Degradation state identification for hydraulic pumps using modified hierarchical decomposition and image processing

Mo-chao Pei¹, Hong-ru Li¹ © and He Yu¹

Abstract
Monitoring the degradation state of hydraulic pumps is of great significance to the safe and stable operation of equipment. As an important step, feature extraction has always been challenging. The non-stationary and nonlinear characteristics of vibration signals are likely to weaken the performance of traditional features. The two-dimensional image representation of vibration signals can provide more information for feature extraction, but it is challenging to obtain sufficient information based on small-size images. To solve these problems, a method for feature extraction based on modified hierarchical decomposition (MHD) and image processing is proposed in this paper. First, a set of signals decomposed by MHD are converted into gray-scale images. Second, features from accelerated segment test (FAST) algorithm are applied to detecting the feature points of the gray-scale image. Third, the real part of Gabor filter bank is used to convolve the images, and the responses of feature points are used to calculate histograms that are regarded as feature vectors. The method for feature extraction fully acquires the multi-layered texture information of small-size images and removes the redundant information. Furthermore, support vector machine (SVM) and nondominated sorting genetic algorithm II (NSGA-II) are introduced to conduct feature selection and state identification. NSGA-II and SVM can conduct the joint optimization of these two goals. The details of the proposed method are validated using experimental data, and the results show that the highest recognition rate of our proposed method can reach 100%. The results of the comparison among the proposed method, local binary pattern (LBP), and one-dimensional ternary patterns (1D-TPs) certify the superiorities of the proposed method. It obtains the highest classification accuracy (99.7%–98%) and the lowest feature set dimension (13–10).

Keywords
modified hierarchical decomposition, image processing, hydraulic pump, degradation state identification

Introduction
The heart of hydraulic system is hydraulic pump, which has the advantages of high pressure and high efficiency. As a component for transferring and converting energy, hydraulic pumps are widely used in the fields of transportation and industry.¹ Failure to troubleshoot hydraulic pumps in time will threaten the safe operation and reliability of the hydraulic system and even cause terrible accidents and economic losses.² Therefore, degradation state identification for hydraulic pumps is conducive to the reliable and long-term operation of the system. However, the structure of hydraulic pump is not as simple as other rotating machinery such as bearings. The thermal-solid-fluid coupling existing inside the hydraulic pump makes its fault information hidden.³ Consequently, a novel method needs to be proposed for feature extraction and applied to the degradation state identification for hydraulic pumps.

The traditional features of time–frequency domain, time domain, and frequency domain are widely used in the fields of fault diagnosis and prognosis. Applying frequency domain analysis and time domain analysis to non-stationary signals is not effective.⁴ Although the time–frequency analysis can be used to extract the features of a non-stationary signal more effectively, its stability will be weakened in the presence of a signal with weak fault characteristics and nonlinear characteristics.⁵ In recent years, some novel methods for feature extraction have been proposed in the fields of fault diagnosis and condition monitoring. Hong-ru Li et al.⁶ extracted the...
was proposed based on permutation entropy and MHD. Cheng et al. proposed a feature indicator called hierarchical entropy based on sample entropy and hierarchical decomposition. Since then, some scholars had combined the hierarchical decomposition with other complexity indicators and proposed some novel feature extraction methods. With the support of hierarchical decomposition, Bing Han et al. developed the single-scale Lempel-Ziv complexity into hierarchical Lempel-Ziv complexity and applied it to fault feature extraction for rotating machinery. In the process of hierarchical decomposition, the length of time series is shortened. Modified hierarchical decomposition (MHD) proposed by Yongbo Li et al. overcome this drawback. The moving-difference and moving-averaging procedure were used to improve hierarchical decomposition, and modified hierarchical permutation entropy was proposed based on permutation entropy and MHD. Cheng Yang et al. extended MHD to multiple scales, and proposed hierarchical multi-scale symbolic dynamic entropy, which is a tensor feature extraction method based on MHD, multi-scale analysis and symbol dynamic entropy. The above studies related to the hierarchical decomposition show that it has the advantage of obtaining the information of signals from high- and low-frequency components, so that it can reflect the fault information more accurately than traditional multi-scale analysis.

