Failure Prediction for Multi-function Machine based on Vibration Data Analysis

Kazuki Kobayashi　Masatoshi Sekine

The decrease in the labor force due to the seriously declining birthrate and aging population in Japan is expanding the need to improve productivity using AI (Artificial Intelligence) and IoT (Internet of Things) technologies[1]. In particular, there is great demand for an AI technology that can quickly and automatically detect an anomaly or event that indicates a failure of a machine.

OKI is working on failure sign detection using a unique vibration analysis algorithm that utilizes machine learning, and experiments have been conducted with various manufacturing machines[2]. According to OKI's experience, the machines to be subjected to anomaly detection are categorized into two types; a monotonically operating machine whose vibration pattern does not change often, and a machine whose operation varies over time resulting in frequent changes in vibration pattern (hereinafter multi-function machine) (**Figure 1**). The former machine group corresponds to a ball screw of a machine tool or a large pump for air conditioning, and the latter corresponds to a robot arm or a metal processing apparatus.

Detecting an anomaly in a multi-function machine is difficult with conventional anomaly detection method due to reasons described later. In this article, a new anomaly detection method for multi-function machines is introduced.

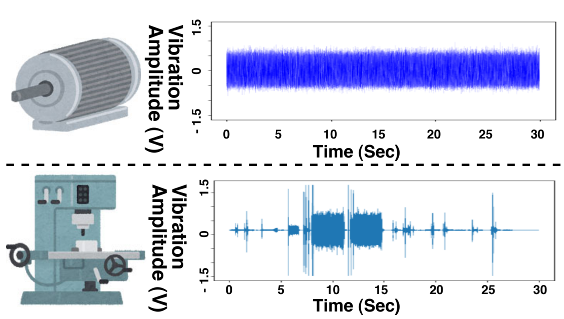


Figure 1. Machines Targeted for Anomaly Detection and Their Vibration Waveforms (Top: Single function machine, Bottom: Multi-function Machine)

Difficulties with Anomaly Detection of Multi-function Machines

In a multi-function machine, the normal vibration state differs for each operation. Therefore, criteria (decision models) to determine the normality/abnormality of the machine are prepared for each operating pattern, and anomaly detection is performed by switching the models as appropriate. The following three problems arise if anomaly detection is tried with the conventional method.

- **Problem 1**

It takes many man-hours to extract and classify vibration data. In addition, there is also a problem that the annotation (meaning) results for the same data often differ when working with multiple persons because the criteria for segmentation and classification are not unified.

- **Problem 2**

When there are many operating patterns, it is a huge burden for people to manually select the suitable feature values for each operation and prepare the decision model.

- **Problem 3**

In anomaly detection targeted at manufacturing machines, abnormal data may not be collected, and it may not be possible to detect an anomaly with a "supervised learning," which requires both normal and abnormal data.

Features of Developed Method

In this section, the features of the new vibration anomaly detection method (hereinafter, developed method) that solves the three problems presented above are described.

First, to extract the feature values of the vibration data, the developed method utilizes deep learning and automates the selection of the feature values. At this time, using deep learning with a "unsupervised learning" enables anomaly detection without requiring abnormal data. Additionally, when creating a model for extracting feature values from the vibration data (hereinafter, feature value extraction model) training data is provided using a method that does not require extraction and classification of vibration data for each operating pattern.

Next, based on the extracted vibration feature values and the operating pattern information specified from the control information, a model for estimating the degree of anomaly (hereinafter, anomaly estimation model) is created. In the developed method, the operating pattern is specified by acquiring the control information of the machine, and the decision model is successively switched according to the result. Examples of control information include control signals of the machine and time elapsed after the start of an operation.

Detailed methods for model creation and anomaly estimation are described in the following sections.

