High-speed vision measurement of vibration based on an improved ZNSSD template matching algorithm

Jian Luo\textsuperscript{a,b}, Bingyou Liu\textsuperscript{a,b}, Pan Yan\textsuperscript{c} and Xuan Fan\textsuperscript{d}

\textsuperscript{a}College of Electrical Engineering, Anhui Polytechnic University, Wuhu, People’s Republic of China; \textsuperscript{b}Key Laboratory of Advanced Perception and Intelligent Control of High-End Equipment, Ministry of Education, Wuhu, People’s Republic of China; \textsuperscript{c}Anhui Dar Intelligent Control System Stock Co., Ltd., People’s Republic of China; \textsuperscript{d}Wuhu HIT Robot Industrial Technology Research Institute Co., Ltd., People’s Republic of China

ABSTRACT
This paper proposes an improved zero-mean normalization sum of squared differences (ZNSSD) algorithm to solve the problem of the inability of traditional structural measurement to extract high-frequency vibration signals. In the proposed technique, the high-speed image sequence of target vibration is captured by a high-speed camera. Then, the ZNSSD template matching algorithm with subpixel accuracy is introduced to process the captured images in the computer. Additionally, a modified search algorithm, the ZNSSD template matching algorithm based on image pyramid (ZNSSD-P), is proposed to significantly reduce the computation time and increase efficiency. Then, a jumping ZNSSD template matching algorithm based on image pyramid (J-ZNSSD-P) is proposed to further improve the efficiency of the ZNSSD-P algorithm. Vibration signals were extracted with Grating Ruler Motion Platform and sound barriers. Results show that the vibration signal extraction method has high precision and efficiency.

1. Introduction
Vibration, a common, natural phenomenon that exists in daily life and industrial production, is also an important topic in the field of mechanics research (see Zhang, 2019). The inevitability of vibration in engineering practice might cause harmful effects. Therefore, many foreign and domestic scientists have conducted related studies on vibration measurements that are widely used in various engineering practices (see Peng et al., 2020).

Vibration measurement algorithms are divided into contact and non-contact types. Contact-type vibration measurement algorithms require the use of contact sensors. An additional mass effect is present given that the contact vibration measurement can change the structural characteristics of the contact. Therefore, measurement error may occur when measuring objects with a light structure. Additionally, several sensors must be installed on the measured structure surface when measuring vibration data for large-scale structures. However, this setting will consume a large amount of manpower and material resources (see Poozesh et al., 2017). In addition to the problems mentioned above, the contact measurement algorithm is restricted by the measurement environment and other factors. In comparison, the contactless measurement algorithm is not affected by the influence of the load or the stabilization limitations of the sensor.

Non-contact measurement includes algorithms based on vision (see Aoyama et al., 2018; Wu et al., 2018) and the Doppler effect (see Liu et al., 2015), see With the rapid development of vision sensing, computer, and image processing technologies, the vision-based vibration measurement method has gradually become the mainstream non-contact measurement method (see Guo & Zhu, 2016; Wang, 2016). In recent years, extensive applications of vision measurement technology have been conducted in large-scale structure vibration detection, and satisfactory experimental results have been achieved (see Bell et al., 2012). For example, (2003) used a high-precision camera to measure the vertical displacement of the pier of Vincent Thomas Bridge in Los Angeles, California, United States of America through the frequency spectrum analysis of the displacement signal. The first two-order modal frequencies of the bridge vibration were obtained, thereby indicating that visual measurement is an effective way for completing the vibration measure-
measurement of large-scale structures. Feng and Feng (2017) used the camera to monitor the structural displacement response, successfully measured the natural frequency and mode shape of the object, and confirmed the effectiveness of the method by measuring the displacement response, successfully measured the natural frequency of large-scale structures. Feng and Feng (2017) used the camera to monitor the structural displacement measurement of large-scale structures. Feng and Feng (2017) combined high-speed camera technology with ultraviolet (UV) band-pass filter to obtain high-contrast ultraviolet images under different welding conditions. They also studied the relationship between the characteristics of the flame and spatter during welding and welding quality and realized the monitoring of working process stability in metal welding. (2016) applied high-speed vision technology to the vibration displacement measurement of underwater environment, extracted the amplitude information of underwater structures under seismic excitation through high-speed video, analyzed the effect of underwater environment on structural vibration, and calculated the natural frequency and damping of the structure. (2016) of the University of Science and Technology of China conducted in-depth research on the information extraction algorithm for high-speed visual measurement systems and applied it to the vibration analysis of forklift steering wheel systems, sound signal measurement and recovery, and bolt damage detection.