Inspired by the rapidly developing image processing technology, some scholars also applied it to extracting fault features. Kaplan et al. applied texture analysis to diagnosing bearing faults. The one-dimensional (1D) vibration signals were converted into two-dimensional (2D) gray-scale images. Local binary pattern (LBP), a classic texture descriptor, was used to extract image features. Hao Zheng et al. used a corner detection method called features from accelerated segment test (FAST) to identify the feature points in gray-scale images. They then used the unoriented scale-invariant feature transform (unoriented-SIFT) to describe the feature points. Melih Kuncan et al. developed the local ternary patterns for image processing into one-dimensional ternary patterns (1D-TP), which was used to extract bearing fault features from vibration signals. 1D-TP can be considered as a variant of image processing technology. Since deep learning can learn the deep abstract features of data and has a strong data expression, it has also been widely used in the field of mechanical fault diagnosis. Although deep learning can process 1D signals, images have been more widely used as its input. Jianyu Wang et al. proposed a bearing fault diagnosis model transferred from an AlexNet model. The images created by eight time-frequency analysis methods, including short-time Fourier transform, and fast kurtogram, were used as the input of the model. Yongliang Bai et al. proposed an algorithm to represent the frequency spectrum characteristics of vibration signals in image form, and a deep neural network (DNN) classified the images to diagnose the faults of wheelset-axlebox assemblies. Duy-Tang Hoang et al. converted the vibration signals into gray-scale images, and a deep convolutional model automatically extracted features from the images and recognized the bearing faults. Compared with 1D signal, 2D image representation has three main advantages: (1) For machinery working in a noisy environment, vibration signals are usually added with noise. However, when the data is transformed into images, the added noise is considered as the illumination of the light to the image. Therefore, the effect of noise is suppressed. (2) Image representation can provide a comprehensive, detailed, and nonlinear description of the data. (3) From the perspective of human vision, 2D images are of course easier to distinguish. With a simple transformation of the signal, people can more easily classify each signal from the images.

Although the application of image processing technology is novel and effective in the fields of fault diagnosis and prognosis, there are still several problems that need to be considered. First, an image often contains a large number of pixels in the field of image processing, while a fault sample is likely to contain a small number of data points in the field of fault diagnosis, which makes a 1D signal into a small-size image. It contains less information, and its local and global information needs to be considered comprehensively. Second, it is challenging to obtain sufficient information based on small-size images. Third, deep learning requires a large number of data samples to meet good recognition results. To extract the degradation features of hydraulic pumps more effectively, it is promising to combine hierarchical decomposition with image feature extraction. In this paper, an approach based on MHD, FAST algorithm, and Gabor filter is proposed to extract features. FAST is a feature point detection algorithm widely used in computer vision. The points with large difference in gray value of surrounding pixels are identified as feature points by FAST algorithm; thus, FAST is a local feature detector and suitable for signals with non-stationary characteristics. The role of the Gabor filter in this paper is to describe the detected feature points. As a powerful tool in the field of computer vision, the Gabor filter proposed by Dennis Gabor is one of the best approaches for texture analysis. Because of its high sensitivity to describing directional features, it is suitable for calculating the gradient of feature points. Wei Jia et al. used histogram of oriented gradients (HOG) descriptor improved by the Gabor filter to extract palmprint features. In this paper, a Gabor filter bank with one scale and several directions is applied to all feature points, and the response histogram is used as the feature vector of a gray-scale image.

In this study, after feature extraction, nondominated sorting genetic algorithm II (NSGA-II) and support vector machine (SVM) are introduced to conduct feature selection and identify different degradation states of a hydraulic pump. Overall, a new strategy based on MHD, FAST algorithm, Gabor filter, NSGA-II, and SVM is proposed to identify degradation states of hydraulic pumps. The results of two experiments show that the proposed method is effective and reasonable. Compared with LBP and 1D-TP, it can obtain the highest classification accuracy with the fewest features.

The following contents are also presented: The details of model establishment for the proposed degradation state identification method are presented in the second section. Experimental test results and discussion content are given in the third section. The fourth section summarizes the full paper and gives conclusions.

**Model establishment**

**Feature extraction**

*Modified hierarchical decomposition (MHD)*. Compared with the conventional hierarchical decomposition, MHD proposed
by Yongbo Li has a better stability.9 The calculation process of MHD includes four steps as follows:

Step 1: For a given time series \( Q\{q(i), i = 1, 2, \ldots, L\} \), an averaging operator \( O_0 \) and a high operator \( O_1 \) can be expressed according to equation (1) and equation (2)

\[
O_0(q) = \frac{q(i) + q(i + 1)}{2} \quad i = 1, 2, \ldots, L - 1
\]

\[
O_1(q) = \frac{q(i) - q(i + 1)}{2} \quad i = 1, 2, \ldots, L - 1
\]

Step 2: The operator \( O_j^h (j = 1 \text{ or } 0) \) at the hierarchical layer \( h \) can be calculated according to equation (3)

\[
O_j^h = \begin{bmatrix}
\frac{1}{2} & 0 & \cdots & 0 & \frac{(-1)^j}{2} & 0 & \cdots & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \cdots & 0 & \frac{(-1)^j}{2} & \cdots & 0 & 0 & 0 \\
\vdots & & & & & & & & & \\
0 & 0 & 0 & \cdots & 0 & \frac{(-1)^j}{2} & \cdots & 0 & 0 & 0 \\
\end{bmatrix}_{(2^{h-1})\times(2^{h-1})}
\]

Step 3: For a given positive integer \( h \) and unique vector \( [v_1, v_2, \ldots, v_h] \), the integer \( e \) is expressed according to equation (4)

\[
e = \sum_{m=1}^{h} 2^{h-m} v_m
\]

where \( \{v_m, m = 1, \ldots, h\} \in \{1, 0\} \) determines the selection of high or averaging operator at the \( m \)th layer.

