Machine Monitoring Using Fuzzy-Neural Networks

Kuo-Cheng Ting¹, Tzu-Yu Lin¹, Yi-Chung Chen^{2*}, Jia-Ching Ying¹, Don-Lin Yang¹, and Hsi-Min Chen¹

¹ Department of Information Engineering and Computer Science, Feng Chia University, Taichung, Taiwan, R.O.C.

² Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Yunlin, Taiwan, R.O.C.

(Received 8 December 2017; Accepted 11 January 2018; Published on line 1 June 2018)

*Corresponding author: chenyich@yuntech.edu.tw

DOI: [10.5875/ausmt.v8i2.1686](https://doi.org/10.5875/ausmt.v8i2.1686)

Abstract: In response to the rapid pace of technological change, many big manufacturers are increasingly looking towards solutions based on plant informatization and Industry 4.0 concepts. However, in the context of Taiwan, such options are off limits to many small and medium-sized firms due to limited scale and capital. This paper proposes a plant informatization approach which can be implemented by smaller manufacturers through using add-on sensor systems to monitor production equipment. An accelerometer is installed on existing machinery to collect vibration data, which is subjected to feature extraction to create a monitoring model through implementing the LDA algorithm and the fuzzy neural networks. Experimental results indicate the resulting model can be effectively used to detect abnormal machinery operations.

Keywords: Abnormality identification; feature extraction; accelerometer

Introduction

The fourth Industrial Revolution focuses on leveraging Internet of Things (IoT) technologies and big data analysis to improve the value chain of entire manufacturing plants along with manufacturing efficiency and quality. For example, sensors installed on production machinery can collect various types of processing data and allow for continuous and automated operation monitoring. Plant managers can thus anticipate machine wear and breakage, thus preventing unnecessary downtime, improving product yield and enhancing plant capacity. However, implementation of such solutions requires sensors be embedded in the machinery. Retrofitting older equipment entails considerable expense. Most factories in Taiwan are traditional small and medium-sized factories without the financial resources to implement such solutions.

This paper proposes a method by which an accelerometer is integrated into existing production machinery to measure machine signals. This approach

allows manufacturers to automatically monitor older machinery, thus effectively extending the lifespan of plant equipment and significantly reducing costs. However, equipment fault diagnosis still poses considerable challenges due to a lack of manufacturing parameter information for older machinery. Chen et al. proposed a method of diagnosing equipment fault integrating subjective technician experience with neural networks [5], but the resulting accuracy detection is insufficiently high for effective implementation. This paper uses Fuzzy neural networks [3][7][9][11] which can describe data distributions more accurately through the multi-dimensional Gaussian function as shown in Fig.1.

Experiments using true machine data are conducted to validate the effectiveness of the approach. Data are collected from a Taiwan spring manufacturer producing high-quality springs on machinery in use for over ten years. Given the firm's high quality standards (only five failures allowed for every 10,000 springs produced), there is strong potential for implementation of Industry 4.0 technologies.

The remainder of this paper is organized as follows.

Section 2 reviews the relevant literature, while section 3 introduces the research methods. Experimental results are summarized in section 4, and conclusions and future directions are provided in section 5.

Related works

Feature Computation

Feature computation uses mathematical formulas to output the data of different sensor data points; the different signals can be used to express different data features. We adapt the Windows method [2] and integrate feature computation to extract spring production machinery data features and increase model accuracy.

Kuo-Cheng Ting received the BE degree from the Department of Information Engineering and Computer Science, Feng Chia University, Taichung, Taiwan (R.O.C) in 2004. He is currently working toward the Ph.D. degree and his main research interests in Industry 4.0 and artificial intelligences.

Tzu-Yu Lin is now the M.S. students of the Department of Information Engineering and Computer Science, Feng Chia University, Taichung, Taiwan (R.O.C). His main research interests in Industry 4.0 and artificial intelligences.

Yi-Chung Chen received the B.S. and M.S. degrees in Electrical Engineering from National Cheng Kung University, Tainan, Taiwan, in 2007 and 2008, and the Ph.D. degrees in Department of Computer Science and Information Engineering from National Cheng Kung University, Tainan, Taiwan, in 2014. Since then, he joined the faculty of Feng Chia University and worked as the Assistant Professor for three years. He is currently an Assistant Professor in the Department of Industrial Engineering and Management, National Yunlin University of Science and Technology. His research interests include Industry 4.0, databases, recommendation system, social network analysis and artificial intelligences.