Variational Autoencoder

In the developed method, it is assumed that supervised learning cannot be utilized for abnormal detection, therefore "unsupervised" deep learning is to used to realize a detection method that does not require abnormal data. The variational autoencoder[3] (VAE), which is the "unsupervised" deep learning method adopted in the developed method, will be briefly explained.

VAE is a type of autoencoder[3] (AE) that is trained to reproduce input data at its output. In AE, compression (encode) is performed by extracting feature values from the input data, and the same dimensionality data as the input data is output (decode) from compressed feature values. At that time, by using the input data as the correct one, it becomes possible to train with "unsupervised leaning."

On the other hand, in VAE, assuming that input data follows a probability distribution of average μ and variance σ^2, μ and σ^2 are extracted from the input data, and learning is performed so as to output the same data as the input data (**Figure 2**). Thus, by assuming that input data follows the probability distribution, it is possible to train a structured, continuous latent space as compared with using AE.

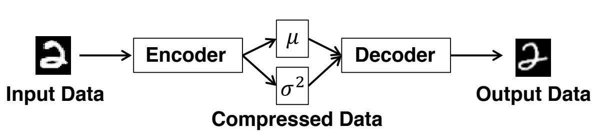


Figure 2. Variational Autoencoder

Model Creation and Anomaly Estimation Methods

In this section, methods for creating feature value extraction model and anomaly estimation model, and anomaly estimation method are described separately in the training and inference stages. Here, it is assumed that the elapsed time after the machine starts an operation is used as the control information.

(1) Training stage

Figure 3 shows the process at the training stage. The vibration data for training is divided into one data for creating the feature value extraction model and another data for creating the anomaly estimation model. The procedure for creating the feature value extraction model is as follows.

- Vibration data for creating the feature value extraction model is Fourier-transformed to calculate a spectrogram. (a-1)

A spectrogram is vibration power information classified by time and frequency and obtained by calculating vibration power for each frequency in short time intervals. In the following, the dimensionality in the time axis direction is T.

- Average value in the time axis direction of the spectrogram obtained in a-1 is calculated. (a-2)
- Result obtained in a-2 is input into an untrained feature value extraction network and the model is trained. (a-3)

Here, the average value in the time axis direction of the spectrogram is calculated in step a-2 and provided as training data for the feature value extraction network in step a-3, thereby making it unnecessary to extract and classify each operating pattern in advance. In addition, the obtained feature value extraction model is used in the inference stage and in creating the anomaly estimation model.

The procedure for creating the anomaly estimation model is as follows.

- Vibration data for creating the anomaly estimation model is Fourier-transformed to calculate a spectrogram. (b-1)
- The spectrogram from b-1 is extracted according to time t (t=1, 2,..., T) then input into the feature value extraction network obtained in a-3. The resulting output from the encoder section (μ and σ^2 in **Figure 2**) is used as the feature value at time t. (b-2)

- The result obtained in b-2 is provided as training data for the anomaly estimation model. In the case of estimation based on statistical distances, the feature value y_{tk} for each time t extracted from training vibration data X_k (k = 1, 2,..., K) is used and the feature value distribution of the normal state is trained for each time t. K is the number of training data for the anomaly estimation model. (b-3)

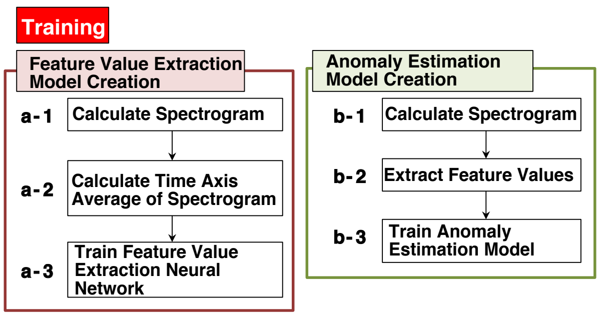


Figure 3. Training Stage Process

(2) Inference stage

Anomaly estimation method using the two models obtained in the learning stage will be described (**Figure 4**).

