

Research Article

Vision-Based Faint Vibration Extraction Using Singular Value Decomposition

Xiujun Lei, Jie Guo, and Chang'an Zhu

Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China (USTC), Hefei, Anhui 230027, China

Correspondence should be addressed to Jie Guo; guojiegj@mail.ustc.edu.cn

Received 23 March 2015; Revised 18 June 2015; Accepted 1 July 2015

Academic Editor: Chih-Cheng Hung

Copyright © 2015 Xiujun Lei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Vibration measurement is important for understanding the behavior of engineering structures. Unlike conventional contact-type measurements, vision-based methodologies have attracted a great deal of attention because of the advantages of remote measurement, nonintrusive characteristic, and no mass introduction. It is a new type of displacement sensor which is convenient and reliable. This study introduces the singular value decomposition (SVD) methods for video image processing and presents a vibration-extracted algorithm. The algorithms can successfully realize noncontact displacement measurements without undesirable influence to the structure behavior. SVD-based algorithm decomposes a matrix combined with the former frames to obtain a set of orthonormal image bases while the projections of all video frames on the basis describe the vibration information. By means of simulation, the parameters selection of SVD-based algorithm is discussed in detail. To validate the algorithm performance in practice, sinusoidal motion tests are performed. Results indicate that the proposed technique can provide fairly accurate displacement measurement. Moreover, a sound barrier experiment showing how the high-speed rail trains affect the sound barrier nearby is carried out. It is for the first time to be realized at home and abroad due to the challenge of measuring environment.

1. Introduction

Vibration measurement is an important step for understanding the behavior of engineering structures. Traditional measurement devices, such as accelerometer [1] and global positioning systems [2, 3], have been widely used in the industry. However, these contact-type measurements have several limitations. For example, the introduction of a sensor may affect the behavior of structures. In many situations, such as remote measurement, high temperature, and magnetic fields, installation of a sensor is difficult and the measurement accuracy may be interfered.

Therefore, significant interest has been given to the development of noncontact measuring techniques, such as speckle photography [4], hologram interferometry [5], and laser Doppler vibrometer [6]. Meanwhile, the vision-based technique offers an effective alternative to noncontact displacement measurement. This technique benefits from the progress in image acquisition and computer vision with the advantages of remote measurement, nonintrusive characteristic, and no mass introduction. Hence, this technique has

been applied in several areas, including structures monitoring [7–11], human motion [12], and underwater measurement [13]. Lee and Shinozuka [10] used image processing techniques, such as texture recognition and projections of the captured image, to accomplish real-time displacement measurement of a flexible bridge. Jurjo et al. [11] presented an experimental methodology for the dynamic analysis of slender structures based on the digital image acquisition and processing techniques. Trigo et al. [13] implemented a combined computer vision and adaptive Kalman filter approach to identify the coordinates of control points in underwater environment. Reference [14] of Park et al. realized 3D displacement measurement model for health monitoring of structures using a motion capture system.

In essence, these vision-based methodologies are based on a series of image processing techniques for realizing video vibration extraction, including edge detection, image segmentation, and feature tracking. However, effective features are complex and difficult to find. In addition several parameters and thresholds must be adjusted in each step of the process to ensure a good performance. In this work, an

alternative methodology is proposed, which introduces the singular value decomposition (SVD) methods for vibration extraction. The SVD method is a common matrix-factorization algorithm in linear algebra that is equivalent to coordinate transform in high dimensional space and can therefore reveal the internal data structure and achieve dimensionality reduction. The SVD method has been widely used in many fields, such as principal component analysis [15, 16], noise reduction [17], faint signal extraction [18, 19], and machine condition monitoring [20].

This work explores the SVD method for video vibration extraction, and the vibration-extracted SVD-based algorithms are presented. Methodology of SVD-based extraction algorithm is investigated in detail in Section 2. A motion simulation of a black circle is introduced to show how the selection of parameters affects the measurement precision. The experiments, sinusoidal motion of a point and train excitation of sound barrier are development. In experiment part, the only needed preparation in advance is one adhesive marker having a negligible mass with respect to the vibrating structure. The additional marker especially can be omitted when the structure has distinct surface patterns, such as textures or edges. In addition, a modification for improving the matching accuracy is reported. The procedure for the SVD-based algorithm is simple, practical, and pervasive and requires little expert interference. The performance of the algorithm is validated by both simulation and experimental examples. The results show that it can accurately extract vibrating signals of large civil structures.

2. Methodology of SVD-Based Extraction Algorithm

SVD is a common matrix-factorization algorithm in linear algebra that can be defined as follows. For a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, two orthogonal matrices $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m] \in \mathbb{R}^{m \times m}$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ exist to meet the following form:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \quad (1)$$

where \mathbf{D} is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, $\mathbf{D} = [\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_q), \mathbf{O}]$, or its transposition decided by $m < n$ or $m > n$, while \mathbf{O} is a zero matrix $q = \min(m, n)$. The diagonal entries σ_i ($i = 1, 2, \dots, q$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_q > 0$) are known as the singular values of \mathbf{A} . The m columns of \mathbf{U} and the n columns of \mathbf{V} are called the left-singular vectors and right-singular vectors of \mathbf{A} , respectively.

To reveal the essence of SVD, (1) can be converted to another form of column vectors \mathbf{u}_i and \mathbf{v}_i :

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_q \mathbf{u}_q \mathbf{v}_q^T, \quad (2)$$

where $\mathbf{u}_i \in \mathbb{R}^{m \times 1}$, $\mathbf{v}_i \in \mathbb{R}^{n \times 1}$, and $i = 1, 2, \dots, q$. According to the SVD theory, the vectors \mathbf{u}_i ($i = 1, 2, \dots, q$), orthonormal to one another, form the orthonormal basis of the m -dimensional space; the vectors \mathbf{v}_i , orthonormal to one another, form the orthonormal basis of n -dimensional space. From (2), the SVD breaks a matrix into simple and meaningful pieces, whereas the magnitudes of σ_i ($i = 1, 2, \dots, q$)

represent the energy distribution of the pieces. The front singular values are large, which point to where the interesting information is, whereas the latter values are nearly close to zero, which can be regarded as noise. Thus, the former k ($k \leq q$) pieces can be truncated as the approximation of the original matrix as follows:

$$\mathbf{A} \approx \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T. \quad (3)$$

The SVD provides an approach for factoring the matrix into a series of linear approximations that expose the underlying structure and discover the redundancy of the matrix. As a result, dimensionality reduction and information extraction can be realized. In this work, the SVD technique is applied to analyze motion video and extract the vibration signals behind the video. Unlike traditional image processing techniques, the SVD-based extraction algorithm does not need any image manipulation, such as edge detection, image segmentation, and feature tracking, which require the adjustment of several parameters and threshold values. The extraction procedure mainly contains two steps: matrix decomposition and image projection. A set of orthonormal image bases (OIB) is obtained by decomposing a matrix created from the former video frames, which exposes the vibrating structure of the video. The vibration information is then revealed in some component signals after projecting all the frames on the OIB.

From the brief procedure, the SVD-based extraction algorithm is simple and straightforward. Based on the intrinsic characteristics of SVD, the algorithm is robust to noise and can always obtain smooth extracted results. The algorithm principle is explained in Section 2.1. More details about the use of the algorithm, such as the meaning of the component signals and the discussion of the algorithm parameters, are discussed in Sections 2.2 and 2.3.

2.1. Principle of the SVD-Based Algorithm. Given a video containing n frames, each frame is an $H \times W$ image, denoted by F_j , $j = 1, 2, \dots, n$. Before applying the SVD-based extraction algorithm, a scope is first selected in the video image. The selection of the scope must ensure that the video in the scope always focuses on the moving object or feature. Therefore, the information hidden in the scope is all about the vibration and SVD can reveal the reliable signals.

Suppose that the height of the scope is h and the width is w . The subimages truncated from the video frames are denoted by R_j , $j = 1, 2, \dots, n$. The subimages to column vectors are transformed and the former m vectors are combined into a global matrix, which can be described as $\mathbf{G}_{m \times (w \times h)} = [R_1(:, R_2(:, \dots, R_m(:)]^T$ using MATLAB code for explaining conveniently.

