# A Pixel Array Framework of Vision-based Remote Vibration Measurements for Modal Analysis and Condition Monitoring

**Miaoshuo Li**

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy

School of Computing and Engineering

The University of Huddersfield

April 2022

# COPYRIGHT STATEMENT

# DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institutes of learning.

In addition, following the guidance of university, some published materials are incorporated in this thesis. Specifically, [1]-[5] of the list of publications are declared as my own written works. The details of reproduction are stated in the preface of each chapter and annotated at that specific paragraph.

*Miaoshuo Li*

# ABSTRACT

Vision-based Remote Vibration Measurement (VRVM) has been actively researched in recent years, because its non-contact and full-field property can fulfil non-intrusive monitoring of large infrastructures and measuring vibration of structure that is difficult to install sensors. As data collected from camera are images rather than vibration response, there is a necessary process of extracting vibration signal from video footages, however, existing VRVM methods provide low accuracy in acquiring vibration responses, which is due to its inherent weakness.

Targeting at addressing this key limit, this PhD study dedicates to develop a novel framework based on the principle of array signal processing (ASP). In recent decades, ASP has advanced rapidly on extensive applications including radar, sonar, seismology, acoustic and so forth. In this study, a pixel array signal processing (PASP) framework is established by analogy between the pixel array (selected pixels of image) and other advanced arrays in relevant area. The proposed framework consists of three phases: The first is to constitute a pixel array by selecting characteristic pixels from reference frame of video. The second is to implement array signal processing paradigm to obtain vibration responses with high signal-to-noise ratio (SNR). Finally, it is to extract the diagnostic features or modal parameters from the filtered signals. Under this framework, a group of methods are developed to meet requirements of different scenarios, and thus lead to following key findings and novel contributions:

It has been derived in theory that the maximum gain of Signal-to-noise ratio (SNR) is a determined value for a given video dataset, which is equal to the number of array elements (characteristic pixels). It suggests that higher number of the characteristic pixels should be used for achieving higher SNR vibration responses.

Subsequently, a mode-shape-based adaptive spatial filtering (MASF) approach, a method based on singular value decomposition value (SVD) and a data independent spatial filtering (DISF) approach are proposed respectively. In particular, the proposed MASF was mathematically proved to be a least mean square filter (LMSF). Moreover, the gain of MASF was verified through simulation on synthetic video footages, in which MASF enhanced SNR by times of pixel number (about 2000) to identify weak mode from noisy images.

Finally, extensive experiments were carried out for fulfilling two common scenarios, including: (1) an experimental modal analysis for a free-free beam with high frequency (*up to 6,389Hz*) modes but small responses (*at micrometre lever*), along with (2) the condition monitoring of a multistage gearbox with mesh frequency (*up to 1,145Hz*); and (3) a free-fixed wind turbine blade with low frequency (*up to 250Hz*) modes but high responses (*over 0.2 meter*), along with (4) the condition monitoring of a reciprocating compressor with rotating frequency (*about 15Hz*). These results validate the outperformance of proposed methods in accuracy, efficiency, and robustness, which is highly conducive to development of the vision-based modal analysis and condition monitoring.

# ACKNOWLEDGMENT

This PhD project is sponsored by the fee-waiver scholarship scheme from the University of Huddersfield (UoH) and a scholarship from China Scholarship Council (CSC). I would like to give my grateful acknowledgements to the UoH and CSC for their financial support to this research.

I would like to appreciate my main supervisor, Professor Fengshou Gu sincerely, who gave me perceptive guidance in research and warm-hearted help in life. He helped me adapt to my Ph.D. study in UK quickly and supported me to acquire rich research achievements.

In the meanwhile, I am grate to my co-supervisor, Professor Andrew D. Ball. His encouragement technical support helped me overcome difficulties in research.

I also would like to thank my colleagues in the Centre for Efficiency and Performance Engineering (CEPE), School of Computing and Engineering at the UoH. I had an unforgettable experience with these nice peers.

I also would like to appreciate Prof. Feng Gao and Dr. Yongjia Xu et al in the EPSRC Future Metrology Hub at the UoH, thank them for providing research equipment generously, which is important to my research achievement.

Special thanks extended to my wife Renren Chen and my 5-years old son Yixuan Li, for their understanding, support, and encouragement during the PhD study.

Finally, I would like to thank my parents and other relatives for their long-term understanding and support of my study.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ASF | Adaptive Spatial Filtering |
| ASP | Array Signal Processing |
| CM | Condition Monitoring |
| CSLDV | Continuous Scanning Laser Doppler Vibrometer |
| CW | Continuous Wave |
| DISF | Data-independent Spatial Filtering |
| EMA | Experimental Modal Analysis |
| ESPI | Electronic Speckle Pattern Interferometry |
| FEA | Finite Element Analysis |
| FEM | Finite Element Model |
| FMCW | Frequency-modulated Continuous WAVE |
| LDV | Laser Doppler Vibrometer |
| LMSF | Least Mean Square Filter |
| LSCF | Least-squares Complex Frequency Domain |
| MAC | Modal Assurance Criterion |
| MASF | Mode-shape-based Adaptive Spatial Filtering |
| MEMS | Micro-electromechanical Systems |
| OMA | Operational Modal Analysis |
| PAF | Pixel Array Framework |
| PASP | Pixel Array Signal Processing |
| RMS | Root Mean Square |
| SHM | Structural Health Monitoring |
| SINAD | Signal to Noise and Distortion Ratio |
| SLDV | Scanning Laser Doppler Vibrometer |
| SNR | Signal to Noise Ratio |
| SVD | Singular Value Decomposition |
| TSA | Time Synchronous Average |
| UAV | Unmanned Aerial Vehicle |
| VRVM | Vision-based Remote Vibration Measurement |
| VVM | Vision-based Vibration Measurement |
| WIR | Window-based Image Registration |

# LIST OF NOMENCLATURE

| | |
|---|---|
| $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ | Reference frame |
| $\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})$ | Predicted sharp image |
| $\boldsymbol{b}(\boldsymbol{x},\boldsymbol{y})$ | Point spread function |
| $\boldsymbol{F}(\boldsymbol{x})$ | One-dimension representation of reference frame |
| $\boldsymbol{p}(\boldsymbol{x})$ | One-dimension representation of predicted sharp image |
| $\boldsymbol{b}(\boldsymbol{x})$ | One-dimension representation of point spread function |
| $\boldsymbol{P_s}$ | Power of signal |
| $\boldsymbol{P_d}$ | Power of distortion |
| $\boldsymbol{P_n}$ | Power of noise |
| $\boldsymbol{I_A}(\boldsymbol{t})$ | Intensity signal from sensor array |
| $\boldsymbol{F}'$ | Image gradient of reference image |
| $\boldsymbol{u}(\boldsymbol{t})$ | Sensor elements of displacement |
| $\boldsymbol{\varphi_i}$ | Mode shape vector |
| $\boldsymbol{z_i}(\boldsymbol{t})$ | Modal displacement |
| $\boldsymbol{\Phi}$ | Mode shape matrix |
| $\boldsymbol{\eta}(\boldsymbol{t})$ | Noise vector |
| $\boldsymbol{W}$ | Weighted matrix |
| $\hat{\boldsymbol{z}}(\boldsymbol{t})$ | Estimator of modal displacement matrix |
| $\boldsymbol{w_i}$ | Weight vector |
| $\hat{\boldsymbol{z}}_{\boldsymbol{i}}(\boldsymbol{t})$ | Estimator of modal displacement |
| $\boldsymbol{SNR}$ | Signal to noise ratio |
| $\boldsymbol{G_a}$ | Array gain |
| $\boldsymbol{m}$ | Number of sensor elements |
| $\boldsymbol{n}$ | Desired order of modes |
| $\boldsymbol{w_{ini}}$ | Initial value of weight vector |
| $\boldsymbol{w_{opt}}$ | Optimizer of weight vector value |

# LIST OF PUBLICATIONS

[1]. **Miaoshuo Li,** Guojin Feng, Rongfeng Deng, Feng Gao, Fengshou Gu, and Andrew D. Ball. "Structural Vibration Mode Identification from High-Speed Camera Footages using an Adaptive Spatial Filtering Approach." ***Mechanical Systems and Signal Processing*** Volume 166,2022,108422, ISSN 0888-3270.

[2]. **Miaoshuo Li,** Guojin Feng, Fengshou Gu, and Andrew Ball. "Investigating into Minimum Detectable Displacement Signal in Image-Based Vibration Measurement." In International Conference on Maintenance Engineering, pp. 882-894. Springer, 2020.

[3]. **Miaoshuo Li**, Robin Appadoo, Wei Hu, Fengshou Gu, and Andrew Ball. "Condition Monitoring of Reciprocating Compressor Based on Acoustic Imaging." In Advances in Asset Management and Condition Monitoring, pp. 977-983. Springer, Cham, 2020.

[4]. **Miaoshuo Li**, Fengshou Gu, Andrew Ball, and Lin Song. "A homography transformation-based video stabilization method for structural health monitoring using an unmanned aerial vehicle." In Sixteenth International Conference on Condition Monitoring and Asset Management, p. 130. British Institute of NDT, 2019.

[5]. **Miaoshuo Li**, Xianghong Wang, Fulong Liu, Xiaoli Tang, Fengshou Gu, and Andrew D. Ball. " Remote monitoring of wind turbine blades based on high-speed photogrammetry." Condition Monitoring and Diagnostic Engineering Management (COMADEM2018): 321

[6]. Fulong Liu, Jiongqi Wang, **Miaoshuo Li**, Fengshou Gu, and Andrew D. Ball. "Operational modal analysis of Y25 bogie via stochastic subspace identification for the condition monitoring of primary suspension systems." In Proceedings of the 13th International Conference on Damage Assessment of Structures, pp. 166-181. Springer, Singapore, 2020.

[7]. Fulong Liu, Jiongqi Wang, **Miaoshuo Li**, Fengshou Gu, and Andrew D. Ball. "Operational modal analysis in the presence of pulse train and harmonics based on SSI." In Proceedings of the 32nd Condition Monitoring and Diagnostic Engineering Management (COMADEM2019).

# 1 INTRODUCTION

*In the first section, the research status of vision-based vibration measurements is investigated through an extensive literature review. To overcome the weakness of existent methods, this research project aims to propose a novel framework based on the principle of array signal processing, which has found extensive applications in radio antenna, seismic arrays, microphones, accelerometers, and telescopes. Accordingly, the research aim and objects of pixel array framework are proposed.*

## 1.1 Background of non-contact vibration measurements

It is well known that dynamic characteristics plays an important role on structural health monitoring[1], [2], and mechanical design & manufacturing[3], [4]. Additionally, vibration signal is also a key manner of monitoring the operating condition of rotating machinery[5]–[7]. Currently, the vibration measurements for condition monitoring or modal test mainly relies on piezoelectric accelerometer[8], [9], however, the accelerometer also possesses some disadvantages, such as it is inaccurate at low frequency, the sensor's loading have great impact on the lightweight structure, sensing network is complex and expensive for large-scale infrastructure.

For above reasons, the non-contact measure techniques have attracted extensive attention in the past several decades, which include the laser Doppler vibrometer (LDV) [10], [11], radar-based vibration measurement (radar displacement monitoring) [12], vision-based vibration measurement, electronic speckle pattern interferometry (ESPI) [13], [14] and holographic interferometry [15], [16]. Amongst them, LDV, radar displacement monitoring and vision-based vibration measurement are three most active fields over the last decade, because they are advantageous in different scenarios according to the particular requirement of test. Their research status and characteristics are outlined respectively as follow.

- Laser Doppler Vibrometry

Laser Doppler Vibrometry (LDV) was invented just after a few years of invention of laser, which transmits a focused laser beam and receives a scattered light from vibrating object, then the Doppler shift between the incident light and reflected light is used to measure accurate displacement, which is also called as the Mach–Zehnder interferometer [17]. Nowadays, LDV has been widely used in structural health monitoring, Micro-electromechanical systems (MEMS), rotating machinery, hearing, and acoustics [18].

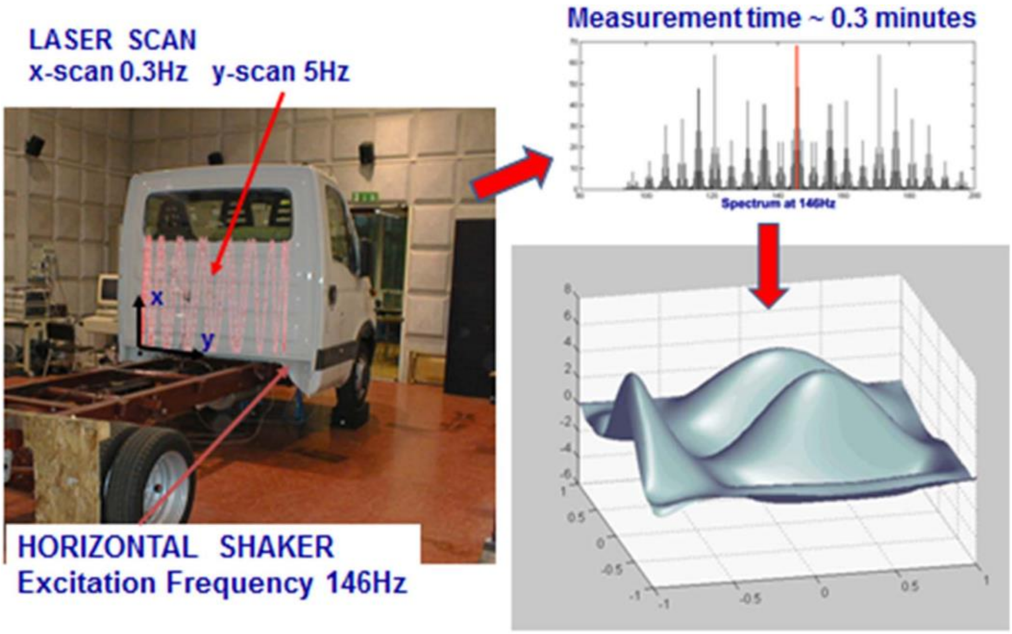



Figure 1-1 CSLDV on a vehicle cab showing scan pattern (left), measured velocity spectrum (top right) and ODS reconstructed from polynomial fit (bottom right) [18], [19]

In particular, the invention of scanning laser Doppler vibrometer (SLDV) boosted its application on the experimental modal analysis (EMA). SLDV can measure an area of structure point by point, in which the orthogonal scanning mirrors controlled by computer are used to relocate the direction of laser beam [18]. In this way, SLDV can reach a much higher spatial resolution than accelerometer array, which can only measure very limited number of points.

Figure 1-1 shows a modal experiment based on a continuous scanning laser Doppler vibrometer (CSLDV) with single frequency excitation. The spectrum displayed the excitation frequency and its harmonic sidebands spaced by the scan frequencies. The deflection pattern is generated by the related frequency components. The precision of CSLDV can be limited by signal drop-outs, primarily related to surface quality, and speckle noise, especially at higher scan speeds. That shows the effect of speckle noise, which is a major disadvantage of the LDV. Due to this optical phenomenon, the laser signal at a few points is missed by ‘speckle drop-out’, and the intensity of noise is related to the roughness of surface [20].

- Microwave radar-based vibration measurements

On the other hand, the radar-based measurement earn also more attention, which can track small displacement with high sensitivity[12], [21]. The principle of radar-based measurement is that transmits electromagnetic signals at specific frequencies and receives echoes to measure the displacement and velocity based on the baseband signal processing. The radar wave can be generally categorized as the single-tone continuous wave (CW) and frequency-modulated continuous wave (FMCW). Unlike CW radar, FMCW radar do not have DC offset problem and can distinguish multiple targets from static clutters. In 2010, a stepped-frequency continuous wave (SFCW) radar is proposed to measure the displacement of civil engineering structure [22]. In 2014, a hybrid radar system combines the FMCW and interferometry mode to achieve a universal performance [23].

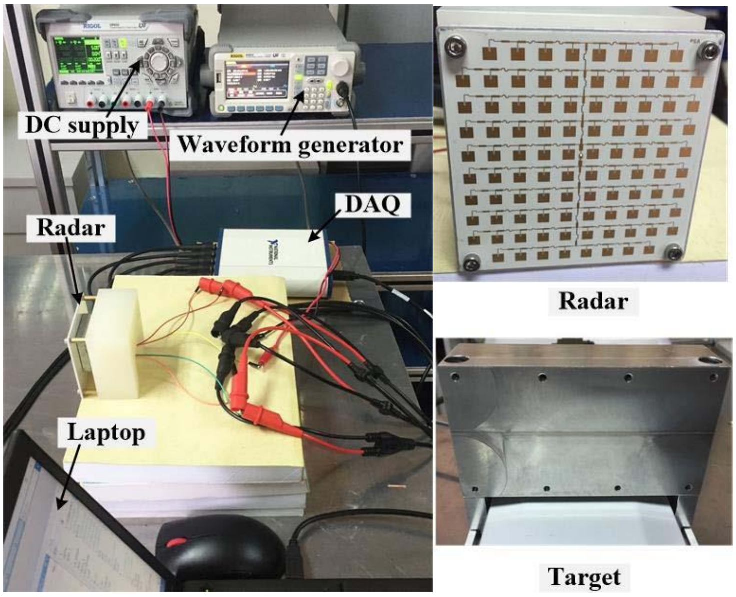


Figure 1-2 The setup of radar-based displacement measurement of a rectangle target [24].

To meet real-time displacement measurement the approximated maximum likelihood approach is proposed in 2017, which can estimate the phase value of beat frequency [24]. The experimental result shows the error is less than 0.025mm. In addition, a parameterized de-alternating (PDA) method is proposed to detect multiple targets, in which the beat frequency is up to 1583.3Hz, but the algorithm cannot separate two similar distance, because of the limited frequency resolution [25]. Furthermore, the radar-based measurement is used to detect vital sign including the heartbeat and the breathing conditions under a very low SNR [26], [27].

However, the accuracy of radar-based measurements are influenced by multi-point scattering, especially when the distance of multiple points to the radar is less than the distance resolution, the beat frequencies will be aliased. In addition, the sampling rate of displacement is limited by the length of sweep circle[26].

- Vision-based vibration measurement (high-speed camera)

Comparably , the high-speed camera is increasingly used to measure vibration with the high spatial resolution [1], which is often called as vision-based [28]–[31], camera-based[32], [33] vibration measurements or sometimes photogrammetry [34] . Like other forementioned non-contact measurement [35], [36], the vision-based methods can be substituted for piezoelectric accelerometer in the scenarios[37], where installing accelerometer is difficult, or the loading of accelerometer affects the characteristics of lightweight object[35]. Furthermore, the high-speed camera can measure vibration at dense spatial points simultaneously at pixel level, which cannot be reached by radar-based methods and scanning laser Doppler vibrometer (SLDV).

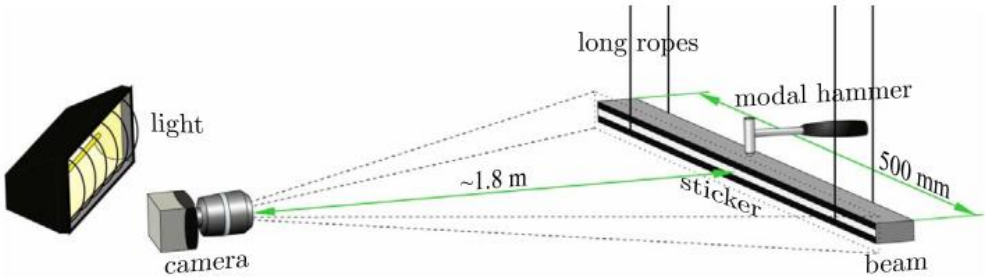



Figure 1-3 A schematic of vision-based modal analysis on a sticker impacted by a hammer [38]

Thus, a large number of research on vision-based vibration measurement arises on modal analysis, health structural monitoring, and condition monitoring of rotating machinery [39]–[41], particularly, the cases of application will be detailed in next section. Inevitably, it also faces a serial of challenges such as the accuracy in high frequency range, robustness to light condition and computational cost.

To compare three techniques, accuracy of LDV is limited by the surface roughness of target; Microwave radar sensing is difficult to distinguish multiple targets with the similar distance to receiver; Vision-based methods require high-quality and stable light condition. Overall, In the competition of three technologies, none of them show absolute superiority or inferiority. In fact, the demand of actual test determines which is the most appropriate measurement, for example, the surface quality of the measured object, the amplitude and frequency range of vibration signal, the measurement distance, and lighting conditions, etc.

On the other hand, the three technologies can be not only competitive, but also complementary. One typical analogy is that autonomous vehicle (AV) generally equipped with cameras, microwave radar, lidar and even ultrasonic sensing to detect objects simultaneously. Similarly, it is possible to fuse laser, radar, vision-based method, and even microphone arrays together to improve the range, spatial resolution, and accuracy of vibration measurements.

## 1.2 Background of vision-based vibration measurements

### 1.2.1 Applications of vision-based vibration measurements

As vision-based vibration measurement has attracted considerable attention in recent decades, [42], [43], a variety of applications of vision-based measurement is introduced to exhibit the superior performance in some typical scenarios [1].

#### 1.2.1.1 Experimental modal analysis (EMA)

According to the application scenarios and conditions, the camera-based modal analysis can be generally classified into two main branches: operational modal analysis (OMA) and experimental modal analysis (EMA) [9],[44]. EMA is a test method for characteristics of a structural vibration which has been developed over the last 50 years [45]. In EMA, the excitations are artificially designed to act on the test specimen, and both the input forces and the responses are collected to identify the modal parameters. Generally, that process is based on the curve fitting methods of frequency response functions (FRFs) in frequency domain or impulse response functions (IRFs) in the time domain [46]. The temporal parameter (natural frequency and damping ratio) and spatial (mode shapes) parameter are two elements of modal properties.

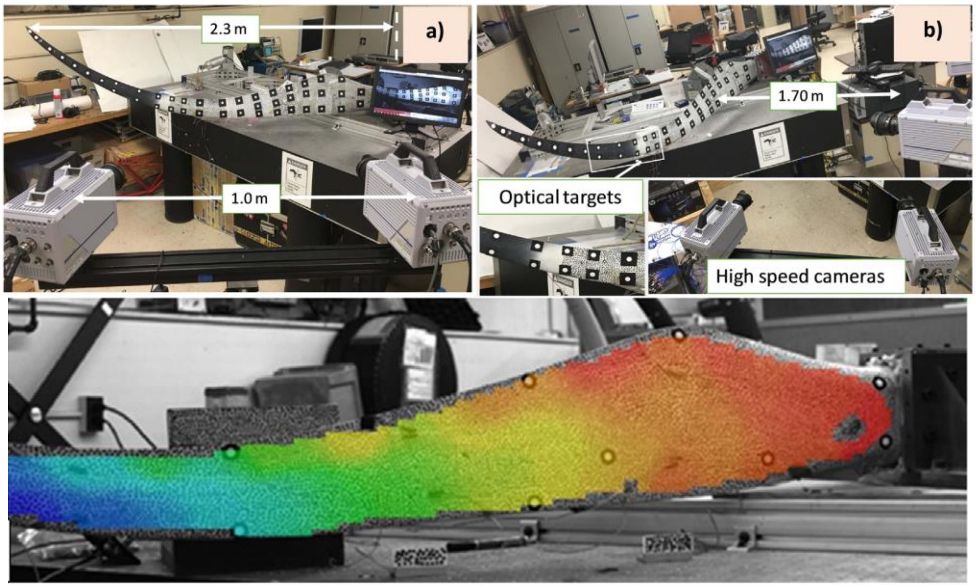



Figure 1-4 The stereo cameras are used to measure the mode shape of a wind turbine [47]

In essence, all the temporal parameters can be identified from response of one single point. However, the spatial parameters can only be identified by responses of multiple points. If accelerometers are used, cost and setup time will increase greatly with the number of measurement points.

In contrast, vision-based measurement can collect thousands of points on object simultaneously without mounting any sensor, as a result, the mode shape can be obtained densely on pixel level. Therefore, a great number of studies on vision-based EMA was conducted [43], [47]. To take just one example, a vision-based EMA for wind turbine blade is carried out by using a pair of high-speed cameras. As shown in Figure 1-4, the obtained mode shape is with a high accuracy and resolution [46].

#### 1.2.1.2 Operational modal analysis (OMA)

Alternatively, operational modal analysis, also known as ambient modal identification, is a method that identifies modal parameters under operating condition and without artificial excitation [48], [49]. Because conducting EMA can be challenging for large-scale infrastructure, in that case, the excitation is often unknown and operational boundary conditions can change the modal properties.

In OMA, the excitation is often unknow and assumed as steady-state broadband random, so that only measuring the output vibration of a structure is required [48]. As the camera can collect the output vibration remotely, The camera-based OMA has been widely applied to the structural health monitoring of large-scale structures in civil engineering [29], [50], such as bridges [33], infrastructures [32] and wind turbines [51]–[53].

One case is shown in Figure 1-5, the mode of wind turbine blade is identified by using a vision-based operational modal analysis (OMA) [54] .The tests were conducted on a pitch controlled, variable speed wind turbine with a rotor diameter and tower height of 80 m. A system consists of four CCD cameras and retro-reflective markers to provide higher visibility.

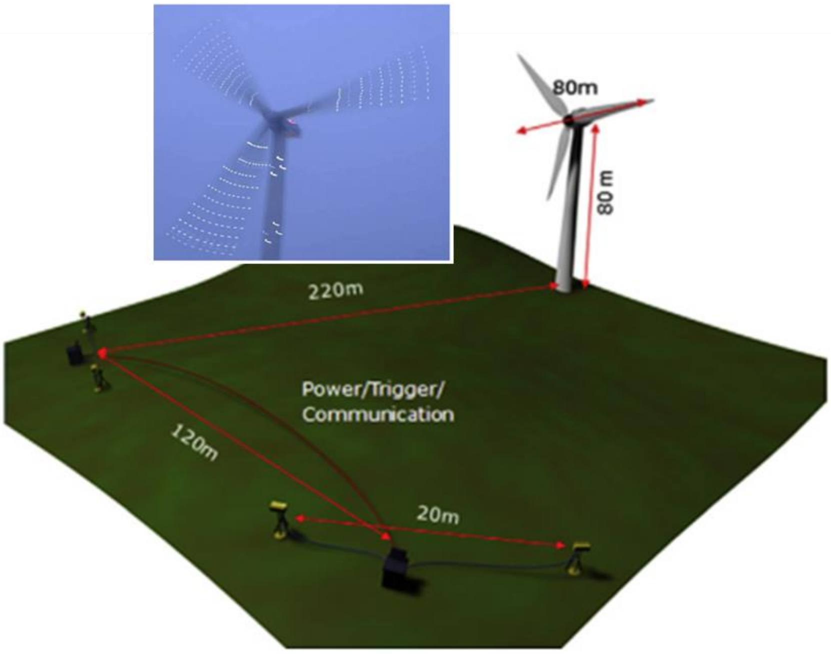



Figure 1-5 A remote shooting system for operational modal analysis of wind turbine blades, on the blades and tower, a set of reflective markers are installed in advance [54].

The result shows measurement error was in the range of 5mm or 1/16,000 of field of view. The deformations on the turbine could be measured with an average accuracy of 25 mm from a measurement distance of 220 m. This case shows the capacity of vision-based method to monitor remote targets.

#### 1.2.1.3 Other pioneering applications

Except for EMA and OMA, more and more new applications of vision-based method are found, due to the non-contact and full-field property. Consequently, an increasing number of outcomes on this topic from academic institutes and commercial companies are published on top journals in recent decades. Some typical examples are taken in last five years as follows.

- Vibro-impact absorber

To measure the vibration of a free ball vibro-impact absorber, the digital image correlation (DIC) combined with the high-speed camera is employed to capture the motion of absorber [35]. The vibro-impact absorber generally consists of two components, first one is a main structure window with a cavity, the second one is free oscillating mass which are interacted during the impact. However, the behaviour of ball and the contact area cannot be measured by classical accelerometer.

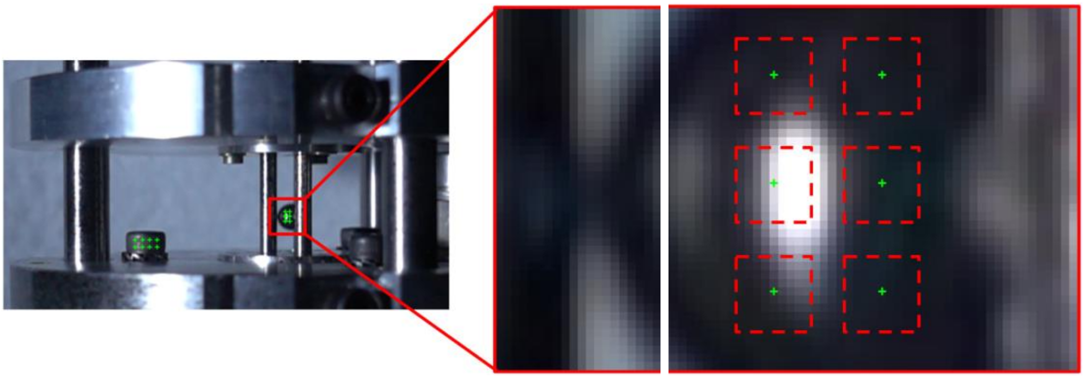


Figure 1-6 Photograph of vibro-impact absorber, the motion of ball is tracked using DIC [35]

As shown in Figure 1-6, the vibro-impact phenomenon of ball can be captured by vision-based method. That overcomes the difficulties encountered to measure the displacement of free oscillating mass with conventional sensors.

- Marine propeller blades underwater

The vision-based method was employed to monitor dynamic response of rotational structure under water [55], by carrying out extra calibration of cameras and correction of light refraction. In this test, 3D displacement fields of marine propeller blades in operation is measured using a pair of stereo cameras. This research shows advantage of remote monitoring for underwater rotating structure.

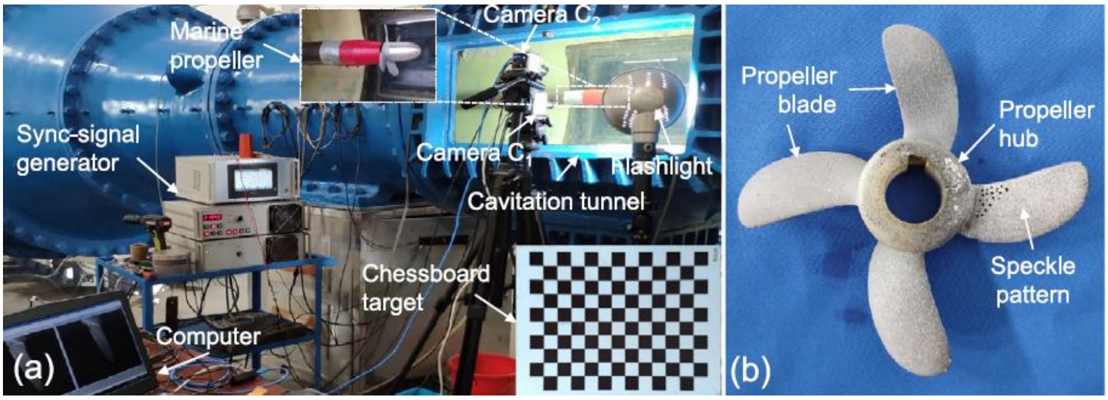



Figure 1-7 The setup of vision-based measurement of a marine propeller blade under water [55]

- Local damage detection

The vision-based measurement was used to localize the damage[51], [56]. The 3D-DIC can measure anomalies of mode shapes and curvature mode shapes caused by local damage [57]. A razor cuts in circular and rectangular membranes with different boundary conditions is detected. Based on 3D-DIC, the locations of damage can be identified in the first two or three mode shapes.

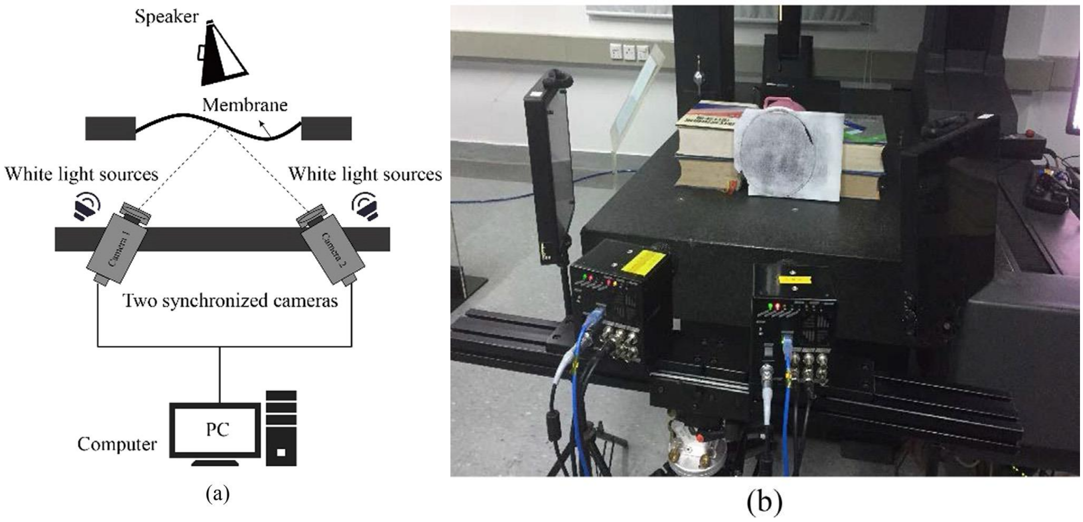



Figure 1-8 The setup of 3D-DIC that include a pair of cameras and the light source, which collect images from two perspective synchronously [57]

A novel method based on 3D-DIC is proposed to detect local damage without excitation signal and any added mass in membranes under different boundary conditions.

#### 1.2.1.4 Challenges in applications

Inevitably, all applications face specific challenges in practice. For example, light condition and the image resolution in long-distance shooting is a big limitation in OMA. In contrast with OMA, the sampling frequency and extracting the small displacement amplitudes of modes at high frequency (generally about micrometre [58]) is also a great challenge for EMA. Because the object tested in EMA [59] is relatively small and generates higher frequency components. Thus, the analysis of dynamic behaviour in higher frequencies plays an important role in EMA. The other challenges also includes complex calibration for camera, computational cost etc.

### 1.2.2 Existing methodology of vision-based vibration measurements

The accelerometer collects vibrational acceleration directly, in contrast, the data collected from camera is a frame sequence, which requires a process of extracting vibrational motion from image. Currently, this extracting processing mainly uses a principle of window-based image registration[60], which belongs to an important subject 'motion estimation' in computer vision [61]. Based on this principle, three algorithms are most developed and used, i.e., digital image correlation[62]–[64], phase-based motion magnification (or optic flow) [65]–[67], and gradient-based optic flow[68], [69]. In this section, the theory, and applications of these algorithms are introduced and analysed respectively.

#### 1.2.2.1 Digital image correlation

The number and quality of published works displays that the digital image correlation (DIC) is the most common approach for vision-based modal analysis currently [1], [44], [70]–[72], which aims at measuring the 2D or 3D strain or displacement based on maximizing the correlation coefficient of given subset between frames [70], [73]. The speckle pattern is usually painted onto the object surface or that is offered by the texture of the specimen's material to enable discrimination between different groups of pixels (subset). DIC technique is used to measure the full-field stress and strain of wind turbine blades were explored [74], [75] and results from DIC agree with strain gauge data well. Recently, the Laser-light speckle is associated with DIC to avoid painting on surface of object [76] .

Helfirck et al. [44] measured the mode shape of a dryer which show the mode shape in high frequency range cannot be measured when the displacement is less than the noise floor [77]. Although DIC involves a large amount of computation and limited by the noise floor, it is proved that DIC is a viable option for EMA and other vibration applications [72].

Based on a systematic uncertainty assessment of DIC, the result shows DIC can provide good results and a high (subpixel) displacement resolution, typically quoted at a hundredth of a pixel [78]. DIC can also be used to estimate the strain over a certain gauge length. The gauge test of 50 pixel, an uncertainty of 0.2 pixel on displacement estimate produce a maximum uncertainty in strain about 0.8%, however, the corresponding uncertainty can decline to 0.02 pixel while reducing the 7-pixel motion to 3.5 pixel.

A high spatial density is another advantage of full-field measurement. A research showed model updating for localized parameter based on DIC and high-speed camera has a better performance than accelerometer [79].

For rotating structures, the vibration analysis is carried out by digital image correlation using image measured by cameras, which is applicable for constant and varying speed[80]. The crucial process is removing rigid body motion components from the primary responses. For this purpose, the translation vector and rotation matrix is calculated, which can denote the rigid transformation between positions in captured two neighbour images. The results of the experiments show that a higher accuracy and clearer spectra of response through getting rid of rigid body motions. If the vibration is subtle, the vibration with single frequency is more identifiable. The results also show the misalignment of mode shape is the shortage of this method.

By tracking the location of a particular marker, the vibration of a rotating beam is measured [81], as an electro-dynamic shaker excites the beam. But the requirement of test rig is very high, and the density is low due to the number of markers on structure is very few. In another case, a rotating wind turbine blade are measured using 3D-DIC, the motion contains both in-plane and out-of-plane motion. The balanced and unbalanced turbine is compared, the balanced turbine showed two harmonic components on out-of-plane spectra. With adding weight, sub-harmonics occurs successively.

#### 1.2.2.2 Phase-based motion magnification

The phase-based motion magnification [66], [82] or phase-based optical flow [51], [83] is another approach to estimate displacement for modal analysis, [65], [84], to which, the surface preparation is not needed [85]. The advantage is the phase of image is much more robust than original intensity data, such as it is not prone to be contaminated and it can determine motion field more accurately. Meanwhile, its computation cost is lower than 3D DIC and 3DPT, by avoiding the image registration in two stereo cameras. The local phase and amplitude are determined by filtering an image with the complex steerable pyramids, and then the constant phase contours are used to estimate the displacement or magnify the motion [67], [86].

Chen et al. first identified the modes with phase-based motion magnification. Through cantilever beam and pipe test, modes were extracted and visualized. The phase-based motion magnification can be associated with DIC, 3D point tracking and other optical methods [87]. In [47], the modes of a 2.3 meter wind turbine blade are extracted by combining motion magnification and DIC, which shows the phase-based motion magnification can improve the signal to noise ratio (SNR) of the estimated displacement. However, the procedure of loading local steerable filter is complex and time-consuming, and the determination of parameters in the algorithm is difficult, especially the selection of frequency band to be magnified. Because the motivation of research is to magnify motion to human vision [88].

Furthermore, a Hilbert phase-based motion estimation is proposed, in which, the phase variation and physical motion is connected based on the Hilbert transform [89]. In addition, the original video is decomposed into mono-component signals by using the Butterworth band-pass filter and peak-picking method. In simulated video, the identified correlation coefficient can reach 99.55%. A simply supported beam is tested for experimental validation to make comparison with conventional phase-based method, the Hilbert phase-based motion estimation showed an obvious superiority.

#### 1.2.2.3 Gradient-based optical flow

As a universal algorithm, the gradient-based optical flow and its variants are also widely used in image-based modal analysis [90], [91]. Besides a typical one LK-optical flow[68], the simplified gradient-based optical flow is developed recently [27], which takes a subset or a pixel as an independent sensor and linearizes the relation between displacement and intensity. However, compared with conventional optical flow method, the accuracy improvement of this approach is not quite noticeable. Afterward, a hybrid identification method that combines a high-speed camera with accelerometers is developed [26], nevertheless, the mode shape extracted is noisy and not reliable.

### 1.2.3 Evaluation of window-based image registration (WIR)

All above approaches can be summarized into a window (a local kernel) based image registration method[61], because all that take advantage of a neighbourhood operation, which could either be a subset (window/patch) in DIC or gradient-based optical flow, or be a local kernel for convolution with image in phase-based methods[61].

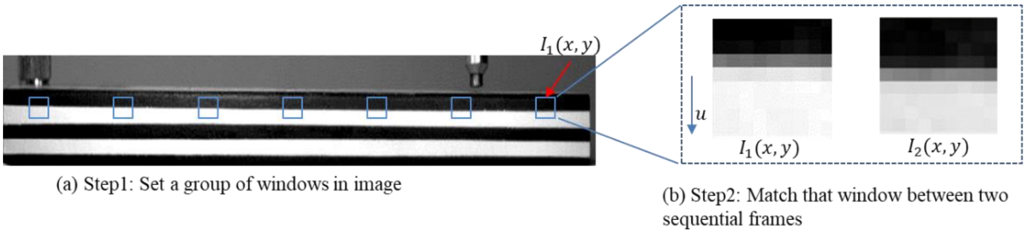



Figure 1-9 Schematic of window-based image registration, its two steps are (a) and (b) respectively

Figure 1-9 illustrate the basic procedures of window-based image registration, which first selects a set of windows and then match them individually between the adjacent frames. The difference of three algorithms is the norms used for registration as listed in Table 1-1.

The inherent difficulty of these methods is the determination of window size. If the window size is too large, the displacement inside a window is aliased by averaging. Instead, if the window size is too small, the obtained result will be very noisy. In addition, the estimation of displacement in each window is completely irrelative, which ignores the high correlation between responses of a dynamic system [92].

Table 1-1 Different norms used in three common window-based image registration

| Algorithm | Norm of image registration |
|---|---|
| Digital image correlation (DIC) | $\max C(u,v) = \sum_x \sum_y (I_1(x+u, y+v) I_2(x,y))$ [77] |
| Phase-based Motion magnification | $A_\theta(x,y)e^{j\phi_i(x,y)} = (G_2^\theta + \mathrm{j}H_2^\theta) \otimes I_i(x,y)$ [83] <br> $\min E(u,v) = \|\phi_1(x+u, y+v) - \phi_2(x,y)\|^2$ |
| Gradient-based optical flow | $\min E(u,v) = \|I_1(x+u, y+v) - I_2(x,y)\|^2$ [68] |

In this rigid framework based on image processing, the bottleneck of measure accuracy is difficult to be broken through. In this sense, this research project devotes to introduce a novel framework based on array signal processing. Because the vibration is a distinctive motion, for which, motion on all spatial elements is correlated, that is a special situation cannot be considered, during computing in an image registration method.

## 1.3 Research motivation

Based on analysing pros and cons of window-based image registration methods, this Ph.D. project proposes a novel framework to improve accuracy, and computational efficiency of vision-based vibration measurements (VVM), so that can pave a way for broadening the applications of using high-speed camera for full-field vibration measurements.

As the defects of window-based image registration is evaluated in last section, the desired framework should be designed:

- To utilize all characteristic pixels to extract displacement rather than isolated windows.
- To enhance SNR of desired signal to a detectable degree, especially in high frequency range.
- To obtain dense spatial estimation of mode shape on pixel level, instead of results from sparse windows.
- To improve the computational efficiency of algorithms.

To achieve these goals, this study conducts a cross-disciplinary research, in which the methodology of array signal processing (ASP) is originally introduced into VVM based on the similarity between proposed pixel array and other sensor arrays in respect to the basic underlying principle. In the proposed framework, a signal pixel or a small subset serve as a displacement sensor [38], and then a set of pixels in image is taken as a sensor array to improve signal quality. It should be noted that here the sensor specifies a unit measuring displacement. The history and principle of ASP is detailed in next section, and chapter 2 demonstrates how pixel array framework is developed by an analogy with other ASP applications.

## 1.4 Background of array signal processing

As an important area in the field of signal processing, array signal processing (ASP) focus on processing spatial signals sampled by a sensor array. The history of array signal processing dates back to World War II's antennas array. Nowadays, it has found wide applications to most sensor arrays, such as radar applications [93], wireless communications [94], acoustics [95] and so on [96]. The same underlying principle of ASP can govern different kinds of wave fields, which commonly includes: 1) Acoustic waves including sound waves, 2) electromagnetic waves including microwave and light, 3) mechanical waves in solids including vibrations.

### 1.4.1 Applications of array signal processing

In this section, the microphone array antenna array are taken to instantiate applications of ASP on acoustic wavefield and microwave wavefield.

#### 1.4.1.1 Antenna array

An antenna array is a set of individual antennas to transmit or receive radio signals jointly. The antenna array (sometimes called phased array) often takes advantage of the specific phase relationship to enhance the power of radio wave in a desired direction, this gain of signal is also known as directivity [97]. For example, radar system utilize electronic scanning arrays, in which a digital phase shifter controlled by a computer can steer the direction without moving the antennas, as shown in Figure 1-10.

The development of phased array antennas was originally started from radar systems for military use, however, that has found uses in many new areas that require directional antennas, like Bluetooth, wireless communication, satellite communications, and wireless local area network (WLAN). It can be expected in future that the scope of applications will boom in such as vehicle radar, spanning IoT, and even wearable electronics, since these antenna arrays become smaller and more intelligent [98]–[101].

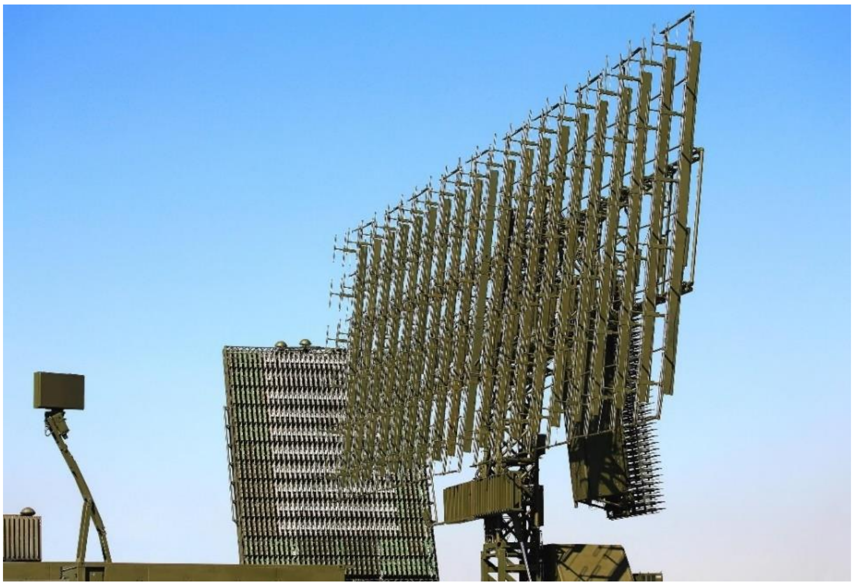


Figure 1-10 Phased array antennas of radar system [99]

However, from perspective of principle behind these applications, the methods of antenna array signal processing (AASP) like spatial filtering works in a similar way, and even the objectives are basically similar, which are to enhance signal from direction of interest, or to find the direction of signal source.

