Infrastructure Monitoring from an In-Service Light Rail Vehicle

FINAL RESEARCH REPORT

Jacobo Bielak (PI), Hae Young Noh (PI), George Lederman,
Siheng Chen, James Garrett, Jelena Kovačevic

Contract No. DTRT12GUTG11

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT

In this project, we explore a data-driven approach for monitoring rail infrastructure from the dynamic response of a train in revenue-service. Presently, rail inspection is performed either visually or with dedicated track geometry cars. We examine a more economical approach where rail inspection is performed by analyzing vibration data collected from an operational passenger train. The high frequency with which passenger trains travel each section of track means that faults can be detected sooner than with dedicated inspection vehicles, and the large number of passes over each section of track makes a data-driven approach statistically feasible. With the financial and logistical support of the Technologies for Safe and Efficient Transportation University Transportation Center, we have deployed a test-system on a light-rail vehicle and have been collecting data for the past two years. The collected data underscores two of the main challenges that arise in train-based track monitoring: the speed of the train at a given location varies from pass to pass and the position of the train is not known precisely. We explore which feature representations of the data best characterize the state of the tracks despite these sources of uncertainty (i.e., in the spatial domain or frequency domain), and we examine how consistently change detection approaches can identify track changes from the data. We show the accuracy of these different representations, or features, and different change detection approaches on two types of track changes, track replacement and tamping (a maintenance procedure to improve track geometry), and two types of data, simulated data and operational data from our test-system. The sensing, signal processing, and data analysis we

present could facilitate safer trains and more cost-efficient maintenance in the future. Moreover, this approach is quite general and could be extended to other parts of the infrastructure, including bridges.

1. Introduction

While continuous monitoring of engineered assets has become cost effective in fields like manufacturing and aerospace, the majority of civil assets are still monitored by visual inspection [1,2]. This project focuses on railroad infrastructure, which is monitored by a combination of visual inspections and specialty track inspection vehicles, known as track geometry cars [3]. Inspection cars are expensive to operate, and as such, they are used to inspect the tracks infrequently [4]. A number of researchers have propose using vibration-based monitoring to assess the track profile as a lower-cost alternative to the type of optical sensors currently employed on track geometry cars [5–10]. All of these studies assume that the sensors are placed on the axle of the train so that they have a direct mechanical connection to the track. In addition, they assume the position of the train is known, typically achieved by requiring that the inspection car operate at a constant speed.

In this project, we explore a more economical approach: using operational trains and allowing the sensors to be placed inside the cabin [11]. By using operational trains, no dedicated monitoring vehicles are required, track closures are avoided, and tracks can be interrogated more frequently. The high frequency of interrogation not only means that issues can be flagged more rapidly, but also more passes over areas of interest can be collected, meaning greater statistical confidence about the

state of the track. Lastly, by allowing the sensors to be placed in the cabin, the sensors can be installed more easily and can be protected from the external environment.

There are numerous challenges that must be overcome for this method to become practical, two of which we explore in detail in this project. The first is that a train's speed over a section of track differs with each pass, so methods robust to train speed must be found for comparing data between passes. This is particularly challenging when the sensors are in the train's cabin, because the train's suspension filters the roughness from the track differently depending on speed. The second challenge is that we do not know the precise location of the train due to GPS noise, so position uncertainty must be considered in attempting to detect track changes from the vibration signal.

We learned about these challenges through our long-term monitoring project in partnership with the Port Authority of Allegheny County. We have placed accelerometers and a GPS antenna on an operational light-rail vehicle and have been collecting data over the last two years. We have focused on determining which analysis techniques are best suited for train-based monitoring and understanding why certain techniques work better than others. To gain insight into the vehicle-based monitoring problem, we also modeled a simplified version of the train-track interaction using a single degree-of-freedom oscillator travelling over a rough track. This parametric simulation allowed us to explore the effects of variable speed and position uncertainty in a controlled setting before validating on the light-rail data.

Simulation has three main benefits. First, we can generate much more data than we could collect from instrumented vehicles. Although we have been collecting data from the light-rail system for two years, there are relatively few recorded maintenance events in the rail-network each year that we can use to test our approach. With simulation, we can rapidly generate hundreds of track changes. Second, we can simulate a wider range of parameters than the narrow band we have observed i our operational system. For example, the light-rail vehicle has a natural frequency that only varies slightly (due to temperature/passenger loading), but we can simulate a much wider range of natural frequencies to ensure that the analysis techniques we find are general. Third, in the data we have collected, we do not have the ground-truth of the train's position. Through simulation, we can study the effect of position error, and propose techniques that work well for the level of error we expect in operational systems.

While our ultimate objective is to build a complete track-monitoring system, we focus in this project on two main components required for automated track monitoring: meaningful feature extraction from the raw vibration signals and detection of track-changes from these features. For the first component, we examined four different features, and then used supervised classification to determine which one provides the most reliable information about the state of the track. The fact that the train's speed varies between each pass (or the speed of the travelling damped oscillator in simulation) made comparing the data from multiple passes challenging. The signals are recorded in the time domain, but given that track changes are localized in space the spatial domain is often more useful for detecting

track changes. We examined features based on the time-domain signal, spatial-domain signal and frequency-domain signal as well as features based on the energy in the signal. We show the discriminate power of these features on data from our parametric simulation as well as on the light-rail dataset.

The second component in our automated detection system, change detection, is a challenging task because we do not know a-priori the type of change we are trying to detect [12]. The most closely related study of automated track anomaly detection from the vibration signal of a train is that done by Molodova et al. [13] where a system for detecting track squats (a type of rail-surface indentation) was proposed. In the study, a detection event was triggered by a vibration signal above a pre-defined threshold. We aim to build a broader detection system where anomalies are defined as changes relative to historical behavior. This ensures that areas with consistently high vibrations (like track switch gear or joints) are not labeled as anomalies, while changes, even in areas with low vibration-amplitude, are detected. For example, we will show that our method detects changes in track geometry due to tamping that can have safety implications despite their small amplitude. This historical detection technique allows for the monitoring of an entire network without manually tagging problematic areas as the method in Molodova et al. requires.

To perform change detection, we experimented with common methods like cumulative sum chart control (CUSUM) [14] and generalized likelihood ratio (GLR) [15], as well as a simplistic Haar filter [16,17] borrowed from the field of signal

processing and computer vision. We report the performance of these approaches both on our simulated data as well as on the light-rail dataset.

Overall, we find that monitoring rail infrastructure from the vibrations recorded in operational trains appears to be a promising technique. As a validation, we were able to detect relatively subtle changes, such as the changes due to tamping, using the simplistic yet robust proposed method. In this project, all of the light-rail data was obtained from a single instrumented vehicle, although a larger number of instrumented vehicles could lead to most accurate assessments of the infrastructure. Given the projected growth of connected vehicles, the proposed method could become an even more relevant method for infrastructure inspection.

2. Rail-Monitoring Method

Our proposed rail-monitoring method uses data collected from accelerometers inside a train to identify changes over time in rail infrastructure. In thinking about how to analyze the data, our first goal was to understand how track roughness, filtered by the train's suspension, produces vibrations in the train's cabin. To do this we used the simplest possible model, a travelling damped oscillator as shown in Fig. 1 and conducted a simulation study as will be described in the following section. We sought to determine how best to detect changes in a section of track as the oscillator traveled repeatedly over it. This same type of model has been used previously in the rail-monitoring literature [8,18], but we extended that work to include variable train speed and differing levels of position uncertainty [19].

2.1 Feature Selection Simulation

We conducted a simulation of an oscillator travelling over a rough track to understand which features, when extracted from its dynamic response, are sensitive to track changes, but robust to speed and position uncertainty. The framework of the simulation is shown in Fig. 2. First we simulated a track profile and a change in that profile, for various types of track repairs. We then simulated the oscillator passing over the roughness 100 times before and after the track-repair, where each pass over the track had a unique speed profile, and extracted relevant features from the dynamic response of the oscillator. Finally, we quantified how well we could differentiate between the data before and after the repair. We repeated the steps of the flow chart in Fig. 2 for oscillators of different natural frequencies and damping ratios. The goal was not only to understand the behavior of the oscillator for parameters most similar to those of the light-rail system, but also to gain insight into how the results would vary for different rail systems so as to make the results more general.

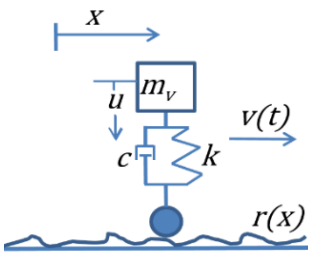

Fig. 1. Traveling oscillator moving over a rough track.

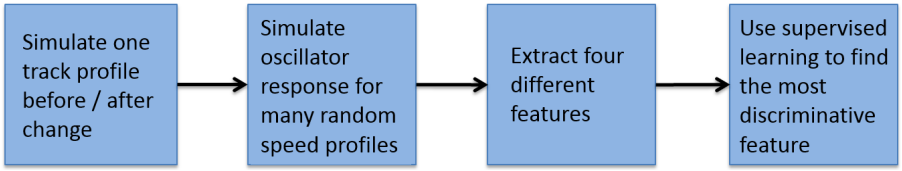

Fig. 2. Flow chart of the simulation. Note that this process was repeated for the three types of track changes, for a variety of damping ratios, natural frequencies, and position uncertainty levels.

