ABSTRACT
The acoustic data remotely measured by handy type microphones are investigated for monitoring and diagnosing the rotational machine integrity in nuclear power plants. The plant operator’s patrol monitoring is one of important activities for condition monitoring. However, remotely measured sound has some difficulties for precise diagnosis or quantitative judgment of rotating machine anomaly, since the measurement sensitivity is different in each measurement, and also, the sensitivity deteriorates comparing with an attached type sensor. Hence, in the present study, several signal feature extraction methods are compared in order to compensate the sensitivity variation and deterioration and to find both sensitive and robust monitoring algorithms. These pre-processed signal feature patterns are classified by kernel based principal component analysis (KPCA) and probabilistic neural network (PNN). KPCA is useful to increase the sensitivity and robustness for classification of normal and abnormal states. Also, PNN is useful to classify the known states and unknown states. The developed algorithms are applied to the acoustic data measured by mockup test facility of rolling bearing type rotating machine which can simulate normal and abnormal states. Here, several kinds of abnormal state data were measured, which consist of inner and outer race defects (Large, Middle and Small defects), and, bearing ball defect as well as normal operation data. It is shown that the developed algorithms can effectively classify the above several kinds of normal and abnormal data.

INTRODUCTION
Condition based maintenance (CBM) is one of important activities for both reliable and efficient nuclear power plant operation. Recently, the new maintenance regulation has started in Japan to allow the utilities to decide optimum maintenance and operating periods for the plants. This helps the flexible and efficient plant operation, however, at the same time, the responsibility of utilities increases to assure high plant reliability. From this point of view, the condition monitoring activities will be more and more important.

In the present paper, the acoustic data remotely measured by handy type microphones are investigated for monitoring and diagnosing the rolling bearing type rotational machine integrity in nuclear power plants. The plant operator’s patrol monitoring is one of important activities for the above-mentioned condition monitoring. However, remotely measured sound has some difficulties for precise diagnosis or quantitative judgment of rotating machine anomaly, since the measurement sensitivity is varied in each patrol, and also, is less than the attached type accelerometer. Hence, the author will investigate the various kinds of acoustic signal processing methods for increasing their sensitivity of anomaly monitoring. Furthermore, the robustness of the monitoring to the operating condition change is also discussed, since both the sensitivity and robustness are very important in practical applications. In the author’s previous paper [1], the signal pre-processing method was proposed, where the acoustic signal is normalized based on the fundamental oscillation period which is extracted by a zero-crossing interval of filtered acoustic signal. Then, the individual periodic pattern is converted into the same length data and used for pattern recognition. In this study, it was noticed that the sensitivity and robustness of the monitoring are largely depends on this pre-processing algorithm. So, effectiveness of signal pre-processing, or, in other words, signal feature extraction algorithms will be examined again by using the mockup test facility of rolling bearing type rotational machine which can simulate various kinds of anomaly.

The extracted signal feature patterns are classified by using both the conventional principal component analysis (PCA) and the kernel-based PCA (KPCA) which are able to reduce the dimension of feature pattern vector and extract an effective feature space for classification and also visualization.
The reason why different kinds of signal processing methods are examined is to find both sensitive and robust methods for acoustic monitoring of rotating machines. Since the skilled human ability is superior to the machine learning capability from viewpoints of anomaly detection sensitivity, it is desired to investigate how to catch up the human ability by machine learning technologies. Furthermore, it should be noted that human experts also have the robust classification capability which can discriminate abnormal states from normal ones even though the machine is operated in different conditions. So, both sensitive and robust machine learning capability are going to be pursued by using the above signal processing methods, or, some heuristic data mining methods.

In the next step of classification, extracted small dimension’s feature vectors are modelled by the probabilistic neural network (PNN) in order to classify the normal and abnormal states. Since the PNN can classify the known or learned states and so-called ‘don’t know state’, it is useful for automated and quantitative judgment for rotating machine condition monitoring.

The developed algorithm is applied to the acoustic data measured by mockup test facility. Here, several kinds of known abnormal state data were measured, which consist of inner and outer race defects, or, bearing ball defects, as well as normal operation data in different rotating speed conditions [2]. The author will discuss how the classification capability increases by introducing the new signal processing methods. Furthermore, robust classification capability to the operating condition change will be also discussed.

SIGNAL PROCESSING METHODS

Feature extraction methods

For anomaly detection using acoustic signals, three kinds of steps should be considered: (1) feature extraction from acoustic signals, (2) classification of normal and abnormal states, and, (3) evaluation of classification results. In step (1), various kinds of methods have been investigated in the speech or speaker recognition research area [3]. Among them, the following three feature patterns are adopted and compared with each other.

