Research Article

The Application of Orthogonal Wavelet Transformation: Support Vector Data Description in Evaluating the Performance and Health of Bearings

Weipeng Li1,2, Yan Cao1, Lijuan Li1, and Siyu Hou1

1School of Mechanical Electrical Engineering, Xi’an Technological University, Xi’an 710000, China
2School of Intelligent Manufacturing, Nanyang Institute of Technology, Nanyang 473000, China

Correspondence should be addressed to Yan Cao; caoyan@xatu.edu.cn

Received 22 November 2021; Revised 30 January 2022; Accepted 24 February 2022; Published 18 March 2022

Academic Editor: Chun Wei

Support vector data description (SVDD) is common supervised learning. Its basic idea is to establish a closed and compact area with the objects to be described as integrity. The described objects are all included within the area or as far as possible. In contrast, other objects are excluded out of the area as far as possible. The inherent nature and laws of data are subsequently revealed, thereby distinguishing the operation state of the machine. In this paper, an orthogonal wavelet transformation-support vector data description (OWTSVDD) is proposed to evaluate the performance of bearings, where the peak-to-peak value of detail signal is extracted through orthogonal wavelet transformation as the set of test samples, thus solving the distance $R_z$ from the set of test samples to the center of the sphere. Based on $HI = R_z^2 - R^2$, its distance to the hypersphere is calculated to judge whether it belongs to the normal state training samples. Finally, the performance and health of bearings are evaluated with HI. According to the classification of two sets of experimental data of rolling bearings, the proposed method better reflects the degeneration of bearing’s performance compared with the (SVDD) HI value without extraction of characteristic value, being entirely able to evaluate the entire life cycle of bearings from normal operation to fault and degradation. The HI evaluation result based on experimental data in Xi’an Jiaotong University is consistent with the life-cycle vibration signal of bearings, providing a scientific basis for production and equipment management and improving the prognostics technology-centered prognostics and health management (PHM).

1. Introduction

Intelligent fault diagnosis and prognostics evaluation [1, 2] are essential factors of the development and application of machine learning in mechanical fault diagnosis technology [3, 4]. In supervised learning [5, 6], the process of utilizing parameters of a set of sample adjustment classifiers with known classification to enable it to reach the required performance is referred to as supervised training or tutored learning. A supervised learning algorithm analyzes the training data to generate inference and can be used for mapping out new living examples. SVM is a machine learning-based supervised learning theory enabling the classification of data based on the minimization of structural risks [7]. There are two classification problems in engineering practice: relatively simple linear separable problems and linear inseparable problems. When the optimal separating hyperplane is yielded in a linearly separable problem, SVM mostly solves the problem [8]. Through the nonlinear kernel function, SVM projects the nonlinear separable data to a high-dimensional space and turns it into a linearly separable problem, thereby achieving the linear distinction of high-dimensional nonlinear data [9]. As a derivative of SVM, the support vector data description (SVDD) proposed by Tax et al. improves the machine learning theory and has been promoted in fault diagnosis [10]. The basic idea of SVDD is to establish a closed and compact area $\Omega$ with the objects to be described as integrity, whereby the described objects are all included within the area $\Omega$ or as far as possible. In contrast, other objects are
excluded out of $\Omega$ as far as possible. SVDD can find the optimal solution of a function set in a circumstance of limited sample quantity, without seeking the optimal solution when the samples tend towards infinity. In the practice of minimizing empirical risks, as SVDD balances both empirical risks and confidence interval to achieve the optimal classification effect in a circumstance of limited samples, it is widely applied in single classification problems [11, 12]. The development of intelligent bearings promotes the acquisition of a large number of bearing operating state monitoring data, which provides a data resource basis for the bearing remaining service life (RUL) prediction method studied in this topic. There are two main classes of existing RUL prediction methods, based on statistical lifetime model methods and state monitoring-based data-driven methods. The model-based approach works well when the established model can accurately describe the degradation process of the machine. However, in practice, it is quite difficult to build an exact degradation model, as the failure mechanism of machines is diverse or not straightforward. Instead, data-driven methods do not need to know the machine failure mechanism. Machine learning is becoming increasingly attractive in data-driven RUL predictions, and can build diagnostic models directly based on raw sensor data, thus escaping the complex process of manually extracting features. The derivative of SVM (support vector data description SVDD) (SVM) has been applied in the field of fault prediction and health management (PHM), and has achieved good results. The average prediction accuracy can reach 80%, greatly improving the reliability and stability of instruments and equipment work, but this level still has room for improvement from practical industrial applications.