#### 1.4.1.2 Microphone array

A microphone array is a set of acoustic transducers arranged in a particular geometry. The applications includes systems for enhance voice, surround sound technologies, acoustic source localization, robotic navigation and so forth [95], [102]–[106]. Accordingly, the functionalities of microphone array signal processing (MASP) can be mainly classified as the sound enhancement and source localization. An application of sound source localization in condition monitoring is acoustic camera (or acoustic imaging) as shown in Figure 1-11. The acoustic camera consists of 64 microphones arranged as spiral, which can localize the sound source within a small range as shown by color contour. The location is derived by applying the adaptive spatial filtering on 64 channels of acoustic signals.

Figure 1-11 An acoustic camera is used to conduct condition monitoring of a compressor [107]

Actually, a large number of methods for MASP are borrowed and generalized from the method for AASP. However, the issues in MASP cannot be well tackled based on the borrowed AASP methods. Because there are still many distinctions between the acoustic signal collected by microphone array and the radio wave received by antenna. For example, the microphone array usually processes a broadband signal, or the environmental noise and the signals are highly non-stationary.

### 1.4.2 Conceptual framework of array signal processing

Although various applications are found, many textbooks and theories focus on the common fundamental principle without a reference to any particular application [96], [108], [109]. Because the mathematic framework of ASP is the same across different applications. In theory, the array signal model can be generalized as following expression [108],

$$\boldsymbol{x}(t) = \boldsymbol{A}(\theta)\boldsymbol{s}(t) + \boldsymbol{\eta}(t) \tag{1-1}$$

$\boldsymbol{x}(t)$ represents the snapshot data received by array elements [108], which denotes signals emitted from several signal sources. $\boldsymbol{A}(\theta)$ is the array manifold matrix determined by the direction-of-arrival (DOA) $\theta$ of signal source, and $\boldsymbol{\eta}(t)$ represents the noise[110].

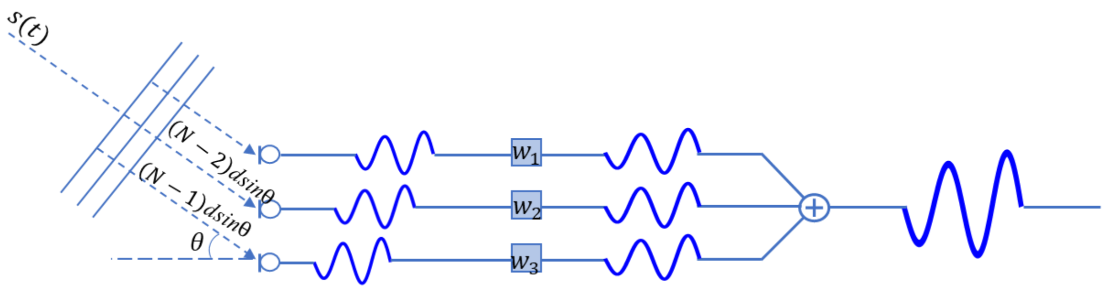



Figure 1-12 A schematic of that microphone array enhances acoustic signals based on spatial filtering

As signal model shown in Figure 1-12, received signals of one signal source $\boldsymbol{s}(t)$ have different phase, which can be attributed to that $\boldsymbol{s}(t)$'s arrival time on each sensor is different. There are two major functionalities by processing this model, first is estimation of spatial parameters of wavefield and the second is enhancement of some signals of interest [103], [110]. The spatial filtering (beamforming) is the central methodology to achieve these goals, which is a processor to operate weighted sum of array elements [111],

$$\boldsymbol{y}(t) = \boldsymbol{w}^H \boldsymbol{x}(t) \tag{1-2}$$

in which $\boldsymbol{w}^H$ is the conjugate transpose of weight vector, $\boldsymbol{y}(t)$ is the filtered signal. Overall, there are two major categories data-independent or adaptive spatial filtering, that is classified depending on whether the statistics of signal are considered in forming the spatial filters [110].

Figure 1-12 shows an example of spatial filtering on microphone array. The delay of arrival time is a parameter relating to the angle $\theta$ (direction) of signal source [112]. The weight vector is designed to shift phase of signals to compensate temporal delay from a selective direction, also known as a 'delay and sum array'[113]. it is to achieve enhancement of desired signal from selective direction and suppression of noise and interference[102], [114], [115].

The adaptive spatial filtering is also widely used to estimate the parameters [116], [117]. In addition, another important parameter estimation is subspace of eigenstruture method, such as the MUSIC (MUltiple SIgnal Classification) and ESPRIT (estimation of signal parameters via rotational invariant techniques)[108], [118].

## 1.5 Research aim and objectives

In order to meet the increasing demand for precision and applicability of full-field vibration measurement in modal analysis and condition monitoring, this PhD research aims to develop a novel framework based on the principle of array signal processing that is targeted at extracting weak vibration signal more accurately and robustly from high-speed camera footages, which can fulfil the detection of diagnostic feature (or identify modal parameters).

To achieve this research aim, the main objectives are identified and prioritized as follows:

- To set up a novel framework that can incorporate from measuring vibration to extracting features (or identifying parameters), which can accomplish the vision-based modal analysis and vision-based condition monitoring.
- To develop algorithms that can enhance SNR of vibration signal, especially for weak signal in high frequency.
- To develop algorithms that can achieve accurate and dense estimation of mode shape.
- To carry out simulation study by programming and using commercial software, in order to verify theoretical calculation and guide experiments.
- To design and implement a set of experiments, in order to validate proposed methods and expand applications of proposed methods.

## 1.6 Position of proposed method in research field

To clarify the status of proposed method, a schematic is drawn to illustrate where the proposed method is positioned at the research field, as shown in Figure 1-13. Firstly, the objective of research is clearly the same to conventional frameworks, which is to estimate weak vibration by using a high-speed camera. The conventional frameworks can be tracked back to the topic of motion estimation in the discipline of computer vision.

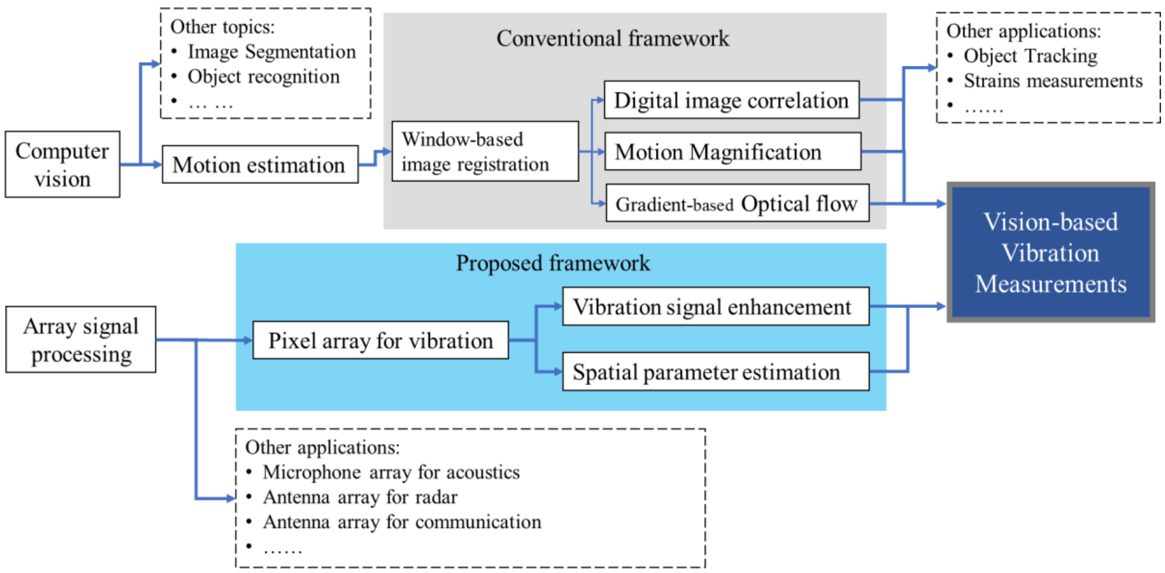



Figure 1-13 A schematic of where the position of proposed method is in research field.

Nevertheless, the proposed framework exploits the principle from a completely different discipline, i.e., array signal processing. That depends on an insight on the inherent similarity between pixel array and other sensor arrays, even though the camera is used in vibration measurements. This cross-discipline research drive the improvement of algorithm performance and has been verified by experiments, which is detailed throughout this thesis.

## 1.7 Thesis structure

This thesis is organised into nine chapters to present the work done to achieve the research aim and objectives. In Figure 1-14, the structure of this thesis is illustrated by proposed key equation and diagram, and the contents of each chapter are briefly listed below:

*Section 1 Introduction*

***Chapter 1*** – The first chapter presents outlook of this project and current status of this field.

*Section 2 Proposed Methodology*

***Chapter 2*** – This chapter describes overview of proposed methodology based on analysis of signal model.

***Chapter 3*** – This chapter describes methods for forming a pixel array from an image.

***Chapter 4*** – This chapter proposes the spatial filtering method to enhance signal.

***Chapter 5*** – This chapter details estimate of mode shape by using iterative and non-iterative methods.

*Section 3 Modelling and Simulation*

***Chapter 6*** – This chapter begins simulation study, which includes the finite element analysis and synthetic video analysis.

*Section 4 Experiments*

***Chapter 7*** – This chapter records two experiment of modal analysis on a free-free stainless beam and wind turbine blade.

***Chapter 8*** –This chapter conducts two experiment of condition monitoring on a reciprocating compressor and a gearbox.

*Section 5 Summary*

***Chapter 9*** – The conclusions, achievements and novelty of the research and suggestions for future work are presented in the last chapter.

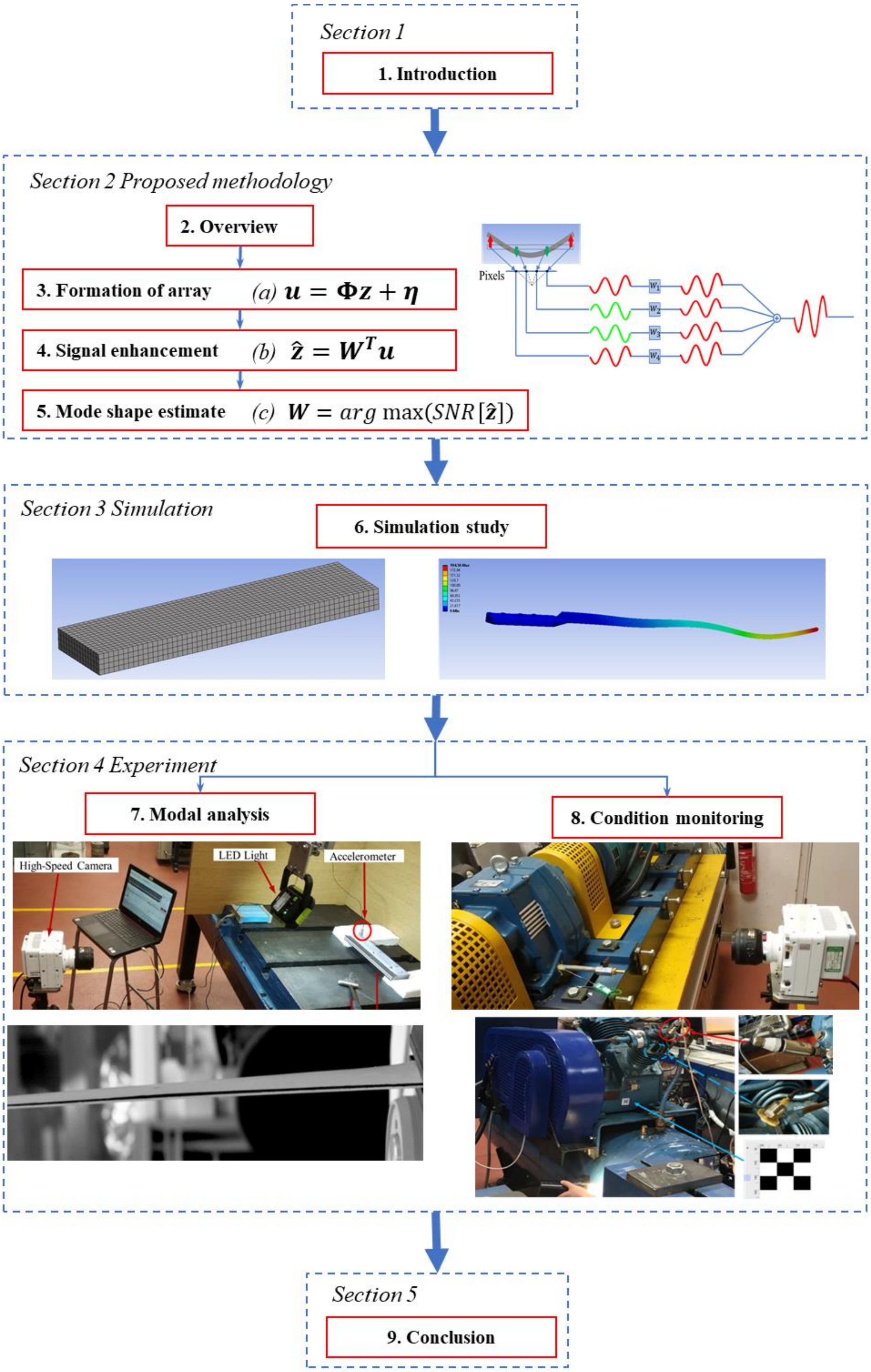



Figure 1-14 A schematic structure of the thesis, (a) is Eq. (2-1), which represents the model of the formed pixel array signal $\boldsymbol{u}$, (b) is Eq. (2-2), which denotes enhancement of modal signal $\boldsymbol{z}$ based on spatial filtering, (c) is Eq. (4-6), which denotes the estimate of mode shape based on the optimization of weight vectors $\boldsymbol{W}$.

In addition, the published materials are incorporated in thesis following the guidance of university. [1]-[5] in the list of publications are declared as my own written works, the details of reproduction are annotated in the introduction of each chapter and cited at that specific paragraph.

## 1.8 Key finding

This chapter illustrates the research status thoroughly by reviewing important literatures about this PhD project. Afterwards, the research aim and objectives are proposed aiming at the defects on existing methods.

# 2 OVERVIEW OF PIXEL ARRAY FRAMEWORK

*This chapter sketches out the key conceptions and workflow of proposed pixel array framework (PAF). The research dedicates to surpass accuracy of conventional vision-based vibration measurements by bringing in the principle of array signal processing. First, two main targets in ASP (signal enhancement and spatial parameter estimation) is analysed from perspective of pixel array. Subsequently, to demonstrate the position of PAF in research field, an analogy analysis is carried out between pixel array and microphone array. Furthermore, three steps of workflow are outlined, which are corresponding to the content of Chapter 3 to Chapter 5 respectively. Finally, an overall comparison with conventional methods is drawn to illustrate the deep reason why the PAF can achieve these advantages.*

*It should be noted that some paragraphs in this chapter are expansion or reproduction of the published materials* [1] in list of publications*, such as section 2.1.1 and 2.2, and the specific paragraphs are annotated by a citation.*

**Highlight:**

- The novel framework based on array signal processing is proposed.
- The signal model differs between pixel array and conventional array. The conventional array has phase shift generally but pixel array has amplitude variation.
- The proposed methodology specializes the algorithm of array signal processing to address particular issue in this research.

## 2.1 Conceptual framework of pixel array processing

### 2.1.1 Array signal model of pixel array

The most significant distinction between pixel array and other sensor arrays is that the signal amplitude changes over different pixels, instead of phase shift. Figure 2-1 shows a resonant object, the mode shape shows a form of standing wave [92]. Each pixel can capture a local element of this standing wave, all elements oscillating in phase, but their amplitudes are different. That implies that the scale and direction (positive or negative) of collected signals are different. The signal model should obey the principle of modal superposition [9], [119],

$$\boldsymbol{u}(t) = \boldsymbol{\Phi}(x)\boldsymbol{z}(t) + \boldsymbol{\eta}\,(t) \tag{2-1}$$

in which, $\boldsymbol{u}(t) \in \mathbb{R}^m$ is the output of pixels that collects the responses at dense multiple points, $\boldsymbol{z}(t) = [z_1(t), z_2(t) \dots z_n(t)]^T \in \mathbb{R}^n \boldsymbol{z}(t)$ denotes a group of modal signals at $n$ orders of natural frequencies [111], as shown in Figure 2-1, the array structure matrix $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_n] \in \mathbb{R}^{m \times n}$ denotes mode shape with $n$ orders, i.e., the amplitude of modal signal on that spatial point, and finally, $\boldsymbol{\eta}(t) \in \mathbb{R}^m$ is the random noise components.

Besides the resonant object, Eq. (1-1) can also conceptualize any other vibrating objects. The conception of $\boldsymbol{\Phi}$ can be broadened to be the spatial element's amplitude of its corresponding signal $\boldsymbol{z}_i$, $\boldsymbol{z}_i$ can be a wideband signal, not merely a mode. For the case that $\boldsymbol{\Phi}$ is not a mode shape, $\boldsymbol{\phi}_i$ denotes amplitude of a travelling wave $\boldsymbol{z}_i$ in every spatial points.

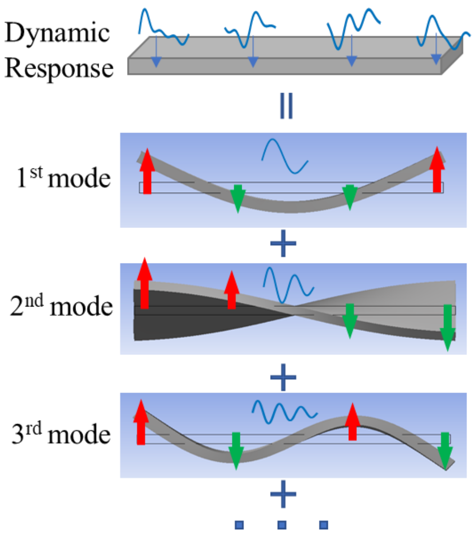



Figure 2-1 Schematics of modal superposition of a free-free beam, the dynamic response are superposed by several modes.

This array signal model demonstrates the signals received by pixel array contains the same wave $\boldsymbol{z}(t)$ and $\boldsymbol{\Phi}(x)$ determines different amplitudes of spatial elements.

### 2.1.2 Vibration signal enhancement based on spatial filtering

Enhancing the vibration signal to be detected based on spatial filtering is first goal of this research. According to the definition of spatial filtering (beamforming) [111], it means a weighted combination of array signals, which is illustrated as Figure 2-2,

$$\hat{\boldsymbol{z}}(t) = \boldsymbol{W}^T \boldsymbol{u}(t) \tag{2-2}$$

in this pixel array case, $\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n] \in \mathbb{R}^{m \times n}$ is the weight vectors. As $\boldsymbol{z}(t)$ is desired signal to enhance, the output $\hat{\boldsymbol{z}}(t)$ of this filter is designed to be an estimator of $\boldsymbol{z}(t)$, because the goal of spatial filter is to enhance modal signal $\boldsymbol{z}(t)$.

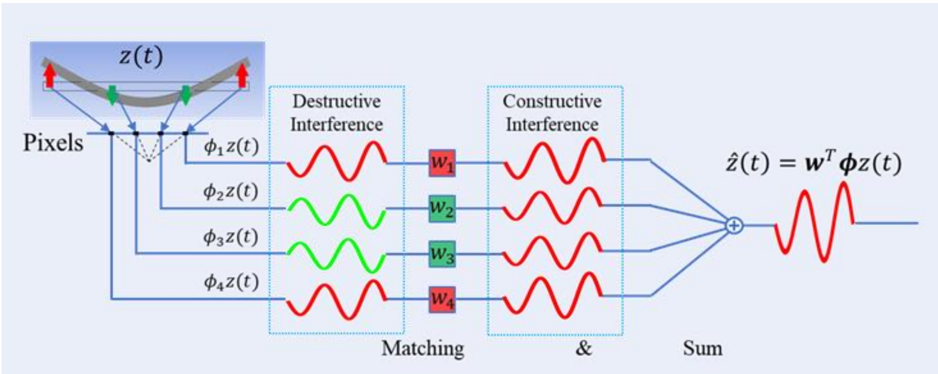



Figure 2-2 A schematic of that pixel array enhances vibration signal

The proposed pixel array is very similar to that of other sensor arrays like microphone array as shown in Figure 1-12 but with difference in array weight formulation. Because the difference in array signal model, there is no phase shift, but amplitude shift. Comparatively, as shown in Figure 2-2, the weight vector for the pixel sensor array is designed to change amplitude of signals along the mode shape, which can be described as a 'matching and sum array'[92]. Intuitively, the weight vector is similar to flip 'green waves' to 'red waves' to unify the direction of vibration before sum in Figure 1-12.

### 2.1.3 Mode shape estimation

Estimating mode shape $\boldsymbol{\Phi}(x)$ is another goal in vision-based vibration measurements, which should be incorporated into the issue of parameter estimation in ASP. According to the property of vibration, $\boldsymbol{\Phi}(x)$ can be either a stand wave in modal analysis or the propagation for a travelling wave in rotating machinery.

According to the theory of ASP, the adaptive spatial filtering, and subspace methods are two effective tool for parameter estimation [108], [118], which were used to estimate the direction of arrival $\boldsymbol{A}(\theta)$ successfully. In this study, aiming at the signal model Eq. (1-1), a novel ASF based on mode shape searching and a subspace method based on SVD are developed to estimate the mode shape.

### 2.1.4 Analogy analysis between pixel array and other sensor arrays

The analogy analysis between the pixel array and other arrays are laid in Table 2-1.

Table 2-1 Comparison between the pixel array processing and conventional array processing

| | **Conventional array processing** | **Pixel array processing** |
|---|---|---|
| **Signal model** | $\boldsymbol{x} = A(\theta)\boldsymbol{s} + \boldsymbol{\eta}$ | $\boldsymbol{u} = \boldsymbol{\Phi}(x)\boldsymbol{z} + \boldsymbol{\eta}$ |
| Array signal | Sensors' output $\boldsymbol{x}$ | Pixel's output $\boldsymbol{u}$ |
| Spatial matrix | Phase shift $\boldsymbol{A}$ of DOA $\theta$ | Mode shape $\boldsymbol{\Phi}$ along pixel coordinate $x$ |
| Signal of interest | Signal transmitted from a source $\boldsymbol{s}$ | Modal displacement $\boldsymbol{z}$ |
| **Methodology** | $\boldsymbol{y} = \boldsymbol{W}^T\boldsymbol{x}$ | $\hat{\boldsymbol{z}} = \boldsymbol{W}^T\boldsymbol{u}$ |
| Spatial filtering | Linear combination of array signals | Linear combination of array signals |
| Parameter estimation | Estimate Direction of arrival $\theta$ | Estimate mode shape $\boldsymbol{\Phi}$ |
| **Schematic** | Figure 1-12 | Figure 2-2 |

In terms of two array signal models, they both can be expressed by using a linear matrix multiplication. While the difference exists in the physical meaning of equation, in which the spatial matrix can be either the direction of arrival $\boldsymbol{A}(\theta)$ or mode shape $\boldsymbol{\Phi}(x)$. However, in essence, the basic underlying principle is the propagating wavefield identically, regardless of the signals are collected from whether an acoustic wavefield $\boldsymbol{x}(t)$ or a vibration wavefield $\boldsymbol{u}(t)$ on solid. This same fundamental principle makes the application of ASP can cut across different disciplines.

As the example of microphone array shown in Figure 1-12, the weight vector is designed to shift phase of signals to compensate temporal delay from a selective direction, also known as a 'delay and sum array' [102], [114], [115] [113]. Comparatively, the pixel array is known as a 'matching and sum array', they all can achieve enhancement of desired signal and suppression of noise.

## 2.2 Workflow of pixel array framework

As the conceptions of array signal processing is introduced, this research proposes a novel framework for vision-based vibration measurements [92]. The proposed framework mainly includes three key modules: pixel sensor array formation, array signal processing, and the feature extraction for modal identification or condition monitoring, which is illustrated in Figure 2-3.

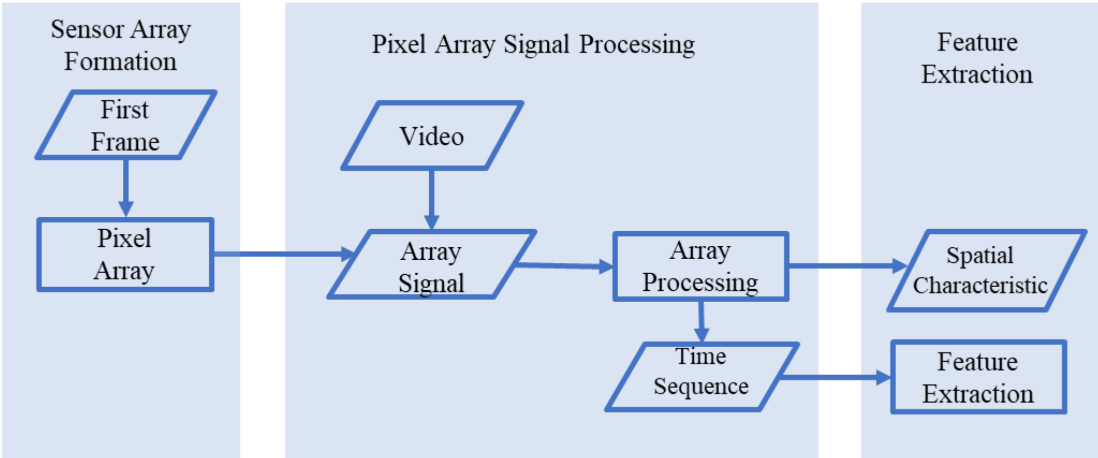



Figure 2-3 Schematic of the proposed framework, which is comprised of array formation, array processing and feature extraction

First, to form a sensor array by selecting pixels sensitive to displacement in a reference image [111]. By taking each pixel or a subset as a displacement sensor, the performance as criterion for these sensors.

Second, to process the signal matrix collected by array. One of proposed methods is an adaptive filter to match mode shape and enhance modal displacement. Based on proposed array signal model, as illustrated in Figure 2-4 (a), the oscillating intensity captured by pixels at different space elements are combined to enhance each modal displacement. According to modal superposition, each dynamic response can be decomposed into several modes [111], as shown in Figure 2-4 (b). In each mode, all space elements are oscillating with the same frequency and in phase. As shown in Figure 2-4 (b), its mode shape $\boldsymbol{\Phi}$ displays a form of standing wave, in which the motion direction is opposite across nodes, as denoted by the red and green arrow, respectively. The weight vectors $\boldsymbol{W}$ are thus designed for this structure, which is set to match the mode shape, so that the constructive waves will be generated.

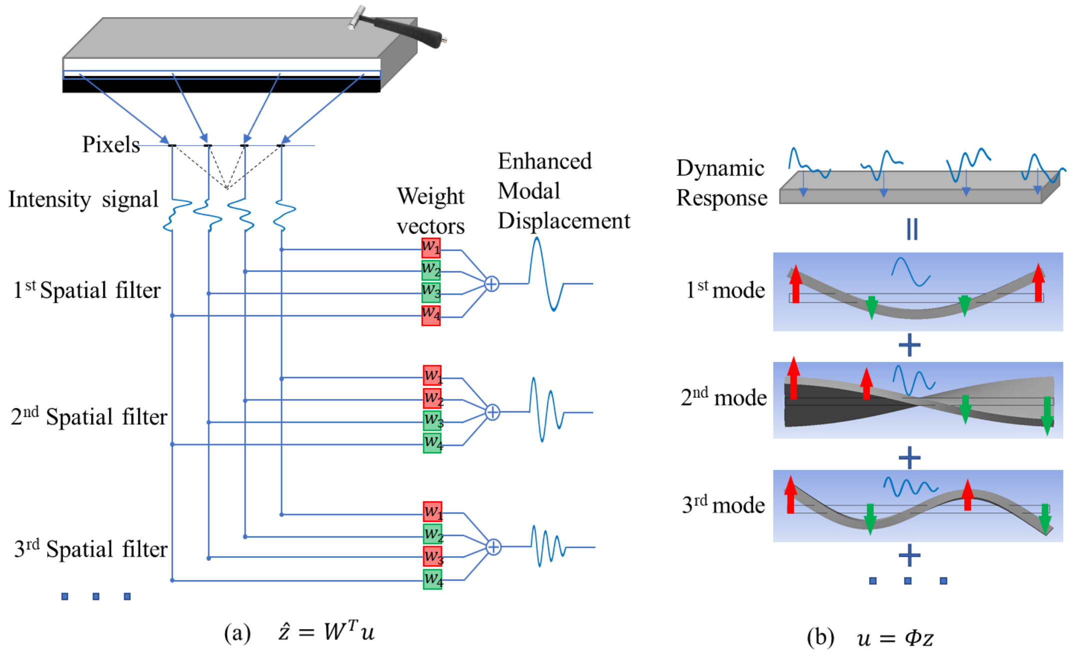



Figure 2-4 Schematic diagram of the proposed method, (a) Illustration of using a pixel sensor array to enhance modal displacement signal and (b) the principle of modal superposition [111]

Third, to extract feature from proposed time sequence and spatial characteristics for purpose of condition monitoring or modal analysis. In modal analysis, that is modal identification, in which a frequency domain method, the Least-squares rational function (LSRF) is used to identify modal parameters including the natural frequency and damping ratio. Alternatively, the feature extraction method for diagnosis such as time-synchronous signal average is used for condition monitoring as one instance.

## 2.3 Contrast with window-based image registration

This section compares the theory and the properties of pixel array framework (PAF) with the window-based image registration, as shown in Table 2-2.

Table 2-2 Comparison between the spatial filtering and window-based methods

| | **Pixel array framework** | **Window-based image registration** |
|---|---|---|
| **Range of application** | Mechanical vibration | No constraint of motion form |
| **Principle** | Array signal processing | Computer vision |
| **Displacement sensors** | One pixel | One window (subset) |
| **Operation on sensors** | To sum them up coherently | To use them incoherently |

It can be seen in Table 2-2, that the prime difference between the two methods is the principles behind them. The local operator used in window-based methods is versatile for different forms of motion such as measuring strain or deformation. However, in modal analysis, all space elements oscillate around an equilibrium point at the same frequencies. Aiming at this distinctive motion, the PAF adds the correlated signals coherently to extract and enhance vibration information.

## 2.4 Key findings

This chapter overlooks overall framework of this research, which is detailed in following three chapters, i.e., formation of sensor array, signal enhancement and mode shape estimate. According to above analysis, the performance of array processing is highly potential to surpass image registration in theory. Therefore, the project aims to propose a novel approach by specializing algorithms of array signal processing to address particular issue in vision-based vibration measurement.

# 3 PIXEL ARRAY FORMATION BASED ON EDGE DETECTION

*Generally, the structure of conventional sensor arrays is determined at the stage of hardware design. However, the proposed pixel array is distinctive because it needs determining the array based on the particular captured image. In formation of this array structure, only pixels located at the edge are selected, because these characteristic pixels are sensitive to estimate displacement. Therefore, an edge-detection-based formation approach is proposed for pixel array framework in this chapter.*

*First, in order to pick out the pixels that are sensitive to vibration signal, the transfer function between pixel intensity and displacement is investigated, which is influenced by the point spread function (PSF). Second, the influence of displacement amplitudes (motion is small or large) on characteristic of pixel-equivalent sensor is analysed. Therefore, the Euler method and Lagrange method are developed for large and small motion. Meanwhile, the obtained signal quality is evaluated based on SINAD.*

*It should be noted that some paragraphs in this chapter are expansion or reproduction of the published materials* [1] and [2] in list of publications, *such as section 3.1 and 3.2, and the specific paragraphs are annotated by a citation.*

**Highlight:**

- A transfer function is used to indicate the relationship between pixel intensity and displacement.
- Point spread function can blur the image and non-linearize transfer function.
- Euler Description is applicable to small displacement within the scope of gradient area.
- Lagrange Description is applicable to large displacement over the gradient vanishing area.

## 3.1 Equivalent displacement sensor by pixels

As sensors are fundamental elements for collecting signals in of ASP, to determine the characteristics of sensor and array structure (geometry of sensor locations) are necessary before the ASP. For the proposed pixel array, not all pixels of image are effective to measure displacement. For example, the motion have no impact on the intensity of a pure color region without obvious texture. Because those pixels' intensity are equal to other neighbour pixels i.e., the image gradient is zero. In addition, it is important for the equivalent sensor (pixel) to figure out its transfer function between displacement (input) and pixel intensity (output) and the sensor characteristics in the formation of pixel array [92].

### 3.1.1 Point spread function

The basic principle of vision-based vibration measurement is to estimate displacement based on the change of pixel intensity. The effect of displacement to intensity is often simplified as a linear relationship [38]. However, this assumption cannot be strictly satisfied, due to the effect of image blurring[61]. Moreover, the displacement is a pixel-level oscillation in vibration measurement, so the non-linearity cannot be neglected relative to the micro-motion scale indiscreetly [92] .

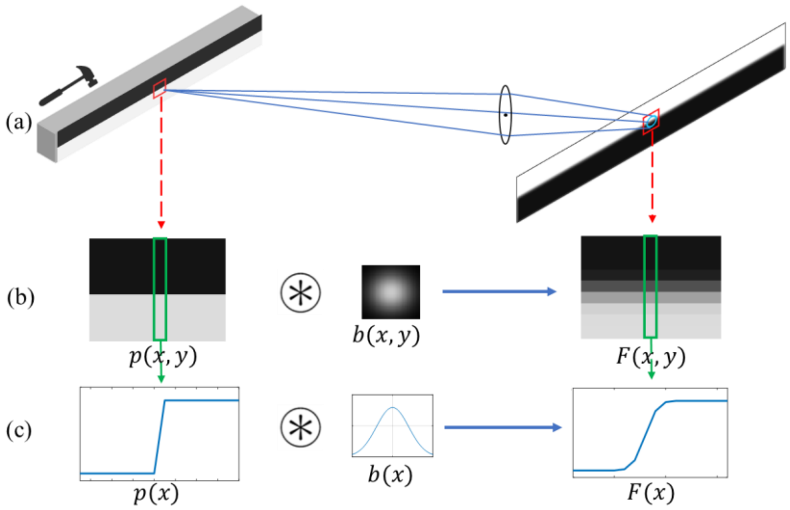



Figure 3-1 (a) The schematic diagram of experiment (b) Illustration of image blurring process and (c) One-dimensional representation of the blurring process

Figure 3-1 shows the process of the blurring in respect of video shooting, image, and function expression respectively. To carry out the theoretical analysis of intensity-displacement relationship, an ideal image is synthesized in a numerical way, in order to simulate a sharp edge on experimental object. In experiment, a sharp edge is often created by painting black-white stripes onto the surface of target object to improve contrast of image. The region shown in Figure 3-1 (b) is the simplified image of experimental object, which only contains a bright and dark area without transitional region, and for simplification, the edge between two areas is set to be horizontal exactly. In this case, each column has the same value, as one column shown in Figure 3-1 (c). The advantage of this simplification is that the intensity-displacement relationship can be investigated under a single coordinate first, before being expanded the conclusion to two-dimension case.

When the pattern edge of object projects on image, the sharp edge will be blurred into a smooth curve due to the defocus aberration [61]. The point spread function (PSF) is often used to describe this blurring effect, the blurred image can be represented by convolution of the predicted sharp image and PSF:

$$F(x,y) = p(x,y) * b(x,y) \tag{3-1}$$

in which, $p(x,y)$ is a sharp edge described by step function, $b(x,y)$ is the PSF kernel, symbol of $*$ represents a convolution operation and $F(x,y)$ is the blurred image [68]. In engineering field, PSF is approximated typically by using a Gaussian profile [120].

To attentively investigate the transfer function of displacement and intensity, one row is picked out of an image to build a one-dimension function $F(x)$. Figure 3-1 (c) illustrates the one-dimensional representation of convolution: $F(x) = p(x) * b(x)$ . The PSF is approximated by the Airy disk, that results in $F(x)$ forming a cumulative distribution function (CDF) of Gaussian distribution. This profile causes the nonlinearity between intensity and displacement.

### 3.1.2 Transfer function between intensity and displacement

When the displacement $u$ needs to be solved in a video footage, at least two frames $F(x)$ and $I(x)$ need to take into consideration. According to the principle of optical flow [68], displacement $u$ is solved by minimizing measure of difference between the translated reference frame $F(x+u)$ and current frame $G(x)$. That is to say, translate the frame $F(x+u)$ to match current frame, that turns out the distance $u$ is:

$$G(x) \approx F(x+u) \tag{3-2}$$

Under pixel-level displacement, this equation can define the relationship between intensity of single pixel and its received displacement [121], where the current intensity $I(x)$ is viewed as output of displacement $u$ [38], its mapping between input $u$ and output $I(x)$ follows $F(x+u)$ . In optical flow methods, the relationship often simplified as,

$$F(x+u) = F(x) + F'(x)u + o(u) \tag{3-3}$$

in which, $F(x)$ and $F'(x)$ are intensity and gradient at pixel $x$ respectively, $o(u)$ represents a higher order infinitesimal of displacement $u$. Sometimes, a more accurate model contains the second derivative is used as,

$$F(x+u) = F(x) + F'(x)u + \frac{F''(x)}{2}u^2 + o(u^2) \tag{3-4}$$

Form perspective of a sensor, and the measured intensity can be taken as a transfer function $F(x+u)$ plus a measurement error. $F(x+u)$ in Eq. (3-4) takes the pixel coordinate $x$ as parameter. Furthermore, $F(u)$ can be expressed by using a cumulative normal distribution function to approximate the Airy disk as point spread function (PSF), that yields,

$$F(u, x_i) = \frac{R}{2}\mathrm{erf}\left(\frac{u + x_i - \mu}{\sqrt{2}\sigma}\right) + Z_o \tag{3-5}$$

In this expression, $erf$ is the Gaussian error function [122], which is the function of cumulative Gaussian distribution. Additionally, $R$ denotes the span of total range, $Z_o$ denotes zero offset at centre point, $x_i$ is the pixel coordinate. $\sigma$ is the variance in error function, which reflects the radius of Airy Disk. $\mu$ denotes the central points of cumulative distribution function (CDF) [123]. Then when the measure noise of camera is considered,

$$I(u, x_i) = F(u, x_i) + \eta \tag{3-6}$$

by which, as the output of transfer function, the intensity value $I(u, x_i)$ on pixel $x_i$ can be obtained for this one-dimension case. In Eq. (3-6), $\eta$ denotes the random image noise.

### 3.1.3 Characteristics of equivalent sensor

To check the omitted nonlinear term $o(u)$ in Eq. (3-3), it can be ignored in the conversion between measured signal (intensity) into wanted signal (displacement) generally. In this case, a linear sensitivity $S = f(x)$ is generally used to simplify the transfer function as illustrated in Figure 3-2.

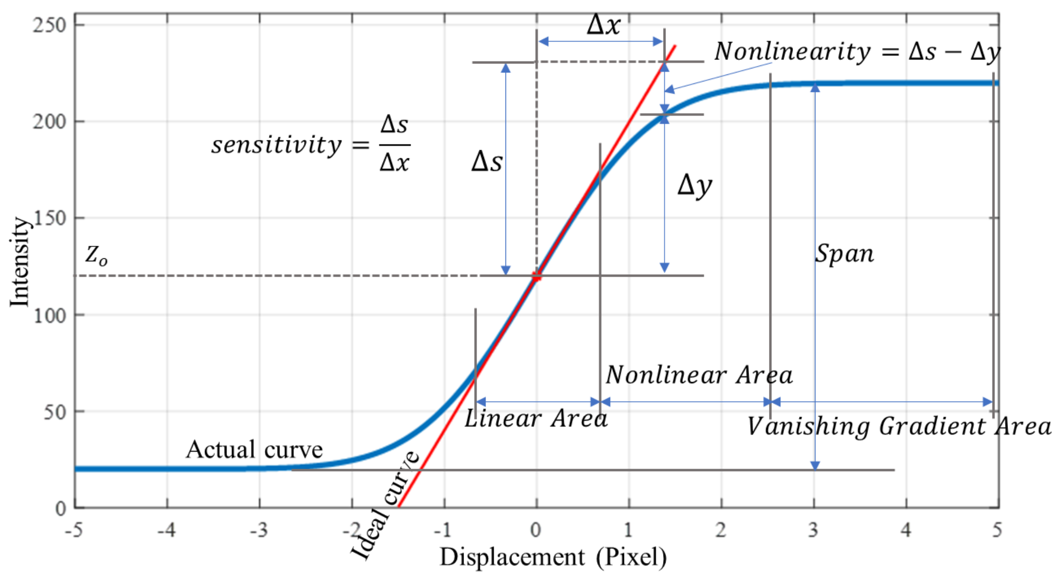



Figure 3-2 Schematic of transfer function of a single pixel, which is not accurate if the displacement is larger than the effective measure range.

In Figure 3-2, the blue one denotes the actual output-input curve of measured intensity, the red line denotes the linearized curve by using sensitivity $S$. The characteristics of this equivalent sensor are expatiated individually as follow.

- Sensitivity $S = f(x_i)$ is a linear transfer function shown as the slope of ideal curve in Figure 3-2. This will determine the strength of signal collected by a pixel $f(x_i)$.
- Nonlinearity $N_L = o(u)$ is deviation of an actual transfer function from the linearized transfer function, which indicate the distortion in collected signal. It can be expanded further into

$$N_L = \frac{f'(x_i)}{2}u^2 + \frac{f''(x_i)}{6}u^3 + \cdots \tag{3-7}$$

- Area division:
  - Linear area refers to distortionless area, in which the error between ideal curve and actual curve is small enough to ignore. It specifies the region with nonlinearity is less than 5%.
  - Nonlinear area refers to area with distortion is too large to ignore, even than linear components, but the gradient is not zero.
  - Vanishing gradient area refers to the region's gradient is zero. If pixel locates at this area, the intensity cannot respond to displacement, so that only noise can be measured.

- The sensor resolution $\Delta x_{min}$ or measurement resolution is the smallest change that can be detected, which is also called quantization. For example, if the color bit is 8, the resolution should be $\Delta x_{min} = 1/256$.
- Drift or zero offset $Z_o$ is defined as the output signal that slowly changes independent of the measured property. It is the initial intensity without displacement at each pixel $Z_o = F(x_i)$ at initial position $x_i$.
- Span of total range $R$ is distance of maximum and minimum measurable value, which depends on the contrast of this area.

The above characteristics reflects the performance of equivalent sensor, especially the sensitivity and non-linearity, because the sensitivity determines the strength of output signal, and the nonlinearity determines the degree of signal distortion, which will be investigated in following sections.

## 3.2 Evaluation of signal quality

According to above analysis, the collected signal consists of desired signal, noise, and distortion, all of which are associated with subset size. A comprehensive indicator is required to evaluate the performance of each pixel sensor. The signal-to-noise and distortion ratio (SINAD) is used as an index because it can involve signal, noise, and distortion together to judge the signal quality.

### 3.2.1 Distortion caused by nonlinearity

Due to the analysis of nonlinearity, its impact on error of signal is investigated here [92]. As one pixel is taken as a sensor measuring displacement, its sensitivity is different at different position. The blue curve shown in Figure 3-3 represents actual output-input curve Eq. (3-8) in a small region of interest (ROI), just like Figure 3-1 (b) and (c),

$$I_C(u; x) = F(x + u) \tag{3-8}$$

in which the noise is set as $\eta = 0$ here, as only is the effect of nonlinearity surveyed now.

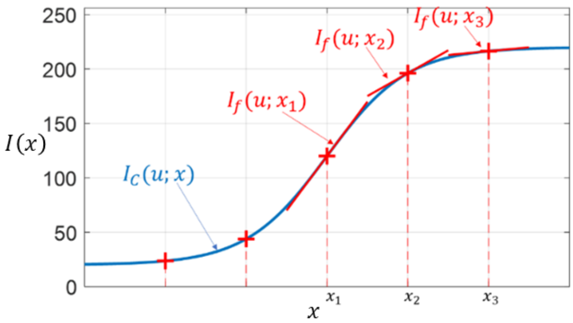



Figure 3-3 Sensor characteristics of pixels at different locations on one edge

The red straight lines are linearised transfer function of different pixels

$$I_f(u;x_i) = F(x_i) + F'(x_i) \times u \tag{3-9}$$

The nonlinearity can be therefore derived by:

$$e_I(u;x_i) = I_C(u;x) - I_f(u;x_i) \tag{3-10}$$

which are shown as the clearance between blue curve and red lines in Figure 3-3. $I_f(u;x_i)$ of three pixels has significant difference of the sensitivity and nonlinearity, that reflects the sensitivity and nonlinearity depending on pixel location $x$ in Figure 3-3.

Subsequently, the static input $u$ is replaced by a dynamic input $u(t)$ to investigate the distortion caused by nonlinear transfer function. In frequency domain, the signal distortion will produce a bunch of unwanted frequency components [124], [125], as shown in Figure 3-4(d), which can be classified as two types: harmonic distortion is the harmonics of rigid mode [61] and intermodulation distortion is sideband of normal mode [62]. Harmonic distortion is many high-order harmonics of fundamental and intermodulation harmonics is sideband of high frequency modes. These distortions will attenuate the original frequency components.

To simulate a response of hammer test, the displacement of a damping vibration signal is set as the input $u(t)$,

$$u(t) = A_1 e^{-r_1 t}\cos(2\pi f_1 t) + A_2 e^{-r_2 t}\cos(2\pi f_2 t) + \cdots \tag{3-11}$$

response $u(t)$can contain $n$ orders of mode theoretically, but here just a rigid mode at $f_1 =$ 3 Hz and a normal mode at $f_2 = 3$ Hz are set, because that is sufficient to reflect two kinds of distortion. Besides, $A_i$ denotes the amplitude of modal signal, $r_i$ denotes the damping ratio of that mode. Then to substitute Eq. (3-11) $u(t)$ into Eq. (3-8) and Eq. (3-5), that yields

$$I_c(t, x_i) = F(u(t), x_i) \tag{3-12}$$

According to Figure 3-4, the nonlinearity exists in the collected signals, Figure 3-4 illustrates the distortion that the nonlinearity imposed on original signal. Figure 3-4(a) represents a region of image including five pixels. It shifts left and right following a displacement $u(t)$ in Figure 3-4 (c). Figure 3-4 (b) shows intensity signals captured at location $x_1$, $x_2$ and $x_3$ with image noise. Additionally, waveforms show an obvious distortion due to the nonlinear profile, this phenomenon can also be found in experimental data.

