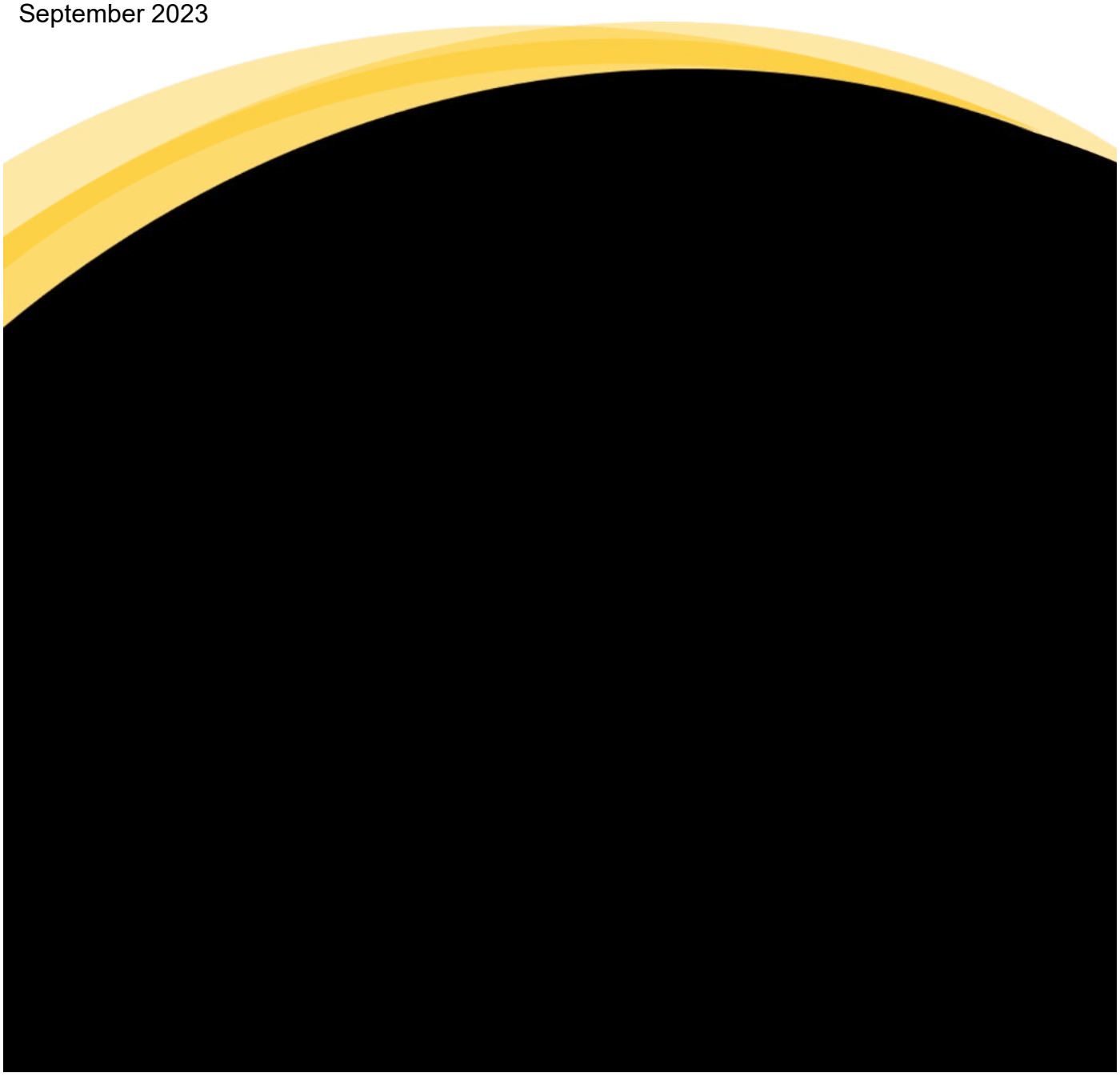
Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction

Project Overview

September 2023



Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction

Project Overview

Prepared by

Zhen Peng, Jun Li, Wensu Chen, Robert Lee, Atif Mansoor, Sergio Banchero, Chao Sun & Sharon Biermann

Keywords

Computer vision, IoT sensor, machine learning, bridge health monitoring

Version control

Final

Project No

iMOVE Project 1-046, Milestone 3

Project steering committee

Raquib Hossain, Main Roads Western Australia

Steve Atkinson, Main Roads Western Australia

Sebastian Davies-Slate, WALGA

Sharon Biermann, The University of Western Australia

Jun Li, Curtin University

Wensu Chen, Curtin University

Chao Sun, The University of Western Australia

Atif Mansoor, The University of Western Australia

Acknowledgment

This research is funded by iMOVE CRC and supported by the Cooperative Research Centres program, an Australian Government initiative.

About PATREC

The Planning and Transport Research Centre (PATREC) is a collaboration between the Government of Western Australia and local universities, constituted to conduct collaborative, applied research and teaching in support of policy in the connected spaces of transport and land use planning. The collaborating parties are: The University of Western Australia, Curtin University, Edith Cowan University, Department of Transport, Main Roads Western Australia, Western Australian Planning Commission and the Western Australian Local Government Association.

Publisher

Planning and Transport Research Centre
The University of Western Australia (M087)
35 Stirling Highway, Crawley, WA 6009
+61 8 6488 3385
patrec@uwa.edu.au

Table of Contents

Outline

This project overview provides a summary of the need for the research, opportunities to address the need, aims and objectives, approach, key findings and recommendations of the project. Details of particular components of the research are provided in technical reports, appended to this summary.

Research need

Regular maintenance programs are essential for the long-term preservation of bridges, and health monitoring plays a vital role in ensuring the effectiveness of these preservation efforts. However, the vast geographical expanse of Western Australia presents a significant challenge for conducting physical inspections, as only 28% of the state's 3,000 bridge structures are located within the Perth metropolitan area. Consequently, Main Roads is limited to infrequent manual inspections, sometimes occurring only once every few years. While timber and timber hybrid bridges, accounting for approximately 40% of the bridge stock, can be easily instrumented to produce health indicators, there are unique challenges associated with instrumenting reinforced concrete and prestressed concrete bridges, which make up about 51% of the stock. This situation increases the likelihood of missing the optimal time window for addressing structural issues, leading to higher maintenance costs and a heightened risk of significant structural failures.

Opportunity

Bridge displacement, vibration and traffic load are three key indicators in bridge health monitoring. Level of displacement is a good descriptor of structural deformation behaviour and an indicator of structural performance and has been used widely for bridge health monitoring. Owing to high instrumentation cost for conducting bridge inspections and vibration tests, computer vision-based methods provide a cost effective and efficient alternative for measuring bridge displacement responses and conducting bridge condition monitoring. IoT technologies can aid in timely data collection resulting in real time bridge health monitoring and facilitating timely generation of warning alarms. Machine learning (ML) techniques have potential to be used for predicting displacement and vibration responses to traffic loading.

As an integrated package, IoT, computer vision and machine learning technologies offer to supplement physical bridge health assessment particularly in remote regional contexts which can be costly, time consuming and unsafe to inspect. Conducting regular, efficient, and reliable bridge health monitoring is essential for the long-term protection of valuable road assets through timely maintenance responses.

Purpose

The aim of the research was to investigate the feasibility of using an integrated package of IoT, computer vision and machine learning technologies to support smart bridge health monitoring and prediction. The specific project objectives were to:

1. Develop IoT technology-based solutions for collecting, visualizing and transmitting vibration response data;
2. Develop computer vision-based solutions for measuring bridge displacement under a range of traffic loads;
3. Develop an integrated proof-of-concept predictive model based on ML techniques for predicting bridge displacement under heavy traffic loads; and
4. Make recommendations on new techniques for efficient bridge monitoring, the wider roll-out of the technology for bridge health monitoring and prediction and suggest the next steps.

Approach and key finding

To deliver on the objectives, using the Stirling Bridge, Fremantle, as the test site, the research was conducted in three parts, each relating to an aspect of the essential indicators for bridge health monitoring: bridge displacement, vibration characteristics and traffic load.

Bridge displacement is important to understand to determine how the structure responds to different traffic loads, especially heavy loads. However, conventional contact-type displacement sensors, such as the linear variable differential transducer (LVDT), require a stationary reference point that is often difficult to find in the field. Additionally, their short measurement range of less than 1 meter limits their application to large-span bridges. As an alternative, a computer vision-based method was tested. Lab tests demonstrated good agreement with LVDT sensor data. Due to the limited range of LVDT, direct on-site validation was challenging, and only indirect checking of the data was possible. The observations were in line with the displacement influence line theory, and the top ten detected displacements correspond to heavy traffic patterns recorded by the traffic camera, indicating the reliability of the collected data. During the test, the physical dimensions of the bridge were needed to translate the number of pixels in the video to real distances. As the side of the bridge was inaccessible for physical measurement, design drawings were used instead. The field test also identified that environmental factors such as wind-induced camera motion and lighting conditions could affect the accuracy of results. The research team proposed possible solutions for future research to address these challenges.

Vibration data can be used to detect variations in the bridge's natural frequency over time, serving as an alarm for potential structural damage. A prototype IoT unit was built to measure bridge vibration using an accelerometer. The data was successfully transferred to the AWS S3 cloud storage over an IoT link, and a dashboard was created for accessing the collected accelerometer data. A comparison with an industry-grade wired sensor showed that the chosen IoT sensor lacked sufficient resolution, leading to the purchase of a better model. Although the field test opportunity was missed, lab tests demonstrated comparable results between the new IoT sensor and a smartphone accelerometer. The prototype has overall demonstrated the cost-effectiveness of IoT for collecting bridge vibration data when IoT data coverage is available. Further considerations include addressing power supply issues in remote areas and determining optimal attachment locations for the unit on the bridge.

The ML part of the project set out to develop proof-of-concept models to predict the bridge's vibration and displacement responses given the observed traffic loads from videos. Both tasks were proven to be difficult, although the vibration prediction achieved comparatively better results. The available literature indicates the inherent difficulty of mathematically inferring bridge displacement from vibration data. Extensive testing was conducted using various ML models, yet the challenge persisted, further confirming the complexity of the task. The 'phantom displacements' in the data also contributed to the difficulty. Nevertheless, the model estimated influence curve seemed reasonable, which suggests that it captured some of the underlining mechanisms. Interestingly, combining the vibration data with traffic videos did not increase accuracy.

Recommendations for further research

The computer vision and IoT components of the project, which concentrated on distinct metrics, both exhibited some potential in this preliminary investigation. There are several areas where further research can be considered to fully realize the potential of the developed techniques in managing bridge infrastructure:

- The IoT-based sensor data transmission to AWS needs further investigation to reduce the transmission rate. Edge computing at the source will reduce the amount of data to transmit to AWS, thus reducing the transmission cost and energy consumption.

- Validate the IoT sensor data against the wired accelerometer in terms of data accuracy and ease of deployment.
- Investigate the possibility of an independent power supply using solar panels etc. because many regional bridges don't have mains electricity.
- Investigate how IoT sensor data can be used to estimate bridge health conditions such as shifts in the fundamental frequency.
- Develop a user-friendly Graphical User Interface (GUI) to integrate and visualize the data collected from IoT sensors, video cameras, and traffic flow surveillance cameras in near-real time.
- Establish bridge performance indicators and safety thresholds with key parameters including structural health, load capacity, and environmental conditions, based on industry standards and best practice.
- Develop a comprehensive database of experience by monitoring more bridges under different conditions in relation to traffic load, environmental conditions, and maintenance history.

The proof-of-concept ML models to predict displacement and vibration from traffic load, served their purpose in exploring the feasibility of the idea. Although the models have academic novelty, this direction is unlikely to yield practical results in the near future. Given the primary objective of devising practical solutions to enhance bridge health monitoring, it is therefore recommended not to pursue further research in this direction.

Before progressing with any further research, engaging stakeholders in discussions becomes crucial to explore the practical implementation of these solutions within Main Roads' operations. This collaboration will facilitate the development of a well-defined path for further advancement.

Technical reports

Details of components of the research are provided in technical (milestone) reports, appended to this summary:

Appendix A: Prototype models for vision-based displacement measurement & IoT sensor data communication and visualization (Milestones 3a-b)

Appendix B: Prototype Machine Learning models for predicting bridge displacement and vibration (Milestone 3c)

PATREC

Australian Government
Department of Industry, Science, Energy and Resources

AusIndustry
Cooperative Research Centres Program

iMOVE

Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction

Final report: Prototype models for vision-based displacement measurement & IoT sensor data communication and visualization

April 2023

APPENDIX A

Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction

Final report: Prototype models for vision-based displacement measurement & IoT sensor data communication and visualization

Prepared by

Zhen Peng, Jun Li, Wensu Chen, Atif Mansoor & Robert Lee

Keywords

Computer vision, IoT sensor, machine learning, bridge health monitoring

Version control

Final report 1.0

Project and Milestone No.

iMOVE Project 1-046: Milestones 3a and 3b

Project steering committee

Raquib Hossain, Main Roads Western Australia
Steve Atkinson, Main Roads Western Australia
Sebastian Davies-Slate, WALGA
Sharon Biermann, University of Western Australia
Jun Li, Curtin University
Wensu Chen, Curtin University
Chao Sun, University of Western Australia
Atif Mansoor, University of Western Australia

Acknowledgment

This research is funded by PATREC and iMOVE CRC and supported by the Cooperative Research Centres program, an Australian Government initiative.

About PATREC

The Planning and Transport Research Centre (PATREC) is a collaboration between the Government of Western Australia and local universities, constituted to conduct collaborative, applied research and teaching in support of policy in the connected spaces of transport and land use planning. The collaborating parties are: The University of Western Australia, Curtin University, Edith Cowan University, Department of Transport, Main Roads Western Australia, Western Australian Planning Commission the Western Australian Local Government Association.

Publisher

Planning and Transport Research Centre
The University of Western Australia (M087)
35 Stirling Highway, Crawley, WA 6009
+61 8 6488 3385
patrec@uwa.edu.au

Project Executive Summary

Integrated IoT, computer vision and machine learning technologies offer a promising supplement to physical bridge health assessment particularly in remote regional contexts which can be costly, time consuming and unsafe to inspect. Conducting regular, efficient and reliable bridge health monitoring is essential for the long-term protection of valuable road assets through timely maintenance responses. This project aimed to investigate the feasibility of using an integrated package of up-to-date technologies, including IoT-based vibration sensors, sensing data transmission and visualization, vision-based techniques for measuring bridge displacement and machine learning techniques for relating the displacement measurement data with traffic load, to support smart bridge health monitoring and prediction. The research produced a proof-of-concept to demonstrate the efficacy and feasibility of an integrated package of technologies for first-level bridge health screening and early warning system, reducing the need for traditional physical inspections and instrumentation.

An IoT accelerometer sensor for vibration acceleration measurement is developed in this project. The feasibility and applicability of the developed IoT based accelerometer sensor for data collection, transfer over IoT link and storage at AWS S3 are investigated. A dashboard is developed to give insight in the collected accelerometer data. The time domain and frequency domain comparisons of IoT accelerometer sensor against smartphone accelerometer show comparable results. This project also develops a target-free computer vision-based approach as an alternative to conventional displacement sensors for measuring bridge displacement responses in a contactless manner. The developed approach involves camera calibration and scale factor determination, natural feature target identification and description, feature matching and tracking. Laboratory tests were conducted to validate the accuracy and efficiency of the developed approach. The efficacy and practicability of the developed approach were further evaluated through field tests on Stirling Bridge in Fremantle, Western Australia. The output of this research provides a low-cost and easy deployable means to estimate bridge dynamic displacements under normal traffic load. Two main environmental factors are identified, i.e., wind and light conditions, which adversely affect the displacement identification accuracy. Furthermore, suggestions were provided to mitigate the effects of wind-induced camera motion and displacement identification errors. The developed technique provides an affordable and easily deployable alternative to conventional contact-type displacement sensor, which can be used for timely bridge health condition assessment.

The benefits of the project include contributing to reducing maintenance, operation costs and risk, and achieving a safe transport infrastructure network, ultimately, increasing productivity.

Contents

1. Project overview

This report (milestones 3a and b) presents the developed methodology, the laboratory and field activities associated with developing and validating IoT and computer vision for smarter bridge health monitoring and prediction. This project has been conducted by University of Western Australia and Curtin University researchers under contract with the iMOVE Australia Limited. All work was performed at the universities and Stirling bridge locations. The period of this project was August 2021 to March 2023. It should be noted that this report covers the IoT sensor development (milestone 3b) and computer vision technique for displacement measurement techniques (milestone 3a). The report for the machine learning component is presented in a separate milestone 3c report.

1.1. Background

In Western Australia there are close to 3000 bridges with Main Roads WA owning the largest proportion of bridge structures (42%), thus is responsible for the routine inspection and maintenance of these assets. Further, Main Roads WA is also responsible for maintaining truck loading standards for publicly accessible roads and bridges, owned primarily by Main Roads WA or local authorities.

While 28% of the state's bridge structures are found within the Perth metropolitan area, with fairly easy access for physical inspection, the vast majority are located at considerable distances away in the South West (33%), Wheatbelt (22%) and in the even more remote areas of the Great Southern, Pilbara, Kimberley, Goldfields Esperance and Gascoyne (17%) [1].

Due to the vast distances involved, Main Roads WA can only conduct manual bridge inspections at very low frequencies, e.g. some bridges are only visited every several years. This could mean missed windows of opportunity for optimal treatment when structural problems occur, which could lead to higher overall maintenance costs and increased risks of major structural failures. It could also mean more disruptions to the freight movement during maintenance.

In addition, while timber/timber hybrid bridges (40% of stock) can be fairly easily instrumented to collect bridge health indicators specific to timber structures, reinforced concrete and prestressed concrete bridges, which form a large and growing proportion of the bridge stock (51 %) in WA, have their own unique set of instrumentation and health evaluation challenges. This research specially considers concrete bridges where the opportunity for solutions which reduce physical instrumentation requirements, further justifies the importance of the proposed research.

Bridge displacement, vibration characteristics and traffic load are three key indicators in bridge health monitoring. Level of displacement is a good descriptor of structural deformation behaviour and an indicator of structural performance and has been used widely for bridge health monitoring. Owing to high instrumentation cost for conducting bridge inspections and vibration tests, computer vision-based methods provide a cost effective and efficient alternative for measuring bridge displacement responses and conducting bridge condition monitoring. IoT technologies can aid in timely data collection resulting in real time bridge health monitoring and facilitating timely generation of warning alarms. The computer vision, IoT and AI based smart decision-making tool enable Main Roads WA to monitor bridge health conditions on an ongoing basis, which does not only save on inspection costs but also increases the efficiency of the maintenance program. It can also generate data that has been previously unavailable for better modelling how bridges respond to traffic loading.

1.2. Project objectives

The overall aim of the project is to demonstrate the feasibility of using a combination of IoT, computer vision, and machine learning technologies to reduce the need for manual bridge inspections and provide ongoing monitoring that supports smart bridge maintenance. In this project, the development, application and integration of IoT technology-based sensor data collection, computer vision methods and machine learning techniques are investigated to enable remote monitoring of bridge structural health to enhance the efficiency of asset monitoring, management and maintenance. The specific project objectives are:

1. To develop IoT technology-based solutions for collecting, visualizing and transmitting vibration response data.
2. To develop computer vision-based solutions for measuring bridge displacement under a range of traffic loads.
3. To develop an integrated proof-of-concept predictive model based on machine learning techniques for predicting bridge displacement under heavy traffic loads; and
4. To make recommendations on new techniques for efficient bridge monitoring wider roll-out of the technology for bridge health monitoring and prediction and suggest the next steps.

2. Methodology and technical details

This section presents the methodologies and technical details for the development of the IOT sensing platform and computer vision-based displacement tracking algorithm. For consistency and integration across all three parts of the work, the model and algorithm development, data analysis and data visualization were implemented in Python.

2.1. IOT sensor-based vibration measurement

The proposed system schematic is shown in Figure 2.1. The functionality of the system is described by a number of layers. The hardware layer which includes the microcontroller, accelerometer, power supply, local storage, RTC, NB-IoT modules etc. A sensing layer which describes the sampling of the accelerometer at a given output data rate (ODR) and saving to local storage (SD card). A transmission layer, which encompasses the transmission of captured data to S3 storage via AWS IoT core. Finally, an access layer which describes how the end user interacts with the data collected – for this project it could include the processing of the transmitted data and displaying to a dashboard.

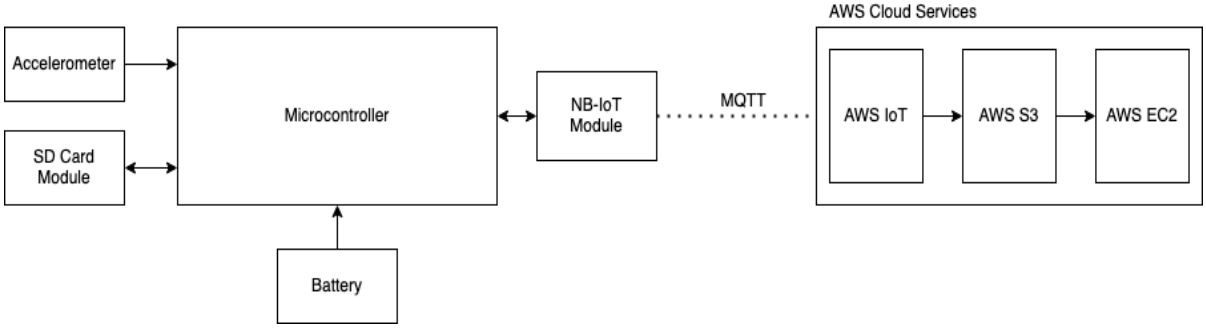


Figure 2.1: The IoT part of the Smart Bridge Health Monitoring system

2.2. Computer vision-based displacement measurement

Existing bridge structures exposed to the operational environment for long service life, are prone to performance degradation owing to material deterioration, natural hazards and human-made loading conditions [2]. Their vertical displacement (deflection) under traffic loads is usually selected as a critical parameter for evaluating bridge performance [3]. Conventional contact-type displacement sensors such as the linear variable differential transducer (LVDT) require a stationary reference point, which is often difficult to be found in the field. Furthermore, the measurement range of traditional displacement sensor is relatively short, which limits its application to large-span bridge structures [4]. To address the limitations of current sensor systems for field applications, the research community has been actively exploring new technologies that can advance the state-of-the-practice in structural health monitoring (SHM). Thanks to the rapid advances in computer vision, the camera-based non-contact vision sensing has emerged as a promising alternative to conventional contact sensors for structural dynamic response measurement and health monitoring. Significant advantages of the vision sensor include its low cost, ease of setup and operation, and flexibility to extract displacements of any points on the structure from a single video measurement.

