In this study, a high-speed camera system is developed to complete the vibration measurement in real time and to overcome the mass introduced by conventional contact measurements. The proposed system consists of a notebook computer and a high-speed camera which can capture the images as many as 1000 frames per second. In order to process the captured images in the computer, the normalized cross-correlation (NCC) template tracking algorithm with subpixel accuracy is introduced. Additionally, a modified local search algorithm based on the NCC is proposed to reduce the computation time and to increase efficiency significantly. The modified algorithm can rapidly accomplish one displacement extraction 10 times faster than the traditional template matching without installing any target panel onto the structures. Two experiments were carried out under laboratory and outdoor conditions to validate the accuracy and efficiency of the system performance in practice. The results demonstrated the high accuracy and efficiency of the camera system in extracting vibrating signals.

© 2015 Optical Society of America

**OCIS codes:** (100.2000) Digital image processing; (120.0120) Instrumentation, measurement, and metrology; (120.7280) Vibration analysis; (150.1135) Algorithms.

http://dx.doi.org/10.1364/AO.54.008198

---

**1. INTRODUCTION**

With the advent of the multimedia age, high-pixel and high-frame-rate cameras can record the location information of structures vibration and accomplish the reappearance of dynamic characteristics through the digital image processing technology [1,2]. Comparing with conventional measurement devices, such as accelerometers [3] and global positioning systems [4,5] widely used in industry, the vision measurement can not only obtain vibration acceleration information but also provide an intuitionistic exhibition of the actual vibration. Moreover, it can avoid the limitation that restricts the application range of traditional contact-type measurement. Noncontact sensors may not change the behavior of structures and can be used in special situations, such as remote measurement, high temperature, and magnetic fields, where the measurement results may be interfered with. Owing to these advantages, research interest is growing in the development of noncontact displacement measurement techniques, such as speckle photography [6], global positioning systems [4], and laser Doppler vibrometers [7]. However, the wide applications of these techniques are hampered due to the high total costs of the systems.

Optical cameras are effective alternatives to the noncontact measurement because the remote measurement and noninterference introduction are feasible. With these devices, all movement information can be shown visually, regardless of whether such movement can be seen or not. Nowadays, vision systems are widely used in engineering monitoring [8–10], underwater measurement [11], and sound passive recovery [12]. A vision technique has been widely developed for structural displacement measurement. Quan et al. [13] realized three-dimensional displacement using two-dimensional digital image correlation. Mudassar and Butt [14] applied improved digital image correlation for in-plane displacement measurement. Fukuda et al. [15] proposed a camera-based sensor system that uses a robust object search algorithm to measure the dynamic displacements of large-scale structures. However, commercial digital cameras have low sample rates, and the movement information cannot be obtain completely if the vibration frequency is more than half of the sample rates according to the sampling principle. Fortunately, high-speed vision systems that can capture images at 1000 frames per second (fps) or even faster have been developed to overcome these restrictions. Davis et al. [12] showed that the vibrations of many everyday objects in response to sound can be extracted to recover audio through high-speed videos. Kirugulige et al. [16] even developed speckle patterns in the crack tip vicinity at rates of 225,000 frames per second, but it is limited.
Conventional image-processing techniques are used for high-speed cameras, which have high requirements for motion-tracking algorithms, but these techniques are complex and cumbersome [16,17]. The standard way to track features between two images is through the template matching in computer vision. The basic template-matching algorithm calculates each position by a distortion function which measures the degree of similarity between the template and the image. The area-based algorithm consists of calculating the correlation function at each position of the search image in a raster scan fashion. Various area-based methods include mean absolute difference, product cross correlation (CC), sequential similarity detection algorithm, sum of absolute difference (SAD), and sum of squared differences (SSD) similarity measure [18]. After template matching is concerned, normalized cross correlation (NCC) is often adopted in similarity measure due to its better robustness [19–21]. Comparing with the other algorithms, NCC is less sensitive to linear changes in the amplitude of illumination in the two compared images. Furthermore, the NCC is confined in the range between 0 and 1; the setting of detection threshold value is much easier than the CC.