However, in engineering applications, given that most fault signals contain nonstationary signals and have both positive and negative data, the traditional signal treatment method is not applicable. Therefore, adopting an appropriate method for feature extraction of signals will correctly reflect the features contained in signals and significantly reduce the data dimension, enabling swift and accurate evaluations of the conditions of equipment by SVDD [13].

Wavelet transformation is an emerging branch of applied mathematics developed in the last 1980s. The multi-resolution characteristic enables wavelet transformation to achieve observation of signals from rough to fine. Its multi-resolution characteristic is also reflected in "zooming" features, allowing it to efficiently perform the localized analysis of nonstationary signals [14, 15]. Feature extraction is a great bottleneck, as noiseless weak fault signals directly influence the accuracy, validity, or abnormal value of aero-engine’s overall vibration fault diagnosis. In predicting the service life of the key component of bearing in fault diagnosis [16, 17], the features should be extracted to reflect the operation conditions of the system [18, 19], and high sensitivity is maintained against abnormal signals [20, 21]. Orthogonal wavelet transformation refers to the selection of orthogonal wavelet function for wavelet transformation, reflecting the local features of signals in the time and frequency domain. It decomposes original signals successively to corresponding local detail signals at each scale, through which the features of original data may be effectively mastered upon analysis of local detail signals. Therefore, a combination of orthogonal wavelet transformation (OWT) [22, 23] and support vector data description (SVDD) is proposed [24, 25]. The OWTSVDD for evaluation of bearing’s performance offers a quantitative health index for conditions evaluation. In the application of OWTSVDD in fault diagnosis, the peak-to-peak value of each detail is extracted with orthogonal wavelet transformation to serve as the training samples of SVDD, therefore solving the distance $R_c$ from the set of test samples to the center of the sphere. Based on $HI = R_c^2 - R^2$, its distance to the hypersphere is subsequently calculated to evaluate and predict the bearing fault. The proposed method is more effective in evaluating the conditions of bearings compared with the SVDD without extraction of characteristic value and is consistent with the life test chart of bearings. Furthermore, the validity of the method has also been proved by different experimental data [26].

2. Support Vector Data Description

Developed by David M. J. Tax et al., support vector data description is a single-value classification method whose theoretical basis originates from support vector data description. While the study is conducted with a condition of sufficient samples in the traditional identification method, the machine learning-based support vector data description is a study targeting the circumstances of limited samples. In the study of fault diagnosis, fault samples are not readily yielded, entailing numerous hindrances in studies. How to carry out the classification of faults and evaluate the operating conditions of equipment with limited samples? Based on the foregoing considerations, Vapnik et al. proposed the principle of statistical learning. By the 90’s, the statistical learning theory was improved, followed by the application of the subsequent SVDD in fault diagnosis [27]. As it already exists, sensor data is combined with machine learning (ML) technologies such as Support Vector Machine (SVM) and BP neural networks to learn the degraded properties of the machine and then use these trained ML models to estimate the RUL.

As a derivative of SVM, SVDD can distinguish abnormal data and normal data such as SVM [28]. SVDD distinguishes data by constructing a hypersphere in high-dimensional space, where the center and radius of the hypersphere can be obtained through a penalty parameter [29]. This method satisfactorily displays the spatial features of vibration data under multivariate factors. The basic idea of SVDD is to establish a closed and compact area $\Omega$ with the objects to be described as an integrity, whereby the described objects are all included within the area $\Omega$ or as far as possible while other objects are excluded out of $\Omega$ as far as possible. Assuming that there are $n$ pieces of data to be described $\{x_i\}$, $i = 1, \ldots, n$, the $n$ pieces of data constitute the learning samples of a single classifier. The goal is to seek a supersphere with minimum volume including all learning samples $x_c$. The supersphere is described as a sphere with a center $a$ and a radius $R$, as illustrated in Figure 1.
The following relationship should be satisfied when seeking a suprasphere containing all learning samples:

$$\min \varepsilon = R^2. \quad (1)$$

Constraint condition:

$$\|x_i - a\|^2 \leq R^2, \quad i = 1, \ldots, n. \quad (2)$$

In view of the fact that not all samples are correctly classified, a specific proportion of misclassified samples is allowed. That is, the distance from the sample $x_i$ to the center of the suprasphere is not necessarily smaller than $R$ in the strict sense. Nonetheless, for a distance greater than $R$, $\xi_i$ should satisfy the following: Penalty. This is similar to binary classification and can be interpreted with the views of statistical learning theory. That is, the empirical risks are not equal to zero, and a specific proportion of misclassified samples is allowed. As depicted in Figure 2, the sample $x_i$ is misclassified as a nontarget sample. To enhance the robustness of classification, by introducing relaxing factor $\xi_i \geq 0$, $i = 1, \ldots, n$, the foregoing equation (1) is changed to

$$\min \varepsilon (R, a, \xi) = R^2 + C_1 \sum_{i=1}^{n} \xi_i + C_2 \sum_{i=1}^{n} \xi_i. \quad (3)$$

