

# A Novel Soft Sensing Based on Wavelet Packet Decomposition

Wang Qiang <sup>1, a</sup>

<sup>1</sup> Dongying Vocational College, Dongying, Shandong 257091, China

<sup>a</sup>122381275@qq.com

**Keywords:** Flow regime identification; Wavelet packet decomposition; support vector machine; Soft sensing.

**Abstract.** In this paper, a new soft sensing method based on wavelet packet decomposition and SVM was put forward. As is known the characteristic of pressure drop is nonlinear and non-stationary. Based on the characteristics that the wavelet packet transform can decompose signals to different frequency bands according to any time frequency resolution ratio, the concept and the algorithm of the wavelet packet energy features are proposed. At the same time, the features are extracted from the differential pressure fluctuation signals of the air-water two-phase flow in the horizontal pipe and the wavelet packet energy features of various flow regimes are obtained. The support vector machine was trained using these eigenvectors as flow regime samples, and the flow regime intelligent identification was realized. The test results show the wavelet packet energy features can excellently reflect the difference between four typical flow regimes, and successful training the support vector machine can quickly and accurately identify four typical flow regimes. So a new way to identify flow regime by soft sensing is proposed.

## Introduction

Now there are two main methods of gas-liquid two-phase flow regime identification. One is visual observation. The other is estimation using the spot flow parameters combining with the flow regime transition rules or the flow regimes charts. Both methods above are inevitable to be affected by subjective factors and difficult to realize objective identification. In this paper, a new soft sensing method based on wavelet packet decomposition and SVM was put forward.

In recent years, many researchers have applied wavelet packet decomposition to signal processing[1].

Wavelet packet decomposition is an available time-frequency method, which can realize the random frequency band without superposition. By wavelet packet decomposition and reconstruction we can separate signals to different frequency bands according to any time frequency resolution ratio, then we can extract wavelet packet energy features. On the other hand, traditional analysis methods such as the Fourier transform[2], time-frequency analysis [3] are not very suitable for feature extraction due to the non-stationary property of the flow regime signals.

The support vector machine (SVM) is a machine-learning algorithm based on the statistical learning theory, which has desirable classification ability even if with fewer samples. SVM provides us with a new method to develop the intelligent flow regimes identification. A novel soft sensing method of flow regime identification based on support vector machine and wavelet packet decomposition is proposed in this paper. The results also prove the soft sensing method is efficient and feasible.

## Wavelet packet algorithm and the feature extraction

### Wavelet packet algorithm

In the analysis of multi-resolution the time frequency windows of different scales have different figures. When the scale is relative small the frequency resolution ratio is bad. When the scale is quite large the time resolution ratio is bad too. This is not applicable to some signals. The wavelet packet method is a more subtle orthogonal decomposition method based on multi-resolution ratio. It can select the frequency band self-adaptively according to the characteristics of the analyzed signals and ascertain the resolution ratio of the signals in different frequency bands[4].

In the analysis of multi-resolution ratio, the scale function  $j(t)$  and the wavelet function  $y(t)$  satisfied the double-scale equation:

$$j(t) = \sqrt{2} \sum_k h(k)j(2t-k) \quad (1)$$

$$y(t) = \sqrt{2} \sum_k g(k)y(2t-k) \quad (2)$$

For the convenience of discussion we mark:  $m_0(t) = j(t)$ ,  $m_1(t) = y(t)$ , And the double-scale equation was rewritten as the following recursion form:

$$m_{2n}(t) = \sqrt{2} \sum_k h(k)m_n(2t-k) \quad (3)$$

$$m_{2n+1}(t) = \sqrt{2} \sum_k g(k)m_n(2t-k) \quad (4)$$

The defined function set  $\{m_n(t)\}_{n \in Z}$  was called the orthogonal wavelet packet of the orthogonal scale function  $j(t)$ . The wavelet packet space is formed by the dilation translation set  $j(t)$ . The operator form of the multi resolution ratio analysis decomposition formula:

$$H[S_k](j) = \sum_k S_k h_{k-2j} \quad (5)$$

$$G[S_k](j) = \sum_k S_k g_{k-2j} \quad (6)$$

Supposing the signal  $f(t) \in U_j^n$ , that is  $f(t) = \sum_k S_k^j m_n(2^{-j}t-k)$  then we get :

$$f(t) = 2^{-\frac{1}{2}} \sum_j H[S_k^j](i)m_{2n}(2^{-j-1}t-i) + 2^{-\frac{1}{2}} \sum_j G[S_k^j](i)m_{2n+1}(2^{-j-1}t-i) \quad (7)$$

This is the formula of wavelet packet decomposition. It illuminates that any signal of wavelet packet decomposition can be decomposed two parts: One is the effect of signal H, that is, the projection on  $\{m_{2n}(2^{-j-1}t-i)\}_{i \in Z}$ ; the other is the effect of signal G, that is, the projection on  $\{m_{2n+1}(2^{-j-1}t-i)\}_{i \in Z}$ .