However, traditional image processing technologies used in high-speed cameras not only have high requirements for motion tracking algorithm, these technologies are also complex and troublesome (see Aqel et al., 2016). The standard method for tracking features between two images is through template matching in computer vision. The basic template matching algorithm calculates each position through a distortion function, which then measures the similarity between the template and the image. The area-based algorithm involves calculating a correlation function at each location of the search image in a raster scan manner. Several area-based methods have been introduced over the years, including mean absolute difference, product cross correlation (CC), sequential similarity detection algorithm, sum of absolute difference (SAD), and sum of variance (SSD) similarity measurement (see Sahani et al., 2011). After considering template matching, due to its better robustness, the zero-mean normalization sum of squared differences (ZNSSD) algorithm is often used in similarity measurements (see Bing et al., 2010; Pangercic et al., 2008). Compared with other algorithms, ZNSSD is less sensitive to the linear change of the illumination amplitude in the two comparison images, making it suitable for actual measurement.

However, a more efficient fast Fourier transform (FFT) cannot be used to directly calculate ZNSSD in the spectral domain. The efficiency of the ZNSSD algorithm is very low, because most thorough search operations require expensive computing costs. In the field of video compression, various algorithms, such as image pyramid decomposition algorithm (see Malik & Soundararajan, 2019), block-matching algorithm (see Muhammad et al., 2020), and so on, have been proposed to improve efficiency.

The sampling frequency of high-speed cameras for structural vibration is relatively high, and the displacement between each image frame is only a few pixels or even just one pixel. Therefore, according to the previous frame in template matching, the matching position can only be found within a very small range. Owing to these characteristics, the current paper proposes an improved ZNSSD for displacement measurement to solve the problem of traditional image processing technology. The improved algorithm can obtain the same results of vibration extraction, but the calculation time of one extraction is much less than that of the ZNSSD algorithm, which provides the potential of online extraction for high-speed captured video. The algorithm can be used to extract displacement by displacement and test displacement on a plane. Finally, the algorithm was applied to the high-speed camera system to realize the vibration measurement. Two experiments were carried out in laboratory and outdoor conditions to verify the performance. The results show the accuracy and efficiency of the camera system.

2. ZNSSD extraction algorithm

The ZNSSD algorithm uses a correlation value to represent the similarity between two images. This algorithm is a matching algorithm to find specific objects in the image. The value of the ZNSSD correlation is calculated
ZNSSD = \[
\frac{\sum_{i=1}^{w} \sum_{j=1}^{h} G(i,j) - \bar{G}(i,j)}{\sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [G(i,j) - \bar{G}(i,j)]^2}} - \frac{\sum_{i=1}^{w} \sum_{j=1}^{h} T(i,j) - \bar{T}(i,j)}{\sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [T(i,j) - \bar{T}(i,j)]^2}}
\]

= \frac{w}{\sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [G(i,j) - \bar{G}(i,j)]^2}} \left( \frac{\sum_{i=1}^{w} \sum_{j=1}^{h} [G(i,j) - \bar{G}(i,j)]^2 - 2 \sum_{i=1}^{w} \sum_{j=1}^{h} [G(i,j) - \bar{G}(i,j)] * [T(i,j) - \bar{T}(i,j)]}{\sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [T(i,j) - \bar{T}(i,j)]^2}} \right) + \frac{\sum_{i=1}^{w} \sum_{j=1}^{h} [T(i,j) - \bar{T}(i,j)]^2}{\sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [T(i,j) - \bar{T}(i,j)]^2}} \right)

