

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**FULL-FIELD DISPLACEMENT ESTIMATION AND
RECONSTRUCTION OF LARGE-SCALE
STRUCTURES BASED ON DIGITAL IMAGE
PHASE**

WANG YUCHAO

PhD

The Hong Kong Polytechnic University

**This programme is jointly offered by The Hong Kong
Polytechnic University and Harbin Institute of Technology**

2026

The Hong Kong Polytechnic University
Department of Civil and Environmental Engineering

Harbin Institute of Technology
School of Intelligent Civil and Ocean Engineering

**FULL-FIELD DISPLACEMENT ESTIMATION AND
RECONSTRUCTION OF LARGE-SCALE
STRUCTURES BASED ON DIGITAL IMAGE
PHASE**

WANG Yuchao

A thesis submitted in partial fulfillment of the requirements for the degree
of **Doctor of Philosophy**

January 2026

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it produces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

WANG Yuchao (Name of student)

ABSTRACT

Displacements are widely used to quantify the deformation performance of large-scale structures such as long-span bridges and high-rise buildings. However, full-field displacement estimation is rarely achieved with traditional sensors and algorithms, thereby limiting the evaluation of deformation performance in large-scale structures. The rapidly developed digital image displacement estimation method has the potential to monitor structure displacements. However, this method faces challenges, including complex environments (extra textures, varying lighting conditions, and noise), low resolution for large-scale structures, and others. This thesis addresses three challenges in digital image-based displacement estimation: complex environments, full-field displacement estimation of large-scale structures, and displacement reconstruction from limited measurements.

First, the adverse effects of complex environments on the structural phase are analyzed and suppressed. The field depth and complex filters are optimized to extract the out-of-the-plane and in-plane local phases, respectively. A Log-Gabor-based local phase extraction method is proposed to estimate displacements under extra textures and varying lighting conditions. The adverse effects of digital image noise on the estimated displacements are analyzed in the complex domain. The structural phase is denoised and reconstructed in the multi-frequency complex domain, thereby improving the accuracy of estimated displacements. An amplitude instability indicator is designed to quantify the noise threshold, achieving phase denoising under unknown noise distributions. Numerical and experimental studies are conducted under different conditions to demonstrate the performance of the present method.

Second, to estimate the full-field displacements of large-scale structures with sufficient measurement range and accuracy, the extracted phase quality under different amplitudes is analyzed. The multi-resolution complex pyramid is constructed based on complex filters, enabling adaptive displacement estimation across multi-scale phases. The robust

discriminative correlation is used to adaptively extract multi-texture targets of large-scale structures. Experimental and in-field studies demonstrate the performance of the proposed method. The high-spatial-resolution displacements and mode shapes of the Hong Kong Tsing Ma Bridge and Guangdong Humen Bridge are estimated.

Third, the digital image phase-based displacements and measured accelerations of the large-scale high-rises are utilized to reconstruct the full-field quasi-static and dynamic displacements, respectively. The pseudo-inverse operation is introduced to reconstruct the mean wind-induced quasi-static displacements, while the stochastic subspace identification is introduced to reconstruct the fluctuating wind-induced displacements. The full-field, high-spatial-resolution interstory drift of the Shenzhen Kingkey 100 high-rise structure is reconstructed by summarizing the quasi-static and dynamic displacements.

PUBLICATIONS

Journal Papers

1. **Wang, Y.**, Hu, W., Teng, J., Xia, Y., 2023. Phase-based motion estimation in complex environments using the illumination-invariant log-gabor filter. Mech. Syst. Signal Process. 186, 109847.
2. **Wang, Y.**, Hu, W., Teng, J., Xia, Y., 2024a. Motion estimation from noisy data with unknown distributions using multi-frame phase-preserving denoising. Mech. Syst. Signal Process. 206, 110924.
3. **Wang, Y.**, Hu, W., Teng, J., Xia, Y., 2024b. Full-field displacement measurement of long-span bridges using one camera and robust self-adaptive complex pyramid. Mech. Syst. Signal Process. 215, 111451.
4. Teng, J., **Wang, Y.**, Xia, Y., Hu, W., 2025. Robust discriminative correlation-based full-field motion estimation of large-scale structures using a single video camera. Eng. Struct. 334, 120224.