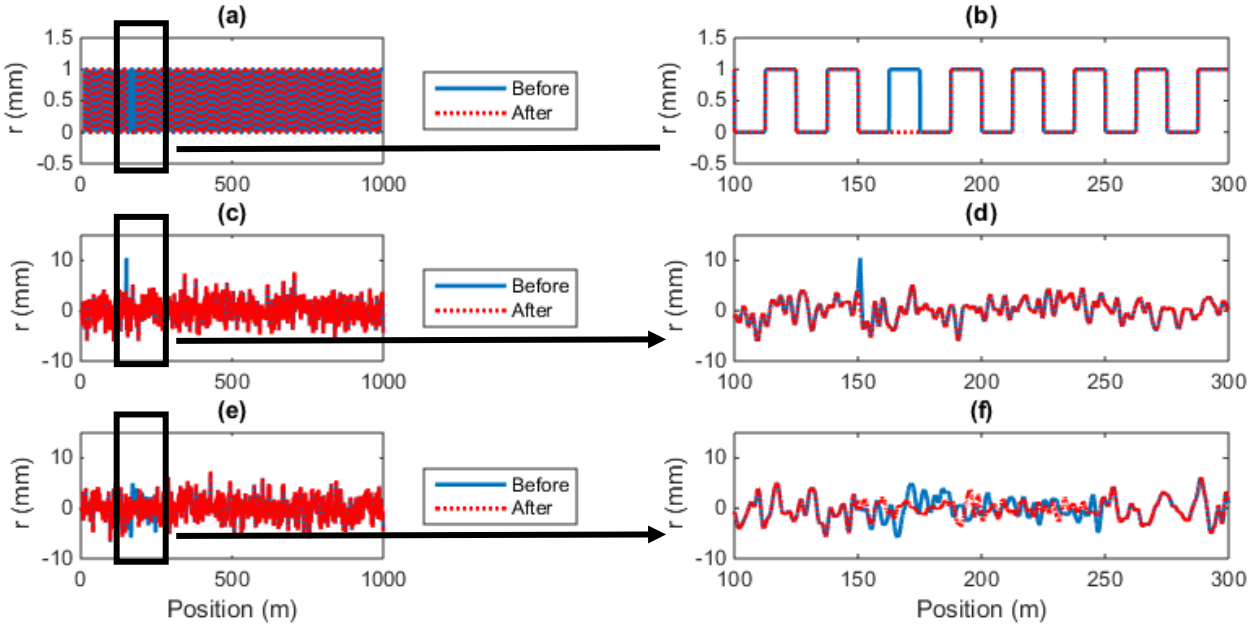

Fig. 3. Three types of roughness changes (left plots), with detail of each (right plots). (a) A toy-model change (used later to provide visual intuition about the simulation) and (b) the detail of the toy-model change. (c) spike change, characteristic of a broken track before and after replacement, with a realistic track roughness and (d) the detail of the spike change at 150m. (e) The tamping change, simulated using a filter with the same smoothing effect of the tamping machine and (f) the detail of tamping which occurs between 150m and 250m.

When simulating the roughness, we generated a 1km section of track for three change types, each of which is shown in Fig. 3. The first track change, in Fig. 3a, is a toy-model of roughness used to visualize the simulation later in the project (greater detail is shown in Fig. 3b). For the second and third track changes (Fig. 3c

and Fig. 3e), we simulated two types of track changes we had observed in practice, a large localized spike which is removed, simulating the replacement of a broken track (detail in Fig. 3d), and a smoothing of a track profile, simulating tamping (detail in Fig. 3f). For each of these realistic changes, we also simulated a realistic track roughness using the parameters found in the literature [4]. For the tamping change, we filtered this track roughness as pe the smoothing effect of the tamping machine in [4], reducing the standard deviation of the track profile from 2mm to 1.5mm over a 200m section of track.

Once we generated the track roughness, the next step was to generate the response of the oscillator; two realizations of this process over the same roughness with different speed profiles are shown in Fig. 4. Fig. 4a shows the speed profiles in the time domain, while Fig. 4b shows them in the spatial domain (i.e. plotted against position). Note when the speed approached zero, no distance is covered, producing the scalloping effect in Fig. 4b, a phenomenon which is common when the train stops at a station. When generating the speed profile, we limited it to be between 0 and 15m/s (35 MPH / 55 KPH) which is the same as that of the light-rail vehicle in our deployment. The toy-model roughness profile is shown in Fig. 4d as a function of position, but the train experiences this profile in the time domain Fig. 4c. We generated the response of the oscillator in Fig. 4e by solving the following differential equation [20]:

$$\ddot{u}(t) + 2\zeta\omega_n(\dot{u}(t) - \dot{r}(t)) + \omega_n^2(u(t) - r(t)) = 0,$$

where ζ is the damping ratio of the oscillator, ω_n is the natural frequency of the oscillator, u is the displacement of the oscillator, and r the track roughness as shown

in Fig. 1. Although the bumps in the track occur at the same location in space, due to the difference in the speed profiles, the oscillator is excited in the two passes at different points in time, as shown by the arrow in Fig. 4e. This posed a challenge when comparing multiple passes in the time domain, so we interpolated the signal into the spatial domain as seen in Fig. 4f. The bumps experienced by the oscillator line up more closely in the spatial domain, but the effects of varying speed are still visible. The wavelengths of the oscillators (the distance over which they occur) varies considerably in Fig. 4f due to the speed (as highlighted in the box), even though the periods of the oscillations (their duration in time) are invariant, as can be seen in Fig. 4e.

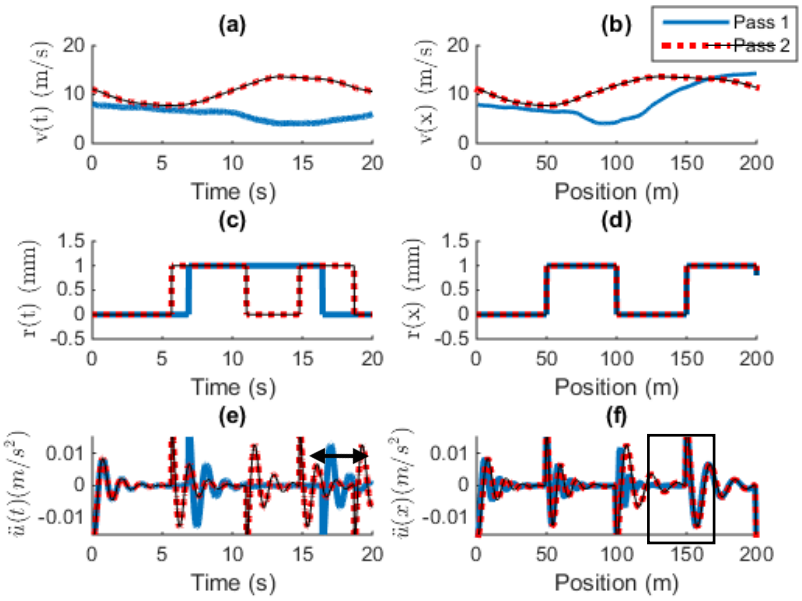

Fig. 4. Two passes over the toy-roughness shown for illustration: (a) the speed profiles in the time domain; (b) speed profiles in the spatial domain; (c) the roughness interpolated in time; (d) the roughness in space; (e) the acceleration of the oscillator in time, the arrow shows the period of misalignment; and (f) the acceleration of the oscillator in space, the square highlights a location where the wavelengths of the oscillations differ. The oscillator in both passes has a frequency of 1.5Hz and a

damping ratio of 0.2.

The variation in Fig. 4 highlights one of the challenges in dealing with variable speed; a second challenge lies in not knowing the exact position of the oscillator, a phenomenon that occurs in practice due to GPS error. In order to plot the signal in the spatial domain (Fig. 4f), we used the position of the oscillator, $x(t)$. In practice, we only know some approximation of the position, $x_\epsilon(t) = x(t) + \epsilon(\mathrm{t})$, where ϵ is the error. Let us assume that this error is normally distributed with zero-mean and standard deviation σ. We show the effect of this error with different standard deviations in Fig. 5. While in Fig. 4f we showed the vibration signal for two passes in the spatial domain, in Fig. 5a we show 200 passes, where each horizontal line is a vibration signal from a single pass, and the color is indicative of the vibration size. Note that we show 100 passes before the track change, and 100 passes after the track change from Fig. 3a. In Fig. 5b and Fig. 5c we show higher levels of position uncertainty by varying the standard deviation of the error, σ. As the position uncertainty grows, detecting changes in the tracks becomes harder, although detection is still possible (even visually) for this type of a trivial track change.

