DETECTING AND CLASSIFYING RAIL CORRUGATION BASED ON AXLE BEARING VIBRATION

K.Alten, A.Fuchs

Center for Mobility Systems
AIT Austrian Institute of Technology GmbH
Giefinggasse 2, 1210 Vienna, Austria

ABSTRACT

Vienna's tram network is regularly surveyed by an inspection tram that uses a laser light section method to measures wear of the rail head and a microphone to detect curve squeal. In order to expand the vehicle's inspection capabilities and encompass more condition indicators for maintenance purposes, the vehicle's accelerometers which were hitherto only employed as an inertial measurement unit are now used to monitor the vibration of the axle bearings as the tram travels across the network. The inspection tram is equipped with four triaxial accelerometers on the axle bearings of the middle bogie and another one on the bogie itself. In the first stage of the current research project, the aim is to automatically detect and classify periodic unevenness of the rail head, known as corrugation, by training a classifier on a range of vibration features. Applying this machine learning algorithm during post-processing allows the condition state of the rail head to be compared across the network.

***Index Terms*—** vibration features, corrugation, machine learning, condition assessment

1. INTRODUCTION

Vienna's public transport operator (Wiener Linien) is responsible for inspecting the city's tramway rails for wear and other defects at regular intervals in order to ensure the limits for track gauge are maintained, which is central to avoiding derailment. In additional to fundamental safety aspects, the regular inspections increasingly focus on serviceability limit states such as comfort criteria in terms of vibro-acoustic emissions for both, passengers and residents living in the vicinity of the tracks. One type of rail defect considered during routine inspection is corrugation, which is a periodic defect on the rail head that can have several causes. It is primarily a result of wheel-rail interaction under the given operating conditions and typically shows a wavelength of 5-15 cm on light railways. Regarding its exact classification and causes, which are not subject of the presented study, interested readers are kindly referred to the available literature [1, 2, 3, 4].

This project was funded by the Austrian Research Promotion Agency

The aim of the presented study is to use MEMS accelerometers, which are currently employed as the inspection vehicle's inertial measurement unit (IMU), for tracking the axle bearings' vibration as the vehicle travels across the network and to use this data to gain information about the rail head, where the dynamic interaction between the wheels and the rails takes place.

In [5], the use of inspection vehicles on a tram track in Budapest was investigated and acceleration spectrograms were employed to distinguish different track defects. Automatic detection was not carried out, as the focus was placed on comparing the results of subjective track rating with static in-situ measurements and rolling stock measurements.

In [6], the authors also investigated the correlation of onboard accelerations with rail surface roughness on a regional trainline in Japan. They found the trailing axle of the instrumented bogie to have a lower detection accuracy than the leading axle. Their results show the feasibility of estimating corrugation from axle box acceleration, but do not give any threshold values or suggest means of automatic detection.

Likewise, the authors of [7] present a feasibility study for a vibration diagnostic tool on subway trains in Milan. They address the sensor and measurement problems to detect short-pitch corrugation, with lesser focus on the signal processing or algorithms for automatic detection.

The application of machine learning algorithms to detect surface conditions using onboard data is investigated in [8]. The authors used velocity and suspension deflection of a motorcycle to classify the road surface in sections of variable length (always using 1 s of data) in real time.

The presented study thus aims to expand upon existing literature by applying machine learning algorithms to onboard data in the urban rail sector to evaluate the rail head condition using vibration data.

2. SENSORS AND DATA PROCESSING

2.1. Hardware

The inspection vehicle running along Vienna's tramway tracks is a former tram car which was modified to contain a server and a data acquisition system triggered by an odometer [9]. Recording starts as soon as the vehicle is set in motion and encompasses a laser light section, one microphone, six triaxial accelerometers and two triaxial gyroscopes. The microphone is sampled at 48 kHz and all accelerometers and gyroscopes at 8192 Hz. Mapping the inspection tram's recordings onto Gauß-Krüger coordinates when locating rail defects or irregularities is based on finding the best match between the tram's path of motion (odometer, gyroscopes) and a network graph of the city's tram lines, containing absolute coordinates at known reference nodes.

2.2. Data analysis

For testing purposes, the features used for automatic detection of rail corrugation are defined using the axle bearings' vertical vibration only. This restriction is acceptable given the nature of the irregularity, which only appears on the running surface of the rails and is thus assumed to have a minimal or even negligible effect on the axles' longitudinal or sideways vibration to the track direction. Furthermore, corrugation typically appears over a considerable length of rail, particularly in bends. Unlike local defects, such as rail breaks, spalling, web defects or squats, corrugation typically extends across tens of meters or more, allowing the rails to be analysed in predefined segments with averaged parameters. These segments (bins) are chosen to be 5 m long, with start and end points being defined according to the network graph.

