NONLINEAR SVM-TYPE AUTOMATIC DICISION ALGORITHM IN NOISY ENVIRONMENT FOR HAMMERING TEST SYSTEM

Kouki Hori* and Nari Tanabe† and Masaya Fujisawa‡
* Tokyo University of Science, Tokyo, Japan
E-mail: 4621520@ed.tus.ac.jp
† Suwa University of Science, Nagano, Japan
E-mail: nari@rs.sus.ac.jp
‡ Tokyo University of Science, Tokyo, Japan
E-mail: fujisawa@rs.tus.ac.jp

***Abstract*—This paper proposes the nonlinear SVM-type automatic decision algorithm in a noisy environment for hammering test system. The proposed method (i) suppresses the noise which is many at low frequency by using a Butterworth type high pass filter, and (ii) extraction of hammering signal from noise signal estimation based on minimum statistics method and Kalman filter algorithm with colored driving source. Furthermore, (iii) finding the optimal discriminant boundary using a nonlinear support vector machine with the first formant and first formant frequency as features obtains from cepstrum analysis results, and (iv) judged and evaluated from an automatic decision algorithm considering time change of feature. The effectiveness of the proposed method was demonstrated by computer simulation.**

I. Introduction

Hammering test is one of the non-destructive tests used to determine whether or not concrete structures, such as bridges and tunnels, are damaged by cracking or peeling. In recent years, there have been accidents in Japan where people have died due to collapsing ceiling panels in tunnels[1]. The cause of this accident was the dropping, loosening, and corrosion of the anchor bolts and cracks of the protective concrete due to insufficient inspection and test by hammering test. After this accident, the importance of hammering test was reaffirmed, and problems with hammeing test began to be pointed out.

The first problem is the decrease in sound volume due to the change of the inspection hammer. Until now, hammering tests have been carried out with a regular hammer, but it is difficult to make accurate judgments because regular hammer inspections are difficult to carry out in high places and narrow areas. To overcome this, a new hammer was developed. The angle and length of the tip of the hammer can be adjusted, enabling inspection in high places and narrow areas. However, this inspection using the hammer also caused the disadvantage of low-hitting volume because the sound was generated by rolling the hammer.

The second problem is that the noise is superimposed on the hammering sound. Most hammering tests in Japan are restricted, but not stopped. As a result, the test environment is one in which vehicles are constantly on the road. In addition, a new wall lifting robot is currently being researched that can climb walls and extract hammering sounds. Because of this, the noise generated by the traveling vehicles that reach the inspectors will increase even more, since there will be no need to even restrict the traffic.

The third problem is the lack of inspectors who can determine the healthy part and damaged part with high accuracy. An inspector with a high ability to distinguish between the sound of healthy parts and the sound of damaged parts of a hammering sound can distinguish between high and low frequencies, large and small hammering amplitudes, and different tones. Conversely, if the above capabilities are not high, it is not possible to determine the healthy part and damaged part with high accuracy.

II. Previous Method

In [2], the effect of the presence or absence of defects on the sound is discussing by conducting a hammering test of actual bridges with clear signs of flaking and delamination. In this previous study, to evaluate the sensitivity of the output to the input, (i) the amplitude ratio is obtained by dividing the maximum sound pressure at each hitting position by the maximum load obtained by the impulse hammer. (ii) The frequency spectrum of sound pressure differs depending on the size and condition of the defect, making it difficult to determine between healthy and damaged parts, and (iii) It was confirmed that it is possible to discriminate between healthy and damaged parts by using Self Organizing Maps.

However, in the actual field, it is difficult to extract accurate features due to the superimposition of noise on the hammering sound. In the previous study, they assumed hammering test on real bridges. However, since the superimposition of noise is not a condition, errors in the determination of healthy and

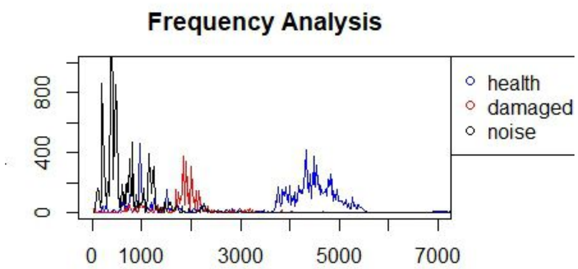

Fig. 1: Frequency analysis

damaged parts are expected to occur in an environment where loud noise exists, such as in actual sites.