However, NCC cannot be directly computed using the more efficient fast Fourier transform in the spectral domain. The efficiency of the NCC algorithm is very low because most computationally expensive and resource-hungry operations require exhaustive search (ES). On the video compression field, various types of block-matching algorithms are proposed to increase the efficiency based on ES [22,23]. These algorithms include three-step search, new three-step search, simple and efficient search, four-step search, diamond search, and adaptive road pattern search. New NCC-based optimization algorithms have been developed to measure mass movements [24–27].

The sampling frequency of the structure vibration by high-speed camera is relatively high, and the displacement between each frame of the image is only a few pixels or even one pixel. Thus the matching position can be found only from a very small range on the basis of the previous frame in template matching. Because of these characteristics, the local search normalized cross correlation (LNCC) is proposed in this study on the displacement measurement to solve the problems of conventional image-processing techniques. The modified algorithm can obtain the same results of vibration extraction, but the computation time of one extraction is far less than that of the NCC algorithm, which provides the potential for online extraction on the high-speed captured video. This algorithm can be used for a point-by-point in displacement extraction and in-plane displacement testing. Finally, the proposed algorithm is applied in the high-speed camera system to realize the vibration measurement. Two experiments were carried out under laboratory and outdoor conditions for performance verification. The results demonstrated the accuracy and efficiency of the camera system in extracting the vibrating signals of structures.

This paper is organized as follows. Section 2 introduces the components and capability parameters of the high-speed camera system. Section 3 presents the NCC-based extraction algorithm and the subpixel accuracy improvement. It then describes the modified LNCC algorithm for template matching and provides a simulation example. Section 4 presents two experiments for performance verification. Section 5 concludes and outlooks.

2. HIGH-SPEED CAMERA SYSTEM

This developed high-speed camera system is shown in Fig. 1(a), which consists of a notebook computer (Intel Core processor 2.9 GHz; 2.75 GB RAM) connected to a camera head with telescopic lenses. The telescopic lenses with large optical zooming capability shown in Fig. 1(b) can be adjusted appropriately to perceive structural motion at different distances. The camera head uses a CCD sensor as the image receiver and can capture 8-bit gray-scale images at high speed. These images are then streamed into the notebook computer through a USB 3.0 interface. The image-processing software on the notebook computer can use the template-tracking algorithm to track a particular fast-moving object and extract motion information. The camera frame rate can reach 1000 fps when the image resolution is set as 300 pixels × 300 pixels.

To measure the displacement of remote structures, a target panel may be installed in advance or specific features are selected on the measuring location of the structure. Then the camera system is ready to capture images from a remote location, and the displacement time history of the structure can be obtained by applying the template-tracking algorithm to the digital video images. If the target panel is unavailable due to the limitation of the measurement environment, the distinct surface patterns of the structure, such as textures or edges, can also be used as matching templates.

3. METHODOLOGY OF THE VIBRATION EXTRACTION ALGORITHM

NCC is a well-known similarity criterion that is more robust than SAD and SSD under uniform illumination changes. A correlation matrix is calculated by shifting a given template pixel-by-pixel across a source image, which provides the information on the degree of matching between the template and the image. The maximum absolute value of the correlation matrix, whose location describes the best matching of the template, is then sought. The NCC coefficients are defined as follows,

![Fig. 1. High-speed camera system: (a) Compositions of the high-speed camera system, (b) the high-speed camera head.](image)
under the template.

The radius of the black circle is 80 pixels, and the gray-scale increases from the center 0 to 1. In Fig. 2(b), the black circle moves a few pixels. Based on these two images, the movement of the black circle can be recovered by applying the NCC method. The first image is used as the source image, and a small scope instead of the whole shifted image is selected as the template. The template in the red box (50 × 50) is cut out from Fig. 2(b), as shown in Fig. 2(c), and the coefficient matrix obtained by performing normxcorr2 is shown in Fig. 2(d), which is displayed as an image. In the coefficient matrix, the location corresponding to the maximum absolute value meets the best matching, which can be used to determine the movement.