Constraint condition:

$$\|x_i - a\|^2 \leq R^2 + \xi_i, \quad \|x_l - a\|^2 \leq R^2 + \xi_i, \quad i, l = 1, \ldots, n. \quad (4)$$

In equation (3), $C_1$ and $C_2$ correspond to the error penalty coefficients of the target and nontarget samples, respectively. Like binary classification, a compromise is made between empirical risks and generalization ability, which is also reflective of the mitigation of structural risks in statistical learning theory.

Equation (3) is optimized to solve the minimum value of $\varepsilon (R, a, \xi)$, thereby working out the values of $R, a, \xi$. This is an extremum problem and a Lagrange function has been constructed as follows:

$$L(R, a, \alpha_i, \gamma_i, \xi_i) = R^2 + C_1 \sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \alpha_i \left[ R^2 + \xi_i - (x_i \cdot x_i - 2a \cdot x_i + a \cdot a) \right] - \sum_{i=1}^{n} \gamma_i \xi_i, \quad (5)$$

where $\alpha_i \geq 0, \gamma_i \geq 0$, which are coefficients of Lagrange; $x_i \cdot x_j$ is the inner product of $x_i$ and $x_j$. For each sample $x_i$, there are Lagrange coefficients of $\alpha_i$ and $\gamma_i$. By solving the partial derivative of $R, a$ and $\xi_i$ in Lagrange equation and by setting it to 0, the following is obtained:

$$\sum_{i=1}^{n} \alpha_i = 1, \quad (6)$$

$$a = \sum_{i=1}^{n} \alpha_i x_i, \quad (7)$$

$$\gamma_i = C - \alpha_i. \quad (8)$$

It can be seen that $a$ is obtained from a linear combination of $x_i$, being a vector having the same dimensionality with $x_i$. By substituting $\alpha_i \geq 0$ and $\gamma_i \geq 0$ into equation (7), the following is obtained:

$$0 \leq \alpha_i \leq C. \quad (9)$$
By substituting equations (6) and (7), the optimized objective function of Lagrange may be written into the following form:

$$L = \sum_{i=1}^{n} \alpha_i (x_i, x_i) - \sum_{i=1,j=1}^{n} \alpha_i \alpha_j (x_i, x_j).$$  \hfill (10)

Equation (10) is a quadratic optimization problem, which can be solved with the existing standard algorithm. Under the constraint of equation (10), the minimal value of (10) is solved to obtain the optimal solution $\alpha_i^*$ of the Lagrange parameter $\alpha_i$. The center of classification supersphere is subsequently solved with equation (6). In practical calculation, the corresponding Lagrange parameter $\alpha_i$ is equal to 0 (greater than 0) only for those samples where equality in equation (3) is established. However, these samples usually account for a small proportion. Where inequality is established in equation (3), the corresponding parameter $\alpha_i$ of those samples will be equal to 0. The samples whose corresponding $\alpha_i$ is not 0 are referred to as support vectors and only this small portion of support vectors determines the value of $a$ and $R$. Regarding other nonsupport vectors, their corresponding $\alpha_i = 0$ is overlooked in the calculation. Therefore, a high calculation efficiency is obtained through this method and proved in the subsequent experiment. When a sample’s corresponding $\alpha_i = 0$, the sample is within the classification boundary; if a sample’s corresponding $\alpha_i \geq 0$, the sample is on or outside the classification boundary.

$C$ is the equation (9) which is the upper boundary of $\alpha_i$ and restricts the influence of samples to the establishment of a classifier. When a sample’s corresponding $\alpha_i$ reaches the upper limit $C$, as shown in equation (9), this sample is considered as a nontarget sample and denied.

