SVM Approximate-based Internal Model Control Strategy

WANG Yao-Nan[1] YUAN Xiao-Fang[1]

Abstract A support vector machine (SVM) approximate-based internal model control (IMC) strategy is presented for the steam valving control of synchronous generators. The proposed SVM IMC strategy includes two main parts: SVM approximate inverse controller and uncertainty compensation in the internal model structure. The SVM inverse controller is derived directly using an input-output approximation approach via Taylor expansion, and it is implemented through nonlinear system identification without further online training. Furthermore, a robustness filter is used for uncertainty compensation in the internal model structure. Simulations show the effectiveness of the SVM IMC strategy for the steam valving control.

Key words Nonlinear control, internal model control (IMC), support vector machines, neural networks, approximate models

Nonlinear internal model control (NIMC) strategy has many industrial applications because of its desirable properties, in particular, good robustness against disturbances, and model mismatch[1]. In the implementation of NIMC, the determination of a plant model represents an important stage of the development. For most practical applications, such as chemical and power plants, it is a difficult task. Furthermore, the inversion of the nonlinear model is shown to play a crucial role. Some researchers have studied analytical and numerical iterative algorithms for the construction of nonlinear operator inverses. In this case, the necessary and sufficient convergence conditions cannot be ensured. This problem is a great handicap in the process control applications[2−3]. Because neural networks (NNs) are capable of approximating any nonlinear dynamics with an arbitrary degree of accuracy, the neural networks based nonlinear internal model control (NN IMC) has attracted much attention[3−5]. NN IMC is an extension of NIMC with NN model as the internal model and NN inverse controller replacing the inverse controller. However, for NN IMC, even though the NN model is available, it is still not easy to design an NN inverse controller. Unfortunately, most neural networks use gradient-based training method like back-propagation, often suffering from the existence of local minima, and it is also not easy to choose a suitable neural networks structure like the number of hidden neurons.

As a novel breakthrough to NNs, support vector machines (SVM)[6−8], a novel machine learning algorithm, has proved to be a powerful alternative in many areas. So far, several theoretical and experimental studies on model identification and control using SVM have been reported in the literature[9−10]. The objective of this paper is to develop an SVM approximate-based internal model control (SVM IMC) strategy for the steam valving control of synchronous generator. In the SVM IMC strategy, SVM approximate inverse controller is derived directly based on a novel input-output approximation via Taylor expansion, and it is implemented straightforward through nonlinear system identification, and furthermore, a robustness filter is used for uncertainty compensation in internal model structure. Because of the innovative approximation, the general SVM training algorithms, such as those available in SVM toolbox, can be used directly.

Received October 27, 2006; in revised form January 11, 2007
Supported by National Natural Science Foundation of China (60375001, 60775047)
1. College of Electrical and Information Engineering, Hunan University, Changsha 410082, P. R. China
DOI: 10.3724/SP.J.1004.2008.00172

1 Plant dynamical model

Due to the fact that steam valving control of synchronous generator cannot only improve the transient stability and dynamic capability but also damp oscillations in power systems, it has become an important area in power system control. A multi-machine system may be described by the following state equation[11]

$$\begin{aligned}&\dot{\delta}_i = \omega_i - \omega_0\\ &\dot{\omega}_i = -\frac{D_i(\omega_i-\omega_0)}{H_i} + \frac{\omega_0}{H_i}(P_{Hi} + C_{Mi}P_{m0i} - \frac{E_{qi}^{'}V_s\sin\delta_i}{X_{d\Sigma i}^{'}})\end{aligned} \tag{1}$$

where it is assumed that excitation control is static; that is, the internal voltage $E_{qi}^{'}(i=1,2,\cdots,n)$ is supposed to be constant, subscript i meaning the ith power-generator unit. δ_i is the power angle of the ith generator, in radian; ω_i the relative speed of the ith generator; P_{Hi} the mechanical power of high-pressure steam; C_{Mi} the distribution factor in middle-pressure steam; P_{m0i} the mechanical input power in p.u.; H_i the inerta constant in second. If we only consider the high-pressure steam valve control and neglect fast valve control, the steam valve control scheme is then described as

$$\dot{P}_{Hi} = -\frac{1}{T_{H\Sigma i}}P_{Hi} + \frac{C_{Hi}}{T_{H\Sigma i}}P_{m0i} + \frac{C_{Hi}}{T_{H\Sigma i}}u_i \tag{2}$$

where $T_{H\Sigma i} = T_{Hgi} + T_{Hi}$ is the equivalent time constant of high-pressure steam; C_{Hi} is the distribution factor in high-pressure steam; u_i is the control signal of steam valve. Therefore, steam valving control is just to regulate u_i for the stabilization of the power angle δ_i in each generator.

Suppose the power angle δ_i is as the output y_i, and the control signal of steam valving u_i as the input. It can then be derived from (1) that

$$y_i^{(3)} = -\frac{D_i}{H_i}\ddot{y}_i + \frac{\omega_0}{H_i}\dot{P}_{Hi} - \frac{\omega_0}{H_i}\frac{E_{qi}^{'}V_s}{X_{d\Sigma i}^{'}}\cos y_i\cdot\dot{y}_i \tag{3}$$

Substituting $\dot{P}_{Hi}$ of (2) into (3) yields

$$y_i^{(3)} = -\frac{\omega_0}{H_i}[\frac{D_i}{\omega_0}\ddot{y}_i + \frac{E_{qi}^{'}V_s}{X_{d\Sigma i}^{'}}\cos y_i \cdot \dot{y}_i - \frac{C_{Hi}}{T_{H\Sigma i}}(P_{m0i} - \frac{P_{Hi}}{C_{Hi}} + u_i)] \tag{4}$$

Substituting P_{Hi} of (1) into (4) yields a nonlinear relation

$$y_i^{(3)} = g(\ddot{y}_i, \dot{y}_i, y_i, u_i) \tag{5}$$

Using the n-order approximation method[12], one has $T\dot{y} = y(k+1) - y(k)$, $T^2\ddot{y} = y(k+1) - 2y(k) + y(k-1)$, $T^3y^{(3)} = y(k+1) - 3y(k) + 3y(k-1) - y(k-2)$, with T being the sampling period. In this manner, (5) can then be described in the discrete system as

$$y(k+1) = f(y(k), y(k-1), y(k-2), u(k)) \tag{6}$$

where $u(k) \in \mathbf{R}$ and $y(k+1) \in \mathbf{R}$ are the control input and system output at time step k and $k+1$, respectively. $f(\cdot)$ in (6) is regarded as an unknown nonlinear mapping, whereas both the relative degree d and plant order n are known. (6) is the so-called nonlinear autoregressive moving average (NARMA) model. For a general discrete-time nonlinear system, the NARMA model is an exact representation of its input-output behavior over the range in which the system operates[13].

