DOI: 10.3901/CJME.2013.05.1041, available online at www.springerlink.com; www.cjmenet.com; www.cjmenet.com.cn

# Feature Extraction of Kernel Regress Reconstruction for Fault Diagnosis Based on Self-organizing Manifold Learning

CHEN Xiaoguang[1], LIANG Lin[1, *], XU Guanghua[1, 2], and LIU Dan[1]

1 *School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China*
2 *State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China*

Received August 24, 2012; revised March 13, 2013; accepted March 15, 2013

**Abstract:** The feature space extracted from vibration signals with various faults is often nonlinear and of high dimension. Currently, nonlinear dimensionality reduction methods are available for extracting low-dimensional embeddings, such as manifold learning. However, these methods are all based on manual intervention, which have some shortages in stability, and suppressing the disturbance noise. To extract features automatically, a manifold learning method with self-organization mapping is introduced for the first time. Under the non-uniform sample distribution reconstructed by the phase space, the expectation maximization(EM) iteration algorithm is used to divide the local neighborhoods adaptively without manual intervention. After that, the local tangent space alignment(LTSA) algorithm is adopted to compress the high-dimensional phase space into a more truthful low-dimensional representation. Finally, the signal is reconstructed by the kernel regression. Several typical states include the Lorenz system, engine fault with piston pin defect, and bearing fault with outer-race defect are analyzed. Compared with the LTSA and continuous wavelet transform, the results show that the background noise can be fully restrained and the entire periodic repetition of impact components is well separated and identified. A new way to automatically and precisely extract the impulsive components from mechanical signals is proposed.

**Key words:** feature extraction, manifold learning, self-organize mapping, kernel regression, local tangent space alignment

## 1 Introduction

Feature extraction is one of the critical components in mechanical fault diagnosis, which is developed through vibration signal analysis and processing. At the viewpoint of identification of conditions and faults, a good feature extraction method should be capable to distinguish different conditions and faults effectively, and helpful to simplify the work of pattern recognition. However, since vibration signals of machinery with fault are often nonlinear and nonstationary, it is difficult to extract effective features. In time domain and frequency domain analysis, statistical indexes, such as peak amplitude, root mean square amplitude, kurtosis and frequency components, are applied to fault diagnosis of rotating machinery[1]. Obviously, these indexes simplify the description of the machine condition, but the selection of index directly affects the pattern recognition. Moreover, most frequency spectrums have similar characteristics, which make misjudgements in detecting machine faults.

Time-frequency analysis can supply both time and frequency information. There are several time-frequency analysis methods, such as orthogonal wavelets, Morlet wavelet transform[2–4]. The selection of time-frequency information still depends on transcendent knowledge and manual intervention. If there is no manual intervention and all the time-frequency components are directly used for pattern recognition, the fault recognition can be seriously influenced by high dimension or strong noise disturbance.

In addition, there are many non-linear factors such as loads, clearance, friction and so on making distinct influences on the signals. The fractal dimensions presented by YANG, et al[5], is adopted to describe the nonlinear behavior of rolling element bearing. The approximate entropy[6] is adopted as a diagnostic tool for machine health monitoring. However, simple parameters can only reflect the overall irregularity of signals, which fail to reveal the details of fractal structures.

Dimensionality reduction is proposed to solve this problem. The traditional dimensionality reduction methods such as principal component analysis and independent component analysis[7–8] are effective only on nonlinear structures. It is difficult to use these methods to discover the nonlinear structure. Since manifold learning concept was proposed and applied successfully by SEUNG, et al[9], in 2000, manifold algorithm has become an active research topic in the field of dimensionality reduction and intelligent

* Corresponding author. E-mail: lianglin@mail.xjtu.edu.cn
This project is supported by National Natural Science Foundation of China(Grant No. 51075323)

pattern recognition. The current manifold algorithms mainly include isometric feature mapping[9], locally linear embedding[10], Laplacian Eigenmaps[11], local tangent space alignment[12], etc. Except for being applied to image processing, manifold learning is also used for fault diagnosis[13–14]. However, how to choose the neighborhood freely without worrying about the topologically instability is a continuous pending problem. The $k$-neighborhood and $\varepsilon$ -neighborhood methods are influenced by curvature, sampling density, and noises. Some extended methods have been proposed to resolve this problem[15–16]. Nevertheless, these methods failed to deal with the sample data without manual intervention.

For the above mentioned reasons, this paper proposes a novel method to extract fault features automatically based on self-organizing manifold learning(SOML). The novelty of the proposed method is combining self-organization neighborhood selection with local tangent space alignment algorithm. Under the non-uniform sample distribution, it can divide the local neighborhoods self-adaptively. Combined with dynamics trajectory of phase space, the proposed method can achieve low-dimensional manifold extraction without manual intervention. Based on the kernel regression of low-dimensional manifold, the signal can be reconstructed.

This paper is organized as follows: the fundamental theories and algorithm of the self-organizing manifold learning are presented in section 2. The feature exaction scheme based on self-organizing manifold learning is given and analyzed in section 3. In section 4, simulated analog signal is adopted to validate effectiveness. In section 5, the fault application is presented to demonstrate the effect of the proposed strategies. Finally, some conclusions are drawn in section 6.

## 2 Overview of SOML Algorithm

### 2.1 Adaptive neighborhood selection

It is obvious that large neighborhoods cause confusions in dealing with the highly twisted and folded manifold. In contrast, small neighborhoods can falsely estimate the relationships between the neighbors, even if the continuous manifold is divided into disjoint sub-manifolds. Due to added noise, the distribution of samples in feature space is usually non-uniform. Therefore, to the sparse distribution of sample data, the neighborhood size must be large enough to ensure overlaps between neighbors. On the other hand, small neighborhoods often represent complex nonlinear distribution. Therefore, the fixed sizes of neighborhoods cannot satisfy the changing manifold structures. Therefore, it is inevitable that the neighborhood size should be selected adaptively with two principles: (1) All of sub-spaces divided by neighborhoods can be connected to construct the topology structure of manifold; (2) There should be enough overlaps between adjacent neighbors, in order to transmit the local information.

