Drive-by Damage Detection in Bridges using Mel-frequency Cepstral Coefficients and Machine Learning

Zhenkun Li[0000-0002-1444-6017] Weiwei Lin[0000-0001-8486-6538]

Youqi Zhang[0000-0003-4555-5061]

Department of Civil Engineering, Aalto University, Rakentajanaukio 4 A, 02150 Espoo, Finland.
zhenkun.li@aalto.fi

Abstract. Drive-by damage detection for bridges has caught much attention in the last decades. A salient advantage of this method is that only a few sensors are instrumented on the passing vehicle instead of monitoring systems on the bridge itself. Damage detection considering the bridge's frequencies extracted from the vehicle's vibration data was a promising way confirmed by scholars worldwide. However, the current research is typically concerned with low frequency responses of the bridge. High frequency responses that contain bridge damage information are ignored. To detect the bridge's damage accurately, both low and high frequency responses of the passing vehicle are considered in this paper. Firstly, the vehicle's frequency responses are utilized as input features to train machine learning models to predict whether the bridge is damaged or not. Then, the efficiency of training is improved by projecting the Hertz scale frequency responses into the Mel scale to reduce the dimensions of inputs, in which the Mel-frequency Cepstral Coefficients (MFCCs) are used to feed machine learning models. To verify the effectiveness of the proposed method, a lab-scale I-shaped simply supported beam and a model car are employed. The results demonstrate that the proposed method is promising for damage detection.

Keywords: Structural health monitoring, Drive-by, Vehicle bridge interaction, Mel-frequency Cepstral Coefficients, Machine learning.

1 Introduction

As one of the most critical infrastructures, bridges have been aging and deteriorating in the past decades. In Europe, a large number of bridges are built before world war II and have served for more than half a century [1]. Monitoring their healthy conditions becomes crucial in the near future [2]. A promising way to detect the damage is to extract the dynamic features from its vibration data [3]. Such dynamic features typically include natural frequencies, modal shapes, damping [4,5], etc. Dynamic features before and after damage can be used as references for damage detection.

Traditionally, sensors are installed on the bridge directly to collect its vibration data, and good results have been obtained [6]. However, to apply this method in practical engineering, a monitoring system is generally needed for a unique bridge. Maintaining such a system will be pretty costly. To overcome this problem, the drive-by damage detection method was proposed by Yang et al. in 2004 [7]. The bridge's fundamental frequency is successfully extracted from the vehicle's vibration data. The drive-by method is economical and easy to operate since it only needs a few sensors installed on the passing vehicles.

Frequencies, as the fundamental property of the bridge, have been commonly researched in the past two decades using the passing vehicle's vibration data [8]. The bridge's natural frequencies, modal shapes, and damping can be identified from the vehicle's vibration data, but the practicability of the method is still limited to low vehicle speed, road roughness, etc. Many difficulties need to be overcome before applying the drive-by method to practical engineering. Due to the development of computer computational capability, machine learning techniques have been adopted in the field of bridge damage detection [9]. However, in current literature, only the peaks or low range of the vehicle's frequency responses are analyzed; high frequency responses are normally ignored because they are easily contaminated by noises.

This paper proposes a strategy utilizing both the vehicle's low- and high-frequency responses. The approach is explored using a steel bridge model and a model car in the laboratory. Mel-frequency cepstral coefficients (MFCCs) are extracted from the vibration data to improve the efficiency of damage detection. The overview of this paper is as follows: Section 2 explains the steps to extract MFCCs and the basic principles of support vector machine (SVM) and logistic regression. Section 3 introduces the experimental setups and damage cases. The results are discussed in Section 4. Finally, this paper is concluded in Section 5.

2 MFCCs, SVM, and logistic regression

2.1 MFCCs

MFCC is a particular cepstrum that has been proved as an effective method in acoustic feature identification. Compared to traditional analysis of frequency-domain responses, MFCCs do not just focus on peaks but a range of frequency responses. It has been proved that MFCCs can perform well in structural health monitoring problems [10,11]. MFCCs are robust when extracting damage-sensitive features under the influence of noises. In this paper, MFCCs are utilized to reduce input dimensions and improve computational efficiency. The original mutual transitions between Hertz and Mel frequency scale are shown in Eq. (1),

$$f_{mel} = 2595 \log_{10}(1 - f_{hz}/700) \tag{1}$$

where f_{mel} is Mel-scale frequencies and f_{hz} is Hertz-scale frequencies. However, due to the reason that the bridge's vibration signals are different from sound signals, the traditional relationship between Mel and Hertz scales has to be reformed before applying to bridge damage detection [12]. Since the bridge's first natural frequencies are within 100 Hz, coefficients in Eq. (1) can be reformed as Eq. (2). The updated relationship between Mel and Hertz frequencies is shown in Figure 1.

