's bandwidth [54]. The effect of interference for structured light cameras and time of flight cameras are shown in Figure 3.5. As seen from Figure 3.5, multiple SL cameras exhibit deteriorating depth values in the resulting 3-D point cloud whereas multiple ToF cameras have little to no interference effect.

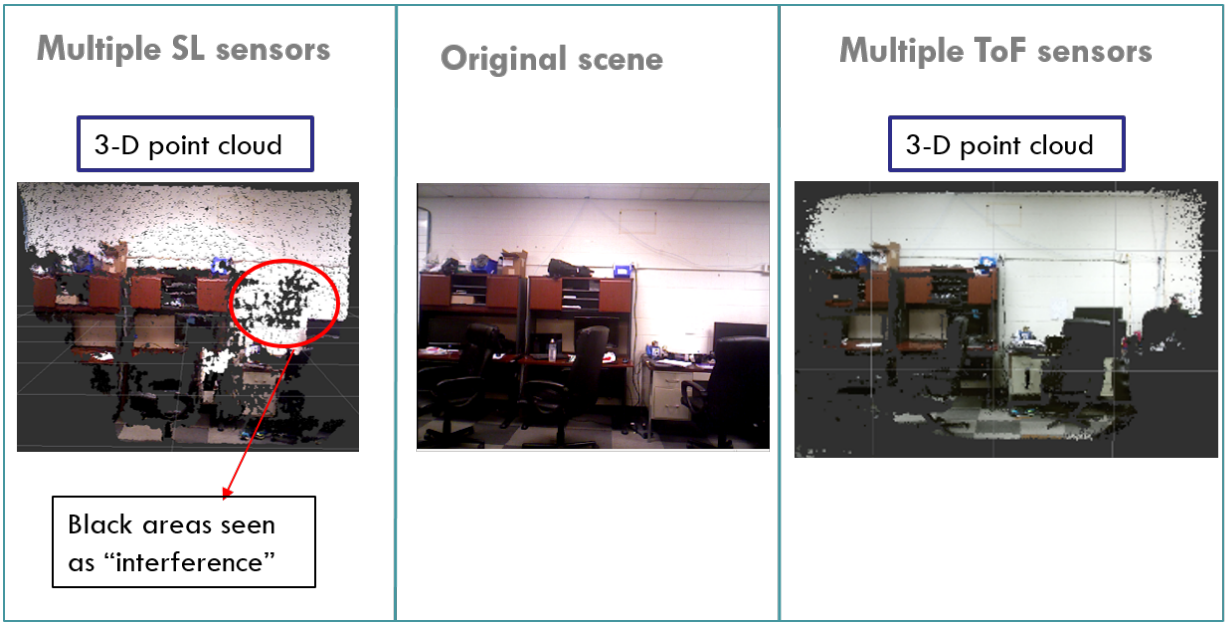


Figure 3.5: Interference effects of SL and ToF sensors

3.2 Local and Global Coordinate systems

With the working principle of the RGB-D camera explained, the coordinate systems of the RGB-D camera can now be discussed. The pose of an RGB-D camera can be referred to as its orientation in its local coordinate frame. The pose of the camera can be broken down into rotation and translation parameters. For a 6-Degrees of freedom Cartesian coordinate system, the pose $\mathbf{p}_k$ contains the rotation matrix, $\mathbf{R}$, about the X,Y and Z axis, and the translation vector, $\mathbf{t}$, in X,Y and Z as shown in Equation 3.6

$$\mathbf{p}_k = \mathbf{R}(\theta_k)^{\{L\}}\mathbf{x}_k^i + \mathbf{t}_k^i \tag{3.6}$$

where ${}^{\{L\}}\mathbf{x}_k^i$ is scan point i at time k of frame L. Once a local pose is established, the question is how we represent a collection of local poses into one global pose or map. Global transformation is a process which converts the local scans into a global coordinate framework. If we consider the first scan as our initial scan in our global map, the transformation matrices between a pair of scans can then be cascaded to put all the scans into the first scan's coordinate system. This formulation is shown below in Equation 3.7

$${}^{\{G/L^k\}}\mathbf{p} = {}^{\{G/L^2\}}\mathbf{p} * {}^{\{L^2/L^3\}}\mathbf{p} * ... * {}^{\{L^{k-1}/L^k\}}\mathbf{p} \tag{3.7}$$

where $\mathbf{G}$ is the global frame taken initially as the first scan, $\mathbf{p}$ is the transformation matrix between a pair of consecutive scans and $\mathbf{L}$ is the local frame.

3.3 Feature Extraction

A point cloud from an RGB-D camera as mentioned in the previous section has 217,088 points in each depth frame. Such a large amount of points may involve high computation times when it comes to post processing. With that notion in mind, a method of correspondence would be required to stitch two consecutive frames together without using all the points in each frame. One method would be to use features from the image and match those features with the next image.

An image can have many features, however there only some certain ones that are distinct or distinguishable in an image. The aim is to find these features for successful correspondences. Studies have shown that selecting corners or blobs as a feature yields successful correspondences. This is because of the fact that moving away from an edge gives high variances in every direction. One computer algorithm that detects these features and describes them in terms of scale and orientation is the Scale Invariant Feature Transform(SIFT) [58, 59]. The advantage of SIFT over other feature detection algorithms is the fact that it can detect correspondences under any scale or rotation change. With this in mind, features from one frame to the next frame overlooking the same observations can now be matched together.

3.3.1 Outlier rejection

There are cases when wrong correspondences are found due to noise or outliers. One of the main reasons for these outliers is due to motion blur. When the object is moving too fast for the RGB-D camera to currently distinguish the features of the image, feature detection will fail as there are no distinct points to choose from. This is one of the limitations of the Kinect v2 as it cannot create a reliable map under high speeds. For this reason, a model

fitting algorithm known as Random Sample Consensus (RANSAC) [60] is used to throw out points which are not within a known threshold distance.

3.4 Point Cloud Registration

The process of aligning and stitching consecutive pairs of local scans of a dataset(provided the current scan has overlapping points with the previous scan) into a map is known as Point Cloud Registration(PCR). The aim of PCR is to find matching correspondences between frames overlooking the same observations so that the frames can be stitched together in a global co-ordinate system. The steps of PCR are outlined in Figure 3.6.

1. Firstly, a pair of 3-D point cloud data sets(${}^{s1}\mathbf{x}_k^i$ and ${}^{s2}\mathbf{x}_k^i$) are input using two RGB-D cameras where s1 and s2 are the respective camera coordinate spaces.

2. Once both datasets have been acquired, correspondences between points of (${}^{s1}\mathbf{x}_k^i$ and ${}^{s2}\mathbf{x}_k^i$) must be established using feature detection such as SIFT. False correspondences are rejected using RANSAC.

3. When all the correspondences have been found between the two point cloud datasets, the matchings scans would need to be transformed onto one coordinate space using registration algorithms explained in Section 3.4.1. This transformation is performed using rotation and translation parameters.

In this thesis, a combination of two established registration algorithms are used: Iterative Closest Point(ICP) [61] and Normal Distributions Transform [62].

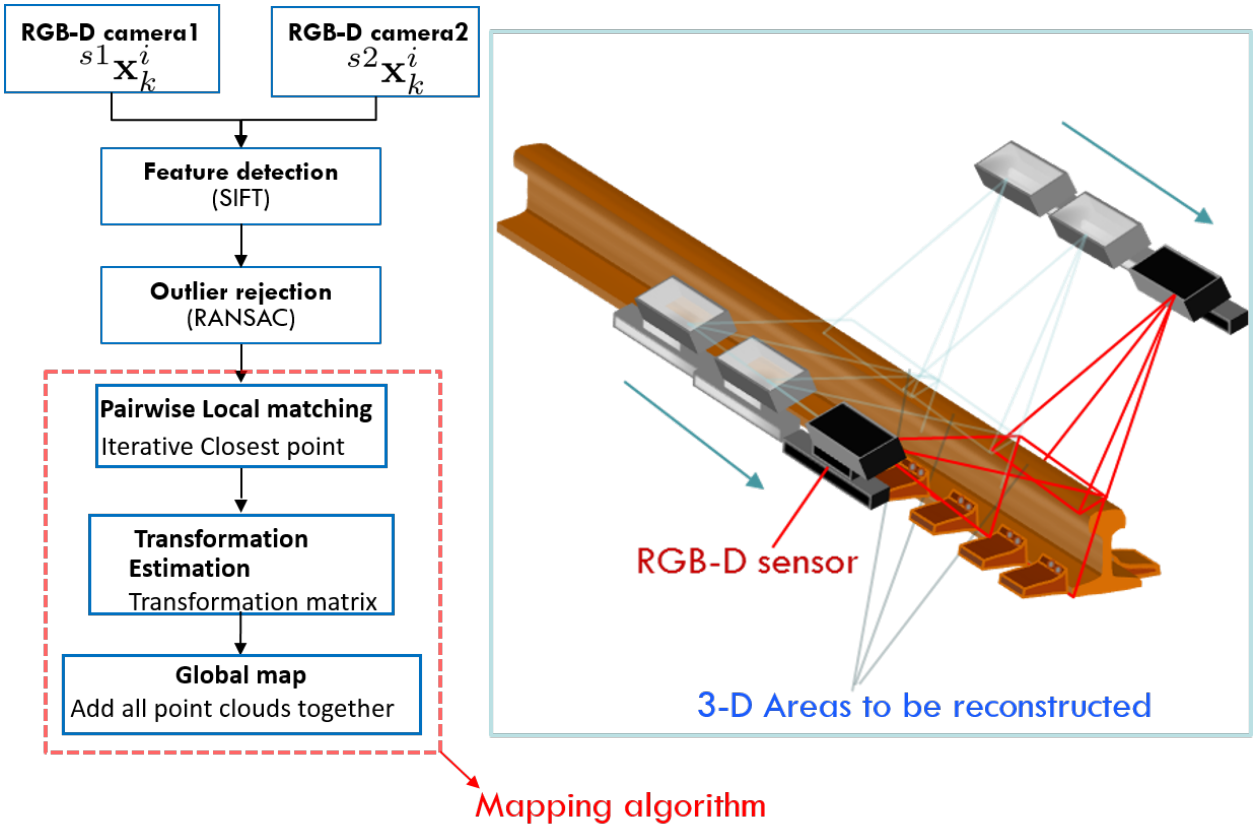


Figure 3.6: Point Cloud Registration

3.4.1 Pairwise registration

Iterative Closest Point

Consider two 2-D local frames containing datasets $\mathbf{x}_{k-1} = \{x_{k-1}^i ... x_{k-1}^N\}$ and $\mathbf{x}_k = \{x_k^i ... x_k^N\}$ respectively where N is the total number of points, $\mathbf{x}_k$ is the source frame, k's, dataset and $\mathbf{x}_{k-1}$ is the target frame, k-1's, dataset. Under the assumption that $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$ have the same points(or more than a 50% overlap) and differ by a combination of a rotation and translation value, one can iteratively transform $\mathbf{x}_k$'s frame to $\mathbf{x}_{k-1}$'s frame. This method is known as Iterative Closest Point(ICP) devised by Zhang [61] and it involves a series of steps as shown in Figure 3.7. The $\mathbf{z}_k$ is the result of transforming $\mathbf{x}_k$'s frame onto $\mathbf{x}_{k-1}$'s frame

and adding those transformed points to form a new dataset $\mathbf{z}_k$ under frame k. The motive of this method is to minimize the sum of the least squared distance error between $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$ given by Equation 3.8

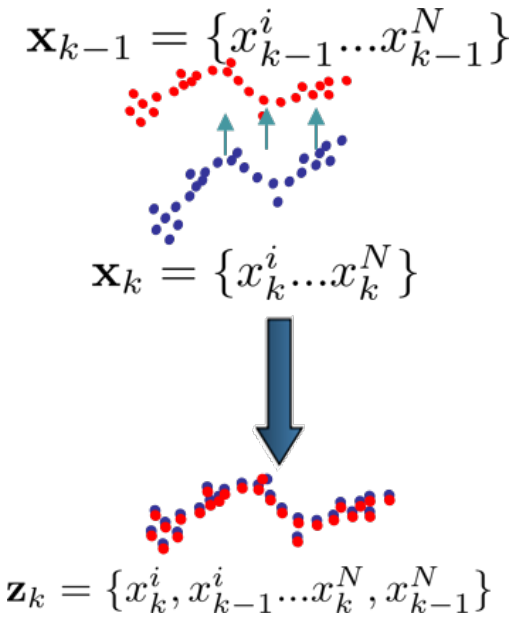


Figure 3.7: Iterative closest point steps

$$E(\mathbf{R}, \mathbf{t}) = argmin\{\frac{1}{N_{x_k}} \sum_{k=1}^{N_{x_k}} ||x_{k-1} - \mathbf{R}x_k - \mathbf{t}||^2\} \tag{3.8}$$

The error of the distance between the source and target vector is squared to eliminate any negative error values. Detailed formulations to find the optimum $\mathbf{R}$ and $\mathbf{t}$ can be found in [61]. ICP is used as the first step on a pair of datasets. Once $\mathbf{x}_k$'s frame is transformed onto the previous frame $\mathbf{x}_{k-1}$, a new pair of frames can then be taken into account where the transformed $\mathbf{x}_k$ frame will now act as the target frame and the incoming time step scan $\mathbf{x}_{k+1}$ will be taken as the source frame. Taking this under consideration, there will exist a small distance error between each pair of target and source frames as the distance error will reach a converging value. When successive pairs of poses are added together in this fashion, naturally , the errors will accumulate as well. For this reason there needs to be a second

step which reduces local errors between frames. This second step is known as the Normal Distributions Transform [62].