- Vibration data for the estimation target is Fourier-transformed to calculate a spectrogram. (c-1)
- The spectrogram from c-1 is input into the feature value extraction network from a-3 for each time t to obtain feature values y_t (t = 1, 2,..., T). (c-2)
- Degree of machine anomaly is estimated using the model obtained in c-2. (c-3)

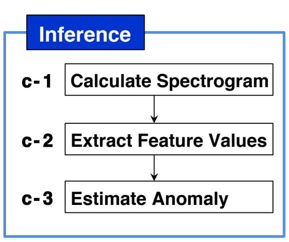
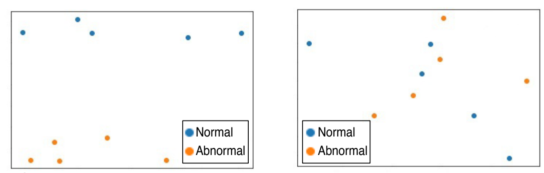


Figure 4. Inference Stage Process

Figure 5 is a distribution example of feature values y_t for a given time t. The figure shows that the distribution of the feature values may or may not separate due to the normality/abnormality of the machine.The reason why such differences occur is that in a multi-function machine, the operation point changes with time.Therefore, it is assumed that even if there is some anomaly in the machine, the distribution of the feature values separates between normal and abnormal during times when the abnormal part is operating, but no difference is seen in the distribution during times when the abnormal part is not operating. Focusing on this distribution difference between normal and abnormal, the degree of anomaly is calculated based on the statistical distance between the feature value distribution y_{tk} of the normal state and the feature value y_t of the evaluation data.



Separation of feature value distribution due to machine normality/abnormality

No separation of feature value distribution due to machine normality/abnormality

Figure 5. Feature Value y_t Distribution for Given Time t (5 Normal/Abnormal Samples Each)

Evaluation Experiment

In order to evaluate the effectiveness of the developed method, an experiment was conducted on a machine (hereinafter, evaluation machine), which OKI produces, containing several motors and mechanical units that operate sequentially. The vibration data was acquired by installing a vibration sensor with a detection frequency of 10 Hz to 15 kHz on the surface of the evaluation machine and operating the machine continuously for about four months. In the experiment, a 30-second sequential operation including multiple motions was evaluated as one sample. **Figure 6** shows the vibration waveform and spectrogram for one of the samples.

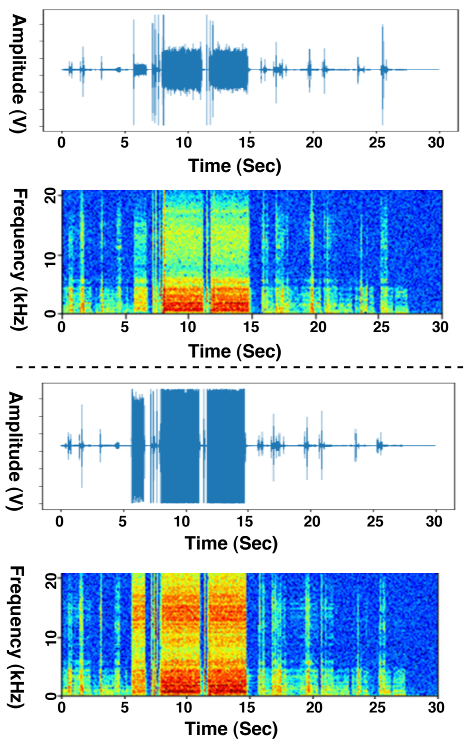


Figure 6. Vibration Waveform and Spectrogram of a 30-second Sequential Operation (Top: First day of experiment, Bottom: After four months)

(1) Model Training

In training, 550 vibration data samples (approx. 4.5 hours' worth) from the first day of the experiment was used as normal data. Five hundred of these samples were used to train the feature value extraction model and the remaining 50 were used to train the anomaly estimation model. The spectrogram was calculated using a window width of 0.1 seconds and a shift width of 0.05 seconds.

