Journal of Theoretical and Applied Information Technology
15 September 2012. Vol. 43 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FEATURE SELECTION OF FREQUENCY SPECTRUM FOR THE BALL MILL LOAD BASED ON INTERVAL PARTIAL LEAST SQUARES

[1, 2] LIJIE ZHAO, [1] XUE FENG, [1] DECHENG YUAN, [1] HUI XIAO
[1] College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110042
[2] State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819

ABSTRACT

Due to highly complex of the grinding mechanism of the ball mill, it is a challenging problem to select the informative frequency spectral features of the response signals. High dimensionality and colinearity of the frequency spectrum are unfavorable to build the effective mill load model in the wet ball mill. Interval Partial Least-Squares Regression (iPLS) is applied to select the feature frequency bands of the shell vibration signal and acoustical signal, which are closely relevant to the parameters of ball mill load. Redundant or irrelevant frequency spectral variables are removed to improve the complexity of ball mill load model and enhance the comprehension of the grinding process using the frequency spectrum features. The experimental results have demonstrated that the performance of the mill load models based on feature spectrum outperforms the full spectrum model for both the shell vibration signal and acoustical signal.

Keywords: *Ball Mill, Mill Load, Feature Selection, Interval Partial Least Square*

1. INTRODUCTION

Parameters and statuses of the ball mill load are very important equipment information and key to control the grinding process [1]. They are closely related to the production rate, product quality, energy consumption in the grinding process[2]. In recent years, with the development of sensors, data processing and communication technology, the analysis and monitoring of the ball mill load based on shell vibration signals with high sensitivity and strong anti-interference have increasingly attracted much focus from both academic and industrial fields [3]. However, comminuting mechanism of ball mill is very complex. Distribution of mineral particles, hardness of ore, slurry viscosity, the number and size of steel ball and other factors directly influence internal impact and grind in the ball mill. It is very difficult to extract the features of the shell vibration signals in the time domain, which are caused by superposition of a series of the impact force and frictions with different intensity and frequencies ranges [4]. Although the vibration and acoustic frequency spectrum contain plenty of information about the mill load, the modeling via full spectrum is difficult to build effectively. This is because the irrelevant and redundant spectral variables maybe cover up the real operation mode and deteriorate the quality of the model due to the hyper-high dimension and high colinearity in the frequency spectrum variables. Therefore, it is very necessary to select the feature spectral bands which are directly relevant to the parameters of the ball mill load.

Parameters of mill load are reflected by different frequency bands of the vibration signals. A sample clustering and kernel principal component analysis was used to select the feature frequency bands and extract the nonlinear features [5, 6]. Although PCA/KPCA can retain the spectrum information as much as possible, the principal components extracted by the PCA/KPCA mainly reflects changes of itself feature-spectrum without considering the influence to parameters of ball mill load, which might result in loss of useful information and instability of model performance. The selected feature frequency bands directly relevant to the parameters of ball mill load by using genetic algorithm- interval partial least square were the suboptimal solutions, due to the random initialization and multi-run of genetic algorithm.

Nogaard et al. [7] proposed a wavelength selection process called interval partial least squares (iPLS) regression. The iPLS algorithm splits the full spectrum into many sub-intervals of equal width, where each sub-interval builds a local PLS regression model. The best regression model based on sub-intervals should require the smallest number of PLS components and produce the lowest

Journal of Theoretical and Applied Information Technology
15 September 2012. Vol. 43 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RMSECV values. iPLS and its extension methods, such as Forward Interval Partial Least Squares (FiPLS), Backward Interval Partial Least Squares (BiPLS) [8], Synergy Interval Partial Least Squares[9] and Moving Window Partial Least-Squares[10], are widely used in the spectral engineering field to analyze various of the spectrum information[11].

Due to high dimensionality and colinearity of the frequency spectrum difficult to build the effective mill load model, a frequency feature selection method of the shell vibration signal relevant to the parameters of ball mill load is proposed based on *i*PLS. Redundant or irrelevant frequency spectral variables are removed to improve the predictive power of ball mill load model and enhance the comprehension of grinding process.