Log-scale Auto-power Spectral Density (log-APSD)
This is a simple frequency spectrum of raw acoustic data. Here, the log-scale pattern of 0-22 kHz frequency range divided by 512 points is used as the feature vector. Here, the frequency axis is treated by the linear scale. Each pattern is evaluated by 0.2 second length acoustic data. Also, in order to compensate the measurement sensitivity variation, each signal is normalized by its RMS (Root Mean Square) value.

Mel-scale Auto-power Spectral Density (Mel-scale-APSD)
Mel-scale is a perceptual scale of frequency pitches which is often used in speech recognition [3]. The frequency scale is converted into the following Mel-scale:

\[ m = 1127 \times \log_e \left(1 + \frac{f}{700}\right) \]  

(1)

Here, the 512 frequency resolutions are compressed into 30 dimensions by this transformation. As for the amplitude, the log-scale APSD is used.

Cepstrum
Cepstrum is defined by the Fourier transformation of log-APSD, and has a time domain scale called by quefrency. If the frequency spectrum \( X(\omega) \) can be assumed as the product of sound source \( G(\omega) \) and transfer function \( H(\omega) \), Cepstrum can be calculated as follows:

\[ c(\tau) = F^{-1} \log|X(\omega)| = F^{-1} \log|G(\omega)| + F^{-1} \log|H(\omega)| \]  

(2)

The second term of right-hand-side equation corresponds to the envelop of frequency spectrum. So, lower frequency \( (\tau) \) parts represent the envelop shape of log-APSD and express vocal tract transfer characteristics. Although the size of Cepstrum vector is the same as log-APSD, the author uses the first 60 components as a feature vector in the present study.
Machine state classification

The above-mentioned feature vectors are used for state classification. Since these feature vectors have a large dimension, and, extensive information about machine states, it might be effective to find essential and low-dimension features from them. For this purpose, PCA and KPCA are used [4,5].

PCA based classification

Given a \( M \) set of \( p \)-dimensional centered observations, \( x(m) \) \((m=1,M)\), which consist of the normalized one-rotation waveform or its Fourier transform, PCA diagnoses the covariance matrix, \( \bar{C} = \frac{1}{M} \sum_{m=1}^{M} x(m) \cdot x(m)^T \). \( (3) \)

To do this, the following eigenvalue equation has to be solved:
\[ \lambda V = \bar{C} V \]. \( (4) \)

Then, \( m \)-th observation, \( x(m) \), is projected onto the \( k \)-th eigenvector, \( V^{(k)} \), as follows:
\[ z_k(m) = V^{(k)T} \cdot x(m) \] \( (5) \)

Here, \( z_k(m) \) is the \( k \)-th score value in PCA. In the present state classification method, appropriately selected small number of score values, \( z_k(m), \ (k=1,K, m=1,M) \), are used to discriminate the machine state.

KPCA based classification [4,5]

KPCA is the extension of PCA to non-linear space and expected to choose more sensitive signal features of abnormal state from observations. First, \( x \rightarrow \Phi(x) \) is assumed to be a possible non-linear mapping. In KPCA, it is enough to define just the dot product of \( \Phi(x) \) to compute the projected score values in non-linear space. To do this, the covariance matrix in non-linear space is defined, instead of Eq. (3), as follows:
\[ \bar{C} = \frac{1}{M} \sum_{m=1}^{M} \Phi(x_m) \cdot \Phi(x_m)^T \] \( (6) \)

The eigenvalue problem is the same as Eq. (4). Here, it is noted that the dimension of \( V \) becomes \( M \) in non-linear space. Since all solutions of \( V \) lie in the span of \{ \( \Phi(x_1) \ldots \Phi(x_M) \) \}, the following relations are satisfied.
\[ \lambda(\Phi(x_m)^T \cdot V) = \Phi(x_m)^T \cdot CV \quad \ (m=1,M) \]
\[ V = \sum_{m=1}^{M} \alpha_m \Phi(x_m) \] \( (7) \)

Here, \( M \)-dimensional vector \( \alpha \) can be computed by solving the following eigenvalue equation:
\[ M \lambda \alpha = K \alpha \] \( (8) \)

The kernel matrix \( K \) is defined as follows:
\[ K(x_i, x_j) = \Phi(x_i)^T \cdot \Phi(x_j) \] \( (9) \)

In the present paper, the following Gaussian kernel function is used:
\[ K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right) \] \( (10) \)