(2) Experiment Results and Considerations

The degree of anomaly was estimated using the developed method with five samples each of vibration data acquired weekly from the first four weeks of experiment and monthly from the first four months. The average degree of anomaly for these five weekly and monthly samples is shown in **Figure 7**. From these results, it can be observed that after a week of experiment, the degree of anomaly becomes large 6 to 15 seconds after the evaluation machine starts operating (indicated by ※ in **Figure 7**). In addition, although there is no significant difference in the degree of anomaly for this time section for the first to the third week, the degree of anomaly becomes larger after the fourth week.

Then the condition of the evaluation machine was investigated, and reason for the results obtained above was considered. As can be seen from the waveforms and spectrograms in **Figure 6**, the evaluation machine begins to exhibit unusual vibrations 6 to 15 seconds after the start of operation. The difference in the amplitude of the generated vibration between the first day of experiment and four months later during the time section in question is apparent. Parts of the machine operating during this time section were investigated, and frictional wear was found in one of the gears.

Taking these facts into account, it is presumed that the degree of anomaly rises during the worn gear's operation, and the anomaly further increased after the fourth week. From this, it is considered that the abnormal parts of the machine can be identified by matching the operation timing of each machinery part with the timing at which the degree of anomaly increases. However, it was not possible to confirm how the degree of gear wear relates with vibration amplitude and degree of anomaly. Future experiment are necessary to match the relationship between anomaly estimation and degree of gear wear.

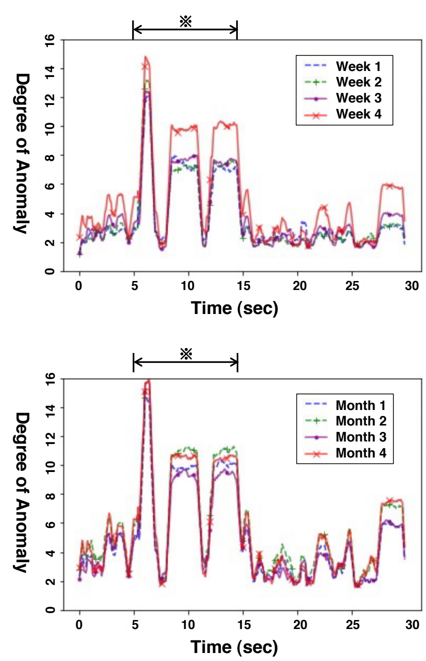


Figure 7. Experiment Results (Top: Weekly Anomaly Changes, Bottom: Monthly Anomaly Changes)

Conclusion

In this article, a new anomaly detection method for multi-function machine was described. Evaluation experiment showed that the proposed method could detect vibration change caused by the wear of a certain gear.

In the future, additional evaluation for the proposed method will be conducted using other multi-function machines. Effort will be made to develop a sophisticated anomaly detection method. ◆◆

References

1) Japan's Ministry of Economy, Trade and Industry: "White Paper on Manufacturing Industries (Monodzukuri) 2017" http://www.meti.go.jp/report/whitepaper/mono/2017/index.html
2) Yusuke Takahashi: "Anomaly Detection using Vibration Analysis with Machine Learning Technology for Industrial IoT System," OKI Technical Review Issue 230, Vol.84, No.2, pp30-33, December 2017
3) Francois Chollet: "Deep Learning with Python," First Edition, pp312-321, 2018, Mynavi Publishing Corporation

Authors

Kazuki Kobayashi, IoT Solution-Business Promotion Department, IoT Platform Division, ICT Business Group

Masatoshi Sekine, AI Technologies R&D Department, Corporate Research & Development Center, Corporate Infrastructure Group

TIPS 【Glossary】

Spectrogram
Represents the time shift in the frequency distribution of vibration power.