

Optical fiber sensing and tensor AP clustering for high-speed road tunnel vehicle detection

Heng Li^{1, 2}, Ziyang Feng², Shaohua Xu², Mingde Zheng², Wentao Zhang³,
Feiyu Zheng^{2 *}

1. GuangXi BeiBu Gulf Investment Group Co., Ltd., Nanning, Guangxi, P.R. China.

2. Guangxi Beitou IT Innovation Technology Investment Group Co., Ltd., Nanning, Guangxi, P.R. China.

3. Guilin University of Electronic Technology, Guilin, Guangxi, P.R. China.

lih@bgigc.com, 84686276@qq.com, xush@bgigc.com, zhengmd@bgigc.com, glietzw@163.com,
zfy@bgigc.com.

*Corresponding author: Feiyu Zheng, E-mail: zfy@bgigc.com

Abstract. Aiming at the problems of signal acquisition, feature extraction and vehicle recognition in highway tunnel vehicle detection, a new tunnel vehicle detection method is proposed by combining optical time-domain reflectometry distributed fiber sensing technology and tensor affine propagation clustering algorithm. Firstly, the distributed optical fiber system designed by optical time domain reflectometry was used to collect the running signals of tunnel vehicles and obtain the measurement data. Secondly, a high-order tensor sample set is constructed by using the spatial resolution of optical fiber as channel number and combining the feature number, time domain and frequency domain. Finally, tensor affine propagation clustering method and other clustering methods are used to test the accuracy. The test results show that the proposed method can better classify vehicles without destroying the original high-dimensional data structure. Meanwhile, the unsupervised clustering algorithm also reduces manual intervention in the identification process, increases the intelligence level of the whole vehicle detection model, and effectively improves the detection accuracy rate of tunnel vehicles.

Keywords: Vehicle detection, optical fiber sensing, tensor AP clustering, signal processing, intelligent transportation.

1. Introduction

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper. Use *italic* for emphasizing a word or phrase. Do not use boldface typing or capital letters except for section headings (cf. remarks on section headings, below).

The tunnel is the “throat” of the highway. If an accident occurs in a tunnel, it will cause serious economic losses and even casualties [1]. It is an important tool for highway operation and tunnel safety to able to monitor the vehicles running in the tunnel in real time and provide early warning before or during an accident. With the deepening of the construction, the span of the tunnel is getting longer and longer, which places higher demands on the long-distance monitoring and monitoring system. There are many studies devoted to long-distance monitoring of tunnel vehicles, such as acoustic emission [2], video [3], microwave [4], infrared [5], etc. Most of the above sensing methods are fixed with position point-type sensing structures that are vulnerable to electromagnetic interference and require real-time on-site power supply, which is difficult to meet the needs of distributed, long-distance, and all-weather highway monitoring.

With the development of measurement and sensor technology, more and more sensors can be used to monitor vehicles [6]. Among them, optical fiber sensing technology has developed rapidly

in recent years, which provides a new principle and method for tunnel vehicle detection. Especially for large and long tunnels with harsh operating environment, complex working conditions and terrain structure, optical fiber sensing presents great advantages [7]. The distributed optical fiber sensing system based on Rayleigh scattering has the advantages of high sensitivity and a wide range of vibration spectrum measurement [8]. Compared with other scattering mechanisms, Rayleigh scattering has a comparative advantage in the application of vibration event measurement, which can realize the real-time measurement of the running state of tunnel vehicles with "passive multi-field and one-line multi-point", and can also transmit optical signals through non-contact mode [9]. In addition, different detection methods have a great influence on the accuracy of vehicle detection. The conventional vehicle detection method based on video image class is difficult to apply to the signal generated by fiber optic vibration system. It is still necessary to research and design a special vehicle identification method for the detection of tunnel vehicle. Affine Propagation (AP) is a method of partitioning around central points [10]. It can find the cluster center adaptively by propagating the information affinity without inputting class information. It is insensitive to the initialization of the similarity matrix, and always achieves deterministic results in iteration.

In this work, we propose a tunnel vehicle method which combines the distributed optical fiber sensing technology of optical time domain reflectometry with a tensor affine propagation clustering algorithm. Firstly, a distributed optical fiber system based on optical time domain reflectometry is used to collect the running signals of the tunnel vehicles and obtain the measurement data. Then, the measurement data is divided and the statistical features are extracted to generate the feature tensor sample set. Finally, the tensor AP clustering method is used to divide the vehicle categories and complete the vehicle detection.

