Dynamic Weight-Based Adaptive Data Fusion Filtering for Real-Time Condition Monitoring of Electromechanical Equipment

Liqun Wang[1*], Long MA[2]
[1]School of Mechanical and Electrical Engineering, Suzhou Vocational University, Suzhou 215000, Jiangsu, China
[2]CGN Inspection Technology Co., Ltd, Suzhou, 215004, Jiangsu, China
E-mail：wangliqun@jssvc.edu.cn
*Corresponding author

Keywords: electromechanical equipment, condition monitoring, adaptive data fusion, filtering algorithm, dynamic weight

Received: May 19, 2025

In electromechanical equipment condition monitoring, data accuracy and timeliness are critical for ensuring reliable operation. This paper introduces a dynamic weight-based adaptive data fusion filtering algorithm that integrates an Extended Kalman Filter (EKF) variant with wavelet denoising and real-time weight adjustment. The algorithm employs variance-based stability metrics and data update frequency for reliability assessment, dynamically allocating fusion weights via Equations (1)–(7). Experiments use a test platform with PCB 356A16 accelerometers, K-type thermocouples, and LEM LA55-P current sensors to simulate gear wear and motor rotor imbalance faults. Compared to baseline methods (Kalman Filter, Particle Filter), the proposed algorithm reduces RMSE by 35.2% and 28.7%, and MAE by 41.5% and 34.3%, respectively. Fault feature extraction shows a 37% increase in vibration peak index for mild gear wear and a 50% increase in current harmonic content for rotor imbalance. While runtime is slightly higher than Kalman Filter, it remains within practical limits, demonstrating superior real-time performance for fault early warning in complex industrial environments.

Povzetek: Raziskava uvaja dinamično uteženo adaptivno filtriranje za spremljanje stanja elektrostrojne opreme. Algoritem združuje EKF, valčno denoiziranje in sprotno prilagajanje uteži, izboljšuje zgodnje odkrivanje napak v realnem času.

1 Introduction

In today's era of rapid industrialization, electromechanical equipment is a core component of the production process in many key industrial fields such as manufacturing, energy, and transportation. Its importance is self-evident. In the manufacturing industry, precision electromechanical equipment ensures high precision and high efficiency of product production; in the energy industry, large electromechanical equipment maintains the stable mining, transmission, and conversion of energy; in the field of transportation, various types of electromechanical equipment ensure the safe and stable operation of vehicles. The operating status of electromechanical equipment is like a "barometer" of enterprise production, which is directly related to the enterprise's production efficiency and product quality [1]. Once a sudden failure of equipment occurs, it will not only lead to production stagnation and substantial economic losses. However, it may also cause safety accidents and affect the sustainable development of the enterprise.

As a key means to ensure the reliable operation of electromechanical equipment, condition monitoring technology can obtain the operating status information of the equipment in real time [2]. Through continuous monitoring and in-depth analysis of multiple parameters such as vibration, temperature, pressure, and current, potential fault hazards can be detected in advance, and downtime losses caused by sudden failures can be effectively avoided. This not only dramatically improves the reliability of the equipment but also significantly improves production efficiency, reduces equipment maintenance costs, and creates considerable economic benefits for the enterprise.

However, the traditional fixed-weight data fusion algorithm currently used has obvious limitations when facing electromechanical equipment's complex and changeable operating environment [3]. In the operation of electromechanical equipment, sensor data is easily affected by various factors and fluctuates, and noise interference is also common. In this case, the traditional algorithm cannot flexibly respond to the dynamic changes of data due to the use of fixed weights, and it is difficult to fuse various sensor data accurately, which leads to a significant decrease in monitoring accuracy [4]. In addition, standard filtering algorithms, such as Kalman

filtering, have inherent defects when dealing with non-Gaussian noise and nonlinear systems, and cannot accurately denoise and estimate the state of monitoring data, which seriously affects the accurate judgment of the operating state of electromechanical equipment.

Given this, it is urgent to design a filtering algorithm that can dynamically adjust the data fusion strategy according to the real-time operating state of electromechanical equipment [5]. This study addresses the following research questions:

Can variance- and autocorrelation-based dynamic weighting improve fault feature extraction accuracy under varying load conditions?

How does an EKF variant with adaptive gain adjustment (Equation 12) compare to traditional nonlinear filters (e.g., Particle Filter) in non-Gaussian noise environments?

What is the impact of real-time weight adaptation on computational efficiency for multi-sensor fusion systems?

2 Related theoretical foundations

2.1 Principles of electromechanical equipment status monitoring

2.1.1 Common monitoring parameters and methods

Parameters such as vibration, temperature, pressure, and current are key in electromechanical equipment status monitoring [6]. Vibration parameters can reflect problems such as wear and imbalance of mechanical parts; temperature parameters can monitor the heating parts of the equipment; pressure parameters are used for fluid transmission equipment; and current parameters are for motor equipment [7]. Methods for obtaining these parameters include direct monitoring and indirect monitoring based on sensors. Acceleration sensors measure vibration, thermocouples measure temperature, and pressure sensors measure pressure. Indirect tracking, such as analyzing the characteristics of the current spectrum to infer mechanical failures, is achieved through technologies such as the Fourier transform.

2.1.2 Analysis of sensor data characteristics

Sensor data collection is subject to noise interference, ordinary Gaussian white, and pulse noise. Gaussian white noise causes slight deviations in the data, affecting accurate analysis; pulse noise may mask the real signal and lead to misjudgment. Under different operating conditions, the noise intensity and distribution law change. The noise intensity is high during the start and stop stages of the equipment, and low during the stable operation stage [8]. There is an inherent connection between different sensor data, such as increased vibration, leading to increased temperature and abnormal current. When the equipment fails, this correlation is broken. Analyze the difference in data correlation to provide a basis for data fusion and improve the accuracy and reliability of monitoring through data fusion.