Here, \( \sigma \) in Eq.(8) has to be normalized so as to satisfy the relation, \( \lambda_s(\alpha^{(k)T} \cdot \alpha^{(k)})=1 \), since the eigenvector, \( V^{(k)} \), have to satisfy \( (V^{(k)T} \cdot \alpha^{(k)})=1 \). After obtaining the eigenvector \( V^{(k)} \) in the non-linear space, the score value can be computed by the dot product of \( \Phi(x) \) and \( \alpha \) as follows:
\[ z_k = V^{(k)T} \cdot \Phi(x) = \sum_{m=1}^{M} \alpha^{(k)}_m (\Phi(x_m)^T \cdot \Phi(x)) \] \( (11) \)

The machine state is classified using appropriately selected small number of score values, \( z_k(m), \ (k=1,K, m=1,M) \). Although the dimension of eigenvector \( V \) in the linear PCA should be less than \( p \), the dimension of observation vector, the dimension of KPCA eigenvector can be extended to the
learning data number $M$. This means the KPCA will have more flexible and sensitive classification capability.

**Heuristic classification**

Since the human can learn the speech recognition ability without any theories, there is a possibility of finding heuristically the good feature vectors for the normal and abnormal state classification. In order to find such good features, the goodness should be defined at first. So, it is assumed that a good classification index is to maximize the following criterion:

$$ C(F_i, F_j) = D_{NA}(F_i, F_j) - D_{NN}(F_i, F_j) - D_{AA}(F_i, F_j) $$  \hspace{1cm} (12)

Here, $F_i$ and $F_j$ are the $i$-th and $j$-th elements extracted from the multi-dimensional feature vector. $D_{NA}(F_i, F_j)$ is the Maharanobis distance between Normal and Abnormal state data defined by

$$ D_{NA}(F_i, F_j) = (\mu^{(i,j)}_A - \mu^{(i,j)}_N)^T (\Sigma^{-1}_A + \Sigma^{-1}_N)^{-1} (\mu^{(i,j)}_A - \mu^{(i,j)}_N). $$ \hspace{1cm} (13)

Here, the symbols ‘N’ and ‘A’ mean normal and abnormal states respectively. Also, the notations $\mu$ and $\Sigma$ express the average and variance of $i$-th and $j$-th elements for normal and abnormal states respectively. To maximize the above criterion means that the abnormal state is allocated away from the normal state and each normal or abnormal states are allocated near in the two dimensional feature space. The author explains the criterion in two dimensional space, but, it is easy to extend it to a higher dimensional space.

To find the optimum feature indices in Eq. (12), some optimization algorithms, such as particle swarm optimization (PSO), should be used. However, in the present study, the whole possible combinations are searched to maximize the criterion, since the computational load was not critical in a two dimensional space case.

**State modeling by PNN and anomaly monitoring [6]**

PNN is defined as a model of a certain state class using mixed Gaussian distribution of class members, $X_{ij}$, as follows:

$$ f_j(X) = \frac{1}{(2\pi\sigma^2)^{\frac{n_j}{2}}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^{n_j} (X - X_{ij})^T (X - X_{ij}) \right] $$ \hspace{1cm} (14)

Here, $X_{ij}$ means the $i$-th member of score vector which belongs to the $j$-th class. The $n_j$ is the number of $j$-th class members and $\sigma$ is the smoothing parameter. The classification can be made by using this probability and a certain threshold level $\varepsilon$ as follows:

$$ f_j(X) \geq \varepsilon \Rightarrow X \in G_j $$

$$ f_j(X) \leq \varepsilon \Rightarrow X \notin G_j $$ \hspace{1cm} (15)

If learning data of $K$ classes exist, it can be discriminated whether a target observation belongs to one of known $K$ classes or an unknown class. The case where the observation doesn’t belong to any known classes is important for practical applications, and, is called as ‘don’t know class’.

For continuous monitoring, such as on-line monitoring or daily patrol monitoring, the following monitoring index, $MI$, which is the log-likelihood of PNN can also be used:

$$ MI(X(t)) = \log(f_{normal}(X(t))) $$ \hspace{1cm} (16)

Here, $X(t)$ is a target feature vector and $f_{normal}(X)$ is a PNN model of the normal state.