In contrast, the author's previous study[3] proposed the hammering test system based on noise suppression using the Kalman filter algorithm with colored driving source and automatic judgement using Support Vector Machine(SVM), a kind of machine learning. However, the noise suppression algorithm of the previous method can suppress noise spectrum that are smaller than hammering spectrum, but it is difficult to suppress noise spectrum that are larger than the hammering supectrum. In addition, there existed many cases of judgement errors due to the existence of overlap in the feature distributions of healthy and damaged parts.

Therefore, in this study, based on the differences in the distributions found in the frequency analysis result of hammering signals recorded from real environments and noise signals, low-frequency noise suppress is performed using a Butterworth type high-pass filter(BW-type HPF), followed by noise estimation using the minimum statistics method and hammering signal estimation using the Kalman filter algorithm with colored driving source. Next, we adopt the first formant as a feature instead of the spectrum peak, which is a feature of conventional methods, and calculate the optimal hyperplane using nonlinear SVM, which is a nonlinear extension of linear SVM. Finally, we use a judgement algorithm that takes the time transition into account, which enables highly accurate noise suppression and automatic judgement making.

III. Proposed Method

A. low frequency noise suppression by BW-type HPF

First, in order to understand the characteristics of hammering signals and noise signals, we prepared sound data with only hammering signals and sound data with only noise signals, and performed frequency analysis as shown in Figure 1.

From Fig 1, it can be seen that there is a large noise in the low-frequency range and hammering sound in the middle to the high-frequency range. As a result of preparing several experimental sound data and frequency analyses, it was found that the inside of bridges and tunnels often had a large noise in the low-frequency band. Therefore, the noise present in the low-frequency band is suppressed by the following third-order BW-type HPF. The observed signal$\hat{x}_{L/R}(n)$ after BW-type HPF is expressed by the following equation.

$$\hat{x}_{L/R}(n) = \sum_{l=0}^{M} a_l x_{L/R}(n-l) - \sum_{l=1}^{W} b_l \hat{x}_{L/R}(n-l) \quad (1)$$

However, $M = W = 2048$ are the feedback forward order and feedback order, respectively, and a_l and b_l are the feedforward coefficient and feedback coefficient, respectively.

B. Residual noise power spectrum estimation by the minimum statistical method[4]

Next, Calculate the observed spectrum after BW-type HPF in the frequency domain. Separate the observed signal$\hat{x}_{L/R}(n)$ after BW-type HPF by a certain short time(frame), and perform Short Time Fourier Transform(STFT) of $\hat{x}_{L/R}(n)$ for each interval. By STFT, we obtain the observed spectrum $\hat{X}_{L/R}(\lambda, k)$ after BW-type HPF. $\hat{X}_{L/R}(\lambda, k)$ is obtained by the following equation.

$$\hat{X}_{L/R}(\lambda, k) = \sum_{\mu=0}^{W-1} \hat{x}_{L/R}(\lambda N + \mu) \cdot h(\mu) \cdot \exp\left(-j\frac{2\pi\mu k}{W}\right) \quad (2)$$

Where λ is the frame number, k is the frequency bin number, W is the DFT frame length ($W = 2048$), and N is the frame interval ($N = W/2 = 1024$). In addition, the Hanning window is used as the window function this time.

For noise estimation, if the residual noise spectrum included in the observed spectrum $\hat{X}_{L/R}(\lambda, k)$ after BW-type HPF fluctuates significantly from frame to frame, it is expected that there will be errors in the judgment of noise estimation. Therefore, we can reduce the noise estimation error by smoothing the observed spectrum $\hat{X}_{L/R}(\lambda, k)$ after BW-type HPF calculated by 2. When $\hat{X}_{L/R}(\lambda, k)$ is smoothed in the frame direction, the smoothed observation power spectrum $\bar{X}_{L/R}(\lambda, k)$ is expressed by the following equation.

$$\bar{X}_{L/R}(\lambda, k) = \alpha \cdot \bar{X}_{L/R}(\lambda-1, k) + (1-\alpha) \cdot |X_{L/R}(\lambda, k)|^2 \quad (3)$$

α is the smoothing factor. It is set to a value in the range $(\alpha = 0.9 \cdots 0.95)$.

