Motion Sequence Decomposition-Based Hybrid Entropy Feature and Its Application to Fault Diagnosis of a High-Speed Automatic Mechanism

Baoxiang Wang, Hongxia Pan * and Heng Du

School of Mechanical and Power Engineering, North University of China, Taiyuan 030051, China; wangbx1990@yahoo.com (B.W.); 652956431@163.com (H.D.)
* Correspondence: hxpan@nuc.edu.cn

Abstract: High-speed automatic weapons play an important role in the field of national defense. However, current research on reliability analysis of automaton principally relies on simulations due to the fact that experimental data are difficult to collect in real life. Different from rotating machinery, a high-speed automaton needs to accomplish complex motion consisting of a series of impacts. In addition to strong noise, the impacts generated by different components of the automaton will interfere with each other. There is no effective approach to cope with this in the fault diagnosis of automatic mechanisms. This paper proposes a motion sequence decomposition approach combining modern signal processing techniques to develop an effective approach to fault detection in high-speed automatons. We first investigate the entire working procedure of the automatic mechanism and calculate the corresponding action times of travel involved. The vibration signal collected from the shooting experiment is then divided into a number of impacts corresponding to action orders. Only the segment generated by a faulty component is isolated from the original impacts according to the action time of the component. Wavelet packet decomposition (WPD) is first applied on the resulting signals for investigation of energy distribution, and the components with higher energy are selected for feature extraction. Three information entropy features are utilized to distinguish various states of the automaton using empirical mode decomposition (EMD). A gray-wolf optimization (GWO) algorithm is introduced as an alternative to improve the performance of the support vector machine (SVM) classifier. We carry out shooting experiments to collect vibration data for demonstration of the proposed work. Experimental results show that the proposed work in this paper is effective for fault diagnosis of a high-speed automaton and can be applied in real applications. Moreover, the GWO is able to provide a competitive diagnosis result compared with the genetic algorithm (GA) and the particle swarm optimization (PSO) algorithm.

Keywords: motion sequence decomposition; hybrid information entropy; wavelet packet decomposition; fault detection of automaton; gray-wolf optimization; support vector machine

1. Introduction

A high-speed automaton is the core equipment of antiaircraft guns, and fault diagnosis of high-speed automatons is vital for defending national security. Different from rotating machinery, a high-speed automaton is a complex dynamic system; intense impact, friction, vibration and beating often appear over the course of movement of components. These factors can easily cause failure of the components of automaton. In addition to heavy background noise, the impacts generated by involved components influence each other significantly. It is worth pointing out here that the motion characteristic is a set of unique impacts that can only be observed in a high-speed automaton. Thus,

fault diagnosis of automatic mechanisms is a challenging task in practical applications. Moreover, because of the dangerousness and complexity of this task, the foremost means for investigation of the reliability of an automaton is simulation analysis. To the best of our knowledge, fault diagnosis of various non-rating components in high-speed automatic mechanisms by employing vibration signals collected from a shooting experiment has not been implemented in previous works. Therefore, further work towards the exploration of an efficient technique to solve this challenging task is urgently necessary.

A shooting action is accomplished by intense impacts of various components of the automaton in accurate time order. In these motion sequences, two components may be involved at the same time point while one component may move at the following time point. Serial shootings will repeat the motion sequences rapidly. The disadvantage of propagation of impact is that it transmits assorted impact signals from various components of the automaton in addition to the signals of interest. This phenomenon makes it difficult to recognize the real state of a high-speed automaton using existing techniques. Moreover, when non-rotating components operate at high speeds or within a complex environment, the weak fault information may be suppressed by huge impacts. Considering the harsh working conditions and the specific motion characteristic, commonly used techniques cannot be applied directly to the vibration signal derived from serial impacts, though traditional time–frequency analysis techniques are effective to solve the nonlinear and non-stationary problem. This paper proposes a two-stage fault diagnosis scheme integrating motion sequence decomposition and modern signal processing to carry out fault diagnosis of automaton. At the first stage, the vibration signal generated by a shooting action is segmented into a series of sub-segments to match the actions of components involved according to their action times. Only the travels in which a fault occurs are considered for automaton fault diagnosis while the rest is ignored. At the second stage, time–frequency approaches are applied on the segmented signals for feature extraction and fault diagnosis.

As mentioned earlier, time–frequency analysis is an extensively applied tool in the field of fault diagnosis, like wavelet packet decomposition (WPD), empirical mode decomposition (EMD), and singular value decomposition (SVD). WPD is a powerful decomposition approach, applications of which involve various stages of fault diagnosis. For example, research on bearing degradation estimation with WPD based energy entropy feature was utilized by Lu [1]. This is a representative example of the application of WPD in feature extraction. In addition, wavelet packet decomposition is also an excellent de-noising method that has been applied widely in fault diagnosis. Cui [2] used wavelet de-noising to preprocess vibration signals for diagnosis of rolling bearing faults prior to feature extraction. Liu [3] employed a wavelet de-noising approach to remove noise in vibration data for increasing the signal to noise ratio (SNR). Tabrizi [4] used wavelet packet decomposition and reconstruction to reduce noise for incipient damage detection. In order to reveal the signature of repetitive transients caused by early faults, Wang [5,6] introduced novel Bayesian inference-based optimal wavelet filtering to extend the infograms [7] based on an entropy metric for fault identification purpose. WPD makes important information available on high-frequency components, and the impacts covering a wide frequency range are assigned large coefficients while the smaller coefficients are set to 0. In this paper, the vibration signal collected from serial impacts is usually contaminated by environmental noise. Thus we adopt WPD to reduce heavy background noise and retain the most useful information for extraction of features.

Information entropy is able to quantify the complexity and uncertainty of non-periodic vibration signals. Furthermore, many contributions [8–12] have considered information entropy as a fault feature for fault diagnosis because the occurrence of faults may change the signal distribution in the time and frequency domains. However, a feature is merely sensitive to a corresponding failure. As a result, it is necessary to estimate the state of machinery from different points of view in order to implement a more reliable fault diagnosis. The vibration signal generated by components of the automaton shows a non-periodic characteristic, while information entropy is a more suitable measure to capture the non-periodic fault information. Therefore, three information entropy features are presented to

comprehensively describe the state information of an automaton. The scheme for the extraction of entropy features information is as follows: (1) EMD adaptively decomposes the de-noised signal into a number of modes to form a data matrix for singular value decomposition. The aim of this procedure is to avoid construction of a trajectory matrix, so that it is not necessary to reconstruct phase space involving extra parameters, such as embedded dimensions and lag time; (2) The three information entropy features including the singular spectrum entropy, the Hilbert time–frequency spectrum entropy, and the Hilbert marginal spectrum entropy, are extracted from the time and frequency domains of IMFs for intelligent fault diagnosis of automatons.