Normal Distribution Transform

The Normal Distribution Transform allows datasets to be represented as probability distributions, thus reducing the effect of sensor noise or outliers. Consider a 2-D case where a point cloud frame is divided into a grid of cells of equal size. Under the assumption that there are at least three points in each cell, the mean and covariance be calculated by Equation 3.9 and 3.10 respectively.

$$\mathbf{q}_k = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_k^i \tag{3.9}$$

$$\mathbf{\Sigma}_k = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_k^i - \mathbf{q_k})(\mathbf{x}_k^i - \mathbf{q_k})^T \tag{3.10}$$

where $\mathbf{x}_k^i$ is the ith 2-D point in frame k's cell and n is the number of points in the cell. Using the mean and covariance, the normal distribution $P(\mathbf{x})$ of the points in each cell can be computed by Equation 3.11.

$$P(\mathbf{x}) \sim \exp(-\frac{(\mathbf{x}_k^i - \mathbf{q}_k)^T\mathbf{\Sigma}_k^{-1}(\mathbf{x_k^i} - \mathbf{q_k})}{2}) \tag{3.11}$$

Thus a piecewise continuous description of the 2-D plane is formed. To prevent discretization, four grids are overlapped with each other such that each 2-D point falls into 4 cells. More details of this formulation can be found in [62]. To perform pairwise registration against grids, the transformation is first computed. The transformation function involves the use of the pose $\mathbf{p}$ which transforms source frame k with dataset $\mathbf{x}_k^i$ to the target frame k-1 with the dataset $\mathbf{x}_{k-1}^i$. An initial estimate of $\mathbf{p}$ can be found using the ICP algorithm. The probability density of the newly mapped points are then computed. Since the probabil-

ity density function of each frame is independent from one another, the joint likelihood can then be found by calculating the product of all the probability density function as shown in Equation 3.12.

$$\psi = \prod_{i=0}^{n} P(T(\mathbf{p}, \mathbf{x}_k^i)) \tag{3.12}$$

where ψ is the joint likelihood of the probability density functions and $T(\mathbf{p}, \mathbf{x}_k^i)$ is the transformation function consisting of a rotation matrix and translation vector. Using the same ideology as Equation 3.12, the score of the registered map can be computed by evaluating the sum of the normal distributions of the transformed target frame's dataset $\mathbf{x}_k^i$ with the source dataset's $\mathbf{q}_k$ and $\mathbf{\Sigma}_k$ as shown in Equation 3.13.

$$score(\mathbf{p}) = \sum_{k=1}^{n} \exp(-\frac{(\mathbf{x}_{k-1}^i - \mathbf{q}_k^i)^T \mathbf{\Sigma}_i^{-1} (\mathbf{x}_k^i - \mathbf{q}_k^i)}{2}) \tag{3.13}$$

The optimal pose $\mathbf{p}$ is then found by determining the minimum of the score. This optimization is found by using Newton's algorithm. Newton's algorithm can be used iteratively to find the pose $\mathbf{p}$ that gives the minimum of the score function. This optimization is performed using Equation 3.14.

$$\mathbf{H}\mathbf{\Delta p} = -\mathbf{g} \tag{3.14}$$

where $\mathbf{H}$ and $\mathbf{g}$ are the Hessian matrix and the transposed gradient vector of $score(\mathbf{p})$ respectively. The increment $\mathbf{\Delta p}$ is then added to the current estimation to give a new $\mathbf{p}$ as shown in Equation 3.15

$$\mathbf{p} \leftarrow \mathbf{p} + \mathbf{\Delta p} \tag{3.15}$$

As such we gain an optimal transformation pose $\mathbf{p}$ that gives the minimum distances between the transformed frame and the source frame. In this thesis, the pose estimate from the Iterative Closest Point is used as an initial guess for the Normal Distribution Transform. The difference between 2-D-Normal Distribution Transform and 3-D-Normal Distribution Transform is the transformation function. Instead of axis/angle notation, 3-D Euler angles are used in the 3-D case. The approach is the same, however, the Hessian matrix $\mathbf{H}$ and gradient vector $\mathbf{g}$ will be computed differently. A detailed formulation is explained in [63]. Once all pairs of consecutive local frames have been matched, the entire dataset can then be converted into a global map using the global transformation formulation described in Section 3.2.

3.5 Summary

In this chapter, a comprehensive overview of three dimensional reconstruction by an RGB-D camera was presented. The interference problem with integrating multiple RGB-D cameras was then shown. The coordinate frames of the RGB-D camera were then discussed as the parameters of a map. Offline mapping was then explained in terms of Point Cloud Registration. Two main areas of this registration process, specifically the Iterative Closest Point and Normal Distributions Transform were then explained in detail. The formulations regarding the Iterative Closest Point followed by an in-depth detail of the Normal Distributions Transform were then explained. The following chapter will present a system overview of the registration process in terms of rail profiling as well as its implementation.

Chapter 4

Rail Profiling by RGB-D Scan matching

This chapter presents a map-based identification approach for detecting railway faults. The proposed map-based identification approach presented in this chapter would be able to detect raised spikes or misplaced ballast on tracks using an inexpensive algorithm approach. The next section describes the design and development of the system to construct that global map. In addition to this, several parameters of that system will be explained and justified in this section.

4.1 Map-based identification approach

As stated in the previous chapter, once the optimal transformation pose has been calculated between each consecutive scan, scans can be converted into global coordinates and be respectively defined as a map in a global system. When such a map has been acquired, it can be used as a reference for new scans. The prior global map is compared with new scans which

are acquired in the same manner to detect any spatial changes between the two colored 3-D point clouds. Figure 4.1 shows an overview of the process. As seen from Figure 4.1, the new scans are stitched together using the pairwise registration process and are subtracted from the previously stored global map. If there are any changes seen in the point cloud, those changed will be highlighted in red to classify a defect.

In terms of computational efficiency, this approach reduces computer memory usage as it does not need to train images to get different orientations. The approach is able to obtain high efficiency due to the way the dataset is organized. During this identification process, both the prior global map data set and the new scan dataset are divided into a tree-based data structure built over an area region or space. A cubic bounding box or node first encapsulates all the points of the dataset. This bounding box is then further divided into eight equal cubes or child nodes. The child nodes are recursively added until there are no points inside each child node also known as a leaf. This technique allows indexing and querying of points faster if their global position is known. If the points stored in the nodes of the new scan's tree structure do not match with the global map's node, those pixels from the new map are then extracted using indexing operations. The operation is quick due to the organization of the tree structure. Figure 4.2 shows the application of this approach on a railway track. The red markings signify defects which in the case of Figure 4.2 can be either in the form of missing spikes or misplaced ballast.

4.2 RGB-D camera depth precision

The RGB-D camera's best working distance was first determined by investigating the uncertainty of the depth values at various close range working distances. With the working distance determined, one can determine the design specifications for the hardware system

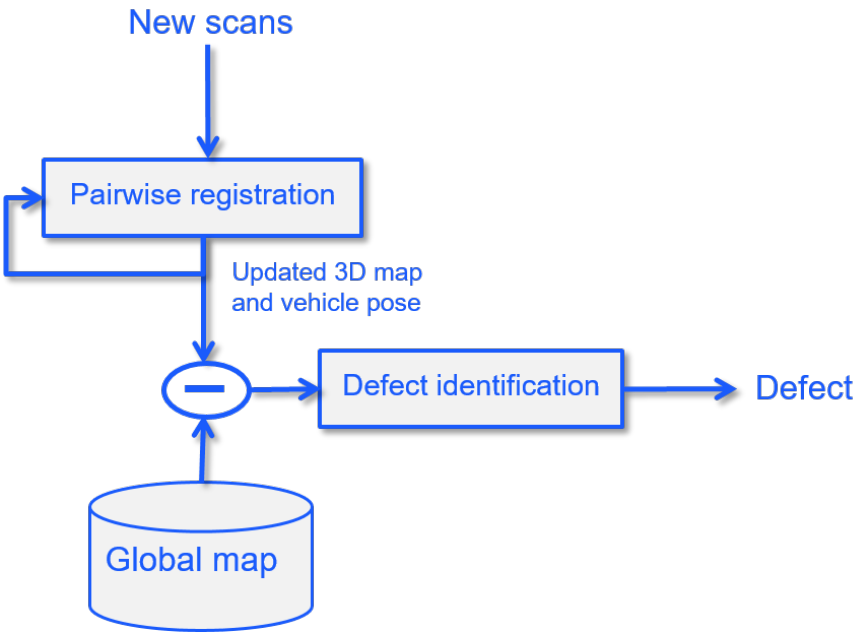


Figure 4.1: Map-based identification approach

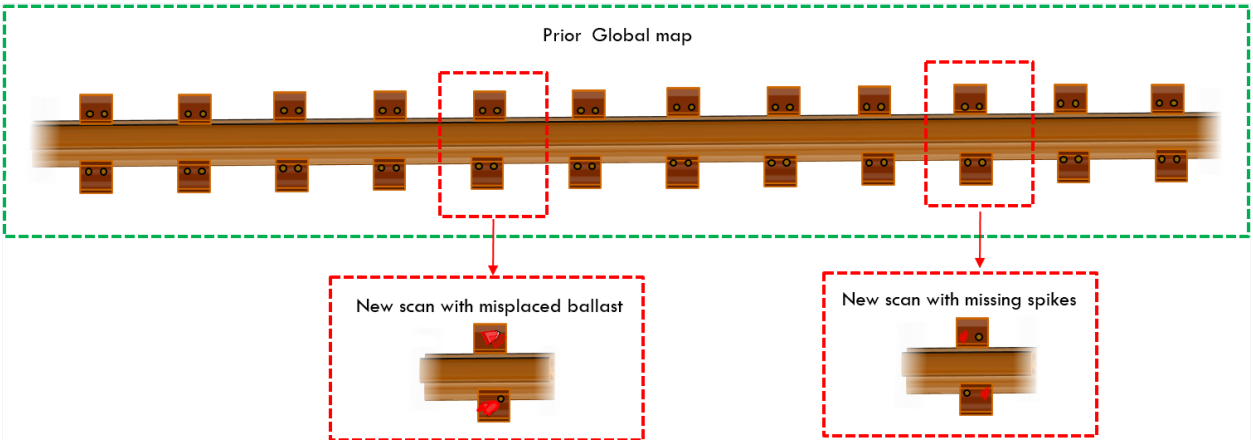


Figure 4.2: Map-based identification concept on rail track

in order to approve the approach. An experiment was performed to investigate the true precision of depth values from the RGB-D camera using the parameters as shown in Table 4.1. The experimental setup is shown in Figure 4.3. In this experiment, the depth values from the RGB-D camera was measured with respect to a single pixel on a flat plate at multiple working distances. The depth valuse from the RGB-D camera were then recorded

Table 4.1: Experimental parameters for RGB-D camera depth precision

Parameters	Value
Flat plate	100mmX100mmx15mm
Perpendicular distance to flat plate	[600,700,800,900,1000]mm
Number of images	450/position
Uniform LED light	30W 500 LED lamp panel

for multiple trials at each working distance. The flat plate was chosen on the basis of the fact that depth measurements from the camera to the pixel will not be affected by outside environment such as roughness, vibration or ambient noise. Close range working distances of 0.6 meters(m) to 1 m were chosen due to the fact that in the actual rail profiling experiment, close range measurements would be needed to capture small geometrical entities of the rail, such as spikes. Finally, uniform LED lighting was provided to create a controlled environment for depth acquisition.

Figure 4.3: Precision of depth values experimental setup

The results of this experiment is shown in Figure 4.4. As seen from the figure, the depth measurements exhibited least standard deviation and maximum precision at 0.6m. The standard deviation increased for large distances, inferring that the camera is not able to produce precise results as the distance increases. One can conclude from these results that at 0.6 m, the RGB-D camera exhibits 1.23 mm(millimeter) standard deviation with a ±0.3mm change. Although this result will not able to detect sub mm faults such as wear on the rail, it will be able to detect mm defects such as raised spikes. For this reason, 0.6 m was chosen as the working distance for the hardware setup of the proposed approach

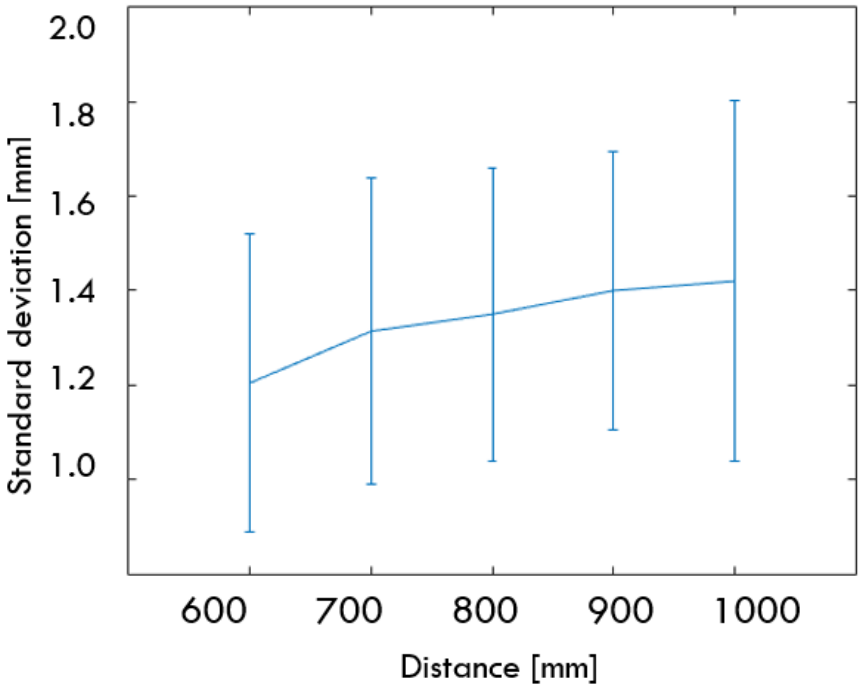


Figure 4.4: RGB-D camera depth precision results

4.3 Mobile rail profiling system overview

A motorized mobile rail profiling system was devised for outdoor purposes as shown in Figure 4.5. A separate computer is needed to operate each RGB-D camera as each one takes more

the 50% of the computer's USB 3.0 controller's bandwidth. In addition, each Kinect v2 requires its own power source. A server/client relationship is built between the mini CPU and the laptop which acts as an external monitor for both computers as well as an interface for operating multiple RGB-D cameras. This system is motorized with the help of high torque low speed DC gear-motors. These gear-motors connect to wheels that are customized to run on rail tracks. Finally, the data from both computers are synchronously stored onto a hard drive where it is then processed offline to create a map of the rail.

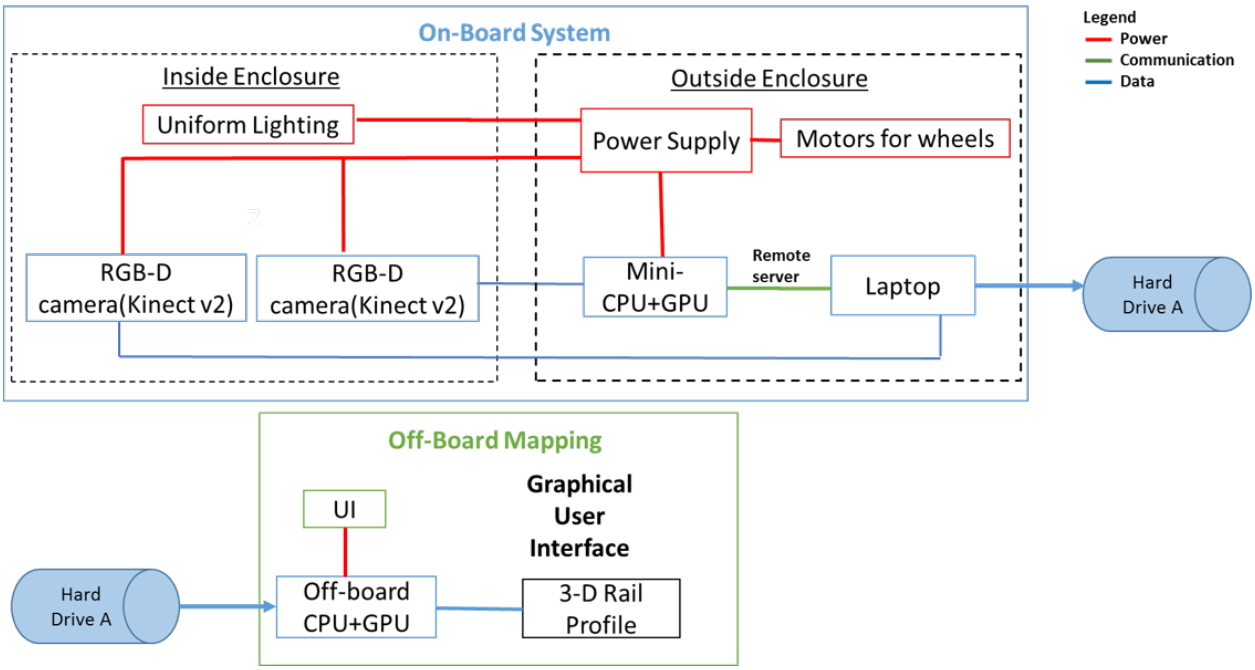


Figure 4.5: Mobile rail profiling system

4.3.1 Rail profiling schematic

A schematic has been presented in Figure 4.6 to break down the system's operations. From the figure, it can be seen that both computers are placed outside the enclosure. This is because there should be no object in the field of view of the Kinect v2s, which are inside the enclosure. The mini-CPU, Kinect v2s and DC LED strip lights are each powered by

a 11.1 battery source. The gear motors are powered in parallel by a 11.1 V battery. A Single-Pole-Single-Throw switch was included for ease of accessibility. The circuit for the gear motors was enclosed in a small switch-box for aesthetics.

