TELKOMNIKA, Vol. 11, No. 7, July 2013, pp. 3704 ~ 3710
e-ISSN: 2087-278X

Soft Sensing Based on Hilbert-Huang Transform and Wavelet Support Vector Machine

Qiang Wang*[1,2], Xuemin Tian[1]
[1]College of Information and Control Engineering, China University of Petroleum,
Qingdao, Shandong 266580, China
[2]Industry Engineering Department, Dongying Vocationnl College,
Dongying,Shandong 257091, China
Corresponding author, e-mail: wq213117@163.com, tianxm@upc.edu.cn

Abstract

At present, much more soft sensing have been widely used in industrial process control to improve the quality of product and assure safety in production. A novel method using Hilbert-Huang transform (HHT) combined with wavelet support vector machine (WSVM) is put forward. Firstly the method analyzes the intrinsic mode function (IMF) obtained after the empirical mode decomposition (EMD), then extracts IMF energy feature as the input feature vectors of the wavelet support vector machine. Based on the wavelet analysis and conditions of the support vector kernel function, a novel multi-dimension admissible support vector wavelet kernel function is presented, which is a multidimensional wavelet kernel, thus enhancing the generalization ability of the SVM. The proposed method is used to build soft sensing of diesel oil solidifying point. Compared with other two models, the result shows that HHT-WSVM approach has a better prediction and generalization.

Keywords*: soft-sensing, Hilbert-Huang transform, empirical mode decomposition, wavelet support vector machine*

1. Introduction

There have been many soft sensing methods. Among them, HHT-WSVM approach method has proven to be a powerful method in soft sensing modelling. Soft sensing [1]-[3] has been widely used in industrial process control to improve the quality of product and assure safety in production. It employs easily-measured variables to predicate process variables to be measured, which is hard to measure directly, through computation and estimation models.

The empirical mode decomposition (EMD) [4], [5] is a technique to decompose a given signal into a set of elemental signals called ‘‘intrinsic mode functions’’ (IMFs). The EMD is the base of the so-called ‘‘Hilbert–Huang transform (HHT)’’ [6] that comprises the EMD and the Hilbert spectral analysis that performs a spectral analysis using the Hilbert transform (HT) followed by an instantaneous frequency computation.

Support vector machine (SVM) [7], [8] has been successfully employed to solve regression problem of nonlinearity and small sample. Based on the wavelet analysis and conditions of the support vector kernel function, a novel multi-dimension admissible support vector wavelet kernel function is presented, which is a multidimensional wavelet kernel, thus enhancing the generalization ability of the SVM.

The purpose of this paper is to apply HHT to WSVM [9]-[13] for feature extraction. The wavelet support vector machine modelling which has a multidimensional wavelet kernel, then enhance the generalization ability of the SVM. The original inputs are firstly decomposed IMF using EMD, then extracts IMF energy feature as the input feature vectors. These new features are then used as the inputs of HHT-WSVM to build soft sensing of diesel oil solidifying point. The simulation test shows that the method is is effective and correct.

2. Hilbert-Huang Transform Theory

2.1. Empirical Mode Decomposition

The theory of Hilbert-Huang Transform presented the concept of the intrinsic mode and

Received January 28, 2013; Revised April 11, 2013; Accepted April 20, 2013

introduced of the method of empirical sifting method. It is very applicable to analyze nonlinear and non-stationary signal which frequency is variable with time. The EMD is a highly adaptive decomposition. It decomposes any complicated signal into so called Intrinsic Mode Functions. Huang have defined IMFs as a class of functions that satisfy two conditions[14]: One is that in the whole data set, the number of extrema and the number of zero-crossings must be either equal or differ at most by one; Another condition of the IMF is that at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero, which means the envelopes defined by the local maxima and minima, respectively, is locally symmetric around the envelope mean. But most of the signals are not IMFs, at any given time, the signal may have more than one oscillatory mode, that is why the signal should be decomposed into IMFs by the sifting process of EMD. While the decomposition is based on three assumptions:

(1) the signal has at least two extrema: one maximum and one minimum.
(2) the characteristic time scale is defined by the time lapse between the extrema.
(3) if the data were totally devoid of extrema but contained only inflection points, then it can be dierentiated once or more times to reveal the extrema.

