Online Detection of Belt Deviation of Belt Weigher Using Extreme Learning Machine

Liang Zhu[1]

School of Mechanical Engineering, Nanjing University of Science & Technology, Nanjing 210094, China
E-mail: `6762065493@qq.com`

Dongbo Li

School of Mechanical Engineering, Nanjing University of Science & Technology, Nanjing 210094, China

Yanqiang Yuan

Nanjing Sanai Industrial Automation Co. Ltd., Nanjing, 211100, China

The belt deviation is one of the important factors which affect the performance of belt weigher. A new online detection of belt deviation of belt weigher based on the existing weighing data and environmental data is developed by using the integration of the ELM and Principle Components AnalysisPCA algorithm to avoid the extra detection equipment. Firstly, the PCA is used to extract features from the original data; secondly, the ELM algorithm is applied to train the SLFNs as the classifier which is developed to identify the belt deviation of belt weigher from the features extracted by PCA; finally, the experiment of the detection of belt deviation is performed on the array belt weigher. The results show that the proposed detection model by using the ELM combined with PCA is able of handling the detection of belt deviation of belt weigher online.

[1]Speaker

1. Introduction

Continuous weighing of bulk materials is the measurement equipment in respect of bulk materials trade, widely used in a variety of agricultural and industrial production and trading places. The belt weigher (BW) is just one type of mass measuring instrument in the family of continuous weighing of bulk materials, which can be used wherever the continuous materials are recorded in conveyor belt system; however, because of the harsh and long time high-load operating environment, the weighing inaccuracy and the equipment accidents occur frequently, which always lead to great economic losses and even exert certain impact on personal safety directly. Surveys have indicated that 70% to 80% of the faults are caused by the belt deviation [1]; therefore, it is necessary to develop a real-time detection system of belt deviation.

Previous researches mostly focused on designing additional sensors, such as array phototransistors, CCD and PSD to detect the belt deviation[2][3].; however, these detection methods can greatly increase the hardware cost and the poor working environment has also exerted seriously bad effect on the sensors. Besides, it is not convenient for the data fusion. As a result, the research of belt deviation only based on the existing BW is of great significance. In order to achieve the detection of belt deviation, the feature of belt deviation should be extracted from the weighing data and environmental data; at meanwhile, the corresponding identification model needs to be put forward to identify the feature. In this sense, the real-time detection of belt deviation based on weighing data and environmental data can be summed up to an online"feature extraction & classification" problem.

With the development of technology, in order to acquire better measurement accuracy, the belt weigher is equipped with more and more weighing units, which make the detection of belt deviation from the mass weighing data and environmental data complex and difficult; therefore, an efficient feature extraction method is significantly important for reducing the dimensionality of the feature of belt deviation and the complexity of subsequent recognition algorithm, and will give rise to considerable savings in term of the operation cost and capital. Besides, the success or not of the feature extraction will affect the identification accuracy of the belt deviation. The principal component analysis (PCA) is a method of dimensionality reduction, in which an orthogonal transformation called K-L transform is applied to map a set of observations of possibly correlated feature into the principal space so that certain principal components of high priority can be selected to enhance the accuracy of classification[4][5].

As to identification model, it can be built by using the machine learning, like logistical regression, softmax regression, artificial neural networks (ANN), etc. According to the estimation of Jaeger, 95% of the literature on ANN is mainly about single hidden layer feedforward networks (SLFNS). SLFNs have no side or back connections to connect two nodes and are generally trained by the back propagation algorithm which suffers from slow convergence and local minimum problem. Aiming at the problems of SLFNs, the extreme learning machine (ELM) was proposed by Huang[6] and has been widely used in all fields[7]. In the networks trained by ELM, the weights from the input layer to the hidden layer and the hidden thresholds are randomly chosen, and then the output weights of networks are analytically determined [8][9].