Step 4: Calculate the component at the node \( e \) of the layer \( h \) in the hierarchical structure according to the following equation

\[
Q_{h,e} = O_0^h \cdot O_1^{h-1} \cdot \cdots \cdot O_1^1 \cdot O_0^1 \cdot Q
\]

To illustrate the MHD, an example is presented in Figure 1. The time series \( Q \) is decomposed into three layers, and the last layer has eight components. Take the case of \( h = 3 \) as an example, as the low-frequency component of \( Q_{2,1} \), \( Q_{3,2} \) is the component at node 2 of layer 3. When \( e = 2 \) and \( h=3 \), the unique vector \( [v_1, v_2, v_3] = [0, 1, 0] \) according to equation (4). Consequently, \( Q_{3,2} \) can be obtained as \( O_1^1 \cdot O_1^1 \cdot O_0^1 \cdot Q \) according to equation (5). \( Q_{3,3} \) is the high-frequency component of \( Q_{2,1} \).

**FAST feature detection**

As shown in Figure 2, a pixel \( p \) is selected from an image, and the calculation method of FAST is as follows: The gray value of \( p \) is represented as \( I_p \), and an appropriate threshold \( t \) is set. There are a series of pixels on a discretized circle with \( p \) as the center and \( r \) as the radius. For instance, there are 16 pixels on the circle with \( r = 3 \) in Figure 2. Any pixel on the circle is represented as \( z \). If the gray value of \( z \) (marked as \( I_z \)) is not greater than the value of \( I_p - t \), \( z \) is darker than \( p \), and the status of \( p \rightarrow z \) (expressed as \( S_{p \rightarrow z} \)) is \( d \). Similarly, the value of \( S_{p \rightarrow z} \) is calculated according to equation (6)

\[
S_{p \rightarrow z} = \begin{cases}
b, & I_p + t \leq I_z \quad \text{(brighter)} \\
s, & I_p - t < I_z < I_p + t \quad \text{(similar)} \\
d, & I_z \leq I_p - t \quad \text{(darker)}
\end{cases}
\]

Equation (6) is applied to all pixels on the circle, and two sets are defined as follows

\[
Y_b = \{z : S_{p \rightarrow z} = b\}
\]

\[
Y_d = \{z : S_{p \rightarrow z} = d\}
\]

The center \( p \) is identified as a feature point if the number of elements in \( Y_b \) or \( Y_d \) is greater than \( n \). The parameter \( n \) is assigned by the user.

**Gabor filter**

In a two-dimensional space, the Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave, which can be expressed as equation (9)

\[
G(x, y; \lambda, \theta, \sigma, \gamma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \exp\left(2\pi i\frac{x\lambda}{\lambda}\right)
\]

its real part is expressed as
\[ G(x, y; \lambda, \sigma, \gamma) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{x' + y'^2}{2\sigma^2} \right) \cos \left( \frac{2\pi x'}{\lambda} \right) \]

\[ \begin{align*}
    x' &= x \cos \theta + y \sin \theta \\
    y' &= -x \sin \theta + y \cos \theta
\end{align*} \]

where \( \hat{\lambda} = -1 \), \( \lambda \) is the wavelength of the sinusoidal function, \( \theta \) controls the direction of the parallel strips in the Gabor kernel, \( \sigma \) is the standard deviation of the Gaussian factor, and \( \gamma \) is the spatial aspect ratio, which controls the ellipticity of the Gabor function. If \( \gamma = 1 \), the image of the Gabor function is circular. A Gabor filter bank can be created by changing \( \theta \) and fixing other parameters. The angle \( \theta_c \) is calculated according to equation (12)

\[ \theta_c = \frac{\pi(c - 1)}{n'}, \quad c = 1, 2, \ldots, n' \]

The proposed feature extraction method

For a given signal segment \( Q\{q(i), i = 1, 2, \ldots, L\} \), the proposed feature extraction method consists of 6 steps. As seen in Figure 3, an example is shown there. In this case, \( Q \) is decomposed into three layers, and the Gabor filter bank has six different directions: \( (0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6) \).

Step 1: Apply MHD to \( Q \). Then, \((2^{hmax+1} - 1)\) hierarchical components can be obtained (\( Q \) is treated as a component), where \( h_{\text{max}} \) is the maximum value of the layer value \( k \), which is also the total number of layers (Figure 3(b)).

Step 2: Transform each hierarchical component into a gray-scale image. First of all, each component is converted into a matrix according to the method shown in Figure 4. When it comes to the size of the matrix, the difference between \( M \) and \( N \) is minimized. In other words, a square matrix is the best result. Note that several data points may be discarded. For example, for a given signal with 4088 data points, \((M, N)\) is set to \((68, 60)\), and 8 points are discarded. Then, all the elements of the matrix are converted to values between 0 and 255 according to equation (13)

\[ \text{New } a_i = \left[ \frac{a_i - \min(A)}{\max(A) - \min(A)} \right] \times 255 \]
converted to uint 8 type, and the gray-scale images are obtained (Figure 3(c)).

Step 3: FAST feature detection: The feature points in each gray-scale image are detected according to the method described in the subsection “FAST feature detection” (Figure 3(d)).