2 SVM IMC strategy

The proposed SVM IMC strategy is based on input-output approximation, which not only avoids complex inverse controller development and intensive computation but also avoids online learning or adjustment. The basic structure of the SVM IMC strategy is described in Fig. 1, which is an extension of NN IMC as shown in Fig. 2. The SVM approximate inverse controller and uncertainty compensation in Fig. 1 are two major parts of this control strategy. Comparing Fig. 1 with the basic structure of NN IMC in Fig. 2, we find that the SVM approximate inverse controller replaced the NN inverse controller and the approximate SVM model was used in uncertainty compensation.

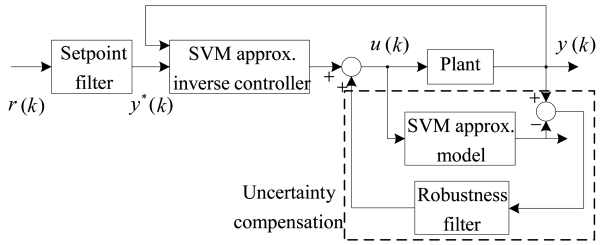


Fig. 1 The structure of SVM IMC

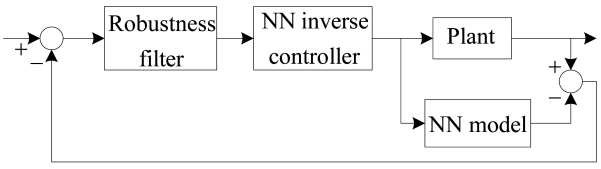


Fig. 2 The structure of NN IMC

2.1 SVM approximate inverse controller

It is known that under certain conditions, an exact input-output representation of the plant is given by the nonlinear autoregressive moving average (NARMA) model in a neighborhood of the equilibrium state. Even assuming that such a model is available, determining the control input resulting in a desired output is no longer a simple task because the output depends nonlinearly on the input. As a consequence, various approximate methods have been proposed in the literature for the determination of the control input[5, 13−15]. In this paper, an approximate method via Taylor expansion is used to design this inverse controller.

2.1.1 Approximate inverse control law

For the NARMA model (6), a Taylor expansion of the system gives

$$\begin{aligned} y(k+1) &= f[Y(k), u(k)] = f[Y(k-1), u(k-1)] + \\ &\sum \frac{1}{r!}\frac{\partial^r f[Y(k), u(k)]}{\partial u(k-1)^r}[\Delta u(k))]^r + \\ &\sum \frac{1}{r!}\frac{\partial^r f[Y(k), u(k)]}{\partial Y(k-1)^r}[\Delta Y(k)]^r \end{aligned} \tag{7}$$

where $Y(k) = [y(k), y(k-1), y(k-2)]$, Δ is the increment operator, and $f[Y(k), u(k)]$ is the output of the system at time k, just equivalent to (7). Here, $f[Y(k-1), u(k-1)]$ is the equivalent denotation of $y(k)$, that is, $f[Y(k-1), u(k-1)] = y(k)$.

Assumption 1[14]**.** The output of system $y(k+1)$ is highly sensitive to the input $u(k)$ in the operating region, that is,

$$|\frac{\partial f[Y(k), u(k)]}{\partial u(k-1)}| \gg |\frac{\sum \frac{1}{r!}\frac{\partial^r f[Y(k), u(k)]}{\partial Y(k-1)^r}[\Delta Y(k)]^r}{\Delta u(k)}| \tag{8}$$

From this assumption, we can drop the third term on the right-hand side of (7) to represent the plant by the following equation:

$$\begin{aligned} y(k+1) &= y(k) + f^1[Y(k-1), u(k-1)] \cdot \Delta u(k) + \\ &R[Y(k-1), u(k-1), \Delta u(k)] \end{aligned} \tag{9}$$

where

$$f^1[Y(k), u(k)] = \frac{\partial f[Y(k-1), u(k-1)]}{\partial u(k-1)} \tag{10}$$

Theorem 1[15]**.** The remainder term $R[Y(k-1), u(k-1), \Delta u(k)]$ in (9) approaches zero at a faster rate than $\Delta u(k)$, and there exits a variable $\epsilon(k) \in (0, +\infty]$ such that

$$|\frac{R[Y(k-1), u(k-1), \Delta u(k)]}{\Delta u(k-1)}| \ll |f^1[Y(k), u(k)]| \tag{11}$$

whenever $|\Delta u(k)| \in [0, \epsilon(k)]$.

Remark 1. According to the Taylor expansion theory, as $|\Delta u(k)| \in [0, \epsilon(k)]$, $R[Y(k-1), u(k-1), \Delta u(k)]$ (Also denoted by R_k in the following text) in (9) is bounded by

$$|R_k| \leq \frac{r_0 \varepsilon^2(k)}{2} \tag{12}$$

with r_0 being a finite positive number.