The radius $R$ of the supersphere may be solved through any support vector $x_k$ through the following equation:

$$R^2 = (x_k \cdot x_k) - 2 \sum_{i=1}^{n} \alpha_i (x_i, x_k) + \sum_{i=1,j=1}^{n} \alpha_i \alpha_j (x_i, x_j).$$  \hfill (11)

To determine whether a new sample $z$ is a target sample, the generalized distance $R_z$ from the sample to the center $a$ of the classifying plane should be calculated first:

$$R_z^2 = \|z - a\|^2 = (z \cdot z) - 2 \sum_{i=1}^{n} \alpha_i (z, x_i) + \sum_{i=1,j=1}^{n} \alpha_i \alpha_j (x_i, x_j).$$  \hfill (12)

Upon judgment of whether $R_z^2 \leq R^2$ is established, it can be determined whether the sample $z$ is a target sample. It may be written into the following form:

$$f_{\text{svda}}(z; a, R) = I[\|z - a\|^2 \leq R^2]$$

$$= I \left( (z \cdot z) - 2 \sum_{i=1}^{n} \alpha_i (z, x_i) + \sum_{i=1,j=1}^{n} \alpha_i \alpha_j (x_i, x_j) \leq R^2 \right).$$  \hfill (13)

Here, the expression of function $I$ is defined as

$$I(A) = \begin{cases} 1, & \text{if } A \text{ is true}, \\ 0, & \text{others}. \end{cases}$$  \hfill (14)

Namely, if $R_z^2 \leq R^2$, the sample $z$ is accepted and determined as a target sample; otherwise, it is rejected and determined as a nontarget sample.

In equations (3) and (5), the parameters $C_1$ and $C_2$ are, respectively, the penalized factors for the classification errors of the two types of samples, and choosing the proportion of different factors can achieve the balance between the two classification errors. These two kinds of classification errors are: the first type of error $\epsilon_1$: misclassify the target sample into nontarget samples, that is, the rejection rate of the target sample, which is called error judgment in the fault diagnosis. The second type of error $\epsilon_2$: misclassify the nontarget sample into the target sample, that is, the acceptance rate of the nontarget sample, which is called a missed judgment in the fault diagnosis.

$$C_1 = \frac{\epsilon_1}{\epsilon_2}.$$  \hfill (15)

In the practical engineering application, the appropriate proportion parameter $C_1/C_2$ can be selected according to the impact of omission and misjudgment on the equipment, so that the loss caused by the classification error to the equipment is minimized.

### 3. Orthogonal Wavelet Transformations

The expansion of the function $z(t)$ in an arbitrary space $L^2 (R)$ under a wavelet basis is called the wavelet transform of function $z(t)$ (abbreviated as WT) [30, 31], and its expression is

$$WT_z(a, b) = \langle z(t), \phi_{a,b}(t) \rangle = \frac{1}{\sqrt{a}} \int_R z(t) \phi \left( \frac{t-b}{a} \right) dt.$$  \hfill (16)

The wavelet basis in wavelet transform is not necessarily an orthogonal basis, but in practical applications, it prefers to find an orthogonal wavelet basis. An important method for constructing orthogonal wavelet bases is referred to as the multiscale analysis.

The multiscale analysis is strictly defined by function space terms, assuming that the subspace sequence $V_m (m \in Z)$ in the space $L^2 (R)$ satisfies the following conditions.

1. Nestededness: $\ldots V_2 \subset V_1 \subset \ldots \subset V_{-2} \subset \ldots$
2. Approximation property: $\cap_{m \in Z} V_m = \{0\}, \cup_{m \in Z} V_m = L^2 (R)$
3. Scalability: $f(t) \in V_n \Leftrightarrow f(2t) \in V_{n-1}$
4. $V_m = \text{span}\{\phi_{m,k} (t), n \in Z\}$ i.e., any level of subspace can be formed by the same function of the corresponding scale by translation
5. $V_{m-1} = V_m \oplus W_m$, that is, any level of subspace can be summarized by the next level of subspace as well as
its orthogonal complement space, and the sequences \( W_m \) do not overlap with each other and are orthogonal systems.

Therefore, the subspace sequences are denoted as \( V_m \) and \( W_m \) multiscale analyzes of the function space. Among them, \( \varphi_{mn} \) is called the scale function, and \( m \) and \( n \) are the scale and translation parameters, respectively.

From the aforementioned definitions, we have

\[
V_0 = V_m \oplus \sum_{i=1}^{m} W_i. \tag{17}
\]

The application of multiscale analyzes in signal processing can be expressed by equation (17). As shown in the following, the arbitrary function \( f(t) \in V_0 \) can be decomposed on the next level of scale and wavelet space.

\[
f(t) = p_1 f(t) + q_1 f(t), \tag{18}
\]

where \( p_1 f(t) = \sum_k C_{1k} \varphi_{1k} \) and \( q_1 f(t) = \sum_k D_{1k} \Theta_{1k} \).