The typical procedure of the vision-based displacement measurement includes:

(1) Video camera setup and calibration. A camera equipped with a lens can be positioned remotely on a tripod for short-term measurements or fixed in place for long-term monitoring. The recorded video can be processed in real-time using image-processing software or stored for post-processing. Prior to use, the camera must be calibrated to establish the geometric relationship between the image coordinates and the corresponding real-world coordinates.
(2) Single or multiple target/ feature detection. Any texture, natural or artificial, on the surface of a structure can be used as a tracking target, provided that it has a distinct pattern that stands out from the surrounding background. However, for accurate pattern matching, a suitable subset with sufficient local texture must be carefully selected for each measurement point.
(3) Feature matching and tracking. To track the motion of a target, its position is identified in a sequence of video images. Advanced vision techniques now offer subpixel tracking accuracy, allowing for precise measurement of even the very small movements.
(4) Displacement extraction. The process of extracting displacement involves converting the structural motion, which is initially measured in pixel units, to physical units such as millimeters, using a scale factor.

In practical applications, ensuring the measurement accuracy of the vision sensor is a significant concern. Measurement errors can arise from any of the procedures mentioned above. Through a review of existing literature and preliminary measurement results, the following potential sources of measurement errors have been identified [5, 6]:

(1) Errors from hardware limitations. Hardware limitations can contribute to measurement errors. The performance of a video camera is typically determined by technical specifications such as image resolution, frame rate, and focal length. The size of the image sensor and its corresponding resolution are related to the field-of-view and expected measurement resolution, given the ratio of the measurement distance and focal length. The maximum resolution of a consumer-class video camera can range from 640x480 to 3840x2160 (4k). According to the Nyquist theorem, the sample rate must be at least twice the highest frequency component of the measured signal to avoid aliasing. A frame rate of 50 fps is typically sufficient to capture the quasi-static and dominant dynamic displacement responses of a bridge. Focal length is a critical hardware parameter, and longer focal lengths provide more detailed zoomed-in views for long-distance targets.

(2) Errors from computer-vision algorithm. When analyzing a video stream, the accuracy, reliability, and efficiency of the feature point detection and matching algorithms used must be evaluated to minimize errors.
(3) Errors from environmental sources. Measurement errors, particularly in field tests, can also result from environmental sources such as camera motion, lighting conditions, and non-uniform air refraction. In field measurements, the camera can be subjected to ambient vibration from wind, passing trains, and other sources, which can cause displacement measurement errors.

A technical pipeline of computer vision-based displacement tracking process is illustrated in Figure 2.2.

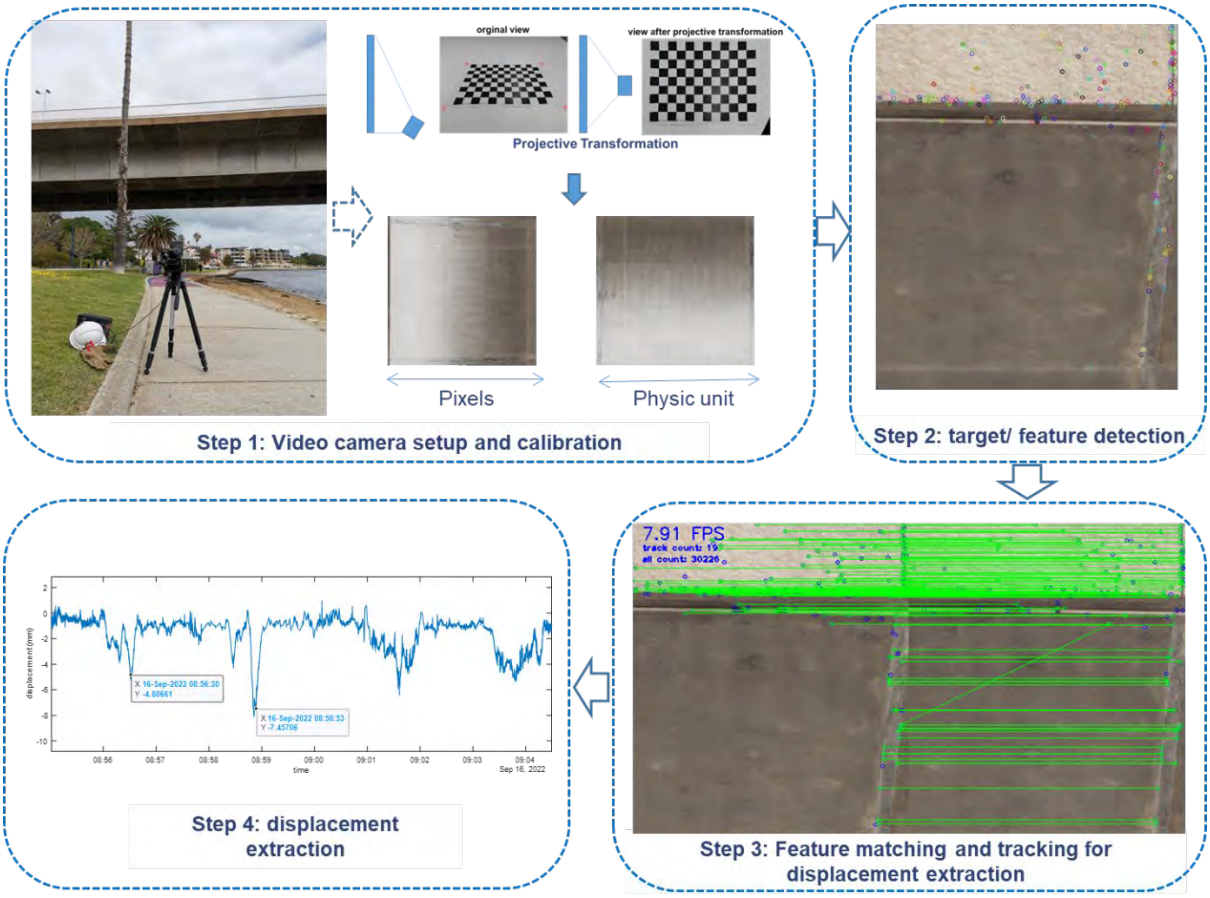


Figure 2.2: Technical pipeline of computer vision-based displacement tracking

2.2.1. Hardware selection

This project includes laboratory and field tests to validate the accuracy of the vision-based displacement tracking algorithm. During the laboratory test, the ground truth data for displacement are measured using a laser displacement sensor, as shown in Figure 2.3(a), and compared to the displacement measured by the vision-based method. During the site visit to Stirling Bridge in Fremantle, it was observed that the distance between the potential camera location and the bridge target component is approximately 15-30 meters. To ensure high measurement resolution, an optical lens with an appropriate focal length was selected to zoom in on the image and obtain an enlarged tracking target or targets. As stated in the project objectives, data collected from IoT

sensors, the vision-based technique, and traffic cameras should be synchronized to train the predictive machine learning model. IoT sensors typically synchronize their system time via the internet, while an offline video camera can synchronize its system time using GPS. Considering these factors, a Sony PXW-FS5 camera (shown in Fig. 2.3(b)) with a Sony ALC-SH135 28-135mm lens and an in-built GPS module is selected for the project. Ref. [7] suggested that the displacement measurement error can be significantly influenced by the motion of the camera induced by wind effect. This problem becomes more severe when using a lightweight compact camera-tripod system with a zoom lens. Therefore, a relatively heavy and rigid tripod is chosen, as depicted in Fig. 2.3(c). Moreover, the battery of Sony PXW-FS5 has a limited lifespan of 10-20 minutes of video recording, which is inadequate to collect sufficient data for training the predictive machine learning model. To address this issue, a portable power station with a capacity of 606Wh is selected, as shown in Fig. 2.3(d), which can extend the continuous video recording time to about 24 hours.

(a) Laser displacement sensor

(b) Sony PXW-FS5 camera

(c) Camera-tripod system

(d) Portable power station

Figure 2.3: hardware selection for vision-based displacement tracking.

2.2.2. Camera calibration and scale factor determination

Accurate camera parameter calibration is essential for computer vision applications. The goal of camera calibration is to establish a precise relationship between a 3D point in the real world and its corresponding 2D projection (pixel) in the image captured by the calibrated camera. The equations that relate 3D point (X_w, Y_w, Z_w) in world coordinates to its projection (u, v) in the image coordinates are shown below

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{P}\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{1}$$

where s is a scale factor. $\boldsymbol{P}$ is a 3×4 Projection matrix consisting of two parts — the intrinsic matrix $\boldsymbol{K}$ that contains the intrinsic parameters and the extrinsic matrix $[\boldsymbol{R}|\boldsymbol{t}]$ that is a combination of 3×3 rotation matrix $\boldsymbol{R}$ and a 3×1 translation $\boldsymbol{t}$ vector.

$$\boldsymbol{P} = \overbrace{\boldsymbol{K}}^{\text{Intrinsic Matrix}} \times \overbrace{[\boldsymbol{R}|\boldsymbol{t}]}^{\text{Extrinsic Matrix}} \tag{2}$$

$$\boldsymbol{K} = \begin{bmatrix} f_x & \gamma & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, [\boldsymbol{R}|\boldsymbol{t}] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \tag{3}$$

where f_x and f_y are the horizontal and vertical focal lengths expressed in pixel units, which are usually the same. c_x and cf_y are the x and y coordinates of the optical centre in the image plane. Using the centre of the image is usually a good enough approximation. γ is the skew between the axes. It is usually 0. The goal of the calibration process is to find the 3×3 matrix $\boldsymbol{K}$, the 3×3 rotation matrix $\boldsymbol{R}$ and the 3×1 translation $\boldsymbol{t}$ vector using a set of known 3D points (X_w, Y_w, Z_w) and their corresponding image coordinates (u, v). The camera calibration can be conveniently conducted using the well-known Python OpenCV camera calibration package. To calibrate the camera, multiple images of a calibration pattern from different angles are needed. Popular calibration patterns include the asymmetric checkerboard and the circular control points.

The camera calibration method described above requires installing a calibration panel on the target structure, which can be difficult to do in the field. Therefore, a more practical calibration method is needed, which can be based on the known physical dimensions of the object's surface. In the case of a bridge structure, the physical dimensions can be obtained from the design drawing. When the camera's optical axis is perpendicular to the objective surface, all points on the surface have the same depth of field, meaning that they can be uniformly scaled down into the image plane. In this case, only one scaling factor (SF) is required. The scale factor can be calculated based on the known physical dimension of the objective surface and its corresponding image dimension in pixels, as follows:

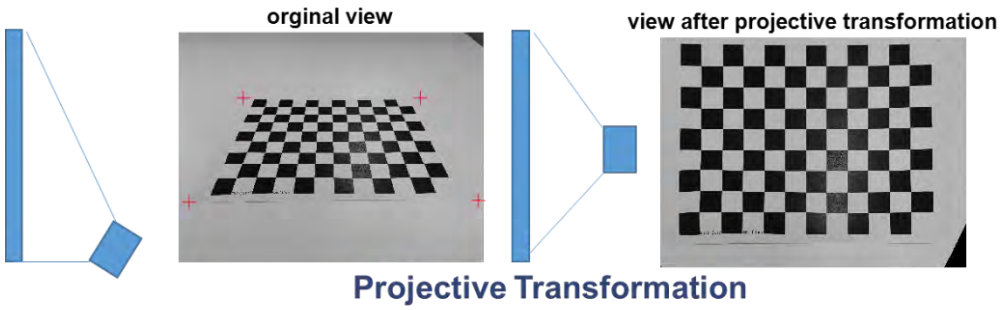
$$SF = \frac{d_{known}}{I_{known}} \tag{4}$$

where d_{known} and I_{known} are the known physical length and the corresponding pixel length at the image, respectively.

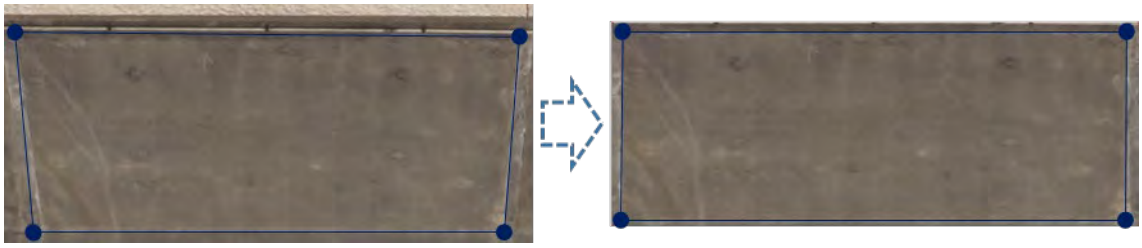
However, the prerequisite of SF given by Equation (4) is the perpendicular of the camera's optical axis to the objective surface. This can pose difficulties in field implementations because even a small misalignment of the angle of the camera can go unnoticed during the experimental setup but can cause errors, especially when the object is relatively far from the camera. In outdoor field tests, it is often inevitable to tilt the camera's optical axis slightly to track the object surface target. However, when the target moves within a plane contained in the 3D structural system. For example, when the traffic load mainly induces bridge vertical motion, the projection relationship can be simplified to a planar projective transformation between a 2D structural plane and a 2D image plane [8]. In this case, a projective transformation can be applied to project the original camera view to a perpendicular view. Figure 2.4 illustrates the projective transformation, which requires the coordinates of four points in world coordinates and image coordinates. As shown in Figure 2.4(b), four corners of a rectangle can be selected by mouse click. With the width and height

of the rectangle in world coordinates and image coordinates, the scale factor can be obtained using Equation (4).

It is important to acknowledge that smartphone cameras and some consumer-grade cameras, such as GoPro, use wide-angle lenses to expand the field-of-view. This often results in distorted images. To address this issue, the lens distortion parameters can be determined during laboratory calibration, and the rectified image can be corrected accordingly. However, if the camera lens does not exhibit any visible distortion, the correction step is not necessarily required. In the current project, the Sony PXW-FS5 camera and its lens are not expected to introduce any image distortion.



(a) checkerboard



(b) bridge component

Figure 2.4: Illustration of projective transformation.

2.2.3. Target detection and tracking algorithm

After determining the scale factor of the video, the motion of each feature point is identified using a target detection and tracking algorithm. Searching for the marker throughout the whole image could be highly time-consuming. To address this, a more efficient approach is to select a Region of Interest (ROI) in the first video frame containing multiple high-quality feature points. Then, the feature points can be detected and matched within the ROI in the subsequent video frames. The ROI can be chosen by using the mouse click on the first video frame in the code implementation.

There are three main categories of target tracking methods used for motion tracking, as summarized in Table 2.1. In practical applications, measurement accuracy is a major concern for the vision sensor. Template matching techniques alone can typically only measure displacements with integer-pixel resolution, as the minimal unit in a video image is one pixel. While pixel-level accuracy is adequate for many applications, higher resolution is often required to measure small structural vibrations, such as the ambient vibration of a short-span concrete bridge. Pixel-level template matching may result in unacceptable measurement errors if the displacement to be measured is in the same order of magnitude as the scaling factor. To improve measurement

accuracy, most tracking methods incorporate interpolation, optimization, and/or up-sampling techniques to achieve subpixel accuracy. Subpixel accuracies reported in many studies vary within orders of magnitude from 0.5 to 0.01 pixel.

Table 2.1. Categories of tracking method and corresponding representative algorithms

Tracking methods	Description	Representative algorithms
Optical flow estimation	Optical flow calculates a velocity for points within the images, and provides an estimation of where points could be in the next image sequence.	Sparse optical flow • Lucas-Kanade algorithm • Sparse RLOF Dense optical flow • Dense Pyramid Lucas-Kanade algorithm • Farneback • DeepFlow • PCAFlow
Area-based template matching	Template matching Searching and finding the location of a template image in a larger image.	• cross-correlation (CC) • normalized cross-correlation • zero-normalized cross-correlation (ZNCC) • The sum of squared differences (SSD) correlation • zero-normalized sum of squared differences (ZNSSD).
Feature point matching	Feature matching refers to finding corresponding features from two similar images based on a search distance algorithm.	Descriptor • Haris corner detection • Shi-Tomasi corner detection • scale-invariant feature transform (SIFT) • Speeded up robust features (SURF) • FAST algorithm for corner detection Matcher • Fast Library for Approximate Nearest Neighbors (FLANN) based matcher • Brute-Force Matching

Optical flow estimation-based tracking methods can lose track of an object when it goes behind an obstacle for an extended period of time or moves too quickly for the tracking algorithm to keep up. In addition, sudden changes in lighting conditions can cause the tracker to lose track of an object. Tracking algorithms can also accumulate errors, causing the bounding box tracking the object to slowly drift away from the object it is tracking. These drifts can be mixed up with quasi-static displacement responses induced by moving vehicle loads. Therefore, optical flow estimation-based methods are not considered in this project. Instead, feature point matching-based techniques, which offer advantages such as geometric and photometric invariances, are extensively adopted for vision-based displacement tracking tasks. It should be noted that the performance of the feature detection algorithm is dependent on factors such as video quality and feature point characteristics. For the laboratory validation and field test in this project, SIFT feature points detection and descriptor, along with FLANN-based matcher, are adopted.

2.2.4. Evaluation matrix

The efficiency and accuracy are two major concerns of the vision-based displacement identification algorithm. The efficiency is evaluated with the analysed frames per second (fps), the accuracy is quantitively evaluated with the correlation coefficient R_c and root mean square error (RMSE), defined as

$$R_c = \sqrt{\frac{N\sum_{i=1}^{N} d_v(i)d_l(i) - \sum_{i=1}^{N} d_v(i)\sum_{i=1}^{N} d_l(i)}{\left[N\sum_{i=1}^{N} d_v^2(i) - \left(\sum_{i=1}^{N} d_v(i)\right)^2\right]\left[N\sum_{i=1}^{N} d_l^2(i) - \left(\sum_{i=1}^{N} d_l(i)\right)^2\right]}} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}[d_v(i) - d_l(i)]^2}{N}} \quad (7)$$

where d_v and d_l represent the vision-based and laser displacement sensor-based measurement, respectively. N is the overall number of sampling points.

3. Preliminary laboratory validation

3.1. IOT sensor-based vibration measurement

The system has been tested on three different platforms during the course of the project, Libelium Waspmote microcontroller board, RAKwireless WisTrio RAK5010, Arduino MKR NB1500 and later with accelerometers of higher resolution. The differences in the hardware are summarised in Table 2.2. Further, the three hardware were placed in the box and used for various experiments. Figure 3.1 shows the encased hardware.

Table 2.2: Summary of hardware differences

	Waspmote	RAK5010	Arduino
Accelerometer	LIS331DLH	LIS331DLH	External
Sampling Rate	100, 400 Hz	100, 200, 400 Hz	100Hz+
SD Card	Built-In	None	External
RTC Module	Built-In	None	Built-In
NB-IoT Module	BG96	BG96	SARA-R410M-02B
Programming	Waspmote API	Arduino	Arduino

Figure 3.1: From left to right, Libelium Waspmote, RAKwireless WisTrio RAK5010 and Arduino MKR NB1500.

The Libelium Waspmote was tested first, but due to minimal support for the Waspmote software libraries and limited SRAM on the microcontroller, it was found not very effective. The RAKwireless WisTrio RAK5010 was tested next for the project. It had connectivity issues over the NB-IoT so it was replaced with Arduino MKR NB1500. The rest of this report concerns the testing and validation of the Arduino MKR NB1500 platform, which used an external SD card shield and the ADXL345 accelerometer.

3.2. Computer vision-based displacement measurement

To validate the target-free vision-based technique developed in this project, laboratory tests were conducted using an experimental setup shown in Figure 3.2. The setup consists of an aluminium column fixed on a shake table and excited by a ground sine-wave with an amplitude of 3mm and frequency of 3Hz. Ground truth displacement at measurement points 1 and 2 were measured using two laser displacement sensors fixed on a steel frame. Since the camera is perpendicular to the aluminium column, the projective transformation procedure is not required. The scale factor between pixel units and physical length units (mm) can be calculated based on prior knowledge of the length of the aluminium column. To compare the accuracy of vision-based displacement identification results using artificial and natural targets, a chessboard mark was attached at points 1 and 2. The laboratory validation utilized SIFT feature point detection and descriptor along with FLANN-based matcher.