Step 4: A Gabor filter bank is created according to the method described in the subsection “Gabor filter”. Then, each gray-scale image is convolved with the real part of the filter bank. Finally, the gradient magnitude \( m_p \) and orientation \( \theta_p \) corresponding to each feature point are calculated according to equations (14) and (15), respectively

\[
m_p = \max [I_p \ast G_c(\theta_c)]
\]

\[
\theta_p = \arg \max_c [I_p \ast G_c(\theta_c)]
\]

where \( c = 1, 2, \ldots, n' \), \( G_c(\theta_c) \) is the real part of the filter bank obtained according to equations (10–12), \( I_p \) is the feature pixel, and \( \ast \) means the convolution operation (Figure 3(e)).

Step 5: Create \( n' \) bins with different \( \theta_c \) and calculate the histogram within each gray-scale image (HI) as follows (Figure 3(f))

\[
HI(\theta_c)_i = HI(\theta_c)_i + m_p \text{ if } \theta_p = \theta_c
\]

Step 6: The histogram corresponding to the signal segment \( Q(HS) \) can be obtained by integrating the histograms of \( (2^{2n_c} + 1) \) gray-scale images as follows (Figure 3(g))

\[
HS = \{HI_1, HI_2, \ldots, HI_{(2^{2n_c}+1)}\}
\]

It is obvious that \( HS \) can be considered as a \( n' \times (2^{2n_c}+1) \) dimensional vector. Feature selection and state recognition based on NSGA-II and SVM.

**Feature selection and state recognition based on NSGA-II and SVM**

After performing feature extraction, a feature pool that contains amounts of information is created. However, the feature pool generally includes redundant information, which will cause the performance of degradation state identification to be impaired. In this paper, NSGA-II and SVM are introduced for joint optimization of feature selection and degradation state identification.

**Nondominated sorting genetic algorithm-II**

NSGA is the earliest multi-objective evolutionary algorithm. NSGA-II is an improved version that addresses the shortcomings of NSGA, such as lack of elitism and high computational complexity. As an algorithm based on nondominated sorting, NSGA-II has the same basic process as the original genetic algorithm. With the properties of an elitist strategy, a fast nondominated sorting procedure, a simple yet efficient constraint-handling method and a parameterless approach, NSGA-II has found increasing application in the fields of fault prognosis and diagnosis for equipment. The procedure of NSGA-II is summarized as follows:

Step 1: Initialize a random population, which consists of \( N \) individuals.

Step 2: Sort the population based on the nondominated sorting procedure and assign ranks to all individuals.

Step 3: Utilize selection, crossover, and mutation operators to create the offspring population \( Q_t \) from the parent population \( P_t \).

Step 4: Combine the offspring population and the parent population to form a whole and sort the whole according to the nondominated sorting procedure.

![Figure 4. Illustration of converting a one-dimensional signal into a two-dimensional matrix.](image-url)
Step 5: Continuously select nondominated solutions from the best ranking \((F_1, F_2, \ldots, F_k, \ldots, F_n)\) to generate a new population \(P_{n+1}\). The selection is not stopped until the selection of a certain \(F_k\) makes the size of \(P_{n+1}\) exceed the population size \(N\). Then select individuals from \(F_k\) according to the crowding distance sorting until the size of \(P_{n+1}\) is equal to \(N\).

Step 6: Go to step 2 if the stop condition is still not met.

For a better understanding, Step 4 and Step 5 are illustrated in Figure 5. For other unfinished details about NSGA-II, refer to Deb et al.\textsuperscript{22}

**Support vector machine**

As one of the classic machine learning techniques, SVM is widely used in pattern recognition. It learns by example to assign labels to objects.\textsuperscript{30} SVM finds an optimal classification hyperplane so that the sum of the distances from this hyperplane to the nearest sample in each class is minimized. Take the binary classification problem as an example, for \(n\) linearly separable samples \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\), \(x \in \mathbb{R}^m\) (where \(m\) is the sample dimension, and \(y_1, y_2, \ldots, y_n\) are classification labels), Finding a hyperplane can be transformed into a solution to the following problem

\[
\begin{align*}
\max & \sum_{i=1}^{n} \lambda_{i} - \frac{1}{2} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{i}\lambda_{j}y_{i}y_{j}x_{i}x_{j} \right) \\
\text{s.t.} & \sum_{i=1}^{n} \lambda_{i}y_{i} = 0, \quad \lambda_{i} \geq 0
\end{align*}
\]  

(18)

As the only variable in problem (18), \(\lambda\) can be solved by quadratic programming. We assume that the optimal solution \(\lambda^* = (\lambda_{1}^*, \lambda_{2}^*, \ldots, \lambda_{n}^*)\). For the classification of nonlinear data, the commonly used method is to express the high-dimensional space through the nonlinear mapping function \(\phi\), and the classification hyperplane is constructed in a high-dimensional space. The kernel function that replaces the dot product operation is defined as

\[
K(x, x') = \phi(x) \phi(x')
\]  

(19)

