Structural dynamic response analysis using deviations from idealized edge profiles in high-speed video

Dashan Zhang
Bo Tian
Ye Wei
Wenhui Hou
Jie Guo

SPIE.

Structural dynamic response analysis using deviations from idealized edge profiles in high-speed video

Dashan Zhang,[a,b,*] **Bo Tian,**[a,b] **Ye Wei,**[c] **Wenhui Hou,**[c] **and Jie Guo**[c]
[a]Anhui Agriculture University, School of Engineering, Hefei, China
[b]Intelligent Agricultural Machinery Laboratory of Anhui Province, Hefei, China
[c]University of Science and Technology of China, Department of Precision Machinery and Precision Instrumentation, Hefei, China

Abstract. In our study, a deviation extraction method is introduced to obtain subtle deviation signals from structural idealized edge profiles. The deviations are employed to reconstruct an analysis matrix that consists of global translations along selected edge profiles, and then a singular value decomposition-based approach is proposed to extract valuable variations from the calculated analysis matrix. To avoid noises from textured edge profiles, a colorization optimization approach is applied to remove variations because of the textures and turn real image stripes into ones that satisfy the constant edge profile assumption more closely in the deviation extraction process. Two practical experiments are conducted to demonstrate the effectiveness and potential applications of our proposed method. The dynamic properties of a lightweight beam and a sound barrier are analyzed successfully by using high-speed videos. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.OE.58.1.014106]

Keywords: noncontact measurement; high-speed vision sensor; structural dynamic response; deviation extraction; singular value decomposition.

1 Introduction

With the development of camera systems and image processing techniques in recent years, vision-based approaches have been implemented as one of the most popular noncontact measurement approaches in areas such as structural health monitoring[1–3] and nondestructive testing.[4–6] Different from traditional contact accelerometers and strain gauges, these burgeoning noncontact alternatives are far more convenient for installation in conditions wherein contact sensors have difficult access and intuitive exhibition of the measurement target are provided. Furthermore, without adding any extra mass on the surface of the measurement target, noncontact devices rarely affect the natural properties of the detected object, making these approaches more suitable for analyzing flexible or lightweight structures.

Compared with common noncontact methods, such as laser Doppler vibrometer, vision-based devices have flexible composition and provide relatively high spatial resolution.[7] The emergence of high-speed imaging sensors has increased the sampling frame rate of cameras considerably, providing more advantage in structural high frequency information observation than common commercial cameras.[8–11] Noises from frame interpolation are technically avoided from the source benefiting from a relatively high sampling rate, and subtle motion details can be recorded in a high-speed video file.

Valuable motion signals hidden in video frames naturally contain vital information on the measurement target. Therefore, motion estimation from a sequence of video images has become a popular problem in computer vision area. Regardless of the technology involved, the most common motion estimation approaches can be grouped as optical flow[12–14] and image template matching methods[15–18] (including digital image correlation). Studies on these algorithms have shown various potential applications in fields, such as modal parameter analysis,[19–21] deformation extraction,[22] and fault diagnosis.[23,24]

Intensity-based motion estimation methods, such as Horn–Schunck[25] and Lucas–Kanade optical flows,[12] calculate the relative spatial and temporal derivative fields by solving the aperture equation at each pixel between consecutive frames. As these techniques theoretically are sensitive to image noise and disturbances, phase information[26,27] is adopted instead of raw pixel intensity values of images to enhance the robustness of intensity-based algorithms when facing illumination variations and noise conditions. However, considering the basic prerequisite transformation from spatial to frequency domain by using complex Gabor or complex steerable filter, the speed of execution of phase-based approaches will be restrained inevitably especially when handling high-speed video frames. As for image template matching algorithms, cross-correlation function may be invalid when facing poor and repetitive textures. In addition, although several subpixel refinement algorithms exist, these template matching methods continue to face difficulty in balancing efficiency and precision especially for subtle subpixel motions in videos.

Subtle motions and deviations in video are usually hard to see with the human eye without any image processing procedure. The emergence of motion magnification techniques allows the small variations in a specific frequency band to be observed intuitively by means of magnifying small intensities or phase variations in the video. In Eulerian video magnification (EVM)[10,28–30] algorithms, subtle variations should be extracted in the timeline and then processed with band-pass filtering before motion magnification. Hence, EVM

*Address all correspondence to Dashan Zhang, E-mail: zhangds@mail.ustc.edu.cn

methods only work properly on whole frames in the video. Compared with EVM algorithms, a deviation magnification (DM)[31] approach reveals and visualizes subtle geometric deviations from their idealized edge profiles in a single image. The DM algorithm extracts deviation signals from intensity variations along the edge of the structure without participation of any image multiscale decomposition, such as Gaussian pyramid and complex steerable pyramid decomposition. Therefore, subtle vibration signals hidden in video frames can be effectively extracted from another perspective.

This study proposes an approach capable of extracting valuable information from intensity variations along structural edge profiles by using a singular value decomposition (SVD)-based method. A deviation extraction algorithm is employed to calculate the deviations from idealized image boundaries in video frames, wherein the deviation signals are used to reconstruct an analysis matrix. As the relative intensity variations along structural edge profiles contain their vibration characteristics, SVD is applied to extract useful vibration signals involved in the analysis matrix. In addition, textures along the edges of textured edge profiles may reduce the intensity continuity of the selected analysis area and introduce errors in the deviation extraction process. Therefore, an image colorization optimization[31,32] method is used to guarantee the constant edge profile assumption.

The rest of the paper is organized as follows. The methodology section discusses theoretical derivations of the deviation extraction algorithm and a simple experiment on a saw blade is conducted to demonstrate the specific deviation extraction process. The following subpart introduces image colorization optimization method, and a simulation test on a sand dune image validates the texture denoising result. The end of the methodology section presents an SVD-based variation extraction approach. Two practical experiments are conducted to demonstrate potential applications and the practical effect in structural vibration analysis using high-speed camera systems. The proposed method and high-speed camera systems thoroughly examine the dynamic responses of a clamped cantilever beam and a sound barrier.

