

Recognition of damage pattern and evolution in CFRP cable with a novel bonding anchorage by acoustic emission

Jingyu Wu^{1,2,3}, Chengming Lan⁴, Guijun Xian^{1,2,3} and Hui Li^{*1,2,3,5}

¹Key Lab of Structures Dynamic Behavior and Control of the Ministry of Education, Harbin Institute of Technology, Harbin, 150090, China

²Key Lab of Intelligent Disaster Mitigation of the Ministry of Industry and Information Technology, Harbin Institute of Technology, Harbin, 150090, China

³School of Civil Engineering, Harbin Institute of Technology, Harbin, 150090, China

⁴School of Civil and Resource Engineering, University of Science & Technology Beijing, Beijing, 100083, China

⁵National Center for Materials Service Safety, University of Science & Technology Beijing, Beijing, 100083, China

(Received August 6, 2017, Revised January 1, 2018, Accepted March 5, 2018)

Abstract. Carbon fiber reinforced polymer (CFRP) cable has good mechanical properties and corrosion resistance. However, the anchorage of CFRP cable is a big issue due to the anisotropic property of CFRP material. In this article, a high-efficient bonding anchorage with novel configuration is developed for CFRP cables. The acoustic emission (AE) technique is employed to evaluate the performance of anchorage in the fatigue test and post-fatigue ultimate bearing capacity test. The obtained AE signals are analyzed by using a combination of unsupervised K-means clustering and supervised K-nearest neighbor classification (K-NN) for quantifying the performance of the anchorage and damage evolutions. An AE feature vector (including both frequency and energy characteristics of AE signal) for clustering analysis is proposed and the under-sampling approaches are employed to regress the influence of the imbalanced classes distribution in AE dataset for improving clustering quality. The results indicate that four classes exist in AE dataset, which correspond to the shear deformation of potting compound, matrix cracking, fiber-matrix debonding and fiber fracture in CFRP bars. The AE intensity released by the deformation of potting compound is very slight during the whole loading process and no obvious premature damage observed in CFRP bars aroused by anchorage effect at relative low stress level, indicating the anchorage configuration in this study is reliable.

Keywords: CFRP cable anchorage; acoustic emission; clustering-based pattern recognition; ultimate bearing capacity

1. Introduction

Cables are very critical components in many structures, such as cable-supported bridges, off-shore platforms, etc. However, the cables are vulnerable to fatigue and corrosion damage. CFRP materials has very good mechanical properties and environmental corrosion resistance, so they are able to be used for fabrication of cables, i.e., the cables are fabricated by CFRP rods, replacing high strength steel wires.

For the weak shear and radial compression of CFRP rods, comprehensive studies (Meier 2012, Noistering 2000, Rizzo and di Scalea 2001, Mei *et al.* 2015, Zhang *et al.* 2014) have been done on the anchorage of CFRP cables. In these studies, the bonding anchorages are the most widely used anchoring configuration for the small generated radial compression that usually causes local fiber bending in CFRP rod. In bonding anchorages, the potting compound, playing the role of transferring the load from CFRP rods to socket, is very critical to cable performance. Meier and Farshad (1996) proposed a kind of resin-based potting compound with gradient stiffness to optimize the bonding

stress distribution within anchoring zone. Zhang *et al.* (2014) selected reactive powder concrete (RPC) grout as potting compound and developed a CFRP ground anchor applied in the Aizhai Bridge in China. Noistering (2000) and Mei *et al.* (2015) improved potting compound injection techniques for the better contact between CFRP rods and potting compound.

In this paper, a novel bonding anchorage was proposed and the performances of CFRP cable using developed anchorage was tested. To understand the damages of CFRP (2014) have been done on the anchorage of CFRP cables. In these studies, the bonding anchorages are the most widely used anchoring configuration for the small generated radial compression that usually causes local fiber bending in CFRP rod. In bonding anchorages, the potting compound, playing the role of transferring the load from CFRP rods to socket, is very critical to cable performance. Meier and Farshad (1996) proposed a kind of resin-based potting compound with gradient stiffness to optimize the bonding stress distribution within anchoring zone. Zhang *et al.* (2014) selected reactive powder concrete (RPC) grout as potting compound and developed a CFRP ground anchor applied in the Aizhai cable during the loading process, acoustic emission (AE) technology which is a kind of nondestructive testing (NDT) method for monitoring the damages in material and structure is employed to detect the damages near the anchoring zone where is also the damage

*Corresponding author, Professor
E-mail: lihui@hit.edu.cn

concentration of CFRP cable. In AE testing, the stress wave released by the material deformation or damage is captured by AE sensor (usually made by lead zirconate titanate, PZT) and stored in waveform (AE signal) by AE instrument. The AE signal contains information about damage pattern and intensity which means the discrete AE signals in time-domain can be used to describe the damage evolution. Here, the damage evolution is referred to the variation of damage extent in a specific pattern during the loading process. In this article, the normalized accumulated AE energy release which is a comprehensive indicator of AE intensity and time-domain distribution is used to describe the damage evolution.

In AE signals analysis methods for the description of damage evolution, the regular parameters of AE signal, such as amplitude, energy, duration, rising time and AE counts are used to describe the damage evolution with single damage mode (Nair and Cai 2010). While, in multi-mode damage monitoring, it is needed to mapping the relationship between the damage patterns and AE signals. As the stress wave released by component material contains frequency that has direct relationship with the fracture mechanism, AE frequency analysis plays an important role in the damage pattern recognition and receives widely applications, e.g., the peak frequency in the spectrum of AE signal is successfully employed to identify the failure modes of FRP materials in the works by Groot *et al.* (1995), Ramirez-Jimenez *et al.* (2004), Godin *et al.* (2004), Gutkin *et al.* (2011).

In our study, the damages inside anchorage are complicated, including the matrix cracking, matrix-fiber debonding, fiber breakage in CFRP bar (de Groot *et al.* 1995) as well as the damage of potting compound resulted by shear-type deformation. The damage patterns of CFRP bar and potting compound may arouse AE signals with similar peak frequency, or the AE peak frequency released from different damage patterns may overlap for the similar component (epoxy resin) in CFPP bar and potting compound. Moreover, the acoustic attenuation effect on AE signal that both frequency and energy characteristics of obtained AE signal will be changed in acoustic propagation aggravate the difficulty of recognizing the damage pattern according to the pre-defined relationship between AE features and damage patterns which is established on the basis of the AE testing in a specific scale. Therefore, more advanced pattern recognition approach is needed.

Recent years, several clustering algorithms (Ding *et al.* 2004, Yang and Nagarajaiah 2013, Alrabea *et al.* 2013), e.g. centroid-based, density-based have been adopted to analyze AE signals. AE clustering analysis is a kind of unsupervised clustering analysis in pattern recognition field, which assumes that the AE signals released by the same type of AE source (damage pattern) have more similarity than from other pattern (Moëvus *et al.* 2008). In AE clustering analysis, AE signal is described by an AE feature vector usually consisting of regular AE parameters and peak frequency (Gutkin *et al.* 2011, Li *et al.* 2016), etc. Clustering algorithms categorize AE dataset (the collection of AE feature vectors) into a few of clusters to make the AE feature vectors have high similarity or correlation within a

cluster and low similarity or correlation between clusters. After that, these clusters have to be labeled with damage patterns according to the prior knowledge about the characteristics of expected damage patterns.

K-means clustering is one kind of centroid-based clustering algorithms and has been comprehensively studied (Ding *et al.* 2004, Xiong *et al.* 2009, Hu *et al.* 2015, Alrabea *et al.* 2013). K-means is the most widely used clustering algorithm for the partitioning of AE dataset in several different applications. Godin *et al.* (2004) combined the K-means and supervised classification learning algorithms including K-Nearest Neighbor (K-NN) and Self-Organizing Map (SOM) to improve the computational efficiency and it was founded the performance of K-means combined K-NN classifier and K-means combined SOM classifier were quite similar. Gutkin *et al.* (2011) compared the performances of three types of clustering-based pattern recognitions: K-means, K-means and SOM, competitive neural network (CNN) and founded the K-means combined SOM classifier can get the better result.