Next, the smoothed residual noise power spectrum $V_{L/R}(\lambda, k)$ is obtained using the minimum statistical method. $V_{L/R}(\lambda, k)$ is obtained by searching for the minimum value of the spectrum in the range of the past L_d frames for $\bar{X}_{L/R}(\lambda, k)$. The $V_{L/R}(\lambda, k)$ is expressed by the following equation.

$$\left.\begin{aligned} V_{L/R}(\lambda, k) = \min(&\bar{X}_{L/R}(\lambda, k), \bar{X}_{L/R}(\lambda-1, k), \\ &\bar{X}_{L/R}(\lambda-2, k), \cdots, \bar{X}_{L/R}(\lambda-L_d+1, k)) \end{aligned}\right\} \quad (4)$$

The minimum statistical method is a widely known method for continuous noise estimation with tracking capability. However, the minimum statistical method has the disadvantage of targeting a relatively high SNR.

Because noise estimation and noise suppression in the frequency domain are generally based on the premise that the probability that the observed power spectrum for each frequency component is the maximum and the noise power spectrum is the minimum is high. In this case, a relatively high SNR is achieved by suppressing the noise power spectrum, which is relatively abundant in the low-frequency region, using a Butterworth-type HPF in advance.

C. Estimation of residual noise coefficient

To calculate the residual noise spectral coefficients, we use the residual noise power spectrum $V_{L/R}(\lambda, k)$ and the smoothed observation power spectrum $\bar{X}_{L/R}(\lambda, k)$.

The ratio of $V_{L/R}(\lambda, k)$ to $\bar{X}_{L/R}(\lambda, k)$ gives the fraction of the residual noise power spectrum included in the smoothed observation power spectrum $\bar{X}_{L/R}(\lambda, k)$ for each frame number and each frequency bin number.

The residual noise power spectral coefficient $Q_{L/R}(\lambda, k)$ is expressed by the following equation.

$$Q_{L/R}(\lambda, k) = \sqrt{osub_{L/R}(\lambda, k)\frac{V_{L/R}(\lambda, k)}{\bar{X}_{L/R}(\lambda, k)}} \tag{5}$$

The $osub_{L/R}(\lambda, k)$ in equation (5) is the oversubtraction coefficient [5].

The over subtraction coefficient applies a correction to the ratio of the smoothed observation power spectrum $\bar{X}_{L/R}(\lambda, k)$ to the residual noise power spectrum for musical noise reduction. By setting a larger proportion of the residual noise spectrum in the observed spectrum $\hat{X}_{L/R}(\lambda, k)$ after BW-type HPF by the over subtraction coefficient, we can prevent residual noise spectrum from being left unestimated in the residual noise spectrum calculation by the full-wave rectification process and reduce the generation of musical noise.
To calculate the over subtraction coefficient, we need the left-right signal to noise ratio $SNR_{L/R}(\lambda, k)$ for each spectral bin. The $SNR_{L/R}(\lambda, k)$ is expressed by the following equation.

$$\begin{aligned} &SNR_{L/R}(\lambda, k) \\ &= 10\log\left(\frac{\bar{X}_{L/R}(\lambda, k) - \min\left(V_{L/R}(\lambda, k), \bar{X}_{L/R}(\lambda, k)\right)}{V_{L/R}(\lambda, k)}\right) \end{aligned} \tag{6}$$

Using the calculated SNR, the over subtraction coefficient $osub_{L/R}(\lambda, k)$ is expressed as.

$$osub_{L/R}(\lambda, k) = \begin{cases} 1.0 & (SNR_{L/R}(\lambda, k) > 20) \\ 4.5 & (SNR_{L/R}(\lambda, k) < -5) \\ 4.0 - \dfrac{3.0}{20.0}\cdot SNR_{L/R}(\lambda, k) & (else) \end{cases} \tag{7}$$

D. Residual noise spectrum estimation by full wave rectification process

The residual noise spectrum is calculated by multiplying the post-BW-type HPF observed spectrum $\hat{X}_{L/R}(\lambda, k)$ by the fraction of the residual noise spectrum contained in the post-BW-type HPF observed spectrum $\hat{X}_{L/R}(\lambda, k)$. Therefore, the residual noise spectrum $|\hat{V}_{L/R}(\lambda, k)|$ is expressed by the following equation using the observed spectrum $\hat{X}_{L/R}(\lambda, k)$ after BW-type HPF and the noise spectrum coefficient $Q_{L/R}(\lambda, k)$.