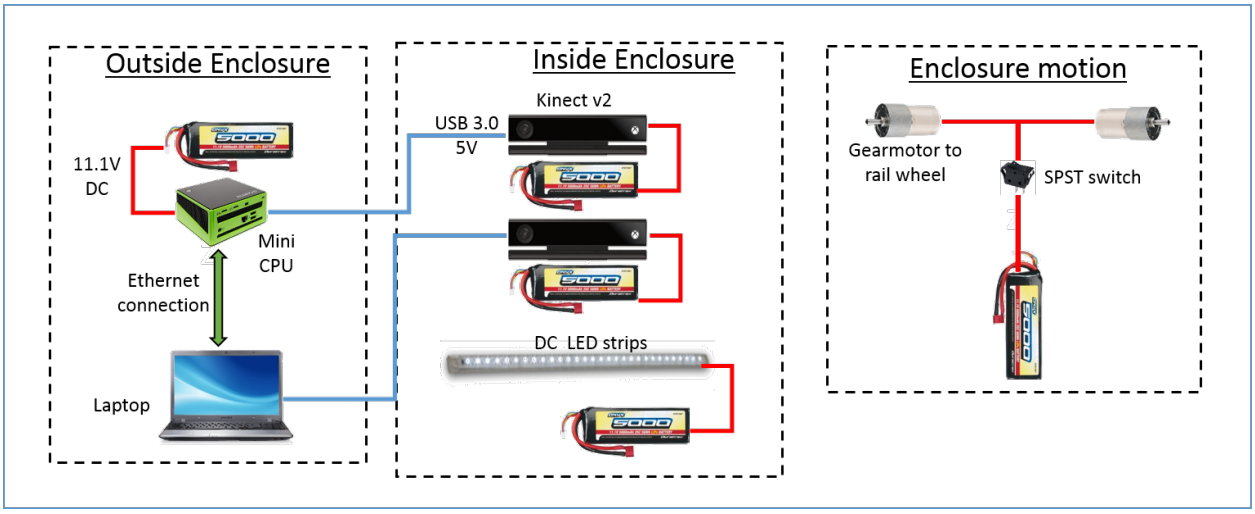


Figure 4.6: Rail profiling schematic

4.3.2 Rail enclosure design

Figure 4.7 shows the concept drawing for the mobile rail profiling enclosure. As stated in Section 3.1, although Kinect v2s do not have as much interference compared to its predecessor, it still would need a protective cover for outdoor environments on sunny days, since there is a rich source of infrared waves that may interfere with the Kinect v2's IR camera. The enclosure housing for this system was built using black painted wooden boards to block sunlight as seen in Figure 4.8. Figure 4.9 shows the blown up view of the rail enclosure. The black box housing in Figure 4.8 encompasses two Kinect v2s mounted on either side of its frame as shown in Figure 4.9. Due to the placement of these cameras, the box was designed to have a larger width than height ratio as seen in Figure 4.7. This is due to the Kinect v2's field of view. The cameras are placed 30 degrees with respect to the rail to allow max-

imum view of the blind spot of the rail. This placement has been validated in a miniature experimental setup explained in the rail enclosure development section. The wheels seen in the conceptual diagram are designed to simulate rail wheels, which are single flanged wheels. These wheels are designed to guide the enclosure even through curved tracks. Finally, uniform LED lighting is planted under the base rim of the enclosure to provide constant lighting for feature detection.

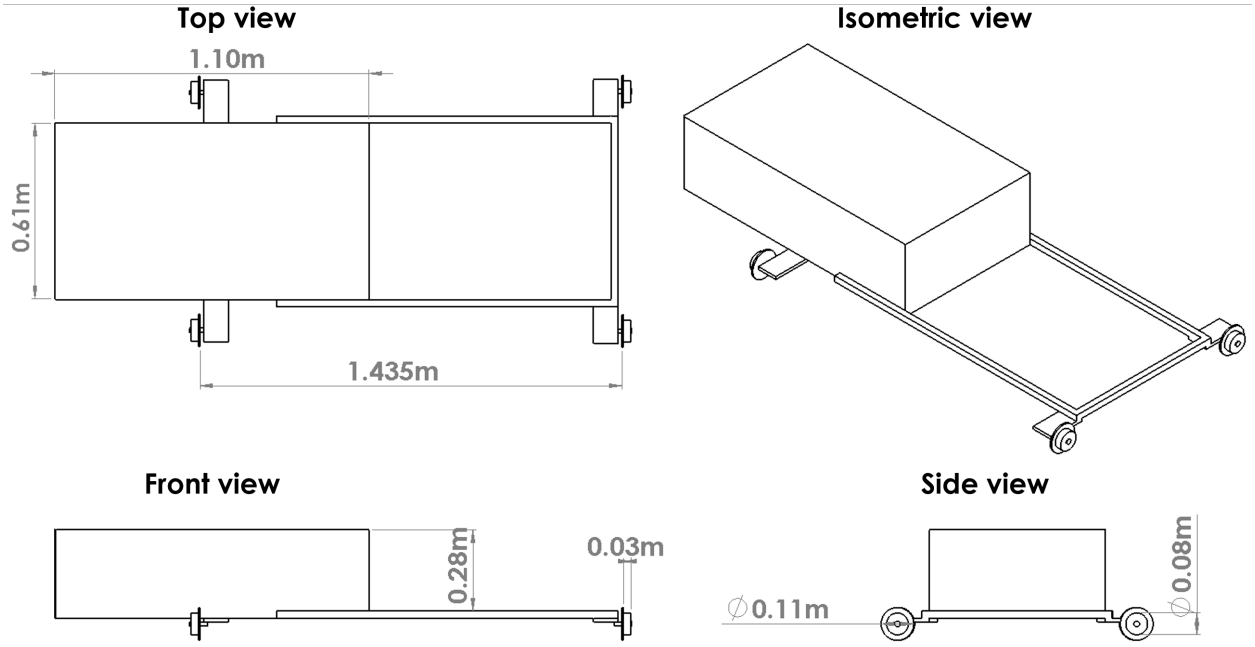


Figure 4.7: Rail enclosure design drawing

4.3.3 Rail enclosure development

Orientation of cameras

The developed rail profiling system is shown in Figure 4.10. For the mobile rail profiling experiment, the orientation of the cameras mattered as most if not all parts of a rail track need to be captured. The experimental setup is shown in Figure 4.11a and Figure 4.11b.

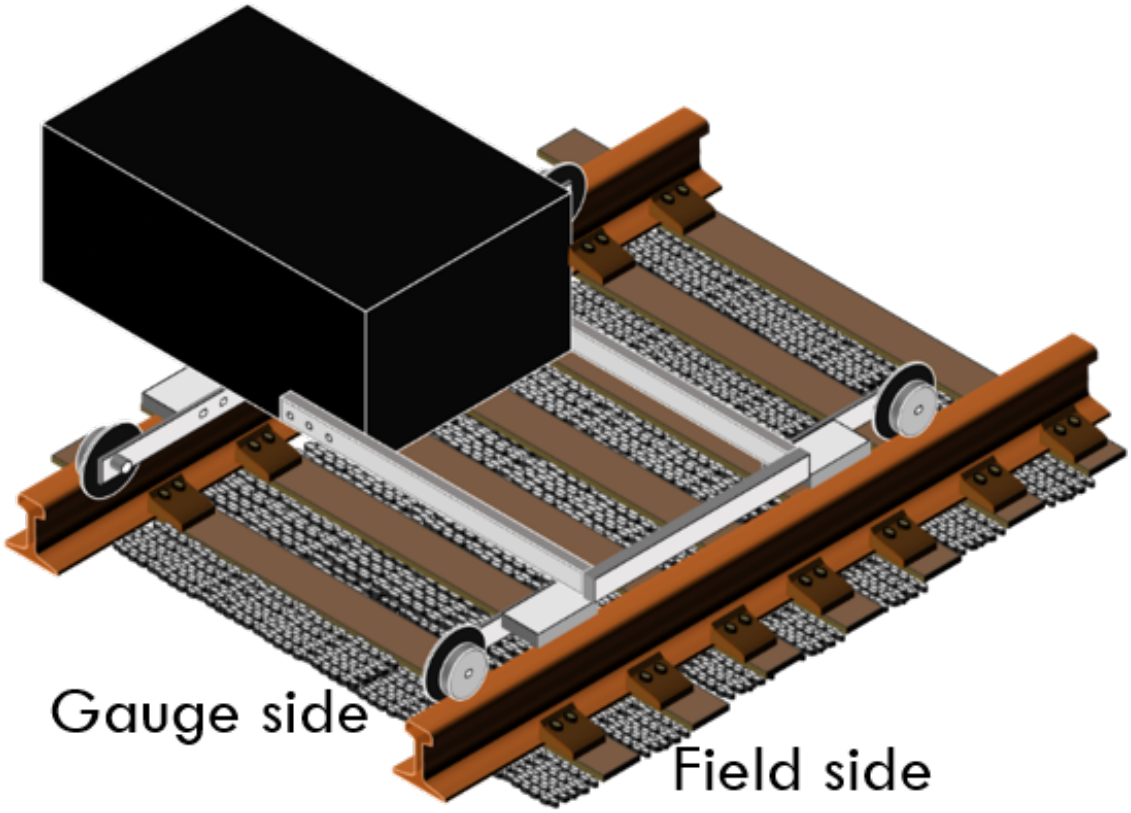


Figure 4.8: Rail enclosure concept

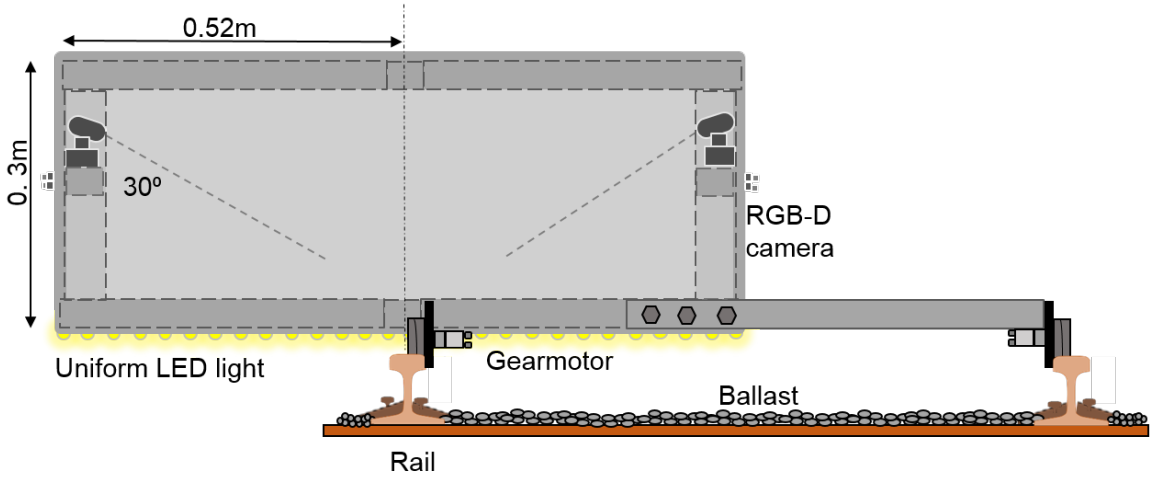


Figure 4.9: Blown up view of rail enclosure

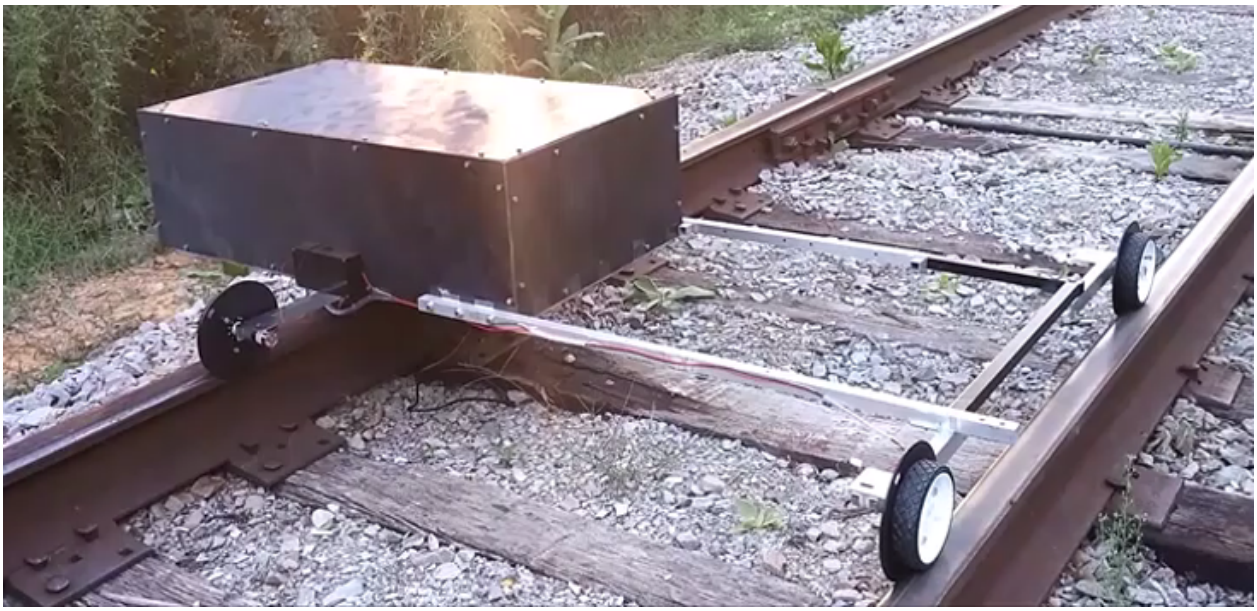
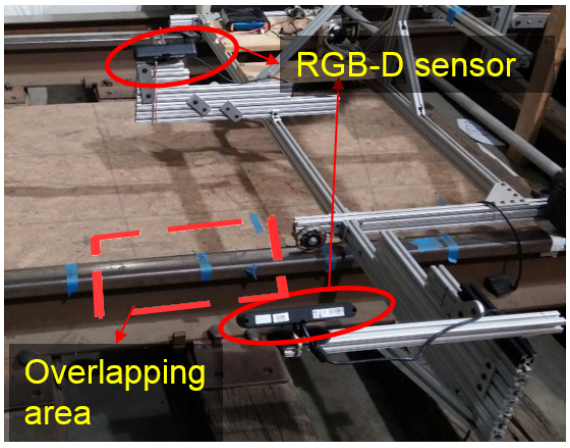
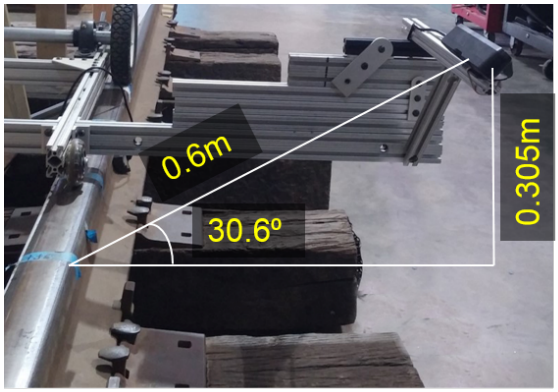


Figure 4.10: Rail enclosure on track

This setup has been tested indoors for the preliminary trials. As seen from Figure 4.11, the RGB-D cameras are placed 0.6m away from the left and right side of the rail for maximum precision. It was confirmed from the design that the cameras would need to be placed at an angle of 30 degrees with respect to the rail to allow maximum view of the blind spot as shown in Figure 4.11b.