Furthermore, it is possible to define the classification index ($CI$) by using the log-likelihood of PNN for the normal state as follows:

$$ L_1 = \min_{j: normal} \log\{f_{normal}(X_{ij})\} $$

$$ L_2 = \max_{j: abnormal} \log\{f_{normal}(X_{ij})\} $$ \hspace{1cm} (17)

$$ CI = L_1 - L_2 $$
Here, the positive and large $CI$ value means good classification performance and less false alarm rate. On the other hand, the negative $CI$ means some abnormal data are classified wrongly into the normal state. This means that the miss-alarm rate increases.

**TEST RESULTS**

**Test facility and measurement**

In order to verify the above mentioned algorithms, the acoustic data were measured using the rolling bearing type test facility shown in Fig. 1. This machine can be operated under several conditions of simulated defects, such as inner and outer race defects (Large, Middle, Small) or ball defects by changing bearing elements, in addition to normal condition’s operation. Some examples of defects are shown in Fig.1. The measurement was made using handy type microphone with digital sampling of 44.1kHz and 10 sec length. The data were measured under different rotating speed, 3000, 2000, 1500, 1000 and 500 rpm.

The examples of auto-power spectral density (APSD) of measured anomaly data are shown in Figs. 2 and 3. Here, two kinds of anomaly data, inner race large and small defects, are compared with the normal state data under 3000 rpm rotating speed conditions. It is seen that the typical resonance frequencies are observed at the 50Hz fundamental mode and it’s higher oscillation modes. Also, it is noted that differences between the normal and anomaly state are not so large except for the high frequency area of large defect case. When vibration data are measured using attached type accelerometer, these differences were clear. However, in the case of the remotely measured acoustic data shown here, the differences become very small. This means that signal processing techniques are important for amplifying the small signal features to distinguish anomaly states.

![Mockup facility of rolling bearing and simulated failures](image1)

![log-APSD of acoustic sound for Normal and Inner race defect](image2)

**Figure 1** - Mockup facility of rolling bearing and simulated failures (Outer and Inner race defects)

**Figure 2** - log-APSD of acoustic sound for Normal and Inner race defect (Large) condition
Figure 3 - log-APSD of acoustic sound for Normal and Inner race defect (Small) condition

**Evaluation of classification performance of PCA, KPCA and a heuristic method**

The purpose of the present classification is to discriminate the abnormal state based on the normal feature vector database. In order to evaluate the classification performance, the normal state database labeled by NN3000 and NN2000 and the abnormal state data labeled by InL3000 and InL2000 are used. Here, numeric labels mean the machine rotating speed in RPM unit. Hence, the 3000 and 2000 RPM conditions have 50Hz and 33Hz fundamental oscillation modes respectively. The label InL means the inner race large defect abnormal state. The reason why two different operating speed data are included is to evaluate the robustness of classification performance to the operating condition change. The appearance of normal state feature vector patterns is different from not only abnormal patterns but also different operating speed normal patterns. The skilled human engineers can discriminate the abnormal state even though the target machine is operated in different speeds. This situation is similar to the human ability of speaker recognition [7]. It can be easily recognized who speaks even if they speak different words. Hence, in this paper, the author tries to discriminate the abnormal states from normal ones, while the target machine is operated in different speeds.

The feature vectors of the above-mentioned four kinds of state data are evaluated by the aforementioned three algorithms for every 0.2 second length sound waveform with 44.1kHz sampling, which include 8820 points digital data. Since the measured data length was 10 seconds, each state has 50 samples of feature vectors.

In the present analysis, the dimension of feature vector is 512 in log-APSD, 30 in Melscale-APSD, or, 60 in Cepstrum. These large-dimension data are projected into the two-dimension space by PCA, KPCA, or, the heuristic method, in order to visualize each state features. After this projection, each state can be modeled by PNN shown by Eq. (14). Then, the monitoring index (MI) or the classification index (CI) can be evaluated by Eqs. (16) and (17).

The classification results by log-APSD using PCA and KPCA are shown in Fig. 4. Here, the principal components are evaluated only from the normal state data, NN3000 and NN2000. And the score values of the 1st and 2nd principal components are plotted. It is shown that the discrimination of normal and abnormal state is difficult in the case of PCA, but, is easy in the case of KPCA. Similar conclusions are obtained in Fig. 5, where the results of Mel-scale-APSD are shown.

On the other hands, the results of Cepstrum in Fig. 6 are a little bit different from other feature extraction cases. In this case, the plot of 1st and 3rd principal component results was shown. Here, both PCA and KPCA are able to discriminate the abnormal states from the normal ones. Furthermore, the differences of operation condition are less noticeable. This suggests that the Cepstrum and KPCA provide the most robust capability, and, almost the same sensitivity to other methods, for the rotating machine acoustic monitoring. From practical viewpoints, this robustness would be very important characteristics.