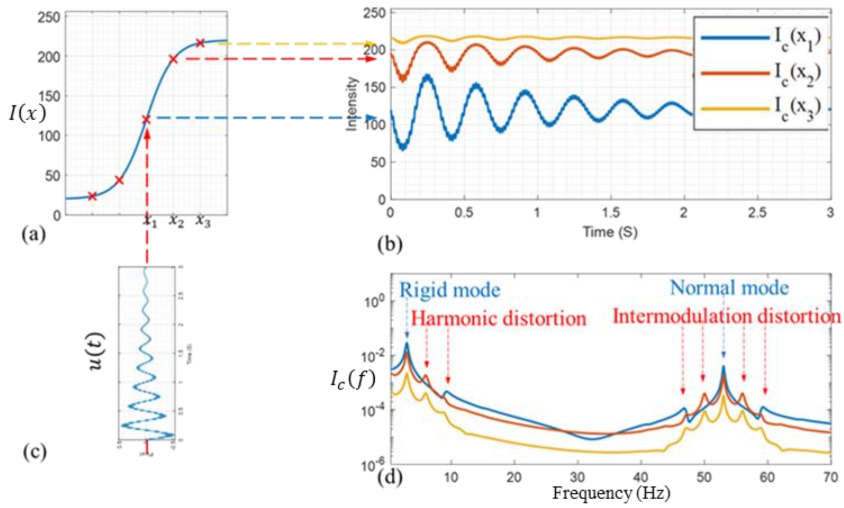



Figure 3-4 Intensity signal of different pixels captured under motion (a) ROI of image, (b) response: intensity signal, (c) excitation displacement signal and (d) spectrum of the real output signal $I_C(t)$

### 3.2.2 Image Noise

The image noise is determined by many factors comprehensively, which is described by a noise model, in which the quantization error is studied particularly [126]. Because the displacement is possible to result in change of intensity under one color-bit. It is potential to be a great influence of image noise.

The error of quantization can be either periodical or random, which is determined by characteristics of the input signal. When the input is a periodic signal, the quantization will produce a periodic error, which distributes at certain frequencies. In contrast, when the input signal contains random component, the quantization error will be random. The quantization error is uniformly distributed between −1/2 LSB and +1/2 LSB, hence it is related to the resolution $Q$, which is the color bits of pixel in this image case. The signal-to-quantization-noise-ratio (SQNR) is used to indicate the image noise,

$$SQNR \approx 1.761 + 6.02Q\ (dB) \tag{3-13}$$

where the input signal is a full-scale sinusoidal wave.

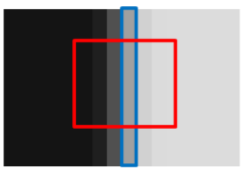
### 3.2.3 Signal-to-noise and distortion ratio

Based the above analysis, the components consist of the signal, distortion, and noise, in all of which, the signal is desired component, others are unwanted components [92]. To consider them comprehensively, the signal to noise and distortion (SINAD) is adopted to evaluate the performance of each pixel, because it allows to take all the signal strength, noise influences and nonlinear distortion into account [127]. That can be expressed as,

$$\boldsymbol{SINAD} = \frac{P_s}{P_d + P_n} \tag{3-14}$$

where the power of linear component $P_s = E[I_f(t)^2]$ represents the power of signal, the nonlinear error $P_d = E[e(t)^2]$ represents the power of distortion, and $P_n = E[\eta(t)^2]$ is noise power [128].

By calculating SINAD of $x_1, x_2, x_3$, it yields 86.5 2.15 0.57, respectively. Furthermore, it shows that trend of SINAD reaches the maximum at the centre of edge and gradually decreases to zero towards both sides. In this area, only a part of pixels is sensitive to displacement. For example, if the pixel has no pattern, its SINAD will be zero or very small. Obviously, this pixel has no contribution to measure displacement.



(a) $F(x,y)$

$$\begin{bmatrix} 0 & \dots & 0 & 0.83 & 2.66 & 86.5 & 2.15 & 0.57 & 0 & \dots & 0 \\ & \dots & & & & \dots & & & & \dots & \\ 0 & \dots & 0 & 0.83 & 2.66 & 86.5 & 2.15 & 0.57 & 0 & \dots & 0 \\ & \dots & & & & \dots & & & & \dots & \\ & \dots & & & & \dots & & & & \dots & \\ 0 & \dots & 0 & 0.83 & 2.66 & 86.5 & 2.15 & 0.57 & 0 & \dots & 0 \end{bmatrix}$$

(b) $Sinad(x,y)$

$$\begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & \dots & & \dots & & \dots & \\ & \dots & & \dots & & \dots & \\ & \dots & & \dots & & \dots & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{bmatrix}$$

(c) $A_s(x,y)$

Figure 3-5 Comparison of SINAD between the proposed method and window-based method, (a) pixels selected based on two kinds of methods on image, (b) SINAD of selected pixels and (c) array structure based on edge detection

Figure 3-5 compares the signal quality of selected by using SINAD with the window-based registration, Figure 3-5 (a) is the transposed region of synthetic image in Figure 3-1 (b), in that each row is equal to $I(x)$ in Figure 3-4 (a). The blue box represents the pixels selected by using SINAD, the red box represents a subset often used in window-based methods. The average SINAD of the red box is far smaller than that of the blue box because it contains many pixels with small or zero SINAD, these pixel collects poor signal even no signal components. Comparably, the pixels in blue box are selective about high performance. That makes difference of efficiency in signal acquisition between the proposed method and conventional methods. As shown in Figure 3-5 (b), the pixel at the centre of edge possesses the largest SINAD, which exhibits the quality of displacement signal obtained from each individual pixel is uneven.

In summary, the edge detection-based formation approach can get all sensitive pixels involved in computation, meanwhile exclude all useless pixel from computation to improve the signal quality in data acquisition. To select these pixels, the Sobel edge detector is applied on the reference image, which yields a binary matrix $A_s(x, y)$ as shown in Figure 3-5 (c). That denotes the spatial coordinates of pixel sensors in the array, known as the array structure of sensor array [100]. The pixels with value one in $A_s(x, y)$ should be selected as elements of the sensor array to record the intensity signals while other pixels will be removed.

## 3.3 Euler description vs Lagrange description

The Lagrangian and Eulerian description (specification) are terminology in fluid dynamics originally, which refers to two ways in observing the flow field. It is brought in to represent a similar methodology in computer vision [129], [130]. Specifically, the Eulerian description is a way of looking at fluid motion that focuses on specific locations. In contrast, the Lagrangian description is a way that the observer follows individual fluid particle.

In this research, this conception is brought in to differentiate manners of observing motion in image. Eulerian description refers to looking at displacement on one fixed pixel or subset, which is found to be suitable for small displacement. In contrast, Lagrangian description focuses on measuring position of one object, which is more applicable to large displacement.

## 3.4 Equivalent sensor based on Euler description

### 3.4.1 Transfer function of subset

In addition, the scale of displacement $u$ is another important factor. If $u$ is too large and get access to gradient vanishing area, displacement will become not detectable. To tackle this problem, a subset can play the role of a sensor in place of single pixel. Its sensitivity $S_{sup}$ can be therefore obtained as follow,

$$S_{sup} = \frac{1}{n}\sum_{i=1}^{n} S(x_i) = \frac{R}{\sigma n\sqrt{2\pi}}\sum_{i=1}^{n} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \tag{3-15}$$

which means sensitivity of subset. As shown in Figure 3-6 (b), the sensitivity decreases while the subset consists of more pixels, but the scope of linear area gets broad.

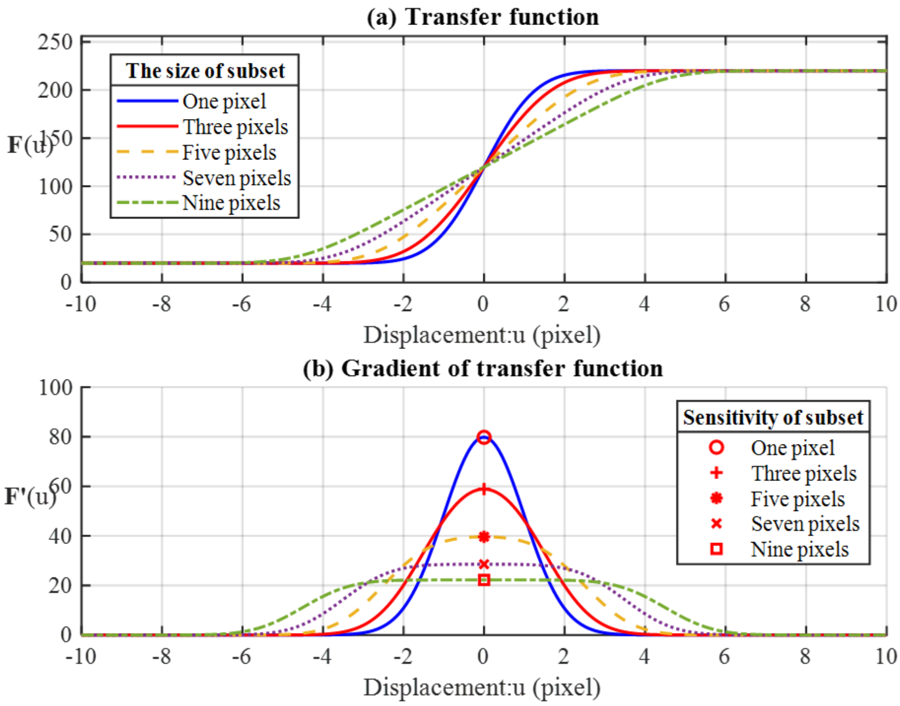



Figure 3-6 (a) Transfer function of subset with different size, the expand of linear area at the exchange of decreasing sensitivity, (b) The gradient of transfer function, which decide the marked sensitivities are

In a similar way, the nonlinearity of subset can be derived.

$$\begin{aligned} N_L &= \frac{R}{\sqrt{8\pi}\sigma n}\Big[\sum_{i=1}^{n} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}\Big]u^2 \\ &+ \frac{R}{\sqrt{72\pi}\sigma n}\Big[\sum_{i=1}^{n}((x_i-\mu)^2-1)e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}\Big]u^3 + \cdots \end{aligned} \tag{3-16}$$

This equation demonstrates the nonlinearity shrinks with increase of subset size, therewith the linear area transfer function is broadened as shown in Figure 3-6 (a). Moreover, the noise level will go down due to average, which will be discussed in following section. That means the sensitivity, nonlinearity, and noise level decrease with respect to subset size simultaneously. It turns out a trade-off exists in the signal power, distortion power and noise power, which will determine the subset size jointly.

### 3.4.2 SINAD under different displacement amplitude

However, the increasing displacement amplitude will influence SINAD, because the linear area is narrow. Once the displacement exceeds this area, the transfer function changes dramatically with gradient vanishing. To broaden the linear area, a subset can be used to be that equivalent sensor by combining several pixels aside the edge centre in place of an individual pixel. Figure 3-7 shows the relationship between amplitudes of exciting displacement and three components in the intensity response.

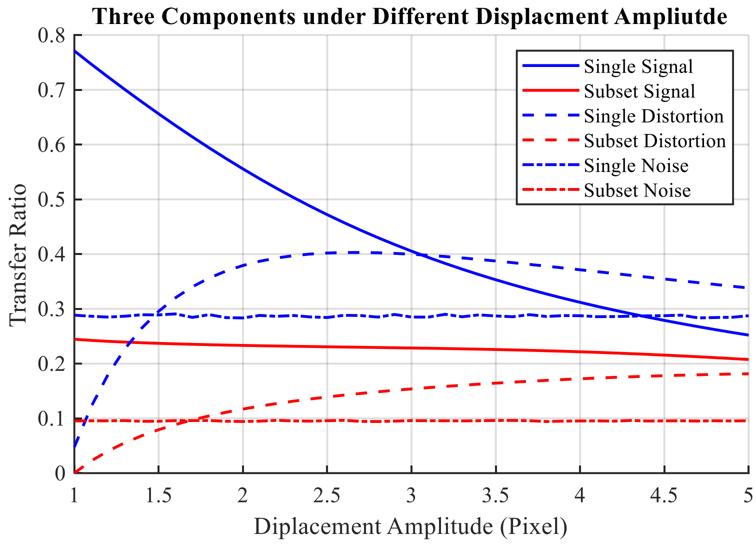



Figure 3-7 Displacement's influence on single pixel and subset, as shown, the single pixel is more sensitive to amplitude of input signal

The set of blue lines is from a single pixel, and the red lines are derived from a subset consisting of nine pixels, which are pixels in a line perpendicular to the edge as shown in Figure 3-7. For the single pixel, the distortion increases significantly due to its narrow linear area, and that reduce the signal components simultaneously. Instead, in subset case, the signal keeps strong, and the distortion is small relatively, because of its wider linear area. In addition, the noise floor is a value irrelated to amplitude of excitation signal.

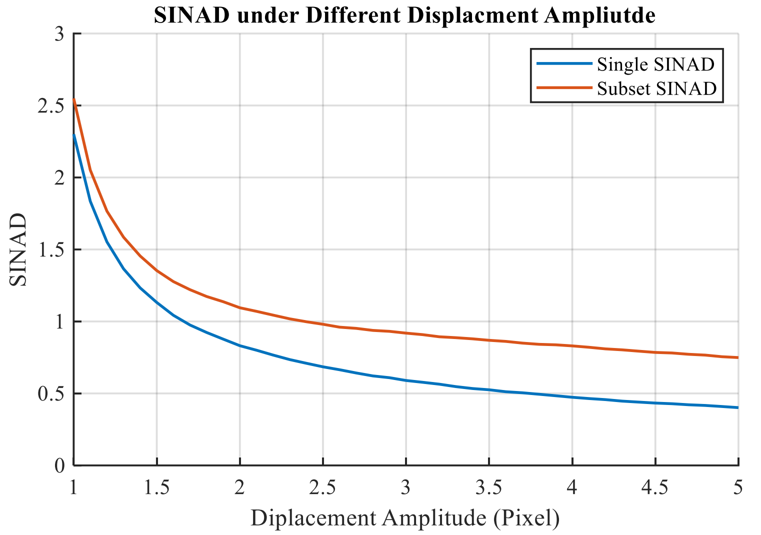



Figure 3-8 Displacement's influence on SINAD of single pixel and subset

It turns out that the SINAD of a subset of nine pixels is higher than a single pixel, because the displacement is larger than one pixel, which has been beyond the linear area of one pixel as shown in Figure 3-8.

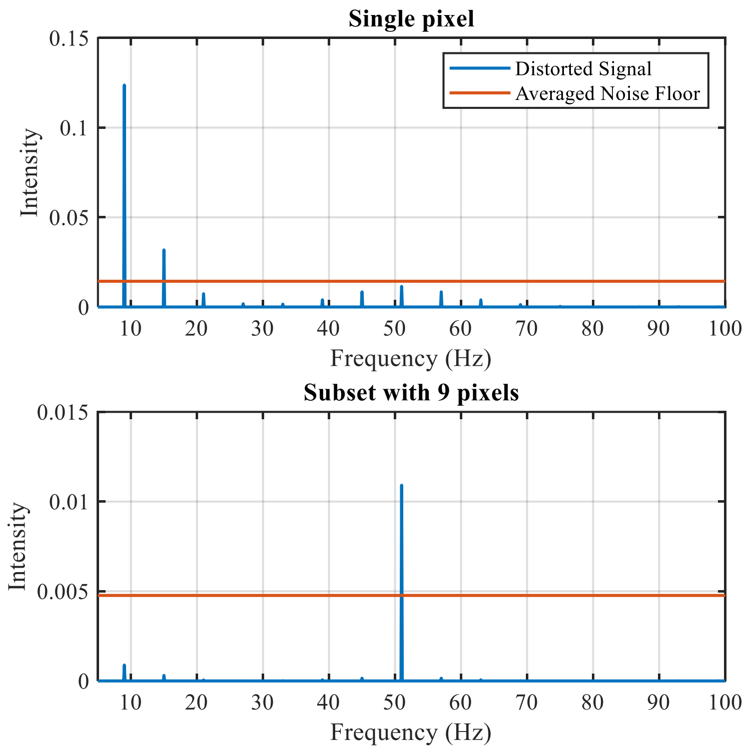



Figure 3-9 The spectra of signal from a single pixel and a subset, the single pixel is distorted severely (upper figure), in contrast, the subset keeps high fidelity (lower figure).

At this large displacement situation, the distorted signal can even be submerged in noise floor, as comparison that Figure 3-9 shows. For the single pixel, its natural frequency has been distorted into several sidebands. In contrast, the natural frequency keeps steady, because it can be taken as a large subset, so that the displacement cannot exceed the measure range of that.

### 3.4.3 SINAD and size of subset

However, what is the optimal size of subset, this section studies the relationship between subset size and SINAD. The subset refers to a one-dimension pixel array, which is aligned with perpendicular the edge. Figure 3-10 illustrates signal, distortion and noise varies with respect to subset size under a displacement with an amplitude of five pixels. Even through the signal, distortion and noise components are decreasing, their decay rates are different, that turns out to the SINAD reaches maximum when subset radius is equal to displacement amplitude, which is marked by a red circle.

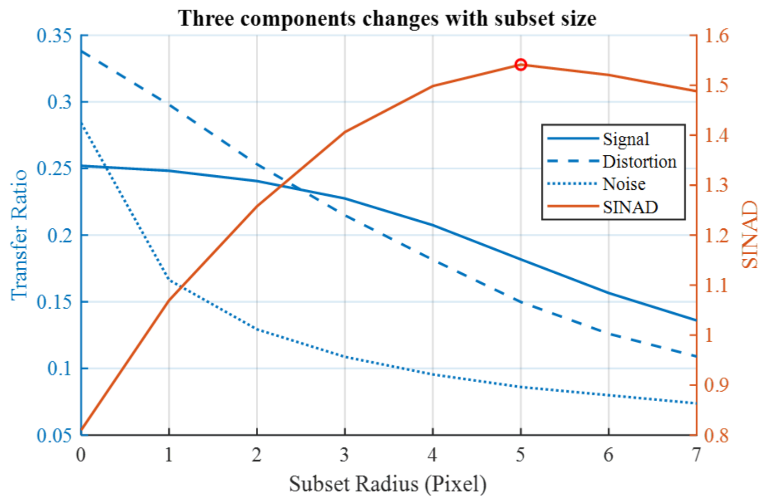



Figure 3-10 Subset size’s influence on SINAD, the red dot represents the max SINAD.

That shows a correlation between subset size and displacement amplitudes, i.e., the subset should be able to cover the displacement. As the displacement is being increased from one pixel to eight pixels, the result supports this correlation holds true.

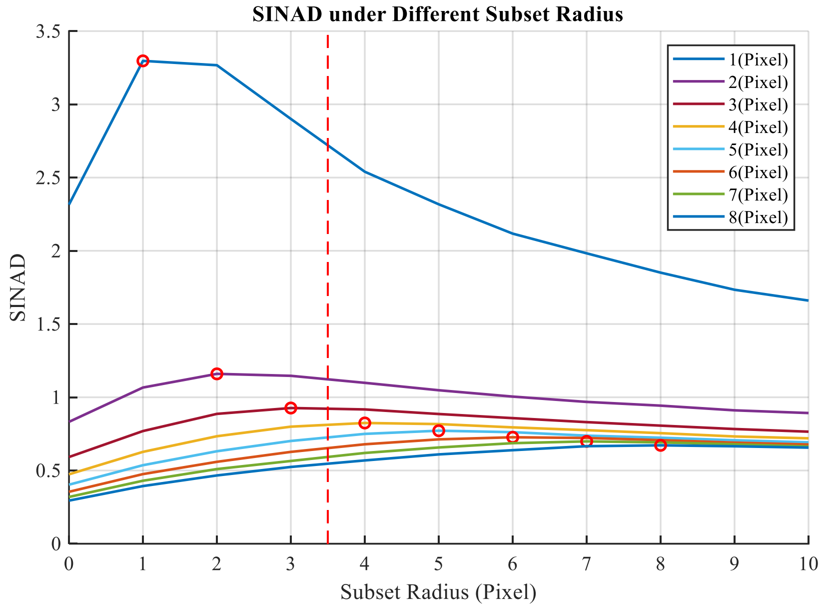



Figure 3-11 Subset size's influence on SINAD under different displacement amplitude.

In essence, as shown in Figure 3-11, SINAD declines continuously with the growth of displacement amplitude. Actually, when the subset radius is larger than radius of point spread function (PSF), the equivalent sensor of subset has no advantage than feature tracking methods (i.e., Lagrange description). In this simulation, the radius of PSF is around three pixels, so if displacement is larger than three pixels, the Lagrange description should be chosen in principle. Additionally, the SINAD will not decline with displacement amplitude.

## 3.5 Equivalent sensor based on Lagrange description

If the amplitude of displacement gets access to gradient vanishing area, an overlarge subset will contain some ineffective pixels in gradient vanishing area, and further lead to a poor performance in signal quality. For this kind of motion, displacement measured by tracking position of a certain point will get higher performance, for example, using corner detectors or sparse optical flow can track motion of a feature. One instance is modal analysis of wind turbine blades, the displacement of blade tip is often larger than ten or more pixels. Thus, a coarse-to-fine edge detection approach is developed for this large displacement case.

First, the procedures of smoothing noisy image are illustrated in Figure 3-12, by using an experimental image (Figure 3-12 (d)),

$$I_s(x) = I(x) * g(x) \tag{3-17}$$

the one row of image $I(x)$ (Figure 3-12 (a)) is convoluted by using a gaussian kernel $g(x)$ (Figure 3-12 (b)), that yields a smoothed image $I_s(x)$. This operation aims to get rid of the local gradient caused by noise, meanwhile remain the high gradient of edge as shown in Figure 3-12 (c).

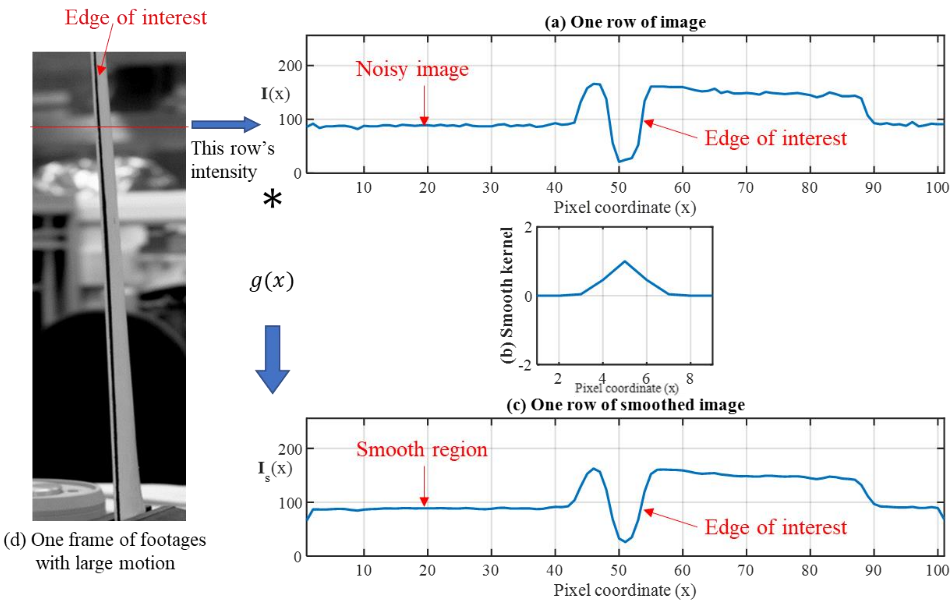



Figure 3-12 The original image is convoluted by a Gaussian kernel in order to suppress the noise in image

Consequently, the image gradient can be used to detect the edge. In order to obtain the subpixel level location of edge, the first gradient $I_s'(x)$ and second gradient $I_s''(x)$ of image is derived. Thus, the subpixel edge is determined as the point where the second gradient is zero $I_s''(x) = 0$, as shown in Figure 3-13.

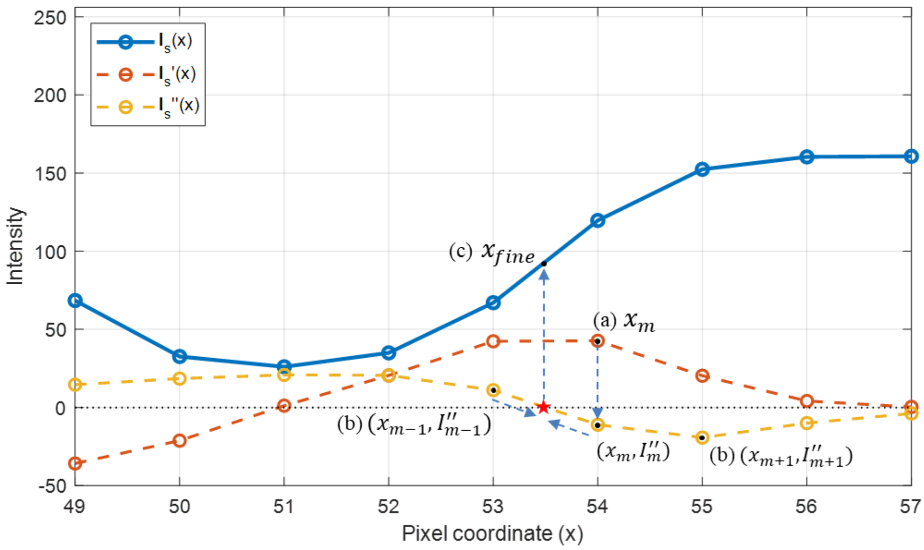



Figure 3-13 Schematic of process how the subpixel edge centre is determined, which uses the first and second order of image gradient

The process of determining $I_s''(x) = 0$ is shown in Figure 3-13. First, the maximum $I_s'(x)$ is used to determine the location of pixel-level edge $x_m$, (Figure 3-13 (a)). Then the pixel $x_{m-1}$ and $x_{m+1}$ beside $x_m$ is used to interpolate the subpixel edge $x_{fine}$, the interpolation can be expressed as,

$$x_{fine} = x_m + \frac{1 - s_{m+1}}{2} x_{m+1} + \frac{1 - s_{m-1}}{2} x_{m-1} \tag{3-18}$$

in which $x_1 = \frac{I_s''(x_m)}{I_s''(x_m) - I_s''(x_{m+1})}$, $x_2 = \frac{I_s''(x_m)}{I_s''(x_m) - I_s''(x_{m-1})}$ presents the distance between zero point and former pixel, $s_{m+1} = \frac{I_s''(x_m) I_s''(x_{m+1})}{|I_s''(x_m) I_s''(x_{m+1})|}$, $s_{m-1} = \frac{I_s''(x_m) I_s''(x_{m-1})}{|I_s''(x_m) I_s''(x_{m-1})|}$ denotes if or not the neighbouring pixel cross zero. Through above steps, the subpixel accuracy of edge can be obtained, and then the change of $x_{fine}$ in each frame is made up of the displacement signal.

This example can reflect the distinction between Euler description and Lagrange description. The Euler description uses the intensity at fixed pixel to estimate displacement. In contrast, the Lagrange description uses the position of a feature point to estimate displacement. They are appropriate for small and large displacement, respectively.

## 3.6 Key findings

In VVM, the formation of pixel array is not direct like other sensor arrays. In conventional cases, the array structure is fixed, but the pixels of an image cannot be used directly. Because it is not all pixels of image are useful to estimate displacement. In this section, an edge detection-based formation approach is proposed for formulating the pixel array, that can get all sensitive pixels involved in computation meanwhile exclude all useless pixel from computation to improve the signal quality in data acquisition. Furthermore, the influence of displacement amplitude on pixel array formation is investigated, the key findings are listed individually.

- A transfer function between intensity and displacement is formulated based on point spread function.
- The transfer function is further divided into linear area, nonlinear area and gradient vanishing area.
- Two classes of methods to form a sensor array are proposed based on Euler description and Lagrange description respectively.

- Applicable scope of two options is evaluated according to SINAD. Overall, small displacement (within its Airy disk) prefers Euler description, otherwise Lagrange description get priority.

# 4 ENHANCEMENT OF MODAL DISPLACEMENT BASED ON SPATIAL FILTERING

*After formulation of a pixel array, this section focuses on signal enhancement on modal displacement signal based on the principle of ASP. Because a major challenge in vision-based vibration measurement is extracting the weak displacement signal from noisy footages, especially in high-frequency range, considering that the displacement of motion is most insensitive in high-frequency among the variables that are used to describe a vibration (displacement, velocity, and acceleration). Furthermore, the upper bound of array gain is found according to Cauchy–Schwarz inequality, which is of importance to find out what the minimum detectable signal (MDS) is theoretically, when the noise floor and image size was already determined in a given image. Finally, a constraint condition is imposed to make implementation of adaptive filtering practicable.*

*It should be noted that some paragraphs in this chapter are expansion or reproduction of the published materials* [1] and [2] in list of publications, *and the specific paragraphs are annotated by a citation.*

**Highlight:**

- The SNR gain before and after spatial filtering relates to number of array elements.
- The SNR gain is also influenced by matching degree between weight vector and mode shape.
- Max array gain or minimum detectable signal demonstrates the limit of capacity of displacement estimate exists for a given image sequence.
- Imposing constraint on Euclidean norm of weight vector can equalize functionality of signal power and SNR in spatial filtering.

## 4.1 Spatial filtering based on maximization of SNR

From the relationship described in Figure 3-4, the displacement can be converted from intensity signal following the Eq. (4-1) [92],

$$\boldsymbol{u}(t) = \frac{\boldsymbol{I}_A(t) - F}{F'} \tag{4-1}$$

in which $\boldsymbol{I}_A(t)$ denotes a vector of intensity signal of pixel sensors selected by using array structure $A_s(x, y)$, $F$ is pixel's intensity in reference image and $F'$ is image gradient that is equivalent to the slope in Figure 3-4(a). Many existing methods can be used to solve the image gradient $F'(x)$, the Sobel operator is chosen in this study.

According to modal superposition, each dynamic response can be decomposed into several modes [111], as shown in Figure 2-1. In each mode, all space elements are oscillating with the same frequency and in phase. Its mode shape $\boldsymbol{\Phi}$ displays a form of standing wave, in which the motion direction is opposite across nodes, as denoted by the red and green arrow, respectively. The weight vectors $\boldsymbol{W}$ are thus designed for this structure, the weight vector is set to match the mode shape, so that the constructive waves will be generated.

As a result of conversion, $\boldsymbol{u}(t)$ becomes a vector of displacement signals collected by the sensor array. The dynamic response $\boldsymbol{u}(t)$ can be further regarded as a linear combination of multiple modes according to the principle of mode superposition [131]. Taking into inevitable noise influences, $\boldsymbol{u}(t)$ can be expressed in a matrix form:

$$\boldsymbol{u}(t) = \boldsymbol{\Phi}\boldsymbol{z}(t) + \boldsymbol{\eta}(t) \tag{4-2}$$

where $\boldsymbol{u}(t) \in \mathbb{R}^m$ is raw displacement signals from a sensor array with $m$ elements. $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_n] \in \mathbb{R}^{m\times n}$ is mode shape matrix with $n$ modes, in which $\boldsymbol{\varphi}_i \in \mathbb{R}^m$ is a mode shape vector. $\boldsymbol{z}(t) = [z_1(t), z_2(t) \ldots z_n(t)]^T \in \mathbb{R}^n$ is a vector representing modal displacement, i.e., displacement under modal coordinate, $\boldsymbol{\eta}(t) \in \mathbb{R}^m$ is a white Gaussian noise vector subject to $\boldsymbol{\eta} \sim N\left(0, {\sigma_\eta}^2\right)$. It is noted that the modal displacement $\boldsymbol{z}(t)$ is the target to be estimated in this study.

Conventional methods use measured $\boldsymbol{u}(t)$ to identify modal parameters directly, but window/subset derived $\boldsymbol{u}(t)$ usually have low SNR and such direct method often result in poor results. Considering there are a large number of channels/pixels available, an array signal processing scheme can be employed for enhancing the targeted components $\boldsymbol{z}(t)$. According to spatial filtering approach, the estimator of modal displacement $\hat{z}_i(t)$ can be obtained by applying weight vectors $\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n] \in \mathbb{R}^{m\times n}$ to the measured response $\boldsymbol{u}(t)$ [111], which yields:

$$\hat{\mathbf{z}}(t) = \boldsymbol{W^T u}(t) = \boldsymbol{W^T \Phi z}(t) + \boldsymbol{W^T \eta}(t) \tag{4-3}$$

in which $\hat{\mathbf{z}}(t) \in \mathbb{R}^n$ is an estimator of $\boldsymbol{z}(t)$. As illustrated in section 2.3, the structure of array signal processing shows that each output $\hat{z}_i(t)$ is a linear combination of all array elements. By multiplying $\boldsymbol{W}$, the displacement signals are transformed from a *m*-dimension spatial coordinate into *n*-dimension modal coordinate.

To examine the effect of $\boldsymbol{w}_i$ more conveniently on a particular estimator $\hat{z}_i(t)$, Eq. (4-3) is rewritten as

$$\hat{z}_i(t) = \sum_{j=1}^{n} \boldsymbol{w_i}^{\boldsymbol{T}} \boldsymbol{\varphi_j}\, z_j(t) + \boldsymbol{w_i}^{\boldsymbol{T}} \boldsymbol{\eta}(t) \tag{4-4}$$

It shows that $\hat{z}_i(t)$ is composed by two parts: $\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_j\, z_j(t)$ due mainly to the modal contribution and noise $\boldsymbol{w}_i{}^T \boldsymbol{n}(t)$ due to noise influence. To extract the weak displacement signal, the weight vector $\mathbf{w}_i$ needs to be designed so that the ratio between signal component and noise component i.e., SNR of $\hat{z}_i(t)$ is maximized. Due to the signal components of interest in $\hat{z}_i(t)$ is only $z_i(t)$, which turns out that other modal signals should be ignored in $\hat{z}_i(t)$ and hence result in the $i^{th}$ term $\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i z_i(t)$ alone. Therefore, the SNR of array output $\hat{z}_i(t)$ can be expressed as:

$$SNR[\hat{z}_i(t)] = \frac{E\{[\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i\, z_i(t)]^2\}}{E\{[\boldsymbol{w}_i{}^T \boldsymbol{\eta}(t)]^2\}} \tag{4-5}$$

where the weight vector $\boldsymbol{w}_i$ is the variable of this equation. Taking output $SNR[\hat{z}_i(t)]$ as an objective function and $\boldsymbol{w}_i$ as a variable, it enables an optimization problem to formulate as:

$$\boldsymbol{w}_i = arg \max_{\boldsymbol{w}_i} SNR[\hat{z}_i(t)] \tag{4-6}$$

The SNR of $\hat{z}_i(t)$ is the ratio between power of signal and variance of noise:

$$SNR[\hat{z}_i(t)] = \frac{E\{[(\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i)\, z_i(t)]^2\}}{E\{[\boldsymbol{w}_i{}^T \boldsymbol{\eta}(t)]^2\}} \tag{4-7}$$

Eq. (4-7) can be reformed into following equation,

$$SNR[\hat{z}_i(t)] = \frac{E\{(\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i)^2 z_i(t)^2\}}{E\{\boldsymbol{w}_i{}^T \boldsymbol{\eta}(t)\boldsymbol{\eta}(t)^T \boldsymbol{w}_i\}} \tag{4-8}$$

Due to the expectation is with respect to the time $t$, other components can be taken out of expectation. Meanwhile, $\boldsymbol{R}_\eta = \boldsymbol{\eta}(t)\boldsymbol{\eta}(t)^T$ denotes the covariance matrix of noise [111].

$$SNR[\hat{z}_i(t)] = \frac{(\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i)^2 E\{z_i(t)^2\}}{\boldsymbol{w}_i{}^{\boldsymbol{T}} \boldsymbol{R}_\eta \boldsymbol{w}_i} \tag{4-9}$$

in which, signal power can be presented by $\sigma_{z_i}{}^2 = E\{z_i(t)^2\}$. Then because $\boldsymbol{\eta}(t)$ is Gaussian noise, its covariance matrix is $\boldsymbol{R}_\eta = \sigma_\eta{}^2 \boldsymbol{I}$.

$$SNR[\hat{z}_i(t)] = \frac{(\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i)^2}{\boldsymbol{w}_i{}^T \boldsymbol{w}_i} \frac{\sigma_{z_i}{}^2}{\sigma_\eta{}^2} \tag{4-10}$$

According to Cauchy–Schwarz inequality,

$$SNR[\hat{z}_i(t)] \leq \frac{(\boldsymbol{w}_i{}^T \boldsymbol{w}_i)(\boldsymbol{\varphi}_i{}^T \boldsymbol{\varphi}_i)}{\boldsymbol{w}_i{}^T \boldsymbol{w}_i} \frac{\sigma_{z_i}{}^2}{\sigma_\eta{}^2} \tag{4-11}$$

By cancelling the same item of $\boldsymbol{w}_i{}^T \boldsymbol{w}_i$, the maximum SNR of the output is a value independent of the weight vector.

$$SNR[\hat{z}_i(t)] \leq (\boldsymbol{\varphi}_i{}^T \boldsymbol{\varphi}_i) \ \frac{\sigma_{z_i}{}^2}{\sigma_\eta{}^2} \tag{4-12}$$

in which, $\sigma_{z_i}{}^2 = E\{z_i(t)^2\}$ represent signal power and $\sigma_\eta{}^2$ is the variance of noise. It shows that if and only if $\boldsymbol{w}_i$ and $\boldsymbol{\varphi}_i$ are linearly dependent, SNR is maximized. According to Cauchy–Schwarz inequality, it can be shown that $SNR[\hat{z}_i(t)]$ has its upper bound. In another word, the upper bound of SNR can be reached, while $\boldsymbol{w}_i$ paralleling with mode shape vector $\boldsymbol{\varphi}_i$, i.e., it yields that $\widehat{\boldsymbol{\varphi}}_i = \boldsymbol{w}_{opt}$, because they are dimensionless. In this way, a SNR-maximized $\mathbf{w}_i$ can be found.

## 4.2 Constraint form of optimization problem

The maximization of SNR establishes between mode shape and modal displacement signal [92]. That pave the way for estimate mode shape and obtain desired signal, however, SNR is difficult to be estimated accurately from experimental data. To make implementation of proposed method feasible, a constraint optimization is transformed from Eq. (4-5) and (4-6).

With a unit weight vector $\boldsymbol{w}_i$, i.e., $\|\boldsymbol{w}_i\|^2 = 1$, the denominator in the middle term of Eq. (4-10) is eliminated, which turns Eq. (4-5) into a constraint optimization problem:

$$\begin{aligned} &\max_{\boldsymbol{w}_i} \quad SNR(\boldsymbol{w}_i) = (\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i)^2 \, \frac{{\sigma_{z_i}}^2}{{\sigma_\eta}^2} \\ &s.t. \quad\quad \|\boldsymbol{w}_i\|^2 = 1 \end{aligned} \tag{4-13}$$

in which $\|\boldsymbol{w}_i\|$ is the Euclidean norm (also called the vector magnitude, Euclidean length, or 2-norm) of vector $\boldsymbol{w}_i$ [132], [133]. To unfold it, that can be expressed as,

$$\|\boldsymbol{w}_i\|^2 = \sum_{k=1}^{m} \left(\boldsymbol{w}_{i,k}\right)^2 = (\boldsymbol{w}_i^T \boldsymbol{w}_i) \tag{4-14}$$

In addition, the noise power $\sigma_\eta^2$ is constant unrelated to $\boldsymbol{w}_i$ as displayed in Eq.(4-10). Therefore, $SNR(\boldsymbol{w}_i)$ is determined by its part of signal power $P(\boldsymbol{w}_i) = (\boldsymbol{w}_i{}^T \boldsymbol{\varphi}_i)^2 {\sigma_{z_i}}^2$ only. This means that $P(\boldsymbol{w}_i)$ can substitute $SNR(\boldsymbol{w}_i)$ to serve as the objective function. In practice, $P(\boldsymbol{w}_i)$ can be easily obtained by measuring the power spectral density at frequency of interest. After the optimization problem is fully defined based on Eq.(4-13), the mode-shape-based adaptive spatial filter (MASF) is established. Then a searching scheme based on node/antinode is put forward as one example of many potential numerical solutions.

## 4.3 Array gain

To evaluate the performance of array, the conception of array gain is introduced, in order to exclude the influence of input signal, which implies the gain of SNR before and after spatial filtering [118].

$$G_a = \frac{SNR[\hat{z}_i(t)]}{\overline{SNR}_{in}} \tag{4-15}$$

where $\overline{SNR}_{in}$ is defined as the average input SNR over all pixels.

In order to find the array gain, the input SNR of each pixel is calculated as:

$$SNR_{in} = \frac{E\left\{\left[\varphi_{ij}\ z_i(t)\right]^2\right\}}{E\{[\eta(t)]^2\}} \tag{4-16}$$

The $\varphi_{ij}$ can be taken out of expectation, and then substitute ${\sigma_{z_i}}^2$ and ${\sigma_\eta}^2$ into Eq. (A8).

$$SNR_{in} = \frac{{\varphi_{ij}}^2 {\sigma_{z_i}}^2}{{\sigma_\eta}^2} \tag{4-17}$$

The mean power of input signal is:

$$\overline{SNR}_{in} = \frac{1}{m}\sum_{j=1}^{m} {\varphi_{ij}}^2 \left(\frac{{\sigma_{z_i}}^2}{{\sigma_\eta}^2}\right) \tag{4-18}$$

It can be written into a form of inner product of $\boldsymbol{\varphi_i}$.

$$\overline{SNR}_{in} = \frac{1}{m}(\boldsymbol{\varphi_i}^{\boldsymbol{T}}\boldsymbol{\varphi_i})\left(\frac{{\sigma_{z_i}}^2}{{\sigma_\eta}^2}\right) \tag{4-19}$$

The array gain is defined as the gain of SNR between output and input signals.

$$G_a = \frac{SNR[\hat{z}_i(t)]}{\overline{SNR}_{in}} \tag{4-20}$$

Substituting Eq. (4-12) and Eq. (4-19) into Eq. (4-20) yields

$$G_a \le \frac{(\boldsymbol{\varphi_i}^{\boldsymbol{T}}\boldsymbol{\varphi_i})\ E\{z_i(t)^2\}}{E\{\eta(t)^2\}} \Big/ \frac{1}{m}(\boldsymbol{\varphi_i}^{\boldsymbol{T}}\boldsymbol{\varphi_i})\ \frac{E\{z_i(t)^2\}}{E\{\eta(t)^2\}} \tag{4-21}$$

By cancelling all the same terms, only $m$ remains.

$$G_a \le m \tag{4-22}$$

In addition, the analogous relationship between array gain $G_a$ and number of array elements $m$ can be easily found in many kinds of sensor arrays [118].By excluding the same terms in Eq. (4-21), it shows that the upper bound of array gain is $m$, where $m$ is the number of array elements. Eq. (4-22) shows that the maximum array gain $G_a$ depends only on the number of array elements $m$, it is often expressed in dB.

$$G_{dB} \le 10\ \log_{10}(m) \tag{4-23}$$

In general, increasing number of array element $m$ enables a higher SNR output, which can be achieved conveniently by this video-based measurement [92].

## 4.4 Minimum detectable signal

Furthermore, the lower bounder of detectable signal using image-based measurement can be determined. Minimum detectable signal is an index that is defined by a least power of a signal that is detectable or discernible from the noise floor [126]. Based on analysis in Section 3.5.1, the noise floor is composed of quantization noise, image noise that can be expressed as,

$$N_f = 10log_{10}\left(N_i + N_Q\right) \tag{4-24}$$

in which the $N_i$ and $N_Q$ represent the image noise and quantization noise respectively. Thereafter, the MDS can be seen as a value resulting from the noise floor, array gain and the required minimum $SNR_{min}$ jointly.

$$SNR_{min} = (MDS + G_{dB}) - N_f \tag{4-25}$$

In particular, $SNR_{min}$ specifies threshold that the minimum signal to be recognized from its noise floor [19]. Based on the above descriptions, multiple factors (noise intensity, array gain and required SNR) will determine MDS of the system. If Eq. (4-25) converts to (4-26), MDS can be analysed.

$$MDS = SNR_{min} + N_f - G_{dB} \tag{4-26}$$

Firstly, the color depth determines the scale of quantization noise, the light source and camera parameters determine the random noise. If noise floor gets higher, MDS will be larger, i.e., the weak signal cannot be recognized. Array gain is determined by the estimation method, if array gain get lager, MDS becomes smaller. Finally, the required SNR is adapted based on requirement of post-processing method or a human operator.

## 4.5 Key findings

The core component of algorithm is proposed in this chapter, by which the spatial propagation of signal (mode shape in modal analysis) and the time sequence of desired signal can be obtained in principle. The conclusions can be summarized as follow.

- The upper bounder of array gain is determined by number of array element (effective pixels selected).
- The array gain is affected by weight vector, while the pixel array is finalized.

- Consequently, the minimum detectable signal (MDS) can be determined by considering the array gain and noise floor.
- Finally, signal power can replace SNR in optimization problem by imposing a constraint condition to make the algorithm executable.

# 5 ESTIMATE OF MODE SHAPE BASED ON SPATIAL PARAMETER ESTIMATION

*As the optimization problem has been defined in Chapter 4, this chapter aims to estimate spatial parameter, specifically mode shape. Two typical methodologies of spatial parameter estimation in ASP is employed to estimate the mode, which are the adaptive filtering method by a process of maximizing the output signal power and subspace method. In comparison, the adaptive filtering can obtain higher SNR, which is acquired in exchange of pre-setting the elementary shape of mode shape and an iterative process. In terms of algorithm properties, it is proved that the proposed adaptive filter is least mean square filter, and the objective function is equivalent to modal assurance criterion, which are crucial property to guide the future development.*

*It should be noted that some paragraphs in this chapter are expansion or reproduction of the published materials* [1] in list of publications*, such as section 5.1.3.1 and 5.1.4, and the specific paragraphs are annotated by a citation.*

**Highlight:**

- An adaptive spatial filtering based on node/antinode searching is developed for estimate of mode shape.
- The proposed adaptive spatial filtering is a least mean square filter.
- SVD based method can estimate mode shape and natural frequency without iteration, but capability of noise suppression is low relatively.

## 5.1 Adaptive filtering method

### 5.1.1 Properties of proposed algorithm

The optimization problem is defined completely in Chapter 4; however, estimating mode shape requires to solve the optimization problem in adaptive spatial filtering. To tackle that, some crucial properties of proposed algorithm are investigated first.