Using the four axles on the tram's middle bogie, the time domain data is extracted from the onboard binary records and downsampled to 2 kHz. This choice was based on the sensors' frequency range and the need for efficiency in post-processing, as a complete survey record of over 30 km length at a full 8192 Hz sampling rate would amount to several gigabyte of data when converted to text format.

The following list of features is computed for each acceleration channel and split into 5 m bins. Within each bin, the mean value is computed for further processing.

1. squared amplitude
2. fast-weighted level ($L_{acc,F}$ in dB re. $1\text{x}10^{-6} m/s^2$)
3. fast-weighted variable bandpass level ($L_{acc,F,vBP}$)
4. intensity ratio(IR)
5. fast-weighted one-third octave levels $L_{acc,F,i}$

The one-third octave bands range from 1.2 Hz to 500 Hz and IR is computed as the ratio between $L_{acc,F,vBP}$ and $L_{acc,F}$. The variable bandpass is a velocity dependent filter, which extracts the signal that is produced by the typical corrugation wavelengths (5-15 cm). Rather than converting the signal from time to space domain (using the tram's velocity) and performing an interpolation to achieve regular sampling intervals, the bandpass filter is applied on the short-time Fourier transform (STFT, window length 512 samples and overlap 480). The bandpass filter is a two-dimensional mask, where each frame's average velocity is used to compute the bandpass' frequency limits. After applying the mask, the spectrum undergoes an inverse STFT to yield the bandpass filtered time domain signal for computing the $Fast$ weighted level.

2.3. Regression model

Unlike the optical sensor system, which is triggered at constant space intervals of 1 cm, the acceleration sensors are triggered at constant time intervals and the recorded vibrations are a function of the vehicle speed and track construction. Gaining information about the running surface and its irregularities from axle bearing acceleration thus requires the compensation of these dependencies prior to extracting features that are a potential condition indicator.

The binned, averaged features listed above are thus checked for their dependency on vehicle speed, curvature and vehicle acceleration. For each survey record, a linear regression model was defined for each dependent feature versus these independent operating conditions. Nearly all features were found to have a logarithmic dependency on the tram's velocity as shown in Fig.1, whereas curvature or vehicle acceleration yielded no apparent trend.

Using these regression models, the expected value of each feature in each bin was computed and subtracted from the measured one. These relative deviations are included in the overall feature catalogue, providing 62 features per channel.

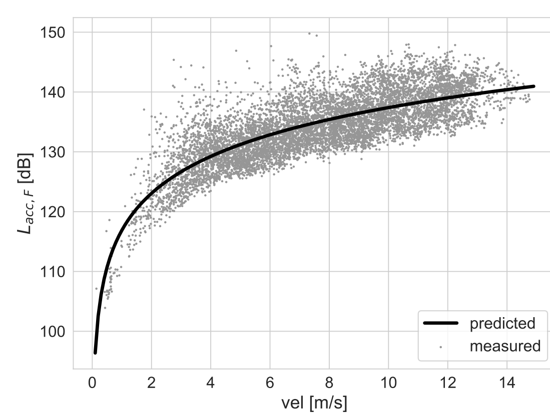

Fig. 1. Measured and predicted acceleration levels in each bin.

3. SUPERVISED LEARNING

3.1. Signal preconditioning

The challenge at hand is a classification task, which aims to allocate a corrugation category to each bin in the network graph. These condition categories are represented by a value ranging from 1 to 5, where 1 indicates no corrugation and 5 very strong corrugation. The corrugation extent of each bin was determined by maintenance personnel during in-situ inspections of around 9 km of track and the class distribution of the available data is shown in Table 3.1.

Table 1. Bin number per corrugation class.

Class	1	2	3	4	5
Bin count	630	339	347	152	220

One concern, as seen in the table, is the sample size for strong corrugation, which is very small compared to the number of bins with weak corrugation. Overfitting thus poses a particular problem for corrugation categories 4 and 5, as their bin counts are considerably smaller than the number of available features. Furthermore, a very limited number of strongly corrugated track sections account for most of the bins in these two classes, meaning that independent variables such as track geometry or vehicle dynamics tend to be very similar across these neighbouring bins. Oversampling techniques such as SMOTE [10] together with undersampling of the majority classes are currently being tested to resolve this problem, but even for classes with larger samples, the number of features available still surpasses the number of observations.

