

Self-validating Pneumatic Actuator Fault Diagnosis Based on Relevance Vector Machine

Zhigang Feng and Ru Wang

School of Automation, Shenyang Aerospace University, Shenyang, Liaoning, China
fzg1023@yeah.net

Abstract

In order to solve the fault diagnosis problem of self-validating (SEVA) pneumatic actuator, an actuator fault diagnosis approach based on relevance vector machine (RVM) regression modeling and relevance vector machine (RVM) multi-classifier is proposed. The RVM regression is used to establish the normal models of the SEVA pneumatic actuator. The residuals generated by comparing the output of the models and the actual SEVA actuator are used as the nonlinear features. Then, the structure of the RVM for multi-classification is designed using k-meaning clustering methods, which is used as fault classifier to identify the condition and fault pattern of the SEVA actuator. The proposed approach is verified using fault data generated by DABLib model and actuator data from Lublin Sugar Factory and compared with support vector machine (SVM) fault diagnosis approach. The results indicate that the proposed approach overcomes the drawbacks of SVM and resolves the small sample and nonlinear problem in SEVA pneumatic actuator fault diagnosis.

Keywords: *Self-validating (SEVA) pneumatic actuator; fault diagnosis; Relevance vector machine (RVM) regression; Relevance vector machine (RVM) multi-classifier; Support vector machine (SVM); Residuals; Feature extraction*

1. Introduction

Pneumatic actuator relies on compressed air as its power, which can be simple to realize rapid linear circular movement, with its simple structure, convenient maintenance, no pollution, and can be used in all kinds of bad working environment such as an explosion protection requirements, dust or wet condition, is widely used in the field of industrial automation. It plays an important role in the process industry [1]. The reliability of the actuator directly affects the reliability of the whole control system. In 1999, researchers from Department of Engineering Science, Oxford University, proposed the concept of Self-validating (SEVA) actuator [2], which was based on the self-validating sensor technology [3]. Compared with conventional actuator, Self-validating actuator is a new generation of actuator, which can not only give the corresponding output according to the input control variable but also can implement self fault detection and diagnosis, and give the accuracy of the control output and status of the actuator itself and other information. It will become an important development direction of actuator.

In this paper, the fault diagnosis approaches are researched for a self-validating pneumatic actuator based on three component pneumatic valve which is developed by our research group. Figure 1 shows the traditional and self-validating pneumatic actuator structure diagram. The traditional pneumatic actuator [4] consists of three parts: pneumatic servo motor, positioner and control valve, as shown in Figure 1 (a). The self-validating pneumatic actuator, as shown in Figure 1 (b), also consists of three parts: pneumatic servo motor, self-validating unit and control valve. There are one temperature sensor, two pressure sensor and one flow rate sensor are integrated in the valve. The self-

validating unit acquires the stem position, flow rate, upstream pressure, down stream pressure and temperature and control value to implement all self-validating functions including fault diagnosis approaches and give self-validating outputs.

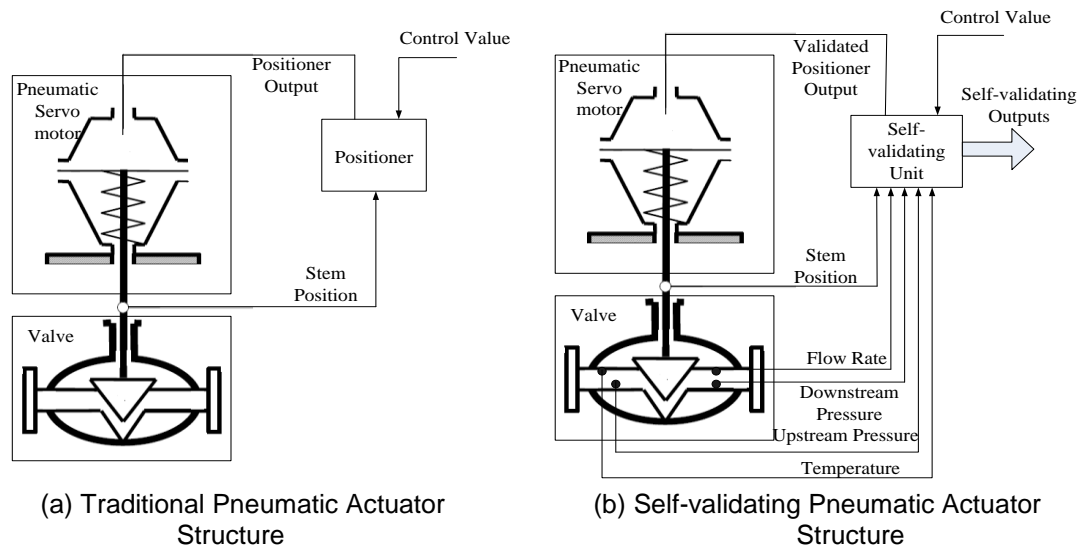


Figure 1. Traditional and Self-validating Pneumatic Actuator Structure Diagram

All kinds of pneumatic actuator fault diagnosis methods have been proposed, mainly divided into three categories, namely, the methods based on the analytical model, neural network and signal processing [5]. These methods have themselves advantages and the combination of many different kinds of intelligence technologies will be a trend of actuator fault diagnosis. In reference [1] and [6], based on the highly nonlinear mapping capabilities, neural network is used to fit the normal model of the actuator, then the residuals are generated by comparing the model output and the actual output of the actuator, finally, the fuzzy inference is used to determine the fault types. Although these methods achieved good results, but there are many disadvantages of neural network, such as easy to fall into local minima, over learning problems, need a lot of training samples and the network structure is difficult to determine. In order to overcome these disadvantages, reference [4] proposed a fault diagnosis method for pneumatic actuator based on least square support vector machine (LS-SVM) regression and support vector machine (SVM) multi-classifier. The SVM employs a linear function in high dimensional feature space as hypothesis space and makes good predictions using small training samples. This is a highly effective mechanism for avoiding over-fitting. However, despite its success, we can identify some significant and practical disadvantages of SVM [7]:

- (1) The kernel function must satisfy Mercer's condition. That is, it must be the continuous symmetric kernel of a positive integral operator;
- (2) Only a single point estimate can be achieved without any uncertainty information;
- (3) Although relatively sparse, the number of support vectors grows linearly along with the increase of training samples size, which increases the computational complexity;
- (4) It is necessary to estimate some insensitive parameters which generally entail extra calculation and setting of parameters.

To overcome above drawbacks, a probabilistic learning framework termed relevance vector machine (RVM) has been originally introduced by Tipping [8]. RVM is a nonlinear pattern recognition model with simple structure based on Bayesian Theory and Marginal Likelihood. The key feature of RVM is that as well as offering excellent performance of prediction and generalization. It improves the inadequacy of SVM [9, 10]. Therefore, this approach has been successfully applied in many fields, such as: speech

classification [10], fatigue crack growth estimation [11], and so on, but has not yet been applied to actuator fault diagnosis. In this paper, the RVM regression is used to fit the normal model of the SEVA actuator, the RVM multi-classifier is used to identify the fault status of the SEVA actuator. Figure 2 illustrates the architecture of SEVA actuator fault diagnosis based on RVM regression and RVM multi-classifier.

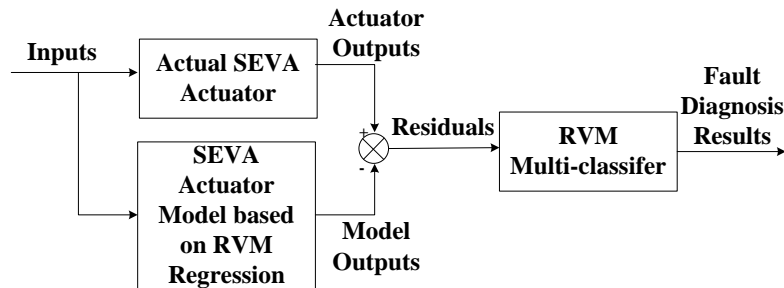


Figure 2. Architecture of Pneumatic Actuator Fault Diagnosis based on RVM Regression and RVM Multi-classifier

In this paper, Section 2 describes the structure and common faults of SEVA pneumatic actuator. Section 3 introduces the theory of RVM regression and SEVA pneumatic actuator modeling method based on RVM regression. Section 4 introduces the theory of RVM classification and RVM multi-classifier establishing method for SEVA pneumatic actuator diagnosis. To verify the effectiveness and superiority of the RVM approach, Section 5 presents a case study with DABLib data and Lublin Sugar Factory data comparing the performance of RVM with SVM. Section 6 includes the final conclusions.