$$f_{mel} = 5\, ln(1 - f_{hz}/5) \tag{2}$$

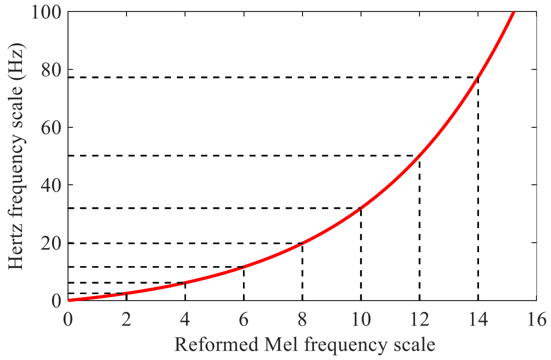


Fig. 1. Relationship between reformed Mel scale and Hertz scale

To extract MFCCs from the passing vehicle's vibration data, there are five steps to be executed: (1) Data preprocessing; (2) Fast Fourier Transform (FFT); (3) Mel Filterbank; (4) Logarithm; (5) Discrete Cosine Transform (DCT). Before transferring the original signals from time-domain to frequency-domain, the Hann window is utilized to avoid frequency spectrum leakage induced by signal truncation, and FFT is employed to transform the windowed frame into frequency responses. For each energy spectrum, Eq. (2) is used to project it into Mel frequency scale. Then the energy spectrum in the Hertz scale is convolved with Mel filter banks. After that, a logarithm and a DCT are applied for each bank to obtain the final MFCCs. This process can be represented in Eq. (3),

$$\boldsymbol{m}_i = D\left(ln\left(Mel(|F(\boldsymbol{x}_i)|^2)\right)\right), i = 1,2,3,\cdots,n \tag{3}$$

where $\boldsymbol{x}_i$ is the original accelerations, and $\boldsymbol{m}_i$ is the vector of MFCCs. F is the Fast Fourier Transform, and Mel represents the transformation from Hertz scale to Mel scale. D means Discrete Cosine Transform. After the above steps, the final MFCCs can be obtained.

2.2 SVM and logistic regression

SVM. SVM is a popular classifier in machine learning, and it has been utilized in structural health monitoring due to its good explainable property. The basic principle of SVM is to maximize the margin between two classes using a hyperplane. If the datasets own m samples presented by $\{(\boldsymbol{x}_i, y_i), i = 1,2,3, \dots, m; y_i \in \{1, -1\}\}$. For a linear classification problem, the optimal hyperplane can be represented as

$$\boldsymbol{w} \cdot \boldsymbol{x}_i + b = 0, \tag{4}$$

where $\boldsymbol{w}$ is the weight vector, $\boldsymbol{x}_i$ represents all data points on the optimal hyperplane, and b is the bias. The proposed problem can be solve by introducing standard Lagrange multiplier method that can be found in reference [13]. When the two classes cannot be linearly separated, the kernel function K is introduced. K can project all features to a higher space and make two classes linearly separable. General kernel functions and their hyperparameters can be found in Table 1.

Table 1. Kernel functions and hyperparameters.

Kernel	Formulars	Hyperparameters
Linear	$K(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$	C
Polynomial	$K(\boldsymbol{x},\boldsymbol{y}) = (\gamma(\boldsymbol{x} \cdot \boldsymbol{y}) + r)^d$	C, γ, r, d
Sigmoid	$K(\boldsymbol{x},\boldsymbol{y}) = \tanh\,(\gamma(\boldsymbol{x} \cdot \boldsymbol{y}) + \boldsymbol{r})$	C, γ, r
Radial basis function (RBF)	$K(\boldsymbol{x},\boldsymbol{y}) = e^{-\gamma\lVert\boldsymbol{x}-\boldsymbol{y}\rVert^2}$	C, γ

