Monitoring the Health Status of Rolling Bearings Based on Adaptive mean-shift

Xiaoqing Xue¹*, Bo Chen¹, Yong Zhang² and Ting Liu¹

¹ School of Physics & Electronic Engineering, Xianyang Normal University, Xianyang, Shaanxi, 712000, China
² Xianyang Hengtong Electric Power Design Co., Ltd, Xianyang, Shaanxi, 712000, China
*Corresponding author’s e-mail: xuexiaoqing0601@dingtalk.com

Abstract: In order to monitor the structural health condition of rolling bearings in real time, an adaptive mean-shift algorithm was proposed to replace the fixed step size in the Mean Shift algorithm with adaptive variable step size. Wavelet packet energy entropy was used to extract the characteristics of rolling bearing vibration signals, and a series of centroids were clustered by adaptive mean-shift after dimensionality reduction. The centroid in normal state was used as the reference centroid to calculate the offsets of all centroids (including reference centroid) and reference centroid. Experiments show that the adaptive mean-shift algorithm converges quickly and has high accuracy. The centroid offset can be used to monitor the structural health effectively.

1. Introduction
Rolling bearings are one of the most widely used parts in various rotating machinery, and whether their operating conditions are normal directly affects the performance of the entire machine. Equipment condition monitoring and fault diagnosis is a process of pattern recognition, and cluster analysis is a method of pattern recognition. In 1967, J. McQueen proposed the k-means clustering algorithm, which repeatedly divides the data into k clusters according to a certain distance function given a set of data points and the required number of clusters k [1]. In 2011, C.T.Yiakopoulos et al. proposed the use of k-means algorithm for fault identification. This algorithm must specify the number of categories in advance, which is easily affected by outliers and has poor self-regulation ability [2]. In 1975, Fukunaga et al. proposed Mean Shift algorithm, which is an iterative process based on density function gradient estimation and pattern search [3]. In 1995, Cheng Yizong defined a family of kernel functions in the Mean Shift algorithm, and set the weight coefficients of the sampling points, which expanded the scope of application of the Mean Shift algorithm [4]. In 2012, Zhao Yijie used the Mean Shift algorithm to monitor the structural health of rolling bearings. This algorithm is greatly affected by the initial value of the core radius and has limited accuracy[5].

The Mean Shift algorithm is a statistical iterative algorithm, which gets the kernel density estimation function according to the sampling points, starts from the starting point, iteratively searches for the most densely sampled area in the feature space, and gradually shifts along the direction where the density of the sample points increases the fastest to the local maximum of the kernel density estimation function [6]. Since the step length of each data point is the same every time, the Mean Shift algorithm can achieve good accuracy, but the calculation cost is high. The adaptive mean-shift
algorithm improves it, and does not need to determine the number of categories in advance. The bandwidth is adaptively determined according to the characteristics of the data distribution, and the sampling point estimation mode is used to design the adaptive estimator. The adaptive estimator connects each data point with the kernel function of different scales. When the kernel function satisfies certain conditions, the iterative process converges to the extreme point. The convergence and accuracy of the algorithm are verified through experiments.

2. Principle
Assume that in the d-dimensional Euclidean Space, each data point is associated with a wideband value. The sampling point estimator, that is, the adaptive kernel density estimation of X points in the feature space is:

$$\hat{f}_k(x) = \frac{1}{nh^d} \sum_{i=1}^{n} k\left(\frac{x - x_i}{h_i}\right)$$

(1)

where \(k(x)\) is called the nuclear prototype function and the radially symmetric kernel function \(K\) satisfies the following boundary conditions:

$$K(x) = c_{\text{K}}k(\|x\|) \quad \|x\| \leq 1$$

(2)

where \(c_{\text{K}}\) is a standardized constant and makes \(k(x)\) the integral equal to 1. Substituting formula (2) into formula (1), the density estimation formula expressed by the nuclear primitive shape function is obtained:

$$\hat{f}_k(x) = \frac{c_{\text{K}}}{nh^d} \sum_{i=1}^{n} k\left(\frac{x - x_i}{h_i}\right)$$

(3)

The density gradient estimator is the derivative of the density estimation formula.