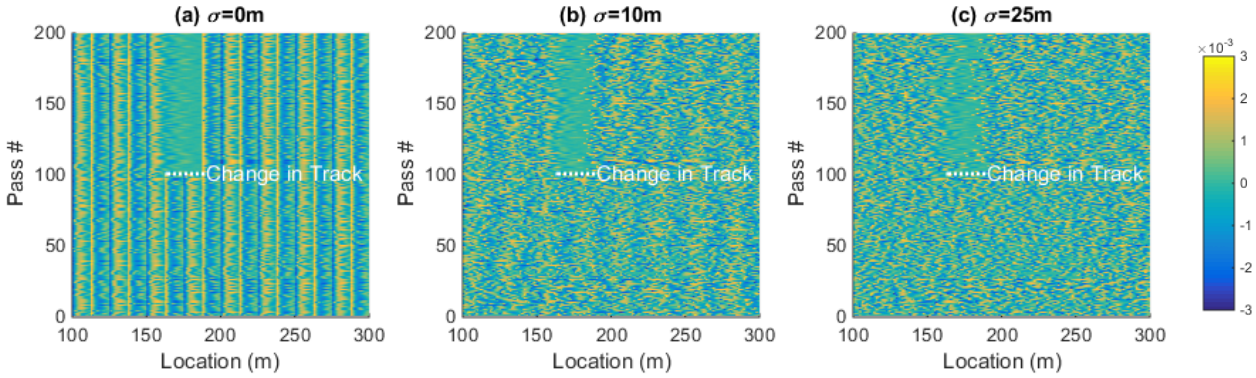

Fig. 5. Effect of position uncertainty. Each row of the above plots shows one pass of the oscillator over the toy-model roughness shown in Fig. 3. Passes 1-100 correspond to the "before" roughness, while passes 101-200 correspond to the "after" roughness. The color along each line represents the acceleration of the oscillator in m/s² but has been truncated at current bounds to show greater clarity. Note that each pass has a unique speed profile. (a) Shows the response of the oscillator in space with no position uncertainty, (b) shows the response with added Gaussian uncertainty, with zero mean and standard deviation of $\sigma = 10\text{m}$ (c) shows the oscillator with $\sigma = 25\text{m}$.

In Fig. 5 we show the vibration signal plotted spatially as a more intuitive signal representation given our interest in particular locations of track. To quantify how well this representation portrays track condition, we consider it as a feature, and compare it with three other features, time-frequency, spatial-frequency and signal-energy, to see which provides the greatest discrimination of track condition. Each is described in the equations below where $F[\cdot]$ denotes the Fourier Transform, $\ddot{u}_n$ is the vector of collected data from nth pass, f_n is the feature vector, and x_ϵ is the position vector with added noise ϵ.

Time-frequency $f_n = F[\ddot{u}_n]$

Spatial-frequency $f_n = F[\ddot{u}_n|_{x_\epsilon}]$

Spatial-amplitude $f_n = \ddot{u}_n|_{x_\epsilon}$

Signal-energy $f_n = \ddot{u}_n{}^2|_{x_\epsilon}$

where $x_\epsilon = x + \epsilon\,;\ \epsilon = N(0, \sigma)$

The motivation for having two types of frequency-based features is the prevalence of frequency-based features in the literature [21–24] or features related to frequency, such as wavelets [25]. "Time" here means the raw signal, whereas "spatial" denotes the signal has been linearly interpolated spatially given its speed

profile. The signal-energy feature is similar to the spatial-amplitude, but it is squared prior to interpolation. This squaring makes a difference because it makes the feature have a non-zero mean. Prior to classification, the spatial-amplitude and signal-energy features were averaged over a 25m section of track. This step increased robustness to position uncertainty for both features, but can filter out some zero-mean oscillations.

To evaluate the features, we used supervised classification; the goal was to see which features allow for the clearest indication of track change. Here we used a support vector machine classifier with a linear kernel (we choose a simple model to avoid overfitting given the relatively few passes we use in our experiments with operational data). We selected 150 passes out of the original 200 for training (the remaining 50 are for testing), and we repeated for 50 fold-cross validation. High classification accuracy means that the data is more easily separable, i.e. the feature is useful. The features were long (even in this simulation, 10^4 samples), so we select only the 50 most discriminative indices of the feature. To accomplish this, we use the technique described in [23], where we find the mean signal for both classes from the training data, then define the most discriminative indices as those with the greatest difference between the two mean signals. In total, to explore feature selection, we simulate 180,000 passes of the oscillator over 1km of simulated track (3 types of track changes × 6 damping ratios × 5 natural frequencies × 10 levels of position uncertainty × 200 passes each).

2.2 Feature Selection Simulation Results

The classification results from the simulation are shown in Fig. 6 and Fig. 7. This was a binary classification, so 50% accuracy would mean the labels are random, while higher accuracy means more consistency in determining the state of the track given the selected features. We report accuracy while varying the level of position uncertainty (Fig. 6a), the oscillator's damping ratio (Fig. 6b) and the oscillator's natural frequency (Fig. 6c). In Fig. 6a, we kept the damping and frequency ratio constant at values we observe in the light-rail system ($\zeta = 0.2$; $\omega_n = 3\pi/2$) [26]. We found that time-frequency features provide low accuracy at all levels of position uncertainty, spatial-amplitude and spatial-frequency fall in accuracy as the position uncertainty increases, and signal-energy offer relatively high accuracy at all levels of position uncertainty. It is not surprising that time-frequency performed badly because it is sensitive to speed changes, which we varied between each run. Spatial-amplitude and spatial-frequency require interpolation by x_ϵ, so they are sensitive to increases in the uncertainty ϵ. Signal-energy also relies on interpolation by x_ϵ, but is less sensitive to position error as it represents the level of roughness or excitation caused by the track, rather than the specific state of the oscillator (it is always positive independent of the oscillator).

We also investigated the effect of varying the damping ratio (Fig. 6b) and the natural frequency (Fig. 6c) if the location uncertainty is kept constant at $\sigma = 7$m. For this level of position uncertainty, signal-energy is the only feature with strong discriminative power. When the damping ratio was zero, the oscillations kept growing over the course of a pass, so the localized track change are undetectable

even from the signal-energy feature. As the damping ratio is increased, signal-energy peaks at $\zeta = 0.2$ then decreases as the damping ratio grows because the changes are too localized and difficult to detect given the position uncertainty. For variation of the natural frequency (Fig. 7c), the stiffer the oscillator, the more it transmits large track changes to the sprung mass, leading to easier detection with the signal-energy feature for high values of the natural frequency. In the case of tamping in Fig. 7c, the reverse trend occurs. Here, a stiffer oscillator leads to more localized changes and because tamping is effectively many small smoothing changes over a section of track, these localized changes are more difficult to detect.

Overall, the signal-energy feature outperforms the other features as it is robust to position uncertainty for both types of track changes. This is particularly convenient as the signal-energy feature could potentially be universally applied in this type of monitoring system. In industry, the most common parameter for measuring track geometry is the standard deviation of the track profile, and signal-energy is related to this parameter, as both involve squaring their respective data points. The vibration signal is effectively a measure of the track profile, albeit filtered by the train's suspension, so the strong performance of signal-energy shows that the traditional parameters for track monitoring can inspire new features that are useful for our statistical track monitoring system.

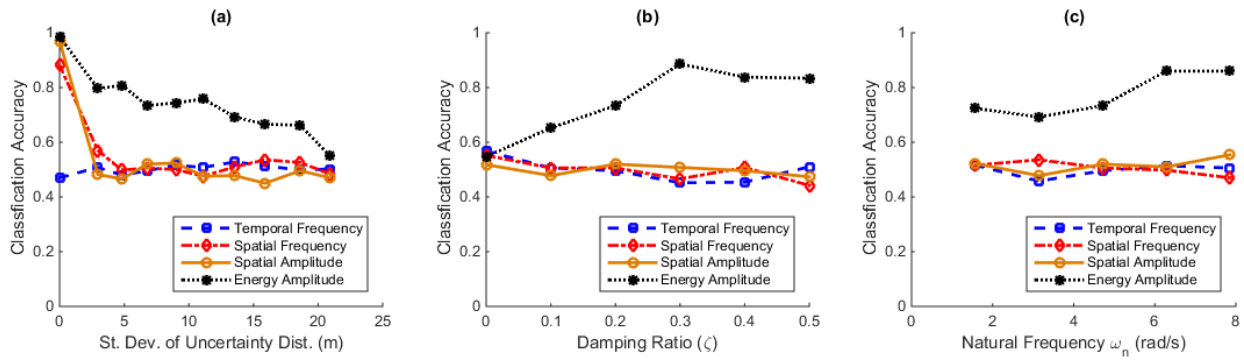

Fig. 6. Classification accuracy for spike change. (a) Effect of position uncertainty for oscillator with $\zeta = 0.2$ and $\omega_n = 1.5\pi$ (b) Effect of varying damping ratio while uncertainty $\sigma = 7m$ and $\omega_n = 1.5\pi$. (c) Effect of varying natural frequency while $\zeta = 0.2$ and uncertainty $\sigma = 7m$.

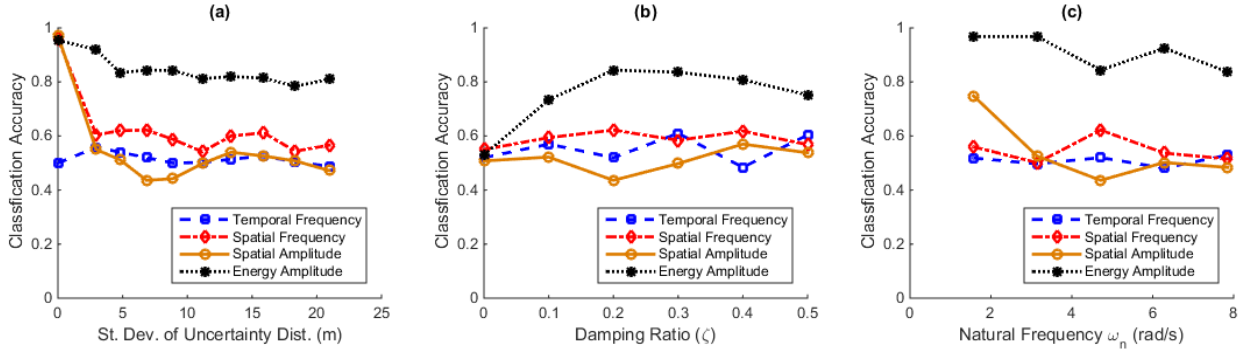

Fig. 7. Classification accuracy for tamping change. (a) Effect of position uncertainty for oscillator with $\zeta = 0.2$ and $\omega_n = 1.5\pi$ (b) Effect of varying damping ratio while uncertainty $\sigma = 7m$ and $\omega_n = 1.5\pi$. (c) Effect of varying natural frequency while $\zeta = 0.2$ and uncertainty $\sigma = 7m$.