Intelligent fault diagnosis has been an attractive research direction in recent years; many meaningful works [13–18] on the implementation of intelligent fault diagnosis in rotating machinery have been reported. It is generally known that the impact of associated parameters on the performance of fault diagnosis classifiers, such as support vector machine (SVM) and the artificial neural network (ANN), is significant. Another challenge related to fault diagnosis of high-speed automatic mechanisms is the lack of samples. For rotating machinery, the number of samples extracting from the vibration signal is based on the working time of rotating components. For a high-speed automaton, the number of samples depends on the number of shots fired. In practice, the number of bullets used for an experiment is finite. Hence, in this paper, for validating the effectiveness of the proposed work, SVM is employed to recognize the various states of a high-speed automaton because of its impeccable ability to classify small samples. To improve the identification ability of the fault classifier, the foremost choice for determining optimal parameters to obtain a higher classification rate is algorithm optimization. Ao [19] proposed an artificial chemical reaction optimization algorithm to optimize SVM parameters for the purpose of bearing fault diagnosis. Cerrada [20] used a genetic algorithm (GA) to enhance the identification accuracy of the random forest classifier. Certainly, particle swarm optimization (PSO) is also an efficient approach for improving the reliability of fault diagnosis [1]. However, for these algorithms, it is observed that search agents only update their states towards the best solution obtained so far without designing a specialized mechanism for local optima avoidance. To exert the superiority of SVM in small sample classification, this paper introduces a gray-wolf optimization (GWO) algorithm to enhance the performance of SVM in fault diagnosis of an automaton.

This paper is organized as follows: Section 2.1 introduces the motion sequence decomposition (MSD) scheme. Section 2.2 introduces the three information entropy features prepared for automaton fault diagnosis. Section 2.3 briefly reviews the theoretical background of SVM and GWO. Section 3 demonstrates the effectiveness of the proposed approach in practical applications. Eventually, conclusions are drawn in Sections 4 and 5. The flowchart of the proposed work is shown in Figure 1.
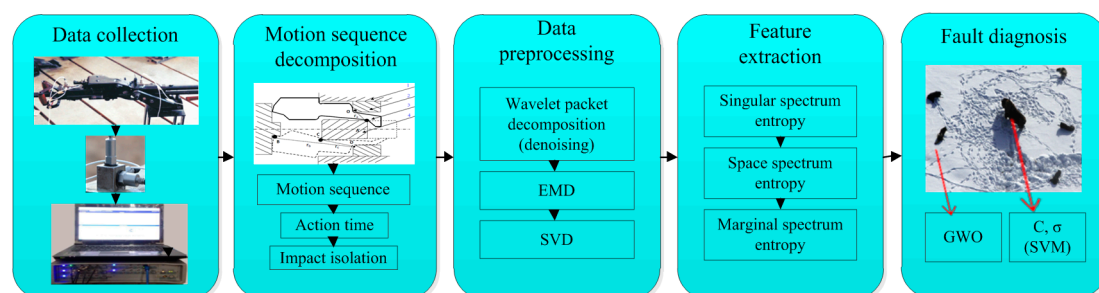

Figure 1. Flowchart of the proposed work for fault diagnosis of a high-speed automaton.

2. Methods

2.1. Motion Sequence Decomposition

Fault diagnosis of a high-speed automatic mechanism is a challenging task in practical applications, principally due to its complex motion procedures as well as its harsh working environment, including high temperatures, high pressure, and high impact. Considering that plenty of state information on

the non-rotating component can be detected by estimating the vibration signal of a specific component, it is essential to investigate the entire working procedure of an automatic mechanism as well as the corresponding action times. As listed in Table 1, the motion procedure is broken into a number of single actions according to the motion time of each component. The total time for accomplishing one shooting is 113.45 ms (The action time of each travel is highlighted in grey background). Since the function of each component in a shooting is known a priori, we can easily determine the travels in which the faults occur. Different from rotating machinery, the function realization of an high-speed automatic mechanism mainly relies on impacting the involved components sequentially, and the impact signal propagates throughout all components of high-speed automaton with rapid attenuation. For example, as shown in Figure 2, four components impact each other in a certain order to accomplish a shooting purpose. Two points, A and B, are impacted as component 4 moves to the right. Inversely, component 4 will impact point C as it returns. Corresponding to their action times, we find that these actions exist at travel 2 and travel 18, respectively. Then the vibration signal consisting of a series of impacts is associated with the decomposed actions. Only the signals derived from the faulty travels are segmented for fault diagnosis of an automatic mechanism. Figure 3 gives an example to illustrate this point. We observe that a shooting arouses a series of impulses acting on the components involved, and each impact duration declines to 0 g quickly during a single shooting. Therefore, in order to enhance the reliability of the fault diagnosis of an automatic mechanism, the impacts generated by the fault should be separated from the signal as its amplitude decreases to approximately 0 g.

Table 1. Travels and action time of each shooting.

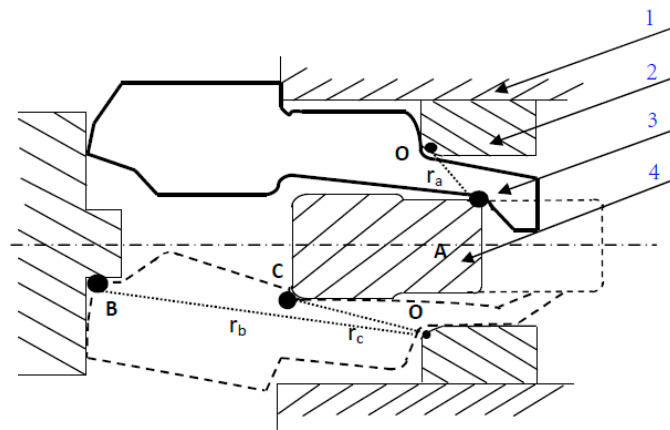| Travels		Action Time
	Entire travel	0 113.45 (ms)
Motion 1	Travel 1	2.65
	Travel 2	3.55
	Travel 3	3.88
	Travel 4	6.07
	Travel 5	11.77
	Travel 6	13.37
	Travel 7	19.9
	Travel 8	34.9
	Travel 9	39.9
	Travel 10	42.84
Motion 2	Travel 11	48.44
	Travel 12	55.27
	Travel 13	58.27
	Travel 14	66.17
	Travel 15	72.77
	Travel 16	92.77
	Travel 17	109.77
	Travel 18	111.05
	Travel 19	113.45