#### 5.1.1.1 Least mean square filter

This section will prove that the proposed method is a least mean square filter, in order to connect the error and signal power together. First, the error of $\boldsymbol{w}$ is defined as the difference between $\boldsymbol{w}$ and $\pm\boldsymbol{\phi}$. Since $\boldsymbol{\phi}$ is a dimensionless, positive, or negative sign of $\boldsymbol{\phi}$ have no difference in presenting mode shape.

$$\boldsymbol{e} = \begin{cases} \boldsymbol{w} - \boldsymbol{\phi} & if\ cos\theta > 0 \\ \boldsymbol{w} + \boldsymbol{\phi} & if\ cos\theta < 0 \end{cases} \tag{5-1}$$

The sign is determined by initial $\boldsymbol{w}$ is set close to $\boldsymbol{\phi}$ $(cos\theta > 0)$ or $-\boldsymbol{\phi}$ $(cos\theta < 0)$, as shown in Figure 5-1.

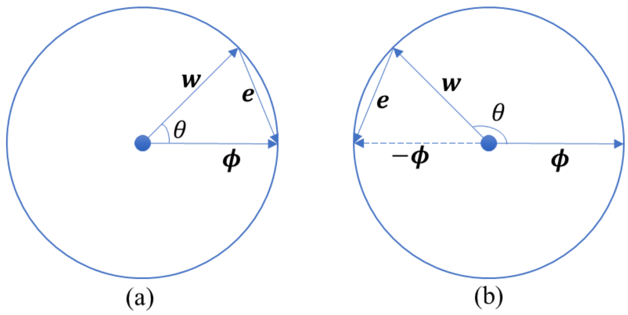



Figure 5-1 Schematic of relationship among weight vector $\boldsymbol{w}$, error vector $\boldsymbol{e}$ and mode shape vector $\boldsymbol{\phi}$ under (a), (b) two potential directions

As the principle of least squares, the square of error is defined as the cost function.

$$||\boldsymbol{e}||^2 = ||\boldsymbol{w} \pm \boldsymbol{\phi}||^2 \tag{5-2}$$

That can be expanded as follow.

$$||\boldsymbol{e}||^2 = ||\boldsymbol{w}||^2 + ||\boldsymbol{\phi}||^2 \pm 2\boldsymbol{w}^T\boldsymbol{\phi} \tag{5-3}$$

The inner product arises in Eq. (5-3), which is also in the objective function of the proposed adaptive filtering.

$$\boldsymbol{w}^T\boldsymbol{\phi} = \pm\frac{1}{2}(\|\boldsymbol{w}\|^2 + \|\boldsymbol{\phi}\|^2 - \|\boldsymbol{e}\|^2) \tag{5-4}$$

Square this term, so that it will be equal to the proposed objective function,

$$(\boldsymbol{w}^T\boldsymbol{\phi})^2 = \frac{1}{4}(\|\boldsymbol{w}\|^2 + \|\boldsymbol{\phi}\|^2 - \|\boldsymbol{e}\|^2)^2 \tag{5-5}$$

in which $\|\boldsymbol{w}\|^2$ and $\|\boldsymbol{\phi}\|^2$ are constants, as a result, maximizing $(\boldsymbol{w}^T\boldsymbol{\phi})^2$ is equivalent to minimize $\|\boldsymbol{e}\|^2$. The equation below is established constantly,

$$\|\boldsymbol{w}\|^2 + \|\boldsymbol{\phi}\|^2 \geq \|\boldsymbol{e}\|^2 \tag{5-6}$$

which can also be derived from geometry in Figure 5-1. Moreover, $\boldsymbol{w}$ is likely to converge on $\boldsymbol{\phi}$ or $-\boldsymbol{\phi}$ based on (5-1). But there is no difference between $\boldsymbol{\phi}$ and $-\boldsymbol{\phi}$ in principle, that depends on the initial value. By now, the proposed filter has been proved to be a least mean square filter (LMSF), so all beneficial property of LMSF can be used to design the solution.

#### 5.1.1.2 Modal assurance criterion

The modal assurance criterion (MAC) is a statistical indicator that is most sensitive to large difference and relatively insensitive to small differences in the mode shapes. This yields a good statistic indicator and a degree of consistency between mode shapes[134].

One normalized MAC is defined as the normalized scalar product of the two sets of vectors $\boldsymbol{w}$ and $\boldsymbol{\phi}$. The resulting scalars are arranged into the MAC matrix

$$MAC(\boldsymbol{w}, \boldsymbol{\phi}) = \frac{|\boldsymbol{w}^T\boldsymbol{\phi}|^2}{(\boldsymbol{w}^T\boldsymbol{w})(\boldsymbol{\phi}^T\boldsymbol{\phi})} \tag{5-7}$$

$\boldsymbol{w}$ denotes the exerted weight vector, which is the supposed estimator of mode shape. $\boldsymbol{\phi}$ denotes actual mode shape to be measured. In this sense, the result of objective function is equivalent to MAC. This equation provides an approach of model updating [135], [136]. If $\boldsymbol{w}$ is derived from finite element analysis, $\boldsymbol{w}$ can be updated based on the measured data $\boldsymbol{\phi}$ directly.

### 5.1.2 Principle of adaptive filtering

As proved in section 5.1.1.1, maximizing $(\boldsymbol{w}^T\boldsymbol{\phi})^2$ is equal to minimizing $\|e\|^2$. Following the strategy shown in Figure 5-2, the result of $(\boldsymbol{w}^T\boldsymbol{\phi})^2$ can substitute error $\boldsymbol{e}$ as the feedback to update the weight vector $\boldsymbol{w}$.

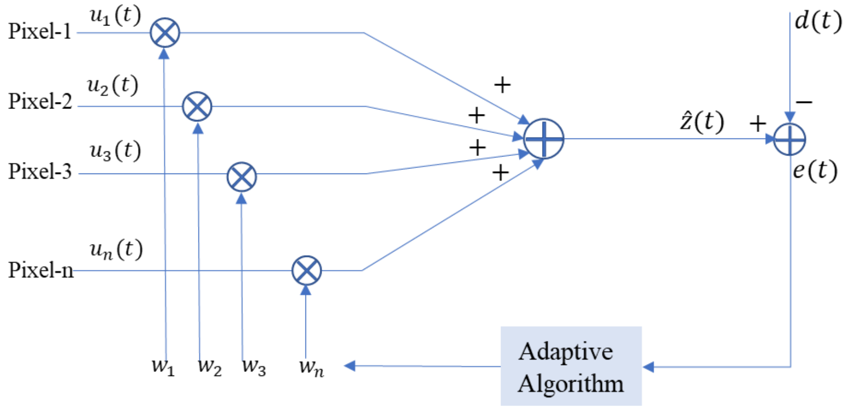



Figure 5-2 Schematic of a typical adaptive spatial filter, the error is replaced by signal power, which has been proved equivalent

In conclusion, the proposed algorithm operates under a standard paradigm of adaptive spatial filtering, in which the signal power (equivalent to error $e(t)$) serves as the feedback, and then the weight vector is updated to approach to actual mode shape $\boldsymbol{\phi}$ in a close loop.

### 5.1.3 Approximation of mode shape of continuous structure

In physics, mode shape of a continuous structure is often a type of normal mode presenting a form of standing wave. In addition, the location of nodes and antinodes shows where the strongest and weakest vibration occurs, which highly influences the performance and safety of the object under test. Generally, standing waves are thought to be produced by interference between waves reflected back and forth at natural frequencies, so that nodes and antinodes can be viewed as the characteristics of a standing wave. In three-dimensional scenario, it generally exists in forms of nodal lines, nodal diameters, or nodal circle of common engineering structures, such as turbine blades, aircraft wings, crane structures and so forth [92].

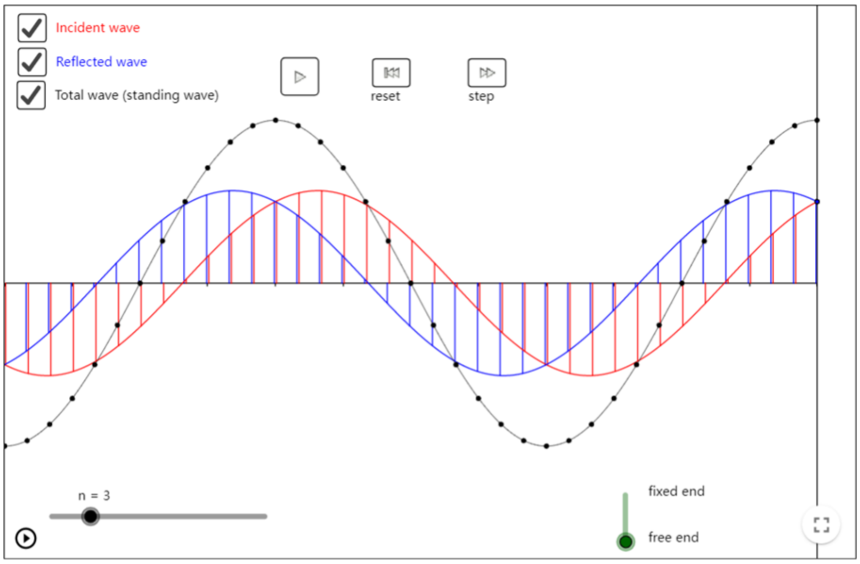



Figure 5-3 A schematic of how a standing wave is generated by two travelling waves from opposite direction, the dynamic process can be observed by animation in [137].

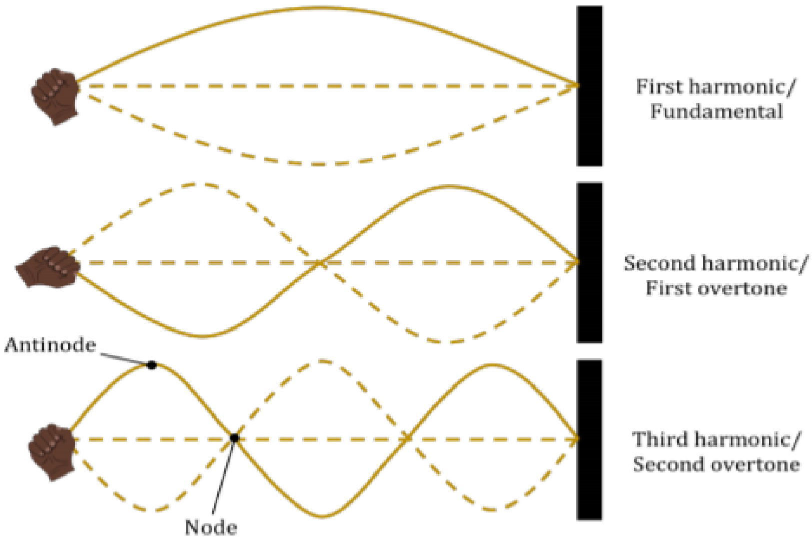



Figure 5-4 The harmonics of standing wave, the node number is increasing with order of harmonics [138], [139]

In addition, the location of nodes and antinodes shows where the strongest and weakest vibration occurs, which highly influences the performance and safety of the object under test.

#### 5.1.3.1 Approximation using a sinusoid-based piecewise function

As mode shape can be taken as a wave signal, the sinusoidal function is chosen for numerically approximating the desired mode shape [92]. The nodes and antinodes split the weight vector $\boldsymbol{w}$ into $j$ pieces $w_j$ to formulate a piecewise function and each piece can be denoted by

$$w_j(y) = A_j \cos\left(\frac{2\pi}{\lambda_j} y + \psi_j\right) \qquad 0 < y < \frac{\lambda_j}{4} \tag{5-8}$$

in which $y$ is local coordinate inside each piece of function, $A_j$, $\lambda_j$, $\psi_j$ is the undetermined amplitude, wavelength, and initial phase, respectively. Its derivative can be obtained as

$$\frac{dw_j}{dy} = -\frac{2\pi A_j}{\lambda_j} \sin\left(\frac{2\pi}{\lambda_j} y + \psi_j\right) \qquad 0 < y < \frac{\lambda_j}{4} \tag{5-9}$$

Following the physical property of wave, the mode shape and its changing rate should be continuous, therefore, the constructed function and its derivative will be continuous as well. This means the amplitudes should be identical on both sides of a node/antinode, thereafter, the boundary condition is imposed on the endpoints of every piece as following,

$$w_j(\frac{\lambda_j}{4}) = w_{j+1}(0) \tag{5-10}$$

Likewise, its derivative also accords with continuity,

$$\dot{w}_j(\frac{\lambda_i}{4}) = \dot{w}_{j+1}(0) \tag{5-11}$$

Naturally, the value of $w_j$ at a node should be zero as well as its derivative at antinode should be zero, and then the boundary condition can be further simplified at a node.

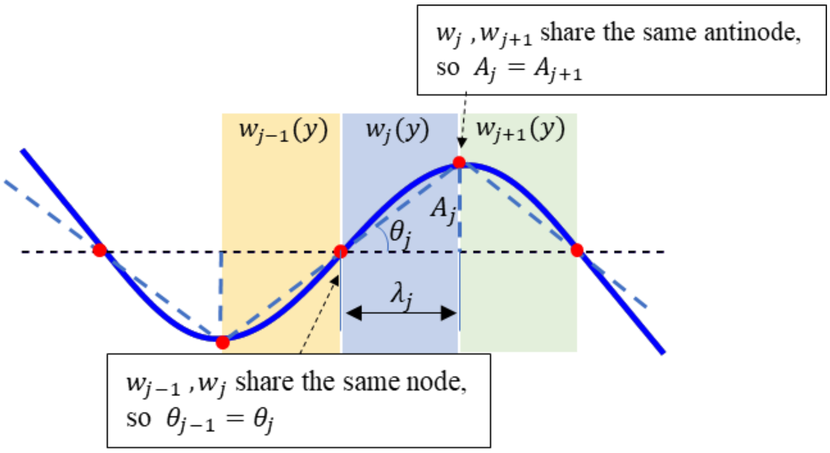



Figure 5-5 A schematic of connection between pieces of piecewise function, it should be noted that the condition imposed on node and antinode are different.

Thus, there exists a relationship for two adjacent amplitudes $A_j$ and $A_{j+1}$, as shown below. This represents the parameters on both side of a node accord with this equation, namely they share the same included angle $\theta_j$. Similarly, at the antinode, two pieces share one amplitude,

$$if\ \ w_j(0) = 0,$$

$$\begin{cases} \dfrac{4A_j}{\lambda_j} = \dfrac{4A_{j-1}}{\lambda_{j-1}} = \tan\theta_j \\ A_j = A_{j+1} \end{cases} \tag{5-12}$$

Likewise, if the $w_j$ starts with an antinode and end up with a node, the relationship is like this, it should be noted that the subscription is different from Eq. (5-12)

$$if\ \ w_j\left(\frac{\lambda_j}{4}\right) = 0,$$

$$\begin{cases} \dfrac{4A_j}{\lambda_j} = \dfrac{4A_{j+1}}{\lambda_{j+1}} = \tan\theta_j \\ A_j = A_{j-1} \end{cases} \tag{5-13}$$

By imposing the above boundary conditions, the neighbouring pieces are coupled by the shared $A_j$ or $\theta_j$, and then based on the Eq. (5-12), $A_j$ can be calculated by

$$A_j = \frac{1}{4}\lambda_j \tan(\theta_j) \tag{5-14}$$

Then, the local coordinate and local phase in each piece can be derived from the global coordinate.

$$x = y - l_j \tag{5-15}$$

$$\psi_j = \frac{\pi}{2}(j-1) \tag{5-16}$$

in which the $l_j$ is the location of node/antinode, which is explained in detail in section 5.1.3.2. Correspondingly, $x$ is under global coordinate. Then substitute all variables into the Eq. (5-8),

$$w(x) = \frac{tan(\theta_j)\lambda_j}{L}\cos\left(\frac{2\pi}{\lambda_j}(x - l_j) + \frac{\pi}{2}(j-1)\right)\ \ l_j < x < l_{j+1} \tag{5-17}$$

in which $x \in (0, L)$ and $L$ is the length of object. By using this function and considering specific boundary condition, the profile of arbitrary mode shape can be approximated numerically, if only the mode shape shows a form of standing wave.

#### 5.1.3.2 Representation of free-free object

The elementary shape of each mode shape can be described by a piece of sinusoidal piecewise function with parameters of all locations $l_j$ of nodes/antinodes

$$l_j = \{0, d_1, a_1, d_2, a_2, \dots, L\} \in (0, L) \tag{5-18}$$

$j \epsilon \mathbb{N}^n$ is number of all node/antinodes, $L$ is the length of object, $d_k$ and $a_k$ denotes location of nodes and antinodes respectively. Because of the constraint in Eq. (5-12), these parameters are coupling as follows,

$$\frac{tan\ (\theta_k)}{tan\ (\theta_k) + tan\ (\theta_{k+1})} = \frac{a_k - d_k}{d_{k+1} - d_k} \tag{5-19}$$

$\theta_k$ denotes included angle its $tan\ (\theta_k)$ is the ratio of amplitude $A_{j+1}$ to a quarter of wavelength $\lambda_j$, which is one factor of the amplitude of each piece. Hence, there is

$$h_j = \{tan\ (\theta_1), tan\ (\theta_1), tan\ (\theta_2), tan\ (\theta_2), \dots\} \tag{5-20}$$

where $h_j$ is hence defined as a coefficient dependent of $a_k$, $d_k$ and $d_{k+1}$. Between the adjacent node and antinode $l_j$ and $l_{j+1}$., the wavelength $\lambda_j$ is determined by their distance of them,

$$\lambda_j = 4(l_{j+1} - l_j) \tag{5-21}$$

Then considering the free-free boundary condition, a piecewise function for twist mode shapes in one-dimension can be constituted as

$$w(x) = \frac{h_j \lambda_j}{L} \cos\left(\frac{2\pi}{\lambda_j}(x - l_j) + \frac{\pi}{2}(j-1)\right) \qquad if \quad l_j < x < l_{j+1} \tag{5-22}$$

in which, $x \in (0, L)$ and $w(x)$ is the weight vector to be optimized.

Likewise, a piecewise function for bending mode shape is defined with undetermined parameters $l_j$ to be optimized, other parameters are dependent of $l_j$, and coupling due to continuity and smoothness condition.

$$w(x) = \begin{cases} \frac{\pi h_1 \lambda_1}{2L}\left(1 - \frac{x}{l_2}\right) & if \quad x < l_2 \\ \frac{h_j \lambda_j}{L} \cos\left(\frac{2\pi}{\lambda_j}(x - l_j) + \frac{\pi}{2}(j-1)\right) & if\ l_j < x < l_{j+1} \\ (-1)^n \frac{\pi h_n \lambda_n}{2L}(x - l_n) & if \quad x > l_{n-1} \end{cases} \tag{5-23}$$

In this numerical expression, the mode shape can be represented by variable node location.

#### 5.1.3.3 Representation of cantilever object

In comparison, the mode shape of cantilever object can be expressed similarly. The fixed end is a node, and another end is an antinode constantly.

$$w(x)=\begin{cases}\dfrac{h_j\lambda_j}{2L}\left[\cos\left(\dfrac{4\pi}{\lambda_j}x\right)-1\right] & if\ x<l_2\\ \dfrac{h_j\lambda_j}{L}\cos\left(\dfrac{2\pi}{\lambda_j}(x-l_j)+\dfrac{\pi}{2}j\right) & if\ l_j<x<l_{j+1}\\ (-1)^{\frac{n+1}{2}}\dfrac{\pi h_n\lambda_n}{2L}(x-l_n) & if\quad x>l_{n-1}\end{cases} \tag{5-24}$$

These two representations demonstrate that the sinusoid-based piecewise function can represent mode shape under different boundary condition.

#### 5.1.3.4 Representation of 2D and 3D mode shape

The representation of 2D or 3D mode shape is similar to 1D case, which needs to $x$ axis and $y$ axis is equal to the 1D situation. However, due to their different geometry, the matrix $M$ is applied to rotate or transform the coordinate.

$$\boldsymbol{\Phi}(x,y)=\boldsymbol{\phi}(\boldsymbol{A}\boldsymbol{x})\boldsymbol{\phi}(\boldsymbol{A}\boldsymbol{y})^T \tag{5-25}$$

In a similar way, the 2D mode shape can be defined in a polar coordinate system,

$$\boldsymbol{\Phi}(\theta,r)=\boldsymbol{\phi}(\boldsymbol{A}\boldsymbol{\theta})\boldsymbol{\phi}(\boldsymbol{A}\boldsymbol{r})^T \tag{5-26}$$

#### 5.1.3.5 Approximation error

The sinusoidal function can approximate the profile of mode shape basically, since the sinusoid function can compose any periodical function based on the Fourier expansion. However, the error exists inevitably because the mode shape is often nonstandard sinusoid function. Because the inner force influences the local deformation, which is also related to the geometry and boundary condition of object. That leads to the curvature cannot be matched at every point. However, the profile shows a high consistency overall as shown in Figure 5-6.

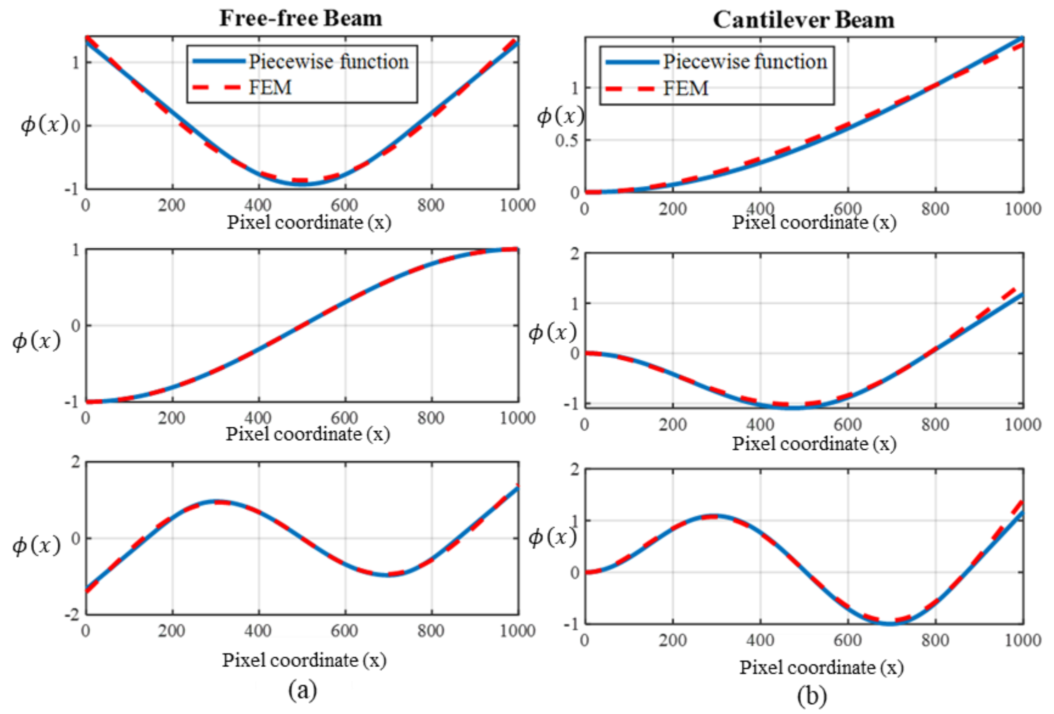



Figure 5-6 Comparison between numerical expression and FEM

To evaluate the performance of the proposed algorithm, the difference of weight vector and the mode shape obtained from FEA is calculated. The errors of free-free beam and cantilever beam are shown in Figure 5-7. The largest error is less than 10%, and more importantly, it does not bias the estimation of mode shape on the whole.

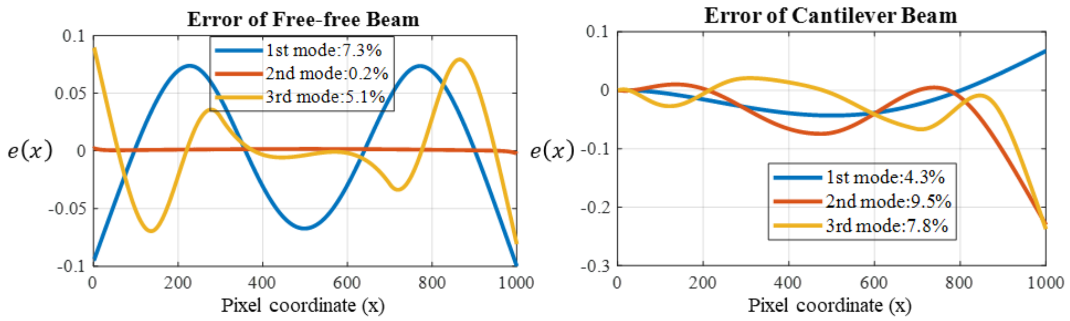



Figure 5-7 Approximation error of free-free beam and cantilever beam

### 5.1.4 Node/antinode searching approach

Based on the above analysis, the location of nodes/antinodes can denote a mode shape approximately by a few parameters [92]. By moving location of each node in the weight vector $\boldsymbol{w}_k$ along direction of increasing signal power $P_k$, the maximal $P_k$ can be obtained.

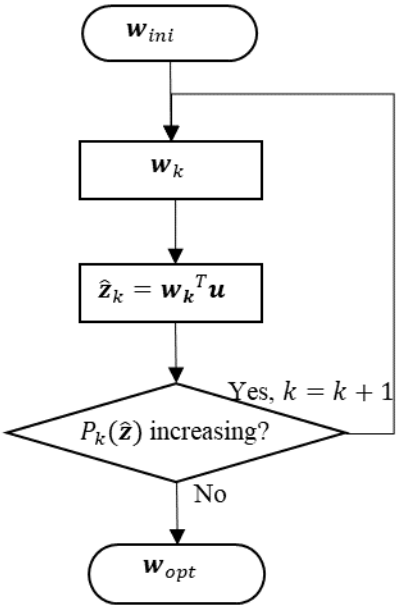



Figure 5-8 Workflow of Node/antinode searching approach

In this expression, the number of variables can be reduced significantly, from number of pixels to number of nodes and antinodes, which makes the optimization problem feasible. Based on this piecewise function, the workflow shown in Figure 5-8 of optimization process can be described as follows:

- Set $w(x)$ using node/antinode location $l_i$ as parameter.
- Set initial value of node/antinode location $\boldsymbol{l}_{ini}$, and obtain $\boldsymbol{w}(\boldsymbol{l}_{ini})$
- Apply $\boldsymbol{w}(\boldsymbol{l}_{ini})$ on the input signals to obtain the output signal $\hat{\boldsymbol{z}}$, and calculate its power $P(\boldsymbol{w}_{ini})$ at a frequency range where the mode lies in.
- Determine the iteration direction based on initial value and the first iteration, which moves the node by one-pixel distance and compares the power $P(\boldsymbol{w}_i)$ with initial one, if the power increases, it is taken as the direction for the following iterations, otherwise, the opposite direction will be used.
- Move the node/antinode $\boldsymbol{l}_k$ and obtain its $P(\boldsymbol{w}_k)$. and repeat this step along the direction of increasing $P(\boldsymbol{w}_k)$
- Terminate when $P(\boldsymbol{w}_k)$ starts to drop, and then determine the optimal weight vector $\boldsymbol{w}_{opt}$ which has the maximum $P(w)$.

## 5.2 Subspace of eigenstructure method

### 5.2.1 Singular value decomposition of array signal model

The node/antinode searching method can obtain mode shape with high accuracy but it requires knowing elementary shape of mode shape, for example, a standing wave. However, some situation cannot meet this assumption, for example, when the boundary condition is uncertain, or the structure is not a continuous object. To improve the robustness of PASP framework, and eigen-decomposition method is developed, which transforms the collected input signals matrix $\boldsymbol{U}$ using singular value decomposition (SVD). In this way, the mode shape and the time sequence of modal signal can be extracted from $\boldsymbol{U}$.

$$\boldsymbol{u}(t) = \boldsymbol{\Phi}\boldsymbol{z}(t) + \boldsymbol{\eta}(t) \tag{5-27}$$

The modal superposition is still the model of signal, then the covariance matrix $\boldsymbol{R}_Z$ of modal signal can be written as,

$$\boldsymbol{R}_Z = \boldsymbol{z}(t)\boldsymbol{z}(t)^T = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1n}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \cdots & \sigma_{nn}^2 \end{bmatrix} \tag{5-28}$$

As the modal signals $\boldsymbol{z}(t)$ are narrowband signals at different frequencies, obviously they are vectors of an orthogonal basis [140] that yields,

$$\sigma_{ij}^2 = \begin{cases} 0 & if\ i \neq j \\ \sigma_{z_i}^2 & if\ i = j \end{cases} \tag{5-29}$$

The orthogonality of this basis can be proven by Fourier series [141], that means covariance of different modes are zeros, except the variance of that mode signal. Thus, Eq. (5-28) can be simplified as

$$\boldsymbol{R}_Z = \boldsymbol{\Sigma}_z^2 = \begin{bmatrix} \sigma_{z_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{z_n}^2 \end{bmatrix} \tag{5-30}$$

The signal matrix can be therefore converted to an orthogonal matrix.

$$\left(\boldsymbol{\Sigma}_z^{-1}\boldsymbol{z}(t)\right)\left(\boldsymbol{\Sigma}_z^{-1}\boldsymbol{z}(t)\right)^T = \boldsymbol{\Sigma}_z^{-2}\boldsymbol{\Sigma}_z^2 = \boldsymbol{I} \tag{5-31}$$

Moreover, according to the definition, the mode shape are a set of eigenvectors that yields,

$$\boldsymbol{\Phi}^H\boldsymbol{\Phi} = \boldsymbol{I} \tag{5-32}$$

Therefore, if SVD acts on $\boldsymbol{u}(t)$,

$$\boldsymbol{u}(t) = \boldsymbol{\Phi}\boldsymbol{\Sigma}_z(\boldsymbol{\Sigma}_z^{-1}\boldsymbol{z}(t)) + \boldsymbol{\eta}(t) \tag{5-33}$$

The components of decomposition are corresponding to mode shape $\boldsymbol{\Phi}$, normalized modal signal $\boldsymbol{\Sigma}_z^{-1}\boldsymbol{z}(t)$ and the square root of covariance matrix $\boldsymbol{\Sigma}_z$ respectively, but the noise $\boldsymbol{\eta}(t)$ will influence the estimate accuracy.

### 5.2.2 Robustness and applicability

The SVD-based method does not need any prior knowledge of mode shape and the iteration process for optimizing. Hence, its applicable to discrete system and uncertain boundary conditions.

However, the noise resistance of eigenvalue-based method is lower relatively, when the signal power is smaller than noise, the modal signal cannot be extracted. In addition, if the measured mode shape is not orthogonal to each other, for instance, the twist and bending mode from 2D perspective. In this case, the mode shape obtained from SVD will bias against actual real mode shape.

## 5.3 Discussion

The SVD-based method and adaptive spatial filtering method are proposed, the distinction of two methods is summarized as following table

Table 5-1 Comparison between SVD and ASF method

| | **Singular Value Decomposition** | **Adaptive Spatial Filtering** |
|---|---|---|
| **Control** | Open loop | Close loop |
| **Orthogonality** | Required to mode shape | Not required |
| **Prior knowledge** | Not required | Prior knowledge to represent mode shape |
| **Signal band** | Narrowband | Not necessary |
| **Computational Efficiency** | Low in high-dimension matrix | Low in high-dimension weight vector |

Overall, these two methods are suitable for different situations. The SVD-based method needs higher SNR but the result without bias if mode shape is orthogonal. In comparison, the adaptive spatial filtering needs parametric representation of mode shape in advance and a close-loop optimization.

## 5.4 Key findings

To estimate the mode shape from pixel array, two classes of methods are proposed in this section, i.e., the adaptive filtering method and SVD based method and then their pros and cons are compared. In summary, the adaptive filtering can obtain higher SNR in exchange of the known elementary shape of mode shape and an iterative process. In contrast, the SVD-based method does not rely on prior knowledge and iteration, but it is more prone to noise.

In addition, it is proved that the proposed adaptive filter is least mean square filter, and the objective function is equivalent to modal assurance criterion, which are crucial properties to developing more efficient algorithm in PASP framework.

# 6 MODELLING AND SIMULATION STUDY

*The numerical simulation is studied to verify the proposed approaches which contains the finite element analysis and synthetic data analysis. Firstly, the finite element models (FEM) of test objects in experiment are built, which is used to predict the test result, especially simulate the cracks of wind turbine blade with different length. Secondly, a synthetic video is constructed in order to implement PASP on simulated data. Finally, the effectiveness of proposed methods is verified, especially the theoretical array gain (AG) and minimum detectable signal (MDS) have got proved.*

*It should be noted that some paragraphs in this chapter are expansion or reproduction of the published materials* [1] and [2] in list of publications. *For example*, [1] is expanded in *section 6.1.1 and 6.2, as well as* [2] is reproduced in *6.2.5.1. The specific paragraphs are annotated by a citation.*

**Highlight:**

- The FEM of wind turbine blades and a free-free beam are built to predict the result of experiment.
- The effectiveness of proposed method for modal identification is verified through simulation.
- The AG and MDS found in theory are verified in simulation.

## 6.1 Finite element analysis

### 6.1.1 Free-free beam

To predict the experiment, the finite element model (FEM) of a free-free beam is formulated by using a FEA software ANSYS. The procedures contain material setting, meshing, imposing boundary condition and result analysis. Table 6-1 shows the material's parameter of stainless steel used in simulation.

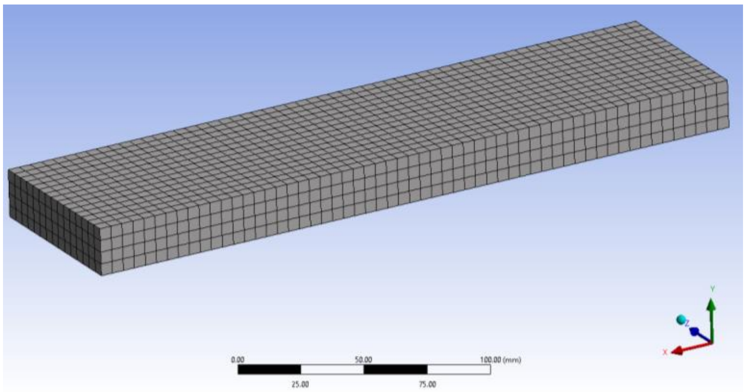


Figure 6-1 the imposed mesh of finite element model.

Figure 6-1 displays the mesh of model, which is important to the accuracy and computational cost. The Hex element is chosen from multiple trails, and the element length is 5mm. The boundary condition stays empty because the boundary condition of experiment is free-free. Figure 6-1 shows the simulated natural frequencies, which can used to compare with experimental result in section 7.1.

Table 6-1 properties of beam's material

**Stainless steel**

| | |
|---|---|
| Density | $7.75 \times 10^{-6}\ kg/mm^2$ |
| Young's modulus | $1.93 \times 10^5\ MPa$ |
| Poisson's ratio | 0.31 |
| Bulk modulus | $1.69 \times 10^5\ MPa$ |
| Shear modulus | $7.36 \times 10^4\ MPa$ |
| Isotropic secant coefficient of thermal expansion | $1.7 \times 10^{-5}\ \ 1/℃$ |
| Compressive ultimate strength | $0\ MPa$ |
| Compressive yield strength | $207\ MPa$ |

| | |
|---|---|
| Tensile ultimate strength | 586 $MPa$ |
| Tensile yield strength | 207 $Mpa$ |

The natural frequencies are listed Table 6-2, and its corresponding mode shapes are in Figure 6-2 [92].

Table 6-2 The FEA result of natural frequencies

| **Order** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|
| **Natural frequency (Hz)** | 1,252.7 | 2,570.9 | 3,362.4 | 5,255.2 | 6,345.7 |

These modes all take the form of standing waves, the blue areas represent nodal lines where the vibration is zero nearly.

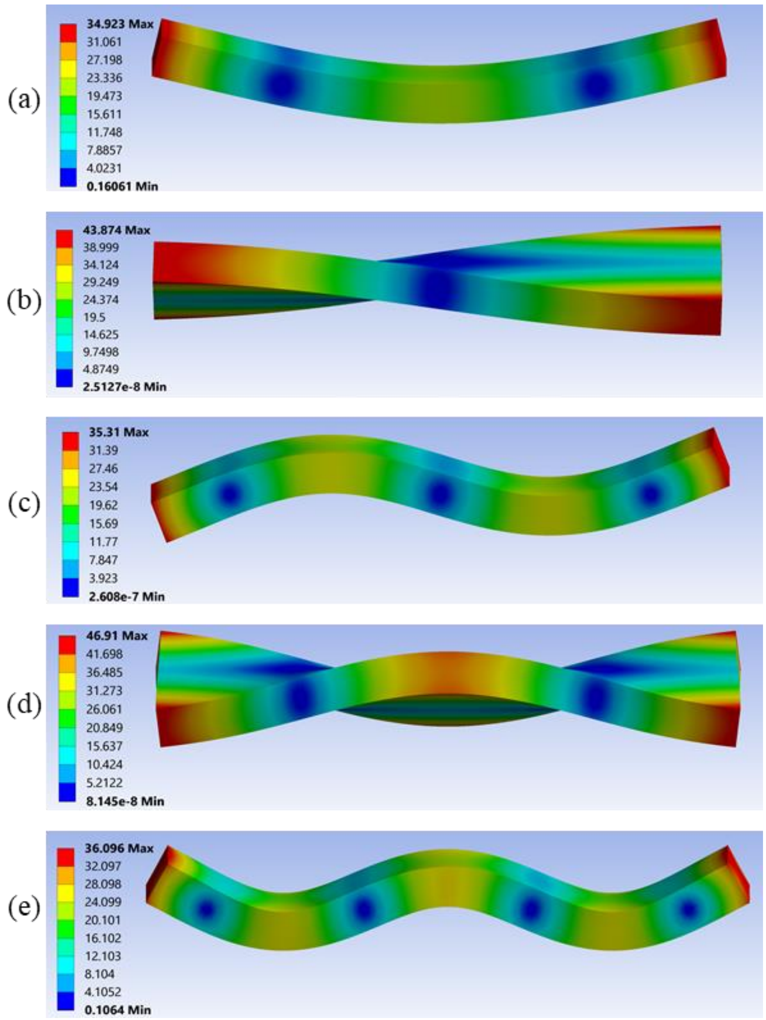



Figure 6-2 Some simulated mode shape, left hand side lists bending mode shapes, right hand side lists twisting mode shapes.

### 6.1.2 Wind turbine blade

For the test of wind turbine blade, several FEMs are formulated, in particular a crack is set on the model in order to simulate the damage on blade and analyse its influence on modal parameters.

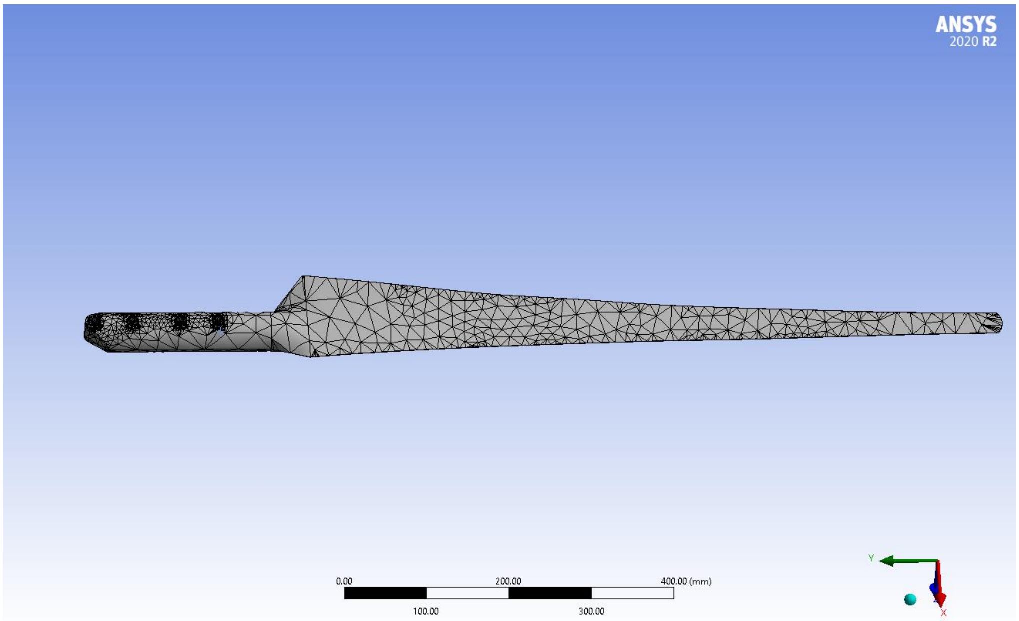



Figure 6-3 The imposed mesh of FEM of wind turbine blade

Due to the curved shape of blade, the shape of mesh is set as Tetrahedrons as shown in Figure 6-3. The boundary condition is a fixed support imposed on the axial and radial surfaces of screw holes, to simulate the experimental boundary condition as shown in Figure 6-4.

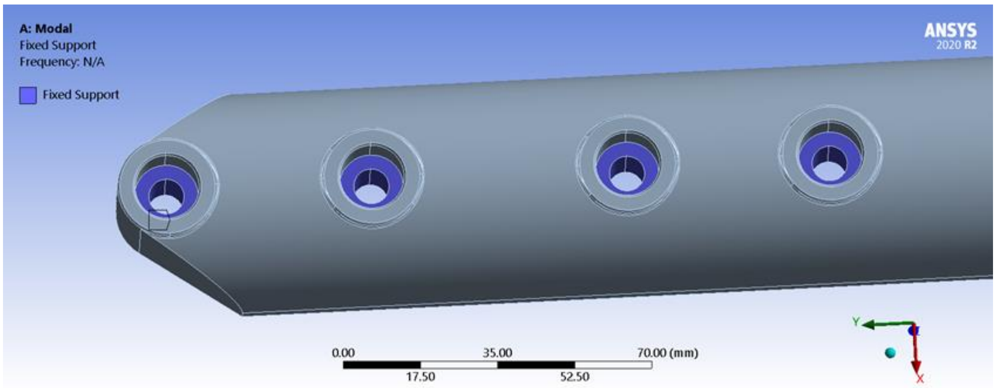



Figure 6-4 The fixed boundary condition is imposed on the area marked by blue color

The material is set as a composite material containing glass fibre, whose property is listed in .

Table 6-3.

Table 6-3 property of wind turbine blade's material

**Epoxy S-Glass UD**

| | |
|---|---|
| Density | $2.6 \times 10^{-6}\ kg/mm^2$ |
| Young's Modulus X direction | $5 \times 10^4\ MPa$ |
| Young's Modulus Y direction | 8,000 $MPa$ |
| Young's Modulus Z direction | 8,000 $MPa$ |
| Poisson's Ratio XY | 0.3 |
| Poisson's Ratio YZ | 0.4 |
| Poisson's Ratio XZ | 0.3 |
| Shear Modulus XY | 5,000 $MPa$ |
| Shear Modulus YZ | 3,846.2 $MPa$ |
| Shear Modulus XZ | 5,000 $MPa$ |

The result of simulation is listed in Table 6-4 and Figure 6-5 respectively. The root of wind turbine blade is fixed, where can be taken as the fixed end of cantilever beam as shown in Figure 6-5.

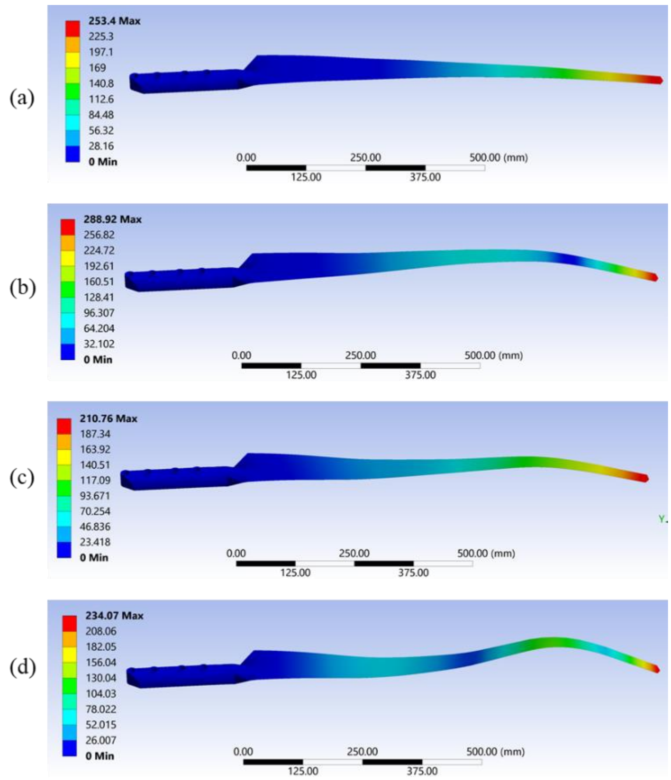


Figure 6-5 Mode shape under fix support at root, (a)first order, (b)second order, (c)third order, (d) forth order

Table 6-4 The FEA result of natural frequencies

| Order | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **Natural frequency (Hz)** | 10.8 | 29.6 | 53.2 | 65.7 |

### 6.1.3 Crack of wind turbine blade

Since modal parameters are important index to detect structural damage in structural health monitoring, different cracks are set on blade model to simulate different degrees of damage in this section as shown in Figure 6-6.

(a)

(b)

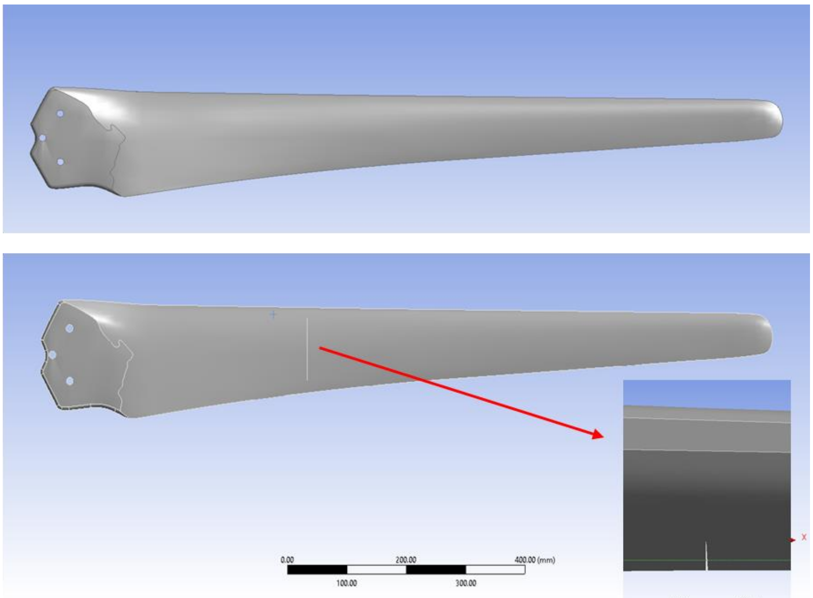



Figure 6-6 FEM of (a) the intact wind turbine blade and (b) simulated crack on the region with high damage risk

The crack was set about 30% away from the root as shown in the Figure 6-6, where is considered as a location prone to damage [40], [142], [143]. Then the depth of crack is changed successively to observe difference of mode shape. As the mode shape is a non-dimensional value, the result $\boldsymbol{\Phi}_i$ exported from ANSYS was normalized by its Euclidean norm $\|\boldsymbol{\Phi}_i\|$,

$$\bar{\boldsymbol{\Phi}}_i = \boldsymbol{\Phi}_i / \|\boldsymbol{\Phi}_i\| \tag{6-1}$$

The normalized mode shapes $\bar{\boldsymbol{\Phi}}_i$ with different depth of crack as shown in Figure 6-7, but the difference is too little to show in chart, thus that are amplified in two red windows.