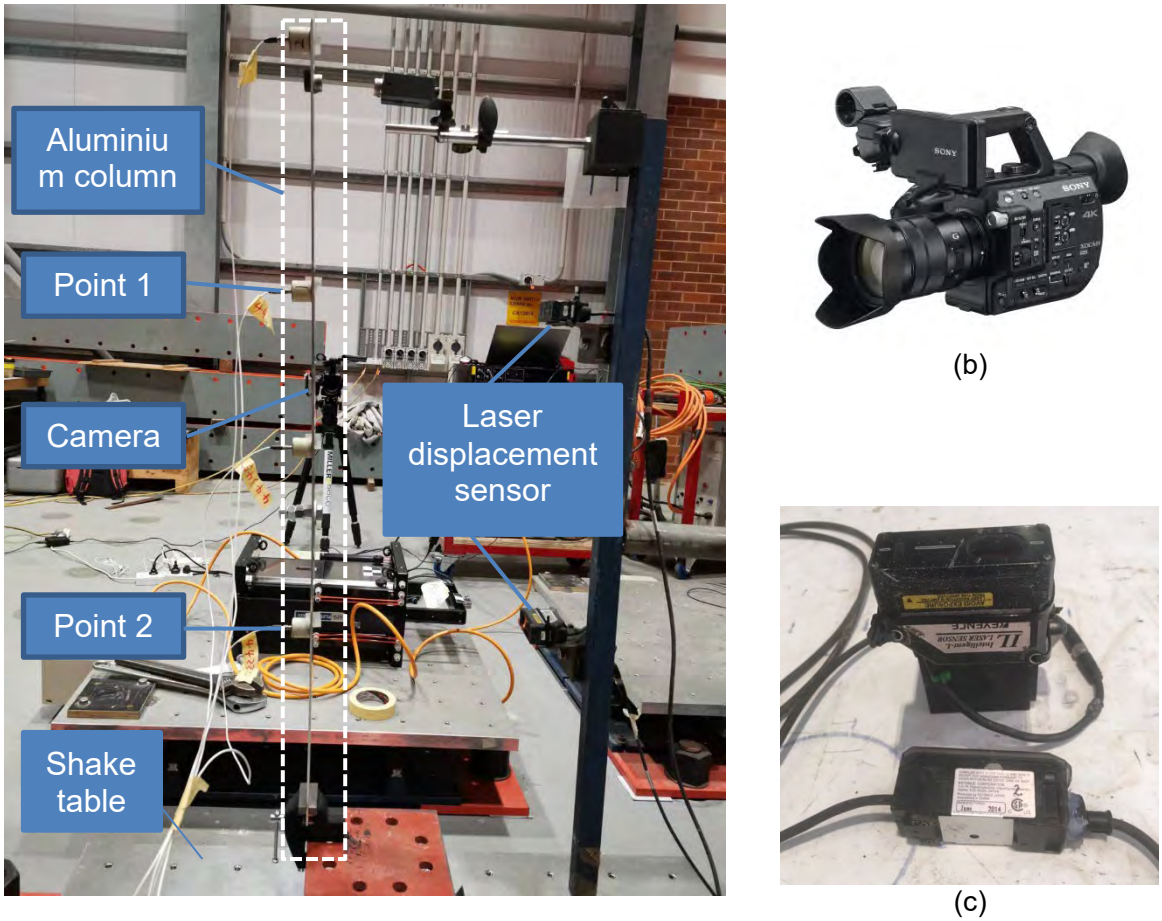
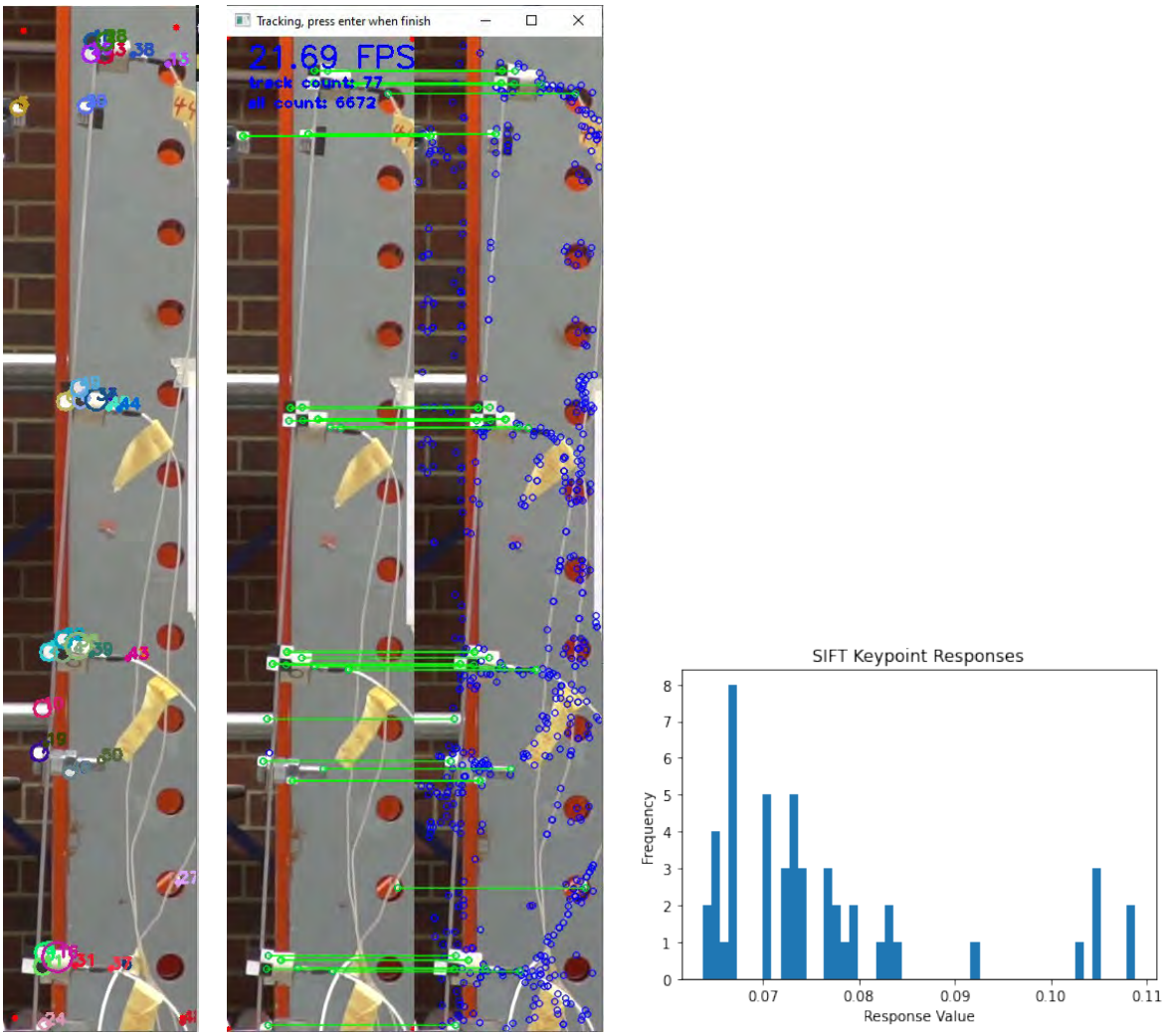


Figure 3.2: (a) Experiment setup, (b) Video camera and (c) Laser displacement sensor

3.2.1. Comparison between artificial and natural target

It is worth emphasizing that the ability to precisely measure bridge displacements using non-contact computer vision sensors with natural targets presents new monitoring possibilities that do not require the installation of artificial targets on the structure or disrupt the operational traffic conditions of the bridge. Therefore, both artificial and natural targets were utilized to identify displacement responses. Figure 3.3 illustrates the detection and matching results of key points along with their corresponding response values. In Figure 3.3(a), it can be seen that distinct key points have been successfully detected in all four artificial targets along the cantilever column. Additionally, natural features on the accelerometer's surface have also been detected, which is used for target-free structural displacement tracking. The first and sequential frames of the video stream are displayed in the left half and right half of Figure 3.3(b), respectively. The horizontal green lines depict the matching between the key points in the first and sequential frames. It can be observed that all feature points have been matched correctly. Figure 3.3(c) presents the response value of a SIFT key point, which indicates the strength of the key point. The response value is determined by the difference between the intensities of the pixels in the neighbourhood of the key point. It is used to filter out weak key points and to retain only strong and distinctive key points. The size of the circle in Figure 3.3(a) is proportional to the response value of each key point.

The response value of a SIFT key point reflects its strength and contrast relative to the surrounding pixels. Higher response values indicate stronger key points, while lower response values indicate weaker ones. Displacement responses at Point 1 identified from both artificial and natural targets are presented in Figures 3.4 and 3.5, respectively. Displacement responses at Point 2 identified from artificial and natural targets are shown in Figures 3.6 and 3.7, respectively. Table 3.1 presents the correlation coefficient and root mean square error for the identified displacement responses. As seen in Figures 3.4-3.7 and Table 3.1, the displacement responses identified from the video align well with those measured by the laser displacement sensor, demonstrating the accuracy of the vision-based displacement identification method. Additionally, the displacement identified from the natural target has a slightly higher RMSE than that from the artificial target, indicating the feasibility of target-free vision-based displacement identification. The algorithm can track motion in the video at an average speed of 21.4414 fps, which is about half of the video's recording speed of 50fps. It is worth noting that this speed is calculated based on the ROI in Figure 3.3, and the tracking speed can be further improved by selecting a smaller ROI that includes only the most distinct key points.



(a) (b) (c)

Figure 3.3: (a) Key points detected in the first frame (b) Key points matching between first and sequential frame (c) Histogram of key points responses value.

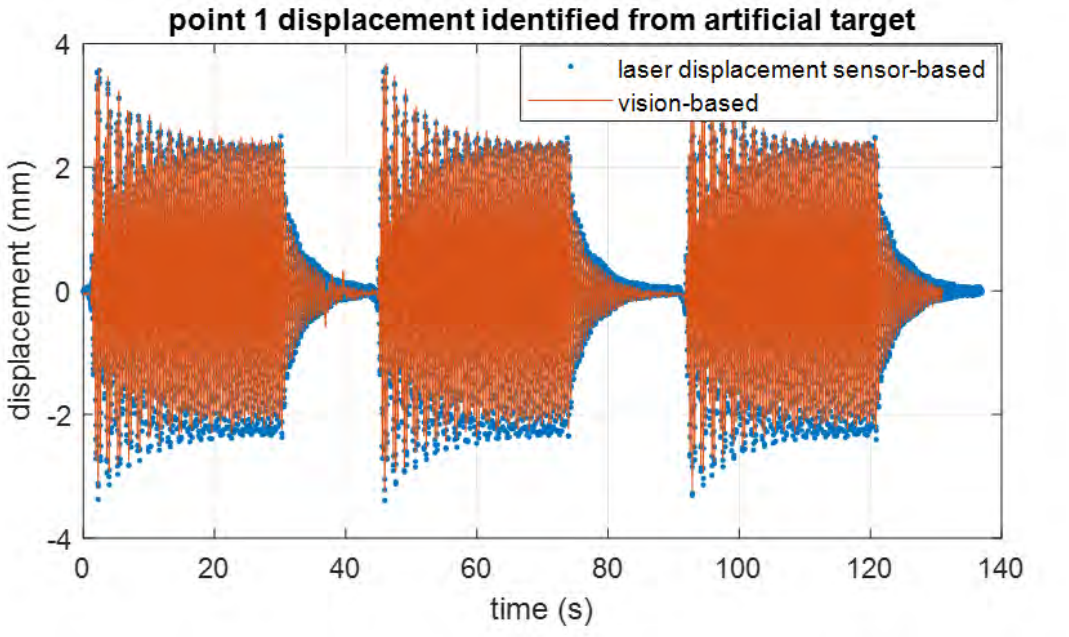


Figure 3.4: Point 1 displacement identified from artificial target.

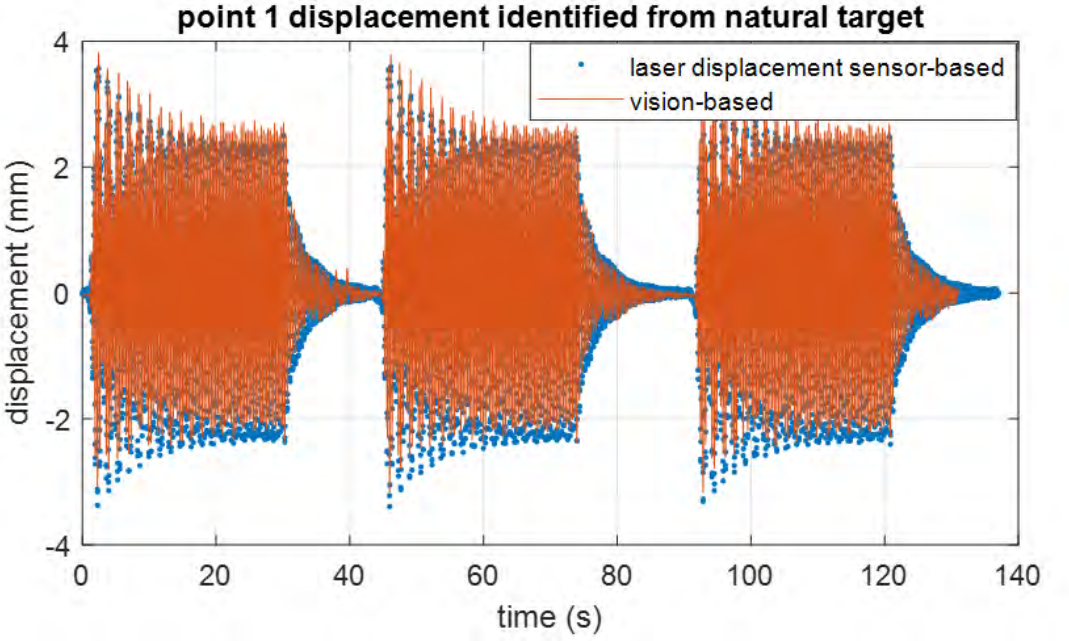


Figure 3.5: Point 1 displacement identified from natural target.

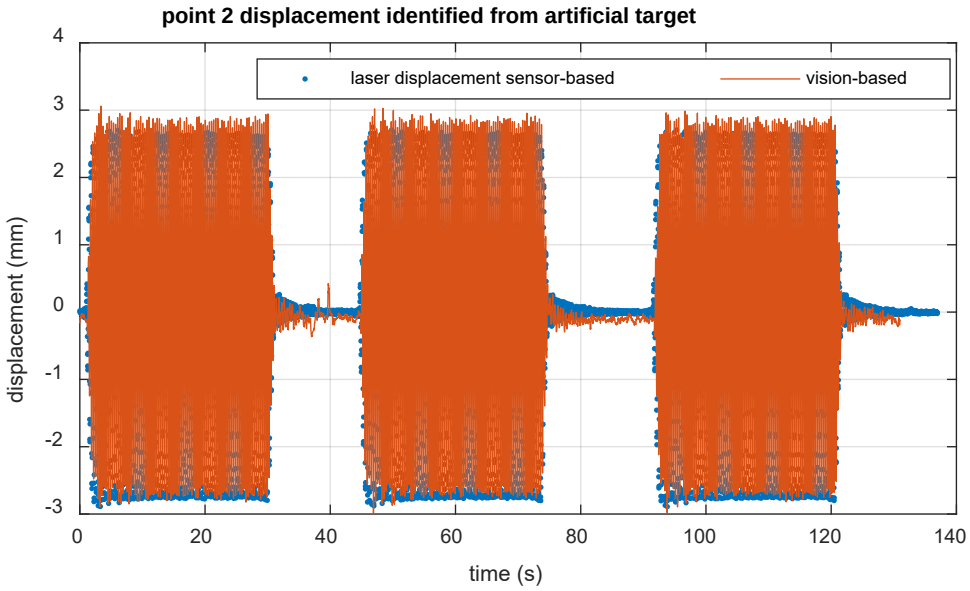


Figure 3.6: Point 2 displacement identified from artificial target.

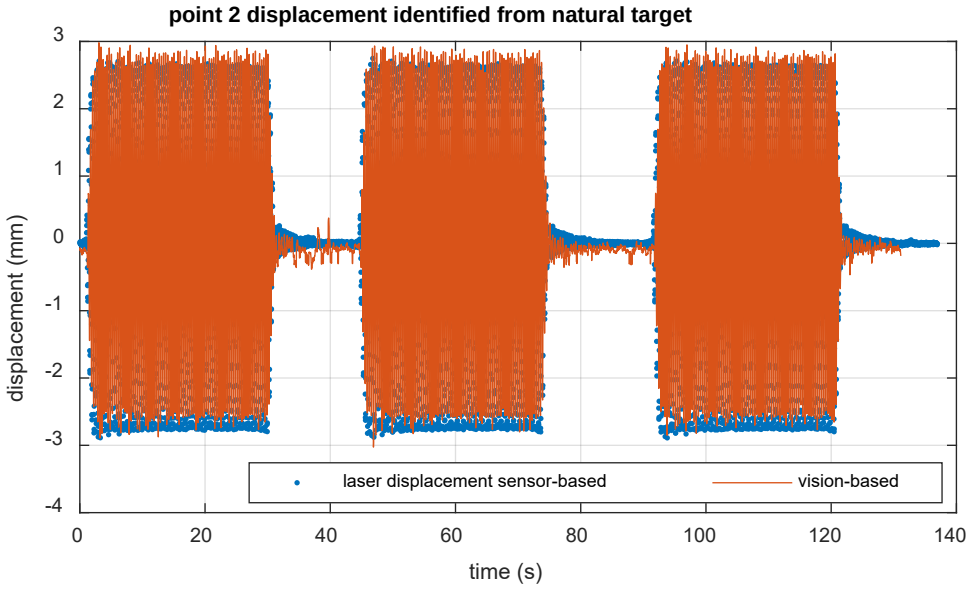


Figure 3.7: Point 2 displacement identified from natural target.

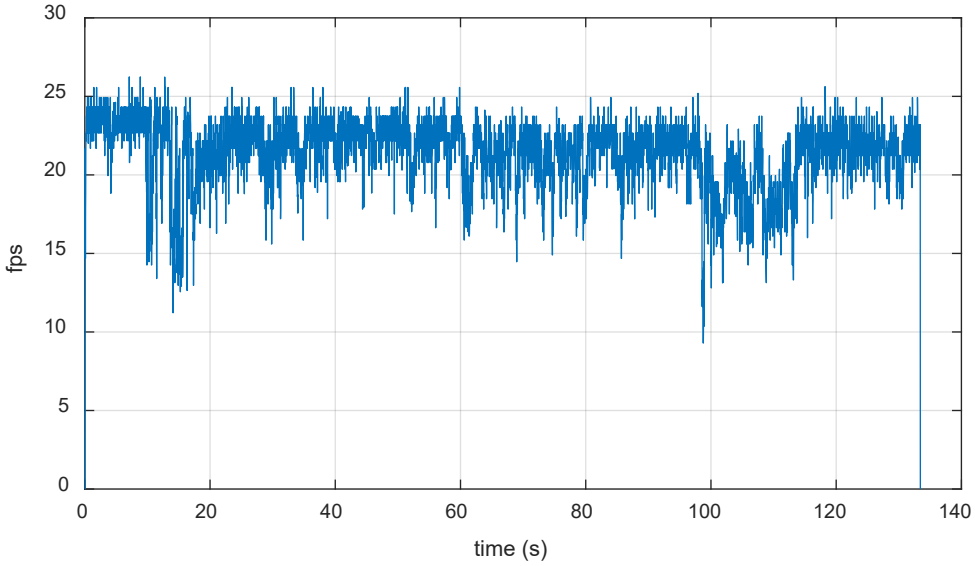


Figure 3.8: Analysed frames per second (fps)

Table 3.1 Error analysis of displacement identification results

		correlation coefficients	root mean square error	analysed frames per second (fps)
Point 1	Artificial target	0.9817	0.2856	21.4414±2.1677
	Natural target	0.9734	0.3843	
Point 2	Artificial target	0.9940	0.1850	
	Natural target	0.9936	0.1960	

Note: the vision-based displacement tracking algorithm are implemented in a HP laptop with AMD Ryzen 7 5800U CPU and 16.0GB RAM.

4. In-situ validation result analysis

During September 14th -16th, 2022, a series of in-situ bridge tests were carried out. The overall view of the field test setup and detailed sensor placement can be seen in Figure 4.1. The Stirling bridge consists of seven spans, with individual lengths from south to north of 23.8m, 81.4m, 75.3m, 69.2m, 63.1m, 54.9m, and 47.2m, as shown in the figure. The data acquisition system (DAQ) for acceleration measurements was placed in the middle section of the second span (81.4m) from the south bridge abutment, as seen in Figure 4.1(a). A Sony PXW-FS5 video camera was placed underneath the bridge, targeted at the same section of the bridge, and powered by a portable power station. Two wired Kistler 8330A3 accelerometers and an IoT accelerometer were installed at the same location to measure the bridge's vertical acceleration responses, as depicted in Figure 4.1(c-d). The wired accelerometers were powered by a portable power station connected to a solar

charger panel. A frame of the video stream recorded by TrafficCam2, installed in the south abutment of Stirling bridge, can be seen in Figure 4.1(e).

(a)

(b)

(c)

(d)

(e)

Figure 4.1. (a) Overall view of field test setup; (b) Video camera; (c) Acceleration data acquisition system; (d) Wired and IoT accelerometer placement (e) Traffic flow on bridge recorded by TrafficCam2.

4.1. IOT sensor-based vibration measurement

4.1.1. Data collection

Testing of the Arduino platform began by stabilising the sampling rate of the ADXL345 accelerometer, which ensured that there was consistent timing between the access of data from the ADXL345 registers and reducing the likelihood of false readings. Next, the timing for the 'flushing' of accumulated samples from the Arduino RAM to the SD permanent storage was optimised, thus saving the data every 10 seconds (1000 samples). These improvements were integrated into software and were used in the experiment at Stirling Bridge in September 2022 as detailed in "Validation and Result" section later.

The transfer of data over Narrow Band IoT to AWS IoT core was also investigated. The Arduino first establishes connection to the Telstra network over NB IoT, then proceeds to connect to AWS IoT Core through MQTT broker (see Figure 4.2). A rule action is implemented in which all messages received to this topic are stored in the Amazon S3 bucket.

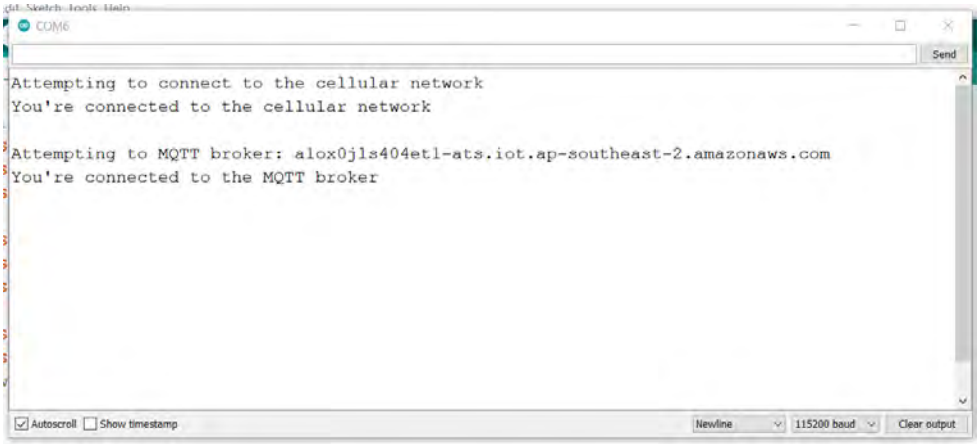


Figure 4.2: Arduino serial monitor readout showing the connection to Telstra network and AWS IoT core

Firstly, a synchronous transfer was tested (each sample is transmitted as it is read) which resulted in a maximum transfer rate of 4Hz. To improve the transfer rate, an asynchronous method was employed, first by testing a 500 sample (5 sec) payload, which was first recorded and saved to the SD prior to sending. However, this method is limited by the message size. Currently, due to the RAM of the Arduino, 2kB messages are the largest size that can be sent reliably, e.g for 30 sec and 1 min (3000 and 6000 samples) of saved accelerometer data – each 2kB message takes approximately 1.5 seconds to send (see Figure 4.3). One minute of 3 axis data requires 80 seconds to send to AWS. Recording and transmitting only Z-axis, one minute of data takes 40 seconds to upload to AWS. Due to AWS access limitation, extensive testing could not proceed.