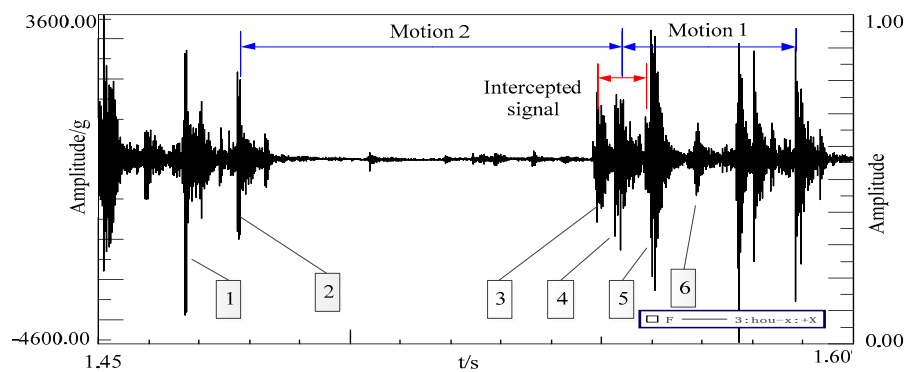**Figure 2.** Partial action sequence diagram.

Figure 3. Vibration signal derived from a single shooting.

As displayed in Figure 3, there are several free travels that exist in the actions of components because there is no obvious impact between impact 2 and impact 3. This means that some actions fail to cause a collision. Moreover, the labeled impacts in Figure 3 are derived from the actions of components in various travels. In practical applications, only the high impacts existing in travel 2 and travel 18 lead to a failure of component 1 and component 3. In this paper, faults are introduced to components 1 and 3 for validation of the proposed work. The details are presented in Section 3. Thus the vibration signal corresponding to travel 2 and travel 18 is selected from each shooting for fault diagnosis.

2.2. Hybrid Information Entropy

Considering that the data collected from a machine gun comprise a high number of non-stationary shocks, the presence of a fault may influence the frequency and energy distribution significantly. Thus, the obtained data are first preprocessed through wavelet packet decomposition (WPD) for investigating the frequency components and energy distribution. Only the most important components are chosen for further estimation. Due to space limitation, for details regarding the theoretical background of WPD, readers may refer to [4].

Empirical mode decomposition (EMD) is a powerful time–frequency analysis technique that has been intensively applied for fault diagnosis [21–23]. It adaptively decomposes the original vibration signal into a series of intrinsic mode functions (IMFs) containing local features. For a vibration signal $Y(t)$, this procedure can be formulated as: $Y(t) = \sum_{i=1}^{n} c_i(t) + r_n(t)$, where $c_i(t)$ and $r_n(t)$ are the IMFs and residue, respectively. After EMD decomposition, the resulting components can show the detailed characteristics of an original input from different points of view.

Entropy is a sensitive quantification index for reflecting the complexity and uncertainty of a system. In this paper, hybrid entropy is employed to extract features for quantifying information changes under different conditions of a machine gun. Given a data matrix $A = [c_1(t), c_2(t), \cdots, c_n(t)]^T$, singular value decomposition on the feature space A is carried out to reflect the intrinsic characteristic of each component [13]:

$$A = U \times S \times V^T, \tag{1}$$

where U and V are orthogonal matrices and $S = \text{diag}[\delta_1, \delta_2, \cdots, \delta_n]$ is a diagonal matrix consisting of elements arranged in descending order. So, the singular spectrum entropy is as follows:

$$\begin{cases} H_{qy} = -\sum_{i=1}^n p_i \log p_i \\ p_i = \delta_i / \sum_{i=1}^n \delta_i \end{cases}. \tag{2}$$

In fact, singular spectrum entropy is a measure of uncertainty of the vibration energy in a system, which can be regarded as a time-domain feature index. However, space spectrum entropy is able to reflect the uncertainty of local energy in this system. After Hilbert transformation, the Hilbert spectrum $H(\omega, l)$ is the representation of the intrinsic mode functions in the time–frequency domain. Similarly, the space spectrum entropy can be calculated by solving singular value decomposition:

$$\begin{cases} H_{kj} = -\sum_{i=1}^n p_i' \log p_i' \\ p_i' = \delta_i' / \sum_{i=1}^n \delta_i' \end{cases}. \tag{3}$$

Subsequently, Hilbert marginal spectrum $h(\omega)$ is calculated to reflect the relationship between the frequency domain and the energy of signal. Then marginal spectrum entropy is defined as:

$$\begin{cases} H_{bj} = -\sum_{i=1}^n p_i'' \log p_i'' \\ p_i'' = h(\omega_i) / \sum_{i=1}^n h(\omega_i) \end{cases}. \tag{4}$$

Marginal spectrum entropy H_{bj} and space spectrum entropy H_{kj} are two metrics for capturing the instantaneous characteristic of modes, so the hybrid entropy feature can assess the real status of a machine gun comprehensively.

2.3. Fault Diagnosis of a High-Speed Automatic Mechanism

2.3.1. Support Vector Machine

Support vector machine (SVM) is a foremost choice for solving small sample learning problems. For classification, SVM maps the original input space into high-dimensional feature space through nonlinear transformation, which has been applied in many fields to solve the nonlinear, over-fitting, high-dimensional pattern recognition problems.

The goal of SVM is to find an optimal hyperplane to identify sample patterns. Take a dataset $\{(x_i, d_i)\}_1^N$, where x_i is the ith input feature vector, d_i is the target value, and N is the total number of training sets. An error function is constructed to transform this aim into solving the minimization of a quadratic function [24]:

$$\begin{cases} \min \frac{1}{2} w^T w + g \sum_{i=1}^N \varepsilon_i^2 \\ \text{s.t.} \quad d_i(w^T \varnothing(x_i) + b) \geq 1 - \varepsilon_i \end{cases}, \tag{5}$$

where ε_i denotes the slack variable, g is a regularization parameter (penalty factor) that controls the complexity of classifier and the number of inseparable sample points, and $\varnothing(x_i)$ is a nonlinear mapping function that maps the original input space into high-dimensional feature space.

The Lagrange function is introduced to determine the solution of Equation (5):

$$J(\boldsymbol{w}, b, \alpha) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \sum_{i=1}^{N}\alpha_i - \sum_{i=1}^{N}\alpha_i d_i\left(\boldsymbol{w}^T\varnothing(\boldsymbol{x}_i) + b\right), \tag{6}$$

where $\alpha_i(i = 1, 2, \cdots, N)$ represent Lagrange multipliers.