To reduce the chance of overfitting, the initial feature space, containing a total of 248 features, is reduced by

(i) assuming the front and rear axle bearings on the same side of the vehicle are redundant for this type of analysis and

(ii) a principal component analysis (PCA) on the annotated and standardised dataset to find the most relevant orthogonal axes in the remaining, transformed feature space.

To this end, all further analyses are restricted to the two front sensors and the PCA showed that the first 30 components account for over 0.93 of the dataset's variance.

3.2. Machine learning algorithms

Prior to training the classifiers, the labelled data was split into a random but stratified set of training (70%) and test data (30%). Stratification ensures that the relative occurrence of each category is maintained within each subset, which is of particular importance given the small sample size of highly corrugated bins.

Out of a choice of the most popular classification methods, logistic regression (LR), random forests (RF) and support vector machines (SVM) were the three classifiers investigated in this project. These models were optimised in terms of classification accuracy (0-1) for the given training scenario using a grid search on their respective hyperparameters. Tuning these high level parameters was performed through a 10-fold cross validation on the training dataset. The achievable accuracy for the test set for each classification method is presented in Table 3.2.

Table 2. Accuracy of the investigated classification methods: logistic regression (LR), random forests (RF) and support vector machines (SVM).

Classifier	LR	RF	SVM
Accuracy	0.5	0.6	0.7

For SVM, a radial basis function kernel with C=10.0 and gamma=0.01 delivered the highest subset classification accuracy, with a score of 0.7 when applied to the test dataset. Herein, C is a regularization parameter also referred to as the soft margin constant. Large values of C lead to greater penalization of misclassification, producing very narrow decision boundaries, whereas small values of C give hyperplanes with broader margins that are better at generalising when applied to new data, but produce more classification errors. Gamma is an inverse-width parameter of the kernel. For small values, the decision boundary is very simple. For larger values of gamma, the complexity of the boundary increases, yielding more accurate classification of the training data but potential overfitting.

For the RF classifier, a forest of 100 trees with a maximum tree depth of 15 and a minimum number of 4 samples required to split an internal node yielded the highest classification score of 0.6.

For LR, the liblinear solver with an L2 penalty and C=0.1 produced the best classification score of only 0.5 on the test dataset. Given the multi-class nature of the problem, this result is better than a randomly picked class but clearly still much lower than desired.

At this point it should be noted that for linear classifiers, adjusting C does not actually have a great impact on the classification accuracy in the case of non-linearly separated data. To visualize the non-linear corrugation class separation, the first two principal components employed in this study are plotted in Fig.2.

One word of caution about this procedure: The 10-fold cross validation randomly selects stratified training and test datasets for optimising the performance of the models. Adjacent bins along the tram tracks may, however, not be fully independent due to only gradually or even un-changing track geometries, track superstructure and driving conditions. Hence there may be considerable correlation between training and testing datasets that could lead to a potential overfit. One way to handle this could see each fold consisting of a labelled dataset of neighbouring bins (assuming the track properties are the same and enough samples from each corrugation class are available). Due to the current lack of annotated data from different regions of the city's network,

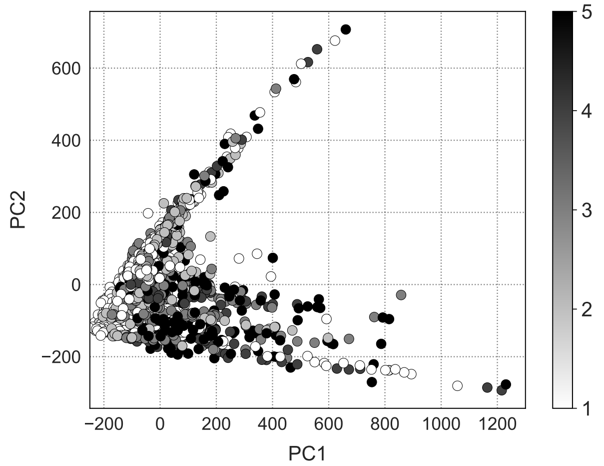

Fig. 2. Non-linear corrugation class separation using the first two principal components (PC).

the authors will follow up on this issue in the next stage of the study when a greater diversity of labelled data in terms of location becomes available.