AE features selection in K-means is critical to clustering quality. It is required to reflect the difference among the AE signals in different damage pattern efficiently for the low accuracy of K-means in high-dimensional AE dataset. Therefore, the redundant AE features should be removed to keep the dimensionality of AE dataset as low as possible. Additionally, principal component analysis (PCA) (Abdi *et al.* 2010) is also used to process AE dataset for the dimensionality reduction, that is K-means clustering is performed in the principal component subspace (Ding *et al.* 2004). PCA-guided K-means has been employed to improve the clustering quality in several different applications. Moëvus *et al.* (2008) selected eight AE parameters which exhibits relatively large variance from 18 AE parameters in AE signal by using agglomerative hierarchical clustering to form AE feature vector and then reduced the dimensionality of AE dataset further by transforming to its principal component subspace. Gutkin *et al.* (2011) selected five AE features: AE amplitude, peak frequency, energy, rising time and duration for clustering and performed clustering analysis in its principal component subspace for the pattern recognition of CFRP material.

The further researches on K-means indicate the imbalanced classes distribution in dataset will affect clustering quality (Xiong *et al.* 2009, Wu 2012). The samples in the dataset can be categorized into majority classes (MA) and minority classes (MI) (Yen *et al.* 2009). The imbalanced dataset means that the number of samples in MA is much larger than the rest classes in MI. The influence of imbalanced dataset to K-means clustering quality is remarkable due to its uniformed size effect (Wu, 2012). This problem can be solved through the improved algorithm for handling imbalanced dataset (Hu *et al.* 2009) or dataset pre-processing approaches (Drummond *et al.* 2003 and Yen *et al.* 2009).

In this article, fatigue and post-fatigue ultimate bearing capacity tests were conducted to evaluate the performances of CFRP cable using developed anchorage and AE testing was adopted to detect the damages of CFRP cable during

loading process. A clustering-based pattern recognition is used to recognize the damage patterns in the post-fatigue ultimate bearing capacity test based on similarities among the AE signals.

2. The novel anchorage for CFRP cable

In most researches on CFRP cable anchorage, the CFRP rods inside anchorage are parallel-arranged. This kind of configuration requires spacing among the CFRP bars for the contact requirement between CFRP bars and potting compound and it inevitably enlarges cable cross-section area.

In this paper, a novel anchorage featured by the diverged arrangement of CFRP bars inside anchorage is developed as shown in Fig. 1. This configuration makes CFRP bars outside anchorage contact closely with each other and reduce cable cross-section area significantly. In the developed anchorage, the used potting compound is resin-based with the high bonding property. Besides, the diverged arrangement of CFRP bars makes a better contact between CFRP bars and potting compound.

These improvements in anchorage configuration is benefit to shorten the bonding length in anchorage. In this study, for achieving the minimalization of anchorage size, the bonding length in proposed anchorage is shortened as much as possible in anchorage design which means the reserved bonding capacity is limited. Therefore, the developed anchorage is much smaller than the current bonding anchorages for CFRP cables with the same or even less bearing capacity (Zhang *et al.* 2014, Mei *et al.* 2015).

In this way, the shortened bonding length in anchorage will expose potting compound to suffer the most serious damage when CFRP cable is loaded. Therefore, if the performance of potting compound in developed anchorage ensures the anchoring requirements in experimental tests, it is no doubt this potting compound is also available in the anchorage with longer bonding length.

3. Experiments

3.1 Fabrication of CFRP bar and cable

The CFRP bar with the nominal diameter of 4 mm was fabricated using pultrusion technique. The pultruded CFRP bar is composed of 19 bunches of Toray T700SC-12K carbon fiber tows with fiber volume fraction approximate 66%. The mechanical properties of Toray T700SC-12K carbon fibers and CFRP bar are listed in Table 1.

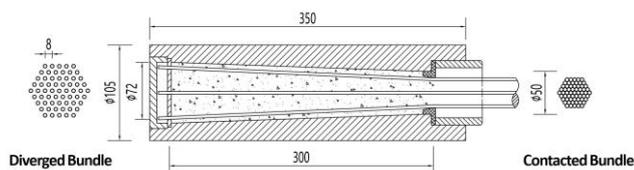


Fig. 1 The novel anchorage for CFRP cables

Table 1 Mechanical properties of Toray T700SC-12K carbon fiber and CFRP bar

	carbon fiber	CFRP bar
Tensile Strength (MPa)	4900	2587
E-Modulus (GPa)	230	130
Elongation (%)	1.8	1.6

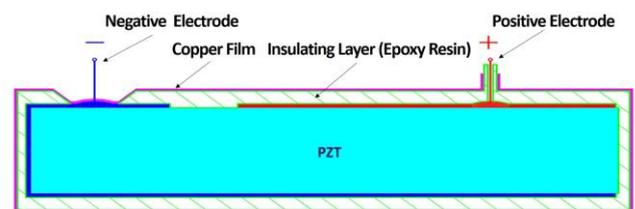


Fig. 2 EMI shielding of PZT patch

In manufacture of CFRP cables, the surface of CFRP bars within anchoring zone should be treated by polishing the resin-enriched layer to avoid the weak interface between CFRP bar and potting compound. A total of two CFRP cables ($\Phi 4-61$) with an effective length of 4 m were manufactured using the developed anchorage in Fig. 1 for the fatigue and post-fatigue ultimate bearing capacity tests.

3.2 AE monitoring system

A DiSP-4/PCI system by Physical Acoustic Corporation was employed to receive and store AE signals. The digital passband filters in DiSP-4/PCI system was set from 10 kHz to 2 MHz that filter friction signals (<10 kHz) and covered the frequency band of all the damage patterns. The preamplifier is set at a gain of 40 dB and the threshold of acquisition trigger was set 50 dB which could filter all the electromagnetic interference (EMI) signals in zero-load state. The sampling rate of AE acquisition system was 5 MHz.

A circular PZT patch with a diameter of 6 mm and a thickness of 1 mm for monitoring the damages. All PZT patches were coated by copper film using magnetron sputtering vacuum coating machine to reduce EMI effect as shown in Fig. 2.

3.3 Component-scale AE testing for damage pattern features

The potting compound and CFRP bar are component materials in bonding anchorage. In this section, the damage pattern features of potting compound and CFRP bars were experimentally obtained.

The potting compound in anchorage mainly suffers shear-type deformation, therefore, the potting compound coupon was loaded using double shear test. In double shear test, the potting compound cylinder ($\Phi 30 \times 90$ mm) was loaded in a speed of 0.5 mm/min and a PZT patch was attached to the lateral surface of the specimen as shown in Fig. 3(a). The CFRP bars in cable are subjected to tensile loading, so tensile loading test of CFRP bars is performed.

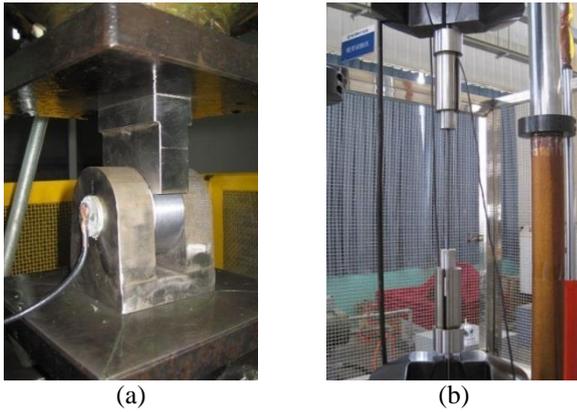


Fig. 3 Double shear test of potting compound coupon (a) and tensile test of a CFRP bar (b)

In the tensile test, the CFRP bar with a length of 150 mm and a diameter of 4 mm was loaded in a speed of 3 mm/min. The PZT patch was attached to the surface of the end of CFRP bar behind anchorage (no strain during tensile loading at this location), as shown in Fig. 3(b).

