Isomap and Deep Belief Network-Based Machine Health Combined Assessment Model

Aijun Yin[1] – Juncheng Lu[1,*] – Zongxian Dai[2] – Jiang Li[1] – Qi Ouyang[3]

[1] Chongqing University, The State Key Laboratory of Mechanical Transmissions, Chongqing, China
[2] No.1 Branch of Chongqing Academy of Metrology and Quality Inspection, Chongqing, China
[3] Chongqing University, College of Automation, Chongqing, China

This paper presents a novel combined assessment model (CAM) for machine health assessment, in which 38 original features of the vibration signal were extracted from time domain analysis, frequency domain analysis, and wavelet packet transform (WPT), following which the nonlinear global algorithm Isomap was adopted for dimensionality reduction and extraction of the more representative features. Next, the acquired low-dimensional features array is input into the well trained deep belief network (DBN) model to evaluate the performance status of the bearing. Finally,after the bearing accelerated degradation data from Cincinnati University were investigated for further research, through the comparison experiments with two other popular dimensionality reduction methods (principal component analysis (PCA) and Laplacian Eigenmaps) and two other intelligent assessment algorithms (hidden Markov model (HMM) and back-propagation neural network (BPNN)), the proposed CAM has been proved to be more sensitive to the incipient fault and more effective for the evaluation of bearing performance degradation.
Keywords: Isomap, dimensionality reduction, deep belief network (DBN), machine health, combined assessment model (CAM)

Highlights
- *A novel effective machine health assessment model was proposed by integrating Isomap into DBN.*
- *Time-domain analysis, frequency-domain analysis, and WPT were adopted to extract the original features.*
- *The optimal DBN model was constructed after a series of experiments.*
- *Comparison experiments with PCA and Laplacian Eigenmaps, HMM and BPNN, proved the validity of the proposed CAM.*

0 INTRODUCTION

Effective machine health assessment provides diverse benefits including improved safety, improved reliability and reduced costs for operation and maintenance of complex manufacturing systems. Monitoring and assessing the degradation trend of some key components, such as bearings, allow the degraded behavior or faults to be corrected before they cause machine breakdown [1]. According to our literature review, it has been discovered that many of the existing works are based on signal-processing methods, such as instantaneous energy density self-similarity, time-frequency entropy, and regularization dimension for health assessment [2] to [6].

However, as the data collected from machines is becoming increasingly greater and the requirements of the speed and accuracy of machine health assessment are becoming higher and higher, there is an urgent need for new assessment methods that can effectively analyse massive amounts of data and automatically provide accurate assessment results. Yu [1] proposed a hidden Markov model (HMM) and contribution analysis-based method to assess machine health degradation. Do and Nguyen [7] utilized adaptive empirical mode decomposition (AEDM) for bearing fault detection and made comparisons with those of

using envelope analysis and the latest version of the EMD. In addition, back-propagation network (BPN) was used by Yan and Guo [8] for on-line bearing performance degradation assessment. More recently, as a new area of machine learning, various deep learning algorithms, such as deep belief networks (DBNs) [9], convolutional neural network [10] and deep neural networks (DNNs) [11] have been applied successfully and rapidly in many different fields. Feng et al. [6] utilized DNNs to implement both fault feature extraction and intelligent diagnosis. Discrete wavelet transform and neural network (NN) were adopted by Li et al. [12] to detect gearbox crack faults. Moreover, Meng et al. [13] proposed a novel hierarchical diagnosis network by collecting DBNs by layer for the hierarchical identification of mechanical systems.

In addition to assessment methods, extracting the representative features of the vibration signal collected from the target machines is another crucial task. There are two kinds of conventional methods for extracting vibration signal features: time domain analysis and frequency domain analysis, which include correlation function analysis, fast Fourier transform (FFT), etc. Recently, some novel analysis techniques, such as wavelet packet transform (WPT) [14], short time Fourier transform (STFT)

[15] and Wigner-Ville distribution (WVD) **[16]** have also been successfully used for the same purpose. However, usually, after a series of transformations the dimensionality of the acquired parameters reaches tens or even hundreds, so it is essential to reduce the dimensionality prior to entering the data into the assessment model. The major idea of dimensionality reduction is that the extracted redundant statistical features in a high-dimensional space are mapped to a few significant features in a low-dimensional space, where these features are used to represent different crack levels **[17]**. Principal component analysis (PCA) is a popular linear dimensionality reduction method that constructs a low-dimensional representation of the data that describes as much of the variance in the data as possible by finding a linear basis of reduced dimensionality for the data, in which the amount of variance in the data is maximal **[18]**. As for non-linear dimensionality reduction methods, in addition to Isometric feature mapping (Isomap), Laplacian Eigenmaps and local linear embedding (LLE) are two other techniques that find a low dimensional data representation by preserving local properties of the manifold **[19]**.

By comparing and analysing the deficiencies of the existing methods, this paper presents a novel combined assessment model (CAM) by combining WPT, Isomap, and DBN to assess the health condition of the target machine; the method includes three steps: (1) extract original features of the vibration signals through time and frequency domain analysis and WPT, (2) reduce the dimensionalities of the high-dimensional arrays by Isomap, (3) enter the acquired low-dimensional arrays composed of significant features into the trained DBN to obtain assessment results. To present a comprehensive comparison and highlight the advantages of the proposed method, two other popular dimensionality reduction methods (PCA and Laplacian Eigenmaps) and two other intelligent algorithms (HMM and back-propagation neural network (BPNN)) have also been introduced to conduct the same experiments.

The rest of the paper is organized as follows: Section 2 presents the proposed combined assessment method (CAM) for assessing machine health, in which all the concepts and algorithms are described clearly. In Section 3, a series of experiments were conducted, including extracting the representative features of the vibration signal and designing the DBN architecture. In Section 4, comparison experiments with the PCA linear dimensionality reduction method and the Laplacian Eigenmaps nonlinear local dimensionality reduction technique, which have commonly been used

for feature extraction were conducted, which proved the advantages of the proposed CAM. The assessment results are analysed in Section 5.