2. Distributed optical fiber sensing system for optical time domain reflectometry

Incoherent light pulses enter the sensing fiber through the ring after being emitted by the laser. Due to the inhomogeneity of the fiber medium and the influence of high-power laser on the transmission medium, the backward-Rayleigh scattered light is constantly generated during the propagation process. The backward-Rayleigh scattered light, which carries external vibration information after passing through the ring is received by the photodetector [11]. Finally, it is collected by the data acquisition card and transmitted to the host computer for data processing, and its optical path diagram is shown in Fig. 1.

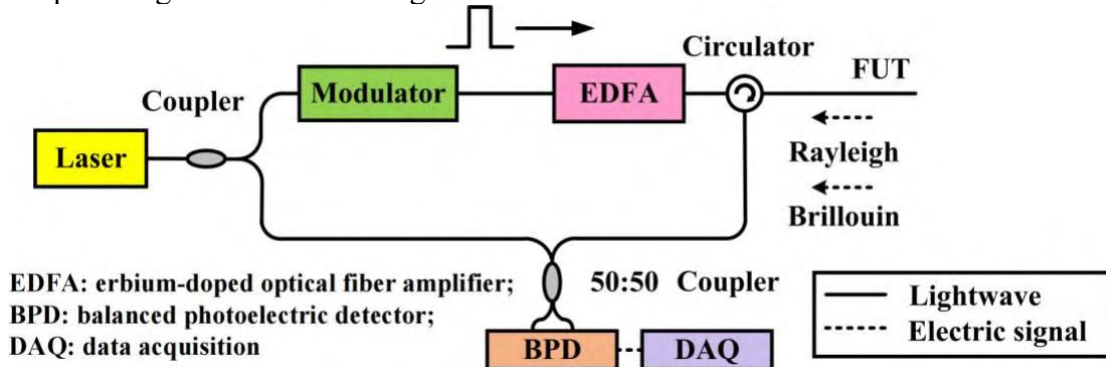


Fig 1 Optical path diagram of the Phi-OTDR system

Each point on the fiber will return the backward-Rayleigh scattered light, so each position on the fiber has a corresponding backward-Rayleigh scattered light intensity, and the light intensity information can reflect the information distribution of the fiber link. If there is a loss, attenuation, break or vibration at a particular point in the fiber link, the back-Rayleigh scattered light intensity generated at that point will change accordingly. By measuring the back-Rayleigh scattered light

intensity at the end of the same fiber cable, the position of the above vibration event in the fiber link can be determined by formula (1).

$$L = vt/2n \quad (1)$$

where L is the position of the optical fiber corresponding to the backward-Rayleigh scattered light detected at a given time, t is the time taken for the backward-Rayleigh scattered light generated by the optical pulse transmitted to L to return to the first end of the optical fiber, v is the speed of light in vacuum, and n is the refractive index of the optical fiber core. In this way, long-distance distributed measurement of optical fiber can be realized. Assuming that P₀ is the optical power of the input pulse of the system, and the power of the back-Rayleigh scattering at the input end of the optical pulse is P_s, it can be expressed as:

$$P_s = \frac{v_g}{2} P_0 S \alpha_s W \quad (2)$$

where V_g is the group velocity of the transmitted light, a term of 1/2 is added to the formula to account for the loss introduced by the coupler in the system, S is the backscatter capture rate of light, alpha_s is the scatter component in the fiber attenuation coefficient, and W is the pulse width of the pulsed light input to the system. The backscatter capture rate S of the light wave can be expressed as follows:

$$S = 0.25(NA/n_0)^2 \quad (3)$$

where NA is the numerical aperture of the fiber, the larger the value, the better the optical receiving effect of the fiber, n₀ is the refractive index of the optical fiber core. The propagation of pulsed light in the fiber will have a certain degree of attenuation, and its attenuation coefficient is alpha, then the peak power obtained by the pulsed light at the position L on the fiber after attenuation can be expressed as:

$$P_L = P_0 e^{-\alpha L} \quad (4)$$

From the above formula, it can be seen that the backward-Rayleigh scattering curve detected by the system has a decreasing trend, which reflects the loss of light transmission in the fiber. When there is a loss, attenuation, break or vibration anywhere in the fiber link, the forward transmission attenuation coefficient will increase rapidly, resulting in a very obvious drop in the backside Rayleigh scattering curve at the above event point. The location of the above event in the optical fiber link can be determined using the formula (1).