2. Structure and Common Faults of SEVA Pneumatic Actuator

As shown in Figure 1(b), the SEVA pneumatic actuator model can be represented by the following Figure 3. In Figure 3, CV represents for control value in Figure 1 (b), P1 represents for upstream pressure in Figure 1 (b), P2 represents for downstream pressure in Figure 1 (b), T1 represents for temperature in Figure 1 (b), X represents for stem position in Figure 1 (b), F represents for flow rate in Figure 1 (b). The inputs of the SEVA pneumatic actuator are CV, P1, P2 and T1, the outputs of the SEVA pneumatic actuator are X and F.

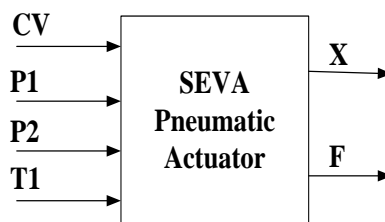


Figure 3. General Scheme of SEVA Pneumatic Actuator Model

The SEVA pneumatic actuator keeping continuous and completing the given functional requirements in a certain period of working time, so that making the given parameter value to maintain within the limits prescribed and this kind of performance is called the trouble-free performance, when the performance is damaged it will generate a fault. Literature [12] summarized the 19 kinds of typical faults, Table 1 shows the details. Fault f1 to f7 belong to the control valve faults, fault f8 to f11 belong to the servomotor faults, fault f12 to f15 belong to the positioner faults, and fault f16 to f19 belong to the general faults or external faults.

Figure 4 to Figure 7 show the actual actuator data from Lublin Sugar Factory without fault and with fault f1, f2 and f15. Figure 8 to Figure 11 show the simulated actuator data using DABLib model without fault and with fault f1, f2 and f15. All the signals are normalize to 0-1, the physical range of P1 and P2 is from 0 to 4000kPa. The physical range of CV and X is 0 to 100%. The physical range of T1 is 0 to 150°C. The physical range of F is 0 to 40 t/h.

Table 1. The Detail Information of the 19 Kinds Faults for SEVA Pneumatic Actuator

Fault No	Description	Fault No	Description
f1	valve clogging	f11	servomotor spring fault
f2	valve or valve seat sedimentation	f12	electro-pneumatic transducer fault
f3	valve or valve seat erosion	f13	stem displacement sensor fault
f4	increase of valve friction,	f14	pressure sensor fault
f5	external leakage (leaky bushing, covers, terminals)	f15	positioner spring fault
f6	internal leakage (valve tightness)	f16	positioner supply pressure drop
f7	medium evaporation or critical flow	f17	unexpected pressure change across valve
f8	twisted servo-motor stem	f18	fully or partly opened bypass valves
f9	servomotor housing or terminal tightness	f19	flow rate sensor fault
f10	servomotor diaphragm perforation	fn	fault free state

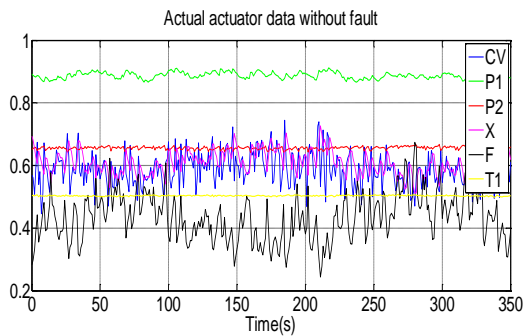


Figure 4. Actual Actuator Data without Fault

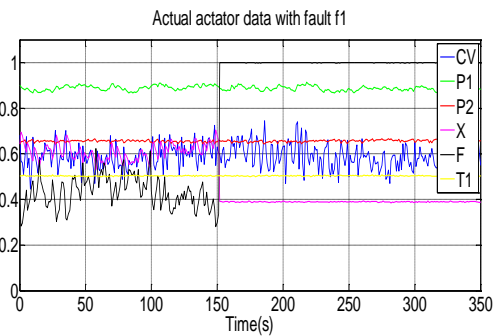


Figure 5. Actual Actuator Data with Fault f1

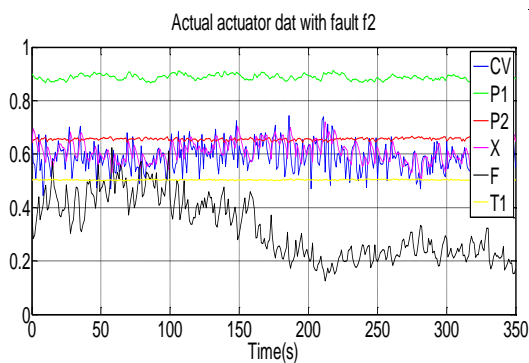


Figure 6. Actual Actuator Data with Fault f2

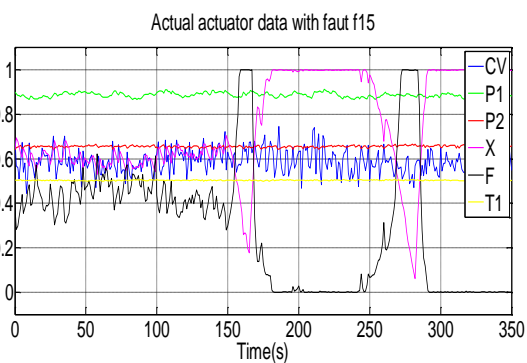


Figure 7. Actual Actuator Data with Fault f15

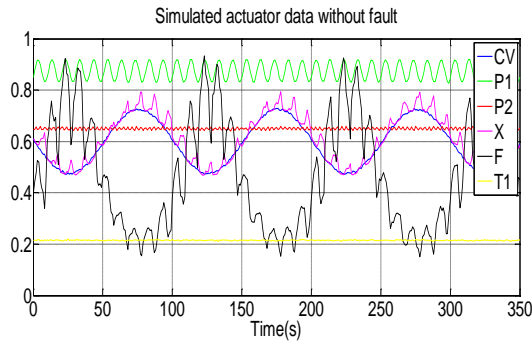


Figure 8. Simulated Actuator Data without Fault

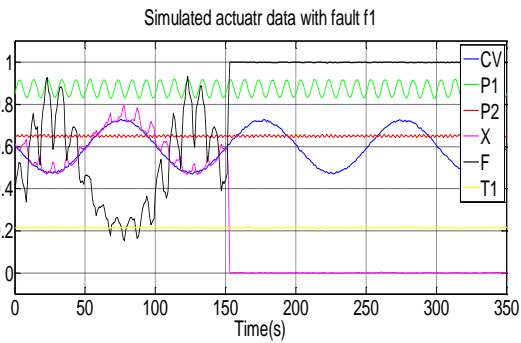


Figure 9. Simulated Actuator Data with Fault f1

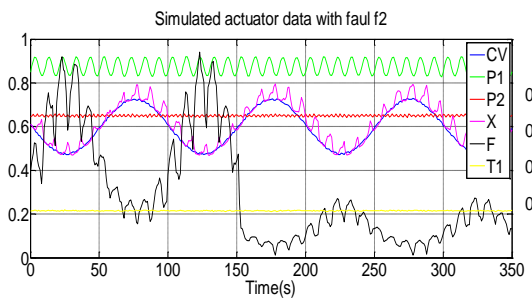


Figure 10. Simulated Actuator Data with Fault f2

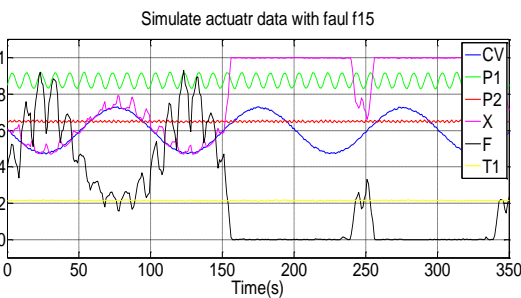


Figure 11. Simulated Actuator Data with Fault f15

From Figure 4 to Figure 11, we can see that when fault occurs, the inputs or the outputs of the actuator will deviate from the normal value, so by modeling the actuator using the inputs and outputs, and comparing the model output with the actual output, we can get the residuals, which can be used as the feature to identify the fault type of the actuator.

