Research Article

A Novel Mittag-Leffler Kernel Based Hybrid Fault Diagnosis Method for Wheeled Robot Driving System

Xianfeng Yuan,[1] Mumin Song,[1] Fengyu Zhou,[1] Zhumin Chen,[2] and Yan Li[1]

[1]*School of Control Science and Engineering, Shandong University, Jinan 250061, China*
[2]*School of Computer Science and Technology, Shandong University, Jinan 250101, China*

Correspondence should be addressed to Fengyu Zhou; zhoufengyu@sdu.edu.cn

The wheeled robots have been successfully applied in many aspects, such as industrial handling vehicles, and wheeled service robots. To improve the safety and reliability of wheeled robots, this paper presents a novel hybrid fault diagnosis framework based on Mittag-Leffler kernel (ML-kernel) support vector machine (SVM) and Dempster-Shafer (D-S) fusion. Using sensor data sampled under different running conditions, the proposed approach initially establishes multiple principal component analysis (PCA) models for fault feature extraction. The fault feature vectors are then applied to train the probabilistic SVM (PSVM) classifiers that arrive at a preliminary fault diagnosis. To improve the accuracy of preliminary results, a novel ML-kernel based PSVM classifier is proposed in this paper, and the positive definiteness of the ML-kernel is proved as well. The basic probability assignments (BPAs) are defined based on the preliminary fault diagnosis results and their confidence values. Eventually, the final fault diagnosis result is archived by the fusion of the BPAs. Experimental results show that the proposed framework not only is capable of detecting and identifying the faults in the robot driving system, but also has better performance in stability and diagnosis accuracy compared with the traditional methods.

1. Introduction

In recent years, the wheeled robots have received a wide range of applications and developments [1–3]. Particularly, in home service area, various kinds of wheeled service robots have become members of our family, such as the elderly companion robot [4] and the sweeping robot [5]. However, robot users are usually nonexpert in robot technology, which means that the faults which occurred in the wheeled robot system may cause serious damage to their life and property. The increasing demand of safety, reliability, and the necessity of low cost have become the bottleneck of wheeled robot applications with current technology. Therefore, it is meaningful to focus on novel fault diagnosis methods, particularly for the man-robot coexistent environments.

Generally speaking, the existing fault diagnosis methods can be classified as the model based and the data driven ones [6, 7]. In the earlier days, the research of model based fault diagnosis methods drew much attention and constituted the mainstream of this field [8, 9]. In [10], based on the mathematical model of the robotic manipulator, Caccavale et al. presented a discrete-time framework for diagnosis of sensors and actuators of robotic manipulators. Using particle filter, Yu et al. [11] proposed a fault-proneness prediction method for robot dead reckoning system. Besides the abovementioned methods, the adaptive observer and some other model based methods have also been designed for fault diagnosis of robot platform or robot manipulator [12, 13]. Those model based fault diagnosis methods are effective and suitable for the diagnosis problem of robot manipulator or robot arm, because robot arm usually works in a structured environment and it is relatively easier to get the accurate mathematical model. While, for wheeled robots, firstly, their working environments are usually dynamic and unstructured, secondly, wheeled robots are usually equipped with various kinds of equipment that are more complex in both hardware and software aspects compared with manipulators. Thus it is hard to get an accurate mathematical model of a wheeled robot working in an unstructured environment, which becomes a restriction of those model based methods.

Moreover, the wheeled robots are well equipped with multiple sensors which implies that large data volumes containing robot running status information are available. Those large data volumes imply difficulties in system modeling, while they provide the required information for data driven based fault diagnosis method.

Principal component analysis (PCA) is a typical representative of the data driven fault diagnosis method. PCA is more suitable for fault detection rather than diagnosis, because it does not use the input-output relationships [14]. Therefore, in order to improve the diagnosis ability, PCA is often used by combining the classifiers, such as the neural network (NN) and the support vector machine (SVM). This hybrid method has been applied in the fault diagnosis of rotating machinery [15], power transmission systems [16], and some other aspects [17, 18]. Applications of PCA could be useful in extracting and interpreting process information from massive data sets, and the pattern recognition techniques could also be used to diagnose the specific running status of the robot.

Nevertheless, there are mainly two problems that exist in the above hybrid diagnosis methods. On the one hand, most of the studies adopted the existing classical kernel (e.g., Gaussian kernel and polynomial kernel) as the kernel function of SVM in their diagnosis methods, while new kernel functions with better classification performance need to be proposed, proved, and applied to the robot fault diagnosis fields. On the other hand, the diagnosed objects are usually complex and with varying degrees of uncertainties. A single PCA model cannot achieve full and complete awareness of the diagnosed object so that the information fusion in data level or decision level is needed to reduce the existing uncertainties.

Mittag-Leffer functions [19, 20] play fundamental roles in fractional calculus, which exhibit intermediate process among exponential function, power function, and polynomial function. Nowadays, fractional calculus has been successfully applied in many aspects, such as the application of fractional Fourier transform in signal processing [21] and the application of fractional order PI controllers [22]. Inspired by fractional calculus, a novel fractional Gaussian kernel named ML-kernel is proposed in this paper, which is a generalization of the traditional Gaussian kernel. The proposed ML-kernel is proved to be positive definite and its diagnosis performance is discussed in this paper. Besides, a hybrid fault diagnosis framework is discussed for robot driving system based on Dempster-Shafer (D-S) fusion and ML-kernel support vector machine (SVM). Multiple PCA models are established to do fault feature extraction and the fault feature vectors are used as the inputs of the ML-kernel SVM classifiers. The ML-kernel SVM classifiers output the preliminary fault diagnosis results which are fused by D-S fusion and the fusion result is taken as the final diagnosis result. Two sets of comparative experiments are carried out to validate the proposed method.

The remainder of this paper is organized as follows. Section 2 briefly introduces the SVM method and the positive definiteness of the presented ML-kernel is also proved in this section. In Section 3, the proposed fault diagnosis framework is described in detail. Section 4 illustrates the architecture of the experimental wheeled robot driving system and the application studies for various fault conditions. Section 5 is devoted to conclusions.