Jia-Ching Ying received the Ph. D. degree in Computer Science and Information Engineering from National Cheng Kung University, Tainan, Taiwan, in 2013. After graduating from National Cheng Kung University, he worked successively in National Cheng Kung University and National Chiao Tung University as a postdoc research fellow. In 2016, he visited University of Southern Queensland, Australia as a visiting academic from January to March. After coming back Taiwan, he was hired by as a Data Scientist in Emotibot Technologies Limited which is committed to innovate the first emotion-aware artificial intelligence robot. He is currently an Assistant Professor in the Department of Information Engineering and Computer Science, Feng Chia University. His initial research interests were data mining and natural language processing, and over time he evolved toward deep learning, big data analytics, spatial trajectory computing, location-based social networking and mobile computing with application.

Don-Lin Yang received a B.E. degree in Computer Science from Feng Chia University, Taiwan, in 1973, an M.S. degree in Applied Science from the College of William and Mary in 1979, and a Ph.D. degree in Computer Science from the University of Virginia in 1985. He worked at IBM Santa Teresa Laboratory from 1985 to 1987 and at AT&T Bell Laboratories from 1987 to 1991. Since then, he joined the faculty of Feng Chia University and is currently a Professor in the Department of Information Engineering and Computer Science. He received 2015 IEET Distinguished Teaching Award. His research interests include data mining, database system, and software engineering. Dr. Yang is a member of the IEEE Computer Society and the ACM.

Hsi-Min Chen received the B.S. and Ph.D. degrees in computer science and information engineering from National Central University, Taoyuan, in 2000 and 2010, respectively. He is currently an Assistant Professor with the Department of Information Engineering and Computer Science, Feng Chia University, Taichung, Taiwan. His research interests include software engineering, object-oriented technology, service computing, and distributed computing.

Fig. 1. Example of Fuzzy neural networks adopting the multi-dimensional Gaussian function to describe the data distribution.

Features Dimension Reduction

The two most typically used dimension reduction methods include (1) combining multiple features into a smaller number of new features better suited to the recognition of a target dataset and (2) directly selecting the features that are most effective for the recognition of a target dataset from among all features. This work focuses on the first approach as it provides better dimensional reduction for accelerometer-related data.

Principal component analysis (PCA) [4] and linear discriminant analysis (LDA) [1] are the two most commonly used algorithms for feature combination. As in the unsupervised dimension reduction method, during the dimension reduction process, PCA will rotate and compress the original data so as to identify the most representative mapping dimension while maintaining the features of the original data. However, PCA cannot be used to determine the data category, thus it is possible that different data categories will be highly superimposed on the same mapping dimension, making LDA a better choice. As a supervised method for dimension reduction, LDA maps the originally high-dimensionality complex data into the computed vectors, so as to close the distance between data of the same category and enlarge the distance between data belonging to different categories.

Classifier

When analyzing the condition of machinery, the reduced dimension results are used for classification requiring a classifier capable of classifying different data categories into different results. At present, the most commonly used classifiers are Support Vector Machine (SVM)[8][10][12], K-nearest Neighbors [12] and Neural Network [12][13]. Neural networks take the reduced dimension features as input, outputting labels of different data categories for machinery diagnosis, i.e. normal operation, imminent fault or faulty.

Frequently asked questions about machine

After understanding the computation of Features, dimension reduction of high-dimension data and application of classifier, we shall learn about how and what category of data shall be collected from the machine for analysis. The frequently asked questions of machine and the corresponding physical signals can be known from the previous research [2], so as to figure out what category of Sensors can be applied to collect the machine signals. For example, machine problems occurring along the principal axis may generate vibration and noise which can be detected using an accelerometer and microphone. Alternatively, reduced motor efficiency along the principal axis might generate vibration, noise, necessitating the additional use of a thermometer.

Algorithms

This section introduces the research problem, method flow, sensors, feature computation and dimension reduction, along with the model training method.

Problem Research Problem

As previously noted, small and medium-sized plants rely on older production machinery, creating problem for even the collection and processing of even very basic data. All problems must be manually reported by onsite personnel, creating difficulties improving the product yield and plant production capacity. Previous studies have proposed machine status monitoring solutions [2], but these solutions still fail to satisfy manufacturer requirements. The present research proposes improvements to this approach by 1. changing the window length, 2. increasing the dimensionality of feature computation, 3. altering the classifier pattern according to the data status, and 4. monitoring current machine status through the FNN inspection.