2. FEATURE SELECTION METHOD OF FREQUENCY SPECTRUM

According to the mechanism of the vibration signal and acoustical signal, a frequency spectral feature of the ball mill outer response method based on iPLS is proposed as depicted in Figure 1.

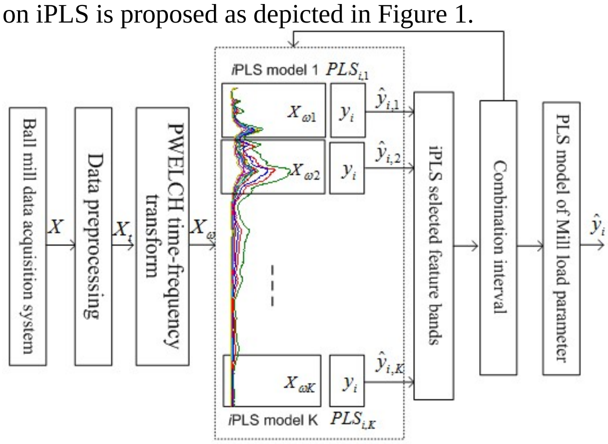


Figure 1: The feature selection strategy of frequency spectrum based on *i*PLS

Feature selection of the frequency spectrum of the shell vibration signals from the ball mill as follows:

(1) First, remove the outliers and noise from the original vibration or acoustical signals X_t. The time domain waveform of shell vibration or acoustical signal are transformed into the frequency domain power spectrum by PWELCH [7], and the power spectrum X_ω are averaged by several rotation periods of frequency spectrum.

(2) Full-spectrum X_ω are split into K subintervals of equal width, $X_{\omega k}, k=1,\cdots,K$.

(3) The local partial least squares regression model $PLS_{i,k}$ of mill load parameter y_i ($i=1,2,3$), such as mineral to ball volume ratio, pulp density and charge volume ratio, is built on each interval $X_{\omega k}$ by using the nonlinear iterative partial least squares algorithm. The $PLS_{i,k}$ model is expressed as follows:

$$\{X_{\omega k}, Y_i\} \xrightarrow{PLS} \{W_{i,k}, P_{i,k}, B_{i,k}, Q_{i,k}\}. \tag{1}$$

where $P_{i,k}=\left[p_{i,k,1},\cdots,p_{i,k,h}\right]\in R^{n\times h}$ and $Q_{i,k}=\left[q_{i,k,1},\cdots,q_{i,k,h}\right]\in R^{n\times h}$ are loading matrix of spectrum interval and output parameters of the ball mill load, respectively. $W_{i,k}$ is the weight matrix and $B_{i,k}$ is the diagonal coefficient matrix of inner model, $B_{i,k}=\left(t_{i,k}^T t_{i,k}\right)^{-1} t_{i,k}^T u_{i,k}$. The output $\hat{y}_{i,k}$ of iPLS mode developed on the *i*-th mill load parameter and the *k*-th spectrum interval is described in Eq. (2):

$$\hat{y}_{i,k} = X_{\omega,k} W_{i,k}\left(P_{i,k}^T W_{i,k}\right)^{-1} B_{i,k} Q_{i,k}^T \tag{2}$$

(4) Prediction performance of the local models developed on spectral subintervals of equal width are compared based on the validation parameter RMSECV (Root Mean Squared Error of Cross Validation) and the other parameter such as r (squared correlation coefficient). RMSE is defined as follows:

$$RMSE_i = \sqrt{\frac{\sum\left(\hat{y}_{i,k}-y_i\right)^2}{N}}, i=1,2,3. \tag{3}$$

where N is the number of samples, y_i is the laboratory measured value and $\hat{y}_{i,k}$ is the predicted value. RMSEC is RMSE calculated from the calibration samples, RMSECV is calculated from the cross-validated samples, and RMSEP is calculated from the independent test set. Correspondingly, the correlation coefficients for these three situations are calculated. The lowest RMSECV of all the local models is the first chosen spectral region.