Nonlinear data is mapped to a high-dimensional space, and the problem (18) can be transformed into the following expression

\[
\begin{align*}
\max & \sum_{i=1}^{n} \lambda_{i} - \frac{1}{2} \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{i}\lambda_{j}y_{i}y_{j}K(x_{i}x_{j}) \right] \\
\text{s.t.} & \sum_{i=1}^{n} \lambda_{i}y_{i} = 0, \quad 0 \leq \lambda_{i} \leq C
\end{align*}
\]  

(20)

where \(C\) is the penalty factor. Consequently, the classification function is expressed as equation (21)

\[
f(x) = \text{sgn} \left[ \sum_{i=1}^{n} \lambda_{i}^*y_{i}K(x_{i}, x) + b^* \right]
\]  

(21)

In this paper, radial basis function (RBF) is selected due to its good performance and universal application.\textsuperscript{31} and the RBF kernel of samples \(x\) and \(x'\) can be expressed as a feature vector defined as follows

\[
K(x, x') = \exp \left( -\frac{\|x - x'\|^2}{2\sigma^2} \right)
\]  

(22)

where \(1/2\sigma^2 = f\) and \(f\) is the kernel parameter.

The proposed method for feature selection and state recognition

NSGA-II has two objective functions, recognition accuracy and the total number of features fed into SVM, in this study. The recognition accuracy is the classification accuracy provided by SVM. The chromosomes of individuals in the population are encoded in binary form, “0” means the corresponding feature is not selected and “1” means it is selected. When NSGA-II runs, the algorithm maximizes the recognition accuracy and minimizes the total number of features simultaneously.

The multiple optimal solutions called “Pareto Frontiers” result from a multi-objective optimization problem.\textsuperscript{33} The Pareto Front with the highest classification accuracy will be selected. In other words, by performing this step, we can get the highest accuracy recognition result and its corresponding lowest-dimensional feature subset. The detailed process is shown in Figure 6.

**Suitable parameters determination**

There are several parameters in the algorithms described above that need to be determined: Hierarchical layer \(h\), radius \(r\) and two thresholds \((t, n)\) of FAST feature detection, the same \((\lambda, \sigma, \gamma)\) and different \((\theta, c, e = 1, 2, \ldots, n)\) of Gabor filter bank, the kernel parameter \(f\) of SVM, and several parameters of NSGA-II. Note that the size of the filter kernel needs to be determined before using a Gabor filter bank. For the selection of \(h\), insufficient decomposition will be caused by too small \(h\), and time-consuming will be caused by too large \(h\). We set \(h = 3\). In our code, the parameter \(f\) of SVM is set to “auto,” which enables \(f\) to be adaptively changed according to the change of feature subset during the operation of NSGA-II. That is, \(f = 1/ N_f\), where \(N_f\) is the number of features. According to the recommendation in Rosten et al.,\textsuperscript{34} we choose \(n = 12\) and \(r = 4\). The more the
angle parameters $\theta_i$ we set, the more the detailed description of the image we obtain. However, this also means more time-consuming. Furthermore, when the number of angles changes in the interval $[4, 6]$, it has little effect on the classification accuracy. Consequently, we take the empirical value as $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) = (0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6)$. According to the recommendation in Zhang et al., we choose $\gamma = 1$. When it comes to the parameters of NSGA-II, they are set to the following relatively balanced values: the total number of evolutionary generations $= 500$, crossover probability $= 0.6$, mutation probability $= 0.6$, and population size $= 30$. Compared with other unset parameters, they have weaker effects on the recognition accuracy. Then, there are four more important parameters, $t, \lambda, \sigma$, and Gabor kernel size (Ksize), that need to be fixed. In fact, the relationship between $\lambda$ and $\sigma$ can be expressed as follows:

$$\frac{\lambda}{\sigma} = \frac{2^b - 1}{2^b + 1} \sqrt{\frac{2}{\ln 2}}$$  \hspace{1cm} (23)

$$\lambda = \frac{2(\sigma + (\sqrt{\ln 2/\pi}))}{\sigma}$$  \hspace{1cm} (24)

where $b$ is the half-response spatial frequency bandwidth of the filter. Specifically, inspired by Bianconi et al., we use equation (24) to express the relationship between $\lambda$ and $\sigma$.

Therefore, we still need to assign values to three independent parameters. The four parameters are selected as follows based on experimental results: $t = 3, \lambda = 2.13, \sigma = 4$, and Ksize $= (11 \times 11)$. The specific details are presented in the section “Experimental validation.”

The proposed degradation state identification method for hydraulic pump

The proposed method consists of the following four steps:

Step 1: Acquire the vibration signals for various degradation states of the hydraulic pump and divide them into several samples.

Step 2: Utilize the method described in subsection 2.1.4 to extract features from each sample. 90 features will be acquired according to the formula $n \times (2^{b_{max} + 1} - 1)$.

Step 3: Integrate the features of all samples into a new dataset, and partition it into a testing dataset and a training dataset with a ratio of 1:1.

Step 4: Utilize the method described in 2.2.3 to identify the different degradation states of the hydraulic pump, and the best feature subset is also obtained.

