2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

KPCA Based Multi-Spectral Segments Feature Extraction and GA Based Combinatorial Optimization for Frequency Spectrum Data Modeling

Jian Tang, Tianyou Chai, Wen Yu, Lijie Zhao, S. Joe Qin

Abstract—**Mill load (ML) estimation plays a major role in improving the grinding production rate (GPR) and the product quality of the grinding process. The ML parameters, such as mineral to ball volume ratio (MBVR), pulp density (PD) and charge volume ratio (CVR), reflect the load inside the ball mill accurately. The relative amplitudes of the high-dimensional frequency spectrum of shell vibration signals contain the information about the ML parameters. In this paper, a kernel principal component analysis (KPCA) based multi-spectral segments feature extraction and genetic algorithm (GA) based Combinatorial optimization method is proposed to estimate the ML parameters. Spectral peak clustering algorithm based knowledge is first used to partition the spectrum into several segments with their physical meaning. Then, the spectral principal components (PCs) of different segments are extracted using KPCA. The candidate input features are serial combinated with mill power. At last, GA with Akaike's information criteria (AIC) is used to select the input features and the parameters for the least square-support vector machine (LS-SVM) simultaneously. Experimental results show that the proposed approach has higher accuracy and better predictive performance than the other normal approaches.**

I. INTRODUCTION

Although wet ball mills have been used widely in many grinding processes, they are often operated at low grinding production rates (GPR) [1]. One of the reasons is the lack of reliable on-line sensors for the mill load (ML) [2]. The mechanical grinding of the ball mill produces strong vibration and acoustic signals which are periodic over a given time interval. In the time domain, useful vibration and acoustical signals are buried in wide-band random noise [3]. The power spectral densities (PSD) of these signals contain information which is directly related to the operating state of grinding [3]. Most existing studies focus on the ML parameters of the dry ball mill [4], few on the wet ball mill are reported. Partial least squares (PLS) and principal component regression (PCR) models are used to estimate pulp density (PD) and particle size [3]. The characteristic frequency sub-bands of the axis vibration and mill acoustical signals are used to construct these soft sensor models. However, the axis signal is disturbed by the transfer system. Some studies show that the acoustic signals contain more information of operating parameters than the axis vibration signal, but the acoustic signals have usually crosstalk with adjacent mills [5].

Recently studies show that the on-contact vibration system has at least twice the resolution of the traditional acoustical signal system [6]. The studies of the shell vibration signal for semi-autogenous (SAG) mill show that the shell vibration is a good indicator of the PD and viscosity [7]. A genetic algorithm-partial least squares (GA-PLS) approach has been proposed to select the characteristic frequency sub-bands for different ML parameters based on the PSD of the shell vibration signal [8]. However, it is difficult to interpret the physical meaning of the selected sub-bands and can lead to information loss because of some unselected sub-bands.

The optimal number of features for classification and regression depends on the training data. The key problem to construct a model via high-dimensional spectrum data is the "the Hughes phenomenon" and the "curse of dimensionality" [9]. When a structure vibration is transformed into a frequency spectrum, the modes of vibration and the large cyclical excitation forces are evidenced by peaks in the spectrum [10]. Experimental results of the shell vibration spectrum verify this theory [8]. Therefore, another soft sensor approach is proposed based on the multi-spectral segments principal component analysis (PCA) and support vector machines (SVM) [11]. However, the partition of the spectral segments is manual, PCA cannot extract nonlinear features, and SVM has to solve a quadratic program (QP) problem. Moreover, although the shell vibration is more sensitive than the shell acoustical and mill power signals, studies show that these signals are mainly related to specific parameters. For example, shell vibration relates to PD [12], shell acoustics relates to mineral to ball volume ratio (MBVR) [13], and mill power relates to charge volume ratio (CVR) [14]. Different spectral segments of the same signal contain redundant and complementary information of ML parameters, and principal components (PCs) do not take into account the correlation between inputs and outputs [15]. Although the first several PCs might be

This work was supported in part by the Natural Science Foundation of China (No. 61020106003), Natural Science Foundation for Post-doctoral Scientists of China (No. 20100471464), Natural Science Foundation of China (No. 60874057), and State Administration of Foreign Experts Affairs Special Program for Elite Overseas Experts.

Jian Tang and Tianyou Chai are with the State Key Laboratory of Synthetical Automation for Process Industries (Northeastern University), Shenyang, China (Email freeflytang@gmail.com).

Wen Yu is with the Departamento de Control Automatico, CINVESTAV-IPN, Mexico City, Mexico.

Lijie. Zhao is with the College of Information Engineering, Shenyang University of Chemical Technology, China.

S. J. Qin is with The Mork Family Department of Chemical Engineering and Materials Science, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 (Email: sqin@usc.edu).

able to properly explain the frequency spectrum, they may have less correlation with the ML parameters. When the frequency spectrum PCs with higher prediction performance and lower correlation is used to construct soft sensor models, they can cause worse prediction performance [16]. Studies show that the input feature subsets influence the appropriate model parameters and vice versa, especially in SVM based modeling [17].