$$|\hat{V}_{L/R}(\lambda, k)| = \begin{cases} \hat{X}_{L/R}(\lambda, k)\cdot Q_{L/R}(\lambda, k) \\ \quad (\hat{X}_{L/R}(\lambda, k) \leq \hat{X}_{L/R}(\lambda, k)\cdot Q_{L/R}(\lambda, k)) \\ \hat{X}_{L/R}(\lambda, k)\cdot Q_{L/R}(\lambda, k) - \hat{X}_{L/R}(\lambda, k) \\ \quad (else) \end{cases} \tag{8}$$

E. Estimate the variance value of residual noise signal

The residual noise signal $\hat{v}_{L/R}(n)$ is calculated by performing an inverse short time fourier transform(ISTFT) on the residual spectrum $|\hat{V}_{L/R}(\lambda, k)|$ using the phase of the observed spectrum $\hat{X}_{L/R}(\lambda, k)$ after BW-type HPF using phase of $\hat{X}_{L/R}(\lambda, k)$. The residual noise signal $\hat{v}_{L/R}(n)$ obtained for each frame by the inverse short-time Fourier transform is multiply added (OverLap Add) to the residual noise signal $\hat{v}_{L/R}(n)$ at intervals of N samples.

Calculate the variance of the residual noise signal $\sigma^2_{\hat{v}_{L/R}}(n)$. The $\sigma^2_{\hat{v}_{L/R}}(n)$ is expressed by the following equation using the residual noise signal $\hat{v}_{L/R}(n)$.

$$\sigma^2_{\hat{v}_{L/R}}(n) = \frac{1}{N-1}\sum_{\lambda=1}^{N-1}\hat{v}^2_{L/R}(n-\lambda) \tag{9}$$

Where N is the number of samples used to calculate the variance value.

F. estimate hammering signal for the kalman filter with colored driving source.

The left and right hammering signals are estimated by performing noise suppression on a state-space model consisting of a state equation representing the temporal variation of the hammering signal and an observation equation representing the observation signal superimposed on the hammering signal and noise signal, using the variance value of the noise signal and the Kalman filter algorithm with colored driving source. In the Kalman filter with colored driving source, the $2N$ dimensional state vector is defined as.

$$\left.\begin{aligned} \boldsymbol{d}(n+1) = [&d_L(n+1),\ d_L(n),\ \cdots,\ d_L(n-N+2), \\ &d_R(n+1),\ d_R(n),\ \cdots,\ d_R(n-N+2)]^T \end{aligned}\right\} \tag{10}$$

The equation of state for the temporal change of the hammering signal is as follows.

[state equation]

$$\boldsymbol{d}(n+1) = \Phi \boldsymbol{d}(n) + \boldsymbol{\delta}(n+1) \tag{11}$$

Now, the $2N \times 2N$matrix state transition matrix Φ and the $2N$ dimensional driving source vector $\boldsymbol{\delta}(n+1)$ are expressed as follows.

$$\left.\begin{aligned}
\Phi &= \begin{bmatrix}
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
1 & 0 & \cdots & \cdots & 0 & \vdots & \ddots & & & \vdots \\
0 & \ddots & \ddots & & \vdots & \vdots & & \ddots & & \vdots \\
\vdots & & \ddots & \ddots & \vdots & \vdots & & & \ddots & \vdots \\
0 & \cdots & 0 & 1 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\
\vdots & \ddots & & & \vdots & 1 & 0 & \cdots & \cdots & 0 \\
\vdots & & \ddots & & \vdots & 0 & \ddots & \ddots & & \vdots \\
\vdots & & & \ddots & \vdots & \vdots & & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & 1 & 0
\end{bmatrix} \\
\boldsymbol{\delta}(n+1) &= \begin{bmatrix} d_L(n+1), 0, \cdots, 0, d_R(n+1), 0, \cdots, 0 \end{bmatrix}^T
\end{aligned}\right\} \tag{12}$$

The drive source vector $\boldsymbol{\delta}(n+1)$ consists of only the hammering signals $d_L(n+1)$ and $d_R(n+1)$, as shown in equation (12), and thus constitutes a colored driving source.

Note that the state transition matrix Φ in Eq. (12) is a matrix that expresses the relationship with the next time and is a shift matrix consisting of only 0 and 1. It can be inferred that when the Kalman filter algorithm with colored driving source is implemented in a noise suppression application, real-time performance will not be sacrificed in its operation.