1 THE PROPOSED CAM

1.1 Techniques for Original Features Extraction

Extracting the representative features of the vibration signal collected from the target machines is a crucial task for the work of health assessment. To illustrate the general applicability of the proposed method, both conventional and modern feature extraction methods will be employed.

1.1.1 Time and Frequency Domain Analysis

Time domain analysis directly based on the time-domain signal to analyse and give results is the simplest and most direct analysis method, which is especially effective when the signal contains simple harmonic, periodic or transient pulse components. Time domain analysis mainly includes probability analysis, time domain average synchronous method, correlation function analysis, and analysis of the extracted features of the time domain waveform. Frequency domain analysis converts the signal in the time domain to the frequency domain with the help of Fourier transform and then determines the type and degree of the fault according to the characteristics of the frequency distribution and the trend of the signal. It mainly includes spectrum analysis, cepstrum analysis, and envelopment analysis.

In the proposed method, 11 time-domain original features, such as crest factor, peak value, variance value, etc., and 13 original features, such as standard deviation frequency, spectrum kurtosis, first order centre, etc. are extracted.

1.1.2 WPT

WPT, an extension of wavelet transform (WT) that can provide level-by-level decomposition, is introduced to extract the original features in this study. WPT can decompose a time-domain signal S into several levels of wavelet packet (WP) nodes to construct the structure of a WP tree, in which each level stands for a frequency resolution. The WP coefficients of a function S can be computed via **[20]**:

$$C_{j,k}^i = \left\langle S, W_{j,k}^i(t) \right\rangle, \quad i = 1, 2, ..., \tag{1}$$

where the integers i, j and k in $W_{j,k}^i(t)$ represent the modulation, scale and translation parameters, respectively. Each coefficient of $C_{j,k}^i$ measures a specific frequency band content indexed by the scale parameters i and j. Therefore, the definition of wavelet packet node energy can be:

$$e_{j,i} = \Sigma\left(C_{j,k}^i\right)^2, \qquad (2)$$

which measures the signal energy contained in some specific frequency band controlled by i and j. Each (i,j) will be referred to as a wavelet packet node in the following. Each value of the node energy of wavelet packet can be regarded as an individual feature element and serve as a robust preliminary exploration of the specific features of the machine vibration signals, from which the useful information for health assessment purposes could be abstracted.

After selection, 14 WPT original features are extracted in this paper.

1.2 Isomap Dimensionality Reduction

The sample data of high dimensional space (D dimensional) is in a low dimensional manifold (L dimensional, $L \leq D$); the manifold structure retains the geometric characteristics of the original data, and L is the intrinsic dimensionality of the sample [21]. As a nonlinear global dimensionality transformation technique, Isomap finds a mapping solution after a series of conversions through which the geodesic distance between input points in the original space can be represented by a Euclidean distance in the projection space as accurately as possible [22], of which the effect can be generally expressed as:

$$M^D \overset{Isomap}{\Rightarrow} M^L, \left(L \leq D\right),$$

where M^D and M^L stand for the D-dimensional original features and projected L-dimensional features, respectively. The main solution procedures are summarized as follows:

Step 1: Establish matrix T which is composed of the geodesic distances between each two data points in M^D according to the neighborhood graph P and shortest-path algorithms proposed by Dijkstra [23], where $M^D = [o_1, o_2, ..., o_M]^T$.

Step 2: Update the entries g_{ij} and double-center matrix G by Eq. [9], where $G = T^2$, $T\left(o_i, o_j\right)$ denotes shortest path between o_i and o_j in graph P.

$$g_{ij} = -\frac{1}{2}\left(g_{ij} - \frac{1}{N}\sum_l g_{il} - \frac{1}{N}\sum_l g_{jl} + \frac{1}{N^2}\sum_{lm} g_{lm}\right). \quad (3)$$

Step 3: Solve the eigenproblem in Eq. (4) and form the columns of matrix V with d primary eigenvectors $v_k(1 \leq k \leq L)$ that correspond to d maximal eigenvalues $\delta_k(1 \leq k \leq L)$, respectively.

$$Gv = \delta v. \qquad (4)$$

Step 4: Compute the entries p_i of matrix M^L which is the low-dimensional representation of M^D by:

$$p_i = \sqrt{\delta_k}\, v_k^i, \qquad (5)$$

where v_k^i is the ith element of the kth primary eigenvector of the matrix V.

Step 5: Construct L-dimensional embedding M^L of the original D-dimensional data matrix M^D with p_i through:

$$M^L = \left[p_1, p_2, \cdots p_M\right]^T.$$

The main procedures can be presented in Fig. 1.
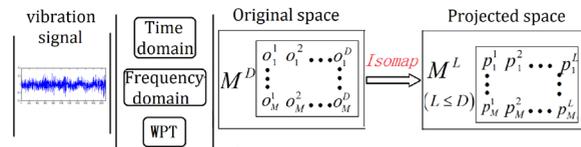
Fig. 1. *Flow chart of Isomap space transformation*

As described in Section 2.1, in this paper, a high-dimensional $n \times 38$ original features array is obtained from the vibration signal. To obtain reasonable dimensionality, a maximum likelihood estimation (MLE) algorithm is introduced to evaluate the intrinsic quantity of the projected dimensionalities before the array is transformed to an $n \times m$ one ($m < 38$) by Isomap.

1.3 DBN-Based Assessment Model

As a probabilistic generative model that can effectively capture the typical information from raw data with various non-linear transformations and approximate complex non-linear functions, DBN is highly appropriate for machine health assessment. Next, the construction and training of DBN model will be introduced.

1.3.1 DBN Architecture

DBN is constructed by stacking a sequence of restricted Boltzmann machines (RBMs) layer by layer [8], as shown in Fig. 2, layer 1 (input layer V) and layer 2 (hidden layer H_1) form RBM$_1$, layer 2 (hidden layer H_1) and layer 3 (hidden layer H_2) form

RBM$_2$, and so on. Each RBM consists of a hidden and a visible layer, both of which are composed of binary stochastic units that only connect with the units from different layer but do not connect with the ones within the same layer.