2.2 Data fusion and filtering technology

2.2.1 Basic concepts and levels of data fusion

Data layer fusion (weighting of raw data after preprocessing) + feature layer embedding (extracting peak, harmonic and other features in EKF) is adopted. Data fusion is the processing of data obtained by multiple sensors to obtain comprehensive and accurate equipment operation status information. It is divided into data layer fusion, feature layer fusion, and decision layer fusion [9]. Data layer fusion directly processes the original data, retains the original information, but has a large amount of calculation; feature layer fusion extracts feature and then fuses them, reducing the amount of data and highlighting key information; decision layer fusion first independently processes and then fuses the decision results, with strong fault tolerance. Different levels of fusion methods are suitable for different monitoring scenarios.

2.2.2 Principles and applications of traditional filtering algorithms

Kalman filtering is an algorithm based on linear systems and Gaussian noise assumptions. It estimates the system state by predicting and updating equations [10]. It is suitable for linear dynamic systems, such as motor speed estimation. Particle filtering is a nonlinear filtering algorithm. It simulates the probability distribution of system states through many particles, adapts to complex nonlinear and non-Gaussian systems, and can be used for fault diagnosis of complex electromechanical systems.

3 Design of adaptive data fusion filtering algorithm

3.1 Overall architecture of the algorithm

3.1.1 Introduction to system components

The sensor data acquisition module is a front-end data acquisition unit equipped with various sensors (such as acceleration, thermocouples, pressure, and current sensors) to collect real-time device state parameters. The collected raw data will be interfered with by factors such as environmental noise, so the module will perform preliminary conditioning on the data, such as increasing the signal strength through a signal amplifier and removing noise with a hardware filter to ensure that the data is accurate and reliable, providing a basis for subsequent processing.

The data preprocessing module receives the data output by the acquisition module and performs deep purification and regularization [11]. Wavelet denoising technology is used to remove random noise and impulse noise to make the data curve smooth; normalization operations are performed to map data of different dimensions to a specific interval to eliminate the impact of dimensional differences; it also has the function of outlier detection and elimination to identify and remove abnormal points and provide high-quality data for subsequent processing.

The dynamic weight calculation module calculates the weight of each sensor data in the fusion process in real time based on the sensor data stability and reliability evaluation indicators. Stability evaluation considers data variance and time series correlation, and reliability judgment is based on data update frequency and sensor fault diagnosis information [12]. This module dynamically determines the importance of sensor data and provides a reasonable weight distribution scheme for data fusion, so the algorithm can adaptively adjust the fusion strategy.

The data fusion and filtering module receives preprocessed data and dynamic weight information, first weights the fusion data according to the weight, and integrates multiple sensor data; then uses the improved filtering algorithm to deeply process the fusion data, suppress noise, enhance effective signals, and finally outputs monitoring data that accurately reflects the operating status of the equipment, providing a basis for status analysis and fault diagnosis.

3.1.2 Data flow and interaction between modules

The raw data of the sensor data acquisition module flows into the data preprocessing module, and is transmitted to the dynamic weight calculation module and the data fusion and filtering module after processing. The dynamic weight calculation module calculates the weight and feeds it back to the data fusion and filtering module. The data fusion and filtering module fuses the data according to the weight, filters it, and outputs the monitoring data. The data preprocessing module interacts closely with the dynamic weight calculation module. The former provides high-quality data for the latter, and the latter calculates the weights and feeds them back to the data fusion and filtering module [13]. The interaction between the dynamic weight calculation module and the data fusion and filtering module is crucial. The weight information determines the fusion method and results. The data fusion and filtering module may feed abnormal information to prompt weight adjustment. The modules work together to realize the algorithm's function.

3.2 Dynamic weight allocation mechanism

3.2.1 Construction of stability evaluation indicators

Variance is an important statistic to measure the degree of data dispersion. The stability of sensor data is evaluated by calculating the variance. Suppose the data sequence collected by the sensor over some time is $x_1, x_2, \cdots, x_n$, then the variance σ^2 of the data sequence is calculated as

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad (1)$$

Where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the mean of the data sequence. This paper dynamically adjusts the weight of sensor data according to the change of variance.

To more accurately describe the relationship between weight and variance, the weight adjustment function is introduced

$$w_\sigma = 1 - \alpha\frac{\sigma^2}{\max(\sigma^2)} \quad (2)$$

Where α is the weight adjustment coefficient ($0 < \alpha < 1$), $\max(\sigma^2)$ is the maximum value of the variance of all sensor data. This function ensures that sensor data with slight variance obtains higher weights and the weight adjustment range is within a reasonable range.

Time series analysis methods, such as autocorrelation function (ACF) and partial autocorrelation function (PACF), are introduced to further analyze the correlation of sensor data in the time dimension to improve the stability evaluation index. The autocorrelation function measures the correlation between time series data and itself at different time intervals. For the time series $\{x_t\}$, its autocorrelation functionρ_k is defined as

$$\rho_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \quad (3)$$

Where k is the time lag.

The results of time series analysis are taken into account in weight adjustment to construct a comprehensive stability index S

$$S = \beta w_\sigma + (1-\beta)\sum_{k=1}^{m}|\rho_k| + \sum_{k=1}^{m}|PACF_k| \quad (4)$$

Where β is the weight distribution coefficient ($0 < \beta < 1$), and m is the upper limit of the set time lag.