2.3 Change Detection Simulation

Having established that among the different criteria examined, signal-energy represents the most robust feature, in this section w explore how we can achieve our second goal: to determine when a change occurs in the tracks.

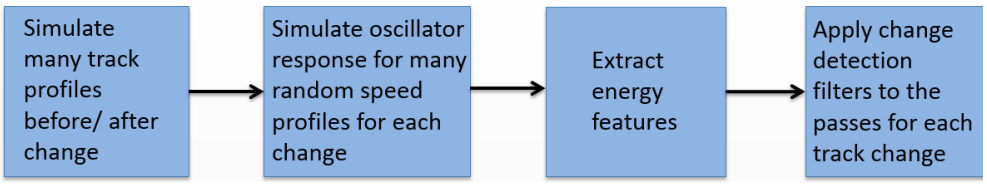

Fig. 8. Flow chat of change detection simulation. Note that we repeat this for different track changes and different levels of position uncertainty.

We conducted a simulation to study change detection approaches by following the flow chart shown in Fig. 8. First we generated many passes of the train before and after a track change (varying the speed profile as in the previous simulation study). We computed the signal-energy feature from the vibration signal, then extracted the data at a particular track location across all passes. Although a number of change-detection approaches exist, many of them are designed for finding statistical change in a scalar quantity. In this application, we compared data from different passes of a moving vehicle, so each pass was a vector. To format the data in a way that we could use these general change detection approaches, we considered the value of the signal-energy feature at a specific location on the track, i, over all passes, n, by building a matrix, $F(n,i)$.

$$F(n,i) = \begin{bmatrix} \cdots & f_1 & \cdots \\ \cdots & f_2 & \cdots \\ & \vdots & \\ \cdots & f_N & \cdots \end{bmatrix}$$

Using this formulation, we experimented with three different change detection filters: cumulative summation control chart (CUSUM), generalized likelihood ratio (GLR) and a Haar filter.

All three change-detection methods effectively detect when there is a change in the mean-value of the feature at a particular location of track compared to that value

for a window (set number) of previous passes. Details of each will be presented briefly to explain how they are applied within the two dimensional feature matrix.

1. Cumulative summation control chart (CUSUM) estimates the mean value over a window of previous passes (known as a sliding window), then compares the current (or most recent value) to the historic mean [15]. If there is a succession of passes all deviating from the mean in one direction, it will trigger a detection event because it is likely that the mean has changed. Mathematically, the mean, $\widehat{\Theta}$, is estimated from historical data over a window of $w + 1$ previous points.

$$\widehat{\Theta}(n, i) = \frac{1}{w + 1} \sum_{k=n-w}^{n} F(k, i)$$

We then find how the data from the current pass differs from this mean.

$$\delta(n, i) = F(k, i) - \widehat{\Theta}(n, i)$$

Finally we add this difference, δ, to a running sum of the differences.

$$g(n, i) = g(n - 1, i) + \delta(n, i)$$

If the data varies from the mean consistently in one direction (positive or negative) intuitively the data has changed because the mean no longer represents the data. We say a change has occurred when the magnitude of g exceeds a threshold, h, $|g(n, i)| > h$.

2. The generalized likelihood ratio (GLR) looks for a change in the underlying distribution, and quantifies the log-likelihood that the recent data is derived from the historical distribution [15]. If the recent data appears dissimilar from the historical data, then the approach will trigger a detection event.

Mathematically, let us call the first window of data y_0, and the second window of data y_1.

$$y_0(n,i) = F(n-w-1:n-1,i)$$

$$y_1(n,i) = F(n:n+w,i)$$

The distribution of y_0 is defined by H_0 and the distribution of y_1 by H_1. We want to see the likelihood ($L[\cdot]$) that y_1 comes from the same distribution as y_0, H_0, versus the likelihood it comes from H_1.

$$g(n,i) = 2\log\frac{L[y_1(n,i)|H_1(n,i)]}{L[y_1(n,i)|H_0(n,i)]}$$

We say a change has occurred when g exceeds a threshold, h, i.e $g(n,i) > h$. In other words, beyond some threshold, the data is so much more likely to have come from a new distribution than from the historical distribution that we say a change in the data has occurred.

3. The Haar filter finds the difference between the sum of recent data and the sum of historical data, triggering a detection event if the difference is too large. Mathematically, the difference between the two windows of data of width $w+1$ can be written as

$$g(n,i) = \left|\sum_{k=n}^{n+w} F(k,i) - \sum_{k=n-w-1}^{n-1} F(k,i)\right|$$

As in previous approaches, we say a change has occurred whe $g(n,i) > h$. This approach tends to find locations where a step change has occurred between the two windows of data.

Each of the approaches relies on a sliding window of data: shorter-sized windows have the potential for detecting changes more rapidly after they occur, while longer windows allow for greater statistical significance; here we chose a window size of 20 passes, which would allow our system to detect a change within a week (assuming a few passes over the track per day). The results of applying each of these filters to the simulated changes can be seen in Fig. 9 as will be discussed in more detail in the next section.

2.4 Change Detection Simulation Results

In this section, we discuss the performance of the three change-detection approaches described previously. Results of each approach applied to one example of each type of simulated track change are shown in Fig. 9. Successful change-detection filters should produce a high value after the change occurs at the same location in the track, and successful detection has been indicated with a red-box. Note that the change is not necessarily detected immediately, as it takes several passes after the change to build up statistical significance that a change has indeed occurred. All three approaches detected the track change on the left in Fig. 9b-d (with an ideal threshold), while only the Haar filter was able to detect the tamping change in Fig. 9h.

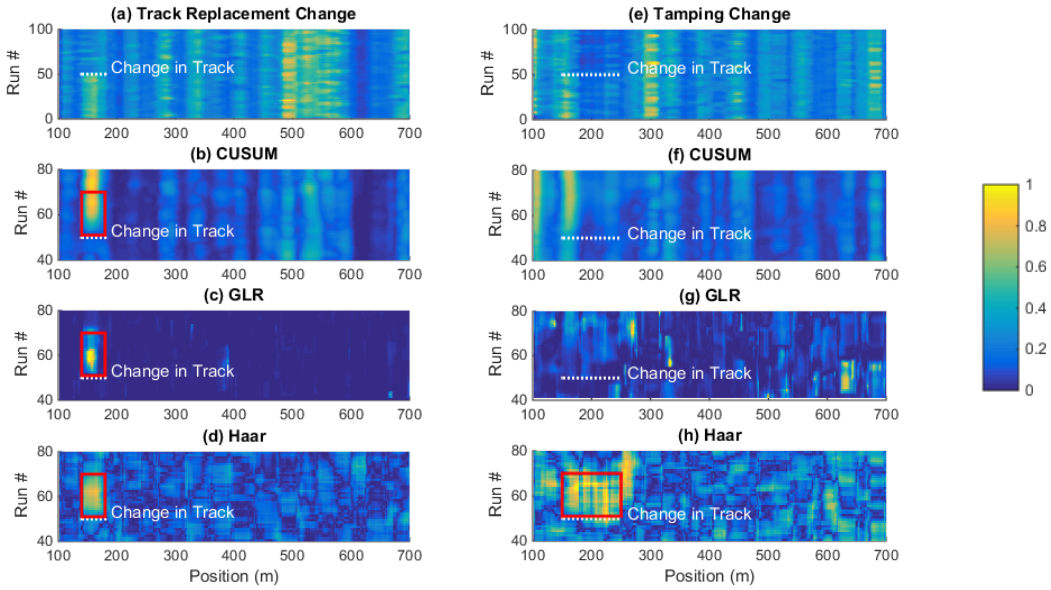

Fig. 9. The change detection filters applied to the data. (a-d) show one example of a track replacement change, where the spike (due to a broken track) is removed at 150m, while (e-h) show a tamping change between 150 and 250m. Values in each figure (shown in color) have been normalized on [0,1]. The red-boxes indicate the true-positive events that could be detected with an ideal threshold. There are no boxes in (f) and (g) as the methods failed to detect the change. The data shown in this figure has no position uncertainty.

CUSUM is perhaps the most versatile of the methods, as it could detect a gradual change in the mean over time, while GLR and Haar work best for abrupt changes in the data. In our case, we assume the track change occurs over a short period of time, so the change is complete before the train passes over that section of track again. Given this assumption, GLR and Haar might be expected to perform better than CUSUM. GLR is more complex than the Haar filter as it involves not only the mean, but also the standard deviation in calculating the distribution of the data before and after the change. A change in variance could trigger a detection event, but it is

unlikely in our application that high variance would be due to an infrastructure change, so GLR is less robust for this application. It is the simplicity of the Haar filter that enables it to consistently detect tamping along the section of track where it was simulated in Fig. 9h.