2 Theory and Algorithms

2.1 Deviation Extraction from the Edge Profile

The deviation extraction process is derived as follows. Consider an apparent simple straight-line edge feature, which always exists in captured subimage $I(x, y, t)$ in timeline t and has a subtle deviation signal $f(x, t)$ at every location x along the straight line. The edge profile at location x can be defined as follows:

$$E_x = I(x, y, t). \tag{1}$$

Assuming the deviation signal $f(x, t) = 0$ for an arbitrary x value, the edge profile $E_x(y, t)$ will remain constant and is rewritten as follows:

$$E_x(y, t) = E(y + f(x, t), t). \tag{2}$$

As the deviation $f(x, t)$ is considered to be subtle in reality, $E(y, t)$ is estimated by using the mean value of all available edge profiles in $I(x, y, t)$. Therefore, subimage $I(x, y, t)$ is then expressed as the sum of the edge profiles and independent image noises $n(x, y, t)$:

$$I(x, y, t) = E(y + f(x, t), t) + n(x, y, t). \tag{3}$$

Considering that $f(x, t)$ is small, subimage $I(x, y, t)$ can be approximated with a first-order Taylor expansion when higher components are ignored:

$$I(x, y, t) \approx E(y, t) + f(x, t)E'(y, t) + n(x, y, t). \tag{4}$$

The average pixel value over x is as follows:

$$\frac{1}{N}\sum_x I(x, y, t) \approx E(y, t) + E'(y, t)\frac{1}{N}\sum_x f(x, t)$$
$$+ \frac{1}{N}\sum_x n(x, y, t), \tag{5}$$

where N represents the pixel number in direction x. Considering the noise remainder item $\frac{1}{N}\sum_x n(x, y, t)$ has less variance than the original one and average of $f(x, t)$ is also a small value, Eq. (5) is approximated using another Taylor expansion:

$$E(y, t) + E'(y, t)\frac{1}{N}\sum_x f(x, t) \approx E\left(y + \frac{1}{N}\sum_x f(x, t)\right). \tag{6}$$

Accordingly, the average edge profile approximates the common edge profile up to a constant shift. The item $\frac{1}{N}\sum_x f(x, t)$ can be regarded as a global translation of $f(x, t)$ and is removed with a follow-up bandpass filter. Therefore, using the translated equation properly approximates the original edge profiles $E(y, t)$, and least square estimation method calculates the deviation signal $f(x, t)$:

$$\arg\ \min \sum_y (I(x, y, t) - E(y, t) - f(x, t)E'(y, t))^2, \tag{7}$$

leading to the final approximate solution of

$$f(x, t) \approx \frac{\sum_y(I(x, y, t) - E(y, t)E'(y, t))^2}{\sum_y E'(y, t)^2}. \tag{8}$$

An experiment is performed to demonstrate the effectiveness of the deviation extraction algorithm. Figure 1 shows a carbon tool steel saw blade fixed on a table vice captured with a Canon 70D SLR Camera, and the shape of its jagged edge on the right side is difficult to identify with the naked eye in the original image. To lessen the calculation burden, areas within the yellow rectangle is selected as analysis area for further operations. Our ultimate target is to extract the subtle deviation along the jagged edge.

Figure 2 presents the schematic and extraction results of the experiment. Figure 2(a) shows that the RGB image is first transformed into grayscale, and the jagged edge within the red rectangle region of interest (RoI) is fitted ideally through line segment detector algorithm and marked with a blue line. The size of the RoI is 480×50 pixels and green arrows in Fig. 2(a) show the orthogonal relationship between the blue fitted line and sampling direction. The detected straight-line segments usually require advance rotation for the convenience of sampling edge profiles because edge profiles need to be sampled strictly perpendicular to identified line
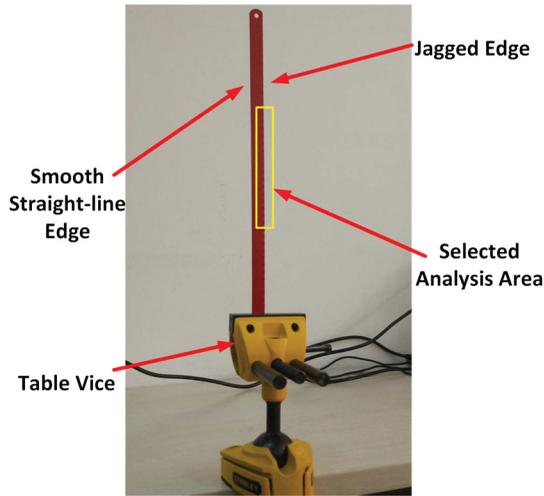
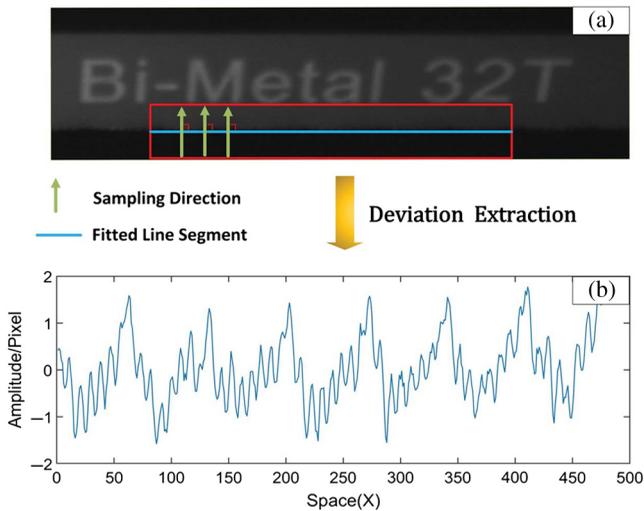**Fig. 1** Experimental setups of the saw blade experiment.