2. SVM Algorithm and the Presented ML-Kernel Function

2.1. Conventional SVM Algorithm. In the past few years, SVM has been one of the most highly studied topics in the machine learning fields, and it has been successfully applied in practice, especially for classification problems (e.g., fault diagnosis) [23, 24]. Based on the statistical theory of VC dimension and structural risk minimization inductive principle, SVM reaches the best compromise between the complexity of modeling and the leaning ability and hunts the best generalization ability. The basic SVM [25] deals with linearly separable two class cases and it can cope with nonlinear problems by introducing kernel functions and slack penalty. Given a training set $S_t = \{x_i, y_i\}_{i=1}^N$, where $x_i \in R^d$ is the ith training input vector, N is the number of training data for SVM, d is the dimension of the input data, and $y_i \in \{-1, 1\}$ is the set of classification tag for training. The optimal hyperplane separating the data can be obtained as a solution to the following optimization problem:

$$\min_{\omega, b, \xi_i} \quad \left(\frac{1}{2} \omega^T \omega + c \sum_{i=1}^N \xi_i \right),$$
$$\text{s.t.} \quad y_i \left(\omega^T x_i + b \right) \geq 1 - \xi_i, \quad \xi_i \geq 0, \ i = 1, 2, \ldots, N, \tag{1}$$

where c is the slack penalty, ω is the adjustable weight vector, b is the offset of the hyperplane, and ξ_i is the distance between the margin and the x_i lying on the wrong side. The equivalent Lagrangian dual problem can be described as follows:

$$\min_{\alpha_i} \quad \left(\sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \alpha_i \alpha_j x_i^T x_j \right),$$
$$\text{s.t.} \quad \sum_{i=1}^N \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, \ i = 1, 2, \ldots, N, \tag{2}$$

where α_i is the Lagrangian coefficient, from which we can obtain $\omega = \sum_{i=1}^N \alpha_i y_i x_i$, $b = y_i - \omega^T x_i$, to solve (1).

The kernel function can map the input vector x into feature space and returns a dot product of the feature space. The linear discriminant function with kernel $K(x_i, x_j)$ is given by the following:

$$f(x) = \text{sgn} \left(\sum_{i,j=1}^N a_i y_i K\left(x_i, x_j\right) + b \right), \tag{3}$$

where $\text{sgn}(\cdot)$ is the signum function.

The fault diagnosis of a robot driving system is a multiple class classification problem, while the conventional SVM was designed for the binary classification problem, so it is not suitable for the fault diagnosis in its original form. A few types of methods for multiclass SVM have been proposed [26]: one against one, one against others, direct acyclic graph,

and so forth. This study employs the "one against one" multiclass SVM. In order to construct the BPAs, we need the probabilistic outputs of the SVM classifiers and the "pairwise coupling" method [27] is used to solve this problem.

2.2. Kernel Function.

The nonlinear pattern recognition problem in fault diagnosis can be transformed into the linear problem in some very high-dimensional feature spaces. The kernel function $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is able to handle any dimension feature spaces without the accurate calculation of $\phi(x_i)$ and $\phi(x_j)$. It has been proven that any function that satisfies the Mercer theorem can be used as a kernel function [28]. Currently, there are three typical kinds of kernel functions:

(1) Polynomial kernel function

$$K\left(x_i, x_j\right) = \left[\left(x_i, x_j\right) + 1\right]^d. \tag{4}$$

(2) Radial basis kernel function (RBF)

$$K\left(x_i, x_j\right) = \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2\delta^2}\right). \tag{5}$$

(3) Sigmoid kernel function

$$K\left(x_i, x_j\right) = \tanh\left(x_i^T x_j + b\right). \tag{6}$$

2.3. Proof of the Positive Definiteness of the ML-Kernel.

As the core of SVM, kernel function and the parameters of the model determine the performance of the SVM algorithm applied to the fault diagnosis system. In this paper, we employ the Mittag-Leffler function as a novel kernel function named as ML-kernel. The Mittag-Leffler function [29] is defined as follows:

$$E_\alpha(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(k\alpha + 1)}, \tag{7}$$

where Γ is the Gamma function and α is an arbitrary positive constant. For $\alpha = 1$, (7) becomes $E_1(z) = e^z$. The presented ML-kernel function can be defined as

$$K\left(x_i, x_j\right) = E_\alpha\left(-\frac{\left\|x_i - x_j\right\|^{2\alpha}}{2\delta^{2\alpha}}\right), \tag{8}$$

where $0 < \alpha \le 1$. When $\alpha = 1$, (8) becomes $E_1(-\|x_i - x_j\|^2/2\delta^2) = \exp(-\|x_i - x_j\|^2/2\delta^2)$, and the ML-kernel is identical to the Gaussian RBF kernel in this condition.

Given a kernel, it is in general straightforward to verify its continuity and symmetry, while the positive definiteness is more important and essential for a kernel. Thus, the proof of the positive definiteness of the proposed ML-kernel is given as below.

For convenience, letting $t = \|x_i - x_j\|^2/2\delta^2$, the ML-kernel (8) can be written as $K(x_i, x_j) = E_\alpha(-t^\alpha)$. In Laplace domain, we can get [30]

$$
\begin{aligned}
K\left(x_i, x_j\right) &= E_\alpha\left(-t^\alpha\right) = \mathscr{L}^{-1}\left\{\frac{s^{\alpha-1}}{s^\alpha + 1}\right\} \\
&= \frac{1}{2\pi i} \lim_{T \to \infty} \int_{\gamma - iT}^{\gamma + iT} \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds,
\end{aligned}
\tag{9}
$$

where γ is a real number that keeps the contour path of integration which is in the region of convergence of $s^{\alpha-1}(s^\alpha + 1)$, $i^2 = -1$, and $\mathscr{L}^{-1}(\cdot)$ denotes the inverse Laplace transformation.

The integration path in (9) can be bended into the equivalent Hankel contour $H_a(\delta)$, which contains three parts: one line L_1 that starts from $(-\infty, -\delta)$, an arc C that encircles the circular disc $|s| = \delta \le \gamma$ counterclockwise, and the other line L_2 that ends at $(-\infty, \delta)$. So, (9) can be written as

$$
\begin{aligned}
E_\alpha\left(-t^\alpha\right) &= \frac{1}{2\pi i} \int_{Ha(\delta)} \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds = I_1 + I_2 + I_3 \\
&= \frac{1}{2\pi i} \int_{L_1} \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds + \frac{1}{2\pi i} \int_C \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds \\
&\quad + \frac{1}{2\pi i} \int_{L_2} \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds.
\end{aligned}
\tag{10}
$$

Along L_1, we have $s = re^{-\pi i} = -r, \delta \le r \le \infty$, and as s goes from $-\infty$ to $-\delta$, r goes from $+\infty$ to δ,

$$I_1 = \frac{1}{2\pi i} \int_{-\infty}^{-\delta} \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds = -\frac{1}{2\pi i} \int_{\delta}^{+\infty} \frac{e^{-rt} r^{\alpha-1} e^{-\alpha\pi i}}{r^\alpha e^{-\alpha\pi i} + 1} dr. \tag{11}$$

Along L_2, we have $s = re^{\pi i} = -r, \delta \le r \le \infty$, and as s goes from $-\delta$ to $-\infty$, r goes from δ to $+\infty$:

$$I_3 = \frac{1}{2\pi i} \int_{-\delta}^{-\infty} \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds = \frac{1}{2\pi i} \int_{\delta}^{+\infty} \frac{e^{-rt} r^{\alpha-1} e^{\alpha\pi i}}{r^\alpha e^{\alpha\pi i} + 1} dr. \tag{12}$$

Along C, we have $s = \delta e^{i\theta}, \theta \in (-\pi, \pi)$, and

$$
\begin{aligned}
I_2 &= \frac{1}{2\pi i} \int_C \frac{e^{st} s^{\alpha-1}}{s^\alpha + 1} ds \\
&= \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{e^{re^{i\theta} t} \left(re^{i\theta}\right)^{\alpha-1}}{\left(re^{i\theta}\right)^\alpha + 1} rie^{i\theta} d\theta = 0.
\end{aligned}
\tag{13}
$$