Differentiate J with w, b, then let partial derivatives of J equal zero to satisfy the optimality condition: $\partial J / \partial w = 0, \partial J / \partial b = 0$. Then the optimization problem is further converted as:

$$\max L(\alpha) = \sum_{k=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j d_i d_j\varnothing(\boldsymbol{x}_i)\varnothing(\boldsymbol{x}_j) \tag{7}$$

$$\text{s.t.} \quad \sum_{i=1}^{N}\alpha_i d_i = 0, \alpha_i \geq 0.$$

Thus a classification hyperplane can be obtained through solving Equation (7). Then replace the inner product $\varnothing(\boldsymbol{x}_i)\varnothing(\boldsymbol{x}_j)$ with kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ for the purpose of feature mapping. The nonlinear decision function is as follows:

$$f(\boldsymbol{x}) = \text{sgn}(\sum_{i=1}^{N}\alpha_i K(\boldsymbol{x}, \boldsymbol{x}_j) + b), \tag{8}$$

where $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is a kernel function derived from Mercer's theorem. Kernel function is a crucial bridge to map the original input space into feature space, while radial basis function is the foremost selection to overcome this challenging issue. Here $K(\boldsymbol{x}_i, \boldsymbol{x}) = \exp\left(-\frac{1}{2\sigma^2}||\boldsymbol{x} - \boldsymbol{x}_i||^2\right)$, in which σ is band width.

It can be visibly observed that the superiority of SVM heavily relies on the kernel function, and the two parameters σ and g may have a significant impact on the performance of the classifier. Therefore, to enhance the reliability and stability of fault diagnosis of a high-speed automaton, a gray-wolf optimization algorithm (GWO) is introduced for improving the identification accuracy of SVM.

2.3.2. Gray-Wolf Optimization Algorithm

A gray-wolf algorithm is a novelty meta-heuristic optimization algorithm imitating the social hierarchy and hunting behavior of gray wolves. Gray wolves generally prefer to live in groups, and the social hierarchy of the gray wolf comprises four levels, as shown in Figure 4 [25]. The first level α, consisting of a male and a female, is the leader level, which is responsible for decision-making. Other wolves in the group should follow the orders of the dominant wolf. The second level is called β. Wolves at this level are the advisors of leaders in terms of decision and management, and they are probably the best candidates to become the dominant wolf when leaders are lacking. The third ranking is called δ. Wolves in this rank obey the previous two levels of wolves and dominate the wolves in the lowest level ω. Moreover, the wolves belonging to this level are responsible for performing real tasks for the group. Hunters assist the higher-level wolves as hunting prey and also find food for the group. Sentinels are responsible for protecting the group. Scouts watch the territory and alarm the group. Then these relationships and behaviors—social hierarchy, encircling prey, hunting behavior, and search for prey—are mathematically modeled for solving optimization problems.
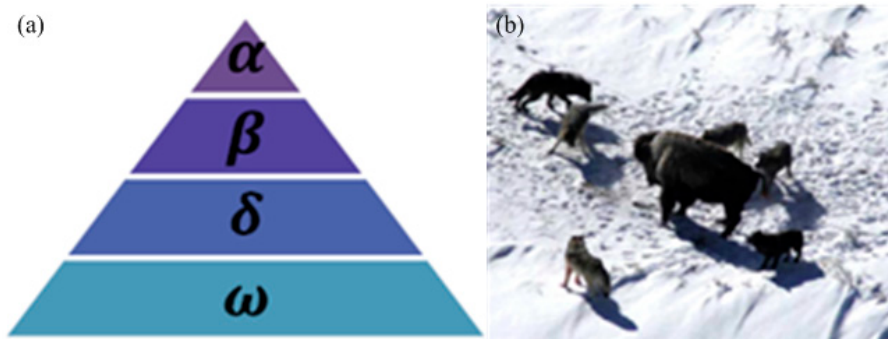
Figure 4. Gray-wolf optimization algorithm: (**a**) social hierarchy of the gray wolf; (**b**) hunting behavior.

- Step 1. Social hierarchy

To model the social hierarchy of gray wolves in designing the optimization algorithm, α is considered as the optimal solution. Thus, β and δ are used as the second and third optimal solutions, respectively. The rest of the candidate solutions are defined to be ω. Obviously, the ω solutions should follow the first three levels according to the relationships depicted in Figure 4a).

- Step 2. Encircling prey

Gray wolves are able to determine the position of prey in space and encircle them. This behavior is formulated as:

$$D = |C \cdot X_p(t) - X(t)| \tag{9}$$

$$X(t+1) = X_p(t) - A \cdot D, \tag{10}$$

where t denotes the iteration number, A and C are coefficient vectors, X_p is the position vector of prey, and X is the position vector of a gray wolf. Here, $A = 2a \cdot r_1 - a$, $C = 2r_2$, in which a is a vector linearly decreased from 2 to 0 in iterations (a can be formulated as: $a = 2 - t \times ((2)/M)$, where M denotes the maximum number of iterations), and r_1 and r_2 are random vectors whose values are derived from [0, 1]. The position of a gray wolf will be updated once a better solution is obtained. Obviously, a gray wolf can reach different positions around the prey in 2D or 3D space by changing the values of coefficient vectors A and C.

- Step 3. Hunting prey

Gray wolves have the ability to determine the position of prey. Since the optimal position of the prey in search space is not known beforehand, the candidate solution is assumed as the target prey. After a better solution is found, the other search agents will update their positions towards the optimal search agent. This behavior can be formulated as:

$$D_\alpha = |C_1 X_\alpha - X|, D_\beta = |C_2 X_\beta - X|, D_\delta = |C_3 X_\delta - X| \tag{11}$$

$$X_1 = X_\alpha - A_1 \cdot D_\alpha, X_2 = X_\beta - A_2 \cdot D_\beta, X_3 = X_\delta - A_3 \cdot D_\delta \tag{12}$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3}. \tag{13}$$

It can be observed that the dominant wolves, α, β, and δ, estimate the position of the prey while the other wolves update their position randomly according to the estimated position of the prey.

- Step 4. Attacking prey

After encircling the prey, gray wolves will attack once the prey stops moving. As mentioned earlier, this phenomenon can be achieved by linearly decreasing the value of a from 2 to 0. In fact, A contains a group of random values defined into $[-2a, 2a]$, so that the value of A will fluctuate with the

decrease in *a*. Defining the values of *A* in [−1, 1], a gray wolf is capable of reaching a new position between the original position and the prey's position.

Currently, search agents update their positions based on position estimation of the three dominant wolves, *α*, *β*, and *δ*. To enhance the global optimization ability of the GWO algorithm, a local optima avoidance mechanism is introduced.