In the adaptive learning methods, self-organizing mapping(SOM) has the ability to divide network nodes adaptively. Using competing-layer neurons to match the center of local neighbors of manifold structures, SOM organize node grids to cover the whole topological structures. Then the local neighbors of high-dimensional manifolds are divided adaptively. Considering the above requisitions, the EM iteration algorithm is adopted to divide local neighbors adaptively[17].

### 2.2 SOML algorithm

The process of the algorithm is shown in Fig. 1, and the implementation procedures of the SOML are detailed as follows.

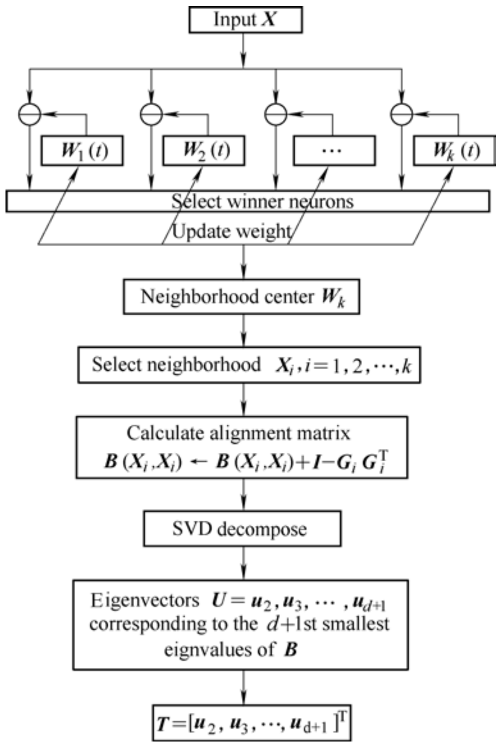


Fig. 1. Schematic diagram of SOML algorithm

(1) Given a set of inputs $\boldsymbol{X}$, the SOM network is adopted to optimize weight $\boldsymbol{W}$ with EM iteration algorithm. Let $p_r^i$ denote the probability that input $x_i$ is assigned to the node with $\boldsymbol{w}_r$. It is constrained by $\sum_r p_r^i = 1$ and $p_r^i$>0. There is a neighborhood function $h_{r,s}$ that corresponds to the control strength between node $r$ and node $s$. Usually it is a decreasing function of the distance between nodes $r$ and $s$. Given the data $\boldsymbol{X}$, the optimal goal is to find the probability assignment $\boldsymbol{P}$ and weight $\boldsymbol{W}$ that minimizes

$$E(\boldsymbol{P},\boldsymbol{W}) = \sum_i \sum_r p_r^i \sum_s h_{r,s} \boldsymbol{D}(x_i, \boldsymbol{w}_s) . \quad (1)$$

So the neighborhood function is defined as

$$h_{r,s} = \exp\left(\frac{-d(r,s)^2}{2\sigma^2}\right), \quad (2)$$

where $d(r; s)$ is the distance between nodes of network, and $\boldsymbol{D}$ is the distance between inputs and weights. The initial weight matrix $\boldsymbol{W}$ is given by

$$\boldsymbol{W} = \mathrm{diag}\left(\left(\sum_q \boldsymbol{Q}\right)^{-1}\right)\boldsymbol{QX}, \tag{3}$$

where $\boldsymbol{Q}=[q_1,\cdots,q_r]$ and $q_i$ is a random value between 1 and $r$. Thus, the location coordinate of topology node is set to element of weight $\boldsymbol{W}$. Therefore, with the EM iteration, weight $\boldsymbol{W}$ is calculated using the prior value of the parameters and then fixed in the maximization for the new value.

(2) Each element of $\boldsymbol{W}$ is set to the center node of local neighbors. To ensure enough overlap, the radius of neighbor is equal to the half of maximal distance between center nodes. According to radius of neighbor, the local neighbors $\boldsymbol{X}_i$ are selected, where $i = 1, 2, \cdots, k$ and $k$ is the number of topology grids.

(3) Compute the $d$ largest eigenvectors $\boldsymbol{g}_1,\cdots, \boldsymbol{g}_d$ of the correlation matrix $(\boldsymbol{X}_i-\overline{\boldsymbol{X}}_i\mathrm{e}^{\mathrm{T}})^{\mathrm{T}}(\boldsymbol{X}_i-\overline{\boldsymbol{X}}_i\mathrm{e}^{\mathrm{T}})$, and set

$$\boldsymbol{G}_i=[\boldsymbol{e}/\sqrt{k},\boldsymbol{g}_1,\cdots,\boldsymbol{g}_d], \tag{4}$$

where $\overline{\boldsymbol{X}}_i$ is the mean of $\boldsymbol{X}_i$.

(4) The alignment matrix $\boldsymbol{B}$ can be computed by carrying out a partial local summation as follows:

$$\boldsymbol{B}(\overline{\boldsymbol{X}}_i,\overline{\boldsymbol{X}}_i)\leftarrow \boldsymbol{B}(\overline{\boldsymbol{X}}_i,\overline{\boldsymbol{X}}_i)+\boldsymbol{B}_i. \tag{5}$$

(5) Compute the 2 to $d$+1 smallest eigenvectors of $\boldsymbol{B}$ and pick up the eigenvector matrix ($\boldsymbol{u}_2, \boldsymbol{u}_3, \cdots, \boldsymbol{u}_{d+1}$), then set $\boldsymbol{T}=(\boldsymbol{u}_2, \boldsymbol{u}_3, \cdots, \boldsymbol{u}_{d+1})^{\mathrm{T}}$.