3.4 Cable-scale AE testing for the recognition of damage pattern and evolution

The fatigue test of CFRP cables was conducted following the PTI specification for the fatigue test of steel cable in cable-stayed bridge. The bending fatigue effect near the anchorage was considered by inserting shim plates between bearing plate and nut to create an angular deviation of 10^{-2} radian. The fatigue loads applied on the two CFRP cables were sinusoidal waveform with mean stress of 700 MPa and constant fatigue stress amplitudes of 200 MPa (cable#1) and 300 MPa (cable#2) by a hydraulic fatigue testing machine for 2 million cycles.

After the fatigue test, a post-fatigue ultimate bearing capacity test on cable#1 and #2 was performed. The two CFRP cables were tested using a horizontal reaction frame, as shown in Fig. 4. The cables were loaded stepwise with a loading interval of 400 MPa. Each loading step was divided into load-ascending stage and load-holding stage and the cable was loaded alternately until the broken, as shown in Fig. 5. In load-ascending stage, the hydraulic loading equipment was operated in load control mode with a speed of 2.5 kN/s, which made the duration of each load-ascending stage last approximate 120 s.

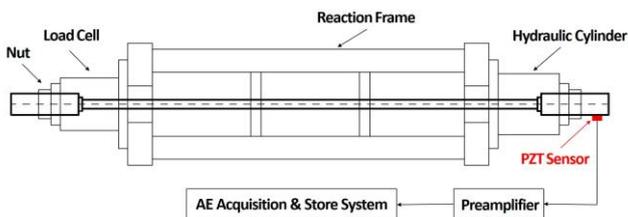


Fig. 4 Schematic of post-fatigue ultimate bearing capacity test and AE monitoring system

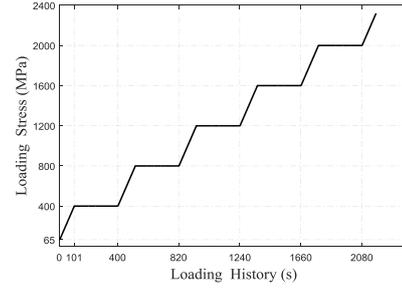


Fig. 5 Stepwise loading scheme

In load-holding stage, the static loading was held for 300 s in order that the damage evolutions at this stress level could be fully developed.

During the fatigue test and ultimate bearing capacity test of CFRP cable, the PZT patch was attached on the out-surface of anchorage as shown in Fig. 4. Therefore, the damages inside anchorage where is also the damage concentration of CFRP cable were detected and stored.

4. Clustering-based damage pattern recognition

Clustering-based pattern recognition is used to correlate AE signals with damage pattern. In this approach, AE feature extraction for clustering analysis, processing of imbalanced AE dataset, K-means combined K-NN for AE dataset partitioning and damage pattern labeling are the main sequential steps and they are elaborated in the following section.

4.1 AE feature for clustering

Wavelet packet transform (WPT) (Wang *et al.* 2006, Khamedi *et al.* 2010) is adopted to extract the frequencies and energy information from the nonstationary AE signal $f(t)$, as shown in Fig. 6.

Each wavelet packet component $f_j^i(t)$ at decomposition level j is expressed as

$$f_j^i(t) = \sum_{k=-\infty}^{+\infty} c_{j,k}^i \psi_{j,k}^i(t) \Delta t \quad (1)$$

where $C_{j,k}^i$ is the wavelet packet coefficients, $\psi_{j,k}^i$ is the wavelet packet by using Db4 mother wavelet. After j -th decomposition, the AE signal can be reconstructed by a summation of all the wavelet packet components as follows

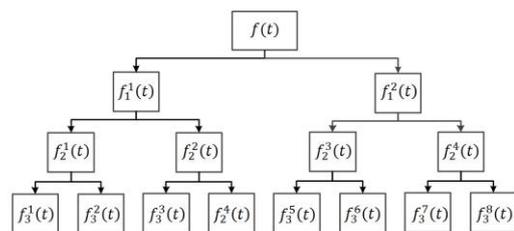


Fig. 6 Wavelet packet transform (WPT) binary tree

$$f(t) = \sum_{i=1}^{2^j} f_j^i(t) \quad (2)$$

The wavelet-based energy E_j^i with peak frequency ranging from $f_s \cdot 2^{j-1} \cdot (i-1)$ to $f_s \cdot 2^{j-1} \cdot i$, can be expressed as in logarithm form

$$\log(E_j^i) = \log\left(\sum f_j^i(t)^2\right) \quad (3)$$

where f_s is the sampling frequency of AE acquisition system.

It is noted that the frequency bandwidth decreases with the decomposition level in WPT. In this paper, the decomposition level of AE signal is determined by the distribution of AE peak frequency because WPT may change the frequency characteristics near the cut-off frequencies of $f_s \cdot 2^{j-1} \cdot i$, which means that the cut-off frequencies at the j -th decomposition level should be far away from the AE peak frequency concentration ranges.

4.2 Pre-processing of imbalanced AE dataset

This imbalanced classes distribution in AE dataset is resulted by the inherent characteristics of multi-modes damage evolution in material as well as the acoustic attenuation. In this study, under-sampling approach in MA is used to deal with the imbalanced AE dataset for its less disturbance to the original dataset compared with over-sampling (Drummond *et al.* 2003). In under-sampling approach, a subset of MA is extracted and then combined with MI to set up a training dataset. In training dataset, the size of MA subset plays an important role to the clustering quality which can be evaluated by MI's F-measure as follows

$$M_p = \frac{2 \times P \times R}{P + R} \quad (4)$$

where M_p is MI's F-measure for evaluating clustering quality, P is the precision rate and R is the recall rate defined as

$$P = \frac{T_p}{T_p + F_p} \quad R = \frac{T_p}{T_p + F_N} \quad (5)$$

where T_p is the size of MI samples which is correctly identified as MI sample in clustering result, F_p is the size of MA samples in the training dataset that is falsely identified as MI sample in clustering result and F_N is the size of MI samples that is falsely identified as MA sample in clustering result. In this paper, the size of MA subset in training dataset is parametric discussed to make MI's F-measure reach maximum which means the clustering analysis can recognize the MI samples with high accuracy.

Moreover, the MA subset selection approaches in training dataset also influence clustering quality to some extent. In this paper, random under-sampling and cluster-based under-sampling (Yen *et al.* 2009) are adopted to select MA subset in training dataset. In random under-sampling, the MA subset is randomly selected. In cluster-based under-sampling, the original dataset is first divided

into K clusters using regular clustering algorithm and the sizes of MA subset in each cluster are decided as follow

$$SS_{MA}^i = (m \times S_{MI}) \times \frac{S_{MA}^i / S_{MI}^i}{\sum_{i=1}^K S_{MA}^i / S_{MI}^i} \quad (6)$$

where SS_{MA}^i is the size of MA subset in cluster i , m is the size ratio of MA subset to MI in training dataset, S_{MA}^i and S_{MI}^i are the sizes of MA and MI in cluster i , and the S_{MI} is the size of total MI. After the size of MA subset in each cluster is calculated by Eq. (6), the MA samples in each cluster can be randomly or distance-based selected including cluster-based randomly selected approach, cluster-based most far and most near selected approaches which are decided according to the distance to MI samples in the same cluster (the details of the cluster-based approaches can be referred in the publication by Yen *et al.* 2009).

4.3 K-means combined K-NN for AE partitioning

K-means is used to partition the dataset by minimizing the objective function J_K as follows

$$J_K = \sum_{k=1}^K \sum_{\mathbf{x}_i \in C_k} (\mathbf{x}_i - \mathbf{m}_k)^2 \quad (7)$$

where \mathbf{X}_i is the feature vector belonging to cluster C_k , \mathbf{m}_k is the mean feature vector of cluster C_k and K is the cluster number.