- Step 5. Search for prey

In this phase, in order to perform a global search, $|A| > 1$ is utilized to update the position of a gray wolf according to the randomly chosen search agent. The objective of the inverse operation is to allow other search agents to update the position of an optimal solution. Moreover, vector *C* comprises random values derived from [0, 2]. This vector is applied to randomly weighting the position of the prey so as to enhance or weaken the effect of the position of the prey on the defined distance. Different from vector *A*, *C* is set to provide random values for the purpose of emphasizing exploration in the entire iteration process. For $|A| < 1$, the search agents tend to converge and the iterations are devoted to exploitation. For $|A| > 1$, the search agent will diverge from the current prey and the iterations attempt exploration. For further details with respect to the gray-wolf optimization (GWO) algorithm, readers may refer to [25]. In this paper, to enhance the capability of SVM for fault diagnosis of automatic mechanism, GWO is utilized to determine the optimal parameters of SVM. It is worth pointing out that the fitness function is defined as the training accuracy of SVM. That is, the aim of the GWO algorithm is to find the optimal parameters *σ* and *g* corresponding to a maximization of the training accuracy of the classifier. The scheme for parameters optimization is summarized as follows:

Parameters setting (e.g., population size, number of iterations) and gray wolf population initialization
Initialize the best position vectors: X_α, X_β, X_δ
Calculate the fitness of each search agent
When $t < M$ (maximum number of iterations)
For each agent
Take the parameters *σ* and *g* as the search agent and training SVM model, respectively, by using training data
Calculate the fitness (training accuracy) of each search agent
Update the optimal search agents *σ* and *g*
End for
Update *a*, *A*, *C*
$T = t + 1$
End while
Get the optimal parameters *σ* and *g* and test the trained SVM by using testing data

---

3. Experiments and Results

3.1. Experimental Setup

In order to investigate the effectiveness of the proposed work, components 1 and 3 (as shown in Figure 2) are seeded with faults to simulate different working conditions with crack sizes. The details are summarized as follows: the first is a crack fault of depth 2 mm, which is set near the fillet of component 3. The second one is a crack fault of depth 1.5 mm, which is also introduced to component 3 at the other position. The third crack fault of depth 1.5 mm is introduced at the fillet of component 1. Figure 5 shows the details on the faults. Hence, taking the above faults and normal conditions into account, four states of high-speed automaton are simulated for the validation of the proposed work. In accordance with different conditions, experiments on five serial shootings are conducted for data collection. The sampling system of the vibration signal used in this study is based on the LMS Test Lab system, as displayed in Figure 6. Two accelerometers are mounted on the automatic mechanism to collect vibration signals generated from five serial shootings with a sampling rate of 204.8 kHz. Figure 7 depicts the experimental setup.
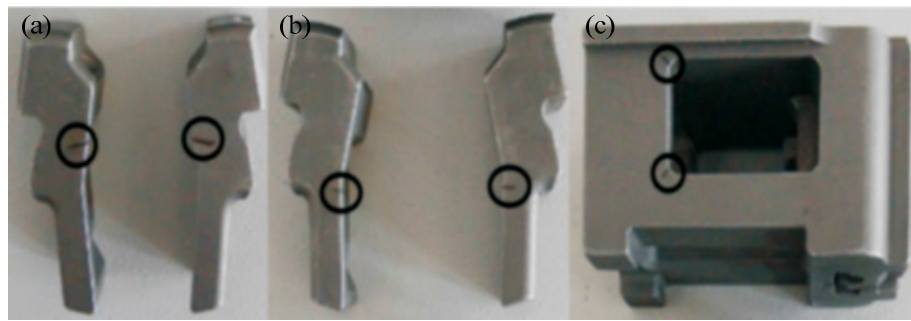
Figure 5. Three faults considered in this study: (**a**) fault 1; (**b**) fault 2; (**c**) fault 3.
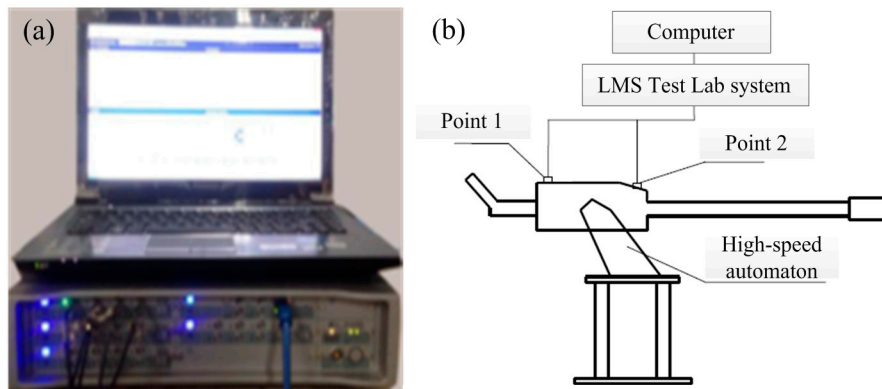

Figure 6. Data collection system: (**a**) LMS Test Lab system; (**b**) schematic figure of data collection.
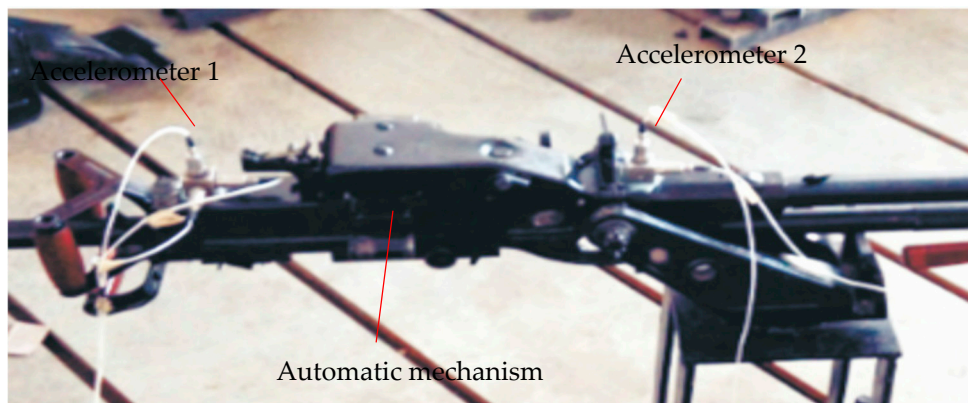

Figure 7. High-speed automatic mechanism.