= 2 - \frac{\sum_{i=1}^{w} \sum_{j=1}^{h} [G(i,j) - \bar{G}(i,j)] * [T(i,j) - \bar{T}(i,j)]}{\sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [G(i,j) - \bar{G}(i,j)]^2}} \sqrt{\sum_{i=1}^{w} \sum_{j=1}^{h} [T(i,j) - \bar{T}(i,j)]^2}

(1)

where, \( \bar{G} = (\sum_{i=1}^{w} \sum_{j=1}^{h} G(i,j))/(w \times h) \); \( \bar{T} = (\sum_{i=1}^{w} \sum_{j=1}^{h} T(i,j))/(w \times h) \), \( T(i,j) \) are the grey value of \( (i,j) \) in the template image \( T \); \( G(i,j) \) is the grey value of \( (i,j) \) in the image \( G \) to be matched. Only considering the direct relationship between the corresponding grey levels will make this correlation description very sensitive to illumination change and image noise. The zero-mean normalization operation of the image can reduce the sensitivity of the correlation description function to illumination and noise, that is, ZNSSD function is insensitive to the linear change of image grey and has strong anti-interference ability.

Given a matching image \( G \) and a template image \( T \), the area of \( G \) template \( T \) is scanned using the area of template \( T \) until the largest correlation with \( T \) or the smallest error is determined. As shown in Figure 1, \( T \) with a size of \( W \times H \) overlays on \( G \) with a size of \( W \times H \). The image is scanned by the full picture, and the correlation coefficient of \( T \) and the corresponding sub-region of \( G \) are calculated to obtain a description template for \( T \) and \( G \). The minimum value of the correlation coefficient matrix of the matching degree of each region is in the region of \( G \), with the highest matching degree with \( T \).

A simple simulation example is provided to validate the performance of the ZNSSD algorithm. In the simulation, let the black circle in Figure 2 move along a pre-determined path. The ZNSSD algorithm is then applied to extract the path coordinates. For conciseness, the motion equations, whose shape is an ellipse, are expressed as follows:

\[
x(t) = 10 \cos(2\pi ft), y(t) = 6 \sin(2\pi ft)
\]

where the vibration frequency \( f = 1 \) Hz and the sampling frequency \( f_s = 50 \) Hz. In this study, black circle simulation experiments were conducted to verify the speed and accuracy of the algorithm. The ZNSSD algorithm successfully extracted the motion signals in the horizontal and vertical directions. The horizontal signal has a wave shape similar to the simulation input \( x \), whereas the vertical signal is similar to input \( y \). Thus, the results of the ZNSSD algorithm match well with the real input.

Given that the algorithm can only determine the extraction effect of the entire pixel displacement, the displacement in the actual measurement is not exactly the entire pixel. Therefore, the extraction result of the pixel can greatly affect the positioning accuracy of the vision measurement system and may not obtain a good application effect in precision measurement. To address the above mentioned issue, we used the surface fitting
Figure 2. Simulation example based on the ZNSSD extraction algorithm. Where, Figure a is the selected matching area; Figure b is the display of the region of interest selected in Figure a; Figure c is the vibration displacement in the X direction extracted by the ZNSSD algorithm; Figure d is the vibration displacement in the Y direction extracted by the ZNSSD algorithm.

(see Gu et al., 2020) algorithm to improve the extraction accuracy.

3. Modified algorithm

3.1. Subpixel accuracy improvement

A smooth surface exists near the maximum point within a correlation coefficient matrix. If the best matching position of the template is not in the position of the entire pixel, the extreme point ancestor of the surface will not appear in this exact position, but near it. The location of the extreme point is the best matching position of the image. Therefore, if the surface fitting is completed by the correlation coefficient matrix, the ideal peak point of the surface can be directly calculated in accordance with the surface function obtained through fitting to obtain a high-precision subpixel displacement estimation. The surface fitted by the correlation coefficient in different cases exhibits great differences. To reduce the calculation amount and increase the practicability of the fitting surface, we select the correlation coefficients of the maximum position and the surrounding eight points.