The objective function J_K is dependent on the cluster centroids initialization which is usually obtained through many iterations to converge a smaller J_K . However, the searching for the initial cluster centroids is time consuming and may get a local optimal objective rather than the global optimal objective J_K . Recently, an enhanced hierarchical K-means approach which combines PCA and cluster centroids optimization is proposed (Alrabea *et al.* 2013) by calculating the partitioning position in principal component subspace as follows

$$M = \frac{\sum_{i=1}^n \sum_{j=1}^i D_j}{n} \quad (8)$$

where n is the total number of samples in dataset, D_j is the squared Euclidean distance between adjacent samples in principal component subspace, and M is the boundary of two clusters in the first principal component subspace. In this approach, K-means is carried out in a binary hierarchical framework and the optimal cluster number is determined by the empirical Davies-Bouldin criterion as follows

$$D_B = \frac{1}{K} \sum_{i=1}^K \max_{i \neq j} \left\{ \frac{d_i + d_j}{D_{i,j}} \right\} \quad (9)$$

where D_B is Davies-Bouldin value with the cluster number K , d_i is the cluster radius of cluster C_i , $D_{i,j}$ are the Euclidean distance between the centroids of clusters C_i and C_j .

The clustering of dataset is performed by combing Eqs. (8) to (9). First, the dataset is categorized into two clusters by using Eq. (8), then the dataset in each cluster can be further categorized into two clusters ($K=3$, one of two clusters is further categorized or $K=4$, both clusters are further categorized). For each cluster number K , the D_B is calculated. Therefore, the relationship between cluster number K and D_B is established and the optimal cluster number will make D_B reach minimum (Wang *et al.* 2006 and Li *et al.* 2016).

The clustering results with the optimal cluster number are further used as the classifier for the pattern identification of the samples by using supervised K-nearest neighbor classification (K-NN) (Godin *et al.* 2004), that is we randomly select a sample in AE dataset, and then determine to which cluster this sample belongs according to the distance of this sample to the nearest K samples in pattern classifiers.

4.4 Partition-based pattern labeling

After AE dataset partitioning, it is needed to label these partitioned classes of AE signals with the expected damage patterns (Moervis *et al.* 2008) based on the priori-knowledge about the potential damage patterns. In this study, the damage pattern labeling is based on the frequency and time sequence characteristics of the AE signals obtained from component-scale AE testing in section 3.3.

The procedure of clustering-based damage pattern recognition is shown in Fig. 7. The frequency and energy of wavelet components of AE signal are brought in AE feature vector by using WPT to form AE dataset. In AE dataset, the MA and MI are defined and under-sampling is adopted in MA to obtain the MA subset in training dataset. The size of MA subset in training dataset on clustering quality is parameter for discussion and the optimal training dataset that make the MI's F-measure reach maximum is determined. PCA-guided hierarchical K-means combined with Davies-Bouldin criterion partitions the optimal training dataset and the clustering results are treated as the pattern classifiers for the classification of MA samples by using supervised K-NN algorithm.

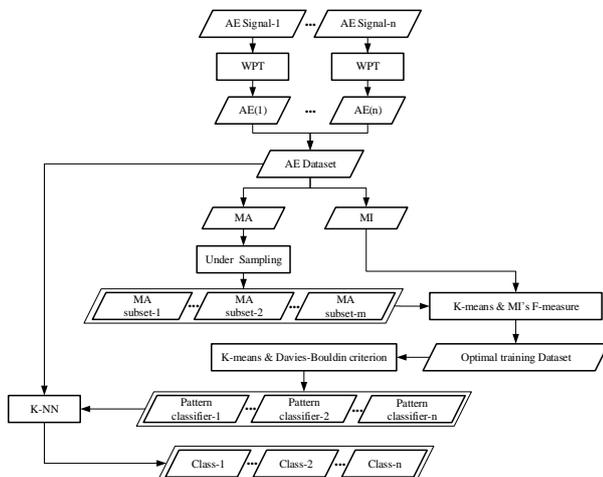


Fig. 7 Clustering-based damage pattern recognition

Finally, the partitioned AE signals are labeled with the expected damage patterns according to the prior knowledge about the AE peak frequencies and time sequence characteristics of the damage pattern.

5. Results and discussion

5.1 Damage pattern features of potting compound and CFRP bar

The failure mode of potting compound and CFRP bars are obtained through the double shear loading test and tensile loading test are shown in Figs. 8(a) and 8(b). It can be seen the potting compound fails in typical shear-type, while CFRP bar fails in tensile-type.

The peak frequency and energy of AE signals from potting compound coupon in double-shear test and CFRP bar in tensile test are shown in Figs. 9(a) and 9(b), respectively. It can be seen from Fig. 9(a) that the AE peak frequency distribution of potting compound is in the range of [50 150] kHz. The number of AE hits increases with loading and there exists a sudden high AE energy releases at the occurrence of brittle shear-type failure. Therefore, this damage process can be divided into the shear deformation of potting component (with low AE energy release) and the cracking coalescence along the interface between epoxy matrix and quartz granule inside potting compound (with high AE energy release). Thus, the damage status of potting compound can be measured by the accumulated AE energy release. For CFRP bars, Fig. 9(b) indicates that the frequency mainly is in the ranges of [70,150] kHz (blue points), [180, 300] kHz (red points), and [350, 500] kHz (black points), respectively, which correspond to the damage patterns of matrix cracking, matrix-fiber debonding and fiber breakage based on the conclusions from de Groot *et al.* (1995), Ramirez-Jimenez *et al.* (2004), Gutkin *et al.* (2011) and Li *et al.* (2016).

From the perspective of loading history, it can be seen from Fig. 9(a) that there is no correlation between AE peak frequency and loading history, which means that the shear deformation of potting compound occurs in the entire loading process. For CFRP bars, the AE peak frequency within the range of [70,150] kHz appears in whole loading process, implying that the matrix cracking also exists during the whole loading stage; while the fiber-matrix debonding

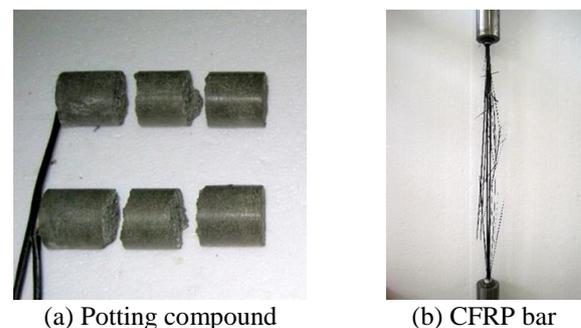
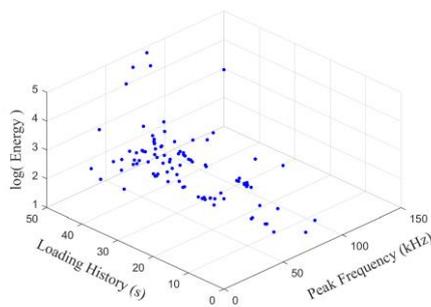
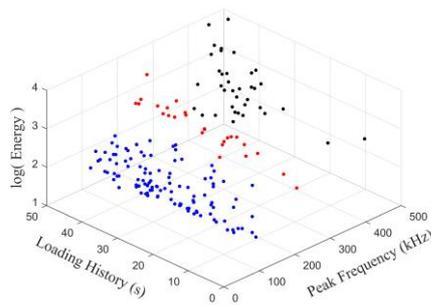


Fig. 8 Shear-type failure of potting compound coupon (a) and tensile-type failure of CFRP bar (b)



(a) Potting compound



(b) CFRP bar

Fig. 9 Peak frequency and energy of AE signals

([180, 300] kHz) mainly occurs at the mid and later loading stages; and fiber failure ([350, 500] kHz) occurs at the last loading stage.