An optical pulse is transmitted in a sensing fiber of length L until the backward Rayleigh scattered signal generated at the end of the fiber is received by the photodetector. The time taken is $tp=2L/vg$, where vg is the propagation speed of light in the fiber. The firing interval between two pulses should be greater than or equal to tp, the maximum repetition frequency of the light pulse $f_{rep} = 1 / t_p = v_g / 2 L$, the maximum frequency response of the system is $f_{max} = f_{rep} / 2$. It can be concluded that the frequency response range of the system is limited by the length of the optical fiber. In long-distance detection, with the increase of the transmission distance, the signal intensity of the detected back-Rayleigh scattered light also decrease, and the random fluctuation of the back-Rayleigh scattered signal will be caused by the frequency drift of the laser, the polarization fading and the electromagnetic noise of the photodetector. To reduce the influence of random noise on detection, the method of adding and averaging signals is generally used to perform simple noise reduction processing, but this will reduce the frequency response range of the system. Assuming that the collected Rayleigh scattering signal is processed N times on average, the frequency response range of the system will be reduced to f_{max}/N . Therefore, the frequency response range of the system is mainly determined by the fiber length and the signal processing method.

3. Tensor affine propagation clustering

Cluster analysis is a kind of research on how to classify data or objects according to their own characteristics, using multivariate statistical methods for comprehensive classification. AP clustering algorithm is a novel clustering algorithm proposed by Canadian scholars in 2007.

Recently, Wei et al. improved it into a tensor AP clustering algorithm [12]. At the beginning of the AP algorithm, all data points are considered as potential class representative points, and automatic clustering is realized through the "message passing" mechanism between each data point. The advantage of this algorithm is that it is fast and efficient, and the AP algorithm does not need to define the number of clusters in advance. In vehicle recognition, it can be assumed that the vehicle category does not need to be known in advance, thus reducing the input of prior knowledge and increasing the intelligence of the whole pattern recognition. The AP algorithm transmits two types of information between each data point: Availability and Responsibility. After many iterations, m excellent class representative points are naturally generated as cluster centers. The message is transmitted according to a simple calculation to find the conditions that make the energy function $E(c)$ reach the minimum:

$$E(c) = - \sum_1^{Num} S(i, c_i) \quad (5)$$

Where: Num is the number of samples; c_i is the representative of point i ; S is the similarity between point i and point c_i , and S is less than 0. Given a data point $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ ($N > 1$), and utilize the x to denote the vector form representation of \mathcal{X} . Therefore, the element $\mathcal{X}_{i_1 i_2 \dots i_N}$ ($1 \leq i_j \leq I_j, 1 \leq j \leq N$) in \mathcal{X} is corresponding to x_l , i.e., the l th element in x , where $l = i_1 + \sum_{j=2}^N (i_j - 1) \prod_{o=2}^{j-1} I_o$ ($2 \leq j \leq N$). Then Tensor distance between two tensors \mathcal{X} and \mathcal{Y} can be represented as:

$$d_{TD} = \sqrt{\sum_{l,m=1}^{I_1 \times I_2 \times \dots \times I_N} g_{lm} (x_l - y_l)(x_m - y_m)} \quad (6)$$

In Eq. (6), g_{lm} is the metric coefficient [14]. The value σ is a regularization parameter and g_{lm} reflects the inherent interaction in various coordinates for high-order space.

$$g_{lm} = \frac{1}{2\pi\sigma^2} \exp \left\{ \frac{-\|p_l - p_m\|^2}{2\sigma^2} \right\} \quad (7)$$

The location distance between the element $\mathcal{X}_{i_1 i_2 \dots i_N}$ (corresponding to x_l) and the element $\mathcal{X}_{i'_1 i'_2 \dots i'_N}$ (corresponding to x_m), which is defined as:

$$\|p_l - p_m\| = \sqrt{(i_1 - i'_1)^2 + (i_2 - i'_2)^2 + \dots + (i_N - i'_N)^2} \quad (8)$$

