A fast high-resolution vibration measurement method based on vision technology for structures

Ki-Sung Son*, Hyeong-Seop Jeon, Gyung-Sun Chae, Jae-Seok Park, Se-Oh Kim

SAE-AN Engineering Co., RM910, Byuksan Digital Valley II, 184, Gason Digital 2-ro, Geumcheon-gu, Seoul, 08501, Republic of Korea

ABSTRACT

Various types of sensors are used at industrial sites to measure vibration. With the increase in the diversity of vibration measurement methods, vibration monitoring methods using camera equipment have recently been introduced. However, owing to the physical limitations of the hardware, the measurement resolution is lower than that of conventional sensors, and real-time processing is difficult because of the use of extensive image processing. As a result, most such methods in practice only monitor status trends. To address these disadvantages, a high-resolution vibration measurement method using image analysis of the edge region of the structure has been reported. While this method exhibits higher resolution than the existing vibration measurement technique using a camera, it requires significant amounts of computation. In this study, a method is proposed for rapidly processing considerable amounts of image data acquired from vision equipment, and measuring the vibration of structures with high resolution. The method is then verified through experiments. It was shown that the proposed method can fast measure vibrations of structures remotely.

1. Introduction

Periodic inspections of surface degradation and analysis of vibration data are performed for structures at industrial sites to monitor their status and diagnose safety. Vibration in a structure is one of the factors that affect its structural safety; as such, the main method to examine safety is through measurement and analysis of the vibration. However, although conventional vibration measurement methods are useful for obtaining information on the status of the structure, the system is complex because of the utilization of several sensors, and requires an experienced inspector. Furthermore, it is difficult to measure vibration if the inspector cannot easily access or attach the sensor, as in the case of high-rise or high-temperature structures. Vibration monitoring methods using camera equipment have recently been introduced to address these difficulties. Vibration measurement using cameras enables monitoring over a large area at the same time.

However, owing to the physical limitations of the camera hardware, the measurement resolution is lower than that of conventional sensors. In addition, as vast amounts of image data must be processed, it is difficult to perform real-time processing and rapidly display the measurement results. Consequently, rather than performing quantitative measurements, vibration monitoring using cameras at industrial sites only monitors status trends. Techniques for enhancing visual vibration perception for vibration monitoring have recently been proposed [1–4].

A method for quantitatively measuring vibration at high resolutions was also introduced. It involves detecting real edges through second-order differential calculations of the structure’s edge images and tracking the detected edges, thereby measuring the vibration [5]. This method addresses the hardware resolution limitations of image pixels in the form of an integer grid. However, owing to degraded resolution caused by the contrast ratio of the edge regions of the structure and extensive computations, it is difficult to display immediate results. Therefore, a method to address these drawbacks is proposed in this study. The proposed method involves a technique for enhancing the measurement resolution by detecting features such as lines and edges, and a high-speed computation technique that simplifies the image processing algorithm. The proposed method is then verified through experiments to measure the accuracy and the results of field tests.

* Corresponding author.
E-mail address: infoson@sae-an.co.kr (K.-S. Son).
2. Vision-based vibration measurement method

The vibration measurement method employing a vision sensor requires a process of tracking the measurement points by continuously analyzing the image data. The measurement points can be referred to as feature elements such as points, lines, surfaces, and colors that appear in the image. Normally, the conventional vision-based vibration measurement method involves selecting high-frequency regions of the image where brightness changes rapidly with changes in the measurement points, such as the sharp lines and edges of an object.

The most basic method for vision-based vibration measurement involves tracking the amount of movement of a measurement point in units of integer pixels. To select the feature region, the user can involve tracking the amount of movement of a measurement point based on vibration measurement method involves selecting high-resolution compared with the conventional method.

2.1. Basic theory

As shown in Fig. 1, the spot energy is typically represented by a Gaussian probability distribution. This phenomenon is the same as when a vision sensor receives light energy in the form of dots and lines. When the center of the Gaussian probability distribution is ??, the largest value is in the center ??; therefore, the maximum value must be found to detect the center.

Hence, vibration can be measured by detecting and tracking the center ?? of the feature image showing a spot energy distribution such as points and lines.

As shown from the points and lines, the edge region exhibits a blur phenomenon owing to light scattering. As shown in Eq. (4), the second-order derivative function. Generally, mask convolution method of second-order derivative method is used to derive the same results as the first-order derivative function. For the edge region, the second-order mask convolution method is used for the second-order derivative function. Generally, mask convolution method of the 2D image is as follows. When the 2D image is \( f(x, y) \), a mask of size \( m \times n \) is \( w(m, n) \), and the result is \( g(x, y) \), which can be expressed in Eq. (3).

\[
g(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x + s, y + t)
\]

where, \( a = (m - 1)/2 \), \( b = (n - 1)/2 \).