Statistics on the peak frequency of AE signals is conducted to obtain the kernel probability density estimation of the AE peak frequency distributions of potting compound and CFRP bar, as shown in Fig. 10. It is clearly seen that there exists overlap in peak frequencies caused by shear deformation in potting compound (including cracking coalescence) and matrix cracking in CFRP bar. Statistics on the peak frequency of AE signals is conducted to obtain the kernel probability density estimation of the AE peak frequency distributions of potting compound and CFRP bar, as shown in Fig. 10. It is clearly seen that there exists overlap in peak frequencies caused by shear deformation in potting compound (including cracking coalescence) and matrix cracking in CFRP bar.

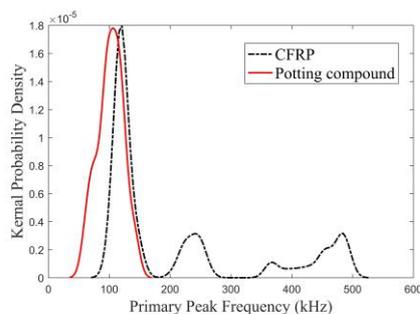


Fig. 10 Kernel probability density of the AE peak frequency of potting compound and CFRP bar

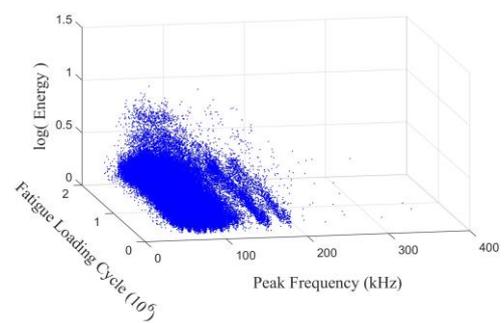
5.2 AE signal analysis during the fatigue loading test

Fig. 11 is the peak frequency and energy distribution of AE signals during fatigue tests. Most of AE signals' peak frequency are under 150 kHz, indicating that the damages of fiber-matrix debonding and fiber fracture in CFRP bars is quite small. In addition, it is worth noted that the distribution of AE hits and energy along time domain (with the peak frequency under 150 kHz) is homogeneously during the whole fatigue loading process, indicates these AE signals are released by the deformation of materials rather than damage.

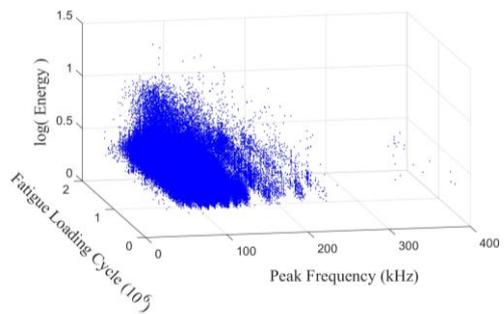
5.3 Damage pattern recognition using AE signal

In the post-fatigue ultimate bearing capacity test of CFRP cables, there is no sudden load drop before cable fracture. The relationship between cable elongations at different load-holding stages and tensile loading is shown in Fig. 12(a). As shown that CFRP cable behaves at linear stage. The nominal stress causing first CFRP bar breakage in cable#1 and #2 are 2362 MPa and 2283 MPa, respectively, which are little lower than the mean fracture strength of CFRP bar in Table 1 due to the non-uniform strength distribution in various CFRP bars. The breakage of CFRP bars locates at approximate 15~20 cm to the mouth of socket, as shown in Fig. 12(b), indicating the anchorage in this study is reliable.

The relationship between AE peak frequency and loading history of cables are shown in Fig. 13, which provides an overview on the evolution of multiple damage patterns inside the anchorage and the CFRP bars.

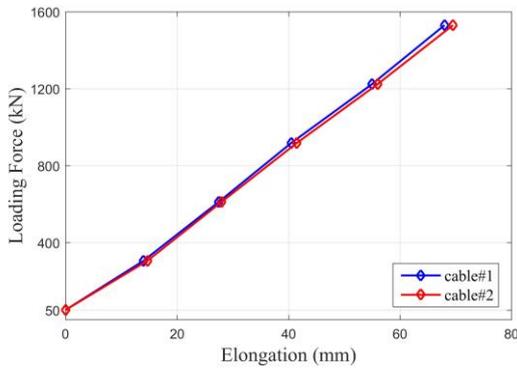


(a) cable#1



(b) cable#2

Fig. 11 Peak frequency and energy of AE signals vs. fatigue loading cycles



(a) CFRP cable's elongation



(b) Fracture mode

Fig. 12 CFRP cable's elongations at load-holding stages vs. loading forces (a) and fracture mode of CFRP cable (b)

As shown in Fig. 13, most of AE signals are released at the load-ascending stages due to the growth of strain. Comparison with the experimental results of damage pattern features of potting compound and CFRP bar as shown in Fig. 8, the AE signals in Fig. 13 with the peak frequency range of [170 300] kHz and above 300 kHz are released by fiber-matrix debonding and fiber fracture in CFRP bar. For the AE signals with peak frequency below 170 kHz, the peak frequency denotes matrix cracking in CFRP bar and shear deformation of potting compound, which implies that the two damage patterns release the same peak frequency and the overlap of peak frequency makes it be impossible to classify these AE signals based on its peak frequency distribution only.

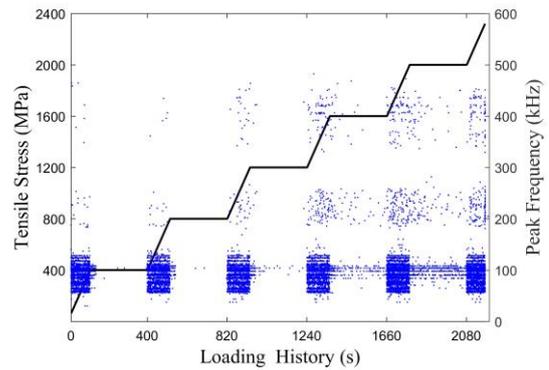
To further recognize the damage pattern by using the AE signals, the AE signal is decomposed by WPT to extract both frequency and frequency-related energy characteristics. It can be seen from Fig. 13 that there are almost no AE signals with peak frequency ranging from 270 to 330 kHz. Therefore, the AE signal is decomposed three levels with 8 wavelet packet components and the first and second cut-off frequencies are 312 kHz and 624 kHz, respectively. Fig. 14(a) is a representative AE signal. The wavelet packet components and their energy are calculated by using Eqs. (1) and (3) and shown in Figs. 14(b) and

14(c). The frequency spectrum of f_3^2 which contains the highest energy and $f(t)$ are calculated by FFT and shown in Fig. 14(d).

