Fault diagnosis of rotating machinery equipped with multiple sensors using space-time fragments

Xunshi Yan [a,b,c,*], Zhe Sun [a,b,c], Jingjing Zhao [a,b,c], Zhengang Shi [a,b,c], Chen-an Zhang [d]

[a] *Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 100084, China*
[b] *The Key Laboratory of Advanced Reactor Engineering and Safety, Ministry of Education, Beijing 100084, China*
[c] *Collaborative Innovation Center of Advanced Nuclear Energy Technology, Beijing 100084, China*
[d] *State Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China*

ARTICLE INFO

ABSTRACT

The vibration signals captured by multiple sensors can be fused and provide rich information to distinguish faults of rotating machinery. However, previous studies mostly regard multiple signals as individual signals and ignore the coupling relationship between signals resulting in a loss of information. To overcome the above problem, this paper proposes a new multi-sensor data fusion algorithm for identifying faults. First, space-time fragments are constructed to combine multiple signals together considering the space and time relationship between signals. Second, histograms of multi-channel shaft orbit based on space-time fragments are extracted to describe faults. Third, K-nearest neighbor is selected as the classification method. The experiments are carried out on a rig of rotating machinery supported by active magnetic bearings and demonstrate the effectiveness of our proposed algorithm.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Rotating machinery plays an irreplaceable role in modern industry. With the rapid development of technology, rotating machinery is becoming larger, more precise and more automatic. At the same time, the safety of rotating machinery is gaining increasing attention. In particular, the failure of large-scale high-speed rotating machinery will result in serious economic losses, and even casualties. Hence, intelligent fault diagnosis of rotating machinery has become a research hot topic in recent years.

In rotating machinery, vibration sensors are arranged along the rotor and used to collect vibration signals which are important for fault diagnosis. Generally, the vibration signal is regarded as a time series, and the faults are viewed as abnormal events occurring in the time series. The analysis and processing of vibration signals are the main strategies to diagnose faults, primarily including time domain methods, frequency domain methods, and time-frequency domain methods. A brief overview is given below. We refer the interested reader to Refs. [1,2] and references therein for a comprehensive summary.

Basic transformations of signals have been used as major methods of describing faults in the early research, such as kurtosis [3], Fourier transform, autoregressive [4], Walsh transform [5] and power spectral density [6], most of which belong to time or frequency domain analysis and have a clear physical meaning in the field of signal processing. Many researchers have

---

combined multiple of the above basic features for a comprehensive description of faults [7]. Since the vibration signal is generally nonstationary, the effect of traditional frequency domain methods such as the Fourier transform method is limited, and the time-frequency domain method is introduced to the research, such as short-time Fourier transform [8], wavelet transform [9] and empirical mode decomposition [10], which integrate the time and frequency information in the description of faults.

In recent years, with the advancement of machine learning, the technologies of rotating machinery fault diagnosis have been greatly promoted. Locally linear embedding [11], sparse representation [12], autoencoder [13] are utilized to project the vibration signals into a nonlinear space for processing, gaining more distinctive feature representation. Deep learning models such as convolutional neural network (CNN) and recurrent neural network (RNN) improve the accuracy of recognition without the need for manually extracting features. In Ref. [14], the wavelet coefficients of vibration signals are composed of 2-D images which are as the input of CNN. According to Ref. [15], a new fully connected layer is added to CNN to solve the fault data imbalance problem. The training procedure of CNN is changed in Ref. [16] by introducing the dropout and small patches under noisy conditions and different working loads. In Ref. [17], angle synchronous averaging is performed on the training data which improved the effect of CNN. RNN with an autoencoder is used in Ref. [18] to deal with multiple time sequence vibration data.

Most previous studies have focused on analysis of the vibration signal captured by single sensor. However, in engineering applications, a series of vibration sensors are arranged at different positions along the rotor, such as in the position of the shaft, gearbox, bearings, etc. Particularly in certain rotating machinery with a long rotor, a number of vibration sensors are intentionally arranged at certain intervals to monitor the running state of the system. Using vibration signals captured by multiple sensors to diagnose faults, it can obtain more comprehensive fault information, avoid the shortcomings of limited information captured by single sensor, and acquire complementary information of sensors at different positions to achieve better recognition results. According to reference [19], the research on the fusion of multi-sensor data is carried out at three levels: the data level, feature level and decision level.

At the data level, the original signals are directly combined, and features are extracted from the combined data, which requires that the measured physical quantities are the same type [20,21]. This kind of fusion algorithms are simple, but often breaks the relationship in time dimension between original signals. At the feature level, features are extracted from original signals and a comprehensive expression is formed by the optimization method, which is used as the input of the classifier or decision layer [22]. The form of feature fusion usually requires an ingenious design to synthesize multiple channels of data sources and describe the faults. At the decision level, the fusion of vibration signals proceeds via the classification algorithm, often combining results of multiple classifiers to obtain better results. The decision level fusion algorithms [23–26] are widely studied, and many other methods of multi-source data fusion in other fields [27] can be introduced to fault diagnosis.

The above fusion algorithms treat channels of vibration signals as independent data sources, simply put the data together in form or fuse channels of signals mathematically, without considering the coupling relationship between signals and causing information loss. Generally, the vibration signals captured by multiple sensors contain information in time domain, and also have a correlation in space domain. When a fault occurs, sensors can obtain fault vibration signals with different amplitudes and phases, due to the difference in distance and angle between sensors and the position of fault occurred. For instance, when two vibration sensors are placed orthogonally and perpendicular to the axis of rotor, there is an amplitude correlation and a fixed phase difference between the two channels of vibration signals. Therefore, there exists coupling relationship between multi-channel vibration signals, and the coupling relationship has some characteristics as follows.

First, the combination of vibration amplitude from multi-channel signals at the same time should be deemed as the running state of rotating machinery at that moment. It will lead to information loss if the combination is considered separately. Second, based on the first point, there exist a time relationship between the vibration amplitude combinations of rotating machinery at two adjacent moments. Without considering this relationship, information will also be lost. Third, the response of the fault lasts for a period of time generally. One channel of vibration signal at a certain moment and another channel of vibration signal at other moment may be related due to the existence of a certain fault. Therefore, multi-channel vibration signals during a period of time need to be considered as a whole.

The above three points describe characteristics of the coupling relationship between vibration signals and are usually ignored in multi-sensor data fusion. To comprehensively describe faults, this paper proposes a new perspective named space-time fragments in the data level fusion that effectively organizes the signals captured by multiple sensors, and designs a histogram feature based on space-time fragments, which integrates the coupling relationship between the signals.

The structure of this paper is organized as follows. Section 2 introduces the concept of space-time fragments. Section 3 details the histogram representation of faults based on space-time fragments. The classification method and flow of algorithm are presented in Section 4. The experiments are described in Section 5 to verify the effectiveness of the proposed algorithm. The conclusions and outlooks are summarized in Section 6.

2. Space-time fragments

Generally, multiple sensors are arranged at different positions in rotating machinery and provide more information for monitoring and diagnosing the system. However, when dealing with multi-sensor information fusion problems, traditional methods often treat vibration signals as time series, processing them separately in time domain, frequency domain or time-frequency domain, and then fusing the results through a classification algorithm, which ignores the coupling relationship of multi-channel
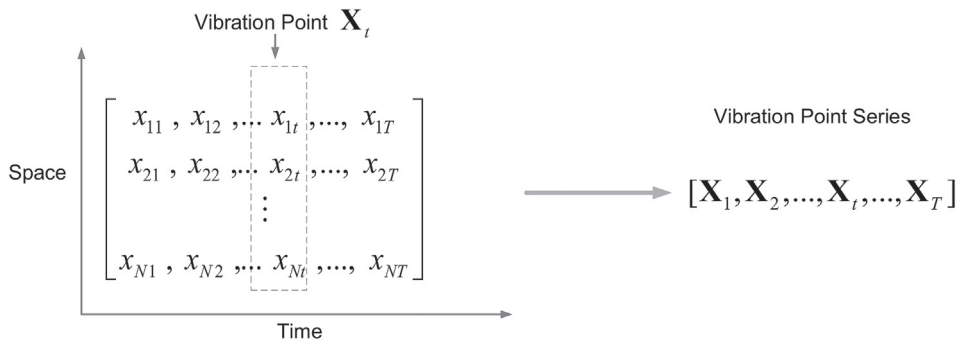
Fig. 1. The description of space-time fragments.

signals. If the coupling relationship of multi-channel signals can be mined, the performance of diagnosis task will be improved.

Inspired by space-time points in computer vision [28–30], we propose a novel idea of preserving the coupling information between multi-channel signals, and design a new fusion strategy called space-time fragments at the data level.