Experimental validation

A hydraulic pump test platform shown in Figure 7 was set up for data acquisition, which consists of a cooling system, a control system, a signal monitoring, acquisition and display system, a pressure regulating system, and a drive system.
The hydraulic pump for this study is an axial piston pump with the following parameters: type: L10VS028DFR, displacement at the rated working condition: 28 ml/r, rated pressure: 22 MPa, and rated rotation speed: 1480 r/min. Three acceleration sensors are, respectively, installed in three mutually orthogonal directions as Figure 8. They acquire vibration signals at a sampling frequency of 50 KHz, each sampling lasts for 1s, and the interval between two samplings is 30 s.

The single loose boot is studied in this paper because the loose boot is one of the common fault patterns of hydraulic pump. The gap between the plunger and the boot will increase when the loose boot occurs. To acquire vibration signals close to the actual situation, the normal plungers are replaced with failed plungers obtained after equipment maintenance. Five different degrees of the loose boot, as shown in Figure 9, are considered. Vernier caliper is used to measure the maximum radial distance between the boot and the plunger under the five different degrees. The five measurements considered as the loose degree are 0.12 mm, 0.18 mm, 0.3 mm, 0.42 mm, and 0.64 mm, respectively. Including the normal state which is considered as a special degradation state, a total of six different degradation states are considered. The data of each degradation state is divided into 100 samples, and each sample consists of 4095 data points. The time domain waveforms of several samples are shown in Figure 10. Furthermore, one sample of each degradation state is decomposed into two sub-signals using MHD, and their time domain waveforms and gray-scale image representations are shown in Figure 11. When the figure is analyzed, it is difficult to separate the degradation states from each other using the time domain waveforms. However, all images appear to be differentiated from each other. Therefore, the conclusion that the image representation of the signal is useful is reinforced.

**Experiment 1: Selection of parameters \( t, \lambda, \sigma, \) and **Ksize**

In this experiment, 600 data samples are directly converted into 600 images for testing. The threshold \( t \) affects the number of detected feature points. One sample of each degradation state is taken to analyze the relationship between \( t \) and the number of feature points, as shown in Figure 12. It can be seen that the feature points decrease with the increase of \( t \). Too few feature points lose too much information and too many feature points lead to time-consuming and acquisition of useless information. Therefore, our goal is to choose a greater \( t \) while guaranteeing the recognition accuracy. In addition, the number of feature points in each state is less than 200 except for the loose degree 0.64 mm when the \( t \) value is greater than 20, and there are less than 100 feature points in the normal state, which will lose too much information. Therefore, the \( t \) value should be selected in the interval [1, 20].

The Ksize value should be generally an odd number so that the filter kernel is centered on a pixel. As the Ksize value increases, neighbors in a larger neighborhood of a pixel participate in the convolution operation on the pixel. Some unrelated neighbors may participate if the Ksize value is too great. Likewise, some useful neighbors may be discarded if the Ksize value is too small.

![Figure 8. Layout of the three vibration sensors.](image)

![Figure 9. Five different degrees of loose boot: (a) 0.12 mm; (b) 0.18 mm; (c) 0.3 mm; (d) 0.42 mm; and (e) 0.64 mm.](image)
Figure 10. Time domain waveforms of several vibration signal samples.

Figure 11. Time domain waveforms and image representations of the sub-signals: (a) normal; (b) loose degree: 0.12 mm; (c) loose degree: 0.18 mm; (d) loose degree: 0.3 mm; (e) loose degree: 0.42 mm; and (f) loose degree: 0.64 mm.
The parameter $\sigma$ should match the texture boundary. The window overlap different texture regions if $\sigma$ is too great. However, a too small $\sigma$ causes window-position perturbations and makes the output unstable.\(^{39}\)

To verify the relationship between these three parameters and classification accuracy, we assign values to them for testing. The result is shown in Figure 13. It can be seen that the distribution of scattered points is not messy but regular. The scattered points with high classification accuracy are clustered together. The following phenomena have been noticed: (1) When the value changes in a range greater than 10, Ksize has little effect on accuracy. The scattered points with high classification accuracy are basically in this range. (2) The scattered points with high classification accuracy are mainly gathered in the open interval $\sigma = (2.5, 5)$. (3) The ideal value of $t$ is in the interval $[1, 7.5]$.

In order to finally determine the values of the parameters, we set the value ranges of the three parameters as $\sigma = [2, 6]$, $t = [1, 7]$, and Ksize = $[11, 16]$; their incremental step sizes are all 1. The test result is visualized as shown in Figure 14. The two points with the highest classification accuracy can be easily found (surrounded by red dotted boxes). Therefore, the ideal parameter settings are as: $t = 3$, $\sigma = 4$, Ksize = $(11 \times 11)$, and $\lambda = 2.13$ according to the equation (24).

**Experiment 2: Verification for the necessity of modified hierarchical decomposition**

The following four methods are used to identify the degradation states of the hydraulic pump.

Method 1: The steps of Method 1 are the same as those described in subsection 2.4 except that the MHD is not...
applied. That is, only the 6 features of the original signal are extracted.