1675821603597	-	February 8, 2023, 10:00:04 (UTC+08:00)	2.0 KB	Standard
1675821604790	-	February 8, 2023, 10:00:05 (UTC+08:00)	2.0 KB	Standard
1675821606071	-	February 8, 2023, 10:00:07 (UTC+08:00)	2.0 KB	Standard
1675821607361	-	February 8, 2023, 10:00:08 (UTC+08:00)	2.0 KB	Standard
1675821608632	-	February 8, 2023, 10:00:09 (UTC+08:00)	2.0 KB	Standard
1675821609924	-	February 8, 2023, 10:00:10 (UTC+08:00)	2.0 KB	Standard
1675821611193	-	February 8, 2023, 10:00:12 (UTC+08:00)	2.0 KB	Standard
1675821612474	-	February 8, 2023, 10:00:13 (UTC+08:00)	2.0 KB	Standard
1675821613751	-	February 8, 2023, 10:00:14 (UTC+08:00)	2.0 KB	Standard
1675821615030	-	February 8, 2023, 10:00:16 (UTC+08:00)	2.0 KB	Standard
1675821616318	-	February 8, 2023, 10:00:17 (UTC+08:00)	2.0 KB	Standard
1675821616951	-	February 8, 2023, 10:00:17 (UTC+08:00)	607.0 B	Standard

Figure 4.3: From AWS S3 Bucket. Each of the messages are 2kB, except the last, and are spaced 1-2 seconds apart.

For visualization, a dashboard is developed using the Dash/Plotly python library. The current working version of the dashboard is shown in Figure 4.4.

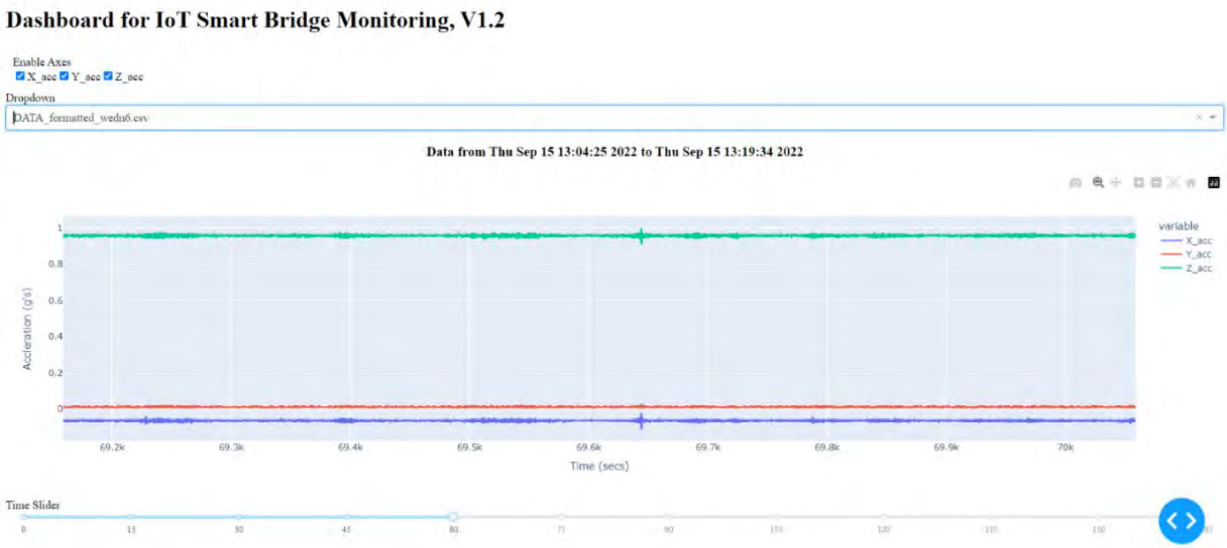


Figure 4.4: Dashboard created to visualise acceleration data using the Dash/Plotly libraries

Different axis selections can be made to display 3 Hours data. The dashboard subtitle is automatically updated whenever the file or time frame is changed. A ‘zoom to selection’ tool and panning tool allows users to quickly inspect data over a finer time/acceleration range. The time slider allows users to quickly pan through the 3Hr data file in 15 min windows.

4.1.2. Validation and results

The bridge testing was conducted over several days in September 2022 at the Sterling Bridge in Perth, WA. The IoT sensor, as well as wired accelerometers were placed inside the girder boxes in the centre of a span, to maximise the bridge vibrations. Data was collected (saved locally) without issue. Due to issues with signal strength for the wired sensors, 30-40 mins of data was collected roadside (see Figure 4.5).

Figure 4.5: Sensors placed roadside at Stirling Bridge. Highlighted in red is the IoT prototype.

Using the data collected, the responses of the IoT sensor (ADXL345) and the industrial wired sensor were compared. A time domain analysis proved to be difficult due synchronization issues in the two data sets. A frequency domain analysis was also conducted, and it was found that the ADXL345 has a much higher noise floor, but there are some aspects of the spectrum (<5Hz) that are particularly comparable. These results are shown in Figure 4.6. Of particular interest to the determination of bridge health is the frequency spectrum and corresponding fundamental frequency. On the basis of results, it was decided to acquire sensors with higher resolutions.

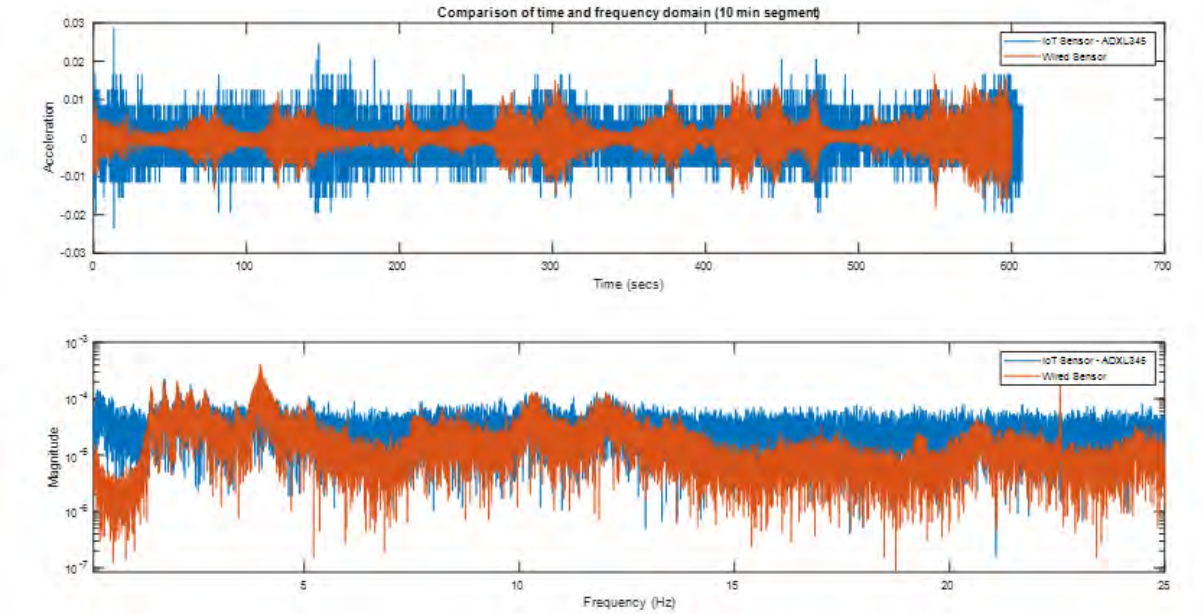


Figure 4.6: The time and frequency domain comparison of the IoT prototype and wired sensor from the roadside testing at Stirling Bridge.

In the next part of the project, the MMA8451 (with resolution of 0.25mg/LSB) and the KX132 (with resolution of ~0.1 mg/LSB) accelerometers have been tested. In particular, the MMA8451 has a resolution of 0.25mg/LSB. This sensor was compared with the onboard accelerometer of smartphone in a basic lab environment. The corresponding time and frequency domain results are shown in Figure 4.7. Across the 0-50Hz range, the spectrums have similar spectral characteristics

(most notably the oscillations in the 0-3Hz range)– much more so than the comparison between ADXL345 and wired sensors. In addition, the MMA8451 has a marginally lower noise floor than the smartphone sensor.

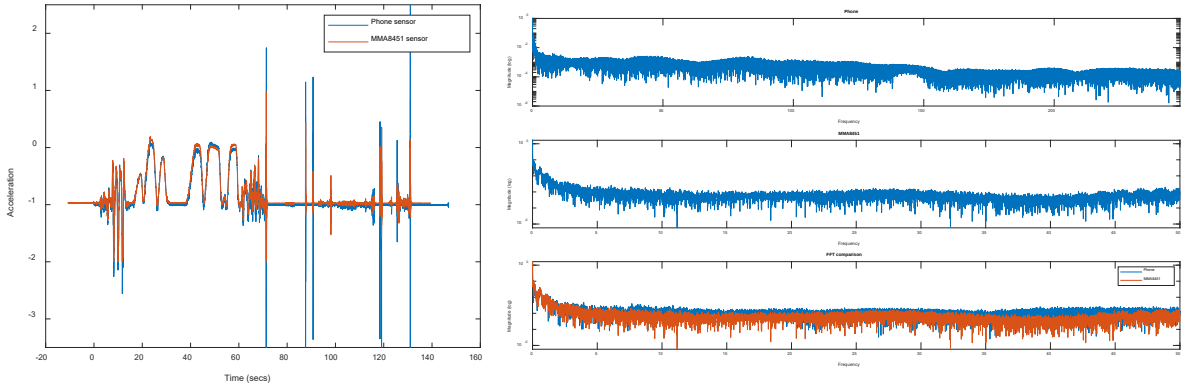
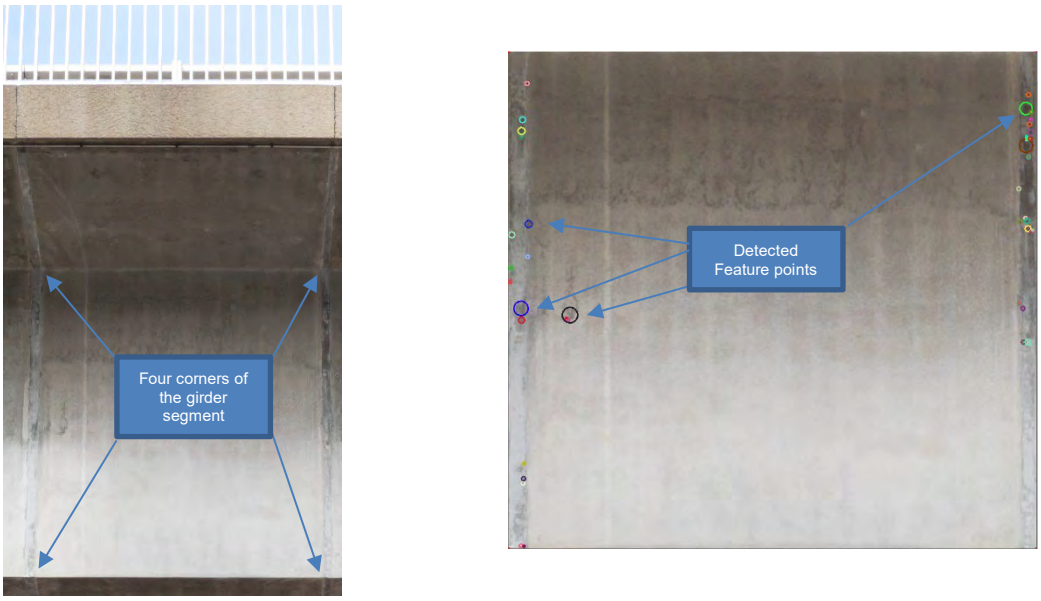


Figure 4.7: The time domain (left) and frequency domain (right) comparison of the MMA8451 sensor and Smartphone accelerometer.

4.2. Computer vision-based displacement measurement

The video recorded by the camera beneath the bridge is processed frame-by-frame to extract the dynamic displacement responses of the bridge, using the steps outlined in Figure 2.2. The first step is to determine the projective transformation matrix and scaling factor from the first video frame. The transformation matrix is used to create the projection from the 3D structural coordinate system to the 2D image plane, while the scaling factor is used to convert camera pixels to real-world coordinate units (mm). Figure 4.8(a) and Figure 4.8(b) show the first frame of a video and its projective transformation, respectively. To obtain the projective transformation, the coordinate of four points in the original frame are obtained by interacting with Python OpenCV, as shown in Figure 4.8(a), and the transformation matrix is calculated based on the coordinate of these four points in physical coordinates. In this report, the four corners of the girder segment side wall are selected. According to the design drawing of Stirling bridge, the width and height of the girder segment side wall are about 3048mm and 3305mm, respectively. The frame view after projective transformation is shown in Figure 4.8(b). The scale factor between the pixel unit and physical engineering unit (mm) is then calculated as 3.3605 mm/pixel. The feature points in the first frame, detected by the SIFT algorithm after projective transform, are shown in Figure 4.8(b). The size of the circle is proportional to the response value of the key point, and the key points are mainly distributed on the edge of the frame with a distinct texture. The feature points between any two adjacent frames are matched using the Fast Library for Approximate Nearest Neighbours (FLANN)-based matcher, and the matching results are presented in Figure 4.9, showing that all feature points have been correctly matched during the tracking process.



(a) (b)

Figure 4.8: Camera calibration (a) original frame; (b) the frame after projective transformation

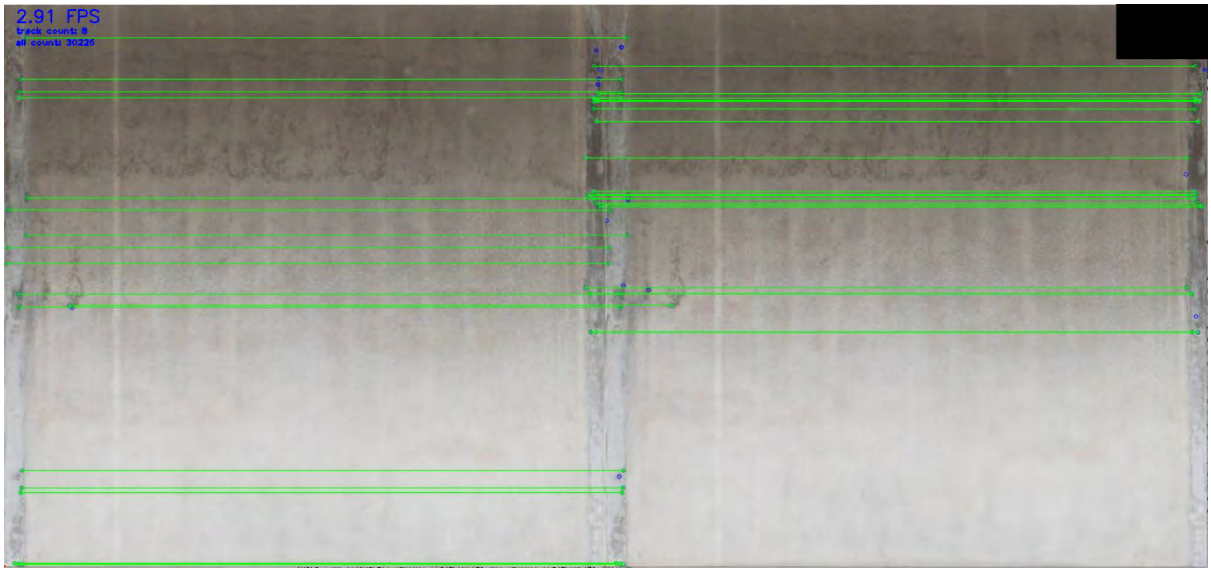


Figure 4.9: key points matching between first and sequential frames.

The bridge dynamic displacement subjected to moving vehicle loads can be continuously extracted from the video stream by following the procedures presented in Figure 2.2. In laboratory, the accuracy of computer vision-based displacement identification accuracy can be verified by comparing it with the LVDT displacement sensor. However, the measurement range of commercial LVDT sensor is usually limited within 1 meter, which is difficult to be installed in-situ to measure the dynamic displacement responses of long-span bridge. The above limitation of traditional LVDT displacement sensor motivated researchers to develop computer vision-based displacement

tracking techniques. Instead, an alternative solution is to reconstruct the dynamic displacement by double integration of the acceleration responses. The acceleration-based displacement reconstruction has been defined as a minimization problem with Tikhonov regularization with the following objective function

$$\prod(\bar{u}) = \frac{1}{2}\int_{T_1}^{T_2}\left(\frac{d^2\bar{u}}{dt^2} - a\right)^2 dt + \frac{1}{2}\beta^2\int_{T_1}^{T_2}\bar{u}^2 dt \tag{8}$$

where $\bar{u}$ and a represent the reconstructed displacement and measured acceleration responses, respectively. β is the Tikhonov regularization factor. Eq. (8) can be solved based on the fact that the first variation of the function should be zero. Detailed derivation process of solving Eq. (8) is available at Ref. [9]. Figure 4.10 shows the accuracy function of acceleration-based displacement reconstruction subjected to different target accuracy α_T. As can be observed in Figure 4.10, the reconstructed displacement is inaccurate in the lower frequency region. This fact can also be explained from frequency domain. In frequency domain, the frequency response function (FRF) of double integrator is -ω^2, where ω represents frequency. Therefore, the FRF between displacement responses and acceleration responses can be expressed as

$$H_{\bar{u}a} = \left|\frac{FFT(\bar{u})}{FFT(a)}\right| = \frac{1}{\omega^2} \tag{9}$$

As suggested by Eq. (9), the reconstructed displacement amplitude can be infinite for the frequency component close to 0, which is manifested as a trend or drift pattern in the time domain. The vehicle-induced bridge displacement response consists of two parts, namely, quasi-static component and modal vibration component. For a long span bridge considered in this project, the vehicle-induced bridge displacement response is dominated by the quasi-static component, which is mainly distributed in the lower frequency region. As a result, it is challenging to accurately reconstruct the dynamic displacement from acceleration measurement. The above statement is also supported by Ref. [9].

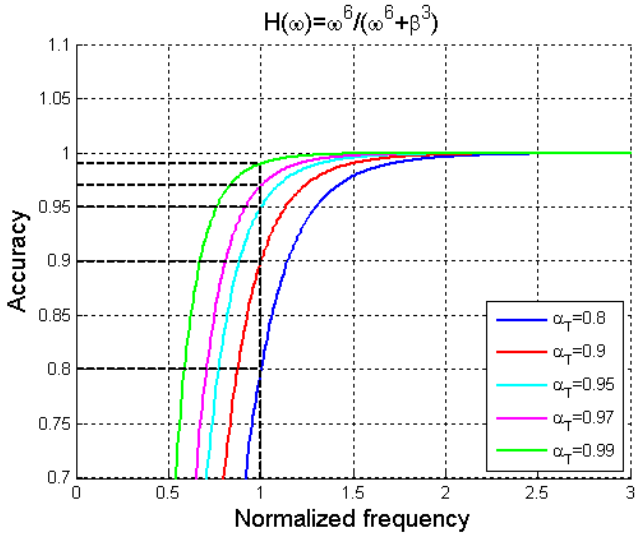
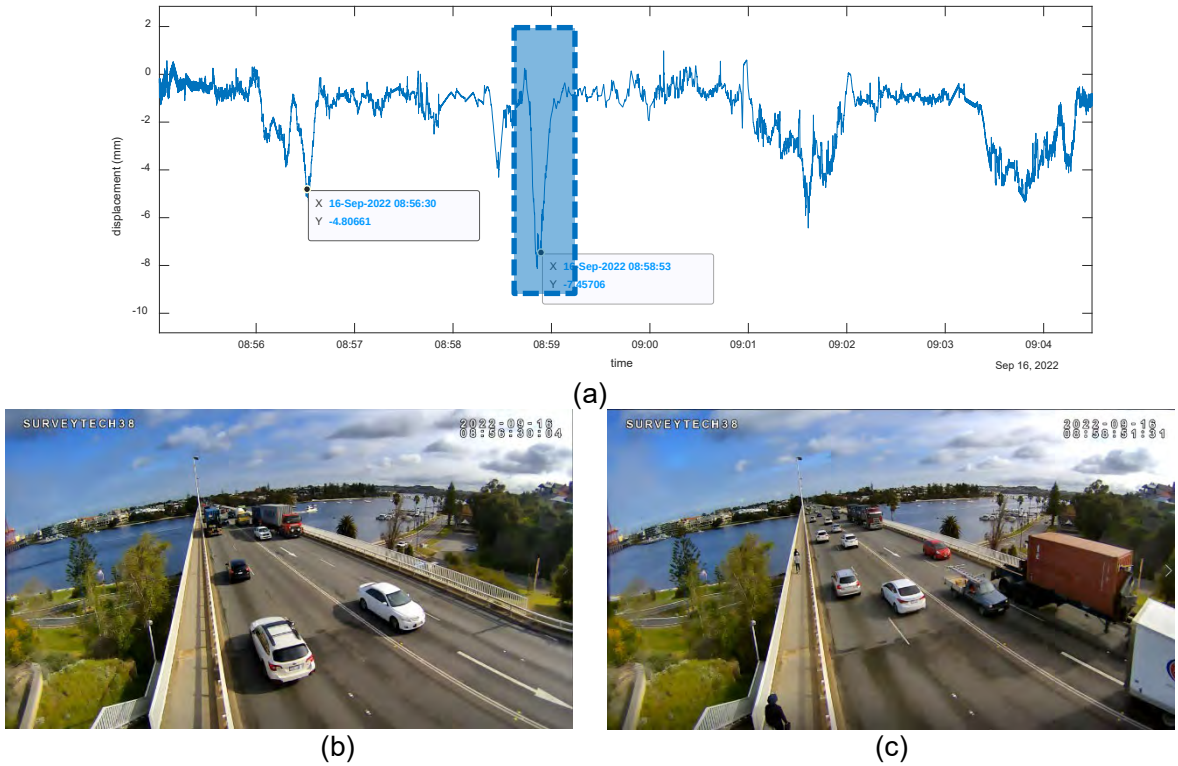


Figure 4.10: Accuracy function for different target accuracy

Since the ground truth of bridge displacement responses subjected to traffic load is not available, the researcher attempted to indirectly verify the rationality and correctness of the developed target-free computer vision-based displacement tracking algorithm. Firstly, the identification results are compared with the traffic flow recorded from the traffic camera. Secondly, the displacement

influence line theory is introduced to further validate the rationality and correctness of the displacement identification results.