Fig. 2 (a) Schematic and (b) extraction results of the saw blade experiment.

segments and fitted line segments are not always idealized horizontally or vertically in an actual situation. In this particular case, the sampled rectangle is rotated clockwise at 3.2743 deg to become horizontal through bilinear interpolation method. Figure 2(b) shows the output deviation result without any denoising approach. An unnoticeable subtle jagged shape is extracted successfully from the saw blade. The extracted deviation signal has undergone a cyclical change with a maximum amplitude of no more than three pixels. Benefitting from the extreme short sampling interval of high-speed imaging, the global translation items between adjacent frames are negligibly small in the high-speed video, and the relative vibration signal perpendicular to the fitted line is presented as follows:

$$R_f = f(x, t) - f(x, t-1). \tag{9}$$

Sampling processes will be more complicated than approximate line segments for other simple curved boundaries, such as circular and ellipse edges. These kinds of edges

need to be idealized and fitted using methods, such as Hough transformation, to obtain their outlines and center coordinates. With the fitted curve boundaries and their center coordinate information, the RoIs are generated and unrolled to satisfy the orthogonal sampling requirement based on their polar equations:

$$\begin{cases} x = a\cos\theta, (a - m \le a \le a + n) \\ y = b\sin\theta, (b - m \le b \le b + n) \end{cases}, \quad (a, b \in \mathbf{N}^+), \tag{10}$$

where $[m, n]$ is the sampling length. As the sampling directions change continually with the change of parameter θ, the obtained relative deviations between frames can only represent subtle motions of their boundaries in different orientations. Therefore, relative deviations of curved elements should not be used as accurate vibration signals because motions perpendicular to the fitted curves are commonly not assumed to be uniform in practical measurement conditions. However, frequency characteristics of these data are still valuable. Accordingly, we discuss mainly vibration signal extraction cases from a structure's idealized straight-line contours.

2.2 Discussion of Textured Edge Profile

Deviation extraction algorithm displays stable performance when the constant assumption of tested edge profile is satisfied. In the experiment above, the idealized edge profile of the tested saw blade is clear and constant. However, the assumption of a constant edge profile may be invalid with textured contours. The textured edge profile can be determined when the edge of the idealized line-segment is not clear due to pixel variations. In practical measurement, the textured part near structural contour can invalidate the assumption of a constant edge profile and influence the deviation extraction result. Thus, an image colorization optimization method is applied to remove unwanted variations and avoid unpredictable measuring error due to texture.

The colorization optimization algorithm works in YUV color space, where Y refers to the monochromatic luminance channel related to pixel intensity, and U and V are the chrominance channels that encode the color. The transformation between RGB and YUV color spaces is as follows:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{11}$$

For a given Y channel, this algorithm assumes that two neighboring pixels should be similar in chrominance channels if they have similar intensities. As U and V are relatively independent, the values in channel U can be estimated by minimizing

$$\arg\ \min\left(\sum_r \left(U(r) - \sum_{s \in N(r)} w_{rs} U(s)\right)^2\right), \tag{12}$$

where $U(r)$ is the pixel of position r in channel U, $N(r)$ is the set of neighboring pixels around $r(s \in N(r))$, w_{rs} refers to the weighting function that evaluates the intensity similarity between $Y(r)$ and $Y(s)$. The least squares problem is
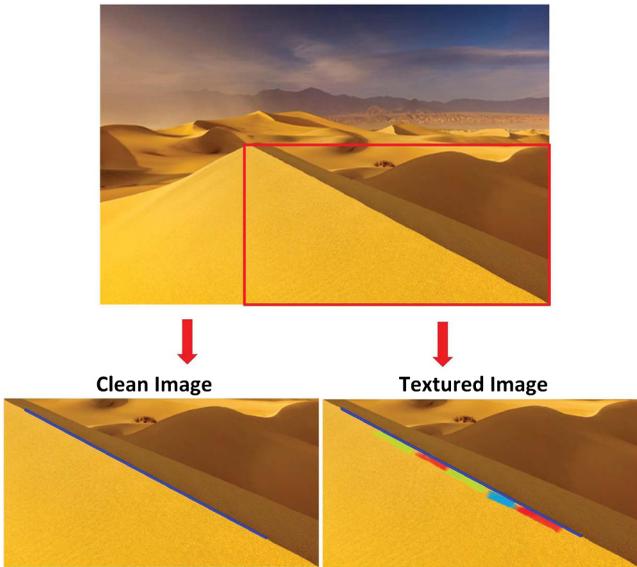
Fig. 3 Clean and textured sand dune straight-line boundaries.

minimized by solving a sparse linear system, and the constraint condition can be presented by marking the colors automatically in both sides close to the identified boundary.

In this section, a simulation test is shown to demonstrate the effect of the denoising process. Figure 3 shows that the sand dune boundaries within the red rectangular region are clear straight-line contours combined with subtle fluctuations. The employed idealized line segment is marked with blue lines in the clean and textured images. A series of color stripes are painted manually along the blue line's contour to create noise textures.

Figure 4 compares the sampling conditions of the clean and created textured images. The selected subimage is rotated counterclockwise at 28.1093 deg before sampling, and the size of sampling area is 480×18 pixels. The constraint condition of colorization optimization is obtained by marking the black and white colors two pixels away from the fitted blue line. Figure 4(b) shows the ellipse regions with textures along the edge that lead to obvious noises and will finally result in deviation extraction errors. Figure 4(c) shows the sampling area, where the denoising algorithm was
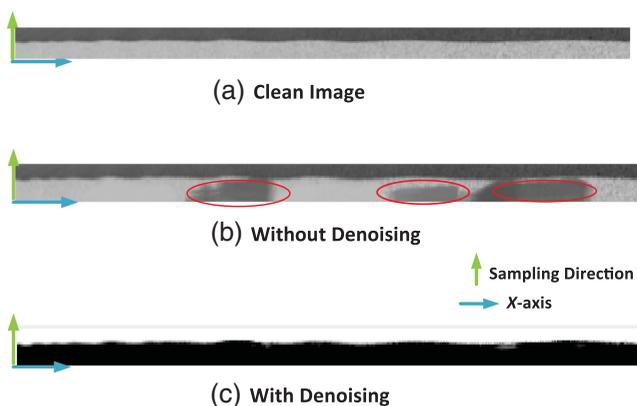

Fig. 4 Sampling results of the (a) clean image, (b) textured image, and (c) textured image with denoising.
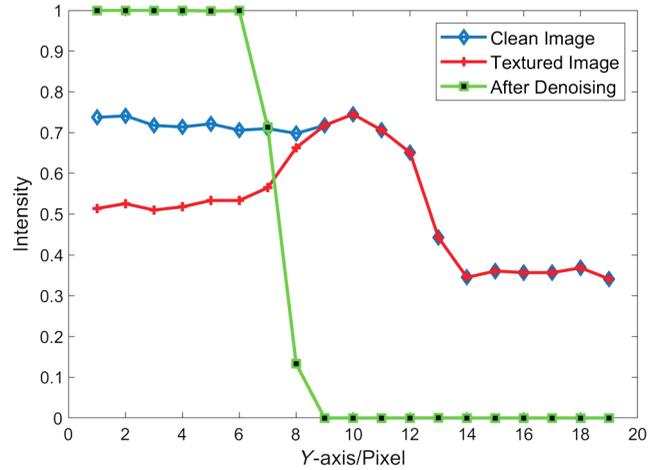

Fig. 5 Edge profile results ($x = 300$) in the simulation test.

conducted and most of the noise textures were removed successfully.