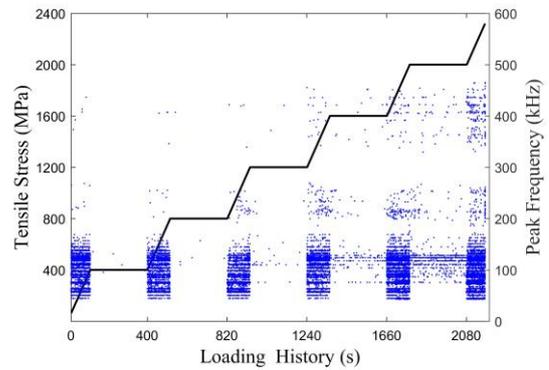
It can be seen from Fig. 14(d) that the peak frequency of f_3^2 (defined as primary peak frequency in this study and calculated by Fast Fourier Transform) is very close to the peak frequency of entire AE signal $f(t)$, which means that the primary peak frequency contains the same damage information as the peak frequency in $f(t)$. Therefore, the primary peak frequency and corresponding logarithmic energy are brought in AE features for clustering analysis. In addition, the logarithmic energies of f_3^1, f_3^2, f_3^3 for describing energy distribution in frequency domain are also included in AE feature vector for clustering analysis. Thus, the feature vector of the i -th AE signal in time sequence is written as follows

$$\mathbf{AE}_{(i)} = [P_{prim}, \log(E_{prim}), \log(E_3^1), \log(E_3^2), \log(E_3^3)] \quad (10)$$

Therefore, the AE dataset, which is a collection of AE feature vectors, is a $n \times 5$ matrix. Before performing PCA, the elements of each variable have to be pre-processed to make the variables centered (mean value equal to 0) and standardized to unit norm as below



(a) cable#1

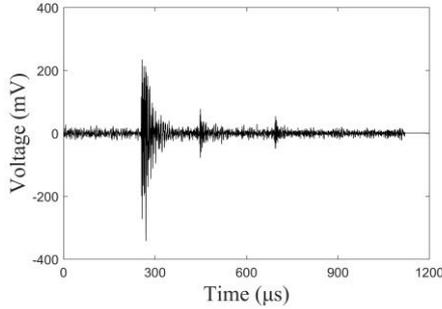


(b) cable#2

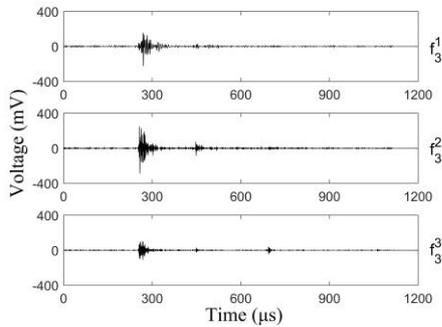
Fig. 13 AE peak frequency distribution vs. loading history of cables

$$\overline{x_j^i} = \frac{x_j^i - \mu}{\sqrt{\sum_1^n (x_j^i - \mu)^2}} \quad (11)$$

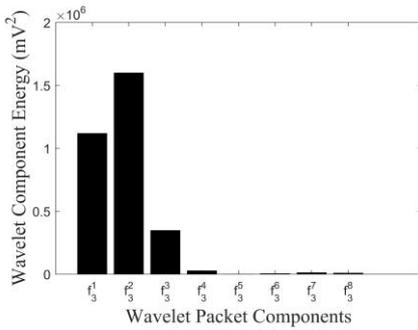
where x_j^i is the value of variable j in $\mathbf{AE}_{(i)}$, μ is the mean of variable j , n is the number of AE signals.



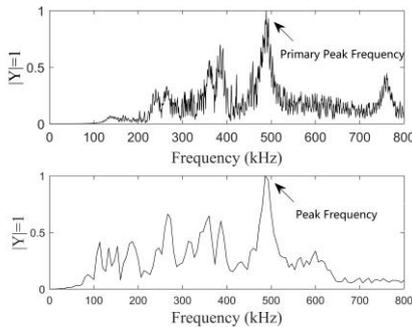
(a) Time-history of a representative AE signal



(b) Wavelet components f_3^1 , f_3^2 , f_3^3 with frequencies of [0, 312] kHz, [312, 624] kHz, and [624, 936] kHz



(c) Wavelet component energy distribution



(d) Comparison of primary peak frequency and peak frequency

Fig. 14 Wavelet packet-based feature of a representative AE signal

Table 2 The number of AE signal in various primary peak frequency ranges

Primary peak frequency (kHz)	G1 <170	G2 [170 300]	G2 >300
cable#1	51664	447	325
cable#2	56479	463	359

The principal components of the pro-processed AE feature dataset matrix are obtained by singular value decomposition (SVD) as follows

$$\overline{\mathbf{AE}} = \mathbf{USV}^T, \quad \mathbf{F} = \mathbf{US} \quad (12)$$

where $\overline{\mathbf{AE}}$ is pre-processed AE feature dataset, \mathbf{U} is the left singular vector, \mathbf{V} is right singular vector, \mathbf{S} is diagonal matrix of singular values, \mathbf{F} is the principal components.

The numbers of AE signals according to the primary peak frequency during entire loading process of cable#1 and #2 are listed in Table 2. It is observed that the amount of AE signals with primary peak frequency above 170 kHz only accounts for approximate 1.5% of total AE signals number, which indicates that there exist severe imbalanced-classes distributions in AE datasets and it will result in serious deviation between clustering results and true classes due to the uniformed size effect of K-means. Therefore, the AE datasets must be pre-processed to reduce the imbalanced-classes degree to ensure the clustering quality. the uniformed size effect of K-means. Therefore, the AE datasets must be pre-processed to reduce the imbalanced-classes degree to ensure the clustering quality.

In Table 2, G1 is categorized into MA samples, while G2 and G3 are grouped into MI samples. After performance PCA on $\overline{\mathbf{AE}}$, the first two principal components are plotted in Fig. 15. As shown, that the boundaries between G2 and G3 are distinct, while there exist slight overlaps near the boundaries between G1 and G2. The inherent structures of G1 are not clear and needed to be further investigated.

On the other hands, it is worth noting that the number of kept principal components for K-means clustering may affect clustering results to some extent. In this study, the number of kept principal components is determined according to the clustering quality measured by the MI's F-measure. It is founded that the clustering quality improves with the increase of kept principal components. But the clustering results are almost unchanged when the number of kept principal components is no less than 3. Therefore, the first four principal components which contains more than 90% of total variation in $\overline{\mathbf{AE}}$ are kept for clustering analysis.

Fig. 16 show the influences of the size and selection approach of MA subset on clustering quality. From Fig. 16, the variation tendency of MI's F-measure with the size of MA subset for cable#1 and cable#2 are quite different. This is attributed to the different overlap degree near the boundary between G1 and G2 as shown in Fig. 15. The overlap degree in cable#1 is lower than cable#2. The relatively clear boundary in cable#1 increases the robust against the influence of imbalanced-classes distribution, i.e. keeps K-means clustering with a relatively high accuracy in a wider range of the ratio of MA subset to MI than cable#2.

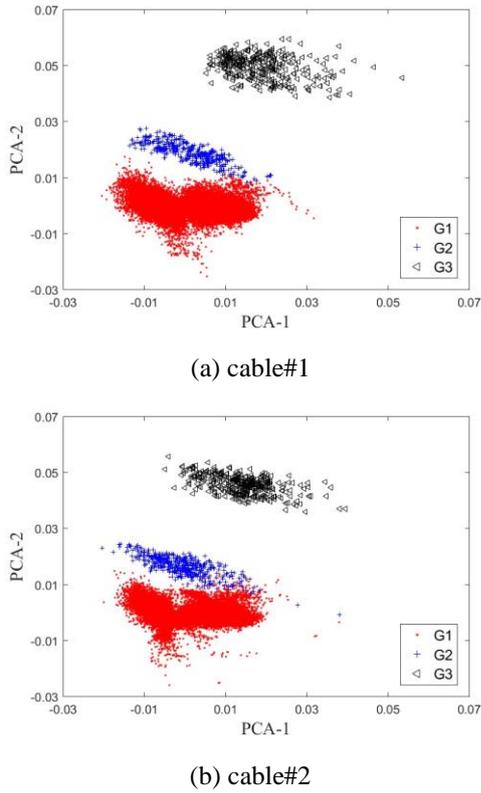


Fig. 15 The first two principal components of AE features

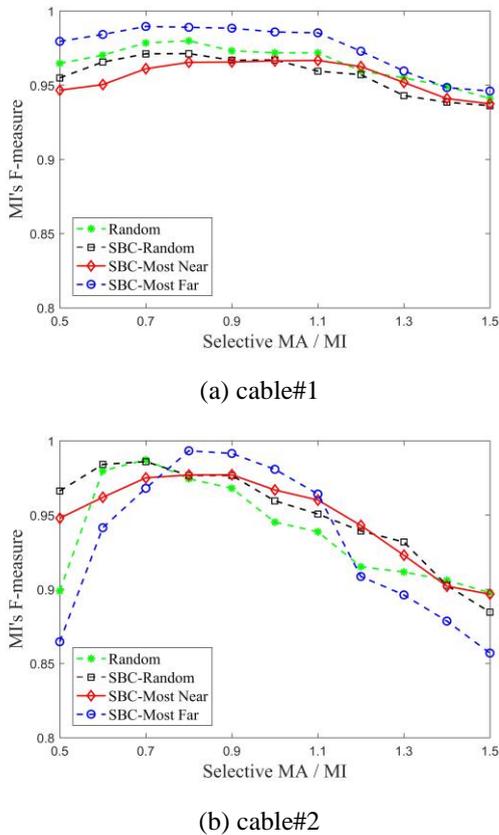


Fig. 16 The influences of the size and selection approach of MA subset on MI's F-measure

