Phase-based Non-contact Vibration Measurement of High Speed Magnetically Suspended Rotor

Cong Peng[1] *Member*, *IEEE*, Cong Zeng[1], Yangang Wang[2], *Member IEEE*

[1] College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China
[2]School of Automation, Southeast University, Nanjing, 210096, China

Abstract—**Video based measurement, as one of the most common non-contact measurement methods for industrial inspection, has been developed rapidly. Digital video cameras offer the benefit of low cost, high automation and are capable to make the simultaneous full-field measurement. This paper focuses on studying the micro-vibrational signal extraction for high speed rotating machinery with digital cameras. An improved phase based motion extraction and learning based video magnification are proposed for measuring the micro-vibration with high frequency and small amplitude. The phase based motion extraction is improved by directly transforming the phase variations into the displacement without the computation of phase gradient. Furthermore, the learning-based motion magnification is utilized to amplify and qualitatively measure the micro-vibration in specified frequency bands. The phase-based motion extraction is a quantitative measurement, while the learning-based video magnification is a qualitative measurement. Experimental results for micro-vibration measurement rising from a magnetically suspended motor system validate the effectiveness of the proposed method.**

Index Terms—Non-contact measurement, vibration analysis, vibration measurement, high-speed magnetically suspended motor, optical flow.

Corresponding author: Yangang Wang; postal address: sipailou 2#, Southeast University, Nanjing, China; telephone number: 86-15110264815 and E-mail address: ygwangthu@gmail.com.

This work was supported by the National Natural Science Foundation of China (No. 61703203 and 61806054), Natural Science Foundation of Jiangsu Province (BK20170812B and K20180355), Fundamental Research Funds for the Central Universities (No. NS2019020) and Foundation of Southeast University (No. 1108000216 and 1108007121).

The paper has not been submitted to any other journal simultaneously.

I. Introduction

The vibration characterized with high frequency and small amplitude is an important factor affecting the working condition and service life of the high speed rotary equipment [1]-[4]. The conventional method to evaluate the vibration performance is realized by vibration sensors, such as the accelerometer. Sensors are reasonably arranged on the equipment according to certain rules. Through effective incentive method, the structural vibration parameter information (such as natural frequency, vibration mode and damping) can be accurately obtained. However, the conventional vibration testing system is complex with low spatial resolution. Moreover, the physical properties, such as the quality, stiffness, damping, and natural frequency, can be influenced by the addition of sensors, which affects the accuracy and objectivity of testing results.

Another dominate non-contact vibration measurement strategy, including the laser vibrometer-based and the digital camera-based measurement method, has attracted much attention due to the increasing requirements of accurate and convenient measurement. The laser vibrometer system can measure structural vibration without any mass-loading effects [5]-[7]. In addition, high spatial resolution can also be achieved for the scanning laser vibrometer system. However, the laser vibrometer system is relatively expensive. Moreover, for structures with large or high-frequency displacements, the measurement results from laser vibrometer systems are not always satisfactory. Compared with the high-cost laser device, the vibration performance can be measured by the digital camera. The digital image can be used to extract quantitative and qualitative information of the vibration. Huang [8] *et al.* proposed an advanced object-tracking algorithm, called the template-corner algorithm. Base on this algorithm, the vibration of wind tunnel can be measured using only one camera. Zhong [9] *et al.* proposed a vibration measurement system based on Non-Projection fringe vision. The system consists of an artificial linear varying-density fringe pattern (LVD-FP) as a sensor and a high-speed camera as a detector. It can accurately measure the instantaneous angular speed and accurately track the instantaneous speed of rotating machinery. Digital image based measurement technology has many advantages, such as non-contact, wide measurement range, full-field measurement, no change in the vibration characteristics of the measured object, etc. In addition, various digital image processing methods are investigated in order to cooperate with the digital camera.

Considerable effort has been made to the 3D digital image correlation (DIC) method for camera-based system [10]-[13].

Wang [10] *et al.* built a structural vibration measurement system based on 3D DIC by using single high-speed CMOS camera. The system can accurately measure the multi-order natural frequencies of structures and analyze the vibration modes of complex structures. German company, Dante, developed a series of measurement systems for structural vibration [11]. Electronic speckle pattern interferometry (ESPI) and 3D DIC were combined to measure deformation, strain field and vibration data of complex structures. This kind of system was portable, flexible and simple to operate. However, it generally required a speckle pattern to mount on the surface of the structure. In addition, the vibration information was obtained through a large number of matching calculations. Considering the large calculation cost of 3D DIC, it is not feasible to process the collected image data in real time, especially for high frequency.

In recent years, the researchers attempted to consider the combination of digital cameras with optical flow methods [14]-[16]. Optical flow method can extract motion information only according to the natural structure of the image without any surface preparation. Traditional optical flow method is based on image intensity, which is very sensitive to noise and interference. When the scene changes abrupt, it would lead to the error of estimating the motion field. Caetano [17] *et al.* developed a vision system based on the correlation and optical flow method. The vision system was utilized to observe the vibration of civil engineering structures and could provide an important data source to achieve a more complete monitoring. Aoyama [18] *et al.* proposed a multithread active vision system that can obtain the optical flow of multiple feature points on a bridge model. This enabled simultaneous measurement of vibrations at multiple points on civil engineering structures. A high-frame-rate optical flow system [19] was developed to estimate the optical flow at high frame rate via an improved optical flow detection algorithm. However, the accuracy measured by the intensity-based optical flow method is relatively low for the micro-motion, which is difficult to distinguish from noise. Moreover, the low computational speed of image intensity limits its application in practical systems, especially in embedded system.

Phase-based optical flow method, based on the phase-based computer vision algorithm, is recently proposed and considered as one of the most effective vibration measurement method [20]-[26]. Fleet *et al.* [20][21] introduced the theory of computer vision based on phase. Phase-based optical flow method is no longer based on the original pixel strength values, but through the analysis of image phase changes to extract motion. The phase information of the image is more robust than the intensity of the image due to the change of contrast and scale. Chen *et al.* [22] extended the phase-based optical flow method to modal identification on simple structures, a cantilever beam and a pipe. Sarrafi *et al.* [24] combined the phase-based optical flow to carry out structural health monitoring of wind turbine blades based on vibration. Moreover, phase-based optical flow method is also combined with motion magnification technique to generate the color and motion changes that can be observed by the naked eye [27]-[32]. Start with the Euler perspective, Wu *et al.* [26] proposed a method called Eulerian video magnification. Spatially filtering the video sequence, then performing time-domain bandpass filtering to extract the interested change signals and amplify them. Wadhwa *et al.* [27] improved this technology and proposed phase-based video motion processing. When amplifying the action, the noise will not be amplified, but the noise will be shifted, so that a better amplification effect can be achieved. Zhang *et al.* [30] put forward a method of video acceleration amplification to address the limitation of magnification when large motion occurs.

However, the conventional phase based optical flow used for the previous measured object is not feasible for the micro-vibration rising from high-speed rotary machines. On one hand, different from the vibrations with low frequency and large amplitude in the majority of the measured system, the micro-vibration due to the high-speed rotor characterized with high frequency and small amplitude. On the other hand, the phase based optical flow needs the computation of flow vectors and phase gradient, which leads to increasing of time complexity and noise sensitivity. Moreover, to the author's knowledge, there have been almost no practical results for the phased based motion extraction method used for the micro-vibration in the high-speed rotary machines.