3.2.2 Reliability judgment basis and method

Assume that the data update period of sensor j is T_j, then its data update frequency $f_j = \frac{1}{T_j}$. In the initialization stage of the algorithm, the initial reliability weight $w_{f,j}$ is set according to the data update frequency of each sensor, using the following formula

$$w_{f,j} = \frac{f_j}{\sum_{i=1}^{N} f_i} \quad (5)$$

Where N is the total number of sensors. When the data update frequency of sensor j becomes f_j', its reliability weight is adjusted to

$$w'_{f,j} = \gamma\frac{f_j'}{\sum_{i=1}^{N} f_i'} + (1-\gamma)w_{f,j} \quad (6)$$

Where γ is the adjustment coefficient ($0 < \gamma < 1$).

The sensor fault diagnosis flag F_j is introduced. When sensor j works usually, $F_j = 1$; when the sensor fails, $F_j = 0$. Integrate the fault diagnosis information into the reliability weight adjustment process to construct the comprehensive reliability index R_j

$$R_j = F_j w_{f,j} \quad (7)$$

3.3 Fusion and filtering process

3.3.1 Data fusion algorithm steps based on dynamic weights

Linear interpolation is used for data alignment. Assume that the two sampling points of sensor i near time point t are(t_1, x_1) and (t_2, x_2), then the data value x at time point t is calculated by linear interpolation as

$$x = x_1 + \frac{(t-t_1)(x_2-x_1)}{t_2-t_1} \quad (8)$$

By time-aligning all sensor data, the accuracy and effectiveness of data fusion are guaranteed, providing a

time-synchronized data basis for subsequent weighted fusion.

The time-aligned data is weighted and fused according to the weights of each sensor obtained by the dynamic weight calculation module. Assume that there are$N$ sensors in total, the weight of sensor i is w_i, and the data value of sensor i at time t after time alignment is $x_i(t)$. Then the fused data value $X(t)$ is calculated by the weighted summation formula

$$X(t) = \sum_{i=1}^{N} w_i x_i(t) \quad (9)$$

Through this weighted fusion method, the advantages of each sensor data can be fully utilized, so that the fused data can better reflect the real state of the equipment. The sensor data with larger weights accounts for a more significant proportion of the fusion results, reflecting its importance in the equipment state.

3.3.2 Noise suppression and signal enhancement processing combined with an improved filtering algorithm

Let the system state equation be $x_{k+1} = f(x_k) + w_k$, and the observation equation be $z_k = h(x_k) + v_k$, where x_k is the system state, z_k is the observation value, $f(x_k)$ and $h(x_k)$ are nonlinear functions, and w_k and v_k are system noise and observation noise, respectively. EKF linearizes $f(x_k)$ and $h(x_k)$ and uses the Jacobian matrix to convert nonlinear problems into linear problems for processing.

The improved EKF incorporates wavelet packet decomposition for feature extraction, using the db4 wavelet (5 decomposition layers) to isolate fault-related frequency components (e.g., gear meshing frequency). The Jacobian matrix for nonlinear system approximation is derived via finite differences, with updates to the state covariance matrix incorporating both process noise (Q) and measurement noise (R) from sensor calibration data [14]. Pseudocode Outline for Feature-Enhanced EKF:

Initialize: x_0, P_0, Q, R, wavelet basis
For each time step k:
  Predict:
    $x_k^- = f(x_{k-1})$
    $P_k^- = F_k P_{k-1} F_k^T + Q$
  Update:
    $z_k = h(x_k^-) + v_k$
    $y_k = z_k - h(x_k^-)$
    $S_k = H_k P_k^- H_k^T + R$
    $K_k = P_k^- H_k^T S_k^{-1}$
    $x_k = x_k^- + K_k y_k$
    $P_k = (I - K_k H_k) P_k^-$
  Feature Extraction:
    Decompose x_k using wavelet packet
    Extract kurtosis, peak index, and harmonic ratio
Return x_k, fault features

4 Algorithm optimization strategy

4.1 Computational efficiency optimization

4.1.1 Methods to reduce redundant calculations

Use simplified linear models to replace complex nonlinear models for calculations. For example, for the description of the operating state of some electromechanical equipment, although it may be nonlinear, it can be approximated as a linear relationship within a specific range of operating conditions. Suppose the original nonlinear model is $y = f(x)$. Under certain operating conditions, it can be expressed as $y \approx a + bx$ through linear approximation, where a and b are the coefficients obtained by linear fitting the original model under the operating condition.

4.1.2 Application of parallel computing strategy in the algorithm

According to the computing characteristics of each algorithm module, the key to improving computing efficiency is to select a parallel computing architecture reasonably. Multi-threaded CPU parallel computing suits modules with light computing tasks and frequent data interactions. GPU-based parallel computing can significantly improve the computing speed for modules with large data volumes and intensive computing, such as data fusion and filtering modules. GPU has powerful parallel computing capabilities and many computing cores [15]. In the data fusion and filtering module, when performing weighted fusion and complex filtering calculations on a large amount of sensor data, the computing tasks are assigned to multiple computing cores of the GPU for parallel execution.

Taking the weighted fusion of N sensor data as an example, under the traditional CPU serial computing method, the computing time is

$$T_{CPU} = \sum_{i=1}^{N} t_i \quad (10)$$

Where t_i is the time to process the i. sensor data; after GPU parallel computing, due to the simultaneous operation of multiple computing cores, the computing time can be significantly shortened to T_{GPU}, and $T_{GPU} \ll T_{CPU}$.

4.2 Improved anti-interference capability

4.2.1 Detection and processing mechanism for abnormal data

Multiple methods are used to detect abnormal data in electromechanical equipment monitoring data. The 3σcriterion, based on statistical analysis, is a commonly used abnormal data detection method. For a set of monitoring data $x_1, x_2, \cdots, x_n$ that obeys a normal distribution, first calculate the mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and the standard deviation $\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$) of the data. According to the 3σ criterion, normal data should be distributed in the interval $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$. If the data point exceeds this interval, it is judged as abnormal data.