In this paper, a novel soft sensor method based on spectral peak clustering, kernel PCA (KPCA), GA and least square-support vector machine (LS-SVM) is proposed to estimate the ML parameters. The spectral peak clustering algorithm is first used to partition the spectrum into several segments with different physical meaning. Then, the spectral kernel PCs (KPCs) of different segments are extracted using KPCA, which are combined with mill power, they are the candidate input features. Finally, GA with the Akaike's information criterion (AIC) is used to optimize the input features and parameters of LS-SVM simultaneously. A successful application is demonstrated on the grinding process of a laboratory-scale wet ball mill.

II. Soft Sensing Approach for ML Parameters

Based on the analysis of the previous section, we use KPCA based multi-spectral segments feature extraction and GA based Combinatorial optimization to estimate the ML parameters. The approach consists of data processing, spectral segment partition, nonlinear spectral feature extraction, combinatorial optimization for input features, and parameters of soft sensor models. The structure is illustrated in Fig. 1. In Fig. 1, superscripts t and f represent the time domain. Subscripts V, A and I represent shell vibration, acoustic and mill power signals, respectively. $x_{org}=\{x_{\rm V}^{\rm t},x_{\rm A}^{\rm t},x_{\rm I}'^{\rm t}\}$ are the original signals; $x_{pro}=\{\mathbf{x}_{\rm V}^{\rm f},\mathbf{x}_{\rm A}^{\rm f},\bar{x}_{\rm I}^{\rm t}\}$ represent the signals after data processing; $\{\mathbf{x}_{\rm V1}^{\rm f},\cdots,\mathbf{x}_{{\rm V}d}^{\rm f},\cdots,\mathbf{x}_{{\rm VD_V}}^{\rm f}\}$ and $\{\mathbf{x}_{\rm A1}^{\rm f},\cdots,\mathbf{x}_{{\rm A}d}^{\rm f},\cdots,\mathbf{x}_{{\rm AD_A}}^{\rm f}\}$ represent the frequency spectral segments; $\{\mathbf{t}_{\rm V1}^{\rm f},\cdots,\mathbf{t}_{{\rm V}d}^{\rm f},\cdots,\mathbf{t}_{{\rm VD_V}}^{\rm f}\}$ and $\{\mathbf{t}_{\rm A1}^{\rm f},\cdots,\mathbf{t}_{{\rm A}d}^{\rm f},\cdots,\mathbf{t}_{{\rm AD_A}}^{\rm f}\}$ represent the nonlinear spectral features of each segment; $\mathbf{x}=[\mathbf{t}'^{\rm f}_{\rm V1},\cdots,\mathbf{t}'^{\rm f}_{{\rm V}d},\cdots,\mathbf{t}'^{\rm f}_{{\rm VD_V}},\mathbf{t}'^{\rm f}_{\rm A1},\cdots,\mathbf{t}'^{\rm f}_{{\rm A}d},\cdots,\mathbf{t}'^{\rm f}_{{\rm AD_A}},\bar{x}_{\rm I}'^{\rm t}]$ represents the optimized input features. c and γ represent the optimized parameters of soft sensor models. $y=\{\varphi_{mb},\varphi_{wm},\varphi_{bmw}\}$ and $\hat{y}=\{\hat{\varphi}_{mb},\hat{\varphi}_{wm},\hat{\varphi}_{bmw}\}$ represent the real and estimates values of ML parameters, respectively.

A. Spectral Segments Partition

In order to avoid the arbitrariness of manual partition and realize the automatic partition, a spectral peak clustering approach is employed. The local peaks of the spectrum are obtained with the method in [18,19]. The definitions of the mass B_m and centroid B_c for the local peaks are:

$$B_m=\sum_{n=n_1}^{n_2}x_n^{\rm f},\qquad B_c=\left(\sum_{n=n_1}^{n_2}n(x_n^{\rm f})^2\Big/\sum_{n=n_1}^{n_2}(x_n^{\rm f})^2\right) \tag{1}$$

where n_1, n_2 denote the frequency range of the local peak and n_c is the central frequency of the local peak. The local peak can be denoted with parameters $<n_1,n_2,n_c,B_m,B_c>$. The local peaks are treated as samples to be clustered, which will be clustered into several peak classes under some criteria. One peak class is one mode of the spectrum, i.e., a spectral segment. The peak class is represented with parameters $\{C_{n_1},C_{n_2},C_{B_m},C_{B_c}\}$. The distance between the peak class and the local peak is defined as [19]:

$$D(B,C)=B_mC_{B_m}\big/(B_c-C_{B_c})^2\ . \tag{2}$$

The modified spectral peak clustering algorithm is described as follows:

1) Specify the number of the peak class N_C. Specify the collection of local peaks $B=\{B_z,z=1,2,...,N_B\}$, where N_B is the number of local peaks and $B_Z=\{n_{z1},n_{z2},n_{zc},B_{zm},B_{zc}\}$ is the zth local peak.