The NCC method provides a simple and intuitive way to find the movement between two images. For a vibration video containing a series of images, the vibration signals can be easily extracted when the NCC operation is repeated between every two images. This method is called the NCC-based extraction algorithm. The algorithm procedure is listed as follows:

1. A fixed scope in which the video images are all focused on the moving object is selected.
2. The video images in this scope are truncated as template series and the first video frame is used as the source image.
3. The coefficient matrix between the source image and each template is calculated, and the location corresponding to the maximum absolute value is considered as the best match.

Finally, the x and y coordinates of these locations represent the movements in the horizontal and vertical directions. Notably, the movements obtained by applying this algorithm actually represent pixels. The real displacement can be obtained when the distance a pixel represents is known. Thus, the NCC-based extraction algorithm provides an effective approach to realize displacement measurement.

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

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

where the vibration frequency \( f = 1 \text{ Hz} \) and the sampling frequency \( f_s = 50 \text{ Hz} \). In this study, black circle simulation experiments are introduced in order to verify the speed and accuracy of the algorithm. The mean time cost of each iterations was introduced to evaluate the computation efficiency of the algorithms because the efficiency is only determined by each frame processing time cost. In other words, the high-speed system requested each frame processing time cost short enough to satisfy the high sample frequency. 1 Hz is randomly selected and high-frequency vibration simulation should be more accurate.

The implementation procedure of the NCC-based algorithm is illustrated in Fig. 3. The extracted results are shown in Fig. 4. The NCC-based 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 NCC-based algorithm match well with the real input.

The only deficiency of this method is that the extracted values are always integers because the result is decided by the location of the best-matching element in the coefficient matrix. This deficiency may affect the extracting accuracy and obtain a
Simulation results by applying the NCC-based algorithm.

Fig. 3. Procedure illustration of the NCC-based algorithm to the simulation example.

Step-type wave shape. The step-type effect is especially obvious when the vibration is small, such as when only several pixel movements occur. Tracking results are with an average tracking error of 0.5 pixel, which is relatively large considering the 10-pixel amplitude. To address this deficiency, filtering or fitting methods can be employed to wipe off the steps and smooth the wave shape. In this study, a subpixel localization algorithm is proposed to improve the extracting accuracy.

B. Subpixel Accuracy Improvement

In template matching, the NCC coefficients describe the correlation between the template and the source image. Given the limitation of matrix operations, the template is shifted pixel-by-pixel. Thus, the coefficient matrix calculated by NCC is only a sampling of the correlation plane, and the location corresponding to the maximum absolute value is a simple and rough estimation of the best matching. The best-matching point is located in the region around the maximum absolute value. Therefore, the region can be used for a more accurate subpixel localization and obtain decimal-level accuracy.

The subpixel localization algorithm is realized by cutting out a 3 × 3 region around the maximum absolute value to fit a two-dimensional quadratic surface, whose extreme point is considered the desired estimation. A region bigger than 3 × 3 may obtain more accurate estimation, but the complexity and quantity of computation also increase. Suppose the equation of the quadratic surface is expressed as Eq. (3) and the 3 × 3 matrix A is described as Eq. (4):

\[
f(x, y) = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2, \tag{3}
\]

\[
A = \begin{bmatrix} f(-1,1) & f(0,1) & f(1,1) \\
 f(-1,0) & f(0,0) & f(1,0) \\
 f(-1,-1) & f(0,-1) & f(1,-1) \end{bmatrix}
\]

\[
x = \begin{bmatrix} 0.30 & 0.62 & 0.30 \\
 0.75 & 0.97 & 0.51 \\
 0.65 & 0.78 & 0.20 \end{bmatrix}. \tag{4}
\]

The coefficients \( a_i (i = 0, \ldots, 5) \) can be estimated by substituting the matrix values into Eq. (3). The extreme point can then be easily calculated by solving the following equations,

\[
a_1 + a_3 \hat{y} + 2 a_4 \hat{x} = 0, \quad a_2 + a_3 \hat{x} + 2 a_5 \hat{y} = 0, \tag{5}
\]

where \( \hat{x} \) and \( \hat{y} \) are the coordinates of the extreme point, which is regarded as a better estimation of the best-matching location.