ACKNOWLEDGEMENTS

Six years have elapsed since I composed the acknowledgments for my Master's thesis at Harbin Institute of Technology, Shenzhen (HITSZ). Throughout this period, I have navigated some life choices and diverse experiences. I am profoundly grateful to have reached the threshold of graduation, owing to the steadfast support and guidance from my supervisors, colleagues, friends, and family.

Foremost, I extend my deepest gratitude to Professor Xia Yong for his exceptional mentorship and support during my joint Ph.D. programme at The Hong Kong Polytechnic University (PolyU). Professor Xia treated me with remarkable kindness, akin to that of his own student, ensuring my academic and personal well-being in Hong Kong. His advice to adopt a progressive research approach was instrumental during my initial phase of the project. This guidance culminated in the successful validation of my methods on the Tsing Ma Bridge. Furthermore, Professor Xia dedicated meticulous effort to revising my manuscripts and presentation materials line by line. The copious annotations in his feedback serve as tangible markers of my scholarly development, which I shall preserve with great care.

I am profoundly indebted to my supervisor at HITSZ, Professor Teng Jun, for granting me the opportunity to pursue my Ph.D. Professor Teng provided invaluable counsel on defining my doctoral research direction and conceptual framework, significantly enhancing the academic rigor and practical relevance of my dissertation. His recommendation for the PolyU joint training programme substantially broadened my expertise, perspective, and professional competencies. Crucially, throughout both my Master's and Ph.D. studies, Professor Teng consistently emphasized a problem-driven research philosophy. This principle has fundamentally anchored my approach to problem identification, research design, communication, and scholarly presentation.

I offer sincere thanks to my co-supervisor in HITSZ, Professor Hu Weihua, for his meticulous guidance in refining my research methodology. Professor Hu adeptly synthesized the research foundations of HITSZ and PolyU, assisting me in crystallizing key research questions, identifying innovative contributions, and establishing logical coherence across dissertation chapters. On the technical front, he facilitated access to essential equipment and experimental sites, enabling the practical implementation of my theoretical work. Under the joint supervision of Professors Teng and Hu, I was privileged to engage with numerous major engineering projects, which provided invaluable platforms for the empirical validation of my research.

Regarding career decisions. I am deeply grateful to Professors Xia, Teng, and Hu for their extraordinary patience, understanding, and unwavering support during this challenging period. I aspire to approach my future career with resolute confidence and integrity, upholding a commitment to principled conduct and meaningful contribution, regardless of the professional path I pursue.

I also extend my sincere appreciation to Professors Li Zuohua and Lu Wei of the HITSZ research group. Professor Li provided substantial support during my Master's studies and the initial phase of my Ph.D., actively encouraging my exploration of novel research directions. Professor Lu significantly enhanced my presentation skills through his guidance in structuring and preparing research reports. The contributions of both Professors Li and Lu were indispensable to the development and completion of my doctoral dissertation.

My journey over these six years would not have been possible without the companionship and assistance of fellow students. I am particularly grateful to those who enriched my two years in Hong Kong: Senior colleagues Drs. Li Lingfang, Wang Xiaoyou, Wu Wanglin, Ding Zhenhao, Tian Wei, Zhang Lu, Li Yixian, Wang Fei; and peers Shan Yushi, Zhang Yuxuan, Hu Yunxin, and Lu Tian. During the height of the pandemic in Hong Kong, I am especially thankful to my fellow joint-training student Hong Shunfa at PolyU for our mutual support and camaraderie.

I also thank senior HITSZ colleagues Tang Dehui and Liu Nanxi, alongside Xu Zengmao, Yuan Cheng, Deng Quanxue, Gong Yunxuan, Yang Qitao, and Ning Jiafei, for their support at various stages. My gratitude extends to junior colleagues Tu Shanjiu, Yang Sen, Ma Longjun, Tang Wenchao, and other Master students within the research group for their assistance.