Given the relatively gentle change in the correlation coefficient surface, we use the bivariate quadratic polynomial to obtain the surface. Assuming that the matching position of the entire pixel is \((x_0, y_0)\), then the fitting surface can be expressed as follows:

\[
f(x, y) = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 \quad (3)
\]

Once the correlation coefficient fitting surface is determined, the peak value of the surface can be used to calculated the location of the extreme point of the quadratic fitting surface.

\[
x = \frac{2a_1a_5 - a_2a_4}{a_4^2 - 4a_3a_5}, \quad y = \frac{2a_2a_3 - a_1a_4}{a_4^2 - 4a_3a_5} \quad (4)
\]

Below is a simple example to illustrate the calculation process and the actual effect of the surface fitting sub-pixel algorithm. We considered a \(3 \times 3\) correlation coefficient matrix in the following equation:

\[
A = \begin{bmatrix}
  f(-1, 1) & f(0, 1) & f(1, 1) \\
  f(-1, 0) & f(0, 0) & f(1, 0) \\
  f(-1, -1) & f(0, -1) & f(1, -1)
\end{bmatrix}
\]
Figure 3. Example of the quadratic surface fitting: (a) a matrix around the best-matching location; (b) the fitting quadratic surface is displayed in the image form and the triangle marker is the better estimation of the best-matching location.

\[
\begin{bmatrix}
0.50 & 0.63 & 0.26 \\
0.75 & 1.12 & 0.42 \\
0.55 & 0.84 & 0.36 \\
\end{bmatrix}
\]

(5)

Figure 3 depicts the correlation coefficient matrix, and the value is represented by the difference in colour. The quadratic surface fitting algorithm is used to fit the surface. The values in the figure are the sampling values fitted at the corresponding positions. The results show that these values are extremely close to the original nine values. The peak coordinates of the fitting surface in Figure 3 are \((-0.1605, 0.1208)\). After the surface fitting operation, the subpixel calculation result of the optimal matching position is \((x_0 - 0.1605, y_0 + 0.1208)\). The positioning accuracy obtained through surface fitting is 0.005–0.01 pixels, and it has the characteristics of high noise resistance and high computational efficiency. Thus, the technique can be applied in actual image matching.

3.2. Matching speed improvement

3.2.1. ZNSSD-P template matching algorithm

The ZNSSD template matching algorithm is used to match and recognize the k-frame image. First, the k-frame image is decomposed into a pyramid, and the corresponding layer target is searched from the top image \(G_k^n\) with low resolution. Second, the target centre coordinate \(F_k^m(i_n, j_n)\) of this layer is determined using the ZNSSD algorithm and then mapped to the high resolution image \(G_k^{n-1}\) in the next layer. The new target centre coordinate \(F_k^{m-1}(2i_n, 2j_n)\) is obtained. Finally, the ZNSSD algorithm is used to search the target of \(G_k^{n-1}\) in the nearby mapping area \(Q_k^{n-1}\). Thus, \(F_k^{m-1}(i_n-1, j_n-1)\), which is the revised target centre coordinates. Continuously cycle the above process, coordinates \(F_k^0(i_0, j_0)\) is subsequently determined (as shown in Figure 4).

\[
G_k^{d+1}(i, j) = \sum_{p=-2}^{2} \sum_{q=-2}^{2} W(p, q)G_k^d(2i + p, 2j + q)
\]

(6)

For \(0 < d < n\), \(W(p, q)\) is a window function with low-pass characteristics that can be defined as follows:

\[
W(p, q) = \frac{1}{256}
\begin{bmatrix}
1 & 4 & 6 & 4 & 1 \\
4 & 16 & 24 & 16 & 4 \\
6 & 24 & 36 & 24 & 6 \\
4 & 16 & 24 & 16 & 4 \\
1 & 4 & 6 & 4 & 1
\end{bmatrix}
\]

(7)

Equations (8)–(10) is the process of finding target points in the n-th layer, as follows:

\[
G_n(i, j) = \sum_{i=1}^{w\times2^{-n}} \sum_{j=1}^{h\times2^{-n}} G(i, j)
\]

(8)

\[
T_n(i, j) = \sum_{i=1}^{w\times2^{-n}} \sum_{j=1}^{h\times2^{-n}} T(i, j)
\]