The energy of a joint configuration of the visible and hidden units is given:

$$E\left(\boldsymbol{v},\boldsymbol{h};\theta\right)=-\sum_{i=1}^{V}\sum_{j=1}^{H}w_{ij}v_{i}h_{j}-\sum_{i=1}^{V}b_{i}v_{i}-\sum_{j=1}^{H}a_{i}h_{j}, \qquad (6)$$

where matrix $\boldsymbol{w}$ denotes the weights between visible and hidden layer, vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are the biases of hidden units h_j and visible units v_i, while $\theta = \{\boldsymbol{w},\boldsymbol{b},\boldsymbol{a}\}$ are the model parameters.
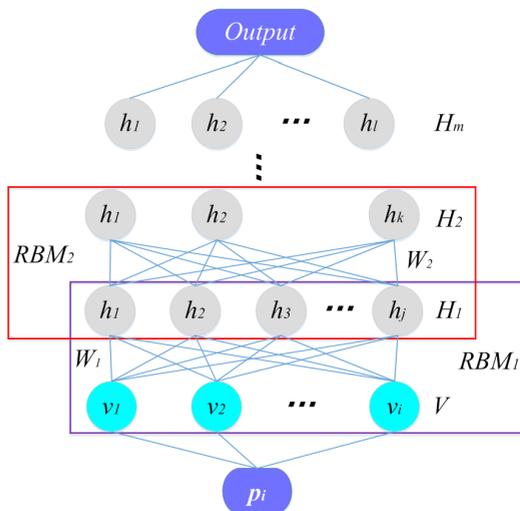
Fig. 2. *Schematic architecture of DBN*

To further clarify the working process of DBN, take single RBM_1 as an example. $\forall i,j,v_i \in \{0,1\}$, $h_j \in \{0,1\}$. Each component of input array X corresponds to a node of the visible layer, by this way X is input into a RBM_1. After a series of calculations, an output array Y of which each component corresponds to a node of the hidden layer is obtained.

1.3.2 DBN Model Training

The distinctive architecture makes it feasible to train DBN by training a series of RBMs with the contrastive divergence (CD) algorithm. The primary training procedure can be summarized as: each RBM layer is trained by using the activation probabilities of the sub-network RBM as the input training data, while the output serves as the input to the next RBM layer. After the unsupervised RBM pre-training, the first layer is fed with the raw input data and is Gaussian-binary restricted Boltzmann machines (GB-RBM) for

real-valued input; the others are binary or Bernoulli–Bernoulli RBM. Finally, the updating rule for the parameter is given:

$$\mathbf{W} \leftarrow \mathbf{W} + \varepsilon_w \left(v_i h_j^{\;0} - v_i h_j^{\;1}\right), \qquad (7)$$

$$\mathbf{a} \leftarrow \mathbf{a} + \varepsilon_a \left(\left\langle h_j \right\rangle^0 - \left\langle h_j \right\rangle^1\right),$$

$$\mathbf{b} \leftarrow \mathbf{b} + \varepsilon_b \left(\left\langle v_i \right\rangle^0 - \left\langle v_j \right\rangle^1\right), \qquad (8)$$

where ε_w, ε_a and ε_b denote the learning rates of weights, hidden bias and visible bias, respectively. Details of the training process can be found in literature [24].

For machine health assessment tasks, after the generative pre-training, other typical discriminative, learning procedures which can effectively fine-tune the weights will be combined to improve the performance of DBN. A very effective way that has been successfully confirmed for implementing discriminative fine-tuning is to add an extra layer of variables that denote the desired labels after the last RBM. Then, the back-propagation algorithm similar to that in the standard back propagation neural network will be introduced to adjust all the network weights.

Multiple samples for training and testing are obtained by dividing the dataset, which are further divided into small "mini-batches" for efficient training [22]. To meet the application requirements, the first 2000 min data N^L of the transformed dataset M^L in projected space, which are all collected under normal conditions are used as training data to train and fine-tune the normal DBN model, that is:

$$\text{Training data:} \quad N^L = \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_N \end{bmatrix} = \begin{bmatrix} \mathrm{p}_1^1 & \cdots & \mathrm{p}_1^L \\ \vdots & \ddots & \vdots \\ \mathrm{p}_N^1 & \cdots & \mathrm{p}_N^L \end{bmatrix}.$$

1.4 Assessment

The specific structure and nonlinear learning procedures of DBN make it very effective for obtaining the intrinsic characteristics from a large number of data. After obtaining the normal DBN model, the entire feature dataset M^L that contains normal data as well as fault data are used as testing data to assess the machine whole life health state, that is:

$$\text{Training data:} \quad M^L = \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_M \end{bmatrix} = \begin{bmatrix} \mathrm{p}_1^L & \cdots & \mathrm{p}_1^L \\ \vdots & \ddots & \vdots \\ \mathrm{p}_M^1 & \cdots & \mathrm{p}_M^L \end{bmatrix},$$

where $N^L \in M^L$.

The principal work for DBN-based machine health CAM can be summarized as follows: (1) Pre-process the original vibration signal data, such as eliminating the abnormal data. (2) Extract features of the datasets with Isomap and divide them into training and testing subsets. (3) Build the assessment model based on deep learning and DBN theories. (4) Train DBN with the training feature data N^L to obtain a normal DBN assessment model. (5) Apply the normal DBN model and the entire dataset M^L for machine health assessment. It should be noted here that steps (3) and (4) will be repeated until the number of epochs is reached. Since being trained with the feature data collected under normal conditions, when the entire dataset is input into the DBN assessment model, its output is the relative probability value of each input belonging to normal condition which is determined by DBN, and its changing trend can reflect the running condition of the tested machine.

Fig. 3 shows the procedures of the proposed CAM.

2 EXPERIMENTS AND ANALYSIS

2.1 Test Rig and Data

The experiment executed bearings run-to-failure tests un-der permanent load was conducted on a special test rig, as shown in Fig. 4, in order to make comparisons with the existing methods and highlight the advantages of the proposed method in this paper; the test data from literature [25] are adopted for the following further research. There are four double row bearings that are force lubricated in the test rig, and the rotation speed of the shaft is applied onto a radial load was kept constant by an alternating current (AC) motor coupled to it via rub belts.