3. SEVA Pneumatic Actuator Modeling based on RVM Regression

In order to get the residuals, first of all, the actuator model should be established. In this paper, the relevance vector machine regression is firstly used to establish the actuator model.

3.1. Theory of RVM Regression

Relevance vector machine (RVM) is simply a specialization of a sparse bayesian model which utilizes the same data-dependent kernel basis [8]. The key feature of RVM is that the inferred predictors are exceedingly sparse in that they contain relatively few “relevance vectors”, as well as offering good generalization performance.

Given a data set of input-target pairs $\{\mathbf{x}_i, t_i\}_{i=1}^n$, considering scalar valued target functions only. When its input is a d-dimensional vector \mathbf{x}_i , we want to make predictions for the output $y(\mathbf{x})$ at any arbitrary test point \mathbf{x} . We follow the standard probabilistic formulation and assume that the targets are samples from the model with additive noise [13]:

$$t_i = y(\mathbf{x}_i; \mathbf{w}) + \varepsilon_i \quad (1)$$

Where ε_i are independent samples from some noise process. The relevance vector machine assumes that the basis function set consists of kernel function at each training points:

$$y(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^n \omega_i K(\mathbf{x}, \mathbf{x}_i) + \omega_0 \quad (2)$$

Where $K(\mathbf{x}, \mathbf{x}_i)$ is some kernel function and $\mathbf{w} = (\omega_0, \omega_1, \dots, \omega_n)^T$ are the weights of the linear model.

From the engineering viewpoint, a common assumption for the noise ε_i are to be independent, identically distributed, following a Gaussian distribution with zero mean and variance σ^2 [14]:

$$p(\varepsilon_i | \sigma^2) = N(\varepsilon_i | 0, \sigma^2) \quad (3)$$

Thus,

$$p(t_i) = N(t_i | y(x_i; \mathbf{w}), \sigma^2) \quad (4)$$

Where the notation $N(t_i | y(x_i; \mathbf{w}), \sigma^2)$ specifies a Gaussian distribution over t_i with mean $y(x_i; \mathbf{w})$ and variance σ^2 . Due to assumption of independence of the t_i , the likelihood of the complete data set can be written as [8, 14]:

$$p(\mathbf{t} | \mathbf{w}, \sigma^2) = \prod_{i=1}^n N(t_i | y(\mathbf{x}_i; \mathbf{w}), \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\|\mathbf{t} - \Phi \mathbf{w}\|^2}{2\sigma^2}\right) \quad (5)$$

Where $\mathbf{t} = [t_1, t_2, \dots, t_n]^T$, Φ is a $n \times (n+1)$ dimension matrix.

$$\Phi_{N \times (N+1)} = \begin{bmatrix} 1 & K(\mathbf{x}_1, \mathbf{x}_1) & K(\mathbf{x}_1, \mathbf{x}_2) & \dots & K(\mathbf{x}_1, \mathbf{x}_N) \\ 1 & K(\mathbf{x}_2, \mathbf{x}_1) & K(\mathbf{x}_2, \mathbf{x}_2) & \dots & K(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & K(\mathbf{x}_N, \mathbf{x}_1) & K(\mathbf{x}_N, \mathbf{x}_2) & \dots & K(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} \quad (6)$$

To control the complexity of the model and to avoid over-fitting, a zero-mean Gaussian prior probability distribution is defined over every ω_i with α_i , and the likelihood of \mathbf{w} is written as [8, 14]:

$$p(\mathbf{w} | \mathbf{a}) = \prod_{i=0}^n N(\omega_i | 0, \alpha_i^{-1}) = \prod_{i=0}^n \frac{\alpha_i}{\sqrt{2\pi}} \exp\left(-\frac{\alpha_i \omega_i^2}{2}\right) \quad (7)$$

Where hyper-parameters vector $\mathbf{a} = [\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n]^T$ controls how far from zero each weight is allowed to deviate.

Consequently, using Bayesian posterior inference, the posterior over \mathbf{w} could be computed as:

$$p(\mathbf{w} | \mathbf{t}, \mathbf{a}, \sigma^2) = \frac{p(\mathbf{t} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \mathbf{a})}{p(\mathbf{t} | \mathbf{a}, \sigma^2)} = N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (8)$$

Where the mean and covariance values are $\boldsymbol{\mu} = \sigma^{-2} \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{t}$, $\boldsymbol{\Sigma} = (\sigma^{-2} \boldsymbol{\Phi}^T \boldsymbol{\Phi} + \mathbf{A})^{-1}$, $\mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_n)$. When $\alpha \rightarrow \infty$, $\mu_i = 0$. Under the Bayesian framework, the parameters α_i can be estimated by maximizing the marginal likelihood $p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2)$, which is obtained by integrating out the weights \mathbf{w} :

$$p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Omega}|^{-\frac{1}{2}} \exp\left(-\frac{\mathbf{t}^T \boldsymbol{\Omega}^{-1} \mathbf{t}}{2}\right) \quad (9)$$

Where $\boldsymbol{\Omega} = \sigma^2 \mathbf{I} + \boldsymbol{\Phi} \mathbf{A}^{-1} \boldsymbol{\Phi}^T$. Setting the derivative of the marginal likelihood with respect to $\log \alpha_i$ to zero, the following update formula for α_i is given [8]:

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \quad (10)$$

Where $\gamma_i = 1 - \alpha_i \boldsymbol{\Sigma}_{i,i}$, $\boldsymbol{\Sigma}_{i,i}$ is the diagonal elements of the matrix $\boldsymbol{\Sigma}$. Sparsity is achieved because most of the parameters α_i are estimated to very large values, thus pruning the corresponding basis functions by forcing their weights to become zero. In the optimization process, the vector from the training set that associates with the remaining nonzero weights is called the relevance vector (RV) [8].

Furthermore, setting the derivative of the marginal likelihood with respect to $\log \sigma^2$ to zero, we obtain the following update formula for σ^2 [13]:

$$(\sigma^2)^{new} = \frac{\|\mathbf{t} - \boldsymbol{\Phi} \boldsymbol{\mu}\|^2}{N - \sum_{i=0}^n r_i} \quad (11)$$

At convergence of the hyper-parameter estimation procedure, the predictions can be made based on the posterior distribution over the weights, conditioned on the maximizing the values $\boldsymbol{\alpha}_{MP}$ and σ_{MP}^2 . We can then complete the predictive distribution for a new value, \mathbf{x}_* :

$$\begin{aligned} p(t_* | \mathbf{t}) &= \int p(t_* | \mathbf{w}, \boldsymbol{\alpha}, \sigma^2) p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) d\mathbf{w} d\boldsymbol{\alpha} d\sigma^2 \\ &\approx \int p(t_* | \mathbf{w}, \boldsymbol{\alpha}, \sigma^2) p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) \delta(\boldsymbol{\alpha} - \boldsymbol{\alpha}_{MP}) \delta(\sigma^2 - \sigma_{MP}^2) d\mathbf{w} d\boldsymbol{\alpha} d\sigma^2 \\ &= \int p(t_* | \mathbf{w}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) d\mathbf{w} = N(t_* | y_*, \sigma_*^2) \end{aligned} \quad (12)$$

Where

$$y_* = \boldsymbol{\mu}^T \boldsymbol{\Phi}(\mathbf{x}_*) \quad (13)$$

$$\sigma_*^2 = \sigma_{MP}^2 + \boldsymbol{\Phi}(\mathbf{x}_*)^T \boldsymbol{\Sigma} \boldsymbol{\Phi}(\mathbf{x}_*) \quad (14)$$

$$\boldsymbol{\Phi}(\mathbf{x}_*) = [1, K(\mathbf{x}_*, \mathbf{x}_1), K(\mathbf{x}_*, \mathbf{x}_2), \dots, K(\mathbf{x}_*, \mathbf{x}_N)]^T \quad (15)$$

y_* is the predicted mean value, σ_*^2 is the predicted variance value, which is the sum being comprised of two variance components: the estimated noise on the data and that due to the uncertainty in the prediction of the weights. In practice, we only need compute the value $\boldsymbol{\mu}$ for the purposes of point prediction, y_* and return $\boldsymbol{\Sigma}$ if required to compute the prediction interval $[y_* - \sigma_*, y_* + \sigma_*]$. Then the upper and low limits on certain confidence level can be achieved.