2) Specify the range of the peak class, in which search the

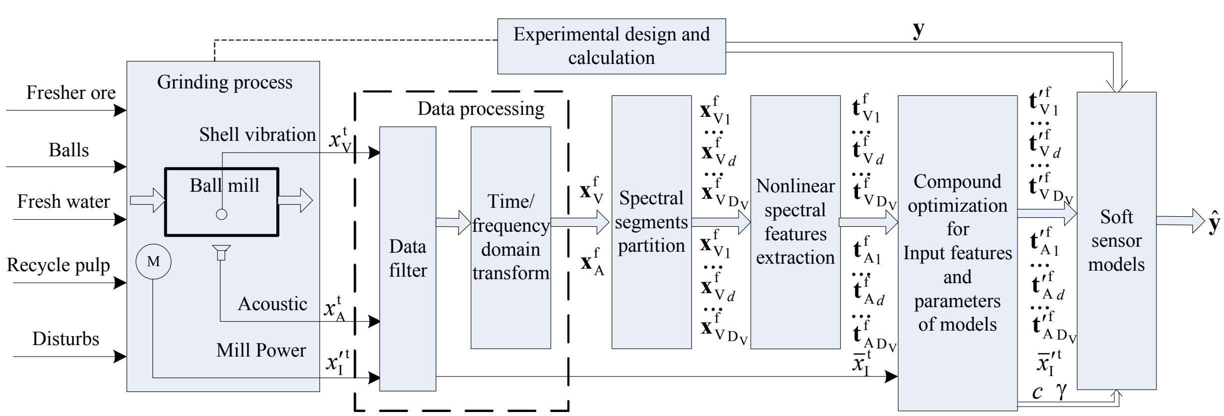

Fig. 1. Structure of the proposed soft sensor method for ML parameters

local peak with maximum mass as the initial peak class. All the peak classes are arranged with the ascending order of the frequency. This is represented as $C=\{C_r, r=1,2,...,N_c\}$, where $C_r=\{C_{rn_1}, C_{rn_2}, C_{rB_m}, C_{rB_c}\}$ is the rth peak class.

3) Calculate the centroid distance ($B_{zc}-C_{rB_c}$) of zth local peak to every initial peak class. Then determine the relative location of the zth local peak to peak class.

4) If the zth local peak is in the left the first peak class, then the local peak is combined into the first peak class.

5) If the zth local peak is in the middle of two initial peak classes, then calculate the distance between the local peak and the neighboring two peak classes with (5). Combine the local peak to the neighboring peak class with a large value, and then obtain a new peak class.

6) If the zth local peak is in the right of the last peak class, then the local peak is combined into the last peak class.

7) Recalculate the mass C_{B_m} and centroid C_{B_c} for every peak class. Repeat step 4) to step 6) until all the local peaks are grouped into one of the peak classes.

By the above algorithm, the frequent spectrum can be grouped into several segments, which can be rewritten as:

$$\begin{cases}\mathbf{x}_{\mathrm{V}}^{\mathrm{f}}=[\mathbf{x}_{\mathrm{V}1}^{\mathrm{f}},\cdots,\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}},\cdots\mathbf{x}_{\mathrm{VD_V}}^{\mathrm{f}}] & d=1,\cdots,\mathrm{D_V}\\ \mathbf{x}_{\mathrm{A}}^{\mathrm{f}}=[\mathbf{x}_{\mathrm{A}1}^{\mathrm{f}},\cdots,\mathbf{x}_{\mathrm{A}d}^{\mathrm{f}},\cdots\mathbf{x}_{\mathrm{AD_A}}^{\mathrm{f}}] & d=1,\cdots,\mathrm{D_A}\end{cases} \quad (3)$$

where $\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}}$ and $\mathbf{x}_{\mathrm{A}d}^{\mathrm{f}}$ are the dth spectral segments of vibration and acoustical spectrum, respectively; D_{V} and D_{A} are the numbers of the spectral segments.

A. Nonlinear Spectral Features Extraction

The objective is to find the projected nonlinear spectral features with maximum variance in a kernel induced feature space. Assume the number of the training samples is L. First, maps the spectral segments $\{(\mathbf{x}_{Vd}^{\mathrm{f}})_l\}_{l=1}^{L}$ and $\{(\mathbf{x}_{Ad}^{\mathrm{f}})_l\}_{l=1}^{L}$ into a higher dimensional space; i.e. mapping $\Phi:(\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}})_l\to\Phi((\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}})_l)$ and $\Phi:(\mathbf{x}_{\mathrm{A}d}^{\mathrm{f}})_l\to\Phi((\mathbf{x}_{\mathrm{A}d}^{\mathrm{f}})_l)$ respectively. Then a linear PCA algorithm is performed. At last we can obtain a nonlinear PCA in the original input space [xx].

For illustration, the extraction of spectral KPCs is derived for $\mathbf{x}_{Vd}^{\mathrm{f}}$. Instead of an explicit nonlinear mapping, the kernel trick is used; i.e. $<\Phi(\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}})_l)\cdot\Phi((\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}})_m>=\mathbf{K}_{\mathrm{V}d}^{\mathrm{f}} \quad l,m=1,2,\cdots L$. The centralized of the $\mathbf{K}_{\mathrm{V}d}^{\mathrm{f}}$ is given by:

$$\widetilde{\mathbf{K}}_{\mathrm{V}d}^{\mathrm{f}}=(\mathbf{I}-\frac{1}{L}1_L1_L^{\mathrm{T}})\mathbf{K}_{\mathrm{V}d}^{\mathrm{f}}(\mathbf{I}-\frac{1}{L}1_L1_L^{\mathrm{T}}) \quad (4)$$