In a high-speed automatic weapon, the initial shooting requires human intervention, but the remaining four shots occur automatically. Therefore, to remove the influence of human intervention from the diagnosis result of automaton fault, the last four shots of each condition are selected for fault diagnosis purposes. For the resulting signals, we split them into four single shootings and then pick the faulty travels (the sampling time is 10 ms) from each shooting for automaton fault diagnosis according to the motion sequence decomposition scheme proposed in Section 2. As depicted in Figure 8, the isolated signals have a similar appearance and the impacts occur within a very short space of time.
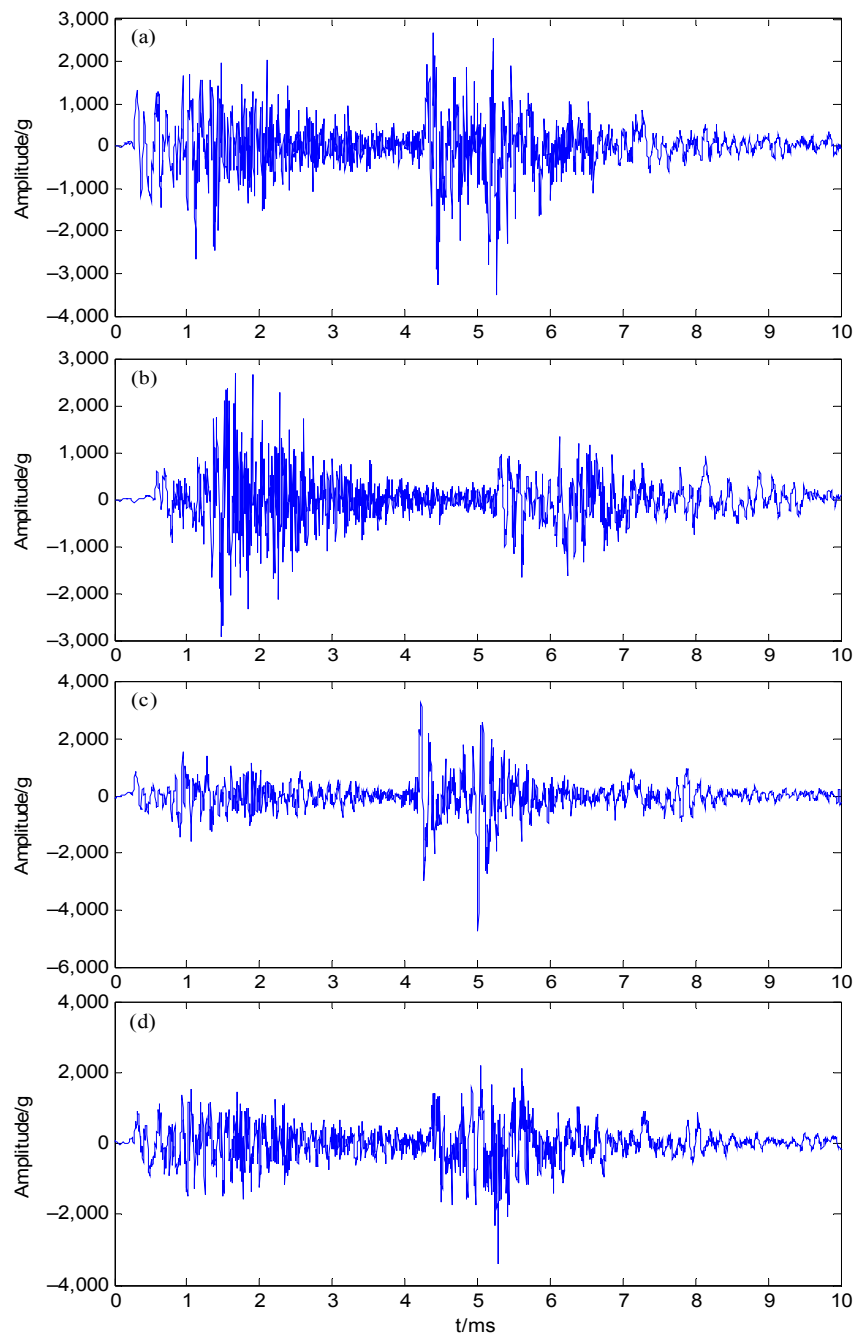
Figure 8. Row vibration signals collected from a high-speed automaton: (**a**) fault 1; (**b**) fault 2; (**c**) fault 3; (**d**) normal state.

3.2. Experimental Results

Because the data were recorded with a high sampling rate (204.8 kHz), a "db4" wavelet is used to decompose the impact vibration signal into a tree including six-layer decomposition. Then the normalized wavelet energy feature is calculated to investigate the detailed energy distribution of the impact signal. Figure 9 shows the normalized energy of a six-layer wavelet packet decomposition of each state. It is generally known that the selection of wavelet functions is difficult for different practical applications. Daubechies-family wavelets are often used for defect detection and fault diagnosis because of their advantages in terms of orthogonality and computational simplicity [26]. In general, a proper decomposition level can be determined in practical applications depending on

the real conditions, as has been demonstrated by many published works [1,26,27]. It can be seen in Figure 9 that the energy of the four conditions is mainly concentrated in the low-frequency segment (bands 1–8). Moreover, the energy in the low-frequency segment is much larger than that in the high-frequency segment (bands 35–40). Evidently, Figure 9 suggests that frequency bands 3 and 7, with higher amplitudes, should be chosen for signal reconstruction because the modes of interest have the potential to retain dominant information about the signal. As an example, the reconstructed signal of fault 1 is presented in Figure 10. Singular value decomposition on the matrix whose rows are IMFs is performed to calculate the three information entropy features. In total, we get three features and 16 samples from feature extracting procedure. Each state contains four samples. Half of the total samples are presented to SVM optimized by GWO for model training, while all samples are used for testing. The initialization parameters of GWO, such as the gray wolf population N, maximum number of iterations M, and search space [lb, ub] are as follows: $M = 200$, $N = 20$, [lb, ub] = $[10^{-2}, 10^3]$. The task of GWO is to find the maximization of SVM accuracy. Furthermore, two meta-heuristic algorithms, the genetic algorithm (GA) and the particle swarm optimization (PSO) algorithm, are also examined for comparison. For GA, the generation and the population size are set at 200 and 20, respectively. The generation gap is 0.9. For PSO, the iteration count and the numbers of particles are set to 200 and 50, respectively ($c_1 = 1.5$, $c_2 = 1.7$, inertia weight $W = 1$). The search space of the algorithms is the same as that of the GWO. We ran the compared algorithms 15 times to investigate their exploration ability. The average optimal solution of 15 runs is listed in Table 2. All the algorithms were performed with Matlab-2014a. The SVM tool in LIBSVM-3.18 library [28] was employed to identify automaton faults. For each run, the optimal identification rate is obtained using 5-fold cross-validation.
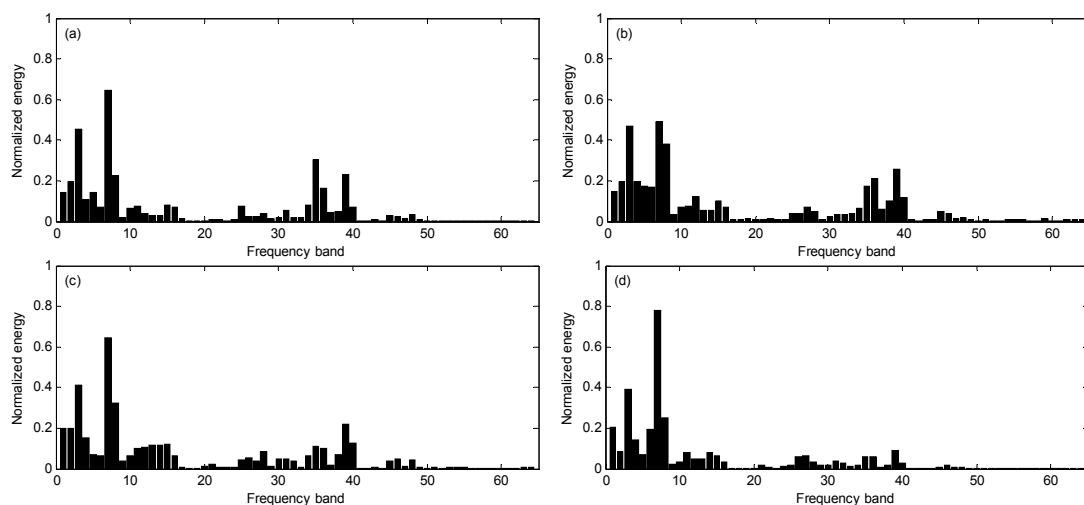