Method 2: The HOLGabor descriptor described in Jia et al.21 is modified with FAST feature detection and used for feature extraction. The gradients of the detected feature points instead of all pixels are calculated. The gradient magnitudes of the detected feature points in the cell are used to vote into the orientation histogram. The other steps of this method are the same as those of Method 1. According to the results of multiple tests, cell size = \((30 \times 30)\) pixels and block size = \((2 \times 2)\) cells are selected in this method. The image corresponding to each sample is divided into 4 cells, and 24 features are acquired.

Method 3: The proposed method described in subsection 2.4. 90 features are acquired.

Method 4: Apply Method 2 after applying the MHD to the original signal. 360 features are acquired.

From the above description of the four methods, it can be seen that the main differences between Method 1 and Method 2 are as follows: The image window in Method 2 is divided into small spatial regions called “cells,” and several adjacent “cells” are merged as a “block.” However, the image window in Method 1 is not divided into cells. Each method is implemented 10 times to reduce random effects. The highest classification accuracy obtained in each run is shown in Figure 15. Table 1 gives the statistical results of accuracy.

The following information is drawn from Figure 10 and Table 1: The classification accuracy obtained by Method 2 is higher than that obtained by Method 1. What is more noticeable is that the classification accuracy obtained by the proposed method is significantly higher than that obtained by Method 4. The proposed method achieves the highest classification accuracy (100–98.7%). There are 3 reasons for the result. First, in Method 2, more detailed information can be obtained since the image is divided into cells. However, each sample only consists of 4095 data points, which results in a smaller gray-scale image. Therefore, the contribution of dividing the image into cells to the classification accuracy is limited. Second, dividing the image into cells and applying the MHD together will improve the classification accuracy, but it will also produce more redundant information. Third, for samples with a small number of data points, not dividing the image into cells based on MHD and FAST feature detection is more beneficial to obtain useful information.

### Table 1. Classification accuracy statistics of the 4 methods in experiment 2.

| Method | Max   | Min   | Mean  |
|--------|-------|-------|-------|
| 1      | 0.88  | 0.867 | 0.876 |
| 2      | 0.9   | 0.88  | 0.888 |
| 3      | 1     | 0.987 | 0.997 |
| 4      | 0.982 | 0.967 | 0.977 |

**Experiment 3: Comparison among the proposed method, LBP, and 1D-TP**

As mentioned in Kaplan et al.,11 when LBP is applied to texture feature extraction, all non-uniform patterns are assigned a bin in the histogram, and each uniform pattern is assigned a bin. We set 8 neighbors for the central pixel. In this case, 59 features are acquired. We set 8 neighbors for the central point, threshold \(\beta = 4.5\) in 1D-TP. In this case, two feature sets, 256 low features and 256 up features, are obtained, respectively.13 For a fair comparison, the four feature sets obtained by the three methods should have the same dimension. In this experiment, to reduce the dimensionality of feature sets generated by the other two methods except for LBP, principal component analysis (PCA)40 is introduced. So that all feature sets are 59-dimensional. To facilitate the comparisons of the identification results, the method for feature selection and state recognition described in subsection 2.2.3 is also used after LBP (simplified into LBP-NSGA-II-SVM) and 1D-TP (simplified into 1D-TP-NSGA-II-SVM), respectively. 15 run times are conducted to reduce random effects. The classification accuracy and the corresponding number of features obtained in each run are shown in Figure 16. Table 2 gives the statistical results of accuracy, and the minimum number of features corresponding to the highest accuracy and the lowest accuracy are also shown in parentheses after them. First of all, the proposed method achieves the highest identification accuracy (99.7–98%). Second, it is more evident that the proposed method requires the least features.
Third, other methods get lower identification accuracy (99.3–95.8%), and they require more features (20–13). Therefore, the superiority of the proposed method is reinforced.

There are three reasons for the results. First of all, the proposed method for feature extraction combines the merits of FAST and Gabor filter. It can comprehensively reflect the local and global information of images. Second, the MHD enhances the integrity of the mined information. Third, in contrast, LBP and 1D-TP are insufficient in reflecting the information of signals.

Summary

To effectively extract the degradation features of hydraulic pumps, a novel method based on MHD and image feature extraction is proposed. The original signals are decomposed and transformed into gray-scale images. FAST algorithm and Gabor filter are applied to the images, and the histograms constructed by the filter response of the feature points are considered as feature vectors. They can fully reflect the local and global information of small-size samples. The effectiveness of the novel method and its superiorities in degradation feature extraction are validated using the experimental data. Furthermore, a new strategy for identifying the degradation states of hydraulic pumps was proposed based on our feature extraction method, NSGA-II and SVM. The experimental results show that the six different degradation states of the hydraulic pump can be successfully identified by the proposed method. This paper has 4 major contributions as follows:

1. The combination of FAST feature detection and Gabor filters can effectively obtain local and global information of the image.
2. MHD was introduced to increase the ability to obtain information.
3. A novel method based on our feature extraction method, NSGA-II and SVM was proposed, which was applied to degradation state identification for hydraulic pumps.
4. The superiority of the method proposed in this paper was demonstrated with experimental signals.