Final results can be obtained by integration of the components For a given signal $s(t)$, the IMFs, constitutive components of the signal $s(t)$,can iteratively extracted as follows:

(1) Find all the points of local maxima and all the points of local minima in the signal.
(2) Create the upper envelope $v_1(t)$ by spline interpolation of the local maxima and the lower envelope $v_2(t)$ by spline interpolation of the local minima of the input signal.
(3) Calculate the mean of the upper envelope and the lower envelope.

$$m = \frac{1}{2}[v_1(t) + v_2(t)] \tag{1}$$

(4) Subtract the envelope mean signal from the input signal to yield the residual.

$$s(t) - m = h \tag{2}$$

(5) Iterate on the residual h until it satisfy the stop criterion. The stop criterion is to check if the residual from step 4 is an IMF or not. Then it is defined as c_1, the first IMF component from the data:

$$c_1 = h \tag{3}$$

After the IMF is found, define the residue as the result from subtracting this IMF from the input signal.

$$s(t) - c_1 = r \tag{4}$$

(6) Repeat the sifting process from step 1 to step 5 many times with the residue as the input signal so that all the IMFs can be extracted from the signal $s(t)$. We can obtain $c_2, c_3, \ldots$ in turn. The sifting process can be stopped by any of the following predetermined criteria: either when the component or the residue becomes so small that it is less than the predetermined value of substantial consequence, or when the residue, becomes a monotonic function from which no more IMF can be extracted. So when the EMD is finished, the original input signal $s(t)$ can be expressed as the following formula :

$$s(t) = \sum_{i=1}^{n} c_i + r \tag{5}$$

That is to say the signal $s(t)$ may decompose IMF which are $c_1, c_2, \ldots, c_n$ and a residual which is r . Where the number of IMF is n .

2.2. The Hilbert Spectrum

The Hilbert transform is often applied to the analysis of the linear and nonlinear system. For an arbitrary time series $X(t)$, the Hilbert Transform, $Y(t)$ is defined as

$$Y(t)=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{X(\tau)}{t-\tau}\mathrm{d}\tau \tag{6}$$

With this definition, $X(t)$ and $Y(t)$ form the complex conjugate pair, so we can have an analytic signal, $Z(t)$ as

$$Z(t)=X(t)+\mathrm{i}Y(t)=a(t)\mathrm{e}^{\mathrm{i}\theta(t)} \tag{7}$$

in which $a(t)=[X(t)^2+Y(t)^2]^{1/2}$ (8)

$$\theta(t)=\arctan Y(t)/X(t) \tag{9}$$

The restrictive conditions of IMF guarantee that the meaningful instantaneous frequency can be provided by this definition. After performing the Hilbert transform on each IMF component, the data $X(t)$ can be expressed in the following form:

$$H(w,t)=\mathrm{Re}\sum_{i=1}^{n}a_i(t)\mathrm{e}^{\mathrm{j}\int w_i(t)\mathrm{d}t} \tag{10}$$

Here the residue r is missed on purpose, for it is either a monotonic function, or a constant. This frequency-time distribution of the amplitude is designated as the Hilbert amplitude spectrum or simply Hilbert spectrum. In order to build soft sensing by EMD, each mode energy which has already decomposed is defined as follow:

$$E_i=\sum_{i=1}^{n}\left|a_i(t)\right|^2 \tag{11}$$

3. Wavelet Support Vector Machine

3.1. Support Vector Machine Regression

Support Vector Machine (SVM) method was proposed by Vapnik in 1995 [7]. It is powerful for the problem with small sample, nonlinear, high dimension, and local minima. The basic idea of the SVM regression is to map the input data into a feature space via a nonlinear map. In the feature space, a linear decision function is constructed. Set the training sample $\{x_i,y_i\}_{i=1}^{n}, x_i\in R$ as the input vector , and set $y_i\in R$ as the corresponding export value. The number of training sample is *n*. In SVM, the regression function is approximated by the following function:

$$f(x)=\omega\cdot\phi(x)+b \tag{12}$$

where ω denotes the weight vector, $\phi(x)$ denotes the high-dimensional feature space and b denotes the bias term. The following formulae can be obtained by minimizing the risk function to coefficient w and b:

$$\min_{w,b,\xi}\frac{1}{2}\|w\|^2+C\sum_{i=1}^{n}(\xi_i+\xi_i^*)$$
$$s.t.\begin{cases}y_i-w\cdot\phi(x_i)-b\le\varepsilon+\xi_i\\ w\cdot\phi(x_i)+b-y_i\le\varepsilon+\xi_i^*\\ \xi_i\ge 0\\ \xi_i^*\ge 0\\ i=1,2,\dots,n\end{cases} \tag{13}$$

where ξ_i and ξ_i^* are slack variables and ε is the accuracy demanded for the approxiation.

This constrained optimization problem is solved using the following Lagrangian form:

$$\max_{\alpha,\alpha^*} \left\{ \begin{array}{l} L_D = -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_i^*)K(x_i \cdot x_j) \\ -\varepsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i) \end{array} \right\} \quad (14)$$

$$s.t. \quad \begin{cases} \sum_{i=1}^{n}\alpha_i = \sum_{i=1}^{n}\alpha_i^* \\ \alpha_i, \alpha_i^* \in [0, C] \\ i = 1, 2, \cdots, n \end{cases}$$

Finally , the corresponding regression function Eqn. (12) can be directly expressed as follows :

$$f(x) = \omega \cdot \phi(x) + b = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i \cdot x_j) + b \quad (15)$$

$K(x_i \cdot x_j)$ is t he nuclear function , and $K(x_i \cdot x_j) = \phi(x_i) \cdot \phi(x_j)$.

3.2. Wavelet Support Vector Machine (WSVM) Algorithm

The procedure of **Wavelet support vector machine (WSVM)** can be shown from Figure 1 and described as:

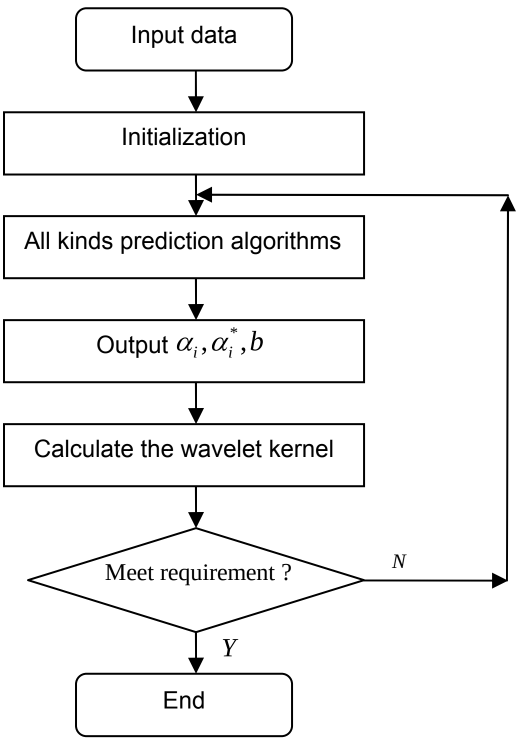

Figure 1. Framework of Wavelet support vector machine

Step 1: Inputing data after preprocessing.

Step 2: Initializing the parameters of WSVM, which are Lagrangian multiplier α_i α_i^* threshold b and so on.

Step 3: Calculating the objective function through training samples. Obtaining α_i α_i^* b through improved SMO algorithm.

Step 4: Calculating the wavelet kernel through the followed formula.

$$f(x)=\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)\prod_{i=1}^{d}\Psi[\frac{x^j-x_i^j}{a_j}]+b \quad (16)$$

Step 5: Calculating the error function until stop conditions are satisfied, otherwise returns step 3.

4. Simulation and Analysis

4.1. Processing Procedure

Figure 2 is the flow chart of soft sensing using HHT and WSVM.