From (8)∼(12), the input-output approximate model using Taylor expansion can be derived by neglecting the remainder R_k and therefore (9) can be rewritten as follows.

$$y(k+1) \approx y(k) + f^1[Y(k-1), u(k-1)] \cdot \Delta u(k) \quad (13)$$

Obviously, $f^1[Y(k-1), u(k-1)]$ is a nonlinear function, and, here, we consider the case where $f^1[Y(k-1), u(k-1)]$ exists but is unknown. Therefore, an SVM network will be used to estimate $f^1[Y(k-1), u(k-1)]$ from the input-output plant signals. Here, the approximation model via SVM approach is referred to as $\hat{f}^1[Y(k-1), u(k-1)]$. Using (13), the control law can be determined directly because the increment $\Delta u(k)$ of the control signal appears linearly.

Before the determination of control law, we give the following assumptions:

Assumption 2[16]**.** $0 < |f^1[Y(k-1), u(k-1)]| \leq g_0$ with g_0 as a finite positive number.

Assumption 3[5]**.** $0 \leq |\Delta u(k)| \leq \varepsilon(k)$ with $\varepsilon(k)$ as a finite positive number.

Assumption 4. $|y^*(k) - y^*(k-1)| \leq \Delta r$ with Δr as a finite positive number.

Assumption 5. $f^1[Y(k-1), u(k-1)] = \hat{f}^1[Y(k-1), u(k-1)] + \xi(k)$. $\xi(k)$ is an approximation error and $|\xi(k)| \leq \bar{\xi}$ with $\bar{\xi}$ being a small positive number; $0 \leq |\frac{\xi(k)}{\hat{f}^1[Y(k-1), u(k-1)]}| \leq \xi_0$ with ξ_0 as a finite positive number smaller than 1.

Remark 2. In Assumption 4, the plant output is required to slowly track the varying reference signals because the power angle of synchronous generator cannot change very fast.

Remark 3. Assumption 5 show that modelling errors are bounded by finite quantities, and the errors can be made finitely small using the general SVM approach with appropriate parameters[7].

Using the input-output approximation model derived in the former section, the SVM approximate inverse control law can be determined straightforwardly from (13) as follows

$$u(k) = u(k-1) + \Delta u(k) \quad (14)$$

$$\Delta u(k) = \frac{y^*(k+1) - y(k)}{\hat{f}^1[Y(k-1), u(k-1)]}, \text{for} |\Delta u(k)| \leq \varepsilon(k)$$

$$\Delta u(k) = \varepsilon(k) \cdot \text{sign}(\Delta u(k)), \ \text{for} |\Delta u(k)| > \varepsilon(k)$$

where $y^*(k+1)$ are the desired trajectory, sign$(\cdot)$ is the signum function, and $\varepsilon(k)$ is not unique because it can be set at any positive value less than the one known to satisfy (11). This control law requires precise information about $\hat{f}^1[Y(k-1), u(k-1)]$ to compute $\Delta u(k)$. In this paper, we consider the case in which $\hat{f}^1[Y(k-1), u(k-1)]$ is derived from SVM network output as in Section 4. The control law of (14) will be practical only if $\hat{f}^1[Y(k-1), u(k-1)] \neq 0$, and this is conducted under Assumption 2. A key feature of this approach is that the inverse model controller can then be simply designed in terms of SVM approximate.

Define control error as

$$e(k) = y^*(k) - y(k) \quad (15)$$

From (9), we obtain

$$e(k+1) = y^*(k+1) - y(k+1) = y^*(k+1) - y(k) - f^1[Y(k-1), u(k-1)] \times \Delta u(k) - R_k - \nu_k \quad (16)$$

where ν_k is modeled as the effect of uncertainties or disturbances, $|\nu_k| \leq \nu_0$ with ν_0 as a finite positive number.

2.1.2 Stability of approximate control law

The robustness of the stability and the performance for the control law (14) are given in Theorem 2.

Theorem 2. Under Assumptions 2∼5, using the control law (14), the solution of error system (16) is uniformly ultimately bounded (UUB)[17] for all k with ultimate bound $\lim_{k\to\infty} |e(k)| \leq \frac{|b_1|+2|b_2|}{2b_0}$, where $b_0 = \beta(k)\alpha(k)[2-\beta(k)\alpha(k)]$, $b_1 = 2[1-\beta(k)\alpha(k)][[1-\beta(k)\alpha(k)] \cdot \Delta r - (R_k + \nu_k)]$, $b_2 = [[1-\beta(k)\alpha(k)] \cdot \Delta r - (R_k + \nu_k)]$, and $0 < \alpha(k) \leq 1$, $\beta(k) = \frac{f^1[Y(k-1), u(k-1)]}{\hat{f}^1[Y(k-1), u(k-1)]}$, Δr and R_k, ν_k are defined in Assumption 4 and (16), respectively.

Proof. Define a variable $\alpha(k)$ where $0 < \alpha(k) \leq 1$ for all k. The control law (12) is equivalently expressed as

$$\Delta u(k) = \frac{y^*(k+1) - y(k)}{\hat{f}^1[Y(k-1), u(k-1)]} \alpha(k) \quad (17)$$

where $\alpha(k) = 1$, if $|\Delta u(k)| \leq \varepsilon(k)$, and $0 < \alpha(k) < 1$, if $|\Delta u(k)| > \varepsilon(k)$.

Using (17), (16) can be presented as

$$\begin{aligned} & e(k+1) = y^*(k+1) - y(k) - \\ & f^1[Y(k-1), u(k-1)] \cdot \Delta u(k) - R_k - \nu_k = \\ & y^*(k+1) - y(k) - [\hat{f}^1[Y(k-1), u(k-1)] + \xi(k)] \times \\ & \frac{y^*(k+1) - y(k)}{\hat{f}^1[Y(k-1), u(k-1)]} \alpha(k) - R_k - \nu_k \end{aligned} \quad (18)$$

Let $\frac{f^1[Y(k-1), u(k-1)]}{\hat{f}^1[Y(k-1), u(k-1)]} = \beta(k)$. Then,

$$e(k+1) = [y^*(k+1) - y(k)](1 - \beta(k) \cdot \alpha(k)) - R_k - \nu_k \quad (19)$$