where $\mathbf{I}$ is an L-dimensional identity matrix and 1_L represent the vectors whose elements are ones, with length L. KPCA circumvents the KPC by a dual eigendecomposition problem for kernel Gram matrix $\widetilde{\mathbf{K}}_{\mathrm{V}d}^{\mathrm{f}}\boldsymbol{\alpha}_{\mathrm{V}d}^{h_{\mathrm{v}}}=\mathrm{L}\lambda_h\boldsymbol{\alpha}_{\mathrm{V}d}^{h_{\mathrm{v}}}=\hat{\lambda}\boldsymbol{\alpha}_{\mathrm{V}d}^{h_{\mathrm{v}}}$, in which $\boldsymbol{\alpha}_{\mathrm{V}d}^{h_{\mathrm{v}}}=(\boldsymbol{\alpha}_{\mathrm{Vq}}^{h_{\mathrm{v}}},\boldsymbol{\alpha}_{\mathrm{V}2}^{h_{\mathrm{v}}},\cdots,\boldsymbol{\alpha}_{\mathrm{VL}d}^{h_{\mathrm{v}}})^{\mathrm{T}}$ is the normalized eigenvector associated with the h_{v}th largest eigenvalues. Then the $h_{\mathrm{V}d\max}$ most significant KPC (the number of the KPC) in feature space take the form of

$$\mathbf{v}_{\mathrm{V}d}=[\mathbf{v}_{\mathrm{V}d}^{1},\cdots,\mathbf{v}_{\mathrm{V}d}^{h_{\mathrm{V}d\max}}]^{\mathrm{T}}=[\Phi((\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}})_1),\cdots,\Phi((\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}})_L)]\mathbf{A}_{\mathrm{V}d}^{\mathrm{f}}, \quad (5)$$

where $\mathbf{v}_{\mathrm{V}d}^{h_{\mathrm{v}}}$ is the h_{v}th columns of $\mathbf{v}_{\mathrm{V}d}$ and $\mathbf{A}_{\mathrm{V}d}^{\mathrm{f}}$ is the matrix with the column of $\boldsymbol{\alpha}_{\mathrm{V}d}^{h_{\mathrm{v}}}$, $h_{\mathrm{v}}=1,2,\cdots,h_{\mathrm{V}d\max}$.

The h_{v}th score vector corresponding to $\mathbf{x}_{Vd}^{\mathrm{f}}$ is:

$$\mathbf{t}_{\mathrm{V}d}^{\mathrm{f}}(h_{\mathrm{v}})=\mathbf{v}_{\mathrm{V}d}\cdot\Phi(\mathbf{z}_{\mathrm{V}d}^{\mathrm{f}})=\sum_{l=1}^{L}(\alpha_{\mathrm{V}d}^{h_{\mathrm{v}}})_l\widetilde{\mathbf{K}}_{\mathrm{V}d}. \quad (6)$$

The extracted nonlinear spectral features, namely score vectors, of $\mathbf{x}_{\mathrm{V}d}^{\mathrm{f}}$ and $\mathbf{x}_{\mathrm{A}d}^{\mathrm{f}}$ can be represent as:

$$\begin{cases}\mathbf{t}_{\mathrm{V}d}^{\mathrm{f}}=[\mathbf{t}_{\mathrm{V}1}^{\mathrm{f}},\cdots,\mathbf{t}_{\mathrm{V}h_{\mathrm{v}}}^{\mathrm{f}},\cdots\mathbf{t}_{\mathrm{V}h_{\mathrm{V}d\max}}^{\mathrm{f}}] & h_{\mathrm{v}}=1,\cdots,h_{\mathrm{V}d\max}\\ \mathbf{t}_{\mathrm{A}d}^{\mathrm{f}}=[\mathbf{t}_{\mathrm{A}1}^{\mathrm{f}},\cdots,\mathbf{t}_{\mathrm{A}h_{\mathrm{v}}}^{\mathrm{f}},\cdots\mathbf{t}_{\mathrm{A}h_{\mathrm{A}d\max}}^{\mathrm{f}}] & h_{\mathrm{A}}=1,\cdots,h_{\mathrm{A}d\max}\end{cases} \quad (7)$$

where, $h_{\mathrm{A}d\max}$ is the largest numbers of the KPC for $\mathbf{x}_{\mathrm{A}d}^{\mathrm{f}}$. The values of the $h_{\mathrm{V}d\max}$ and $h_{\mathrm{A}d\max}$ is decided by the thresh value of the contribution ratio of the KPCs. At last, the extracted features for $\mathbf{x}_{\mathrm{V}}^{\mathrm{f}}$ and $\mathbf{x}_{\mathrm{A}}^{\mathrm{f}}$ are given as follows:

$$\begin{cases}\mathbf{t}_{\mathrm{V}}^{\mathrm{f}}=[\mathbf{t}_{\mathrm{V}1}^{\mathrm{f}},\cdots,\mathbf{t}_{\mathrm{V}d}^{\mathrm{f}},\cdots\mathbf{t}_{\mathrm{VD_V}}^{\mathrm{f}}] & d=1,\cdots,\mathrm{D_V}\\ \mathbf{t}_{\mathrm{A}}^{\mathrm{f}}=[\mathbf{t}_{\mathrm{A}1}^{\mathrm{f}},\cdots,\mathbf{t}_{\mathrm{A}d}^{\mathrm{f}},\cdots\mathbf{t}_{\mathrm{AD_A}}^{\mathrm{f}}] & d=1,\cdots,\mathrm{D_A}\end{cases} \quad (8).$$