4. Vehicle detection method based on optical fiber sensing and tensor clustering

4.1 The flow of the proposed method

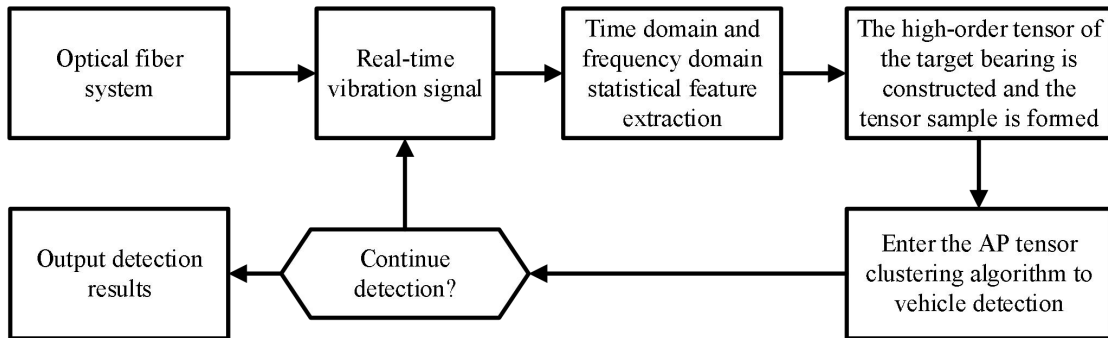


Fig.2 Flow chart of proposed method for high-speed road tunnel vehicle detection

The vehicle detection method proposed is shown in Figure 2. Firstly, the vibration signal of the target vehicle is collected by optical fiber vibration system. Secondly, the statistical features of the vibration signals are extracted in the time domain and frequency domain. Thirdly, the spatial resolution of optical fiber vibration system is used as the number of channels, and the third-order tensor is constructed by combining the characteristics of time domain and frequency domain to form tensor samples of different vehicles. Finally, the obtained tensor samples are input into the AP

tensor clustering algorithm for vehicle detection, and the labels of different vehicle categories are obtained, and the vehicle detection results are output.

4.2 Statistical features and tensor samples

Accurate vehicle detection requires obtaining a feature space with rich vehicle information [14]. Since the time domain analysis is simple and the frequency domain analysis is intuitive, and both can provide vehicle information from different directions, so we adopts the time domain statistical features and frequency domain statistical features to extract the original vibration signals, and then obtain the feature space. The adopted time domain characteristics and frequency domain characteristics are shown in Table 1.

Table 1 Time features (TF) and frequency features (FF)

Num.	formula	Num.	formula
TF_1	$\sum_1^N x(n)/N$	FF_1	$\sum_1^K s(k)/K$
TF_2	$\sqrt{\sum_1^N (x(n) - TF_1)^2 / N - 1}$	FF_2	$\sqrt{\sum_1^K (s(k) - FF_1)^2 / K - 1}$
TF_3	$(\sum_1^N \sqrt{ x(N) } / N)^2$	FF_3	$\sqrt{\sum_1^K (s(k) - FF_1)^3 / K (\sqrt{FF_2})^3}$
TF_4	$\sqrt{\sum_1^N x(n)^2 / N}$	FF_4	$\sum_1^K (s(k) - FF_1)^4 / K \cdot FF_2$
TF_5	$\max x(n) $	FF_5	$\sum_1^K f_k s(k) / \sum_1^K s(k)$
TF_6	$\sum_1^N (x(n) - TF_1)^3 / (N - 1) TF_1^3$	FF_6	$\sqrt{\sum_1^K (f_k - FF_5)^2 s(k) / K}$
TF_7	$\sum_1^N (x(n) - TF_1)^4 / (N - 1) TF_1^4$	FF_7	$\sqrt{\sum_1^K f_k^2 s(k) / \sum_1^K s(k)}$
TF_8	TF_5 / TF_4	FF_8	$\sqrt{\sum_1^K f_k^4 s(k) / \sum_1^K f_k^2 s(k)}$
TF_9	TF_5 / TF_3	FF_9	$\sum_1^K f_k^2 s(k) / \sqrt{\sum_1^K s(k) \sum_1^K f_k^4 s(k)}$
TF_{10}	$N \cdot TF_5 / \sum_1^N x(n) $	FF_{10}	FF_6 / FF_5

In Table 1, $n=1,2,...N$. where N is the number of vibration data points. And $k=1,2,...K$. where K is the number of spectral lines, and represents the frequency value of the k spectral lines. After obtaining the above time domain and frequency domain features, a third-order tensor sample representing a single set of signals is formed by using different space points, different feature numbers, time domain and frequency domain. The tensor sample is input into the AP tensor clustering algorithm to obtain vehicle detection results.