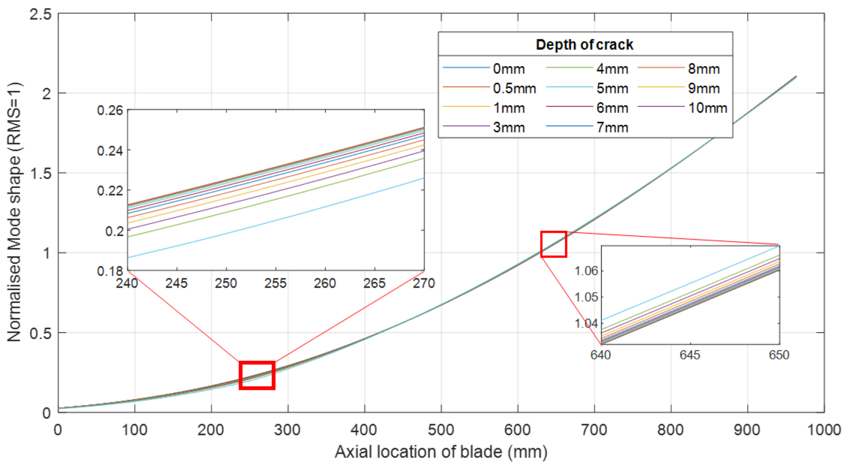



Figure 6-7 The normalized mode shapes with different depth of crack, the difference caused by crack is amplified in two red windows

$$\zeta_i = \frac{\overline{\boldsymbol{\Phi}}_i - \overline{\boldsymbol{\Phi}}_1}{\|\overline{\boldsymbol{\Phi}}_1\|} \times 100\% \tag{6-2}$$

Actually, the norm of exported mode shape is basically the same but with a little numeral error. Next, we used this equation to compare the degree of mode shape change.

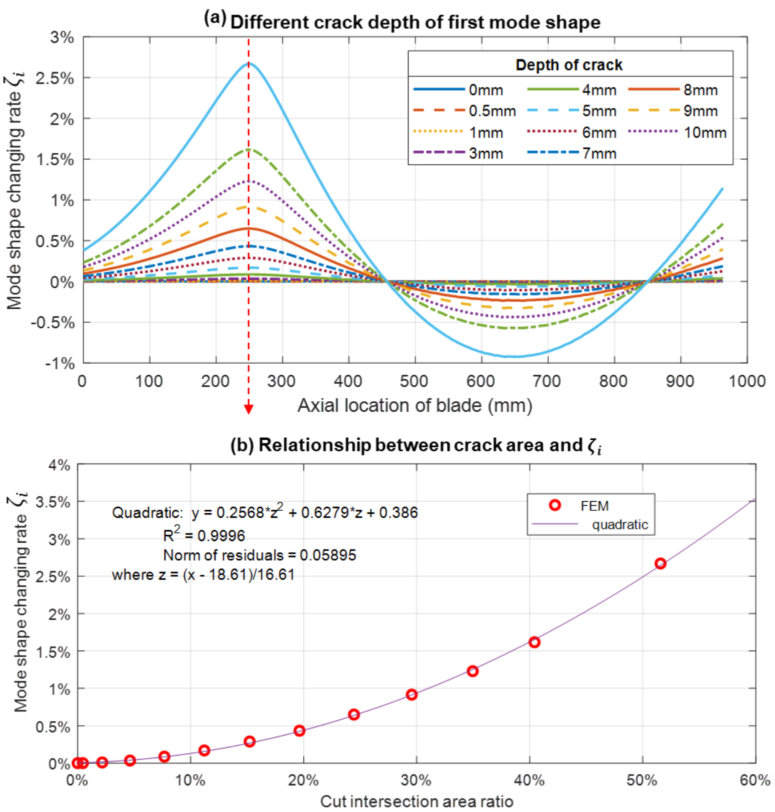



Figure 6-8 (a) The change of mode shape with different crack depth (b) The relationship between cut intersection area and the change of mode shape, which is well fitted by a quadratic function.

The result in Figure 6-8 (a) show a correlation between crack depth and change of mode shape. The lines represent different depth in centimetre from *0cm* to *12cm*. Furthermore, a quadratic function is used to fit this correlation as shown in Figure 6-8 (b), which shows the maximum variation (the position marked by a red dashed line in Figure 6-8 (a)) of mode shape is basically equal to the quadratic of the cross-sectional area caused by the crack.

The simulation results show the possibility of detecting cracks based on mode shape. However, the small crack cannot change mode shape distinguishably. For example, a 0.5cm crack changes the mode shape less than one thousandth based on the relationship shown in Figure 6-8 (b). However, the stress around crack will be doubled by this small crack, because even small crack can have a significant impact on the stress distribution of structure. In this sense, the vision-based method still faces significant challenges on damage detection.

## 6.2 Modal identification using simulated signal

### 6.2.1 Setup of synthetic footages

To verify the proposed method, a simulated dataset is obtained by a synthetic video footage in MATLAB [92]. Figure 6-9 shows a reference image where the black-white stripe pattern is painted on one face of a free-free beam. The image is blurred by a $5 \times 5$ PSF kernel and the region of interest has $25 \times 1000$ pixels. The colour depth is set as 8-bits and the sampling rate is set as 1,500 fps. In addition, the resulted images are further blurred by a noise floor of 5.66 dB white noise.

Figure 6-9 One representative frame in synthetic video footages

The oscillation of the pixel intensity was simulated in accordance with the response of dynamic system, in which four damping signals were set to simulate the response of free-free beam under impact excitation. The characteristics are detailed in Table 6-5, in which the rigid body mode at 3 Hz with amplitude of 1/2 pixel is also included to simulate real test sensoria when the beam lay on a soft base. The amplitudes of other three modes were set lower by a factor of ten, corresponding to a SNR that varies by 20 dB, which allows the performance of proposed algorithm to be evaluated sufficiently under different SNRs [126] so that it can be reliable applied to process images captured by high speed camaras.

| Modes | Rigid | $1^{st}$ | $2^{nd}$ | $3^{rd}$ |
|---|---|---|---|---|
| Frequency | 3 Hz | 73 Hz | 277 Hz | 457 Hz |

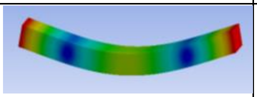
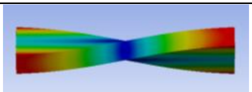
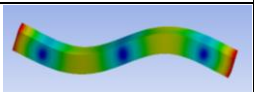
| Amplitude | 1/2 pixel | 10/256 pixel | 1/256 pixel | 0.1/256 pixel |
|---|---|---|---|---|
| SNR | 9.7 dB | -11.7 dB | -31.8 dB | -51.8 dB |
| Mode shape | Rigid body |  |  |  |

Table 6-5 Setting of dynamic characteristics

An array structure $A_s(x,y)$ is constructed by applying edge detection on reference image of synthetic video. As shown in Figure 6-10 (a), two white rows represent one value in $A_s(x,y)$ that are 2000 selected pixel sensors of array, data matrix $\boldsymbol{u}(t)$ is constructed to be of 2000×6000. Figure 6-10 (b) shows a typical signal of $\boldsymbol{u_i}(t)$ at pixel $i = 4$. As the pixel is located at the red rectangle in Figure 6-10(a), the selected pixel is likely to have the highest output as the three modes all produce the highest responses. However, the spectrum of the signal in Figure 6-10(b) shows just modal response at the 1st mode at 73Hz, but it is not possible to find the responses from two high orders of modes at 277Hz and 457Hz. This shows that the noise level of the raw signal is too high, making it difficult to resolve the two higher modes directly.

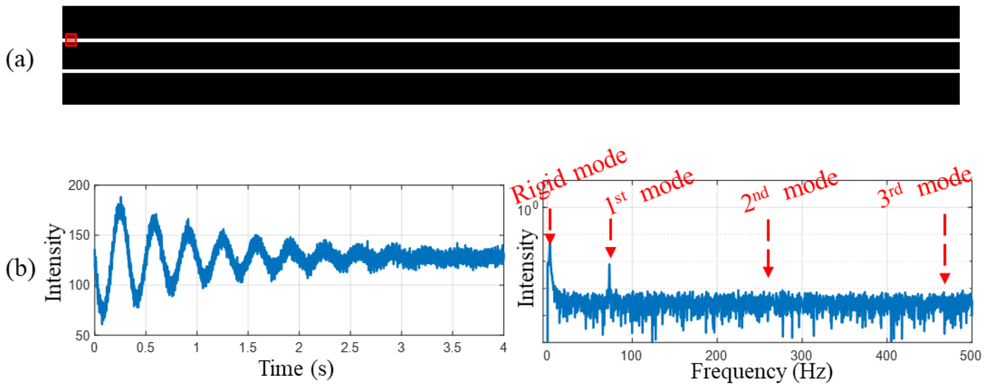



Figure 6-10 Information in simulated footages, includes (a) sensor array $A_s(x,y)$ obtained by edge detection (b) intensity signals of one array element and its spectrum

### 6.2.2 Adaptive spatial filtering

Performing the procedures described before, the results can be seen from Figure 6-11, every modal displacement $\hat{z}_i(t)$ is strengthened individually based on its corresponding weight vector. Furthermore, the power of modal displacement $P(w_i)$ is increasing when the weight vector is approaching to the pre-set mode shape. The convergence rate is related to the number of the nodes, and the error of mode shape is mainly affected by the noise level. For example, for the first and the second mode, the correct nodes can be always found with insignificant errors, that means the error of node location is under one pixel. For the third mode shape, the average error $Err(\boldsymbol{\varphi}_i)$ is up to 8.32% in ten sets of Monte Carlo test, but it still allows capturing the main profile features and therefore makes the amplitude at 457Hz being clearly visible. The estimated result shows smoothness is because the used sin-based function is smooth originally. The estimate error exists on the location of nodes, as well as on different curvature between estimated mode shape and real value.

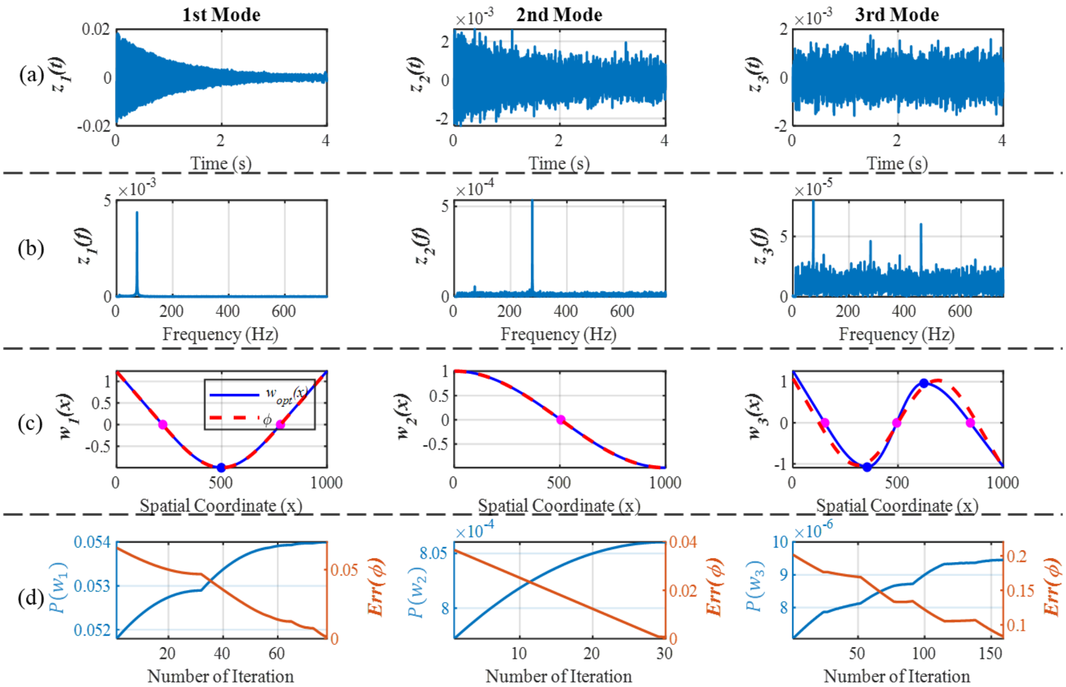
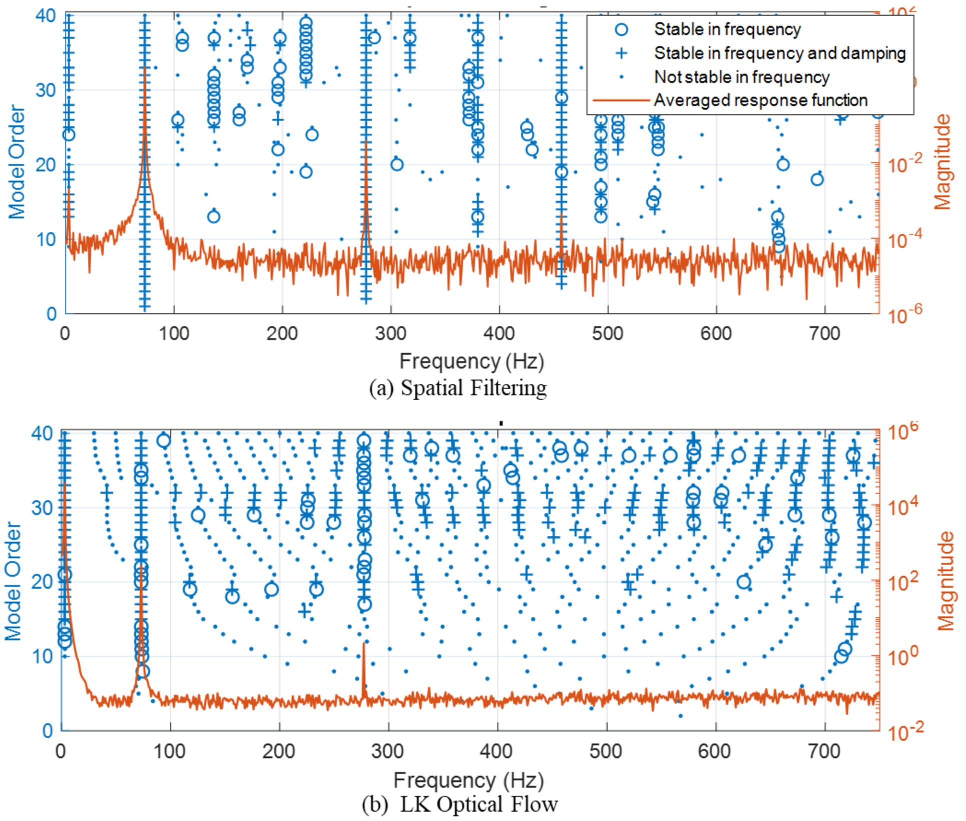



Figure 6-11 Three obtained modes using spatial filtering (a) output signal (b) spectrum of output signal (c) optimum weight vector $\boldsymbol{w}_{opt}$ and pre-set mode shape $\boldsymbol{\varphi}_i$ (d) signal power $P(w_i)$ and the error of estimated mode shape $Err(\boldsymbol{\varphi}_i)$ changes with iteration

LK optical flow as a typical window-based method is compared with spatial filtering using the same video footage. 20 windows with size of 10×10 pixels are selected uniformly on image.

### 6.2.3 Modal identification

The FRFs from spatial filtering and LK optical flow are used as the system response to identify the modal parameters using LSRF, respectively. Figure 6-12shows the stabilization diagrams, it can be observed that both methods can identify the first mode stably. However, when the amplitude decreases under 1/256 pixel in the second order, the stability of LK method declines greatly. In the third order, the LK method could not already identify the mode anymore, but spatial filtering could still get a good result.



(a) Spatial Filtering

(b) LK Optical Flow

Figure 6-12 Stabilization diagrams from results of (a) Spatial filtering and (b) LK optical flow, respectively

As legends shown in Figure 6-2, the ‘*o*’, ‘+’ and ‘·’ represent the stable results identified in different model order. If results are consistent at a frequency, that means the identified natural frequency is reliable [9].

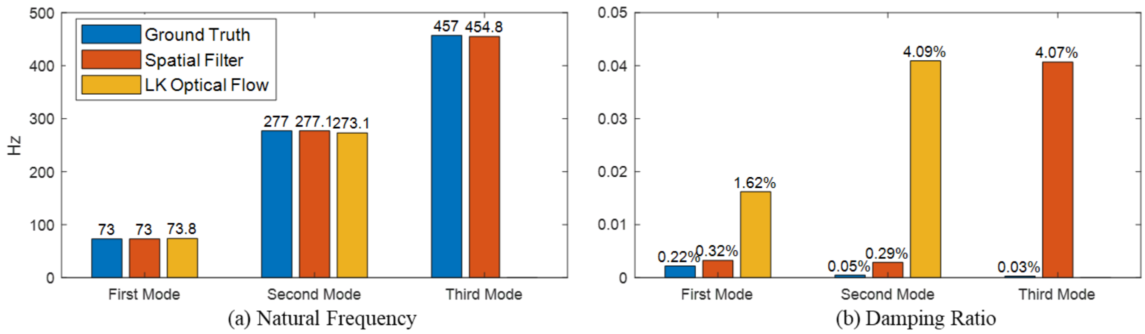



Figure 6-13 Natural frequency and damping ratio identified using Spatial filtering and LK optical flow

Figure 6-13 compares the identified natural frequencies and damping ratio. The signals from spatial filtering can successfully identify the third mode with a SNR of -51 dB, while LK optical flow cannot identify it. Moreover, the identification error of spatial filtering is smaller than that of LK optical flow method. Since video footage is the same and both results are identified by LSRF, the difference of them depends thoroughly on SNR of signals obtained by spatial filtering and LK optical flow.

### 6.2.4 Robustness of mode shape estimation

To verify the robustness of the proposed method, it was applied to match different mode shapes which were created by setting the beam with different boundary conditions with different elasticities of end supports. The initial value $\boldsymbol{w}_{ini}$ is still the same in three cases, however, the optimizer $\boldsymbol{w}_{opt}$ can approach to the different mode shapes $\boldsymbol{\varphi}$ through iterations, as shown by the first two modes in Figure 6-14. Although when SNR is too low in the third mode (-51.8 dB), the error of mode shape estimate turns visible. It demonstrates that the algorithm relies on the prior knowledge of solution forms i.e., standing waves, but the specific parameters of solutions are determined by measuring the change in signal power. In this numerical way, the profile of mode can be approximated even on complex continuous structures.

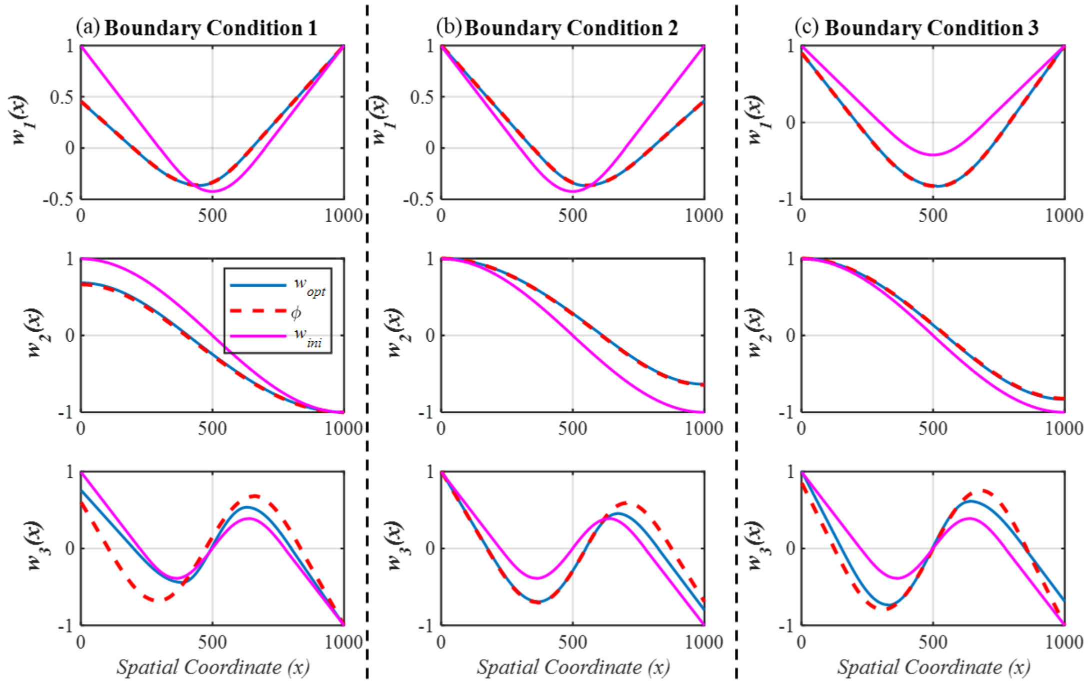



Figure 6-14 $\boldsymbol{w}_{opt}$ estimation under different boudary conditions (a), (b) and (c)

### 6.2.5 Verification of signal quality

#### 6.2.5.1 Quantization noise

In image noise, one component is the quantization error, which can be either periodic or random [126]. To observe the quantization error, an intensity variation caused by a displacement with two frequency components, is calculated from Eq. (7),

$$I(t) = F(a_1 \cos(2\pi f_1 t) + a_2 \cos(2\pi f_2 t) + n(t)) \tag{6-3}$$

in which, $f_1 = 3Hz, f_2 = 457Hz,\ a_1 = 1pixel,\ a_2 = 1/256\, pixel$. The Figure 6-15 (a) is the pure signal without noise, the quantization errors are regularly distributed over multiples of 3Hz, making the original signal difficult to identify. Because it is not evenly distributed in the frequency domain, it cannot be considered as noise for average denoising. Then, the noise is gradually added to the signal, the error gradually changes from the modulation peaks to the uniformly distributed noise in the frequency domain.

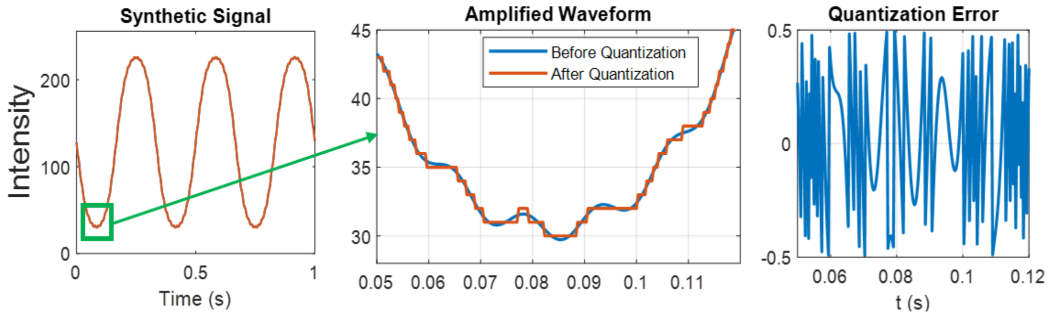



Figure 6-15 Schematic of quantisation error

The upper diagrams are waveforms, the local of signal is amplified to present the quantized signal and quantization errors. Figure 6-16 shows the signal spectra with 0dB, -80 dB and -60 dB noise added, respectively.

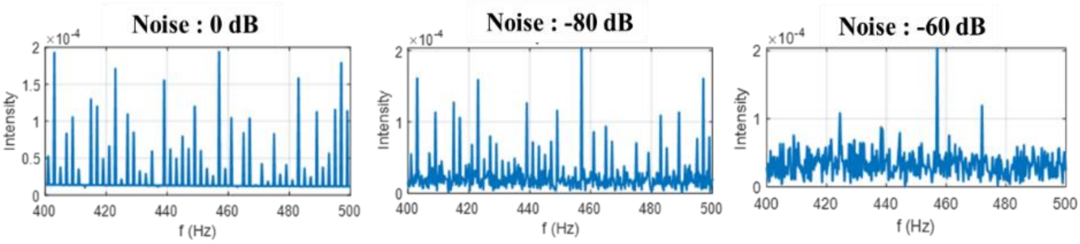



Figure 6-16 The testing signals with different level of noise.

From the time domain diagram of the quantization error (middle diagrams), it can be seen that the error amplitude always falls in the range between 0 and 0.5, and the amplitude is basically the same, but the signal with random term is easier to resolve, and also easier to process.

Likewise, the damped sine wave is compared with the standard sine wave in Figure 6-17 since it is a common type of signal in modal analysis. It turns out that the damped amplitude has an equivalent effect to noise dither. Because the variation of amplitude increases the randomness of error.

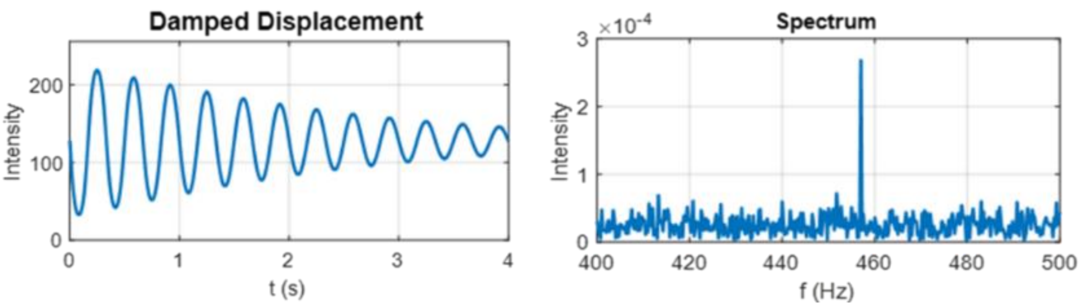



Figure 6-17 The test using undamped displacement and damped displacement.

In summary, the SQNR of each set of signals above is roughly equal to the theoretical value 49.9 dB obtained with the difference just being the randomness. In practice, the periodic error cannot occur generally since the random term always exists. Thus, the quantization noise can be taken as a gaussian noise, whose power is merely related to Q.

#### 6.2.5.2 Verification of array gain

Another simulation is implemented to examine array gain $G_a$ of Eq. (4-23) under different number of pixels $m$ [92]. Figure 6-18 shows $G_a$ increase trend as the number of array elements increase from 400 to 4000 with a step of 400, which was obtained when the input noise is 5.6dB and other parameters remain unchanged for the synthetic datasets underlying.

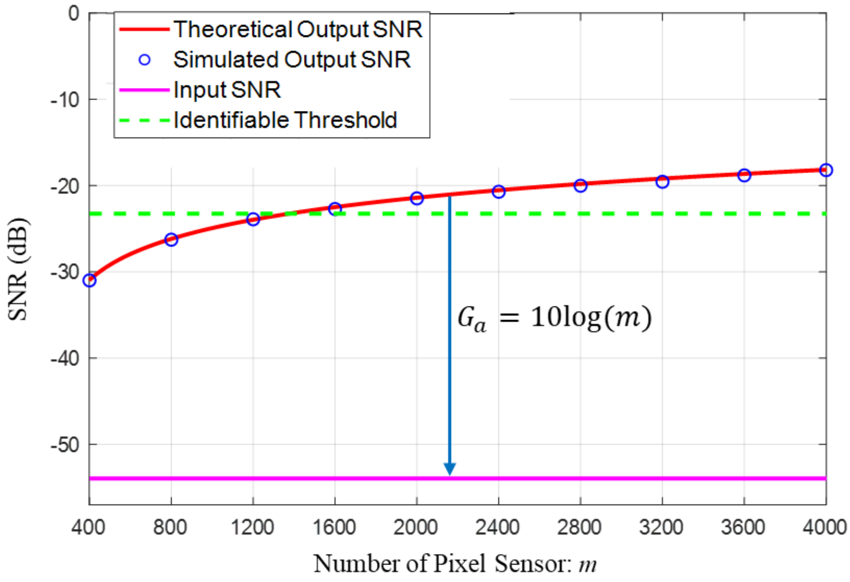



Figure 6-18 SNR gain of spatial filtering with respect to number of sensor elements

According to Eq. (4-23) , the distance between magenta and red line is exactly the array gain $G_a$. That shows a great consistency between theory and simulation. It demonstrates that increasing array elements is an effective way to improve output SNR. It also explains how the third mode of 1/2560 is extracted, because signal is enhanced by about 2000 times that results in its SNR excesses the identifiable threshold. In the synthetic image, ROI has a total of 25 rows of pixels, but only 2 rows at the centre of edge are utilised. In practice, without changing any other shooting conditions, just spraying more high-contrast patterns can add more array elements effectively.

#### 6.2.5.3 Verification of Minimum Detectable Signal

Figure 6-19 shows a synthetic image for image-based modal analysis, which is constructed by white and black stripes while contaminated by gaussian noises. Suppose the required SNR is -20dB, so far, all the items in MDS have been determined. The MDS obtained by simulation is compared with the theoretical value. The number of channels is adopted from 200 pixels to 1000 pixels, the color depth is considered to be 8 bits, 12 bits and 16 bits.

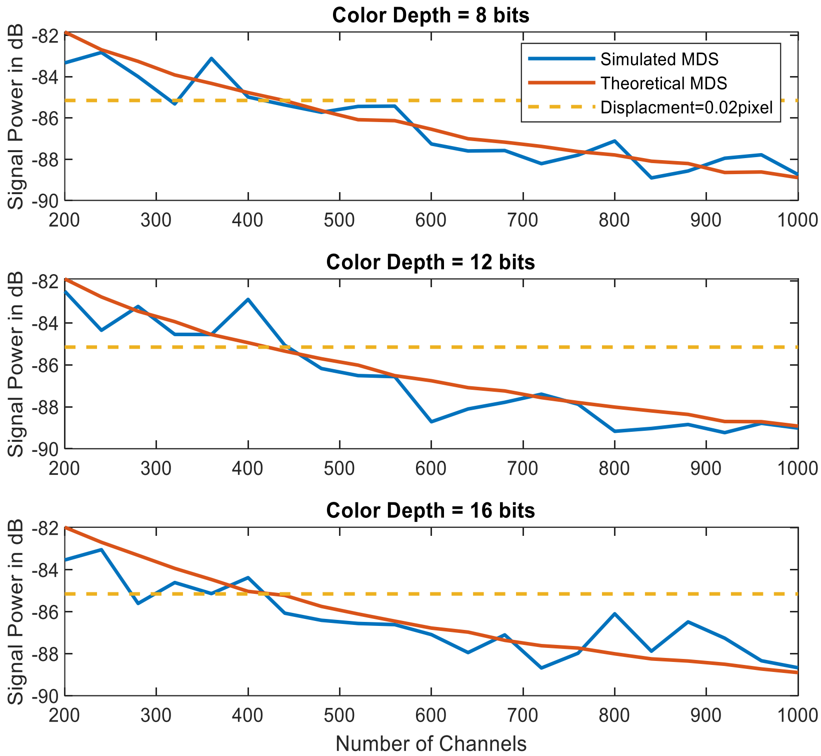



Figure 6-19 Input-output curve of an equivalent intensity-displacement sensor.

It can be observed from Figure 6-19 that the simulation results are basically consistent with the theoretical ones. The yellow line is the power of a standard sinusoidal signal with an amplitude of 0.02 pixels, which can be effectively detected with 500 pixels or more being used. Secondly, the color depth has little influence on the results, because the quantization noise accounts for a very small pro-portion compared to the overall noise. The simulation results show that the MDS estimation is reliable, and MDS can be predicted before the experiment and improved by adjusting the parameters.

## 6.3 Key findings

This section mainly covers simulation works including finite element model and signal processing using synthetic video data. Specifically, the FEM of wind turbine blades, and a free-free beam are built to predict the result of experiment. The result shows change of modal parameters caused by crack is not significant, that is very challenging to detect crack from mode shape. On the other hand, the effectiveness of proposed method for modal identification is verified through simulation, and the obtained AG and MDS is verified through simulation.

# 7 APPLICATION OF VISION-BASED VIBRATION MEASUREMENT TO MODAL ANALYSIS

*Two experiments of vision-based modal analysis are carried out to validate proposed methods in this section. First is a free-free beam, another is wind turbine blade, mode of both which are identified by a hammer test and captured by a high-speed camera. Afterwards, the proposed PASP methods are compared with a typical window-based method, LK optical flow. The beam test shows the proposed adaptive spatial filtering method has significant advantage on output signal SNR. The wind turbine blade test shows the SVD-based method can obtain accurate mode shape, however, the difference caused by small crack is too subtle to detect.*

*It should be noted that some paragraphs in this chapter are expansion or reproduction of the published materials* [1], [4] and [5] in list of publications*, such as* [1] in *section 7.1 and* [4] & [5] in *7.2.1, and the specific paragraphs are annotated in a form of citation.*

**Highlight:**

- Compared with LK optic flow (a typical conventional window-based method), the proposed adaptive spatial filtering method has higher performance on SNR gain.
- SVD-based method can obtain mode shape of wind turbine blades, which is more accurate than conventional LK optic flow.
- The change caused by crack cannot be detected by vision-based method when crack is very small.

## 7.1 Modal analysis of free-free beam

### 7.1.1 Experimental Setup

To further verify the proposed method, an experimental study was carried out, which measures the vibration responses of a fee-free rectangular beam with the dimension of $284 \times 20 \times 75$mm and stainless-steel [92]. The layout of the test setup is shown in Figure 7-1. Two flexible porous material blocks are used to support the beam, avoiding too large amplitudes of the rigid mode. An impact hammer is used to excite the free-free beam with two white strips as illustrated in Figure 7-1 (b). A Phantom (SN24006) high-speed camera was equipped with Nikon 35mm 1:1.8G lens, and the sampling rate was 14,000 fps and the resolution was $96 \times 1024$ (height×width) for multiple tests. An LED light provides a high luminous intensity and a uniform lighting. In the meantime, the acceleration signal was also collected by a piezoelectric accelerometer for benchmarking with the results with camera-based measurements. In addition, a finite element model (FEM) was built in ANASYS software to mainly validate the mode shapes obtained by the proposed method.

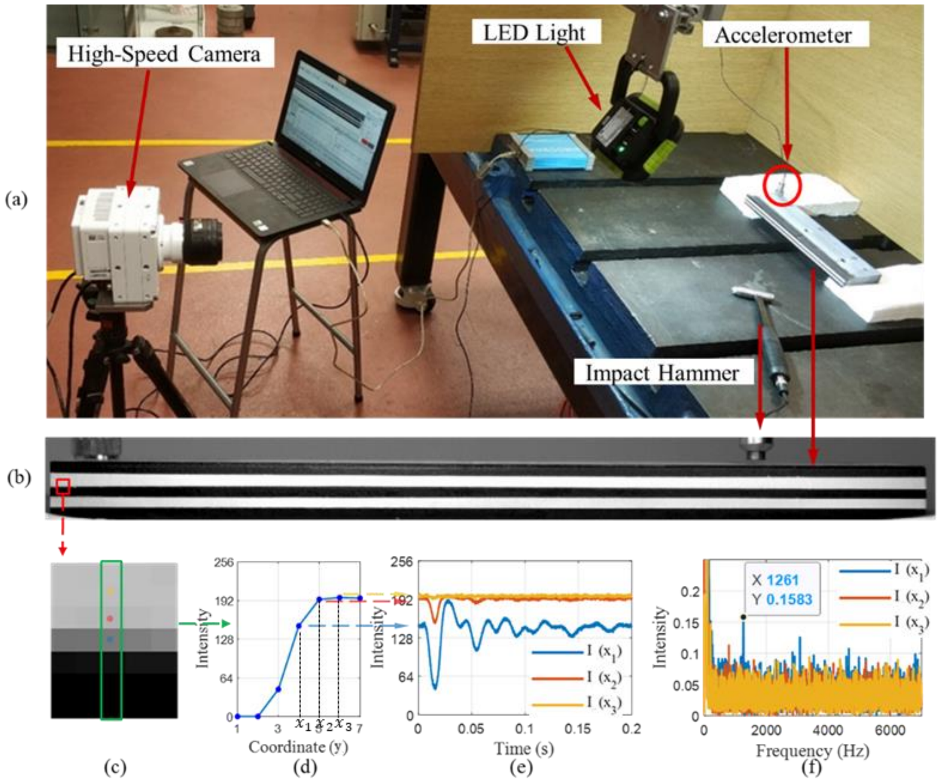



Figure 7-1 Experiment data, (a) setup of experiment (b) one frame in experimental video, (c) blurred edge in red rectangle, (d) one column of blurred image, (e) intensity signals of three pixels in that column (f) spectrum

Figure 7-1 (c) shows the blurring effect by amplifying a small region of image in Figure 7-1 (b), which validates the analysis made based the simulated image. The intensity profile Figure 7-1 (d) also agrees with the predicted $F(x)$. Besides, Figure 7-1 (e) shows the intensity at pixel dotted in Figure 7-1 (d), which is analogous to the distortion of simulated signal. Figure 7-1 (f) shows the spectrums from the outputs of three pixels, except for first mode at around 1261 Hz, all other modes are submerged in noise floor.

Applying the edge detection to the reference image of Figure 7-1 (b), according to analysis in section 3.4, 4000 pixels are selected to formulate the sensor array. As a result, a data matrix of 4000×3000 is constructed as $\boldsymbol{u}(t)$, meaning that the temporal sample points are 3000 which covers the entire transient responses caused by the impact.

### 7.1.2 Implementation of Nodes/antinodes Searching Scheme

Figure 7-2 shows four groups of weight vectors generated. In each group, the differences are generated by changing one parameter of the expression, which is one of parameters to be optimized. Once the node location is fed into the programmed function, the corresponding weight vectors will be produced as shown in Figure 7-2.

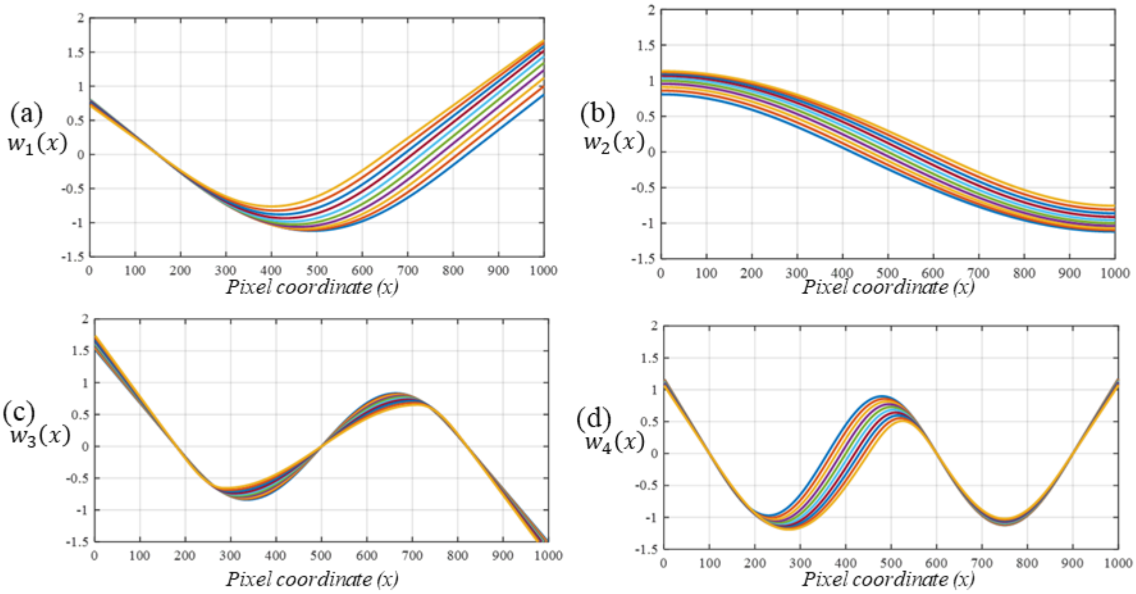



Figure 7-2 A schematic of weight vectors that change continuously, from (a) to (d), the node number of $w_i$ is different. The set of lines in each figure represents a processing of that the piecewise function changes with moving node location.

The operation of spatial filtering is $\hat{\boldsymbol{z}}(t) = \boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{u}(t)$ . Ten signals in $\boldsymbol{u}(t)$ with $1000\ pixels \times 3000\ time\ sequence$ are shown in Figure 7-3.

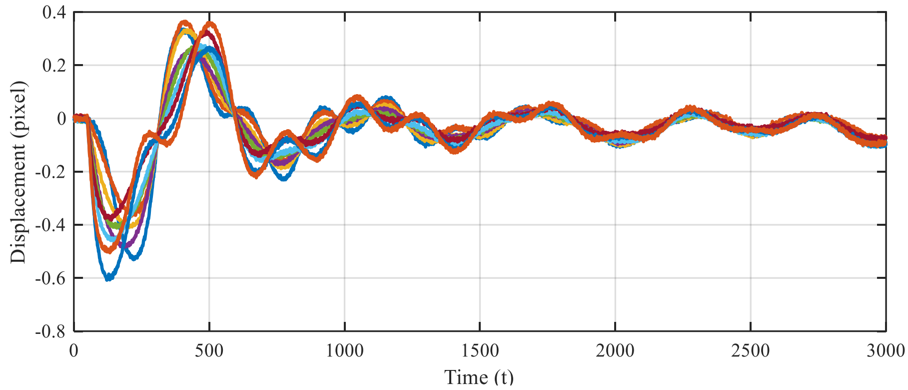


Figure 7-3 Ten examples in input signal $\boldsymbol{u}(t)$ matrix of experimental footages

The power spectral density (PSD) of output signal $\hat{\boldsymbol{z}}(t)$ are shown in subplot (c) from Figure 7-4 to Figure 7-9 respectively. The initial nodes are set arbitrarily at 170, 425, 800 in pixel coordinate, which are optimized one by one. The default searching direction is set arbitrarily towards left. First, move first node to 169, the output signal power shows decline. That indicates this change enlarges error of mode shape estimate, so that the moving direction is adjusted to the right (Figure 7-4 (a)). Then the first node of weight vector goes back to 170 and continue to move along right direction until the signal power declines again at location 257, that indicates the last location 256 can enable signal experience strongest constructive interference (Figure 7-4 (b)). In this process, the signal power increases smoothly from 1.04e-4 to 1.25e-4 (Figure 7-4). This smooth curve accords with theoretical prediction of $P(\boldsymbol{w}_i) = ({\boldsymbol{w}_i}^T \boldsymbol{\varphi}_i)^2 {\sigma_{z_i}}^2$, because when $\boldsymbol{w}_i$ is set as a unit vector, $P(\boldsymbol{w}_i)$ should be a cosine function of included angle between $\boldsymbol{w}_i$ and $\boldsymbol{\varphi}_i$.

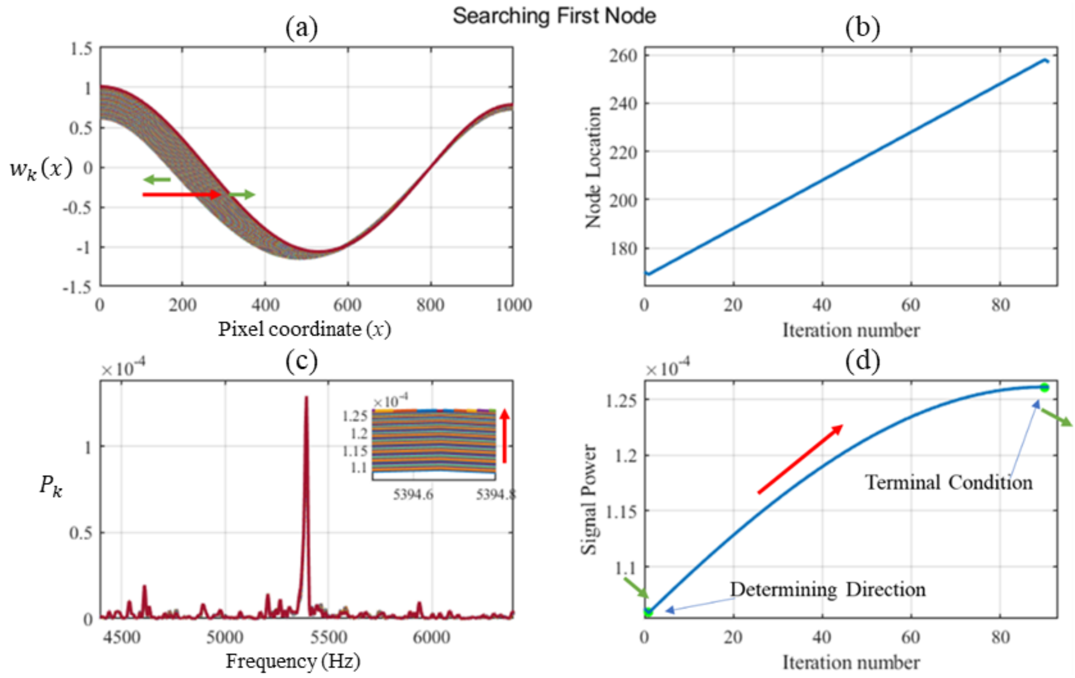



Figure 7-4 First searching step: (a) Weight vector, (b) node location changes with iterations, (c) PSD of output signal and (d) peak of PSD at natural frequency changes with iterations

Then, the program searches the second node location. Because the default direction makes signal power increase in first iteration, the following steps keep moving in this direction until termination condition occurs, i.e., when the signal power declines.