In this work, an improved phase based motion extraction method is proposed to evaluate the micro-vibration rising from the high-speed magnetically suspended rotor system. In order to improve signal to noise ratio (SNR), the video downsampling and Gaussian blur are introduced in the proposed motion extraction method. Moreover, the learning-based motion magnification is cooperated with phase information to amplify the small imperceptible motions in specified frequency bands. It is adopted to effectively enables visualization of the micro-vibration. There are three main contributions of the work. First, the quantitative and qualitative measurement for rotary machine micro-vibration by phase-based motion extraction. Then, the computation cost can reduce since the phase gradient does not need to compute anymore. In addition, the robustness of the motion computation is greatly improved since the noise in the phase gradient is removed in theory.

II. THEORY OF VIBRATION FOR ROTATING MACHINERY

In this section, we first develop the dynamic model of the magnetic bearing-rotor system. Then we analyze the micro-vibration source and add the main vibration source to obtain the complete dynamic model of the bearing-rotor system. The dynamics analysis of the magnetic bearing-rotor system is the base for the vibration measurement.

A. Dynamics modeling of the rotating machinery

The magnetically suspended rotor is considered in this work as the typical high-speed rotating machinery. According to the Newton's law, the dynamical model of the magnetic bearing-rotor system is presented as,

$$M\ddot{\xi}_i + G\dot{\xi}_i = F \tag{1}$$

where F is the generalized force, M is the diagonal mass matrix, and G is skew-symmetric gyroscopic matrix. To describe the dynamics of the rotor, it is convenient to introduce the generalized coordinates ξ_g, the bearing coordinates ξ_b, and the sensor coordinates ξ_s. The coordinates can be transformed and have the relationship as follows,

$$\xi_g = T_b\xi_b \tag{2}$$

$$\xi_g = T_s\xi_s \tag{3}$$

where T_b and T_s are the transformation matrices. Then the generalized force vector F can be obtained from the magnetic force vector F_m by the coordinates transformation

$$F = T_bF_m \tag{4}$$

where F_m is the magnetic force and can be linearized in the neighborhood of an operating point as

$$F_m = K_\xi\xi_b + K_i i \tag{5}$$

where K_ξ and K_i are the parameter matrices for displacement stiffness and current stiffness, i is the control current for the active magnetic bearings.

In the work, in order to stabilize the rotor system, the decoupled PD controller and the cross feedback algorithm are applied as composite control system for the active magnetic bearing. Then the control currents are presented as,

$$i = -K_A K_S D_i \left[K_P \xi_g + K_D \dot{\xi}_g \right] \tag{6}$$

where K_P and K_D are the proportional and derivative gain matrices. K_A and K_S are the power amplifier and sensor gain matrices. D_i is the current distribution matrix, which converts the control current from the controller into the currents acting on the magnetic bearing coils. It is noted that D_i is a pseudoinverse of the coordinates transformation T_b. The pseudoinverse is calculated by the Moore-Penrose matrix inverse, which is given as,

$$D_i = {T_b}^T \left(T_b {T_b}^T \right)^{-1} \tag{7}$$

Substituting (1), (2), and (4)-(7) into (3), the complete dynamic model of the closed-loop magnetic bearing-rotor system is

$$\mathrm{M}\ddot{\xi}_i + G\dot{\xi}_i + T_b K_i K_A K_S {T_b}^T \left(T_b {T_b}^T \right)^{-1} K_D \dot{\xi}_g + \left[T_b K_i K_A K_S {T_b}^T \left(T_b {T_b}^T \right)^{-1} K_P - T_b K_\xi {T_b}^T \left(T_b {T_b}^T \right)^{-1} \right] \xi_g \tag{8}$$

B. Micro-vibration analysis

The main source causing the micro-vibrations in the rotating machinery is the rotor mass imbalance. The mass imbalance is composed of the static imbalance and the dynamic imbalance. The static imbalance leads to the offset of the mass center to the geometric center and generates the synchronous micro-vibration forces. On the other hand, the dynamic imbalance, which causes the synchronous micro-vibration torques, results in the misalignment of the principal axis and the geometric axis of the rotor. That is, both the static imbalance and dynamic imbalance are caused by the non-colinear relationship of the geometric axis and the inertial axis, which is shown in Fig. 1.

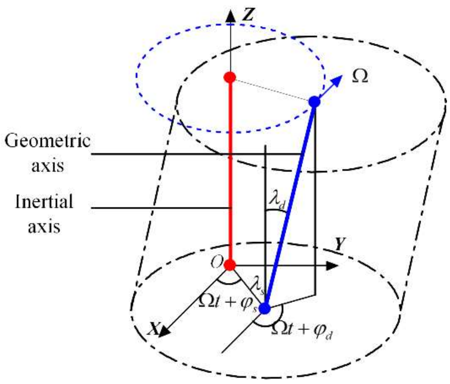

Fig. 1 The non-colinear relationship between the geometric axis and the inertial axis.

In order to describe the mass imbalance disturbance, the displacement of the geometric axis can be presented by the inertial axis with the misalignment component,

$$\xi_g = \xi_i - \Delta\xi \tag{9}$$

with

$$\Delta\xi = \lambda cos(\Omega t + \varphi) \tag{10}$$

where λ and φ are the amplitude and the initial phase of the static imbalance, respectively.

Substituting (9) into (8), the complete micro-vibration dynamical model for the magnetic bearing-rotor is presented as

$$\mathrm{M}\ddot{\xi}_i + G\dot{\xi}_i + T_b K_i K_A K_S T_b{}^T\left(T_b T_b{}^T\right)^{-1} K_D\left(\dot{\xi}_i - \Delta\dot{\xi}\right) + \left[T_b K_i K_A K_S T_b{}^T\left(T_b T_b{}^T\right)^{-1} K_P - T_b K_\xi T_b{}^T\left(T_b T_b{}^T\right)^{-1}\right] \cdot (\xi_i - \Delta\xi) = 0 \tag{11}$$

It can be observed that the mass imbalance enters the closed-loop system inevitably. The micro-vibration of the rotating machinery is characterized with the rotational speed. Thus, the frequency of the micro-vibration is varied with the rotating speed. In addition, the measurement of magnetic bearing rotor system is different from that of the traditional bearing rotor system. The vibration of traditional bearing rotor system is caused by the lubricating oil in the bearing. Its harmonic characteristics are very messy and irregular, which is not conducive to the performance verification of the video-based measurement method. Compared to the traditional bearing rotor system, the vibration characteristics of magnetic bearing rotor system are quite clear due to the definite vibration sources.

III. PHASE-BASED MOTION EXTRACTION FOR VIBRATION MEASUREMENT AND MOTION MAGNIFICATION

In this section, an improved phase-based displacement extraction algorithm is proposed to extract the vibration displacement signal of the magnetically suspended motor. Fig. 2 shows the overall workflow of this work. First, the motion in the spatial domain of the image can be extracted by calculating the local phase changes in the frequency domain. Then, in order to visualize the vibration process, the learning-based motion magnification technique is used to magnify the very small motion with high frequency.