Define the Lyapunov candidate $V(k) = e^2(k)$. Then,

$$\begin{aligned} & V(k+1) - V(k) = e^2(k+1) - e^2(k) = \\ & \{[y^*(k+1) - y(k)][1 - \beta(k)\alpha(k)] - R_k - \nu_k\}^2 - e^2(k) = \\ & \{[y^*(k) + \Delta r - y(k)][1 - \beta(k)\alpha(k)] - R_k - \nu_k\}^2 - e^2(k) = \\ & \{[e(k) + \Delta r][1 - \beta(k)\alpha(k)] - (R_k + \nu_k)\}^2 - e^2(k) = \\ & \beta(k)\alpha(k)[\beta(k)\alpha(k) - 2] \cdot e^2(k) + \\ & 2[1 - \beta(k)\alpha(k)][[1 - \beta(k)\alpha(k)] \cdot \Delta r - (R_k + \nu_k)] \cdot e(k) + \\ & \{[1 - \beta(k)\alpha(k)] \cdot \Delta r - (R_k + \nu_k)\}^2 = \\ & - b_0 e^2(k) + b_1 e(k) + b_2^2 \end{aligned} \quad (20)$$

According to Assumption 5 and (17), $b_0 = \beta(k)\alpha(k)[2 - \beta(k)\alpha(k)] > 0$.

We have

$$b_1^2 + 4b_0 b_2^2 = 4\{[1 - \beta(k)\alpha(k)] \cdot \Delta r - (R_k + \nu_k)\}^2 = 4b_2^2 \geq 0 \quad (21)$$

Therefore, it follows from (20) that $V(k+1)-V(k)<0$, when $e(k)>\frac{b_1+\sqrt{b_1^2+4b_0b_2^2}}{2b_0}$ or $e(k)<\frac{b_1-\sqrt{b_1^2+4b_0b_2^2}}{2b_0}$.

From (21), we have $b_1^2+4b_0b_2^2=4b_2^2$. Hence, $V(k+1)-V(k)<0$ when

$$e(k)>\frac{b_1+|2b_2|}{2b_0} \text{ or } e(k)<\frac{b_1-|2b_2|}{2b_0} \tag{22}$$

Therefore, we can conclude that the solutions of (18) are uniformly ultimately bounded for all k with ultimate bound $|e(k)|_{k\to\infty}\leq\frac{|b_1|+2|b_2|}{2b_0}$. □

2.2 Uncertainty compensation in internal model structure

2.2.1 Uncertainty compensation

In practice, uncertainties are inevitable. In this paper, uncertainty attenuation can be achieved via internal model with a robustness filter as shown in Fig. 1. Then,

$$y(k+1)=y(k)+f^1[Y(k-1),u(k-1)](\Delta u(k)+\Delta u_c(k))+R_k+\nu_k \tag{23}$$

where $\Delta u_c(k)$ is the increment output of the robustness filter. Because $\zeta(k)=z^{-1}(R_k+\nu(k))$ and $\Delta u_c(k)=-\frac{F(z)\zeta(k)}{\hat{f}^1[Y(k-1),u(k-1)]}$, (23) implies

$$\begin{aligned}y(k+1)=&y(k)+f^1[Y(k-1),u(k-1)]\Delta u(k)-\\&\beta(k)F(z)\zeta(k)+R_k+\nu_k\end{aligned} \tag{24}$$

where $F_1=1-\beta(k)z^{-1}F(z)$.

Therefore, by suitably choosing filter $F(z)$, the remainder term R_k and disturbance ν_k can be attenuated to some extent[5]. Under the control architecture as shown in Fig. 1, the equivalent control law can be expressed as

$$u(k)=u(k-1)+\frac{[y^*(k+1)-y(k)]\alpha(k)-F(z)\zeta(k)}{\hat{f}^1[Y(k-1),u(k-1)]} \tag{25}$$

This control law consists of the SVM approximate inverse control and uncertainty compensation and, thus, combines the advantages of both inverse control and nonlinear IMC.

Remark 4. The control law in (25) implies the difference between the structure of SVM IMC and that of NN IMC. The plant output y_k is used in the SVM approximate inverse control law. In the mean time, the output of the robustness filter is the direct input to the plant. Furthermore, the SVM approximate inverse controller and the uncertainty compensation can be designed separately according to the control law (25).

2.2.2 Robustness of stability and performance

Under the SVM IMC strategy, the following result of the robust stability and performance is common as stated in Theorem 3.

Theorem 3. Under Assumptions 2 $\sim$ 5, using the control law (25), the solution of error system (16) is uniformly ultimately bounded (UUB)[17] for all k with ultimate bound $\lim_{k\to\infty}|e(k)|\leq\frac{|c_1|+2|c_2|}{2b_0}$, where $b_0=\beta(k)\alpha(k)[2-\beta(k)\alpha(k)]$, $c_1=2[1-\beta(k)\alpha(k)][[1-\beta(k)\alpha(k)]\cdot\Delta r-F_1(R_k+\nu_k)]$, $c_2=[[1-\beta(k)\alpha(k)]\cdot\Delta r-F_1(R_k+\nu_k)]$, $\alpha(k)$, $\beta(k)$, Δr and R_k, ν_k are the same as defined in Theorem 2, and F_1 is presented in (24).