Method flow

The operating flow can be divided into several steps: (1) We use an add-on accelerometer to collect vibration data from the machinery, (2) this data is then preprocessed to normalize the influence of gravity, (3) the processed data are used for feature computation, using Windows within a fixed range to output signal data to the various data point categories, (4) due to the use of different feature computation categories, the original training data is transformed into high-dimensionality data which requires LDA dimension reduction, so as to sequence data with high feature content, and (5) data

with high feature content is used for FNN training while outputting the neural parameters for the data identification of the machine status.

In addition, when new test data are being inputted, the processing flow of the new data will be identical to that of the original data and the known vectors with reduced dimensions shall be input into the test data flow to identify machine status.

Sensors data collection

In this research, the accelerated velocities of the X, Y and Z axes are respectively collected to measure machine vibrations. Due to the impact of earth gravity, the vibration data cannot be directly processed as collected. High-pass filtering is first used to filter out the effects of gravity, leaving the vibration signals generated by the actual principal axis.

Features computation

The Windows method is most frequently used to process vibration signal data collected from accelerometers. However, machinery used to produce springs of different lengths produce different degrees of vibration, thus the sequence data must be cut and processed for features computation using the superimposed Window method. This results in the output features being unable to truly represent the motion period of a spring, thus the Window coverage is modified by using the actual motion period of each vibration signal for feature computation, thus outputting the true data features. We then followed [13] to retrieve the following data features: mean value, energy value, mean square root, mean absolute error, standard deviation, crest value, period of finished spring product and variable square value.

Dimension Reduction of Features

Feature computation of the accelerometer data will produce at least twenty-four dimensions, but neural training for high-dimension data is inefficient because not every dimension could completely represent the features of the original data. What's more, how to decide the dimension to be reduced. A limitation of LDA is that the reduced dimension must be the number of categories minus one; therefore, two dimensions are reduced in this paper, and the LDA can map the high-dimension complex data to the computed vectors while distinguishing among the different categories of data variability. Using the sparse matrix mentioned in [1] we obtain the co-variance matrix S_B between different categories and the variance matrix S_W within various categories, where S_B is the predicted discrete vectors near the mean value while S_W is the sparse value of

various predicted vectors.

$$S_B = \sum_{a=1}^N n_a (m^a - m)(m^a - m)^T, \quad (1)$$

$$S_W = \sum_{a=1}^N n_a \sum_{i=1}^{n_a} (X_i^a - m^a)(X_i^a - m^a)^T. \quad (2)$$

where N represents the number of categories, n_a represents the sample number in category a , $X_i(a)$ represents the i sample in the No. a category, $m(a)$ represents the mean vector value of the sample in category a and m represents the mean vector of all data. LDA computes the high-dimension into a mapping vector $\mathbf{w}$, so as to store the category integrity under the low dimension. The following formula can be maximized through $\mathbf{w}$, and the co-variance between the different categories can be maximized and the co-variance within the same category can be minimized, which means the $J(\mathbf{w})$ in Eq. (3) can be maximized. After computing $\mathbf{w}$, the data can be mapped into $\mathbf{w}$ and the new coordinates can be discovered to achieve reduce the high dimension to the lower one.

$$J(\mathbf{w}) = \frac{\mathbf{w}^T S_B \mathbf{w}}{\mathbf{w}^T S_W \mathbf{w}}. \quad (3)$$

Training and identification of fuzzy neural networks

The fuzzy neural networks (FNN) used here are a combination of fuzzy logic and neural networks. As FNN is constructed using fuzzy logic concepts, it uses a non-binary non-binary description, making FNN more applicable for the proposed data model. The model also has the learning capacity of neural networks, ensuring accurate learning. As shown in Fig. 2, FNN is divided into four layers: an input layer, fuzzy layer, rule layer and output layer. The reduced dimensionality feature data is sent to the input layer, which is then sent to the fuzzy layer for membership degree computation by the Gaussian function. Afterwards, the output of the fuzzy layer is subjected to multiplicative movement to obtain the movement response. Finally, all responses of the rule layer are integrated by the output layer to determine the FNN output data. We now introduce the equations of the proposed FNN, where the input and the output of the i th neuron in j th layer are respectively u_i^j and o_i^j . First, the equation of the input layer should be:

$$O_i^1 = u_i^1 = r. \quad (4)$$

where r is the input of the FNN. Next, the equations of the membership layer and rule layer should be:

$$O_{ij}^2 = \exp \left\{ -\frac{(O_i^1 - m_{ij})^2}{(\sigma_{ij})^2} \right\}, \quad (5)$$

$$O_i^3 = \prod u_i^3, \quad (6)$$

where m and σ are respectively the mean and the standard deviation of the Gaussian function. Finally, the equation of the output layer should be

$$o_j^4 = \sum_i^m u_{ij}^4 w_i^4. \quad (7)$$

FNN is a form of supervised learning to reduce the variability between the output value and the expected value, thus the variability between the two must be computed. FNN continuously corrects the weighting according to the variability. Therefore, the error function E is defined to express the learning qualities of FNN as shown below:

$$MSE = \frac{1}{n} \sum_1^n (\text{observed}_i - \text{predicted}_i)^2, \quad (8)$$

where the training output data and actual output data are used for computation. After computing the E value, the error function E can be used to deliver and correct the set weight, as shown in Eq. (5).

$$\Delta w = -a \times \frac{\partial E}{\partial w}, \quad (9)$$

where Δw represents the weight correction, a is the learning rate, ∂w is the weight and ∂E is the output variability. Finally, we can obtain all equations for updating the parameters of FNN by using Eq. (9).

$$m(k+1) = m(k) + \xi ew \prod o^2 \frac{1}{\sigma}, \quad (10)$$

$$\sigma(k+1) = \sigma(k) + \xi ew \prod o^2 \left[\frac{o^1 - m}{\sigma^2} \right], \quad (11)$$

$$w(k+1) = w(k) + \xi eo^3, \quad (12)$$

Fig. 2. The FNN structure used in this paper.

Experimental simulation

Experimental simulations were run using vibration signals from machinery in the spring manufacturing plant along with the data set used in [2][5][6]. Table 1 presents 13 types of spring data, of which 1 type indicates a normal spring, and 6 types each indicate vibrations caused by an excessively short or long springs. As abnormalities cannot be predicted in advance, the vibration data for this data was recorded manually with constant search and production times for production runs of 6 springs. Each data category was divided into 80% training and test sets in an 80/20 ratio. Experiments were conducted on a PC running Windows10 64bit with an Intel i7 processor and RAM 16G, using Matlab as the compiler.

Data Dimension Reduction

For dimension reduction, this paper uses LDA which improves the variability of different data categories and reduces feature variability within single categories. The standard for LDA dimension reduction is the number of data categories minus one; therefore, given three categories in the experimental data sets, the number of data feature dimensions must be reduced to two. Figures 3 and 4 respectively show the dimensions before and after reduction. Prior to reduction, the features of the three data categories are highly intersecting, but this is reduced in the two-dimensional data used for LDA.

Table 1. Spring data sets used in the experimental simulations.

Clustering	Spring length (mm)
Overly-short defective products	171, 181, 187, 188, 189, 190
Non-defective products	191
Overly-long defective products	192, 193, 194, 195, 210, 230

Fig. 3. Spring data set prior to dimension reduction (taking the first three signs of all dimensions).

Fig. 4. Spring data set reduced to three dimensions.

Machine Status Monitoring

Table 2 shows that previous research method [2] only correctly identifies short and long springs 50% of the time, but the use of FNN increases identification accuracy to 89%.

Conclusion

Machinery vibration obtained through add-on sensors is subjected to pre-processing and then inputted into FNN for training and then used for machine monitoring, providing traditional small and medium-sized plants with a cost-effective means to monitor plant equipment status.

Acknowledgment

This work was supported in part by the Ministry of

Science and Technology of Taiwan, R.O.C., under Contracts MOST 106-2119-M-224-003.

Table 2. Results of identification degree.