(a) Mobile rail profiling area of interest



(b) Orientation of RGB-D camera

Figure 4.11: Indoor rail experiment setup

Motorized operation

One challenge for this enclosure was to make sure it does not slip off the rail tracks during operation. For this case, special track wheels were designed to stay on the tracks as shown in Figure 4.12. A plastic disc was screw-mounted to a pneumatic tire so that it acts as a guard for a single flanged track wheel. The pneumatic tire would be facing the field side of the railway track. As such, we ensure that if the enclosure tends to move left/right, the other track wheel will prevent it from doing so. Each rear wheel of the enclosure is connected to a 10 RPM gearmotor. The low speed of the motor was taken into consideration to make sure there was no motion blur present when the camera was acquiring data.

Server and client communication for multiple camera operation

Since the Kinect v2 can operate only on one computer, another computer had to be installed to account for the second camera. A server and client relationship was established to synchronize the time between the computers. A crossover cable established the local connection between the two computers. The synchronization was performed in software using a robot interface known as Robot Operating System(ROS).

Data storage

The captured RGB and depth images from the RGB-D camera are recorded by ROS and stored in the mini CPU. Since each RGB and depth image contain about 217,088 points and the fact that the RGB-D camera streams at a default rate of 30 FPS, overtime the computer will develop high memory and bandwidth consumption. Bandwidth is calculated using Equation 4.1, where it can be seen that bandwidth has a direct relationship to frame rate. Frame rate has an inverse relationship with respect to the Frame Interval(FI) under a

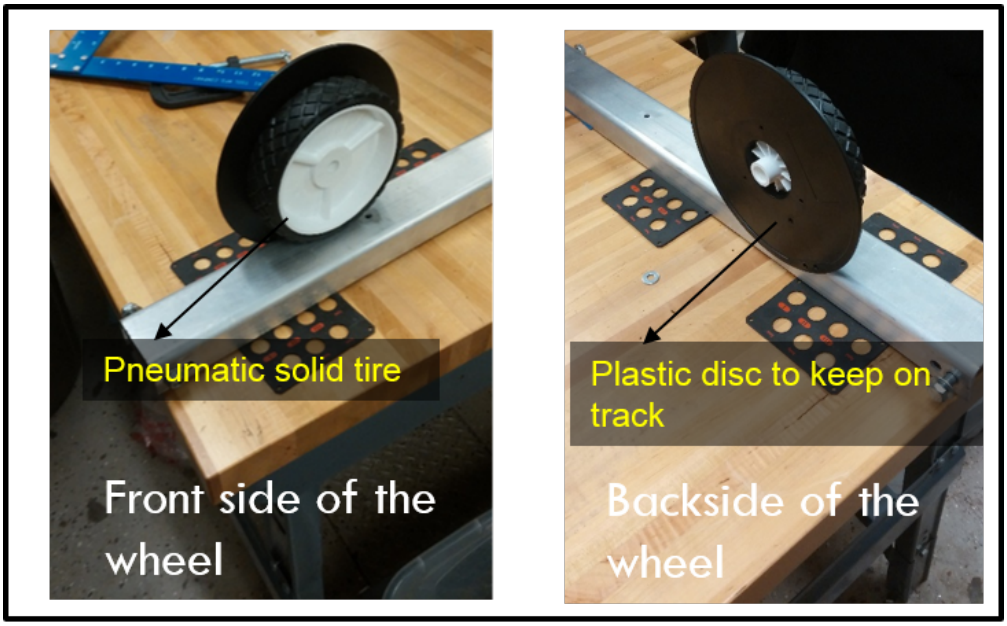


Figure 4.12: Rail wheel

constant velocity. The FI is the distance between each successive frame as shown in Figure 4.13. In this case, the constant velocity is chosen to be small due to motion blur. As such, the frame rate of the RGB-D camera needs to be reduced to such an extent that there is at least a 50% overlap between each successive frame for pairwise matching. With this in mind, the number of data points have been reduced as well as memory consumption.

$$Bandwidth(MB/s) = \frac{brf}{1,0248,576} \tag{4.1}$$

where b is the bytes per pixel, r is the resolution and f is the frame rate.

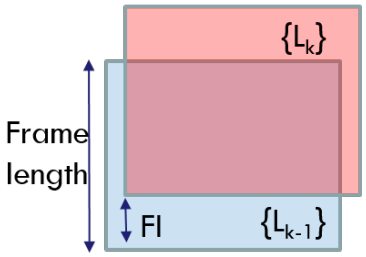


Figure 4.13: Frame interval

Table 4.2: Experimental parameters for long rail profiling

Parameters	Value
Length	100 meters
Speed	0.1 m/s
Frame interval	15cm/frame
RGB-D Frame rate	2 FPS
Location	Huckleberry Trial, Christiansburg, Virginia,USA
Time of day	7 PM EST

4.4 Design of experiment

Keeping the parameters of the developed system in mind, a design of experiment was made to measure long rail profiles in an outdoor environment. Table 4.2 lists the experimental parameters for the long rail profiling experiment. One thing to note here is the slow speed of 0.1m/s. This is due to the fact that the camera is not able to detect features faster than this speed as it induces motion blur and makes it challenging for the camera to match data. Taking this slow speed into account, the frame rate has been reduced from the default rate of 30 FPS to 2FPS for less memory consumption. Figure 4.14 shows the experimental conditions for this railway experiment and the camera's view of the track. The next chapter will describe the results acquired from this experiment.

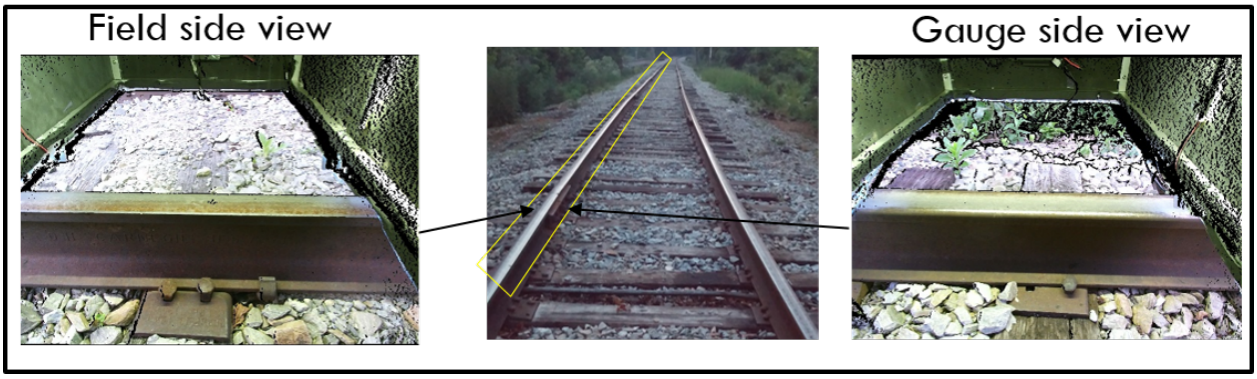


Figure 4.14: Long rail profiling experimental conditions

4.5 Summary

This chapter explained the map-based identification approach used for identifying rail defects and how a tree-based data structure makes the map-based algorithm computationally efficient. A motorized rail profiling system was then designed and implemented based on this approach. It was found that the orientation of the cameras played a big role when it comes to acquiring scans of the blind spot of the rail. Several other parameters such as lighting, motorized operation , multiple camera operation were then discussed in the following sections. Finally, a design of the experiment was shown for long rail profiling. The next chapter will describe the results acquired from this experiment.

Chapter 5

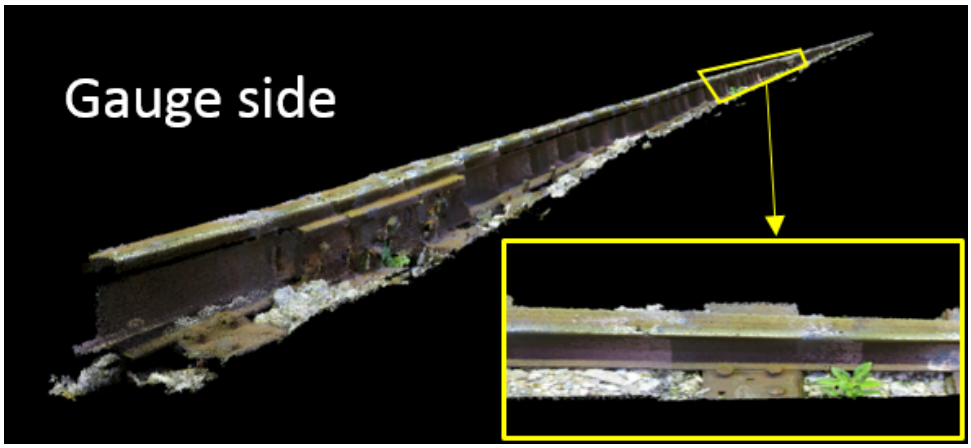
Rail Profiling experimental results

The effectiveness of the proposed system was investigated and evaluated through three steps. The first step investigates the color and depth registration of a 100 m reconstructed rail profile. The second step is where the geometrical properties of the rail components were investigated using multi-spatial resolution techniques. The final step tests the effectiveness of the proposed system by detecting raised spikes and rail components. The proposed approach was performed in a software framework for robots called Robotic Operating System (ROS). All visual results were extracted from an executable script in ROS known as a node.

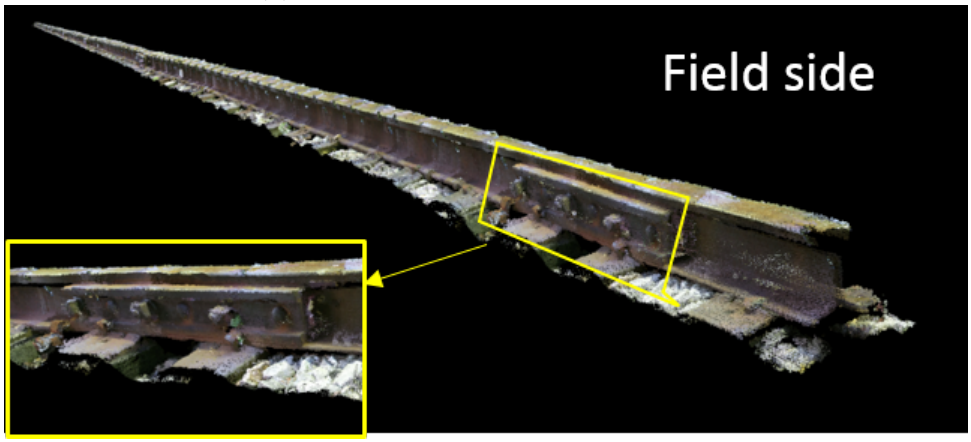
5.1 3-D reconstructed rail profile

Figure 5.1a and Figure 5.1b show the reconstructed RGB-D profiles of the gauge and field side of the rail respectively. As seen from each figure, the RGB-D camera is able to capture the rail structure and its components such as tie plates and ballast. A small portion of the track from Figure 5.1a is focused to show the ability of profiling vegetation along the tracks. Figure 5.1b shows an enlarged view of the fishplate to visually inspect the alignment of the

stitched map. As seen from the figure, no alignment issues can be seen on the fishplate as the bolts and nuts are set in the correct position.



(a) Reconstructed gauge side



(b) Reconstructed field side

Figure 5.1: The reconstructed RGB-D rail profile.

5.1.1 Map building approach reliability

The uncertainty in the depth values were investigated to determine the reliability of the pairwise registration and map building approach. A 20 m section of the railhead was reconstructed for validating the reliability of this approach. Depth measurements of the railhead were collected in the longitudinal direction and compared with a pre-defined parametric ref-

Table 5.1: Experimental parameters for map building uncertainity

Parameters	Value
Length	20 m
Speed	0.1 m/s
Sections	Longitudinal: [Complete, Left, Center, Right]
Number of maps	10

erence profile of the railhead. The deviations of the depth values from the reference profile for each 2m segment of the 20m section were taken into account to investigate the reliability of the approach. The longitudinal profile was further divided into left, center and right sections to see the changes in the transverse direction as shown in Figure 5.2. Ten runs of measurements were performed to create ten reconstructed maps of the railhead. The deviations from each map were then averaged to give a mean of the deviations for the entire longitudinal section as well as the left, center and right sections of the longitudinal profile. The experimental parameters are summarized in Table 5.1, while the sections of the 20m longitudinal profile is shown in Figure 5.2.