5. Engineering experiment testing

In order to prove the performance of the proposed intelligent detecton method, an OTDR-based fiber optic vibration signal acquisition pilot was established on the expressway to collect the signals of vehicles passing through the tunnel. The OTDR system Settings used in this test are shown in Fig 3. A Narrow Linewidth Laser (NLL) with a bandwidth of 3K is used as the light source, and the frequency shift of the light source is less than 20 MHz. The light source has a wavelength of 1550 nm and a maximum output power of 20 dBm. The light source is first output to a 90:10 optical coupler, which divides the laser into two paths: one is a photosensitive path, and the other is a reference path. In the sensor path, the light source is modulated by an Acousto-Optic Modulator (AOM), resulting in a detection pulse of 10 ns (corresponding to a spatial resolution of 1m). The AOM is driven by an arbitrary waveform function generator. An Erbium-Doped Fiber Amplifier (EDFA) for optical power amplification can compensate for the loss of optical transmission in the fiber. The amplified pulsed light is filtered using a Bragg grating to remove spontaneous noise

removal before being sent to the sensing fiber. The laser pulses were then sent through the circulator (to a single-mode fiber (1 km in length in this experiment)).

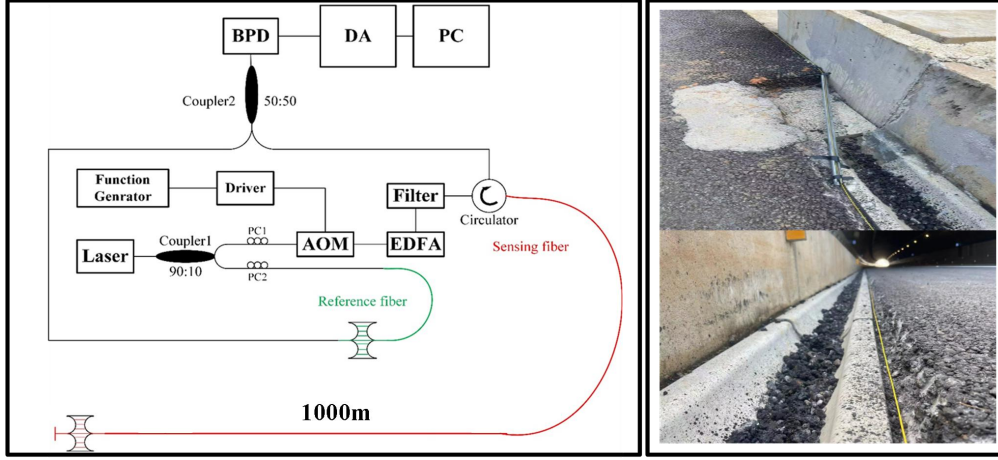


Fig. 3 OTDR system schematic diagram and tunnel optical fiber

We artificially divide the vehicles into three categories, Class A, Class B and Class C. Firstly, different vibration signals of three types of vehicles are obtained, and the obtained vibration signals are divided into signal fragments consisting of 4096 points, which are used as the feature extraction samples. If 60 samples are taken from each of the three vehicles, the total number of samples is 180. The original vibration signal is shown in Figure 4 below. Secondly, 10 time domain features and 10 frequency domain features are calculated in the original vibration signal (Table 1).