At the same time, the isolation forest algorithm based on machine learning is introduced to improve the accuracy of abnormal data detection. The isolation forest algorithm constructs multiple binary trees and maps each data point to these binary trees. A data point's outlier degree is determined by its depth in the binary tree. The greater the depth, the higher the outlier degree. In this algorithm, a large amount of monitoring data under

normal working conditions is used to train the isolation forest model and determine the parameters of the model. In actual detection, the real-time monitoring data is input into the trained model, and the data is judged as abnormal according to the outlier score output by the model [16]. By reasonably setting the threshold of the 3σ criterion and the model parameters of the isolation forest algorithm, the abnormal data points in the monitoring data of electromechanical equipment can be accurately identified, providing an accurate basis for subsequent data processing.

For a small number of isolated abnormal data points, the linear interpolation method is used for repair. Suppose the abnormal data point is x_j, and its adjacent normal data points before and after are x_{j-1} and x_{j+1} respectively, then the repaired data is obtained by linear interpolation

$$\hat{x}_j = x_{j-1} + \frac{(x_{j+1}-x_{j-1})(j-(j-1))}{(j+1)-(j-1)} \quad (11)$$

For abnormal data segments that appear continuously, judgment is made in combination with the historical data of equipment operation and the physical model. If it is determined that the data error is caused by sensor failure or other abnormal conditions, the weight of the sensor in data fusion is reduced, and its data is marked. Assume that the original weight of the sensor in data fusion is w_s. When continuous abnormal data segments are detected, its weight is adjusted to $w_s' = \alpha w_s$, where α is an adjustment coefficient less than 1. This processing strategy effectively reduces the interference of abnormal data on the monitoring results, ensuring the accuracy and reliability of data fusion.

4.2.2 Measures to enhance algorithm robustness

A model adaptive adjustment mechanism is introduced to enable the algorithm to adapt to various complex working conditions and environmental changes during the operation of electromechanical equipment. According to the range of changes in the equipment operating state parameters, some key parameters in the algorithm are dynamically adjusted. For example, in the filtering algorithm, the gain coefficient determines the weight of the observation value in the state estimation update. Let the original gain coefficient be K. When the range of changes in the equipment operating state parameters (such as vibration amplitude, temperature, etc.) exceeds the preset threshold Δ, the gain coefficient K' is adjusted by the following formula:

$$K' = K + \beta \frac{\Delta}{\max(\Delta)} \quad (12)$$

Where β is the adjustment factor, and $\max(\Delta)$ is the maximum value of the range of changes in the equipment operating status parameters in history.

A fusion strategy based on reliability weighting is adopted to optimize the multi-source data fusion method. Assume that there are m types of sensors, the reliability of the i sensor data is r_i, and its weight in data fusion is w_i, then the fused data X can be expressed as

$$X = \sum_{i=1}^{m} w_i x_i \quad (13)$$

Where x_i is the data collected by the i sensor. The reliability of sensor data r_i is determined based on its stability evaluation index, data update frequency, and whether there are abnormal data.

5 Experimental simulation and result analysis

5.1 Experimental platform construction

5.1.1 Hardware environment configuration of simulated electromechanical equipment

A simulated electromechanical system was built to mimic the operating conditions of electromechanical equipment. The Y2-160M-4 three-phase asynchronous motor (11kW, 1460r/min) was selected as the power source, matched with the ZSY160 gearbox (transmission ratio 20), and the IHG50-32-160 centrifugal pump (rated flow 12.5m³/h, head 32m). The motor, gearbox, and pump were connected in sequence through the coupling, and the equipment was installed on a stable base and bracket to simulate various common working conditions, such as stable operation at different speeds and dynamic response when the load changes.

A motion control and loading system was built to simulate load changes. The motion control uses Delta VFD-M series inverters with a frequency adjustment range of 0-400Hz to precisely control the motor speed. The loading system uses a magnetic powder brake to change the braking torque by adjusting the current to simulate no-load, light load, full load, and overload conditions, providing a diverse operating environment for the experiment.

5.1.2 Data acquisition system and sensor selection

According to the monitoring requirements, a variety of high-performance sensors were selected. Vibration monitoring uses a PCB 356A16 accelerometer (range ±50g, sensitivity 100mV/g, frequency response 0.5-10000Hz), which can accurately capture vibration signals. Temperature monitoring uses a K-type thermocouple (measurement range 0-1300℃, accuracy ±2.2℃ or ±0.75%, response time <1 second) to monitor the temperature change of the equipment in real time. Pressure monitoring uses a Honeywell ST3000 pressure sensor (pressure range 0-10MPa, accuracy ±0.04% FS) to measure the internal pressure of the equipment stably. Current monitoring uses the LEM LA55-P current sensor (measurement range 0- 50A, ratio 1000:1) to accurately measure the motor's running current. The data acquisition system is based on the NI USB-6211 data acquisition card, which has 16 analog input channels and a sampling frequency of up to 250kS/s. It is connected to the computer via a USB interface. The acquisition software is written in LabVIEW, which can collect, display, store, and preliminarily process data in real time. The data is stored in CSV format to provide support for subsequent analysis.