Finally, I owe my deepest gratitude to my parents. They have provided an exceptional foundation for my growth. Throughout my adulthood, from selecting my undergraduate institution, deciding to transfer from Electrical Engineering to Civil Engineering, choosing to resign from my job to pursue a Ph.D., to determining my current career path, they have consistently respected my decisions and offered unwavering support, despite any pressures they may have faced. It is my sincere hope to fulfill my responsibilities and repay their boundless devotion in the years ahead.

CONTENTS

CERTIFICATE OF ORIGINALITY	i
ABSTRACT.....	iii
PUBLICATIONS	v
ACKNOWLEDGEMENTS	vii
CONTENTS	xi
LIST OF FIGURES	xv
LIST OF TABLES	xxi
LIST OF SYMBOLS	xxiii
CHAPTER 1 INTRODUCTION	1
1.1 Research background	1
1.2 Research objectives	5
1.3 Thesis organization	6
CHAPTER 2 LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Displacement estimation under complex environments.....	9
2.2.1 Contact displacement estimation methods.....	10
2.2.2 Non-contact displacement estimation methods	13
2.3 Full-field displacement estimation and reconstruction of large-scale structures	26

2.3.1 Full-field displacement estimation of large-scale structures.....	26
2.3.2 Full-field displacement reconstruction of large-scale structures	29
2.4 Summary and challenges	31

CHAPTER 3 PHASE-BASED DISPLACEMENT ESTIMATION UNDER COMPLEX ENVIRONMENTS35

3.1 Introduction.....	35
3.2 Phase and subpixel displacement projection in complex environments.....	36
3.2.1 Phase-based subpixel displacement estimation.....	36
3.2.2 Mechanisms of complex environment effects on the extracted phase...37	
3.2.3 LGME method in complex environments.....	45
3.3 Performance of LGME in complex numerical environments.....	49
3.3.1 Performance evaluation in complex 1D signals.....	49
3.3.2 2D signals: in-plane extra textures and illumination variance	51
3.3.3 2D signals: out-of-the-plane extra texture occlusions.....	56
3.4 Performance of LGME in complex experimental environments.....	59
3.4.1 In-plane extra textures and illumination variations.....	59
3.4.2 Out-of-the-plane extra texture occlusions	67
3.5 Summary.....	71

CHAPTER 4 DISPLACEMENT ESTIMATION FROM NOISY DATA WITH UNKNOWN DISTRIBUTIONS USING PHASE-PRESERVING DENOISING.....73

4.1 Introduction.....	73
4.2 Phase-preserving denoising under unknown noise distributions.....	74
4.2.1 Effects of noise on phase-based displacement estimation	74
4.2.2 Denoising and phase preservation using the complex pyramid	76
4.2.3 Multi-frame phase-preserving denoising (MPD)	83
4.3 Numerical examples	84
4.3.1 Data generation	84
4.3.2 Phase denoising performance	85
4.3.3 Displacement denoising performance	90
4.4 Experimental studies.....	92

4.4.1 Factors affecting the noise and experimental setup	92
4.4.2 Displacement denoising under different ISO levels	95
4.5 Summary	100
 CHAPTER 5 FULL-FIELD DISPLACEMENT ESTIMATION BASED ON COMPLEX PYRAMID AND DISCRIMINATIVE CORRELATION. 103	
5.1 Introduction	103
5.2 Multi-scale displacement estimation based on multi-resolution complex pyramid	104
5.2.1 Phase extraction of multi-scale displacements	104
5.2.2 Multi-resolution complex pyramid based on complex filters	108
5.2.3 Multi-resolution complex pyramid-based displacement estimation	109
5.2.4 Experimental study	113
5.3 In-field study 1: Full-field displacement estimation of the Tsing Ma Bridge	119
5.3.1 Measurement environment	119
5.3.2 Filter optimization	121
5.3.3 Linear measurement pixel selection	121
5.3.4 Displacement results	124
5.3.5 Full-field displacement results	127
5.4 In-field study 2: Multi-scale displacement estimation of the Kingkey 100 high- rise structure	131
5.4.1 Measurement environment	131
5.4.2 Torsional mode-induced axis rotation and baseline correction	133
5.4.3 Displacement results	134
5.5 Multi-type target displacement estimation using discriminative correlation	136
5.5.1 Multi-type target detection based on discriminative correlation	136
5.5.2 Compact model	139
5.5.3 Compact samples	141
5.5.4 Experimental study on the four-story steel frame	144
5.6 In-field study 3: Multi-type target displacement estimation of the Humen Bridge	148
5.7 Summary	152