(5) Develop PLS models for all possible combinations of the first chosen spectral region and the rest of the sub-intervals one by one. RMSECV is calculated for each combination region. The combination of two intervals with the lowest RMSECV is selected as the second spectral feature.

(6) This procedure isn't stopped until the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RMSECV increases with the number of interval combination.

3. EXPERIMENT RESULTS AND DISCUSS

3.1 Descriptions Of Experimental Ball Mill

The experiments were performed on a laboratory scale lattice-type ball mill (XMQL-420×450) with the drum of 460 mm in diameter and 460 mm in length, as shown in Fig.1. The vibration signals of the ball mill were measured by a vibration acceleration sensor with sampling frequency 51,200Hz. The acoustic sensor was installed in the distance 1/3 from the inlet of ball mill. The experimental ball mill has maximum ball load of 80 kg, pulverizing capacity of 10 kg per hour and a rated revolution of 57 per minute. The grinding experiments are done by adding the steel balls of different size(diameter of 30, 20 and 15mm), copper ore and water into the ball mill, which have been homogenized, and last for one minute.

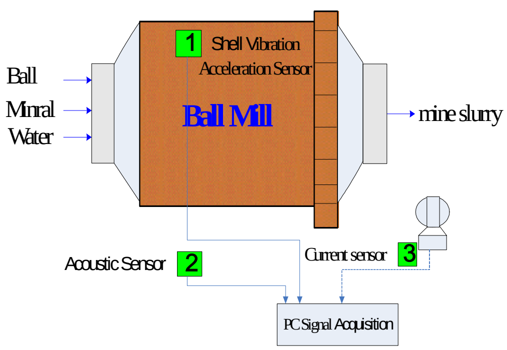


Figure 2: Signal acquisition from the experimental ball mill

3.2 Feature Selection Of Vibration Frequency Spectrum

In the application of the *i*PLS algorithm, the number of intervals, the number of interval combination, the number of the local PLS models built on each interval are determined by the lowest RMSECV. Table 1 shows the results of feature selection of vibration frequency spectrum which are closely relevant to the parameters of mill load, such as mineral to ball volume ratio, charge volume ratio and pulp density. RMSECV is calculated for each combination of different intervals and the number of intervals. Take the mineral to ball volume ratio for instance, the spectral region is chosen as the first interval with comb=[1] when the RMSECV is the lowest among all the local PLS models developed on the single interval. The chosen interval in combination with all the remaining intervals, the lowest RMSECV is chosen as the second interval combination, comb=[1 13]. The spectral feature selection stops when the RMSECV doesn't decrease with the addition of the number of combination intervals. Similarly, the combination of charge volume ratio with the lowest RMSECV is comb=[18 15 14 13 17 1 16] and combination of pulp density is comb=[7 10 8 9].

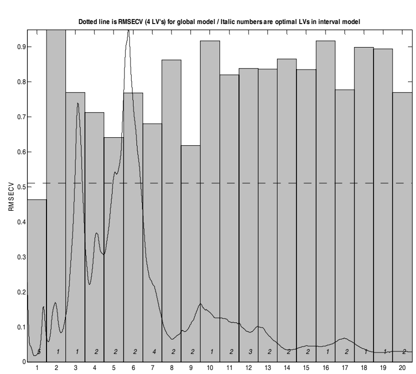
RMSECV of each interval models (bars) and full-spectrum model (dotted line) of the first selected interval of the vibration signals are shows in Figure 3. In Figure 3 (a) to (c), the vertical axis represents RMSECV of all the local PLS models for the mineral to ball volume ratio, charge volume ratio and pulp density, respectively. The number of interval (20) is expressed in abscissa axis. The numbers on the bar graph below represent the optimal number of latent variables of the local PLS models. It can be seen from Figure 3, RMSECV of local PLS model built on the first selected feature frequency band is less than the full-spectrum PLS model in the choice of the three mill load parameters of vibration signals.