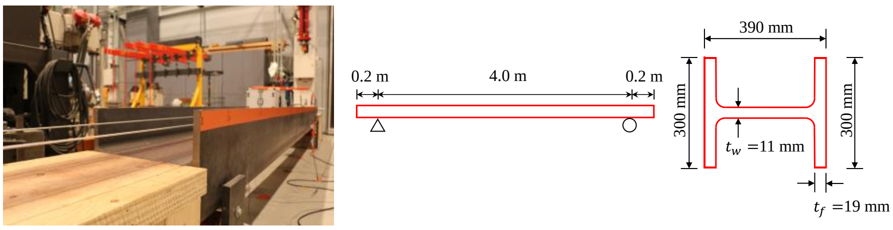
Logistic regression. Logistic regression is a good linear classifier in machine learning. It is based on linear regression that can be represented as Eq. (5),

$$z = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n = \boldsymbol{\theta}^T \boldsymbol{x} \tag{5}$$

where $\boldsymbol{\theta}$ is the coefficient vector for the linear regression model. $\boldsymbol{x}$ are features ($x_0 = 1$). Utilizing Eq. (5), all features are transformed to a continuous value z. However, when the labels are different classes. It cannot be used to calculate the loss between true class and predict value. To solve this problem, the linking function, Sigmoid ($g(z) = 1/(1 + e^{-z})$), is employed to project all values into 0-1. The projected value $g(z)$ is regarded as the possibility belonging to class 1, so the possibility for class 0 can be obtained as $1 - g(z)$. The loss for the logistic regression model is deduced by Maximum Likelihood Estimation (MLE) which can be found in reference [14]. To overcome the overfitting problem, l_2 regularization is utilized in this paper.

3 Lab-scale experiments

3.1 Experimental setups



(a) Bridge and car **(b)** Boundary conditions **(c)** Cross-section

Fig. 2. Model of the vehicle bridge interaction system

To verify the proposed method, a lab-scale experiment is performed. In the experiment, an I-shaped simply supported HEA400 beam made by Q355 is utilized. The beam's length is 4.4 m in total, and its support length at each end is 0.2 m. The total mass of the beam is 550 kg. The beam, its cross-section, and the boundary conditions can be found in Figure 2.

A model car driven by a remote unit is utilized to simulate a vehicle on the bridge. The car's mass is 9.462 kg, with the front axle of 4.315 kg and the rear axle of 5.147 kg. Two accelerometers are installed on the car's front and rear axles to collect vibration data. The sampling frequency is set as 10 kHz. The experiment is performed in the structural laboratory at Aalto University with environmental noises.

3.2 Bridge's damage

Damages in experiments are normally simulated by stiffness reduction, but it is difficult to find real damage in engineering. Since this paper analyzes the vehicle's frequency responses and the bridge's natural frequencies are determined by its stiffness and mass matrices, the method of adding mass to the beam is adopted [15]. There are three cases in total: 0 kg (intact), 2 kg (minor damage), and 20 kg (large damage), which are summarized in Table 2. The mass is added by a hook weighing 2 kg. The damage degree is defined as the ratio of added mass to the beam's mass.

Table 2. Damage cases.

Cases	Intact	Minor damage	Large damage
Added mass	0 kg	2 kg	20 kg
Position	-	0.5 L[1]	0.4 L
Runs	563	67	66
Damage degree	0.00 %	0.73%	4.00 %

3.3 Machine learning training and testing

In this paper, the sci-kit learn package [16] in Python 3.8 environment is employed to train the SVM and logistic regression models. Both frequency responses and MFCCs are utilized as input features. Before training, all features are normalized to eliminate the influence of different scales. For each training, the same samples are selected from intact cases and damaged cases to avoid data imbalance problems. For example, when identifying large damages, 66 samples are randomly selected from the intact cases. This process is executed ten times to circumvent the occasionality of selection from intact cases. The 5-fold cross-validation (CV) strategy is employed to obtain the testing accuracy for damage detection. In this case, if the accuracy is near 50%, it means that the machine learning model cannot classify the intact and damaged cases.

4 Experimental results and discussions

4.1 Frequency-based SVM

As mentioned before, only using low frequency responses is normally hard to determine the bridge's healthy conditions. In this paper, high frequency responses of the vehicle's vibration are explored. Firstly, the large damage case and health damage case are utilized for analysis. Firstly, all kernels' hyperparameters are set as: $C = 1.0, \gamma = 0.01, r = 0, d = 3$ for the SVM model. The 5-folder CV accuracy results with respect to used frequency responses are shown in Figure 3.