(9)

\[
ZNSSD = \left[ \frac{\sum_{i=1}^{w\times2^{-n}} \sum_{j=1}^{h\times2^{-n}} G_n(i, j) - \bar{G}_n(i, j)}{\sqrt{\sum_{i=1}^{w\times2^{-n}} \sum_{j=1}^{h\times2^{-n}} [G_n(i, j) - \bar{G}_n(i, j)]^2}} \right]
\]
\[\sum_{i=1}^{\frac{h \times 2^{-n}}{n}} \sum_{j=1}^{\frac{h \times 2^{-n}}{n}} \left( T_n(i,j) - \bar{T}_n(i,j) \right)^2 \]

\[\sum_{i=1}^{\frac{h \times 2^{-n}}{n}} \sum_{j=1}^{\frac{h \times 2^{-n}}{n}} \left( T_n(i,j) - \bar{T}_n(i,j) \right)^2 \]

where, \( G_n(i,j) \) represents the \( n \)-th layer of the matching image, and \( T_n(i,j) \) represents the \( n \)-th layer of the template image, \( G(i,j) \) represents the original image, \( T(i,j) \) represents the template image.

The calculation times are as follows:

\[N = \frac{1}{2^n} \times WH + 8n.\]

Compared to the calculation of \( W \times H \) times under the global search strategy, the image pyramid construction significantly reduces the matching calculation times and improves the matching efficiency.

\[m = \text{Round}((k_1(s - p) + k_2c)),\]

where \( m \) is the number of skip layers, \( k_1 \) and \( k_2 \) are the scale coefficients, \( s \) is the similarity score between the target and the template in the current layer (full score of 100), \( p \) is the matching skip threshold, \( c \) is the position of
the current pyramid layer, and Round represents \( k_1(s - p) + k_2c \) rounded off to the nearest integer.

The new target centre \( F_{n-m}^{k-1}(2^mi_n,2^mj_n) \) is obtained by mapping \( F_{k}^{n}(i_n,j_n) \) to the \( n - m \) pyramid image \( G_{n-m}^{k} \) after the jump. In the mapping area \( H_{n}^{n-m} \) centered on the coordinate \( F_{n-m}^{k} \), the ZNSSD algorithm is used to re-search the target of this layer. Thus, \( F_{n-m}^{k}(i_{n-m},j_{n-m}) \) can be quickly found, which is the target centre coordinate of the corrected image \( F_{n-m}^{k} \). Then, the process is continued until \( F_{0}^{k}(i_0,j_0) \) is obtained. If \( n - m < 0 \) appears in the calculation process, it jumps to the bottom image; if \( m < 0 \) appears in the calculation process, it does not jump and maps directly to the next layer. In this paper, the values of \( k_1, k_2, \) and \( p \) are 0.02, 0.25, and 70, respectively.

### 3.2.3. General flow chart of the J-ZNSSD-P matching algorithm

After capturing the high-speed image sequence of a vibrating object by using a high-speed camera, the first frame of the image is taken, and the region of interest (ROI) is selected. Subsequently, the J-ZNSSD-P algorithm is used to search for the target and determine its centre coordinates. The surface fitting algorithm is then utilized

![Image](image.png)

**Figure 5.** Overall flowchart of the vibration extraction.
to improve the extraction accuracy. The specific search process is shown in Figure 5.

### 3.3. Performance analysis of the modified algorithm

To verify the performance of the modified algorithm, we conducted the improved black circle experiments that adopted the same test conditions in Subsection 3.1. The experimental results are shown in Figure 6 and Table 1.

In the test, this paper selects four methods for comparison, which are the ZNSSD matching method, J-ZNSSD-P matching method and UCC subpixel matching algorithm. UCC algorithm is an advanced subpixel image registration technology, which allows the resolution to be adjusted by changing the up sampling factor. The up-sampling factors of one integer pixel and 0.01 pixel in UCC algorithm are set to 1 and 100 respectively. Figure 6 shows the results of vibration extraction by the above four methods. The results of ZNSSD and UCC(usfac = 1) are shown in Figure 6(a) and (c), and the results of J-ZNSSD-P and UCC(usfac = 100) are shown in Figure 6(b) and (d). From Figure 6, we can see that the matching accuracy of J-ZNSSD-P and UCC (usfac = 100) is significantly higher than the other two methods, and the extracted curve is smoother.