Next, from the equation (10), define the 2-dimensional observation vectors for the observation signals $x_L(n+1)$ and $x_R(n+1)$, where the noise signals $v_L(n+1)$ and $v_R(n+1)$ are superimposed on the hammering signals $d_L(n+1)$ and $d_R(n+1)$ at the next time, as follows.

$$\boldsymbol{x}(n+1) = [x_L(n+1), x_R(n+1)]^T \tag{13}$$

In this case, the observation equation is expressed as follows.

$$\boldsymbol{x}(n+1) = \boldsymbol{m}^T \boldsymbol{d}(n+1) + \boldsymbol{v}(n+1) \tag{14}$$

Now, the $2 \times 2N$ observation transition vector $\boldsymbol{m}^T$ and noise vector $\boldsymbol{v}(n+1)$ are respectively as follows.

$$\left.\begin{aligned}
\boldsymbol{m}^T &= \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{array}\right] \\
\boldsymbol{v}(n+1) &= \begin{bmatrix} v_L(n+1) \\ v_R(n+1) \end{bmatrix}
\end{aligned}\right\} \tag{15}$$

By solving equation (11) and equation (14), unlike the usual Kalman filter [7][8], we can estimate the best estimated hammering signal $\hat{\boldsymbol{d}}(n+1|n+1)$ by sequential state and observation updates without parameter estimation. The best estimated hammering signal $\hat{\boldsymbol{d}}(n+1|n+1) = \left[\hat{d}_L(n+1), \hat{d}_R(n+1)\right]^T$ is computed by noise suppression using the Kalman filter algorithm with colored driving source.

G. Feature extraction

Calculate the hammering cepstrum $c_{L/R}(\tau)$ by applying an ISTFT to the logarithmic hammering spectrum obtained from the hammering signal $\hat{d}_{L/R}(n+1)$ using the STFT. If the Fourier transform is denoted by the symbol $\mathcal{F}$, the hammering cepstrum $c_{L/R}(\tau)$ is as follows.

$$c_{L/R}(\tau) = \mathcal{F}^{-1}\left[\log \breve{D}_{L/R}(\lambda, k)\right] \tag{16}$$

The quefrency $q(\tau)$ is as follows.

$$q(\tau) = 1000 \cdot \sum_{\tau=0}^{W} \frac{0.046}{W}\tau \tag{17}$$

However, since it is multiplied by 1000, the dimension is $[ms]$, and 0.046 is the number of seconds of DFT frame length W.

Next, the low quefrency part of the hammering cepstrum is liftered. Liftering is an operation to separate the various quefrency components. The cepstrum in the low quefrency part represents the hammering spectral envelope (the general shape of the hammering spectrum), while the cepstrum in the high quefrency part represents the fine structure of the hammering spectrum. In this case, we wanted to extract the hammering spectral envelope, so we applied a low pass lifter to the hammering cepstrum as follows.

$$\hat{c}_{L/R}(\tau) = \begin{cases} c_{L/R}(\tau) & (5 < q(\tau) < 41) \\ 0 & (else) \end{cases} \tag{18}$$

Finally, calculate the hammering formants $p_{L/R}(\lambda)$ by performing a STFT on the hammering cepstrum $\hat{c}_{L/R}$ after the low pass lifter. Calculate the feature vector $\boldsymbol{\vartheta}_\lambda$ by calculating the first formant p_λ and the first formant frequency f_λ from the hammering formants $p_{L/R}(\lambda)$ at all frame numbers. $\boldsymbol{\vartheta}_\lambda$ is the following equation.

$$\boldsymbol{\vartheta}_\lambda = \{f_\lambda, p_\lambda\} \tag{19}$$

H. Generating teacher data

In Support Vector Machine (SVM) [12], supervised data is required to determine the initial position of the hyperplane and the support vector. This teacher data $\boldsymbol{r}_\lambda$ is represented by the following equation.

$$\boldsymbol{r}_\lambda = \{\boldsymbol{\zeta}_\lambda, \boldsymbol{\vartheta}_\lambda\} = \{\boldsymbol{\zeta}_\lambda, (f_\lambda, p_\lambda)\} \tag{20}$$

Now, ζ_λ is the class label, with $+1$ being the healthy part and -1 being the damaged part.