It can be seen from Figure 3 that with the increase of used frequency responses, the accuracy increases as well ("Linear kernel"). When the selected frequency range is

[1] L: the beam's span length (4.0 m, the supportive length is not included)

0~200 Hz, all four kernels' accuracy is relatively low. If 0~400 Hz frequency responses are utilized for analysis, the CV accuracy becomes higher ("Linear kernel"). After more than 400 Hz frequency responses are employed, the accuracy becomes stable. Also, it can be seen that when "Polynomial", "RBF", and "Sigmoid" kernels are used, the accuracy is poor. This is because these three kernels' hyperparameters normally need to be adjusted to suit the special problem, and a grid search strategy is needed to improve their accuracy. However, with the increase of selected frequency responses (input features), the computation will become heavier. Grid search requires enormous amounts of computational resources and even cannot be achieved. Thus, dimension reduction is needed for frequency responses.

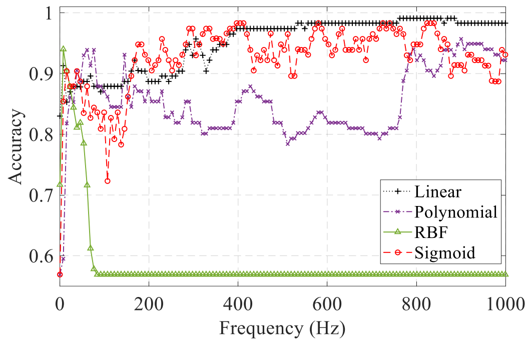


Fig. 3. Accuracy with respect to selected frequency range

4.2 MFCCs-based SVM

To improve the computational efficiency, the vehicle's frequency responses are transformed to MFCCs using Eq. (3). An important step before the transformation is to select a suitable number of Mel filter banks. More filter banks mean that the frequency responses are divided finely, but the computational load will also increase. If few Mel banks are utilized, the damage sensitive features may not be extracted successfully. The CV accuracy with respect to the number of Mel filter banks can be found in Figure 4. It is worth noting that the accuracy is computed using all kernels with the grid search strategy, and only the best accuracy is selected to plot.

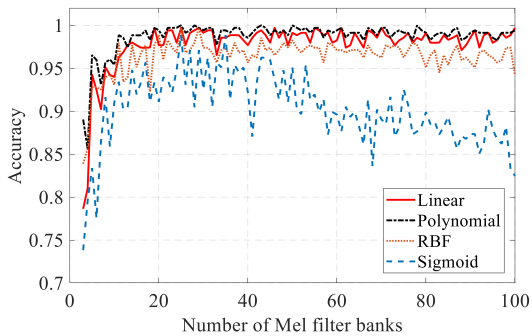


Fig. 4. Accuracy with respect to selected filter banks

It can be seen from Figure 4 that when the filter banks are few (3-10 banks), the accuracy is low. When 20 Mel filter banks are utilized, the CV accuracy can reach a relatively high point, after which the accuracy becomes stable. In this paper, 25 filter banks are selected. Also, it can be seen that the "Polynomial" kernel can perform the best in the bridge damage detection problem, and the "Linear" kernel is the second winner. However, the "Polynomial" kernel owns four hyperparameters as shown in Table 1. Finding its best performance with grid search will cost much time. Therefore, the "Linear" kernel is selected for analysis in this paper.

4.3 MFCCs-based logistic regression

From the above analysis, we can see that the SVM model with a "Linear" kernel can identify the bridge's condition with high accuracy. Thus, it can be inferred that in the high dimension space of MFCCs, the intact and damaged cases are linearly separable. Logistic regression, as a good linear classifier, is tested in this section.

For the l_2 regularization hyperparameter C, smaller C means greater penalty on the logistic regression model's weights. In this paper, $C = 1.0$ is selected. When 25 MFCCs are employed, the 5-fold CV accuracy of damage detection can reach 100 %. It means that the intact and damage cases can be linearly separately when MFCCs are utilized as damage indicators.