$$\hat{\nabla}f_k(x) = \nabla\hat{f}_k(x) = \frac{2c_{\text{K}}}{nh^d} \sum_{i=1}^{n} (x - x_i)k\left(\frac{x - x_i}{h_i}\right)$$

(4)

Assume \(g(x) = -k'(x)\) and suppose that the finite point derivative of the nuclear prototype function in \(x \in [0, \infty)\) does not exist, and the other point derivatives exist. Then we use \(g(x)\) as the prototype function, the kernel function is defined as:

$$G(x) = c_{\text{G}}g(x)$$

(5)

Incorporating equation (5) into equation (4):
Comparing formula (3) and formula (6) can prove:

\[
\hat{m}_g(x) = C \frac{\hat{\mathbf{V}}f_k(x)}{f_0(x)} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_i^{d-2}} \left( \sum_{i \neq j} g(x) \right) - x
\]

where \( C \) is a constant, and formula (7) is called the mean shift vector. Replacing \( h_i \) in all the above formulas with a fixed value \( h \) is the process of the Mean Shift algorithm. It can be seen that the bandwidth in the Mean Shift algorithm is fixed, while the bandwidth in the adaptive mean-shift algorithm is continuously adaptively adjusted according to the distribution characteristics of the data.

3. Experiment and result

3.1 Comparison of Mean Shift and adaptive mean-shift

The source of the experimental data is the American Electrical Engineering Laboratory. The device is shown in Figure 1. It consists of a 1.5KW motor, a torque sensor/encoder, a fan and a controller. Measuring point 1 is the base, and measuring point 2 is the housing. The motor load is controlled by the fan, and the data is collected through the vibration acceleration sensor. The sampling frequency is 12KHz. In the experiment, the fan-end rolling bearing has 4 states: normal state, outer ring failure, inner ring failure and rolling element failure.

![Figure 1. Experimental diagram of fan rolling bearing failure.](image)

Use wavelet packet energy entropy to reduce the dimensionality of the original data. Feature extraction is performed every 2048 data, 59 groups of extracted feature data sets are selected for each
operating state, and 236 data sets are clustered using Mean Shift and adaptive mean-shift algorithms. The kernel function is Gaussian kernel function, and the kernel radius is selected from 21 values in [0.03, 0.01]. The results are shown in Table 1.

Table 1. Clustering results of Mean Shift and adaptive mean-shift algorithm under the same core radius.

| Nuclear radius | Mean Shift | Accuracy | Adaptive mean-shift | Accuracy |
|----------------|------------|----------|---------------------|----------|
| 0.03           | 3          | 75%      | 3                   | 75%      |
| 0.029          | 3          | 75%      | 3                   | 75%      |
| 0.028          | 3          | 75%      | 3                   | 75%      |
| 0.027          | 3          | 75%      | 3                   | 75%      |
| 0.026          | 3          | 75%      | 3                   | 75%      |
| 0.025          | 4          | 77.54%   | 3                   | 75%      |
| 0.024          | 4          | 83.9%    | 3                   | 75%      |
| 0.023          | 4          | 97.88%   | 3                   | 75%      |
| 0.022          | 4          | 97.88%   | 3                   | 75%      |
| 0.021          | 4          | 96.19%   | 3                   | 75%      |
| 0.02           | 4          | 94.49%   | 3                   | 75%      |
| 0.019          | 5          | 92.37%   | 3                   | 75%      |
| 0.018          | 6          | 90.25%   | 4                   | 97.46%   |
| 0.017          | 7          | 85.17%   | 4                   | 98.30%   |
| 0.016          | 8          | 81.78%   | 4                   | 99.58%   |
| 0.015          | 8          | 78.81%   | 4                   | 100%     |
| 0.014          | 9          | 72.46%   | 4                   | 100%     |
| 0.013          | 17         | 63.56%   | 4                   | 100%     |
| 0.012          | 25         | 55.51%   | 4                   | 100%     |
| 0.011          | 39         | 47.03%   | 4                   | 100%     |
| 0.01           | 60         | 39.47%   | 4                   | 100%     |

Table 1 shows Mean Shift has an accuracy of more than 90% within the core radius [0.018, 0.023]. Outside this interval, regardless of whether the core radius increases or decreases, the accuracy will decrease. The accuracy of adaptive mean-shift in the core within the radius [0.01, 0.018] is over 97%, the classification number is 4, the accuracy within [0.018, 0.03] is 75%, and the classification number is 3. The results show that the adaptive mean-shift algorithm is less affected by the core radius, and the result is relatively stable.