Journal of Theoretical and Applied Information Technology
15 September 2012. Vol. 43 No.1

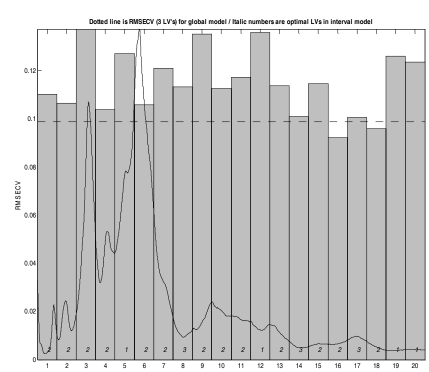
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table I
Feature selection of the shell vibration spectrum closely relevant to the mill load parameters

Mill load parameters	Interval number	Segments	PC	Combination number(Comb)	Interval region	*RMSECV*	r_{cv}	*RMSEP*	r_p
Mineral to ball volume ratio	10	5	5	[6 4 1 5 2 3]	5001-6000,3001-4000, 1-1000,4001-5000, 1001-2000,2001-3000	0.5114	0.8089	0.8039	0.7409
	10	10	6	[4 6 5 2 1 3]	3001-4000,5001-6000, 4001-5000,1001-2000, 1-1000,2001-3000	0.5423	0.7884	0.8907	0.7129
	20	5	5	[16 11 17 10 14 9 13 1]	7501-8000,5001-5500, 8001-8500,4501-5000, 6501-7000,4001-4500, 6001-6500,1-500	0.7704	0.6688	0.2908	0.9521
	20	10	5	[16 11 17 10 14 9 13 1]	7501-8000,5001-5500, 8001-8500,4501-5000, 6501-7000,4001-4500, 6001-6500,1-500	0.7590	0.6420	0.2908	0.9521
Charge volume ratio	10	5	5	[6 10 7 8 9 2 3]	5001-6000,9001-10000, 6001-7000,7001-8000,8001-9000,1001-2000,2001-3000	0.1024	0.7542	0.0545	0.9348
	10	10	5	[10 4 8 9 5 2 3]	9001-10000,3001-4000,7001-8000,8001-9000,4001-5000,1001-2000,2001-3000	0.1249	0.7269	0.0716	0.8827
	20	5	4	[18 15 14 13 17 1 16]	8501-9000,7001-7500, 6501-7000,6001-6500, 8001-8500,1-500,7501-8000	0.1379	0.6933	0.0498	0.9509
	20	10	4	[18 15 14 13 17 1 16]	8501-9000,7001-7500, 6501-7000,6001-6500, 8001-8500,1-500,7501-8000	0.1402	0.6344	0.0498	0.9509
Pulp density	10	5	4	[7 10 8 9]	6001-7000,9001-10000, 7001-8000,8001-9000	0.1605	0.7796	0.0601	0.9684
	10	10	4	[7 10 9 8]	6001-7000,9001-10000, 8001-9000,7001-8000	0.196	0.7533	0.0601	0.9684
	20	5	4	[19 17 9]	9001-9500,8001-8500, 4001-4500	0.1912	0.7507	0.0774	0.9630
	20	10	6	[11 17 9]	5001-5500,8001-8500, 4001-4500	0.2449	0.7376	0.0550	0.9692

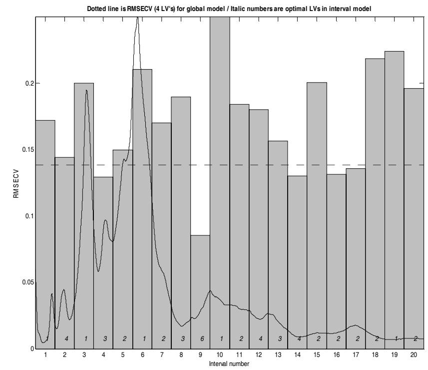


(a)