Proof. Choosing the Lyapunov function as $V(k)=e^2(k)$, one has

$$\begin{aligned}&V(k+1)-V(k)=e^2(k+1)-e^2(k)=\\&\{[y^*(k+1)-y(k)][1-\beta(k)\alpha(k)]+\\&\beta(k)F(z)\zeta(k)-R_k-\nu_k\}^2-e^2(k)=\\&\{[y^*(k)+\Delta r-y(k)][1-\beta(k)\alpha(k)]-\\&F_1(R_k+\nu_k)\}^2-e^2(k)=\\&\{[e(k)+\Delta r][1-\beta(k)\alpha(k)]-F_1(R_k+\nu_k)\}^2-e^2(k)=\\&\beta(k)\alpha(k)[\beta(k)\alpha(k)-2]\cdot e^2(k)+\\&2[1-\beta(k)\alpha(k)][[1-\beta(k)\alpha(k)]\cdot\Delta r-F_1(R_k+\nu_k)]\times\\&e(k)+\{[1-\beta(k)\alpha(k)]\cdot\Delta r-F_1(R_k+\nu_k)\}^2=\\&-b_0e^2(k)+c_1e(k)+c_2^2\end{aligned} \tag{26}$$

Similar to (20), $b_0=\beta(k)\alpha(k)[2-\beta(k)\alpha(k)]>0$. Similar to (21), $c_1^2+4b_0c_2^2=4c_2^2\geq0$. Hence from (26), $V(k+1)-V(k)<0$, when $e(k)>\frac{c_1+|2c_2|}{2b_0}$ or $e(k)<\frac{c_1-|2c_2|}{2b_0}$.

Therefore, one concludes that the solutions of (18) are uniformly ultimately bounded for all k with ultimate bound $|e(k)|_{k\to\infty}\leq\frac{|c_1|+2|c_2|}{2b_0}$. □

Remark 5. Comparing Theorem 3 with Theorem 2, we will find that the error systems for both control law (25) and control law (14) are uniformly ultimately bounded (UUB) with ultimate bounds $\frac{|c_1|+2|c_2|}{2b_0}$ and $\frac{|b_1|+2|b_2|}{2b_0}$, respectively. Comparing c_1, c_2 with b_1, b_2, we can conclude that a suitable robustness filter $F(z)$ can compensate uncertainty or disturbance and at the same time improve the tracking accuracy. As the modeling error is larger at high frequencies, intuitively the addition of a low-pass filter adds robustness characteristics to the control architecture. The robustness filter is typically a first-order one with unit gain and frequency response is tuned to eliminate high-frequency noise introduced by measurement devices[3].

3 Implementation of SVM IMC

The theorems and algorithms presented in the previous sections require an approximate model $\hat{f}^1[Y(k-1),u(k-1)]$ of $f^1[Y(k-1),u(k-1)]$ for the plant. Clearly, those results are of limited values if the model $\hat{f}^1[Y(k-1),u(k-1)]$ of the plant is unknown. How to use SVM to obtain $\hat{f}^1[Y(k-1),u(k-1)]$ required by the control law (14) and (25).

3.1 SVM based function approximation

Over with ANN and standard SVM[6], least squares SVM(LS-SVM)[7−8]has the following advantages: no number of hidden units has to be determined, no centers have to be specified for the Gaussian kernel, and less parameters have to be prescribed, so LS-SVM is used here for the model identification.

Let $\{x_t,y_t\}_{t=1}^N$ be the set of input/output training data with input x_t and output y_t. Consider the regression model $y_t=f(x_t)+e_t$, where x_t are deterministic points, f is a smooth function and e_t are uncorrelated errors. To estimate the nonlinear f, the following model is assumed.

$$f(\boldsymbol{x})=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x})+b \tag{27}$$

where $\boldsymbol{\varphi}(\boldsymbol{x})$ denotes an infinite dimensional feature map.

The regularized cost function of the LS-SVM is given as

$$\min J(\boldsymbol{\omega}, e) = \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} + \boldsymbol{\gamma}\frac{1}{2}\sum e_t^2 \tag{28}$$

$$\text{s.t. } y_t = \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}(x_t) + b + e_t,\ t = 1, \cdots, N \tag{29}$$

To solve this constrained optimization, a Lagrangian is constructed

$$L(\boldsymbol{\omega}, b, e; \alpha) = J(\boldsymbol{\omega}, e) - \sum \alpha_t(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}(x_t) + b + e_t - y_t) \tag{30}$$

with α_t being the Lagrange multipliers. The conditions for optimality are given by

$$\frac{\partial L}{\partial \omega} = 0,\ \frac{\partial L}{\partial b} = 0,\ \frac{\partial L}{\partial e_t} = 0,\ \frac{\partial L}{\partial \alpha_t} = 0 \tag{31}$$

Substituting (27) $\sim$ (30) into (31) yields the following set of linear equations:

$$\begin{bmatrix} 0 & I_N^{\mathrm{T}} \\ I_N & \Omega + \gamma^{-1}I_N \end{bmatrix} \cdot \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix} \tag{32}$$

with $\boldsymbol{y} = [y_1, \cdots, y_N]^{\mathrm{T}}$, $I_N = [y_1, \cdots, 1]^{\mathrm{T}}$, $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_N]^{\mathrm{T}}$, $\Omega_{ij} = K(x_i, x_j)$.

The resulting LS-SVM model can be evaluated at a new point x_*

$$\hat{f}(x_*) = \sum \alpha_t K(x_*, x_t) + b \tag{33}$$

where M is the number of support vectors (SVs), $K(\cdot, \cdot)$ is kernel function, and α_t, b are the solutions to (32).

As one of the most popular kernel functions in machine learning, Gaussian kernel function is selected for controller design in this paper. It takes the following form:

$$K(x_i, x_j) = \exp(\frac{-\|x_i - x_j\|^2}{2\sigma^2}) \tag{34}$$

where σ denotes the kernel (bandwidth) parameter.

As the training of LS-SVM is equivalent to a linear programming problem, LS-SVM method can realize global optimization effectively. Moreover, the learning results decide the number of SVs and thus the nodes of hidden layer of SVM network are selected. It is well known that SVM generalization performance depends on a good setting of hyperparameters and the kernel parameters. Bayesian evidence framework is an effective way for parameters optimization of LS-SVM regression, and this approach is described in detail in [7]. According to the Bayesian evidence theory, the inference is divided into three distinct levels. Training of the LS-SVM regression can be statistically interpreted in level 1 inference. The optimal regularization parameter can be inferred in level 2. The optimal kernel parameter selection can be performed in level 3.