In this paper, we propose a real-time detection system of belt deviation which uses PCA and ELM to identify the belt deviation from the weighing data and environmental data. The remainder of this paper is organized as follows. Section 2 provides a short summary of PCA. In Section 3, a small overview of ELM is given. In Section 4, a case study is performed and the proposed method is detailed. Experiment results are discussed and the following conclusion of this work is depicted.

2. Principle Component Analysis

Principal component analysis is a multivariate statistical method to study multivariable correlation and data compression, which was first proposed by Pearson[10]. PCA algorithm can be summarized as follows:

Step1. Pre-treat the original datasets: scaling. Let the $k \times n$ matrix X represent the n datasets of data in which each dataset consists of k variables:

$$X = \left[x^1, x^2, ..., x^n \right] = \begin{bmatrix} x_1^1, x_1^2, ..., x_1^n \\ x_2^1, x_2^2, ..., x_2^n \\ \vdots \ \ \vdots \ \ \ddots \ \ \vdots \\ x_k^1, x_k^2, ..., x_k^n \end{bmatrix} \tag{2.1}$$

Then the scaled datasets can be formulated as

$$X = \left[x^{1*}, x^{2*}, ..., x^{n*} \right] = \begin{bmatrix} x_1^{1*}, x_1^{2*}, ..., x_1^{n*} \\ x_2^{1*}, x_2^{2*}, ..., x_2^{n*} \\ \vdots \ \ \vdots \ \ \ddots \ \ \vdots \\ x_k^{1*}, x_k^{2*}, ..., x_k^{n*} \end{bmatrix} \tag{2.2}$$

where x_j^{i*} ($i=1,...,n$ $j=1,...,k$) represents the scaled value, and is calculated by

$$x_j^{i*} = \begin{cases} 0, & \|x_j - 1_n\|_2 \\ (x_j^i - \bar{x}_j)/\|x_j - 1_n\|_2, else \end{cases} , \quad \bar{x}_j = 1/n \cdot \Sigma_{i=1}^n x_j^i$$

$$x_j = \left[x_j^1, x_j^2, ..., x_j^n \right]^T , \quad 1_n = \left[1,1,...,1 \right]^T_{n \times 1} \tag{2.3}$$

Step2. Calculate the covariance matrix. The covariance matrix R is calculated as

$$R = \left(X^* \right) \left(X^* \right)^H \tag{2.4}$$

Step3. Determine the eigenvalues and the eigenvectors of covariance matrix R. According to the quantities, the eigenvalues of R can be sorted as $\overset{1}{\lambda} \geq \overset{2}{\lambda} \geq ... \geq \overset{k}{\lambda} \geq 0$, and $\overset{1}{p}, \overset{2}{p}, ..., \overset{k}{p}$ represents the corresponding eigenvectors. The eigenvalues and eigenvectors can be gained by the singular value decomposition of the covariance matrix R.

Step4. Determine the scores. Pick up the eigenvectors $\overset{1}{p}, \overset{2}{p}, ..., \overset{s}{p}$ corresponding to the s largest eigenvalues $\overset{1}{\lambda} \geq \overset{2}{\lambda} \geq ... \geq \overset{s}{\lambda} > 0$ to form the projection matrix L, and finally for any observation x of k variables the score Y can be formulated as,

$$Y = L \cdot x = \left[p_1, p_2, ..., p_s \right]^T \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{bmatrix}$$

$$\tag{2.5}$$

3. Extreme Learning Machine

Extreme learning machine is an emerging learning technique proposed for generalized SLFNs to overcome the challenging issues faced by BP learning algorithm [11]. Unlike the traditional SLFNs, the training speed of ELM has been greatly improved, and the training only generates the unique optimal solution avoiding falling into local optimum[12][13].