B. GA-based Combinatorial Optimization

The extracted nonlinear features of vibration and acoustical spectrum combined with mill power are used as the candidate input features, which are represented as:

$$\mathbf{x}_{\mathrm{all}}=[\mathbf{t}_{\mathrm{V}}^{\mathrm{f}},\mathbf{t}_{\mathrm{V}}^{\mathrm{f}},\bar{x}_{\mathrm{I}}^{\mathrm{t}}] \quad (9)$$

The least square-support vector machines (LS-SVM) simplifies the QP problem to solve a set of linear equations by changing the loss function in SVM to a sum of squared errors [xxi]. Therefore, the LS-SVM algorithm with the RBF kernel functions is used to build the nonlinear models for ML parameters. As different ML parameters are related to different spectral features, it is necessary for different LS-SVM models to select different inputs. At the same time, the selection of input features is a problem of multi-source information fusion. Proper fusion can improve the prediction and robustness of soft-sensor models. GA is a promising alternative to conventional heuristic methods, which can deal with large search spaces efficiently and is not prone to local optimal solutions. The simultaneous selection of the input features and LS-SVM parameters is considered as a combinatorial optimization problem, which can be solved by GA and other intelligent optimization algorithms.

We define f_{ea} as the parameter for GA to select, which is represented as:

$$
\begin{aligned}
&f_{ea}=\{f_{\mathrm{V}},\ \ f_{\mathrm{A}},\ \ f_{\mathrm{I}}\} &&,\quad where\\
&f_{\mathrm{V}}=\{f_{\mathrm{V1}},\cdots,f_{\mathrm{Vd}},\cdots,f_{\mathrm{VD_V}}\}, &&0\le f_{\mathrm{V}d}\le h_{\mathrm{V}d\max} \qquad (10)\\
&f_{\mathrm{A}}=\{f_{\mathrm{A1}},\cdots,f_{\mathrm{Ad}},\cdots,f_{\mathrm{AD_A}}\}, &&0\le f_{\mathrm{A}d}\le h_{\mathrm{A}d\max}\\
&f_{\mathrm{I}}=\{h_{\mathrm{I}}\} && h_{\mathrm{I}}=\{0,1\}
\end{aligned}
$$

are the parameters of vibration spectrum, acoustical spectrum and mill power respectively; $f_{\mathrm{V}d}$ and $f_{\mathrm{A}d}$ are the numbers of KPCs for dth spectral segment.

Here, we denote the regularization parameter and the radius of the RBF as c and γ, respectively. The objective of the Combinatorial optimization is to select f_{ea}, c and γ by minimizing the error between the output of the LS-SVM model $\hat{\mathbf{y}}$ and the true value $\mathbf{y}$. In order to improve the prediction accuracy of the soft sensor model, Akaike's information criteria (AIC) is used as the fitness function of GA, which is represented as:

$$\Phi_{\zeta}(f_{\mathrm{ea}},c,\gamma)=L_{\mathrm{val}}\cdot\ln\left(\left(\frac{1}{L_{\mathrm{val}}}\sum_{l=1}^{L_{\mathrm{val}}}[\mathbf{y}(l)-\hat{\mathbf{y}}(l)]^2\right)\Big/L_{\mathrm{val}}\right)+2\cdot(p+1) \qquad (11)$$

where L_{val} is the number of the test samples and p is the number of the selected input features.

The selected input features of the training samples can be represented as:

$$\mathbf{x}=[\mathbf{t}'^{\mathrm{f}}_{\mathrm{V}1},\cdots,\mathbf{t}'^{\mathrm{f}}_{\mathrm{V}d},\cdots,\mathbf{t}'^{\mathrm{f}}_{\mathrm{VD_V}},\mathbf{t}'^{\mathrm{f}}_{\mathrm{A}1},\cdots,\mathbf{t}'^{\mathrm{f}}_{\mathrm{A}d},\cdots,\mathbf{t}'^{\mathrm{f}}_{\mathrm{AD_A}},\bar{x}'^{\mathrm{t}}_{\mathrm{I}}] \qquad (12)$$

Denoting the new sample as $\mathbf{z}$, the resulting prediction models for ML parameters based on LS-SVM is:

$$\hat{y}=\sum_{l=1}^{L}\beta_l\mathbf{K}(\mathbf{x}_l,\mathbf{z})+b\,. \qquad (13)$$

III. Laboratory Experiment Test

A. Experiment System

The experiments are performed on a laboratory scale ball mill (XMQL-420×450). The vibration signal is picked up by an accelerometer located on the middle of the shell, and the acoustic signal and the mill power are also recorded. In order to find the impact of every possible load and operating parameters on the shell vibration at different grinding conditions, several assumptions on the mill operating conditions are made. Due to limited space, experimental detail is described in [8] and omitted here.

B. Application Results

The following parameters are used to calculate the PSD using the Welch's method: the data length is 32,768, the section number is 32 and the overlap fraction length is 512. The detail of the PSD for different grinding conditions is given in [8]. The curves of the vibration frequency spectrum for building models are shown in Fig.2.