### 2.3 Robustness analysis

Through the EM iterative process, $\boldsymbol{W}$ approaches to a reasonable topology network which covers the whole structure, and the results satisfy the first principle. On the other hand, it is obvious that the connected distance between the adjacent nodes is used to be the radius of neighborhood, so as to ensure enough overlaps between neighbors. Thus, the second principle of neighborhood selection is also satisfied. Based on the adaptive neighborhood, the low-dimensional embedding can be extracted effectively. The residual variance[18] is adopted to estimate the representation from high-dimensional structure to embedded space, and it is defined as

$$R(k)=1-\rho^2_{\boldsymbol{D_X D_T}}, \tag{6}$$

where $\rho$ is correlation coefficient, $k$ is neighborhood size, $\boldsymbol{D_X}$ and $\boldsymbol{D_T}$ are matrixes of Euclidean distances (between pairs of points) in $\boldsymbol{X}$ and $\boldsymbol{T}$, respectively.

The smaller the residual variance is, the better high-dimensional data are compressed in the embedded space. Since node neighborhood is divided adaptively by the grid of network for competition mechanism, EM iteration can overcome the limitation of fixed neighborhood algorithms and get a better residual variance.

## 3 Feature Extraction Based on SOML

### 3.1 Schematic diagram

For feature extraction in machinery fault diagnosis, the SOML algorithm is adopted to explore the geometric distribution properties embedded in the high-dimensional space. Thus, on the basis of the principles above, a new approach of feature extraction method based on adaptive manifold learning is proposed. First, high-dimensional observation space is built with phase space reconstruction, and then map the space phase data into a low-dimensional feature space by SOML, and estimate the intrinsic distribution of samples to gain the embedding manifold structure. Finally, the signal is reconstructed by kernel regression strategy. The schematic diagram of the feature extraction method based on adaptive manifold learning is shown in Fig. 2.

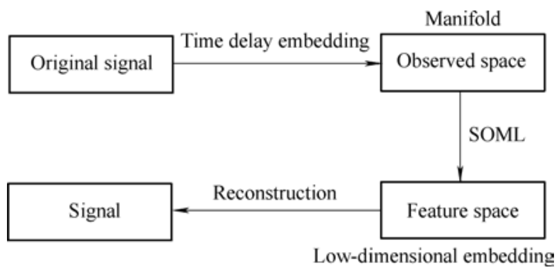


Fig. 2. Schematic diagram of feature extraction strategy

### 3.2 Observed space reconstruction

Generally speaking, $m$-dimensional delay embedding space is equivalent to the original unobserved state space of the dynamics system. Taking the long-time steady-state response of the variable $\boldsymbol{x}$ and performing the time delay embedding transformation lead to the trajectories in the embedding space which comprise a geometric object called a manifold. Given proper parameters, a manifold can be reconstructed by embedding a time series into a high-dimensional space, in which the topological structure and nonlinear characteristics hidden in the low-dimensional time series can be easily extracted. For a time series $x_1,x_2,\cdots,x_N$, the delay vectors in the embedded phase space are given by $\boldsymbol{Y}_i=[x_i,x_{i+\tau},\cdots,x_{i+(m-1)\tau}]$, where $i=1, 2, \cdots, N$, $m$ is the embedding dimension, and $\tau$ is the delay time.

### 3.3 Embedding dimension

At the viewpoint of geometry, the vibration data of machinery in the same state has the same geometric properties in space distribution or topological structure. Their mapping points in the low-dimensional embedded space can be distributed in embedding manifolds or in its neighbors. Obviously the relationship between embedding dimension $d$ of manifold and embedding of dynamics system is defined as

$$d = \min(\boldsymbol{m}), \tag{7}$$

where $\boldsymbol{m}$ is estimated from $E_1(m)$ and $E_2(m)$ curve5 calculated[19].

### 3.4 Signal reconstruction based on kernel regression

Feature extraction is regarded as a mapping from initial data to a feature space with a special function. By manifold learning, an embedding manifold $f:\boldsymbol{T}\subset R^d \to R^m$ can be found. Then the inverse function $F_f(x)$ is considered as a feature extraction function, where the feature $t = F_f(x)$.

For high-dimensional data $\boldsymbol{X}$, an excellent estimate form $F_f(x)$ is provided. Therefore, the purpose is to minimize the reconstruction error:

$$E(f) = \int_{R^n} \left\| x - f(F_f(x)) \right\|^2 p(x)\mathrm{d}x, \tag{8}$$

where $\|\bullet\|^2$ stands for the 2- norm of a matrix, and $f(F_f(x))$ is a reconstruction point of the feature $F_f(x)$ in manifold. It is worth noting that the embedding space has no explicit formulation to correspond the nonlinear model. Thus the signal is reconstructed with statistic regression strategy through the coordinate of low-dimensional embedding. Because the kernel regression[20] adopts the non-parametric statistics to evaluate the probability density function, it is suitable that the signal is reconstructed by the embedding manifold with kernel regression. Thus, for a given kernel function $K(x)$ and sample data $\boldsymbol{Y}$, the reconstruction signal based on the Nadaraya-Watson kernel regression estimator is defined as

$$f(\boldsymbol{t}) = \sum_{i=1}^{N} \frac{K(\boldsymbol{t}-\boldsymbol{t}_i)}{\sum_{j=1}^{N} K(\boldsymbol{t}-\boldsymbol{t}_j)} \boldsymbol{Y}_i \,. \tag{9}$$

As an effective tool, the reconstruction procedure is described as follows: the training sample $\boldsymbol{t}_i$ is selected from low-dimensional embedding $\boldsymbol{T}$, and the corresponding original data $\boldsymbol{Y}_i$ is conformed at the same time. Then the remainder of $\boldsymbol{T}$ is used as the reconstruction sample $\boldsymbol{t}$. According to the $\boldsymbol{Y}_i$ and the $\boldsymbol{t}_i$, the function $f(t)$ is calculated with Eq. (9).

## 4 Experiments Verifications and Discussion

To verify the capability of feature extraction of the proposed method, the nonlinear Lorenz system was adopted for test and described as

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \alpha(y-x) \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = (\gamma - z)x - y, \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = xy - \beta z \end{cases} \tag{10}$$

where parameters $\alpha=16$, $\beta=4$, $\gamma=45.92$, and the 8 dB white noise was also mixed. According to the evaluation index proposed in Eq. (7), $E_1(m)$ and $E_2(m)$ curves were calculated , respectively, where the time delay $\tau=1$. The result is shown in Fig. 3.