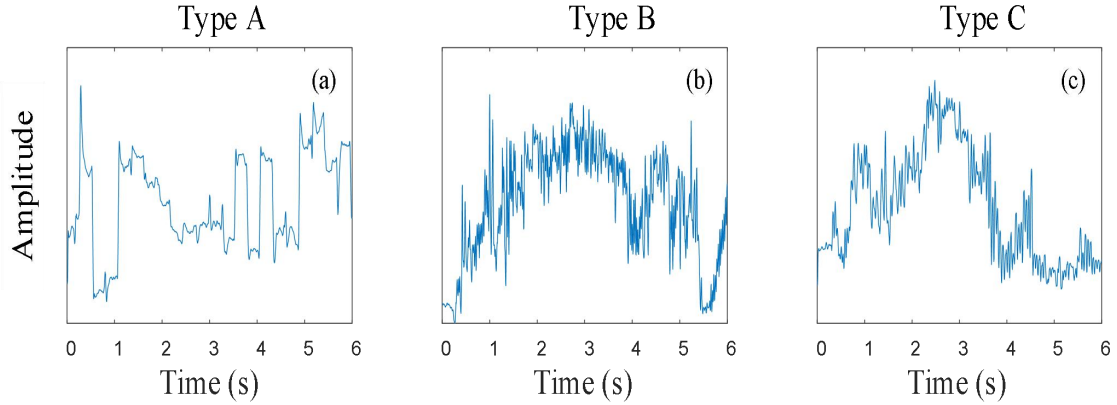


Fig.4 Different types vibration signals of vehicles

Fig. 5 shows the tensor sample constructed in this work. After feature extraction, tensor samples for pattern recognition are constructed from vibration points, different features and different domains obtained by different spatial resolutions of optical fiber vibration, and then sent to the proposed AP tensor clustering algorithm for vehicle detection. The result is shown in Figure6. In addition, these tensor samples are also used for vehicle detection by K-means clustering [13] and density peak search clustering (PDS) [14]. The comparison of results is shown in Table 2. It can be seen that the performance of the other algorithms is slightly worse than the AP tensor method used in this paper.

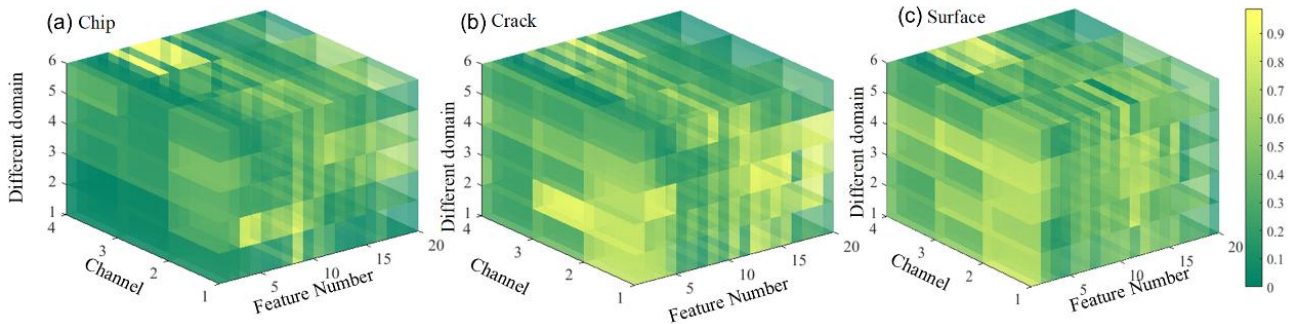


Fig.5 Tensor samples

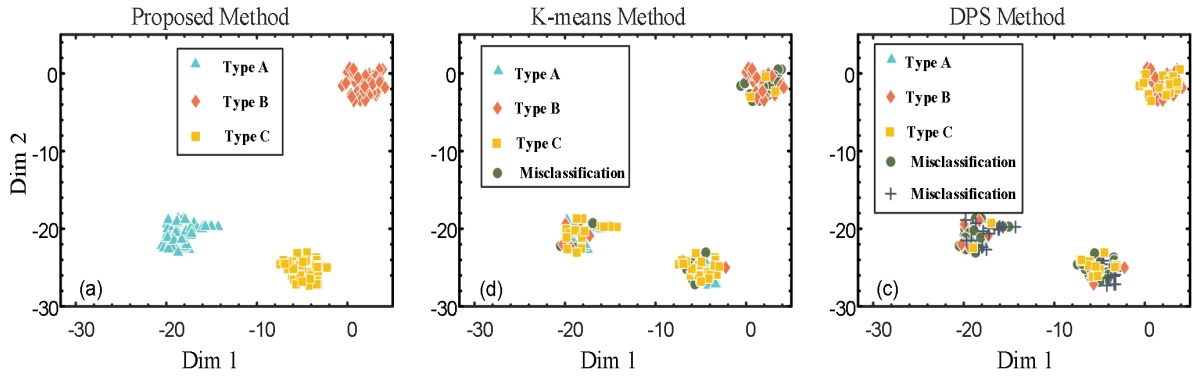


Fig.6 Visual station display of different method recognition results

Table 2 Actual test result

Approach	Vehicles Number	Correct identification	Accuracy
Proposed method	358	351	98.04%
K-means method	369	321	86.99%
DPS method	336	271	80.65%