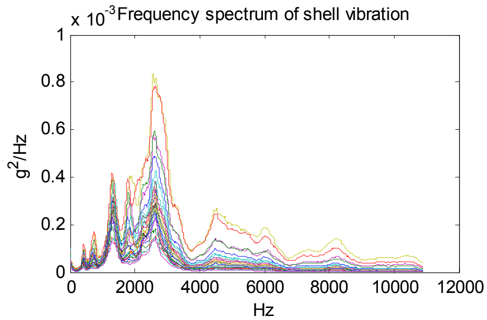

Fig.2. Curves of the vibration frequency spectrum for building models

TABLE I
PERCENT VARIANCE CAPTURED BY PLS ALGORITHM

	LV#	Vibration spectrum X-Block	Vibration spectrum Y-Block	Acoustic spectrum X-Block	Acoustic spectrum Y-Block	Mill power X-Block	Mill power Y-Block
M	1	94.99	10.66	51.48	57.30	100	45.79
B	2	1.94	51.17	12.19	30.06	NaN	NaN
V	3	0.69	16.83	5.50	8.16	NaN	NaN
R	4	0.14	13.98	4.89	2.81	NaN	NaN
	1	95.14	46.22	12.91	70.71	100	0.93
P	2	1.27	34.74	47.82	7.76	NaN	NaN
D	3	1.07	10.26	7.66	13.80	NaN	NaN
	4	1.51	1.65	8.12	3.28	NaN	NaN
C	1	95.11	36.09	52.19	41.82	100	34.56
V	2	1.87	36.97	8.30	35.81	NaN	NaN
R	3	0.91	8.96	5.43	13.35	NaN	NaN
	4	0.94	7.85	6.57	4.37	NaN	NaN

Because the PLS algorithm aims to maximize covariance between the input and output data using a few latent variables (LVs), it is used to analyze the relationships between the full frequency spectrum, mill power and ML parameters. The statistical results are shown in Table I.

Based on TABLE I, the percent variance captured by the 1st LV shows that the vibration frequency spectrum has the largest correlation with PD, 95.14% to 46.22%; the acoustical frequency spectrum has the largest correlation with MBVR, 51.48% to 57.30%; the mill power also has the largest correlation with MBVR, 45.79%. It is shown that the information fusion and feature selection are necessary to construct the effective ML parameters models.

Fig. 2 shows that we can partition the vibration frequency spectrum into at least three segments. The vibration spectrum is composed of 60 small peaks. Using the proposed spectral peak clustering algorithm, the vibration spectrum is partitioned into four segments, which are 102-2385Hz (Vibration Low Frequency, VLF), 2385-4122Hz (Vibration Medium Frequency, VMF), 4122-7227Hz (Vibration High Frequency, VHF) and 7600-11000Hz (7227-11000Hz, Vibration High-high Frequency, VHHF), respectively. Using the same steps, the acoustical frequency spectrum is partitioned into three segments, which are 1-1071Hz (Acoustical Low Frequency, ALF), 1071-1688Hz (Acoustical Medium Frequency, AMF) and 1688-2073Hz (Acoustic High Frequency, AHF), respectively.

TABLE II
CONTRIBUTION RATIOS OF KPCS FOR DIFFERENT SPECTRAL SEGMENTS

Signals	Spectral segments	KPC # 1	2	3	4	5
Vibration	VLF	88.26	5.271	3.023	1.704	0.0696
	VMF	98.31	1.162	0.0244	0.0137	0.0046
	VHF	99.80	0.114	0.029	0.015	0.010
	VHHF	99.85	0.0798	0.0335	0.0098	0.0079
Acoustic	ALF	85.20	7.278	3.558	1.290	1.199
	AMF	90.02	4.645	2.015	1.352	0.0895
	AHF	87.23	7.804	1.782	1.375	0.0598

TABLE III
PERFORMANCE ESTIMATION OF DIFFERENT MODELING

	Approach	Data sets[a]	Parameters[b]	RMSSE
MBVR	PCR	V+A+I	Number of LVs:6	0.2471
	PLS	V+A+I	Number of LVs: 4	0.2546
	GAPLS	V+A+I	Number of LVs:8	0.2502
	This paper	PCA_V	{(42825,450),(4,1,2,1),(#,#,#),(#)}	0.1767
		$KPCA_V$	{(28225, 536),(2,3,0,1),(#,#,#),(#)}	0.2128
		PCA_A	{(31507, 327),(#,#,#),(1,2,2),(#)}	0.1808
		KPC_A	{(46855, 520),(#,#,#),(1,1,3),(#)}	0.1811
		$I+PCA_{VA}$	{(21758, 596),(2,1,0,1),(2,1,0),(1)}	0.1637
		$I+KPCA_{VA}$	{(26406, 398),(1,1,0,2),(2,2,0),(0)}	0.1325
PD	PCR	V+A+I	Number of LVs: 6	0.1922
	PLS	V+A+I	Number of LVs: 4	0.1479
	GAPLS	V+A+I	Number of LVs:12	0.1365
	This paper	PCA_V	{(42825, 450),(4,1,2,1),(#,#,#),(#)}	0.1401
		$KPCA_V$	{(38948, 454),(1,1,2,2), (#,#,#),(#)}	0.07995
		PCA_A	{(49123, 415),(#,#,#),(3,3,5),(#)}	0.1486
		KPC_{AA}	{(49846, 300), (#,#,#),(1,1,0), (#)}	0.2241
		$I+PCA_{VA}$	{(48482, 301),(1,3,0,0),(1,2,0),(0)}	0.1123
		$I+KPCA_{VA}$	{(20946, 426),(2,1,0,2),(1,1,2),(0)}	0.07249
CVR	PCR	V+A+I	Number of LVs: 6	0.2184
	PLS	V+A+I	Number of LVs: 7	0.1826
	GAPLS	V+A+I	Number of LVs: 5	0.1598
	This paper	PCA_V	{(44677, 365),(7,1,0,0), (#,#,#),(#)}	0.1793
		$KPCA_V$	{(49767, 468),(7,1,0,0),(#,#,#),(#)}	0.1491
		PCA_A	{(27968, 597),(#,#,#),(3,3,2),(#)}	0.1985
		$KPCA_A$	{(49312, 321),(#,#,#) (2,3,0), (#)}	0.2534
		$I+PCAV_A$	{(20268, 563),(1,4,0,0),(1,2,0),(0)}	0.1343
		$I+KPCA_{VA}$	{(42870, 461),(1,4,0,0),(1,1,0),(0)}	0.1187