Hence, we can obtain

$$I_1 + I_3$$

$$= \frac{1}{2\pi i} \int_\delta^{+\infty} \left[-\frac{r^{\alpha-1} e^{-\alpha\pi i}}{r^\alpha e^{-\alpha\pi i} + 1} e^{-rt} + \frac{r^{\alpha-1} e^{\alpha\pi i}}{r^\alpha e^{\alpha\pi i} + 1} e^{-rt} \right] dr$$

$$= \frac{1}{2\pi i} \int_\delta^{+\infty} \left[\frac{r^{\alpha-1}}{r^\alpha + e^{-\alpha\pi i}} - \frac{r^{\alpha-1}}{r^\alpha + e^{\alpha\pi i}} \right] e^{-rt} dr \qquad (14)$$

$$= \frac{1}{2\pi i} \int_\delta^{+\infty} \frac{r^{\alpha-1} \left(e^{\alpha\pi i} - e^{-\alpha\pi i} \right)}{r^{2\alpha} + r^\alpha \left(e^{\alpha\pi i} + e^{-\alpha\pi i} \right) + 1} e^{-rt} dr$$

$$= \frac{1}{\pi} \int_\delta^{+\infty} \frac{r^{\alpha-1} \sin(\alpha\pi)}{r^{2\alpha} + 2r^\alpha \cos(\alpha\pi) + 1} e^{-rt} dr.$$

Letting $\delta \rightarrow 0^+$, we have

$$E_\alpha \left(-t^\alpha \right) = I_1 + I_2 + I_3$$

$$= \frac{1}{\pi} \int_\delta^{+\infty} \frac{r^{\alpha-1} \sin(\alpha\pi)}{r^{2\alpha} + 2r^\alpha \cos(\alpha\pi) + 1} e^{-rt} dr > 0. \qquad (15)$$

Therefore $K(x_i, x_j) = E_\alpha(-\|x_i - x_j\|^{2\alpha}/2\delta^{2\alpha}) = K(x_j, x_i)$, and, for all $a \in \mathbb{R}$, we have $\sum_{i,j=1}^n a_i a_j K(x_i, x_j) \geq 0$. In other words, the proposed ML-kernel is symmetrical and positive definite. Therefore, the proof is complete.

3. Fault Diagnosis Method Based on ML-Kernel SVM and D-S Fusion

As shown in Figure 1, there are two main processes of the proposed approach, namely, the training process and the fault diagnosis process. Before the application of the proposed approach, the initial samples should be obtained from the laboratory experiments. In the training process, multiple PCA models are set up based on the data sampled in the normal and faulty states. Then, those models are used to do fault feature extraction and the ML-kernel SVM classifiers are trained. In the diagnosis process, new sampled data are normalized firstly. Secondly, the PCA models established in the training process are applied to do fault feature extraction. The fault feature vectors are then served as the inputs of the trained ML-kernel SVM classifiers, respectively, and the probabilistic outputs of the classifiers are taken as the preliminary fault diagnosis results. The BPAs are constructed based on the preliminary fault diagnosis results and the confidence values calculated from the confusion matrix. To reduce the uncertainties of the preliminary diagnosis results, the D-S fusion algorithm is introduced for decision fusion and the final diagnosis results are given based on the fusion results. The proposed approach is elaborated in detail as follows.

3.1. Data Preprocessing and Establishment of Multiple PCA Models.
Suppose that there are $h + 1$ kinds of robot running states represented as $\{S_0, S_1, \ldots, S_h\}$. S_0 represents the normal condition and $S_1, \ldots, S_h$ represent h kinds of faulty states.

Given that the robot driving system is equipped with m sensors and we get n groups of samples in each running state, so the original sampled data sets can be written as $D_{\text{all}} = [D_0, \ldots, D_k, \ldots, D_h]^T$, $D_k \in R^{n \times m}$ ($k = 0, 1, \ldots, h$). D_0 represents the samples under normal condition S_0 and D_k ($k = 1, \ldots, h$) represent the samples in the kth kind of faulty state S_k. To establish PCA models, several steps are introduced.

Step 1 (data normalization). To reduce the influence of different dimensions of the sensors, the training data should be normalized before establishing the PCA models. For a data set of n observations and m process variables $D_k \in R^{n \times m}$ ($k = 0, 1, \ldots, h$), we can get the normalized data matrix $\overline{D}_k$ by

$$M_j = \frac{1}{n} \sum_{i=1}^n d_{ij},$$

$$\sigma_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(d_{ij} - M_j \right)^2}, \qquad (16)$$

$$\overline{x}_{ij} = \frac{d_{ij} - M_j}{\sigma_j},$$

where M_j and σ_j are the mean and standard deviation, respectively, of the jth variable, d_{ij} is an element of matrix D_k, and $\overline{x}_{ij}$ is an element of the normalized data matrix $\overline{D}_k$.

Step 2 (singular value decomposition). Consider

$$C_k = P_k \Lambda_k P_k^T, \qquad (17)$$

where C_k represents covariance matrix of $\overline{D}_k$, Λ_k is a diagonal matrix containing the eigenvalues of C_k in decreasing order, and P_k is orthogonal and contains the eigenvectors of C_k.

Step 3 (determine the loading matrix according to the number of PCs). Given $\beta_i = \lambda_i / \sum_{j=1}^m \lambda_j$, the number of principal components (PCs) l is determined to satisfy the equation $\beta_1 + \beta_2 + \cdots + \beta_l \geq \mu$, where μ is a constant and usually required to be bigger than 0.85 [31].

The loading matrix $\widehat{P}_k = [p_1, p_2, \ldots, p_l]$ consists of the former l eigenvectors of the covariance matrix and $\overline{D}_k$ can be decomposed as

$$\overline{D}_k = T_s \widehat{P}_k^T + E, \qquad (18)$$

where $T_s = \overline{D}_k \widehat{P}_k$ is named as score matrix and E is the residual matrix.