Fig. 1 shows the form of a space-time fragment. Each sensor captures one channel of the vibration signal and a space-time fragment is created according to all channels of signals in a period of time. Intuitively, the space-time fragment is a data matrix. Each row of the space-time fragment is a channel of signal and represents the time information. The column denotes different sensors and consists the space information of the sensor layout. One column is the vibration amplitude combination of signals acquired at different positions, defined as vibration point (VP). The VP is a physical quantity that characterizes the current running state of the rotating machinery at one moment. Substantially, unlike the traditional data-level fusion methods that connect signals directly, a space-time fragment is a time series of VPs along the time axis, which remains the timing relationship of the original signals and the space layout relationship of the sensors.

As a type of data-level fusion method, the space-time fragment meets the requirements of the three characteristics of the aforementioned coupling relationship information. Its basic unit is VP, which preserves the space information of different sensors at the same time. The space-time fragment also considers the time relationship of VPs so that the time information of the sensors at different moments is fully utilized. The entire matrix consisting of multi-channel signals in the space-time fragment is processed overall; thus, the relationship between any two sensors at different moments is included. Hence, compared to previous data-level fusion methods, space-time fragments mine coupling information of multi-channel signals comprehensively, and prepare organized data for feature selection without information loss.

3. Histogram of multi-channel shaft orbit

The space-time fragment is a data-level fusion method that organizes the original signals and reflects the coupling relationship between multi-channel vibration signals but is not a feature vector and cannot be directly used as input to the classification methods. It is necessary to extract features to remove redundant information, while preserving the main characteristics of the space-time fragment. In this section, a new variation based on space-time fragments that generalizes the two-dimensional shaft orbit to a high-dimensional space, is represented by a histogram, which describes the faults of rotating machinery and has sufficient ability to distinguish different classes of faults. The details are as follows.

3.1. Shaft orbit

The shaft orbit is a fault description method for multi-sensor data fusion at the feature level and takes into account the space-time coupling relationship [31]. The description involves vibration signals captured by two sensors placed orthogonally in a radial plane. In the normal state, the shape of the shaft orbit is circular or approximately circular. When an abnormality or failure occurs, the shaft orbit is distorted into other shapes, corresponding to different classes of faults. Intelligent methods can classify faults by recognizing the shape of the shaft orbit.

Fig. 2 shows the formation of shaft orbit. The vibration sensors $S1$ and $S2$ are placed orthogonally in a certain radial plane of the rotor of the rotating machinery and collect two channels of vibration signals. The vibration amplitudes of signals at the same time constitute vibration points in the x_1-x_2 coordinate plane, which are connected to the shaft orbit. In essence, the shaft orbit is an expression of a space-time fragment given two sensors, reflecting the space-time coupling relationship between two channels of signals.

The shaft orbit has an intuitive physical meaning for describing faults of rotating machinery [5]. However, the shaft orbit can only describe two channels of signals captured by sensors placed orthogonally in the same radial plane, and does not work if the number of sensors is larger than 2. The shaft orbit cannot be directly utilized as a kind of feature and requires discriminative features to be described further.
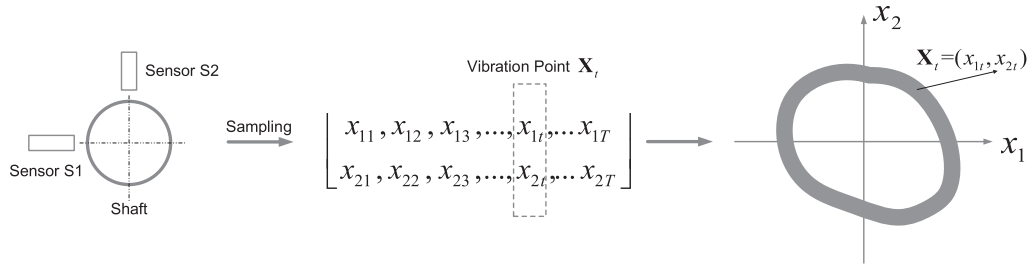
Fig. 2. The formation of a shaft orbit.

3.2. Multi-channel shaft orbit

Based on the space-time fragment, we extend the concept of the shaft orbit to the high-dimensional space, that is, the case of two orthogonally installed sensors is extended to that of multiple sensors. Multi-channel shaft orbit (MCSO) is constructed to mine the coupling relationship between vibration signals.

Suppose N vibration sensors are arranged in rotating machinery. In a period of time, N channels of vibration signals are sampled, and the space-time fragment **stf** is constructed as:

$$\mathbf{stf} = [\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_N]^T$$
$$= [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_t, \ldots, \mathbf{X}_T] \tag{1}$$

where T is the number of sampling points in the time interval, $\mathbf{S}_i = [x_{i1}, x_{i2}, \ldots, x_{iT}]$ represents the ith channel of signal and $\mathbf{X}_t = (x_{1t}, x_{2t}, \ldots, x_{Nt})$ denotes the VP at sampling point t.

An N-dimensional space is built and at a certain moment, VP $\mathbf{X}_t$ is thought as a coordinate point in the N-dimensional space. x_{ij} is the projection component of VP $\mathbf{X}_t$ on the x_i axis, denoting vibration amplitude of the signal captured by the ith sensor at the jth sampling point in the space-time fragment. The VPs are connected together and form a curve in N-dimensional space, which is defined as multi-channel shaft orbit. Similar to the shaft orbit, the shape is an important characteristic of MCSO, which is essentially the relative position distribution of VPs.

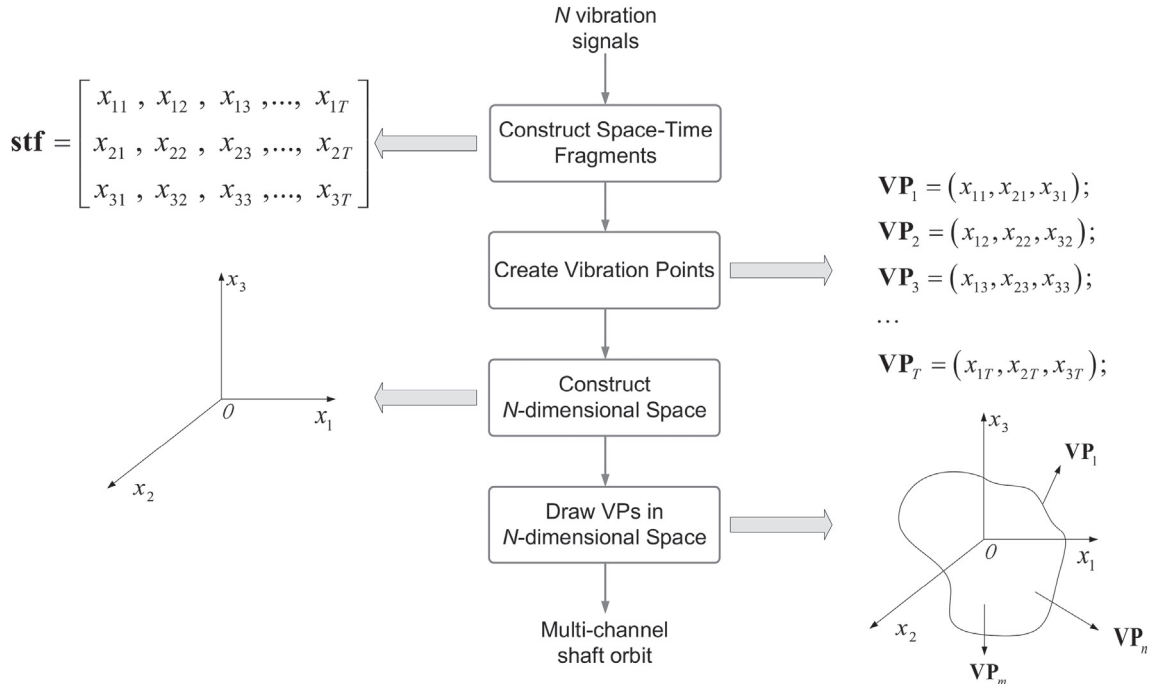Fig. 3 shows the flowchart of formation of MCSO, and a concrete example with $N = 3$ is given.

Fig. 3. The formation of an MCSO and an example when $N = 3$.

As the transition between data-level and feature-level methods for multi-sensor data fusion, MCSO is a concrete manifestation of a space-time fragment in a high-dimensional space and transforms a data matrix into a curve and thus retains all characteristics of the coupling relationship between multi-channel signals. The MCSO description is also an extension of traditional shaft orbit in the case of multiple sensors, and its shape is an important representation of rotating machinery faults. Similar to shaft orbit, MCSO is not a feature vector, and the shape of MCSO cannot be visually recognized. Compared to the space-time fragment, MCSO is much closer to the description of faults but it is necessary to extract distinguishing features further.

3.3. Histogram of multi-channel shaft orbit

MCSO is a variation of the space-time fragment, and its geometric shape represents characteristics of the faults. The features of MCSO should express its geometric information. Therefore, we design histogram of multi-channel shaft orbit (HMCSO) to distinguish faults of rotating machinery.