The matrix $\mathbf{G}_{m \times (w \times h)}$ is then decomposed by SVD, and matrices \mathbf{U} , \mathbf{D} , and \mathbf{V} are then obtained. In the diagonal matrix \mathbf{D} , the former k significant singular values are reserved, whereas the rest are set to zero. Thus, the decomposition can be expressed as follows:

$$\begin{aligned} \mathbf{G}_{m \times (w \times h)} &= \mathbf{U}_{m \times k} \cdot \mathbf{D}_{k \times k} \cdot \mathbf{V}_{(w \times h) \times k}^T \\ &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T, \end{aligned} \quad (4)$$

where $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$ and $k < m$; the definitions of \mathbf{U} , \mathbf{D} , and \mathbf{V} are similar to (1). In (4), the vectors \mathbf{v}_i ($i = 1, 2, \dots, k$) are orthonormal to one another and form the orthonormal basis of $(w \times h)$ -dimensional space. Given the length of the vector \mathbf{v}_i is $w \times h$, \mathbf{v}_i can be rewritten as matrix form, denoted by $I_{i,w \times h}$. Thus, the matrices $I_{i,w \times h}$ ($i = 1, 2, \dots, k$) form a set of OIB. According to the SVD theory, this set of OIB may reveal some characteristics of the video. Given that the subimages forming the matrix $\mathbf{G}_{m \times (w \times h)}$ all reflect the vibration of the same object, the OIB may expose the underlying vibration structure of the video, which is verified later by the motion simulation of black circle.

Corresponding to the i th component, the vector $\sigma_i \mathbf{u}_i$ is a coordinate of the projection of the global matrix $\mathbf{G}_{m \times (w \times h)}$ on the basis vector \mathbf{v}_i (or $I_{i,w \times h}$):

$$\begin{aligned} \mathbf{G} \cdot \mathbf{v}_i &= (\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T) \cdot \mathbf{v}_i \\ &= \sigma_i \mathbf{u}_i \mathbf{v}_i^T \cdot \mathbf{v}_i = \sigma_i \mathbf{u}_i \end{aligned} \quad (5)$$

or

$$\begin{aligned} \mathbf{G} \cdot \mathbf{v}_i &= [R_1(\cdot), R_2(\cdot), \dots, R_m(\cdot)]^T \cdot \mathbf{v}_i \\ &= [R_1 \otimes I_{i,w \times h}, R_2 \otimes I_{i,w \times h}, \dots, R_m \otimes I_{i,w \times h}]^T \\ &= [\sigma_i u_{i1}, \sigma_i u_{i2}, \dots, \sigma_i u_{im}]^T, \end{aligned} \quad (6)$$

where $\sigma_i \mathbf{u}_i = [\sigma_i u_{i1}, \sigma_i u_{i2}, \dots, \sigma_i u_{im}]^T$ and the operation \otimes is defined as follows: given two matrices of the same dimension, \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B} = \sum_i \sum_j a_{ij} b_{ij}$. Thus, $\sigma_i u_{ij}$ is the projection of R_j on $I_{i,w \times h}$ and $\sigma_i \mathbf{u}_i$ is the projection signal of the m subimages on $I_{i,w \times h}$, which reveals the vibration information of the former m frames on the i th principal direction.

Similar to (6), all the subimages R_j ($j = 1, 2, \dots, n$) can be projected on the set of OIB, and the i th projection signal can be considered as principal component $C_i \in \mathbb{R}^{n \times 1}$ ($i = 1, 2, \dots, k$):

$$C_{ji} = R_j \otimes I_{i,w \times h}, \quad (7)$$

where $C_i = [C_{1i}, C_{2i}, \dots, C_{ni}]^T$. These component signals reflect the vibration information of the whole video.

The procedure of the SVD-based extraction algorithm is listed as follows:

- A fixed scope with the size of $h \times w$ is selected from the video image, in which the subimage always focuses on the moving object or feature.
- The subimages in the scope is truncated from the video, and the former m subimages are combined to form a global matrix $\mathbf{G}_{m \times (w \times h)}$.
- The SVD is then applied to $\mathbf{G}_{m \times (w \times h)}$, and the key k singular values are retained with the corresponding orthonormal vectors \mathbf{v}_i ($i = 1, 2, \dots, k$) reshaped as $h \times w$ matrices that represent a set of OIB.
- All the subimages are projected on the set of OIB and the k component signals are obtained, which reflect the vibration information of the video.

For easy understanding, the procedure is illustrated in Figure 1 by applying the method to a simulation example. The SVD-based algorithm is then applied to extract the path coordinates. For conciseness, the motion equations, whose shape is an ellipse, are expressed as follows:

$$\begin{aligned} x(t) &= 10 \cos(2\pi ft) \\ y(t) &= 6 \sin(2\pi ft), \end{aligned} \quad (8)$$

where the vibration frequency $f = 1$ Hz and the sampling frequency $fs = 50$ Hz. The algorithm parameters are set as $m = 50$ and $h = w = 50$. The extraction results with their spectra are shown in Figure 2.

The first component signal has small variation around a high level, which may be assumed as the brightness information of the subimages. The second and third components have the same spectrum peak as the input frequency, indicating that these two signals may be related to the video vibration. Compared with the input motion in (8), the second component signal is similar to the x input, and the third is similar to the y input. Moreover, the other components vibrate faster with high harmonic spectra, and their amplitudes are small compared with the former components. Thus, the latter components can be regarded as disturbance terms introduced by the computational accuracy error, which can be explained by the fact that the former three components account for more than 98% of the total signal energy. In practice, only the former three singular values and their corresponding orthonormal vectors are calculated to save computing time, which can be conveniently realized by the function `svds`.

The results show that the SVD-based extraction algorithm is similar to dimension reduction. Although the pixels of the video image can be considered as features reflecting the video vibration, the extraction algorithm can remarkably reduce the feature dimension and obtain only the k component signals, which reveal the vibration information hidden behind the video. OIB has an important role in the process and deeply exposes the underlying vibration structure of the video. The brief processes and the results demonstrate that the SVD-based extraction algorithm is efficient and effective. Similar to the NCC-based algorithm, the SVD-based method is robust to different lighting conditions because it directly extracts the brightness information as the first component signal.

However, with regard to the actual physical significance of the components, an explicit conclusion based on the theoretical analysis above is difficult to achieve, while the NCC-based algorithm can be used to help explain the process from the numerical simulation perspective.

2.2. Significance Analysis of the Component Signals. From the decomposition in (4), the global matrix $\mathbf{G}_{m \times (w \times h)}$ is composed of k component matrices, which can be expressed as

$$\begin{aligned} \mathbf{G}_{m \times (w \times h)} &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T \\ &= \mathbf{G}_{1,m \times (w \times h)} + \mathbf{G}_{2,m \times (w \times h)} + \dots + \mathbf{G}_{k,m \times (w \times h)}, \end{aligned} \quad (9)$$

where $\mathbf{G}_{i,m \times (w \times h)} = \sigma_i \mathbf{u}_i \mathbf{v}_i^T$, $i = 1, 2, \dots, k$. Equation (9) shows that each component matrix is equivalent to a

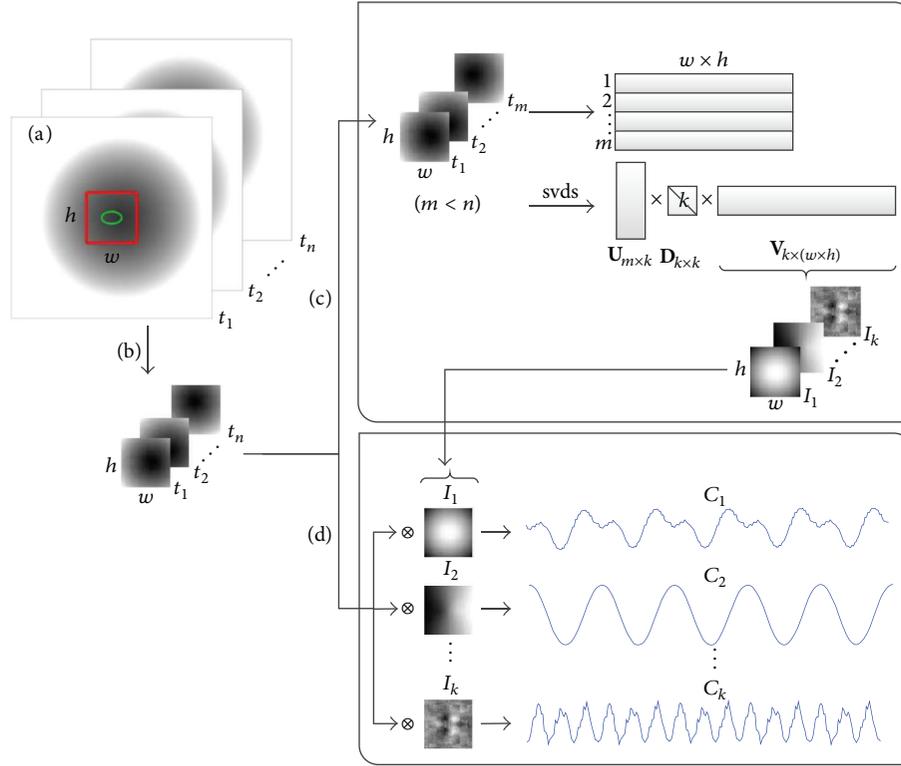


FIGURE 1: Procedure illustration of the SVD-based algorithm to the simulation example.