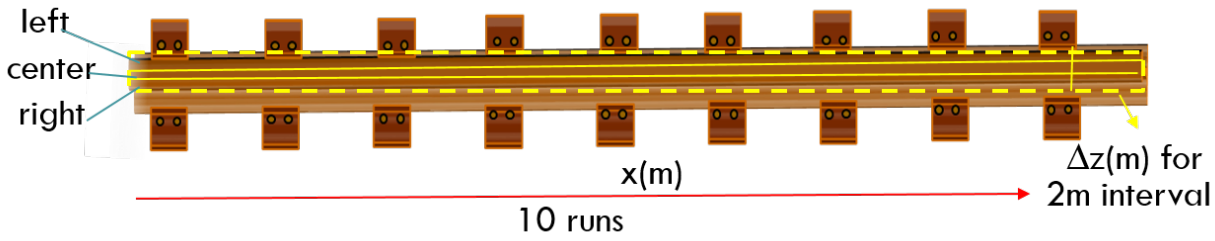


Figure 5.2: Sections of the longitudinal profile

Figure 5.3 shows the standard deviations of the longitudinal profile as well as the left, center and right sections of the longitudinal profile, whereas Figure 5.4 shows all the sections under one graph. Table 5.2 displays the mean of the standard deviations for each section of the ten reconstructed maps. As seen from Figure 5.3, the maximum deviations occur on the

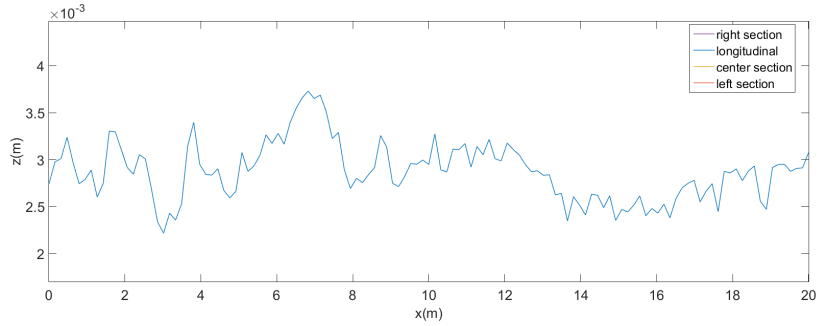
Table 5.2: Standard deviations of depth values of longitudinal profile

Longitudinal Section	Standard deviation(mm)
Whole	2.81
Left	3.12
Center	2.33
Right	3.01

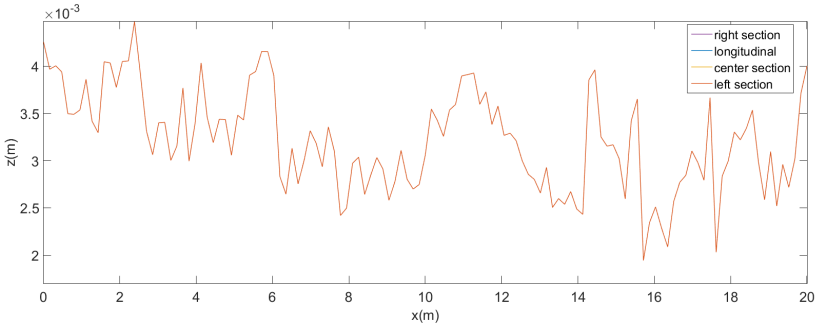
left and right edges of the rail, whereas the minimum deviations occur on the center section of the longitudinal profile. This is due to the fact that there are more overlaps between frames on the center section, whereas the left and right sections are edges of the frames which contains sparse points. Therefore, depth deviations are least when there are multiple scan overlaps.

5.1.2 Pairwise registration errors

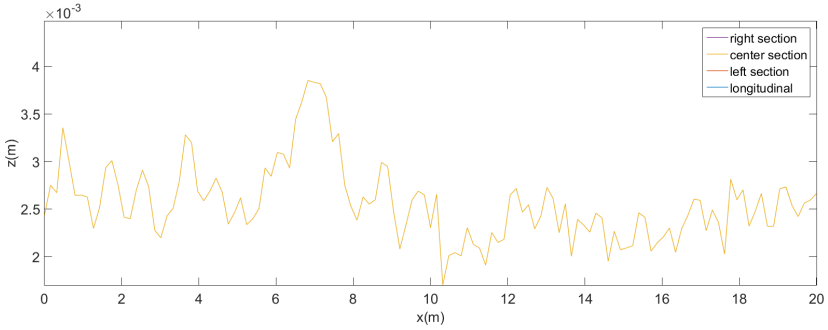
As mentioned in Chapter 3, there will be errors between the target frame and the source frame using the registration algorithms ICP and Normal Distributions Transform. The local errors will exist when the map is converted into a global co-ordinate framework. For this reason, the map building algorithm was run twice to get an initial estimate of the error and remove its effects when the map building algorithm is run a second time. A quantitative study was performed to see the relation of rotation and translation errors in local matching overtime. The pairwise errors were then compared with the corrected global map. Figure 5.5 shows the translation and rotation errors developed after pairwise registration and the effect of the errors after global correction. As seen from Figure 5.5a and Figure 5.5b, the errors between a pair of frames increases monotonically as the 3-D reconstructed rail profile's length increases. This is due to the fact that the errors are subject to local frames and when these local frames are stitched, the errors will accumulate accordingly. The manual



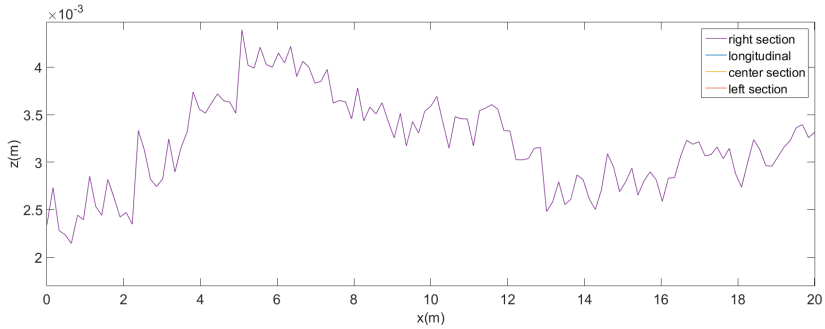
(a) Whole longitudinal profile



(b) Left section of longitudinal profile



(c) Center section of longitudinal profile



(d) Right section of longitudinal profile

Figure 5.3: Standard deviation of depth values of longital profile of railhead

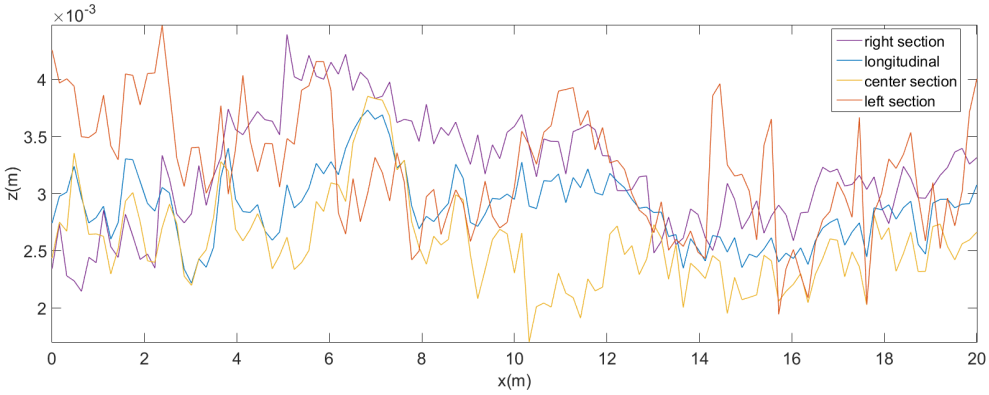
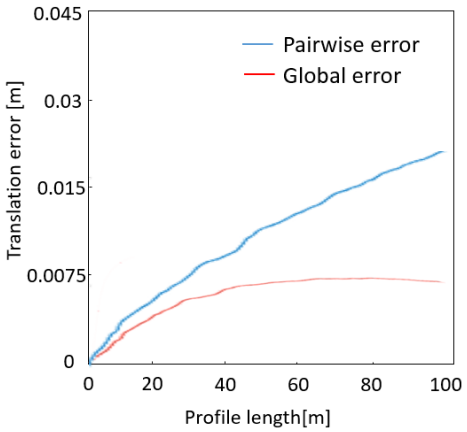


Figure 5.4: Standard deviation of all sections

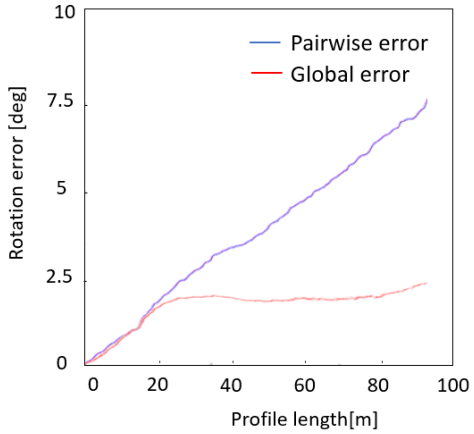
global correction shown in red in Figure 5.5a and Figure 5.5b takes into account the errors between the frames and reduces the translation and rotation errors to an average of 0.0069m and 2.2 degrees respectively. One thing to keep in mind is that even though the errors were accounted for in the global correction manually, these are only used as an estimate to improve the pairwise matching algorithm. A more robust automatic global registration method would be needed in the future to account for all the scans.

5.2 Multi-spatial resolution mesh

Using the 3-D reconstructed rail profile, a mesh was created for the 100 m profile. Figure 5.6 shows the reconstructed mesh for the 100 m shown in different resolutions with the lowest resolution displayed in Figure 5.6a and highest displayed in Figure 5.6c. From Figure 5.6a, we see the top view of the rail profile in mesh format. The irregular bumps seen periodically throughout the profile are the fishplates which are on either side of the rail. Figure 5.6b shows the mesh for a 10 m segment at a higher resolution. In this case, the ties can be distinguished from the ballast as large bumps. The head of the rail is a concentrated mesh which implies the area of overlap between the two scans. Figure 5.6c displays the highest



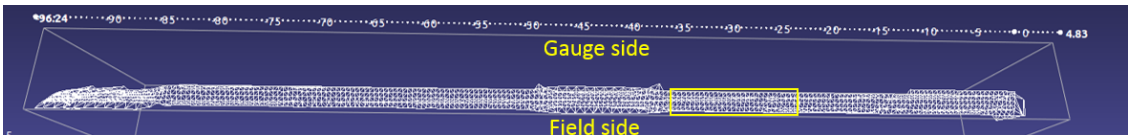
(a) Translation error



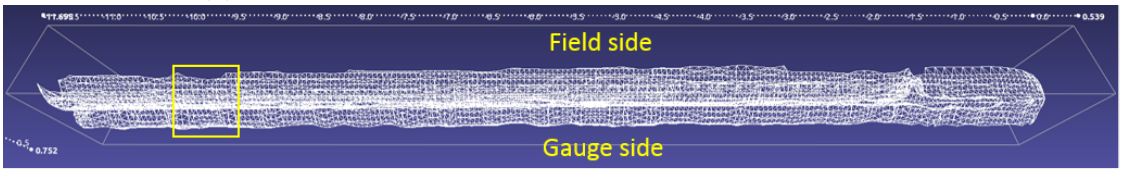
(b) Rotation error

Figure 5.5: Accumulated translation and rotation errors of pairwise registration

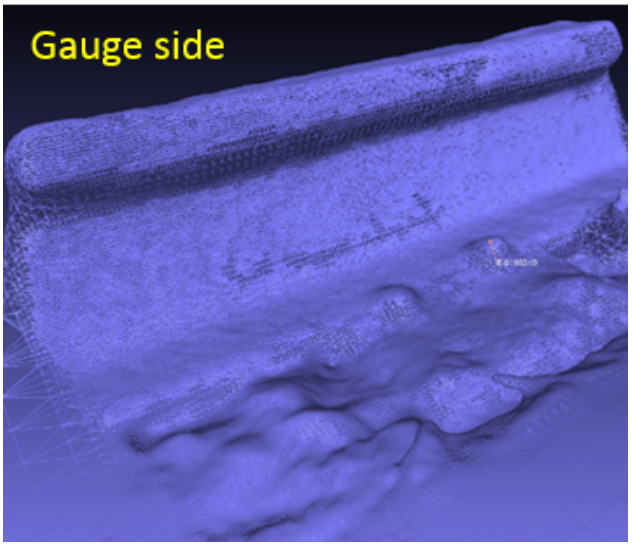
resolution the mesh could achieve with the ability to distinguish raised spikes. The mesh can also be used to geometrically profile obstructions on the track such as vegetation as shown in Figure 5.7. With this in mind, one can inspect large scale deformations such as track deformation on the lowest resolution and small scale defects such as raised spikes on the highest resolution.



(a) Top view of the 100 m reconstructed mesh rail profile



(b) Zoomed in segment of 10 m mesh

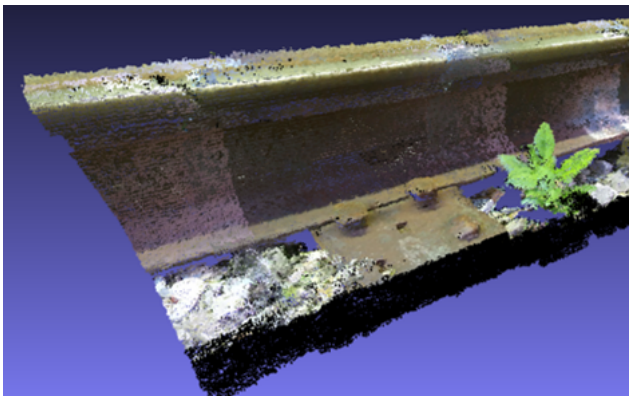


(c) Zoomed in 0.5 m segment of tie plate

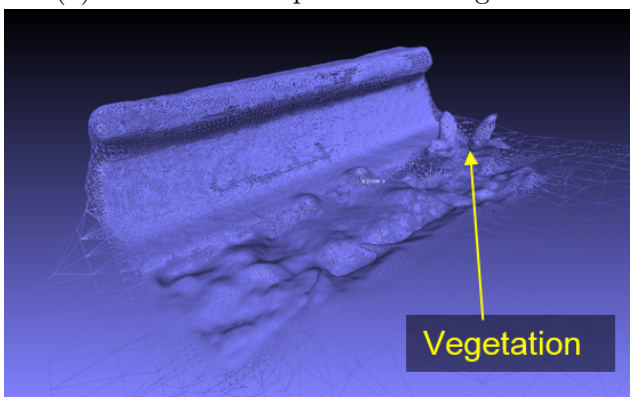
Figure 5.6: The reconstructed mesh of the rail in different resolutions.

5.3 Defect detection

Certain features such as raised or missing spikes and misplaced can be detected with the help of the global map. After the global map was acquired for 100 m, certain artificial faults were introduced onto the components(misplaced ballast on railhead, spikes, etc.) as shown in Figure 5.8a. A new scan was acquired for the defects and compared with the global map. By removing the new scan from the global map, differences in the new scan can be seen as red patches as shown in Figure 5.8b.



(a) Reconstructed profile with vegetation



(b) Reconstructed mesh with vegetation

Figure 5.7: Local measurement results

5.4 Summary

This chapter described the results achieved from the long rail profiling experiment. Firstly, the RGB-D profiles for 100 m were reconstructed. Views of the field and gauge side were shown to show different captured aspects of the rail such as fishplates and vegetation. A quantitative analysis on the map building uncertainty and pairwise registration errors was then conducted. In addition to this, the multi-mesh resolution profiles of the rail were created to show the amount of detail preserved as you increase or decrease resolution. It was determined that ties and ballast become more prominent when zoomed further into the mesh. Finally, the map based identification results were shown. The next chapter will

(a) Scan with raised spikes and misplaced ballast

(b) Spike and ballast change identification

Figure 5.8: Defect detection results

provide the application of RGB-D profiling to the road industry and its results.