Fig. 9 shows just one example of each type of change, but we quantified the success of each approach over the 100 changes we simulated. We defined a true positive as a change detected within 20 passes and within 25m of a true change. A false positive was a change detected outside of this window. We report false positives and false negatives (true positives that were not detected) for a range of thresholds in Fig. 10a, because selecting the appropriate threshold itself can be difficult. Methods for selecting the best threshold are beyond the scope o this project, so we report the error level assuming the ideal threshold was selected in order to see which change detection technique has the most potential. Assuming false positives, *FP*, and false negatives, *FN*, are equally bad, we select the threshold level that will minimize the larger of the two errors, which occurs where the false negative and false negative curves intersect as show in Fig. 10a. In other words, we want to select a threshold, h, where

$$h = \underset{h}{\operatorname{argmin}}(\max[FN, FP]).$$

In total, for understanding the trade-off in the change detection approaches, we simulated 200,000 passes (2 types of track changes × 100 track changes × 100 passes × 10 levels of uncertainty).

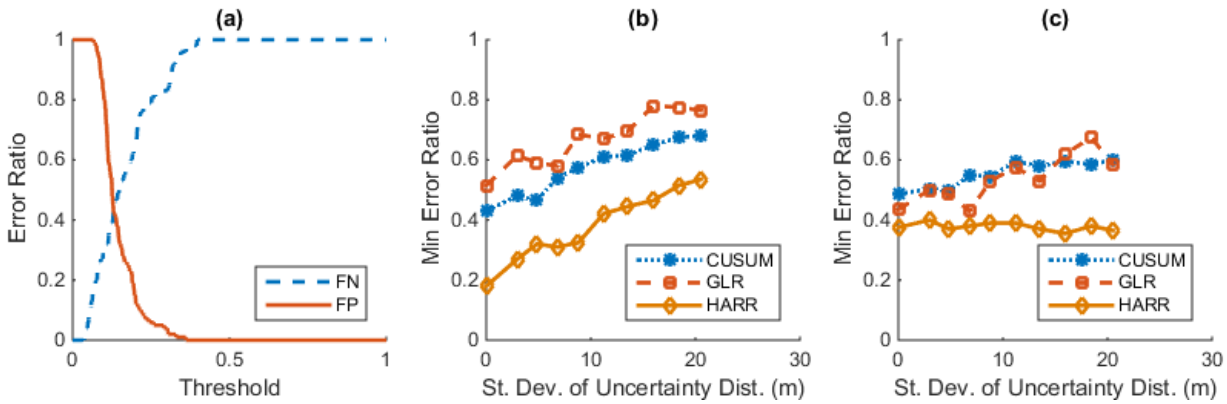

Fig. 10. Change detection results from the simulation. (a) A typical plot of false negatives (FN) and false positives (FP) as the threshold is varied, shown with data for the 100 examples of this type. In this case, the plot is shown for CUSUM with no position uncertainty, and where the two lines cross, there is 43% error of both types. The data has been normalized on [0,1] so the threshold spans the whole range. (b) Here we show this minimum error for all three approaches and all position uncertainty levels for the spike change. (c) The minimum error for the tamping change.

In Fig. 10b we report the lowest error (i.e. ideal threshold) for all change detection approaches on the track change simulation. While the error increased for all methods as position uncertainty increases, the Haar filter consistently performed the best. Fig. 10c shows the same error quantification but for the tamping change simulation, and again the Haar filters outperforms the other methods. Because tamping occurs over a larger section of track, position uncertainty matters less as we saw in the classification results shown in Fig. 7a, so it is consistent that the Haar filter remains relatively unaffected by position uncertainty. It is interesting that even though all three methods are applied to signal-energy data (which itself was robust to position uncertainty when classified) the GLR and CUSUM approaches are sensitive to position uncertainty. One reason for this may be that the classifier used

in generating Fig. 7a, an SVM, was more robust to position uncertainty than the CUSUM or GLR methods.

3. Validation on the Light-Rail Vehicle Dataset

The previous sections provide a general understanding of which features and detection filters work well to detect changes in a rough track from the dynamic response of a travelling oscillator. In this section, we investigate whether the same findings hold true for our test-system deployed on a light-rail vehicle.

3.1 The Light-Rail Vehicle Dataset

We have collected an extensive dataset over the last two years; this i the first study to our knowledge on vibration-based track monitoring in which data was collected over such a long period while the trains were in operational service. We provide details about our instrumentation, about two infrastructure changes of interest, and, finally, we assess the validity of the approaches on this dataset.

The sensors on the train were installed through our partnership with the Port Authority of Allegheny County; the location of the sensors can be seen in Fig. 11. We placed two uni-axial accelerometers inside the cabin of the train to measure vertical vibrations, one tri-axial accelerometer on the central wheel truck, and a BU-353 GPS antenna mounted just under the roof. In this project, we only use data from the accelerometer mounted inside the cabin above the front wheel-truck, a Vibra-

Metrics model 5102 sensor. The train car is a CAF light-rail vehicle that has an empty mass of 40 metric tons, and a length of 27m.

We have collected data from hundreds of complete passes through the 30km rail network, sampling continuously from the GPS at 1Hz and from the accelerometer at 1.6kHz. Details of this deployment can also be found in a previous conference proceeding [11]. One of the challenges in analyzing the data has been handling the varying speed and the position uncertainty, issues which inspired the above simulations.

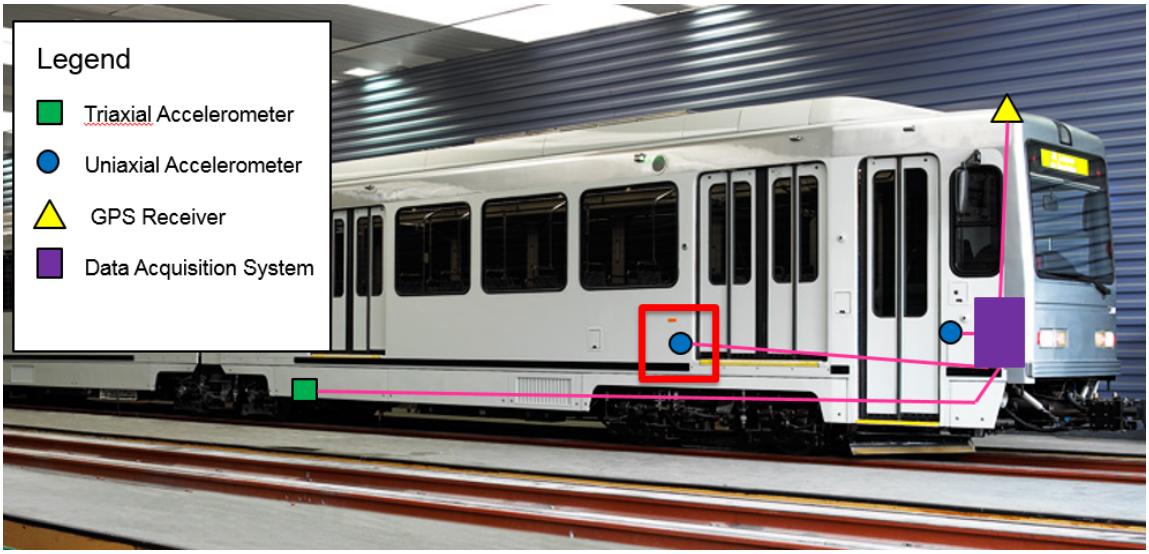

Fig. 11. Instrumentation of light rail vehicle. We placed a GPS antenna just below the roof the train, the data acquisition system in a cabinet near in the conductor, accelerometers inside the main cabin and on the central wheel truck. In this project we only used data from the accelerometer highlighted with the red box, which measures vertical vibrations inside the cabin of the train.

4.1 Trac Change in the Light-Rail Dataset

In September 2014, the owner of the light-rail system replaced the track in an old road-crossing. We use this known maintenance activity, where faulty track was replaced with good track, to test our signal analysis pipeline. Fig. 12a and b show two characteristic passes of the train over the 1km section of track. In the second pass the train stops at two stations, the first at 200m and the second at 520m, whereas in the first pass it only stops at on of the stations (at 200m). The train stops when there are passengers to pick up or let off, which poses a challenge when compared data between passes. One hundred passes over this section of track are shown in Fig. 12c, where the color indicates the value of accelerations recorded. Although difficult to see in Fig. 12, there is a high frequency vibration event when crossing a road at 220m, an event which no longer occurs after the 50th pass, when the repair is done. It is much easier to see the change in Fig. 13 because signal-energy is a better indicator of track state. As the train crosses the road at 220m, high signal-energy can be seen in Fig. 13a. This spike is absent in Fig. 13b or after pass #50 in Fig. 13c due to the repair.

Note as well that the first station stop in Fig. 12(a) and Fig. 13(a) occurs just after 200m, whereas in (b) it occurs just before 200m. This is likely due to the orientation of the train. The GPS is at one end of the train, and the train always stops at the same location within the station. If the orientation of the car changes, the position can differ by 27m, the length of the car. Due to the sensor's location on the train, it will interrogate the tracks near the stations at different speeds. A sensor at the front of the train will travel quickly over tracks at the beginning of the station,

and slowly over tracks towards the end of the station, and the reverse for a sensor at the back of the train, adding to the difficulty of comparing data directly between different runs.