The identification results corresponding to some representative traffic patterns are presented in Figure 4.11 and Figure 4.12. Overall, the bridge displacement responses subjected to operational conditions are mainly induced by the traffic load. As evidenced by Figure 4.11 and Figure 4.12, the bridge displacement responses reach the valley value at the time instant when the heavy vehicle passes the measurement point (the middle of second span from the south abutment). It is interesting to notice that the shape of displacement curve corresponding to traffic patterns from a different direction is different. In particular, when the heavy vehicles are mainly distributed on the Fremantle-Perth city direction traffic line (as highlighted with the blue box in Figure 4.11 and Figure 4.12), the displacement at the left side of the valley value is larger than that of the right side. In contrast, when heavy vehicles are mainly distributed on the Perth city-Fremantle direction traffic line (as highlighted with the green box in Figure 4.12), the displacement at the right side of the valley value is larger than that of the left side. The above phenomenon can be explained by the bridge influence line theory [10]. An illustration of the influence line of measurement section subjected to moving load is presented in Figure 4.13. When vehicle passes the adjacent span (the first and third) of the second span, the moving load induce a positive influence on the displacement responses of the measurement section. Since the third span is three times longer than the first span, the positive influence when vehicle moves on the third span is much more significant than that of the first span. As a result, the bridge displacement responses slightly increase, and then decrease to the valley value when vehicle passes the bridge from Fremantle to Perth city direction. This phenomenon can indirectly validate the correctness and accuracy of the proposed target-free displacement tracking algorithm.

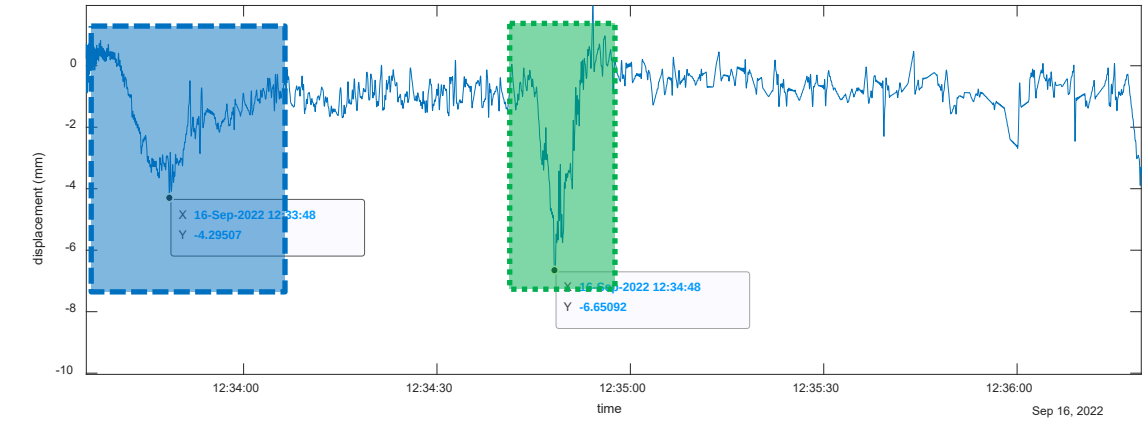


(a)

(b)

(c)

Figure 4.11: (a) bridge dynamic displacement between 8:55-9:05 Sep 16, 2022; (b) traffic pattern at 08:56:30 Sep 2022: heavy vehicles distributed on both direction traffic line; (c) traffic pattern at 08:58:53 Sep 2022: heavy vehicles mainly distributed on the Fremantle-Perth city direction traffic line.



(a)

(b) (c)

Figure 4.12: (a) bridge dynamic displacement between 12:33-12:37 Sep 16, 2022; (b) traffic pattern at 12:33:48 Sep 2022: heavy vehicles mainly distributed on the Fremantle-Perth city direction traffic line; (c) traffic pattern at 12:34:48 Sep 16 2022: heavy vehicles mainly distributed on the Perth city-Fremantle direction traffic line.

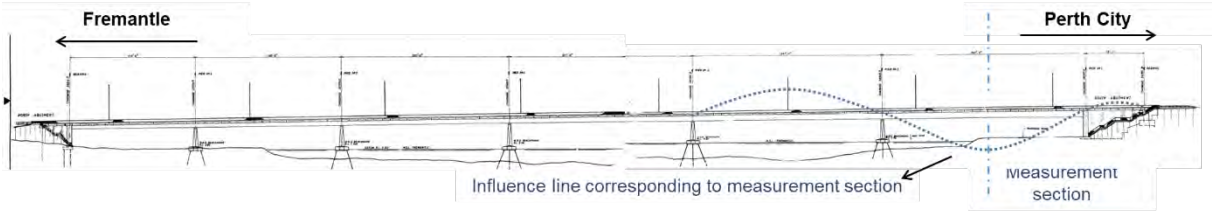


Figure 4.13: illustration of the influence line of measurement section subjected to moving load.

Figure 4.14 shows the histogram of bridge displacement responses during the time period of 08:54:47 to 13:54:09 on September 16, 2022, along with their 95% and 99% confidence intervals. Figure 4.15 presents the traffic pattern corresponding to the top 10 identified displacement responses. The largest displacement response, with a value of 10.39 mm, was identified at 13:16:56. Our analysis reveals that heavy trucks were present on the bridge deck when the largest displacement responses were observed. It should be noted that the vehicle-induced peak displacement is influenced by various factors such as vehicle weight, speed, type, road roughness, and so on. In the future study, it is suggested to collect additional data to calculate the displacement amplification factor (DAF) and analyse the effect of vehicle type, vehicle speed on DAF.

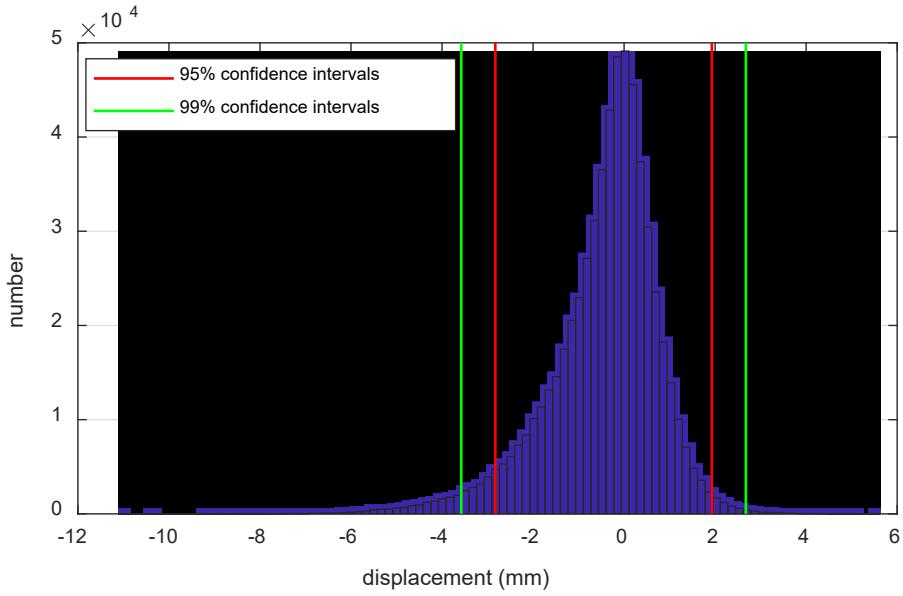


Figure 4.14: histogram of bridge displacement response during 08:54:47-13:54:09 Sep 16 2022.

(a) 10.39mm

(b) 8.46mm

(c) 8.27mm

(d) 8.11mm

(e) 7.91mm

(f) 7.90mm

(g) 7.62mm

(h) 7.57 mm

(i) 7.46 mm

(j) 7.18 mm

Figure 4.15: traffic pattern corresponding to top 10 identified displacement.

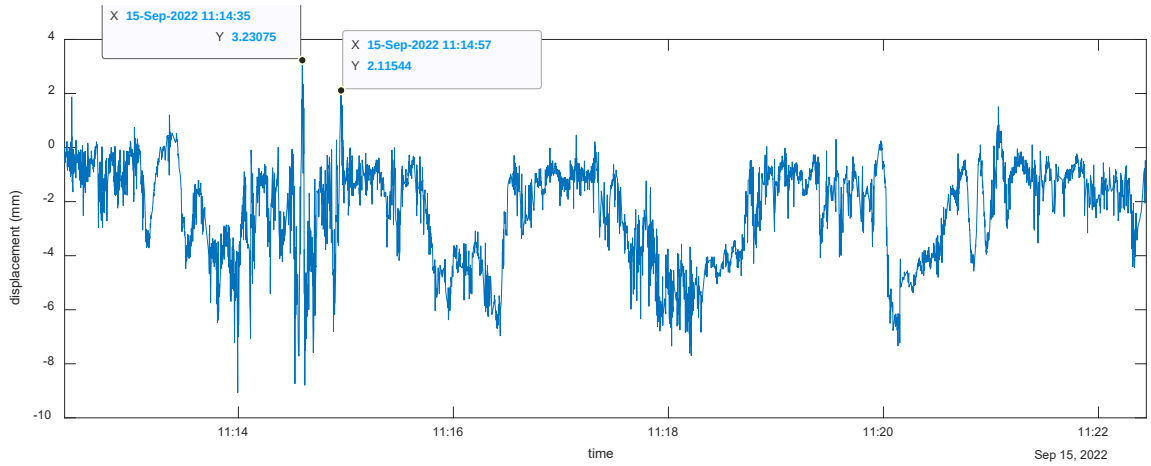
4.2.1. The effect of wind-induced camera motion

Wind-induced camera motion can significantly affect the accuracy of vision-based displacement tracking. When the camera is subjected to wind-induced motion, the captured images can be blurry or distorted, making it challenging to track object displacement accurately. To evaluate the impact

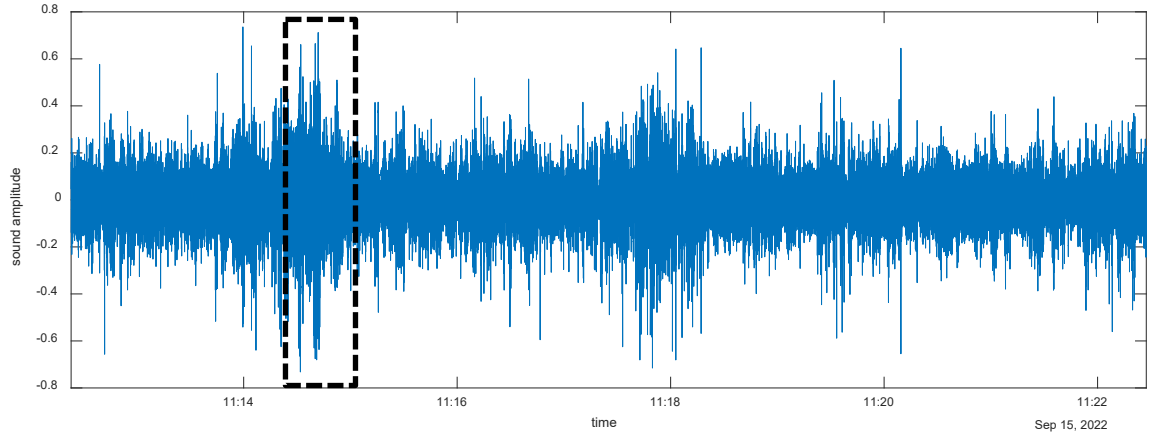
of wind-induced camera motion on bridge displacement identification, we analyzed the video recorded between 11:12am-12:22pm on September 15, 2022, a time period characterized by relatively strong winds during the field test. Figure 4.16 shows the displacement identification results, sound wave, and traffic condition corresponding to this period. The sound wave was extracted from the video, and positive displacement peaks were observed in the vision-based displacement identification result at two instances, 11:14:35 am and 11:14:57 am, respectively. However, sudden upward movement of the bridge during operational conditions is unlikely. Figure 4.16(b) shows that the amplitude of the sound wave corresponding to the highlighted black box is greater than that at other times, indicating strong wind during this period. Moreover, heavy vehicles were not present on the bridge deck during these two instances, as shown in Figure 4.16(c-d), making it highly likely that the positive displacement peaks were induced by wind-induced camera motion. The displacement identification error caused by camera motion is of the same order of magnitude as that caused by vehicle-induced bridge displacement.

In the field application, the following suggestions are given to alleviate the wind induced camera motion and displacement identification error:

(i) Relatively heavy and solid camera tripod can be used.
(ii) A camera case can be used to cover the video camera, and thus avoid the wind effect.
(iii) Developing signal processing technique to remove the camera motion induced displacement identification error.



(a)



(b)

(c) (d)

Figure 4.16: (a) bridge dynamic displacement; (b) the audio wave and (c) traffic condition corresponding to 11:12am-12:22pm Sep 15, 2022.

4.2.2. The effect of light condition

Light conditions have a significant impact on the accuracy of vision-based motion tracking systems that use cameras to track the motion of objects. The quality and clarity of the images captured by the camera are critical factors that affect the accuracy of the tracking. In low light conditions, images captured by the camera may be grainy, noisy, or blurry, making it difficult for the system to accurately track object motion. Similarly, in very bright light conditions, the camera may capture overexposed images with limited detail and contrast, making it challenging for the system to track motion accurately. To evaluate the impact of light conditions, we compared the quantity and quality of SIFT feature points detected under three representative lighting conditions: front lighting, backlighting, and direct sunlight on the camera lens. We selected two instances for each lighting condition and present the SIFT feature points detection results and the corresponding response values in Fig. 4.17 and Fig. 4.18, respectively. Our findings indicate that the best feature detection results were achieved under structural front lighting conditions, followed by backlighting. However, very few feature points could be detected when the camera lens was under direct sunlight.

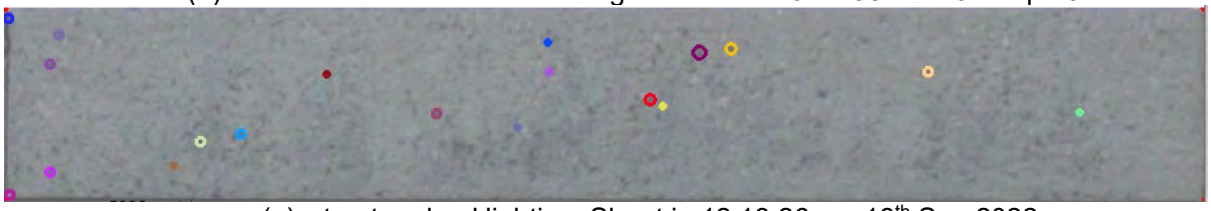
To avoid such situations, we recommend adjusting the shooting angle, shielding the camera, or avoiding shooting during strong direct sunlight in field applications.

(a) Structure front lighting. Shoot in 08:54:47 am 16th Sep 202

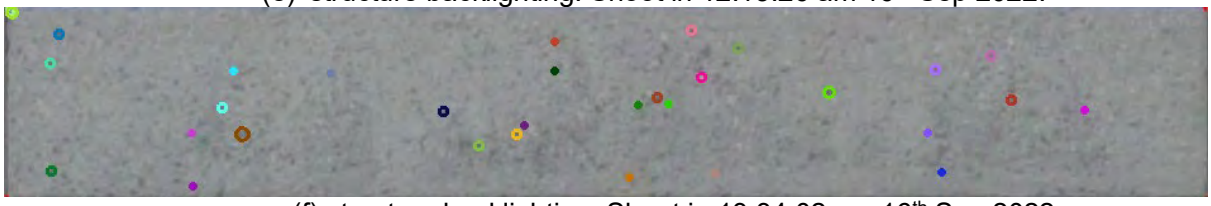
(b) Structure front lighting. Shoot in 09:25:00 am 16th Sep 2022.

(c) camera lens under direct sunlight. Shoot in 09:52:26 am 16th Sep 2022.

(d) camera lens under direct sunlight. Shoot in 10:12:35 am 16th Sep 2022.

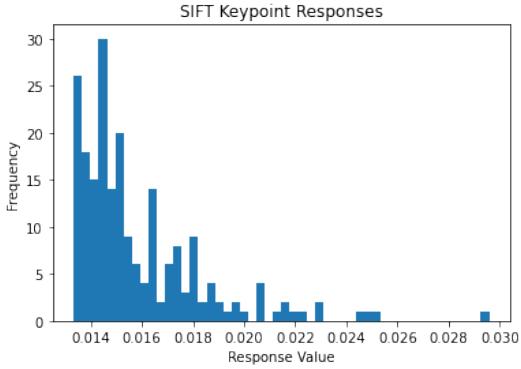


(e) structure backlighting. Shoot in 12:13:26 am 16th Sep 2022.

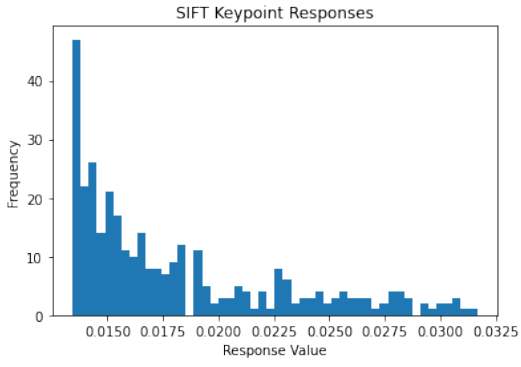


(f) structure backlighting. Shoot in 13:34:02 am 16th Sep 2022.

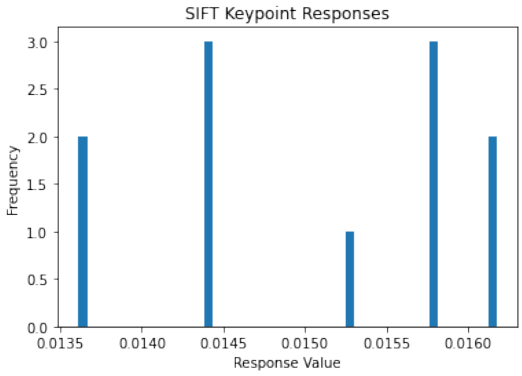
Figure 4.17: feature points detected from (a-b) Structure front lighting condition; (c-d) camera lens under direct sunlight condition; and (e-f) structure backlighting condition.



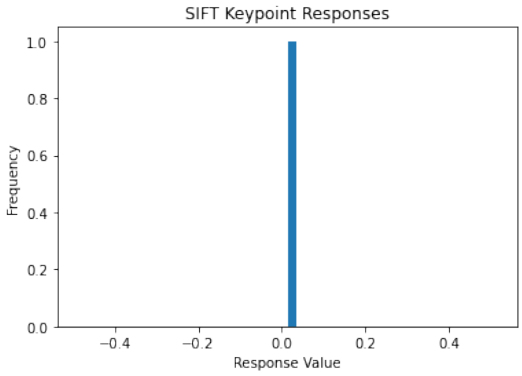
(a) Structure front lighting. Shoot in 08:54:47 am 16th Sep 2022.



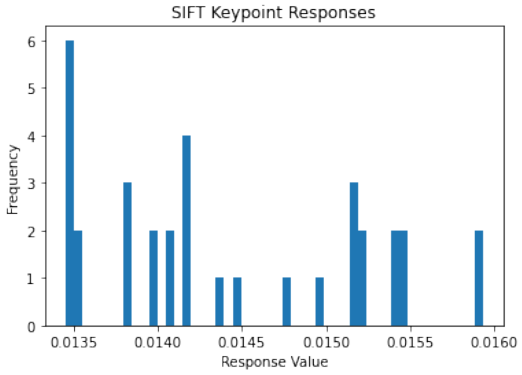
(b) Structure front lighting. Shoot in 09:25:00 am 16th Sep 2022.



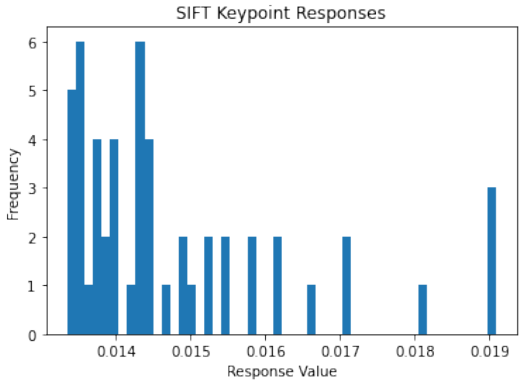
(c) camera lens under direct sunlight. Shoot in 09:52:26 am 16th Sep 2022.



(d) camera lens under direct sunlight. Shoot in 10:12:35 am 16th Sep 2022.



(e) structure backlighting. Shoot in 12:13:26 am 16th Sep 2022.



(f) structure backlighting. Shoot in 13:34:02 am 16th Sep 2022.

Figure 4.18: response value of feature points detected from (a-b) Structure front lighting condition; (c-d) camera lens under direct sunlight condition; and (e-f) structure backlighting condition.

5. Conclusions

This research project has developed IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction. The IoT sensor phototype, computer-vision-based displacement identification algorithm and predictive machine learning model are validated via laboratory testing, and full-scale field tests on a real highway bridge. The findings in relation to the IoT and computer vision components of the project have been presented in this report and are summarized as follows:

5.1. IOT sensor-based vibration measurement

The project demonstrated the applicability of cost effective IoT based accelerometer sensors for data collection, transfer over IoT link and storage at AWS S3. A dashboard is developed to give insight in the collected accelerometer data. The time domain and frequency domain comparison of high resolution IoT accelerometer against smartphone accelerometer show comparable results, though a thorough comparison against wired accelerometer is needed.

5.2. Computer vision-based displacement measurement

Researchers have developed a target-free compute-vision-based approach as a substitute for conventional displacement sensors to measure bridge displacement responses in a contactless manner. This method involves camera calibration and scale factor determination, natural feature target identification and description, feature matching and tracking.

Laboratory tests results verified that the developed target-free vision-based method can accurately and effectively identify structural displacement responses by using the natural feature on the structural surface. In particular, the correlation coefficient and RMSE corresponding to the worst identification result is 0.9734 and 0.3843, respectively. The average analysed frames per second by using the developed algorithm is about 21 fps.