As a simple example, suppose that the 3 × 3 matrix A takes the values in Eq. (4), as shown in image form in Fig. 5(a). The above subpixel localization algorithm is then applied to fit the quadratic surface and calculate the extreme point. In Fig. 5(b), the fitting surface is displayed as an image form and the extreme point is marked by the triangle marker. The extreme point is a better estimation of the best-matching location. This improvement is used on the simulation in Subsection 3.A, and the results are shown in Fig. 4. The obtained wave shapes
are much smoother and match well with the actual inputs compared with the simulation conducted without using the improvement, and the average tracking error reduces to 0.1 pixel with the same vibration amplitude. At slight cost, the subpixel localization algorithm obtains a significant promotion.

Therefore, the NCC-based extraction algorithm is simple, effective, and requires less expert interference. The extracted results have clear physical meanings that represent the coordinate values in the horizontal and vertical directions. The subpixel localization algorithm improves the extracting accuracy further.

C. Modified Algorithm

Although the accuracy and robustness of the NCC-based algorithm perform well in image matching, the action principle is to calculate the correlation coefficients between the matching template and the whole image in all positions for the optimal matching point. Computational quantity will rapidly increase along with the increase of the target image size. The large computational time needed is one of main causes to restrict the application of the image matching process in high-speed camera systems. Figure 6 shows the special relationship that exists between each adjacent frame of the high-speed camera in which the characteristics of small-scale displacement determine that the useful search range is only within a small region around the previous frame-matching position. Therefore, a local search algorithm based on gradient descent search theory is proposed. This optimization algorithm only needs to calculate the correlation coefficient of several positions which are close to the previous frame image matching place between the searching image and the matching template. Such closeness greatly suppresses the number of search points, which reduces the computational cost to a large extent. The modified algorithm is also implemented by MATLAB language for this work.

1. Implementation of the Modified Algorithm

The distribution map of cross-correlation coefficients between the template and the search image is a contour map, which is a gradient descent graph with the matching position as the center in the template image matching process. In this contour map [Fig. 5(b)], the essence of the modified algorithm is to find the local optimal method. A judging matrix is built as shown in Eq. (6):

$$A_{i+1} = \begin{bmatrix} 
\gamma(x_i-1,y_i+1) & \gamma(x_i,y_i+1) & \gamma(x_i+1,y_i+1) \\
\gamma(x_i-1,y_i) & \gamma(x_i,y_i) & \gamma(x_i+1,y_i) \\
\gamma(x_i-1,y_i-1) & \gamma(x_i,y_i-1) & \gamma(x_i+1,y_i-1) 
\end{bmatrix}$$

(6)

Search Image

Fig. 6. Matching schematic diagram between search image and template image.

$${A}_{i+1}$$ is as the same as matrix $$A$$ in Subsection 3.B; it is the first step matched detection matrix. $$(x_i,y_i)$$ is the location of the previous frame template matching. $$\gamma(x,y)$$ is the correlation coefficient between the matching template and the searching image in position $$(x,y)$$. It needs to be noticed that the matching template is selected from the first frame so that the matching position of the first frame is known. Each step of search direction is determined by the motion vector of gradient raise. The steps of the algorithm are listed as follows:

1. With the position of the previous frame template matching $$(x_i,y_i)$$ taken as the search window center, $$A$$ matrix is used to judge whether $$A_{i+1}(2,2) = \text{max}(A_{i+1})$$. If yes, find the matching position, end the search, and process to Step 3; otherwise, go to Step 2.

2. With the position of $$\text{max}(A_{i+1,k})$$ ($$k$$ indicates the number of iterations) in step 1 taken as the search window center, three or five search points are added and the $$A$$ matrix is applied to judge whether $$A_{i+1,k}(2,2) = \text{max}(A_{i+1,k})$$. If yes, find the matching position, end the search, and process to Step 3; otherwise, continue to Step 2.