Figure 5 presents the comparisons of the edge profiles when $x = 300$. The curves of the clean and textured images indicate that textures along the edge will cause intensities on both sides of the fitted line blur into each other and lower their difference, causing the contour to become unclear and the extracted deviation signal is contaminated. Colorization optimization process has the opposite effect and can remove variations due to textures. Consequently, the constant edge profile assumption is satisfied.

Figure 6 shows the extracted deviation signals, and Table 1 presents their quantitative analysis results. Simulation data are processed using MATLAB R2018a on a machine with a single 4.00 GHz processor (Intel Core i7-6700k) and 16 GB RAM without overclocking. The deviation signal extracted from the textured image after denoising is consistent with the clean image and their correlation coefficient is 0.9986. However, elapsed time also increases because of the colorization optimization process. In practical measurement, avoiding textured areas along the contours is better.

2.3 SVD-Based Variation Extraction

Relative deviations of adjacent frames in the video can be calculated using the steps discussed above. Global translations are reconstructed into the following analysis matrix formed with subtle variations of the edge profile:

$$\mathbf{G}_{(t-1) \times N} = [f(x, 2) - f(x, 1), \ f(x, 3) - f(x, 2), \ \ldots f(x, t) - f(x, t - 1)]^*, \tag{13}$$

where t is the number of video frames, N is the pixel number in x direction, and $*$ is the matrix transposition. In practical measurement, each row of matrix $\mathbf{G}$ is composed of useful subtle pixel variations and noises. In this section, an SVD-based approach is employed to extract these subtle vibrations.

The SVD is a factorization of a matrix in linear algebra, which is broadly applied in signal processing and statistics. For a given real matrix $\mathbf{G} \in R_{(t-1) \times N}$, the SVD decomposes the matrix in the following form:
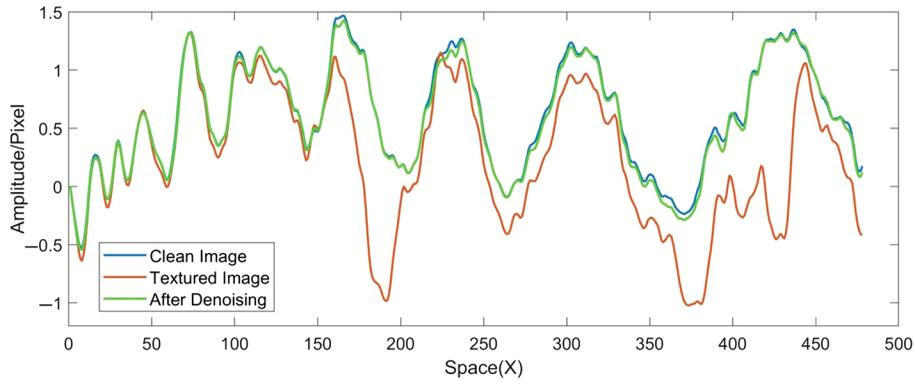
Fig. 6 Extracted deviation signals in the simulation test.

Table 1 Quantitative analysis of the extracted deviation signals in the simulation test.

	Corr-coefficient	Elapsed time (s)
Clean image	1.0000	0.0899
Textured image	0.6267	0.0960
Image after denoising	0.9986	0.5408

$$\mathbf{G} = \mathbf{U}\Sigma\mathbf{V}^* = \sigma_1\mathbf{u}_1\mathbf{v}_1^* + \sigma_2\mathbf{u}_2\mathbf{v}_2^* + \ldots + \sigma_q\mathbf{u}_q\mathbf{v}_q^*, \qquad (14)$$

where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{t-1}]$ is an $(t-1) \times (t-1)$ unitary matrix, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N]$ is an $n \times n$ unitary matrix, $\Sigma = [\text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_q)]$, $(q = \min(t-1, N))$ is an $(t-1) \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal.

SVD theory states that vectors in matrices $\mathbf{U}$ and $\mathbf{V}$ are orthonormal to one another in each vector group and form orthonormal bases of $(t-1)$-dimensional and N-dimensional spaces, respectively. Decomposition breaks a matrix into pieces based on the descending order of the singular values. Given the magnitudes of singular values indicate the energy distribution of the decomposed pieces, matrix $\mathbf{G}$ can be approximated by reserving former k singular values $\sigma_i (i = 1, 2, \ldots, k)(k < q)$. The low-rank approximation of matrix $\mathbf{G}$ takes the form of

$$\mathbf{G} \approx \sigma_1\mathbf{u}_1\mathbf{v}_1^* + \sigma_2\mathbf{u}_2\mathbf{v}_2^* + \ldots + \sigma_k\mathbf{u}_k\mathbf{v}_k^*. \qquad (15)$$

Given the corresponding singular values of noises are normally small, the summation of former k elements is re-created as final vibrations. In real life, the signal extraction process utilizes the data compression function of SVD. Figure 7 shows the flowchart of the overall variation extraction process.

3 Experimental Verification

The proposed method extracts subtle motions from the matrix reconstructed by the relative deviation signals from the image's idealized edge profile using an SVD-based approach. In this section, two experiments are conducted to demonstrate the effectiveness and potential applications of the proposed method.

3.1 Sound-Induced Subtle Vibration Analysis

Structural modal parameters, such as resonant frequencies and elasticity (Young's modulus), are important evaluating indicators in structural safety inspection. For lightweight structures, the added mass introduced by traditional contact sensors will inevitably affect the final modal test result. In the first experiment, the proposed approach was applied to estimate the material properties of the lightweight clamped saw blade from subtle motions in the video.