For n arbitrary different samples $(x^i,\ t^i)$, where $x^i=[x_1^i,x_2^i,...,x_k^i]^T\in\Re^k$, $t^i=[t_1^i,t_2^i,...,t_l^i]^T\in\Re^l$, and the output function of ELM for generalized SLFNs is

$$f(x^i)=\Sigma_{j=1}^{L}\beta_j^T\cdot h(w_j^T\cdot x^i+b_j) \tag{3.1}$$

where β_j is the output weights from the jth hidden node to the output node; $w_j\in\Re^k$ and $b_j\in\Re$ are the weights and thresholds from the input nodes to the jth hidden node. The linear function is chosen as the excitation function of the output layer in ELM.

The standard SLFNs with L hidden nodes using the activation function $h(x)$ can approximate to these n samples with zero error which means $\Sigma_{j=1}^{n}\|f(x^i)-t^i\|_2=0$, i.e., there exist β_j, w_j and b_j so that

$$H\beta=T \tag{3.2}$$

where

$$H(w_{1,.}...,w_L,x_{1,.}...,x_n,b_{1,.}...,b_L)=\begin{bmatrix} h(w_1^T\cdot x^1+b_1) & \cdots & h(w_L^T\cdot x^1+b_1) \\ \vdots & \vdots & \vdots \\ h(w_1^T\cdot x^n+b_1) & \cdots & h(w_L^T\cdot x^n+b_L) \end{bmatrix}_{n\times L} \tag{3.3}$$

$$T=[t^1,t^2,...,t^n]_{n\times l}^T \tag{3.4}$$

$$\beta=[\beta_1,\beta_2,...,\beta_n]_{L\times l}^T \tag{3.5}$$

therefore, the training of the SLFNs can be completed by finding the least-square solutions $\tilde{\beta}$ of $H\beta=T$, in which the complex iterative computation as well as the local extreme problem can be avoided. The least-square solutions $\tilde{\beta}$ can be calculated as,

$$\tilde{\beta}=H^+T \tag{3.6}$$

where H^+ is the Moore-Penrose generalized inverse matrix of H. Further, $\tilde{\beta}=H^+T$ is the sole solution and has the smallest norm among all the least-squares solutions of $H\beta=T$:

$$\|\tilde{\beta}\|=\|H^+T\|\leqslant\|\beta\| \tag{3.7}$$

According to the theory of the generalization ability, the smaller the weights, the better generalization ability of SLFNs will be. Hence, $\tilde{\beta}=H^+T$ can make the network to have great generalization performance; besides, the threshold b of the hidden layer and the weights w connecting the input layer and the hidden layer can be chosen randomly[14].

4. Case Study

4.1 Acquisition of Experiment Data

In this section, we will conduct a case study on the 3# array belt weigher (ABW) which can recycle the materials in the BW test center of Nanjing Sanai Industrial Automation Co. Ltd. The ARM processor STM32F103ZET6 is applied to sample the real-time data, and then transmit the real-time data through the RS485 bus to the upper PC which receives the real-time data by using the serial communication program developed in MATLAB. The sampling frequency of the real-time data is 10Hz. Then, the belt deviation can be detected by the model based on ELM and PCA, which are also developed in MATLAB, as soon as the real-time data is collected.
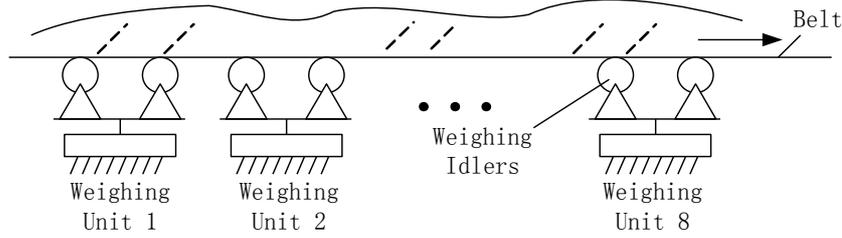
Figure 1: Schematic Diagram of ABW

As shown above, there are 8 weighing units in 3# ABW in general. These weighing units are always designed far enough away from the loading point to avoid the weighing inaccuracy caused by the impact of the load, especially as to the sudden large materials[15]. During each experiment, the max total amount of materials through any ABW is 50t, and Table 1 shows the major parameters of 3# ABW:

Width of belt	Idler spacing	Thickness of belt	Groove angle of idler
1000mm	1200mm	12mm	30°

Table 1: Parameters of 3# ABW

We gather the data in respect of 8 weighing units and the environmental data without any belt deviation in the early stage, in which the 3# ABW runs at the flows of no-load, 300t/h and 600t/h. The data of 8 weighing units with belt deviation at three different flows and the environmental data are gathered in the second stage. before building the detection model of belt deviation, all the gathered data are subdivided into the learning sets (training and validation) and the test sets.

4.2 Establishment of Detection Model

As described in Section 1, in order to realize the intelligent detection of belt deviation of ABW, save the hardware cost and improve the efficiency of detection model, a scheme is proposed by using PCA and ELM algorithm, as shown in Fig. 2:
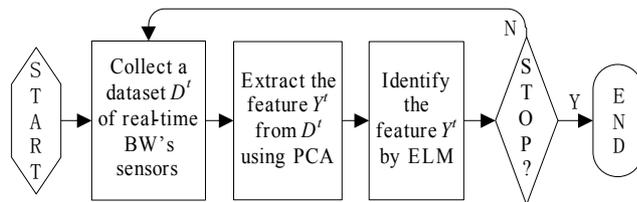
Figure 2: Process of Online Detection of Belt Deviation

During the experiments, we selected 15797 effective samples containing 7900 learning samples and 7897 test samples. Because the belt deviation mainly leads to the linear offsets of weighing unit data, all the distributions of the samples are unique with linear structure; therefore, PCA, which is a linear method, is more suitable to the online extracting of feature Y and runs faster than other manifold learning methods such as ISOmap, MDS, LE, NPE and LLE etc. The input of ELM is the feature Y extracted by PCA. The ELM is adopted as the online classifier owing to its rapid training. The detection model of belt deviation can be established in accordance with the following steps:

(1) Using the learning datasets to calculate the projection matrix L as depicted in Section 2;

(2) Convert all collected datasets to the features Y with the projection matrix;

(3) Set the number of hidden nodes according to the dimension of the feature Y;

(4) Select the threshold b of the hidden layer and the weights w connecting the input layer and the hidden layer randomly with uniform distribution;

(5) Calculate the output weights β from the hidden node to the output node as depicted in Section 3;

4.3 Experiments by Using ELM in Combination with PCA

In this subsection, we firstly take the raw learning samples as the input and train the SLFNs by ELM directly; secondly, we apply PCA to convert all the samples to the features *Y*, and then train the SLFNs by ELM, BP and RBF with the learning samples; finally, we test the models of ELM, PCA-ELM, PCA-BP and PCA-RBF with the test samples of three different flows containing the data of eight weighing units and the environmental data. The results are summarized and compared in Tables 2- 3:

Algorithm	Average training time of 2172 samples	Average test accuracy(%)	Average testing time of 2369 samples	Number of nodes
ELM	26.48ms	86.94	8.75ms	20
PCA-ELM	30.42ms	92.24	8.43ms	20
PCA-BRNN	63.80s	92.6	27.36ms	20
PCA-RBF	6.354	89.98	26.42	60

Table 2: Comparison of Different Models on the 300t/h dataset

Algorithm	Average training time of 7900 samples(ms)	Average test accuracy(%)	Average testing time of 7897 samples	Number of nodes
ELM	47.77	88.1	30.36	20
PCA-ELM	55.24	93.5	28.43	20
PCA-BRNN	213508.592	91.12	52.24	20
PCA-RBF	81386.15	82.89	64.98	60

Table 3: Comparison of Different Models on All Datasets

Table 2 shows the comparison of different models which are all trained and tested by the 300t/h dataset. Table 3 shows the comparison of different models at different flows. Both the comparisons show that the PCA-ELM model spends far less training time and testing time than those of PCA-BRNN or PCA-RBF. PCA-ELM also obtains the second best average test accuracy with respect to the flow of 300t/h and the best on all datasets, which indicates the great generalization ability. In addition, during the training process, we also find that the RBF and BRNN take up far more memories than that the ELM does when the quality of learning samples is very huge.