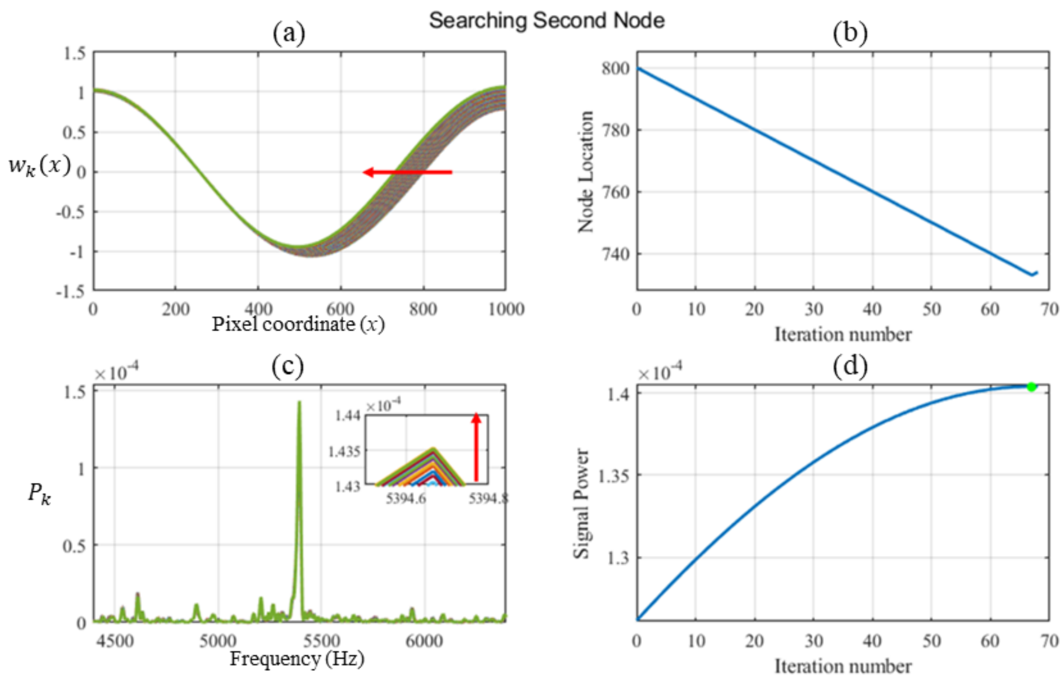



Figure 7-5 Second searching step: (a) weight vector (b) Node location changes with iterations (c) PSD of output signal (d) Peak of PSD at natural frequency changes with iterations

On basis of the obtained location of two nodes, the antinodes are searched in the same way.

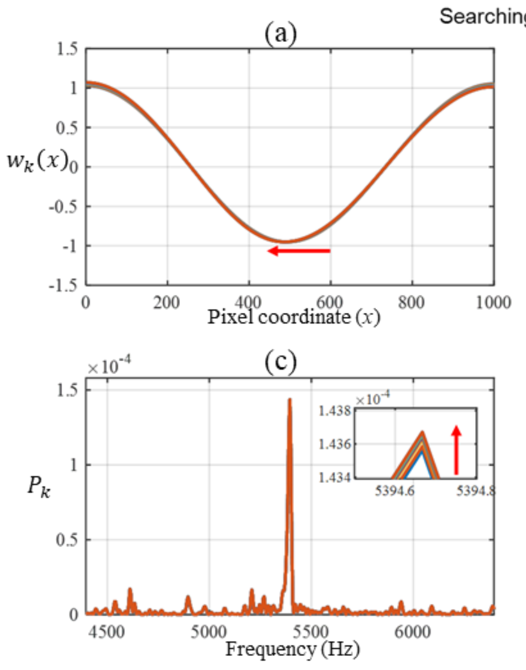



Figure 7-6 Third searching step: (a) weight vector (b) Node location changes with iterations (c) PSD of output signal (d) Peak of PSD at natural frequency changes with iterations

In order to test the robustness, another initial weight vector is set as 350, 513, 675 to execute the same program. At this run, first node location is optimized to 262, which is slightly different from 256. This phenomenon is caused by noise influence. However, the results of two tests can converge in a very close range, no matter the initial location is on the left or the right side of optimized result. Moreover, the required moving directions are different in two tests, however, a very consistent result shows up. That illustrates the initial value has not influenced on the result, instead the noise level will affect the accuracy more.

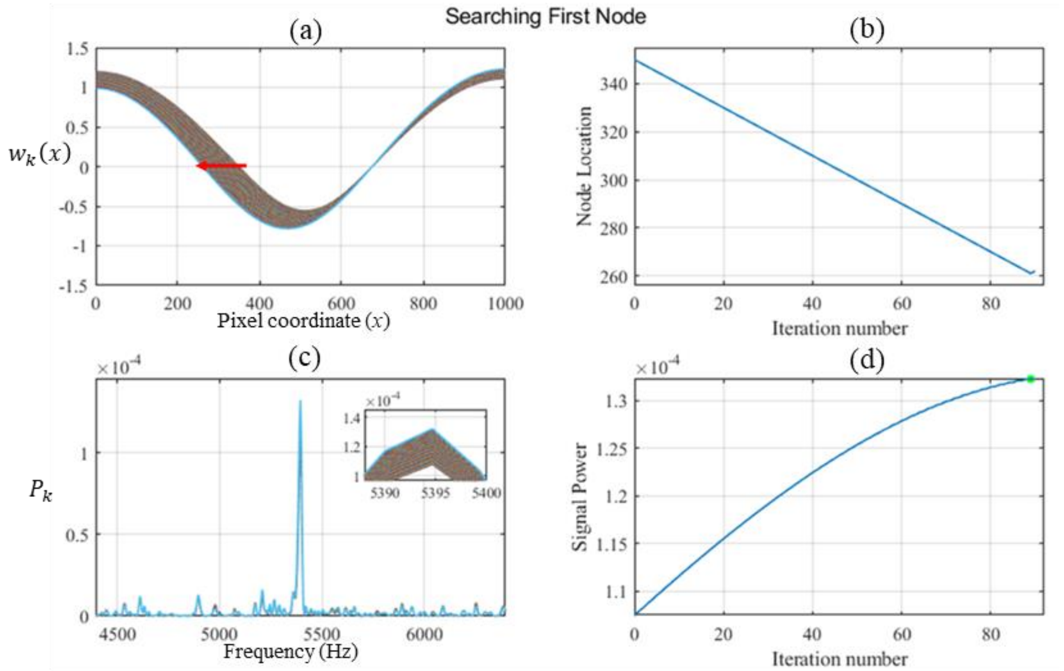



Figure 7-7 First searching step: (a) weight vector (b) Node location changes with iterations (c) PSD of output signal (d) Peak of PSD at natural frequency changes with iterations

Correspondingly, the second node is optimized to 733. Even the default searching direction (left) is not the wanted one (right), the signal power can adjust the searching direction adaptively by using the first iteration result.

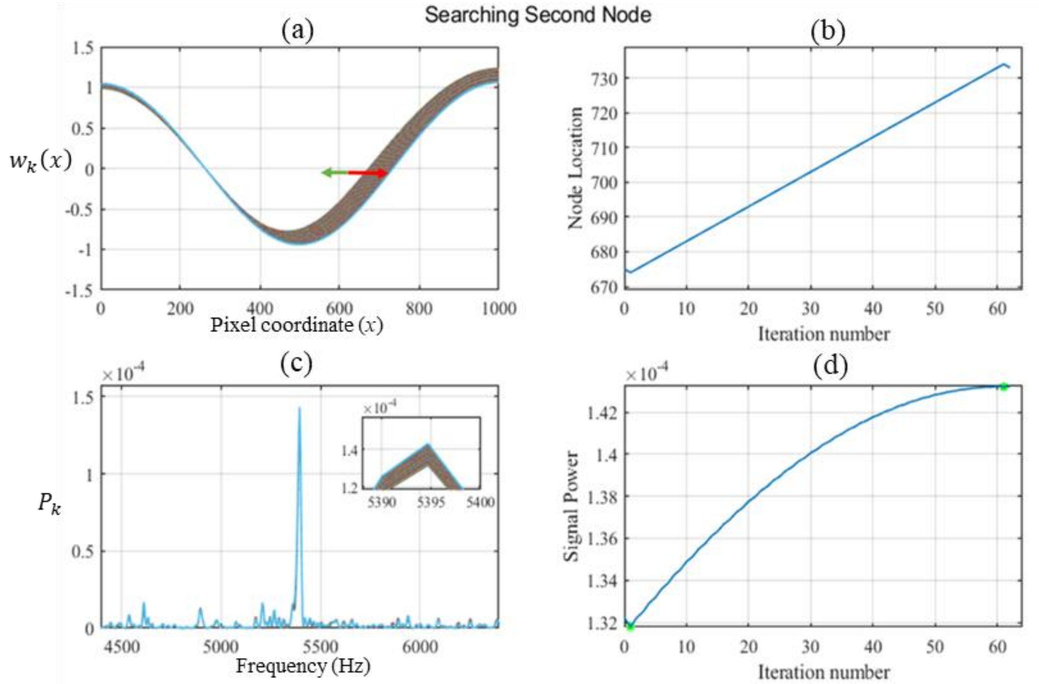



Figure 7-8 Second searching step: (a) weight vector (b) Node location changes with iterations (c) PSD of output signal (d) Peak of PSD at natural frequency changes with iterations

Then the antinode can be obtained in the same way.

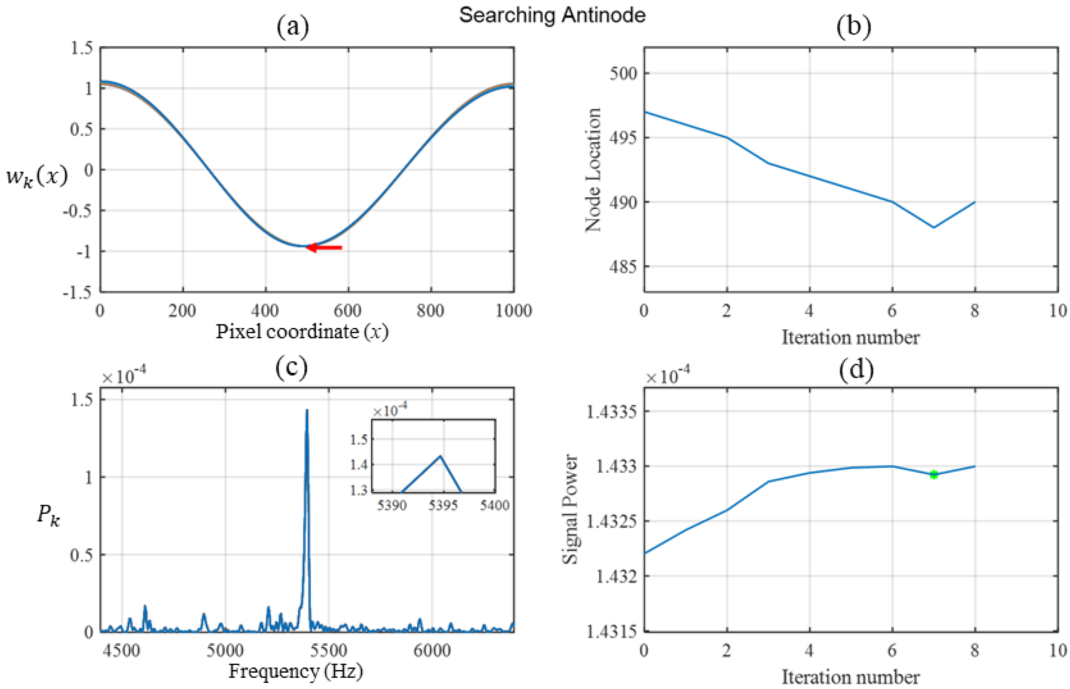



Figure 7-9 Third searching step: (a) weight vector (b) Node location changes with iterations (c) PSD of output signal (d) Peak of PSD at natural frequency changes with iterations

For the purpose of comparison, FEM is obtained from ANSYS, whose node location are at 242, 500, 757 in a coordinate from one to one thousand.

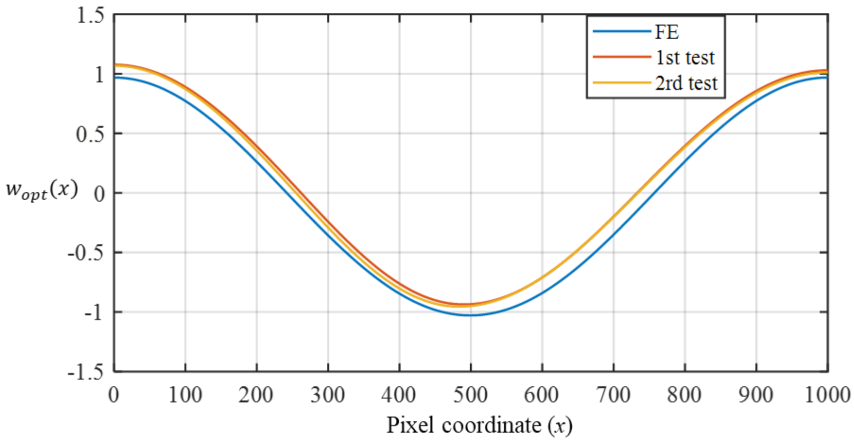



Figure 7-10 Comparison of twice test results and FEM

Considering that the FE model also has error with measurement, and the optimized result can converge steadily from different initial values, this result is reliable and acceptable. All other modes are found by going through the same program with just modifying the input node number.

### 7.1.3 Result Analysis

Comparing the proposed method with FEM shows the estimated $\boldsymbol{w}_{opt}$ and $\hat{\boldsymbol{z}}_i(t)$ generally match the prediction in FEM. The beam can show both bending modes and the twist modes when hitting on a conner of the beam. These modes can be observed by the 3D mode shapes shown in Figure 7-11. However, when shooting one face (2D perspective of the beam it cannot fully separate the twisting (the 2nd and 4th) modes from bending (1st, 3$^{rd}$, and 5th) modes. As a result, mode mixing can occur, especially, the modes with stronger response in low frequency range will be more likely to appear in the high frequency modes. As shown in Figure 7-11, the 2$^{nd}$ mode (twisting) appears in the 3$^{rd}$ and 5$^{th}$ mode (bending), and the 1$^{st}$ mode (bending) shows up in the 4$^{th}$ mode (twisting). However, such model mixing can be easily recognised and separated out as their frequency values are the same across different modes. Furthermore, multiple tests have shown that $\boldsymbol{w}_{opt}$ always converges in a small range, and the selection of initial weight vector does not affect the final result.

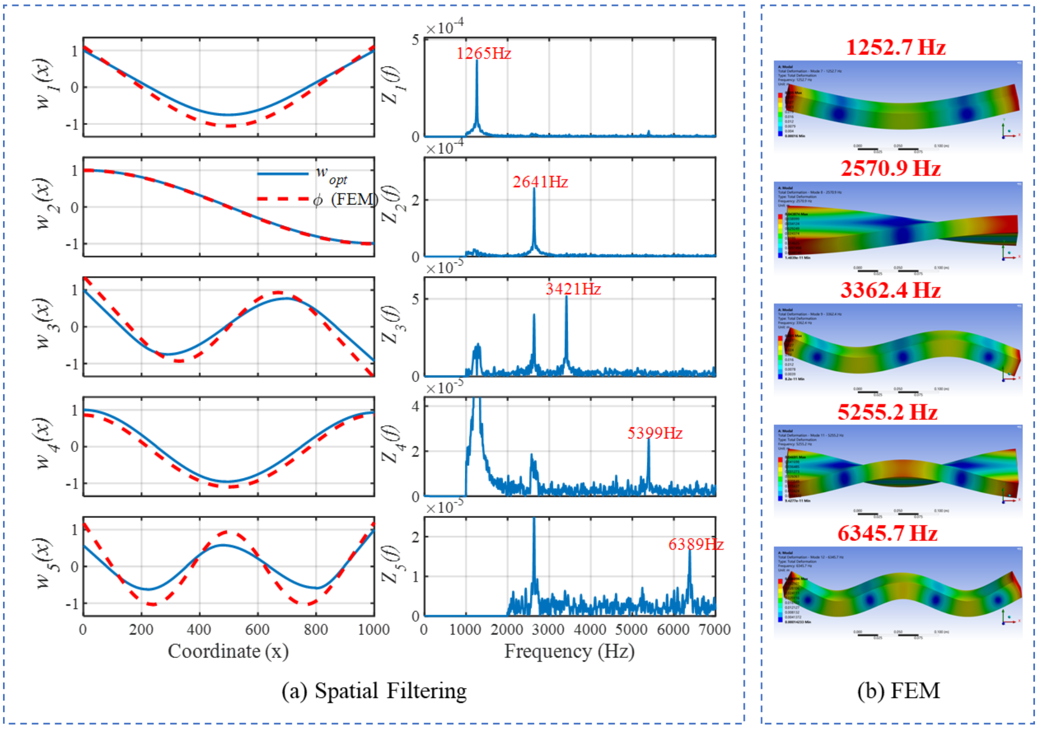



Figure 7-11 Comparison of vibration mode identification result from (a) spatial filtering and (b) FEM

In addition, the resulted displacement spectrum is compared with the LK optical flow and the result of accelerometer in Figure 7-12.The spectrum of spatial filtering is a sum of five outputs in $\widehat{\boldsymbol{Z}}(f)$. It can be seen that the peaks of all five modes are detectable in spatial filtering, however, LK optical flow cannot detect the fourth and fifth modes due to its limited noise suppression capability.

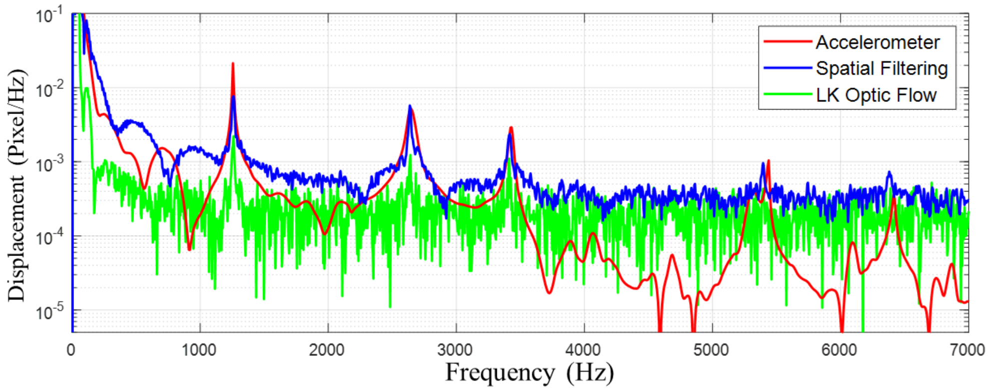



Figure 7-12 Comparison of spectra obtained from accelerometer, spatial filtering and LK optical flow.

The LSRF [144] is still used to identify the modal parameters. The results from accelerometer, spatial filtering, LK optical flow and FEM are compared in Figure 7-13. The LK optical flow merely detect first three modes and miss last two modes at high frequency. By contrast, the spatial filtering detected all five modes in frequency range and highly consistent with the acceleration.

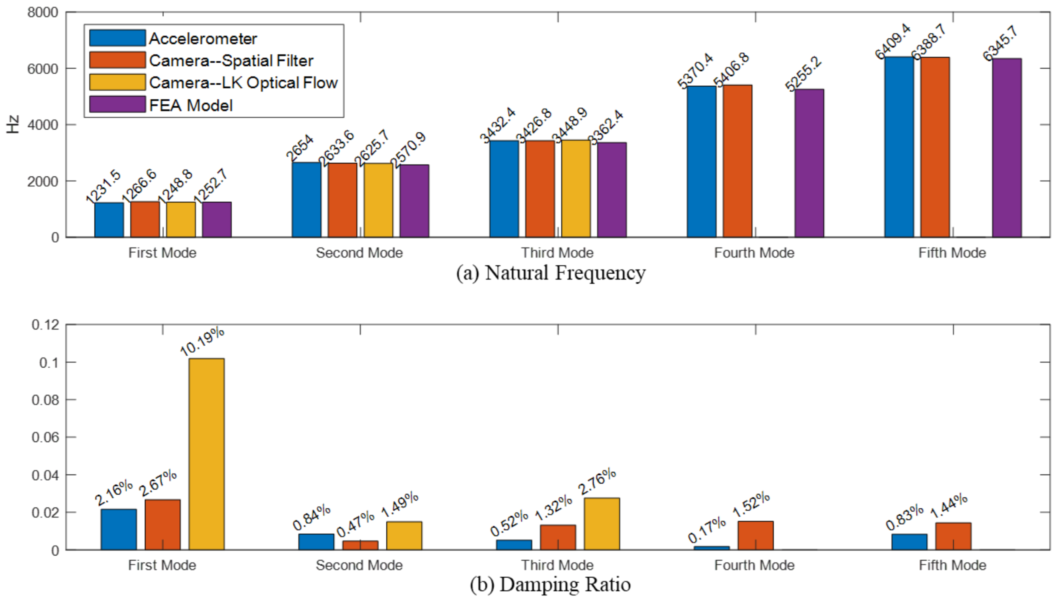



Figure 7-13 Comparison of modal parameters identified in experiment: (a) the natural frequency and (b) damping ratio

The natural frequencies of acceleration, spatial filtering, optical flow, and finite element are compared together. A consistency of the five detected frequencies proves that the output of spatial filtering with high SNR can ensure that weak modes in high frequency can be identified accurately. The error of damping ratio in spatial filtering is still acceptable and it shows great advantage than LK method.

### 7.1.4 Discussion

Above are the detailed procedures about how the mode shape is obtained in the program. In summary, the prior knowledge of mode shape is required about the node number, the property of standing wave and boundary condition, which can be obtained through theoretical analysis. But the exact node location is obtained by measuring signal power of experimental data (PSD of output signal at corresponding natural frequency). FEM is just used to compare with them. The result showing smoothness is because the employed sinewave-based function is smooth originally. Actually, the error shows on the location of nodes and also on the curvature between the estimated and real value.

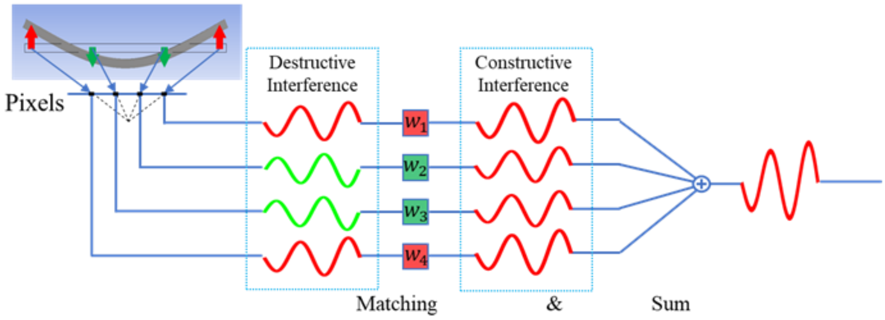



Figure 7-14 Spatial filtering for a pixel sensor array

The weight vector is used to match mode shape in this research. As shown in Figure 7-14, weight vector can change the destructive components into constructive components to enhance the signal power. Even though the mode shape is unknown at the beginning, an inaccurate assumption of weight vector can be applied to the input signals and generate a part of constructive components. By changing the weight vector gradually, the signal power will increase or decrease with respect to the variation of constructive degree in each step. If the weight vector gets close to the real mode shape, the signal power will increase. Otherwise, the signal power will drop. Thus, the power of output signal can indicate the matching degree between the weight vector and the mode shape. In this sense, the proposed algorithm just uses a common principle of adaptive filtering, in which its desired signal is each mode. As shown in Figure 7-14, when most destructive components turn into constructive, the mode shape can be found eventually. This is a typical optimization problem in adaptive filtering, once this problem is defined, it must have solutions. The proposed node/antinode searching method is not the unique solution, not applicable to all situation and can be further improved in the future.

As the effectiveness of the proposed method has been verified by simulation and experiment, the principle, estimation precision and computational efficiency of MASF compares with the window-based image registration, as shown in Table 7-1.

Table 7-1 Comparison between the MASF and window-based image registration

| | **Mode-shape-based adaptive spatial filtering** | **Window-based image registration** |
|---|---|---|
| **Range of application** | Mechanical vibration | No constraint of motion form |
| **Principle** | Array signal processing | Image registration |
| **Equivalent displacement sensor** | One pixel | One window (subset) |
| **Output signal** | Displacement under modal coordinate | Displacement under space coordinate |
| **Total of pixel usage** | Number of array elements: $m$ | Total pixels in all windows |
| **Number of outputs** | Expected modal orders: $n$ | Number of windows |
| **Maximum gain** | Number of array elements: $m$ | Not explicit |

It can be seen in Table 7-1, that the prime difference between the two methods is the principles behind them. The local operator used in window-based methods is versatile for different forms of motion such as measuring strain or deformation. However, in modal analysis, all space elements oscillate around an equilibrium point at the same frequencies. Aiming at this distinctive motion, the method based on pixel array framework can extract and processes vibration information more pertinently so as to reduce redundancy of image more efficiently.

Regarding to the computational efficiency, MATLAB is used to implement this algorithm on a laptop with Intel Core i7-7700HQ CPU @ 2.80 GHz processor and 16G GB RAM. Both the proposed method and LK optical flow are carried out to extract displacement signal using the same experimental dataset, the proposed method costs 29.7s totally, by contrast, LK optical flow takes 101.9 s [92].

### 7.1.5 Conclusions of current experiment

The experiment demonstrates the proposed adaptive spatial filtering can reach higher SNR than LK optic flow, which is a classic window-based image registration method. all five modes within frequency range are identified accurately, with the highest frequency being up to 6388 Hz, which cannot be identified with LK optical flow method. The fundamental reason for such an improvement can be attributed to two factors: firstly, the formation of sensor array takes pixel's sensitivity to displacement into account, so that the signal quality is guaranteed in data acquisition; secondly, spatial filtering can make modal displacements experience constructive interference by utilizing the correlation of all dynamic responses.

## 7.2 Modal analysis of wind turbine blade

In this section, experimental modal analysis of wind turbine blades is carried out by using conventional method and proposed method under different condition. Compared with stainless beam, the frequency range is lower, but the displacement amplitude is much larger, thus the Lagrange description and SVD is adapted.

### 7.2.1 LK optical-flow-based method

In this experiment, the LK optical flow is employed to extract displacement, four markers are tracked to obtain the displacement of blade.

#### 7.2.1.1 Experimental setup & instrumentation

In this research, an experimental rig consists of a set of blades from a 2kW wind turbine and an air excitation system [145]. The model of wind turbine is iSTA Breeze I-2000, it has 3 composite blades with length of 107 cm and weight 650 g. To obtain vibration modal properties (frequency, damping ratio and shape) of these lightweight blades in laboratory, the blades is fixed on a frame shown in Figure 7-15 and excited by air pulses. In this way, the measurement was carried out which are more similar to the real operating conditions. In addition, 7 black markers are made on the trailing edge as the objective to be tracked by a remote high-speed camera for obtaining the flap-wise vibration as illustrated in Figure 7-15.

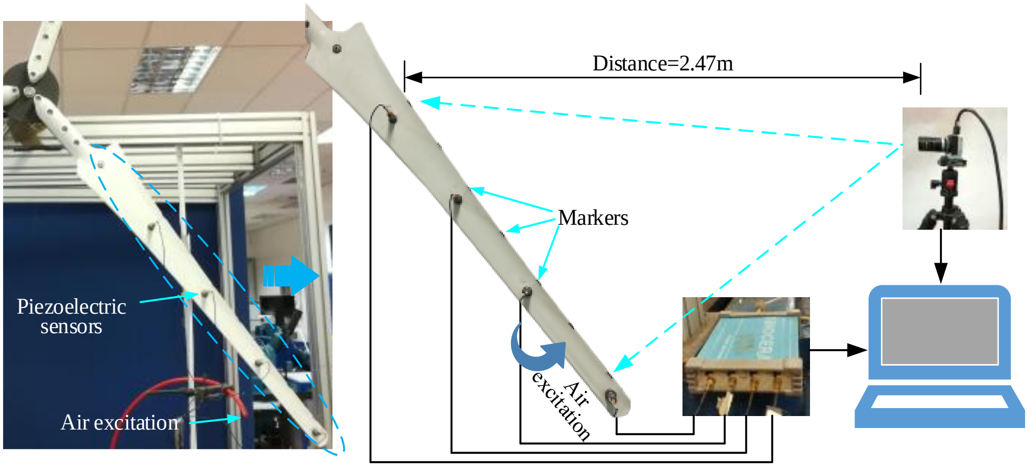



Figure 7-15 Schematic diagram of experimental rig and instrumentation of the high-speed camera and the accelerometers

The air excitation system is made up of a compressor to generate high pressure air, an air pressure regulator gauge, and a solenoid valve. This system can produce an air impulse to excite blades, being like natural excitation due to a strong wind turbulence. In combination of the high-speed camera set at 2.47 metres away from the blades, this test system can measure the vibration modes in a remote way and expected to be more accurate compared conventional contact measurements.

The video measurement system consists of a high-speed industrial camera with a single lens. The high-speed camera uses a CMOS sensor, the resolution is 640×480 pixel and the maximum fps is at 600. In this experiment, frames per second was set to 200 fps for low computational cost and found is sufficient to identify the vibration properties. The captured videos then are processed in Python & OpenCV and MATLAB based on the forementioned method in previous sections.

Meanwhile, for the verification of the photogrammetric results obtained, acceleration responses were also collected by 4 piezoelectric accelerometers. These sensors are mounted at points near location of markers.

#### 7.2.1.2 Vibration measurement based on LK optical flow

Unlike accelerometers measuring the acceleration straightforwardly, vision-based method needs to detect and track markers in the video, get the trajectory of markers' motion, and then convert trajectory of pixel in successive images into physical displacement in space, finally solve the accelerations through differential calculation. The signals are acquired by a four-channel dynamic measurement system which is hosted by laptop in Figure 7-16. The vibration signal extracted from image includes three steps. First, to verify the accuracy of optical measurement. While the camera is capturing the video, piezoelectric accelerometer is collecting acceleration at the same time. Because the traditional contact sensor is precise and frequency range is wide, the vibration signals collected from accelerometers can be regarded as a good reference to verify the error of the vision-based method.

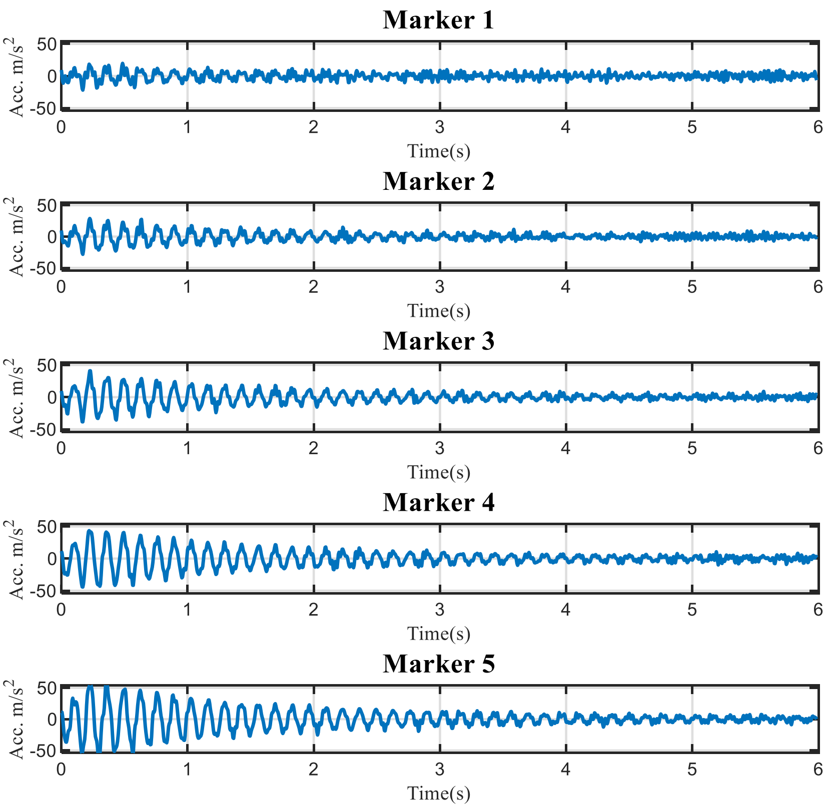



Figure 7-16. Acceleration of 5 markers acquired from photogrammetry

The second step, the baseline condition is obtained. All used piezoelectric accelerometers are removed from the blades. since mass of the sensors inevitably has impact on the dynamic characteristics of the blades. So that using camera we can obtain the more real dynamic characteristics of healthy condition,

The third step, according to other researchers' previous research [17], the trailing edge deformation and real load distribution at root end are two important factors leading to failure. Thus, at the non-destructive condition, loosening the screws at the end root and adding mass near the trailing edge are implemented to simulate the typical fault on the wind turbine blades. The aim is clear that judge the fault using photogrammetry.

The comparison between vision-based methods and accelerometer signals in the time domain and frequency domain is shown in Figure 7-17. Sampling rate of piezoelectric sensor is 1.5k Hz, that is significantly higher than vision-based method (200 Hz).

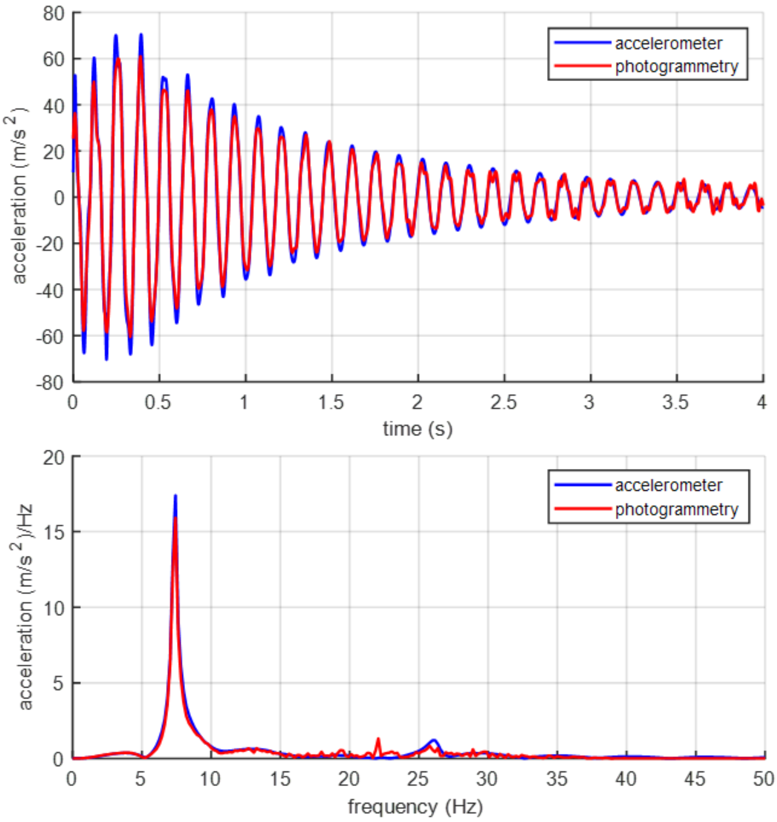



Figure 7-17. Comparison of comparison between vision signals and accelerometer signals in Time domain and frequency domain

Apparently more high frequency components can be acquired by the accelerometers. Meanwhile, the resolution of CMOS sensor also has the influence on the result. Despite these, the amplitude and frequency values are well agreeable. Given that large-scale wind turbine blades have lower natural frequency, in this case the vision-based method should be very effective.

#### 7.2.1.3 Modal Identification

It is well validated that the stochastic subspace identification (SSI) is reliable, so that it is implemented to identify the modal parameters [146]. The stabilization shown in Figure 7-18 exhibits the process of identifying the modes over different model orders.

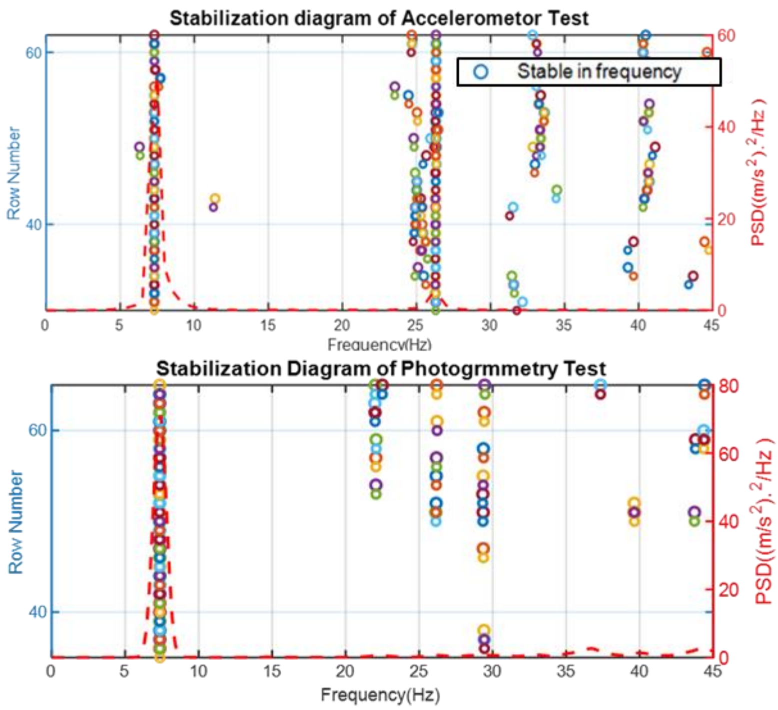



Figure 7-18. The stabilization diagram of the method of accelerometer and photogrammetry

As shown in legend, '*o*' represents the stable results identified in different model order, the color of '*o*' is random just for display effect. A consistent result means the identified natural frequency is reliable [9]. To calibrate the results, the stabilization diagram and the results are provided in Figure 7-18 and Figure 7-19 respectively for the case of normal blade conditions.

The identified natural frequency, damping ratio, and the modal shape are used to implement a condition monitoring method, which are shown in Figure 7-19 and Table 7-2. Although, there are different spurious modes in these two diagrams, both they successfully identify the same modes.

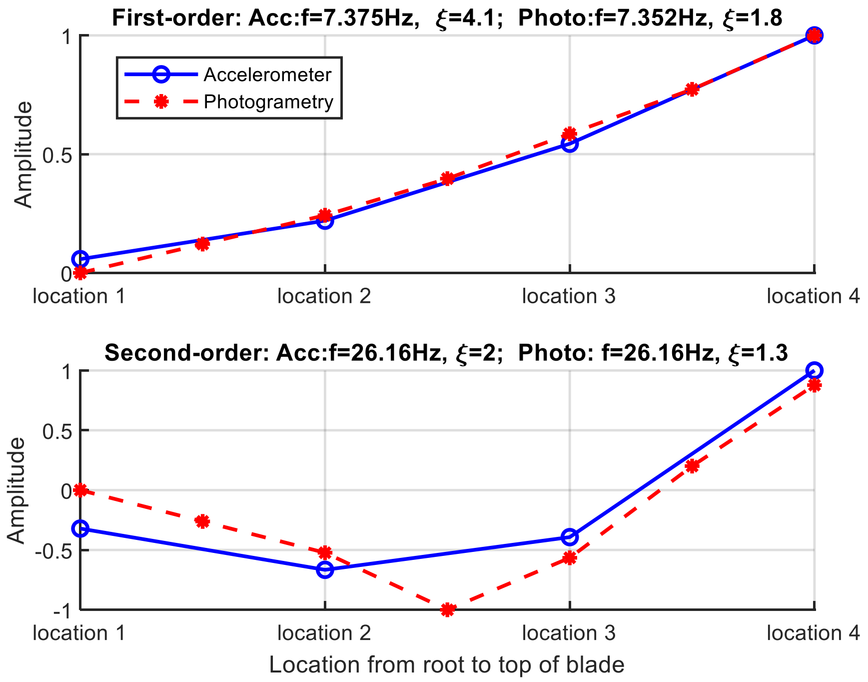



Figure 7-19. The mode shape for the first and second modes

In Figure 7 19, the differences for first and second natural frequencies is 0.31% and 0.1% respectively. However, the error of damping ration is more obvious, which is common problem in using output-only methods. Moreover, two mode shapes are consistent with each other for the two methods. These then confirms that the vision-based system shows acceptable performance.

#### 7.2.1.4 Damage Detection based on Modal Parameters

Because the modal parameters of wind turbine blades can be accurately extracted from photogrammetry. The damage detection based on vision method can be implemented logically. The artificial faults are set to check out if this method is stable and sensitive to some common fault of the wind turbine blades. All in all, the change of natural frequency can indicate the faults, regulation of damping ratio is not very clear. Modal shapes keep feature of first and second order mode but cannot see a significant change.

At the beginning, all sensors are removed from the blade to test the parameters of baseline condition, so that the faulty cases could have a standard. Since part mass are subtracted, the natural frequency increase naturally, comparing with the above verification test.

After that, the sensors as 3 extra masses (11g per sensor) are gradually added on the blade, the change of frequency are discovered clearly, especially when the third mass block is added near the end of the blade, where often accumulates most snow or ice and cause most enormous imbalance and lead to fault easily.

Table 7-2 Difference of modal parameters caused by fault

| | | First order | | | | Second order | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Natural frequency | Change rate | Dampi ng ratio | Change rate | Natural frequency | Change rate | Dampi ng ratio | Change rate |
| **Baseline** | | 8.79Hz | 0% | 1.4% | 0% | 29.73Hz | 0% | 0.9% | 0% |
| **Icing or frozen** | 1 | 8.63Hz | -1.82% | 1.0% | -28.57% | 28.87Hz | -2.89% | 1.3% | 44.44% |
| | 2 | 8.34Hz | -5.12% | 1.6% | 14.29% | 28.65Hz | -3.63% | 1.2% | 33.33% |
| | 3 | 7.50Hz | -14.68% | 1.4% | 0% | 26.83Hz | -9.75% | 1.0% | 11.11% |
| **Loosen** | | 8.61Hz | -2.05% | 1.2% | -14.29% | 29.26Hz | -1.58% | 1.4% | 55.55% |

Finally, the screw which fix blades on the shaft was loosened, this part always bears huge stress. Therefore, this fault is chosen to simulate early crack or loosening. From table 1 it can be seen that the change in natural frequency is distinctive, which allows for an assessment of blade conditions [145].

#### 7.2.1.5 Modal Identification using Unmanned Aerial Vehicle

Because of the non-contact property of vision-based measurement, increasing studies on remote structural health monitoring based on camera on an unmanned aerial vehicle (UAV) [147]–[151]. To test the feasibility of monitoring on-site wind turbine, the flap-wise vibration of a wind turbine blade is measured using the UAV and post-processed by the stabilization algorithm [152].

The blade is fixed by a clamp and excited by a wood hammer. The black-white square markers are stuck on the clamp and the wall as the background points. The UAV is hovering about 1.5m in front of the blade in Figure 7-20.

Figure 7-20 Hovering UAV captures blade

Figure 7-21 shows the trajectories of feature points tracked by the optical flow method. The background trajectories are used to work out the displacement drift, and foreground trajectories are compensated to obtain the accurate displacement of the blade. Because of the diffuse noise frequency, band pass filter cannot achieve good effect. But the homography matrix can represent motion between each frame. The homography transformation is used to perform video stabilization. The procedures are: 1) 4 still points in frame are tracked to get their trajectory. 2) Relying on these trajectories, the homography matrices are estimated. 3) Then the frame can be reconstructed and stabilized using the homography matrix. In this reconstructed video, displacement drift is suppressed by estimated transformation.

Figure 7-21 Trajectories of feature points

The spectrum as shown in Figure 7-22 is derived from the displacement signal, the first mode is extracted successfully, which is 9.612 Hz. The residual of the drift is about 0.38 Hz, but since it is distant from the inherent frequency of blade, it can be separated from the mode.

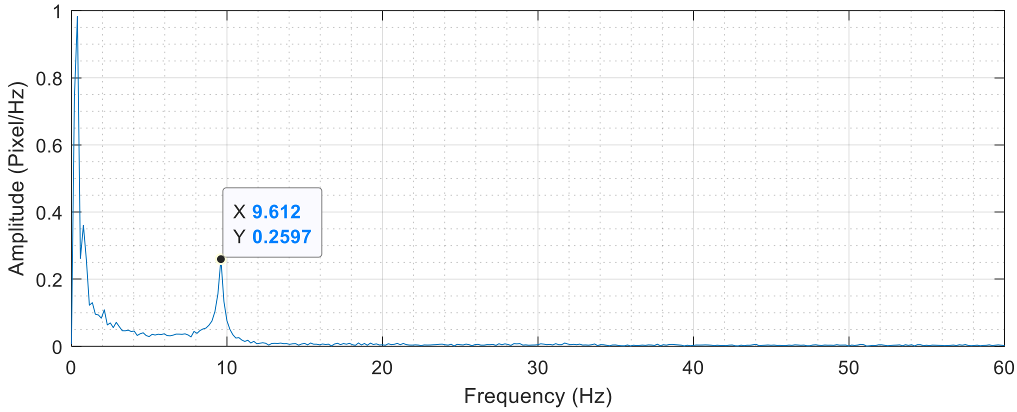


Figure 7-22 Spectrum of wind turbine blade displacement

In summary, the first mode of a 1.5m wind turbine blade are identified by an UAV. That result validates the video stabilization is effective for modal identification. But the second and higher order modes are missed. The factors are multiple, including the ego-motion, subtle displacement relative to resolution and the sample rate of camera. These influence factors will be addressed in the future research [152].

#### 7.2.1.6 Conclusions of current experiment

In this research, modal parameters identified using a vision-based method are validated by the result of the traditional piezoelectric sensor-based method. The difference of the natural frequency is less than 0.31%. Furthermore, the healthy case and the faulty cases are differentiated by the image signals distinctly. The mass change of blades as small as 11g can be detected, which is shown more significantly when the added mass is more distant from the root.

In conclusion, high-speed camera combines with the air impulse excitation is an effective way to detect the mass addition and loose fault. However, the spatial density and the frequency range still need further improvement for crack detection.

### 7.2.2 Array signal processing method

In previous section, the method used in vision-based measurement is a typical window-based method, LK optical flow. It can obtain the natural frequency and mode shape. However, its spatial density has only five points and the frequency band is narrow. Therefore, the framework based on array signal processing is used to make up for the above deficiencies. According to the analysis in section 3.3 and section 5.2, the LaGrande description and SVD is selected to deal with condition.

#### 7.2.2.1 Experiment setup

In this experiment, a wind turbine blade segment of 465 mm is fixed on a cramp as shown in Figure 7-23, one high-speed camera is used to capture motion of blade. The strip on the leading edge is painted to increase the contrast, in order to enhance the feature for tracking. A Phantom (SN24006) high-speed camera equipped with Nikon 35mm 1:1.8G lens was employed in this test. The sampling rate and the resolution were set as 500 fps and $1280 \times 128$ respectively.