3.2. Feature Extraction and SVM Training.
During the process of PCA, the orthogonal loading matrix $\widehat{P}_k$ can be considered as the main features of the original training data set. So, we can do data dimensionality reduction and feature extraction at the same time using the following equation:

$$F_k = \left(\overline{D}_{\text{all}} \cdot \widehat{P}_k \right) \in R^{(h+1)n \times l}, \quad k = 0, 1, \ldots, h, \qquad (19)$$
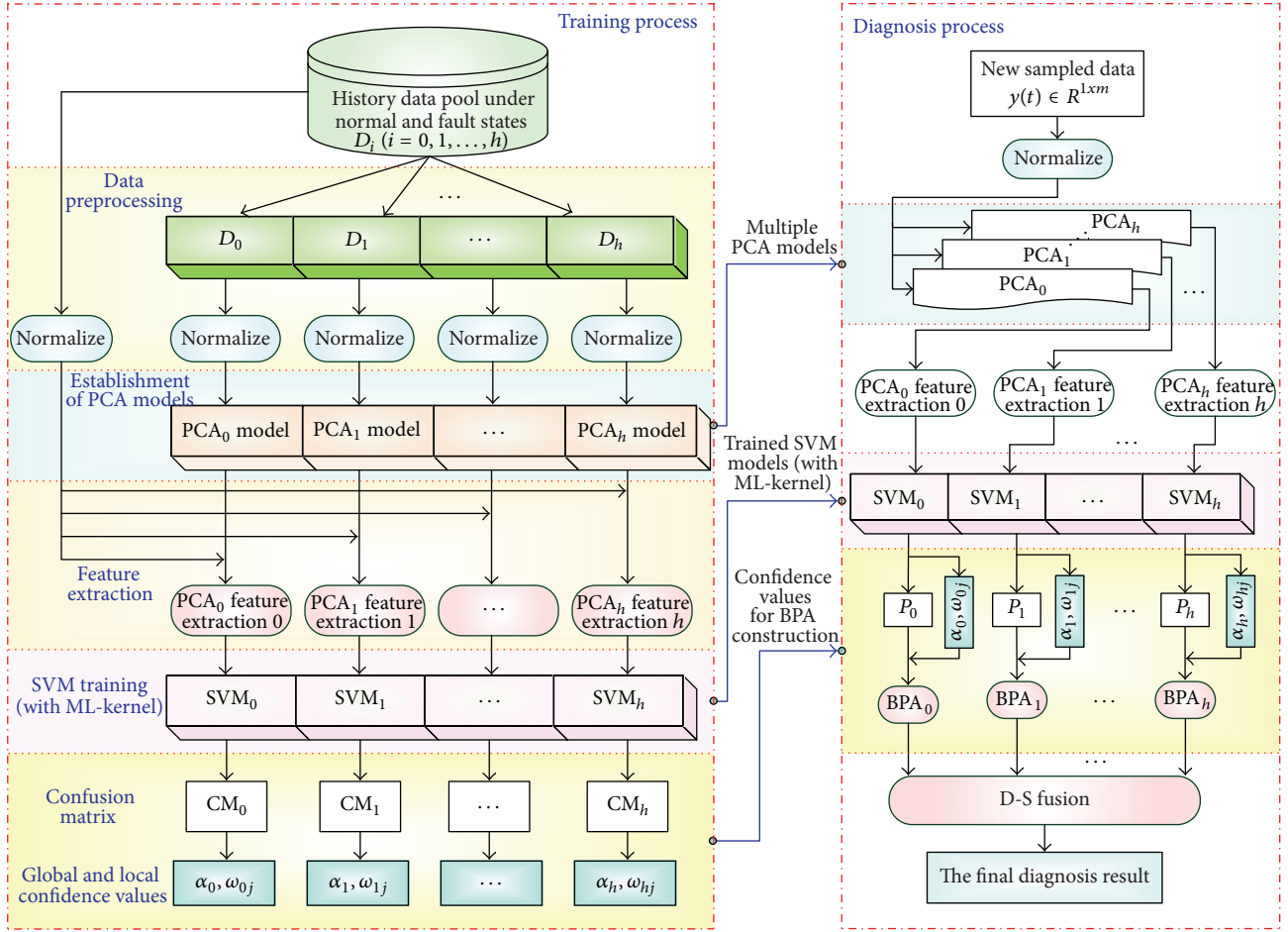
FIGURE 1: Flow diagram of the proposed fault diagnosis approach.

where $\overline{D}_{all} \in R^{(h+1)n \times m}$ is the normalized training data sets and $\widehat{P}_k \in R^{m \times l}$ is the loading matrix of the kth PCA model. l is the number of principal components and F_k is used as the final training sets of SVM_k $(k = 0, 1, \ldots, h)$.

A novel ML-kernel presented in Section 2.3 is applied as the kernel function of SVM_k and particle swarm optimization (PSO) [32] is adopted to tune the parameters c, δ, and α. Here, c is the slack penalty and δ and α are two parameters of the ML-kernel.

3.3. Decision Fusion.
To reduce the uncertainties and imprecisions of the preliminary fault diagnosis results, D-S fusion is introduced in the proposed fault diagnosis framework. The determination of BPAs is the first and most important step in evidence theory. In our approach, we construct BPAs based on the probabilistic outputs of the PSVM classifiers and their confidence values.

Step 1 (calculation of the confidence values). The average classification accuracy of SVM_k can be calculated by

$$\overline{a}_k = \frac{\sum_{p=1}^{h+1} \left(c_{pp} \cdot N_p \right)}{N}, \tag{20}$$

where c_{pp} is the diagonal elements of the confusion matrix of SVM_k, $N_p = n$ is the number of samples under the pth kind of fault condition, and $N = (h + 1)n$ is the total number of training samples in the training set F_k. So, we can get $\overline{a}_k = \sum_{p=1}^{h+1} c_{pp}/(h + 1)$, which can be used as the global confidence of SVM_k.

The qth column vector of the confusion matrix c_q $(q = 1, 2, \ldots, h + 1)$ indicates the local confidence for the qth kind of fault and the local confidence can be calculated by

$$\omega_{kq} = \frac{c_{qq}}{\sum_{p=1}^{h+1} c_{pq}}, \quad q = 1, 2, \ldots, h + 1. \tag{21}$$

Then we can incorporate the local confidence ω_{kq} into the probabilistic output of SVM_k and after normalization we can get

$$f'_{kq} = \frac{\omega_{kq} f_{kq}}{\sum_{q=1}^{h+1} \omega_{kq} f_{kq}}, \tag{22}$$

where $0 \leq f_{kq} \leq 1$ is the output of SVM_k.

Figure 2: Experimental robot.

Step 2 (construction of BPAs and D-S fusion). In our approach, the BPAs are defined as

$$m_k\left(\varnothing, S_0, S_1, \ldots, S_h, \Theta\right)$$
$$= \left(0, \overline{a}_k f'_{k0}, \overline{a}_k f'_{k1}, \ldots, \overline{a}_k f'_{kh}, 1 - \overline{a}_k\right). \quad (23)$$

It can be indicated from (23) that the frame of discernment $P(\Theta) = \{\varnothing, S_0, S_1, \ldots, S_h, \Theta\}$. Here, $\varnothing$ denotes the empty set, and S_h represents the hth kind of running condition of the robot. It is obvious that $\sum_{A \in P(\Theta)} m_k(A) = 1$; $m_k(\varnothing) = 0$. With BPAs, we use a fast fusion algorithm based on the matrix analysis [33] to accomplish D-S fusion algorithm.

4. Implementations on Wheeled Robot

A real application of robot driving system fault diagnosis is selected to illustrate the aforesaid theories and the proposed diagnosis framework. The experimental robot and its fault diagnosis problem are described briefly, followed by the discussions of the three key components in the proposed diagnosis framework, namely, data collection and preprocessing, feature extraction and SVM training, and decision fusion. In addition, several groups of contrast experiments are given in this section.