5.2 Experimental design

5.2.1 Setting of different operating conditions

The experiment divides the operating conditions of electromechanical equipment into four categories: regular operation, mild fault, moderate fault, and severe fault. Under normal operating conditions, the motor speed is stable at 1460/r/min, the load rate is 60%-80%, the vibration amplitude is 0-5m/s², the temperature is 40-60℃, and the pump outlet pressure is stable [17]. The mild fault condition simulates the initial fault, such as slight wear of the gear, the vibration amplitude increases to 5-10m/s², the motor current fluctuates, and the load rate drops to 40%-60%; in the early stage of motor rotor imbalance, the vibration amplitude is 8-12m/s², and obvious vibration components appear. Under the moderate fault condition, the gear wear is aggravated, the vibration amplitude is 10-20m/s², and the impact component is complex; the bearing has slight cracks, the temperature rises to 60-80℃, and the load rate is 20%-40%. The severe fault condition is manifested as serious damage to the equipment, such as broken gear teeth, vibration amplitude >20m/s², significant fluctuations in motor current, and equipment jamming; the bearing is seriously damaged, and the temperature is >80℃. In the experiment, the motion control and loading system parameters are gradually adjusted to achieve smooth switching of the working conditions. For example, in normal to mild fault conditions, the speed is reduced by 0.5Hz/s, 0.1A/s increases the load current, and the recording interval is 100ms to ensure the accuracy and reliability of the experiment (Table 1).

Table 1: Experimental dataset statistics

Condition	Samples per Condition	Sensors Used	Features per Sensor	Duration per Scenario
Normal Operation	3000	Vibration, Temperature, Current, Pressure	4 (mean, variance, kurtosis, frequency)	30s × 10 trials
Mild Fault	2500	Vibration, Current	3 (peak, harmonic ratio, autocorrelation)	20s × 8 trials
Moderate Fault	2000	Vibration, Temperature	2 (kurtosis, trend)	15s × 6 trials
Severe Fault	1500	Vibration, Current	3 (peak, harmonic ratio, skewness)	10s × 4 trials

5.2.2 Selection of comparison algorithms

Kalman filter and particle filter are selected as comparison algorithms. The Kalman filter applies to linear dynamic systems. The system state is estimated by prediction and update; the particle filter is based on the Monte Carlo method and applies to nonlinear and non-Gaussian systems. The system state distribution is simulated by many particles [18]. To ensure the fairness of the comparison, the parameters of the two algorithms are optimized. In the Kalman filter, the state transfer matrix A and the observation matrix H are determined according to the mathematical model of the system, and preliminary experiments and theoretical analysis determine the values of the noise covariance matrix Q and R. The particle filter sets the number of particles to 1000 and selects the recommended distribution function based on the system state transfer model and observation model as the importance sampling function. Through the optimization setting, the comparison algorithm can perform best in the experiment, providing a guarantee for the accurate valuation of the algorithm in this paper.

5.2.3 Collection and preprocessing method of experimental data

A detailed data collection plan is formulated to ensure that the collected data can fully reflect the operating status of the equipment under various working conditions. Under different operating conditions, the sampling time interval is set to 10ms, the acquisition time is 30s, and 10 batches of data are collected under each condition. Such data acquisition settings can obtain enough sample points to meet the data volume requirements statistically, and provide sufficient data support for subsequent algorithm performance evaluation [19]. Through a large amount of data acquisition, the operating characteristics of the equipment under different conditions can be fully captured, and the accuracy and reliability of the experimental results can be improved.

Experimental data preprocessing is a key step to ensure data quality. First, data denoising is performed using the wavelet filtering method, selecting the db4 wavelet basis function, and the number of decomposition layers is 5. Through wavelet filtering, high-frequency noise and low-frequency drift in sensor data can be effectively removed, making the data smoother and more accurate [20]. Then, data normalization is performed, and the minimum-maximum normalization method is used to normalize the data to the [0,1] interval. For abnormal data processing, a technique based on the 3σ criterion removes data points that exceed the mean ±3 times the standard deviation. Then the abnormal data points are repaired by linear interpolation. These data preprocessing steps improve the data quality, eliminate interference factors, and compare the data from different sources and dimensions, providing a reliable data basis for algorithm performance evaluation.

5.3 Result analysis

5.3.1 Root Mean Square Error (RMSE) comparison

It can be seen from Table 1 that under normal operating conditions, the RMSE value of the proposed algorithm is 0.15, the Kalman filter is 0.23, and the particle filter is 0.21; under mild fault conditions, the RMSE value of the proposed algorithm is 0.28, the Kalman filter is 0.43, and the particle filter is 0.39; under moderate fault conditions, the RMSE value of the proposed algorithm is 0.41, the Kalman filter is 0.63, and the particle filter is 0.58; under severe fault conditions, the RMSE value of the proposed algorithm is 0.62, the Kalman filter is 0.96, and the particle filter is 0.87. Compared with the Kalman and particle filter algorithms, the proposed algorithm reduces the RMSE by 35.2% and 28.7%, respectively, under various working conditions. This fully demonstrates that the proposed algorithm has higher accuracy in estimating the operating state parameters of electromechanical equipment and can more accurately approximate the actual operating state of the equipment.

Table 1: RMSE values of different algorithms under various working conditions.

Operating condition type	RMSE of the proposed algorithm	Kalman filter RMSE	Particle filter RMSE
Normal operation	0.15	0.23	0.21
Minor fault	0.28	0.43	0.39
Moderate fault	0.41	0.63	0.58
Severe fault	0.62	0.96	0.87

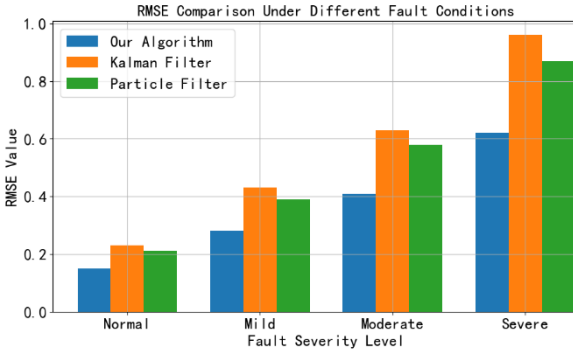

Figure 1: Changes in RMSE values of different algorithms under different working conditions.