The training steps for RVM regression show as follows:

- (1) Initializing $\{\alpha_i\}$ and σ^2 .
- (2) Calculating μ and Σ by using $\Sigma = (\sigma^{-2} \Phi^T \Phi + \mathbf{A})^{-1}$ and $\mu = \sigma^{-2} \Sigma \Phi^T \mathbf{t}$.
- (3) Re-estimating α_i^{new} and $(\sigma^2)^{new}$ by using formula (10) and (11).
- (4) Repeating steps (2) and (3) until satisfying the convergence conditions or the maximum number of iterations has been reached.
- (5) The samples with non-zero weights are the relevance vectors. Meanwhile the model parameters of RVM can be calculated.
- (6) By using formula (13) and (14), the predicted mean value y_* and the predicted variance value σ_*^2 can be calculated.

3.2. SEVA Pneumatic Actuator Modeling based on RVM Regression

According to the principle of pneumatic actuator action, it is known that there are certain function relationship between the inputs and outputs of the actuator. As shown in Figure 3, the function relationship can be represented as equation (16).

$$(\hat{X}, \hat{F}) = f(CV, P1, P2, T1) \quad (16)$$

The function relationship which is also called the model of the actuator can be established through some mathematical algorithms using the inputs and outputs of the actuator. In this paper, the RVM regression is used to establish the model. But, there is only one output of RVM regression, so the two outputs of actuator are separately modeled using equation (17) and (18).

$$\hat{X} = f_1(CV, P1, P2, T1) \quad (17)$$

$$\hat{F} = f_2(CV, P1, P2, T1) \quad (18)$$

When fault occurs, fluid flow F or valve stem displacement X will be abnormal, the function relationships above will change. Actuator faults can be detected by detecting these changes. The intuitive method of fault diagnosis is to analyze the residuals of outputs of fault actuator and normal actuator. From the above two models, there are two residuals can be calculated by equation (19) and (20):

$$r_1 = X - \hat{X} = X - g_1(CV, \sqrt{P1 - P2}, T1) \quad (19)$$

$$r_2 = F - \hat{F} = F - g_2(CV, \sqrt{P1 - P2}, T1) \quad (20)$$

The actuator models of g_1 and g_2 are established using the actuator data without fault. Suppose there are NS groups of normal actuator data, each group has NN points. The actuator modeling procedure using RVM regression show as following:

(1) Generating training samples and testing samples. Suppose NST groups of normal data are selected as training data, NSE groups of normal data are selected as testing data, where $NST+NSE=NS$. The training inputs of the two RVM regression model are both $X T_{NST \cdot NN \times 3}$. The training outputs of RVM regression for g_1 is $T X_{NST \cdot NN \times 1}$, The training outputs of RVM regression for g_2 is $T F_{NST \cdot NN \times 1}$. The testing inputs of the two RVM regression model are both $X T_{NSE \cdot NN \times 3}$. The testing outputs of RVM regression for g_1 is $E X_{NSE \cdot NN \times 1}$, The training outputs of RVM regression for g_2 is $E F_{NSE \cdot NN \times 1}$.

$$\begin{aligned}
 XT_{NST \times 3} &= \begin{bmatrix} CV_{11} & P_{11} & T_{11} \\ \vdots & \vdots & \vdots \\ CV_{1NN} & P_{1NN} & T_{1NN} \\ \vdots & \vdots & \vdots \\ CV_{NST1} & P_{NST1} & T_{NST1} \\ \vdots & \vdots & \vdots \\ CV_{NSTNN} & P_{NSTNN} & T_{NSTNN} \end{bmatrix} \quad TX_{NST \times 1} = \begin{bmatrix} X_{11} \\ \vdots \\ X_{1NN} \\ \vdots \\ X_{NST1} \\ \vdots \\ X_{NSTNN} \end{bmatrix} \quad TF_{NST \times 1} = \begin{bmatrix} F_{11} \\ \vdots \\ F_{1NN} \\ \vdots \\ F_{NST1} \\ \vdots \\ F_{NSTNN} \end{bmatrix} \quad (21)
 \end{aligned}$$

$$\begin{aligned}
 XE_{NSE \times 3} &= \begin{bmatrix} CV_{11} & P_{11} & T_{11} \\ \vdots & \vdots & \vdots \\ CV_{1NN} & P_{1NN} & T_{1NN} \\ \vdots & \vdots & \vdots \\ CV_{NSE1} & P_{NSE1} & T_{NSE1} \\ \vdots & \vdots & \vdots \\ CV_{NSENN} & P_{NSENN} & T_{NSENN} \end{bmatrix} \quad EX_{NSE \times 1} = \begin{bmatrix} X_{11} \\ \vdots \\ X_{1NN} \\ \vdots \\ X_{NSE1} \\ \vdots \\ X_{NSENN} \end{bmatrix} \quad EF_{NSE \times 1} = \begin{bmatrix} F_{11} \\ \vdots \\ F_{1NN} \\ \vdots \\ F_{NSE1} \\ \vdots \\ F_{NSENN} \end{bmatrix} \quad (22)
 \end{aligned}$$

(2) Turning the kernel function parameters and training the RVM regression model. In this paper, Gaussian kernel function is select, $K(\mathbf{x}, \mathbf{z}) = \exp(-\|\mathbf{x} - \mathbf{z}\|^2 / 2\sigma^2)$. The parameter σ affects the RVM regression performance. In order to get the optimal performance, genetic algorithm (GA) is used to select the optimal kernel parameter. The means square error (MSE) of the regression model is used as the fitness function. The detail information about GA can be found in reference [15]. After finding the optimal parameter σ , the actuator models of g_1 and g_2 are trained using the optimal parameter σ with the training samples and testing samples established in (1).

(3) Calculating the residuals using the established models. For a new group of actuator data, suppose the data have NT points, the inputs of the actuator models of g_1 and g_2 are both $XI_{NT \times 3}$. Taking the $XI_{NT \times 3}$ into the actuator models of g_1 and g_2 , the prediction of X and F are $\hat{X}_{NT \times 1}$ and $\hat{F}_{NT \times 1}$. The residual $r1_{NT \times 1} = X_{NT \times 1} - \hat{X}_{NT \times 1}$, $r2_{NT \times 1} = F_{NT \times 1} - \hat{F}_{NT \times 1}$.

In this paper, the simulated actuator data is generated using DABLib toolbox, which is a classic pneumatic valve fault simulation model [12]. It can produce various types and intensities fault data, and also can import the actual industrial data into the simulation model. The model can generate the fault as in Table 1. For the simulated actuator data, the actuator is driven by low-frequency sine wave, pressure signals P1 and P2 are driven by the sine functions with appropriate amplitudes and frequencies. Each fault runs 190 times, each time runs 350 seconds, and eventually we got the 190 samples for each actuator status.

The actual actuator data is come from Lublin Sugar Factory which is produced in trouble-free operation. Though importing CV, P1, P2, X, F and T1 into DABLib model, and then adding 19 kinds of fault to simulate the model, the actual actuator data with faults can be generated. For the actual actuator data, each fault runs 90 times, each time runs 350 seconds, and eventually we got the 90 samples for each actuator status.

The actuator model of g_1 and g_2 based on RVM regression are trained using the 190 without fault samples of the simulated actuator data. The optimal parameter σ is 0.0305 using GA. Figure 12 to Figure 15 show the residuals of the simulated actuator data without fault and with fault f1, f2 and f15. Figure 16 to Figure 19 show the residuals of the actual actuator data without fault and with fault f1, f2 and f15. From these figures, we can see that the modeling accuracy fulfills the fault diagnosis requirements. If actuator

fault occurs, the r_1 or r_2 will abnormal, so the residuals are selected as the feature for actuator fault diagnosis.