Figure 8 shows the experimental setup. The lightweight saw blade shown in Fig. 1 was clamped using a table vice and then excited by a linear ramp of frequency airwaves. The dimensions of the clamped saw blade are $0.29 \times 0.0126 \times 0.00065$ m, and Young's modulus and density are 2.06×10^{11} N·m^{-2} and 7.85×10^3 kg·m^{-3}, respectively. An audio file with a frequency band from 10 to 500 Hz was played by a loudspeaker from ~0.1 m away from the clamped saw blade. The volume of the sound was set to 80 dB. Figure 9 presents the waveform and short-time Fourier transform spectrogram of the input excitation sound. When air fluctuations hit the saw blade, these fluctuations may cause small forced vibrations on the surface of the vibrating object. Vibrations caused by the fluctuations
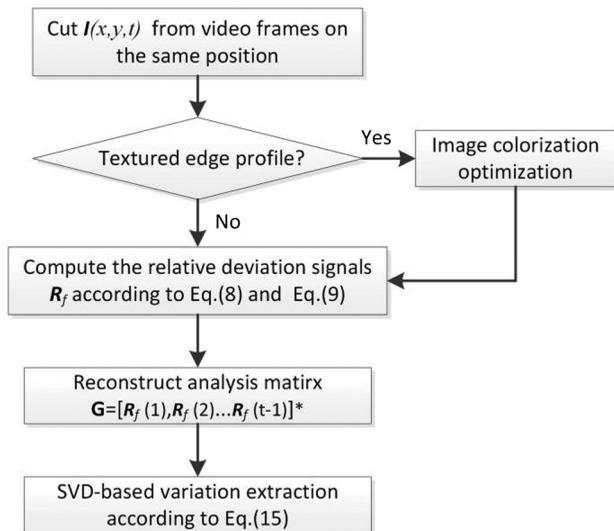

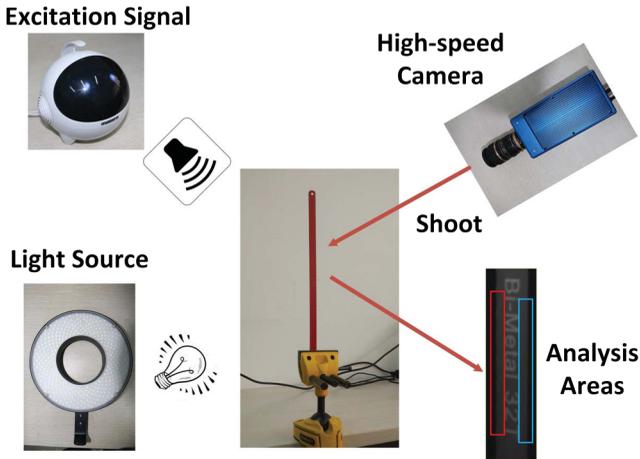**Fig. 7** Flowchart of the variation extraction process.

Fig. 8 Setups of the beam property estimation experiment.

were found to be in the order of one hundredth to one thousandth of a pixel using a phase-based optical flow. Small vibrations motivated by the excitation sound signal were recorded by the high-speed camera system (Mode-5KF10M, Agile Device, Inc.) at 500 fps with a resolution of 580×180 pixels. A USB 3.0 interface ensured data transmission between high-speed camera and computer, and a LED light source was applied to provide adequate illumination.

Two analysis areas were selected within the blue and red boxes in Fig. 8 to calculate the deviations of the contours. The image size for both areas was 350×30 pixels, and the smooth side of the saw blade was located on the right of the blue rectangle. Considering the acceptable clean boundary condition in this case, the analysis areas in the first experiment were not subjected to image colorization optimization. SVD decomposition result indicated signals corresponding to the first two singular values occupy over 97% of energy in the analysis matrix. Thus, the rest of the

components were considered noises and ignored in our final data. The average elapsed time per frame was ~0.04 s.

Figure 10 shows the waveforms and frequency spectrum results. Four peaks were observed at 6.37, 40.16, 113.10, and 221.60 Hz in their spectrograms. These peaks can be considered the first four resonant frequencies of the clamped saw blade in the modal test. Note the first peak 6.37 Hz is smaller than the lowest frequency (10 Hz) of the input excitation sound because of the presence of signal components below 10 Hz and the relatively high sensitivity of the lower modes.

The clamped saw blade could be considered as a cantilever beam, the theoretical resonant frequencies are estimated according to Euler–Bernoulli beam theory as follows:

$$f_n = \frac{3.52\alpha}{2\pi l^2} \sqrt{\frac{EI}{\rho A}}, \qquad (\alpha = 1, 6.27, 17.55, 34.39 \dots)$$

(16)

where f_n is the resonant frequency for the n'th mode, E refers to the Young's Modulus, I is the moment of inertia of the cantilever beam, ρ refers to the density, A is the cross-sectional area, and l is the length of the beam. Table 2 compares theoretical resonant frequencies and peaks found in the spectrograms. The resonant frequencies obtained from the proposed method are consistent with theoretical results. In real life, if the dimensions of a beam structure are already known, the experimental modal frequencies can help estimate properties, such as elasticity.