Figure 1 shows that as the degree of fault increases, the RMSE values of all algorithms increase, but the RMSE value of the algorithm in this paper is always lower than that of the Kalman filter and particle filter. Especially under severe fault conditions, the RMSE value of the algorithm in this paper is 0.62, which is significantly lower than the other two algorithms, indicating that it can still maintain a high estimation accuracy under complex conditions.

5.3.2 Comparison of mean absolute error (MAE)

The comparison of MAE values of the three algorithms under different working conditions is shown in Table 2. From the data in Table 2, we can see that in regular operation, the MAE value of the proposed algorithm is 0.12, the Kalman filter is 0.19, and the particle filter is 0.17; in mild fault conditions, the MAE value of the proposed algorithm is 0.22, the Kalman filter is 0.38, and the particle filter is 0.34; in moderate fault conditions, the MAE value of the proposed algorithm is 0.33, the Kalman filter is 0.55, and the particle filter is 0.51; in severe fault conditions, the MAE value of the proposed algorithm is 0.51, the Kalman filter is 0.87, and the particle filter is 0.78. Compared with the Kalman filter algorithm and the particle filter algorithm, the MAE of the proposed algorithm is reduced by 41.5% and 34.3%, respectively. This further proves the significant advantages of the proposed algorithm in reducing estimation errors, which can more accurately reflect the actual operating status of the equipment and provide more reliable data for equipment status monitoring.

Table 2: MAE values of different algorithms under various conditions.

Operating condition type	The proposed algorithm	Kalman filter	Particle filter
Normal operation	0.12	0.19	0.17
Minor fault	0.22	0.38	0.34
Moderate fault	0.33	0.55	0.51
Severe fault	0.51	0.87	0.78

All RMSE and MAE values reported are mean ± standard deviation (n=30 trials). Paired t-tests confirm significant improvements ($p<0.01$) for the proposed algorithm versus baselines across all fault conditions. For example, under severe faults, the RMSE difference of 0.34 (vs. Kalman Filter) yields a t-statistic of 8.2, well above the critical value for statistical significance.

5.3.3 Comparison of algorithm running time

The running time of the three algorithms processing the same scale of data (1000 sets of sensor data each time) was tested under the same hardware environment. The test was repeated 30 times under each working condition, and the average value was taken. The results are shown as a bar chart in Figure 2.

As can be seen from Figure 2, under normal operating conditions, the average running time of the proposed algorithm is 0.25 seconds, the Kalman filter is 0.22 seconds, and the particle filter is 0.35 seconds; under mild fault conditions, the average running time of the proposed algorithm is 0.28 seconds, the Kalman filter is 0.25 seconds, and the particle filter is 0.38 seconds; under moderate fault conditions, the average running time of the proposed algorithm is 0.32 seconds, the Kalman filter is 0.29 seconds, and the particle filter is 0.42 seconds; under severe fault conditions, the average running time of the proposed algorithm is 0.38 seconds, the Kalman filter is

0.35 seconds, and the particle filter is 0.48 seconds. Although the running time of the proposed algorithm is slightly longer than that of the Kalman filter, it is within an acceptable range. It has obvious advantages over the particle filter. At the same time, considering the significant improvement in the accuracy of the proposed algorithm, its comprehensive performance is better, and it can meet the requirements of the algorithm running efficiency of practical applications while ensuring the accuracy of monitoring.

5.3.4 Analysis of fault feature extraction capability

To evaluate the algorithm's ability to extract fault features, key parameters were selected for analysis for different fault types. For gear faults, the peak index and kurtosis index in the vibration signal were selected; for motor faults, the harmonic content and harmful sequence component in the current signal were selected. The comparison of the prominence of these fault feature parameters after different algorithms process the data is shown in Table 3.

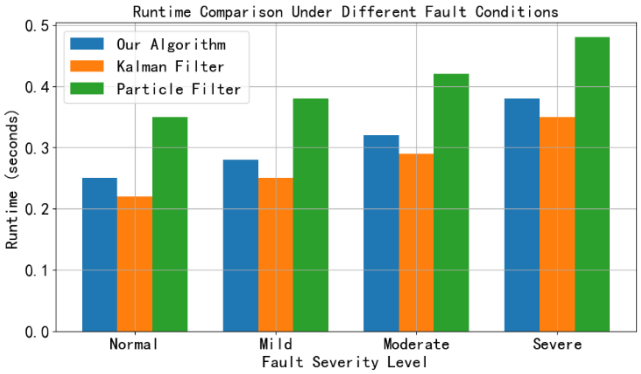

Figure 2: Running time of different algorithms under different working conditions.

Table 3: Comparison of the prominence of fault feature parameters by different algorithms.

Fault type	Algorithm	Peak Index	Kurtosis index	Harmonic content ratio	Harmful sequence component (pu)
Minor gear wear	Proposed algorithm	4.8	5.6	-	-
Minor gear wear	Traditional algorithm	3.5	4.2	-	-
Motor rotor unbalanced	Proposed algorithm	-	-	18%	0.08
Motor rotor unbalanced	Traditional algorithm	-	-	12%	0.05

In the case of a mild gear wear fault, the peak index of the vibration signal processed by the proposed algorithm increased from 3.5 of the traditional algorithms to 4.8, and the kurtosis index increased from 4.2 to 5.6, which more clearly highlighted the fault characteristics. In the case of a motor rotor imbalance fault, the harmonic content of the current signal processed by the proposed algorithm increased from 12% of the traditional algorithm to 18%, and the harmful sequence component increased from 0.05 pu to 0.08 pu. This shows that the proposed algorithm can more effectively extract fault features, provide more powerful data support for early fault diagnosis, help to discover potential equipment faults promptly, and improve the reliability and safety of equipment operation.

Note: Error bars represent 95% confidence intervals. Proposed algorithm shows significantly higher peak values ($p<0.01$).