vibration subvideo containing m frame images whose sizes are $w \times h$. In other words, the vibration in the original video can be decomposed into k vibration components in the subvideos, whereas each subvideo corresponds with the corresponding singular value or the component signal. The physical significance of the component signals can be realized by reconstructing the subvideos and analyzing these components through the NCC-based algorithm. However, the reconstructed subvideos showed that obtaining the obvious vibration intuitively is difficult. The reason is that the first component signal, varying slightly on a high level, is a reflection of image average brightness, so the corresponding subvideo contains all brightness information, whereas the other subvideos only display extra difference about vibration free from brightness, thereby bringing difficulty to the significance analysis. Therefore, we reconstruct the subvideos by another way: $sV_{i,m \times (w \times h)} = \mathbf{G}_{1,m \times (w \times h)} + \mathbf{G}_{i,m \times (w \times h)}$, $i \geq 2$, and consider $sV_{i,m \times (w \times h)}$ as the i th subvideo. The reconstitution procedure is illustrated as in Figure 3.

After obtaining the subvideos, the SVD-based algorithm is performed by considering the first frame of the original video $F_{1,H \times W}$ as the source image and the subvideo frames as the template images. The results of the simulation example in Section 2.1 are shown in Figures 4(a) and 4(b).

The results showed that the SVD-extracted signals in the x -direction are constant, except for the second subvideo, whose x -direction signal is a standard cosine signal similar to that of the input x in (8). The signals in the y -direction

are constant, except for the third subvideo, whose y -direction signal is a standard sine signal similar to that of the input y in (8). Therefore, the second and third components signals are related to the simulation inputs x and y , respectively, and both reflect the vibration information of the original video, whereas the other components are not related to the vibration status. In practice, only the former three component signals are needed in performing the SVD-based extraction algorithm because the first component signal represents the brightness variation of video, while the second and third components reveal all the vibration information.

The trajectories of the second and third subvideos are plotted in Figure 4(c). These two trajectories are overall orthogonal, which can be interpreted by the decomposition irrelevance of SVD. Thus, the SVD-based extraction algorithm provides two orthonormal image bases: $I_{2,w \times h}$ and $I_{3,w \times h}$. When the original video is projected on these two bases, the vibration in original video is decomposed into two orthogonal vibration, which are simple and easy to analyze and contained in the subvideos.

To further exhibit the algorithm effect, another simulation example with a complicated vibrating path is provided with the inputs

$$\begin{aligned} x(t) &= 10 \cos(2\pi ft) \\ y(t) &= 6 \sin(2\pi ft) + 8 \sin(2\pi \cdot 2.7 ft). \end{aligned} \quad (10)$$

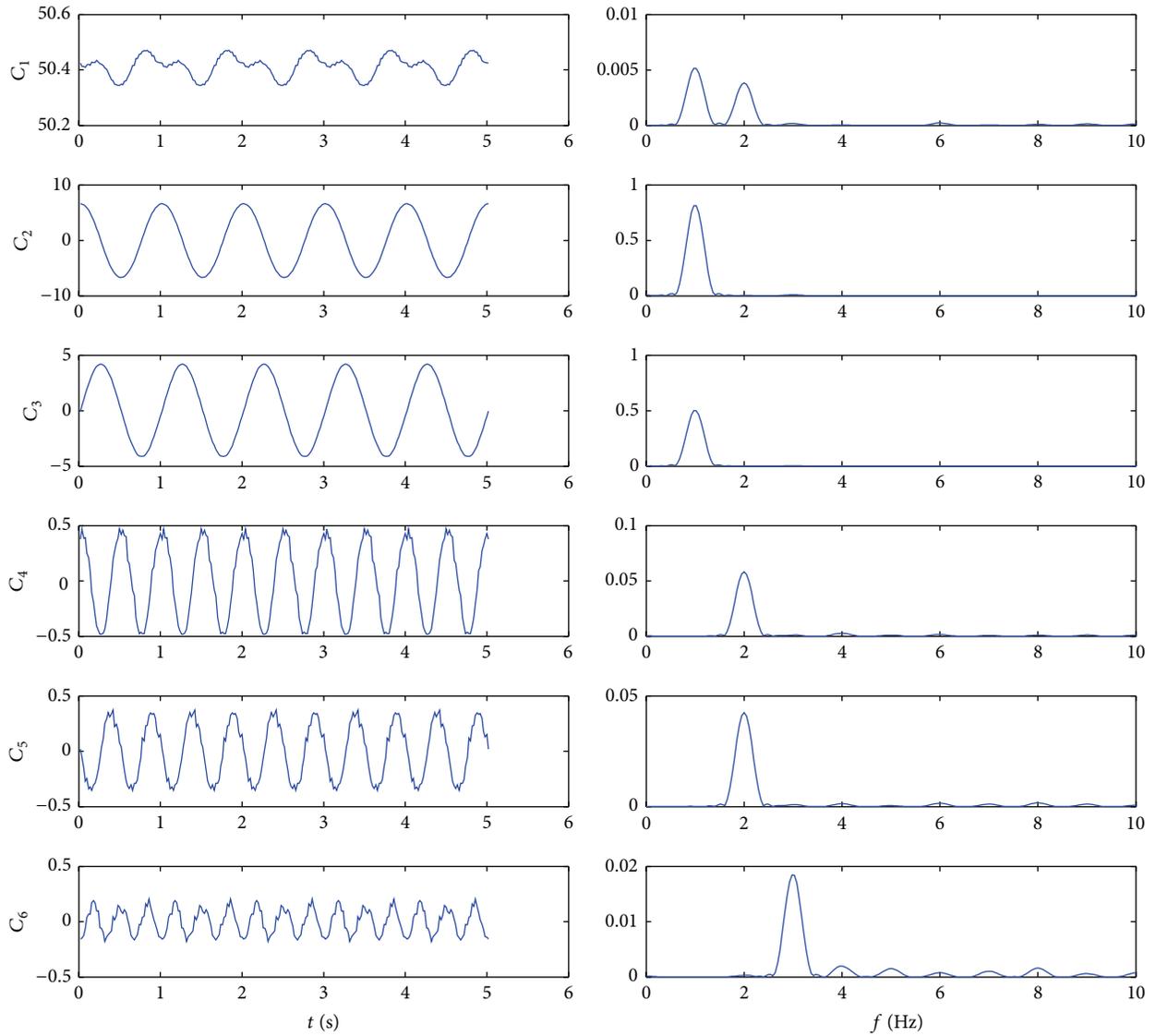


FIGURE 2: Simulation results by applying the SVD-based algorithm.

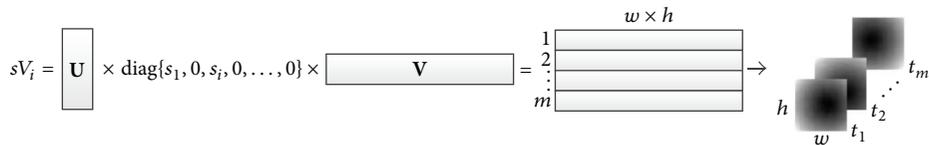


FIGURE 3: Reconstitution of the component subvideos from the decomposition of SVD.

The results obtained by applying the SVD-based algorithm for significance analysis are shown in Figure 5. Only the second and third components reflect the video vibration information, whereas the others are not related to the vibration. Although the complexity of the inputs increases, the orthogonality between the trajectories is still satisfied. The only difference is that the two trajectories are globally rotated with an angle compared with the simple example.

2.3. Selection of Algorithm Parameters. Several parameters must be chosen in applying the SVD-based algorithm, including the size and shape of the scope and the count of frames m used in the decomposition of SVD. Instead of using width w and height h to describe the scope, the scope area $S = (w \times h)$ and width-height ratio $\rho = (w/h)$ are used as the representation of size and shape. The selection of these parameters influences the performance and computing time.