Quantitative contrast results regarding tracking error and computation time are given in Table 1. The table indicates that with the improvement of subpixel resolution level from 1 to 0.01 pixel, the absolute average horizontal error of the UCC algorithm reduces from 0.3029 to 0.0982 pixel, and the absolute average vertical error reduces from 0.2398 to 0.0881 pixel. Meanwhile, the time consumed increases from 11.4 to 18.44 ms/frame. When subjected to the requirement of template size, the UCC algorithm does not have the capability of giving dual attention to both accuracy and high efficiency. The analysis of ZNSSD algorithm shows that the accuracy and efficiency of ZNSSD algorithm are lower than the other three methods, and the accuracy and efficiency of J-ZNSSD-P algorithm is the highest among the four algorithms. Moreover, the improved algorithm has a wide range of potential applications in the vibration extraction of high-speed digital cameras.

### 4. Experimental results and analysis

#### 4.1. Experiment on a grating ruler motion platform

In order to test the effectiveness of the improved J-ZNSSD-P algorithm in the actual displacement extraction, a visual measurement experiment based on the grating ruler platform was carried out. This is a manual operation platform equipped with a grating ruler sensor, as shown in Figure 7. The grating ruler is a commonly used high-precision displacement sensor that works according to the principle of physical formation of the moiré fringe.

Figure 6. Performance comparison of the results extracted.
Table 1. Errors and time consumption comparisons in the simulation test.

| Algorithm     | Max error (pixel) | Absolute average error (pixel) | $T_{\text{total}}$ (s) | $T_{\text{avg}}$ (ms) |
|---------------|-------------------|--------------------------------|------------------------|-----------------------|
| J-ZNSSD-P     | 0.3112            | 0.0892                         | 0.0713                 | 0.0606                | 1.39                  | 5.56                   |
| ZNSSD         | 0.6858            | 0.6932                         | 1.2480                 | 0.5537                | 4.86                  | 19.44                  |
| UCC($\text{usfac} = 1$) | 0.5360            | 0.3346                         | 0.3029                 | 0.2398                | 2.85                  | 11.4                   |
| UCC($\text{usfac} = 100$) | 0.3397            | 0.1120                         | 0.0982                 | 0.0881                | 4.61                  | 18.44                  |

It has the characteristics of large detection range, high detection accuracy, and fast response speed.

The grating ruler is composed of a grating reading head and a scale grating, which is fixed on the translation table. The grating reading head is connected with the rotary screw through the connecting plate. When the screw is rotated manually, the connecting plate and grating reading head synchronously move horizontally. In the experiment, we used high-speed camera vision system and grating ruler to measure the platform motion simultaneously, and the measurement results are compared to verify the measurement effect of the visual displacement extraction algorithm. As shown in Figure 7, the grating ruler grid distance used in the experiment is 0.02 mm, the resolution is 1 $\mu$m, the accuracy is $\pm 3$ $\mu$m, and the sampling frequency is 20 Hz.

Given that the surface of the connecting plate is smooth and there is no obvious tracking object, a circular target with a diameter of 20 mm is pasted as the shooting target of the visual system, as shown in the figure. The high-speed camera is 3 m away from the platform, and the video is recorded at the rate of 200 fps, and the captured image size is 160 $\times$ 160 pixels, one of which is shown in Figure 7(a). The position of the red box in the figure is selected as the tracking template, and the template size is 50 $\times$ 50 pixels. In using the target for pixel calibration, the diameter of the circle occupies about 104.7 pixels, and the pixel resolution is about 0.191 mm/pixel.

The results obtained after using the J-ZNSSD-P algorithm to extract the displacement of the target motion video are shown in Figure 7. As a comparison, the measurement result of the grating ruler motion is also drawn in the figure. It can be seen that the displacement extracted by the improved algorithm is highly consistent with the measurement results of the grating ruler and that the average error are 0.037 mm (figure b) and 0.027 mm (figure c).
Figure 8. Sound barrier experiment (a) Experimental environment; (b) Selected ROI; (c) Displacement and frequency spectrum of sound barrier.