Further in-situ validation results revealed that the vision-based displacement subjected to traffic load aligns well with the traffic pattern and is explainable by the bridge displacement influence line theory. However, the accuracy of vision-based displacement identification can be affected by environmental factors such as wind-induced camera motion and light conditions.

To mitigate the effects of wind-induced camera motion and displacement identification errors, the following are recommended: i) use a relatively heavy and solid camera tripod; ii) utilize a case to cover the video camera to avoid wind effects; iii) develop signal processing techniques to eliminate camera motion-induced displacement identification errors. To mitigate the inaccuracy resulted from poor light conditions, it is recommended to adjust the filming angle, shield the camera, and avoid filming during the times of intense direct sunlight.

6. Recommendations for further research

Based on the outcomes of this research effort, there are several areas where further research can be considered to fully realize the potential of the developed techniques in managing bridge infrastructure:

- The IoT based sensor data transmission to AWS needs further investigation to reduce the transmission time. The development of ML model on the basis of IoT sensor data in comparison with the wired accelerometer also needs further research.
- Graphical User Interface (GUI) for efficient decision making: A GUI can be developed to integrate and visualize the data collected from IoT sensors, video cameras, and traffic flow surveillance cameras in near-real time. This can provide a user-friendly interface for bridge inspectors and engineers to interpret the data and make informed decisions about bridge maintenance and repair strategies.
- Definition of bridge performance indicators and safety thresholds: Bridge performance indicators and safety thresholds should be defined to estimate the performance of bridges and facilitate decision-making. This can involve identifying key parameters such as structural health, load capacity, and environmental conditions, and establishing appropriate performance indices and safety thresholds based on industry standards and best practices.
- Campaign monitoring of bridges in different conditions: Despite the good results obtained from this research project, monitoring of more bridges under different conditions is necessary to develop a comprehensive database of experience. This can involve monitoring a larger number of bridges in Western Australia and collecting data on various factors such as traffic loads, environmental conditions, and maintenance history to further validate and refine the developed methodologies.

In summary, the results of this project provide a solid foundation for the development of integrated IoT and computer vision for smarter bridge health monitoring and prediction. Further research and expansion of this project can enhance the applicability of the developed methodologies to a wider population of bridges in Western Australia and beyond. The areas of research mentioned above can contribute to the advancement of bridge infrastructure management practices and enable more effective decision-making for ensuring the safety and performance of bridge structures.

References

[1] E. Makita, Replacement of timber bridges in Western Australia, Australian Small Bridges Conference, 9th, 2019, Surfers Paradise, Queensland, Australia, 2019.
[2] Z. Peng, J. Li, H. Hao, Structural damage detection via phase space based manifold learning under changing environmental and operational conditions, Engineering Structures, 263 (2022) 114420.
[3] B.F. Spencer Jr, V. Hoskere, Y. Narazaki, Advances in computer vision-based civil infrastructure inspection and monitoring, Engineering, 5 (2019) 199-222.
[4] Y.H. Hong, S.G. Lee, H.S. Lee, Design of the FEM-FIR filter for displacement reconstruction using accelerations and displacements measured at different sampling rates, Mechanical Systems and Signal Processing, 38 (2013) 460-481.
[5] C.-Z. Dong, F.N. Catbas, A review of computer vision–based structural health monitoring at local and global levels, Structural Health Monitoring, 20 (2021) 692-743.
[6] D. Feng, M.Q. Feng, Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection–A review, Engineering Structures, 156 (2018) 105-117.
[7] D. Ribeiro, R. Calçada, J. Ferreira, T. Martins, Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system, Engineering Structures, 75 (2014) 164-180.
[8] Y. Xu, J.M. Brownjohn, Review of machine-vision based methodologies for displacement measurement in civil structures, Journal of Civil Structural Health Monitoring, 8 (2018) 91-110.
[9] B. Spencer, F. Gomez, J. Park, H. Yoon, F. Moreu, Reference-Free Estimates of Railroad Bridge Displacement Under Revenue Service Traffic, United States. Department of Transportation. Federal Railroad Administration, 2020.
[10] Z. Peng, J. Li, H. Hao, Long-term condition monitoring of cables for in-service cable-stayed bridges using matched vehicle-induced cable tension ratios, Smart Struct. Syst, 29 (2022) 167-179.

Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction

Final report part B: prototype Machine Learning models for predicting bridge displacement and vibration

June 2023

APPENDIX B

Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction

Final report part B: prototype Machine Learning models for predicting bridge displacement and vibration

Prepared by

Sergio Banchero, Chao Sun

Keywords

Computer vision, IoT sensor, machine learning, bridge health monitoring

Version control

Draft 1.0

Project No

iMOVE Project 1-046: Milestone 3b

Project steering committee

Raquib Hossain, Main Roads Western Australia

Steve Atkinson, Main Roads Western Australia

Sebastian Davies-Slate, WALGA

Sharon Biermann, The University of Western Australia

Jun Li, Curtin University

Wensu Chen, Curtin University

Chao Sun, The University of Western Australia

Atif Mansoor, The University of Western Australia

Acknowledgment

This research is funded by iMOVE CRC and supported by the Cooperative Research Centres program, an Australian Government initiative.

About PATREC

The Planning and Transport Research Centre (PATREC) is a collaboration between the Government of Western Australia and local universities, constituted to conduct collaborative, applied research and teaching in support of policy in the connected spaces of transport and land use planning. The collaborating parties are: The University of Western Australia, Curtin University, Edith Cowan University, Department of Transport, Main Roads Western Australia, Western Australian Planning Commission and the Western Australian Local Government Association.

Publisher

Planning and Transport Research Centre
The University of Western Australia (M087)
35 Stirling Highway, Crawley, WA 6009
+61 8 6488 3385
patrec@uwa.edu.au

Table of Contents

Executive summary[1]

Note: the main body of this report only includes the Machine Learning parts of the project but the project executive summary covers all three components.

Regular maintenance programs are essential for the long-term preservation of bridges, and health monitoring plays a vital role in ensuring the effectiveness of these preservation efforts. However, the vast geographical expanse of Western Australia presents a significant challenge for conducting physical inspections, as only 28% of the state's 3,000 bridge structures are located within the Perth metropolitan area. Consequently, Main Roads is limited to infrequent manual inspections, sometimes occurring only once every few years. While timber and timber hybrid bridges, accounting for approximately 40% of the bridge stock, can be easily instrumented to produce health indicators, there are unique challenges associated with instrumenting reinforced concrete and prestressed concrete bridges, which make up about 51% of the stock. This situation increases the likelihood of missing the optimal time window for addressing structural issues, leading to higher maintenance costs and a heightened risk of significant structural failures.

This research aims to explore alternative approaches for monitoring the health condition of concrete bridges. It consists of three components, each emphasising one of the three essential indicators for bridge health monitoring: bridge displacement, vibration characteristics, and traffic load. The Stirling Bridge was chosen as the test site by the stakeholders.

Bridge displacement is important to understand how the structure responds to different traffic loads, especially heavy loads. However, conventional contact-type displacement sensors, such as the linear variable differential transducer (LVDT), require a stationary reference point that is often difficult to find in the field. Additionally, their short measurement range of less than 1 meter limits their application to large-span bridges. As an alternative, a computer vision-based method developed by the Curtin team was tested. Lab tests demonstrated good agreement with LVDT sensor data. Due to the limited range of LVDT, direct on-site validation was challenging, and only indirect checking of the data was possible. The observations were in line with the displacement influence line theory, and the top ten detected displacements correspond to heavy traffic patterns recorded by the traffic camera, indicating the reliability of the collected data. During the test, the physical dimensions of the bridge were needed to translate the number of pixels in the video to real distances. As the side of the bridge was inaccessible for physical measurement, design drawings were used instead. The field test also identified that environmental factors such as wind-induced camera motion and lighting conditions could affect the accuracy of results. The research team proposed possible solutions for future research to address these challenges.

Vibration data can be used to detect variations in the bridge's natural frequency over time, serving as an alarm for potential structural damage. A prototype IoT unit was built to measure bridge vibration using an accelerometer. The data was successfully transferred to the AWS S3 cloud storage over an IoT link, and a dashboard was created for accessing the collected accelerometer data. A comparison with an industry-grade wired sensor showed that the chosen IoT sensor lacked sufficient resolution, leading to the purchase of a better model. Although the field test opportunity was missed, lab tests demonstrated comparable results between the new IoT sensor and a smartphone accelerometer. The prototype has overall demonstrated the cost-effectiveness of IoT for collecting bridge vibration data when IoT data coverage is available. Further considerations

[1] The main body of this report only covers the Machine Learning part of the project but the project executive summary covers all three components.

include addressing power supply issues in remote areas and determining optimal attachment locations for the unit on the bridge.

The Machine Learning part of the project set out to predict the bridge's vibration and displacement responses given the observed traffic loads from videos. Both tasks were proven to be difficult, although the vibration prediction achieved comparatively better results. The available literature indicates the inherent difficulty of mathematically inferring bridge displacement from vibration data. We conducted extensive testing using various Machine Learning models, yet the challenge persisted, further confirming the complexity of the task. The 'phantom displacements' in the data (see Figure 15) also contributed to the difficulty. Nevertheless, the model estimated influence curve (Figure 14) seems reasonable, which suggests that it has captured some of the underlining mechanisms. Interestingly, combing the vibration data with traffic videos did not increase the accuracy.

The primary objective of Main Roads is to devise a practical solution that enhances bridge health monitoring. The computer vision and IoT solutions concentrate on distinct metrics, and both exhibit promising potential in this preliminary investigation. As we progress, engaging stakeholders in discussions becomes crucial to explore the practical implementation of these solutions within Main Roads' operations. This collaboration will facilitate the development of a well-defined path for further advancement.

1. Overview

This report summarises the progress made on the machine learning (ML) part of milestone 3 of the iMOVE project 1-046 (Integrated IoT, computer vision and machine learning technologies for smarter bridge health monitoring and prediction).

1.1. Background

Existing bridge structures exposed to the operational environment for long service life, are prone to performance degradation owing to material deterioration, natural hazards, and human-made loading conditions [1]. Their vertical displacement under traffic loads is usually selected as a critical parameter for evaluating bridge performance [2]. Conventional contact-type displacement sensors such as the linear variable differential transducer (LVDT) require a stationary reference point, which is often difficult to find in the field. Furthermore, the measurement range of traditional displacement sensor is relatively short, which limit its application to large-span bridge structures [3].

To address the limitations of current sensor systems for field applications, the research community has been actively exploring new technologies that can advance the state of the practice in structural health monitoring (SHM). The Curtin team focuses on field-testing the camera-based noncontact vision sensing solution, which has emerged as a promising alternative to conventional contact sensors for structural dynamic response measurement and health monitoring. Significant advantages of the vision sensor include its low cost, ease of setup and operation, and flexibility to extract displacements of any points on the structure from a single video measurement.

Meanwhile, the UWA teams focus on proof-of-concept dashboard visualisation of vibration data transmitted over IoT communication links and ML models predict bridge displacement for different levels of traffic load.

1.2. Project Objectives

The overall aim of the project is to demonstrate the feasibility of using a combination of IoT, computer vision, and machine learning technologies to reduce the need for manual bridge inspections and provide ongoing monitoring that supports smart bridge maintenance. In this project, the development, application and integration of IoT technology-based sensor data collection, computer vision methods and machine learning techniques will be investigated to enable remote monitoring of bridge structural health to enhance the efficiency of asset monitoring, management and maintenance. The specific project objectives are:

1. To develop IoT technology-based solutions for collecting, visualizing and transmitting vibration response data;
2. To develop computer vision-based solutions for measuring bridge displacement under a range of traffic loads;
3. To develop an integrated proof-of-concept predictive model based on machine learning techniques for predicting bridge displacement under heavy traffic loads; and
4. To make recommendations on new techniques for efficient bridge monitoring, the wider roll-out of the technology for bridge health monitoring and prediction and suggest the next steps.

2. Data Collection

To establish the relationship between traffic loading and bridge vibration that it causes, their corresponding datasets needed to be collected at the same time.

2.1. Vibration Data

Two wired Kistler 8330A3 accelerometers and an IoT accelerometer were installed at the same location to measure the bridge's vertical vibration responses. Details can be found in Section 3.2 of the Milestone 2a report.

2.2. Traffic Data

Traffic survey company SurveyTech was subcontracted by the UWA team to collect videos using their mast-mounted cameras. Figure 1 shows that four cameras were originally planned, plus two sets of tube counters were originally planned. However, given that this is still a prototyping stage, only CAM3 and CAM4 were deployed to cut down the cost. CAM1, CAM2 and tube counters could be reconsidered if the current data suggest there is a benefit to using them.

Figure 1 Traffic camera locations

The cameras were set up at 7:30 pm Tuesday 13th September 2022 and removed at 6:30 am Saturday 17th September 2022.

3. Machine Learning Models

3.1. Data Processing and Cleaning

Captured video data was processed using UWA's video analytics (VA) software to extract the following information:

- Vehicle counts;
- Vehicle class: our VA software groups vehicles into five categories, Austroads Class 1, Class 2, Class 3 – 5, Class 6 – 9, Class 10 – 12;
- Vehicle location (on screen)
- Vehicle location (on the bridge plane, in Euclidean distance units)
- Vehicle speed

Vehicle locations on the bridge surface are estimated using a homographic transformation. To do this, a homography matrix is estimated by selecting and matching a set of key features on the camera frame (e.g., the edges of some dashed-line road markings) and on a calibrated aerial image. After calibrating the matrix, it is possible to approximate the location on the bridge, along its longitudinal and traversal axes, of an object detected by the camera. Note that there are several factors that can introduce errors in this approximation: The homography transformation does not

factor in the non-linear distortion presented by the camera, and the vehicle location on the ground is estimated from a rectangular bounding box on the screen coordinates, which can be affected by the vehicle shape, size, and relative position and angle to the camera.

Figure 2 – Key points used to estimate the homographic transformation

Due to the nature of the scene, which is a long bridge being recorded with a low-altitude fixed camera, vehicles that are further from the camera than a certain limit cannot be reliably detected by the detection and tracking algorithm. This is due to the relative vehicle size on the screen, the resolution of the camera, and an increasing number of occluding objects as the vehicles move further away from the camera. However, vehicles might still affect the bridge displacement and vibration even from a considerable distance from the measuring point. To mitigate the unreliable detection at further distances, an extrapolation process takes place to estimate the position of vehicles after they were last detected (in the case of vehicles moving away from the camera) and before they are first detected (in the case of vehicles moving towards the camera). The extrapolation method assumes that vehicles stay in their first or last detected lane and travel at the average lane speed. Vehicle positions in time are then estimated until they reach a maximum distance (e.g. 200m away from the bridge displacement measurement point).

The processed data were then fed into the ML models to predict bridge vibration and displacement. Only a subset of the data was selected based on the results and available “target” data, as accelerometer and computer vision displacement data were only produced for a shorter period than the video data.

Our analysis found that the IoT accelerometer did not have enough precision for our modelling so we switched to the wired Kistler 8330A3 accelerometer data. It meant we had a significantly less amount of data to work with because the wired sensors were only run for several hours in total. The limited data means we could not fully calibrate the model but it was sufficient for developing prototypes. The IoT team has ordered a higher-precision sensor that will be used for the next stage of data collection.

3.2. Bridge Vibration Prediction

A variety of Machine Learning (ML) models was tested. These include Random Forest, Scalar Vector Regression, and Polynomial Regression. Random Forest performed the best therefore this is the model used for the comparison in this report. The prototypes consist of different data inputs which were tested on all the ML models explained above.

Prototype 1 – 10 features

Prototype 2 – 25 features

Prototype 3 – 576 features

3.2.1. Accelerometer Data Processing

The wired accelerometer collected data at a rate of 100 Hz. The data was saved into individual files of 1-hour increments creating a total of 60,000 sensor readings per file.

To reduce the impacts of noise in the sensor data, Short Term Fourier Transform (STFT) was performed on the accelerometer data. It is more appropriate for analysing frequency components that vary over time than trying to predict the exact acceleration at any given time. Two main parameters are modifiable, the segment time and the overlap. The segment time is how long each segment in the transform is. The overlap is how much of the previous signal should the next one overlap.

For prototype 1 and 2, the segment time of 2 seconds and overlap time of 1 second was chosen. For prototype 3, the segment time of 5 seconds and an overlap of 2.5 seconds were chosen.

As the sensor sampling frequency of 100 Hz, it gives an output of 0 to 50 Hz in 0.5Hz increments. An example output of the STFT on a 60,000 acceleration data sample can be seen in Figure 3. The lighter the colour, the more ‘energy’ at that frequency level.

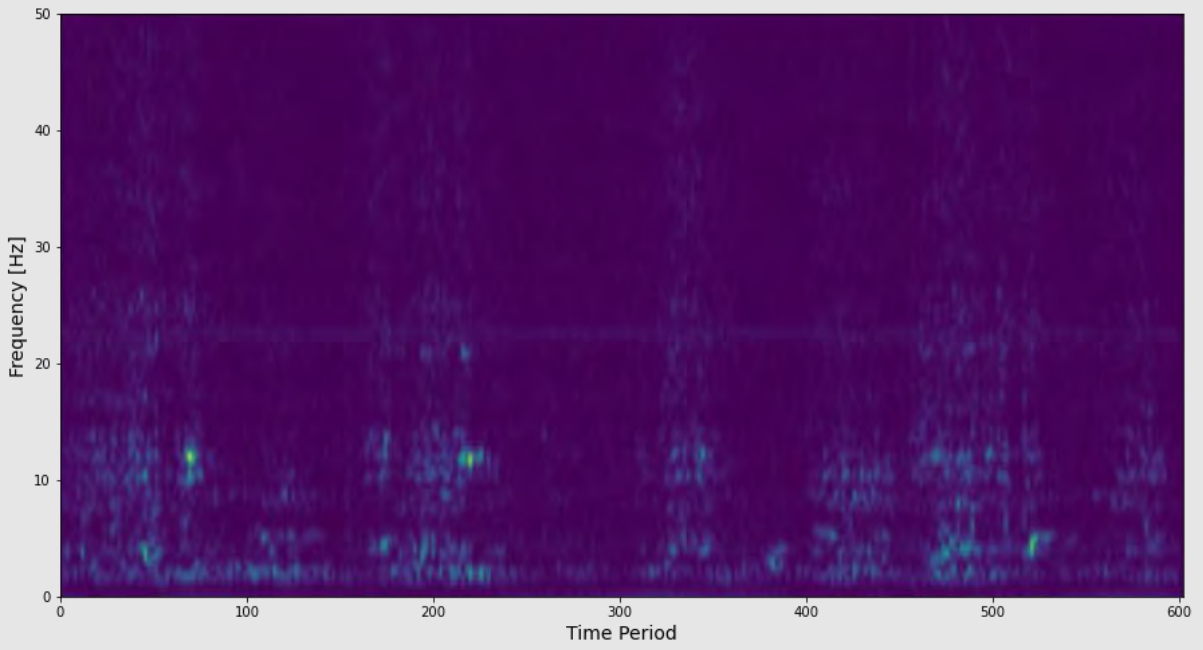


Figure 3 - Output of the Short-Term Fourier Transform

As can be seen, there is a lot of variation in the frequency around 10 – 15 Hz.

Prototypes 1 & 2

The output for these prototypes was the values at the 12 Hz frequency. The output of an STFT is a complex number. To make for an easier prediction, the absolute value of the radius was calculated and used as the expected output.

Prototype 3

Instead of using just one frequency, frequencies between 5 and 30 Hz were used. For every time step, the output of the STFT between those frequencies was summed up, and this value was normalised between 0 and 1. This value was called '*Displacement Energy*'.

3.2.2. Vehicle Analytics Data Analysis

As explained above, there are five vehicle groups predicted from the VA software. Table 1 shows the Austroads classes and their correlation to the input data class number.

Table 1 - Vehicle Groups and Associated Austroads Classes

VA Vehicle Groups	Austroads Class Number/s
0	1
1	2
2	3-5
3	6-9
4	10-12

Acceleration Data vs Vehicle Data

Figure 4 shows the normalised acceleration values of five consecutive data files (light blue). Performing the STFT and performing the Displacement Energy calculations results in the black line.

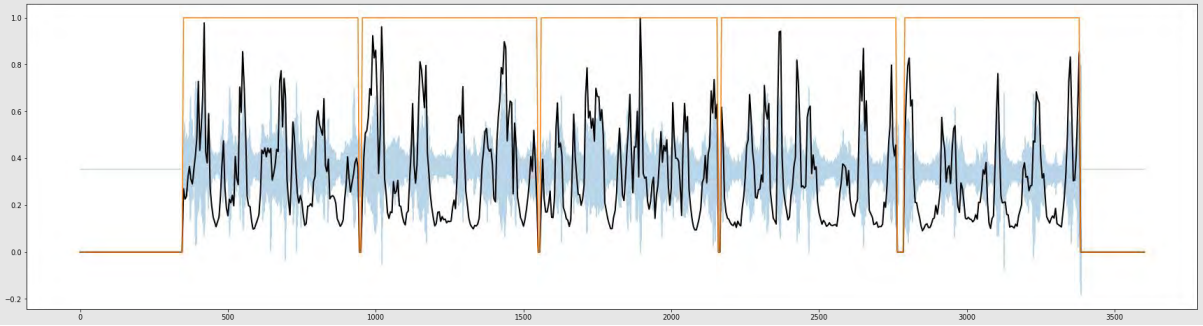


Figure 4 - Accelerometer and STFT Graph

As expected, the STFT follows the acceleration data but is smoother.

Analysing the vehicle data can give an idea of the correlation and potential lag between the vehicles and the displacement energy. As Figure 5 shows, there is generally a relationship between an increase in displacement energy when there are more vehicles. It also shows that at some points, the displacement energy lags the number of vehicle counts. This could be either due to differences in classes (less number of vehicles but a higher proportion of trucks) or that vehicles from earlier in the detection have a longer effect on the bridge.