3.2 Vibration Analysis of Sound Barrier

The sound barrier is a type of platy structure installed on both sides of the railway to protect inhabitants from noise pollution. The sound barrier may suffer from strong suction caused by high-speed trains passing by in practical working conditions. The vibrations induced by the wind may lead to structural damage and decrease the working life of sound
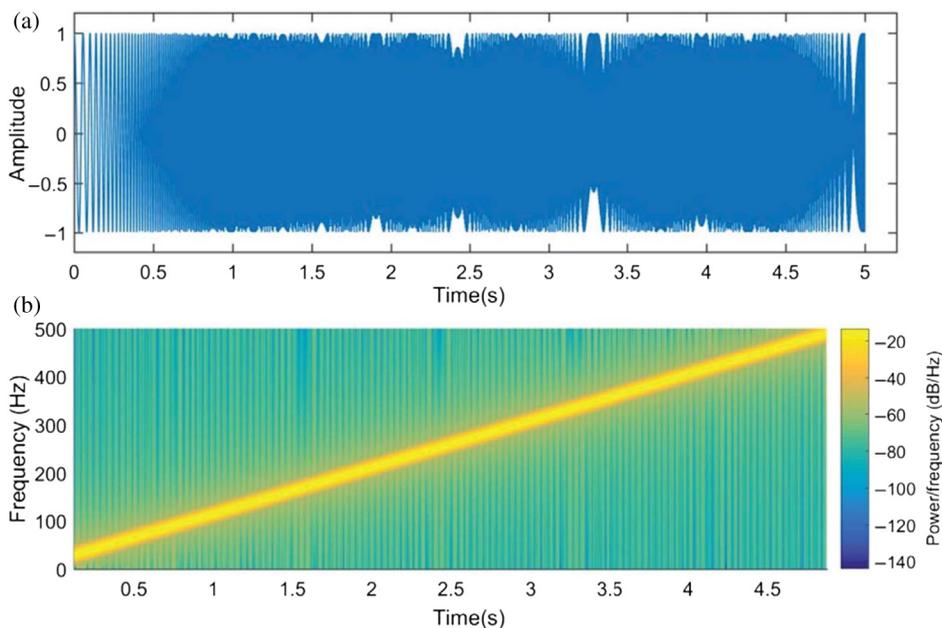

Fig. 9 (a) Waveform and (b) spectrogram of the input excitation signal.
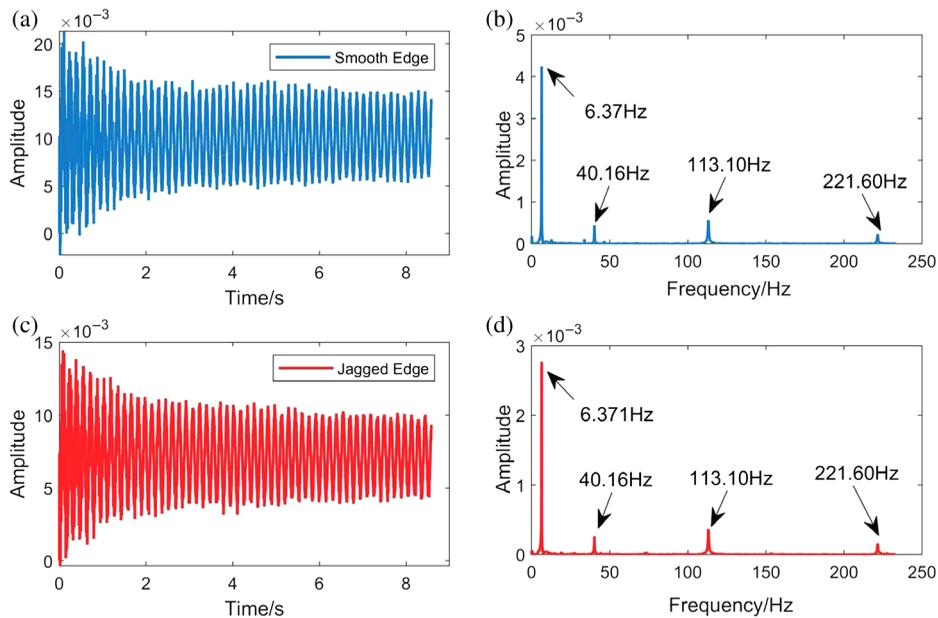
Fig. 10 (a)–(d) Vibration extraction results in the clamped beam experiment.

Table 2 Comparisons of theoretical and extracted resonant frequencies.

	Mode1	Mode2	Mode3	Mode4
Theoretical method (Hz)	6.41	40.17	112.43	220.31
Proposed method (Hz)	6.37	40.16	113.60	221.60

barriers along high-speed railways. These structures are usually installed on suspension viaducts and contact sensors are difficult to install. Vision-based noncontact method is a suitable alternative to analyze their dynamic responses when a train passes by.

Figure 11 shows the setup of the second experiment. Field test was conducted near KunShan railway station, which is an important railway hub in East China. The high-speed camera system (Mode-5KF04M, Agile Device, Inc.) was
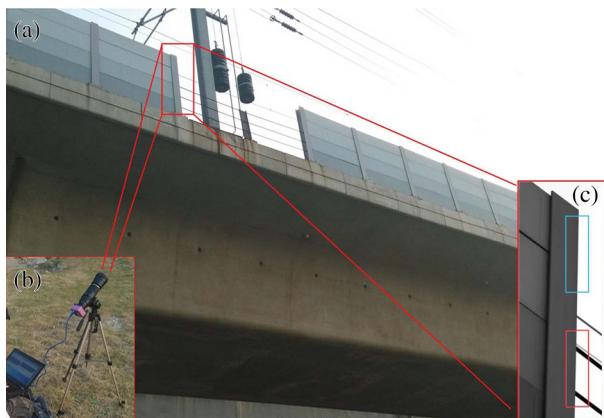

Fig. 11 Sound barrier experiment setups on (a) the experiment environment, (b) experimental devices and (c) video image and selected analysis areas.

placed below the viaduct in a distance. The height of the railway bridge is around 12 m, and the height of the sound barrier is 2.15 m. The distance between the camera head and the pier is ~28 m, and the distance between the camera head and sound barrier was estimated at 30 m. Measurement error rate caused by the camera tilt angle (about 26 deg) is ~0.6% to 0.8%,[33] which is considered acceptable. Hence, the errors have limited influence on frequency information analysis. The frame rate was set to 232 fps, and the resolution of the image was 660×1880 pixels during the experiment. Figure 11(c) presents selected areas for deviation extraction in which a clean straight-line boundary within the blue box and a textured boundary within the red box. The dimension of the subimages was 120×500 pixels. Colorization optimization algorithm before deviation extraction process was conducted on the textured subimages.