(b)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

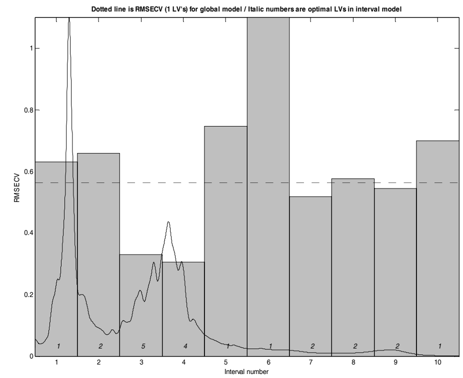


(c)

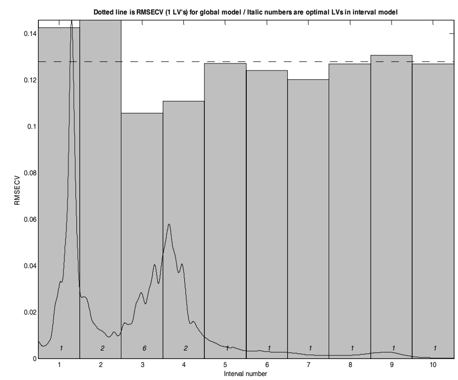
Figure 3: Shell vibration signal feature selection of the mill load parameters (a) Mineral to ball volume ratio (b) Charge volume ratio (c) Pulp density

3.3 Feature Selection of Acoustic Frequency Spectrum

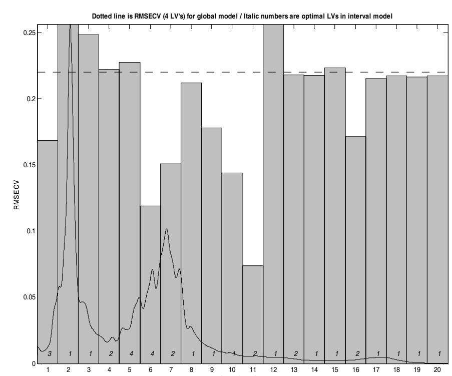
Feature selection of the acoustic signals is similar to feature selection of the shell vibration signals. The best interval combination of the mineral to ball volume ratio is comb=[7 9 6 5 3 1 4] when the full spectrum is split into 10 intervals and five latent variables of the first chosen region. The best interval combination of the charge volume ratio is comb=[10 2 6 5 3] when the full spectrum is split into ten intervals and two latent variables of the first chosen region. The best interval combination of the pulp density is comb=[12 10 9 5 1 11] when the full spectrum is split into 20 intervals and five latent variables of the first chosen region. The first feature band selection of the acoustic signal for the mill load parameters are shown in Fig.4. It can be seen from Figure 4, RMSECV of the first selected feature band model of acoustic signal for three mill load parameters are less than their full spectrum PLS model.



(a)



(b)



(c)

Figure 4: Acoustic signal feature selection of the mill load parameters (a) Mineral to ball volume ratio (b) Charge volume ratio (c) Pulp density

3.4 Comparisons of Feature-Spectrum Model and Full-Spectrum Model

Performance comparison of full spectrum PLS model and feature spectrum PLS model is shown in Table 2. From Table 2, prediction error of the vibration and acoustic feature spectral model is less than their full spectrum model. Full-spectrum

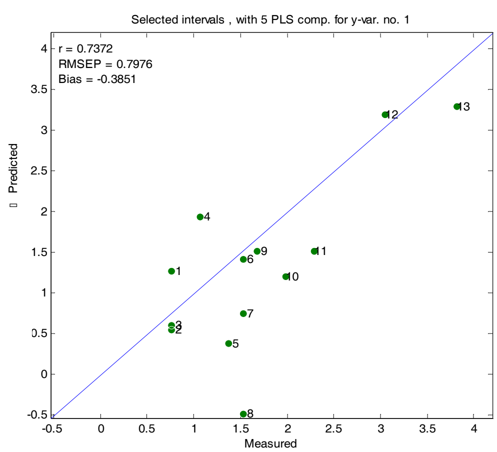
Journal of Theoretical and Applied Information Technology
15 September 2012. Vol. 43 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