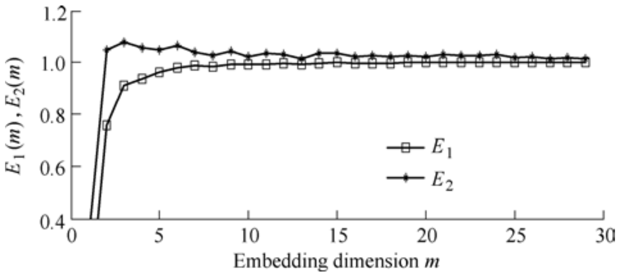


Fig. 3. $E_1(m)$ and $E_2(m)$ curves of the Lorenz system

Of course the minimum embedding dimension is 3, since $E_1(m)$ and $E_2(m)$ curves are slowly increasing when $m > 2$. Two dimension projection of the phase space is shown in Fig. 4, where $m=30$ and $\tau=1$. Due to the mixed noise, the dynamics trajectory is difficult to be identified. The extraction result of Lorenz system by the SOML technique is shown in Fig. 5, which is the projection of three dimension embedding manifold.

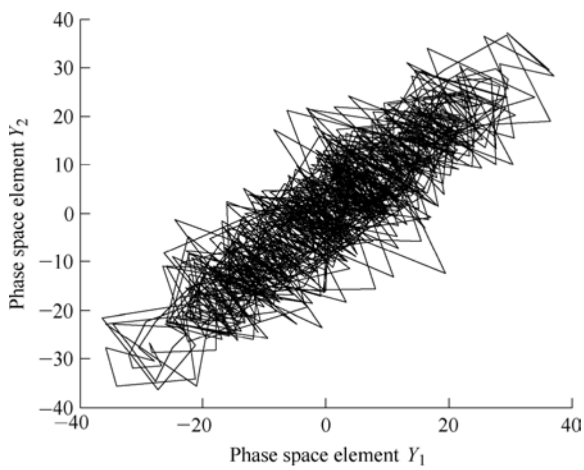


Fig. 4. Phase space of Lorenz system with added noise

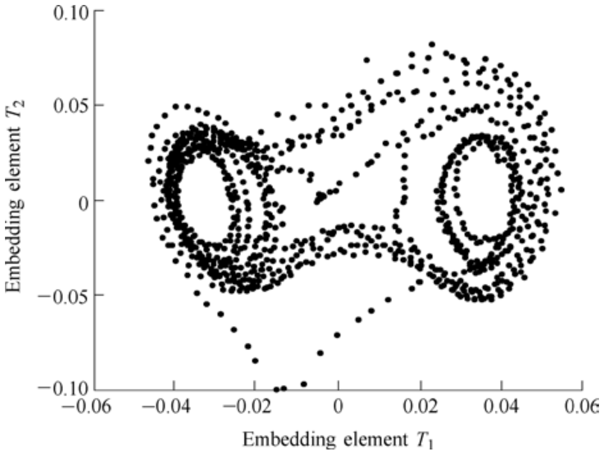


Fig. 5. Projection of three-dimensional embedding manifold

It is found that the embedding manifold structure is similar to the Lorenz system in Fig. 5. Then, the embedding manifold is reconstructed to the phase space, which is shown in Fig. 6. The reconstructed result indicates that the level of noise variance decreased to 10.5 dB, and the dynamics trajectory of chaotic behavor was extracted precisely.

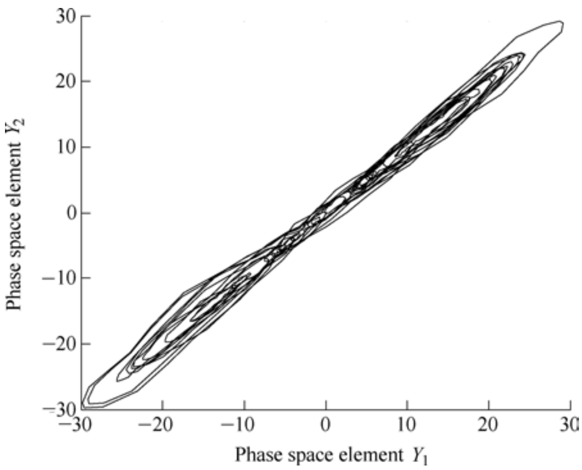


Fig. 6. Reconstructed phase space of the Lorenz system

With the EM iteration, the neighborhood sizes learned from the SOM network is shown in Fig. 7. It is easy to figure out that the neighborhood sizes of nodes fluctuate are between 4 and 78 and the average is 38 due to the varying sampling distribution.

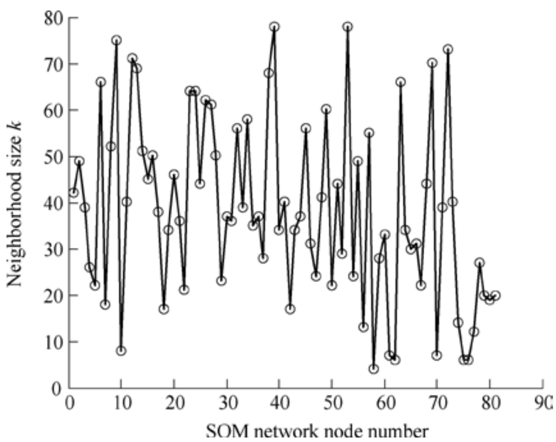


Fig. 7. Neighborhood size learned from SOM network

In order to verify the capability of the proposed method, different neighborhood sizes were adopted by LTSA to extract the low-dimensional embedding, where the neighborhood size $k$ is set to 10, 20, 30, and 40 respectively. The experimental results with different $k$ are shown in Fig. 8. It is clear that the curve of the reconstructed phase space with fixed neighborhood size may reflect the whole structure of dynamics trajectory. However, there are deformations in outline and internal trajectory to some extent. Even if the neighborhood size is equal to the average of adaptive neighborhood size (i.e. $k = 40$), the reconstructed result shown in Fig. 8(d) is still not as good as SOML method displayed in Fig. 6. It indicates that the SOM neighborhood selection strategy is more effective than the fixed neighborhood idea.