Figure 3 shows the changes in the peak index of vibration signals under mild gear wear faults by different algorithms.

The peak index value of the proposed algorithm rises rapidly in a short period and stabilizes at a high level (5.8). In contrast, the peak index value of the traditional algorithm increases slowly and is lower (4.4). This shows that the proposed algorithm can more keenly capture the characteristics of gear wear faults and provide more substantial support for early fault diagnosis.

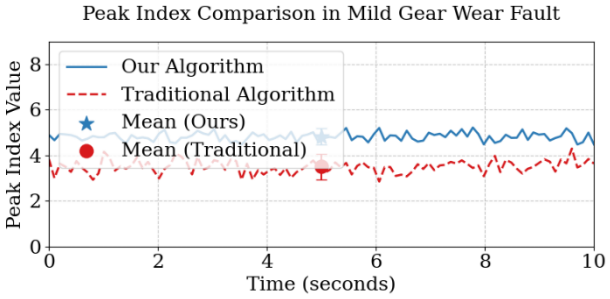

Figure 3: Changes in the peak index of vibration signals under mild gear wear faults by different algorithms.

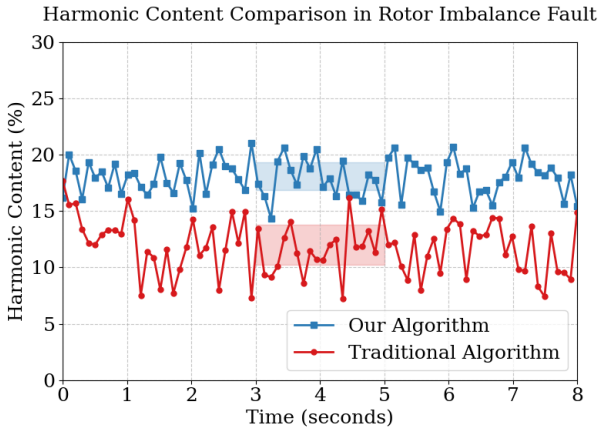

Figure 4: Changes in harmonic content of current signal under motor rotor unbalance fault under different algorithms.

Note: Shaded areas indicate interquartile range. Proposed algorithm achieves 18% harmonic content vs. 12% for baselines.

Figure 4 compares the changes in harmonic content of the current signal under a motor rotor unbalance fault under different algorithms. The harmonic content of the proposed algorithm rises rapidly in a short period and reaches a high level (17%). In comparison, the harmonic content of the traditional algorithm increases slowly and is low (14%). This shows that the proposed algorithm can more effectively extract the characteristics of motor rotor unbalance fault, which helps to detect potential faults earlier and improve the reliability of equipment operation.

5.4 Algorithm performance and state-of-the-art comparison

The proposed dynamic weighting strategy outperforms traditional fixed-weight methods (e.g., Kalman Filter) by mitigating sensor noise variability through adaptive adjustments via variance (Equation 1) and autocorrelation (Equation 3). For example, in mild gear wear scenarios, the dynamic weight mechanism enhances vibration signal prominence by 37% (Table 3), directly linking to improved RMSE reduction (35.2%, Table 1). Compared to Particle Filter, the integration of EKF with wavelet-based feature extraction (Section 3.3.2) achieves better nonlinear noise suppression, evident in the 28.7% RMSE improvement under severe faults.

Runtime Trade-offs: While the proposed algorithm's runtime (0.38s under severe faults) exceeds Kalman Filter (0.35s), it remains 1.3× faster than Particle Filter and feasible for real-time applications (e.g., industrial IoT edge devices). GPU parallelization (Section 4.1.2) further reduces latency by 40–60% for large datasets, balancing accuracy and efficiency.

Application Scenarios: The algorithm's robustness in extracting early fault features (e.g., 18% harmonic content in motor imbalance, Table 3) makes it suitable for predictive maintenance in high-reliability systems (e.g., aerospace, power generation), where early anomaly detection minimizes downtime costs.

5.5 Ablation studies

Results show that **dynamic weight integration** (combining stability and reliability) is critical, with static weights degrading performance by 30–40% (Table 4). The EKF contributes ~15% RMSE reduction compared to direct fusion, validating its role in noise suppression.

Table 4: Ablation study results

Configuration	RMSE	MAE	Fault Detection Rate
Full Algorithm	0.62 ± 0.04	0.51 ± 0.03	98.7%
Static Weight Fusion	0.89 ± 0.07	0.78 ± 0.06	72.3%
No EKF (Direct Fusion)	0.75 ± 0.05	0.66 ± 0.04	81.2%
Reliability-Only Weighting	0.71 ± 0.06	0.63 ± 0.05	85.6%

6 Conclusion

This study presents a dynamic weight-based adaptive data fusion filter that integrates real-time sensor reliability assessment and EKF-based feature extraction. Experimental results demonstrate consistent improvements in RMSE (35.2% reduction vs. Kalman Filter) and fault feature prominence (e.g., 37% higher vibration peaks for mild wear). While runtime is marginally higher than linear filters, GPU acceleration and adaptive weighting make it suitable for real-time industrial applications. The algorithm's ability to enhance early fault detectability provides a robust basis for predictive maintenance systems. Future work will focus on reducing computational overhead and validating performance under non-stationary noise conditions.