Figure 12 shows the final extracted waveforms and their frequency spectrum results. SVD decomposition results indicated that signals corresponding to the first three singular values occupy over 95% of energy in the analysis matrix. Thus, the other components were regarded as noises and ignored in our final data. Three obvious peaks—10.42, 21.07, and 45.77 Hz—were found in their Fourier spectrums. The textured edges are closer to the surface of bridge, and the moment of the train arrival is determined in the signal from the clean edge deviations and unclear in the signal from the textured edge deviations. Studies have indicated that the natural frequencies of the railway bridge are relatively low and rarely found from train-induced dynamic responses, wherein these extracted vibrations are considered dominated by excitation frequencies associated with the passing of the high-speed train.[34,35] The three peaks can be regarded as excited by the pulsed wind from the train because the excitation frequency of the wind load caused by the train carriages ranges from ~2 to 4 Hz, and the natural frequencies of the tested sound barrier (less than 5 Hz) are far from the excitation frequency of the train. These three peaks can be regarded excited by the pulsed wind from train locomotive.[36]
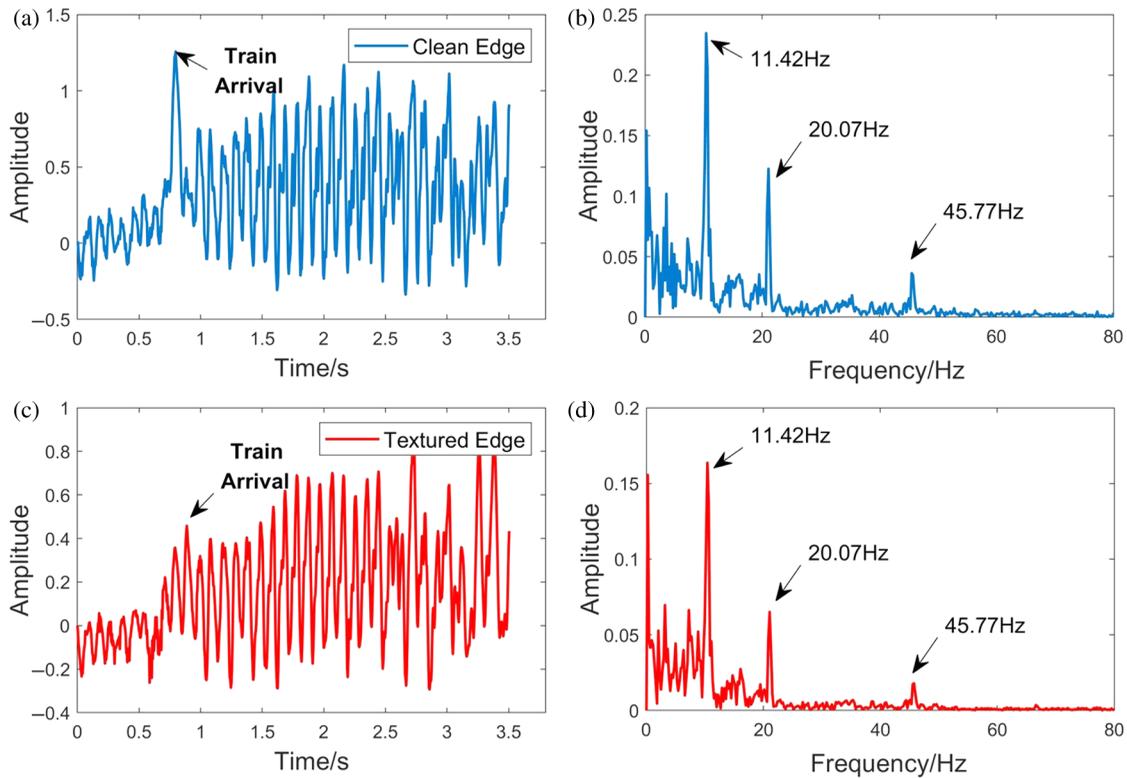
Fig. 12 (a)–(d) Vibration extraction results in the sound barrier experiment.

The three peaks can be considered the characteristic frequencies of the sound barrier.

4 Conclusions

In this study, a deviation extraction algorithm is proposed to measure variations along the edge profiles of an image. The derivation of the proposed algorithm is first introduced theoretically, and an experiment is conducted on the jagged edge of a saw blade to validate the effectiveness of the deviation extraction method. To avoid noises from textured edge profiles in an image, a colorization optimization approach is applied to remove variations due to textures and transform real image stripes into ones that satisfy more closely the constant edge profile assumption in the deviation extraction process. The simulation test on a sand dune image shows the negative effect of a textured boundary and compares the deviation extraction results before and after denoising. Quantitative analysis shows the extracted signals after colorization optimization maintain high correlation coefficient with the signals extracted from the clean boundary. The calculated global translations are reconstructed as an analysis matrix and an SVD-based method is proposed to extract useful subtle variations from the analysis matrix. Two experiments were conducted in the verification process to demonstrate further the potential applications of the proposed method. Vibration characteristics of a lightweight cantilever beam and a sound barrier were analyzed by using signals obtained by the proposed approach.

Considering the proposed method works on the pixel intensities, sudden illumination variations in practical measurement conditions may lead to unclear edge profiles and influence the deviation extraction process. The variations extracted using the proposed SVD-based method only reflect the tendency of the structure's vibration, wherein these signals are considered unsuitable as the true vibrations of the target. Future studies should focus on improving the tolerance of illumination and determine the proper relationship between actual vibration signals and extracted variations.

References

1. D. You, X. Gao, and S. Katayama, "Monitoring of high-power laser welding using high-speed photographing and image processing," *Mech. Syst. Signal Proc.* **49**(1–2), 39–52 (2014).
2. D. Feng and M. Q. Feng, "Experimental validation of cost-effective vision-based structural health monitoring," *Mech. Syst. Signal Proc.* **88**, 199–211 (2017).
3. D. Feng and M. Q. Feng, "Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection a review," *Eng. Struct.* **156**, 105–117 (2018).
4. J. W. Park et al., "Vision-based displacement measurement method for high-rise building structures using partitioning approach," *NDT & E Int.* **43**, 642–647 (2010).
5. G. Dobie et al., "Visual odometry and image mosaicing for NDE," *NDT & E Int.* **57**(8), 17–25 (2013).
6. Y. J. Cha, J. Chen, and O. Buyukozturk, "Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters," *Eng. Struct.* **132**, 300–313 (2017).
7. P. L. Reu, D. P. Rohe, and L. D. Jacobs, "Comparison of DIC and LDV for practical vibration and modal measurements," *Mech. Syst. Signal Proc.* **86**, 2–16 (2017).
8. Q. Zhang and X. Su, "High-speed optical measurement for the drumhead vibration," *Opt. Express* **13**(8), 3110–3116 (2005).