[a] The subscript 'V', 'A' and 'I' indicates the shell vibration , acoustical and mill power signals respectively.

[b] The parameters are defined as {(c , γ), (f_{V1} , f_{V2} , f_{V3} , f_{V4}), (f_{A1} , f_{A2} , f_{A3}),(f_I)},which are the parameters of the LS-SVM models, the numbers of the spectral PCs of the vibration and acoustical spectrum, the status of mill power respectively. '#' indicates that the value of the parameters is not included in the soft sensor model.

The results in TABLE II show that the variances captured by KPCA with the same RBF kernel for different spectral segments are different. It is necessary to select different KPCs in different spectral segments for construct effective ML parameters soft sensor models.

In this paper, the initial numbers of the KPCs are selected as 6, 4, 2, 2, 8, 8 and 8 for VLF, VMF, VHF, VHHF, ALF, AMF and AHF, respectively. These spectral features are serial combination with mill power as the candidate features with (12). The length of bit strings for encoding c , γ and f_{ea} are 20, 20 and 19 respectively. After encoding the candidate input features and LS-SVM parameters, the Combinatorial optimization based on GA is made. Considering the random initialization problem of GA's population, the simulation is running at least 20 times.

For comparison, the soft sensor models of different approaches and data sets are built simultaneously. The numbers of the LVs for PCR and PLS algorithm are decided with leave-one-out cross validation method. The prediction results and statistical parameters are listed in TABLE III, in which RMSSE is the root mean square relative error of the testing data set. The real and estimated curves with PLS, PCR and the proposed method are shown in Fig. 3~ Fig. 5.

The models based on GA-PLS have better prediction accuracy than PCR and PLS, but the former has the longest training time. The results show that it is important to select

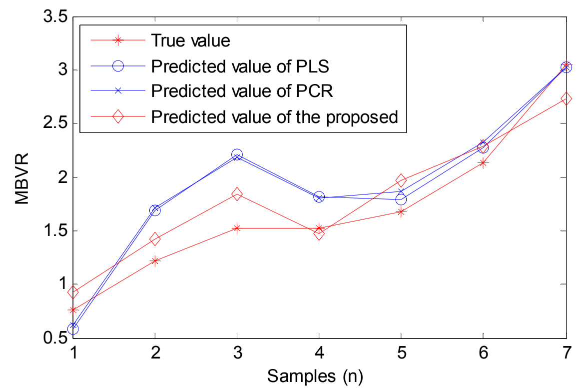

Fig. 3. Prediction result of the MBVR

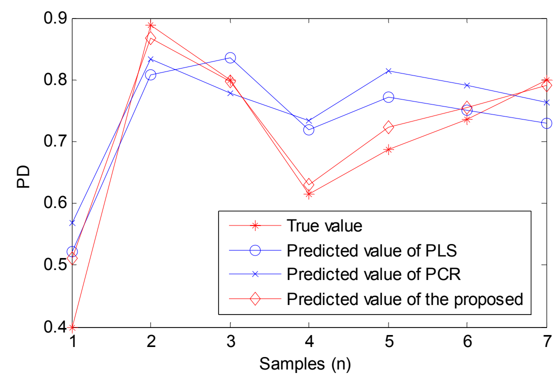

Fig. 4. Prediction results of the PD

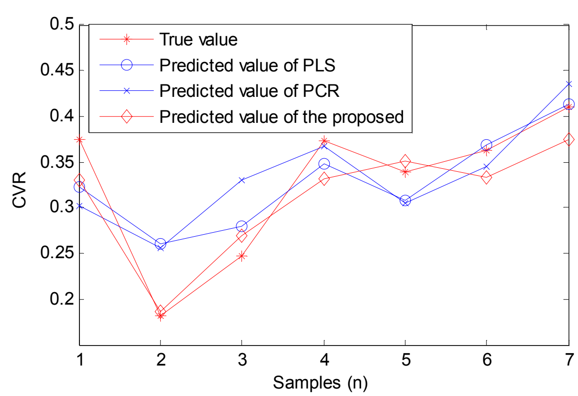

Fig. 5. Prediction results of the CVR

different input features for different ML parameters, and the sensitivity of the ML parameters to the frequency spectral segments and mill power are different. For example, the f_{ea}

for MBVR, PD and CVR are {(1,1,0,2),(2,2,0),(0)}, {(2,1,0,2),(1,1,2),(0)} and {(1,4,0,0),(1,1,0),(0)} respectively. The results based on different data set also shows that only vibration or acoustical spectrum is insufficient to model all the ML parameters satisfactorily. The vibration spectrum has the highest sensitivity with the PD soft sensor model, the RMSSE of which is 0.07995, but for the proposed approach, the RMSSE is 0.07249. The results consist with the mechanism analysis, which also shows the superiority of shell vibration in modeling the ML parameters. But the contribution of mill power is not as important as in the industry-scale ball mill, the reason maybe come from the laboratory-scale ball mill which is too small to reflect the ML.