4.1. Description of Experimental Robot. As shown in Figure 2, we use the wheeled service robot developed by our research group as the experimental platform. This robot is driven by two differential wheels and it is equipped with various kinds of sensors such as one gyroscope (L3GD20), two incremental encoders, two temperature sensors (DS18B20),

Table 1: Fault position and its common fault modes.

Fault categories	Fault position	Fault mode	Tag
Normal condition	None	None	S_0
Mechanical faults	Left wheel	Low pressure	S_1
	Right wheel	Low pressure	S_2
	Left coupling	Loosening	S_3
	Right coupling	Loosening	S_4
Sensor faults	Left encoder	Pulse loss	S_5
	Right encoder	Pulse loss	S_6
	Gyroscope	Constant drift	S_7

current detecting circuits, and voltage detecting circuits. The architecture of the driving system is shown in Figure 3.

In general, faults which occurred in a wheeled robot driving system can be divided into two categories: mechanical faults and sensor faults. In fact, each of the two categories can be subdivided into many small classes. However, only a few typical kinds of high risk faults often occur in the actual course of using the robot [11]. In this paper, we mainly focus on the diagnosis of 7 common kinds of high risk faults and the normal condition S_0 is treated as a special kind of "fault." As shown in Table 1, the fault space can be defined as $S_{err} = \{S_0, S_1, \ldots, S_7\}$.

In order to achieve the effective detection and diagnosis of the faults presented in Table 1, the fault symptom space must be determined, which means that we should select the available and useful sensor signals in the robot driving
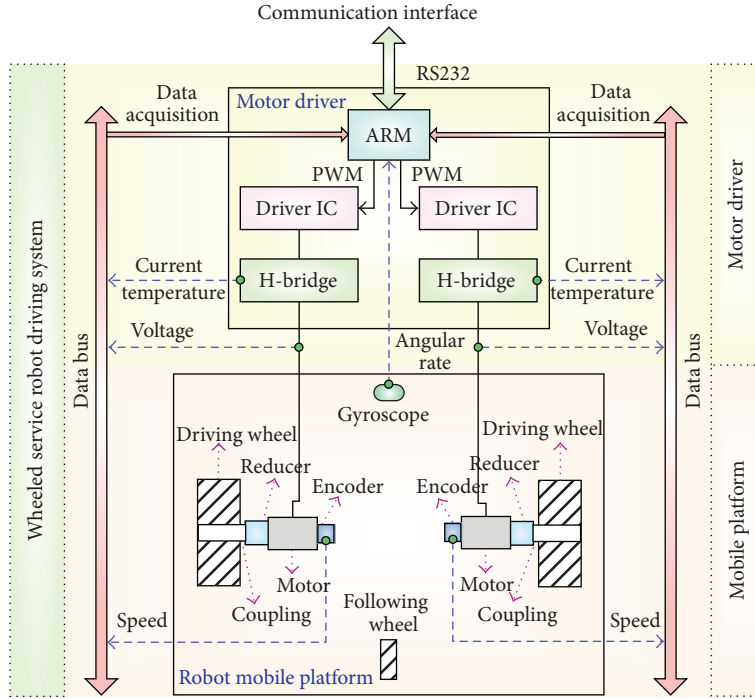
FIGURE 3: Architecture of wheeled robot driving system.

system. It can be indicated from Figure 3 that the working status of the robot driving system is closely related with motor speed, driving voltage, armature current, angular rate, and temperature of H-bridge. Considering that the temperature of H-bridge is easily affected by environment temperature, we use the change rate of the H-bridge temperature as characteristic signal and the final fault symptom space can be defined as {left wheel speed $\mathbf{V}_l$, right wheel speed $\mathbf{V}_r$, left motor driving voltage $\mathbf{U}_l$, right motor driving voltage $\mathbf{U}_r$, left motor armature current $\mathbf{I}_l$, right motor armature current $\mathbf{I}_r$, change rate of the left H-bridge temperature $\mathbf{T}_l$, change rate of the right H-bridge temperature $\mathbf{T}_r$, angular rate $\mathbf{W}$}.

4.2. Data Collection and Preprocessing. The robot motion controller (ARM chip) is responsible for data collection and uploading. In our experiments, we sample 200 sets of data under each of the running conditions (S_0–S_7), respectively. So, the raw data sets can be marked as $D_{\text{all}} = [D_0, \ldots, D_k, \ldots, D_7]^T \in R^{1600 \times 9}$ and $D_k \in R^{200 \times 9}$ denotes data sampled under the kth running condition. For simplification without losing generality, 100 sets of data in each D_k are randomly selected as the original training samples $X_k \in R^{100 \times 9}$ and the remaining 100 sets of data are used as the original testing samples $Y_k \in R^{100 \times 9}$. With the normalized $\overline{X}_k$ ($k = 0, \ldots, 7$), 8 PCA models are established by (17) and (18). The cumulation variance proportion of the PCs for each PCA$_k$ model is shown in Figure 4 and the threshold value μ is set to 0.85. We can get $l_k = [5, 5, 5, 4, 5, 5, 5, 5]$, which represents the optimal number of PCs for PCA$_k$ ($k = 0, \ldots, 7$).

4.3. Feature Extraction and SVM Training. The normalized training data set $\overline{X}_{\text{all}} = [\overline{X}_0, \ldots, \overline{X}_7]^T \in R^{800 \times 9}$ is projected
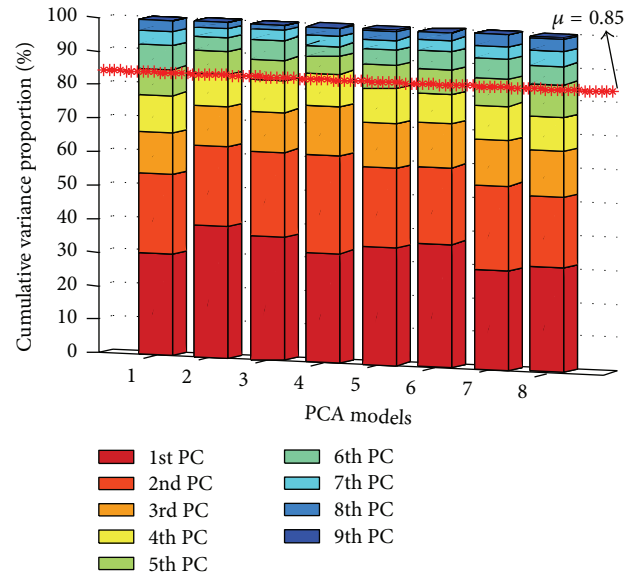

FIGURE 4: Cumulation variance proportion of the PCs.

onto the principal component subspace of each PCA$_k$ model and we can get the feature vectors $F_k \in R^{800 \times l_k}$, ($k = 0, \ldots, 7$) by (19). Then F_k is used to train SVM$_k$ with 5-fold cross validation and PSO algorithm for parameters optimization.