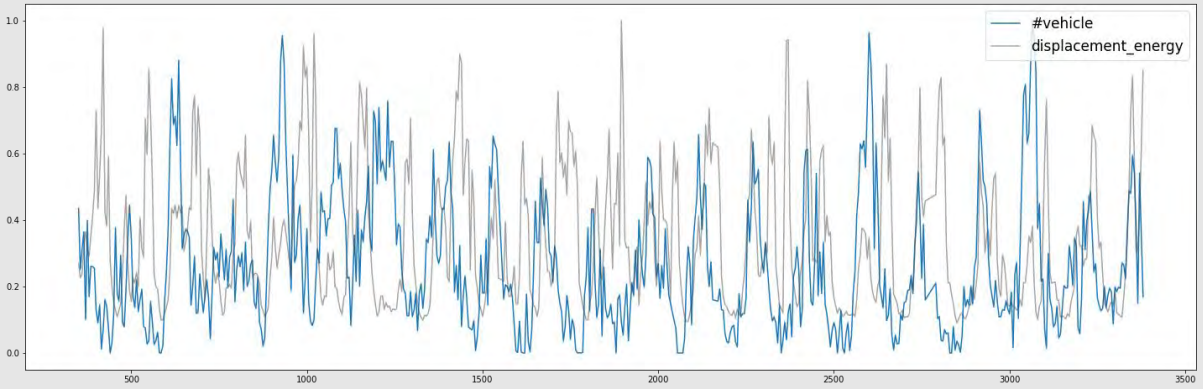


Figure 5 - Vehicle Counts and Displacement Energy

The location of the vehicle detection would also influence the vibration size and hence displacement energy. It is hypothesised that vehicles detected closer to the sensor would have a greater effect on the displacement energy than vehicles further away. Figure 6 shows the location of vehicle detections. The sensor is located on the right side (the exact location relative to the camera is unknown).

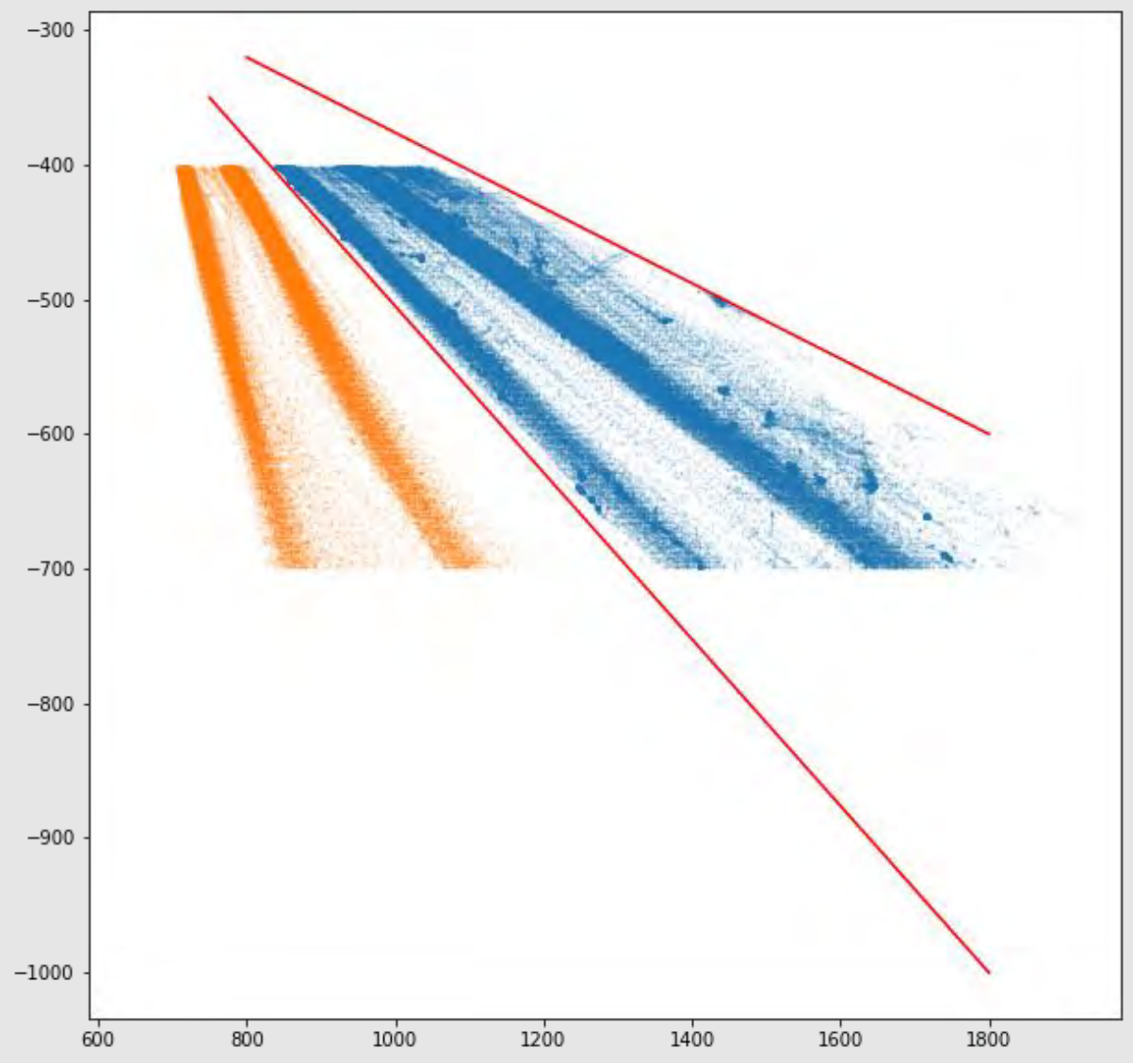


Figure 6 - Location of detected Vehicles

The hypothesis that the closer the vehicle is to the sensor, the larger the displacement energy effect can also be extended so that the detections do not just have a left or right, but also which section they are in. 5 sections per side of the road were chosen for our analysis. Figure 7 shows the sections of the road.

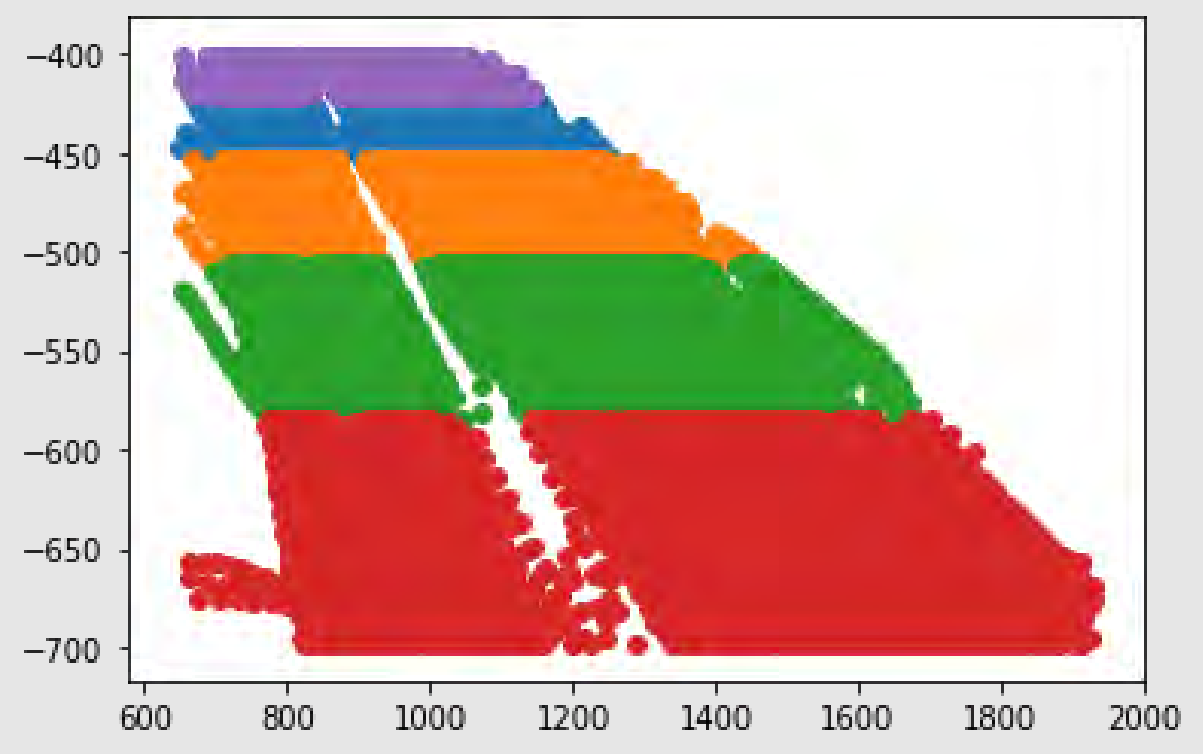


Figure 7 - Location of Vehicles in Segments

3.2.3. Vehicle Analytics Data Inputs

Prototype 1

For this prototype, for each period, all the speeds of each detected object are added to their specific class array. This array is then used to determine the number of vehicles in the class and the average speed of the vehicle class. This is done for each class and for every time step in the sensor file. An example output is below. As such, there is a feature size of 10 (5 classes x 2 variables per class)

Prototype 2

For this prototype, the data was put into 'bins'. For each time step and each class, detections are counted into bins based on their speed. The bin maximum value and the size of the bins can be varied. In our prototyping, a bin size of 25 pixels per second was used with a maximum bin size of 100 pixels per second. If the pixel per second value is above 100, it is placed in another bin. Based on the chosen parameters, the data input size would be 25 (5 classes with 5 bins each).

Prototype 3

Based on all the features discussed in Section 3.2.2 the data for this prototype was just the count of vehicles and the average speed of the vehicle based on the following criteria:

- Which group of vehicles (5 options)
- Which section along the road (5 options)
- Which side of the road (2 options)

This gives 25 possible class-section options and 10 possible class-side options. Additionally, total per class (5), per section (5), per side (2) and overall total (1) were used. Each option has 2 values (vehicle count and average speed). Furthermore, as the analysis suggested, the input at previous time points may affect the current time point, therefore, 5 previous points were also used as lagged variables (lagging 1, 2, 3, 5 and 8 time slots). The total number of features is therefore 576 ((25 +10 + 5 + 5 + 2 + 1) x x 2 x (5 + 1).

3.2.4. Results

Comparison

Table 2 shows that Prototype 3 has the highest R^2 value, which indicates better performance. R^2 was chosen as a comparison value as Mean Squared Error (MSE) could not be used. This is because the output for prototypes 1 & 2 is the frequency value whereas the output for prototype 3 is the normalised 'displacement energy'.

Table 2 - Comparison Between Prototypes

Prototype Number	R^2
Prototype 1	-0.07
Prototype 2	-0.06
Prototype 3	0.52

Prototype 1

576

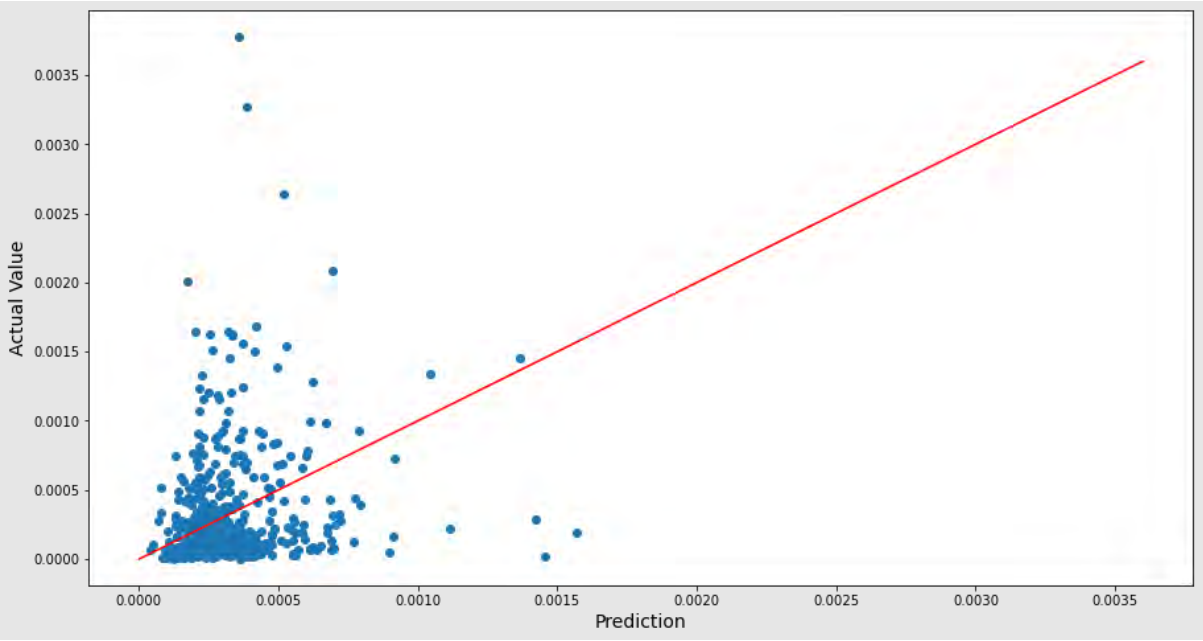


Figure 8 - Actual vs Predicted for Prototype 1

Figure 8 shows the underfitting of the model. Most of the predictions are between 0 and 0.0005. The figure shows that a lot of these predictions are not close to the correct line (red line).

Prototype 2

As with Figure 8, Figure 9 shows the underfitting of the model. It is slightly better at predictions with the majority of predictions between 0 and 0.0075. The figure shows that a lot of these predictions are not close to the correct line (red line).

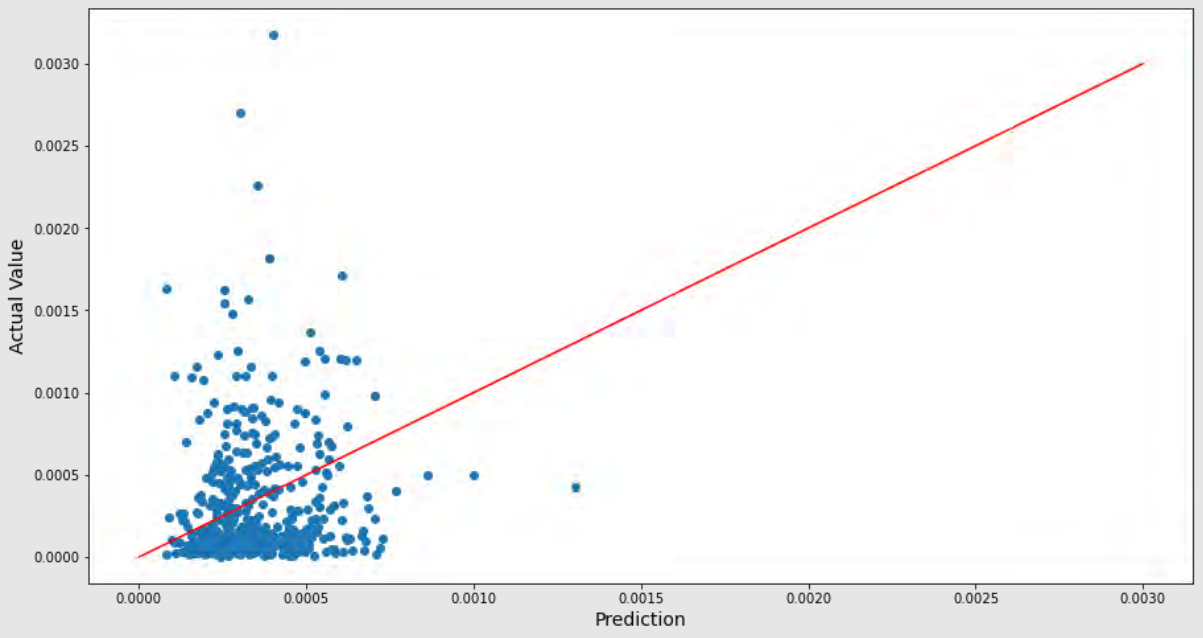


Figure 9 - Actual vs Predicted for Prototype 2

Prototype 3

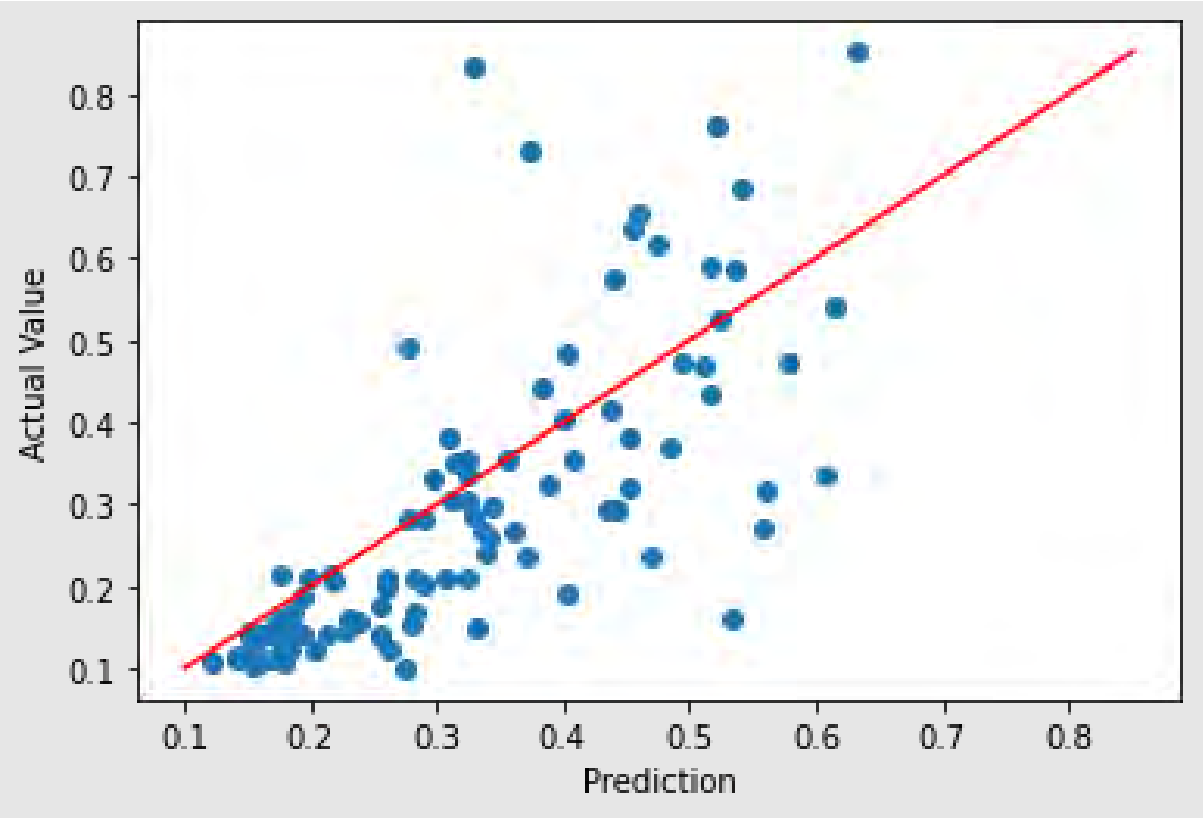


Figure 10 - Actual vs Predicted for Prototype 3

Figure 10 shows a better-performing model. Due to the difference in output, there is not a direct comparison between Figure 10 and Figure 8 & Figure 9. What can be seen is that the distribution of predictions vs actual value is spread out and not clustered like in Figure 8 & Figure 9. This shows that the model is not underfitting and suggested that more features (discussed in Section 3.2.5) could potentially improve the predictions.

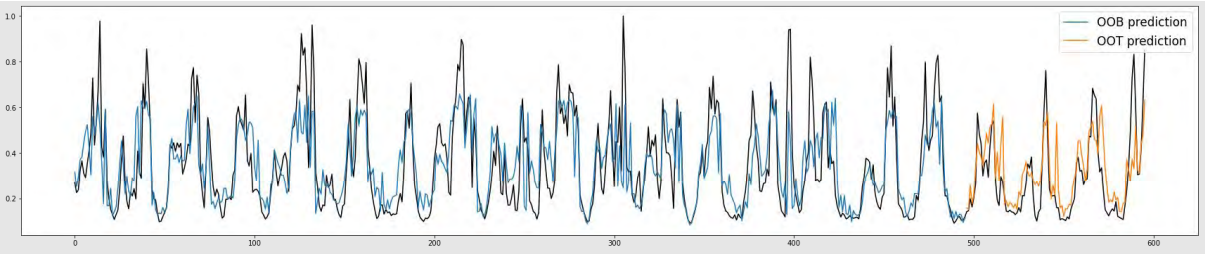


Figure 11 - Prototype 3 Time Series Predictions

Figure 11 shows the prediction of the trained Random Forest model from Prototype 3. The blue line is the estimation of the displacement energy by using the out-of-bag input values. The orange line is the estimation of the displacement energy based on the unseen values that were not used for training. As can be seen, both lines follow the expected values closely. It can be seen that the peaks are often understated by the predictions. This may be due to a lack of instances of high peak points as they are less frequent than the 'normal' vibration'. For example, a limited amount of times class 3 or 4 were in the video. Another possible reason is that it could be underfitting, and therefore predicting a lower value to reduce the error. This is further discussed in Section 3.2.5.

3.2.5. Discussion

Although not perfect, the ML model with 500 features (Prototype 3) performed significantly better than the initial ML models (Prototype 1 & 2). It suggests that our analysis of the positioning of vehicles and the lagging effects of vehicles is correct. We believe that the model can be improved with the addition of extra data points and new information such as wind speed and vehicle mass.

Wind speed is a factor that places a load on the bridge, which in turn, would cause displacement. A gust of wind could happen at a time when there is no significant movement of vehicles on the bridge. This could result in an acceleration value that is higher than predicted by the model as it does not know that information.

Vehicle mass is also a major factor in the displacement of the bridge. The vehicle mass is assumed to be learned by the ML model through the differentiation of classes. The ML model could be improved by deriving a better estimation of mass. A possible way to do this would be to have more classes of vehicles. If our Video Analytics program could identify each Austroad class and further separate large vehicles into empty vs. loaded, it may be able to learn a better approximation of mass. For this to work, a significantly more amount of data would need to be collected to ensure there are sufficient samples of each class as input.

Further to the above, in general, more data could improve the accuracy of the model. As events which have large spikes (Figure 4) are not as frequent as low-level vibration, there are relatively few data points to correctly learn how to predict such events. Having more data allows for more instances of large spikes in vibration and therefore a greater ability for the ML model to correctly learn how to predict such large vibration events.

3.3. Displacement Prediction

The objective of this effort is to estimate the average displacement (vs a reference point) in a discrete time window given the status of the bridge as observed by one or multiple fixed cameras recording traffic conditions.