3.2 Implement of SVM IMC

Because SVM is a universal approximator, it is regarded as a convenient way to model a nonlinear input-output mapping. As is well known, system identification can be carried out using either parallel models or series-parallel models[13]. In general, most identification schemes use the series-parallel model, and it is also used in this paper. The series-parallel model utilizes the plant output to obtain the estimate at time, and the series-parallel model is used to predict future values of the output using current values of inputs and outputs. SVM is used to approximate the input-output representation (6) as shown in Fig. 3 and is often called SVM NARMA model. Using the general SVM learning algorithm, we can reach the approximate model of (6) as

$$\hat{y}(k+1) = \hat{f}[Y(k), u(k)] = \sum \alpha_t K[(Y(k), u(k)), (Y(t), u(t))] + b \tag{35}$$

and therefore we can get the approximate nonlinear function in the control law, that is,

$$\hat{f}[Y(k-1), u(k-1)] = \frac{\partial \hat{f}[Y(k-1), u(k-1)]}{\partial u(k-1)} = -\frac{u(k-1)}{\sigma^2}\sum \alpha_t K[(Y(k-1), u(k-1)), (Y(t), u(t))] \tag{36}$$

where variables are the same as in (33) and (34).

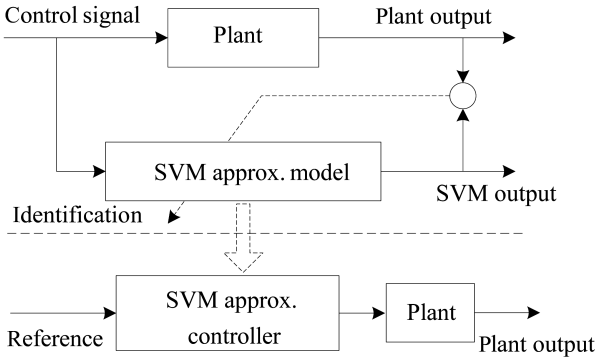


Fig. 3 SVM input/output representation for controller design

4 Simulations

It is assumed that the designer has sufficient prior information about excitation control and the creation of a suitable set of training data. The following procedure is designed to generate a set of data that adequately represents the dynamics of the plant for the anticipated range of input and output signals. Control signals $u(k)$ that vary in magnitude between 0 and 1 are applied to the plant. This range of control signal magnitudes is partitioned into 20 equal and non-overlapping subranges. A control signal $u(k)$ is semi-randomly generated, and at each instant k is subjected to the condition that $|\Delta u(k)| = |u(k) - u(k-1)| = 0.05$. This control signal is applied to the plant and the input-output data pairs obtained $[Y(k), u(k)]$ are recorded. The unknown values of $f^1[Y(k-1), u(k-1)]$ are estimated by (36).

Because the steady operating state of the power angle δ is $\delta \in (0, 90)$, the system is reset whenever appropriate with the following algorithm.

If $\max[y(k), y(k-1), y(k-2)] \geq 90$ or $\min[y(k), y(k-1), y(k-2)] \leq 0$, set $u(k) = 1$; else set $u(k) = u(k-1) - 0.05$.

The system is reset in this manner to keep the state of the plant in its anticipated operating region. Special care is taken to ensure that approximately 30 input-output data pairs are obtained when the magnitude of the con-

trol signal is within each of 20 sub-ranges. Using the general SVM learning algorithms in Section 4, we can reach the approximation model $\hat{f}^1[Y(k-1), u(k-1)]$ and it is used in the control law (14) and the control law(25). In our experiment, the training dataset for SVM consisted of $20 \times 30 = 600$ samples, Gaussian kernel is selected and the regularization parameter and kernel parameter are determined within Bayesian framework. LS-SVM parameters are $\sigma = 1.15$ and $\gamma = 200$. The robustness filter $F(z)$ is selected as $F(z) = \frac{1-r_1}{1-r_1 z^{-1}}$ with $r_1 = 0.75$.

To validate and verify the performance of the SVM IMC strategy, computer simulations have been carried out on a typical power system with three synchronous generators as in [11], and the structure of this system is shown in Fig. 4, parameters and initial values of Generator 1 and Generator 2 are given as $\omega_0 = 314.159$, $D_1 = 5$, $D_2 = 3$, $H_1 = 8$, $H_2 = 10.2$, $C_{M1} = 0.7$, $C_{M2} = 0.72$, $C_{H1} = 0.3$, $C_{H2} = 0.28$, $P_{m01} = 0.82$, $P_{m02} = 0.8$, $T_{H\Sigma 1} = 0.398$, $T_{H\Sigma 2} = 0.4$. Particular parameters of the transmission lines and other components are the same as in [11], and Generator 3 is taken as reference $E_{q3}^{'} = 1\angle 0^\circ$.

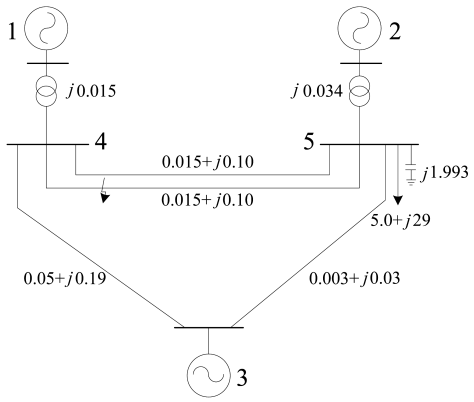


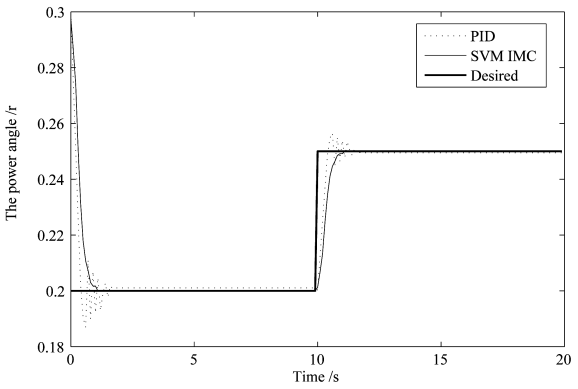
Fig. 4 Typical power system with three generators

Example 1. Control performance of set-point tracking. Here, the SVM IMC strategy is compared with conventional PID controller (denoted by PID). The reference trajectory represents a reduction of δ_d in Generator 1 from 0.3 to 0.2 initially, maintaining the level at 0.2 closely, and an increase of δ_d to 0.25 at $t = 10$ s. Fig. 5 (a) shows the performance of the SVM IMC strategy (solid line) and PID controller (dashed line), Fig. 5 (b) shows the control errors of the SVM IMC strategy and PID controller. As is evident in Fig. 5, the SVM IMC strategy can regulate the desired set-point quickly and smoothly.