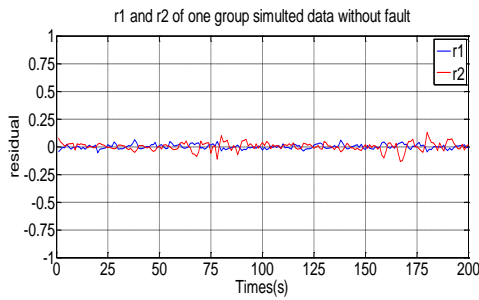


Figure. 12. r_1 and r_2 of Simulated Data without Fault

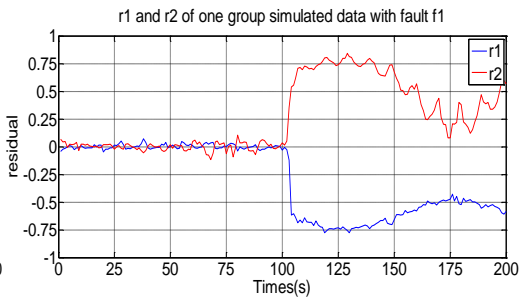


Figure. 13. r_1 and r_2 of Simulated Data with Fault f1

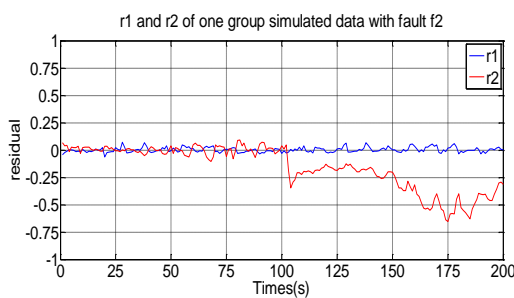


Figure. 14. r_1 and r_2 of Simulated Data with Fault f2

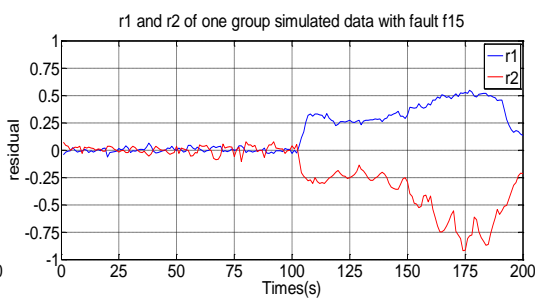


Figure. 15. r_1 and r_2 of Simulated Data with Fault f15

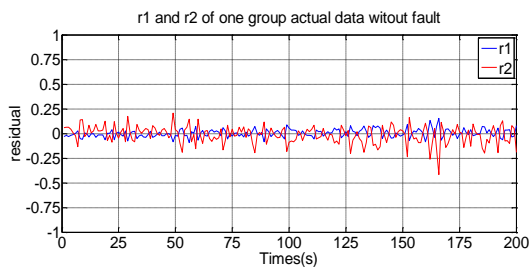


Figure. 16. r_1 and r_2 of Actual Data without Fault

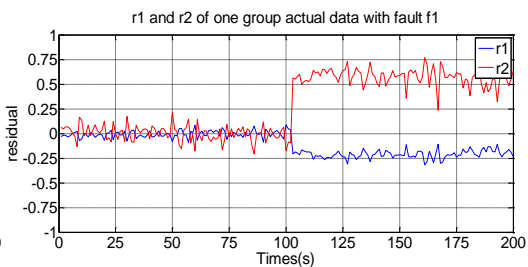


Figure. 17. r_1 and r_2 of Actual Data with Fault f1

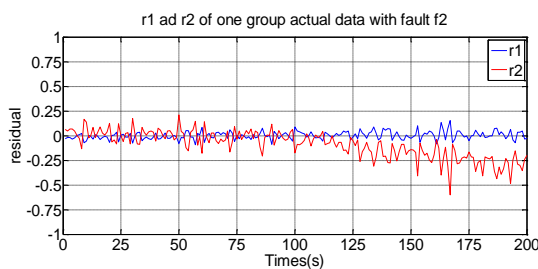


Figure. 18. r_1 and r_2 of Actual Data with Fault f2

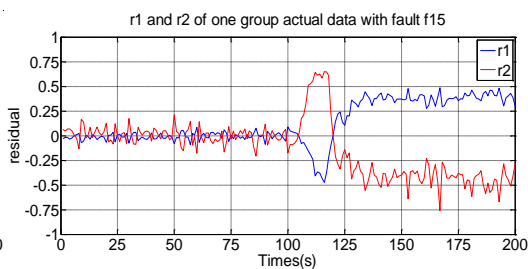


Figure. 19. r_1 and r_2 of Actual Data with Fault f15

After analyzing all the residuals of the 19 kinds of faults and fault-free state, we found that the residuals r_1 and r_2 of f_n , f_4 , f_5 , f_8 , f_9 and f_{14} are similar to each other, the residuals r_1 and r_2 of f_1 , f_7 , f_{10} and f_{16} are similar to each other, the residuals r_1 and r_2 of f_2 and f_{19} are similar to each other, the residuals r_1 and r_2 of f_6 and f_{13} are similar to

each other. It is difficult to distinguish them just using r1 and r2, so in this research, we just consider the status identification of f1, f2, f3, f6, f11, f12, f15, f17, f18 and fn. In the future, we will research on extracting new feature to identify the other fault status.

4. SEVA Pneumatic Actuator Fault Diagnosis based on RVM Multi-classifier

In order to identify the fault type of the SEVA pneumatic actuator, a multi-classifier should be established. In this paper, the relevance vector machine classification is expanded to multi-classification and firstly used to fault diagnosis of SEVA pneumatic actuator.

4.1. Theory of RVM Classification

The relevance vector machine for classification and the relevance vector machine for regression have the similar framework [8]. For two-class classification, it is desired to predict the posterior probability of membership of one of the classes given the input \mathbf{x} . The logistic sigmoid link function $\sigma(y) = 1 / (1 + e^{-y})$ is applied to $y(\mathbf{x})$ to link random and systematic components, and generalize the linear model. Following the definition of the Bernoulli distribution, the likelihood is written as [16, 17]:

$$p(\mathbf{t} | \mathbf{w}) = \prod_{i=1}^n \sigma\{y(x_i; \mathbf{w})\}^{t_i} [1 - \sigma\{y(x_i; \mathbf{w})\}]^{1-t_i} \quad (23)$$

Where the target $t_i \in \{0, 1\}$.

The likelihood is complemented by a prior over the weights in the form of

$$p(\mathbf{w} | \mathbf{a}) = \prod_{i=1}^n N(w_i | 0, \alpha_i^{-1}) \quad (24)$$

Where $\mathbf{a} = [\alpha_0, \alpha_1, \alpha_3, \dots, \alpha_n]^T$ is the hyper-parameters introduced to control the strength of the prior over its associated weight.

Then, given a new test point \mathbf{x}_* , predictions are made for the corresponding target t_* , in terms of the predictive distribution:

$$p(t_* | \mathbf{t}) = \int p(t_* | \mathbf{w}, \mathbf{a}) p(\mathbf{w} | \mathbf{t}, \mathbf{a}) p(\mathbf{a} | \mathbf{t}) d\mathbf{w} d\mathbf{a} \quad (25)$$

Where $p(\mathbf{w} | \mathbf{t}, \mathbf{a})$ is the posterior weight distribution, and $p(\mathbf{a} | \mathbf{t})$ is referred to as evidence.

The weights cannot be analytically obtained, and therefore, a Laplace's approximation procedure is used.

(1) For the current fixed \mathbf{a} , values of, the 'most probable' weights \mathbf{w}_{MP} are found, giving the location of the mode of the posterior distribution. Since $p(\mathbf{w} | \mathbf{t}, \mathbf{a}) \propto p(\mathbf{t} | \mathbf{w}) p(\mathbf{w} | \mathbf{a})$, this is equivalent to finding the maximum of the following equation:

$$\log p(\mathbf{t} | \mathbf{w}) p(\mathbf{w} | \mathbf{a}) = \sum_{i=1}^n [t_i \log y_i + (1 - t_i) \log(1 - y_i)] - \frac{1}{2} \mathbf{w}^T \mathbf{A} \mathbf{w} \quad (26)$$

For the most probable weights \mathbf{w}_{MP} , with $y_i = \sigma\{y(x_i; \mathbf{w})\}$ and $\mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_n)$ being composed of the current values of \mathbf{a} . This is a penalized logistic log-likelihood function and requires iterative maximization. The iteratively reweighted least-squares algorithm can be used to find \mathbf{w}_{MP} .