We define the angle between VP $\mathbf{X}_t$ and axis x_i as follows:

$$A_{X_t,x_i} = \arccos \frac{x_{it}}{d_t} \tag{2}$$

where d_t is the distance between $\mathbf{X}_t$ and the original point.

According to the angle between VPs in space and axis x_i, the space is divided into M subspaces on average, and $R_{x_i}(j)$ represents the jth subspace, which satisfies:

$$R_{x_i}(j) = \{P \mid (j-1)\frac{\pi}{M} \leq A_{P,x_i} < j\frac{\pi}{M}, P \in R^N, 1 \leqslant j \leqslant M\} \tag{3}$$

That is, all the VPs that satisfy Eq. (3) fall into subspace $R_{x_i}(j)$.

Upon further expansion, the space is separately divided according to coordinate axes $x_1, x_2, \ldots, x_N$, and the entire space is divided into M^N subspaces. Each subspace can be represented as

$$R(s_1, s_2, \ldots, s_N) = \{P \mid (s_i - 1)\frac{\pi}{M} \leq \alpha_{it} < s_i\frac{\pi}{M}, 1 \leq i \leq N, P \in R^N\} \tag{4}$$

where $1 \leq s_1, s_2, \ldots, s_N \leq M$.

The calculation procedure of HMCSO is as follows.

First, determine $R(s_1, s_2, \ldots, s_N)$ where VP $\mathbf{X}_t$ in the space-time fragment **stf** belongs to according to the angle $\alpha_t = (A_{X_t,x_1}, A_{X_t,x_2}, \ldots, A_{X_t,x_N})$ between VP X_t and axes $x_1, x_2, \ldots, x_N$.

Second, compute the distance between VPs in the space-time fragment with the geometric center defined as:

$$d_j = \left\| \mathbf{X}_j - \mathbf{X}_0 \right\| \tag{5}$$

where $j = 1, 2, \ldots, T$ and $\mathbf{X}_0$ is the geometric center. d_j is defined as the vibration amplitude of VP X_j.

Third, calculate the average distance between VPs in the subspace and geometric center

$$h_k = \frac{1}{z_k} \sum_q d_q^k \tag{6}$$

where d_q^k denotes the distance between VPs falling into the kth subspace with the geometric center, $k = 1, 2, \ldots, M^N$. z_k represents the number of VPs falling into the kth subspace.

Finally, construct HMCSO as follows:

$$\mathbf{H} = [h_1, h_2, \ldots, h_{M^N}] \tag{7}$$

where $\mathbf{H}$ is the feature vector of HMCSO.

3.4. Discussion of HMCSO

In the process of forming HMCSO, the high-dimensional space is divided into several subspaces and the mean of vibration amplitudes of VPs within each subspace are selected as bin values to generate a histogram feature vector, which characterizes the fluctuation of the vibration amplitude in the high-dimensional space. Next, we introduce two special cases of HMCSO to explain characteristics of HMCSO in detail.

As two vibration sensors are orthogonally located to each other in the radial plane of the rotor in rotating machinery, MCSO degenerates into a traditional shaft orbit, and the N-dimensional space is also reduced to an x_1x_2 plane composed of two axes. The two channels of signals constitute a space-time fragment. According to the mechanism of HMCSO, the plane is divided into several subspaces. The average amplitude of VPs falling into the corresponding subspace is calculated, which constitutes the HMCSO feature. When these average amplitudes have a small variance, it means that the shaft orbit is circular; that is, the system is in a normal state, whereas if the shaft orbit exhibits a deformed shape, the latter indicates that a fault has occurred.

Fig. 4 specifically shows an example HMCSO with two vibration sensors, and we elaborate on the calculation of HMCSO. Set M to 4, that is, divide the plane into four subspaces according to a certain coordinate axis. Fig. 4(a) and (b) show the division
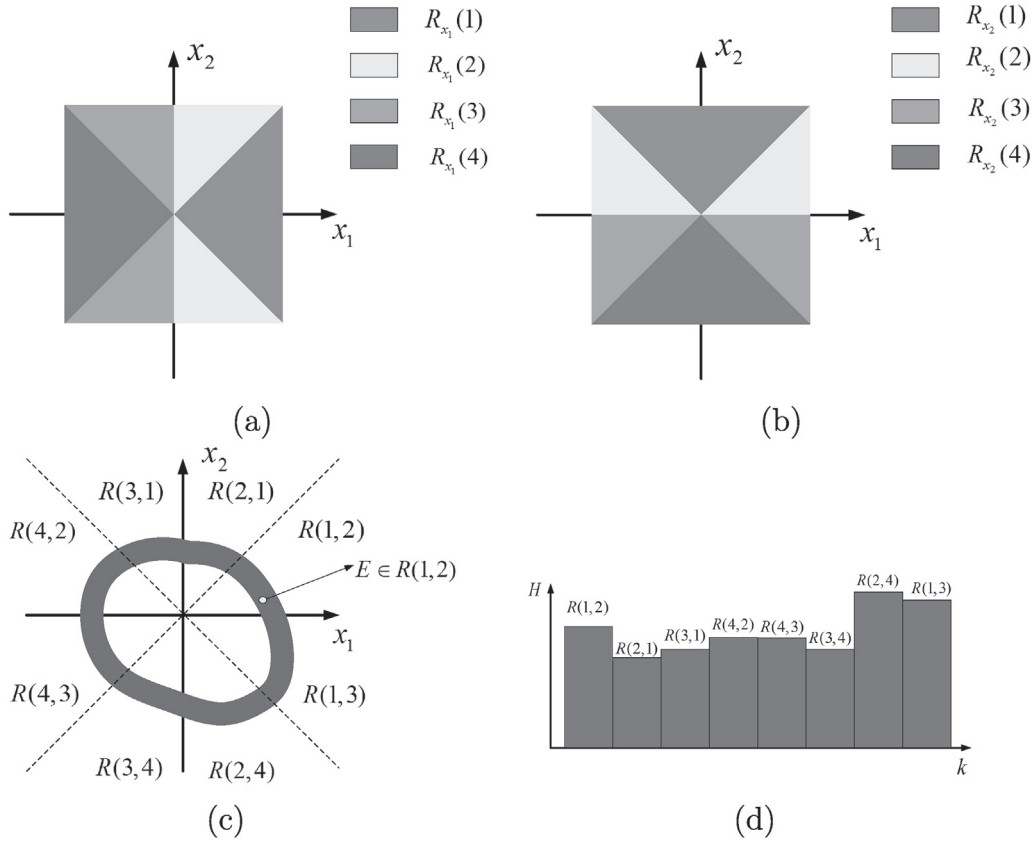
Fig. 4. An example of HMCSO with two vibration sensors where $N = 2$ and $M = 4$. (a) and (b) show the divided subspaces according to axes x_1 and x_2, (c) is the MCSO in $x_1 - x_2$ space and (d) shows the HMCSO in two-dimensional space.

results separately according to axes x_1 and x_2. Hence, the plane is divided into 16 subspaces that are represented by $R(s_1, s_2)$. Selecting VP E as an example in Fig. 4(c), $E \in R(1, 2)$ indicates that E simultaneously falls into the subspace $R_{x_1}(1)$ relative to x_1 and the subspace $R_{x_2}(2)$ relative to x_2. The angle A_{E,x_1} ranges from 0 to $\pi/4$. Similarly, the angle formed by E and axis x_2 ranges from $\pi/4$ to $\pi/2$. Fig. 4(d) shows the average value of the vibration amplitude of the VPs falling into subspaces. Obviously, $R(2, 4)$ corresponds to the largest bin value and indicates the vibration amplitude is larger than those in other subspaces, while the vibration amplitude in $R(2, 1)$ is the smallest of all subspaces. The variation of bin values correspond exactly to the shape of the shaft orbit. Therefore, HMCSO represents the shape of MCSO to infer the class of faults. We also observe that 16 subspaces should have been generated, but in fact, only 8 subspaces actually exist. The bin value corresponding to 8 non-existent subspaces is set to 0, and the histogram is a sparse feature vector with 16 dimensions.

When the number of sensors is 3, that is, $N = 3$, the space $x_1 x_2 x_3$ is divided into several subspaces, which is similar to that when $N = 2$. M is set to 4 and the subspaces according to x_1 are shown in Fig. 5. $R_{x_1}(1)$ and $R_{x_1}(4)$ are separate cones(see Fig. 5(a) and (d)). $R_{x_1}(2)$ presents positive half of the space R^3 by removing $R_{x_1}(1)$, and $R_{x_1}(3)$ is the other half of the space when $R_{x_1}(4)$ is removed(see Fig. 5(b) and (c)). The division of subspaces according to x_2/x_3 is analogous to that according to x_1. The entire space is divided into 4^3 subspaces so that the histogram vector has a dimension of 4^3.

Base on the two-dimensional and three-dimensional cases, the calculation of HMCSO in an N-dimensional space is demonstrated. It can also be inferred that HMCSO can characterize the distribution of VPs and represent the shape of MCSO.