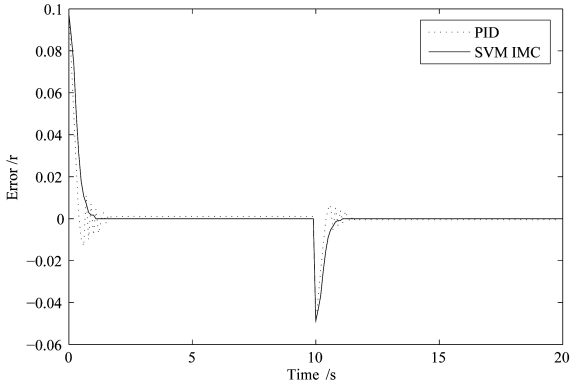
To investigate the robustness of the controller, the system parameters of Generator 1 are supposed to be changed; that is, variable values of $T_{H\Sigma}$ are changed from 0.398 to 0.45. The control performances of SVM IMC strategy and PID controller in this condition are shown in Fig. 6. Fig. 6 (a) shows the performance of the SVM IMC strategy (Solid line) and PID controller (Dashed line). Fig. 6 (b) shows the control errors of the SVM IMC strategy and PID controller. The performance of the PID controller clearly degrades as the system parameter changes. However, the SVM IMC strategy still gives a good robustness. Fig. 6 shows that SVM IMC strategy has better robustness than PID controller when system parameter varies.

Example 2. Control performance at three phase short fault condition. To further show the competitiveness of the SVM IMC strategy, its performance is compared with that of other controllers. Here, comparisons are carried out between three different controllers: conventional PID controller (CON1), decentralized H_∞ controller[11] (CON2), and the SVM IMC strategy. In Figs. 7 and 8, CON1 (dashed line), CON2 (solid line), SVM IMC strategy (bold wide solid line denoted) show the dynamic performances of these three controllers.

Suppose a three phase symmetrical short fault is forced on Bus 4, Figs. 7 and 8 show two kinds of different cases, that is, relative short time fault (Fault starts at 0.1 s and terminates at 0.2 s) and long time fault (Fault starts at 0.1 s and terminates at 0.45 s). It can be observed from Fig. 7 that power system comes back to former operating state faster and better controlled by SVM IMC than controlled by conventional PID controller and decentralized H_∞ controller when fault lasts a relative short time. It can be observed from Fig. 8 that if fault lasts for a long time, conventional PID controller will lead to instability, and decentralized H_∞ controller takes a very long time to stabilize system with numerous oscillations, whereas SVM IMC strategy takes a much short time to stabilize system with few oscillations. Both simulations show that the SVM IMC strategy has good dynamic performance and damping ability.