Chapter 6

Road profiling by RGB-D Scan matching

This chapter proposes the application of RGB-D profiling to the road industry. Profiling of the roads can come of great use as most tire companies are interested in knowing the roughness of the road. By determining the roughness, one can determine the IRI and can subsequently grade the quality of the road. The next section describes the system design and development of the road profiling system followed by the design of experiment for long road profiling.

6.1 Mobile road profiling system overview

Figure 6.1 shows a system overview of the mobile road profiling system. The system is divided into four sections: Vehicle Enclosure, On-Board Vehicle, Off-Board Mapping and Graphical User Interface. As a brief introduction, the vehicle enclosure houses the RGB-D sensors which transmit the data to the mini-CPU positioned on board the vehicle. The data

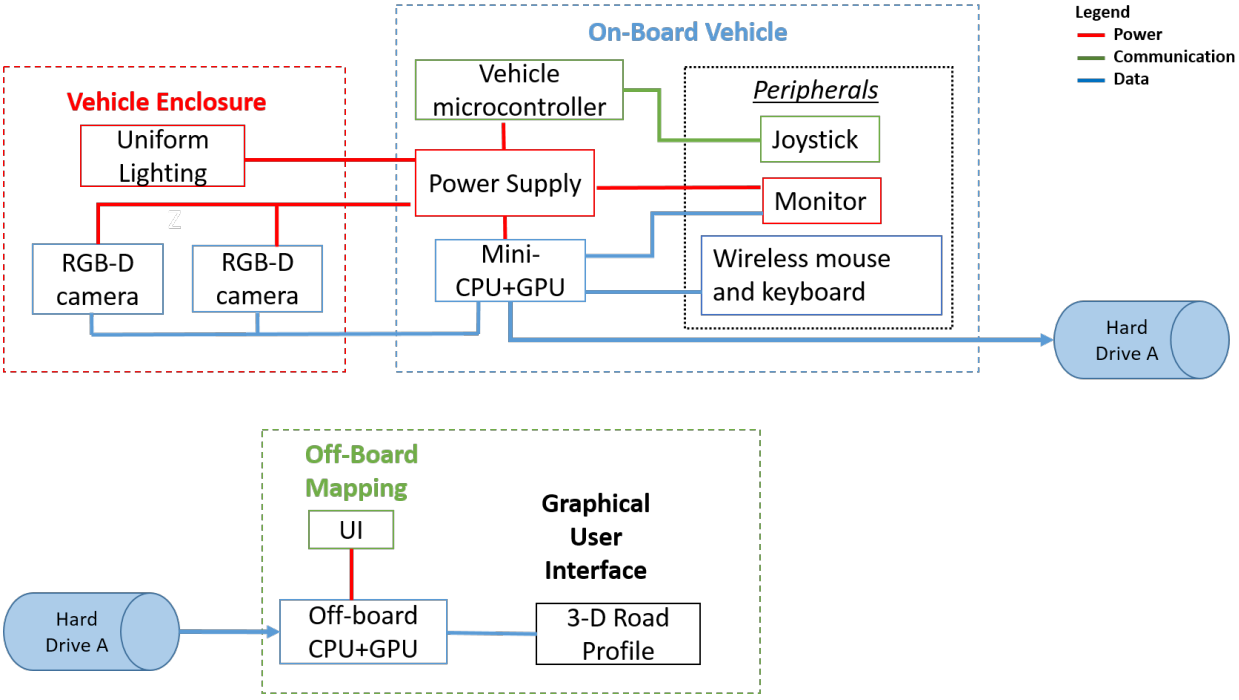


Figure 6.1: Mobile road profiling system

from the mini-CPU is then stored in an external hard drive for off-board mapping using a desktop workstation. Finally, the 3-D constructed road profile is shown in a Graphical User Interface(GUI). Each of these sections will be explained briefly.

6.1.1 Vehicle Enclosure

Figure 6.2 shows the concept drawing for the mobile road profiling enclosure. The road profiling enclosure had a similar set-up as the rail profile enclosure with two key differences. One of the differences was that fact that the enclosure did not need to have four wheels as this could be attached to the back of an electric vehicle using a hitch. The second difference was the orientation of the RGB-D cameras. In this enclosure, the cameras are facing downwards normal to the surface of the road. With this in mind, an RGB-D camera was mounted to the top right and top left halves of the enclosure as shown in Figure 6.3.

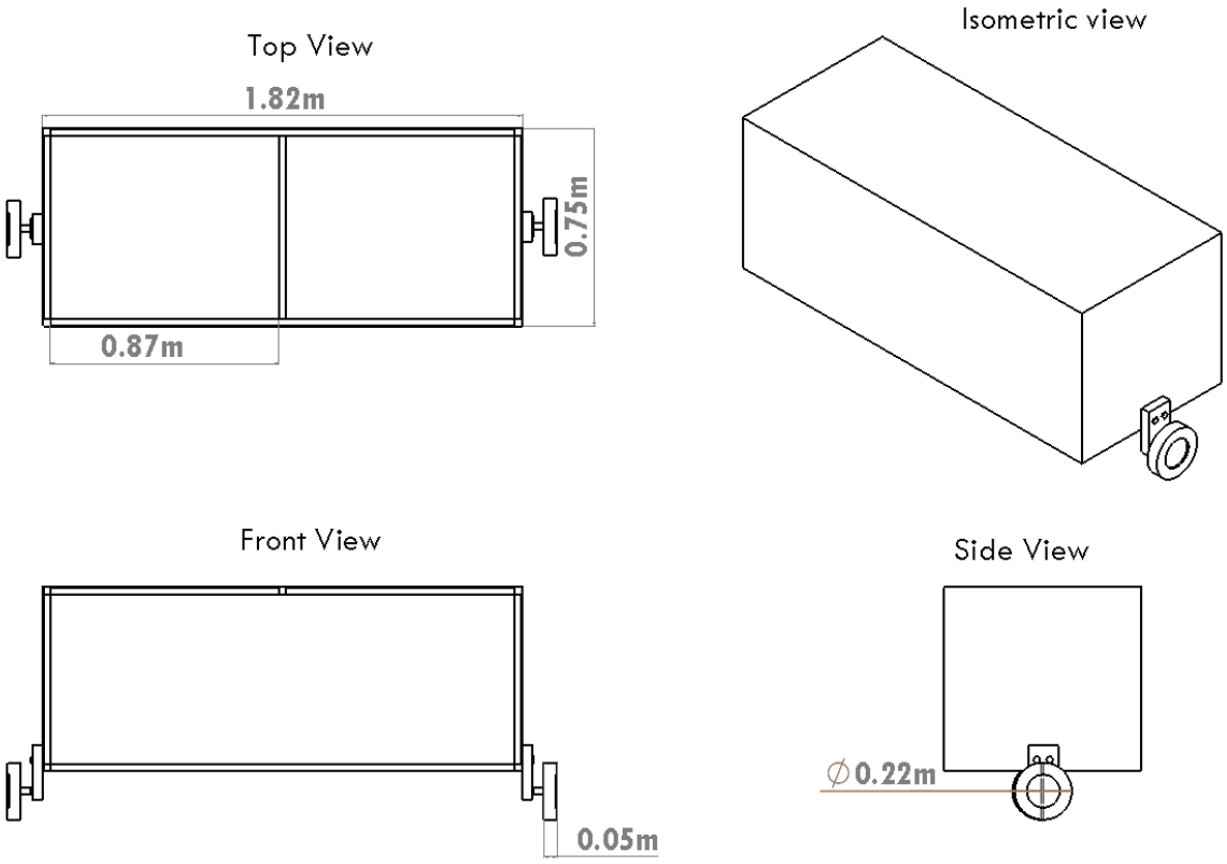


Figure 6.2: Road enclosure design drawing

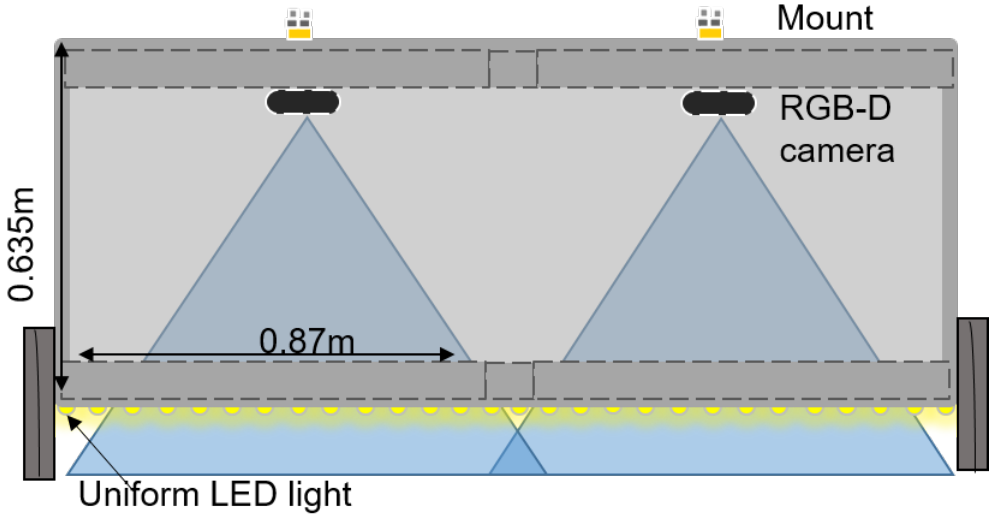


Figure 6.3: Blown up view of enclosure

6.1.2 On-Board Vehicle

The three main areas that influence the on-board vehicle system are as follows: Power supply, teleoperation and data storage.

Power schematic

To provide long runs of experiments, the vehicle is powered electrically by heavy duty batteries. These batteries also power the CPU and the lights by the means of an inverter. The inverter converts the DC voltage supplied by the batteries to AC voltage for the lights and the mini-CPU. The constraint here is that the voltage must be distributed equally to the mini CPU and the lights as shown in Figure 6.4.

Teleoperation

One constraint for ICP as mentioned in Chapter 3.2.1 is that it needs a very good initial estimate for a transformation matrix to avoid being stuck in a local minimum. For this case, some form of odometry is needed to measure the speed of the vehicle. As such, the electric vehicle is operated by a remote controller for constant throttle. With an appropriate Inertial Measurement Unit(IMU) and an odometer, it is possible to measure the position of the vehicle. Using this measurement, it is possible to give an initial estimate for the transformation matrix for each frame. The remote controller is connected to the black box which houses the vehicle mainframe computer.

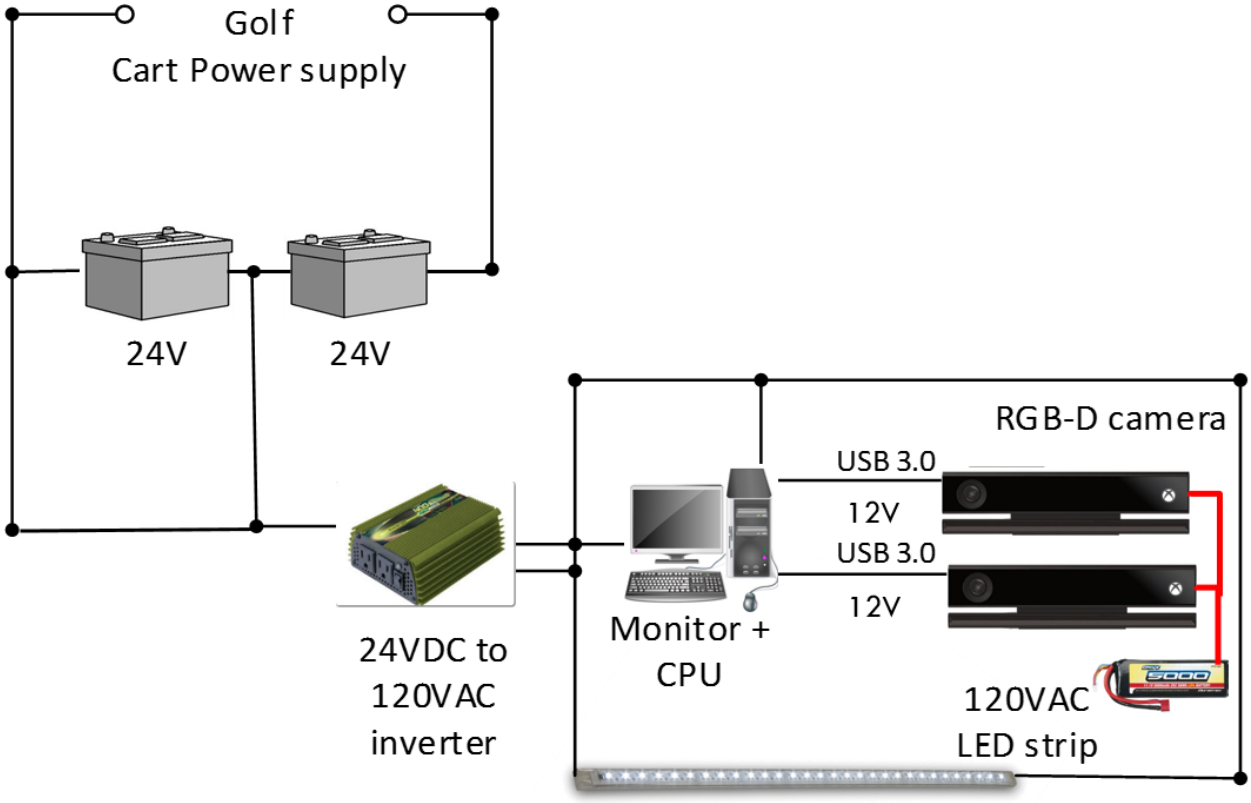
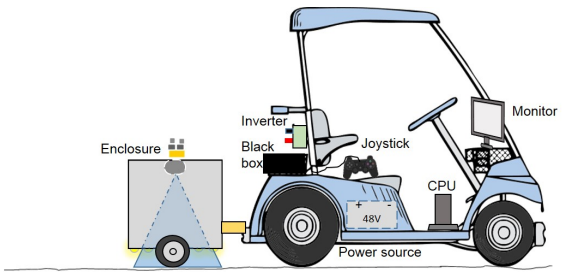


Figure 6.4: Power supply schematic



(a) On-Board System concept

(b) Developed On-Board system

Figure 6.5: On-Board System concept and developed version.

6.1.3 Design of experiment

The long road profiling experimental parameters were chosen on the same basis as the long rail profiling experiment. Table 6.1 lists the parameters used for this experiment whereas

Table 6.1: Experimental parameters for long road profiling

Parameters	Value
Length	100 meters
Speed	0.4 m/s
Frame interval	25cm/frame
RGB-D Frame rate	5 FPS
Road material	Coarse Concrete
Time of day	830 PM EST

Figure 6.6 shows the experimental conditions of the test site.