Category	Spring length (mm)	Overall identification rate (BPNN) [5]	Overall identification rate (FNN)
Overly-short defective products	171 181 187 188 189 190	50%	89.7%
Non-defective products	191	90%	90%
Overly-long defective products	192 193 194 195 210 230	51%	92.3%

References

- [1] W.-S. Chen, Y. Zhao, B. Pan, and B. Chen, "Supervised kernel nonnegative matrix factorization for face recognition," *Neurocomputing*, vol. 205, pp. 165-181, Sep. 2016.
doi: [10.1016/j.neucom.2016.04.014](https://doi.org/10.1016/j.neucom.2016.04.014)
- [2] Y.-M. Chen, "A Prototype System of Production Scheduling and Machine Tool Monitoring with Add-on Sensors" Master thesis, Dept. of Information Engineering and Computer Science, Feng Chia University, Taichung, Taiwan, 2016.
- [3] P. Eklund and F. Klawonn, "Neural fuzzy logic programming," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 815-818, Sep. 1992.
doi: [10.1109/72.159071](https://doi.org/10.1109/72.159071)
- [4] X. Kang, X. Xiang, S. Li, and J. A. Benediktsson, "PCA-Based Edge-Preserving Features for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7140-7151, Sep. 2017.
doi: [10.1109/TGRS.2017.2743102](https://doi.org/10.1109/TGRS.2017.2743102)
- [5] C.-J. Kuo, K.-C. Ting, Y.-C. Chen, D.-L. Yang, and H.-M. Chen, "Automatic machine status prediction in the era of industry 4.0: Case study of machines in a spring factory," *Journal of Systems Architecture*, vol. 81, no. 11, pp. 44-53, Nov. 2017.
doi: [10.1016/j.sysarc.2017.10.007](https://doi.org/10.1016/j.sysarc.2017.10.007)
- [6] C.-J. Kuo, K.-C. Ting, and Y.-C. Chen, "State of product detection method applicable to Industry 4.0 manufacturing models with small quantities and great variety: An example with springs," in *proceeding of IEEE Int. Conf. on Applied System Innovation (ICASI 2017)*, Sapporo, Japan, May 13-17, 2017.
doi: [10.1109/ICASI.2017.7988251](https://doi.org/10.1109/ICASI.2017.7988251)
- [7] K.-C. Kwak and D.-H. Kim, "Adaptive Neuro-Fuzzy networks with the aid of fuzzy granulation," *IEICE Trans. Information and Systems*, vol. E88-D, no. 9, pp. 2189-2195, Sep. 2005.
- [8] J. López, S. Maldonado, and M. Carrascob, "Double regularization methods for robust feature selection and SVM classification via DC programming," *Information Sciences*, vol. 429, pp. 377-389, March 2018.
doi: [10.1016/j.ins.2017.11.035](https://doi.org/10.1016/j.ins.2017.11.035)
- [9] A. K. Paul, P. C. Shill, M. R. I. Rabin, and K. Murase, "Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease," *Applied Intelligence*, pp. 1-18, Sep. 2017.
doi: [10.1007/s10489-017-1037-6](https://doi.org/10.1007/s10489-017-1037-6)
- [10] Z.-W. Sun, Z.-W. Guo, X.-P. Wang, J. Liu, and S.-Y. Liu, "Fast Extended One-Versus-Rest Multi-label SVM Classification Algorithm Based on Approximate Extreme Points," in *proceeding on Int. Conf. on Database Systems for Advanced Applications (DASFAA)*, Suzhou, China, March 27-30, 2017, pp. 265-278.
doi: [10.1007/978-3-319-55753-3_17](https://doi.org/10.1007/978-3-319-55753-3_17)
- [11] J. Tang, F. Liu, Y. Zou, W. Zhang, and Y. Wang, "An Improved Fuzzy Neural Network for Traffic Speed Prediction Considering Periodic Characteristic," *IEEE Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2340-2350, Sep. 2017.
doi: [10.1109/TITS.2016.2643005](https://doi.org/10.1109/TITS.2016.2643005)
- [12] J.-Y. Yang, Y.-P. Chen, G.-Y. Lee, S.-N. Liou, and J.-S. Wang, "Activity Recognition Using One Triaxial Accelerometer: A Neuro-Fuzzy Classifier with Feature Reduction," *Lecture Notes in Computer Sciences*, vol. 4740, pp. 395-400, Jan. 2007.
doi: [10.1007/978-3-540-74873-1_47](https://doi.org/10.1007/978-3-540-74873-1_47)
- [13] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using Acceleration Measurements for Activity Recognition: An Effective Learning Algorithm for Constructing Neural Classifiers," *Pattern Recognition Letters*, vol. 29, no. 16, pp. 2213-2220, Dec. 2008.
doi: [10.1016/j.patrec.2008.08.002](https://doi.org/10.1016/j.patrec.2008.08.002)