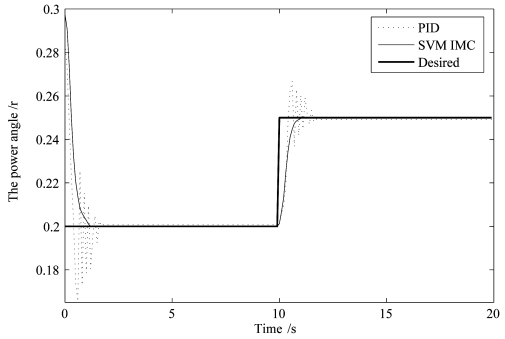


(a) Control performances

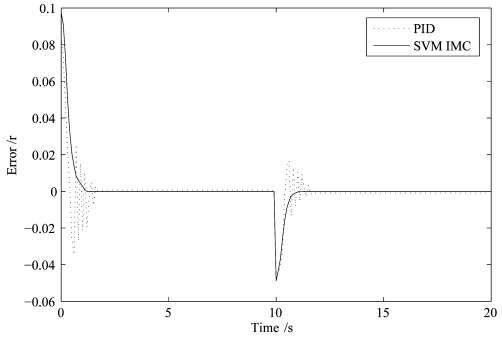


(b) Control errors

Fig. 5 Control performances of set-point tracking

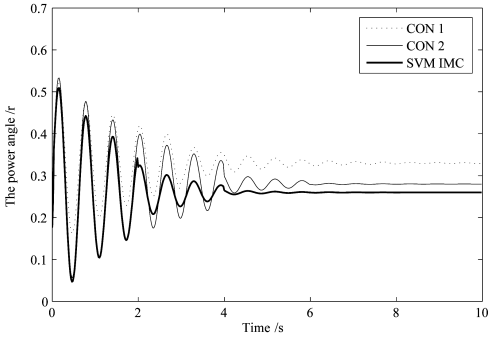


(a) Control performances

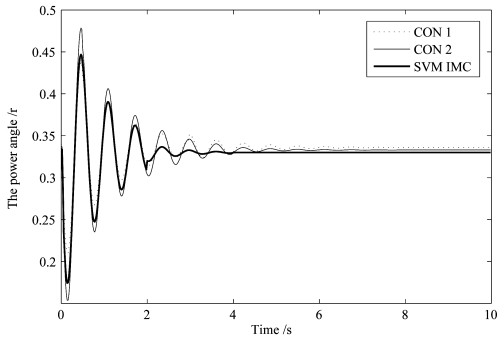


(b) Control errors

Fig. 6 Control performances for set-point tracking with system parameter varying

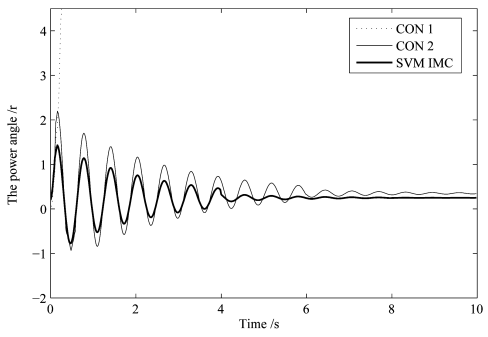


(a) Oscillations of δ_1 for different controllers

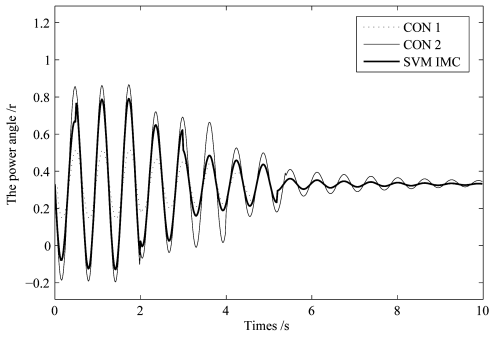


(b) Oscillations of δ_2 for different controllers

Fig. 7 Dynamic responses under relative short time three-phase symmetrical short fault condition



(a) Oscillations of δ_1 for different controllers



(b) Oscillations of δ_2 for different controllers

Fig. 8 Dynamic responses under long time three-phase symmetrical short fault condition

5 Conclusion

This paper has presented a novel SVM IMC strategy as an alternative to IMC methods. In the SVM IMC strategy, only a general identification technique is involved and the amount of computation is quite small because only one SVM network requires to be trained for both model approximation and control formulation without further training. The training of SVM is equivalent to solving a linearly constrained quadratic programming problem, which can lead to a global optimal solution, whereas NN IMC may not converge to global solutions because of the inherent algorithm design of NNs. Simulations show the effectiveness of the SVM IMC strategy for the steam valving control of synchronous generator.

References

1 Morari M, Zafiriou E. *Robust Process Control*. New Jersey: Prentice Hall, 1989

2 Economou C G, Morari M, Palsson B O. Internal model control: extension to nonlinear system. *Industrial and Engineering Chemistry, Process Design and Development*, 1986, **25**(2): 403−411

3 Nahas E P, Henson M A, Seborg D E. Nonlinear internal model control strategy for neural network models. *Computers and Chemical Engineering*, 1992, **16**(12): 1039−1057

4 Rivals I, Personnaz L. Nonlinear internal model control using neural networks: application to processes with delay and design issues. *IEEE Transactions on Neural Network*, 2000, **11**(1): 80−90

5 Han X L, Hua D. An approximate internal model-based neural control for unknown nonlinear discrete processes. *IEEE Transactions on Neural Networks*, 2006, **17**(3): 659−670

6 Vapnik V N. *The Nature of Statistical Learning Theory*. Berlin: Springer-Verlag, 1995

7 Gestel V T, Suykens J A K, Baestaens D E, Lambrechts A, Lanckriet G, Vandaele B. Financial time series prediction using least squares support vector machines within the evidence framework. *IEEE Transactions on Neural Networks*, 2001, **12**(4): 809−821

8 Suykens J A K, Lukas L, Vandewalle J. Sparse approximation using least squares support vector machines. In: Proceedings of IEEE International Symposium on Circuits and Systems. Geneva, Switzerland: IEEE, 2000. 757−760

9 Chan W C, Chan C W, Cheung K C, Harris C J. On the modelling of nonlinear dynamic system using support vector neural networks. *Engineering Applications of Artificial Intelligence*, 2001, **14**(2): 105−113

10 Suykens J A K, Vandewalle J, Moor B D. Optimal control by least squares support vector machines. *Neural Networks*, 2001, **14**(1): 23−35

11 Xi Zai-Rong, Cheng Dai-Zhan. Decentralized steam valving controller for nonlinear multi-machine power systems. *Automation of Electric Power Systems*, 2002, **26**(21): 7−11 (in Chinese)

12 Ge S S, Zhang J, Lee T H. Adaptive MNN control for a class of non-affine NARMAX systems with disturbances. *Systems and Control Letters*, 2004, **53**(1): 1−12

13 Narendra K S, Mukhopadhyay S. Adaptive control using neural networks and approximate models. *IEEE Transactions on Neural Networks*, 1997, **8**(3): 475−485

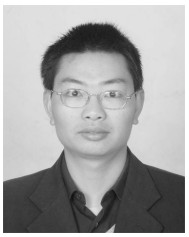
14 Adetona O, Sathananthan S, Keel L H. Robust adaptive control of nonaffine nonlinear plants with small input signal changes. *IEEE Transactions on Neural Networks*, 2004, **15**(2): 408−416

15 Adetona O, Garcia E, Keel L H. A new method for the control of discrete nonlinear dynamic systems using neural networks. *IEEE Transactions on Neural Networks*, 2000, **11**(1): 102−112

16 Levin A U, Narendra K S. Control of nonlinear dynamical systems using neural networks-part II: observability, identification, and control. *IEEE Transactions on Neural Networks*, 1996, **7**(1): 30−42

17 Spooner J T, Maggiore M, Ordonez R, Passino K M. *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques*. New York: Wiley, 2002. 568

WANG Yao-Nan Received his Ph. D. degree from Hunan University. He was a postdoctor at National University of Defense Technology and Alexander von Humboldt Stiftung. Currently, he is a professor in College of Electrical and Information Engineering, Hunan University. His research interest covers intelligent control, intelligent image processing, and intelligent robotics. E-mail: yaonan@hnu.cn



YUAN Xiao-Fang Ph. D. candidate at College of Electrical and Information Engineering, Hunan University. He received his bachelor degree from Hunan University in 2006. His research interest covers intelligent control and neural networks. Corresponding author of this paper.
E-mail: yuanxiaof@21cn.com