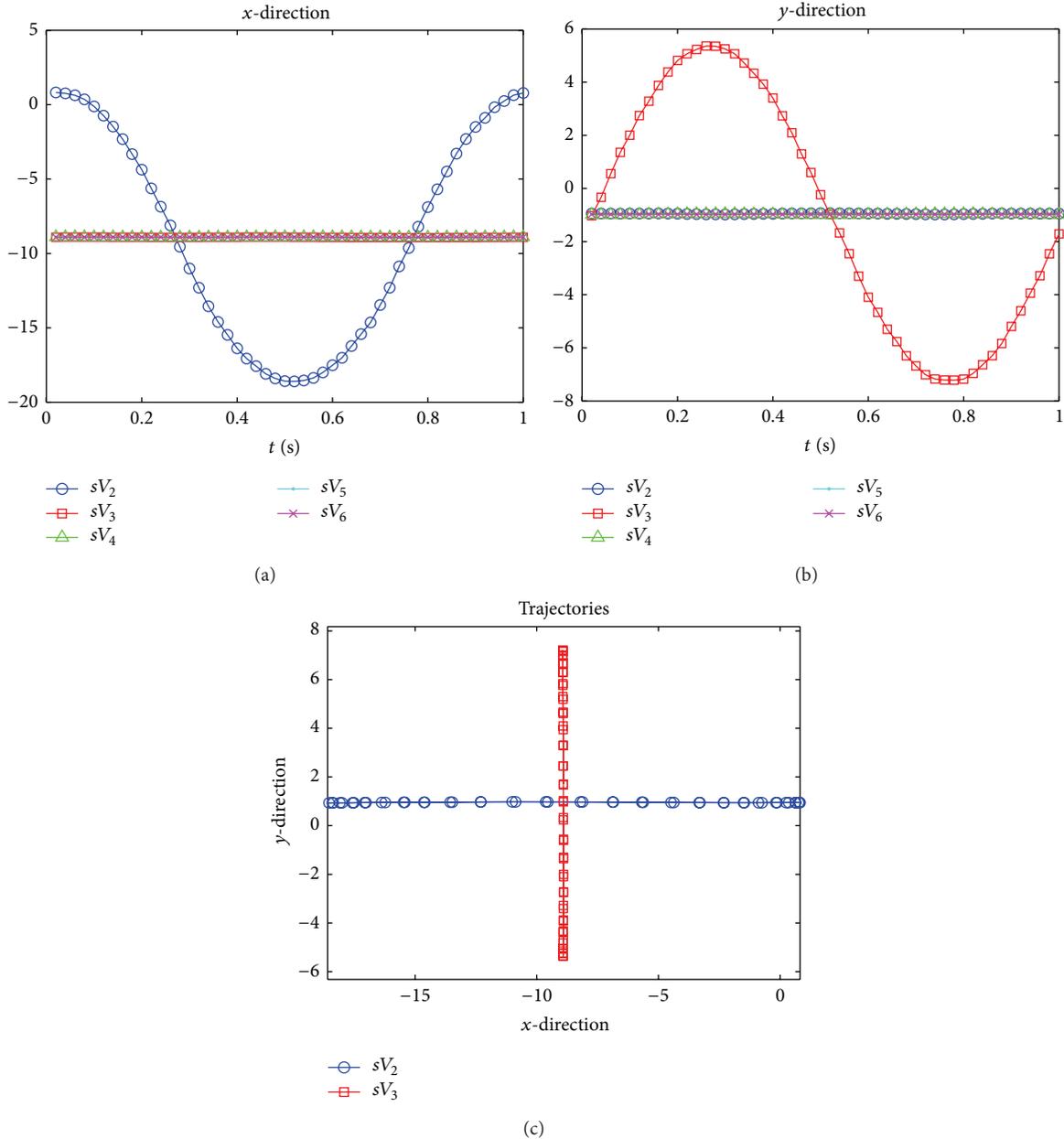


FIGURE 4: Significance analysis by applying the SVD-based algorithm: (a) x -direction vibration of the subvideos; (b) y -direction vibration of the subvideos; and (c) trajectories of the second and third subvideos.

The extent of the influence of these parameters is discussed in this section.

Considering the simulation in Section 2.1 as an example, and by performing the SVD-based algorithm in the case of one parameter varying and the others fixed, the algorithm performance along the varying parameter is obtained. The curve of the performance provides an idea on the selection of algorithm parameters. The variation ranges of the parameters are shown in Table 1.

Two criteria are proposed for evaluating the performance: the Correlation Coefficient (CC), used to determine the

TABLE 1: Variation ranges of the parameters.

	m	S	ρ
1	5–100	900 = (30 × 30)	1
2	50	3 × 3–100 × 100	1
3	50	900 = (30 × 30)	0.05–20

waveshape matching degree, and the Root-Mean-Square Ratio (RMSR) between the extracted signal and real input,

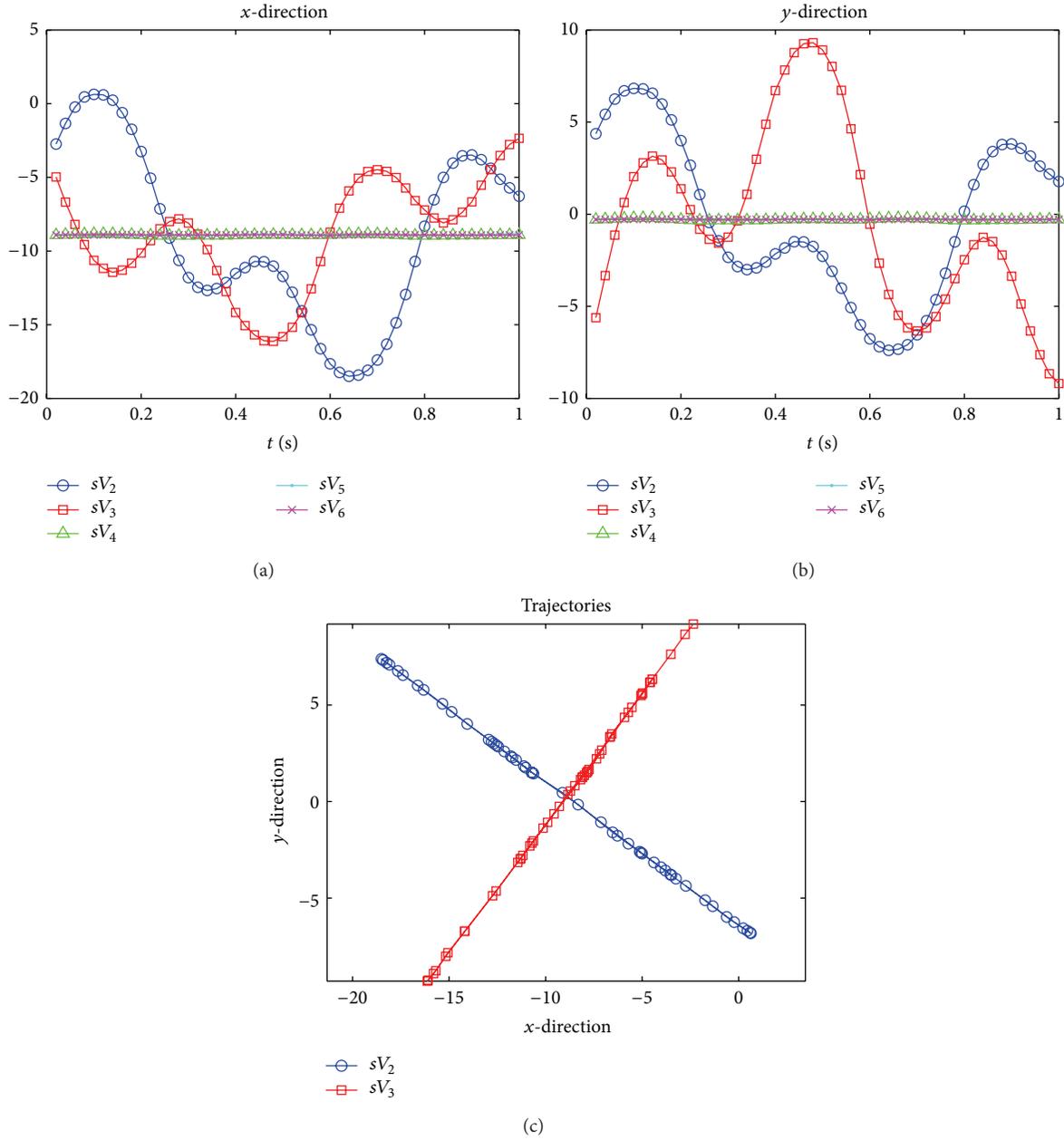


FIGURE 5: Significance analysis of another complicated simulation: (a) x -direction vibration of the subvideos; (b) y -direction vibration of the subvideos; and (c) trajectories of the second and third subvideos.

used to measure the recovering accuracy, which are defined as follows:

$$\begin{aligned} \text{CC}(s, r) &= \frac{(s - \bar{s})^T (r - \bar{r})}{\sqrt{(s - \bar{s})^T (s - \bar{s})} \sqrt{(r - \bar{r})^T (r - \bar{r})}} \\ \text{RMSR}(s, r) &= \frac{\sqrt{(s - \bar{s})^T (s - \bar{s})}}{\sqrt{(r - \bar{r})^T (r - \bar{r})}}, \end{aligned} \quad (11)$$

where $s \in \mathbb{R}^{n \times 1}$ is the extracted signal; $r \in \mathbb{R}^{n \times 1}$ is the real input; \bar{s} is the mean value of s ; and \bar{r} is the mean value of r .