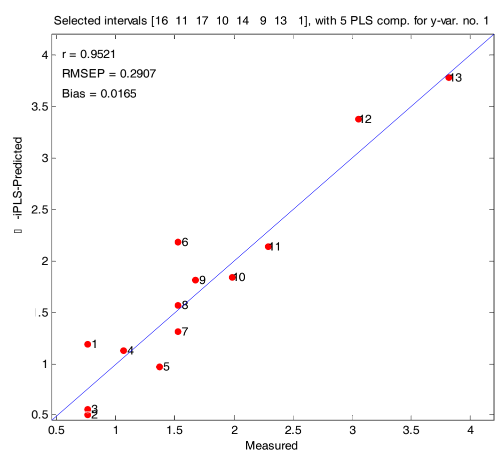
model performance of the shell vibration signals and acoustic signals are compared for three parameters of the mill load. Only the ball volume ratio parameter of the mill load, prediction RMSE of the full spectrum model of acoustic signal is less than the full spectrum model of the shell vibration signal. For the charge volume ratio and pulp density parameters of the mill load, the performance of shell vibration full-spectrum model are superior to the acoustic full-spectrum model. Model results of the full spectrum and the feature spectrum for the shell vibration and acoustic signals are shown in Figure 5. Figure 5 (a) and (b) represent the full spectrum and feature spectrum PLS model of the shell vibration signal, and (c) and (d) represent the full spectrum and feature spectrum PLS model of the acoustic signal. It can be seen from Figure 5, prediction results of the feature spectrum model are superior to the full spectrum for the shell vibration signal and acoustic signal, and performance of feature spectrum model of shell vibration signal is better than the acoustic signal feature spectrum model.

Table II
Performance comparisons of iPLS model and full-spectrum PLS model

Model Data	Modeling method	mineral to ball volume ratio		charge volume ratio		pulp density	
		RMSEP	r_p	RMSEP	r_p	RMSEP	r_p
vibration	Full-spectrum PLS	0.7976	0.7372	0. 1314	0.8117	0. 0582	0.9471
vibration	Feature-spectrum *i*PLS	0.2907	0.9521	0.0550	0.9692	0.0498	0.9509
acoustic	Full-spectrum PLS	0.4500	0.8623	0.2164	0.2459	0.2277	0.5758
acoustic	Feature-spectrum *i*PLS	0.3579	0.9326	0.1091	0.8808	0.1073	0.7255

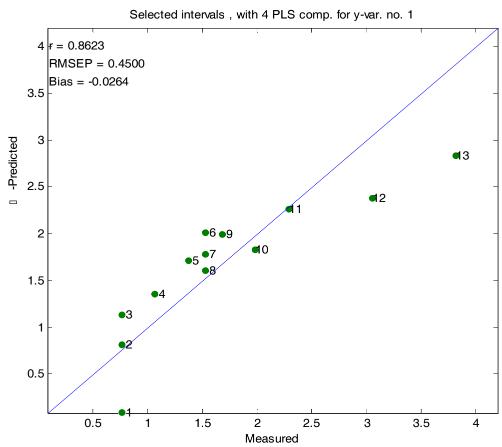


(a)

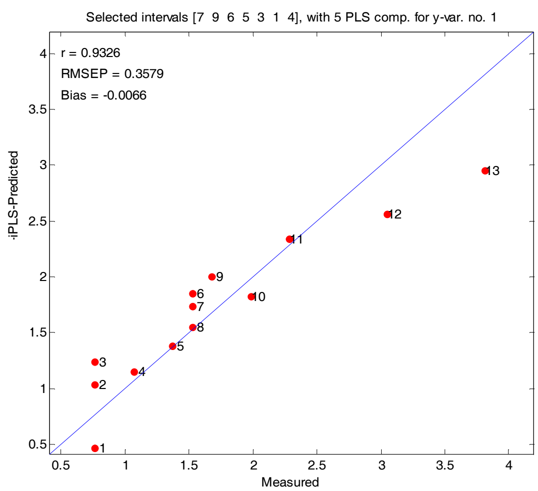


(b)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



(c)



(d)

Figure 5: Predictions of full-spectrum model and feature-spectrum model (a) vibration full-spectrum model; (b) vibration feature-spectrum model; (c) acoustical full-spectrum model; (d) acoustical feature-spectrum model.