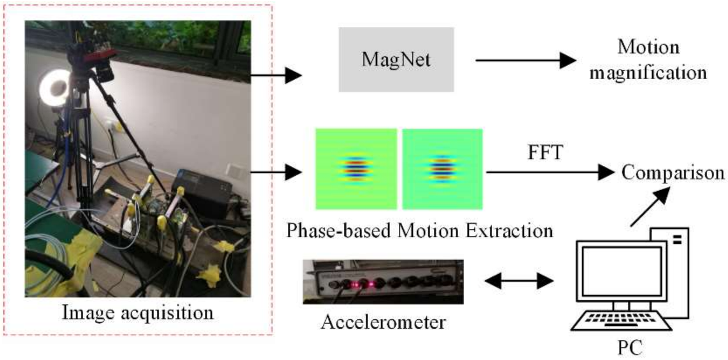

Fig. 2 Workflow of video-based vibration measurement.

A. Phase-based Motion Extraction

Video consists of a series of images. Images usually contain two domains, the spatial domain and the temporal domain. The spatial domain corresponds to the intensity value of each pixel in the single image, and the temporal domain corresponds to the relationship between the images of video and time. In addition, images in spatial domain can be decomposed into amplitude signals and phase signals by specific filters. This is similar to the process of decomposing accelerometer signals using Fourier transform or wavelet transform. As indicated by the Fourier transform theorem, any motion in the spatial domain causes the variations of phase in frequency domain. In order to estimate local motion, the 2D Gabor filter is utilized to transform the images of video in spatial domain into the frequency domain.

For a video, $I(x,y,t_0)$ denotes the frame of the video at the time t_0 with the resolution of $M \times N$. x and y represent the horizontal and vertical pixel coordinates of the image (an image can be regarded as a two-dimensional matrix, each element of the matrix is called pixel). The local amplitude and local phase information of the image is computed by spatially band passing with the Gabor filter as follows,

$$A_\theta(x,y,t_0)exp\left(i\phi_\theta(x,y,t_0)\right) = I(x,y,t_0)\otimes(G_\theta + iH_\theta) \tag{12}$$

where $A_\theta(x,y,t_0)$ represents the spatial local amplitude and $\Phi_\theta(x,y,t_0)$ represents the spatial local phase. $\otimes$ denotes the convolution operator. G_θ and H_θ are the real and imaginary parts of the Gabor filter respectively as,

$$\begin{cases} G_\theta = \ exp\left(-\frac{{x_\theta}^2+\gamma^2{y_\theta}^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x_\theta}{\lambda}+\psi\right) \\ H_\theta = \ exp\left(-\frac{{x_\theta}^2+\gamma^2{y_\theta}^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x_\theta}{\lambda}+\psi\right) \end{cases} \tag{13}$$

where $x_\theta = x\cos\theta + y\sin\theta$ and $y_\theta = -x\sin\theta + y\cos\theta$. θ determines the orientation of the Gabor filter, ranges from $0°$ to $360°$. $0°$ represents the horizontal direction and $90°$ represents the vertical direction. γ and ψ denote the wavelet parameter and phase offset of the cosine function. σ denotes the standard deviation of the Gaussian function, which determines the size of the acceptable region of the Gabor filter kernel. The 2D Gabor filter can also be expressed as,

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = exp\left(-\frac{x_\theta{}^2+\gamma^2 y_\theta{}^2}{2\sigma^2}\right) exp\left(i\left(2\pi\frac{x_\theta}{\lambda}+\psi\right)\right) \tag{14}$$

Then we show how phase signals can be converted into displacement signals by spatial band passing with Gabor filters.

Equation (12) can also be defined by the integral,

$$F(x,y,t) = \iint_{-\infty}^{+\infty} I(u,v,t)g(x-u,y-v;\lambda,\theta,\psi,\sigma,\gamma)\,dudv \tag{15}$$

where $F(x,y,t)$ denotes the frequency domain of the frame at time t. $g(\square)$ function represents the Gabor filter with the parameter $(x,y,\lambda,\theta,\psi,\sigma,\gamma)$.

Fig. 3 shows the real and imaginary pairs of 2D Gabor filter oriented at 0°, 45°, 90° and 135°.

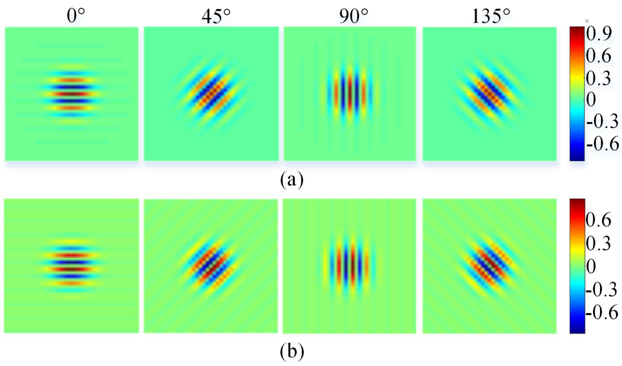

Fig. 3 Gabor filter in four different orientations. (a) The real parts. (b) The imaginary parts.

Assuming that $I(x,y,t_0)$ as the image intensity of the frame at the time t_0 and spatial location (x,y). The local motion $(\Delta x,\Delta y)$ occurs in the frame, then the image intensity of the next frame at the time $t_0+\Delta t$ is expressed as $I(x+\Delta x,y+\Delta y,t_0+\Delta t)$.

For a specific case, $\theta=0$ is selected as the orientation of the Gabor filter, and the motion in horizontal direction will be extracted. The spatial variables then can be simplified to $x_\theta=\mathrm{x}$ and $y_\theta=\mathrm{y}$. The frame at time t and the frame at time $t+\Delta t$ can be transformed into the frequency domain as,

$$F(x,y,t_0) = \iint_{-\infty}^{+\infty} I(u,v,t)e^{\left(-\frac{(x-u)^2+\gamma^2(y-v)^2}{2\sigma^2}\right)}e^{i\left(2\pi\frac{x-u}{\lambda}+\psi\right)}\,dudv \tag{16}$$

$$F(x,y,t_0+\Delta t) = \iint_{-\infty}^{+\infty} I(u+\Delta x,v+\Delta y,t+\Delta t)e^{\left(-\frac{(x-u-\Delta\ )^2+\gamma^2(y-v-\Delta\ )^2}{2\sigma^2}\right)}\cdot e^{i\left(2\pi\frac{x-u-\Delta}{\lambda}+\psi\right)}\,d(u+\Delta x)d(v+\Delta y) \tag{17}$$

Equation (16) and (17) can be reorganized by putting the phase term which are independent of integral variables outside the integral. The new representation is expressed as,

$$F(x,y,t_0) = e^{i\left(-2\pi\frac{x}{\lambda}\right)} \iint_{-\infty}^{+\infty} I(u,v,t) e^{\left(-\frac{(x-u)^2+\gamma^2(y-v)^2}{2\sigma^2}\right)} e^{i\left(2\pi\frac{u}{\lambda}+\psi\right)} \, dudv \tag{18}$$

$$F(x,y,t_0+\Delta t) = e^{i\left(-2\pi\frac{x+\Delta x}{\lambda}\right)} \iint_{-\infty}^{+\infty} I(u+\Delta x, v+\Delta y, t+\Delta t) \cdot e^{\left(-\frac{(x-u-\Delta\ )^2+\gamma^2(y-v-\Delta\ )^2}{2\sigma^2}\right)} e^{i\left(2\pi\frac{u}{\lambda}+\psi\right)} \, dudv \tag{19}$$