Figure 9. Energy distribution of each state after wavelet packet decomposition: (**a**) normal state; (**b**) fault 1; (**c**) fault 2; (**d**) fault 3.
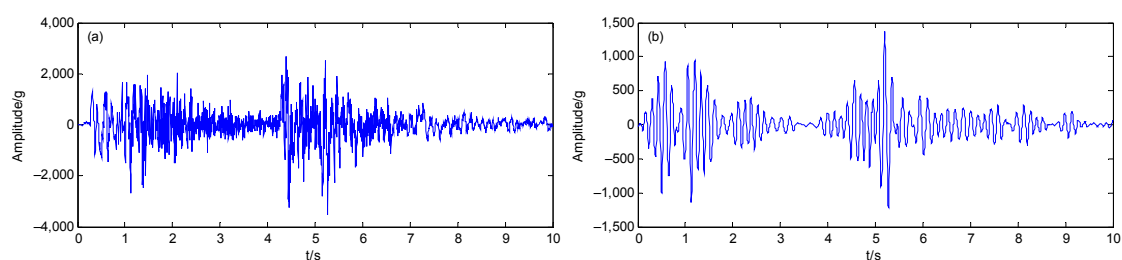

Figure 10. Impact signal: (**a**) original impact signal; (**b**) reconstruction of (**a**).

Table 2. Exploration ability of algorithms compared.

Method	Training Samples	Testing Samples	Optimal Fitness (%)
GA-SVM	8	16	75
PSO-SVM	8	16	75
GWO-SVM	8	16	87.5

We observe from Table 2 that SVM-GWO provides the highest identification accuracy (87.5%) for the four states of a high-speed automatic mechanism. The optimal classification accuracies of PSO-SVM and GA-SVM reach 75%. This means that GWO has a stronger ability to avoid local optima compared with GA and PSO during model training. Based on this, the average identification rate for training and testing set is presented in Table 3. From Table 3, we find that the testing accuracies of GA-SVM, PSO-SVM, and GWO-SVM are 81.25%, 87.5%, and 87.5%, respectively. Like GA and PSO, the diagnosis result of SVM using GWO is satisfactory. Although GWO fails to show obvious superiority in terms of classification accuracy compared with PSO and GA, it may be used as an alternative because of its fewer initialization parameters and stronger exploration capability.

Table 3. Diagnosis results using three approaches.

Method	Training Accuracy (%)	Testing Accuracy (%)
GA-SVM	100	81.25
PSO-SVM	87.5	87.5
GWO-SVM	100	87.5

4. Discussion

This paper proposes a motion sequence decomposition technique to isolate useful segments of impact signals collected from action sequences of a high-speed automaton. The major findings are as follows:

(1) It is possible to accurately estimate the state of non-rotating components in a high-speed automaton by dealing with the vibration signal recorded from the shooting experiment. The result obtained above reveals the effectiveness of the proposed work to diagnose the state of an automaton. At present, many published works have been presented for fault diagnosis of rotating machinery, such as rolling bearing, gearbox, etc. Researchers have paid little attention to fault diagnosis of non-rotating machinery. It is worth pointing out that, to the best of our knowledge, no published works using a hybrid entropy feature have approached the issue of distinguishing the states of high-speed automaton based only on impact signal.

(2) Motion sequence decomposition is an efficient approach to fault diagnosis in high-speed automatic mechanisms. Also, it is of great significance for fault diagnosis of non-rotating machinery. From Section 2, we understand that the impact signal of high-speed mechanism shows a specific characteristic different from common signals. Moreover, a shooting action comprises 19 motion travels, and the impact signal that is sampled includes superposed signals coming from these travels as well as components involved. In addition, due to the "three-high" working condition, extraction of the fault features is more complicated in practical applications. That is, this paper accomplishes a challenging task.

(3) As displayed in Table 4, for each fault, we find that the feature indexes are unstable. This presents a great challenge to fault diagnosis of a high-speed automaton. The main reason may lie in the fact that the mass of dynamics system (automaton) tends to decrease along the shooting time. That is, the shooting experiment starts with five bullets, while the mass of the system varies during the five shootings. Moreover, the amount of gunpowder in each bullet is different, which also affects the response of the dynamic system.

(4) GWO is introduced to improve the performance of SVM in the fault identification of an automaton. Experimental results confirm that GWO has a stronger ability to avoid local optima while exploring the best performance of SVM compared with the widely used algorithms GA and PSO (Table 2). Although GWO fails to show obvious superiority, it provides a competitive diagnosis result in this study (Table 3). It can be used as a potential alternative for parameter optimization.

Table 4. Hybrid entropy features.

State	H_{kj}	H_{bj}	H_{qy}
Normal conditions	8.1267	7.6934	1.5944
	7.8538	7.5349	1.5688
	8.1505	7.7227	1.8123
	7.8150	7.5249	1.3778
Fault 1	7.2681	6.9814	1.1907
	7.5189	7.2021	1.2030
	7.2469	7.0287	0.9091
	7.2067	7.0129	0.8338
Fault 2	7.9746	7.3326	0.8424
	7.8506	7.1696	0.7401
	7.8968	7.4239	1.1333
	7.8181	7.1364	0.7521
Fault 3	7.9501	7.7056	1.8466
	8.1020	7.6833	1.7437
	8.0773	7.7977	2.2783
	8.0250	7.6825	1.5842

5. Conclusions

Aiming at providing an effective approach to fault diagnosis of a high-speed automatic mechanism, this paper proposes a motion sequence decomposition technique combining modern signal processing methods to diagnose non-rotating machinery fault. A modern signal processing technique, wavelet packet decomposition (WPD), is applied to remove noise and redundancy. Three information entropy features from the segmented vibration signal are calculated with the help of EMD. The gray-wolf optimization algorithm is introduced to improve the performance of SVM. To sum up, the following conclusions can be obtained:

(1) The proposed work in this paper is effective for fault diagnosis of a high-speed automatic mechanism and the results obtained are satisfactory.