\( p_1 f(t) \) is the approximation part, and \( q_1 f(t) \) is the detailed part. The approximation part is subsequently further decomposed repeatedly to yield the approximation and detail parts on any scale. The iterative formula is

\[
p_{m-1} f(t) = p_m f(t) + q_m f(t)
= \sum_k C_{mk} \varphi_{mk} + \sum_k D_{mk} \Theta_{mk}, \tag{19}
\]

where \( C_m = H C_{m-1} \) and \( D_m = G C_{m-1} \), \( H \) is the low-pass filter. For each decomposition, the sampling of \( p_m f(t) \) is twice as sparse as the original, and the resolution gets coarser while the waveform gets smoother. \( G \) is the mirror high-pass filter of \( H \), and the bandwidth is also reduced by a factor of two each time. After \( m \) decompositions, the following is obtained:

\[
f(t) = p_m f(t) + \sum_{j=1}^{m} q_j f(t). \tag{20}
\]

The formula is the low-frequency global information of the function, and the second term is the corresponding local detail information \( f(t) \) on each scale from \( V_0 \) to \( V_{m-1} \), obtained by successive decompositions [32]. The aforementioned decompositions are depicted in Figure 3.

### 4. The Method of Prognostics and Evaluation of Bearing’s Service Life

The vibration signals are decomposed using the foregoing orthogonal wavelet transformation. The features of \( f(t) \)'s corresponding local signals on each scale are then extracted to serve as a feature vector. Finally, orthogonal wavelet transformation-support vector data description (OWTSVDD) is used for prognostics and evaluation of bearing’s service life, thereby achieving intelligent fault diagnosis. The specific steps are shown in Figure 4:

**Step 1.** The orthogonal wavelet transformation at \( m \) scale is performed for bearing’s original normal signals to extract \( f(t) \)'s corresponding local detail signals at all scales from number \( V_0 \) to number \( V_{m-1} \). The analysis is shown in Figure 3, where \( V_0 \) is the original signal.

**Step 2.** Regarding the extraction of features of bearing’s normal signals in step 1, the feature vectors of peak-to-peak value at the first five scales are used. Assuming that the corresponding peak-to-peak value of \( W_i \) \( (i = 1, 2, \ldots, m - 1) \) is \( X_{pp}(i) = \max(W_i) - \min(W_i) \), \( (i = 1, 2, 3 \ldots, m - 1) \), the constructed feature vector \( T \) is as follows:

\[
T = [X_{pp}(1), X_{pp}(2), X_{pp}(3), \ldots, X_{pp}(m-1)]. \tag{21}
\]

**Step 3.** The feature vector constructed in step 2 is used to serve as SVDD training samples, thereby solving \( R \) and constructing the SVDD prognostics and evaluation model.

**Step 4.** A detailed analysis of signals is performed to test the samples using orthogonal wavelet transformation. The peak-
to-peak value at each scale is then extracted to serve as the test sample set, thus solving the distance $R_z$, from the test sample set to the center of the sphere. Based on $HI = R_z^2 - R^2$, its distance to the hypersphere is calculated to determine whether it is in normal conditions.

**Step 5.** Setting $HI = \varepsilon$, the greater the $\varepsilon$, the further the deviation from normal conditions and the severer the degeneration.

**5. Experimental Analysis**

**5.1. Experimental Conditions.** The bearing's fault data is analyzed on the rolling bearing test bench and the experimental data is from the electrical engineering laboratory of US Case Western Reserve University [33]. It aims to verify the validity of the OWTSVDD bearing performance evaluation model. The bearing is a 6205-2RS deep groove ball bearing, installed on the drive end on the right side of the motor to support the motor shaft. The rolling bearing's multichannel fault signals are collected synchronously by 3 acceleration sensors whose installation locations are depicted in Figure 5. Under normal conditions, the sampled frequency of bearing was 12 kHz [34]. When the rotor load is 1Hp, the fault samples of the inner bearing race’s pitting at different depths of 0.1778, 0.3556, 0.5334, and 0.7112 mm were collected, respectively. 50 sets were collected under the foregoing four circumstances and the sampling length of each set of data was 512. First, under the condition of no OWT feature extraction, the prognostics and evaluation of performance degeneration were conducted for bearings with different fault levels longitudinally (the development from minor faults to major faults was considered as an overall process) with support vector data description. 50 sets of samples were collected from bearings in normal conditions to serve as the training samples of the SVDD model. The fault samples collected in the foregoing four circumstances were tested and the observed changes of HI are illustrated in Figure 6.