5. Conclusion

In this paper, in order to detect the belt deviation of BW from the existing data containing the weighing data and the environmental data, we propose an online detection model of belt deviation by using the ELM combined with PCA, which is used to extract features of the belt deviation from the weighing data and environmental data. The results in Tables 2-3 imply that the PCA can improve the performance of the following and reduce the dimension of the samples, thereby greatly reducing the complexity of the following classification. After that, the ELM algorithm is introduced to train the classifier. The results in Tables 2-3 indicate that the ELM spends much less training and testing time and obtains great performance with relatively fewer hidden nodes. Above all, the results show that the model based on the PCA-ELM is very suitable for the online detection of belt deviation of BW, the application of the model based on PCA-ELM will save a lot of hardware cost by replacing the existing belt deviation detection equipment and greatly improve the quality of belt deviation detection.

References

[1] J.-W. Zhang , P. H. Lou. *Automatic Detection and Hydraulic Correction Technology of Belt Deviation*. Mechanical Engineer. 2008(10): 26-28(In Chinese)

[2] D. Jie. *Application research to PSD in the laser displacement detecting system*. Infrared and Laser Engineering, 2007, 36(z1): 281-284(In Chinese)

[3] Z. Liang, D. B. Li, et al. *Research and design of detection system for belt deviation based on FPGA+MCU*. Machine tool & hydraulics, 2014(7): 86-89,95(In Chinese)

[4] C.Y. Yang, T.Y. Wu. *Diagnostics of gear deterioration using EEMD approach and PCA process. Measurement,* 2015(62): 75-87

[5] L. Tao, TIAN Hongxiang, GUO Wenyong. *Application of PCA to Diesel Engine Oil Spectrometric Analysis.* Spectroscopy And Spectral Analysis, 2010, 30(3): 779-782

[6] G. -B. Huang, Q.-Y. Zhu, Chee-Kheong Siew. *Extreme learning machine: Theory and applications*. Neurocomputing, 2006, 70(1): 489-501

[7] A. A. Mohammed, R. Minhas, Q.M. Jonathan, M.A. Wu, Sid-Ahmed. *Human face recognition based on multidimensional PCA and extreme learning machine*. Patten Recognition, 2011, 44(7): 2588-2597

[8] A. Grigorievskiya, Y. Miche, et al. *Long-term time series prediction using OP-ELM*. Neural Networks, 2014, 51(1): 50-56

[9] M. N. Elliackin, Figueiredo, B. Teresa, Ludermir. *Investigating the use of alternative topologies on performance of the PSO-ELM.* Neurocomputing, 2014, 127(1): 4-12

[10]K. Pearson, *On lines and planes of closest fit to systems of points in space*, Philos. Mag. 2 (6) (1901) 559–572.

[11]G.-Bi. Huang. *Learning capablity and storage capacity of two-hidden-layer feedforward networks*, IEEE Transactions on Neural Networks, 2003, 14(2): 274-281

[12]G. Huang, G.g-B. Huang, S. J. Song, K. Y. You. *Trends in extreme learning machines: A review.* Neural Networks, 2015, 61(1): 32-48

[13]G.-Bi. Huang. *An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels*. Cognitive Computation, 2014, 6(3): 376-390

[14]Y. Qi. *Approach of EEG detection based on ELM and approximate entropy. Chinese Journal of Scientific Instrument*, 2012, 3(3): 514-519(In Chinese)

[15]V. K. Donis, A. E. Rachkovskii, and V. M. Sin. *How the conveyor belt length affects belt weigher accuracy. Measurement Techniques*, 2004, 47(2): 163-167