The parameters and operation condition of the bearings are displayed in Table 2. To collect the vibration signal effectively, each bearing was installed with a PCB 353B33 high sensitivity quartz ICP accelerometer on the bearing housing. There are four channels: channels 1 to 4 correspond to bearings 1 to 4, respectively.

Collection of the data was facilitated with NI DAQ Card 6062E. When the test-to-failure experiment was completed, a dataset which consists of 984 individual vibration signal data files was obtained. The files of ASCII format contain all the information of the four channels. Each file contains 20480 data points with the sampling rate of 20 kHz, and the recording interval of the files is 10 minutes. At the end of the test experiment, failure occurred in the outer race of bearing 1.
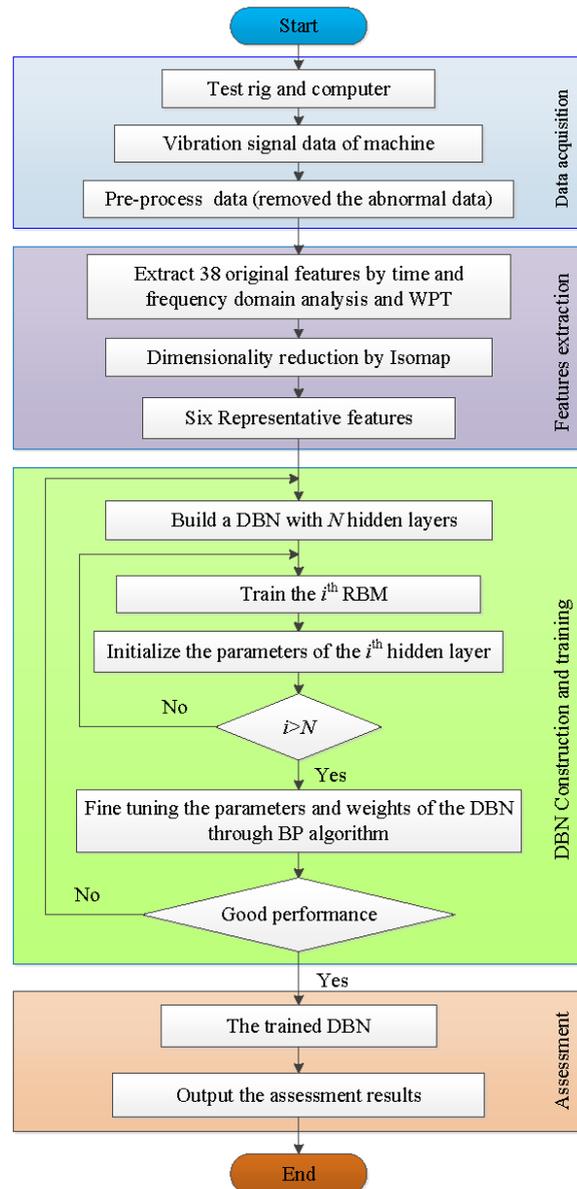
Fig. 3. *Flowchart of the proposed CAM*

Table 2. *Bearing parameters and operation condition*

Type	Number	Ball diameter [mm]	Contact angle [deg]	Rotation speed [RPM]	Radial load [kN]
ZA-2115	4	10	0	2000	26.671
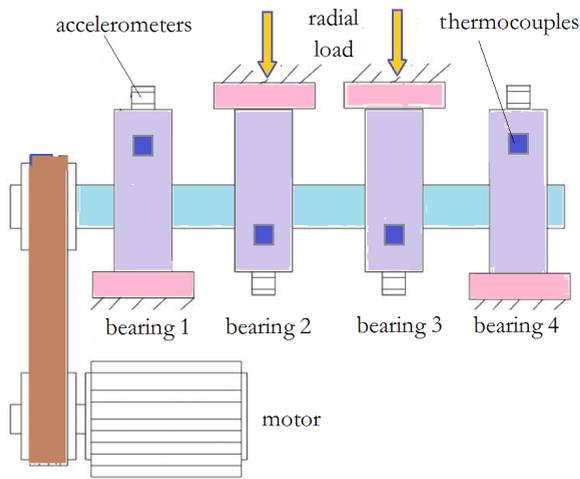
Fig. 4. *Test rig*

2.2 Features Extraction

In view of the outer race failure occurred in bearing 1 at the end of the test experiment, the data of channel 1 were chosen for subsequent experiments. Firstly, pre-processing the data of the bearing vibration signal datasets was carried out, during which three documents that were detected to have abnormal data points were removed. Then the techniques of time domain analysis, frequency domain analysis, and WPT were employed to extract the 38 original features, and a 9810×38 original features array was obtained. Next, the intrinsic quantity of the dimensionalities was calculated with the algorithm of MLE, and the answer was 6. Then the Isomap non-linear dimensionality reduction method was employed to further extract the more representative features; thus, the 9810×38 high-dimensional original features array was converted into a 9810×6 low-dimensional one, which contains more representative features. Fig. 5 indicates that four features (features 1, 2, 3 and 5) appeared to become irregular at about 5300 and all these 6 features have a significant mutation around the point of 7000. The phenomena described above may suggest that the bearing tested in the experiment could have a fault in the vicinity of these abnormal points.

Although extracted from the same test dataset, the six feature trends in Fig. 5 still exhibit significant inconsistency; therefore, these features are not supposed to be applied individually for machine health assessment as they are not able to provide the consistent and accurate degradation pattern. A more feasible method is to construct an effective model that can fuse all the information of these representative features to obtain reliable and accurate machine health assessment results.

In the next experiments, the 9810×6 features array obtained above will be used to achieve two functions: the former 2000×6 subpart will be applied for training and tuning the assessment model, while the whole dataset will be used to test the model and get the health assessment results of the machine during the entire test period.