The result was observed for evaluation of equipment's conditions using SVDD without feature extraction. The test result HI of the first 50 sample points was observed when the pitting was 0.1778 mm: the deviation of fault from the equilibrium position was about 1.4, indicating a relatively smaller HI amplitude. For sample points of 51–100, the deviation reached about 0.8 when the pitting was 0.3556 mm, indicating a greater deviation of HI from the equilibrium position. For points of 101–150, the deviation returned to about 1.5 when the pitting was 0.5334 mm, the change of amplitude was negligible. For sample points of 151–2000, the deviation of samples from the equilibrium position was 1.5, which remained unchanged. From the foregoing changes of HI, it failed to assess its relationship with the pitting fault of bearings. That is, the HI value cannot reflect the degeneration of bearing’s performance.

Moreover, the evaluation and analysis were subsequently performed with OWTSVDD bearing performance evaluation model. Under the same experimental conditions, 50 sets of the inner bearing race’s pitting at depths of 0.1778, 0.3556, 0.5334 and 0.7112 mm were collected, respectively. 512 points were collected for each data set. The 50 sets of samples in normal conditions were taken. After feature extraction through orthogonal wavelet transformation, the vectors served as the training samples of the OWTSVDD bearing performance evaluation model. The test of samples was subsequently conducted after feature extraction of pitting signals in foregoing four circumstances through orthogonal wavelet transformation, thus yielding the changes of HI as depicted in Figure 7.

As inferred from Figure 7, starting from the 0.1778 mm pitting of bearing’s inner race, for 0–50 sample points, the HI is constantly at an amplitude with about 0.75 deviation from the normal state; for 51–100 sample points, when the fault is 0.3556 mm pitting, the amplitude becomes instable and the HI starts to deviate from the equilibrium position of 0.75 and expands dramatically. For 101–150 sample points, when the fault is 0.5334 mm pitting, the amplitude experienced accelerated changes and the bearing’s performance started to degenerate. For 151–200 sample points, when the fault is 0.7112 mm pitting, the bearing’s performance further degenerated. When pitting exceeds 0.7112 mm, the bearing’s performance is even more degenerated. It can be inferred that HI is directly proportional to the exacerbated degeneration of the bearing’s pitting fault. In other words, HI significantly reflects the degeneration of bearing’s performance.

It can be concluded from the comparison between Figures 6 and 7 that, the OWTSVDD bearing performance evaluation model has been established with the feature vectors extracted through orthogonal wavelet transformation is more effective in evaluating bearing’s performance. As calculated, the health evaluation parameter HI is of more definite significance and better monitors the degeneration of bearings. From Figures 8 and 9 of decomposition of 0.3556 and 0.5334 mm signals through orthogonal wavelet transformation, it also can be seen that the regularity of d5 and d4 higher scales experienced an abrupt change from 0.3556 to 0.5334 mm.

**5.2. Collection of Different Experimental Data.** To further verify the OWTSVDD bearing performance evaluation model, verification and analysis were performed for the bearing’s service life evaluation model based on different experimental data of rolling bearing. The data was sourced from the XJTU-ST bearing [35]. According to the relevant description [36–38], used in the study of bearing degradation, the horizontal vibration signals provide more significant information than vertical vibration signals. Therefore, the experiment adopted horizontal vibration signals, and the test bench is depicted in Figure 10. The experiment was started at a fixed speed and the accelerated degeneration test of rolling bearing was carried out under different working conditions. When the maximum amplitude of a horizontal or vertical signal exceeded $A^{10}$, it was considered that the bearing became invalid and the service life test was terminated, whereby it was the maximum amplitude of a vertical vibration signal under normal working conditions.
Figure 5: Rolling bearing test bench.

Figure 6: Evaluation chart of OWT without feature extraction.

Figure 7: Evaluation chart of feature extraction through orthogonal wavelet transformation.
the experiment, the tested bearing may incur numerous faults such as outer race breakage, outer race fault, inner race fault, and rolling part fault. The model of test rolling was LDK UER204. The acceleration signals were collected in continuous windows, with a sampling frequency of 25.6 kHz. At each sampling time, a total of 32768 data points (namely, 1.28 s) were recorded. The sampling cycle was 1 min and the data (Table 1) representing 3 different loads were considered.
In this paper, the first set of data under condition 1 was adopted as the training set; namely, the life-cycle data of bearing 1_5.