6.2 Experimental results

The application of RGB-D profiling for road elevation measurements was evaluated in three steps. Firstly, the roughness of the road was extracted by using the information from the reference plane as stated in the Literature Review section. The second step examines the geometrical properties of the road by visualizing the RGB-D profile as well as the multi-spatial resolution meshes. The final step is to evaluate local measurement results and investigate scan matching reliability. The proposed approach was performed in a software framework for robots called Robotic Operating System (ROS). All visual results were extracted from an executable script in ROS known as a node.

6.2.1 Roughness of the profile

As mentioned in the Literature Review and the Three Dimensional Reconstruction and Offline Mapping section, the reference plane will need to be extracted from the scan-matched data. Figure 6.7a shows a small portion of the scan matched data of a road with gravel. For

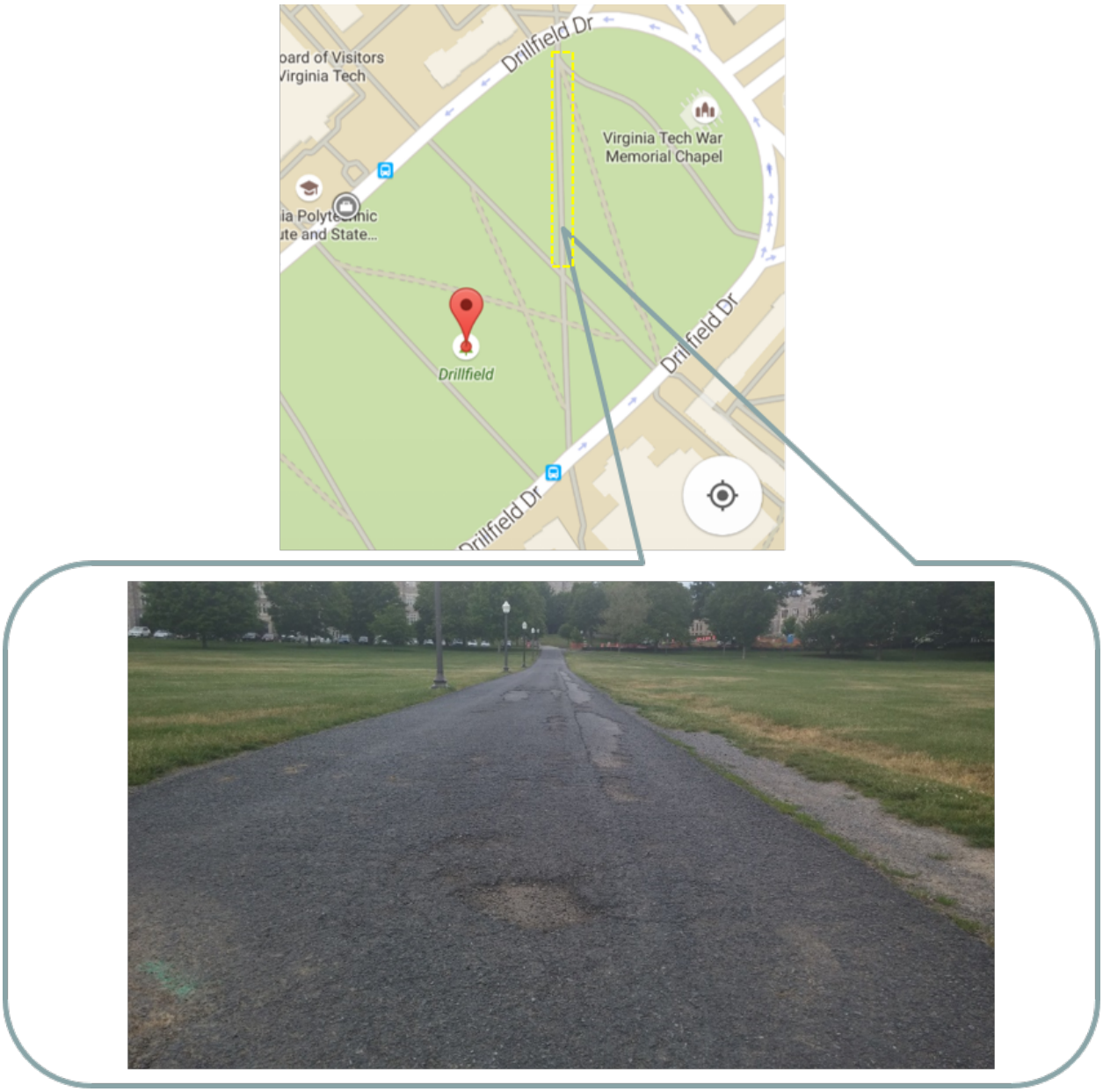


Figure 6.6: Road test site condition

the reference, a road surface with no gravel was taken into account. A plane is fitted to the road with no gravel to create a reference plane as shown in Figure 6.7b. The equation for the fitted plane is shown below in Equation 6.1

$$z(x, y) = 1.778 - 0.02767x - 0.042y \tag{6.1}$$

where x and y are the transverse and longitudinal lengths of the road in meters respectively and z is the height from the camera to the road in meters.

With these two figures in mind, the depth or z values of the reference profile are subtracted from the scan matched data to give the depth relative to the reference as shown in Figure 6.7c. The relative measurements consequently gives the roughness of the road. This technique is used for each scan to create long road profiles. One thing to note here is that reference plane is a consequence of using inertial sensors and must be artificially created and removed from the data to get a road profile. Further investigation must be performed in the future to validate this reference profile against other inertial sensors such as laser line scanners.

6.2.2 RGB-D road profile and multi-resolution results

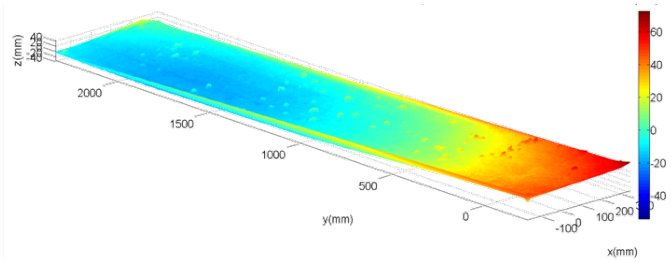
Using the roughness for each scan portrayed in the previous step, the RGB-D profile is constructed for 100 meters as shown in Figure 6.8. To investigate the geometrical properties of this RGB-D profile, multi-spatial resolutions of the road are created as shown in Figure 6.9. As seen in Figure 6.9a, a 50 centimeters(cm) grid resolution was used to display the large amount of data. Features such as potholes are evident as they can be seen as distinct circles. As the mesh is zoomed further into a 10 m segment with 1 cm resolution, we see cracks forming longitudinally along the road. If we zoom in even further into a 1 meter segment of 1 millimeter(mm) resolution, the geometrical properties of the pothole can then be clearly established.

6.2.3 Local measurement results

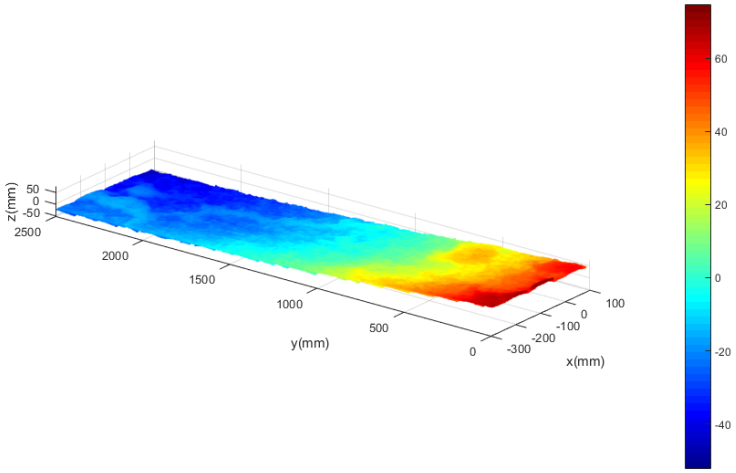
Figure 6.10 visualizes the scan matching process by comparing the real environment with the reconstructed one. Figure 6.10a shows the scene that needs to be reconstructed. Figure 6.10b shows the reconstructed gravel and rocks in a 3-D format. A gray-scale colorbar has been used for visualization purposes with white color indicating high z values, whereas black color indicates low z values. The reconstructed gravel scene is a combination of two scans as shown in Figure 6.10b. The purpose of showing the comparison with the real and reconstructed road profiles is to measure the scan matching reliability by qualitative evaluation. In terms of qualitative sense, the reconstructed scene offers high reliability as the position of the rocks on the reconstructed scene are positioned in the same orientation as the real world scene.

6.3 Summary

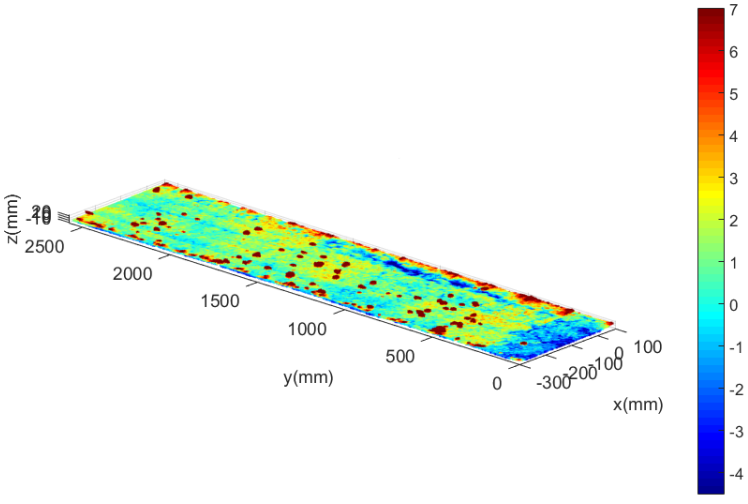
This chapter described the applicability of road profiling using RGB-D cameras. Firstly, the system design and development were investigated followed by the experimental parameters of the long road profiling experiment. Quantitative results were shown for roughness of the test site. Furthermore, the multi-mesh resolution profiles of the road were created to show the amount of detail preserved as you increase or decrease resolution. Finally, some local measurement results were shown for gravel on the road. The next chapter will summarize the conclusions brought about from RGB-D profiling as an inspection method.



(a) Original scan matched data



(b) Fitted reference plane



(c) Roughness of road profile

Figure 6.7: The extraction of roughness data

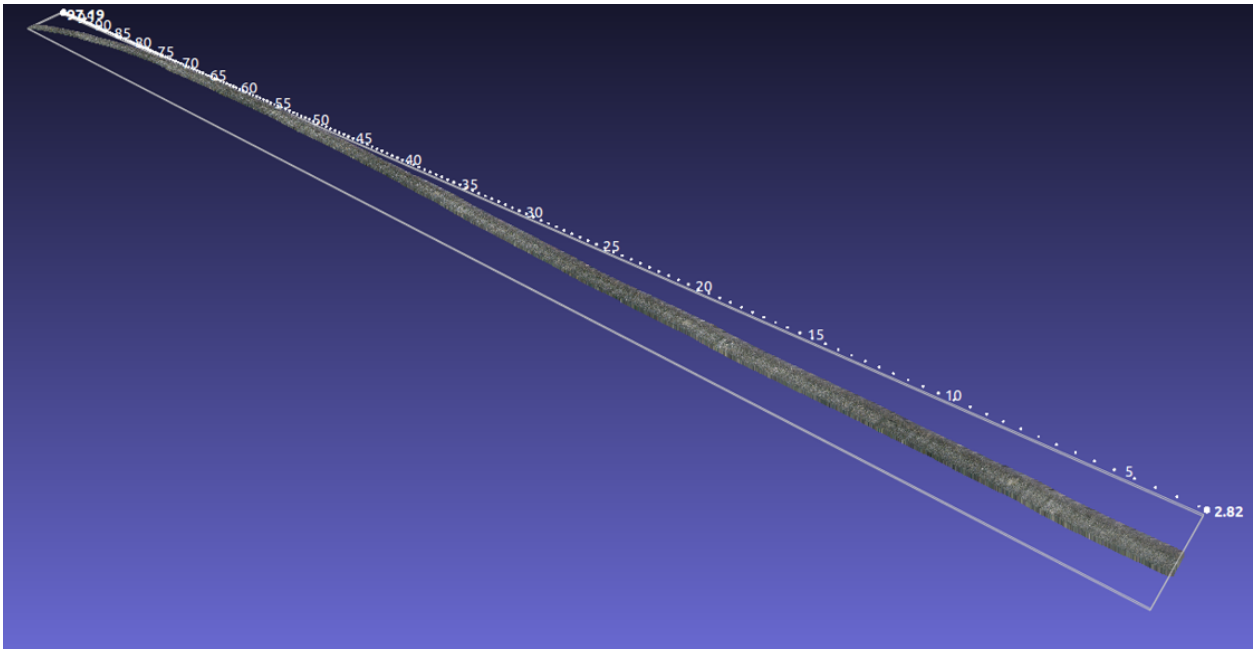
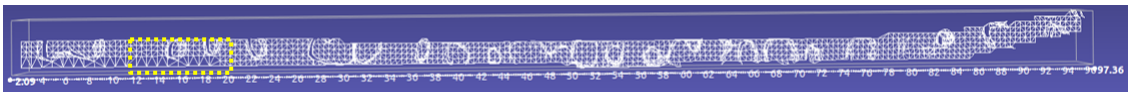
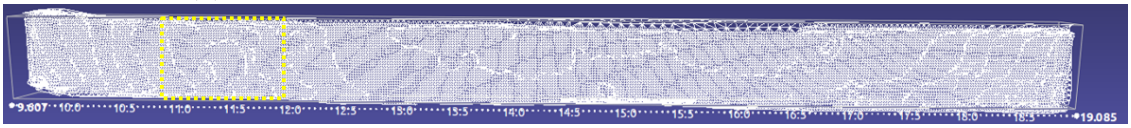


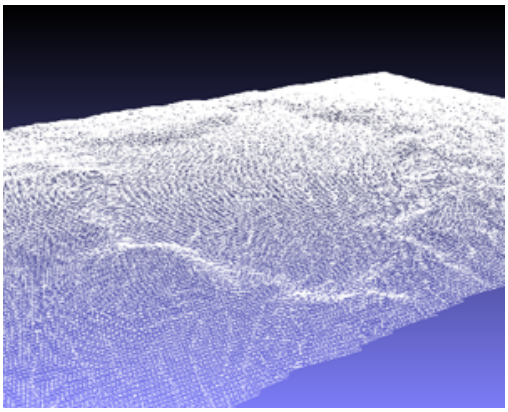
Figure 6.8: Perspective view of the RGB-D profile



(a) Top view of 100 meters reconstructed mesh road profile



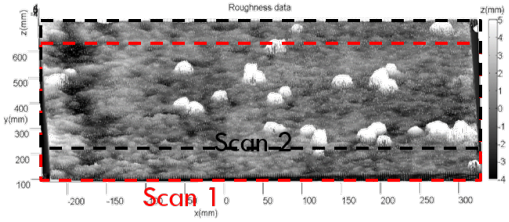
(b) Zoomed in segment of 10 meters mesh



(c) Zoomed in 1 meter segment of pothole

Figure 6.9: The reconstructed mesh of the road in different resolutions.