2.3 Design of DBN Structure

The numbers of input, hidden nodes, and hidden layers are the most critical parameters in the design of neural networks. As for the DBN-based assessment model, the number of input nodes corresponds to the number of the dimensionality of the array composed of extracted features data, of which the value is 6. Since the one-step-ahead assessment method is chosen, one output node is enough to meet the requirements.
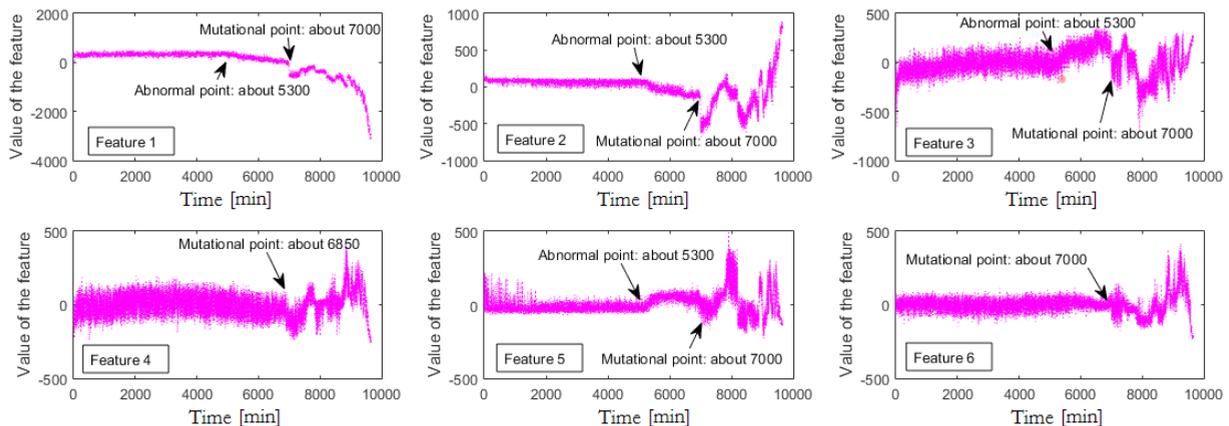
Fig. 5. *The 6 features extracted by Isomap-based dimensionality reduction*

The number of hidden nodes determine the ability of neural networks to capture nonlinear patterns from the input data. On the one hand, neural networks with too few hidden nodes may not be competent enough to model the data. On the other hand, too many hidden nodes may bring about over-fitting problems and result in inferior assessment performance. As a generative neural network model with multiple hidden layers, DBN is more powerful in acquiring the complicated relationship from the input data. Therefore, the changeable number of hidden layers and hidden nodes provides the designer of DBN a significant amount of freedom, which also brings many problems. To date, there has been no mature means in theory for determining the number of hidden layers and nodes, which remains an intractable task for the design of DBN structure.

In this paper, the numbers of hidden layers and nodes of the DBN were selected by a series of experiments. The data for training DBN were divided into two parts: 80% of them were applied for training, while the rest for validation. Given that the numbers of hidden layers of DBNs [13] that had been successfully applied to solve various problems are all more than two, hence, the constructed DBN was initialized with three hidden layers. After several attempts, it was discovered that when the hidden nodes of the three hidden layers are set as 100–50–50, respectively, the evaluation result would be better than other design patterns. Analogously, after comparing with the same analysis method, the architectures of the DBNs with 4, 5 and 6 hidden layers that realized the most satisfactory evaluation results are 100–100–50–20, 100–100–50–50–10 and 100–100–100–50–20–20, respectively. The assessment outputs of these optimal architectures of the DBNs described above are displayed in Fig. 6, which shows that the DBN with 4 hidden layers could acquire better assessment
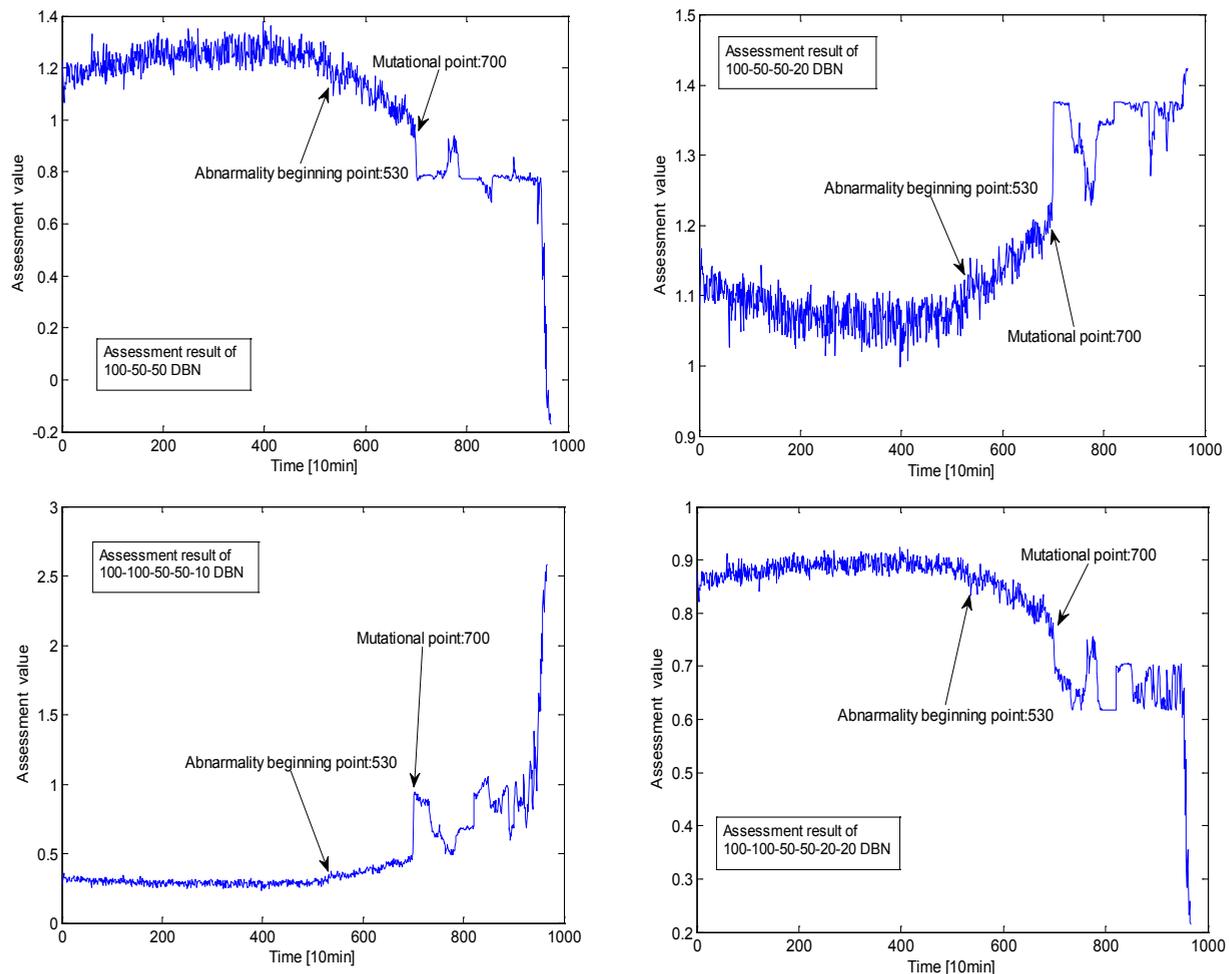
Fig. 6. *The assessment results of the 4 different optimal DBNs*