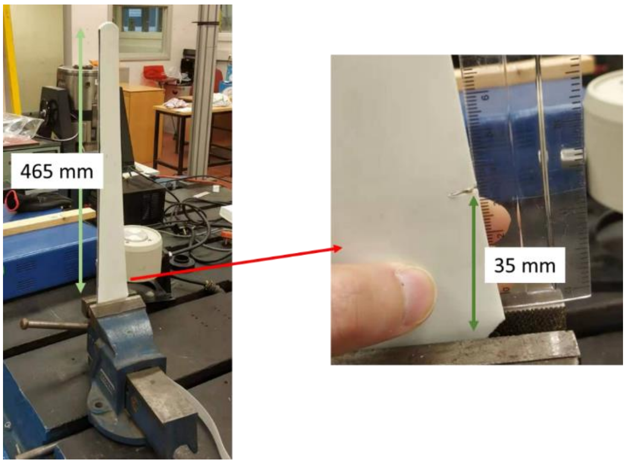



Figure 7-23 Experiment setup of wind turbine blade and the manual crack as damage

Totally, three damage cases are set to test the diagnosis ability of vision-based method, which include intact blade as baseline, a 5 mm crack and an 8 mm crack at 35 mm from root. An impact hammer is used to excite the blade.

Figure 7-24 One frame from experimental video in experimental modal analysis of wind turbine blade

From the frame of experimental video Figure 7-24 One frame from experimental video, the blade has a high-contrast, however the background is complex.

#### 7.2.2.2 Pixel array processing and result analysis

To extract the displacement from video footages, the coarse-to-fine edge detection approach is for this large displacement case. The algorithm is detailed in Section 3.5, firstly the image is smoothed by using a gaussian kernel, and then image gradient is calculated to detect the edge, where has the largest gradient. Subsequently, its second order gradient is calculated to refine the edge position. The subpixel estimate of edge is determined at where the second gradient is zero. It can be known from the above that the greatest image gradient value indicates position of edge, thus the second order gradient is taken, and its zero value is used as the subpixel estimate of edge.

Through above steps, the subpixel accuracy of edge can be obtained, and the displacement of every row will form a matrix of input signal. The mode shape and time sequence of each mode are obtained by SVD of the input signal matrix. There are three modes are decompressed as shown from Figure 7-25 to Figure 7-27.

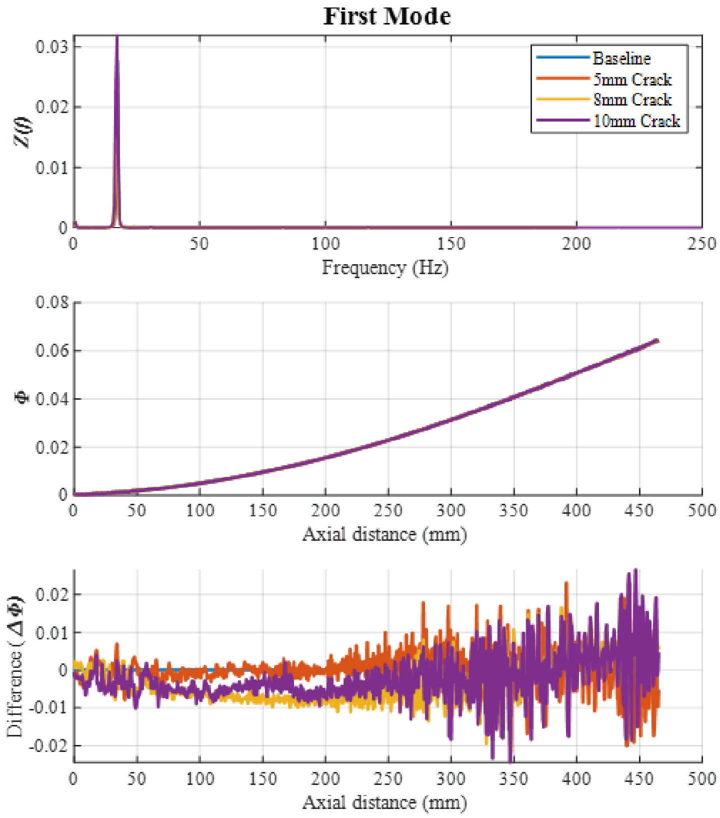



Figure 7-25 The first mode of three different cases. The spectrum, mode shape and the difference of mode shape are shown in upper, middle, and lower figures.

The first order of natural frequency is 17.6 Hz, even though its spectrum is clean, three cases cannot be distinguished, as well as the mode shape. In middle figure, to compare mode shape of three cases directly, the mode shape is overlapped completely. If the difference of mode shape is calculated, that show a noisy and crossed value.

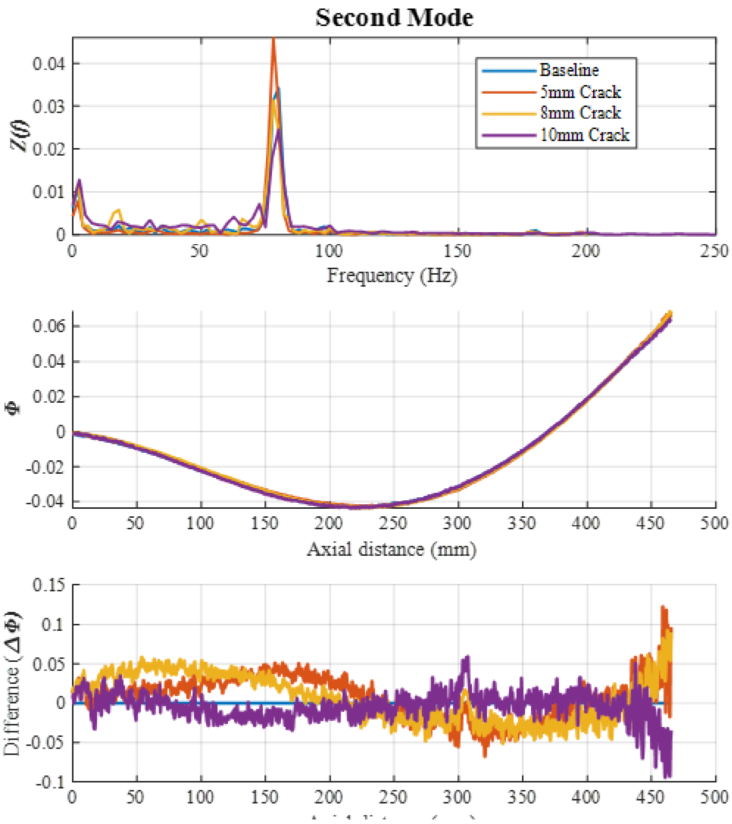



Figure 7-26 The second mode of three different cases. The spectrum, mode shape and the difference of mode shape are shown in upper, middle, and lower figures

As for the second mode at 78.3Hz, the natural frequency and mode shape is still very close to each other, and the difference of mode shape show a noisier contour, due to lower SNR in higher frequency.

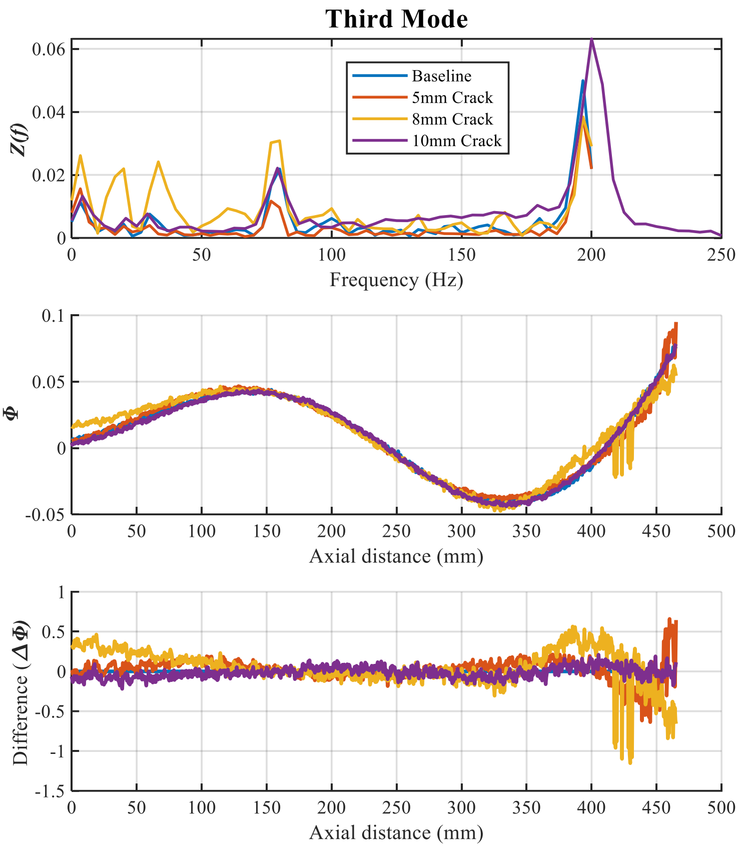



Figure 7-27 The spectrum, mode shape and the difference of mode shape are shown in upper, middle and lower figures

The third mode is 199.3Hz, that also show a similar graph. Because the mode is very weak, the spectrum and mode shape get noisier. Likewise, the crack cannot be detected by frequency or mode shape.

#### 7.2.2.3 Conclusions of current experiment

Based on above analysis, the mode shape and natural frequency can be identified based on SVD of array signal matrix accurately. Compared with LK optical flow, the density of mode shape is improved significantly. Meanwhile, the weak mode is identified up to 200 Hz, which cannot be obtained through LK optical flow.

However, the small crack detection remains large challenges, because the mode shapes are affected by crack too subtly.

## 7.3 Key findings

This section uses two experiment to validate proposed method. Compared with conventional window-based method, the proposed adaptive spatial filtering method has higher performance on SNR gain. SVD-based method can obtain mode shape of wind turbine blades, which is more accurate than result calculated from LK optic flow. However, the change caused by crack cannot be detected by vision-based method, when crack is very small.

# 8 APPLICATION OF VISION-BASED VIBRATION MEASUREMENT TO ROTATING MACHINERY CONDITION MONITORING

*Two condition monitoring experiments are conducted to validate the proposed methods in this chapter. The vibration of a reciprocating compressor and a gearbox are measured by the vision-based method, and the fault features are extracted based on the vibration signals. The experiment on compressor validates the LK optical flow can extract the features in low frequency band, like rotating frequency, but not valid in higher frequency range. In contrast, the experiment on gearbox verifies the proposed data-independent spatial filtering combined with time-synchronous averaging can manage to extract the sideband of two mesh frequencies. Even though many aspects of proposed methods can be further improved, this early-stage research shows the promise of vision-based condition monitoring in the future.*

*It should be noted that some paragraphs* in *section 8.1 are expansion or reproduction of the published materials* [3] of list of publications, *and the specific paragraphs are annotated in a form of citation.*

**Highlight:**

- The vibration of a reciprocating compressor and a gearbox are captured by a camera, and the fault features are extracted from image data.
- The experiment of compressor shows the LK optical flow can extract the features in low frequency band, like rotating frequency, but not valid in higher frequency range.
- In contrast, the experiment on gearbox shows the proposed method combined with Time-synchronous averaging can manage to extract the sideband of two mesh frequencies.

## 8.1 Condition monitoring of a reciprocating compressor

### 8.1.1 Experimental setup

The experiment was carried out on a two-stage single acting reciprocating compressor (Figure 8-1). Which is powered by a 2.5KW three-phase induction motor. The motor drives the pistons to reciprocate in cylinders of a V-shaped chamber. The first stage cylinder with the second stage cylinder is connected by an intercooler coil. At the end of air flow, the compressed high-pressure air is stored in a tank with a maximum capacity of 1.38Mpa [107] .

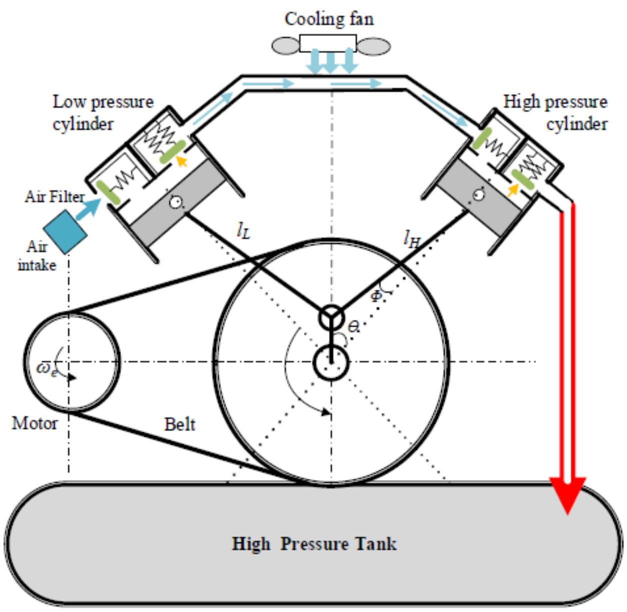



Figure 8-1 Schematic diagram of the two-stage reciprocating compressor

To investigate vision-based condition monitoring, camera and conventional instrument are set up jointly.

On one hand, one dynamic pressure transducer is mounted on the top of cylinder and an accelerometer is installed on top of the compressor chamber as shown in Figure 8-2. On the other hand, the shooting equipment is a smart phone, which resolution is 4k (3840×2160), and the frame rate is 30 per second. An LED provides sufficient lighting. A checkbox marker is stuck on chamber for tracking with side length of 4.7 mm.

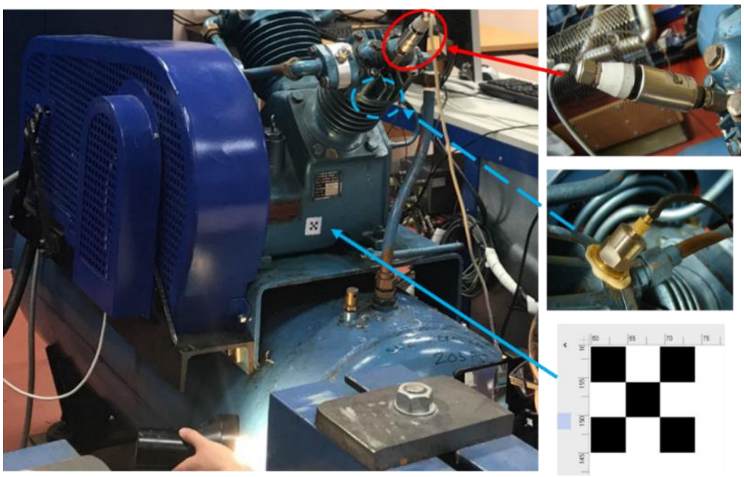


Figure 8-2 Setup of the test rig, the dynamic pressure sensor, accelerometer, and the marker are highlighted in the picture.

Two operating conditions are set. First, a healthy condition is set as the base line. Second, a discharge valve leakage (DVL) is employed manually. The cylinder pressure and vibration are collected to distinguish two cases.

### 8.1.2 Result analysis

#### 8.1.2.1 Cylinder pressure and vibration

The Figure 8-3 exhibits the pressure of Baseline and DVL respectively the cylinder in single stroke, which is most direct indicator of leakage. The DVL would delay the intake valve opening moment and advance the discharge valve opening moment. In each cycle, the transported air volume and pressure increase will be less than the normal condition. Consequently, the efficiency of the machine would decrease. That will further result in the delay of vibration signal [153].

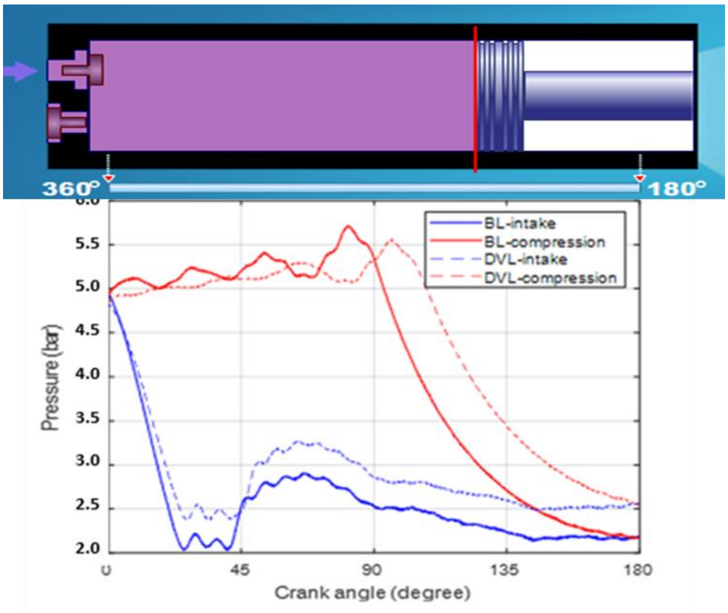



Figure 8-3 Air pressure of 2-stage cylinder in angular domain

Next, the acceleration signal is collected to analyse the change of dynamic characteristics caused by DVL. The Figure 8-4 shows the time synchronized average of vibration signals and its spectrum in single stroke, which shows amplitude of the high frequency components are modulated by the rotational speed, and the rotational frequency exists in the meantime. But the distinction between baseline and DVL is not obvious. To extract effective indictor, more advanced algorithm is required to analyse the modulation components, like modulation signal bispectrum (MSB). To simplify this process in a non-contact measurement, the vision-based method is developed.

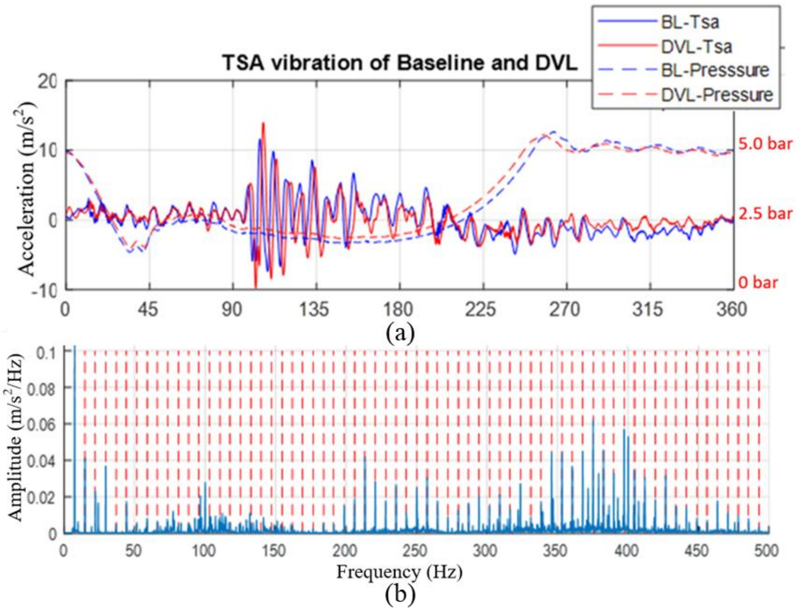



Figure 8-4 (a) Dynamic pressure and vibration in angular domain (b) spectrum of vibration signal

It is clear in Figure 8-4 (b) the significant peak is distributed uniformly, that is because all vibration is modulated by the rotational frequency (Figure 8-5). Furthermore, the rotational frequency of compressor is very stable because the motor and load are fixed for a specific compressor. It can be seen in Figure 8-5, the rotational frequency varies within 0.3 Hz and 0.05 Hz difference between baseline and DVL in the range of the 50 psi and 100 psi, which is equivalent to 3 rpm. Therefore, it can be thought basically that the frequency of rotation is a determined feature for a compressor, although the leak of valve will decrease the rotational frequency 0.05 Hz [107].

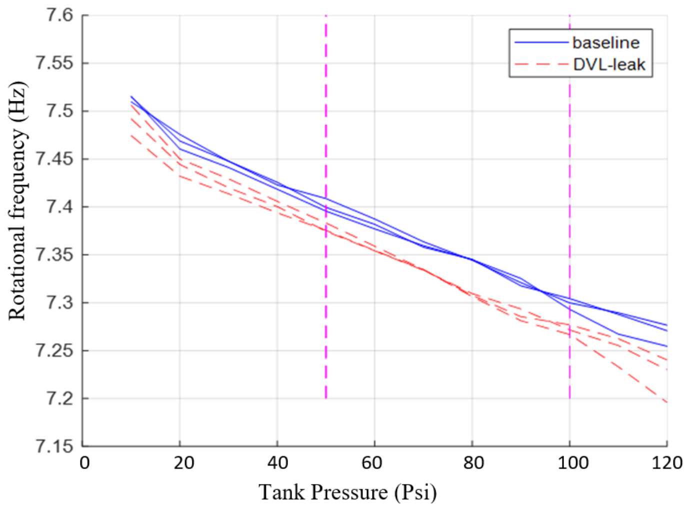



Figure 8-5 Change of rotational frequency with increase of tank pressure

#### 8.1.2.2 Vision-based measurement based on LK optical flow

As the characteristics of compressor is known based on above analysis, the aim is to study how to extract the features based on the video. Firstly, the subpixel corner detection[154] is implemented to find the features, which calculates the inner product of corner point and surrounding pixels to refine the accuracy. Afterwards, the LK sparse pyramid optical flow[155], [156] is adopted to track the corners. Total 16 corners are detected and tracked steadily (figure 5.3). A block size is around $36.5 \times 36.5$ pixel, so that a millimetre-to-pixel conversion can be calculated as 0.1288 mm/pixel. This conversion is related to the shooting distance or the focal distance.

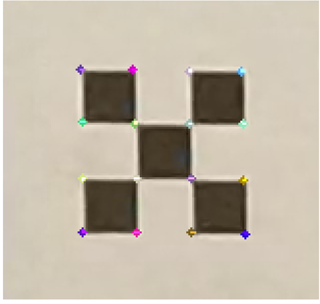


Figure 8-6 Feature points detected in real frame

The displacement of feature points in time domain is shown in Figure 8-7. The amplitude in the range of ±1 pixel, which means the displacement is less than ±0.12 mm.

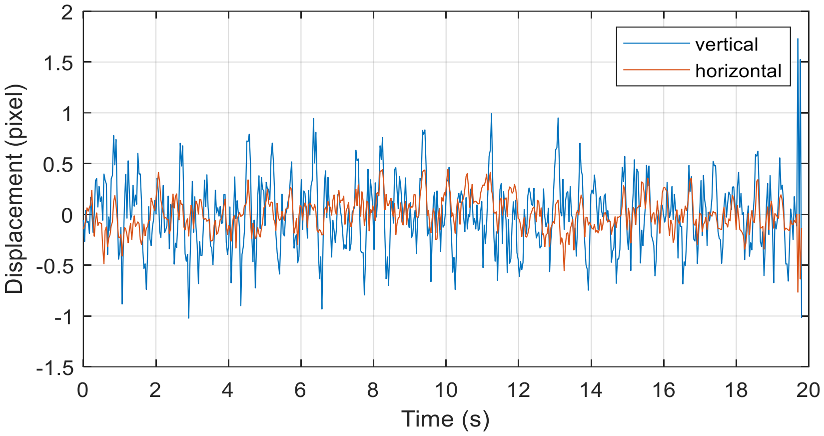



Figure 8-7 Displacement signal based on LK optic flow

Although the frequency band is limited, the rotational frequency and its second order harmonic can be extracted using this method as shown in Figure 8-8, because the rotational frequency can only exist within this range.

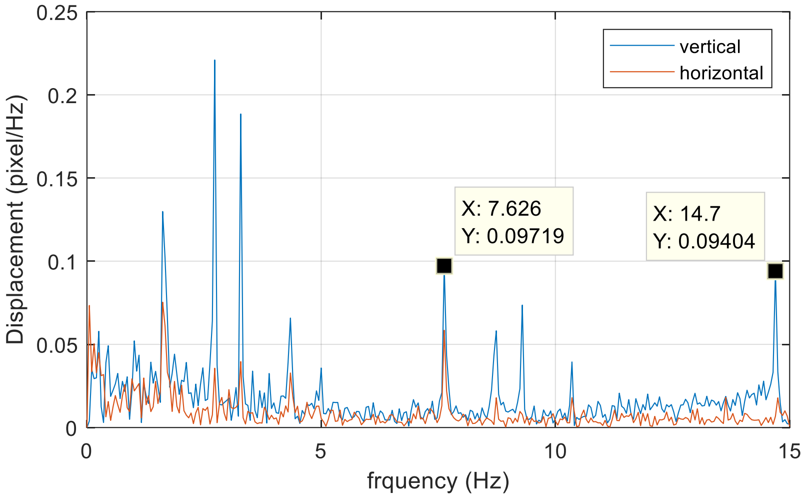



Figure 8-8 Spectrum of displacement signal based on LK optic flow

The advantage of this method is it can use a relative low sample frequency method to monitor basic operating condition of reciprocating compressor.

### 8.1.3 Conclusions of current experiment

In this early-phase experiment, the displacement is difficult to detect in high frequency, not only because of the limitation of camera’s sample rate, but also the subtle displacement amplitudes. Thus, the accuracy of current algorithm can meet the requirement of measuring rotating frequency and first order harmonics, which can reflect some basic condition of machine.

In conclusion, even though the conventional vision-based vibration measurement can extract the feature signal in low frequency such as rotational frequency around 7.6 Hz. The interference signal and noise will interfere the feature extraction and will affect the accuracy the robustness of fault diagnosis. Considering this situation, the hardware and processing method require further improvement. For this reason, the spatial filtering method is carried out and tested on a gearbox.

## 8.2 Condition monitoring of gearbox

By contrast with the modal signal, the vibration of rotating machinery is a wideband signal, which leads to that the signal power in Eq. (4-13) is difficult to obtain. Thus, a data independent spatial filtering method is used to extract key features for operating condition of gearbox.

### 8.2.1 Background of mesh frequency and its sidebands of gearbox

The mesh frequency and their sidebands are important signature correlating to most fault of gearbox, which can be caused by modulation due to gear pitting and misalignment. The rotation frequency of gear and pinion holds following relationship,

$$f_{Gear} = f_{Pinion} \times \frac{Number\ of\ pinion\ teeth(N_p)}{Number\ of\ gear\ teeth(N_g)} \tag{8-1}$$

which can be rewritten in the equivalent form:

$$f_{mesh} = f_{Pinion} \times N_p = f_{Gear} \times N_g \tag{8-2}$$

$f_{mesh}$ is the tooth-mesh frequency, also called gear-mesh frequency, is the rate at which gear and pinion teeth periodically engage. If gearbox has gear tooth pitting or shaft misalignment, the distributed sidebands will appear in both sides of mesh frequency.

$$f_{sideband,Pinion} = f_{Mesh} \pm m \times f_{Pinion} \tag{8-3}$$

The sidebands of gear fault frequency appear at different rotating frequencies according to the location of faults.

$$f_{sideband,Gear} = f_{Mesh} \pm m \times f_{Gear} \tag{8-4}$$

in which $m$ is the order of harmonics. That will show multiple sidebands in spectrum, which can reveal the location of faults reliably. This experiment aims to test whether the proposed vision-based method can detect the sideband [157].

### 8.2.2 Experimental setup

A Phantom (SN24006) high-speed camera equipped with Nikon 35mm 1:1.8G lens was employed in this experiment. The sampling rate was 6,000 fps and the resolution was 512×576, simultaneously an accelerometer is used to capture the vibration for comparison as shown in Figure 8-9.

Figure 8-9 Experimental setup of condition monitoring of gearbox by using a high-speed camera.

One frame of captured video is shown as Figure 8-10, which show a low brightness due to short exposure time, that will lead to a weak signal power of vibrational displacement estimated from this dark image. In this sense, the potential improvement of experiment is the enhancement of light source. In this experiment, there the tooth number of two pairs of gear are59 13 47 58 respectively, thus it will turn out two mesh frequency based on the Eq. (8-2). The first-stage mesh frequency $f_{m1} = 316.75$Hz and the second-stage mesh frequency $f_{m2} = 1145.1$Hz. The rotating frequency of three shaft (one gear frequency and two pinion frequencies Eq. (8-2)) are measured at $f_{r1} = 5.36Hz, f_{r2} = 24.37Hz\ and\ f_{r3} = 19.74Hz$.

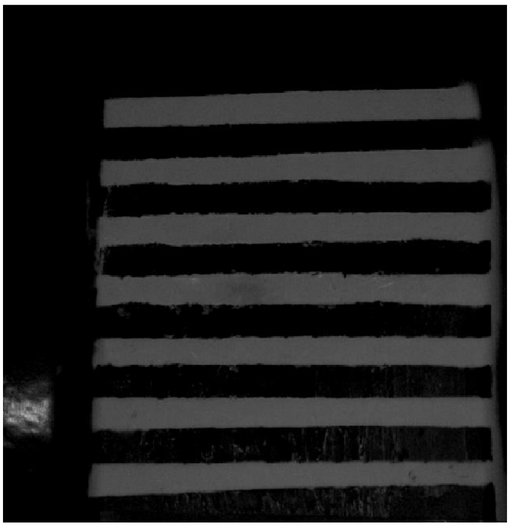


Figure 8-10 One frame of experimental video

The mesh figure (Figure 8-11) shows intensity signals of one image. The intensity of white stripe just reaches about 120, which is lower than a half of maximum, which shows a huge potential of signal strength improvement. According to the analysis in section 3.1.3, the brightness determines the sensitivity of displacement directly.

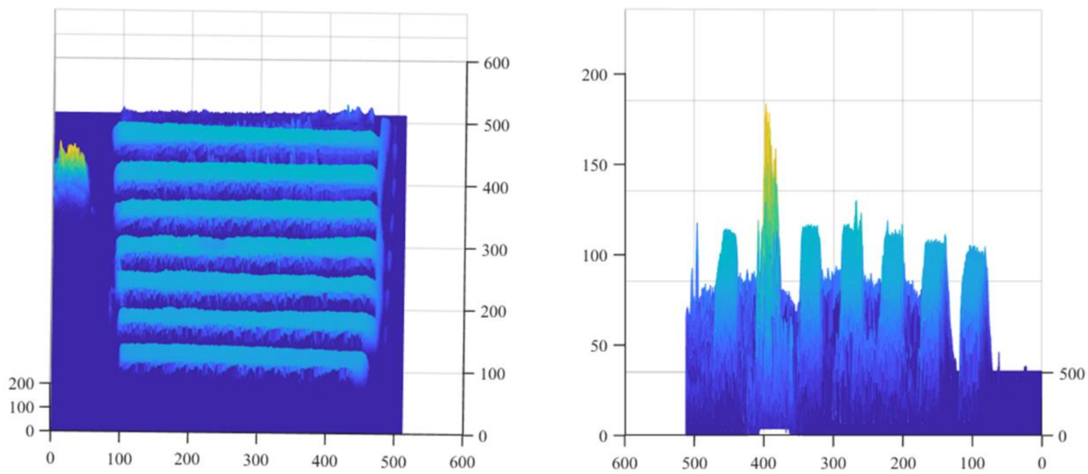



Figure 8-11 Intensity matrix observed in a 3D perspective.

## 8.2.3 Result analysis

### 8.2.3.1 Formation of sensor array

As described in 3, the edge detection method is used to formulate the pixel array. To remove the interference in image (such as shadow, reflection, and defect of painted pattern), the pre-processing to form the pixel array is carried out through following steps,

- Calculate the image gradient on the axis perpendicular to strip.
- Select the pixels, whose gradient is larger than a selected threshold.
- Use dilate operation in order to connect disconnected edges.
- Use erode operation in order to remove spare pixels of edges.
- Remove the small objects from binary image.
- Obtain the image gradient of pixels in this binary image.
- Find the peaks on the edge and choose the edge breadth.

Thereafter, a pixel array is formulated, each pixel of which has high sensitivity to vibration.

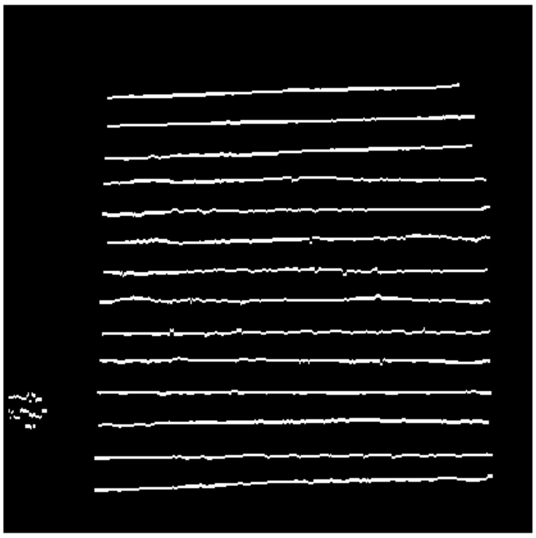


Figure 8-12 The pixel array selected according to the image gradient, which will be used to estimate vibration.

It should be noted that p, then the dimension of array will be reduced from image size to the number of pixels on the edge, in this case, there are 5137 pixels in pixel array.

#### 8.2.3.2 Data independent spatial filtering (DISF)

The vibration signal of rotating machine is different from modal test, which can still be represented by $\boldsymbol{u}(t) = \boldsymbol{\Phi}\boldsymbol{z}(t) + \boldsymbol{\eta}(t)$ i.e., the Eq. (2-1). $\boldsymbol{\Phi}$ cannot call the mode shape now, but it is a value that is equivalent to mode shape in equation. In this case, $\boldsymbol{\phi}_i$ denotes every spatial element's amplitude of the signal $\boldsymbol{z}_i$. Likewise, $\boldsymbol{z}_i$ is no longer to be the modal signal, it is a wideband signal here.

Considering $\boldsymbol{\phi}_i$ generated from the impact and vibration of rotating machine, the wave propagation tends to be traveling instead of standing. $\boldsymbol{\phi}_i$ is thus taken as rigid-body mode shape and approximated by a constant. According to Eq.(4-13), $\boldsymbol{w}_i = \boldsymbol{\phi}_i$ should still hold on, and a vector with uniformed elements $\boldsymbol{w} = [1,1, \ldots ,1]^T$ is used as weight vector. Since the weight vector is a pre-set value based on property of travelling wave independent of collected data, it is called a data independent spatial filtering (beamforming) [111]. Figure 8-13 shows the output signal went through this filtering and the feature components emerged in spectrum.

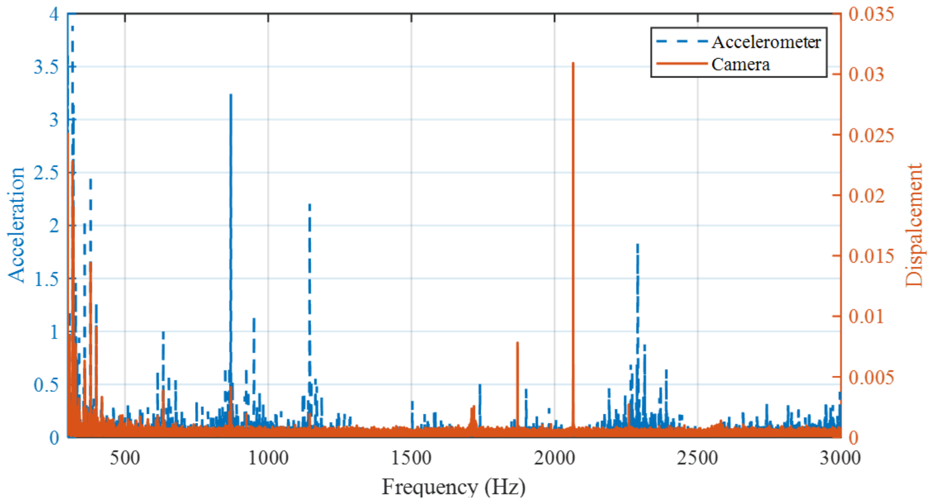



Figure 8-13 Spectrum of acceleration and displacement from image.

The spectrum of displacement and acceleration show different tendency in Figure 8-13, concretely, the displacement is sensitive in lower frequency, but attenuates fast with increase of frequency. Moreover, a component at 2083Hz and sideband occurs in displacement spectrum based on vision method, which cannot find in acceleration. It is likely to arise from motion of camera, but that did not affect the characteristic frequency.

Overall, the signal component in frequency range is detectable basically. To obtain more accurate fault characteristics, the side band of two mesh frequency will be further enhanced based on Time-synchronous Averaging (TSA).

### 8.2.3.3 Time-synchronous averaging

Full machinery component interaction arises in a every single rotation period. Time-synchronous averaging (TSA) is a rotation-based average, which is averaging over uniform rotation angles or complete rotations. It is conducive to supress noise, disturbance, or periodic signal content that is not coherent with the rotation. Based on different shaft frequency, its corresponding sideband of mesh frequency can be detected more clearly. For example, the sideband $f_{m1} \pm nf_{r1}$ around first-stage mesh frequency $f_{m1} = 316.75$Hz is averaging over the first shaft frequency $f_{r1} = 5.36Hz$ as shown in Figure 8-14(a). Likewise, the second rotating shaft frequency $f_{r2} = 24.37Hz$ is used to enhance its corresponding sideband $f_{m1} \pm nf_{r2}$, however, no significant sidebands are detected. It is consistent with acceleration signal, which shows that the gear on this shaft generates few sidebands.

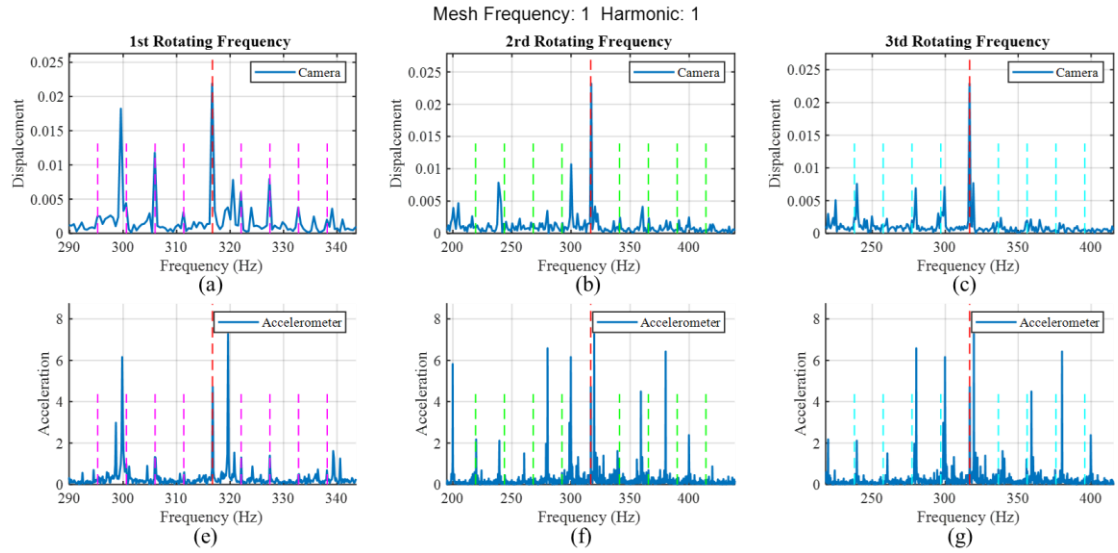



Figure 8-14 TSA spectrum in the frequency range around the first mesh frequency of first-stage gears for three rotating speeds

The second harmonic of first-stage mesh frequency $2f_{m1} = 633.5\ Hz$ is obvious. In addition, the first rotating sideband$2f_{m1} \pm nf_{r1}$ around $633.5\ Hz$ can be detected in Figure 8-14 (a).

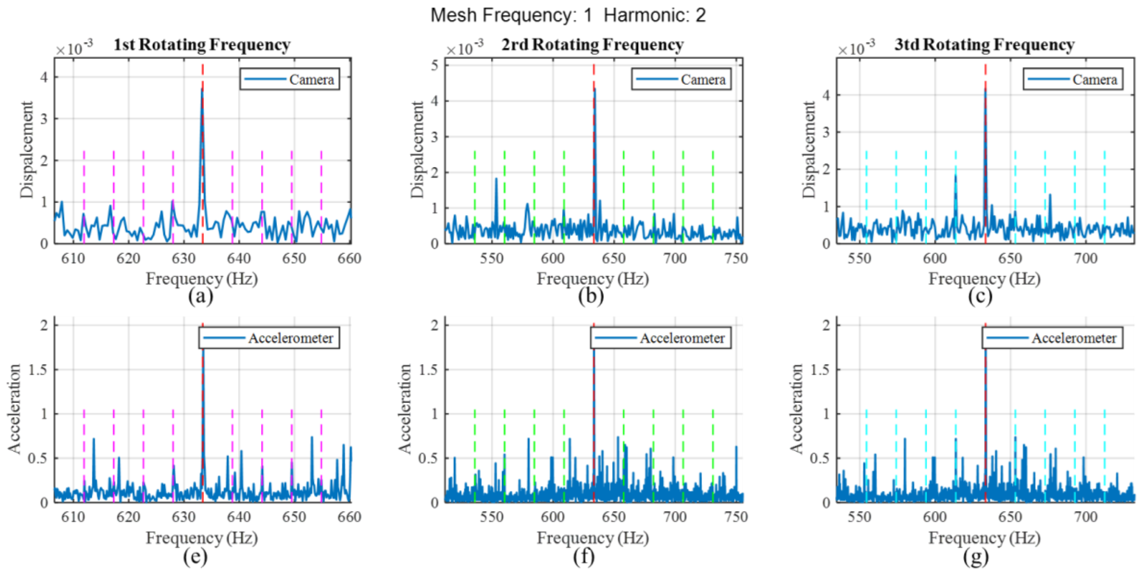



Figure 8-15 TSA spectrum in the frequency range around the second mesh frequency of first-stage gears for three rotating speeds

As shown in Figure 8-16 (b) and (c), the sidebands $f_{m2} + nf_{r2}$, $f_{m2} + nf_{r3}$ of second-stage mesh frequency at $f_{m2} = 1145.1$Hz are still recognizable, though no longer obvious. The sideband is weak even in acceleration spectrum, that shows the difficulty to extract this weak feature.

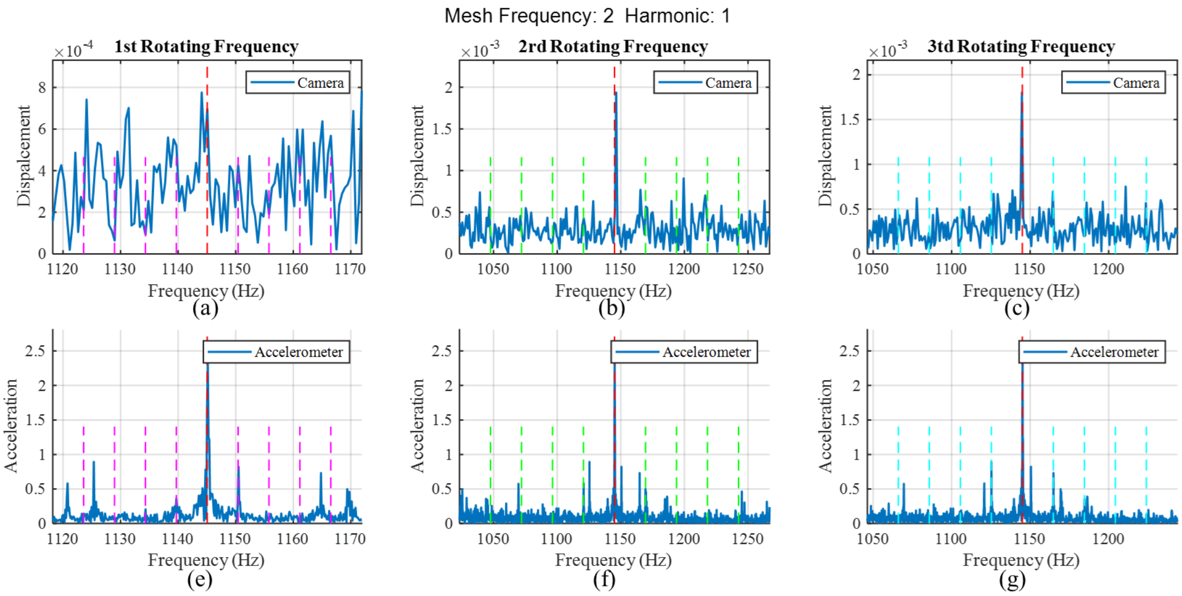



Figure 8-16 TSA spectrum in the frequency range around the first mesh frequency of second-stage gears for three rotating speeds

Finally, the second harmonic of second-stage mesh frequency $2f_{m2} = 2290.2$ cannot be found as shown in Figure 8-17. Because the high frequency component in displacement spectrum is far smaller than in the low frequency range. That shows the feature has gone out of the detectable range of vision-based method in this experiment.

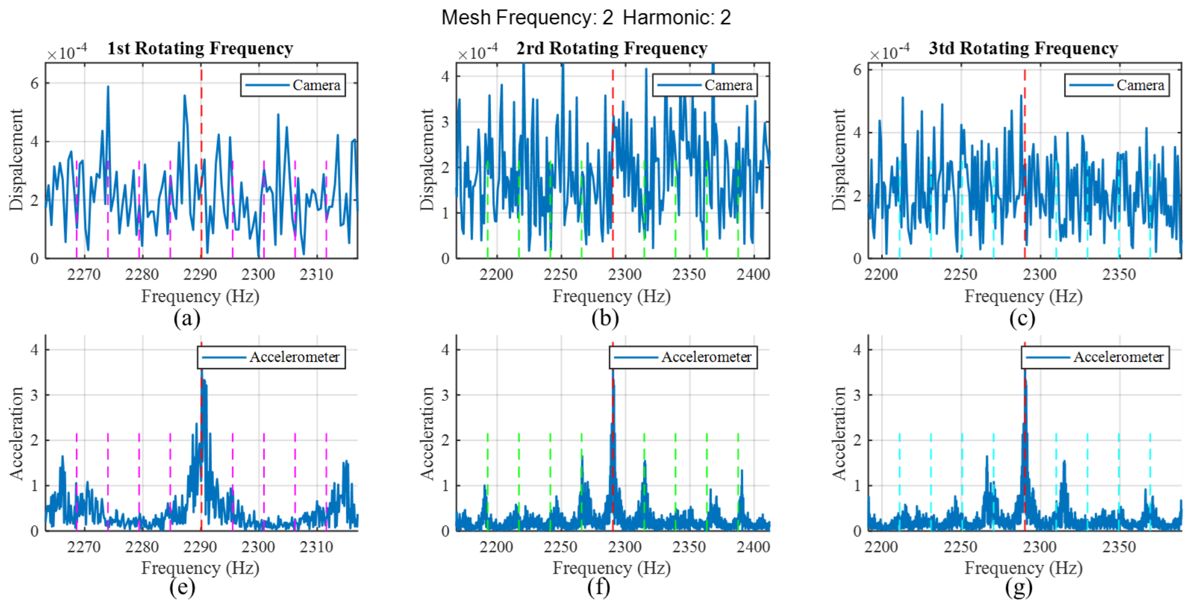



Figure 8-17 TSA spectrum in the frequency range around the second mesh frequency of second-stage gears for three rotating speeds

### 8.2.4 Conclusions of current experiment

In this section, by using the spatial filtering and TSA jointly, the sidebands of first harmonic of first-stage mesh frequency at 316.75Hz are recognized very clear. Decreasingly, the sideband of second-stage mesh frequency at 1145.1Hz can only be recognized barely. However, the second order of second mesh frequency at 2274Hz is too weak to detect.