The CC ranges from 0 to 1, where 0 means no correlation and 1 indicates perfect correlation.

2.3.1. Influence of the Parameter m . Let the parameter m range from 5 to 100, whereas the scope is set as $900 = (30 \times 30, \rho = 1)$, as shown in the first line of Table 1. Applying the SVD-based algorithm, the results are shown in Figure 6. The red curves marked by circle markers are the CC and RMSR criteria between the second component signal and the input x , whereas the blue curves marked by square markers are the criteria between the third component signal and the input y . Figure 6(c) displays the computing time of different m .

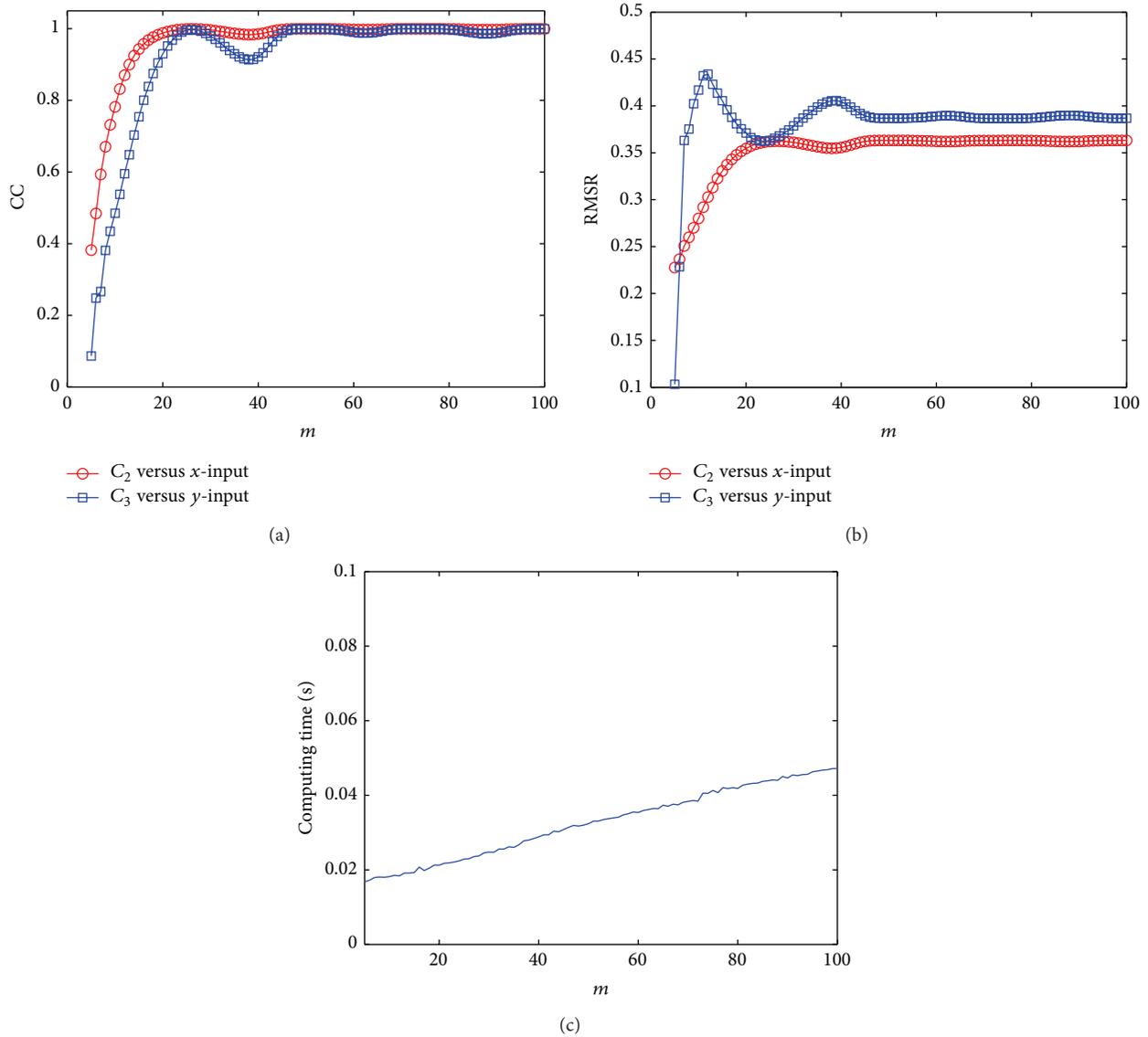


FIGURE 6: Influence of the parameter m : (a) CC criterion of the second and third components signals with the x - and y -inputs; (b) RMSR criterion of the second and third components signals with the x - and y -inputs; and (c) computing time of different m .

When m is small, the CC and RMSR increase gradually. After a value about 46, CC gets close to 1 and RMSR tends towards stability, indicating that large m is needed to ensure good performance. The tendency of RMSR, which is not to 1, indicates that the amplitude of extracted signal is unequal to the real input. Thus, the magnitude of the extracted signals does not have physical significance. Figure 6(c) shows that the computing time increases linearly along with m . Thus, the selection of m needs balance between the computing time and the algorithm performance.

2.3.2. Influence of the Parameter S . The area S changes under the condition of $m = 50$ and $\rho = 1$; that is, the scope is square. Suppose that the side length of square scope ranges from 3 to 100. The results are shown in Figure 7.

Figure 7 shows that the CC criterion remains unchanged around 1 and the RMSR increases linearly along with the side length, indicating that the scope area has little effect on the algorithm performance, except for enlarging the signal amplitude. However, the scope area has a significant impact on computing time, which increases quadratically.

2.3.3. Influence of the Parameter ρ . Assuming $S = 900$ and $m = 50$, the width-height ratio ρ changes from 0.05 to 20. Notably, when ρ is small, the corresponding relationship changes to the extracted results: the second component corresponds to the input y instead of x , whereas input x is reflected by the third component. Thus, adjustments are made by inverting the order of the two components at small ρ in Figures 8(a) and 8(b).