I. Calculation of hyperplane by nonlinear SVM

The ultimate goal of SVM is to calculate the hyperplane that properly separates the healthy part and damaged part. In order to compute the appropriate hyperplane, the teacher data should update $\boldsymbol{w}$ as an error if it is closer than a certain value $h > 0$ from the hyperplane so that it can correctly identify noises smaller than h. This h is called the margin, and by maximizing it, we can calculate the appropriate hyperplane. Now, the discriminant function $f(\boldsymbol{\vartheta_\lambda})$, which outputs the healthy part as $+1$ and the damaged part as -1, is expressed by the following equation.

$$f(\boldsymbol{\vartheta_\lambda}) = sign\{\boldsymbol{w}^T\phi(\boldsymbol{\vartheta}_\lambda) + b\} \tag{21}$$

However, $\phi(\boldsymbol{\vartheta}_\lambda)$ is a mapping function to a higher dimension.

The optimal discriminative hyperplane for margin maximization can then be obtained by solving the main problem of the inequality constraint optimization problem, defined as.

$$\left.\begin{array}{l}\text{evaluation function}\\ \quad L_p(\boldsymbol{w},\boldsymbol{\xi}) = \displaystyle\min_{\boldsymbol{w},b,\xi}\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{\lambda=1}^{N}\xi_\lambda\\ \text{constraint}\\ \quad \zeta_\lambda(\boldsymbol{w^T}\phi(\boldsymbol{\vartheta}_\lambda)+b) - 1 + \xi_\lambda \geq 0,\ \xi_\lambda \geq 0\end{array}\right\} \tag{22}$$

Since the main problem is a 2-variable minimization problem with complex constraints, we need to consider a dual problem for the main problem and apply Lagrange's undecided multiplier method to the main problem. The solution $\boldsymbol{w}_0$ and b_0 of the optimization problem can be obtained by making the condition called the KKT (Karush-Kuhn-Tucker) condition a necessary and sufficient condition. The dual problem, calculated using the main problem and the KKT condition, is as follows.

$$\left.\begin{array}{l}\text{evaluation function}\\ L(\alpha) = \displaystyle\max_{\alpha}\sum_{\lambda=1}^{N}\alpha_\lambda - \frac{1}{2}\sum_{\lambda=1}^{N}\sum_{j=1}^{N}\alpha_\lambda\alpha_j\zeta_\lambda\zeta_j\phi(\boldsymbol{\vartheta}_\lambda)^T\phi(\boldsymbol{\vartheta}_j)\\ \text{constraint}\\ \qquad 0 \leq \alpha_\lambda \leq C,\ \ \boldsymbol{\alpha}^T\boldsymbol{\zeta} = 0\end{array}\right\} \tag{23}$$

By solving the KKT condition and the duality problem, the weight coefficient of the optimal discriminative hyperplane $\boldsymbol{w}_0$ can be calculated as follows.

$$\boldsymbol{w}_0 = \sum_{\lambda=1}^{N}\alpha_\lambda\zeta_\lambda\phi(\boldsymbol{\vartheta}_\lambda) \tag{24}$$

It can be seen that the optimal solution is represented by a linear combination of the teacher data. The optimal bias term B_0 is computed on the data satisfying $0 < \alpha_\lambda < C$. The optimal bias term B_0 is as follows.

$$b_0 = \frac{1}{N_M}\sum_{\lambda\in M}\left(\boldsymbol{\zeta}_\lambda - \sum_{m\in S}\alpha_m\zeta_m k(\boldsymbol{\vartheta}_\lambda,\boldsymbol{\vartheta}_m)\right) \tag{25}$$

Now, m is an element of the set of support vectors S.

Using equation (24) and equation (25), the optimal identification boundary y is as follows.

$$\begin{aligned} y &= \hat{\boldsymbol{w}}_0^T\phi(\boldsymbol{\vartheta}_\lambda) + b_0\\ &= \sum_{\lambda=1}^{N}\alpha_\lambda\zeta_\lambda\phi(\boldsymbol{\vartheta}_\lambda)\boldsymbol{\vartheta} + \frac{1}{N_M}\sum_{\lambda\in M}\left(\zeta_\lambda - \sum_{m\in S}\alpha_m\zeta_m k(\boldsymbol{\vartheta}_\lambda,\boldsymbol{\vartheta}_m)\right)\end{aligned} \tag{26}$$

However, $k(\boldsymbol{\vartheta}_\lambda, \boldsymbol{\vartheta}_m)$ is the kernel function, and in this case, we applied the Gaussian kernel $k(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}')$ to the kernel function. The Gaussian kernel $k(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}')$ is expressed as.

$$k(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}') = \exp(\frac{\|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'\|}{2\sigma^2}) \tag{27}$$

The Gaussian kernel is a kernel function that gives the same effect as a decision by mapping to an infinite -dimensional space.