The square of the signal's 2 norm is equal to the energy of original signal at the time region. So it is reliable that we use wavelet packet energy spectrum to denote the energy of the original signal. The signal of energy may select the sum of the square of various frequency bands. Supposing the corresponding energy of someone frequency bands is  $E_j$  then we get:

$$E_j = \int |S_j(t)|^2 dt = \sum_{k=1}^N |x_{jk}|^2 \quad (8)$$

Where,  $x_{jk}$  ( $j = 0,1,\dots,15$ ),  $k = (0,1,\dots,n)$  represent the discrete points' amplitude of reconstruction signal  $S_j$  and  $N$  is the sample length of the  $j$  subspace.

### The feature extraction

The experiment introduction and the flow regime signals can be found in literature. In the experiment, differential pressure fluctuation signals of four typical flow regimes can be showed in Fig.1.

In this paper wavelet packet energy is selected as feature extraction. Above all, the collected pressure-difference fluctuation signals are decomposed by the four-layer wavelet packet, and the decomposed signals in various frequency bands are obtained within the frequency domain. Then restructure the coefficient of the wavelet packet decomposition, and extract the signals of various frequency bands range. Then the wavelet packet energy eigenvectors of flow regimes are established. At last the wavelet packet energy can be normalized by the maximal energy various frequency bands range. Fig.2 depicted the typical bubbly flow Primitive signal and Reconstructed signal.

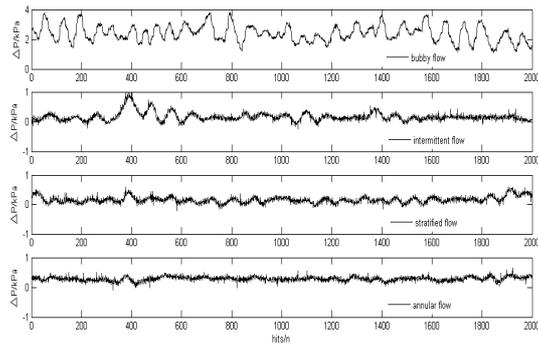


Fig.1 Differential pressure signals

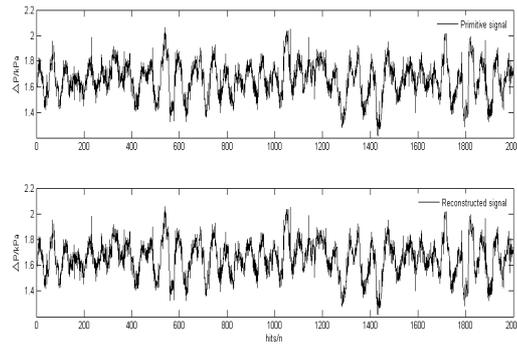


Fig.2 Primitive signal and Reconstructed signal

## Flow regime identification by SVM

### The support vector machine

The support vector machine (SVM) is a machine-learning algorithm based on the statistical learning theory, which has desirable classification ability even if with fewer samples. The basic idea of applying SVM to pattern classification is as follows: first to map the input space into a high dimension feature space via a nonlinear map realized by appropriate kernel functions, and then to seek the optimal linear division in the new space.

For a simple binary classification problem, given a set of training vector belonging to two separate classes:  $(x_1, y_1) (x_2, y_2) \dots (x_N, y_N)$ ,  $x_i \in \mathfrak{R}^n$ ,  $y_i \in \{-1, +1\}$ . If the two classes are linearly separable, the hyperplane that does the separation is:

$$w \cdot x + b = 0 \quad (9)$$

This linear classification is termed as the Optimal Separating Hyperplane(OSH). Since the distance to the closest point is  $1/\|w\|$ , finding the OSH amounts to minimize  $\|w\|$  and the objective function is:

$$\text{Minimize: } f(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w \cdot w), \quad (10)$$

$$\text{Subject to: } y_i(w \cdot x_i + b) \geq 1, i = 1, \dots, N$$

The classification function can be written as:

$$f(x) = \text{sgn}\left(\sum_{i=1}^N a_i y_i x_i \cdot x + b\right) \quad (11)$$

Where  $b$  is the bias and the coefficients  $a_i$  are the Lagrange multipliers associated with constraint:

$$a_i \geq 0, \sum_{i=1}^N a_i y_i = 0 \quad (12)$$

In the presence of noise, SVM introduced slack variables  $x_i \geq 0$ , and set an upper bound on the size of the  $a_i$ . Therefore, the Lagrange multipliers are modified to  $0 \leq a_i \leq C$ ,  $i = 1, \dots, N$ , Where  $C$  is a penalty factor to penalize training errors.