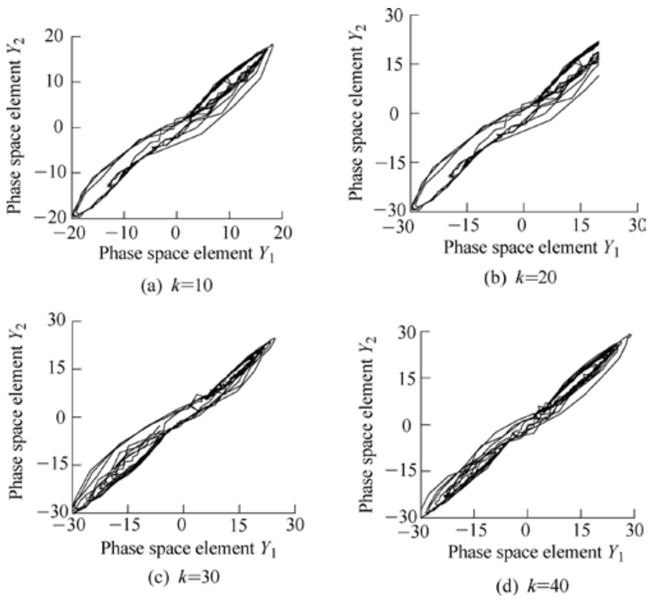


Fig. 8. Reconstructed phase space with different neighborhood

To quantitively evaluate the effectiveness of the proposed method, the residual variance calculated by the low-dimensional embedding of neighborhood size $k$ is shown in Fig. 9, where $k$ changes in the range (5, 100). Fig. 9 shows that the residual variance of the low-dimensional embedding driven by parameter $k = 20$ is less than the other parameters $k$. However, though $k$ is set to 20 to extract the features of the Lorenz system, the result shown in Fig. 8(b) also reflects the large reconstruction errors so that it is difficult to select an optimal fixed parameter $k$ which can satisfy both the minimum residual variance and the reconstruction errors at the same time.

Moreover, if the above indexes(residual variance and reconstruction error) are used criterion to select neighborhood size $k$, each of low-dimensional embedding and signal reconstruction should be calculated with different $k$. Obviously, the training cost will be huge. Conversely, the residual variance calculated by the low-dimensional embedding of SOML algorithm is only 0.018. Compared with Fig. 9, the proposed method obtains a smaller residual variance and a smaller reconstruction error. Therefore, The SOML technique has more ability in extracting features than the fixed parameter method.

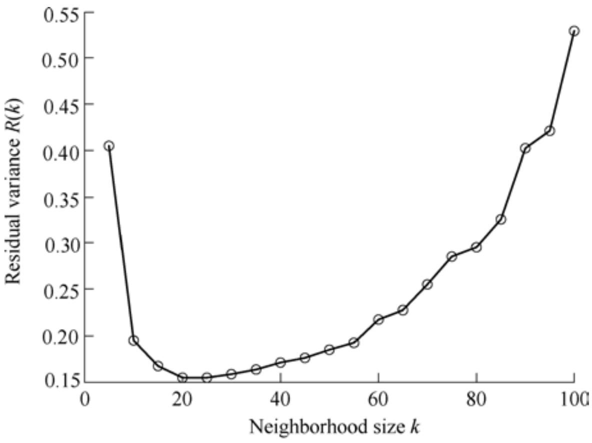


Fig. 9. Curve of residual variance for fixed neighborhood size

## 5 Application in Machinery

Due to the complex structure of the reciprocating machinery, it is difficult to diagnose the dynamic characteristics of engine. In this paper, the abnormal sound signal was recorded by microphone from a Jiefang CA141 automotive gasoline engine with piston pin fault. During the measurements, the revolution speed of the gasoline engine was 1 800 r/min. Although there are same vibration source of the machine for vibration signals and sound signals, the frequency response range of sound sensors is much lower than that of accelerometers. In addition, the performance of abnormal sound caused by faults is usually decided by the crankshaft speed. In the view of audible sound, the sampling frequency is set to 6kHz that the low and middle frequency reflecting part impacts can be highlighted effectively.

The result in Ref. [3] indicates that the abnormal sound is usually aroused by parts impact, caused by defect or wear of parts. Because the impact of parts is a transient process, abnormal sound can be characterized by the presence of periodic repetition of the impulse attenuation components. Fig. 10 shows the waveform of collected sound signal. Obviously, the waveform is more complex compared to the rotational machinery, whose signal should be a pure sine curve. The periodical impulse components are submerged by the noise. So how to extract these periodical features is a key problem to the diagnosis of the engine.

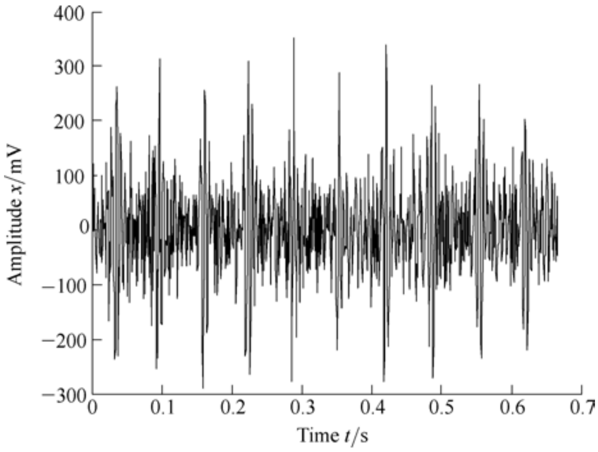


Fig. 10. Original abnormal sound signal

According to the evaluation index proposed in Cao algorithm, the minimum embedding dimension is 3. Then, a high-dimensional manifold is built by phase space reconstruction, where $m = 30$ and delay time $\tau = 1$. Finally, the three-dimensional embedding extracted by SOML is adopted to reconstruct the signal with gauss kernel function, which is shown in Fig. 11. At the view of the whole waveform, it is not difficult to find that the time of the interval appearance just rightly corresponds to the impulse response, so it can be identified as the knock vibration of the piston pin.