5.3. Comparison of Performance Evaluation Based on Different Experimental Data. In the service life test of rolling bearing, the features were extracted using an orthogonal wavelet transformation. The db10 wavelet packet was selected and decomposed into 5 layers. The peak-to-peak feature vectors of the 5th and 4th layers served as the input vectors of the OWTSVDD bearing performance evaluation model. The set of feature vectors with this structure was subsequently used for training the OWTSVDD bearing performance evaluation model. With rolling bearing 1_5 as an example, the sampling frequency was 25.6 kHz; the sampling interval was 1 min; the actual service life was 52 min. Figure 11 illustrates the life-cycle vibration signals of bearing 1_5. At the 35 min, the vibration amplitude of rolling bearing increased significantly compared to normal standard amplitude. This time point marked the start of the degeneration of the bearing and the prognostics of the bearing’s RUL.

50 sets of bearing data were collected at 26, 29, 32, 35, 38, and 41 min, respectively. First, OWT feature extraction, the prognostics and evaluation of performance degeneration were conducted for bearings with different fault levels longitudinally (the development from minor faults to major faults was considered as an overall process) with SVDD bearing performance evaluation model. 50 sample sets were collected from bearings at 2 min to serve as training samples of the SVDD model. The fault samples collected in foregoing four circumstances were subsequently tested, with observed changes of HI depicted in Figure 12.

The result was observed for SVDD bearing performance evaluation model without feature extraction. The test result HI of the first 50 sample points was observed for the bearing’s data at 26 min: the fault was about 0.02 at the equilibrium position and the HI amplitude was negligible. For sample points of 51–100, the over HI was about 0.02 when the bearing had been running for 29 mins; at the 74th sample point, HI significantly deviated from the equilibrium position. For sample points of 101–150, HI changes negligibly around 0.02 when the bearing had been running for 35 mins. For sample points of 151–300, the samples deviated from the equilibrium position for 0.02 when the bearing had been running for 38 mins and 41 mins, which remained unchanged. From the foregoing changes of HI, it failed to figure out its relationship with the pitting fault of bearings. That is, the HI value cannot reflect the degeneration of bearing’s performance.

In the next, the evaluation and analysis were performed with OWTSVDD bearing performance evaluation model. Under the same experimental conditions, 50 sets of data at 26, 29, 32, 35, 38, and 41 min were collected, respectively. 512 points were collected for each data set. The OWT feature extraction, the prognostics and evaluation of performance degeneration were conducted for bearings with different fault levels longitudinally (the development from minor faults to major faults was considered as an overall process) with SVDD bearing performance evaluation model.
classification model. The test of samples was then conducted after feature extraction of pitting signals in the foregoing four circumstances through the orthogonal wavelet transformation, thereby yielding the changes of HI as depicted in Figure 13.

In Figure 13, starting from the data of bearing at 26, 29 and 32 min, it is observed that HI is constantly at a stable amplitude with a deviation around the normal state 0 for 0–150 sample points. A significant amplitude is noted from the 74th sample point; after the 100 sample points, the amplitude starts to drop. For 151–200 sample points, the amplitude becomes unstable according to the data of bearing at 35 min and HI starts to deviate from the equilibrium position of 0.25 following a trend of dramatic expansion. For 201–250 sample points, the amplitude experiences an abrupt acceleration according to the data of bearing at 38 min and the bearing performance degenerates. For 251–300 sample points, the bearing’s performance further degenerates according to the fault data at 41 min. After 42 min, the degeneration of the bearing’s performance is exacerbated. It
is noticed that HI is increased with the aggravation of the bearing’s pitting fault. Namely, HI significantly reflects the degeneration of bearing’s performance. It can be learned from the comparison between Figures 12 and 13 that, the OWTSVDD bearing performance evaluation model established with the feature vectors extracted through orthogonal wavelet transformation is more effective in evaluating the bearing’s performance. V&hehealthevaluationparameter HI as calculated is of more definite significance. Furthermore, it is consistent with the life-cycle vibration signals at 35 min in Figure 11, proving that the OWTSVDD bearing performance evaluation model better monitors the degeneration of bearings. After decomposition of 35 and 38 min signals through orthogonal wavelet transformation, it has been concluded from the comparison of Figures 14 and 15 that the regularity of d5 and d4 at higher scales obviously worsened.

Deep learning methods: the neural network analyzes the original data, and the original data takes orthogonal wavelet transform as the input value of the neural network, among which 80% of the data is used as training samples and 20% of the data is used as test samples. During training, the maximum training is set to 1000 times, and the training results of the neural network are shown in the transverse represent the training times and the ordinate represents the mean square error in Figure 16. When the error of four training times is 0.022051, when the error has converged to a

![Figure 13: Evaluation chart of feature extraction through orthogonal wavelet transformation.](image)

![Figure 14: Wavelet decomposition chart of bearing’s inner race at 35 min.](image)
Figure 15: Wavelet decomposition chart of bearing’s inner ring at 38 min.