6. Conclusions

In this work, a kind of optical fiber vibration system is proposed to obtain the vibration signal, and the vehicle detection model is used to detect the vehicle detection of the vehicle. This method firstly uses the optical fiber vibration system to collect the vibration signal, and the vibration signal is divided into the fiber optic vibration channel, the time domain, the frequency domain and the three-order sample. Secondly, the degree of similarity between the samples is measured by the tensor distance, and the low-dimension structure loss of the high dimensional data is avoided, so that the vehicle sample is more suitable for classification, and then the proposed AP tensor clustering method is used to identify the vehicle samples. The experiment shows that the proposed method has better identification effect, and other methods can better identify the vehicle in the high dimensional sample, and better the characteristics of the vehicle. At the same time, the use of unsupervised methods avoids the training of large numbers, reduces the participation of artificial participation, and increases the degree of intelligence of the model.

Acknowledgements

We acknowledge the support from the Guangxi Key R&D projects (AB22035047 & AB23026038).

References

- [1] Zhou Y, Feng S, Li J. Study on the failure mechanism of rock mass around a mined-out area above a highway tunnel—Similarity model test and numerical analysis[J]. Tunnelling and Underground Space Technology, 2021, 118: 104182.
- [2] Schumacher T, Higgins C C, Lovejoy S C. Detection of vehicles with studded tires using acoustic emission sensors mounted to highway bridges[J]. Journal of transportation engineering, 2010, 136(5): 480-487.
- [3] L. X., H. T., J. X., et al. Fine traffic congestion detection with hierarchical description[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 24439-24453.

- [4] Xu J, Hilker N, Turchet M, et al. Contrasting the direct use of data from traffic radars and video-cameras with traffic simulation in the estimation of road emissions and PM hotspot analysis[J]. *Transportation Research Part D: Transport and Environment*, 2018, 62: 90-101.
- [5] Liu X, Wang J, Li J. URTSegNet: A real-time segmentation network of unstructured road at night based on thermal infrared images for autonomous robot system[J]. *Control Engineering Practice*, 2023, 137: 105560.
- [6] Seo Y, Kim S M. Estimation of greenhouse gas emissions from road traffic: A case study in Korea[J]. *Renewable and Sustainable Energy Reviews*, 2013, 28: 777-787.
- [7] Pendão C, Silva I. Optical fiber sensors and sensing networks: overview of the main principles and applications[J]. *Sensors*, 2022, 22(19): 7554.
- [8] Tosi D, Molardi C, Blanc W. Rayleigh scattering characterization of a low-loss MgO-based nanoparticle-doped optical fiber for distributed sensing[J]. *Optics & Laser Technology*, 2021, 133: 106523.
- [9] Liu H, Ma J, Xu T, et al. Vehicle detection and classification using distributed fiber optic acoustic sensing[J]. *IEEE Transactions on Vehicular Technology*, 2019, 69(2): 1363-1374.
- [10] Frey B J, Dueck D. Clustering by passing messages between data points[J]. *science*, 2007, 315(5814): 972-976.
- [11] Di Sante R. Fibre optic sensors for structural health monitoring of aircraft composite structures: Recent advances and applications[J]. *Sensors*, 2015, 15(8): 18666-18713.
- [12] Wei Z, He D, Jin Z, et al. Density-based affinity propagation tensor clustering for intelligent fault diagnosis of train bogie bearing[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24: 6053-6064.
- [13] Taherkhani N, Pierre S. Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17(11): 3275-3285.
- [14] Kong D, Wu F, Saroglou C. Automatic identification and characterization of discontinuities in rock masses from 3D point clouds[J]. *Engineering Geology*, 2020, 265: 105442.