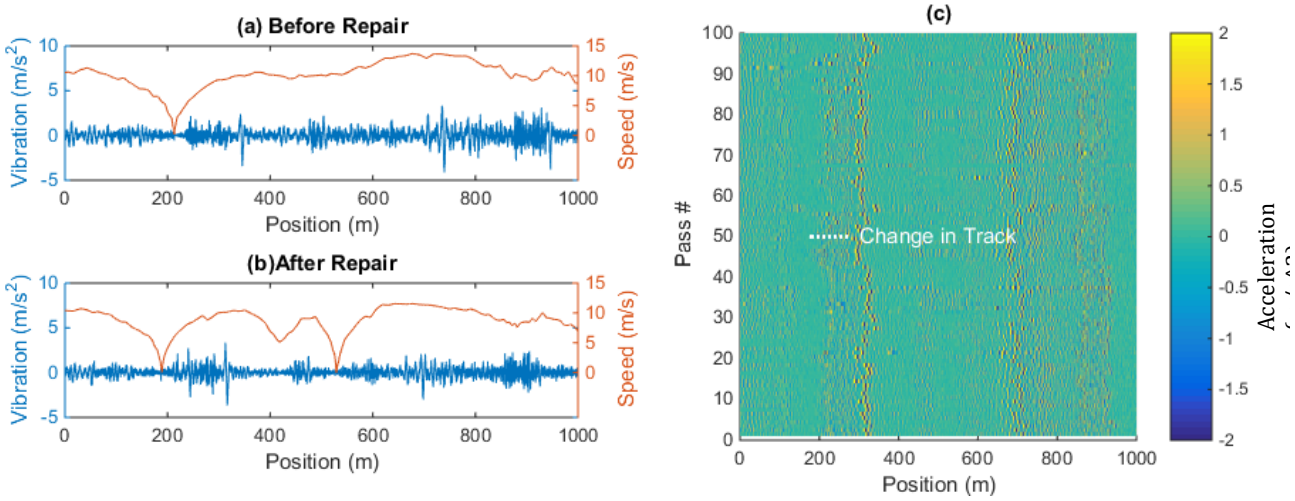

Fig. 12. Spatial signal. (a) A pass before repair showing both the train speed and vertical vibrations from sensor inside the train. (b) A pass after repair. (c) 50 passes before and 50 passes after the repair, where each pass is a horizontal line and the color indicates the value of the vibrations. With the spatial-amplitude feature, the track change is nearly impossible to see.

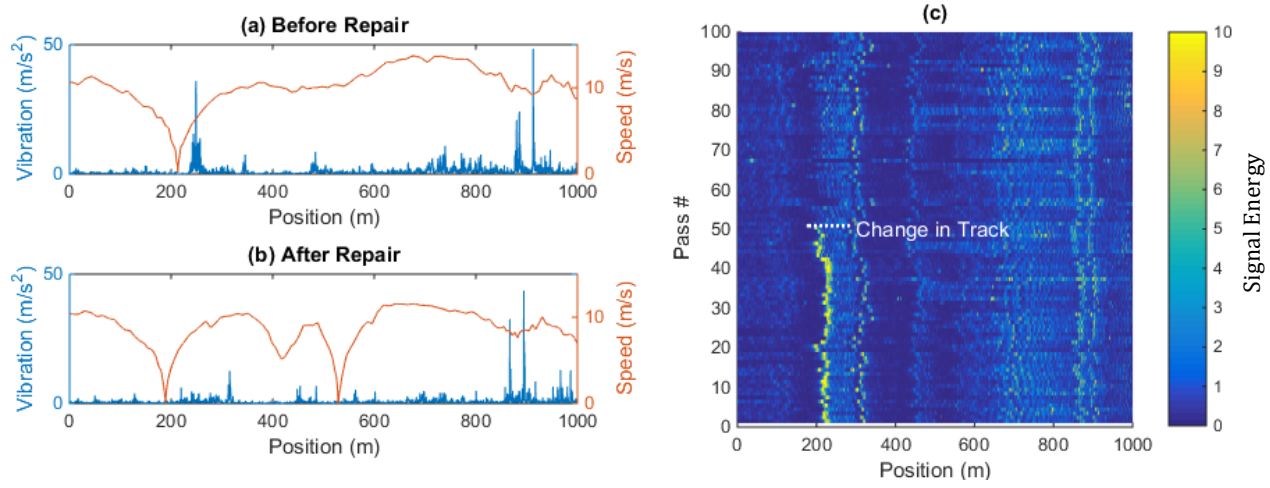

Fig. 13. Signal-energy. (a) pass before repair showing both the train speed and vertical vibrations from sensor inside the train. (b) A pass after repair. (c) 50 passes before and 5 passes after the repair,

where each pass is a horizontal line and the color indicates the magnitude of the signal-energy feature. With the signal-energy feature, the track change is clearly visible.

As in the simulation, we extracted different features from the data and used classification to determine which features allow the infrastructure change to be detected most readily. In a binary classification between the two states, we achieve 91% accuracy in Fig. 14a, drawing data from a 500m section of track shown in Fig. 14c. The second best feature appears to be the time-frequency feature. However, as we saw in the simulation, this feature would not be expected to detect track changes given the train's variable speed. It is far more likely that the feature is detecting temperature differences between the two classes as the data was collected over a year, and Pittsburgh is a temperate region with significant temperature variability. This hypothesis is confirmed by Fig. 14b where we show the adjacent 500m section of track where no work was done. Classification using the frequency-feature allows for 78% accuracy despite the fact that the infrastructure has not changed, whereas classification using the signal-energy feature allows for 67% accuracy (close to random) meaning it is more robust to environmental conditions. Again we use 75% of the data for training, 25% for testing and 50-fold cross validation as in the classification in the simulation study.

(c) Data used for classification

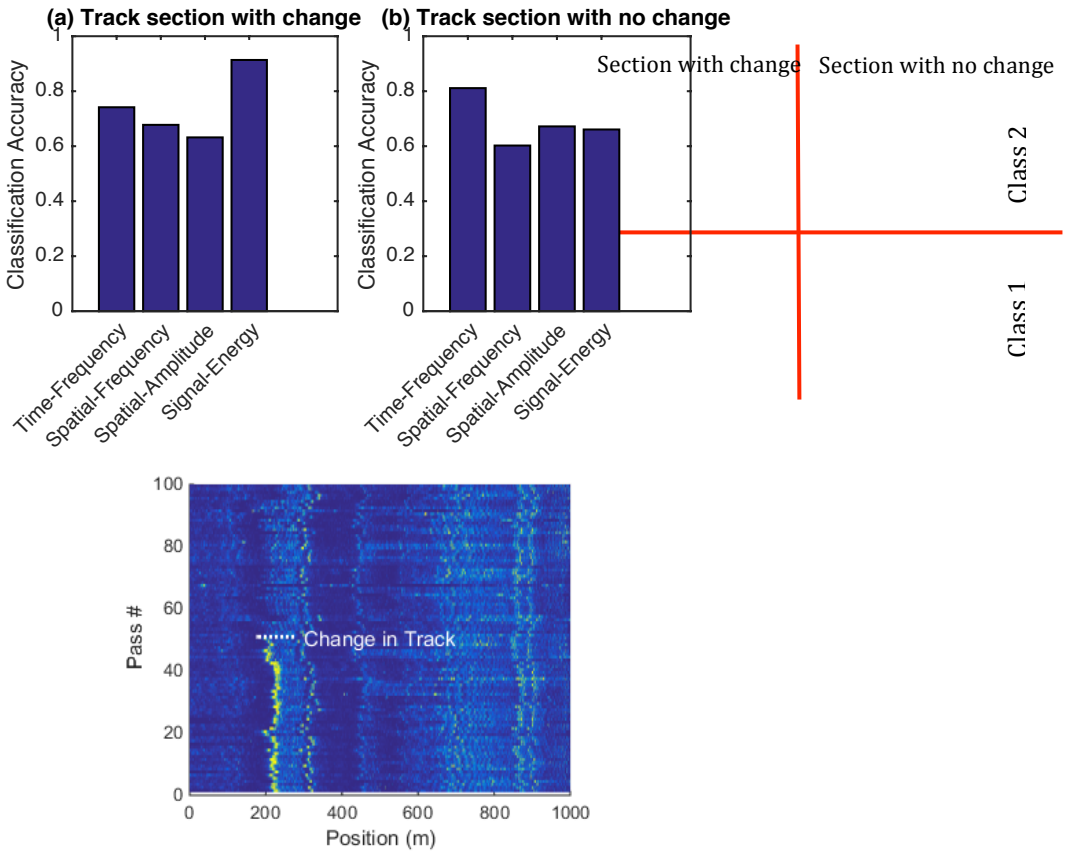

Fig. 14. Classification accuracy of a 500m section (a) of a track where rails were replaced, and (b) of a track nearby where no work was done. (c) Shows where the data for the classification was drawn from and the two classes used in the binary classification. High classification accuracy means the classes are separable, and 50% accuracy means the classes are not separable leading to random classification. Signal-energy is sensitive to track changes because it achieves 91% accuracy when there is a track change, and is close to random when there are no track changes, meaning it is not classifying based on environmental factors.

4.2 Tamping Change in the Light-Rail Dataset

Tamping is an important maintenance procedure used to improve track geometry. The tamping machine (shown in Fig. 15) measures the profile of the track, then adjusts the ballast below the track to produce a smoother, safer ride. One

future goal could be to use our data to optimize tamping schedules. At this early stage though, we are most interested in identifying which features are sensitive to tamping, and which detection approach works best for this feature. In the summer of 2014, three 500m sections of track were tamped as shown in Fig. 16, starting at 1500m, 2000m and 2500m. Unlike the track change in the previous section, the change due to tamping is subtle, but occurs over a much larger section of track. Notice how in the regions with no tamping, the signal is relatively consistent over all 100 passes, while in the areas with tamping, there is more signal-energy (i.e. more bumpy ride) before tamping, and less signal-energy (i.e. smoother ride) after tamping.