Now the phase terms being integrated of (18) and (19) are both $e^{i\left(2\pi\frac{u}{\lambda}+\psi\right)}$. Therefore, the final definite integral of them is identical, which are expressed as φ'. Then, compute the phase angle of (18) and (19),

$$\varphi\big(F(x,y,t_0)\big) = -2\pi\frac{x}{\lambda} + \varphi' \tag{20}$$

$$\varphi\big(F(x,y,t_0+\Delta t)\big) = -2\pi\frac{x+\Delta x}{\lambda} + \varphi' \tag{21}$$

The phase difference is presented as,

$$\varphi\big(F(x,y,t_0+\Delta t)\big) - \varphi\big(F(x,y,t_0)\big) = 2\pi\frac{\Delta x}{\lambda} \tag{22}$$

As can be observed from (22), it is easy to see that the horizontal motion Δx is proportional to the phase difference. Similarly, the motion in other directions are also related to phase differences only by changing the orientation θ of Gabor filter.

However, the conventional phase-based optical flow computes the phase gradient of a spatially bandpassed video to estimate the motion field. Following the Taylor expansion,

$$\varphi\big(F(x,y,t_0)\big)/\varphi t = \frac{\partial\varphi}{\partial x}\Delta x + \frac{\partial\varphi}{\partial y}\Delta y \tag{23}$$

Considering the horizontal direction only, which assumes $\Delta y = 0$, the phase-based optical flow is,

$$\varphi\big(F(x,y,t_0+\Delta t)\big) - \varphi\big(F(x,y,t_0)\big) = \frac{\partial\varphi}{\partial x}\Delta x \tag{24}$$

Different from existing phase-based optical flow method, we make a further step for the above general Taylor expansion equation. Specifically, a concrete φ is used, i.e., the 2D Gabor filter, to obtain a more detailed result, as shown in (22). It is noted that $\frac{\partial\varphi}{\partial x}$ in the existing phase-based optical-flow is replaced by a constant value $\frac{2\pi}{\lambda}$. It does not have to compute the phase gradient in the whole image while the horizontal motion Δx is directly obtained via the proportion to the temporal phase differences.

To improve signal to noise ratio, local amplitude is used by performing a weighted spatial Gaussian blur on the phases. As noise is always low amplitude phase signal, to reduce these meaningless signals, local amplitude is used by performing a weighted spatial Gaussian blur on the phases. For the *N*th frame, the weighted phase signal φ'_N can be computed as,

$$\varphi'_N = (\varphi_N A_N) \otimes h(x, y)/A_N \otimes h(x, y) \tag{25}$$

where A_N and φ_N represent the amplitude and phase signal of the *N*th frame respectively. $h(x, y)$ is 2D Gauss function, which can be expressed as,

$$h(x, y) = exp[-(x^2 + y^2)/\rho^2] \tag{26}$$

The standard deviation of Gaussian filter ρ represents the spatial domain filter widths. The larger the standard deviation, the wider the 2D Gauss image is and the better the filtering effect achieves. This step incurs a small computational cost, but improves the signal-to-noise ratio (SNR) and lowers the noise floor, which can reflect the real signal more.

B. Learning-based Motion Magnification

Learning-based motion magnification is an algorithm for amplifying the visualization of small motion in videos using convolution neural network. Previous motion magnification techniques are based on hand-designed filters, such as complex steerable pyramids[27] or Reisz pyramids[28], which may not be optimal. The objective of learning-based motion magnification is to learn a set of filters directly from examples that can extract the motion signals and then manipulate them to produce magnified frames. In the following, some of important details are described.

To explain the motion magnification problem, we consider a simple case of 1D signal $I(x, t) = f\big(x + \delta(x, t)\big)$ at position x and time t, with the motion $\delta(x, t)$. The magnified signal $\tilde{I}(x, t)$ can be expressed as,

$$\tilde{I}(x, t) = f\big(x + (1 + \alpha)\delta(x, t)\big) \tag{27}$$

for the magnification factor α. To select motion in specify frequency band, a temporal bandpass filter $B(\cdot)$ is used. Then we can amplify the bandpass signal. Fig. 4 illustrates the convolution neural network, which consists of three parts, encoder, manipulator and decoder.

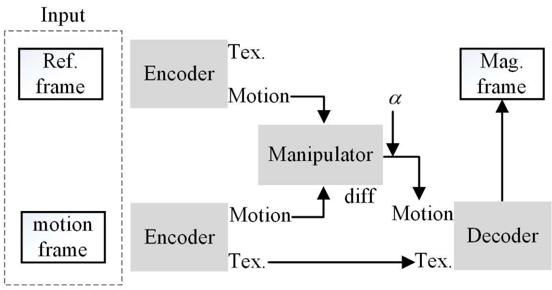

Fig. 4 The overview of magnification network architecture.

The encoder can extract motion representation M_1 for the first frame and M_2 for the second frame. The manipulator extracted the motion between two given frames, and multiplied by magnification factor α, generate motion magnification representation $\widetilde{M_2}$. To improve the quality of the result, set some non-linearity in the manipulator,

$$\widetilde{M_2} = M_1 + r\ \left(\alpha \cdot f(M_2 - M_1\ )\right) \tag{28}$$

where $f(\cdot)$ is expressed as a convolution with 3×3 kernel size and one stride with Rectified linear unit (ReLU) activation function, and $r(\cdot)$ is a 3 × 3 convolution followed by a 3 × 3 residual block.

The motion representation is linear in displacement signal. For the given motion representation $M(t)$ at time t, we isolate the motion of interest with a temporal bandpass filter $B(\cdot)$. Amplified that bandpass signal by the magnification factor α, then add it to the motion $M(t)$ to get the amplified temporal-filtering motion representation,

$$M_{temporal} = M(t) + \alpha \cdot B\ \left(M(t)\right) \tag{29}$$

Finally, the decoder reconstructs the modified representation into the resulting motion magnified frames.

Obtaining the real motion magnified dataset is difficult. Oh *et al.* [32] designed a synthetic dataset that capture small motion well. The texture of synthetic datasets obtained by using real image datasets. The foreground objects come from PASCAL VOC dataset [33]. The image of MS COCO dataset [34] is used as background, and the foreground objects are pasted directly onto the background. The background and the direction of each foreground object are random to simulate local motion. To improve the robustness of the neural network, the intensity of some frames is perturbed artificially. Then some data with low contrast and low texture feature are designed to improve the effect of neural network on low quality videos. This synthetic ground-truth magnified video pairs are available online.

The motion magnification neural network follows supervised training. The loss function is given as,

$$L = L(Y, Y') + \lambda\left(L(T_a, T_b) + L(T'_b, T_{Y'}) + L(M_a, M'_b)\right) \tag{30}$$

where L represents the L_1 norm. Y is the real motion magnified video, Y' is the magnified video output from neural network. T_a, T_b is the texture representation of reference frame and motion frame respectively. T'_b is the texture representation of perturbed motion frame, $T_{Y'}$ is the texture representation of perturbed motion magnification frame. M_a is the motion representation of motion frame, M'_b is the motion representation of setting perturbed motion frame. λ is the regularization weight, there is set to 0.1. The whole model is trained by minimizing the loss function.