In this study, the experimental data comes from a hydraulic pump with prefabricated faults. The purpose of the experiment is to verify the effectiveness of the proposed method in identifying the degradation state of the hydraulic pump. When it comes to condition monitoring of an actual hydraulic pump, the real-time vibration data of the hydraulic pump is acquired first. After dividing the data into some samples, the feature representations are calculated according to the feature extraction method proposed in this paper. Vibration data of the prefabricated faults has extracted features in advance. In this way, two sample sets are obtained. The sample set belonging to the actual hydraulic pump is considered as the testing set. The other is considered as the training set. These two sample sets are fed to the NSGA-II-SVM proposed in this paper to conduct feature selection and state recognition, and the degradation state of the actual hydraulic pump can be identified. As the identification continues, the training set can be continuously updated, which is conducive to the subsequent identification.

In previous studies, some effective feature indicators were proposed by us. Fusing the degradation features proposed in this paper with them is the focus of further studies.

Acknowledgments

The authors gratefully acknowledge the support of National Natural Science Foundation of China (Grant No. 51275524).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project is supported by National Natural Science Foundation of China (Grant No. 51275524).

**ORCID iDs**

Hong-ru Li  [https://orcid.org/0000-0003-1312-5892](https://orcid.org/0000-0003-1312-5892)

**References**

1. Ma Z, Wang S, Zhang C, et al. A load sequence design method for hydraulic piston pump based on time-related markov matrix. *IEEE Trans Reliab* 2018; 67: 1237–1248.
2. Yu H and Li H. A conditional factor VAE model for pump degradation assessment under varying conditions. *Appl Soft Comput* 2021; 100: 106992.
3. Yu H, Li H and Li Y. Vibration signal fusion using improved empirical wavelet transform and variance contribution rate for weak fault detection of hydraulic pumps. *ISA Trans* 2020; 107: 385–401.
4. Chen X, Cheng G, Li Y, et al. Fault diagnosis of planetary gear based on entropy feature fusion of DTCWT and OKFDA. *J Vib Control* 2017; 24(21): 5044–5061.
5. Kumar A and Kumar R. Time-frequency analysis and support vector machine in automatic detection of defect from vibration signal of centrifugal pump. *Measurement* 2017; 108: 119–133.
6. Li H, Sun J and Tian Z. A novel method based upon modified composite spectrum and relative entropy for degradation feature extraction of hydraulic pump. *Mech Syst Signal Process* 2019; 114: 399–412.
7. Jiang Y, Peng CK and Xu Y. Hierarchical entropy analysis for biological signals. *J Comput Appl Math* 2011; 236: 728–742.
8. Han B, Wang S, Zhu Q, et al. Intelligent fault diagnosis of rotating machinery using hierarchical lempel-ziv complexity. *Appl Sci* 2020; 10: 4221.
9. Li Y, Li G, Yang Y, et al. A fault diagnosis scheme for planetary gearboxes using adaptive multi-scale morphology filter and modified hierarchical permutation entropy. *Mech Syst Signal Process* 2018; 105: 319–337.
10. Yang C and Jia M. Health condition identification for rolling bearing based on hierarchical multiscale symbolic dynamic entropy and least squares support sensor tree–based binary tree. *Struct Health Monit* 2020; 20: 151–172.
11. Kaplan K, Kaya Y, Kuncan M, et al. An improved feature extraction method using texture analysis with LBP for bearing fault diagnosis. *Appl Soft Comput* 2020; 87: 106019.
12. Zheng H, Cheng G, Li Y, et al. A new fault diagnosis method for planetary gear based on image feature extraction and bag-of-words model. *Measurement* 2019; 145: 1–13.
13. Kuncan M, Kaplan K, Mi Naz MR, et al. A novel feature extraction method for bearing fault classification with one dimensional ternary patterns. *ISA Trans* 2020; 100: 346–357.
14. Wang J, Mo Z, Zhang H, et al. A deep learning method for bearing fault diagnosis based on time-frequency image. *IEEE Access* 2019; 7: 42373–42383.
15. Yb A, Jya B, Jw B, et al. Image representation of vibration signals and its application in intelligent compound fault diagnosis in railway vehicle wheelset-axlebox assemblies. *Mech Syst Signal Process* 2021; 152: 107421.
36. Zhang Y, Li W, Zhang L, et al. Adaptive learning Gabor filter for finger-vein recognition. IEEE Access 2019; 7: 159821–159830.

37. Wu J, Wei P, Yuan X, et al. A new gabor filter-based method for automatic recognition of hatched residential areas. IEEE Access 2019; 7: 40649–40662.

38. Chen L, Lu G and Zhang D. Effects of different Gabor filters parameters on image retrieval by texture. In: 10th International Multimedia Modelling Conference, Brisbane, Australia, 5-7 January 2004, 2004, pp.273–278. New York: IEEE.

39. Dunn D and Higgins WE. Optimal Gabor filters for texture segmentation. IEEE Trans Image Processing 1995; 4: 947–964.

40. Abdi H and Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat 2010; 2: 433–459.