Taking SVM$_8$ as an instance, Figure 5 shows the distribution of the particles during parameters optimization process using PSO algorithm and the optimal parameters of SVM$_8$ are {$c = 11.086, g = 2.997, \alpha = 0.950$}. The optimal parameters of other SVM models are shown in Figure 6. With the trained SVM$_k$ models, the global and local confidence
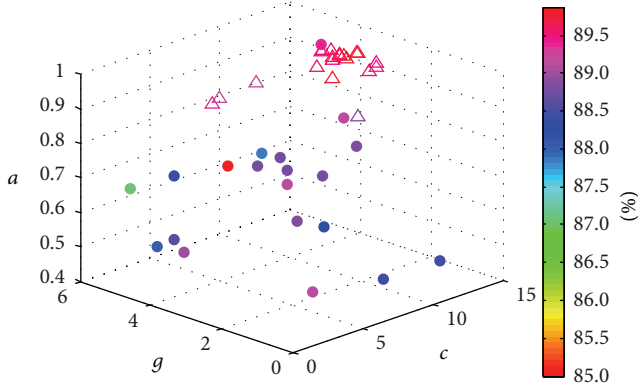
FIGURE 5: Parameters tuning based on PSO algorithm (• denotes the original distribution of the particles and △ denotes the final distribution of the particles).

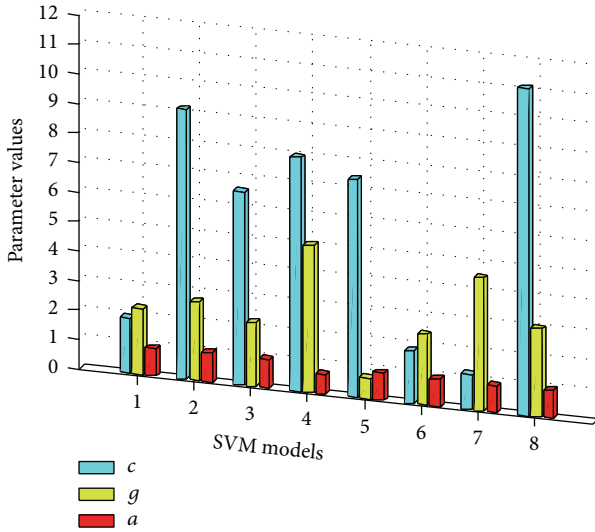FIGURE 7: Confidence values of SVM_k.

FIGURE 6: Optimal parameters of SVM_k.

values $(\bar{a}_k, \omega_{k0}, \ldots, \omega_{k7})$ can be obtained using (20) and (21), respectively, and the confidence values of SVM_k ($k = 0, \ldots, 7$) are presented in Figure 7.

4.4. Decision Fusion. With the global and local confidence values elaborated in Figure 7, we can construct the BPAs for D-S fusion by (23). Taking S_2 and S_4, for example, we get two sets of fusion records randomly and the details are presented in Table 2.

As shown in Table 2, there are three error diagnoses in the 7th, the 14th, and the 16th row, because one single PCA model cannot achieve complete awareness of the robot driving system. While in the proposed approach, multiple PCA models are used to do feature extraction and D-S fusion is applied to fuse the outputs of the ML-kernel PSVM classifiers. Thus, the proposed approach can achieve better awareness of the system and diagnose the faults accurately. Besides, it can be indicated from the 5th and the 7th column that the confidence value (0.959 and 0.995) after D-S fusion
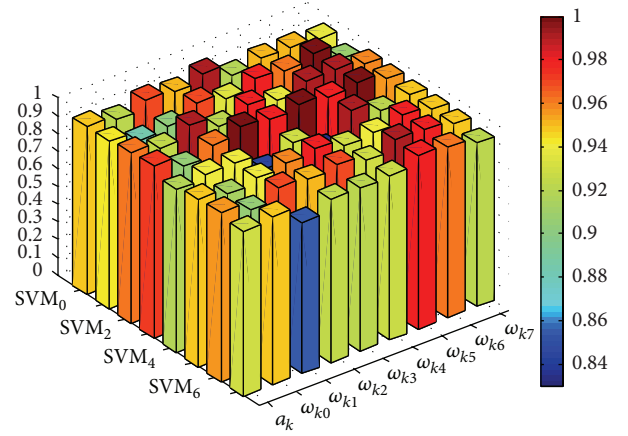
is much bigger than that of any single PCA_k. In other words, the proposed approach can reduce the uncertainties of the diagnosis result efficiently.

The classification accuracy indicates the ability to diagnose the entire categories which are defined as 7 kinds of faulty states $\{S_1, \ldots, S_7\}$ and 1 normal state $\{S_0\}$. In this study, we use the true positive rate as the diagnosis accuracy, which is defined as

$$\text{Acc} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \times 100\%, \tag{24}$$

where N_{FN} is the number of false negatives defined as the number of faults in category k that are not classified as category k and N_{TP} is the number of true positives.

According to Figure 8, the final diagnosis accuracy is 96.75% for the testing samples $Y_{\text{all}} = [Y_0, \ldots, Y_k, \ldots, Y_7] \in R^{800 \times 9}$ (see Section 4.2) in our experiments.

4.5. Contrast Experiments. To further verify the effectiveness of the proposed framework and the ML-kernel, several groups of contrast experiments are conducted, respectively, for comparison.

4.5.1. Evaluation of the ML-Kernel. In order to evaluate the performance of the proposed ML-kernel, we compare the diagnosis accuracy of the ML-kernel with the existing three typical kernel functions based on the proposed hybrid diagnosis framework. For fair comparison, we use the same sets of training samples $X_{\text{all}} = [X_0, \ldots, X_7]$ and testing samples $Y_{\text{all}} = [Y_0, \ldots, Y_7]$ (mentioned in Section 4.2). Besides, PSO and 5-fold cross validation are applied to find the optimal parameters for each kernel function. The experimental results are presented in Table 3.

From the 5th row and the 6th row of Table 3, we can see that the proposed ML-kernel has an identical classification performance compared with the classical Gaussian RBF kernel when $\alpha = 1$. When $0 < \alpha \leq 1$, we can see that the diagnosis ability of the proposed ML-kernel is better than that of the classical Gaussian RBF kernel. From the discussion in Section 2.3, we know that the proposed ML-kernel can be

TABLE 2: BPAs assignments and fusion experiment records.