We tested all four features in a binary classification to discriminate between before and after a tamping event; we found that the signal-energy feature is the best for detecting this change. The classes in the classification and the region from which the data was drawn are shown in Fig. 17c. Signal-energy achieves 90% classification accuracy (shown in Fig. 17a) with a change, and <60% accuracy (effectively random) in Fig. 17b when there is no change. This means that the signal-energy feature is classifying due to the change in the track and not an environmental variable. Frequency and spatial-frequency features on the other hand report 70% accuracy if there is a change (as in Fig. 17a), and 70% with no change (as in Fig. 17b). We can assume the discriminative power stems from temperature changes and not from infrastructure changes. Note that for consistency, we only classified data from a 500m section of track as in the previous section, but the tamping occurs

over a larger region. Classification accuracy can be slightly increased by considering a larger section of track.

Fig. 15. The tamping machine used on the light-rail system.

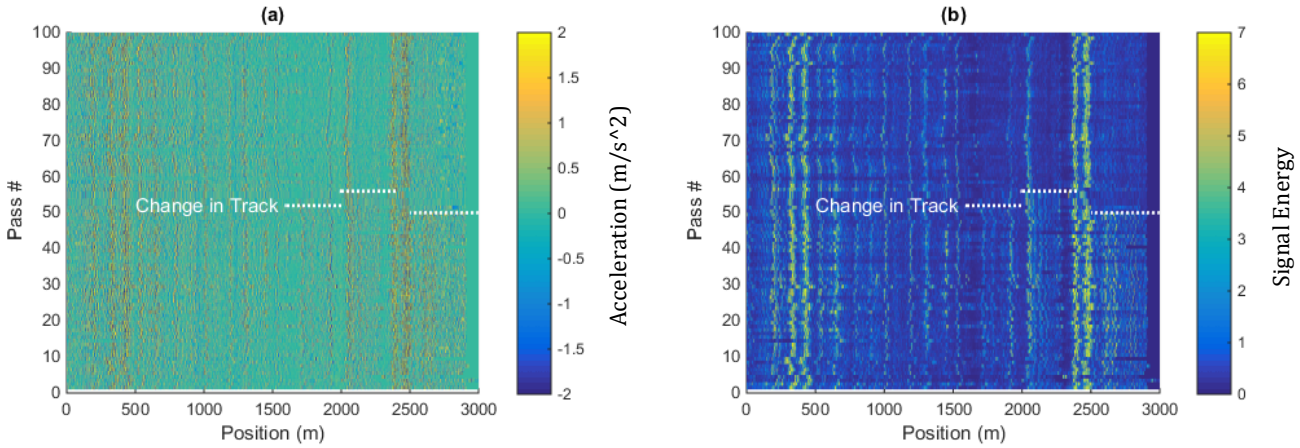

Fig. 16. Region of track where tamping occurs with (a) showing the spatial-amplitude feature and (b) showing the signal-energy feature. The tamping maintenance was done three times on three separate days due to the limitations on how much work the tamping machine can do per day. Note that the peaks between 2400 and 2500m (and between 400 and 500m) are due to switchgear in the tracks.

(c) Data used for classification

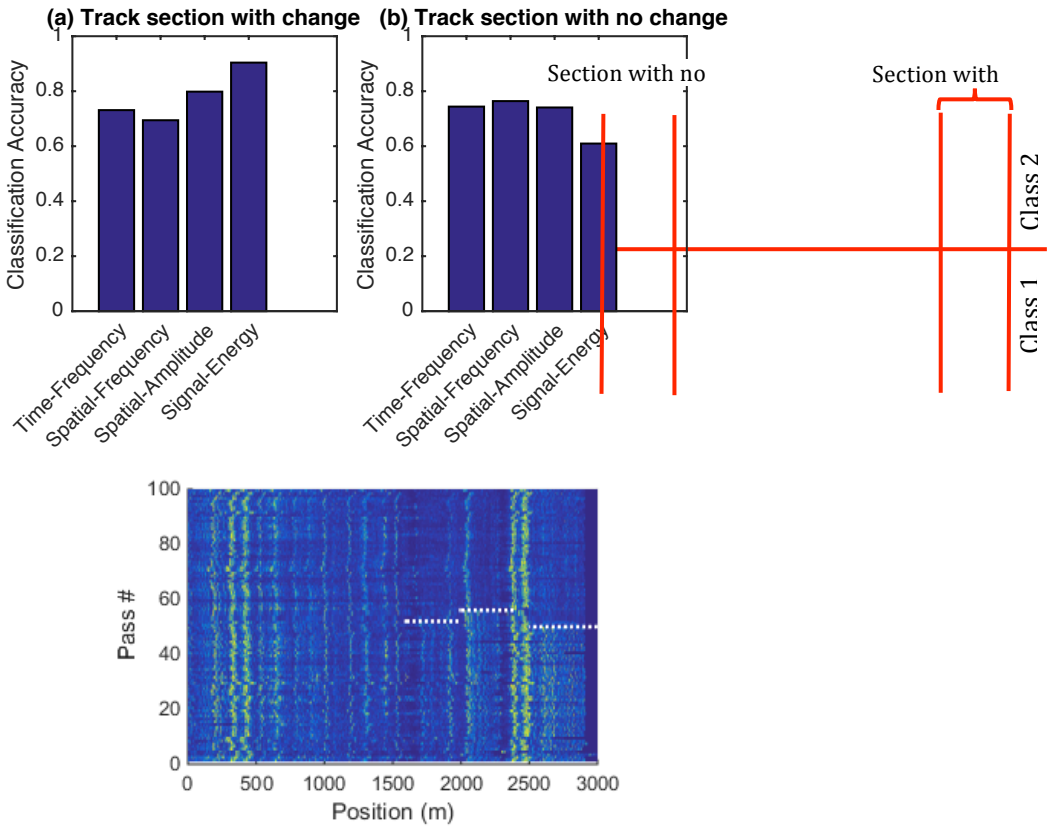

Fig. 17. Classification accuracy of a 500m section (a) of a track where tamping work was done, and (b) of track nearby where no work was done. As in Fig. 14, the high accuracy of the signal-energy feature where there is track work shows it is sensitive to infrastructure changes, and the low accuracy (almost random) where no work has been done shows it is robust to environmental variables. (c) We show the data used for the classification in relation to Fig. 16b, both in terms of which 500m sections of track were used, and how the two classes in the binary classification were defined.

4.3 Detecting Chang in the Light-Rail Dataset

The ultimate goal of the project is to automatically detect changes in the infrastructure. As such, change detection is a vital step. Although we have collected hundreds of passes through the rail-network, we still have relatively few known infrastructure changes. Thus, we do not have sufficient data to rigorously test

different threshold levels as we did i the simulation. Instead, we applied the change detection methods on the operational data and present the results with a basic threshold applied in Fig. 18. This figure mirrors the results in Fig. 9, where the raw data and all three approaches are shown. Figure 19a shows the signal-energy feature for the track replacement change, and Figure 19b for the tamping changes. Fig. 18 c-h show the results of applying the three change detection approaches to each of these two changes, in which the red-boxes indicate the change-detection approach was successful.

For both types of track changes on the light-rail dataset, the Haar filter performed the best, reliably detecting both track and tamping changes (Fig. 18d and h) as predicted by our simulation results. For the track-type change, CUSUM and Haar filter (Fig. 18b and Fig. 18d respectively) could detect a change with zero error given the correct threshold. For the tamping change, CUSUM and GLR fail to detect the change at all. CUSUM fails because the variability in the sections with high energy (like the switchgear at 2500m) are larger than the change from tamping itself, which occurs on low-energy sections of track. It is unclear why the GLR method has false-positives, although it could be due to the GLR's sensitivity to changes in variance as discussed earlier. It is important to note that the Haar filter succeeds in our original goal: detecting changes relative to historical behavior, rather than simply detecting areas of the track that produce large vibrations. The high values around 2500m in Fig. 18a due to switchgear do not affect the detection in Fig. 18h.