Best Validation Performance is 0.022051 at epoch 4

Figure 16: Best validation performance is 0.022051 at epoch 4.
small error, and the establishment of the neural network model is completed. The 30 sets of data at 30 s are selected as the validation sample set. The sample set data is input into the trained BP neural network, and each sample corresponds to a certain fault model. The actual output results of the neural network are compared with the standard signal. The comparison results are shown in the following Table 2.

### Table 2: Comparison of the standard signal and the actual signal feature vectors.

<table>
<thead>
<tr>
<th>Standard signal</th>
<th>Reality signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8962</td>
<td>0.8908</td>
</tr>
<tr>
<td>0.8528</td>
<td>0.8600</td>
</tr>
<tr>
<td>0.8080</td>
<td>0.7878</td>
</tr>
<tr>
<td>0.6510</td>
<td>0.8968</td>
</tr>
<tr>
<td>0.8788</td>
<td>0.8806</td>
</tr>
<tr>
<td>0.8320</td>
<td>0.8785</td>
</tr>
<tr>
<td>0.8711</td>
<td>0.8884</td>
</tr>
<tr>
<td>0.8250</td>
<td>0.8171</td>
</tr>
<tr>
<td>0.8094</td>
<td>0.7928</td>
</tr>
</tbody>
</table>

6. Conclusion

In this paper, an OWTSVDD bearing performance evaluation method has been proposed through the combination of orthogonal wavelet transformation (OWT) and support vector data description (SVDD), where the quantitative health index (HI) for conditions evaluation is given. In the application of the OWTSVDD evaluation method in fault diagnosis, the peak-to-peak value of each detail is extracted through orthogonal wavelets to serve as the training samples of SVDD, thereby solving the distance $R_z$ from the test sample to the center of the sphere. Based on $HI = R^2_z - R^2$, its distance to the hypersphere is calculated to evaluate and predict the bearing fault. Compared with SVDD devoid of feature extraction, this method effectively evaluates the conditions of bearings and is consistent with the life test chart of bearings. Furthermore, the validity of the method is also proved by different experimental data.

The conclusion of this paper is as follows:

1. For bearing's experimental data without feature extraction through orthogonal wavelet transformation, the relationship between changes of HI and that of the bearing's pitting fault cannot be observed and HI is not reflective of degeneration in the bearing's performance.

2. By comparing the two sets of experimental data, the validity of OWTSVDD in fault diagnosis and prognostics has been proved. In the analysis of the two sets of experimental data, OWTSVDD is entirely able to evaluate the entire life cycle of bearings from normal operation to fault and degradation.

3. The experimental data in Xi’an Jiaotong University is consistent with the bearing’s life-cycle vibration signals. When OWTSVDD is used to analyze the experimental data of Xi’an Jiaotong University, the fast expansion in bearing’s fault and performance evaluation at the 35 min is 100% consistent with the expansion of bearing 1_5’s fault at the 35 min in the life-cycle test chart.

4. This method addresses the issue with the bearing’s performance evaluation with small sample size. During the data analysis in the electrical engineering laboratory of US Western Reserve University, under the condition of limited samples, the desired signals were extracted through orthogonal wavelet transformation to serve as input vectors of OWTSVDD and the evaluation effect is 99% consistent with the evaluation effect based on the experimental data in Xi’an Jiaotong University.

5. This index is sensitive to reflect the degeneration of bearing’s performance, providing a scientific basis for production and equipment management and improving the prognostics technology-centered prognostics and health management (PHM).

6. The orthogonal wavelet decomposition chart also verified poorer regularity. It can be inferred that the regularity of $d_5$ and $d_4$ higher scales became markedly poor from 35th to the 38th min, indicating a significant degeneration of bearing’s performance.

OWTSVDD bearing performance evaluation method has been widely applied to online fault diagnosis, artificial intelligence and data processing, etc., markedly boosting the efficiency of fault diagnosis and data classification. In the future, the study will concentrate on the improvement of fault diagnosis technologies based on analysis of vibration signals, acoustic signals, thermal signals and others. The mechanical fault diagnosis method will be further optimized based on analyzing various types of faults and machines.

Data Availability

The data used to support this study can be found in the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Li Weipeng performed the numerical simulation. Li Weipeng and Cao Yan analyzed the theoretical study, analyzed the results of numerical simulation, and prepared the manuscript. Li Lijuan, Hou Siyu reviewed the manuscript.

Acknowledgments

This work was supported by a grant from Shaanxi Province Innovation Capacity Support Program (2018TD-036) and a grant from Shaanxi Province Key R&D Program (2019GY-125).
References