4.2. Vibration analysis of the sound barrier

The feasibility of using the high-speed visual vibration frequency measurement algorithm is confirmed in the first experiment. This section presents the second experiment, that is, the sound barrier experiment. Sound barrier, which is also known as sound wall or noise barrier, is a platform structure designed to protect residents from noise pollution coming from both sides of the railway or highway. Thus, this structure is the most effective method for reducing the impacts of road, railway, and industrial noises. However, with the increase in train speed, gravity and impact force are introduced when high-speed trains pass through the noise barrier. This phenomenon leads to the violent vibration of the barrier, which in turn, can cause material fatigue and loosened assemblies. In addition, when the sound barrier is damaged and falls on the railway track, disastrous consequences might occur.

The traditional measurement algorithm is difficult to use, because installing obstacles in the working train line is inconvenient. To solve this problem, a vision-based measurement algorithm is thus proposed. This equipment mainly includes a notebook and a high-speed camera, which are used to capture the displacement vibration
video at a frame rate of 232 FPS. Figure 8. below displays the picture and displacement spectrum of the sound barrier of a railway section.

Figure 8(a)–(c) show the noise barrier experimental environment map, selected ROI, displacement diagram of the lateral movement of the sound barrier, and spectrum analysis. The frequencies of the first two orders of the sound barrier are 10.27, 20.92, and 45.41 Hz. The dynamic characteristics of the sound barrier are analyzed on the basis of the above results. The findings reveal that the high-speed visual measurement can also be used in practical engineering applications.

5. Conclusion

More information can be obtained by using a high-speed camera to collect videos of vibration object, but the speed of traditional image processing algorithm is lower than that of a camera. This mismatch leads to the low real-time performance of the vibration frequency measurement method based on high-speed vision. To improve the speed and accuracy of the traditional ZNSSD template matching algorithm, an improved ZNSSD template matching algorithm is proposed in this paper. The simulation results show that the accuracy and speed of the improved algorithm are significantly improved compared with the traditional algorithm. Furthermore, two experiments were carried out under laboratory and outdoor conditions to validate the accuracy and efficiency of the Improved ZNSSD algorithm performance in practice. The results demonstrated the high accuracy and efficiency of the proposed algorithm in extracting vibrating signals.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by Preliminary Research of National Natural Science Foundation of Anhui Polytechnic University (no. Xjky02201906), National Natural Science Foundation of Anhui Polytechnic University (no. Xjky02201906), and Wuhu University (no. Xjky02201906), National Natural Science Foundation of Anhui Polytechnic University, and Anhui Engineering University Jiujiang Industrial collaborative innovation special fund project (2021cyxtb2).

References

Aoyama, T., Li, L., Jiang, M., Inoue, K., Takaki, T., & Ishii, I. (2018). Vibration sensing of a bridge model using a multithread active vision system. IEEE/ASME Transactions on Mechatronics, 23(1), 179–189. https://doi.org/10.1109/TMECH.2017.2764504

Aqel, M., Marhaban, M. H., Saripan, M. I., & Ismail, N. B. (2016). Adaptive-search template matching technique based on vehicle acceleration for monocular visual odometry system. IEEE/ASME Transactions on Mechatronics, 21(6), 739–752. https://doi.org/10.11022/tee.22299

Bell, E., Peddle, J., & Goudreau, A. (2012). Bridge condition assessment using digital image correlation and structural modeling.