By using nonlinear SVM instead of linear SVM, complex feature distributions can be handled, thus enabling highly accurate decisions.

J. Decision algorithm considering time transition

Figure 2 shows a plot of the feature values using the sound and damage data as input data.

It can be seen that there are many overlapping features between the sound data of the healthy part and the sound data of the damaged part. This suggests that overlapping parts may be difficult to determine even with the best discriminating boundary, so a new feature that can separate healthy from damaged parts is needed.

As a result of analyzing the time trends of the features, we found that the judgment results of a certain frame are likely to be the same continuously. Therefore, from frame number λ, the dominant one in the constant frame set $\{F_i\}$ is considered as the judgment result as shown in Figure 3.

However, for this decision, the frame numbers from 1 to 5 are used as a constant frameset. Also, the number of seconds in a 1 frame is 0.046 seconds, so even for a fixed set of frames, 0.23 seconds, and the decision hammer is moving only slightly. If the judgment changes moment by moment in these 0.23 seconds, it means that there is a healthy part and a damaged part in a small section of the hammer movement, which is almost impossible. In order to prevent

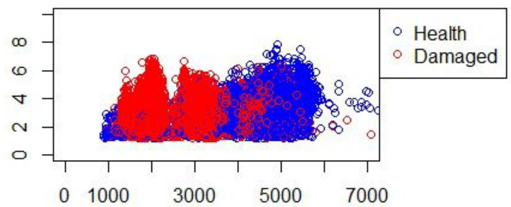

Fig. 2: Plotting features

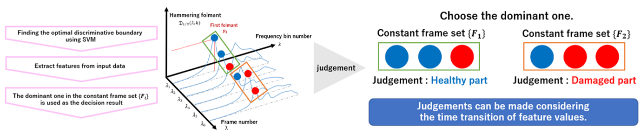

Fig. 3: Decision algorithm considering time transition

the judgment from changing from moment to moment within a very small movement interval, the judgment for the damaged part is canceled the second time when the judgment for the healthy part is made the first time and the judgment for the damaged part is made the second time, and the judgment for the damaged part is made when the judgment for the damaged part is made the third time. In addition, when the judgment of the healthy part and the damaged part is made alternately 3 times, the hammer moving section is judged as unknown as it is extremely difficult to judge.

This enables highly accurate judgments by correcting erroneous judgments that exist in a certain set of frames that are considered to represent residual noise or silence other than strikes in the sound data.

IV. Computer simulation

A. Noise suppression

The spectrograms before and after noise suppression are shown in Figure 4. First, in (i), we can see that there is a large orange noise in the low-frequency range and an orange noise that appears with time. Next, in (ii), we can see that noises that are mostly in the low-frequency range and time-varying noises such as car noises are suppressed.

In conclusion, it can be confirmed that low-frequency noise and time-varying noise are suppressed by the noise suppression method in this paper.

B. Determination of healthy and damaged parts

The cross-validation method was applied to evaluate the performance of the decision results by linear SVM and nonlinear SVM. For the experimental data, we prepared 3 of sound and damage data from 3 bridges for 30 seconds each. In this paper, a total of 7776 data (3888 data in the healthy part and 3888 data in the damaged part) is used as the teacher data. In order to improve the judgement accuracy, it would be better to increase the number of teacher data. However, since the emphasis of this hammering test system is on real-time, we decided to set the number of teacher data so that the learning time would not be a burden to the workers in the field. Using the prepared data, we evaluated the performance of the conventional method [3] and the proposed method.