3. Return the match position $$(x_{i+1}, y_{i+1})$$ and the matrix $$A_{i+1}$$, calculate the subpixel accuracy location by matrix $$A_{i+1}$$ according the method in Subsection 3.B. Thus, the precise displacement of the search image can be obtained from the matching position and the subpixel location.

Figure 7 shows a search procedure of the algorithm, where the points of the numbers 1, 2, 3, and 4 are the additional search points in each iterative process, and the direction where the color deepening is the increasing direction of the distribution map of the cross-correlation coefficients. The entire iterative process begins from the start position and finishes at the end position.

2. Performance Analysis of the Modified Algorithm

To verify stability and speed of the modified algorithm, black circle experiments that adopted the same test conditions in Subsection 3.A were improved and conducted. The experimental results are shown in Fig. 8 and Table 1.

A comparison of the extraction results shown in Fig. 8 indicates that the effect of the modified algorithm is the same as that of the NCC-based algorithm. The similarity is attributed
to the almost identical calculation methods, as well as the same judging criterion and the same correlation coefficient matrix. Thus, the same location positioning of the pixel and the same results in subpixel precision are obtained. Moreover, the subpixel localization is realized by cutting out a $3 \times 3$ matrix around the maximum absolute value to fit a quadratic surface and calculate the extreme point, and this matrix is existed in pixel localization. Only twice matrix divisions are calculated for subpixel localization. So the proposed modifications for subpixel localization influence the computational load weakly.

Displacement between adjacent frames is quite small when such faint vibrations are sampled by a high-speed digital camera. In the Table 1, each row of the LNCC column shows the number of frames ($N$) that needed $n$ (1, 2, or 3) iterations. $t_{\text{mean}}$ indicates the mean time cost for each iteration, and $t_{\text{total}}$ indicates the total amount of time cost for $N$ frames. It clearly shows that each frame extracts the position only by once or twice iteration. In other words, the majority of the correlation coefficient calculations of the needed positions are only 12 or 14, which greatly improves the computational efficiency. $t_{\text{mean}}$ of the NCC column indicates each frame processing time needed, and $N$ indicates the whole image stack is 251. Comparing the time spent on integer pixel extraction of the two algorithms, LNCC is 0.6359s and NCC is 6.1266s in total time cost, and the modified algorithm is 10 times higher than the conventional NCC-based algorithm for single position displacement extraction. In summary, the modified algorithm inherits the robustness and accuracy of the NCC-based algorithm and far exceeds the rapidity of the latter. The potential applications of the modified algorithm used in high-speed digital camera vibrations extraction are numerous.

3. Selection of Algorithm Parameters

Several parameters must be chosen in applying the modified algorithm, and the most important is how the template area is selected. The selection of the parameters influences the performance and computing time. The extent of the influence of the parameter $m$ is discussed in this section.

Consider the simulation in Subsection 3.C.2 as an example. By performing the modified algorithm in the case of the parameter of template size $m$ varying from $10 \times 10$ to $60 \times 60$, this square scope taking feature as center is simple and suitable. The algorithm performance along the varying parameter is obtained. The curve of the performance provides an idea on the selection of algorithm parameter. The results are shown in Fig. 9(a). The red curves are the average error of the $x$ direction, whereas the blue curves are the average error of the $y$ direction. Figure 9(b) displays the computing time of different $m$.

When $m$ is small, the average tracking error decreases gradually. Below a value of about $20 \times 20$, the average tracking error gets close to 0.2 pixel and tends toward stability, which indicates that a large matching template is needed to ensure good performance. Figure 9(b) shows the computing time exponential growth along with $m$. In conclusion, large $m$ is helpful for extracting the vibration signals accurately, but it also increases computing time. Thus, the selection of $m$ needs the balance between the computing time and the algorithm performance.