CHAPTER 6 FULL-FIELD DISPLACEMENT RECONSTRUCTION BY FUSING LIMITED DIGITAL IMAGE PHASE AND ACCELERATION MEASUREMENTS.....	155
6.1 Introduction.....	155
6.2 Full-field displacement reconstruction theory fusing limited phase-based displacements and accelerations.....	156
6.2.1 Pseudo-inverse-based quasi-static displacement reconstruction.....	156
6.2.2 Modal decomposition-based dynamic modal displacement reconstruction.....	159
6.2.3 Full-field displacement reconstruction fusing limited phases and accelerations.....	164
6.3 Performance of DFDR under different numerical cases.....	166
6.3.1 Effects of the number of truncated singular values on quasi-static displacement reconstruction.....	167
6.3.2 Effects of the measurement points on quasi-static displacement reconstruction.....	170
6.3.3 Effects of the excitation frequency on quasi-static displacement reconstruction.....	174
6.3.4 Effects of the excitation type on modal dynamic displacement reconstruction.....	176
6.3.5 Effects of the noise on modal dynamic displacement reconstruction ..	182
6.4 In-field study: Full-field displacement reconstruction of the Kingkey 100 ..	187
6.4.1 Response characteristic analysis under typhoons ..	188
6.4.2 Full-field displacement reconstruction based on DFDR.....	199
6.5 Summary.....	208
CHAPTER 7 CONCLUSIONS AND FUTURE WORK.....	211
7.1 Conclusions.....	211
7.2 Future Work.....	213
REFERENCES	215

LIST OF FIGURES

Figure 1.1 Distribution of super high-rise buildings in China.....	1
Figure 1.2 The number of bridges in China.....	2
Figure 1.3 Organization of Thesis	8
Figure 2.1 Challenges in obtaining full-field displacements	33
Figure 3.1 Accurate phase extraction under extra textures.....	38
Figure 3.2 Gabor filters with different parameters in the spatial and frequency domains	41
Figure 3.3 Adverse effects of extra textures on estimating 1D signal displacements	43
Figure 3.4 Displacements estimated by Gabor filters with different central frequencies	44
Figure 3.5 Disturbed 1D signal and filter envelopes with different parameters.....	45
Figure 3.6 Gabor and Log-Gabor filters in the frequency domain	47
Figure 3.7 Procedure of LGME	48
Figure 3.8 1D signals with different solar-induced time-variant pixel values.....	49
Figure 3.9 Localization indices of different Log-Gabor filters in 1-D signal cases	50
Figure 3.10 Displacements estimated by different methods under complex environments	51
Figure 3.11 Structure to be measured with in-plane extra texture in the image plane ...	52
Figure 3.12 Dataset 2 with in-plane extra textures and illumination variations	52
Figure 3.13 Displacements of the structure to be measured	53
Figure 3.14 <i>Loc</i> indexes calculated on the first frame of the 2D dataset.....	54
Figure 3.15 Estimated displacements with in-plane extra textures	54
Figure 3.16 Displacements estimated by traditional Gabor filters under in-plane extra textures and illumination variations.....	55
Figure 3.17 Displacements estimated by LGME under in-plane extra textures and illumination variations	56
Figure 3.18 Rendered model with extra textures along the light axis	57
Figure 3.19 Dataset frames under different apertures.....	58