IV. EXPERIMENAL SETUP

Fig. 5 shows the side view of the experiment scene. The vison measurement system is made of a high-speed camera mounted with a high-quality optical lens (manufacturer Nikon, focal length 20mm), light sources and a computer for data storage. The camera used in this paper is the 12M180MCX high-speed camera of IO Industries Inc. It can adjust the size of any pixel range below the highest resolution 4096×3076. When the output format and the resolution is reduced, the minimum frame period of the camera (maximum frame rate) can be increased and the range for the exposure times is also re-calculated.

The camera is fixed by a tripod and adjusted to the appropriate viewing position. The vertical distance between the camera and the motor is about 0.8m. Meanwhile, the LED lamp is used to illuminate the magnetically suspended motor to provide sufficient brightness conditions and improve the quality of the captured images. When collecting images, the horizontal distance between the LED lamp and the motor is 2.3m, and the height of the lamp is 1.5m above the ground. In this location, the lamp can illuminate the whole motor. A certain brightness is necessary for the image capture. When the light strength is low, the overall effect of the image is dim and the contrast between foreground and background is not obvious enough.

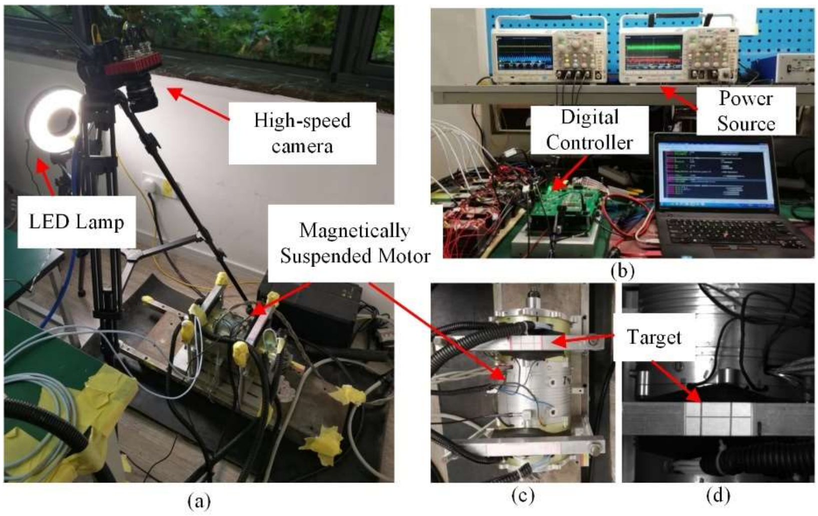

Fig. 5 Experiment setup. (a) The side view of the experiment scene. (b) The control system for the magnetically suspended motor. (c) The enlarged motor scene. (d) The screenshot of the camera.

The probe of the acceleration sensor is vertically installed on the surface of the motor according to its identification direction. Since the motor is necessary to visualize the process of object motion. The magnetic bearing control system is used to stabilize the rotor suspension, and the motor driver is used to rotate the motor and adjust the rotor speed. Fig. 5(b) shows the control system for the magnetically suspended motor, and Fig. 5(c) shows the enlarged motor scene. The rotor speed of the motor is set to 6000rpm, 9000rpm, 12000rpm and 15000rpm (revolutions per minute) in turn. Correspondingly, the frame rate of the high-speed camera is set up to 300fps, 500fps, 600fps and 800 fps (frames per second) to record the sequence of images in turn. All the videos are captured at a resolution of 1024×1024 pixels. An example screenshot from the recorded video is given in Fig. 5(d).

V. EXPERIMENTAL RESULTS

The video sequences of motor at different rotor speeds are captured. We compute the displacement signals and then transform the displacement signals into acceleration signals. The fast Fourier transform (FFT) is performed on the acceleration signal to obtain a frequency spectrum, which is compared with the measurement result of accelerometer. Methods for improving the SNR are discussed. Video magnification is used to visualize micro-vibration.

A. Vibration Analysis

In order to investigate the performance of the camera-based measurement, the vibration frequencies of the magnetically suspended motor are simultaneously measured by traditional accelerometer and high-speed camera. The accelerometer is attached to the side of the motor shell to measure the radial vibration of the magnetically suspended motor. The rotor speed of the motor is set to 6000rpm, 9000rpm, 12000rpm and 15000rpm respectively. The acceleration signals obtained from the accelerometer were transformed into the frequency domain in order to obtain frequency spectrum by performing the fast Fourier transform (FFT).

Fig. 6(a1) shows the measured acceleration signals and the results of the FFT analysis from the accelerometer at four rotor speed cases. According to Fig. 6(b1), the peak frequency of motor is 100Hz, 150Hz, 200Hz and 250Hz respectively at the rotor speed of 6000rpm, 9000rpm, 12000rpm and 15000rpm. The vibration amplitude of the peak frequencies are 0.12804 m/s^2, 0.12116 m/s^2, 0.05748 m/s^2 and 0.08876 m/s^2 respectively. This frequency is the synchronous vibration signal caused by rotor mass imbalance, as the offset of geometric center and the inertial center.

Fig. 6(a2) and Fig. 6(b2) shows the acceleration signal and the FFT result by the video based measurement. All the videos were captured at the frame rate of 1000fps and measured up to 1s. For an image with the resolution of $M \times N$, downsampling is done by a factor of s to obtain an image I with the resolution of $M/s \times N/s$. The neighboring pixels in the size of $s \times s$ is averaged. So the pixels in the $s \times s$ window of the original image become one pixel. The videos are downsampled by a factor of s prior to process to change the scale. With video downsampling, the processing time is reduced and the SNR is improved, which is discussed in detail in the following section. In areas with greater texture or contrast, such as edges, the displacement signals are less noisy. The pixels of the edge of the target are selected as the measurement points. From the downsampled video, the local phase changes between the *N*th frame and the first frame oriented in the horizontal direction for these pixels are recorded. We use spatial Gaussian blur to increase SNR. The standard deviation of Gaussian blur is set to 5. The local phase variations correspond to the displacement signals in units of pixels. Calculate the scale factor of the physical size and pixel width of the target at the same depth in the video frame. The target is 212 pixels tall and 502 pixels wide in the video frame, with 0.189 millimeters per pixel. So the scale factor in this experiment is 0.189. The displacement in units of pixels can be converted to units of millimeters by multiplying this scale factor. Dividing the displacement signals by the number of measurement pixels to obtain the average displacement signal. Laplacian of the Gaussian (LOG) operator is imposed on the average displacement signal to recover acceleration signal. Transform the acceleration signal to the frequency response signal using the fast Fourier transform.