Ground truth displacement values were calculated through the Curtin Team computer vision-based method (see more details in the Milestone 2a report). The results of this method were given to the UWA team as the displacement values. These are then aggregated in discrete time windows of 1 second and used as the target value for ML prototype models to predict displacement.

Displacement models are evaluated using time-stratified 5-fold cross-validation. To elaborate on the process, initially, all available data is sorted by time. Once sorted, the data is then split into 5 folds, each containing an equal number of data points. For each fold, a model is fitted with the remaining 4 folds, which means that the model is trained with 80% of the data. Once the model is fit, it is evaluated using the remaining 20% of the data in that fold. This process is repeated five times, with each fold being used as the evaluation set once. By using this method, we can ensure that the model is evaluated on all available data, while also mitigating any influence that time might have on the results. Metrics reported are the average value between the 5 folds, together with the standard deviation. Evaluation metrics are the coefficient of determination (R^2), and the Mean Squared Error (MSE).

Two models were developed and tested for the displacement prediction problem: A physics-based linear model and a Machine-learning “black box” model.

3.3.1. Models

Physics-based Linear Model

The physics-based linear model (PLM) assumes that the displacement on the vertical axis of a cross-section of the bridge can be estimated as the sum of the individual actions of all the vehicles on the bridge at that given point in time. Individual vehicle actions are modelled as the product between a ***location factor***, which depends exclusively on the vehicle location along the bridge main axis, and a ***weight factor*** that depends exclusively on the estimated vehicle weight.

$$\hat{y}(t) = \sum_{\text{vehicles}} f(x_i) \cdot g(w_i) + y_0$$

For implementation purposes, the location factor is approximated through a regularized piecewise-constant function, i.e., the length of the bridge is divided into D discrete regions and a weight is assigned to each region. In particular, the length of the bridge was divided into 18 sections. The weight factor is estimated through a proxy constant that depends on the vehicle class, an attribute that is provided by the deep-learning detection model employed to extract the traffic conditions from the video footage and can be one of C=5 possible classes (see Data Processing and Cleaning).

Data is then discretised into a sparse 3D occupancy-matrix V where the axes are the time slot, length section, and vehicle class. This matrix is filled with the average vehicle count in each cell. The model parameters to be adjusted are then the weights of each vehicle class (G), the weight for each bridge section (F), and a constant bias y_0.

$$\hat{y}(t) = \sum_{c}^{C} \left(G_c \cdot \sum_{x}^{D} F_x V_{t,x,c} \right) + y_0$$

Parameters G, F and y_0 are optimized utilizing the Nelder-Mead method. The cost function to be minimized is minus R^2 of y and $\hat{y}$, plus a regularization factor over the shape of the location weight term. This regularization factor penalizes the first and second discrete derivates of F_x. This regularisation can also be considered a smoothing factor of the shape of F_x.

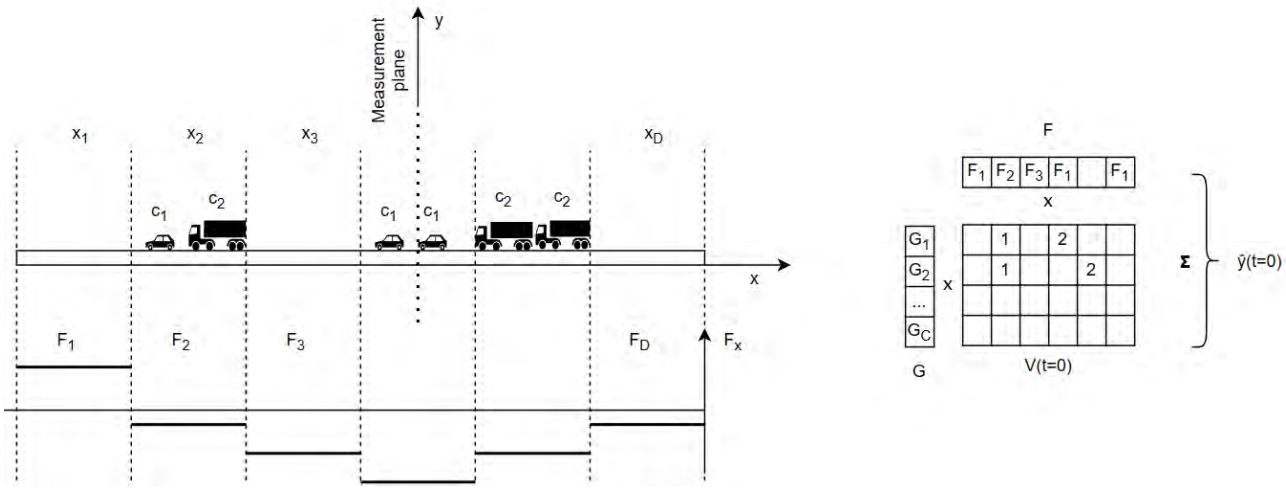


Figure 12 - Diagram of the factors affecting the physics-based linear model, with vehicles from multiple classes on different locations of the bridge (left). Matrix representation of the diagram, showing the occupancy matrix V for the specific time displayed on the diagram, and weighting factors G and F.

Machine Learning Model

In a similar fashion to the vibration prototype three, a machine learning model based on Random Forest was implemented, in this case, trained to predict displacement. Features were obtained by calculating 3 metrics, namely vehicle count, average speed, and the average size of the detection bounding boxes, on combinations of three dimensions. These dimensions are:

- Which group of vehicles (5 options)
- Which section along the bridge (8 options)
- Which side of the bridge (2 options)

Furthermore, the input at previous time points may affect the current time point, and some misalignment between the timestamps of the video and the sensor might exist, therefore, lagged variables were also added. Selected combinations of these dimensions and metrics resulted in a total of 1008 features.

An additional variation of the ML model adding features extracted from the wired accelerometer signal was implemented. In this variation, different frequency components, average values and standard deviation of acceleration measurement per time slot were included as features, resulting in 1021 features in total.

3.3.2. Results

Physics-based Linear Model

The PLM model achieved an R^2 of 0.102 (std 0.063) and an MSE of 0.922 (std 0.190), both of which show a low level of performance. This is visually confirmed in Figure 13, which plots the predicted values against true values, with each point as a measurement. The diagonal line indicates a perfect match between the two so ideally the dots should align with it closely but it is certainly not the case here. Having said that, the 50th percentile regression line does show some alignment with the diagonal, suggesting that the results are not entirely random and the model captures some trend in the data. Furthermore, the other percentile regression lines show that residuals are close to homoschedastic, which is a desirable condition to model the distribution of the true displacement in terms of the expected displacement.

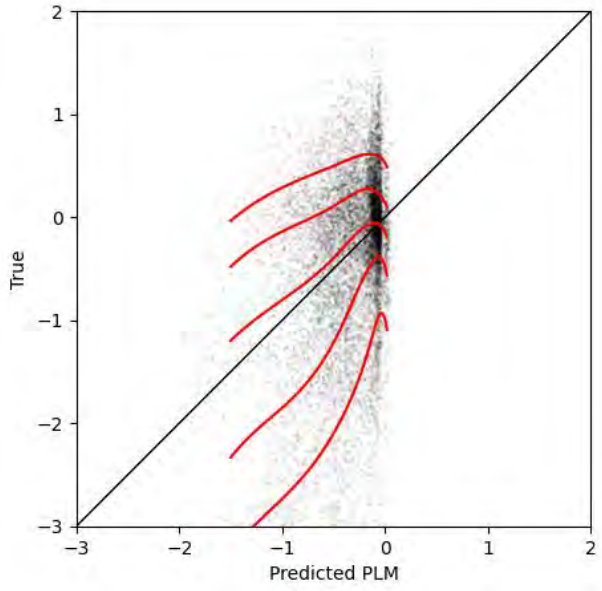


Figure 13 - Scatter plot of predicted vs actual for the PLM model. Red lines indicate the percentile regression splines (df=5) for percentiles 10, 25, 50, 75, and 90

Although predicting the exact displacements at any given time is difficult given the nature of the problem, the model estimated influence curve on the left of Figure 14 seems reasonable, which suggests that it has captured some of the underlining mechanisms. The vertical axis shows how vehicles influence the bridge displacement and the direction of influence, with positive values indicating upward movement and negative values indicating downward movement. The horizontal axis shows vehicle positions along the bridge, with point zero being the location where the displacement was measured. The curve suggests that vehicles will have the most downward pressure on the bridge when they are at point zero, and it reduces as they move away from it. When they move about 100 – 120 metres to the right, which is somewhere in the next span, they move point zero upward instead of downward, which is explainable by the displacement influence line theory. The bar chart on the right of Figure 14 shows that heavier vehicles move the bridge more than light vehicles, which is also sensible.

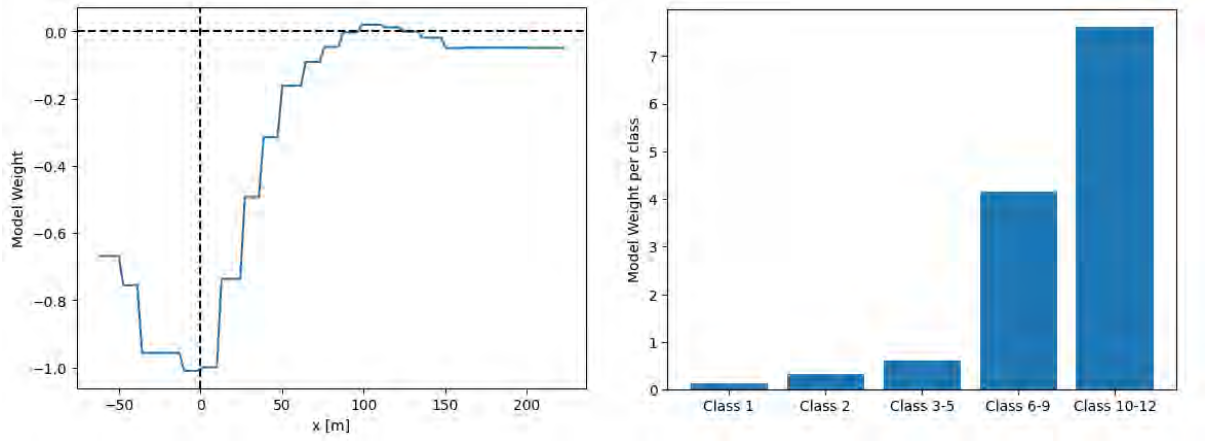


Figure 14 – Estimated vehicle influence on the displacement of the bridge along the longitudinal direction, origin represents the location of the displacement measurement (left). Estimated vehicle influence depends on the classification (right). Both charts show the average parameters between the 5 folds models.

The lines in the distance-time diagram in Figure 15 visualise how vehicles move through time along the bridge. Most observed large displacements on the right-hand side correlate to busier and heavier vehicular movements. However, there are also phantom displacements (shaded in blue) happening when there is no vehicle presence. Some possible explanations are:

- Distant vehicles that are not captured by our traffic cameras could still cause vibration waves mediated by the bridge structure;
- Residual displacements but some displacements are too large to be explained by this;
- Noise in the displacement data, which might be wind induced.

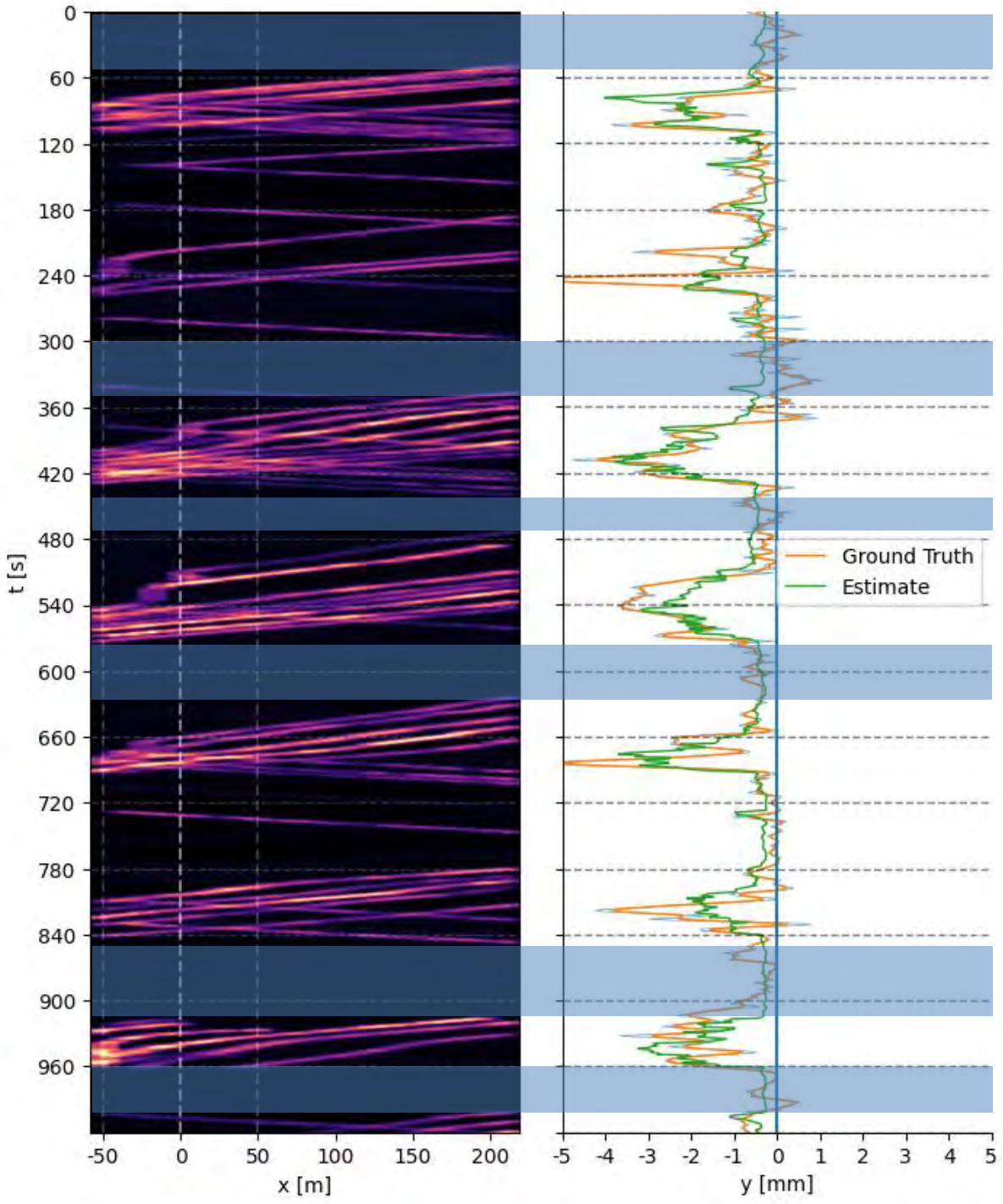


Figure 15 – A distance-time heatmap showing the estimated load, based on the weights assigned to each class, along the bridge throughout time (left); vertical marks show an approximate range of observable vehicles in the camera frame. The true displacement and estimated displacement correspond with the time of the heatmap (right).

Machine Learning Model

The Random Forest model achieved an R^2 of 0.123 (std 0.042) and an MSE of 0.901 (std 0.183). The alternative version including the accelerometer data as additional features achieved an R^2 of

0.124 (std 0.052) and a MSE of 0.936 (std 0.195). Figure 16 has a similar pattern to Figure 13, with no clear alignment between the true and predicted values.

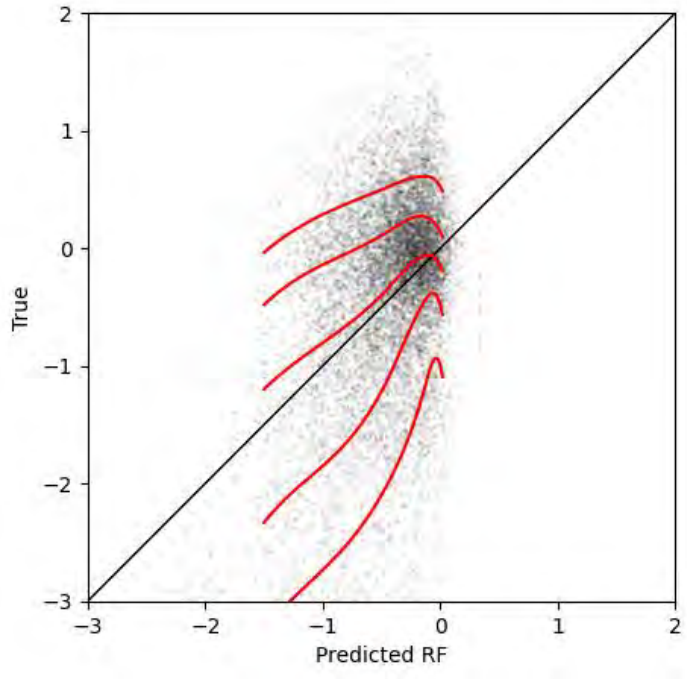


Figure 16 - Scatter plot of predicted vs actual for the Random Forest model. Red lines indicate the percentile regression splines (df=5) for percentiles 10, 25, 50, 75, and 90

Comparison

The two models have close performances according to Table 3 and Figure 17. The latter also shows that the models mostly struggle to predict the large displacements, most of which appear to be abrupt.

Table 3 – Performance comparison of the PLM and RF models

Prototype Number	R^2 (std)	MSE (std)
PLM	0.102 (0.063)	0.922 (0.190)
RF (base version)	0.123 (0.042)	0.901 (0.183)

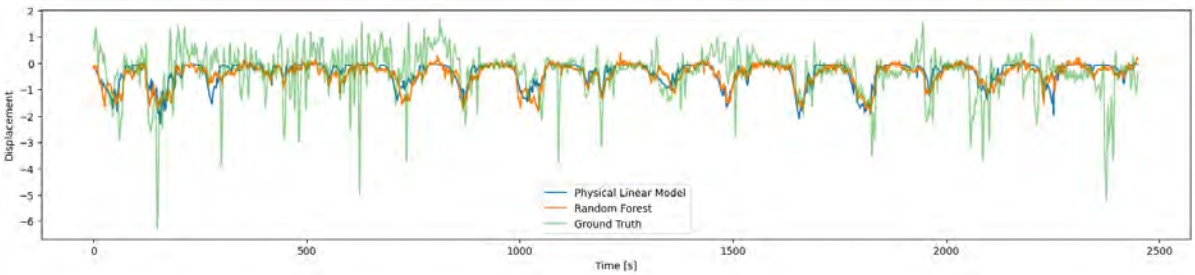


Figure 17 – Time series of one of the 5 folds comparing the ground truth displacement vs both models.

3.3.3. Discussion

It is apparent from the time trend charts that each of the folds presents some structurally different time series data, with considerable low-frequency shifts and different frequency oscillations with varying amplitudes. These large changes in the data along the time dimension make achieving a better model more challenging. A possible solution for this factor could be obtaining more data points to achieve more generalization power in the models. Figure 15 illustrates the fact that noticeable displacement variations appear even when no vehicle is shown in the frame (areas highlighted in blue).

One factor that might add to bias in the models is the effect of vehicles not visible or not identifiable in the frame, either by occlusion or merely by the distance to the camera, on the bridge's displacement. Utilizing multiple cameras along the bridge on both sides of it would greatly reduce or even eliminate this factor, which could be tested in the future.

Another factor is the difficulty in assessing the weight of the vehicles traversing the bridge. Even with perfect Austroads classification, trucks might have widely varying loads. We consider this could be partially mitigated with more labelled data, as it is possible to at least identify the presence or absence of load on the trucks, and the type of load (containers, other vehicles, tanks, etc.). These proxies could provide additional information to the estimation of the load on the bridge.

When comparing the linear model and the machine learning model, even though they are very different in nature, both models perform similarly in terms of performance metrics. It is interesting to note that there were situations in the validation sets where predictions from both models deviate considerably between them. Furthermore, in some cases, the linear model outperformed the machine learning model and vice versa. This leads to the belief that further refinement and combination of both techniques, for instance using stacking or blending, might yield better results.

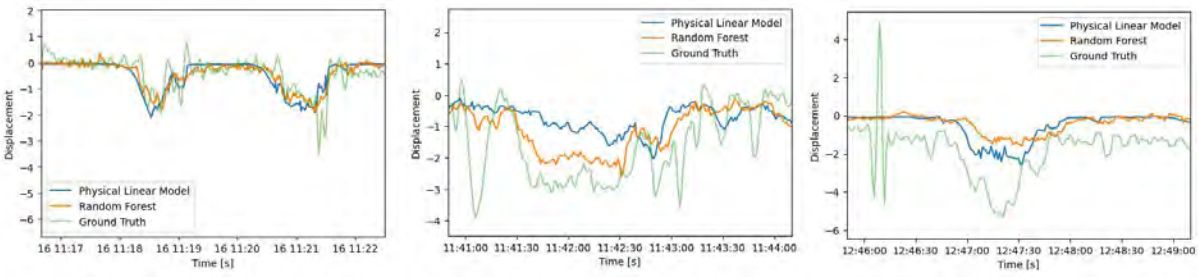


Figure 18 – Zoom in on different situations found on the validation set when comparing models: Both models properly predict the output (left), RF outperforming PLM (middle), PLM outperforming RF (right).

Given the parsimony and interpretability of the linear model, it seems like a sensible choice over the machine learning option for this task, considering their similar performances.

It is worth noting that the parameters of the physical model are well aligned with the expected theoretical values, as seen in Figure 14. The closer the vehicle is to the sensor the more it tends to displace the bridge down, and in the adjacent spans, they tend to displace the bridge up. Additionally, large vehicles (class 3 and above) have a much larger influence that the smaller classes in the estimated displacement.

Adding information from the accelerometer data provided no noticeable improvement in the RF model, (0.001 improvement in R^2 and decreased performance in terms of MSE). Furthermore, the

accelerometer-derived features ranked among the least important in terms of feature importance (based on a decrease in impurity).

References

[1] Z. Peng, J. Li, H. Hao, Structural damage detection via phase space based manifold learning under changing environmental and operational conditions, Engineering Structures, 263 (2022) 114420.

[2] B.F. Spencer Jr, V. Hoskere, Y. Narazaki, Advances in computer vision-based civil infrastructure inspection and monitoring, Engineering, 5 (2019) 199-222.

[3] Y.H. Hong, S.G. Lee, H.S. Lee, Design of the FEM-FIR filter for displacement reconstruction using accelerations and displacements measured at different sampling rates, Mechanical Systems and Signal Processing, 38 (2013) 460-481.