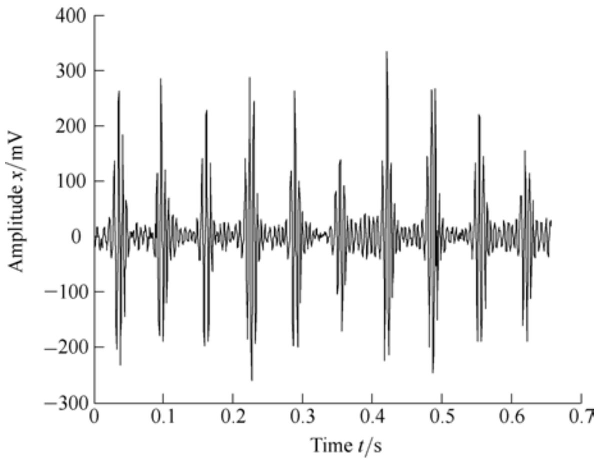


Fig. 11. Reconstructed waveform with SOML

In order to estimate the performance of feature extraction, the reconstructed signal was compared with the original signal, which is shown in Fig. 12. The reconstructed signal based on SOML extracts not only the impulse component accurately, including amplitude and period, but also the small amplitude component, which can reflect other information. Moreover, the LTSA with the fixed neighborhood size was also applied to the same signal to extract the feature components. The plots in Fig. 13 give out the experimental results of residual variance outputs with different neighborhood sizes, where $k$ changes in the range (10, 100).

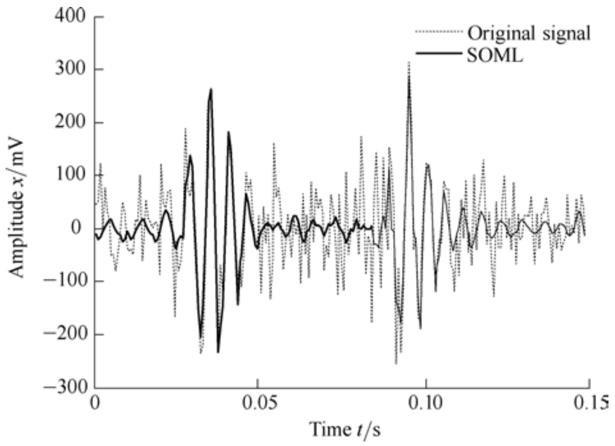


Fig. 12. Comparison of signal between SOML and original signal

As depicted in Fig. 13, the $R(k)$ exhibits its minimal value at $k=20$. Therefore, LTSA with neighborhood size $k=20$ is used to extract the features of the signal. The comparison to original signal is shown in Fig. 14. For LTSA with fixed neighborhood size, except for the impulse feature, the irrelevant noise components were also extracted, such as the noise components at 0.076 s and 0.085 s respectively. Moreover, compared with Fig. 12, the signals between impulse components were also filtered. Obviously, the experiment result shows that the small neighborhood size leads to lack necessary overlaps among neighbors. While for large $k$, the computed tangent space cannot represent the local geometry well. Then the small periodic components are linearized so that they are filtered. Meanwhile, the noise components are emerged because of the lack of necessary overlaps. It also indicates that the nonlinear mapping capability of SOML is superior to that of LTSA.

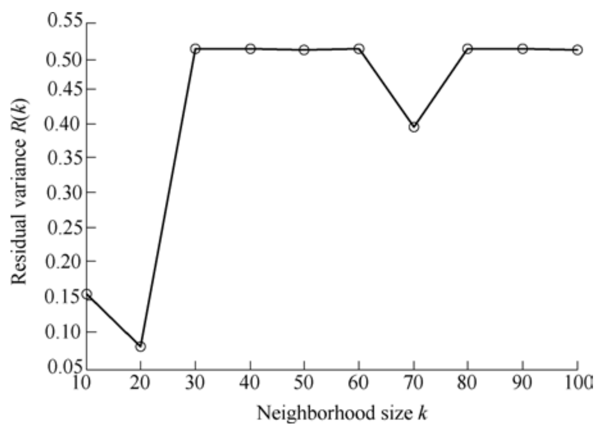


Fig. 13. $R(k)$ curve with different neighborhood size

For the comparison, the continuous wavelet transform (CWT) was also used to extract the impulse components[4], where the wavelet function was Morlet wavelet. Since the amplitude division computed by hard-threshold function is rough, the noise components of the wavelet coefficient which are greater than the threshold cannot be suppressed effectively. Therefore, the soft-threshold is adopted in this paper. To find an optimal wavelet filter which can discover the periodic impulse, the first step is to search the optimal shape factor $\beta$. Increase $\beta$ from 0.1 to 2 and calculate the entropy of the corresponding coefficients. The optimal shape factor $\beta$ leaded by the minimal Shannon entropy relationship is obtained. As depicted in Fig. 15, the entropy exhibits its minimal value at $\beta=0.6$. Therefore, $\beta=0.6$ was selected as the optimal shape factor.