(2) The logistic log-likelihood function can be differentiated twice to obtain the Hessian in the form of:

$$H = \nabla_{\mathbf{w}} \nabla_{\mathbf{w}} \log [p(\mathbf{t} | \mathbf{w}) p(\mathbf{w} | \boldsymbol{\alpha})] = (-\boldsymbol{\Phi}^T \mathbf{B} \boldsymbol{\Phi} - \mathbf{A}) \quad (27)$$

Where $\mathbf{B} = \text{diag}(\beta_1, \beta_2, \dots, \beta_n)$, is a diagonal matrix with $\beta_i = \sigma[y(x_i; \mathbf{w}_{MP})] \{1 - \sigma[y(x_i; \mathbf{w}_{MP})]\}$, and $\boldsymbol{\Phi}$ is the design matrix with $\Phi_{nm} = K(\mathbf{x}_n, \mathbf{x}_m)$ and $\Phi_{n0} = 1$. This result is then negated and inverted to give the covariance $\boldsymbol{\Sigma}$ as show as follows, for a Gaussian approximation to the posterior over weights centered at \mathbf{w}_{MP} :

$$\boldsymbol{\Sigma} = -\mathbf{H}^{-1} = (\boldsymbol{\Phi}^T \mathbf{B} \boldsymbol{\Phi} + \mathbf{A})^{-1} \quad (28)$$

(3) Using the statistics covariance $\boldsymbol{\Sigma}$ and mean \mathbf{w}_{MP} of the Gaussian approximation, the hyper-parameters are updated using the following equation.

$$\begin{cases} \alpha_i^{new} = \frac{\gamma_i}{\mathbf{w}_{MPi}^2} \\ \gamma_i = 1 - \alpha_i \boldsymbol{\Sigma}_{i,i} \end{cases} \quad (29)$$

Where \mathbf{w}_{MPi} is the i -th element of weights \mathbf{w}_{MP} , $\boldsymbol{\Sigma}_{i,i}$ is the i -th diagonal element of the covariance $\boldsymbol{\Sigma}$. During the optimization process, many α_i will have large values, and thus, the corresponding model weights are pruned out, realizing sparsity. The optimization process typically continues until the maximum change in α_i values is below a certain threshold or the maximum number of iterations is reached. We can then complete the predictive classification result for a new value, \mathbf{x}_* :

$$t_* = \sigma(y_*(x_*; \mathbf{w}_{MP})) = \sigma(\mathbf{w}_{MP}^T \boldsymbol{\Phi}(x_*)) \quad (30)$$

The training steps for RVM classification show as follows:

- (1) Initializing $\{\alpha_i\}$.
- (2) Calculating \mathbf{w}_{MP}^{new} and $\boldsymbol{\Sigma}$ by using $\mathbf{w}_{MP}^{new} = \mathbf{w}_{MP} + \Delta \mathbf{w}$ and $\boldsymbol{\Sigma} = -\mathbf{H}^{-1} = (\boldsymbol{\Phi}^T \mathbf{B} \boldsymbol{\Phi} + \mathbf{A})^{-1}$.
- (3) Re-estimating α_i^{new} by using formula (29).
- (4) Repeating steps (2) and (3) until satisfying the convergence conditions or the maximum number of iterations has been reached.
- (5) The samples with non-zero weights are the relevance vectors. Meanwhile the model parameters of RVM can be calculated.
- (6) By using formula (30), the predicted classification result can be calculated.

4.2. RVM Multi-classifier Design for SEVA Pneumatic Actuator Fault Diagnosis

Similar to SVM, RVM is a binary classifier and can process the multi-class problem by the one-against-one (OAO), one-against-all (OAA) or binary hierarchical classification structure methods [17]. For the C -class problem, in one-against-all method, each classifier distinguishes one class from the other $C-1$ classes, and the class labels of the inputs are decided by winner-take-all method. Each classifier needs to be trained on the whole training set, and there is no guarantee that good discrimination exists between one class and the remaining classes. In one-against-one method, the decision is made by majority

voting strategies. This requires training and testing of $C(C-1)/2$ binary classifiers. This approach is prohibitive when C is large.

Thus, we chose a binary hierarchical classification structure in Figure 20. Each node is a RVM binary classifier, coarse separation between classes occurs in the beginning (at upper levels) in the hierarchy and a finer classification decision later (at lower levels). At the top node, we divide the original 10 classes into two smaller groups of classes (macro-classes); this clustering procedure is repeated in subsequent levels, until there is only one class in the final sub-group. This hierarchical structure decomposes the problem into 9 binary sub-problems. For testing, only about $\log_2 10$ classifiers are required to traverse a path from top to bottom.

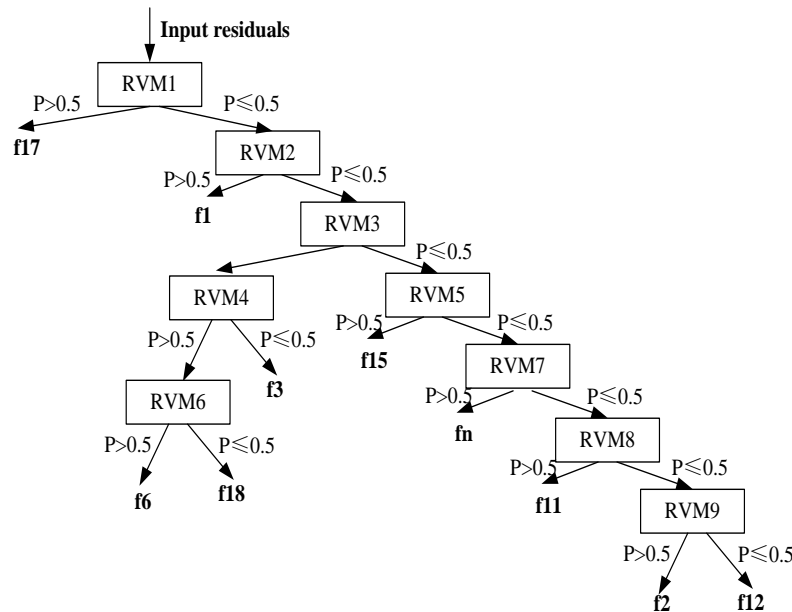


Figure 20. Hierarchical Multi-classifier Structure based on RVM Classification

In our research, standard k -means clustering is used to design the binary hierarchical structure, and Figure 20 shows the result. RVM1 is used to classify f17 vs. f1, f2, f3, f6, f11, f12, f15, f18 and fn. RVM2 is used to classify f1 vs. f2, f3, f6, f11, f12, f15, f18 and fn. RVM3 is used to classify f3, f6, f18 vs. f2, f11, f12, f15 and fn. RVM4 is used to classify f3 vs. f6, f18. RVM5 is used to classify f15 vs. f2, f11, f12 and fn. RVM6 is used to classify f6 vs. f18. RVM7 is used to classify fn vs. f2, f11, f12. RVM8 is used to classify f11 vs. f2 and f12. RVM9 is used to classify f2 vs. f12.

In the training phase, the training samples are grouped according to Figure 20. Then RVM1 to RVM9 are trained using the corresponding group of training samples. After that, by inputting the feature vector into the trained multi-classifier, the fault type of the SEVA actuator can be identified.

5. Experiment and Results

To verify the proposed method, a comparison experiment for fault diagnosis reliability and fault diagnosis efficiency of RVM and SVM is conducted. A comparison experiment of RVM performance under simulated actuator data and actual actuator data is also done. As described in Section 3.2, the 190 groups of simulated actuator data for each state are generated using DABLib toolbox, which is driven by low-frequency sine wave, and pressure signals P1 and P2 are driven by the sine functions with appropriate amplitudes

and frequencies. The 90 groups of actual actuator data for each state are come from Lublin Sugar Factory in which the 19 kinds of fault are added using the DABLib model. In this research, for the 20 kind's state of the SEVA actuator, we just consider the status identification of f1, f2, f3, f6, f11, f12, f15, f17, f18 and fn.

Figure 21 shows the schematic of SEVA pneumatic actuator fault diagnosis method based on RVM. The fault diagnosis steps shows as following:

(1) Establishing the two actuator models g_1 and g_2 by RVM regression using the fault-free data.

(2) Training procedure. First of all, getting all the prediction of the training samples using the two actuator models based on RVM regression, then, generating the residuals of r_1 and r_2 by comparing the prediction outputs and actual actuator outputs, which is used as the input feature vector of the multi-classifier. Finally, training the designed RVM multi-classifier, that is shown in Figure 20.

(3) Fault identification procedure. For testing data of SEVA actuator, first of all, the residuals are generated using the two actuator models g_1 and g_2 , then, the fault status is identified by taking the residuals in to the RVM multi-classifier.