result than a 3-hidden-layer one, for example, the waveform of the former is more stable than that of the latter before the mutational point, and its mutation at time point 700 is more obvious. While the DBN with the architecture of 100–100–50–50–10 performs best in the experiment by virtue of its smoothest, clear and reasonably trended curve, with the number of hidden layers increasing to 6, the performance of DBN had become worse.

The phenomenon above can be explained as that DBN with less than five hidden layers may not have enough power to model the data; however, when there are more than 5 hidden layers the over-fitting problems may result in inferior assessment performance. It is, therefore, the conclusion that the 5-hidden-layer DBN with the architecture of 100–100–50–50–10 is the best model for the work of machine health assessment. (It is worth noting that since the value of each point of the waveforms is a relative one, the direction the waveforms trended (up or down) does not affect the quality of the assessment results.)

3 COMPARISON EXPERIMENTS

3.1 Comparison Experiments of Dimensionality Reduction

Isomap, adopted in this paper, is a nonlinear global dimensionality reduction method; next, the comparison experiments with linear dimensionality reduction method PCA and nonlinear local dimensionality reduction technique Laplacian Eigenmaps that have been used for feature extraction were conducted.

In order to make fair and reasonable comparisons, the most suitable DBN architectures for PCA and Laplacian Eigenmaps are constructed, which are 100–100–50–50–5 and 100–100–50–50–10, respectively. The assessment results of DBNs with these two different dimensionality reduction methods are
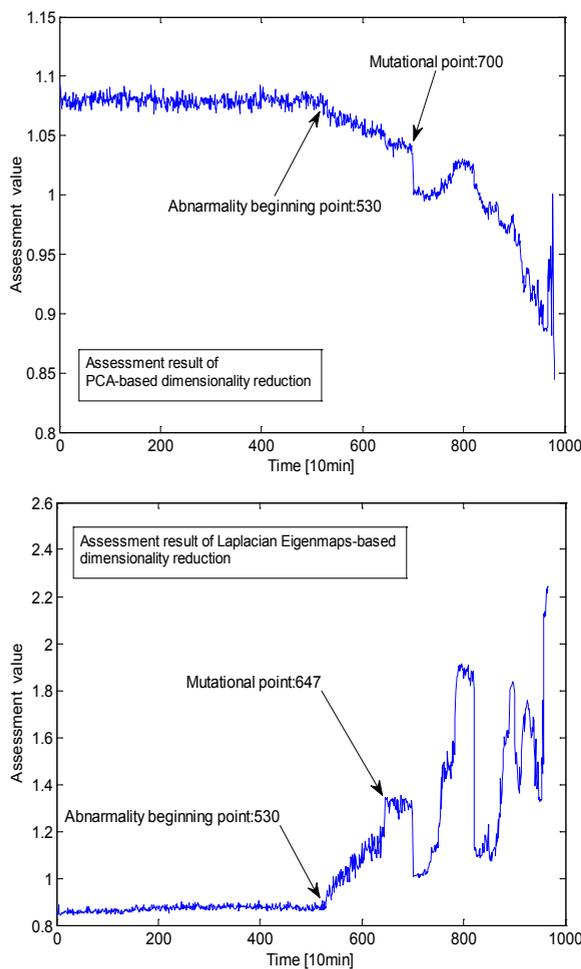
Fig. 7. *The assessment results of PCA and Laplacian Eigenmaps-based dimensionality reduction method*
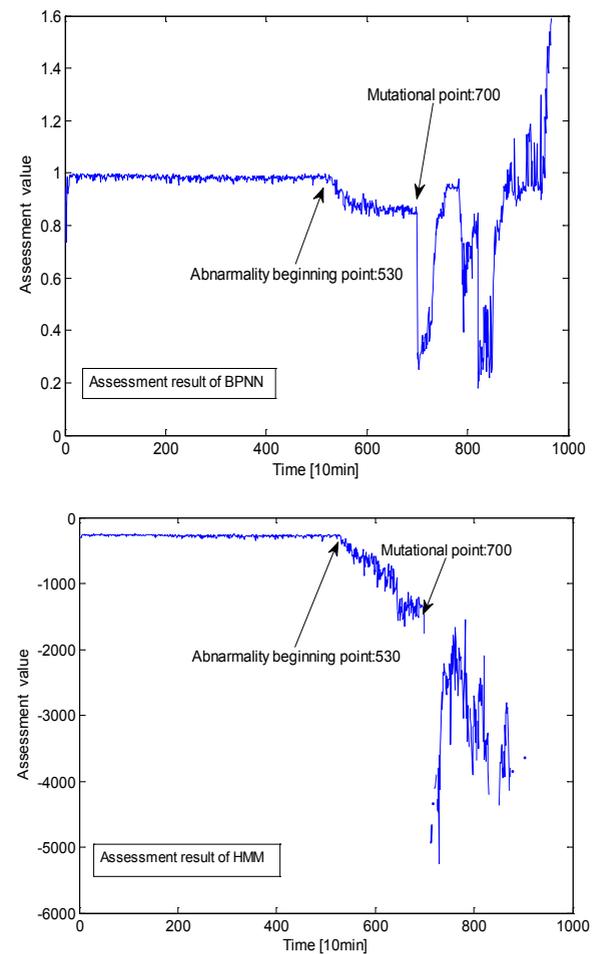
Fig. 8. *The assessment results of BPNN and HMM assessment models*