In Fig. 7, two examples of heuristic classification results are shown, where two frequency’s amplitudes of log-APSD are chosen for the plot. In the left figure, the frequencies, 50Hz and 301Hz were chosen to represent the normal and abnormal states. The author assumed that these frequencies represented fundamental and higher modes of rotation which were excited by the inner race defect.
However, the classification results are not so good. Hence, the optimum frequencies are searched based on Eq. (12). As a result, it can be found the frequency combination of 301Hz and 1.2kHz can maximize Eq. (12). The results are shown in the right-hand-side of Fig. 7. Here, it is shown that the good discrimination of normal and abnormal states is attained. This result suggests us that the high frequency impact sound from inner race defect and rolling balls would be also characteristic features for anomaly monitoring in addition to the higher mode oscillation amplitude.

As a summary, the classification capability are evaluated by the Maharanobis distance defined by Eq. (13), and, shown in Table 1. Here, the larger $D_{NA}$ and smaller $D_{NN}$ or $D_{AA}$ values are preferred for the good classification. The qualitative evaluations are also shown in this table. These results show that the classification performance largely depends on the feature extraction algorithm. Among them, Cepstrum and KPCA provide the good performance for robust and sensitive classification. Furthermore, it is seen that the heuristic method has a good possibility for not only the classification capability but also the knowledge discovery for anomaly condition monitoring.

![Figure 4 - State classification by PCA and KPCA based on log-APSD](image1)

![Figure 5 - State classification by PCA and KPCA based on Mel-scale-APSD](image2)
Figure 6 - State classification by PCA and KPCA based on Cepstrum

Figure 7 - State classification by heuristic method of log-APSD
<table>
<thead>
<tr>
<th>Feature Extraction</th>
<th>Classification</th>
<th>Components</th>
<th>Maharanobis Distance</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>X-axis</td>
<td>Y-axis</td>
<td>DNA</td>
</tr>
<tr>
<td>log-APSD</td>
<td>PCA</td>
<td>1</td>
<td>2</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>KPCA</td>
<td>1</td>
<td>2</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>61.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>58.9</td>
</tr>
<tr>
<td>Mel-scale-APSD</td>
<td>PCA</td>
<td>1</td>
<td>2</td>
<td>3.65</td>
</tr>
<tr>
<td></td>
<td>KPCA</td>
<td>1</td>
<td>2</td>
<td>10.6</td>
</tr>
<tr>
<td>Cepstrum</td>
<td>PCA</td>
<td>1</td>
<td>2</td>
<td>7.27</td>
</tr>
<tr>
<td></td>
<td>KPCA</td>
<td>1</td>
<td>3</td>
<td>6.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>76.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>75.83</td>
</tr>
<tr>
<td>log-APSD</td>
<td>Heuristics</td>
<td>1227 Hz</td>
<td>301 Hz</td>
<td>104.88</td>
</tr>
<tr>
<td>Mel-scale-APSD</td>
<td>Heuristics</td>
<td>13602 Hz</td>
<td>413 Hz</td>
<td>109.54</td>
</tr>
<tr>
<td>Cepstrum</td>
<td>Heuristics</td>
<td>18quefrency</td>
<td>3quefrency</td>
<td>98.65</td>
</tr>
</tbody>
</table>

Table 1 - Evaluation summary of classification

CONCLUSION

The various kinds of signal processing methods have been discussed for condition monitoring of rotating machines using remotely measured acoustic signals. Although the remotely measured sound has advantages for condition monitoring, their monitoring accuracy is less than attached type sensors. To fill the lack of sensitivity, several kinds of advanced statistical signal processing methods were introduced such as PCA, KPCA and PNN, in addition to the heuristic knowledge discovery method. And, it was shown the introduction of these signal processing methods contributed to improvement of both sensitivity and robustness for condition monitoring. In Japan, the new maintenance regulation rule has just be introduced where condition monitoring technologies will be expected to have a more important role. For this future perspective, new technologies for condition monitoring should be pursued more.

ACKNOWLEDGEMENTS

This research was supported by Japan Society of Maintenonology. The author also expresses his thanks to Drs. T. Tamaoki and S. Shimizu in Toshiba Corporation for their kind cooperation with this work.

REFERENCES

6) D.F.Spect, “Probabilistic Neural Networks for Classification, Mapping, or Associative Memory”, *Proc. IEEE Int. Conf. On Neural Networks*, 1998, 1, p.525-531.