For vibration measurement, the mask convolution operation is performed through the following procedure.

1. Acquire images for vibration measurement using a camera.
2. Set the measurement point setting in acquired images.
3. Detect of real-type pixels coordinates by using proposed mask convolution method.
4. Unit conversion between pixel coordinate value and real(mm) value.

To detect the real center ?? from the mask convolution results, a technique using the two coordinate points closest to the zero-crossing point (center ??), as in the existing mean of the multi-interval second-order derivative method is used.

2.2. First-order derivative mask convolution method for line center detection

As explained previously, the lines in the image exhibit a Gaussian probability distribution, and the first-order derivative is used to detect the center of the line. As shown in Eq. (4), the first-order derivative is used to obtain the difference from neighboring pixel values to determine the slope between the pixels. Common techniques for detecting the edges in an image using first-order derivatives include the Sobel, Prewitt, and Roberts methods [13].
where \( x \) is the vertical-axis coordinate value of the image, \( y \) is the horizontal-axis coordinate value, \( f \) is the input image, and \( G \) is the first-order derivative of \( f \). If a general first-order derivative is used, as shown in the aforementioned equation, then the noise of a specific pixel can considerably impact the result.

Therefore, in this study, a derivative using the mean value of the surrounding region is employed, making it robust against random noise (Eq. (5)).

When \( L(x) \) is the first-order derivative function of the one-dimensional image \( I(x) \), \( T \) is a constant for determining the range of the surrounding region to calculate the mean.

\[
G(x) = f(x + 1, y) - f(x - 1, y) \\
G(y) = f(x, y + 1) - f(x, y - 1) \\
\] (4)

The mask convolution operation is highly effective for quickly processing the proposed equation during image processing. When \( T = 3 \), Eq. (5) can be modified as the following Eq. (6).

\[
L(x) = \frac{1}{7} \sum_{n=1}^{T} \left\{ \frac{I(x + n) - I(x - n)}{2n} \right\} \\
\] (5)

During image processing, multiplication and division operations of the entire image only change the scale and do not impact feature detection. Hence, this can be simplified into Eq. (7).

\[
L(x) = -2 \cdot I(x - 3) - 3 \cdot I(x - 2) - 6 \cdot I(x - 1) + 6 \cdot I(x + 1) + 3 \cdot I(x + 2) + 2 \cdot I(x + 3) \\
\] (7)

The multiple at each pixel position is a set of matrices used in the mask convolution operation.

The process is shown in Fig. 3. The features in the image can be detected by generating a mask and executing a convolution operation.

Fig. 4 shows an example of the results from the first-order derivative mask convolution method in a one-dimensional region perpendicular to the line in an image, with black lines on a white background (Fig. 3.).

As a result of mask convolution, the coordinates closest to the zero-crossing point \( P_0, P_1 \) in \( L(x) \) are determined, and the real coordinates \( \hat{x}, \hat{y} \), the center of the line, are detected using Eq. (8).

\[
\mu = \frac{-L(P_0)}{L(P_1) - L(P_0)} + P \\
\] (8)

2.3. Second-order derivative mask convolution method for edge center detection

Existing methods for edge detection using the second-derivatives include the Laplacian and Laplacian of a Gaussian [13].

The Laplacian differential operator that defines the second-order derivative in the image is defined as the sum \( \nabla^2 f \) of the second-order derivatives in the \( x \) direction and \( y \) direction.

\[
\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \\
\] (9)

\[
\frac{\partial^2 f}{\partial x^2} = f(x + 1, y) + f(x - 1, y) - 2f(x, y) \\
\frac{\partial^2 f}{\partial y^2} = f(x, y + 1) + f(x, y - 1) - 2f(x, y) \\
\] where \( x \) is the vertical-axis coordinate value of the image, \( y \) is the horizontal-axis coordinate value, \( f \) is the input image, and \( \frac{\partial f}{\partial y} \) is the second-order derivative of the horizontal axis with respect to \( f \). The
Laplacian operator that defines the second-order derivative in the image uses the second-order derivative to determine the integer coordinates closest to the zero-crossing point as an edge.