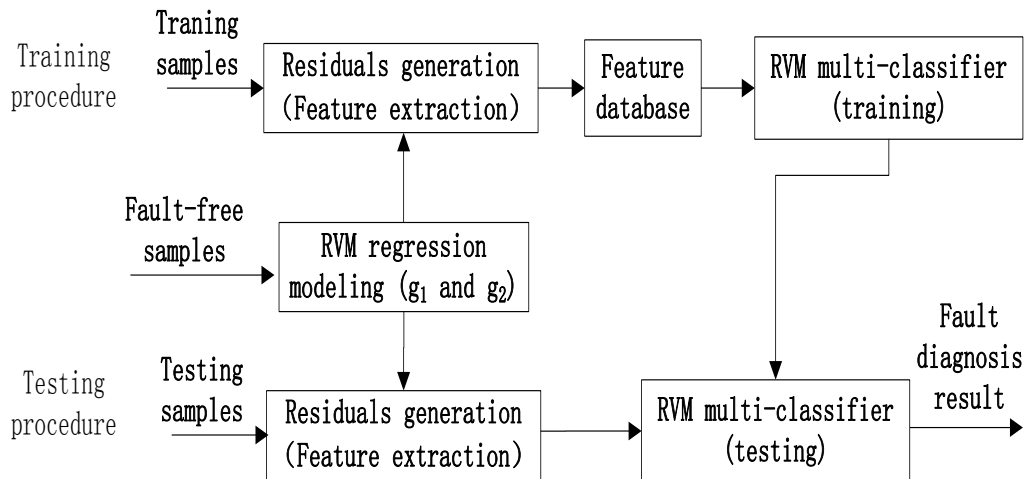


Figure 21. The Schematic of SEVA Pneumatic Actuator Fault Diagnosis Method based on RVM

5.1. Experiment and Results with Simulated Actuator Data

According to the training and the testing procedures of proposed approach described in Figure 21, 20 groups of simulated data are used to establish the two actuator models and train the RVM multi-classifier; the other 170 groups are used for testing. Table 2 shows the detailed fault identification result of the testing samples with RVM method, which indicates that the identification ratio for all fault modes is very high. It shows that all f1, f11, f15, f17, f18 and fn samples are identified as the correct state. Three f2 samples are determined falsely as f15. One f3 sample is determined falsely as f1, One f3 sample is determined falsely as f15. One f6 sample is determined falsely as f3. One f12 sample is determined falsely as f15. Table 3 shows the detailed information for every RVMs in the RVM multi-classifier. It indicates that the relevance vector numbers are far less than the training samples numbers, which gives a good evident for the sparsity of RVM.

Table 2. Fault Identification Results of the Testing Samples (Simulated Data with RVM)

Testing data	Fault Identification results										Correct Num	Identification ratio
	f1	f2	f3	f6	f11	f12	f15	f17	f18	fn		
f1(170)	170	0	0	0	0	0	0	0	0	0	170	100%
f2(170)	0	167	0	0	0	0	3	0	0	0	167	98.24%
f3(170)	1	0	168	0	0	0	1	0	0	0	168	98.82%
f6(170)	0	0	1	169	0	0	0	0	0	0	169	99.41%
f11(170)	0	0	0	0	170	0	0	0	0	0	170	100%
f12(170)	0	0	0	0	0	169	1	0	0	0	169	99.41%
f15(170)	0	0	0	0	0	0	170	0	0	0	170	100%
f17(170)	0	0	0	0	0	0	0	170	0	0	170	100%
f18(170)	0	0	0	0	0	0	0	0	170	0	170	100%
fn(170)	0	0	0	0	0	0	0	0	0	170	170	100%

Table 3. Detailed Information For Every RVMs in the RVM Multi-Classifier (Simulated Data)

Parameters	RVM1	RVM2	RVM3	RVM4	RVM5	RVM6	RVM7	RVM8	RVM9
Sigma (σ) for gauss kernel	2	2	2	2	2	2	2	2	2
RV Num/Train sample Num	9/200	7/180	6/160	2/60	4/100	2/40	3/80	3/60	3/40

In order to verify the advantages of RVM, the comparison of RVM multi-classifier and SVM multi-classifier for simulated actuator data is conducted. For the SVM multi-classifier, the input feature is also the residuals r1 and r2 for RVM multi-classifier. So the hierarchical classification structure of the SVM multi-classifier is the same as the RVM multi-classifier shown as in Figure 20. It only needs to change the RVM binary classifier to SVM binary classifier. The same as RVM multi-classifier, 20 groups of simulated data are used to train the SVM multi-classifier; the other 170 groups are used for testing.

Table 4 shows the detailed fault identification result of the testing samples with SVM multi-classifier, it indicates that except f2 and f6, the identification ratio for the other fault types are all decreased. The overall identification ratio of RVM multi-classifier is 99.60%, the overall identification ratio of SVM multi-classifier is 98.53%. Obviously, for the simulated actuator data, the accuracy of RVM multi-classifier is little higher than SVM multi-classifier. Table 5 shows the detailed information for every SVMs in the SVM multi-classifier. It obviously indicates that, the support vector for each SVMs greatly increased, the sparsity of RVM is much better than SVM. So for online identification, RVM multi-classifier occupies less memory and consumes less CPU time than SVM multi-classifier.

Table 4. Fault Identification Results of the Testing Samples (Simulated Data with SVM)

Testing data	Fault Identification results										Correct Num	Identification ratio
	f1	f2	f3	f6	f11	f12	f15	f17	f18	fn		
f1(170)	169	1	0	0	0	0	0	0	0	0	169	99.41%
f2(170)	0	169	0	0	0	0	1	0	0	0	169	99.41%
f3(170)	0	2	164	0	0	1	1	0	1	0	164	96.47%
f6(170)	0	1	0	169	0	0	0	0	0	0	169	99.41%
f11(170)	0	0	0	0	170	0	0	0	0	0	170	100%
f12(170)	0	0	0	0	0	169	1	0	0	0	169	99.41%
f15(170)	0	4	0	0	0	0	166	0	0	0	166	97.65%
f17(170)	0	5	0	0	0	0	0	165	0	0	170	97.06%
f18(170)	0	4	0	0	0	0	0	0	164	0	170	97.65%
fn(170)	0	2	0	0	0	0	0	0	0	168	170	98.82%

Table 5. Detailed Information Foe every SVMs in the SVM Multi-classifier (Simulated data)

Parameters	SVM1	SVM2	SVM3	SVM4	SVM5	SVM6	SVM7	SVM8	SVM9
Sigma (σ) for gauss kernel	2	2	2	2	2	2	2	2	2
SV Num/Train sample Num	39/200	32/180	21/160	15/60	16/100	16/40	20/80	10/60	12/40

5.2. Experiment and Results with Actual Actuator Data

For further validation of the proposed approach, 20 groups of actual actuator data are used to establish the two actuator models and train the RVM multi-classifier; the other 70 groups are used for testing. Table 6 shows the detailed fault identification result of the testing samples with RVM method, which indicates that the identification ratio for all fault modes is also very high, but a little lower than the simulated data. This is because the actual actuator data have more noise and more uncertain information. Table 7 shows the detailed information for every RVMs in the RVM multi-classifier. It also indicates that the relevance vector numbers are far less than the training samples numbers.

Table 6. Fault Identification Results of the Testing Samples (Actual Data with RVM)

Testing data	Fault Identification results										Correct Num	Identification Ratio
	f1	f2	f3	f6	f11	f12	f15	f17	f18	fn		
f1(170)	68	0	2	0	0	0	0	0	0	0	68	97.14%
f2(170)	0	70	0	0	0	0	0	0	0	0	70	100%
f3(170)	0	0	66	0	0	0	0	0	0	4	66	94.29%
f6(170)	0	0	0	70	0	0	0	0	0	0	70	100%
f11(170)	0	0	0	0	70	0	0	0	0	0	70	100%
f12(170)	0	0	0	0	0	65	0	0	0	5	65	92.86%
f15(170)	0	2	1	1	0	0	59	0	0	7	59	84.29%
f17(170)	0	0	0	0	0	0	0	70	0	0	70	100%
f18(170)	0	0	0	0	0	0	0	0	70	0	70	100%
fn(170)	0	0	0	0	0	0	0	0	0	70	70	100%

Table 7. Detailed Information Foe every RVMs in the RVM Multi-classifier (Actual Data)

Parameters	RVM1	RVM2	RVM3	RVM4	RVM5	RVM6	RVM7	RVM8	RVM9
Sigma (σ) for gauss kernel	2	2	2	2	2	2	2	2	2
RV Num/Train sample Num	3/200	3/180	3/160	2/60	9/100	2/40	7/80	2/60	2/40

In order to verify the advantages of RVM, the comparison of RVM multi-classifier and SVM multi-classifier for actual actuator data is also conducted. The same as RVM multi-classifier, 20 groups of simulated data are used to train the SVM multi-classifier; the other 70 groups are used for testing.