(2) This paper demonstrates that the information entropy feature can be used as an efficient measure of fault information to recognize faults in automatic mechanisms.

(3) The gray-wolf optimization (GWO) algorithm is used to improve the classification accuracy of SVM. Although GWO fails to indicate obvious superiority, it provides a competitive diagnosis result compared with GA and PSO.

(4) The proposed work is of great significance for fault diagnosis of non-rotary mechanical parts. Fault diagnosis of a high-speed automaton is a challenging task in real life, so this subject needs more attention, as well as broader views.

Author Contributions: Baoxiang Wang proposed the method used in this paper and wrote the article. Hongxia Pan performed the shooting experiments, collected experimental data, and improved the proposed approach. Heng Du designed the shooting experiment and made many meaningful comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lu, C.; Chen, J.; Hong, R.; Feng, Y.; Li, Y. Degradation trend estimation of slewing bearing based on LSSVM model. *Mech. Syst. Signal Process.* **2016**, *76*, 353–366. [CrossRef]

2. Cui, H.; Qiao, Y.; Yin, Y.; Hong, M. An investigation of rolling bearing early diagnosis based on high-frequency characteristics and self-adaptive wavelet de-noising. *Neurocomputing* **2016**, *216*, 649–656. [CrossRef]

3. Liu, Z.; He, Z.; Guo, W.; Tang, Z. A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery. *ISA Trans.* **2016**, *61*, 211–220. [CrossRef] [PubMed]

4. Tabrizi, A.; Garibaldi, L.; Fasana, A.; Marchesiello, S. Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine. *Meccanica* **2015**, *50*, 865–874. [CrossRef]

5. Wang, D. An extension of the infograms to novel Bayesian inference for bearing fault feature identification. *Mech. Syst. Signal Process.* **2016**, *80*, 19–30. [CrossRef]

6. Wang, D.; Tsui, K.L. Dynamic Bayesian wavelet transform: New methodology for extraction of repetitive transients. *Mech. Syst. Signal Process.* **2017**, *88*, 137–144. [CrossRef]

7. Antoni, J. The infogram: Entropic evidence of the signature of repetitive transients. *Mech. Syst. Signal Process.* **2016**, *74*, 73–94. [CrossRef]

8. Asr, M.Y.; Ettefagh, M.M.; Hassannejad, R.; Razavi, S.N. Diagnosis of Combined Faults in Rotary Machinery by Non-Naive Bayesian Approach. *Mech. Syst. Signal Process.* **2017**, *85*, 56–70. [CrossRef]

9. Yang, Z.X.; Zhong, J.H. A hybrid EEMD-based sampEn and SVD for acoustic signal processing and fault diagnosis. *Entropy* **2016**, *18*. [CrossRef]

10. Li, Y.; Xu, M.; Wang, R.; Huang, W. A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy. *J. Sound Vib.* **2016**, *360*, 277–299. [CrossRef]

11. Liu, H.; Han, M. A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. *Mech. Mach. Theory* **2014**, *75*, 67–78. [CrossRef]

12. Qian, Y.; Yan, R.; Hu, S. Bearing degradation evaluation using recurrence quantification analysis and Kalman filter. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 2599–2610. [CrossRef]

13. Kang, M.; Kim, J.M. Singular value decomposition based feature extraction approaches for classifying faults of induction motors. *Mech. Syst. Signal Process.* **2013**, *41*, 348–356. [CrossRef]

14. Muruganatham, B.; Sanjith, M.A.; Krishnakumar, B.; Murty, S.A.V.S. Roller element bearing fault diagnosis using singular spectrum analysis. *Mech. Syst. Signal Process.* **2013**, *35*, 150–166. [CrossRef]

15. Zhang, Y.; Zuo, H.; Bai, F. Classification of fault location and performance degradation of a roller bearing. *Measurement* **2013**, *46*, 1178–1189. [CrossRef]

16. Zhou, H.; Chen, J.; Dong, G.; Wang, R. Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model. *Mech. Syst. Signal Process* **2016**, *72*, 65–79. [CrossRef]

17. Tian, Y.; Ma, J.; Lu, C.; Wang, Z. Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine. *Mech. Mach. Theory* **2015**, *90*, 175–186. [CrossRef]

18. Feng, J.; Lei, Y.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process* **2016**, *72*, 303–315.

19. Ao, H.; Cheng, J.; Yang, Y.; Truong, T.K. The support vector machine parameter optimization method based on artificial chemical reaction optimization algorithm and its application to roller bearing fault diagnosis. *J. Vib. Control* **2015**, *21*, 2434–2445. [CrossRef]

20. Cerrada, M.; Zurita, G.; Cabrera, D.; Sánchez, R.V.; Artés, M.; Li, C. Fault diagnosis in spur gears based on genetic algorithm and random forest. *Mech. Syst. Signal Process.* **2016**, *70*, 87–103. [CrossRef]

21. Grasso, M.; Chatterton, S.; Pennacchi, P.; Colosimo, B.M. A data-driven method to enhance vibration signal decomposition for rolling bearing fault analysis. *Mech. Syst. Signal Process.* **2016**, *81*, 126–147. [CrossRef]

22. Liu, Z.; Cao, H.; Chen, X.; He, Z.; Shen, Z. Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing* **2013**, *99*, 399–410. [CrossRef]

23. Xu, H.; Chen, G. An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO. *Mech. Syst. Signal Process.* **2013**, *35*, 167–175. [CrossRef]

24. Yang, D.; Liu, Y.; Li, S.; Li, X.; Ma, L. Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm. *Mech. Mach. Theory* **2015**, *90*, 219–229. [CrossRef]
25. Seyedali, M.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Soft.* **2014**, *69*, 46–61.
26. Yu, J. Health condition monitoring of machines based on hidden Markov model and contribution analysis. *IEEE Trans. Instr. Meas.* **2012**, *61*, 2200–2211. [CrossRef]
27. Yu, J. Local and nonlocal preserving projection for bearing defect classification and performance assessment. *IEEE Trans. Ind. Electron.* **2012**, *59*, 2363–2376. [CrossRef]
28. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intel. Sys. Tech.* **2011**, *2*. [CrossRef]