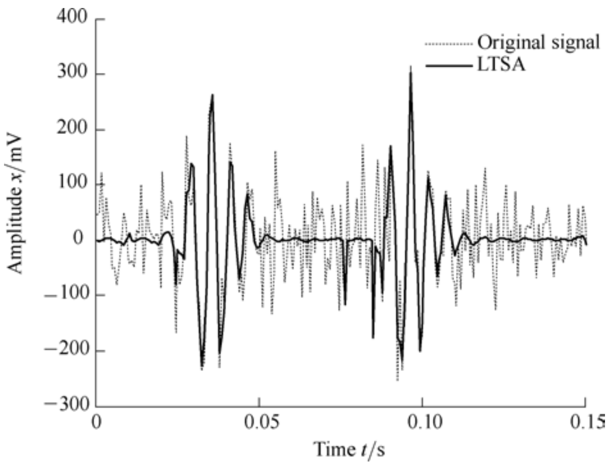


Fig. 14. Comparison of reconstruction signal between LTSA and original signal

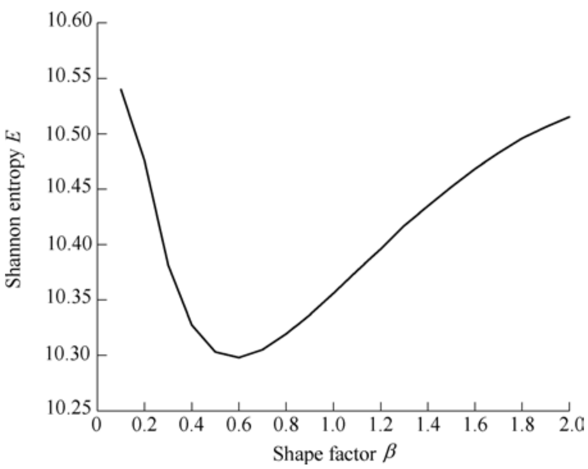


Fig. 15. Relationship between Shannon entropy of the wavelet coefficient and shape factor $\beta$

After the shape factor $\beta=0.6$ is selected, the de-noised signal by applying the Morlet wavelet filter with optimal shape factor $\beta$ is shown in Fig. 16. Comparing the original signal with that after the CWT filter, the main periodic impulse features were extracted basically. However, it is obvious that there are noise components in the second impulse feature, and the whole amplitude of reconstructed signal also decreases.

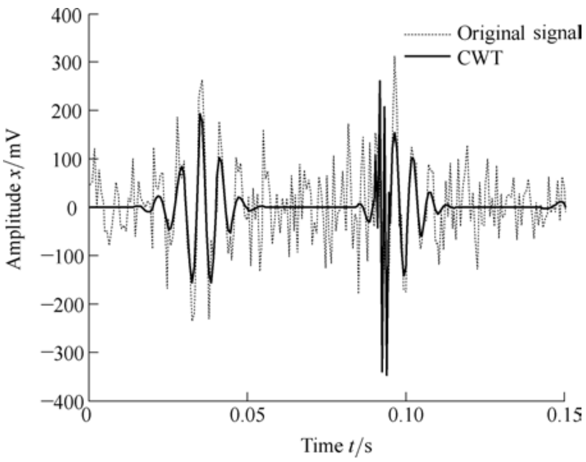


Fig. 16. Reconstructed waveform with wavelet filter

Moreover, the small amplitudes of the signal between periodic impulses are considered as the noise to be completely filtered. The performance of wavelet filter based on de-noising method is influenced by the relative energy levels of signal coefficients and noise coefficients. It also implies that the feature extraction capability of SOML is superior to that of CWT.

In order to further verify the validity of the proposed method in extracting localized defects, measurements from rolling bearing experiments is also considered. In rotating machinery, the failure of rolling bearings can result in the deterioration of machine running conditions. Therefore, it is significant to be able to detect and diagnose the existence accurately and automatically. The measurement was performed on a 308 rolling bearing with rotating speed of 1 600 r/min and sampling frequency of 40 kHz. The vibration signal obtained from a defective bearing with an outer race fault is shown in Fig. 17.

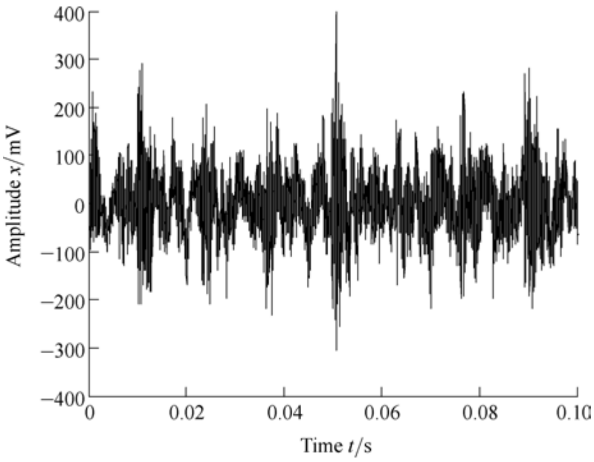


Fig. 17. Vibration waveform of bearing outer race failure

It even gives the impression that the bearing is fault-free, as no impulse can be seen clearly in the signal. According to the evaluation index proposed in Cao algorithm, the minimum embedding dimension is also 3. The reconstructed signal with three-dimensional embedding extracted by SOML is shown in Fig. 18. Fig. 18 shows that the impulses are visible, which correspond to the outer race fault frequencies respectively. The result shows that the reconstructed signal by SOML can effectively extract the impulse features and be able to suppress the disturbance noise well.

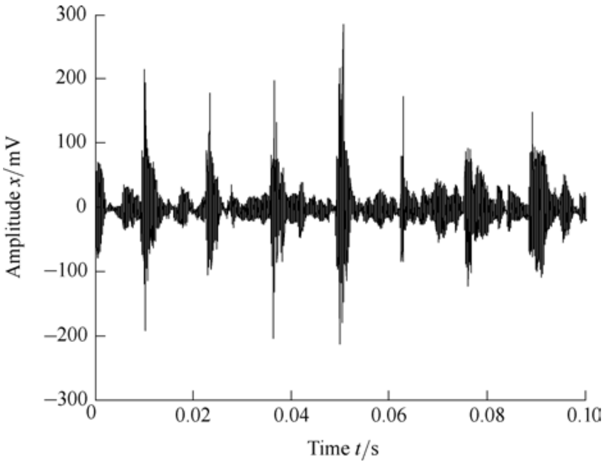


Fig. 18. Reconstruction signal with SOML

## 6 Conclusions

(1) Experiments demonstrate that The SOML technique has more ability in extracting features, and the impulse components can be extracted effectively.