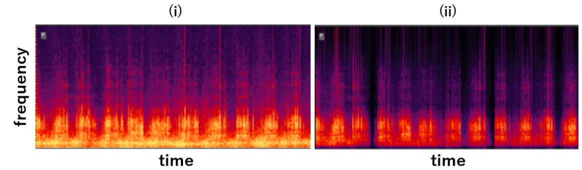

Fig. 4: (i)Before noise suppression, (ii)After noise suppression

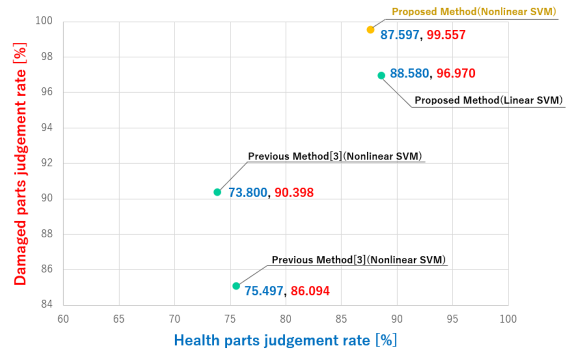

Fig. 5: Performance comparison

The result of the decision is shown in Figure 5.

From Figure 5, when the linear SVM was applied to the previous method, the healthy part was 75.5% and the damaged part was 86.1%. When the nonlinear SVM was applied, the healthy part was 73.8% and the damaged part was 90.4%. In comparison, when the linear SVM was applied to the proposed method, the healthy part was 88.6% and the damaged part was 97.0%. When the nonlinear SVM was applied, the healthy part was 87.6% and the damaged part was 99.6%.

These results show that the proposed method achieves higher decision accuracy than the previous method.

V. Conclusion

This paper proposes a nonlinear SVM-type automatic decision algorithm in noisy environment for hammering test system.

The features of the proposed method are (i) low-frequency noise suppression by a Butterworth type high-pass filter considering the frequency analysis of healthy part sound and damaged part sound data and noise data, and residual noise suppression by Kalman filter algorithm with a colored driving source to achieve high accuracy noise suppression. In addition, (ii) a judgment algorithm considering the time transition is used to realize a general-purpose and highly accurate percussion sound inspection system.

The effectiveness of the proposed method is demonstrated by comparing the spectrograms of the signals before and after noise suppression and by evaluating the generic performance of the cross-validation method for determining healthy and damaged parts.

Future work includes the development of a system that can accurately suppress noise and make automatic judgments even in a tunnel environment and the development of a hammering test system that can be applied to a wall lifting robot.

REFERENCES

[1] Research and Review Committee, "Report of the investigation and review committee on the falling tunnel ceiling plate accident", 2013.

[2] Y.sonoda, K.Kawabata, M.Beppu, Y.Fukui, "A fundamental study on the evaluation method of inner defect of concrete using hammering test data", Structural engineering papers Vol.57A, 2011.

[3] K.Hori, N.Tanabe, T.Furukawa, "Development of support system suitable for sound inspection", The 2020 IEICE Shinetsu Section Conference.

[4] R.Martin, "Spectral Subtraction Based on Minimum Statistics", in *Proc.Eur.Signal Processing Conf*, pp.1182-1185, 1994.

[5] M.Berouti, R.Schwartz, J.Mkhoul, "Enhancement of speech corrupted by acoustic noise", Jan. 2003.

[6] K.Yamashita, S.Ogata, T.Shimamura, "Improved Spectral Subtraction Utilizing Iterative Processing", The IEICE Transactions, A Vol.J88-A No.11 pp.1246-1257, 2005.

[7] M.Hino, "On a plain induction of Kalman's theory of predictive estimation", Tokyo Institute of Technology Civil Engineering Research Report, Technical Report, No.15 Dec. 1973.

[8] S.Adachi, I.Maruta, "Fundmentals of Kalman Filter", Tokyo Denki University Press, Mar. 2013.

[9] Nari TANABE, Toshihiro FURUKAWA, Shigeo TSUJII, "Robust Noise Suppression Algorithm with the Kalman Filter Theory for White and Colored Disturbance", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Volume E91.A, Issue 3, pp.818-829, 2010.

[10] S.Furui, "New Acoustics and Speech Engineering", Modern Science Society, 2006.

[11] C.M.Bishop, "Pattern Recognition and Machine Learning, Information Science and Statistics, Volume 1", Marusan Publishing, 2012.

[12] C.M.Bishop, "Pattern Recognition and Machine Learning, Information Science and Statistics, Volume 2" Marusan Publishing, 2012.

[13] R.Kohavi. "A study of cross-validation and bootstrap for accuracy estimation and model selection", Appears in the International Joint Conference on Artificial Intelligence, 1995.