3.3. Classifying corrugation acrosss the network

The SVM with its optimised hyperparameters (highest classification accuracy out of the three investigated models) is trained using the entire labelled dataset and subsequently applied to the unlabelled dataset. That way, a total of 30 months' worth of records from the inspection tram, covering over 8000 km of track had a corrugation class assigned to them. As Vienna's network is 417 km long, these records already provide multiple coverage of all tracks at regular intervals. Plotting the corrugation extent on a map of the city's tram network allows the identification and localization of hot spots, which are potentially in need of rail grinding to reduce the corrugation depth. Furthermore, the classification of repeated surveys over the same stretch of track will allow the development of corrugation to be monitored over time for better maintenance planning.

One issue that still needs resolving in the current classification procedure is the presence of other rail head irregularities that lead to features which the algorithm may mistake for corrugation. While switches, squats or rail breaks produce impulsive signals with a high transient component that can be differentiated from harmonic signals, other defects such as spalling may be more difficult to detect. The reliability of the classification algorithm will thus need in-situ validations at the current stage.

4. CONCLUSION AND OUTLOOK

In this study, the use of machine learning algorithms on vibration acceleration features was tested to classify the extent of corrugation of tram rails in Vienna. The classification is performed on bins of 5 m length, using vibration signals from the front axle bearings of an unpowered bogie on an inspection tram that covers the entire network biannually. The tram's velocity was found to be the predominant independent variable which affects the magnitude of nearly all investigated features. In order to eliminate the vehicle's influences, it was compensated through the use of a regression model to obtain the expected magnitude of each feature in each bin at the current velocity.

Other independent variables, which will be looked at in the next stage of the study are the type of superstructure (ballasted track versus the prevalent grooved rails in the road surface) and temperature influences.

Currently, the support vector machine defined for this task yields a classification accuracy of only 70%, which the authors aim to improve upon by

(i) incorporating new vibration features that could be indicative of corrugation, such as data from the bogie accelerometer or axle vibration in the lateral direction,

(ii) including acoustic features from the onboard microphone,

(iii) further compensating the influence of the survey vehicle on the data by investigation cepstrum coefficients, and

(iv) improving the data quality of training data by finding more bins (labelled data) with strong corrugation from different origins across the network.

Being able to produce an accurate map of the latest corrugation status would provide the infrastructure operator with an ideal support for decision-making during long- and midterm maintenance planning and support staff during in-situ inspections.

5. ACKNOWLEDGEMENT

The work presented herein is part of a project sponsored by the Austrian Research Promotion Agency. The authors would also like to thank their colleagues at Wiener Linien for their constant support in this work.

6. REFERENCES

[1] Stuart L Grassie, "Rail corrugation: advances in measurement, understanding and treatment," *Wear*, vol. 258, no. Issues 78, pp. 1224–1234, 2005.

[2] Stuart L. Grassie, "Short wavelength rail corrugation: field trials and measuring technology," *Wear*, vol. 191, no. 1, pp. 149 – 160, 1996, 4th International Conference on Contact Mechanics and Wear of Rail-Wheel Systems.

[3] Jn Mandula, Brigita Salaiov, and Mria Koval'akov, "Prediction of noise from trams," *Applied Acoustics*, vol. 63, no. 4, pp. 373 – 389, 2002.

[4] S.L. Grassie, "11 - rail corrugation," in *WheelRail Interface Handbook*, R. Lewis and U. Olofsson, Eds., pp. 349 – 376. Woodhead Publishing, 2009.

[5] Ákos Vinkó and Péter Bocz, "Experimental investigation on condition monitoring opportunities of tramway tracks," *Periodica Polytechnica Civil Engineering*, vol. 62, no. 1, pp. 180 – 190, 2018.

[6] H. Tanaka, M. Matsumoto, and Y. Harada, "Application of axle-box acceleration to track condition monitoring for rail corrugation management," in *7th IET Conference on Railway Condition Monitoring 2016*, 2016.

[7] M. Bocciolone, A. Caprioli, A. Cigada, and A. Collina, "A measurement system for quick rail inspection and effective track maintenance strategy," *Mechanical Systems and Signal Processing*, vol. 21, no. 3, pp. 1242 – 1254, 2007.

[8] Christian Gorges, Kemal ztrk, and Robert Liebich, "Impact detection using a machine learning approach and experimental road roughness classification," *Mechanical Systems and Signal Processing*, vol. 117, pp. 738 – 756, 2019.

[9] R. Wehr, A. Fuchs, and W. Wehr, "Akustische zustandsbewertung von gleisen im straßenbahnnetz," in *Tagungsband DAGA 2017 - 43. Jahrestagung für Akustik*, March 2017, p. 4.

[10] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 1, pp. 321–357, 2002.