Figure 3.20 Displacements estimated by LGME under extra textures along the light axis	59
Figure 3.21 Experimental set-up with in-plane extra textures and illumination variations	60
Figure 3.22 <i>Loc</i> values calculated on the first frame of the captured video	62
Figure 3.23 Case 1: displacements estimated under neighbor extra texture in the image plane	62
Figure 3.24 Zoom-in of Case 1	63
Figure 3.25 Case 2: subtle displacements estimated under neighbor extra texture.....	64
Figure 3.26 Illumination variations in Case 3	65
Figure 3.27 Case 3: displacements estimated under neighbor extra texture in the image plane and illumination variations	66
Figure 3.28 Zoom-in of Case 3	66
Figure 3.29 PSD of Case 3	67
Figure 3.30 Extra texture along the light axis	68
Figure 3.31 Captured structure under different apertures	69
Figure 3.32 Estimated displacements under different apertures	70
Figure 4.1 1D signal and the filtered results by convolving the log-Gabor filter	77
Figure 4.2 1D signal and the amplitude map of the filtered signal	78
Figure 4.3 Responses by applying filters with different central frequencies	79
Figure 4.4 Original signal, complex quadrature filter bank, and amplitude map of the complex pyramid.....	80
Figure 4.5 Procedure of the MPD	83
Figure 4.6 Rendered image and different noisy images	85
Figure 4.7 Clean and 50 noisy images (Gaussian)	85
Figure 4.8 Images denoised by different methods (Gaussian noise).....	86
Figure 4.9 Phase variation maps for different noises	87
Figure 4.10 Gradient variation maps for different noises	87
Figure 4.11 Phase preservation performance under different noises	88
Figure 4.12 Phase preservation performance of denoising methods under different noise levels.....	89
Figure 4.13 Thresholds at the first three levels of the pyramid.....	90
Figure 4.14 Displacement estimated from images denoised by different methods	91

Figure 4.15 Relative errors of estimated motions.....	92
Figure 4.16 Experimental setup	95
Figure 4.17 Relative errors of estimated displacements under ISO 800	96
Figure 4.18 Estimated displacements from videos denoised by different methods under ISO 800	97
Figure 4.19 Relative errors of estimated displacements under ISO 1600	98
Figure 4.20 Estimated displacements from videos denoised by different methods under ISO 1600	98
Figure 4.21 Relative errors of estimated displacements under ISO 3200	99
Figure 4.22 Estimated displacements from videos denoised by different methods under ISO 3200	100
Figure 5.1 Moving 1D signals and corresponding amplitudes	107
Figure 5.2 Amplitude values of the moving 1D signal filtered by different filters	107
Figure 5.3 Procedure of the multi-scale motion estimation based on the multi-resolution complex pyramid	113
Figure 5.4 Experimental setup	114
Figure 5.5 <i>Loc</i> values of Points 1 and 2.....	115
Figure 5.6 Measurement pixels of Point 1	116
Figure 5.7 Displacements from the camera and accelerometers	117
Figure 5.8 PSD of the displacement estimated from the video clip and accelerometer	118
Figure 5.9 Mode shapes estimated from the video clip	118
Figure 5.10 Comparison between the estimated mode shape and the FEM result	119
Figure 5.11 Measurement environment of the Tsing Ma Bridge	120
Figure 5.12 Measurement points	120
Figure 5.13 Localization indices at Points 2 and 5 of Tsing Ma Bridge	121
Figure 5.14 Textures at Point 2.....	122
Figure 5.15 Pixel selection with the linear phase at measurement point 2	123
Figure 5.16 Textures at Point 5.....	124
Figure 5.17 Pixel selection with the linear phase at measurement point 5	124
Figure 5.18 Displacements (mm) from the video clip (Points 1 – 5)	125
Figure 5.19 Displacements from the video clip and the GPS (Point 2).....	126
Figure 5.20 Displacements from the video clip and the GPS (Point 5).....	126