Once the optimal ratio m is determined, the MA subset samples by random and cluster-based selections (including cluster-based random, cluster-based most near and most far selections) are respectively combined with MI samples to form the optimal training datasets. Then, the hierarchal K-means is carried out in the principal component subspace of these optimal training datasets by using Eq. (8) and the optimal cluster number K is determined by the Davies-Bouldin criterion in Eq. (9).

The Davies-Bouldin values on the clustering results in different optimal training dataset with cluster number is shown in Fig. 17. The Davies-Bouldin values by random and SBC-random selections in Fig. 17 are the average that the process of setting up of optimal training dataset and clustering analysis are repeated for 10 times. The optimal cluster numbers in all optimal training datasets are 4. The cluster results are used as pattern classifiers for the pattern identification of the rest MA samples (the MA samples not in the optimal training dataset) by using the supervised K-NN classification, here $K=3$ in this study. After the further supervised classification of the rest MA samples, there are no any rest MA samples falling into MI.

The classification results of the rest MA dataset based on the pattern classifiers (the number of pattern is 2 according to clustering results in Fig. 17) are also evaluated by Davies-Bouldin criterion and shown in Fig. 18. It is concluded from Fig. 18 that the random selection approach achieves the best separation in MA dataset and the corresponding classifications of MA samples are shown in Fig. 19.

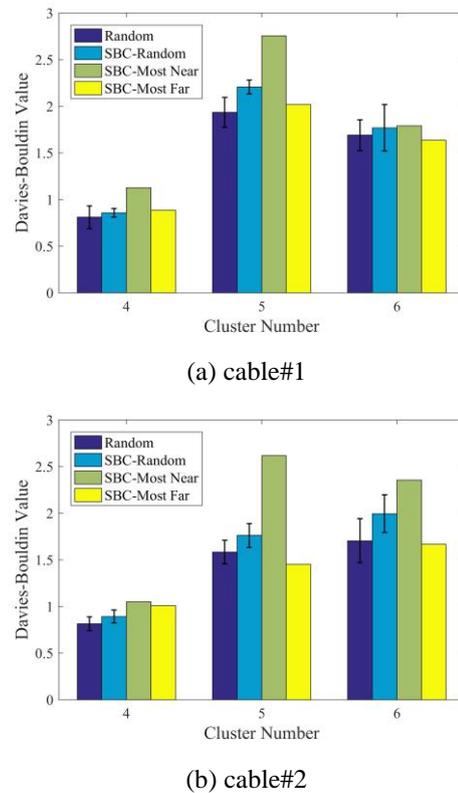


Fig. 17 Cluster number vs. Davies-Bouldin value in training datasets

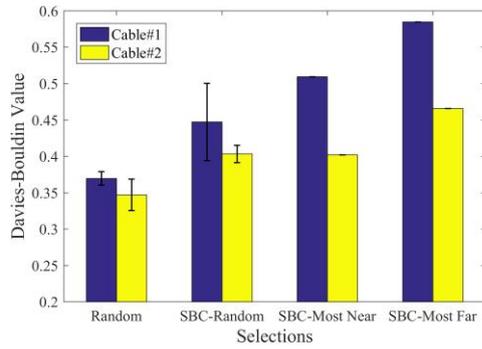
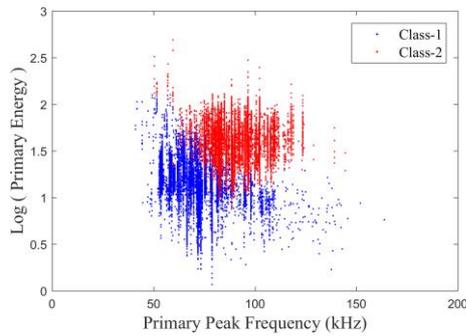
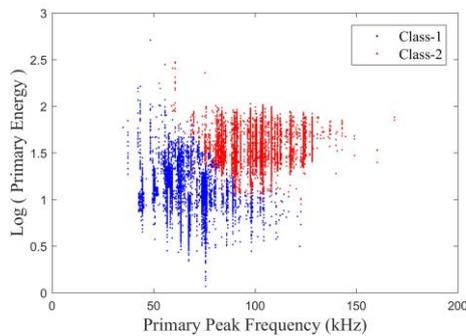


Fig. 18 Davies-Bouldin criterion on the classifications of the rest MA samples



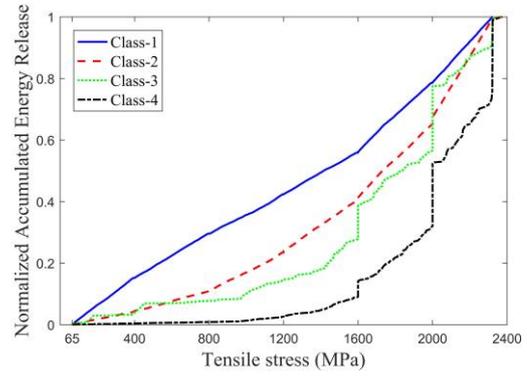
(a) cable#1



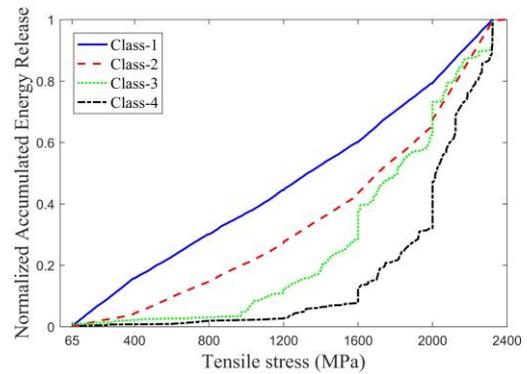
(b) cable#2

Fig. 19 Classification of MA samples based on the pattern classifiers formed by random selection

The multi-mode damage evolution inside anchorage can be measured by the normalized accumulated AE energy releases as shown in Fig. 20. According to the obtained frequency and time sequence characteristics of the damage pattern in section 5.1, the Class-1 is regarded to be released by the shear deformation in potting compound for its linear distribution in all load-ascending stages. Class-2 is release by matrix cracking in CFRP bars. Class-3 and Class-4 are release by the AE signals with primary peak frequency in the ranges of [170 300] kHz and above 300 kHz, respectively. Therefore, they are labeled as matrix-fiber debonding and fiber fracture in CFRP bar.



(a) cable#1



(b) cable#2

Fig. 20 Multi-mode normalized accumulated energy release vs. tensile loading

From Fig. 20, the normalized AE energy releases by damage patterns of fiber-matrix debonding and fiber fracture in CFRP bars which are closely related with material failure are only 0.145 and 0.026 in cable#1, 0.125 and 0.028 in cable#2 at the stress level of 1200 MPa. There is almost no damage at the load-holding stages of 1200 MPa. Thus, it is reasonable to make the conclusion that CFRP bar is reliable under 1200 MPa. For the potting compound, the AE energy release is almost linear distribution and energy release in all the load-holding stages are quite small which means the damage status in potting compound is far from failure in the whole loading process.