Table 8 shows the detailed fault identification result of the testing samples with SVM multi-classifier, it indicates that the identification ratio is decreased than the RVM multi-classifier. The overall identification ratio of RVM multi-classifier for actual data is 96.86%, the overall identification ratio of SVM multi-classifier is 93.28%. And also, the identification ratio of RVM multi-classifier and SVM multi-classifier for the actual data are all decreased than the simulated data. This is because the actual actuator data have more noise and more uncertain information. Meanwhile, the reducing magnitude of the SVM multi-classifier is larger than the RVM multi-classifier, which indicates that the performance of RVM is better than the SVM. Table 9 shows the detailed information for

every SVMs in the SVM multi-classifier. It also obviously indicates that, the support vector for each SVMs greatly increased, the sparsity of RVM is much better than SVM.

Table 8. Fault Identification Results of the Testing Samples (Actual Data with SVM)

Testing data	Fault Identification results										Correct Num	Identification ratio
	f1	f2	f3	f6	f11	f12	f15	f17	f18	fn		
f1(70)	67	0	3	0	0	0	0	0	0	0	67	95.71%
f2(70)	0	68	0	0	0	2	0	0	0	0	68	97.14%
f3(70)	0	0	67	3	0	0	0	0	0	0	67	95.71%
f6(70)	0	0	0	70	0	0	0	0	0	0	70	100%
f11(70)	0	0	0	0	69	0	0	0	0	1	69	98.57%
f12(70)	0	0	0	0	0	66	0	0	0	4	66	94.29%
f15(70)	0	1	5	0	3	9	166	0	0	6	46	65.71%
f17(70)	0	0	0	0	0	0	0	70	0	0	70	100%
f18(70)	0	0	0	0	0	0	0	0	70	0	70	100%
fn(70)	0	0	0	0	0	10	0	0	0	60	60	85.71%

Table 9. Detailed Information For Every SVMs in the SVM Multi-classifier (Actual Data)

Parameters	SVM1	SVM2	SVM3	SVM4	SVM5	SVM6	SVM7	SVM8	SVM9
Sigma (σ) for gauss kernel SV	2	2	2	2	2	2	2	2	2
Num/Train sample Num	52/200	31/180	30/160	12/60	25/100	12/40	21/80	17/60	9/40

6. Conclusions

Fault diagnosis of SEVA pneumatic actuator based on RVM regression modeling and RVM multi-classifier is proposed. Comparison results of RVM method and SVM method using the simulated actuator data by DABLib and actual actuator from Lublin Sugar Factory indicate that the proposed approach overcomes the drawbacks of SVM and resolves the small sample and nonlinear problem in SEVA pneumatic actuator fault diagnosis.

(1) Fault identification ratio of RVM method is better than SVM method, especially for the results of actual actuator data.

(2) Whether the RVM method or the SVM method, the identification ratio for actual actuator data are all poor than the simulated actuator data, this is because the actual actuator data have more noise and more uncertain information.

(3) For the SVM multi-classifier and RVM multi-classifier, the support vectors for each SVMs are far more than the relevance vectors for each RVMs, the sparsity of RVM is much better than SVM. For online identification, RVM multi-classifier occupies less memory and consumes less CPU time than SVM multi-classifier.

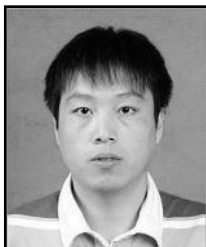
Acknowledgements

The authors would like to thank the financial support of the financial support of the National Natural Science Foundation of China NSFC-61104023.

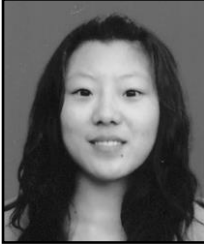
References

- [1] J. Faisel, J. Ron and M. Witczak, "A neuro-fuzzy multiple-model observer approach to robust fault diagnosis based on the DAMADICS benchmark problem", *Control Engineering Practice*, vol. 14, no. 6, (2006), pp. 699 -717.
- [2] J. Yang and D. W. Clarke, "Self-validating Actuator", *Control Engineering Practice*, vol. 7, no. 2, (1999), pp. 249 - 260.
- [3] Z. Feng, Q. Wang and K. Shida, "Design and Implementation of a Self-validating Pressure Sensor", *IEEE Sensors Journal*, vol.9, no.3, (2009), pp. 207 - 218.
- [4] Z. Feng and X. Zhang, "Pneumatic Actuator Fault Diagnosis Based on LS-SVM and SVM", *Chinese Journal of Sensors and Actuators*, vol. 26, no.11, (2013), pp. 1610 - 1616.
- [5] W. Chen and M. Sail, "Actuator fault diagnosis for a class of nonlinear systems and its application to a laboratory 3D crane", *Automatica*, vol. 47, no. 7, (2011), pp. 1435 - 1442.
- [6] Z. Feng, X. Zhang and H. Yang, "Research of Pneumatic Actuator Fault Diagnosis Method Based on GA Optimized BP Neural Network and Fuzzy Logic", *Lecture Notes in Computer Science*, vol. 7952, no. 2, (2013), pp. 578 - 585.
- [7] J. Yan, Y. Liu, S. Han and M. Qiu, "Wind Power Grouping Forecasts and Its uncertainty Analysis Using Optimized Relevance Vector Machine", *Renewable and Sustainable Energy Reviews*, vol. 27, (2013), pp. 613 - 621.
- [8] M. E. Tipping, "Sparse Bayesian Learning and the Relevance Vector Machine", *Journal of Machine Learning Research*, vol. 1, (2001), pp. 211 - 244.
- [9] A. Widodo, E. Y. Kim, J. D. Son, B. S. Yang, A. C. C. Tan, D. S. Gu, *et al.*, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine", *Expert Systems with Applications*, vol. 36, no. 3, (2009), pp. 7252 – 7261.
- [10] M. Matsumoto and J. Hori, "Classification of silent speech using support vector machine and relevance vector machine", *Applied Soft Computing*, vol. 19, (2014), pp. 12 – 19.
- [11] E. Zio and F. D. Maio, "Fatigue Crack Growth Estimation by Relevance Vector Machine", *Expert Systems with Applications*, vol. 39, no. 12, (2012), pp. 10681 – 10692.
- [12] M. Bartys, R. Patton, M. Syfert, *et al.*, "Introduction to the DAMADICA Actuator FDI Benchmark Study", *Control Engineering Practice*, vol. 14, no. 6, (2006), pp. 577 – 596.
- [13] F. Wang, B. Gou and Y. Qin, "Modeling Tunneling-induced Ground Surface Settlement Development Using a Wavelet Smooth Relevance Vector Machine", *Computers and Geotechnics*, vol. 54, (2013), pp. 125 – 132.
- [14] F. He, J. Xu, M. Li and J. Yang, "Product Quality Modeling and Prediction Based on Wavelet Relevance Vector Machines", *Chemometrics and Intelligent Laboratory System*, vol. 21, (2013), pp. 33 – 41.
- [15] J. Yuan, K. Wang, T. Yu and M. Fang, "Integrating Relevance Vector Machines and Genetic Algorithms for Optimization of Seed-separating Process", *Engineering Applications of Artificial Intelligence*, vol. 20, no. 7, (2007), pp. 970 - 979.
- [16] K. Bastani, Z. Kong, W. Huang, X. Huo, *et al.*, "Fault Diagnosis Using an Enhanced Relevance Vector Machine (RVM) for Partially Diagnosable Multi-station Assembly Processes", *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 1, (2013), pp. 124 - 136.
- [17] D. Chao and H. Zhao, "Hyperspectral Image Classification and Application Based on Relevance Vector Machine", *Journal of Remote Sensing*, vol. 14, no. 6, (2010), pp. 1273 - 1284.

Authors



Zhigang Feng, he is an associate professor of Shenyang Aerospace University. He received his Doctor's Degree from Harbin Institute of Technology, P.R. China at 2009. His main research direction includes system fault diagnosis, self-validating sensor and self-validating actuator.



Ru Wang, she received the bachelor degree of Automation from Jilin University of Chemical Technology, China, in 2011. She is currently studying for a master degree in the School of Automation, Shenyang Aerospace University.