3.5. Summary

This section describes the shaft orbit, MCSO and HMCMO. The relationship between the aforementioned and space-time fragment can be summarized. Space-time fragment is a novel perspective for achieving multi-sensor data fusion at the data level and is essentially equivalent to MCSO. MCSO is the embodiment of space-time fragment in high-dimensional space, and transforms the matrix form into a curve form. The traditional shaft orbit is a special case of MCSO with only two sensors. HMCSO is a kind of feature extracted from MCSO, which presents the shape information of MCSO and mine the coupling information between multi-channel signals; this approach can be considered as a feature-level multi-sensor data fusion method.
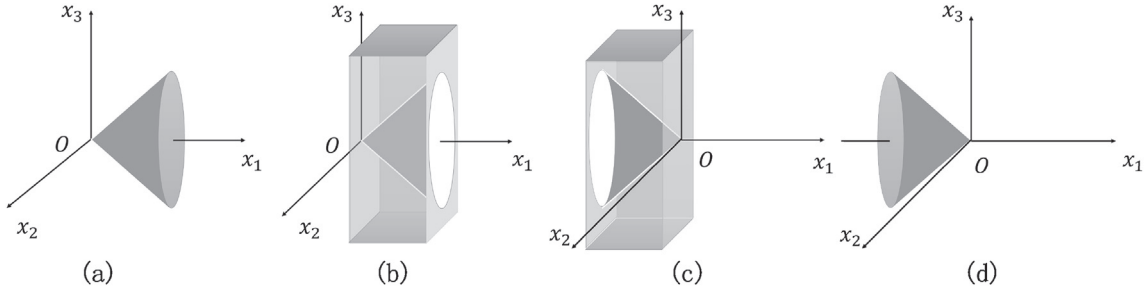
Fig. 5. An example of HMCSO when $N = 3$ and $M = 4$. (a) $R_{x_i}(1)$, (b)$R_{x_i}(2)$, (c)$R_{x_i}(3)$ and (d) $R_{x_i}(4)$.

4. Classification algorithm

Fault diagnosis is a pattern recognition problem that includes two steps: feature extraction and classification. Feature extraction is introduced in the last section and a classification method is proposed in this section. In the study of fault diagnosis, various classification methods in the field of machine learning have been applied, such as Naive Bayes classifier [32], Support Vector Machines(SVM) [33], deep neural networks [2] and AdaBoost [31]. Since this paper focuses on the method of feature extraction, a basic classification algorithm K-nearest neighbor(KNN) is adopted in our experiment.

4.1. K-nearest neighbor

The KNN is a classic but effective pattern recognition method. Its principle is that a sample belongs to the same class as the nearest ones in the feature space. Assume that the training dataset is:

$$T = \left\{ (\mathbf{X}_1, y_1), (\mathbf{X}_2, y_2), \ldots, (\mathbf{X}_{N_{\text{train}}}, y_{N_{\text{train}}}) \right\} \tag{8}$$

where N_{train} is the sample number of training dataset. Each sample $\mathbf{X}_i$ associates with a label $y_i \in \left\{ l_1, l_2, \ldots, l_C \right\}$ and C is the number of classes. Find the K samples nearest to the testing sample $\mathbf{X}_{\text{test}}$ in the training dataset as $N_K(\mathbf{X}_{\text{test}})$ according to the given distance metric. Based on the decision strategy, the label of testing sample is(see Eq. (9)):

$$y_{\text{test}} = \arg \max_j \sum_{\mathbf{X}_i \in N_K(\mathbf{X}_{\text{test}})} I\left(y_i = l_j \right) \tag{9}$$

where $i = 1, 2, \ldots, N_{\text{train}}$ and $j = 1, 2, \ldots, C$. I is an indicator function that equals to 1 when $y_i = l_j$, otherwise 0.

The main parameters of KNN are the distance metric and number of neighbors. The distance metric include Euclidean distance, Mahalanobis distance and Hamming distance. The number of neighbors K is selected based on experience, and voting decisions are made according to the labels of K nearest training samples. In this paper, the Euclidean distance is chosen and set K to 1.

KNN is simple and easy to use, has a clear physical meaning, and does not require a training process. In many applications, the effect of KNN is even better than that of certain sophisticated algorithms. The disadvantage is that in testing process, the new sample should be compared with a large number of training samples, which is time-consuming as the dimension of features increases.

4.2. The proposed algorithm

Fig. 6 shows the flowchart of the proposed algorithm. After vibration signals are sampled by multiple sensors, a space-time fragment constructed, and MCSO is formed in an N-dimensional space. The histogram feature is extracted from MCSO as a feature vector for classification. Different from traditional diagnosis task, the construction of space-time fragment is inserted before the step of feature extraction to achieve multi-sensor fusion at the data level and prepare organized data for feature selection.

5. Experiments

To verify the performance of the proposed algorithm, we use a rig of rotating machinery supported by active magnetic bearing (AMB) for the experiments, and the AMB system is equipped with multiple displacement sensors to construct a fault dataset. AMB is a kind of mechatronic equipment that has the characteristics of no lubrication and no contact with rotor, and is widely used in various fields of industry, such as high-speed motors, compressors, vacuum pumps, etc [34–36].
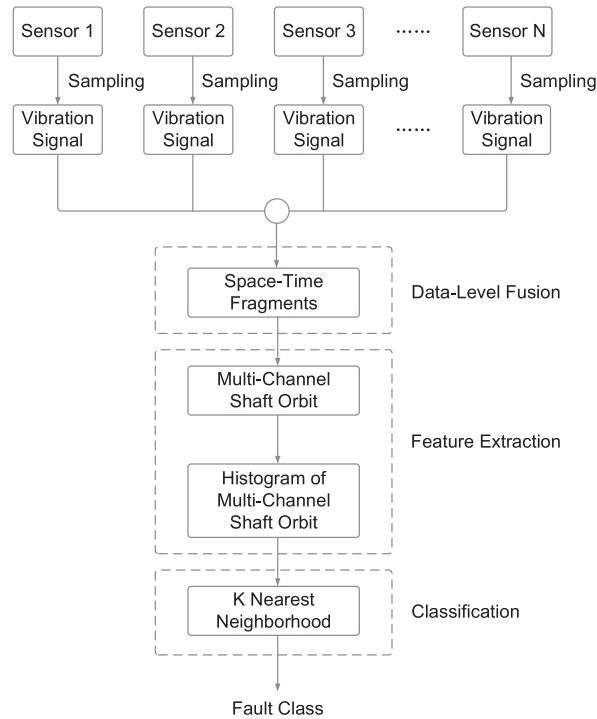
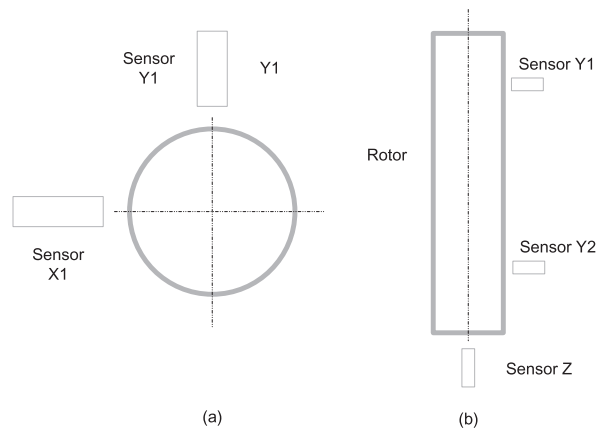Fig. 6. Flowchart of the proposed algorithm.

Fig. 7. The layout of displacement sensors in the test rig, (a) radial (b) axial.

The test rig is vertical with a rotor length of 1053 mm and a motor rotating speed of 6000 rpm. The rig is equipped with five displacement sensors for displacement feedback of the AMB system which also act as vibration sensors to collect vibration data and monitor the running state of the system. The layout of the sensors and physical test rig are shown in Figs. 7 and 8 respectively. The number of sensors N is fixed at 5.