Fault mode	PCA model	BPAs									Outputs
		$m(S_0)$	$m(S_1)$	$m(S_2)$	$m(S_3)$	$m(S_4)$	$m(S_5)$	$m(S_6)$	$m(S_7)$	$m(\theta)$	
S_2	PCA_0	0.004	0.003	**0.463**	0.003	0.446	0.002	0.024	0.004	0.051	S_2
	PCA_1	0.004	0.004	**0.492**	0.006	0.333	0.003	0.095	0.005	0.058	S_2
	PCA_2	0.004	0.003	**0.620**	0.004	0.279	0.002	0.046	0.003	0.039	S_2
	PCA_3	0.001	0.001	**0.708**	0.001	0.151	0.001	0.107	0.002	0.028	S_2
	PCA_4	0.004	0.002	**0.350**	0.002	*0.546*	0.002	0.010	0.003	0.081	S_4
	PCA_5	0.004	0.002	**0.498**	0.004	0.410	0.002	0.025	0.004	0.051	S_2
	PCA_6	0.002	0.002	**0.507**	0.002	0.390	0.002	0.047	0.003	0.045	S_2
	PCA_7	0.003	0.002	**0.609**	0.003	0.301	0.002	0.006	0.003	0.071	S_2
	DS	0.000	0.000	**0.959**	0.000	0.041	0.000	0.000	0.000	0.000	S_2
S_4	PCA_0	0.007	0.004	0.419	0.006	**0.501**	0.002	0.005	0.005	0.051	S_4
	PCA_1	0.005	0.003	0.021	0.004	**0.902**	0.002	0.001	0.004	0.058	S_4
	PCA_2	0.006	0.003	*0.515*	0.005	**0.423**	0.002	0.003	0.004	0.039	S_2
	PCA_3	0.001	0.001	0.022	0.001	**0.942**	0.000	0.003	0.002	0.028	S_4
	PCA_4	0.007	0.003	*0.742*	0.004	**0.153**	0.002	0.002	0.006	0.081	S_2
	PCA_5	0.007	0.003	0.320	0.005	**0.605**	0.002	0.002	0.005	0.051	S_4
	PCA_6	0.004	0.003	0.400	0.003	**0.535**	0.002	0.004	0.004	0.045	S_4
	PCA_7	0.005	0.003	0.363	0.005	**0.544**	0.002	0.002	0.005	0.071	S_4
	DS	0.000	0.000	0.005	0.000	**0.995**	0.000	0.000	0.000	0.000	S_4

TABLE 3: Diagnosis accuracy of the proposed hybrid framework with different kernel functions.

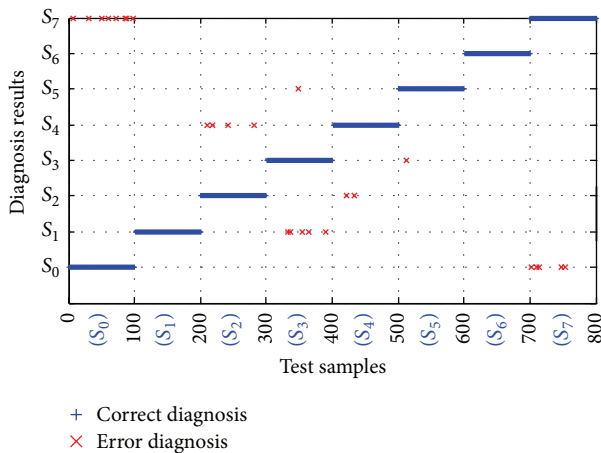| Kernel function	Diagnosis accuracy								
	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	Total
Polynomial kernel	79%	97%	96%	85%	94%	98%	99%	81%	91.51%
Sigmoid kernel	92%	94%	94%	69%	71%	96%	99%	91%	88.25%
Gaussian RBF kernel	81%	97%	95%	86%	92%	99%	100%	84%	91.76%
ML-kernel ($\alpha = 1$)	81%	97%	95%	86%	92%	99%	100%	84%	91.76%
ML-kernel ($0 < \alpha \leq 1$)	92%	100%	96%	94%	98%	99%	100%	95%	96.75%

FIGURE 8: Fault diagnosis result.

regarded as a generalized form of the Gaussian RBF kernel and the experimental results here verify this conclusion again. Table 3 demonstrates that the proposed ML-kernel has the best performance for fault diagnosis of the wheeled robot

driving system, followed by the Gaussian RBF kernel and the polynomial kernel, while the sigmoid kernel has the worst performance in our experiments.

4.5.2. Evaluation of the Proposed Hybrid Diagnosis Framework. The performance of the proposed framework can be evaluated by comparison with traditional nonfusion diagnosis framework. 10 groups of new test data are sampled and each group contains 800 samples which are sampled under each of the running conditions ($S_0, \ldots, S_7$), respectively (100 samples for each running condition). The ML-kernel is adopted as the kernel function of the SVMs both in the proposed framework and in the traditional framework. The experimental result is shown in **Figure 9**, from which we can see that the proposed framework achieves the average accuracy of 94.46% (where the highest diagnosis accuracy and the standard deviation are 97.5% and 1.95, resp.). While, for the traditional framework, the average accuracy is 88.15% (where the highest diagnosis accuracy and the standard deviation are 95.5% and 5.76, resp.), it is clear that the proposed framework achieves better performance both in diagnosis accuracy and in stability, which can be owing to the multiple PCA models and the fusion in decision level.
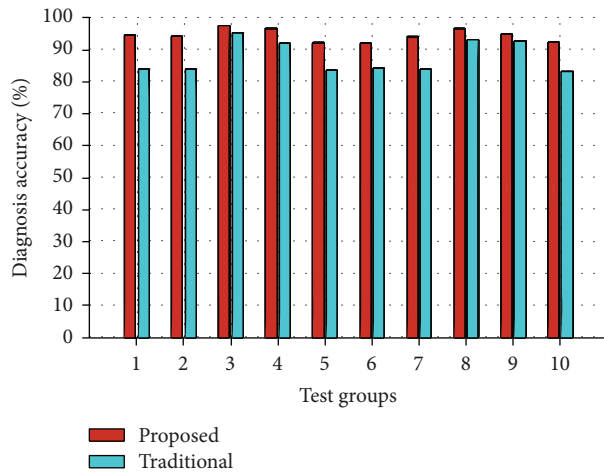
FIGURE 9: Experimental results based on different frameworks.

5. Conclusion

A novel hybrid fault diagnosis framework for wheeled robot driving system is proposed in this paper. The proposed framework is composed of three key components, namely, data collection and preprocessing, feature extraction and SVM training, and decision fusion. Besides, a novel fractional ML-kernel is presented and its positive definiteness and diagnosis ability are discussed in this study. In the proposed framework, multiple PCA models are established to do fault feature extraction firstly. Secondly, the extracted fault feature vectors are used to train the ML-kernel PSVM classifiers with PSO algorithm and cross validation for parameters tuning. Based on the probabilistic outputs and confidence values of those classifiers, the BPAs are constructed. Finally, the BPAs are fused by D-S fusion algorithm that follows the final diagnosis result. In contrast with the earlier studies, the proposed approach can achieve better awareness of the diagnosed system and reduce the uncertainties of the diagnosis result significantly. Through an illustrative application of wheeled robot driving system fault diagnosis, the proposed method is verified as an efficient way of diagnosing the faults in robot driving system and has better performance in stability (standard deviation 1.95) and diagnosis accuracy (highest diagnosis accuracy 97.5%) compared with the traditional methods. In the future, the combination with parallel computing and the cost-sensitive fault diagnosis framework will be studied.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
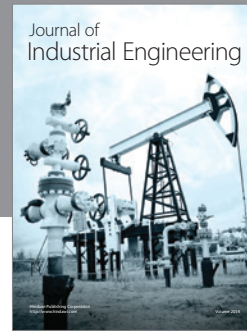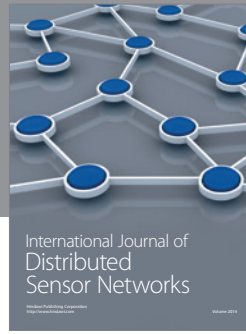
Acknowledgments

References

[1] R. C. Luo and C. C. Lai, "Enriched indoor map construction based on multisensor fusion approach for intelligent service robot," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 8, pp. 3135–3145, 2012.