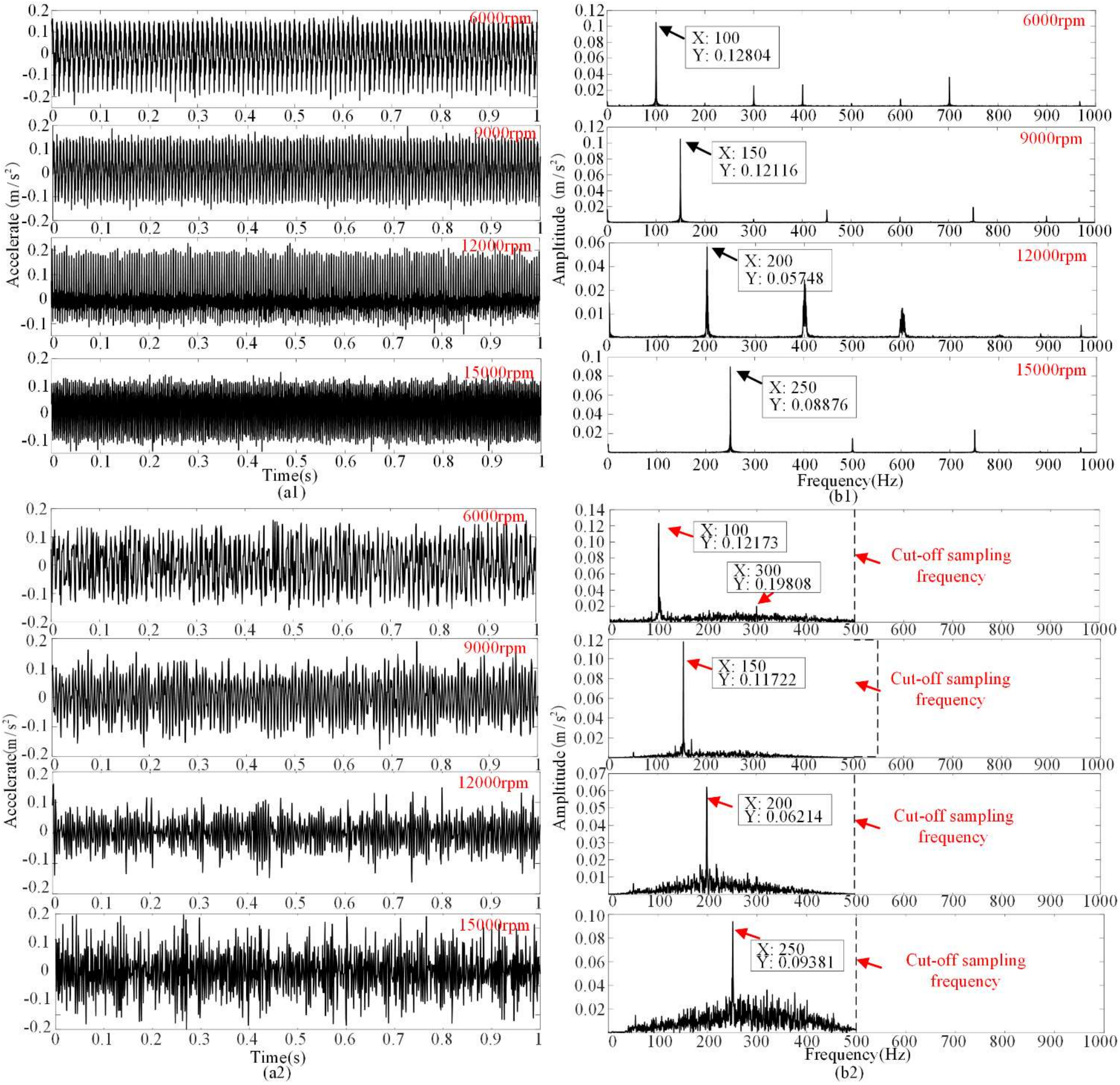

Fig. 6 Vibration measurement of the magnetically suspended rotor at 6000, 9000, 12000, 15000rpm. (a1) The acceleration signal by accelerometer in time domain. (b1) The acceleration signal by accelerometer in frequency domain. (a2) The acceleration signal by camera in time domain. (b2) The acceleration signal by camera in frequency domain.

It can be clearly seen from Fig. 6(b2) that the peak frequencies of 100Hz, 150Hz, 200Hz and 250Hz correspond to 6000rpm, 9000rpm, 12000rpm and 15000rpm were detected, which are consistent with the accelerometer measurement. The vibration amplitude of the peak frequencies from the camera are 0.12173 m/s^2, 0.11722 m/s^2, 0.06214 m/s^2 and 0.09381 m/s^2 respectively, which is close to that from the accelerometer. This confirms that the vibration frequency of high-speed rotating engineering structures can be accurately measured without the need of conventional accelerometers.

When measuring the vibration of the motor at the rotor speed of 6000rpm. The camera was capable of discerning the vibration of the first two peak frequencies of the motor at 100Hz and 300Hz. In this study, due to the limitation of camera frame rate, it is impossible to obtain the higher order frequency doubling vibration signal of motor. If the camera with higher sampling frequency is used, the higher order vibration frequency of the structure can be obtained in principle.

B. Comparisons with 2D DIC

Then, an experiment is formulated to make a comparison between the proposed technique and the state-of-the-art approach, in order to validate the performance of the proposed phase-based measurement method. The pixels on the edge of the target are selected as the measurement points. We measure the displacement signals of the measurement points by applying the 2D Digital Image Correlation and the proposed phase-based motion extraction method to the images of the magnetically suspended motor at 6000rpm. The displacement in units of pixels can be converted to units of millimeters by multiplying the scale factor. Then, the average displacement signals can be converted into acceleration information in order to describe the vibration. Fig. 7 shows the acceleration signal and the FFT result by the accelerometer, the proposed phase-based motion extraction method and the 2D DIC.

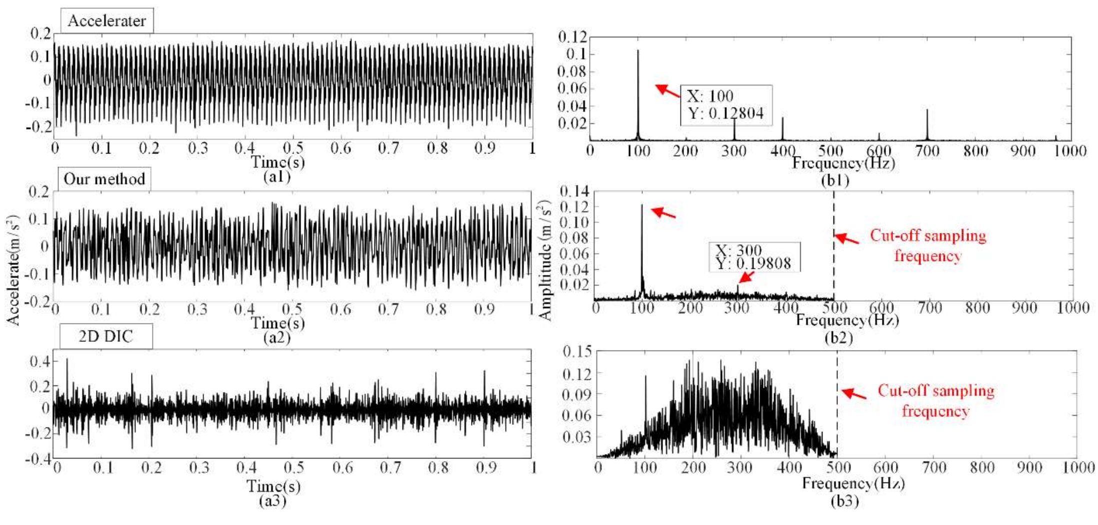

Fig. 7 Vibration measurement of the magnetically suspended rotor at 6000 rpm. (a1) The acceleration signal by accelerometer in time domain. (b1) The acceleration signal by accelerometer in frequency domain. (a2) The acceleration signal by the proposed phase-based extraction method in time domain. (b2) The acceleration signal by the proposed phase-based extraction method in frequency domain. (a3) The acceleration signal by 2D DIC in time domain. (b3) The acceleration signal by 2D DIC in frequency domain.