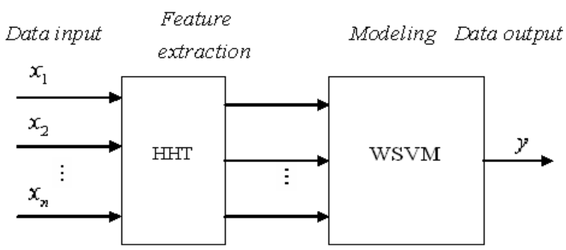

Figure 2. The framework of HHT-WSVM

4.2. Data Selection

The data diesel oil solidifying point was obtained at Soutwest Research Institute (SWRI) on a project sponsored by the U.S. Army. And we can download from [15]. In this article we select freezegatest data as research objective. It includes 401 measured varibles and 250 samples.

The outliers had been removed by 3σ edit rule and three five-point smoothing method [16] beforhead. These 201 samples of pretreated were randomly divided into training data and test data. In which 101 of these samples were used to model training, and the remaining 100 were used as a test data set.

4.3. The Extract Feature of IMF Energy

In this paper, random signal of freezegatest data was decomposed by HHT. In which, EMD not only come from the feature of signal but also overcome scheduling base function which wavelet often select before decomposing. So EMD possessed favorable local adaptability. From above analysis, we can calculate each mode energy according formula (11). At last the each energy can be normalized by the maximal energy.

4.4. Result and Discussion

Root mean squared error (RMSE) , maximal absolute error (MaxAE) and mean absolute error (MeanAE) are used to evaluate the generalization performance of these algorithms, which are as follows:

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \quad (17)$$

$$MaxAE=\max|y_i-\hat{y}_i| \quad (18)$$

$$MeanAE = \frac{1}{n}\left(\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|\right) \tag{19}$$

where y_i and $\hat{y}_i$ are respectively actual value and prediction value.

The Morlet wavelet is selected to construct the SVM kernel function and form the WSVM. Which is as follow:

$$K(x,x') = \prod_{i=1}^{d}\cos[\frac{1.75(x-x')}{a_i}]\exp(-\frac{\left\|x-x'\right\|^2}{2a_i^2}) \tag{20}$$

In order to evaluate the improved algorithm, in this section three typical regression datasets are selected to test it. which are as follows: KPCA-LSSVM method, HHT-SVM method and HHT-WSVM method. The predication error of three soft sensing methods are at below table 1. It can be seen clearly that HHT-WSVM method has better value on RMSE and MaxAE.

Table 1. The Predication Error of Three Soft Sensing Methods

	KPCA-SSVM	HHT-SVM	HHT-WSVM
RMSE	0.6137	0.5037	0.0985
MaxAE	2.0696	4.0616	0.1096
MeanAE	0.8302	0.2537	0.0097

The predication results of three soft sensing methods are at below figures. Compared with Figure 3 and Figure 4, Figure 5 shows that the estimated outputs of soft sensor match the real values and follow the varying trend very well.