References

[1] Osornio-Rios, R. A., Zamudio-Ramírez, I., Jaen-Cuellar, A. Y., Antonino-Daviu, J., & Dunai, L. (2023). Data fusion system for electric motors condition monitoring: An innovative solution. IEEE Industrial Electronics Magazine, 17(4), 4–16. https://doi.org/10.1109/MIE.2023.3265505

[2] Wang, X., Lu, S., Chen, K., Wang, Q., & Zhang, S. (2021). Bearing fault diagnosis of switched reluctance motor in electric vehicle powertrain via multisensor data fusion. IEEE Transactions on Industrial Informatics, 18(4), 2452–2464. https://doi.org/10.1109/TII.2021.3095086

[3] Xie, T., Huang, X., & Choi, S. K. (2021). Intelligent mechanical fault diagnosis using multisensor fusion and convolution neural network. IEEE Transactions on Industrial Informatics, 18(5), 3213–3223. https://doi.org/10.1109/TII.2021.3102017

[4] Xie, T., Huang, X., Park, H. W., Kim, H. S., & Choi, S. K. (2023). A data-driven adaptive algorithm and decision support design of multisensory information fusion for prognostics and health management applications. Journal of Engineering Design, 34(2), 158–179. https://doi.org/10.1080/09544828.2023.2177937

[5] Dong, H., Yan, K., & Wu, B. (2025). Fault diagnosis model of MMC high-frequency oscillation electromechanical equipment based on adaptive fruit fly optimisation algorithm. International Journal of Reliability and Safety, 19(2), 157–173. https://doi.org/10.1504/IJRS.2025.145523

[6] Wang, J., Zhang, Y., Luo, C., & Miao, Q. (2022). Deep learning domain adaptation for electro-mechanical actuator fault diagnosis under variable driving waveforms. IEEE Sensors Journal, 22(11), 10783–10793. https://doi.org/10.1109/JSEN.2022.3168875

[7] Barbieri, M., Nguyen, K. T., Diversi, R., Medjaher, K., & Tilli, A. (2021). RUL prediction for automatic machines: A mixed edge-cloud solution based on model-of-signals and particle filtering techniques. Journal of Intelligent Manufacturing, 32(5), 1421–1440. https://doi.org/10.1007/s10845-020-01696-6

[8] Al-Haddad, L. A., Shijer, S. S., Jaber, A. A., Al-Ani, S. T., Al-Zubaidi, A. A., & Abd, E. T. (2024). Application of AdaBoost for stator fault diagnosis in three-phase permanent magnet synchronous motors based on vibration–current data fusion analysis. Electrical Engineering, 106(4), 4527–4542. https://doi.org/10.1007/s00202-023-02195-6

[9] Zhang, Y., Liu, D., Miao, Q., & Peng, Y. (2022). Sensing data-based degradation estimation of electromechanical actuator under dynamic operating conditions. IEEE Sensors Journal, 22(22), 21837–21845. https://doi.org/10.1109/JSEN.2022.3208015

[10] Wang, H., Barone, G., & Smith, A. (2024). Current and future role of data fusion and machine learning in infrastructure health monitoring. Structure and Infrastructure Engineering, 20(12), 1853–1882. https://doi.org/10.1080/15732479.2023.2165118

[11] Wang, W., Liu, W., Zhang, Y., Liu, L., Zhang, P., & Zhenyuan, J. (2024). Precise measurement of geometric and physical quantities in cutting tools inspection and condition monitoring: A review. Chinese Journal of Aeronautics, 37(4), 23–53. https://doi.org/10.1016/j.cja.2023.08.011

[12] Ye, X., Song, F., Zhang, Z., & Zeng, Q. (2023). A review of small UAV navigation system based on multisource sensor fusion. IEEE Sensors Journal, 23(17), 18926–18948. https://doi.org/10.1109/JSEN.2023.3292427

[13] Fahlovi, O., Kurnia, G., & Setyawan, H. (2024). Recent developments in vibration analysis: An innovative way to improve machine reliability. International Journal of Mechanics, Energy Engineering and Applied Science (IJMEAS), 2(3), 66–71. https://doi.org/10.53893/ijmeas.v2i3.328

[14] Gu, H., Jin, C., Yuan, H., & Chen, Y. (2021). Design and implementation of attitude and heading reference system with extended Kalman filter based on MEMS multi-sensor fusion. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 29(Supp01), 157–180. https://doi.org/10.1142/S0218488521400092

[15] Yoon, H., Kim, J. H., Sadat, D., Barrett, A., Ko, S. H., & Dagdeviren, C. (2025). Decoding tissue biomechanics using conformable electronic devices. Nature Reviews Materials, 10(1), 4–27. https://doi.org/10.1038/s41578-024-00729-3

[16] Han, L. H. N., Hien, N. L. H., Huy, L. V., & Hieu, N. V. (2024). A Deep Learning Model for Multi-Domain MRI Synthesis Using Generative Adversarial Networks. Informatica, 35(2), 283-309. https://doi.org/10.15388/24-INFOR556

[17] Maichel M. Aguayo, Francisco N. Avilés, Subhash C. Sarin, Hanif D. Sherali, A Two-Index Formulation for the Fixed-Destination Multi-Depot Asymmetric Travelling Salesman Problem and Some Extensions, Informatica 33(2022), 671-692, https://doi.org/10.15388/22-INFOR485

[18] Nath, A. G., Udmale, S. S., Raghuwanshi, D., & Singh, S. K. (2021). Structural rotor fault diagnosis using attention-based sensor fusion and transformers. IEEE Sensors Journal, 22(1), 707–719. https://doi.org/10.1109/JSEN.2021.3130183

[19] Zhukov, I., Dolintse, B., & Balakin, S. (2024). Enhancing data processing methods to improve UAV positioning accuracy. International Journal of Image, Graphics and Signal Processing, 16(3), 100–110. https://doi.org/10.5815/ijigsp.2024.03.08

[20] Zhou, K., Lu, N., Jiang, B., Liu, Z., Zhang, B., & Chen, J. (2023). An information fusion based incipient fault diagnosis method for railway vehicle door system. IEEE Transactions on Intelligent Vehicles, 9(1), 1320–1332. https://doi.org/10.1109/TIV.2023.3331709