Four classes of data samples are obtained from the test rig, including normal, unbalance, misalignment, and rub-impact, with a total of 1445 samples. Each sample consists of 15,000 data points in 0.6 s. The number of samples in each class is shown in Table 1, and the signal demos are presented in Fig. 9. The experiments are performed under condition that 150 samples are randomly selected from each class of samples for training and the remaining are reserved for testing. Each experiment are repeated 20 times, and the average accuracy is selected as the final result.
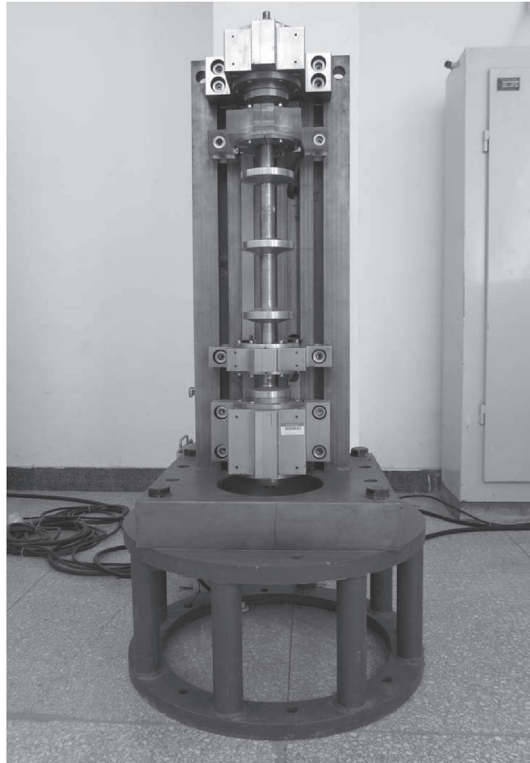
Fig. 8. The rig of the rotating machinery supported by AMB used for experiments.

Table 1
Dataset setting.

Classes	Sample Number	Training Number	Testing Number
Normal	410	150	260
Unbalance	508	150	358
Misalignment	322	150	172
Rub-impact	205	150	55

5.1. Comparison with other fusion algorithms

To illustrate the performance of the proposed algorithm, we compare it to two multi-sensor fusion algorithms.

The first comparison algorithm is voting, which extracts features from each channel of signal and provides results separately, thus the final decision is voted for. The algorithm fuses multi-sensor data at the decision level. The second algorithm is connecting by which joint features are extracted from multi-channel signals and all feature vectors are combined into one long feature vector; classification is conducted according to the long feature vector, and this method fuses multi-sensor data fusion at the feature level. To facilitate comparison, both comparison algorithms use KNN for classification, and short-time Fourier transform is selected as the basic type of feature.

From Table 2, the proposed algorithm is compared with two multi-sensor fusion algorithms. The HMCSO algorithm yields a recognition accuracy of 97.3%, which is much higher than the other two algorithms. The results indicate that HMCSO acquires more information than traditional fusion strategies due to mining of the coupling relationship between multi-channel signals.

Fig. 10 presents confusion matrices of the three algorithms. HMCSO can achieve a greater than 96% accuracy in distinguishing all four kinds of faults and even achieve an accuracy of 100% in both misalignment and rub-impact faults. The voting method exhibits more misjudgments than HMCSO, and 12% of the unbalanced samples are mistakenly considered as normal samples, while 7% of the rub-impact samples are misclassified as unbalance fault. Similarly, in the connecting algorithm, 14% of the unbalanced samples are considered normal samples. The above results show that the HMCSO, which contains a large amount of coupling information between different sensors, not only has an improved recognition accuracy in general, but also correctly classifies each class of samples evenly. The voting and connecting algorithms have varying performances in identifying different
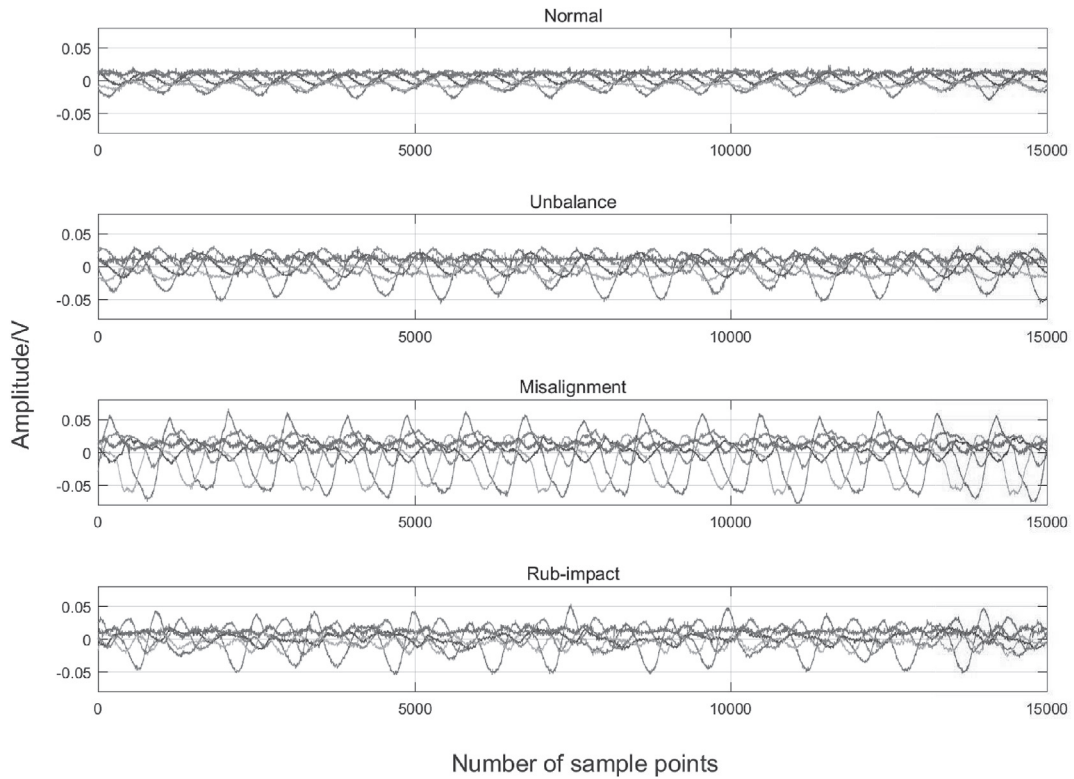
Fig. 9. The signal demo of four classes.

Table 2
Comparison between the proposed algorithm and two other multi-sensor fusion algorithms.

Algorithm	Accuracy
HMCSO	97.3%
Voting	90.3%
Connecting	92.2%

class of faults. Hence, HMCSO is proven to be more comprehensive and balanced to describe and distinguish various classes of faults.

5.2. The relationship between feature dimension and performance of HMCSO

According to the angle between the vibration point and a specific coordinate axis, the space is divided into M subspaces, so that the whole space is divided into M^N subspaces; as a result, the dimension of the feature vector of HMCSO is M^N.

As illustrated in Fig. 11, as M increases, the dimension of HMCSO increases exponentially, while the recognition accuracy of faults steadily increases. When M exceeds 5, the dimension of features still increases sharply, but the accuracy increases slightly even decreases in certain situations because the feature dimension at that point has far exceeded the number of samples. According to the Vapnik-Chervonenkis theory, the model is overfitted which results in the computation time and memory requirement to increase greatly.

The dimension of HMCSO represents the fineness of space division. Intuitively, as the division becomes finer, the shape of MCSO described by the histogram is closer to the true distribution of VPs in the continuous high-dimensional space. Therefore, when applying HMCSO, it is necessary to select appropriate M values to ensure a suitable recognition accuracy while maintaining the time and memory consumption in a proper range by prevent overfitting.