In this study, the basic concept is utilized for detecting the center coordinates of the edge using the second-order derivative as zero-crossing detection, for which the following transformation formula can be employed such as Eq. (10). The difference between the short-term average (STA) value in an edge region of the image and long-term average (LTA) value in a large region can be used to obtain a result in a similar form to the second-order derivative. Compared with the existing mean of the multi-interval second-order derivative method, this technique can derive results rapidly with fewer computations, and it is also strong against noise because it uses the mean value of the surrounding pixels. When the one-dimensional image is $I(x)$, the pixel range for LTA and STA is $LT$ and $ST$, respectively. The result $E(x)$ can be expressed as in Eq. (10).

$$E(x) = \frac{1}{2LT + 1} \sum_{n=-LT}^{LT} I(x+n) - \frac{1}{2ST + 1} \sum_{m=-ST}^{ST} I(x+m)$$  

(10)

A set of matrices for the mask convolution operation is searched in Eq. (10).

When $LT = 3$ and $ST = 1$, this can be expressed as Eq. (11).

$$E(x) = \frac{3 \cdot I(x - 3) + 3 \cdot I(x - 2) - 4 \cdot I(x - 1) - 4 \cdot I(x) - 4 \cdot I(x + 1) + 3 \cdot I(x + 2) + 3 \cdot I(x + 3)}{21}$$  

(12)

The multiple at each pixel position is a set of matrices used in the mask convolution operation.

The process is shown in Fig. 5. The edge region features in the image can be detected by generating a mask and executing a convolution operation.

Fig. 6 shows the result of second-order derivative mask convolution for a one-dimensional image perpendicular to the edge in the edge image. The zero-crossing point can be detected using Eq. (8), and as a result of mask convolution, the center $??$ of the edge can be detected as a real value in $E(x)$.

3. Verification of proposed technique

3.1. Laboratory experiment

To verify the proposed technique, an experiment to evaluate accuracy was performed using an shaker and user markings. As shown in Fig. 7, the target structure was attached to the exciter, and excitation was performed at 1 Hz. The vibration was measured using an industrial charge-coupled device (CCD) camera, which acquired video data at 300 frames/s at a resolution of $640 \times 480$. The lighting was kept constant using a lighting device. The proposed technique was developed in Visual Studio C++.
For an accurate comparison, user markings in a grid pattern were used, and vibration was measured through the line detection mask convolution method. The distance between the measurement point and the camera was set to 2 m.

A 50-mm optical lens was used for the camera, and the data of the laser displacement sensor was acquired at 1024 samples/s at a 16-bit resolution.

Fig. 8 shows the vibration measurements with a 1-Hz excitation. The average peak-to-peak (P-P) of the laser displacement sensor was 200.00 μm, and the root mean square (RMS) for 5 s was 62.65 μm. The average P-P of the proposed method was 203.90 μm, and the RMS for 5 s was 62.33 μm. According to a comparison of the two signals, with an average difference in P-P of 3.90 μm, the proposed method showed an error rate of 1.95% compared with the laser displacement sensor, and an error rate of 0.51% based on a difference of 0.32 μm in RMS.

The size of the unit pixels measured in this experiment was 171 μm/pixel; this was measured using the distance between the top and
bottom of the user-marked region. Hence, the actual vibration displacement size P-P of 200.00 μm can be considered 1.17 pixels in the image.

The laser displacement sensor used in this experiment has an error of 1 μm when measured at a distance of 50 mm. The proposed method showed high accuracy considering what was measured at a distance of 2 m.

3.2. Pipe vibration experiment using test bed

A shaker was attached to the pipe test bed, as shown in Fig. 9, to perform the experiment. The experimental equipment, including the industrial CCD camera, was configured the same way as in Section 3.1. Attaching the grid-pattern user markings to the pipe, the first-order derivative mask convolution method for line detection and second-

---

**Fig. 6.** Example of second-order derivative mask convolution of one-dimensional image perpendicular to edge in image.

**Fig. 7.** Experimental setup for accuracy evaluation of 1-Hz vibration status.
order derivative mask convolution method for edge detection using the upper edge region of the pipe were used for measurement. To accurately examine the proposed method, the laser displacement sensor was installed just above the user-marked region.

Experiments were also conducted to determine the difference in computation time between the existing mean of the multi-interval second-order derivative method, which measures vibration using edge detection, and the proposed second-order derivative mask convolution method.

Fig. 10 shows the vibration measurements using the first-order derivative mask convolution method for the user-marked region. The RMS of the laser sensor was 29.05 μm, and that of the proposed method was 29.59 μm, a difference of 0.54 μm. This represents an error of 1.9% based on the laser sensor, which is within the error range of the laser sensor. Fig. 11 shows the frequency analysis results on a log scale. According to the results, peaks appeared at 24, 36, and 48 Hz, which are the harmonic frequencies of 12 Hz, and the sizes in each peak region were observed to coincide with the laser sensor.