As can be seen in Fig. 7 (a3) and Fig. 7 (b3), both time-domain signal and frequency-domain signal measured by the 2D DIC do not agree with the predicted results. It can be clearly seen from Fig. 7 (b3) that the first peak frequency measured by the 2D DIC is 100Hz, while the second peak frequency cannot be accurately measured. The measured frequency spectrum contains a lot of noises between 200Hz and 500Hz. However, by the proposed phase-based extraction method (shown in Fig. 7 (b2)), the first peak frequency 100Hz and the second peak frequency 300Hz can be accurately measured with the low noise floor. It can be verified that the accuracy measured by the 2D DIC method is relatively low for the micro-motion, which is difficult to distinguish from noise. And our method is more accurate for the micro-motion measurement.

In addition, we compare the processing time of 2D DIC and our phase-based method for the same video (resolution, 1024 × 1024; frame rate, 500fps; recording time, 1 second). The processing time of our method is 124.34s, while the processing time of 2D DIC is 970.11s. This verifies that our method also greatly reduces the computation time.

C. Gaussian Blur and Video Down sampling Comparisons

The measured acceleration signal in frequency domain from the camera measurement has a large noise floor. Changing the sigma of Gaussian filter or downsampling video at different scales may increase the SNR. The comparisons are made to determine the effect of these factors mentioned above on the noise floor and SNR of the measured spectrum.

Fig. 8 shows the comparison of the effects of the Gaussian blur on the measured spectrum. Based on the measurement results at 6000rpm, the comparisons are made to set the sigma of Gaussian blur to 5, 10, 20. With the increase of standard deviation, the noise floor of vibration signals gradually decreases while preserving the peak frequency.

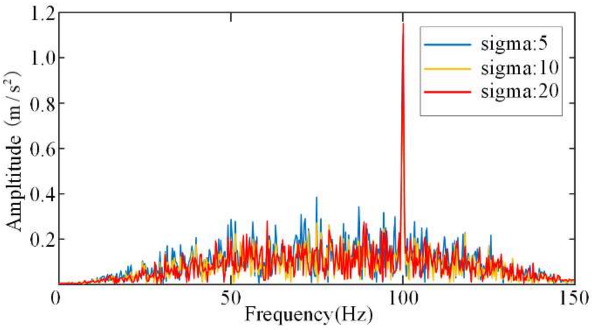

Fig. 8 Comparison of the effects of the standard deviation of Gaussian blur on the acceleration signal in frequency domain.

Fig. 9 shows the comparison of measured spectrum at different scales for downsampling. Previously, the video is downsampled by a factor of two in each dimension. Now for comparison, downsampling each dimension of the video four times or eight times. From the comparison, more spatial downsampling can significantly lower the noise floor and improve the SNR of the measured frequency spectrum without changing the amplitude of vibration peak. More spatial downsampling can reduce the processing time at the same time.

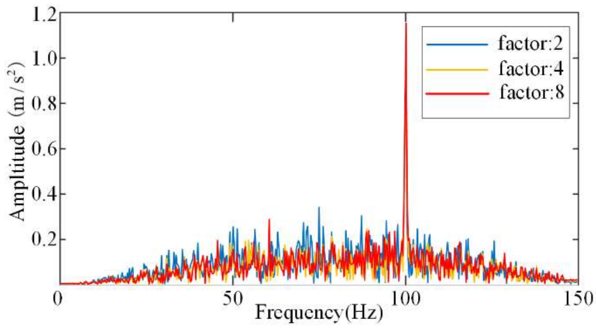

Fig. 9 Comparison of the effects of the video dawnsampling on the acceleration signal in frequency domain.

D. Vibration Visualization

The learning-based motion magnification algorithm is applied to the original captured sequence of images to visualize the micro-vibration of motor. To identify the changes in the resonant frequencies, the center frequencies for motion magnification correspond to the resonant frequencies at different rotor speed measured in previous section. All the width of the filters are selected as 10 Hz and all the amplification factor are set as 20. With specific frequency bands of 95-105 Hz, 145-155 Hz, 195-205 Hz and 245-255 Hz for the sequence of videos of the magnetically suspended motor rotating at 6000rpm, 9000rpm, 12000rpm, 15000rpm, we get the magnified motion at the selected resonant frequency.

For better visualization, we paint a high contrast marker on the surface of motor end cover. Take the motion magnified video at 6000 rpm in frequency band of 95-105 Hz as an example, the screenshots of the marker are shown in Fig. 10(a). Select a region of interest around the marker, as shown in Fig. 10(b1) and Fig. 10(c1). Fig. 10(b2) and Fig. 10(c2) shows a spatiotemporal y-t slice of the video of the single column pixels (the red line on the region of interest) for the (b2) original video frames, and (c2) the motion magnification video.

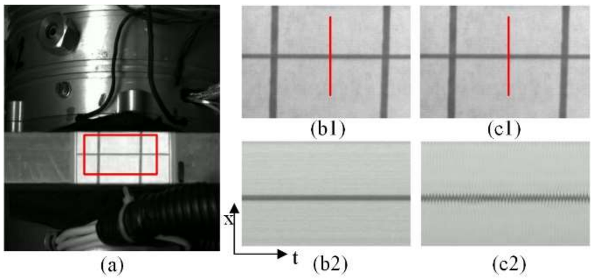

Fig. 10 Motion magnification of the micro-vibration. (a) One of screenshots of captured video. (b1) A region of interest around the marker of the original video. (b2) A spatiotemporal y-t slice of the original video along the red line marked on the region of interest. (c1) A region of interest around the marker of the magnified video. (c2) A spatiotemporal y-t slice of the magnified video along the red line marked on the region of interest.

In the original video, the micro-vibration is invisible, but in the magnified video, the vibration is clear to see. Learning-based motion magnification successfully reveals the motion changes that can not be seen with the naked eye in the video. Through this technology, the effect of visual enhancement can be achieved, and valuable information can be excavated. In conclusion, the quantitative and qualitative measurement is implemented respectively by the phase-based motion extraction and the learning-based video magnification.

VI. CONCLUSION

This paper focuses on the problem of vibration extraction for rotating machinery via a camera-based non-contact measurement. The quantitative and qualitative measurement are studied for a high-speed magnetically suspended motor. An improved phase-based motion extraction method is applied to quantitative measurement of micro-vibration characteristics. In addition, learning-based motion magnification is employed to visualize the micro-vibration which is invisible to the naked eye. We acquire the videos of motor rotating at 6000 rpm, 9000 rpm, 12000 rpm and 15000 rpm, respectively. The motion power spectrums for the acceleration signals are computed and the main synchronous vibration frequency response signals are 100 Hz, 150 Hz, 200 Hz and 250 Hz, respectively. It closely matches those measured signals by the accelerometer. In the future work, the proposed phase-based vibration measurement can be researched to apply for more multipoint measurement situation.