5.3. The relationship between number of sensors and performance of HMCSO

In the test rig, there are four radial sensors, X1, Y1, X2, Y2, and one axial sensor, Z, for measuring the vibration of the rotor system. As shown in Fig. 7, four radial sensors are distributed in two axial planes, and X1/Y1 and X2/Y2 are distributed perpen-
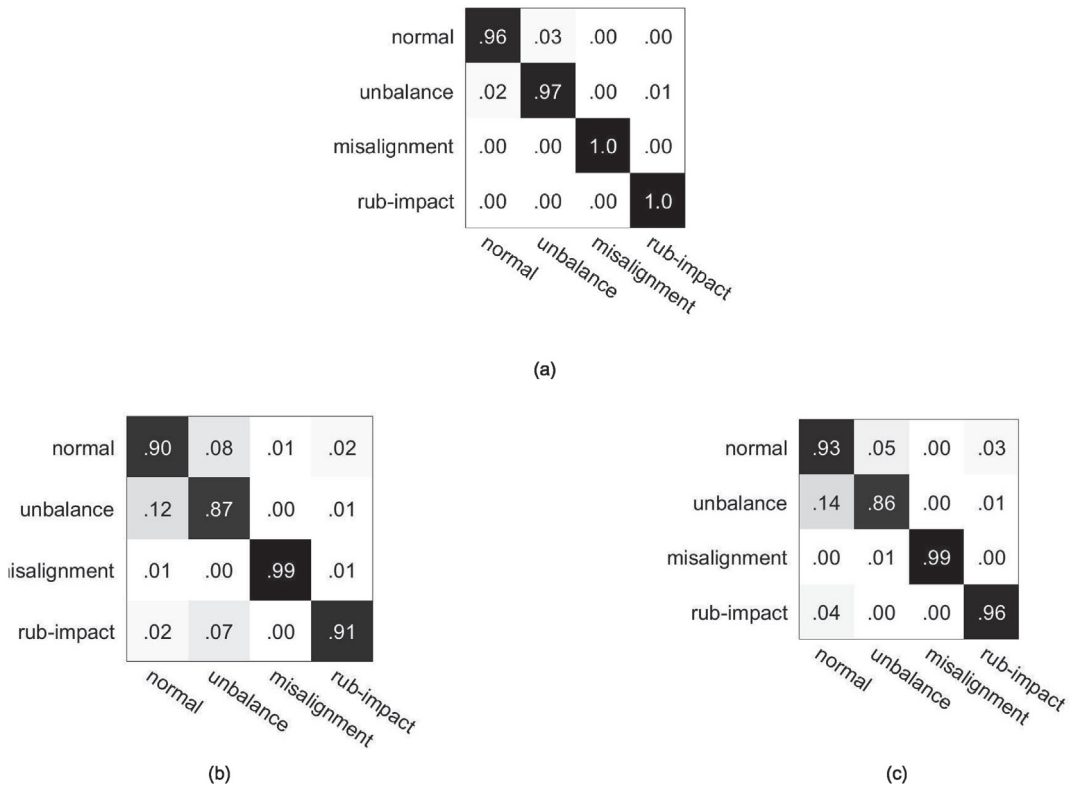
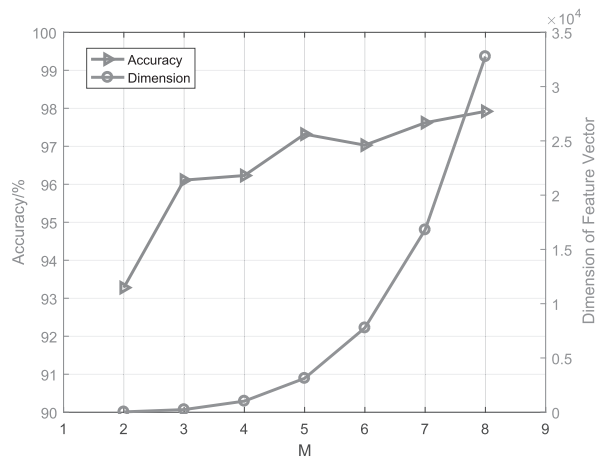**Fig. 10.** Confusion matrices of (a)HMCSO, (b)voting and (c)connecting.

Fig. 11. The relationship between *M* and the performance of HMCSO.

dicular to each other respectively. The signals of two or three sensors are selected to construct the HMCSO feature, each with 10 different combinations, and 4 sensors are selected with 5 different combinations.

Fig. 12 specifically shows the recognition accuracy of HMCSO in the case of using different combinations with various numbers of sensors. When two sensors are used, the performance of each combination is quite different, and the recognition accuracy of the optimal X1/X2 and X1/Y2 combinations are both greater than 93%, while that of the combination Y1/Z is only 61.4%, which indicates that different sensor combinations provide varying amounts of information for identification. The combinations with three or four sensors, including the X1/Y2 combination, always performances better, while the sensor combinations that includes Y1/Z sensor combination always has a relatively low recognition accuracy. It is therefore concluded that coupling information between X1/Y2 has been determined, and that different sensors in a rotor system provide different contributions to fault diagnosis.
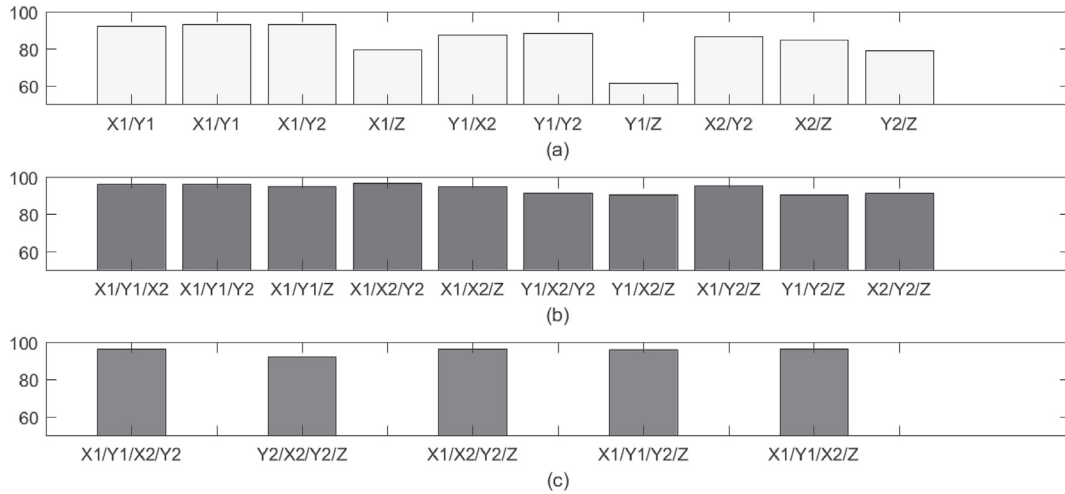
Fig. 12. Performance under different sensor combinations. (a)combinations with two sensors, (b)combinations with three sensors and (c)combinations with four sensors.

Table 3
Average accuracy of HMCSO given different numbers of sensors.

Number of Sensors	Average Accuracy%
2	84.8 ± 9.6
3	94.1 ± 2.6
4	95.4 ± 1.8

Table 3 provides the average and standard deviation of the recognition accuracy given various numbers of sensors. Overall, with the increase in the number of sensors, the average recognition rate of the various combinations gradually increases, indicating that the increase in the number of sensors provides a larger amount of useful information for the diagnosis task, and the variation in the recognition accuracy of different combinations also decreases with the increasing number of sensors. These results show that HMCSO achieves effective fusion of sensor information of "good" sensors and "bad" sensors by mining the coupling relationship between different sensors.

5.4. The effect of noise on the performance of the algorithm

It is important to verify whether HMCSO works when the measured signals are noisy. Although the dataset is collected under realistic working conditions, we add Gaussian white noise to all channels of signals. The signal-to-noise ratio (SNR) is used as a measure of the level of noise. Fig. 13 shows the signal affected by noise with different SNRs. The signals become distorted heavily when SNR is lower than 40 dB and cannot be identified when the SNR is lower than 25 dB.

Fig. 14 shows the recognition accuracy with various SNR values. When the value of the SNR is larger than 30 dB, HMCSO can still achieve an accuracy greater than 90%. As the SNR increases, the accuracy improves and gradually similar to that of the original signal. However, the accuracy of HMCSO sharply decreases as the noise level increases. When the SNR is lower than 5 dB, all of the channels of the original signals are heavily polluted by noise and the classification result becomes random. The denoising method is suggested to introduce under severe noise conditions, which will significantly improve the accuracy of our algorithm.

The above experimental results verifies that HMCSO has strong resistance to noise and can be applied to noise environment. Since HMCSO is in essence a histogram representation which is statistical and resistant to noise. In addition, each bin value of the histogram adopts the average of vibration amplitudes of VPs, which also reduces the effect of noise.

5.5. The effect of sensor fault on the proposed algorithm

Since HMCSO is a characteristic representation of multi-sensor data fusion, when the sensor itself has a fault, the latter will affect the performance of HMCSO. Hence, we classify sensor faults into two categories and discuss the effects of sensor faults on the HMCSO performance.

For the first kind of sensor fault, some sensors are severely affected by noise. Unlike the previous section, in which it is discussed that all sensors are subject to noise, only one or two of the sensors are severely disturbed by noise in this section. Figs. 15 and 16 show the algorithm recognition rate when one or two of the five sensors are subjected to different levels of noise. It can be observed that the algorithm embodies a strong anti-noise ability, and even when SNR is equal to 5 dB, the
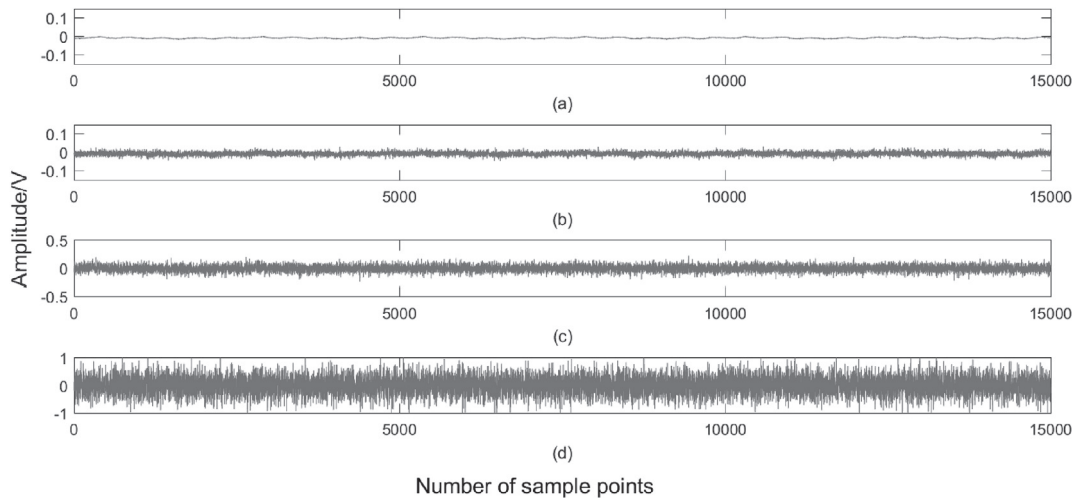
Fig. 13. An example of the signals affected by noise (a) original signal (b) SNR = 40 dB (c) SNR = 25 dB (d) SNR = 5 dB.
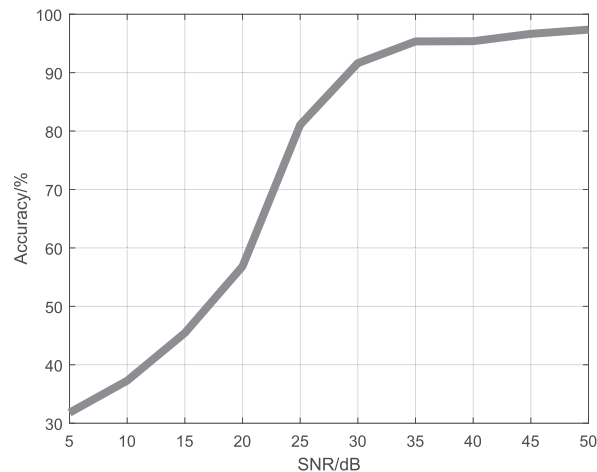