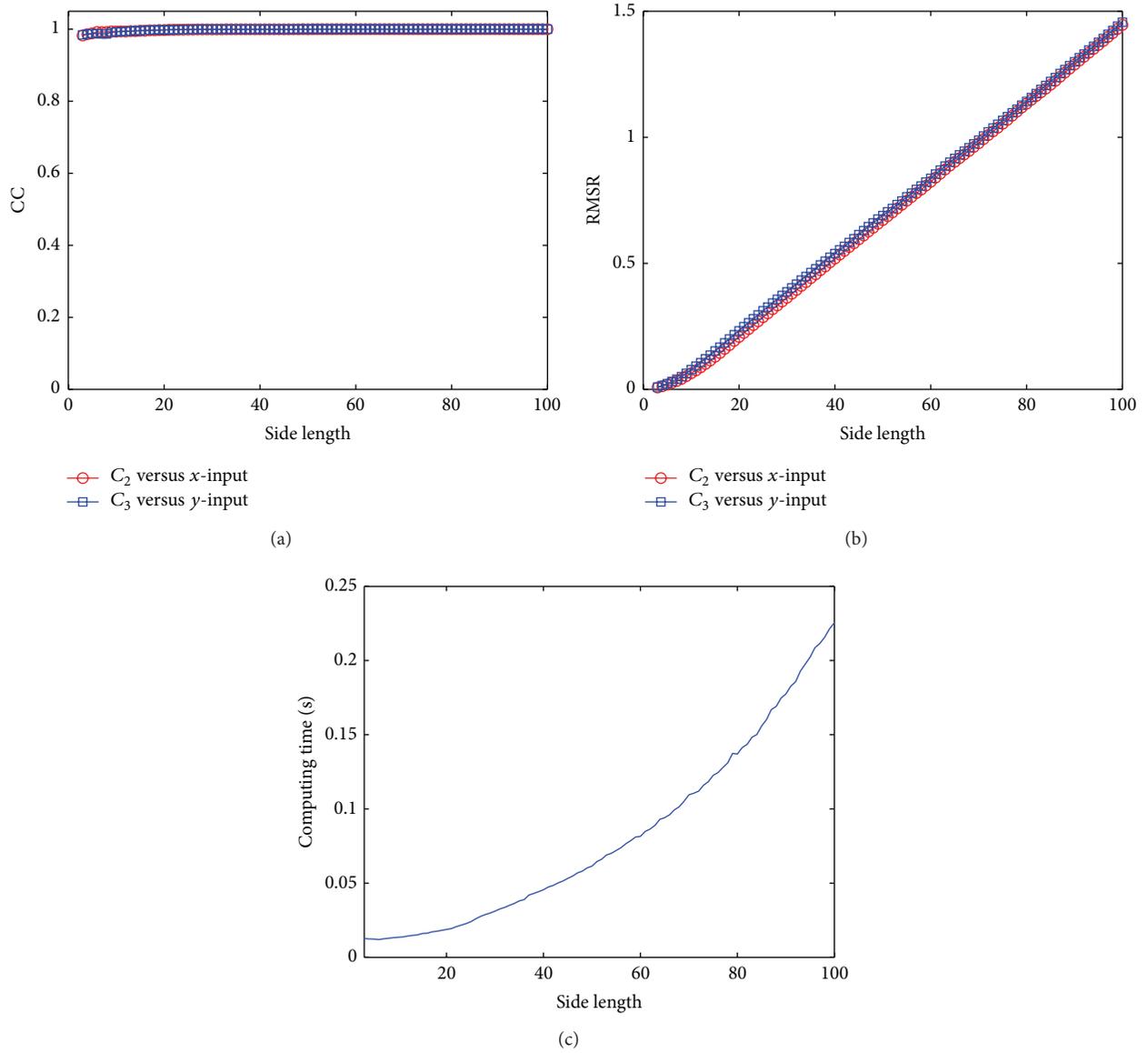


FIGURE 7: Influence of the parameter S : (a) CC criterion of the second and third components signals with the x - and y -inputs; (b) RMSR criterion of the second and third components signals with the x - and y -inputs; and (c) computing time of different S .

The results show that the width-height ratio ρ also has little impact on the CC criterion, except for the inversion of the corresponding relationship. The RMSR result shows that the amplitude of the component signals changes along with ρ . When ρ is small, the vibration signal in the horizontal direction is suppressed so that the extracted result has a small amplitude. Meanwhile, the signal in the vertical direction is promoted to obtain a large amplitude. Thus, the x -direction signal switches to the third component and the y -direction signal switches to the second. When ρ is large, the result is opposite. Therefore, the amplitude of the results is decided by both the area and the width-height ratio; that is, the SVD-based algorithm has directivity. When the scope has emphasis in one direction, the vibration on this direction is easier to extract and may have a larger amplitude. The circular scope

may be more suitable for use despite its complexity, which is not further studied in this work. In addition, the computing time does not significantly change because the size of scope is constant, as shown in Figure 8(c).

In conclusion, (a) the matching degree is only related to m and is related to neither the scope, the size, nor the shape. Large m is helpful for extracting the vibration signals accurately but also increases computing time. (b) The amplitude of the extracted signals is not equal to the practical inputs, which depends on the size and shape of the scope. Large area and long direction increase the amplitude. (c) The computing time is related to m and the scope area, except for the scope shape. The impact of the scope area is greater than m , indicating that the calculated amount of the matrix decomposition in (4) is decided by m and the scope area $h \times w$.

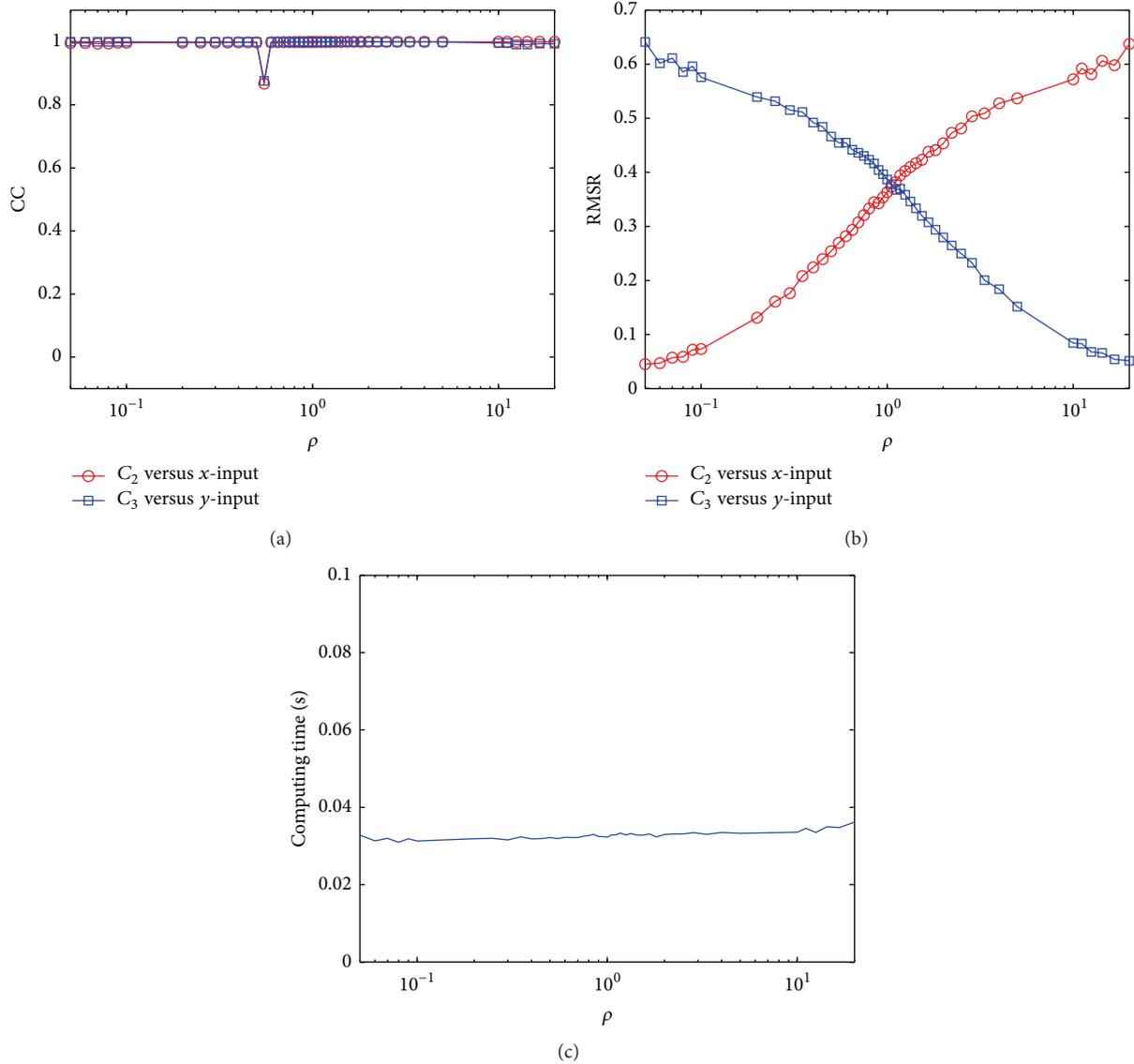


FIGURE 8: Influence of the parameter ρ : (a) CC criterion of the second and third components signals with the x - and y -inputs; (b) RMSR criterion of the second and third components signals with the x - and y -inputs; and (c) computing time of different ρ .

Finally, the selection of parameters can be summarized as follows:

- (a) A larger m is better, but the selection needs balance between the performance and the computing time.
- (b) The size of the scope should not be too large.
- (c) Square scope is simple and suitable.
- (d) To ensure good results, m can be set large at the price of decreasing the area.

3. Experimental Verification

3.1. Grating Ruler Shake Table Test. In order to evaluate the performance of the SVD-based vibration extraction

technique, the pixel level accuracy, and the target track ability of a dynamic response, a laboratory test was carried out using a grating ruler which called incremental grating displacement sensor moving table subjected to sinusoidal excitation. In this experiment, the grating ruler was fixed in the moving table so that it can accurately track the motion signals of moving table. The sampling frequency of the grating ruler is 20 Hz and the grating pitch is 0.02 mm with 1 μ m resolving power. One target was fixed on a table which could be programmed to move in one direction with an arbitrary amplitude and frequency. As shown in Figure 9(a), video camera was placed at about 3 m away from the table and stood at comfortable height so that the shake table exists in camera plane. It should be better that the image plane be parallel to the object plane but not necessary.