displayed in Fig. 7, from which it can be determined that the resulting waveform of PCA-based method has the same abnormality beginning point (at time point 530) and mutational point (at time point 700) with that of the Isomap-based method, but the former fluctuates much more severely before abnormality beginning point and becomes disorganized at the end. While, the resulting waveform of Laplacian Eigenmaps -based method performs better for the two problems mentioned above, but the overly bizarre mutational point (at time point 647) also demonstrated the defect of this method. The analysis and comparisons above surely proved the correctness of the Isomap-based dimensionality reduction method.

3.2 Comparison Experiments of Assessment Models

Next, two other kinds of intelligent algorithms HMM and BPNN that have already been successfully and widely applied in fault diagnosis, data classification, are also introduced to conduct the similar experiments under the corresponding same conditions. As an extension of Markov chains, HMM that contains a finite number of states, where each state generates an observation at a certain time point, is a state-of-the-art technique for model recognition due to its elegant mathematical structure and the availability of computer implementation [1]. While a BPNN model is composed of many idealized layers of nodes and specified by the node characteristics (weights), the learning rules (transfer functions, always the sigmoid function), network interconnection geometry (different layers), and dimensionality (number of layers and nodes), of which the learning feeds back into the model to change the weights of nodes between layers in order to decrease errors between predicted and measured values [26].

In order to make the results of the experiments more persuasive, the dimensionality reduction method is kept the same as Isomap only the assessment models were replaced by BPNN and HMM, respectively. The assessment results of these two models are shown in Fig. 8, which indicates that BPNN can accurately identify the anomaly and mutation at time points 530 and 700, respectively, but it performs disorderedly after the mutational point when the waveform trends to the opposite direction. While the waveform of HMM is very smooth at the beginning, it drops too fast after the abnormality beginning point 530, which does not accord with the actual situation, and the transformation at mutational point 700 is too vigorous so that the waveform graph disconnected there; additionally, its tail waveform is also very confusing.

4 ASSESSMENT RESULTS

After a series of experiments, in which the comparisons with two popular dimensionality reduction methods (PCA and Laplacian Eigenmaps-based) and two intelligent assessment algorithms (HMM and BPNN-based), the proposed CAM in this paper, which employs Isomap and the DBN with architecture of 100–100–50–50–10, was verified to be very effective and accurate for machine health assessment work.
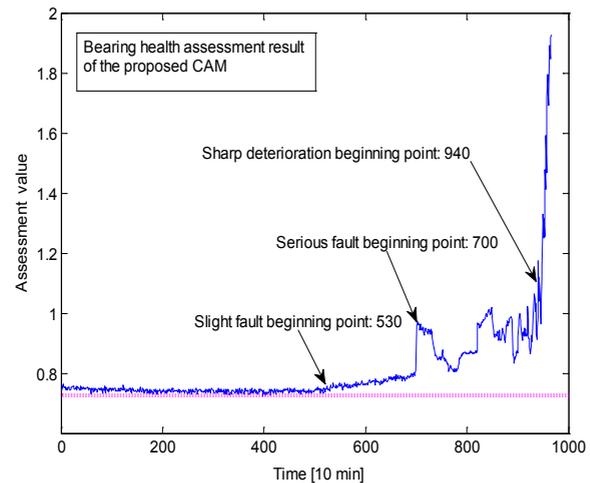
Fig. 9. *Assessment result of the proposed CAM*

As shown in Fig. 9, after a long time of normal running, the proposed CAM detected the early slight degradation of the bearing at time point 530, indicating where the slight faults of wear, pitting or overheat may begin occurring. Then the CAM monitored a mutation signal (at point 700) which presents that crackle, fatigue spalling or some other serious faults might occur there. Finally, it can be observed from the end of the waveform that after the time point 940, the condition of the bearing began to deteriorate so sharply that it could not continue working.

Compared with the real test situations, the assessment results of the proposed method in this paper are very much in line with the actual running conditions of the bearings.

5 CONCLUSIONS

Machine health assessment is playing an increasingly important role, which can provide diverse benefits including improved safety, improved reliability and reduced costs for operation and maintenance of complex manufacturing systems. This paper presents a novel combined assessment method (CAM) for

assessing the health condition of the target machine (rolling-element bearings). To deal with the non-stationary property of the vibration signals collected from machines, three widely used techniques involving time domain analysis, frequency domain analysis, and WPT are adopted to extract 38 original features of the datasets. Then, the Isomap nonlinear global algorithm is adopted for dimensionality reduction and extracting more representative features. Next, the acquired low-dimensional features array is input into the well-trained DBN model to evaluate the performance status of the bearing. Finally, the bearing-accelerated degradation data from Cincinnati University were investigated for further research, through the comparison experiments with two popular dimensionality reduction methods (PCA and Laplacian Eigenmaps-based), and two intelligent assessment algorithms (HMM and BPNN-based), the proposed CAM is proved more sensitive to the incipient fault of rolling bearings and more effective for bearing performance degradation evaluation. In future work, CAM can be applied in other fields for evaluation or classification.

6 ACKNOWLEDGMENT

7 REFERENCES

[1] Yu, J. (2012). Health condition monitoring of machines based on hidden Markov Model and con-tribution analysis. *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 8, p. 2200-2211, **DOI:10.1109/TIM.2012.2184015**.

[2] Loutridis, S. (2006). Instantaneous energy density as a feature for gear fault detection. *Mechanical Systems and Signal Processing*, vol. 20, no. 5, p. 1239-1253, **DOI:10.1016/j.ymssp.2004.12.001**.