Fig. 14. Performance of the proposed algorithm under various noise levels.

algorithm can reach a recognition rate of greater than 80%. The algorithm only fails when the SNR of the two signals is 5 dB.

For the second kind of sensor fault, the abnormal signals caused by the sensor failure mainly includes the random output, zero output and saturation output. Figs. 17 and 18 show the performance under the conditions that one or two sensors are useless. The algorithm performs normally under a zero output of sensors, which is equivalent to constructing the HMCSO with four or three sensors. With a random output of sensors, it is similar to the situation of severe noise interference, and the classification accuracy is still greater than 80%, even though affected to a certain extent. When the two sensors are affected, the classification results lose credibility. The algorithm fails as if one of sensors reaches saturation output because the fault signal is always at the maximum amplitude which dominates the shape of the MCSO so that the other normal signals cannot convey information.

As seen from the above experiments, when one of the five sensors fails, in addition to the situation of saturation output, the HMCSO mostly exhibits an identification accuracy similar to that all sensors are in good condition. When two sensors fail, that is, 40% of the information sources are not functioning sufficiently accurate, the recognition accuracy of the HMCSO is seriously affected. In actual applications, sensor faults can often be detected through a self-test mechanism, and when constructing the HMCSO, the remaining functioning sensors can be used to acquire the signals to construct the HMCSO. As mentioned in the previous section, when the three or four channels of signals are utilized to construct HMCSO, recognition results with high degrees of accuracy can still be obtained.
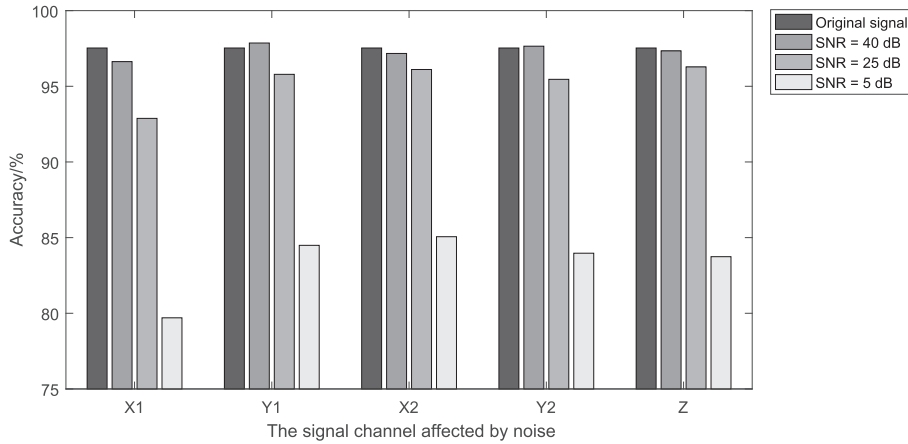
Fig. 15. Performance of the proposed algorithm under the condition of one sensor with the first kind of sensor fault.
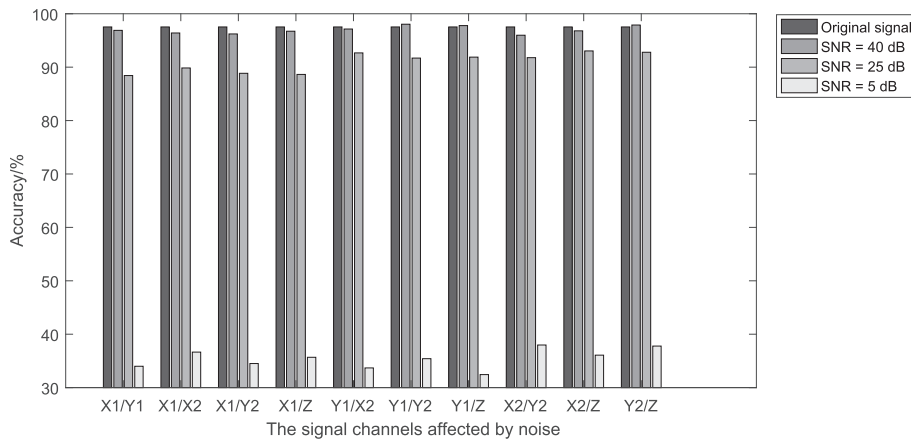

Fig. 16. Performance of the proposed algorithm under the condition of two sensors with the first kind of sensor fault.

5.6. Summary of experiments

In this section, we perform a series of experiments on a test rig of rotating machinery supported by AMB for fault diagnosis, and draw some conclusions.

First, compared with other multi-sensor fusion algorithms, HMCSO has the advantage of providing a high recognition accuracy by mining information of the coupling relationship between different sensors, and classifies different classes of faults evenly.

Second, the performance of HMCSO is related to M. As the value of M increases, the dimension of the feature increases exponentially and the recognition accuracy also increases. However, in practical application, it is necessary to strike a compromise with regards to selection of the M value to ensure a certain level of precision and avoid overfitting.

Third, the number of sensors also affects the performance of HMCSO. In general, the increase in the number of sensors results in improvement of the performance of HMCSO and reduction in the performance variation of different sensor combinations.

Fourth, HMCSO has strong anti-noise ability and can work under heavy noise environment.

Fifth, it is still useful for HMCSO when one of the five sensors occurs a fault in most situations. However, when two sensors fail, HMCSO basically is out of use and the self-test mechanism of sensors is needed to find the sensors in the fault state, then HMCSO is constructed by the signals captured by the remaining functioning sensors.

It should be noted that since the databases used in various previous studies are different, and the performance of algorithms involved is related to finely tuned parameters. Therefore, we focus on the significance of space-time fragment and the validity of HMCSO features in experiments, instead of emphasizing proof that the proposed algorithm performs better than other existing algorithms. In practical applications, the readers can choose to use space-time fragments and HMCSO in other algorithm frameworks. HMCSO can also be used as one kind of feature to fuse with other features to describe faults.
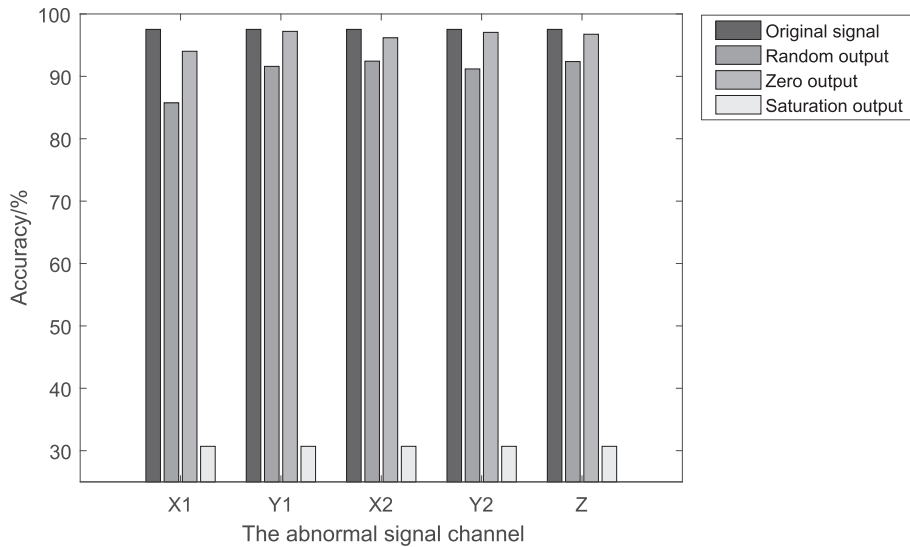
Fig. 17. Performance of the proposed algorithm under the condition of one sensor with the second kind of sensor fault.
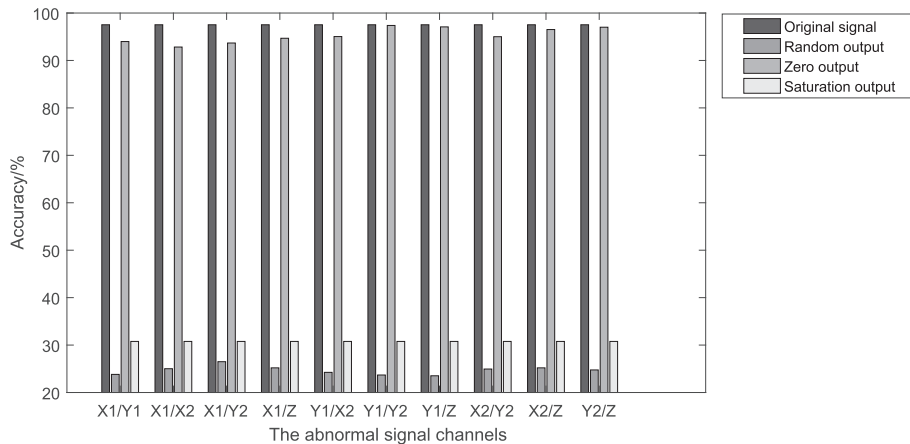