IV. Conclusions

In this paper, a soft sensor approach for estimating the wet ball mill load based on FFT, Clustering, KPCA, GA, and LS-SVM is proposed. This novel approach possesses the following characteristics:

1) Features of the frequency spectrum are extracted;

2) Frequency spectrum is partitioned into segments with physical meaning automatically;

3) The number of nonlinear KPCs for the spectral segments is specified in advance;

4) Information fusion is achieved in multi-source sensors effectively

5) Input features and the LS-SVM parameters are optimized simultaneously.

This approach is successfully applied to a laboratory-scale grinding process with designed experiments, which produce better predictive performance.

References

[1] P. Zhou, T. Y. Chai and H. Wang, "Intelligent optimal-setting control for grinding circuits of mineral processing," IEEE Transactions on Automation Science and Engineering, vol. 6, no. 4, pp. 730-743, 2009.

[2] B. Behera, B. K. Mishra and C. V. R. Murty, "Experimental analysis of charge dynamics in tumbling mills by vibration signature technique," Minerals Engineering, vol. 20, no. 1, pp. 84–91, 2007.

[3] Y. Zeng, E. Forssberg, "Monitoring grinding parameters by vibration signal measurement-a primary application," Minerals Engineering, vol. 7, pp. 495-501, 1994.

[4] G.Q. Si, H. Cao, Y.B. Zhang, L.X. Jia, "Experimental investigation of load behaviour of an industrial scale tumbling mill using noise and vibration signature techniques," Minerals Engineering, vol. 22, no.15, pp. 1289-1298, 2009.

[5] Y. Zeng, E. Forssberg, "Monitoring grinding parameters by signal measurements for an industrial ball mill," International Journal of Mineral Processing, vol. 40, no. 1, pp. 1-16, 1993.

[6] K. Gugel, G. Palcios, J. Ramirez, M. Parra, Improving ball mill control with modern tools based on digital signal processing (DSP) technology, In: IEEE Cement Industry Technical Conference, Dallas, 2003, pp. 311-318.

[7] S.J. Spencer, J.J. Campbell, K.R. Weller, Y. Liu, Acoustic emissions monitoring of SAG mill performance, Intelligent Processing and Manufacturing of Materials, 1999, pp. 936-946.

[8] J. Tang, L.J. Zhao, J.W. Zhou, H. Yue, T.Y. Chai, "Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell," Minerals Engineering, vol. 23, no. 9, pp. 720-730, 2010.

[9] K. Fukunaga, "Effects of sample size in classifier design," IEEE Trans.Pattern Anal. Machine Intell., vol. 11, pp. 873–885, Aug. 1989.

[10] M.H. Richardson, D.L. Formenti, Global curve fitting of frequency response measurements using the rational fraction polynomial method, Proceedings of the 3th International Modal Analysis Conference, Orlando, Florida, 1985, pp. 390-397.

[11] J. Tang, L.J. Zhao, W. Yu, H. Yue, T.Y. Chai, "Soft sensor modeling of ball mill load via principal component analysis and support vector machines," Lecture Notes in Electrical Engineering, vol. 67, pp. 803-810, 2010.

[12] Spencer, S. J., Campbell, J. J., Weller K. R. and Liu, Y. Acoustic emissions monitoring of SAG mill performance. Intelligent Processing and Manufacturing of Materials, Honolulu, HI , USA, 1999, 10-15 July, 936-946.

[13] Z.H. Wang, B.C. Chen, "Present state and development trend for ball mill load measurement," Chinese Powder Science and Technology (Chinese), vol. 7, no. 1, pp. 19-23, 2001.

[14] M. A. Van Nierop and M. H. Moys, "Exploration of mill power modeled as function of load behavior," Minerals Engineering, vol. 14, no. 10, pp.1267-1276, 2001.

[15] J.L. Liu. "On-line soft sensor for polyethylene process with multiple production grades," Control Engineering Practice, vol. 15, pp. 769-778, 2007.

[16] I.T. Jolliffe, Principal Component Analysis, Berlin, Springer Press, 2002.

[17] C. L. Huang, C. J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," Expert Systems with Applications, vol. 31, pp. 231-240, Aug. 2006.

[18] J. Tang, W. Yu, L. J. Zhao, H. Yue and T. Y. Chai, Modeling of operating parameters for wet ball mill by modified GA-KPLS, The Third International Workshop on Advanced Computational Intelligence, SuZhou, China, pp. 107-111, 2010.

[19] Q. W. Ye and T. Q. W, "Optimization modal analysis with PSO based on spectrum segmentation", Chinese Journal of Scientific Instrument, vol. 30, pp. 547-588, Aug. 2009.

[xx] B. Scholkopf, A. Smola and K. R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," Neural Computation, vol. 10, no. 5, pp. 1299-1319, 1998.

[xxi] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers", Neural Processing Letters, vol. 99, no. 3, pp. 293-300, 1999.