(a) Gravel on road



(b) Reconstructed gravel

Figure 6.10: Local measurement scans of gravel.

Chapter 7

Conclusions and future work

7.1 Conclusions

This thesis has presented an approach for a railroad and road inspection method using RGB-D cameras. The inexpensive RGB-D cameras which can reconstruct color and geometric information are able to see consecutive scans due to its large field of view unlike past methods such as laser scanning or computer vision. The 3-D reconstructed global map generated from the proposed approach is used as a map-based identification method to recognize rail defects which reduces computation time unlike Computer Vision. The multiple synchronization of these RGB-D cameras in an outdoor environment was made possible using an enclosure housing system for these cameras. Finally, the applicability of this scan matching approach was also demonstrated for the road industry.

The effectiveness of the proposed inspection method for railroad profiling was investigated and evaluated through three steps. The first step investigates the color and depth registration of the map by inspecting a 100 m reconstructed rail profile. The results showed

that the RGB-D camera was able to capture the rail structure and its components such as tie plates and ballast with no ambiguities. Reliability of this approach was investigated by calculating the uncertainty of the depth values of the 3-D reconstructed profile for multiple runs. The pairwise matching algorithm proved most reliable where there were multiple overlaps such as the center of the railhead. The pairwise registration errors were then investigated. The results showed that the monotonically increasing translation and rotation errors will reduce significantly after global correction. The second step investigates the geometrical properties of the rail components were investigated using multi-spatial resolution techniques. It was determined that ties and ballast which were seen as irregular bumps become more prominent when zoomed further into the mesh. Thus, the preservation of the data has been validated using multi-spatial resolution techniques. The final step investigates the applicability of the proposed system was by detecting of raised spikes and rail components. By removing the new scan from the global map, raised spikes and misplaced ballast could be seen as red patches in the new scan. This proves the applicability of the proposed approach for fault detection without expensive computational effort.

The applicability of the proposed inspection method for road profiling was also investigated and evaluated through three steps. Firstly, quantitative and qualitative results were shown for the roughness of a small portion of the road used in the experiment. It was determined that using this roughness result could give us more information about the IRI. Furthermore, the multi-mesh resolution profiles of the road were created to show the amount of detail preserved as the resolution was increased or decreased. Features such as potholes are evident as they can be seen as distinct circles on the lowest resolution. As the mesh is zoomed further into a 10 m segment with 1 cm resolution, cracks were seen forming along the road. As the mesh is zoomed in further into a 1 m segment of 1 mm resolution, the geometrical properties of the pothole could be clearly established. Finally, some local

measurement results were shown for gravel on the road. In terms of qualitative sense, the reconstructed local measurement offered high reliability as the position of the rocks on the reconstructed scene are positioned in the same orientation as the real world scene.

7.2 Future work

This paper has focused on a new inexpensive inspection method for the rail and road industry using RGB-D camera technology. However, there are some limitations of this technology and much work is still left open. One of the limitations that needs to be investigated is the effect of speed for this profiling approach. As previously stated in this thesis, the RGB-D camera would not be able to match the same features in consecutive scans correctly if there is high speed which causes motion blur. An implementation of an a Field Programmable Gate Array to control the shutter speed of the camera could be helpful with regards to the speed issue. Another work that needs to be investigated is the automatic removal of error from the global map . In this thesis, the errors in the global map were corrected manually after taking into account the errors from the first run. However. these errors can be removed automatically and with more precision using an automatic global optimizer algorithm. It is the author's opinion that once the speed and global registration issues have been solved, the proposed map-based identification technique can be used as an inexpensive inspection method for both the rail and road industry.

Bibliography

[1] MC Duffy. The standard rail section, transverse fissures and reformed mill practice, 1911–1955. *Journal of Mechanical Working Technology*, 4(3):285–305, 1980.

[2] HD Grohmann, Klaus Hempelmann, and Arnold Groß-Thebing. A new type of rcf, experimental investigations and theoretical modelling. *Wear*, 253(1):67–74, 2002.

[3] DF Cannon, K-O EDEL, SL Grassie, and K Sawley. Rail defects: an overview. *Fatigue & Fracture of Engineering Materials & Structures*, 26(10):865–886, 2003.

[4] John E Garnham, David I Fletcher, Claire L Davis, and Francis J Franklin. Visualization and modelling to understand rail rolling contact fatigue cracks in three dimensions. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 225(2):165–178, 2011.

[5] Vassilios A Profillidis. *Railway engineering*. 2000.

[6] Vassilios A Profillidis. *Railway management and engineering*. Ashgate Publishing, Ltd., 2014.

[7] Feilong Liu, Nicholas Dembski, Ahmed Soliman, Giorgio Rizzoni, Brian Thompson, and Bowie Vaughn. A kalman-filter-based multi-sensor terrain profile measurement system: principle, implementation and validation. In *SPIE Defense and Security Symposium*, pages 69650J–69650J. International Society for Optics and Photonics, 2008.

[8] Alfredo Cigada, Federico Mancosu, Stefano Manzoni, and Emanuele Zappa. Laser-triangulation device for in-line measurement of road texture at medium and high speed. *Mechanical Systems and Signal Processing*, 24(7):2225–2234, 2010.

[9] Glenn A Fawcett, Christopher R Bennett, and David G Ollerenshaw. Flexible low-cost data capture technology for road networks. In *Proceedings of the 4th International Conference on Road and Airfield Pavement Technology*, 2002.

[10] John Laurent, Jean François Hébert, Daniel Lefebvre, and Yves Savard. Using 3d laser profiling sensors for the automated measurement of road surface conditions. In *7th RILEM International Conference on Cracking in Pavements*, pages 157–167. Springer, 2012.

[11] Feng Li. A methodology for characterizing pavement rutting condition using emerging 3d line laser imaging technology. 2012.

[12] Clemens Mair and Saeed Fararooy. Practice and potential of computer vision for railways. 1998.

[13] Esther Resendiz, LF Molina, JM Hart, J Riley Edwards, Steven Sawadisavi, Narendra Ahuja, and CPL Barkan. Development of a machine-vision system for inspection of railway track components. In *12th World Conference on Transport Research, Lisbon, Portugal*, 2010.

[14] Ying Li, Charles Otto, Norm Haas, Yuichi Fujiki, and Sharath Pankanti. Component-based track inspection using machine-vision technology. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval*, page 60. ACM, 2011.

[15] Hoang Trinh, Norman Haas, Ying Li, Charles Otto, and Sharath Pankanti. Enhanced rail component detection and consolidation for rail track inspection. In *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, pages 289–295. IEEE, 2012.

[16] E Stella, P Mazzeo, M Nitti, G Cicirelli, A Distante, and TD Orazio. Visual recognition of missing fastening elements for railroad maintenance. In *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on*, pages 94–99. IEEE, 2002.

[17] J Riley Edwards. *Advancements in railroad track inspection using machine-vision technology*. PhD thesis, University of Illinois at Urbana-Champaign, 2009.

[18] Francescomaria Marino, Arcangelo Distante, Pier Luigi Mazzeo, and Ettore Stella. A real-time visual inspection system for railway maintenance: automatic hexagonal-headed bolts detection. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(3):418–428, 2007.

[19] Coenraad Esveld. Modern railway track. 2001.

[20] Luis Ferreira and Andrew Higgins. Modelling aspects of rail track maintenance scheduling. American Society of Civil Engineers, 1998.

[21] AB Allison. Sperry rail service. *Bull Natl Railway Historical Soc*, 33(6), 1968.

[22] Robin Clark. Rail flaw detection: overview and needs for future developments. *NDT & E International*, 37(2):111–118, 2004.

[23] Robert E Green. Non-contact ultrasonic techniques. *Ultrasonics*, 42(1):9–16, 2004.

[24] Jianzheng Cheng and Leonard J Bond. Assessment of ultrasonic ndt methods for high speed rail inspection. In *41ST ANNUAL REVIEW OF PROGRESS IN QUANTITATIVE NONDESTRUCTIVE EVALUATION: Volume 34*, volume 1650, pages 605–614. AIP Publishing, 2015.

[25] Donatella Cerniglia, Gregory Garcia, Semih Kalay, and Fabrizio Prior. Application of laser induced ultrasound for rail inspection. In *Proceedings of the World Congress in Railway Research. Montreal, Canada*, 2006.

[26] Richard M White. Generation of elastic waves by transient surface heating. *Journal of Applied Physics*, 34(12):3559–3567, 1963.

[27] S Krishnaswami. Theory and applications of laser-ultrasonic techniques. *Ultrasonic nondestructive evaluation: Engineering and biological material characterization*, pages 435–494, 2003.

[28] SHANT Kenderian, BORO B Djordjevic, D Cerniglia, and G Garcia. Dynamic railroad inspection using the laser-air hybrid ultrasonic technique. *Insight-Non-Destructive Testing and Condition Monitoring*, 48(6):336–341, 2006.

[29] JE Garnham and JH Beynon. The early detection of rolling-sliding contact fatigue cracks. *Wear*, 144(1):103–116, 1991.

[30] Rainer Pohl, A Erhard, H-J Montag, H-M Thomas, and H Wüstenberg. Ndt techniques for railroad wheel and gauge corner inspection. *NDT & E International*, 37(2):89–94, 2004.

[31] Zenglu Song, Tsutomu Yamada, Hideki Shitara, and Yasushi Takemura. Detection of damage and crack in railhead by using eddy current testing. *Journal of Electromagnetic Analysis and Applications*, 2011, 2011.

[32] Javier García-Martín, Jaime Gómez-Gil, and Ernesto Vázquez-Sánchez. Non-destructive techniques based on eddy current testing. *Sensors*, 11(3):2525–2565, 2011.

[33] C.J. E, H.D. F, and P.R. A. Magnetic particle inspection apparatus and method, January 24 1950. URL http://www.google.com/patents/US2495545. US Patent 2,495,545.

[34] René Heyder and Gregor Girsch. Testing of hsh® rails in high-speed tracks to minimise rail damage. *Wear*, 258(7):1014–1021, 2005.

[35] Yong Li, John Wilson, and Gui Yun Tian. Experiment and simulation study of 3d magnetic field sensing for magnetic flux leakage defect characterisation. *NDT & E International*, 40(2):179–184, 2007.

[36] Steven Sawadisavi. *Machine-vision inspection of railroad track*. PhD thesis, University of Illinois at Urbana-Champaign, 2009.

[37] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.

[38] Alejandro Lorente, David Llorca, Miguel Velasco, José García, and Fernando Domínguez. Detection of range-based rail gage and missing rail fasteners: Use of high-resolution two-and three-dimensional images. *Transportation Research Record: Journal of the Transportation Research Board*, (2448):125–132, 2014.

[39] Zhifeng Zhang, Qibo Feng, Zhan Gao, Cuifang Kuang, Cheng Fei, Zhang Li, and Jinyun Ding. A new laser displacement sensor based on triangulation for gauge real-time measurement. *Optics & laser technology*, 40(2):252–255, 2008.

[40] Zhen Liu, Junhua Sun, Heng Wang, and Guangjun Zhang. Simple and fast rail wear measurement method based on structured light. *Optics and Lasers in Engineering*, 49 (11):1343–1351, 2011.

[41] Jin-ren Lu and Yuan-qing HUANG. Laser triangulation method for surface measurement. *Journal-Xiamen University Natural Science*, 43(1):50–53, 2004.

[42] Gerald F Marshall and Glenn E Stutz. *Handbook of optical and laser scanning*. CRC Press, 2011.

[43] Michael W Sayers and Steven M Karamihas. The little book of profiling. *the Regent of the University of Michigan*, 2, 1998.

[44] Michael Sayers and Steven Karamihas. Estimation of rideability by analyzing longitudinal road profile. *Transportation Research Record: Journal of the Transportation Research Board*, (1536):110–116, 1996.

[45] Rohan Perera, Starr Kohn, and Sohila Bemanian. Comparison of road profilers. *Transportation Research Record: Journal of the Transportation Research Board*, (1536):117–124, 1996.

[46] Hans Prem. A laser-based highway-speed road profile measuring system. *Vehicle System Dynamics*, 17(sup1):300–304, 1988.

[47] Arturo González, Eugene J O'brien, Y-Y Li, and K Cashell. The use of vehicle acceleration measurements to estimate road roughness. *Vehicle System Dynamics*, 46(6): 483–499, 2008.

[48] Michael W Sayers, Thomas D Gillespie, and AV Queiroz. The international road roughness experiment. establishing correlation and a calibration standard for measurements. Technical report, 1986.

[49] Thomas D Gillespie, Michael W Sayers, and Leonard Segel. Calibration of response-type road roughness measuring systems. *NCHRP Report*, (228), 1980.

[50] Andreas Kolb, Erhardt Barth, Reinhard Koch, and Rasmus Larsen. Time-of-flight sensors in computer graphics. In *Proc. Eurographics (State-of-the-Art Report)*, volume 6, 2009.

[51] Carlo Dal Mutto, Pietro Zanuttigh, and Guido M Cortelazzo. *Time-of-flight cameras and microsoft Kinect*. Springer Science & Business Media, 2012.

[52] Thiemo Wiedemeyer. IAI Kinect2. https://github.com/code-iai/iai_kinect2, 2014 – 2015. Accessed June 12, 2015.

[53] Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1):116–124, 2013.

[54] John Sell and Patrick O'Connor. The xbox one system on a chip and kinect sensor. *IEEE Micro*, (2):44–53, 2014.

[55] Annette Payne, Andy Daniel, Anik Mehta, Bradley Thompson, Cyrus S Bamji, Dane Snow, Hirotaka Oshima, Larry Prather, Mike Fenton, Lou Kordus, et al. 7.6 a 512× 424 cmos 3d time-of-flight image sensor with multi-frequency photo-demodulation up to 130mhz and 2gs/s adc. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pages 134–135. IEEE, 2014.

[56] Timo Breuer, Christoph Bodensteiner, and Michael Arens. Low-cost commodity depth sensor comparison and accuracy analysis. In *SPIE Security+ Defence*, pages 92500G–92500G. International Society for Optics and Photonics, 2014.

[57] Nima Rafibakhsh, Jie Gong, Mohsin K Siddiqui, Chris Gordon, and H Felix Lee. Analysis of xbox kinect sensor data for use on construction sites: depth accuracy and sensor interference assessment. In *Constitution research congress*, pages 848–857, 2012.

[58] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.

[59] Paul Scovanner, Saad Ali, and Mubarak Shah. A 3-dimensional sift descriptor and its application to action recognition. In *Proceedings of the 15th international conference on Multimedia*, pages 357–360. ACM, 2007.

[60] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[61] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2):119–152, 1994.

[62] Peter Biber and Wolfgang Straßer. The normal distributions transform: A new approach to laser scan matching. In *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 3, pages 2743–2748. IEEE, 2003.

[63] Martin Magnusson, Achim Lilienthal, and Tom Duckett. Scan registration for autonomous mining vehicles using 3d-ndt. *Journal of Field Robotics*, 24(10):803–827, 2007.