If the two classes are nonlinearly separable, the input vector should be nonlinearly mapped to a high dimensional feature space by an inner-product kernel  $k(x_i, x_j) = f(x_i) \cdot f(x_j)$ . We choose the Gaussian RBF in this experiment, which has the form:

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2s^2}\right) \quad (13)$$

where  $s$  is the width of the Gaussian function.

### Flow regime identification

The diagram of flow regime identification is shown in Fig.3.

In literature[5] which Hsu gave a comparison and pointed out that the “one-against-one” method is more suitable for practical use than other methods. In this experiment “one-against-one” method is selected to classify the four typical flow regimes of gas-liquid flow in the horizontal tube.

In this paper the 120 samples are taken as the training samples, including 30 samples of each type. After finishing the training of the network, we select 80 groups of the data the test samples and

simulate the learning effect of the network, including 20 samples of each type. The result shows: the SVM identified 76 samples correctly. The whole identification ratio reaches 95%. There are sample data and wavelet packet energy of some test samples in table 1.

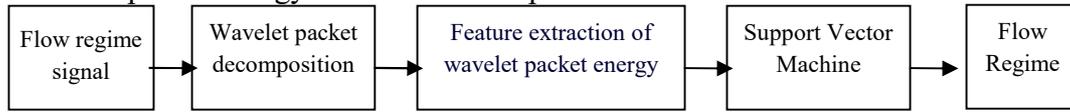


Fig.3 The diagram of flow regime identification

Tab.1 Test sample data and wavelet packet energy of four flow regimes

Wavelet packet energy feature														Flow		
1~16	16~32	32~48	48~64	64~80	80~96	96~112	112~128	128~144	144~160	160~176	176~192	192~208	208~224	224~240	240~256	regime
0.76712	0.74112	0.5697	0.51961	0.49753	0.72123	0.60051	0.60238	0.86314	0.55129	0.68723	0.90609	0.63464	0.52039	0.96732	0.51278	B
0.27131	0.34748	0.45773	0.47075	0.56829	0.87707	0.53698	0.69385	0.73919	0.68837	0.41214	0.73596	0.59717	0.95521	0.66413	0.79555	I
0.11399	0.21577	0.25513	0.20972	0.43544	0.64623	0.29125	0.45784	0.50311	0.38827	0.92579	0.87427	0.61928	0.44731	0.92484	0.39359	S
0.10501	0.091	0.17651	0.08809	0.94709	0.65053	0.42043	0.41539	0.41803	0.92955	0.35951	0.74261	0.97979	0.36729	0.60233	0.38783	A

Annotation: S-stratified flow; I-intermittent flow; B-bubbly flow; A-annular flow

**Conclusions**

- (1) The wavelet packet decomposition is an effective method of differential pressure fluctuation feature extraction. It can extract the energy of the differential various frequency bands and construct the eigenvectors of various flow regimes.
- (2) The flow regime identification method that combining wavelet packet decomposition with SVM is reliable.

**Acknowledgements**

The author are grateful to the support of Project of Shandong Province (J16LC61、2016—xd—039、17SDJ003、201606、J13LD60、17SVE007、2017539).

**References**

[1] Shen M , Sun C F. Method for Extracting Time-Varying Rhythms of Electroencephalogy via Wavelet Packet Analysis, IEE Proceedings-Science : Measurement and Technology , Singapore , 2001

[2] Brett D M , Justin W , Seung H S. BCI Competition 2003-Data Set Ia : Combining Gamma-Band Power with Slow Cortical Potentials to Improve Single-Trial Classification of Electroencephalographic Signals, IEEE Transactions on Biomedical Engineering , Vol.51, No 6, 2004

[3] Schroeder M , Bogdan M , Rosenstiel W , et al . Automated EEG Feature Selection f or Brain Computer Interfces, I n Proc. 1st I nt . I EEE EMBS Conf . Neural Eng, Italy ,2003

[4] Wang Q, Zhou Y L, Cui Y F, etal. Applied Study of EMD and Neural Networks on Flow Regime Identification for Gas-Liquid Two-Phase Flow, Journal of Engineering Thermophysics, Vol.28, No.3, 2007

[5] Hsu, C W, Lin C J. A Comparison of Methods for Multi-Class Support Vector Machines, IEEE Trans on Neural Networks, Vol.13, No.3, 2004