4.4 Damage detection for different damage severity

For multiple classification problems, there are two strategies: one vs. one (OVO) and one vs. rest (OVR). In this paper, OVR is adopted. 66 runs from intact cases are randomly selected, and there are 199 runs together. The 5-folder CV result of logistic regression is 92.9%, and for SVM, the detection accuracy is 93.4%. It can be seen that both these two models can identify the bridge's damage with relatively high accuracy.

5 Conclusions

In this paper, a method utilizing the passing vehicle's vibration data to detect the bridge's damage is proposed. MFCCs, initially used for acoustic recognition, were used to reduce the dimension of frequency responses. Then, two machine learning models, SVM and logistic regression, are employed to detect the bridge's damage. The results show that when MFCCs are utilized, the training efficiency is greatly improved, and the accuracy of damage detection is high. The main conclusions are: (1) Both low and high frequency responses of the passing vehicle contain damage information of the bridge. With the increase of used frequency responses, the accuracy of damage detection using machine learning models increases; (2) Utilizing the extracted MFCCs from original vibration signals, dimensions of input for machine learning models can be reduced greatly, and the computational efficiency is improved accordingly.

Acknowledgment

This study was fully financial sponsored by the Jane and Aatos Erkko Foundation in Finland. Y. Zhang is financially supported by the Academy of Finland (Decision number: 339493). The supports on the vibration tests provided by the staff of the laboratories in the Department of Civil Engineering and the Department of Mechanical Engineering at Aalto University are also greatly appreciated.

References

1. Keeping European bridges safe. 2019. https://ec.europa.eu/jrc/en/news/keeping-european-bridges-safe, last accessed 2022/4/8
2. Fujino Y, Siringoringo DM.: Bridge monitoring in Japan: The needs and strategies. Structure and Infrastructure Engineering; 7(7–8), 597–611 (2011).
3. Hou R, Xia Y.: Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. Journal of Sound and Vibration; 491, 115741 (2021).
4. Zhang J, Qu CX, Yi TH, Li HN.: Damage detection for decks of concrete girder bridges using the frequency obtained from an actively excited vehicle. Smart Structures and Systems; 27(1), 101–114 (2021).
5. Li Z, Hou J, Jankowski Ł.: Structural damage identification based on estimated additional virtual masses and Bayesian theory; 65(45), 1–18 (2022).
6. An Y, Chatzi E, Sim SH, Laflamme S, Blachowski B, Ou J.: Recent progress and future trends on damage identification methods for bridge structures. Structural Control and Health Monitoring; 26(10), 1–30 (2019).
7. Yang YB, Lin CW, Yau JD.: Extracting bridge frequencies from the dynamic response of a passing vehicle. Journal of Sound and Vibration; 272(3–5), 471–493 (2004).
8. Nagayama T, Reksowardojo AP, Su D, Mizutani T.: Bridge natural frequency estimation by extracting the common vibration component from the responses of two vehicles. Engineering Structures, 150: 821–829 (2017).
9. Malekjafarian A, Golpayegani F, Moloney C, Clarke S.: A machine learning approach to bridge-damage detection using responses measured on a passing vehicle. Sensors (Switzerland), 19(18) (2019).
10. Zhang G, Harichandran RS, Ramuhalli P.: Application of noise cancelling and damage detection algorithms in NDE of concrete bridge decks using impact signals. Journal of Nondestructive Evaluation, 30(4): 259–272 (2011). DOI: 10.1007/s10921-011-0114-8.
11. Kong Q, Zhu J, Ho SCM, Song G.: Tapping and listening: A new approach to bolt looseness monitoring. Smart Materials and Structures, 27(7) (2018).
12. Mei Q, Gül M, Boay M.: Indirect health monitoring of bridges using Mel-frequency cepstral coefficients and principal component analysis. Mechanical Systems and Signal Processing, 119: 523–546 (2019).
13. Cortes C, Vapnik V.: Support-vector networks. Machine Learning; 20(3): 273–297 (1995).
14. Cox DR. The regression analysis of binary sequences. Journal of the Royal Statistical Society: Series B (Methodological), 20(2): 215–232 (1958).
15. Zhang Y, Miyamori Y, Mikami S, Saito T.: Vibration-based structural state identification by a 1-dimensional convolutional neural network. Computer-Aided Civil and Infrastructure Engineering, 34(9): 822–839 (2019).
16. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al.: Scikit-learn: Machine learning in Python. The Journal of Machine Learning Research; 12, 2825–2830 (2011).