Fig. 18. Performance of the proposed algorithm under the condition of two sensors with the second kind of sensor fault.

6. Conclusion

In this paper, the concept of space-time fragments is proposed, and vibration signals captured by multiple sensors are fused at the data level. Histogram of multi-channel shaft orbit is designed, as a new kind of feature based on space-time fragment, which fully retains the coupling relationship of multiple sensors. Experiments have verified the effectiveness of HMCSO.

In the future, HMCSO can be merged with other kinds of features to achieve a more discriminating fault description. Moreover, space-time fragments can be studied in greater detail to mine more effective features for use in fault diagnosis.

Acknowledgments

References

[1] R. Liu, B. Yang, E. Zio, X. Chen, Artificial intelligence for fault diagnosis of rotating machinery: a review, Mech. Syst. Signal Process. 108 (2018) 33–47.
[2] S. Khan, T. Yairi, A review on the application of deep learning in system health management, Mech. Syst. Signal Process. 107 (2018) 241–265.
[3] Y. Wang, J. Xiang, R. Markert, M. Liang, Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications, Mech. Syst. Signal Process. 6667 (2016) 679–698.

[4] F. Li, B. Tang, R. Yang, Rotating machine fault diagnosis using dimension reduction with linear local tangent space alignment, Measurement 46 (8) (2013) 2525–2539.
[5] X. Xiang, J. Zhou, C. Li, Q. Li, Z. Luo, Fault diagnosis based on walsh transform and rough sets, Mech. Syst. Signal Process. 23 (4) (2009) 1313–1326.
[6] J. Pieyro, A. Klempnow, V. Lescano, Effectiveness of new spectral tools in the anomaly detection of rolling element bearings, J. Alloy. Compd. 310 (1) (2000) 276–279.
[7] M. Cerrada, R.-V. Snchez, C. Li, F. Pacheco, D. Cabrera, J. Valente De Oliveira, R.E. Vsquez, A review on data-driven fault severity assessment in rolling bearings, Mech. Syst. Signal Process. 99 (2018) 169–196.
[8] L. Ou, D. Yu, H. Yang, A new rolling bearing fault diagnosis method based on gft impulse component extraction, Mech. Syst. Signal Process. 81 (2016) 162–182.
[9] Y. Han, B. Tang, L. Deng, Multi-level wavelet packet fusion in dynamic ensemble convolutional neural network for fault diagnosis, Measurement 127 (2018) 246–255.
[10] R. Yuan, Y. Lv, G. Song, Multi-fault diagnosis of rolling bearings via adaptive projection intrinsically transformed multivariate empirical mode decomposition and high order singular value decomposition, Sensors 18 (4) (2018) 1210.
[11] Y. Zhang, D. Ye, Y. Liu, Robust locally linear embedding algorithm for machinery fault diagnosis, Neurocomputing 273 (2018) 323–332.
[12] G. Chen, F. Liu, W. Huang, Sparse discriminant manifold projections for bearing fault diagnosis, J. Sound Vib. 399 (2017) 330–344.
[13] H. Shao, H. Jiang, Y. Lin, X. Li, A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders, Mech. Syst. Signal Process. 102 (2018) 278–297.
[14] S. Guo, T. Yang, W. Gao, C. Zhang, A novel fault diagnosis method for rotating machinery based on a convolutional neural network, Sensors 18 (5) (2018) 1429.
[15] F. Jia, Y. Lei, N. Lu, S. Xing, Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization, Mech. Syst. Signal Process. 110 (2018) 349–367.
[16] W. Zhang, C. Li, G. Peng, Y. Chen, Z. Zhang, A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load, Mech. Syst. Signal Process. 100 (2018) 439–453.
[17] C. Sobie, C. Freitas, M. Nicolai, Simulation-driven machine learning: bearing fault classification, Mech. Syst. Signal Process. 99 (2018) 403–419.
[18] H. Liu, J. Zhou, Y. Zheng, W. Jiang, Y. Zhang, Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders, ISA (Instrum. Soc. Am.) Trans. 77 (2018) 167–178.
[19] M.S. Safizadeh, S.K. Latifi, Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell, Inf. Fusion 18 (2014) 1–8.
[20] D.T. Hoang, H.J. Kang, Rolling element bearing fault diagnosis using convolutional neural network and vibration image, Cogn. Syst. Res. 53 (2019) 42–50.
[21] M. Xia, T. Li, L. Xu, L. Liu, C.W. De Silva, Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks, IEEE ASME Trans. Mechatron. 23 (1) (2018) 101–110.
[22] J. Liu, Y. Hu, Y. Wang, B. Wu, J. Fan, Z. Hu, An integrated multi-sensor fusion-based deep feature learning approach for rotating machinery diagnosis, Meas. Sci. Technol. 29 (5) (2018) 055103.
[23] Y. Lei, J. Lin, Z. He, D. Kong, A method based on multi-sensor data fusion for fault detection of planetary gearboxes, Sensors 12 (2) (2012) 2005–2017.
[24] Y. Lei, M.J. Zuo, Gear crack level identification based on weighted k nearest neighbor classification algorithm, Mech. Syst. Signal Process. 23 (5) (2009) 1535–1547.
[25] J. Ma, F. Xu, K. Huang, R. Huang, Gnar-garch model and its application in feature extraction for rolling bearing fault diagnosis, Mech. Syst. Signal Process. 93 (2017) 175–203.
[26] Y.T. Ai, J.Y. Guan, C.W. Fei, J. Tian, F.L. Zhang, Fusion information entropy method of rolling bearing fault diagnosis based on n -dimensional characteristic parameter distance, Mech. Syst. Signal Process. 88 (2017) 123–136.
[27] X. Yan, Y. Luo, Recognizing human actions using a new descriptor based on spatialctemporal interest points and weighted-output classifier, Neurocomputing 87 (1) (2012) 51–61.
[28] I. Laptev, On space-time interest points, Int. J. Comput. Vis. 64 (2005) 107–123.
[29] X. Yan, Y. Luo, Action recognition via cumulative histogram of multiple features, Opt. Eng. 50 (1) (2011) 123–128.
[30] X. Yan, Y. Luo, Making full use of spatial-temporal interest points: an adaboost approach for action recognition, in: IEEE International Conference on Image Processing, 2010, pp. 4677–4680.
[31] X. Yan, Z. Sun, J. Zhao, Z. Shi, C.-A. Zhang, Fault diagnosis of active magnetic bearing rotor system via vibration images, Sensors 19 (2) (2019) 144.
[32] N. Zhang, L. Wu, J. Yang, Y. Guan, Naive bayes bearing fault diagnosis based on enhanced independence of data, Sensors 18 (2) (2018) 463.
[33] Z. Su, B. Tang, Z. Liu, Y. Qin, Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine, Neurocomputing 157 (2015) 208–222.
[34] G. Schweitzer, E.H. Maslen, Magnetic Bearings: Theory, Design, and Application to Rotating Machinery, Springer Berlin Heidelberg, 2009.
[35] J. Zhao, Z. Sun, X. Yan, G. Yang, Y. Zhou, X. Liu, Z. Shi, T. Fan, X. Zhang, Helium blower test based on aerodynamic force simulation, Ann. Nucl. Energy 118 (2018) 283–290.
[36] D. He, Z. Shi, X. Yan, J. Zhao, X. Liu, Z. Sun, G. Yang, Study on fault diagnosis system of active magnetic bearing, in: International Topical Meeting on High Temperature Reactor Technology, HTR, American Nuclear Society, Las Vegas, NV, United states, 2016, pp. 56–59.