6. Conclusions

A novel anchorage for CFRP cable was developed and its performances was investigated through fatigue and post-fatigue ultimate bearing tests. Acoustic emission (AE) technique is employed to monitor the damages inside anchorage. Following conclusions are obtained:

- The double-shear test of potting compound and tensile loading test of CFRP bar for obtaining the frequency feature in AE signals show the existence of peak frequency overlapping which make it difficult to discriminate AE signals only through frequency features with peak frequency in the frequency overlapping range. Therefore,

the primary peak frequency and the energies of different wavelet components are extracted by WPT as AE feature vector in unsupervised K-means clustering analysis to build up damage pattern classifiers for the pattern identification of the AE signals.

- In fatigue test, the frequency and energy distribution of AE signals during the loading process is homogeneous and only a small amount of AE signals are released by fiber-matrix debonding and fiber fracture. Thus, it is concluded that the fatigue damage inside anchorage is quite small.
- The PCA-guided hierarchical K-means clustering combined with K-NN is employed to partition the AE signals during the post-fatigue ultimate bearing capacity test. The partitioning results indicate that there are four classes in AE dataset. And they are labeled with the expected damage patterns according to the peak frequency and time sequence characteristics of the damage pattern features of potting compound and CFRP bar.
- The AE testing of the damages inside anchorage during the post-fatigue ultimate bearing capacity test reveals: the multi-mode damage evolution in CFRP bars is in accordance with the mature failure process of CFRP composite, which indicates that the anchorage damage to CFRP bars is small to avoid premature damage at relatively low stress level; For potting compound, there is no obvious damage occurs in potting compound during the whole loading process for the linear increase of the cumulated AE energy released by potting compound which means the potting compound is reliable in this novel anchorage.

Acknowledgments

This study is financially supported by the NSFC (Grant No. 51478039) and the Beijing Nova Program (Grant No. Z15110000315053).

References

- Abdi, H. and Williams, L.J. (2010), "Principal component analysis", Wiley Interdisciplinary Reviews: Computational Statistics, **2**(4), 433-459.
- Alrabea, A., Senthilkumar, A.V., Al-Shalabi, H. and Bader, A. (2013), "Enhancing K-means algorithm with initial cluster centers derived from data partitioning along the data axis with PCA", *J. Adv. Comput. Netw.*, **1**(2), 137-142.
- de Groot, P.J., Wijnen, P.A.M. and Jansen, R.B.F. (1995), "Real-time frequency determination of acoustic emission for different fracture mechanisms in carbon/epoxy composites", *Compos. Sci. Technol.*, **55**(4), 405-412.
- Ding, C. and He, X.F. (2004), "K-means clustering via principal component analysis", *Proceedings of the International Conference on Machine Learning*, Banff, Alberta, Canada.
- Drummond, C. and Holte, R.C. (2003), "C4.5, class imbalance, and cost sensitivity: Why under-sampling beats over-sampling", Workshop on learning from imbalanced datasets. Washington DC: Citeseer.
- Godin, N., Hugué, S., Gaertner, R. and Salmon, L. (2004), "Clustering of acoustic emission signals collected during tensile tests on unidirectional glass/polyester composite using supervised and unsupervised classifiers", *Ndt & E Int.*, **37**(4), 253-264.
- Gutkin, R., Green, C.J., Vangrattanachai, S., Pinho, S.T., Robinson, P. and Curtis, P.T. (2011), "On acoustic emission for failure investigation in CFRP: Pattern recognition and peak frequency analyses", *Mech. Syst. Signal. Pr.*, **25**(4), 1393-1407.
- Hu, Y., Guo, D.F., Fan, Z.W., Dong, C., Huang, Q.H., Xie, S.K., Liu, G.F., Tan, J., Li, B.P. and Xie, Q.W. (2015), "An improved algorithm for imbalanced data and small sample size classification", *J. Data Anal. Inform. Process.*, **3**(3), 27-33.
- Khamedi, R., Fallahi, A. and Oskouei, A.R. (2010), "Effect of martensite phase volume fraction on acoustic emission signals using wavelet packet analysis during tensile loading of dual phase steels", *Mater. Design*, **31**(6), 2752-2759.
- Li, D.S., Hu, Q., Ou, J.P. and Li, H. (2011), "Fatigue damage characterization of carbon fiber reinforced polymer bridge cables: Wavelet transform analysis for clustering acoustic emission data", *Sci. China. Technol. Sci.*, **54**(2), 379-387.
- Li, L., Swolfs, Y., Straumit, L., Xiong, Y. and Lomov, S.V. (2016), "Cluster analysis of acoustic emission signals for 2D and 3D woven carbon fiber/epoxy composites", *J. Compos. Mater.*, **50**(14), 1921-1935.
- Mei, K.H., Li, Y.J. and Lu, Z.T. (2015), "Application study on the first cable-stayed bridge with CFRP cables in China", *J. Traffic Transportation Eng.*, **2**(4), 242-248.
- Meier, U. (2012), "Carbon fiber reinforced polymer cables: Why? Why Not? What If?", *Arab. J. Sci. Eng.*, **37**(2), 399-411.
- Meier, U. and Farshad, M. (1996), "Connecting high-performance carbon-fiber-reinforced polymer cables of suspension and cable-stayed bridges through the use of gradient materials", *J. Comput-Aided. Mater. Design.*, **3**(1-3), 379-384.
- Moevus, M., Godin, N., R'Mili, M., Roubly, D., Reynaud, P., Fantozzi, G. and Farizy, G. (2008), "Analysis of damage mechanisms and associated acoustic emission in two SiCf/[Si-B-C] composites exhibiting different tensile behaviours. Part II: Unsupervised acoustic emission data clustering", *Compos. Sci. Technol.*, **68**(6), 1258-1265.
- Nair, A. and Cai, C.S. (2010), "Acoustic emission monitoring of bridges: Review and case studies", *Eng. Struct.*, **32**(6), 1704-1714.
- Noistering, J.F. (2000), "Carbon fibre composites as stay cables for bridges", *Appl. Compos. Mater.*, **7**(2-3), 139-150.
- Ramirez-Jimenez, C.R., Papadakis, N., Reynolds, N., Gan, T.H., Purnell, P. and Pharaoh, M. (2004), "Identification of failure modes in glass/polypropylene composites by means of the primary frequency content of the acoustic emission event", *Compos. Sci. Technol.*, **64**(12), 1819-1827.
- Rizzo, P. and di Scalea, L.F. (2001), "Acoustic emission monitoring of CFRP cables for cable-stayed bridges, 6th Annual International Symposium on NDE for Health Monitoring and Diagnostics", International Society for Optics and Photonics. 129-138.
- Wang, G., Wang, Z.Z., Chen, W.T. and Zhuang, J. (2006), "Classification of surface EMG signals using optimal wavelet packet method based on Davies-Bouldin criterion", *Med. Biol. Eng. Comput.*, **44**(10), 865-872.
- Wu, J. (2012), *Advances in K-means Clustering A Data Mining Thinking*, Springer Berlin Heidelberg, Germany.
- Xiong, H., Wu, J. and Chen, J. (2009), "K-means clustering versus validation measures: A data-distribution perspective", *IEEE. T. Syst. Man. CY. B.*, **39**(2), 318-331.
- Yang, Y.C. and Nagarajaiah, S. (2013), "Output-only modal identification with limited sensors using sparse component analysis", *J. Sound Vib.*, **332**(19), 4741-4765.
- Yen, S.J. and Lee, Y.S. (2009), "Cluster-based under-sampling approaches for imbalanced data distributions", *Expert. Syst. Appl.*, **36**(3), 5718-5727.
- Zhang, K.Y., Fang, Z., Nanni, A. and Chen, G.P. (2014), "Experimental study of a large-scale ground anchor system with

FRP tendon and RPC grout medium”, *J. Compos. Constr.*,
19(4), 04014073.

BS