[2] J.-S. Hu, J.-J. Wang, and D. M. Ho, "Design of sensing system and anticipative behavior for human following of mobile robots," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 4, pp. 1916–1927, 2014.

[3] C. Tang, C. Zhou, H. Hu, W. Pan, and L. Xie, "A semi-supervised learning system for service robots to recognise human actions," *Advanced Robotics*, vol. 28, no. 13, pp. 907–918, 2014.

[4] S. Xu, D. Tu, Y. He, S. Tan, and M. Fang, "ACT-R-typed human–robot collaboration mechanism for elderly and disabled assistance," *Robotica*, vol. 32, no. 5, pp. 711–721, 2014.

[5] S. L. Smith, M. Schwager, and D. Rus, "Persistent robotic tasks: monitoring and sweeping in changing environments," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 410–426, 2012.

[6] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 11, pp. 6414–6428, 2014.

[7] S. X. Ding, P. Zhang, T. Jeinsch, E. L. Ding, P. Engel, and W. Gui, "A survey of the application of basic data-driven and model-based methods in process monitoring and fault diagnosis," in *Proceedings of the 18th IFAC World Congress*, pp. 12380–12388, Milano, Italy, September 2011.

[8] M. Luo, D. Wang, M. Pham et al., "Model-based fault diagnosis/prognosis for wheeled mobile robots: a review," in *Proceedings of the 31st Annual Conference of IEEE Industrial Electronics Society (IECON '05)*, pp. 2267–2272, IEEE, November 2005.

[9] N.-B. Hoang and H.-J. Kang, "A model-based fault diagnosis scheme for wheeled mobile robots," *International Journal of Control, Automation and Systems*, vol. 12, no. 3, pp. 637–651, 2014.

[10] F. Caccavale, A. Marino, G. Muscio, and F. Pierri, "Discrete-time framework for fault diagnosis in robotic manipulators," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1858–1873, 2013.

[11] L. Yu, M. Wu, Z. Cai, and Y. Cao, "A particle filter and SVM integration framework for fault-proneness prediction in robot dead reckoning system," *WSEAS Transactions on Systems*, vol. 10, no. 11, pp. 363–375, 2011.

[12] M. Van, H.-J. Kang, Y.-S. Suh, and K.-S. Shin, "A robust fault diagnosis and accommodation scheme for robot manipulators," *International Journal of Control, Automation and Systems*, vol. 11, no. 2, pp. 377–388, 2013.

[13] L. M. Capisani, A. Ferrara, A. Ferreira De Loza, and L. M. Fridman, "Manipulator fault diagnosis via higher order sliding-mode observers," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 10, pp. 3979–3986, 2012.

[14] H. Wang, T.-Y. Chai, J.-L. Ding, and M. Brown, "Data driven fault diagnosis and fault tolerant control: some advances and possible new directions," *Acta Automatica Sinica*, vol. 35, no. 6, pp. 739–747, 2009.

[15] S. Dong, D. Sun, B. Tang, Z. Gao, W. Yu, and M. Xia, "A fault diagnosis method for rotating machinery based on PCA and Morlet kernel SVM," *Mathematical Problems in Engineering*, vol. 2014, Article ID 293878, 8 pages, 2014.

[16] J. Ni, C. Zhang, and S. X. Yang, "An adaptive approach based on KPCA and SVM for real-time fault diagnosis of HVCBs,"

IEEE Transactions on Power Delivery, vol. 26, no. 3, pp. 1960–1971, 2011.

[17] F. Wu, S. Yin, and H. R. Karimi, "Fault detection and diagnosis in process data using support vector machines," *Journal of Applied Mathematics*, vol. 2014, Article ID 732104, 9 pages, 2014.

[18] D. You, X. Gao, and S. Katayama, "WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 628–636, 2015.

[19] I. Podlubny, *Fractional Differential Equations*, Academic Press, New York, NY, USA, 1999.

[20] Y. Li, Y. Chen, and I. Podlubny, "Stability of fractional-order nonlinear dynamic systems: lyapunov direct method and generalized Mittag-LEFfler stability," *Computers & Mathematics with Applications*, vol. 59, no. 5, pp. 1810–1821, 2010.

[21] H. Sheng, Y. Chen, and T. Qiu, *Fractional Processes and Fractional-Order Signal Processing*, Springer, Berlin, Germany, 2012.

[22] F. Zhou, Y. Zhao, Y. Li, and Y. Chen, "Design, implementation and application of distributed order PI control," *ISA Transactions*, vol. 52, no. 3, pp. 429–437, 2013.

[23] J.-S. Lee and K.-B. Lee, "An open-switch fault detection method and tolerance controls based on SVM in a grid-connected t-type rectifier with unity power factor," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 12, pp. 7092–7104, 2014.

[24] R. Langone, C. Alzate, B. De Ketelaere, J. Vlasselaer, W. Meert, and J. A. K. Suykens, "LS-SVM based spectral clustering and regression for predicting maintenance of industrial machines," *Engineering Applications of Artificial Intelligence*, vol. 37, pp. 268–278, 2014.

[25] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 2000.

[26] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

[27] T. Hastie and R. Tibshirani, "Classification by pairwise coupling," *The Annals of Statistics*, vol. 26, no. 2, pp. 451–471, 1998.

[28] X. Zhang, X. Liu, and Z. J. Wang, "Evaluation of a set of new ORF kernel functions of SVM for speech recognition," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 10, pp. 2574–2580, 2013.

[29] H. J. Haubold, A. M. Mathai, and R. K. Saxena, "Mittag-Leffler functions and their applications," *Journal of Applied Mathematics*, vol. 2011, Article ID 298628, 51 pages, 2011.

[30] Y. Li, Y. Chen, and I. Podlubny, "Mittag-leffler stability of fractional order nonlinear dynamic systems," *Automatica*, vol. 45, no. 8, pp. 1965–1969, 2009.

[31] A. Alkaya and I. Eker, "Variance sensitive adaptive threshold-based PCA method for fault detection with experimental application," *ISA Transactions*, vol. 50, no. 2, pp. 287–302, 2011.

[32] Y. Bao, Z. Hu, and T. Xiong, "A PSO and pattern search based memetic algorithm for SVMs parameters optimization," *Neurocomputing*, vol. 117, pp. 98–106, 2013.

[33] Q. Li, X. Xu, Q. Yan, and Z. Wang, "A fast data fusion algorithm based on matrix analysis for target recognition in sensor networks," in *Proceedings of the 1st International Conference on Pervasive Computing Signal Processing and Applications (PCSPA '10)*, pp. 468–474, IEEE, Harbin, China, September 2010.