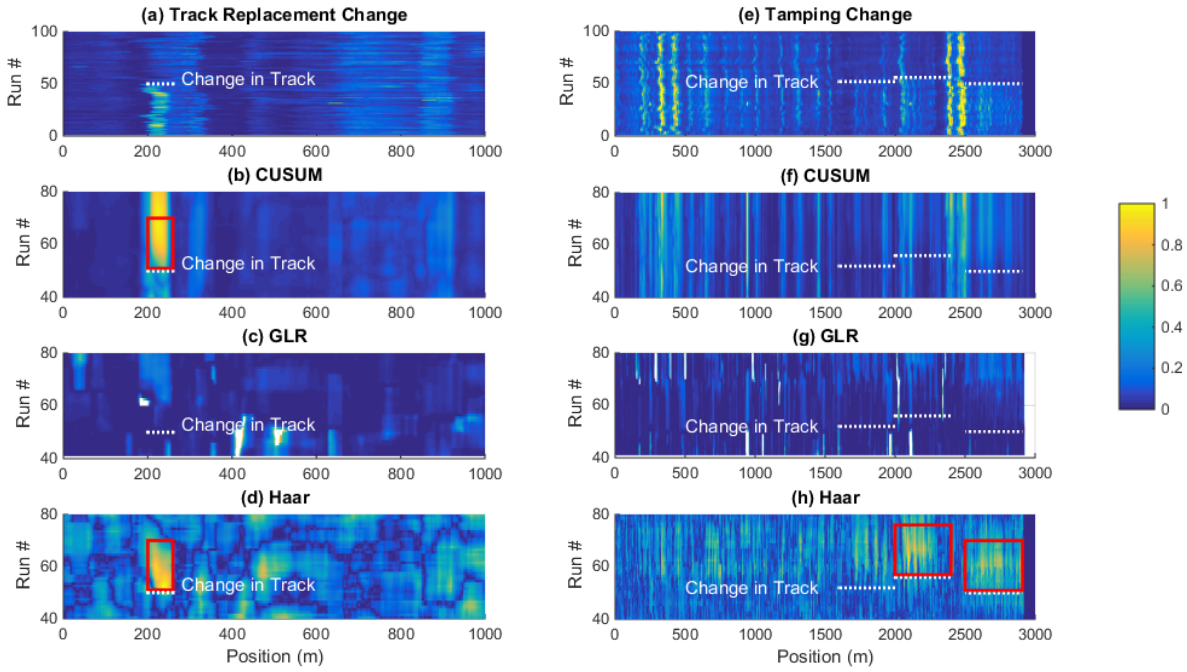

Fig. 18. Change detection on the light-rail dataset for the track replacement change at Potomac Crossing (a)-(d) and tamping change (e)-(h). Panels (a) and (e) show the signal-energy feature while (b)-(d) and (f)-(h) show the respective change detection techniques. The red-boxes indicate true-positive changes that could be detected with an appropriate threshold.

4. Concluding Remarks and Recommendations

In this project, we have proposed a new way to monitor rail infrastructure by recording data from accelerometers in the cabin of a train while it operates in revenue service. To realize such a system, new features must be extracted from the vibration signal that are able to characterize the rails despite the filtering from the train's suspension, the train's changing speed, and the noise in the GPS location. We tested four features, time-frequency, spatial-frequency, spatial-domain and signal-energy on simulated data and found signal-energy to be the best feature for detecting both track changes and tamping changes. We then tested the

same features on the operational data from the light-rail system, and showed that signal-energy was the most sensitive to infrastructure changes and the least sensitive to environmental variability. Both in our simulations and in the operational data, we found that frequency-based features do not work well, despite their widespread use in many structural health-monitoring studies.

In addition to feature extraction, we studied a variety of unsupervised change detection approaches, including CUSUM, GLR, and the Haar filter. We found that the Haar filter outperformed the other approaches on both the simulated data and operational data, as it was particularly robust to position uncertainty.

From the results of this project, it appears that vehicle-based monitoring could be a low-cost approach to monitor our infrastructure; we were able to detect changes on an operational system using just a single sensor on a train in revenue service.

Future work on train-based rail-infrastructure monitoring should look at additional types of changes in rail infrastructure, and more instances of each type. In this project, we considered statistical approaches that largely ignore the underlying physics of the problem; hybrid approaches, which combine statistics with a physics-based approach could provide additional insight.

Because connected vehicles are becoming more and more prevalent, we believe that collecting the requisite data for the proposed type of monitoring will become easier. The more data available, the more accurate the monitoring could be. However approaches capable of fusing data from heterogeneous vehicles and sensors must be developed.

This type of low-cost monitoring system could compliment traditional inspection techniques to provide continuous and objective information to infrastructure owners. This

information would allow for more targeted maintenance interventions, improving the reliability, safety and efficiency of rail transit.

References

[1] C.R. Farrar, K. Worden, Structural Health Monitoring: A Machine Learning Perspective, John Wiley & Sons, 2012.

[2] J.M.W. Brownjohn, Structural health monitoring of civil infrastructure, Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 365 (2007) 589–622. doi:10.1098/rsta.2006.1925.

[3] D. Barke, W.K. Chiu, Structural Health Monitoring in the Railway Industry: A Review, Struct. Health Monit. 4 (2005) 81–93. doi:10.1177/1475921705049764.

[4] C. Esveld, Modern Railway Track: Digital Edition, 3.4 edition, MRT-Productions, 2015.

[5] M. Molodova, Z. Li, R. Dollevoet, Axle box acceleration: Measurement and simulation for detection of short track defects, Wear. 271 (2011) 349–356. doi:10.1016/j.wear.2010.10.003.

[6] P.F. Westeon, C.S. Ling, C. Roberts, C.J. Goodman, P. Li, R.M. Goodall, Monitoring vertical track irregularity from in-service railway vehicles, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 221 (2007) 75–88. doi:10.1243/0954409JRRT65.

[7] Y. Oshima, K. Yamamoto, K. Sugiura, A. Tanaka, M. Hori, Simultaneous monitoring of the coupled vibration between a bridge and moving trains, in: Proc. Fifth Int. IABMAS Conf. Phila. USA 11-15 July 2010, n.d. http://www.crcnetbase.com/doi/abs/10.1201/b10430-114 (accessed November 5, 2012).

[8] J. Real, P. Salvador, L. Montalbán, M. Bueno, Determination of Rail Vertical Profile through Inertial Methods, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 225 (2011) 14–23. doi:10.1243/09544097JRRT353.

[9] M. Bocciolone, A. Caprioli, A. Cigada, A. Collina, A measurement system for quick rail inspection and effective track maintenance strategy, Mech. Syst. Signal Process. 21 (2007) 1242–1254. doi:10.1016/j.ymssp.2006.02.007.

[10] Y. Naganuma, M. Kobayashi, M. Nakagawa, T. Okumura, Condition monitoring of shinkansen tracks using commercial trains, in: 2008 4th IET Int. Conf. Railw. Cond. Monit., 2008: pp. 1–6.

[11] G. Lederman, H. Young Noh, J. Bielak, Rail-infrastructure Monitoring through the Dynamic Response of a Passing Train, in: IWSHM, Stanford, CA, 2015. http://dpi-proceedings.com/index.php/SHM2015/article/view/844/844 (accessed November 9, 2015).

[12] H. Noh, R. Rajagopal, A.S. Kiremidjian, Sequential structural damage diagnosis algorithm using a change point detection method, J. Sound Vib. 332 (2013) 6419–6433. doi:10.1016/j.jsv.2013.07.005.

[13] M. Molodova, Z. Li, A. Nunez, R. Dollevoet, Automatic Detection of Squats in Railway Infrastructure, IEEE Trans. Intell. Transp. Syst. 15 (2014) 1980–1990. doi:10.1109/TITS.2014.2307955.

[14] E.S. Page, Continuous Inspection Schemes, Biometrika. 41 (1954) 100–115. doi:10.2307/2333009.

[15] F. Gustafsson, Adaptive Filtering and Change Detection, 1 edition, Wiley, Chichester New York, 2000.

[16] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2001 CVPR 2001, 2001: pp. I–511–I–518 vol.1. doi:10.1109/CVPR.2001.990517.

[17] M. Vetterli, J. Kovacevic, Wavelets and Subband Coding, CreateSpace Independent Publishing Platform, 2013.

[18] E.G. Berggren, M.X.D. Li, J. Spännar, A new approach to the analysis and presentation of vertical track geometry quality and rail roughness, Wear. 265 (2008) 1488–1496. doi:10.1016/j.wear.2008.01.029.

[19] T. Matarazzo, S. Pakzad, STRIDE for Structural Identification Using Expectation Maximization: Iterative Output-Only Method for Modal Identification, J. Eng. Mech. (2016) 04015109. doi:10.1061/(ASCE)EM.1943-7889.0000951.

[20] R.R. Craig, A.J. Kurdila, Fundamentals of Structural Dynamics, 2nd ed., Wiley, 2006.

[21] G. Lederman, Z. Wang, J. Bielak, H.Y. Noh, J.H. Garrett, S. Chen, et al., Damage quantification and localization algorithms for indirect SHM of bridges, in: IABMAS, Shanghai, CN, 2014.

[22] Y. Oshima, K. Yamamoto, K. Sugiura, Damage assessment of a bridge based on mode shapes estimated by responses of passing vehicles, Smart Struct. Syst. Volume 13, Number 5 (2014).

[23] F. Cerda, S. Chen, J. Bielak, J.H. Garrett, P. Rizzo, J. Kovacevic, Indirect structural health monitoring of a simplified laboratory-scale bridge model, Smart Struct. Syst. 13 (2014) 859–868.

[24] S. Chen, F. Cerda, P. Rizzo, J. Bielak, J.H. Garrett, J. Kovacevic, Semi-Supervised Multiresolution Classification Using Adaptive Graph Filtering With Application to Indirect Bridge Structural Health Monitoring, IEEE Trans. Signal Process. 62 (2014) 2879–2893. doi:10.1109/TSP.2014.2313528.
[25] A. Caprioli, A. Cigada, D. Raveglia, Rail inspection in track maintenance: A benchmark between the wavelet approach and the more conventional Fourier analysis, Mech. Syst. Signal Process. 21 (2007) 631–652. doi:10.1016/j.ymssp.2005.12.001.
[26] Dumitriu, Madalina, Influence of the Suspension Damping on the Ride Comfort of Passenger Railway Vehicles, Sci. Bull. Univ. Plitechnica Buchar. Ser. D. 74 (2012).