4. EXPERIMENTAL VERIFICATION

A. Moving Platform Experiment

The performance of the LNCC-based algorithm in vibration extraction, the pixel-level accuracy, and the target-track ability in dynamic response were evaluated. A laboratory test was carried out using a grating ruler, which is a high-accuracy displacement measurement tool that uses the moiré fringe technology of grating and photoelectric conversion. This tool is also called the incremental grating displacement sensor. In this experiment, the grating ruler was fixed in the moving platform so to accurately track the motion signals of the moving platform. The sampling frequency of the grating ruler was 20 Hz and the grating pitch was 0.02 mm with 1 um resolving power. One cross target was fixed on the platform which could be programmed to move with an arbitrary amplitude and frequency in one direction. Li [29] has proved that NCC algorithm is not

![Fig. 8. Performance comparison of the results extracted by NCC-based and LNCC-based algorithms.](image-url)

![Table 1. Time Cost Comparison Between NCC and LNCC](table-url)

<table>
<thead>
<tr>
<th>$n$ (iterate)</th>
<th>$N$ (frame)</th>
<th>$t_{\text{mean}}$</th>
<th>$t_{\text{total}}$</th>
<th>$t_{\text{mean}}$</th>
<th>$N$ (frame)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>0.0013</td>
<td>0.0325</td>
<td>0.0244</td>
<td>251</td>
</tr>
<tr>
<td>2</td>
<td>203</td>
<td>0.0025</td>
<td>0.5170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>0.0038</td>
<td>0.0865</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total time:</strong> 0.6359s</td>
<td></td>
<td></td>
<td></td>
<td><strong>Total time:</strong> 6.1266s</td>
<td></td>
</tr>
</tbody>
</table>

![Fig. 9. Influence of the template size: (a) average tracking error to $x$ and $y$ direction, (b) computing time of different template size.](image-url)
sensitive to noise and visibility, and the modified algorithm inherits the characteristics of NCC except for efficiency so it is insensitive to noise and robust to different lighting conditions too. Thus, as shown in Fig. 10, the object surface and cross target were illuminated by a white light source that uses optical fiber, and the video camera was placed at about 1 m away from the table and along the optical axis so the shake table can exist in the camera plane. Recorded images were digitized and stored in a computer for further processing.

Camera calibration of the system was necessary before the tests. The camera captured the target panel at 200 fps and detected the motion of the target on a connected computer, as shown in Fig. 10(b). Figure 11(a) revealed that the size of the captured image was $400 \times 300$. The actual size of the pre-designed target panel was calculated, and results showed that 20 mm corresponds to 135 pixels in the captured image. Therefore, 1 pixel corresponded to about 0.148 mm. A cross target captured by the camera was displayed in Fig. 11(b), in which the selected template for the scope was an $80 \times 80$ square. To evaluate the performance of the LNCC-based algorithm in vibration extraction with nontarget, an $80 \times 80$ structural feature image was captured by the camera [Fig. 11(c)]. This image was also selected as the matching template.

The guide screw was driven by a manual control input, and the measurement was taken for 10 s. The displacement time history measured by the high-speed camera system was compared with that measured by the grating ruler, as shown in Fig. 12. The cross target tracking results agreed with the grating ruler perfectly, with an average tracking error of 0.035 mm, which was relatively small considering the 3 mm amplitude. Only a slight difference in vibration extraction was observed between target and nontarget tracking. The time spent on the integer pixel extraction by the two algorithms and two templates was shown in Table 2. The mean time needed for each frames processing by NCC was 0.2011s and the LNCC was only 0.0037s when the vibration was extracted with the target. Compared with the simulation results in Table 1, the average time cost increases because of the increase in size of the matching template. Furthermore, the modified algorithm is 54 times faster than the traditional NCC-based algorithm. The efficiency of LNCC would be more obvious with the increase

### Table 2. Time Cost Comparison of Two Algorithms and Two Templates

<table>
<thead>
<tr>
<th></th>
<th>LNCC</th>
<th>NCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>With target</td>
<td>0.0037</td>
<td>0.2011</td>
</tr>
<tr>
<td>Nontarget</td>
<td>0.0029</td>
<td>0.1538</td>
</tr>
</tbody>
</table>
in the size of the searching image. The extraction time cost by the target was also similar to that of the nontarget.