(2) Through the application of EM iteration algorithm to the non-uniform distribution, the affection of transcendent knowledge and manual intervention can be avoided and the local neighborhoods can be divided adaptively, then combining EM iteration with LTSA, a smaller residual variance and a smaller reconstruction error can be obtained at the same time. Therefore, the feature extraction capabilities can be improved obviously.

(3) The simulation of Lorenz system and the application of an engine with piston pin fault and a bearing with outer race fault show that the feature extraction capability of SOML is superior to that of CWT and LTSA.

### References

[1] LEI Yaguo, HE Zhengjia, ZI Yanyang, et al. New clustering algorithm-based fault diagnosis using compensation distance evaluation technique[J]. *Mechanical System and Signal Processing*, 2008, 22(2): 419–435.

[2] WANG Huaqing, LI K, SUN H, et al. Feature extraction method based on pseudo-wigner-ville distribution for rotational machinery

in variable operating conditions[J]. *Chinese Journal of Mechanical Engineering*, 2011, 24(4): 661–668.
[3] ZHOU Bin, SHEN Yudi, XIAO Yunkui. Feature extraction and diagnosis form the abnormal sound faults of diesel engine based on the continuous wavelet transform coefficients[J]. *Transactions of CSICE*, 2006, 24(4): 363–368. (in Chinese).
[4] LIN Jing, QU Liangshen. Feature detecting and fault diagnosis based on continuous wavelet transform[J]. *Journal of mechanical engineering*, 2000, 36(12): 95–100. (in Chinese)
[5] YANG Junyan, ZHANG Youyun, ZHU Yongsheng. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension[J]. *Mechanical System and Signal Processing*, 2007, 21(5): 2 012–2 024.
[6] ZHANG Shuqing, ZHAO Yunchun, ZHANG Liguo, et al. Mechanical fault diagnosis based on band-phase-randomized surrogate data and multifractal[J]. *Chinese Journal of Mechanical Engineering*, 2011, 24(5): 885–895.
[7] AMAZ M, ROBERT X G. PCA-Based feature selection scheme for machine defect classification[J]. *IEEE Trans on Instrumentation and Measurement*, 2004, 53(6): 1 517–1 525
[8] HYVARINEN A, OJA E. Independent component analysis: algorithms and applications[J]. *Neural Networks*, 2000, 13(425): 411–430.
[9] SEUNG H S, DANIEL D L. The manifold ways of perception[J]. *Science*, 2000, 290(5 500): 2 268–2 269.
[10] ROWEIS S, SAUL L. Nonlinear dimensionality reduction by locally linear embedding[J]. *Science*, 2000, 290(5 500): 2 323–2 326.
[11] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. *Neural Computation*, 2003, 15(6): 1 373–1 396.
[12] ZHANG Zhenyue, ZHA Hongyuan. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment[J]. *Journal of Scientific Computing*, 2005, 26(1): 313–318. (in Chinese).
[13] YANG Jianhong, XU Jinwu, YANG Debin, et al. Noise reduction method for nonlinear time series based on principal manifold learning and its application to fault diagnosis[J]. *Journal of Mechanical Engineering*, 2006, 42(8): 154–158. (in Chinese)
[14] TANG Baoping, LI Feng, QIN Yi. Fault diagnosis based on feature compression with orthogonal locality preserving projection[J]. *Chinese Journal of Mechanical Engineering*, 2011, 24(5): 891–898.
[15] ZHANG Z, WANG J, ZHA H. Adaptive manifold learning[J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 2012, 34(2): 253–265.
[16] YANG L. Building connected neighborhood graphs for locally linear embedding[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(5): 827–831.
[17] NG S K, MCLACHLAN G J. Speeding up the EM algorithm for mixture model-based segmentation of magnetic resonance images[J]. *Pattern Recognition*, 2004, 37(8): 1 573–1 589.
[18] DICK D R, OLGA K, PIETIKAINEN M. Selection of the optimal parameter value for the locally linear embedding algorithm[C]// *Proc. of the 1st Int.Conf. on Fuzzy Systems and Knowledge Discovery*, Orchid Country Club, Singapore, November 18–22, 2002: 359–363.
[19] CAO L Y. Practical method for determining the minimum embedding dimension of a scalar time series[J]. *Physica D*, 1997, (110): 43–50.
[20] TAKEDA H, FARSIU S. MILANFAR P. Kernel regression for image processing and reconstruction[J]. *IEEE transaction on image processing*, 2007, 16(2): 349–366.

## Biographical notes

CHEN Xiaoguang, born in 1988, is currently a PhD candidate at *School of Mechanical Engineering, Xi'an Jiaotong University, China*. He received his bachelor degree from *Xi'an Jiaotong University, China*, in 2010. His research interests include signal process and intelligent fault diagnosis.
Tel: +86-29-83395054; E-mail: chenxiaoguang2008@163.com

LIANG Lin, born in 1973, is currently working at *Xi'an Jiaotong University, China*. He received his PhD degree from *Xi'an Jiaotong University, China*, in 2007. His research interests include fault diagnosis and measurement.
Tel: +86-29-82663707; E-mail: lianglin@mail.xjtu.edu.cn

XU Guanghua, born in 1964, is currently a professor at *Xi'an Jiaotong University, China*. He received his PhD degree from *Xi'an Jiaotong University, China*, in 1995. His research interests include fault diagnosis and brain-computer interface.
Tel: +86-29-82663707; E-mail: xugh@mail.xjtu.edu.cn

LIU Dan, born in 1978, is currently working at *Xi'an Jiaotong University, China*. He received his PhD degree from *Xi'an Jiaotong University, China*, in 2007. His research interests include remote monitoring and intelligent fault diagnosis.
Tel: +86-29-83395054; E-mail: liudan1978@mail.xjtu.edu.cn