9. H. Kim et al., "Visual encoder: robust and precise measurement method of rotation angle via high-speed RGB vision," *Opt. Express* **24**(12), 13375 (2016).
10. A. Davis et al., "The visual microphone: passive recovery of sound from video," *ACM Trans. Graphics* **33**(4), 79 (2014).
11. D. Zhang et al., "Efficient subtle motion detection from high-speed video for sound recovery and vibration analysis using singular value decomposition-based approach," *Opt. Eng.* **56**(9), 094105 (2017).
12. S. Baker and I. Matthews, "Lucas-Kanade 20 years on: a unifying framework," *Int. J. Comput. Vision* **56**(3), 221–255 (2004).
13. J. Guo et al., "Vision-based measurement for rotational speed by improving Lucas-Kanade template tracking algorithm," *Appl. Opt.* **55**(25), 7186 (2016).
14. D. Diamond, P. Heyns, and A. Oberholster, "Accuracy evaluation of sub-pixel structural vibration measurements through optical flow analysis of a video sequence," *Measurement* **95**, 166–172 (2017).
15. W. Wang et al., "Frequency response functions of shape features from full-field vibration measurements using digital image correlation," *Mech. Syst. Signal Proc.* **28**(28), 333–347 (2012).
16. B. Pan and L. Tian, "Superfast robust digital image correlation analysis with parallel computing," *Opt. Eng.* **54**(3), 034106 (2015).
17. D. Zhang, M. Luo, and D. D. Arola, "Displacement/strain measurements using an optical microscope and digital image correlation," *Opt. Eng.* **45**(3), 535–545 (2006).
18. G. Zhu et al., "Sound recovery via intensity variations of speckle pattern pixels selected with variance-based method," *Opt. Eng.* **57**(2), 026117 (2018).
19. J. G. Chen et al., "Modal identification of simple structures with high-speed video using motion magnification," *J. Sound Vibr.* **345**, 58–71 (2015).
20. A. Davis et al., "Visual vibrometry: estimating material properties from small motions in video," in *Comput. Vision and Pattern Recognit.*, pp. 5335–5343 (2015).
21. A. Davis et al., "Visual vibrometry: estimating material properties from small motions in video," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(4), 732–745 (2017).
22. W. Feng et al., "Technique for two-dimensional displacement field determination using a reliability-guided spatial-gradient-based digital image correlation algorithm," *Appl. Opt.* **57**, 2780–2789 (2018).
23. S. Lu et al., "A novel contactless angular resampling method for motor bearing fault diagnosis under variable speed," *IEEE Trans. Instrum. Meas.* **65**(11), 2538–2550 (2016).
24. X. Wang et al., "A computer-vision-based rotating speed estimation method for motor bearing fault diagnosis," *Meas. Sci. Technol.* **28**(6), 065012 (2017).
25. B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.* **17**(1–3), 185–203 (1980).
26. D. J. Fleet and A. D. Jepson, "Computation of component image velocity from local phase information," *Int. J. Comput. Vision* **5**, 77–104 (1990).
27. E. P. Simoncelli et al., "Shiftable multiscale transforms," *IEEE Trans. Inf. Theory* **38**(2), 587–607 (1991).
28. H. Y. Wu et al., "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph.* **31**(4), 1–8 (2012).
29. N. Wadhwa, M. Rubinstein, and W. T. Freeman, "Phase-based video motion processing," *ACM Trans. Graph.* **32**(4), 1–10 (2013).
30. N. Wadhwa et al., "Eulerian video magnification and analysis," *Commun. ACM* **60**, 87–95 (2016).
31. N. Wadhwa et al., "Deviation magnification: revealing departures from ideal geometries," *ACM Trans. Graph.* **34**(6), 1–10 (2015).
32. A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," *ACM Trans. Graph.* **23**(3), 689–694 (2004).
33. D. Feng et al., "A vision-based sensor for noncontact structural displacement measurement," *Sensors* **15**(7), 16557–16575 (2015).
34. M. Q. Feng et al., "Nontarget vision sensor for remote measurement of bridge dynamic response," *J. Bridge Eng.* **20**(12), 04015023 (2015).
35. D. Feng et al., "Model updating of railway bridge using in situ dynamic displacement measurement under trainloads," *J. Bridge Eng.* **20**(12), 04015019 (2015).
36. L. Hermanns, J. G. Gimnez, and E. Alarcn, "Efficient computation of the pressures developed during high-speed train passing events," *Comput. Struct.* **83**(10), 793–803 (2005).

Dashan Zhang received his BS degree in mechanical engineering from Guizhou University in 2012. He received his PhD degree in mechanical engineering from the University of Science and Technology of China, Hefei, China, in 2017. Currently, he is a lecturer with the School of Engineering, Anhui Agriculture University. His research interests include image processing and optical measurement.

Bo Tian received his PhD degree in mechanical engineering from the University of Science and Technology of China, Hefei, China, in 2017. Currently, he is a lecturer with the department of automotive engineering, Anhui Agriculture University. His current research interests include image processing, pattern recognition, and fault diagnosis.

Ye Wei received his BS degree in mechanical engineering from University of Science and Technology of China, Hefei, China, in 2014. Currently, he is a PhD candidate in the Department of Precision Machinery and Precision Instrumentation at the University of Science and Technology of China. His current research interests include optical measurement and image super-resolution reconstruction.

Wenhui Hou received her BS degree in mechanical engineering from Yanshan University. Currently, he is a PhD candidate in the Department of Precision Machinery and Precision Instrumentation at the University of Science and Technology of China. Her current research interests include pattern recognition, machine learning, and nondestructive testing.

Jie Guo received his BS and PhD degrees in mechanical engineering from the University of Science and Technology of China, Hefei, China, in 2010 and 2015, respectively. Currently, he is a postdoctoral fellow with the Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China. His current research interests include machine vision, pattern recognition, machine learning, and fault diagnosis.