[3] Öztürk, H., Sabuncu, M., Yesilyurt, I. (2008). Early detection of pitting damage in gears using mean frequency of scalogram. *Journal of Vibration and Control*, vol. 14, no. 4, p. 469-484, **DOI:10.1177/1077546307080026**.

[4] Loutridis, S.J. (2008). Self-similarity in vibration time series: application to gear fault diagnostics. *Journal of Vibration and Acoustics*, vol. 130, no. 3, p. 569-583, **DOI:10.1115/1.2827449**.

[5] Yu, D., Yang, Y., Cheng, J. (2007). Application of time–frequency entropy method based on Hilbert–Huang transform to gear fault diagnosis. *Measurement*, vol. 40, no. 9-10, p. 823-830, **DOI:10.1016/j.measurement.2007.03.004**.

[6] Feng, Z., Zuo, M.J., Chu, F. (2010). Application of regularization dimension to gear damage assessment. *Mechanical Systems and Signal Processing*, vol. 24, no. 4, p. 1081-1098, **DOI:10.1016/j.ymssp.2009.08.006**.

[7] Do, V.T., Nguyen, L.C. (2016). Adaptive empirical mode decomposition for bearing fault detection. *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 62, no. 5, p. 281-290, **DOI:10.5545/sv-jme.2015.3079**.

[8] Yan, J., Guo, C. (2011). A dynamic multi-scale Markov model based methodology for remaining life prediction. *Mechanical Systems and Signal Processing*, vol. 25, no. 4, p. 1364-1376, **DOI:10.1016/j.ymssp.2010.10.018**.

[9] Tran, V.T., Yang, B.-S., Oh, M.-S., Tan, A.C.C. (2009). Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference. *Expert Systems with Applications*, vol. 36, no. 2, p. 1840-1849, **DOI:10.1016/j.eswa.2007.12.010**.

[10] Hinton, G.E., Salakhutdinov, R.R. (2006). Reducing the dimensionality of data with neural networks. *Science*, vol. 313, no. 5786, p. 504-507, **DOI:10.1126/science.1127647**.

[11] Lecun, Y., Bottou, L., Bengio, Y., Haffner P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, vol. 86, no. 11, p. 2278-2324, **DOI:10.1109/5.726791**.

[12] Li, Z., Ma, Z., Liu, Y., Teng, W., Jiang, R. (2015). Crack fault detection for a gearbox using discrete wavelet transform and an adaptive resonance theory neural network. *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 61, no. 1, p. 63-73, **DOI:10.5545/sv-jme.2014.1769**.

[13] Gan, M., Wang, C., Zhu, C. (2016). Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings. *Mechanical Systems and Signal Processing*, vol. 72-73, p. 92-104, **DOI:10.1016/j.ymssp.2015.11.014**.

[14] He, Q.B. (2013). Vibration signal classification by wavelet packet energy flow manifold learning. *Journal of Sound and Vibration*, vol. 332, no. 7, p. 1881-1894, **DOI:10.1016/j.jsv.2012.11.006**.

[15] Klein, R., Ingman, D., Braun, S. (2001). Non-stationary signals: phase-energy approach theory and simulations. *Mechanical Systems and Signal Processing*, vol. 15, no. 6, p. 1061-1089, **DOI:10.1006/mssp.2001.1398**.

[16] Baydar, N., Ball, A. (2001). A comparative study of acoustic and vibration signals in detection of gear fail-ures using Wigner–Ville distribution. *Mechanical Systems and Signal Processing*, vol. 15, no. 6, p. 1091-1107, **DOI:10.1006/mssp.2000.1338**.

[17] Wan, X., Wang, D., Tse, P.W. Xu, G., Zhang, Q. (2016). A critical study of different dimensionality reduction methods for gear crack degradation assessment under different operating conditions. *Measurement*, vol. 78, p. 138-150, **DOI:10.1016/j.measurement.2015.09.032**.

[18] Wang, H. (2012). Block principal component analysis with L1-norm for image analysis. *Pattern Recognition Letters*, vol. 33, no. 5, p. 537-542, **DOI:10.1016/j.patrec.2011.11.029**.

[19] Borchers, B., Young, J.G. (2007). Implementation of a primal-dual method for SDP on a shared memory parallel architecture. *Computational Optimization and Applications*, vol. 37, no. 3, p. 355-369, **DOI:10.1007/s10589-007-9030-3**.

[20] Hemmati, F., Orfali, W., Gadala, M.S. (2016). Roller bearing acoustic signature extraction by wavelet packet transform,

applications in fault detection and size estimation. *Applied Acoustics*, vol. 104, p. 101-118, **DOI:10.1016/j.apacoust.2015.11.003**.

[21] Hauberg, S: (2015). Principal curves on Riemannian manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1915-1921.

[22] Benkedjouh, T., Medjaher, K., Zerhouni, N., Rechak, S. (2013). Remaining useful life estimation based on nonlinear feature reduction and support vector regression. *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, p. 1751-1760, **DOI:10.1016/j.engappai.2013.02.006**.

[23] Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, vol. 1, no. 1, p. 269-271, **DOI:10.1007/BF01386390**.

[24] Tran, V.T., AlThobiani, F., Ball, A. (2014). An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks. *Expert Systems with Applications*, vol. 41, no. 9, p. 4113-4122, **DOI:10.1016/j.eswa.2013.12.026**.

[25] Lee, J., Qiu, H., Yu, G., Lin, J. (2007). Rexnord Technical Services: Bearing Data Set. Moffett Field, University of Cincinnati. NASA Ames Prognostics Data Repository, NASA Ames *from: http://data-acoustics.com/measurements/bearing-faults/bearing-4/*, accessed on 2016-10-20.

[26] Meng, Z., Xu, Y., Zheng, Y., Zhu, Y., Jia, Y., Chen, S. (2014). Inversion of lunar regolith layer thickness with CELMS data using BPNN method. *Planetary and Space Science*, vol. 101, p. 1-11, **DOI:10.1016/j.pss.2014.05.020**.