Bing, P., Wang, Z., & Lu, Z. (2010). Genuine full-field deformation measurement of an object with complex shape using reliability-guided digital image correlation. Optics Express, 18(2), 1011. https://doi.org/10.1364/OE.18.001011

Feng, D., & Feng, M. Q. (2017). Experimental validation of cost-effective vision-based structural health monitoring. Mechanical Systems and Signal Processing, 88, 199–211. https://doi.org/10.1016/j.jyspp.2016.11.021

Gu, T., Tu, Y., Tang, D., Lin, S., & Fang, B. (2020). A trimmed moving total least-squares method for curve and surface fitting. Measurement Science and Technology, 31(4), 045003. (8pp). https://doi.org/10.1088/1361-6501/ab4ff6

Guo, J., & Zhu, C. (2016). Dynamic displacement measurement of large-scale structures based on the Lucas-kanade template tracking algorithm. Mechanical Systems and Signal Processing, 66-67, 425–436. https://doi.org/10.1016/j.jyspp.2015.06.004

Lei, X. J. (2016). Research on key technologies and applications of high-speed vision system for dynamic measurement. University of Science and Technology of China.

Liu, L. S., Zhang, H. Y., Wang, T. F., Guo, J., & Chen, C. Q. (2015). Doppler spectrum imaging of vibrating target using laser heterodyne detection. Guangxue Jingmi Gongcheng/Optics and Precision Engineering, 23(6), 1508–1515.

Malik, S., & Soundararajan, R. (2019). Llrnet: A multiscale sub-band learning approach for low light image restoration. 2019 IEEE International Conference on Image Processing (ICIP), IEEE.

Muhammad, N., Bibi, N., Kamran, M., Bashir, Y., & Kim, D. G. (2020). Image noise reduction based on block matching in wavelet frame domain. Multimedia Tools and Applications, 79(11), 26327–26344. https://doi.org/10.1007/s11042-020-09158-0

Pangeric, D., Rusu, R. B., & Beetz, M. (2008). 3D-based monocular SLAM for mobile agents navigating in indoor environments. Proceedings of 13th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2008, September 15-18, 2008, Hamburg, Germany. IEEE.

Peng, C., Miu, W. D., & Ceng, C. (2020). 3D vibration measurement method for lightweight beam based on machine vision. Journal of Beijing University of Aeronautics and Astronautics, 47(02), 207–212. https://doi.org/10.13700/j.bj.1001-5965.2020.0211

Poozesh, P., Baqersad, J., Niezrecki, C., Avitabile, P., Harvey, E., & Yarala, R. (2017). Large-area photogrammetry based testing of wind turbine blades. Mechanical Systems and Signal Processing, 86, 98–115. https://doi.org/10.1016/j.jyspp.2016.07.021

Rao, G., Sreekala, R., Kumar, K. S., Gopalakrishnan, N., Muthumani, K., & Iyer, N. R. (2016). Seismic response measurement of an under-water model through high speed camera and feature tracking. Experimental Techniques, 40(1), 83–90. https://doi.org/10.1007/s40799-016-0013-0

Sahani, S. K., Adhikari, G., & Das, B. K. (2011). A fast template matching algorithm for aerial object tracking. 2011 International Conference on Image Information Processing. IEEE.
Wahbeh, A. M., Caffrey, J. P., & Masri, S. F. (2003). A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Materials and Structures, 12*(5), 785–794. [https://doi.org/10.1088/0964-1726/12/5/016](https://doi.org/10.1088/0964-1726/12/5/016)

Wang, Y. D. (2016). Research on real-time vision detection technology for motion feature of high-speed Mechanical system. *Journal of Mechanical Engineering, 52*(2), 82–90. [https://doi.org/10.3901/JME.2016.02.082](https://doi.org/10.3901/JME.2016.02.082)

Wu, J., Jiang, L., Li, C., Gu, Y., & Wang, A. (2018). Study on the remote multi-points vibration measurement method based on single infrared camera. *Chinese Journal of Scientific Instrument, 39*(9), 255–263. [https://doi.org/10.19650/j.cnki.cjsi.J1803685](https://doi.org/10.19650/j.cnki.cjsi.J1803685)

You, D. Y., Gao, X. D., & Katayama, S. (2014). Monitoring of high-power laser welding using high-speed photographing and image processing. *Mechanical Systems and Signal Processing, 49*(1-2), 39–52. [https://doi.org/10.1016/j.ymssp.2013.10.024](https://doi.org/10.1016/j.ymssp.2013.10.024)

Zhang, Y. M. (2019). *Mechanical vibration* (2nd ed.). Tsinghua University Press.