4. CONCLUSION

Aiming at the ultra-high-dimensionality and strong collinearity existing in the spectrum data of the wet ball mill, this paper presents a feature selection method of the shell vibration and acoustic spectrum based on interval partial least squares, and builds full spectrum and feature spectrum PLS model of the mineral to ball volume ratio, charge volume ratio and pulp density of the three mill load parameters. The experimental results show that prediction performance of PLS model based on the feature spectrum is better than the full-spectrum model, and the feature spectrum model based on the shell vibration is superior to the acoustic feature spectrum model. Due to the experiment limitations to small samples of a wide range of operating conditions change, the more experiments should be done to further verify the feature bands.

REFRENCES:

[1] Ning Xu, Junwu Zhou, Qing Wang, "Discussion on Intelligent Controlling of Mineral Processing", *Copper Engineering*, Vol. 107, No. 1, 2011, pp. 54-60.

[2] Ping Zhou, Tianyou Chai, "Intelligent Monitoring and Control of Mill Load for Grinding Processes", *Control Theory & Applications*, Vol.25, No.6, 2008, pp. 1095-1099.

[3] Tianjing Feng, Huangang Wang, Wenli Xu, Ning Xu, "An on-line Mill Load Monitoriog System Based on Shell Vibration Signals", *Ming & Metallurgy*, Vol.19, No.2, 2010, pp. 66-69.

[4] M. A. Van Nierop and M. H. Moys, "Exploration of Mill Power Modeled As Function of Load Behaviour", *Minerals Engineering*, Vol.14, No.10, 2001, pp.1267-1276.

[5] Jian Tang, Lijie Zhao, Heng Yue, Tianyou Chai, "Soft sensor for ball mill load based on multi-source data feature fusion", Journal of *Zhejiang University (Engineering Science)*, Vol.44, No.7, 2010, pp. 1406-1413.

[6] Jian Tang, Xiuping Zheng, Lijie Zhao, Heng Yue, Tianyou Chai, "Soft Sensing of Mill Load Based on Frequency Domain Feature Extraction and Information Fusion", *Chinese Journal of Scientific Instrument*, Vol.31, No.10, 2010, pp. 2161-2167.

[7] L. Norgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, S. B. Engelsen, "Interval Partial Least Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy", *Applied Spectroscopy*, Vol. 54, No.3, 2000, pp. 413-419.

[8] Xiaobo Zou, Jiewen Zhao, Yanxiao Li, "Selection of the Efficient Wavelength Regions in FT-NIR Spectroscopy for Determination of SSC of 'Fuji' Apple Based on BIPLS and FIPLS Models", *Vibrational Spectroscopy*, Vol.44, No.2, 2007, pp. 220-227.

[9] Xiangrong Zhu, Na LI, Xinyuan Shi, Yanjiang Qiao, Zhuoyong Zhang, "Analysis of Geniposide and Total Nitrogen Content in Qingkailing Injection Intermediate by Near

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Infrared Spectroscopy and Synergy Interval Partial Least Squares", *Chemical Journal of Chinese Universities*, Vol.29, No.5, 2008, pp. 906-911.

[10] Jianhui Jiang, R. James Berry, Heinz W.Siesler, Yukihiro Ozaki, "Wavelength Interval Selection in Multicomponent Spectral Analysis by Moving Window Partial Least-Squares Regression with Applications to Mid-Infrared and Near-Infrared Spectroscopic Data", *Analytical Chemistry*, Vol.74, No.14, 2002, pp. 3555-3565.

[11] Xiaobo Zhou, Zeng Zhu, Jiewen Zhao, "Selection of the Efficient Wavelength Regions in Agricultural Product NIR Spectroscopy based on Interval Partial Least-Squares (iPLS)", *Modern Scientific Instruments*, Vol.17, No.1, 2007, pp. 86-88.