REFERENCES

[1] G. Schweitzer and E. H. Maslen, *Magnetic Bearing: Theroy, Design, and Application to Rotating Machinery*. Berlin, Germany: Spinger, 2009.

[2] C. Peng, M. Zhu, K. Wang, Y. Ren, and Z. Deng, "A Two-Stage Synchronous Vibration Control for Magnetically Suspended Rotor System in the Full Speed Range," *IEEE Trans. Ind. Electron.,* vol. 67, no. 1, pp. 480-489, Jan. 2020.

[3] D. Jiang and P. Kshirsagar, "Analysis and control of a novel power electronics converter for active magnetic bearing drive," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 2222-2232, May. 2017.

[4] C. Peng, J. Fang, and P. Cui, "Dynamics modeling and measurement of the microvibrations for a magnetically suspended flywheel," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 12, pp. 3239–3252, Dec. 2015.

[5] H. Khalil, D. Kim, J. Nam, and K. Park., "Accuracy and noise analyses of 3D vibration measurements using laser doppler vibrometer," *Measurement*, vol. 94, pp. 883-892, DEC. 2016.

[6] H. M. Ismail, C. G. Pretty, M. K. Signal, A. C. Amies, M. Haggers, and J. G. Chase, "Laser doppler vibrometer validation of an optical flow motion tracking algorithm," *Biomed. Signal. Proces. Control*, vol. 49, pp. 322 -327, Mar. 2019.

[7] D. Zhan, D. Jing, M. Wu, D. Zhang, L. Yu, and T. Chen, "An accurate and efficient vision measurement approach for railway catenary geometry parameters," *IEEE Trans. Instrum. Meas.,* vol. 67, no. 12, pp: 2841-2853, Dec. 2018.

[8] M. Huang, B. Zhang, and W. Lou, "A computer vision-based vibration measurement method for wind tunnel tests of high-rise buildings," *J. Wind Eng. Ind. Aerodyn.*, vol. 182, pp. 222-234, Nov. 2018.

[9] J. Zhong, S. Zhong, Q. Zhang, Z. Peng, S. Liu, and Y. Yu, "Vision-Based Measurement System for Instantaneous Rotational Speed Monitoring Using Linearly Varying-Density Fringe Pattern," *IEEE Trans. Instrum. Meas.,* vol. 67, no. 6, pp. 1434-1445, Jun. 2018.

[10] W. Wang, J. E. Mottershead, T. Siebert, and A. Pipino, "Frequency response functions of shape features from full-field vibration measurements using digital image correlation," *Mech. Syst. Signal Pr*., vol. 28, no. 2, pp. 333–347, Apr. 2012.

[11] A. Martínez, J. A. Rayas, R. Cordero, and K. Genovese, "Analysis of optical configurations for ESPI," *Opt. Lasers Eng.*, vol. 46, no. 1, pp. 48–54, Jan. 2008.

[12] A. Lavatelli and E. Zappa, "Modeling uncertainty for a vision system applied to vibration measurements," *IEEE Trans. Instrum. Meas*., vol. 65, no. 8, pp. 1818-1826, Aug. 2016.

[13] A. Lavatelli, and E. Zappa, "A displacement uncertainty model for 2-D DIC measurement under motion blur conditions," *IEEE Trans. Instrum. Meas.,* vol. 66, no. 3, pp: 451-459, Mar. 2017.

[14] S. Baker and I. Matthews, "Lucas-Kanade 20 Years On: A Unifying Framework," *Int. J. Comput. Vision*, vol. 56, no. 3, pp. 221-255, Feb. 2004.

[15] B. K. P. Horn, and B. G. Schunck, "Determining optical flow," *Artif. Intell*., vol. 17, no. 1-3, pp. 185-203, Aug. 1981.

[16] T. Brox and J. Malik, "Large displacement optical flow: descriptor matching in variational motion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 500-513, Mar. 2011.

[17] E. Caetano, S.Silva, and J.Bateira., "A vision system for vibration monitoring of civil engineering structures," *Exp. Tech.,* vol. 35, no. 4, pp. 74-82, Aug. 2011.

[18] T. Aoyama *et al.,* "Vibration sensing of a bridge model using a multithread active vision system," *IEEE-ASME Trans. Mechatron.*, vol. 23, no. 1, pp. 179-189, Feb. 2018.

[19] I. Ishii, T. Taniguchi, K. Yamamoto, and T. Takaki., "High frame-rate optical flow system," *IEEE Trans. Circuits Syst. Video Tech*., vol. 22, no. 1, pp. 105–112, Jan. 2012.

[20] D. J. Fleet, and A. D. Jepson, "Computation of component image velocity from local phase information," *Int. J. Comput. Vision*, vol. 5, no. 1, pp. 77-104, Aug. 1990.

[21] D. J. Fleet, and Y. Weiss, "*Optical Flow Estimation,*" in *Handbook of Mathematical Models in Computer Vision*, New York: Springer, 2006.

[22] J. G. Chen, N. Wadhwa, Y. Cha, F. Durand, W. Freeman, and O. Buyukozturk, "Modal identification of simple structures with high-speed video using motion magnification," *J. Sound Vib*., vol. 345, pp. 58-71, Jun. 2015.

[23] Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar, and D. Mascarenas, "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification," *Mech. Syst. Signal Pr*., vol. 85, pp. 567-590, Feb. 2017.

[24] A. Sarrafi, Z. Mao, C. Niezrecki, and P. Poozesh, "Vibration-based damage detection in wind turbine blades using phase-based motion estimation and motion magnification," *J. Sound Vib*., vol. 421, pp. 300-318, May. 2018.

[25] C. Liu, A. Torralba, W. Freeman, F. Durand, and E. H. Adelson, "Motion magnification," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 519-526, Jul. 2005.

[26] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 1-8, Jul. 2012.

[27] N. Wadhwa, M. Rubinstein, F Durand, and W. Freeman, "Phase-based video motion processing," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 1-10, Jul. 2013.

[28] N. Wadhwa, M. Rubinstein, F. Durand, and W. Freeman., "Riesz pyramids for fast phase-based video magnification," in *Proc. IEEE Int. Conf. on Comput. Photogr.*, 2014, pp. 1-10.

[29] M. Elgharib, M. Hefeeda, F. Durand, and W. Freeman. "Video magnification in presence of large motions," in *Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit.*, 2015, pp. 4119-4127.

[30] Y. Zhang, S. L. Pintea, and J. C. V. Gemert. "Video Acceleration Magnification," in *Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit.*, 2017, pp.502-510.

[31] N. Wadhwa *et al.,* "Eulerian video magnification and analysis," *Commun. ACM*, vol. 60, no. 1, pp. 87-95, Dec. 2016.

[32] T. H. Oh *et al.,* "Learning-based Video Motion Magnification," presented at *Eur. Conf. on Comput. Vis*., Germany, Sep. 8-14, 2018.

[33] M. Everingham and J. Winn, "The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Development Kit," *Int. J. Comput. Vision*,, vol.111, no.1, pp. 98-136, 2006.

[34] T. Y. Lin *et al.*, "Microsoft COCO: Common Objects in Context," in *Proc. Eur. Conf. on Comput.*, 2014, pp.740-755.