**B. Sound Barrier Experiment**

An attempt to extract the vibration of the sound barrier is reported to validate the effectiveness of the proposed algorithms in the field. The sound barrier, also called sound-wall or noise barrier, is a platy structure designed to protect inhabitants on both sides of the railway from noise pollution. It is the most effective method of mitigating roadway, railway, and industrial noise sources. However, with the increase in train speed, strong suction and impact are introduced when high-speed trains pass by the sections installed with sound barriers, leading to violent vibrations of the barriers; such vibrations may cause material fatigue and loose assembly. In addition, when the sound barriers are damaged and fall on the railway track, disastrous consequences may occur. Thus, to improve the performance of sound barriers, the vibration course when the high-speed train passes should be first measured. However, traditional measurement methods are difficult to use because of the inconvenience of installing the barriers at the working train line. Therefore, vision-based measurement has been put forward to solve such a problem.

The viaduct installed with sound barriers is located in KunShan City of China, as shown in Fig. 13(a). The experiment was set up below the viaduct, as shown in Fig. 13(b). The distance between the camera and the barrier is about 30 m. The setup mainly includes a PC and a high-speed camera, which are used to capture the displacement vibration video. When a high-speed train passes by, a video containing the vibration of the sound barrier can be shot at 232 fps. One video image is displayed in Fig. 13(c). In practice, the assumption that the barrier deformation is negligible relative to the vibration is acceptable. Although the barrier was shot at a certain elevation, the vibration in the video can be considered as the projection of the actual movements on the imaging plane of the camera, which has no effect on the extraction of vibration properties.

The position of the $50 \times 50$ template is displayed as the red box in Fig. 13(c). The actual size of per pixel is 0.78 mm, which is calculated by using the known physical size of the barrier.

This high image resolution in the case of remote measurement is attributed mainly to the use of a telephoto lens. The vibration displacement and its Fourier spectrum extraction by the modified algorithm applied to the barrier video are shown in Fig. 14. The $y$ direction is the train direction, and $x$ is the vertical direction of the train, which has an airflow to impact sound barrier. The vibration time history is clearly displayed when the train passes, and the moment of train arrival is marked in red. Three obvious spectral peaks can be observed at 10.42, 21.07, and 45.77 Hz, which can be considered as the characteristic frequencies of the sound barrier. Based on these results, the dynamic characteristic of the sound barrier, which was previously difficult to detect, can be grasped now.

**5. CONCLUSIONS**

This paper describes a modified method based on the NCC technique to measure structural vibration using a high-speed digital camera. To meet real-time requirements, the NCC template tracking algorithm in the field of computer vision was introduced and a modified LNCC algorithm was proposed to reduce the required calculation time. Through the optimization, the computation of one iteration that is mainly comprised by a collection of images could be performed rapidly within 0.2 ms even in a common computer. The properties of high efficiency, high precision, and good robustness of the modified algorithm were contributed to its application in the high-speed camera measurement system.

Two experimental studies validated the performance of the proposed algorithm. The motion platform experiment demonstrated the accuracy of the vibration measurement by comparing the results of the high-speed camera system with those of a conventional sensor, namely a grating ruler. The nontarget tracking test can also obtain almost the same results, which proved that the measurement method has the strong practicability even without a target. The proposed method greatly increased the potential applications of the high-speed camera system. An experiment on a sound barrier on a high-speed rail
further displayed the reliability and accurately of extracting the vibrating signals of real-life structures.

The algorithm can only meet the real-time calculations under a frequency sampling below 500 Hz and is still unable to meet the real-time calculations under a super-high-frequency sampling (more than 500 Hz). Further works are improving the efficiency of the algorithm so that vibration extraction online can be realized.

**Funding.** China Postdoctoral Science Foundation (2014M551820); National Basic Research Program of China (2014CB049500); Natural Science Foundation of Anhui Province (Anhui Provincial Natural Science Foundation) (1508085QE83).

**REFERENCES**