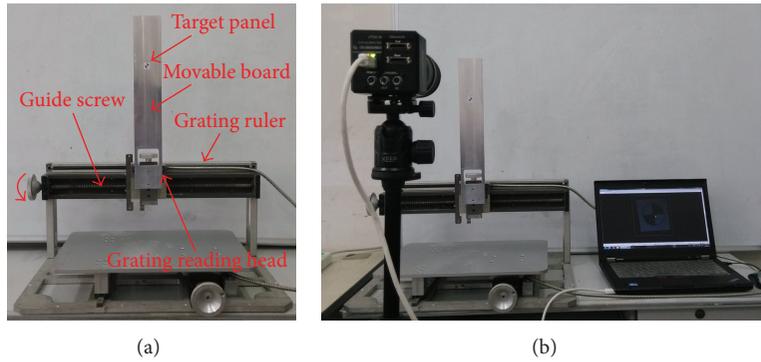


FIGURE 9: Grating ruler shake table test: (a) grating ruler table; (b) test configuration.

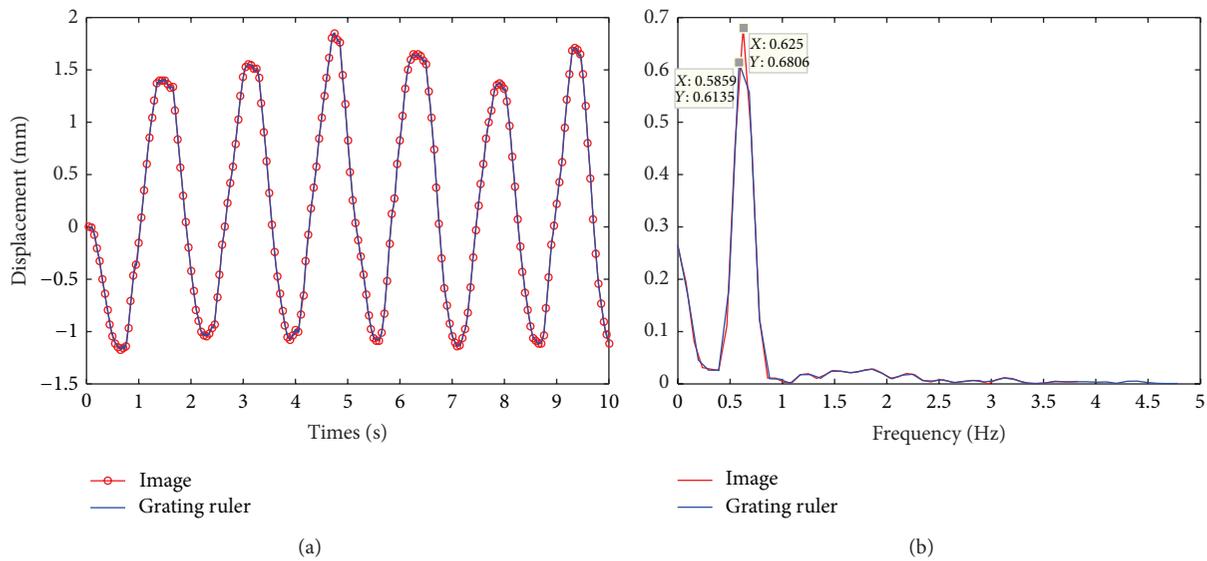


FIGURE 10: Comparison between measured dynamic displacements: (a) time domain; (b) frequency domain.

In this experiment, the camera captures the target panel with 200 frames per second and detects the motion of target on a connected computer as shown in Figure 9(b). The actual pixel size of a digital image was measured by predesigned target panel; the size of a pixel was calibrated such that 1 pixel corresponds to about 0.185 mm. The shake table was programmed to move along x -direction in sinusoidal motion of 1.5 mm amplitude, with a frequency of 0.5 Hz. The whole measurement was taken for 10 seconds. In the whole process of experiment, the displacement obtained by vision system was compared with that measured by the grating ruler. The size of the captured image was 500×500 ; the algorithm parameters are set as $m = 100$; and the scope is square with the size of 50×50 which covers part of the white circular target panel. Some postprocessing is necessary because the vibration signal obtained directly by SVD-based algorithm cannot agree with grating ruler data. Figure 10

shows horizontal displacement time histories measured by the vision-based system and the grating ruler.

The maximum error is 0.15 mm and basically within error limit of 0.1 mm. The extraction precision between these two algorithms is almost the same while the SVD-based algorithm costs less computation time than the NCC-based algorithm. Therefore, both algorithms can successfully accomplish the vibration extraction and obtain the same conclusion. However, the difference is that the SVD-based algorithm concentrates all the vibration information in the second component, whereas the NCC-based algorithm divides the information into two signals. The SVD-based algorithm finds a principal direction of vibration different from the x - and y -directions, and the results of NCC are the projections of the principal direction on x and y similar to rotation of coordinates. Thus, the SVD-based algorithm reveals the vibration essence, whereas the NCC-based algorithm obtains results with actual



FIGURE 11: (a) Viaduct installed with sound barriers; (b) experimental setup; and (c) one image of the shot video.

physical meaning and the unit of results is pixel. Altogether, the vision-based vibration extraction using singular value decomposition is available.

3.2. Measurement of Sound Barrier Experiment. An attempt to extract the vibration of sound barrier is reported to validate the effectiveness of the proposed algorithms in field. 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 has been found that sound barrier is the most effective method of mitigating roadway, railway, and industrial noise sources.

However, with the increase in the train speed, strong suction and impact are introduced when high-speed trains pass by the section installed with sound barriers, leading to violent vibrations of the barriers that may cause the fatigue of materials 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 to install the barriers at the working train line. Therefore, vision-based measurement has been put forward to solve such problem.

The viaduct installed with sound barriers is located in Kunshan, as shown in Figure 11(a). The experiment is set up below the viaduct in Figure 11(b). 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 passed by, a video containing the vibration of sound barrier is shot at 232 frames per second. One image of the video is displayed in Figure 11(c). In practice, the assumption that the deformation of barrier is negligible relative to the vibration is acceptable. Although the barrier is 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.

Applying the SVD-based extraction algorithm to the barrier video, the second and third components signals with

their corresponding spectrum are shown in Figure 12. The algorithm parameters are set as $m = 100$; the scope is square with the size of 50×50 , whose position is displayed as the red frame in Figure 11(c). The second component clearly indicates the moment of train arrival, and three obvious spectral peaks can be observed in its spectrum at 10.42, 21.07, and 45.77 Hz, which can be considered as the characteristic frequencies of the sound barrier. The third component signal indicates the moment of train arrival; however, its spectrum implies that it is more similar to white noise. Thus, the vibration information of the barrier comes together in the second component.

Actually, comparing the proposed method with traditional measurements is difficult because of assembling inconvenience. However, the trends of vibration are obvious when the train arrives and accord with standard signal of sound barriers. The accuracy of SVD-based algorithms is proved in measurement of harmonic motion experiment. In a word, the vision technique can successfully accomplish the vibration extraction of sound barrier.

4. Conclusions

In this study, a simple and effective algorithm based on the SVD methods was investigated for vibration extraction. The algorithms realized the noncontact and remote measurements requiring little expert interference and extra preparation. The SVD-based algorithm skillfully obtains a set of OIB by performing a simple matrix decomposition; the projections of the video frames on this OIB represent the vibration information. Further, through the study of simulation, we found that only the second and third components signals contain the vibration information, which are orthogonal. Thus, the algorithm can be sped up using the SVD function to calculate only the former three singular values. It can be observed obviously that the SVD-based algorithm spends far less time than the traditional NCC-based algorithm in the measurement of harmonic motion experiments. The selection of algorithm parameters was also discussed, providing the foundation for making better use of the SVD-based algorithm. In addition, it is the first time to measure the vibration of high-speed rail trains sound barrier. Both the simulation and experimental study validated the effectiveness of the proposed algorithms. The experimental example also showed the reliable ability to accurately extract vibrating signals in practice. The vision-based techniques can accomplish vibration measurement of large civil structures, where access for installation of conventional instrumentation is difficult.