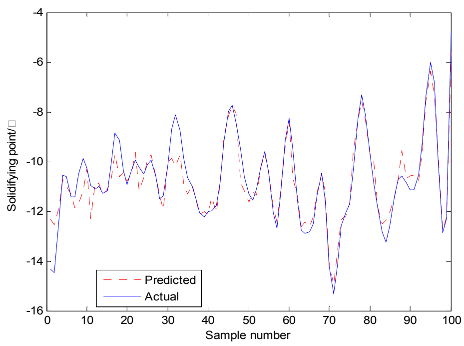

Figure 3. Prediction results based on KPCA-LSSVM

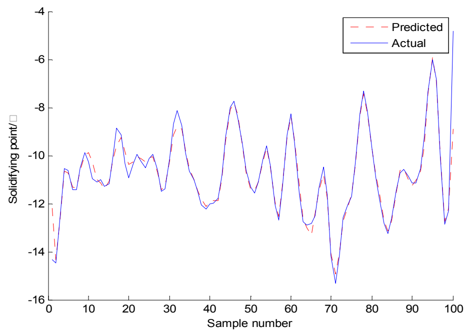

Figure 4. Prediction results based on HHT-SVM

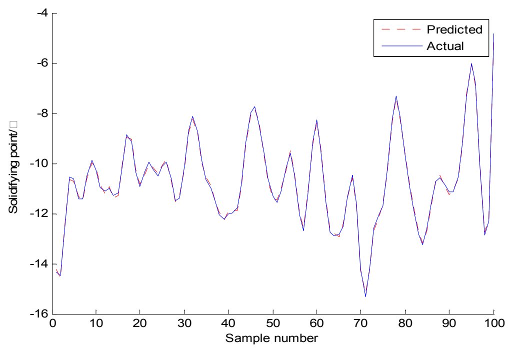

Figure 5. Prediction results based on HHT-WSVM

5. Conclusion

In this article a kind of soft sensing is proposed by combining Hilbert-Huang transform (HHT) combined with wavelet support vector machine (WSVM) is put forward.Firstly the method analyzes the intrinsic mode function (IMF) obtained after the empirical mode decomposition (EMD), then extracts IMF energy feature as the input feature vectors of the wavelet support vector machine. Based on the wavelet analysis and conditions of the support vector kernel function, a novel multi-dimension admissible support vector wavelet kernel function is presented, which is a multidimensional wavelet kernel, thus enhancing the generalization ability of the SVM. The proposed method is used to build soft sensing of diesel oil solidifying point. Compared with other two models, the result shows that HHT-WSVM approach has a better prediction and generalization.

Acknowledgements

This work was financially supported by National Natural Science Foundation of China (51104175) and Natural Science Foundation of Shandong Province of China (ZR2011FM014).

References

[1] WW Yan, HH Shao, and XF Wang. "Soft sensing modeling based on support vector machine and Bayesian model selection". *Computers and Chemical Engineering*. 2004; 28: 1489-1498.
[2] M Shakil, M Elshafei, MA Habib, et al. "Soft sensor for NOx and O2 using dynamic neural networks". *Computers and Electrical Engineering*. 2009; 35: 578-586.
[3] JS Yu, AL Liu, KJ Huang, Soft Sensor Technology in Petrochemical Industry Application, Beijing: Petrochemical Press. 2000.
[4] NE Huang, Z Shen, SR Long, ML Wu, HH Shih, Q Zheng, NC Yen, CC. Tung, and HH Liu, *"The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis"*. in Proc. Roy. Soc. London A. 1998; 454: 903-995.
[5] NE Huang*. "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis*". Proc. Roy. Soc. London A. 2003; 459: 2317-2345.
[6] NE Huang. "*HHT basics and applications-for speech, machine health monitoring, and bio-medical data Analysis*". March 24, 2003.
[7] VN Vapnik. The Nature of statistical Learning Theory. New York: Springer-Verlag. 1995.
[8] M Behzad, KA Eazi, M, Palhang. "Generalization performance of support vector machines and neural networks in runoff modeling". *Expert Systems with Applications*. 2009; 36: 7624-7629.
[9] MH Yang, RC Wang. "DDoS detection based on wavelet kernel support vector machine". *The Journal of China Universities of Posts and Telecommunications*. 2008; 15: 59-63.
[10] XCZhu. "Application of software aging prediction based on wavelet analysis and support vector machines". *Computer Simulation*. 2012; 29: 265-269.
[11] X Ju, CB Liu, F Pan. "Soft sensor study and application based on WSVM". *Journal of Mechanical& Electrical Eng ineering*. 2011; 29: 476-478.
[12] CL GuO, LJ Yu SL Zeng. "Research on urban traffic flow prediction based on LS-WSVM". *Logistics Engineering and Managemen,* 2012; 34: 95-100.
[13] C Chen, LC Cai, CL Liu. "Chaotic time series prediction algorithm base on wavelet support vector machine". *Computer Simulation*. 2012; 29: 223-226,.
[14] NE Huang, Z Shen, SR Long. A new view of nonlinear water waves: *The Hilbert spectrum, Annual Review of Fluid Mechanics*, London, 1999; 31: 417-4579.
[15] Data available at http://software.eigenvector.com/Data/SWRI/index.html
[16] J Wang, X Hu, Application of Mattlab in Vibration Signal Processing. Beijing: China Water Power Press. 4006; 78-82.