3.2 Adaptive mean-shift rolling bearing health monitoring experiment
The experimental data adopts the rolling bearing data of the American Intelligent Maintenance System. The schematic diagram of the tested bearing is shown in Figure 2. Four bearings are installed on a shaft, the rotation speed is 2000RPM, the load is 6000 pounds, and a sensor is installed in each of the vertical and horizontal directions. We collect a set of data every 10 minutes.
In the experiment, the horizontal data of the third bearing is selected as the sampling point. Starting from the first day of data, the original data for 5 consecutive days (5290 minutes) are collected and perform a wavelet packet feature extraction on the first 8192 data of each column, we can get 529 sets of data. Take Gaussian kernel function, the initial radius is 0.01, the number of iterations is 100, we use adaptive mean-shift to cluster these 529 sets of data. The monitoring results of the rolling bearing structure state are shown on the polar coordinate axis as shown in Figure 3.

In Figure 3, the adaptive mean-shift divides 529 groups of data into 180 categories (that is, there are 180 centroids). Take the centroid (that is, the center of the circle) in the normal state of the rolling bearing as the reference, the distance between each centroid and the center of the circle represents the magnitude of the offset. The offset starts from the center of the circle. At first, it changes slowly, and the distance from the center of the circle is small, which indicates that the rolling bearing structure is gradually malfunctioning, but the degree of failure is small. The dense data points indicate that there are more data points with less degree of failure. After the data point pointed by the arrow in the figure (that is, the 147th centroid), the offset is getting larger and larger, which indicates that the degree of structural failure is increasing. Large, sparse data points mean that there are fewer data points with a small degree of failure.

4. Conclusions
In this paper, the adaptive mean-shift algorithm is used to cluster the data of different failure states of rolling bearings, and the convergence and accuracy of the adaptive mean-shift algorithm are verified through comparative experiments with the Mean Shift algorithm. Due to the step length of each data point drift continuously adjust to the optimal value, the calculation speed will be improved, and the
clustering effect will be good. A series of centroids are used to describe the change process of rolling bearing structure from normal state to fault state. The centroid offset can determine whether the structure is faulty and the severity of the fault. The adaptive mean-shift algorithm is an effective way to diagnose the health status of the rolling bearing structure, but the damage forms of mechanical structures are complex and changeable, how to determine the damage location after assessing the damage degree is our further work.

Acknowledgments
This work was supported in part by Shaanxi Provincial Key Research and Development Program 2021NY-213, in part by Xianyang Normal College Youth Key Program XSYGG201605, in part by Special Scientific Research Program of Shaanxi Provincial Department of Education 19JK0928.

References
[1] Macqueen, J. (1967) Some Methods for Classification and Analysis of Multivariate Observations. In: Proc.of Berkeley Symposium on Mathematical Statistic and Probability, America.pp. 281-297.
[2] Yiakopoulos C T, Gryllias K C, Antoniadis I A. (2011) Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. Expert Systems with Applications, 38:2888-2911.
[3] Fukunaga, K., Hostetler, L.D. (1975) The estimation of the gradient of a density function with applications in pattern recognition. IEEE Transactions on Info. Theory, 21:32-40.
[4] Chen, Y.Z. (1995) Mean shift, mode seeking, and clustering. IEEE Trans on Pattern Analysis and Machine Intelligence, 17:790-799.
[5] Zhao, Y.J. (2012) Human security protection techniques using Mean Shift clustering. Xi’an: Chang ‘An University.
[6] Ilan, S., Bogdan, G., Peter, M. (2003) Adaptive Mean shift based clustering in high dimensions. In: IEEE International Conference on Acoustics,Speech and Signal Processing, Hong Kong, China. pp. 456-463.