Further work is under way to research the theoretical explanation of the OIB. Moreover, the obtaining of OIB requires the participation of the former m frames and impedes the real-time extraction of vibration, which needs further improvement.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

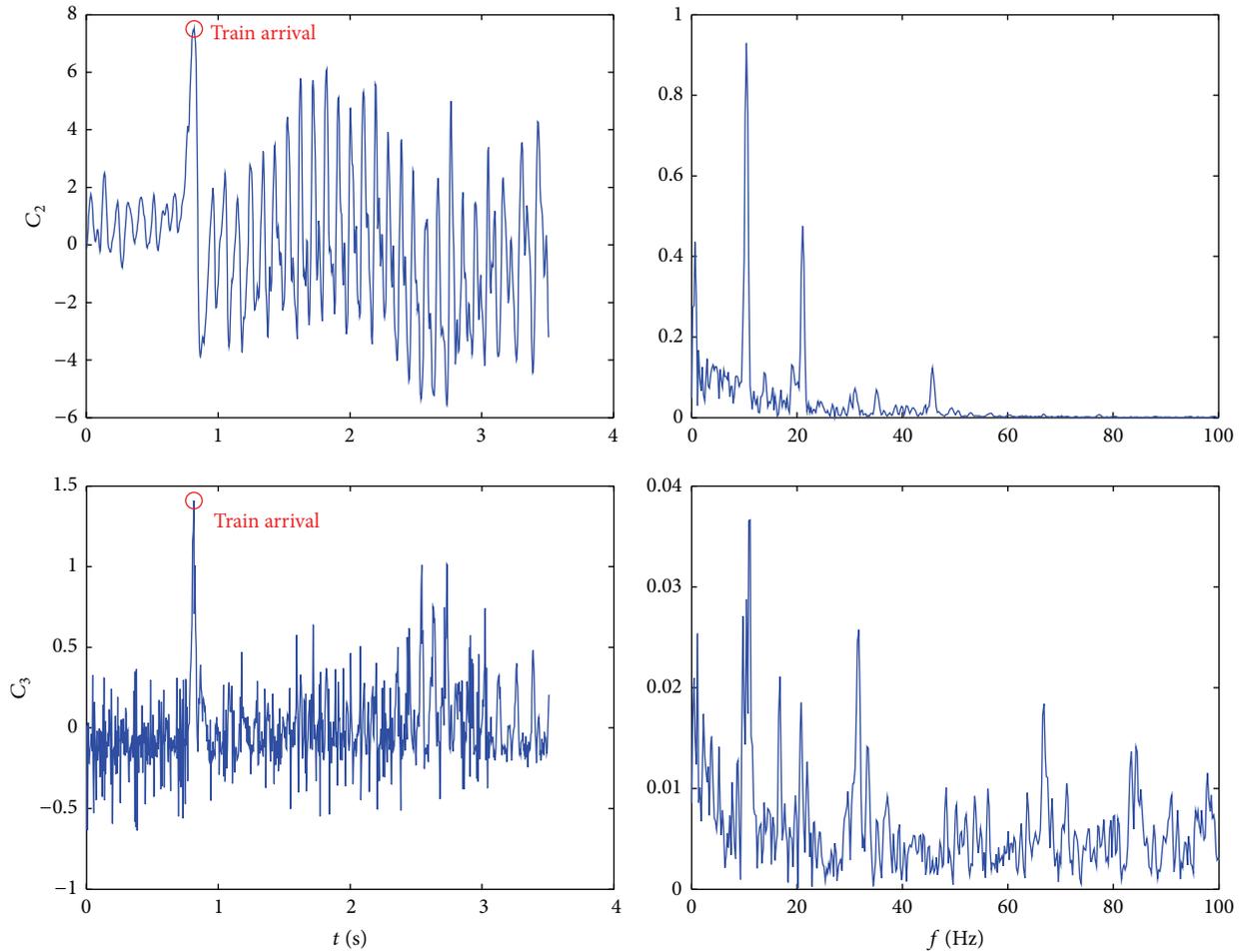


FIGURE 12: Extraction results by applying the SVD-based algorithm to the vibration video of the sound barrier.

Acknowledgments

This work was supported by the National Key Basic Research Program of China (973 Program) under Grant no. 2014CB049500 and the Key Technologies R&D Program of Anhui Province under Grant no. 1301021005.

References

- [1] W. Zhou and D. Chelidze, "Blind source separation based vibration mode identification," *Mechanical Systems and Signal Processing*, vol. 21, no. 8, pp. 3072–3087, 2007.
- [2] M. Meo, G. Zumpano, X. L. Meng, E. Cosser, G. Roberts, and A. Dodson, "Measurements of dynamic properties of a medium span suspension bridge by using the wavelet transforms," *Mechanical Systems and Signal Processing*, vol. 20, no. 5, pp. 1112–1133, 2006.
- [3] L. Xu, J. J. Guo, and J. J. Jiang, "Time-frequency analysis of a suspension bridge based on GPS," *Journal of Sound and Vibration*, vol. 254, no. 1, pp. 105–116, 2002.
- [4] Z. Yuanpeng, Z. Huaning, Z. Wenling, and L. Hefei, "Application of the fourier transform in electronic speckle photography," *Experimental Mechanics*, vol. 42, no. 1, pp. 18–24, 2002.
- [5] J. L. Valin, E. Gonçalves, F. Palacios, and J. R. Pérez, "Methodology for analysis of displacement using digital holography," *Optics and Lasers in Engineering*, vol. 43, no. 1, pp. 99–111, 2005.
- [6] H. H. Nassif, M. Gindy, and J. Davis, "Comparison of laser Doppler vibrometer with contact sensors for monitoring bridge deflection and vibration," *NDT & E International*, vol. 38, no. 3, pp. 213–218, 2005.
- [7] J. Zhao, J. Lin, J. Yao, and J. Ding, "Two-dimensional impact reconstruction method for rail defect inspection," *Mathematical Problems in Engineering*, vol. 2014, Article ID 236574, 9 pages, 2014.
- [8] J.-W. Zhou, D.-H. Liu, L.-Y. Shao, and Z.-L. Wang, "Application of digital image correlation to measurement of packaging material mechanical properties," *Mathematical Problems in Engineering*, vol. 2013, Article ID 204875, 8 pages, 2013.
- [9] B. H. Kim, "Extracting modal parameters of a cable on shaky motion pictures," *Mechanical Systems and Signal Processing*, vol. 49, no. 1-2, pp. 3–12, 2014.
- [10] J. J. Lee and M. Shinozuka, "Real-time displacement measurement of a flexible bridge using digital image processing techniques," *Experimental Mechanics*, vol. 46, no. 1, pp. 105–114, 2006.

- [11] D. L. B. R. Jurjo, C. Magluta, N. Roitman, and P. B. Gonçalves, "Experimental methodology for the dynamic analysis of slender structures based on digital image processing techniques," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1369–1382, 2010.
- [12] P. J. Figueroa, N. J. Leite, and R. M. L. Barros, "Tracking soccer players aiming their kinematical motion analysis," *Computer Vision and Image Understanding*, vol. 101, no. 2, pp. 122–135, 2006.
- [13] F. C. Trigo, F. P. R. Martins, A. T. Fleury, and H. C. Silva Jr., "Identification of a scaled-model riser dynamics through a combined computer vision and adaptive Kalman filter approach," *Mechanical Systems and Signal Processing*, vol. 43, no. 1-2, pp. 124–140, 2014.
- [14] S. W. Park, H. S. Park, J. H. Kim, and H. Adeli, "3D displacement measurement model for health monitoring of structures using a motion capture system," *Measurement*, vol. 59, pp. 352–362, 2015.
- [15] J. de Leeuw, "Principal component analysis of binary data by iterated singular value decomposition," *Computational Statistics and Data Analysis*, vol. 50, no. 1, pp. 21–39, 2006.
- [16] K.-C. Lee, J.-S. Ou, and M.-C. Fang, "Application of SVD noise-reduction technique to PCA based radar target recognition," *Progress in Electromagnetics Research*, vol. 81, pp. 447–459, 2008.
- [17] M. Bydder and J. Du, "Noise reduction in multiple-echo data sets using singular value decomposition," *Magnetic Resonance Imaging*, vol. 24, no. 7, pp. 849–856, 2006.
- [18] A. Al-Zaben and A. Al-Smadi, "Extraction of foetal ECG by combination of singular value decomposition and neuro-fuzzy inference system," *Physics in Medicine and Biology*, vol. 51, no. 1, pp. 137–143, 2006.
- [19] X. Z. Zhao and B. Y. Ye, "Selection of effective singular values using difference spectrum and its application to fault diagnosis of headstock," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1617–1631, 2011.
- [20] C. Cempel, "Generalized singular value decomposition in multidimensional condition monitoring of machines—a proposal of comparative diagnostics," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 701–711, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