Fig. 12 shows the vibration measurements using the second-order derivative mask convolution method for the edge region of the structure. At an RMS of 28.25 μm, there was an 8.01-μm difference with the laser sensor. This corresponds to an error of 2.8% based on the laser sensor, demonstrating a larger error than the method with user markings. Fig. 13 shows the frequency analysis results on a log scale. The sizes in the peak regions were shown to coincide with the laser sensor. Unlike line detection mask convolution, edge detection mask convolution measures vibration based on the edge region images of the structure. As a result, there are more error factors
owing to the contrast ratio and light scattering in the edge region of the structure, although this method is highly applicable to structures in which user markings are difficult to attach.

Next, using the edge regions at the top part of the pipe, the difference in computation speed between the existing mean of the multi-interval second-order derivative method, which measures vibration using edge detection, and the proposed second-order derivative mask convolution method was evaluated.

As region-based analysis techniques, both the existing and proposed methods require a size of at least $15 \times 15$, and the size of the measurement area must be varied according to the color, brightness, light amount, surface homogeneity, and vibration amplitude of the structure, which are characteristics of the actual object to be measured. Therefore, vibration measurement in actual sites using a large area may sometimes be more accurate.

In terms of the CPU specifications of the system for measuring performance in this experiment, a 6-core, 3.7-GHz Intel i7-8700k was used. The computation speed was measured while increasing the measurement point size for vibration measurement to $15 \times 15$, $20 \times 20$, $25 \times 25$, and $30 \times 30$. According to the results, the vibration measurements for each size were within the error range of the measurements from the previous experiment and the 1.0-μm RMS error for both the existing and proposed methods.

Fig. 14 shows a comparison of the average computation speed at 1000 frames. The existing method exhibited an average computation time of 0.642 ms for a measurement point size of $15 \times 15$, while the proposed method was approximately 0.049 ms, a 13-fold difference in processing speed. At a size of $30 \times 30$, the existing method showed a time of 1.367 ms and the proposed method 0.071 ms, a 19-fold difference in processing speed. This demonstrates that the difference in computation speed increases as the measurement point grows in area. Considering the image acquisition time, memory access time, and CPU usage, the existing method could perform real-time measurement of 1 point with a size of $15 \times 15$ at 200 frames/s, while the proposed method could perform real-time measurement of at least 5 points with a size of $15 \times 15$ at 300 frames/s. Thus, the proposed method was able to measure vibration in real time from continuously acquired images.
3.3. Field application

As shown in Fig. 15, a field application experiment was conducted using grid-pattern user markings. The measured object was an air circulation exhaust pump, with vibration generated by the rotational movement of the pump motor. As this is where the vibration problem occurs, a damper was attached to the bottom of the pump to reduce vibration. A 50-mm optical lens was used with a distance between the measured object and camera of approximately 5 m. To measure vibration, images were acquired at 300 frames/s at a resolution of 640 × 480. The proposed method was used to simultaneously measure biaxial vibration in real time from the acquired continuous images.

As shown in Fig. 16, according to the vibration measurements in the vertical direction, the RMS was 0.053 mm, and the vibration frequencies showed peaks at 2.0, 3.2, and 5.2 Hz. In the horizontal direction, the RMS was 0.054 mm, and the vibration frequencies showed peaks at 3.8, 7.8, and 29.8 Hz. This confirmed that the vibration frequencies were below 50 Hz, demonstrating that this method can be applied in the field with even a 200 frames/s camera for large structures such as that in this experiment.

4. Conclusion

In this study, a vibration measurement method using a vision sensor was proposed. To measure vibration through the vision sensor, first- and second-order derivative algorithms were developed to effectively detect the features of the measurement points in the image, such as points, lines, and edges, and a mask convolution method was proposed to quickly compute them.

A variety of experiments were conducted to verify the proposed method. The shaker experiment demonstrated that the proposed method exhibits higher measurement resolution than the laser displacement sensor. The pipe test-bed experiment confirmed the feasibility of measuring vibration through user markings and structure edges as the features, and showed an improvement in processing speed through the mask convolution method. The results of this study also confirmed the application feasibility of the proposed method in the field.

Techniques using cameras and other vision sensor equipment have recently been developed in various fields, among which the proposed technique can be applied to nondestructive inspection and structural status monitoring.

This technique is useful for the difficult to approach like a dangerous area in nuclear power plants.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2019-0-01746, “Development of Vision-based Intelligent Monitoring System for Structural Safety Diagnosis”).

References