In summary, although the vision-based method cannot reach the accuracy of acceleration at this phase, the result still shows the capability of vision-based method to diagnose the fault of gearbox. Furthermore, the improvement on hardware (light source and camera) and algorithm can promise vision-based method to reach the accuracy of accelerometer in the future.

## 8.3 Key findings

The vibration of a reciprocating compressor and a gearbox are captured by a camera, and the fault features are extracted from image data. The experiment of compressor shows the LK optical flow can extract the features in low frequency band, like rotating frequency, but not valid in higher frequency range. In contrast, the experiment on gearbox shows the proposed method combined with time-synchronous averaging can manage to extract the sideband of two mesh frequencies. Even through the methods can be improved in many aspects, this early-stage research shows the promise of vision-based condition monitoring.

# 9 CONCLUSIONS AND FUTURE WORKS

*In previous chapters, the research on theory, simulation, and experiment are detailed. In the last chapter, the key points of above parts are summarized.*

## 9.1 Objectives and achievements

This Ph.D. project mainly developed a novel framework based on array signal processing for vision-based vibration measurement. The main achievements are listed corresponding to research objectives.

- **Objective 1**: To set up a novel framework that incorporates from measuring vibration to extracting features (or identifying parameters), which can accomplish the vision-based modal analysis or condition monitoring.
- **Achievement 1**: A PASP framework is proposed that can achieve the vision-based vibration measurement with high-accuracy, through a connected sequence including forming pixel array, array processing and extract features (modal parameter for modal analysis and diagnostic feature for condition monitoring).

- **Objective 2**: To develop algorithms that can enhance SNR of vibration signal, especially for weak signal in high frequency.
- **Achievement 2**:
  - A (MASF) algorithm is developed, which can enhance SNR by the number of pixels in array, and then the detection of weak signal is achieved in numerical simulation and experiment.
  - An SVD-based method is developed, which can extract weak signal without using iterative process.
  - A data-independent spatial filtering is developed for condition monitoring because the vibration wave propagation of rotating machinery tends to be a travelling wave.

- **Objective 3**: To develop algorithms that can achieve accurate and dense estimation of mode shape.
- **Achievement 3**:
  - A Node/antinode searching scheme is developed to implement the proposed MASF, which can obtain the mode shape by searching the location of node/antinode at a pixel level.

    - The proposed SVD-based method can also realize estimation of mode shape, which is applicable to the mode shape that cannot be represented by its node/antinode.

- **Objective 4**: To carry out simulation study by programming and using commercial software, in order to verify theoretical calculation and guide experiments.
- **Achievement 4**: A group of FEM is built by using Ansys, which manage to predict the experiment result. In addition, a synthetic video is created by coding in MATLAB, which is used to verify the proposed equations and predict the experiment result.

- **Objective 5**: To design and implement a set of experiments, in order to validate proposed methods and expand its applications.
- **Achievement 5**: A group of pioneering vision-based experiments are conducted, the main achievements in four experiments are listed respectively.
    - In modal analysis of free-free stainless beam, the natural frequencies and mode shape up to 6388 Hz are identified, which cannot be achieved by conventional methods.
    - In modal analysis of wind turbine blade, the natural frequencies and mode shape are identified without prior knowledge and iterative process, which justify the robustness and applicability on large displacement response (about 0.2 m).
    - In condition monitoring of reciprocating compressor, a simple setup is combined with conventional LK optical flow to measure the rotating speed of motor, that validates the conventional method can monitor the basic parameter such as the rotating speed but not sufficient for high frequency range.
    - In condition monitoring of gearbox, the high-speed camera is combined with proposed framework to diagnose the fault, in which the mesh frequency up to 1145Hz and their sidebands are detected clearly. That shows the proposed framework not only can conduct modal analysis, but also is effective to wideband signal in condition monitoring.

## 9.2 Conclusions

Overall, this PhD project sets up a novel pixel array signal processing framework for vision-based vibration measurement, which is comprised of pixel array formation, pixel array signal processing and the feature extraction (or parameter identification). This framework tends to take pixels as an array of displacement sensors, rather than an image. Furthermore, a serial of approaches under this framework is proposed to meet requirements of various test scenarios. In addition, the simulation and experiment study were conducted to validate the performance of proposed, which shows the measurement accuracy is highly improved than conventional methods. The conclusions of specific approaches are drawn as follow.

- An edge detection-based formation approach is proposed for formulating the pixel array, that can get all sensitive pixels involved in computation, meanwhile exclude all useless pixel from computation, which leads to improving the signal quality in data acquisition.
- A mode-shape-based adaptive spatial filtering approach is proposed, which can obtain the high-SNR modal displacement by using with high computational efficiency. In particular, a node/antinode searching scheme is proposed based on sinusoid-based piecewise functions, which works as an adaptive filter to match mode shape and enhance modal displacement. It shows that the proposed method can identify the weak natural frequency in the high frequency range and estimate the mode shape of continuous structures. Through theoretical analysis, the upper bounder of array gain is found, which is an explicit value equal to the number of array elements. In simulation, the theoretical array gain is verified, and a weak mode with amplitude of 1/2560 pixel is detected through enhancing its SNR by nearly 2000 times. In experiment of a free-free beam, all five modes within frequency range are identified accurately, with the highest frequency being up to 6388 Hz, which cannot be identified with LK optical flow method.
- An SVD-based method is proposed to identify mode shape that cannot be represented by its node/antinode. In the experimental modal analysis of wind turbine blade, the SVD-based method is used to identify the modal parameters, within the frequency up to *250Hz* and the displacement magnitude about *0.2m*. Compared with LK optical flow, the spatial density of mode shape and bandwidth of natural frequency are improved significantly.
- A data independent spatial filtering is proposed for condition monitoring of rotating machinery. In experiment of multistage gearbox, the proposed spatial filtering combines

with time-synchronous signal average to detect the diagnostic feature up to 1145Hz, i.e., the mesh frequency and its sidebands successfully. The result demonstrates that the PASP framework has sufficient capability of monitoring rotating machinery, and the accuracy is promising to increase with the development of algorithm and hardware in the future.

The reason of improvement can be attributed to that the array signal processing is more pertinent to structural vibration than the conventional image registration method. Because the structural vibration is a distinctive motion that all space elements is oscillatiing subtlely around a equilibrium point at the same frequencies. Aiming at this distinctive motion, the method based on array processing can extract and processes vibration information more effectively so as to reduce redundancy of image more efficiently.

However, there is still a development prospect in the robustness and accuracy of the framework, which will be discussed in the section of future work.

## 9.3 Novelties

- A novel framework is first proposed based on the principle of array signal processing. Compared with the conventional framework of window-based image registration, the proposed framework can break through current limit of SNR by overcoming inherent defect of image registration methods.
- An edge detection method is first developed to select pixels for computation. Compared with selecting pixel from a rectangle window, the quality of acquired signal is improved significantly.
- A mode-shape-based adaptive spatial filtering approach is first developed. Moreover, the maximum gain of ASF is found, which cannot be achieved by conventional methods.
- A node/antinode search scheme is fist developed to solve the optimization problem in adaptive spatial filtering with high computational efficiency.
- The vision-based condition monitoring of gearbox is first carried out. The promising results show the potential of extensive applications to vision-based condition monitoring in the future.

## 9.4 Contribution to knowledge

The contribution of this Ph.D. project to knowledge can be summarized into following two aspects.

- This study innovates framework of VVM completely. The existing methods are based on the window-based image registration, which belongs to the methodology of image processing or computer vision. In contrast, the proposed framework stems from the array signal processing, as a result, enhancement of signal power is more significant than conventional framework. Because the constructive interference of signal can only be achieved based on the array processing, but not the principle of image processing.
- The proposed framework expands theory and applications of array signal processing. As a mature discipline, the array signal processing has been applied to almost all sensor arrays. The application on VVM further expands the conception of ASP, especially the weight vector of spatial filter cannot only achieve the phase shifter but also the amplitude adjustment.

## 9.5 Future works

Many future works will be derived from this new framework, the direction can be categorized into improving robustness and accuracy.

- In terms of robustness or applicability, more complex situations require to consider, such as 1) the uncertain boundary condition, 2) 3D measurement for complex-shape objects, 3) the adaptive spatial filtering for wideband sign and so on.
- Regarding to the further improvement on accuracy, 1) the noise model of image, 2) the convexity of optimization, 3) the adaptive filtering for identifying the damping ratio and 4) upgrading the experiment setup etc. are suggested to investigate in the future.

In summary, this pioneering research can inspire more potential works that are worthy of research in the future.

# REFERENCES


[1] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, 'Photogrammetry and optical methods in structural dynamics – A review', *Mechanical Systems and Signal Processing*, vol. 86, pp. 17–34, 2017, doi: 10.1016/j.ymssp.2016.02.011.

[2] R. Brincker, 'Introduction to Operational Modal Analysis', p. 375.

[3] M. Friswell and J. E. Mottershead, *Finite Element Model Updating in Structural Dynamics*. Springer Science & Business Media, 2013.

[4] C. Rainieri and G. Fabbrocino, *Operational Modal Analysis of Civil Engineering Structures*. New York, NY: Springer New York, 2014. doi: 10.1007/978-1-4939-0767-0.

[5] F. P. García Márquez, A. M. Tobias, J. M. Pinar Pérez, and M. Papaelias, 'Condition monitoring of wind turbines: Techniques and methods', *Renewable Energy*, vol. 46, pp. 169–178, 2012, doi: 10.1016/j.renene.2012.03.003.

[6] C. Scheffer and P. Girdhar, *Practical Machinery Vibration Analysis and Predictive Maintenance*. Elsevier, 2004.

[7] R. B. Randall, *Vibration-based Condition Monitoring: Industrial, Automotive and Aerospace Applications*. John Wiley & Sons, 2021.

[8] A. Brandt, *Noise and Vibration Analysis: Signal Analysis and Experimental Procedures*. John Wiley & Sons, 2011.

[9] J. He and Z.-F. Fu, *Modal analysis*. Oxford ; Boston: Butterworth-Heinemann, 2001.

[10] P. Castellini, M. Martarelli, and E. P. Tomasini, 'Laser Doppler Vibrometry: Development of advanced solutions answering to technology's needs', *Mechanical Systems and Signal Processing*, vol. 20, no. 6, pp. 1265–1285, Aug. 2006, doi: 10.1016/j.ymssp.2005.11.015.

[11] M. S. Allen and M. W. Sracic, 'A new method for processing impact excited continuous-scan laser Doppler vibrometer measurements', *Mechanical Systems and Signal Processing*, vol. 24, no. 3, Art. no. 3, Apr. 2010, doi: 10.1016/j.ymssp.2009.11.004.

[12] D. Zhang, M. Kurata, and T. Inaba, 'FMCW Radar for Small Displacement Detection of Vital Signal Using Projection Matrix Method', *International Journal of Antennas and Propagation*, vol. 2013, p. e571986, Dec. 2013, doi: 10.1155/2013/571986.

[13] O. J. Løkberg, 'Electronic Speckle Pattern Interferometry', in *Optical Metrology: Coherent and Incoherent Optics for Metrology, Sensing and Control in Science, Industry and Biomedicine*, O. D. D. Soares, Ed. Dordrecht: Springer Netherlands, 1987, pp. 542–572. doi: 10.1007/978-94-009-3609-6_36.

[14] W.-C. Wang, C.-H. Hwang, and S.-Y. Lin, 'Vibration measurement by the time-averaged electronic speckle pattern interferometry methods', *Appl. Opt., AO*, vol. 35, no. 22, pp. 4502–4509, Aug. 1996, doi: 10.1364/AO.35.004502.

[15] T. Kreis, *Handbook of Holographic Interferometry: Optical and Digital Methods*. John Wiley & Sons, 2006.

[16] W. Schumann and M. Dubas, *Holographic Interferometry: From the Scope of Deformation Analysis of Opaque Bodies*. Springer, 2013.

[17] D.-M. Chen, Y. F. Xu, and W. D. Zhu, 'Damage Identification of Beams Using a Continuously Scanning Laser Doppler Vibrometer System', *Journal of Vibration and Acoustics*, vol. 138, no. 051011, Jun. 2016, doi: 10.1115/1.4033639.

[18] S. J. Rothberg *et al.*, 'An international review of laser Doppler vibrometry: Making light work of vibration measurement', *Optics and Lasers in Engineering*, vol. 99, pp. 11–22, Dec. 2017, doi: 10.1016/j.optlaseng.2016.10.023.

[19] M. Martarelli and D. J. Ewins, 'Continuous scanning laser Doppler vibrometry and speckle noise occurrence', *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 2277–2289, Nov. 2006, doi: 10.1016/j.ymssp.2005.06.003.

[20] A. B. Stanbridge and D. J. Ewins, 'MODAL TESTING USING A SCANNING LASER DOPPLER VIBROMETER', *Mechanical Systems and Signal Processing*, vol. 13, no. 2, pp. 255–270, Mar. 1999, doi: 10.1006/mssp.1998.1209.

[21] J. Wang, X. Wang, L. Chen, J. Huangfu, C. Li, and L. Ran, 'Noncontact Distance and Amplitude-Independent Vibration Measurement Based on an Extended DACM Algorithm', *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 1, pp. 145–153, Jan. 2014, doi: 10.1109/TIM.2013.2277530.

[22] 'An interferometric radar for non-contact measurement of deflections on civil engineering structures: laboratory and full-scale tests: Structure and Infrastructure Engineering: Vol 6, No 5'. https://www.tandfonline.com/doi/abs/10.1080/15732470903068557 (accessed Apr. 13, 2022).

[23] G. Wang, C. Gu, T. Inoue, and C. Li, 'A Hybrid FMCW-Interferometry Radar for Indoor Precise Positioning and Versatile Life Activity Monitoring', *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 11, pp. 2812–2822, Nov. 2014, doi: 10.1109/TMTT.2014.2358572.

[24] Y. Xiong, Z. Peng, G. Xing, W. Zhang, and G. Meng, 'Accurate and Robust Displacement Measurement for FMCW Radar Vibration Monitoring', *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1131–1139, Feb. 2018, doi: 10.1109/JSEN.2017.2778294.

[25] Y. Xiong, S. Chen, G. Xing, Z. Peng, and W. Zhang, ‘High-precision frequency estimation for FMCW radar applications based on parameterized de-alternating and modified ICCD’, *Meas. Sci. Technol.*, vol. 29, no. 7, p. 075010, Jun. 2018, doi: 10.1088/1361-6501/aac3f6.

[26] Y. Xiong, S. Chen, X. Dong, Z. Peng, and W. Zhang, ‘Accurate Measurement in Doppler Radar Vital Sign Detection Based on Parameterized Demodulation’, *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 11, pp. 4483–4492, Nov. 2017, doi: 10.1109/TMTT.2017.2684138.

[27] Y. Xiong, Z. Peng, C. Gu, S. Li, D. Wang, and W. Zhang, ‘Differential Enhancement Method for Robust and Accurate Heart Rate Monitoring via Microwave Vital Sign Sensing’, *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, Art. no. 9, Sep. 2020, doi: 10.1109/TIM.2020.2978347.

[28] G. Busca, A. Cigada, P. Mazzoleni, and E. Zappa, ‘Vibration Monitoring of Multiple Bridge Points by Means of a Unique Vision-Based Measuring System’, *Exp Mech*, vol. 54, no. 2, Art. no. 2, Feb. 2014, doi: 10.1007/s11340-013-9784-8.

[29] B. F. Spencer, V. Hoskere, and Y. Narazaki, ‘Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring’, *Engineering*, Mar. 2019, doi: 10.1016/j.eng.2018.11.030.

[30] D. Feng and M. Q. Feng, ‘Identification of structural stiffness and excitation forces in time domain using noncontact vision-based displacement measurement’, *Journal of Sound and Vibration*, vol. 406, pp. 15–28, Oct. 2017, doi: 10.1016/j.jsv.2017.06.008.

[31] B. K. Oh, J. W. Hwang, Y. Kim, T. Cho, and H. S. Park, ‘Vision-based system identification technique for building structures using a motion capture system’, *Journal of Sound and Vibration*, vol. 356, pp. 72–85, 2015, doi: 10.1016/j.jsv.2015.07.011.

[32] J. G. Chen, A. Davis, N. Wadhwa, F. Durand, W. T. Freeman, and O. Büyüköztürk, ‘Video Camera–Based Vibration Measurement for Civil Infrastructure Applications’, *Journal of Infrastructure Systems*, vol. 23, no. 3, p. B4016013, Sep. 2017, doi: 10.1061/(ASCE)IS.1943-555X.0000348.

[33] J. G. Chen, T. M. Adams, H. Sun, E. S. Bell, and O. Büyüköztürk, ‘Camera-Based Vibration Measurement of the World War I Memorial Bridge in Portsmouth, New Hampshire’, *Journal of Structural Engineering*, vol. 144, no. 11, p. 04018207, Nov. 2018, doi: 10.1061/(ASCE)ST.1943-541X.0002203.

[34] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, ‘Predicting full-field strain on a wind turbine for arbitrary excitation using displacements of optical targets measured with photogrammetry’, in *Special Topics in Structural Dynamics, Volume 6*, Springer, 2015, pp. 99–114.

[35] R. Chabrier, E. Sadoulet-Reboul, G. Chevallier, E. Foltête, and T. Jeannin, ‘Full-field measurements with Digital Image Correlation for vibro-impact characterisation’,

*Mechanical Systems and Signal Processing*, vol. 156, p. 107658, Jul. 2021, doi: 10.1016/j.ymssp.2021.107658.

[36] I. Amenabar, A. Mendikute, A. López-Arraiza, M. Lizaranzu, and J. Aurrekoetxea, 'Comparison and analysis of non-destructive testing techniques suitable for delamination inspection in wind turbine blades', *Composites Part B: Engineering*, vol. 42, no. 5, Art. no. 5, 2011, doi: 10.1016/j.compositesb.2011.01.025.

[37] H. Lee and D. Han, 'Deformation Measurement of a Railroad Bridge Using a Photogrammetric Board without Control Point Survey', *Journal of Sensors*, vol. 2018, pp. 1–10, Jul. 2018, doi: 10.1155/2018/6851252.

[38] J. Javh, J. Slavič, and M. Boltežar, 'The subpixel resolution of optical-flow-based modal analysis', *Mechanical Systems and Signal Processing*, vol. 88, pp. 89–99, 2017.

[39] C. C. Ciang, J.-R. Lee, and H.-J. Bang, 'Structural health monitoring for a wind turbine system: a review of damage detection methods', *Measurement Science and Technology*, vol. 19, no. 12, Art. no. 12, 2008.

[40] Y. Du, S. Zhou, X. Jing, Y. Peng, H. Wu, and N. Kwok, 'Damage detection techniques for wind turbine blades: A review', *Mechanical Systems and Signal Processing*, vol. 141, p. 106445, Jul. 2020, doi: 10.1016/j.ymssp.2019.106445.

[41] W. Fan and P. Qiao, 'Vibration-based Damage Identification Methods: A Review and Comparative Study', *Structural Health Monitoring*, vol. 10, no. 1, Art. no. 1, Jan. 2011, doi: 10.1177/1475921710365419.

[42] L. Yu and B. Pan, 'Single-camera high-speed stereo-digital image correlation for full-field vibration measurement', *Mechanical Systems and Signal Processing*, vol. 94, pp. 374–383, Sep. 2017, doi: 10.1016/j.ymssp.2017.03.008.

[43] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, 'Photogrammetry and optical methods in structural dynamics – A review', *Mechanical Systems and Signal Processing*, vol. 86, pp. 17–34, 2017, doi: 10.1016/j.ymssp.2016.02.011.

[44] M. N. Helfrick, C. Niezrecki, P. Avitabile, and T. Schmidt, '3D digital image correlation methods for full-field vibration measurement', *Mechanical Systems and Signal Processing*, vol. 25, no. 3, pp. 917–927, 2011, doi: 10.1016/j.ymssp.2010.08.013.

[45] D. J. Ewins, *Modal Testing: Theory, Practice and Application*. John Wiley & Sons, 2009.

[46] C. Wei, M. Jin, H. Jiasheng, Y. Chen, and H. Song, 'A method to distinguish harmonic frequencies and remove the harmonic effect in operational modal analysis of rotating structures', *Mechanical Systems and Signal Processing*, vol. 161, p. 107928, Dec. 2021, doi: 10.1016/j.ymssp.2021.107928.

[47] P. Poozesh, A. Sarrafi, Z. Mao, P. Avitabile, and C. Niezrecki, 'Feasibility of extracting operating shapes using phase-based motion magnification technique and

stereo-photogrammetry', *Journal of Sound and Vibration*, vol. 407, pp. 350–366, 2017, doi: 10.1016/j.jsv.2017.06.003.

[48] B. Peeters and G. De Roeck, 'Stochastic System Identification for Operational Modal Analysis: A Review', *Journal of Dynamic Systems, Measurement, and Control*, vol. 123, no. 4, pp. 659–667, Feb. 2001, doi: 10.1115/1.1410370.

[49] 'System Identification Methods for (Operational) Modal Analysis: Review and Comparison | SpringerLink'. https://link.springer.com/article/10.1007/s11831-012-9069-x (accessed Apr. 14, 2022).

[50] D. Feng and M. Q. Feng, 'Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review', *Engineering Structures*, vol. 156, pp. 105–117, Feb. 2018, doi: 10.1016/j.engstruct.2017.11.018.

[51] 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', *Journal of Sound and Vibration*, vol. 421, pp. 300–318, May 2018, doi: 10.1016/j.jsv.2018.01.050.

[52] N. Najafi and U. S. Paulsen, 'Operational modal analysis on a VAWT in a large wind tunnel using stereo vision technique', *Energy*, vol. 125, pp. 405–416, Apr. 2017, doi: https://doi.org/10.1016/j.energy.2017.02.133.

[53] M. Ozbek, F. Meng, and D. J. Rixen, 'Challenges in testing and monitoring the in-operation vibration characteristics of wind turbines', *Mechanical Systems and Signal Processing*, vol. 41, no. 1, pp. 649–666, Dec. 2013, doi: https://doi.org/10.1016/j.ymssp.2013.07.023.

[54] M. Ozbek and D. J. Rixen, 'Operational modal analysis of a 2.5 MW wind turbine using optical measurement techniques and strain gauges', *Wind Energy*, vol. 16, no. 3, Art. no. 3, 2013, doi: 10.1002/we.1493.

[55] Z. Su, J. Pan, S. Zhang, S. Wu, Q. Yu, and D. Zhang, 'Characterizing dynamic deformation of marine propeller blades with stroboscopic stereo digital image correlation', *Mechanical Systems and Signal Processing*, vol. 162, p. 108072, Jan. 2022, doi: 10.1016/j.ymssp.2021.108072.

[56] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, 'Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks', *Journal of Sound and Vibration*, vol. 388, pp. 154–170, Feb. 2017, doi: 10.1016/j.jsv.2016.10.043.

[57] Y. Hu, W. Guo, W. Zhu, and Y. Xu, 'Local damage detection of membranes based on Bayesian operational modal analysis and three-dimensional digital image correlation', *Mechanical Systems and Signal Processing*, vol. 131, pp. 633–648, Sep. 2019, doi: 10.1016/j.ymssp.2019.04.051.

[58] J. Javh, J. Slavič, and M. Boltežar, ‘High frequency modal identification on noisy high-speed camera data’, *Mechanical Systems and Signal Processing*, vol. 98, pp. 344–351, 2018, doi: 10.1016/j.ymssp.2017.05.008.

[59] A. Sabato, P. Poozesh, P. Avitabile, and C. Niezrecki, ‘Experimental modal analysis of a utility-scale wind turbine blade using a multi-camera approach’, 2018, vol. 1149. doi: 10.1088/1742-6596/1149/1/012005.

[60] D. Robinson and P. Milanfar, ‘Fundamental performance limits in image registration’, *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1185–1199, Sep. 2004, doi: 10.1109/TIP.2004.832923.

[61] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer London, 2010. [Online]. Available: https://books.google.co.uk/books?id=8_2RNQEACAAJ

[62] T. C. Chu, W. F. Ranson, and M. A. Sutton, ‘Applications of digital-image-correlation techniques to experimental mechanics’, *Experimental Mechanics*, vol. 25, no. 3, Art. no. 3, Sep. 1985, doi: 10.1007/BF02325092.

[63] P. Bing, X. Hui-min, X. Bo-qin, and D. Fu-long, ‘Performance of sub-pixel registration algorithms in digital image correlation’, *Meas. Sci. Technol.*, vol. 17, no. 6, Art. no. 6, May 2006, doi: 10.1088/0957-0233/17/6/045.

[64] P. Bing, X. Hui-min, X. Bo-qin, and D. Fu-long, ‘Performance of sub-pixel registration algorithms in digital image correlation’, *Meas. Sci. Technol.*, vol. 17, no. 6, Art. no. 6, May 2006, doi: 10.1088/0957-0233/17/6/045.

[65] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, ‘Phase-based video motion processing’, *ACM Transactions on Graphics*, vol. 32, no. 4, p. 1, Jul. 2013, doi: 10.1145/2461912.2461966.

[66] A. J. Molina-Viedma, L. Felipe-Sesé, E. López-Alba, and F. A. Díaz, ‘3D mode shapes characterisation using phase-based motion magnification in large structures using stereoscopic DIC’, *Mechanical Systems and Signal Processing*, vol. 108, pp. 140–155, Aug. 2018, doi: 10.1016/j.ymssp.2018.02.006.

[67] T. Gautama and M. A. V. Hulle, ‘A phase-based approach to the estimation of the optical flow field using spatial filtering’, *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1127–1136, Sep. 2002, doi: 10.1109/TNN.2002.1031944.

[68] B. D. Lucas and T. Kanade, ‘An Iterative Image Registration Technique with an Application to Stereo Vision’, in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, San Francisco, CA, USA, 1981, pp. 674–679. Accessed: Mar. 26, 2019. [Online]. Available: http://dl.acm.org/citation.cfm?id=1623264.1623280

[69] M. Gulde, L. Kortmann, M. Ebert, E. Watson, J. Wilk, and F. Schäfer, ‘Robust optical tracking of individual ejecta particles in hypervelocity impact experiments’, *Meteoritics & Planetary Science*, 2017, doi: 10.1111/maps.12958.

[70] D. Gorjup, J. Slavič, and M. Boltežar, ‘Frequency domain triangulation for full-field 3D operating-deflection-shape identification’, *Mechanical Systems and Signal Processing*, vol. 133, p. 106287, Nov. 2019, doi: 10.1016/j.ymssp.2019.106287.

[71] 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’, *Mechanical Systems and Signal Processing*, vol. 28, pp. 333–347, 2012, doi: 10.1016/j.ymssp.2011.11.023.

[72] P. L. Reu, D. P. Rohe, and L. D. Jacobs, ‘Comparison of DIC and LDV for practical vibration and modal measurements’, *Mechanical Systems and Signal Processing*, vol. 86, pp. 2–16, 2017, doi: 10.1016/j.ymssp.2016.02.006.

[73] J. Huang, K. Zhou, W. Chen, and H. Song, ‘A pre-processing method for digital image correlation on rotating structures’, *Mechanical Systems and Signal Processing*, vol. 152, p. 107494, May 2021, doi: 10.1016/j.ymssp.2020.107494.

[74] J. Carr, J. Baqersad, C. Niezrecki, P. Avitabile, and M. Slattery, ‘Dynamic stress–strain on turbine blade using digital image correlation techniques part 1: static load and calibration’, in *Topics in Experimental Dynamics Substructuring and Wind Turbine Dynamics, Volume 2*, Springer, 2012, pp. 215–220.

[75] J. Carr, J. Baqersad, C. Niezrecki, P. Avitabile, and M. Slattery, ‘Dynamic Stress–Strain on Turbine Blades Using Digital Image Correlation Techniques Part 2: Dynamic Measurements’, in *Topics in Experimental Dynamics Substructuring and Wind Turbine Dynamics, Volume 2: Proceedings of the 30th IMAC, A Conference on Structural Dynamics, 2012*, R. Mayes, D. Rixen, D. T. Griffith, D. De Klerk, S. Chauhan, S. N. Voormeeren, and M. S. Allen, Eds. New York, NY: Springer New York, 2012, pp. 221–226.

[76] K. Zaletelj, V. Agrež, J. Slavič, R. Petkovšek, and M. Boltežar, ‘Laser-light speckle formation for deflection-shape identification using digital image correlation’, *Mechanical Systems and Signal Processing*, vol. 161, p. 107899, Dec. 2021, doi: 10.1016/j.ymssp.2021.107899.

[77] P. Bing, X. Hui-min, X. Bo-qin, and D. Fu-long, ‘Performance of sub-pixel registration algorithms in digital image correlation’, *Meas. Sci. Technol.*, vol. 17, no. 6, pp. 1615–1621, May 2006, doi: 10.1088/0957-0233/17/6/045.

[78] E. Zappa, P. Mazzoleni, and A. Matinmanesh, ‘Uncertainty assessment of digital image correlation method in dynamic applications’, *Optics and Lasers in Engineering*, vol. 56, pp. 140–151, May 2014, doi: 10.1016/j.optlaseng.2013.12.016.

[79] K. Zaletelj, J. Slavič, and M. Boltežar, ‘Full-field DIC-based model updating for localized parameter identification’, *Mechanical Systems and Signal Processing*, vol. 164, p. 108287, Feb. 2022, doi: 10.1016/j.ymssp.2021.108287.

[80] R. Huňady, P. Pavelka, and P. Lengvarský, ‘Vibration and modal analysis of a rotating disc using high-speed 3D digital image correlation’, *Mechanical Systems and Signal Processing*, vol. 121, pp. 201–214, Apr. 2019, doi: 10.1016/j.ymssp.2018.11.024.

[81] A. Yashar, N. Ferguson, and M. G. Tehrani, ‘Measurement of rotating beam vibration using optical (DIC) techniques’, *Procedia Engineering*, vol. 199, pp. 477–482, Jan. 2017, doi: 10.1016/j.proeng.2017.09.052.

[82] Aral Sarrafi, ‘Wind Turbine Blade Damage Detection Via 3-Dimensional Phase-Based Motion Estimation’, 2017.

[83] J. G. Chen, N. Wadhwa, Y.-J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, ‘Modal identification of simple structures with high-speed video using motion magnification’, *Journal of Sound and Vibration*, vol. 345, pp. 58–71, 2015, doi: 10.1016/j.jsv.2015.01.024.

[84] A. Davis, M. Rubinstein, N. Wadhwa, G. J. Mysore, F. Durand, and W. T. Freeman, ‘The visual microphone: passive recovery of sound from video’, *ACM Transactions on Graphics*, vol. 33, no. 4, pp. 1–10, Jul. 2014, doi: 10.1145/2601097.2601119.

[85] M. Eitner, B. Miller, J. Sirohi, and C. Tinney, ‘Effect of broad-band phase-based motion magnification on modal parameter estimation’, *Mechanical Systems and Signal Processing*, vol. 146, p. 106995, Jan. 2021, doi: 10.1016/j.ymssp.2020.106995.

[86] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, ‘Performance of optical flow techniques’, *Int J Comput Vision*, vol. 12, no. 1, pp. 43–77, Feb. 1994, doi: 10.1007/BF01420984.

[87] Y. Yang *et al.*, ‘Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification’, *Mechanical Systems and Signal Processing*, vol. 85, pp. 567–590, Feb. 2017, doi: 10.1016/j.ymssp.2016.08.041.

[88] N. Wadhwa *et al.*, ‘Motion microscopy for visualizing and quantifying small motions’, *PNAS*, vol. 114, no. 44, Art. no. 44, Oct. 2017, doi: 10.1073/pnas.1703715114.

[89] G. Liu, M. Z. Li, Z. Mao, and Q. S. Yang, ‘Structural motion estimation via Hilbert transform enhanced phase-based video processing’, *Mechanical Systems and Signal Processing*, vol. 166, p. 108418, Mar. 2022, doi: 10.1016/j.ymssp.2021.108418.

[90] J. Javh, J. Slavič, and M. Boltežar, ‘Measuring full-field displacement spectral components using photographs taken with a DSLR camera via an analogue Fourier integral’, *Mechanical Systems and Signal Processing*, vol. 100, pp. 17–27, 2018, doi: 10.1016/j.ymssp.2017.07.024.

[91] J. Javh, J. Slavič, and M. Boltežar, ‘Experimental modal analysis on full-field DSLR camera footage using spectral optical flow imaging’, *Journal of Sound and Vibration*, vol. 434, pp. 213–220, Nov. 2018, doi: 10.1016/j.jsv.2018.07.046.

[92] M. Li, G. Feng, R. Deng, F. Gao, F. Gu, and A. D. Ball, ‘Structural vibration mode identification from high-speed camera footages using an adaptive spatial filtering

approach', *Mechanical Systems and Signal Processing*, vol. 166, p. 108422, Mar. 2022, doi: 10.1016/j.ymssp.2021.108422.

[93] W.-D. Wirth, *Radar techniques using array antennas*, no. 10. IET, 2001.

[94] T. J. Rouphael, *Wireless Receiver Architectures and Design: Antennas, RF, Synthesizers, Mixed Signal, and Digital Signal Processing*. Academic Press, 2014.

[95] J. Benesty, J. Chen, Y. Huang, and J. Dmochowski, 'On Microphone-Array Beamforming From a MIMO Acoustic Signal Processing Perspective', *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, Art. no. 3, Mar. 2007, doi: 10.1109/TASL.2006.885251.

[96] J. Benesty, 'Fundamentals of Signal Enhancement and Array Signal Processing', p. 468.

[97] 'Directivity', *Wikipedia*. Mar. 22, 2022. Accessed: Apr. 19, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Directivity&oldid=1078594183

[98] R. Even, 'How Phased Array Antenna Manufacturers Benefit from Rapid Prototyping'. https://www.nano-di.com/blog/2019-how-phased-array-antenna-manufacturers-benefit-from-rapid-prototyping (accessed Apr. 05, 2022).

[99] R. L. Haupt and Y. Rahmat-Samii, 'Antenna Array Developments: A Perspective on the Past, Present and Future', *IEEE Antennas and Propagation Magazine*, vol. 57, no. 1, Art. no. 1, Feb. 2015, doi: 10.1109/MAP.2015.2397154.

[100] A. Sleiman and A. Manikas, 'Antenna array manifold: a simplified representation', in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*, Jun. 2000, vol. 5, pp. 3164–3167 vol.5. doi: 10.1109/ICASSP.2000.861209.

[101] S. S. Moghaddam and M. S. Moghaddam, 'A comprehensive survey on antenna array signal processing', *Trends in Applied Sciences Research*, vol. 6, no. 6, p. 507, 2011.

[102] D. Lashi, Q. Quévy, and J. Lemeire, 'Optimizing Microphone Arrays for Delay-and-Sum Beamforming using Genetic Algorithms', in *2018 4th International Conference on Cloud Computing Technologies and Applications (Cloudtech)*, Nov. 2018, pp. 1–5. doi: 10.1109/CloudTech.2018.8713331.

[103] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Springer Science & Business Media, 2008.

[104] A. Meyer and D. Döbler, 'NOISE SOURCE LOCALIZATION WITHIN A CAR INTERIOR USING 3D-MICROPHONE ARRAYS', p. 8.

[105] P. Poozesh, K. Aizawa, C. Niezrecki, J. Baqersad, M. Inalpolat, and G. Heilmann, 'Structural health monitoring of wind turbine blades using acoustic microphone array', *Structural Health Monitoring*, vol. 16, no. 4, pp. 471–485, 2017.

[106] J.-M. Valin, F. Michaud, J. Rouat, and D. Létourneau, ‘Robust Sound Source Localization Using a Microphone Array on a Mobile Robot’, *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, vol. 2, pp. 1228–1233, 2003, doi: 10.1109/IROS.2003.1248813.

[107] M. Li, R. Appadoo, W. Hu, F. Gu, and A. Ball, ‘Condition Monitoring of Reciprocating Compressor Based on Acoustic Imaging’, in *Advances in Asset Management and Condition Monitoring*, Cham, 2020, pp. 977–983. doi: 10.1007/978-3-030-57745-2_80.

[108] S. U. Pillai, *Array Signal Processing*. Springer Science & Business Media, 2012.

[109] H. Krim and M. Viberg, ‘Two decades of array signal processing research: the parametric approach’, *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, Jul. 1996, doi: 10.1109/79.526899.

[110] J. Benesty, ‘Fundamentals of Signal Enhancement and Array Signal Processing’, p. 468.

[111] B. D. V. Veen and K. M. Buckley, ‘Beamforming: a versatile approach to spatial filtering’, *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, Apr. 1988, doi: 10.1109/53.665.

[112] J. Li and P. Stoica, ‘Robust Adaptive Beamforming’, p. 30.

[113] J. C. Chen, Kung Yao, and R. E. Hudson, ‘Source localization and beamforming’, *IEEE Signal Processing Magazine*, vol. 19, no. 2, Art. no. 2, Mar. 2002, doi: 10.1109/79.985676.

[114] D. R. Fuhrmann and G. S. Antonio, ‘Transmit beamforming for MIMO radar systems using signal cross-correlation’, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 1, Art. no. 1, Jan. 2008, doi: 10.1109/TAES.2008.4516997.

[115] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, ‘Convex Optimization-Based Beamforming’, *IEEE Signal Processing Magazine*, vol. 27, no. 3, Art. no. 3, May 2010, doi: 10.1109/MSP.2010.936015.

[116] S. S. Haykin, *Adaptive filter theory*, Fifth edition. Upper Saddle River, New Jersey: Pearson, 2014.

[117] J. Li and P. Stoica, ‘Robust Adaptive Beamforming’, p. 30.

[118] P. S. Naidu, *Sensor array signal processing*. Boca Raton, Fla.: CRC Press, 2001.

[119] R. C. Hibbeler, *Engineering Mechanics: Dynamics*. Prentice Hall, 2010.

[120] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, ‘Gaussian approximations of fluorescence microscope point-spread function models’, *Appl. Opt., AO*, vol. 46, no. 10, pp. 1819–1829, Apr. 2007, doi: 10.1364/AO.46.001819.

[121] M. J. Usher and D. A. Keating, *Sensors and Transducers: Characteristics, Applications, Instrumentation, Interfacing*. Macmillan International Higher Education, 1996.

[122] 'Error function', *Wikipedia*. Apr. 15, 2022. Accessed: Apr. 16, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Error_function&oldid=1082784677

[123] 'Cumulative distribution function', *Wikipedia*. Apr. 04, 2022. Accessed: Apr. 22, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Cumulative_distribution_function&oldid=1080969035

[124] J. C. Pedro and N. B. Carvalho, *Intermodulation Distortion in Microwave and Wireless Circuits*. Artech House, 2003.

[125] M. H. Rashid, *Power Electronics Handbook*. Butterworth-Heinemann, 2017.

[126] M. Li, G. Feng, F. Gu, and A. Ball, 'Investigating into Minimum Detectable Displacement Signal in Image-Based Vibration Measurement', in *Proceedings of IncoME-V & CEPE Net-2020*, Cham, 2021, pp. 882–894. doi: 10.1007/978-3-030-75793-9_82.

[127] K. M. Gharaibeh, *Nonlinear Distortion in Wireless Systems: Modeling and Simulation with MATLAB*. John Wiley & Sons, 2011.

[128] W. Kester, 'Understand SINAD, ENOB, SNR, THD, THD + N, and SFDR so You Don't Get Lost in the Noise Floor', p. 8.

[129] N. Wadhwa *et al.*, 'Eulerian video magnification and analysis', *Commun. ACM*, vol. 60, no. 1, Art. no. 1, Dec. 2016, doi: 10.1145/3015573.

[130] S. Cygert and A. Czyzewski, 'Eulerian motion magnification applied to structural health monitoring of wind turbines', *The Journal of the Acoustical Society of America*, vol. 144, no. 3, Art. no. 3, Sep. 2018, doi: 10.1121/1.5067923.

[131] R. C. Hibbeler, *Engineering Mechanics: Dynamics*. Prentice Hall, 2010.

[132] 'Norm (mathematics)', *Wikipedia*. Apr. 02, 2022. Accessed: Apr. 17, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Norm_(mathematics)&oldid=1080664860

[133] 'Vector and matrix norms - MATLAB norm - MathWorks United Kingdom'. https://uk.mathworks.com/help/matlab/ref/norm.html (accessed Apr. 17, 2022).

[134] R. J. Allemang, 'The Modal Assurance Criterion – Twenty Years of Use and Abuse', *SOUND AND VIBRATION*, p. 8, 2003.

[135] J. E. Mottershead and M. I. Friswell, ‘Model Updating In Structural Dynamics: A Survey’, *Journal of Sound and Vibration*, vol. 167, no. 2, pp. 347–375, Oct. 1993, doi: 10.1006/jsvi.1993.1340.

[136] J. E. Mottershead, M. Link, and M. I. Friswell, ‘The sensitivity method in finite element model updating: A tutorial’, *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2275–2296, Oct. 2011, doi: 10.1016/j.ymssp.2010.10.012.

[137] u/mennyaka, ‘Standing wave as a result of reflection’, *GeoGebra*, Nov. 16, 2016. https://www.geogebra.org/m/jbeCtgzx (accessed Apr. 07, 2022).

[138] ‘Standing Wave’. http://www.cmm.gov.mo/eng/exhibition/secondfloor/MoreInfo/2_11_0_StandingWave.html (accessed Apr. 17, 2022).

[139] ‘224 Physics Lab: Standing Waves’. http://science.clemson.edu/physics/labs/labs/224/standwave/index.html (accessed Apr. 05, 2022).

[140] ‘Orthogonal basis’, *Wikipedia*. Mar. 18, 2022. Accessed: Apr. 20, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Orthogonal_basis&oldid=1077835316

[141] ‘Fourier series’, *Wikipedia*. Apr. 16, 2022. Accessed: Apr. 20, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Fourier_series&oldid=1083029985

[142] M. A. S. Shohag, E. C. Hammel, D. O. Olawale, and O. I. Okoli, ‘Damage mitigation techniques in wind turbine blades: A review’, *Wind Engineering*, vol. 41, no. 3, pp. 185–210, 2017.

[143] R. Wu, D. Zhang, Q. Yu, Y. Jiang, and D. Arola, ‘Health monitoring of wind turbine blades in operation using three-dimensional digital image correlation’, *Mechanical Systems and Signal Processing*, vol. 130, pp. 470–483, Sep. 2019, doi: 10.1016/j.ymssp.2019.05.031.

[144] A. Arda Ozdemir and S. Gumussoy, ‘Transfer Function Estimation in System Identification Toolbox via Vector Fitting’, *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6232–6237, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.1026.

[145] M. Li, X. Wang, F. Liu, X. Tang, F. Gu, and A. D. Ball, ‘Remote Monitoring of Wind Turbine Blades based on High-speed Photogrammetry’, *Condition Monitoring and Diagnostic Engineering Management*, p. 321, 2018.

[146] E. Reynders and G. De Roeck, ‘Reference-based combined deterministic–stochastic subspace identification for experimental and operational modal analysis’, *Mechanical Systems and Signal Processing*, vol. 22, no. 3, pp. 617–637, 2008.

[147] Y. Ham, K. K. Han, J. J. Lin, and M. Golparvar-Fard, ‘Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs): a

review of related works', *Visualization in Engineering*, vol. 4, no. 1, Art. no. 1, Jan. 2016, doi: 10.1186/s40327-015-0029-z.

[148] A. Khadka, B. Fick, A. Afshar, M. Tavakoli, and J. Baqersad, 'Non-contact vibration monitoring of rotating wind turbines using a semi-autonomous UAV', *Mechanical Systems and Signal Processing*, vol. 138, p. 106446, Apr. 2020, doi: 10.1016/j.ymssp.2019.106446.

[149] T. Lendzioch, J. Langhammer, and M. Jenicek, 'Estimating Snow Depth and Leaf Area Index Based on UAV Digital Photogrammetry', *Sensors*, vol. 19, no. 5, Art. no. 5, Jan. 2019, doi: 10.3390/s19051027.

[150] G. Pajares, 'Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs)', Apr. 2015. https://www.ingentaconnect.com/content/asprs/pers/2015/00000081/00000004/art00015 (accessed May 23, 2019).

[151] G. Morgenthal and N. Hallermann, 'Quality Assessment of Unmanned Aerial Vehicle (UAV) Based Visual Inspection of Structures', *Advances in Structural Engineering*, vol. 17, no. 3, Art. no. 3, Mar. 2014, doi: 10.1260/1369-4332.17.3.289.

[152] M. Li, F. Gu, A. Ball, and L. Song, 'A homography transformation based video stabilization method for structural health monitoring using a unmanned aerial vehicle: Sixteenth International Conference on Condition Monitoring and Asset Management', *Sixteenth International Conference on Condition Monitoring and Asset Management (CM 2019)*, vol. 1, Aug. 2019, Accessed: Oct. 07, 2021. [Online]. Available: http://www.proceedings.com/49622.html

[153] G. Feng, N. Hu, Z. Mones, F. Gu, and A. D. Ball, 'An investigation of the orthogonal outputs from an on-rotor MEMS accelerometer for reciprocating compressor condition monitoring', *Mechanical Systems and Signal Processing*, vol. 76, pp. 228–241, 2016.

[154] L. Lucchese and S. K. Mitra, 'Using saddle points for subpixel feature detection in camera calibration targets', in *Asia-Pacific Conference on Circuits and Systems*, Oct. 2002, vol. 2, pp. 191–195 vol.2. doi: 10.1109/APCCAS.2002.1115151.

[155] B. D. Lucas and T. Kanade, 'An Iterative Image Registration Technique with an Application to Stereo Vision', in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, San Francisco, CA, USA, 1981, pp. 674–679. Accessed: Mar. 25, 2019. [Online]. Available: http://dl.acm.org/citation.cfm?id=1623264.1623280

[156] J.-Y. Bouguet, 'Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm', p. 9.

[157] R. Zhang, F. Gu, H. Mansaf, T. Wang, and A. D. Ball, 'Gear wear monitoring by modulation signal bispectrum based on motor current signal analysis', *Mechanical Systems and Signal Processing*, vol. 94, pp. 202–213, Sep. 2017, doi: 10.1016/j.ymssp.2017.02.037.