Figure 5.21 Video of Tsing Ma Bridge captured by F/4 aperture under climate change	127
Figure 5.22 Displacements estimated from the proposed method, the traditional phase-based method, and the GPS at the mid span under climate changes.....	127
Figure 5.23 Full-field measurement points	128
Figure 5.24 Displacements of the main cable and girder	129
Figure 5.25 PSD of the displacements from the video clip and GPS stations	130
Figure 5.26 Mode shapes from the video clip and the FEM	131
Figure 5.27 Measurement setup of the Kingkey 100 high-rise building.....	132
Figure 5.28 Illumination variations (per minute) of videos captured on the 96 th floor of the Kingkey 100	132
Figure 5.29 Original horizontal displacement estimated by the phase-based method on the 96 th floor under Typhoon Maliksi	133
Figure 5.30 PSD of the horizontal displacement estimated by the phase-based method on the 96 th floor under Typhoon Maliksi	134
Figure 5.31 Quasi-static N-S displacements estimated by the phase-based method (96 th floor) and the GNSS (top floor) under Typhoon Maliksi.....	135
Figure 5.32 Quasi-static E-W displacements estimated by the phase-based method (96 th floor) and the GNSS (top floor) under Typhoon Yagi	136
Figure 5.33 Basic process of implementing discriminative correlation.....	138
Figure 5.34 Model parameter compacting using the factorized convolution operator .	140
Figure 5.35 Updating process of the Gaussian mixture model	142
Figure 5.36 Procedure of the motion estimation for multi-type targets based on the discriminative correlation.....	144
Figure 5.37 Experimental setup	145
Figure 5.38 Sensors used to measure structural responses.....	145
Figure 5.39 Measured displacements on each floor	146
Figure 5.40 PSD of the displacements estimated from the video	147
Figure 5.41 Displacements estimated by different methods under different amplitudes	148
Figure 5.42 Frame of the surveillance video	149
Figure 5.43 Measurement points of the bridge	149
Figure 5.44 Estimated displacements of the bridge	150

Figure 5.45 PSD of the estimated displacements	151
Figure 5.46 Normalized mode shape magnitudes estimated from the video and the accelerometers (0.3662 Hz)	152
Figure 6.1 Procedure of DFDR.....	166
Figure 6.2 Displacements comparison between the reconstructed and FEM results under different numbers of singular values	168
Figure 6.3 Reconstructed interstory drift ratios under different numbers of singular values	170
Figure 6.4 Cases with different numbers of measurement points.....	171
Figure 6.5 Interstory drift ratios comparison between the reconstructed and FEM results under different numbers of measurement points	172
Figure 6.6 Cases with different measurement point placements	173
Figure 6.7 Interstory drift ratios comparison between the reconstructed and FEM results under different measurement point placements	174
Figure 6.8 Interstory drift ratios comparison between the reconstructed and FEM results	175
Figure 6.9 Reconstructed and simulated accelerations of the 80 th floor under white noise inputs.....	177
Figure 6.10 Reconstructed and simulated displacements of the 80 th floor under white noise	178
Figure 6.11 Reconstructed and simulated accelerations of the 80 th floor under fluctuating wind	180
Figure 6.12 Reconstructed and simulated displacements of the 80 th floor under fluctuating winds.....	181
Figure 6.13 Reconstructed and simulated displacements of the 80 th floor under different- level low-frequency noises	183
Figure 6.14 Reconstructed and simulated displacements of the 80 th floor under different- level high-frequency noises	185
Figure 6.15 Reconstructed and simulated displacements of the 80 th floor under different- level white noises.....	186
Figure 6.16 Kingkey 100 high-rise structure	188
Figure 6.17 Wireless distributed SHM system and sensors.....	188
Figure 6.18 GNSS displacement sensors	189

Figure 6.19 Measurement point layout	190
Figure 6.20 The surveillance camera and laser beam target on the 80 th floor.....	191
Figure 6.21 The interstory deformation measurement point on the 80 th floor.....	191
Figure 6.22 The interstory deformation of the steel frame on the 96 th floor.....	192
Figure 6.23 Frames captured on the 80 th and 96 th floors.....	193
Figure 6.24 Comparison between the estimated and LDV-measured displacements under Typhoon Maliksi	195
Figure 6.25 PSD of measured displacements and accelerations under Typhoon Maliksi	197
Figure 6.26 Quasi-static and dynamic displacements measured by the video and the LDV.....	198
Figure 6.27 Reconstructed and measured quasi-static displacements of Kingkey 100 top floor under typhoons	200
Figure 6.28 Maximum N-S full-field quasi-static horizontal displacement reconstruction of Kingkey 100 under Typhoon Maliksi	201
Figure 6.29 Maximum E-W full-field quasi-static horizontal displacement reconstruction of Kingkey 100 under Typhoon Yagi.....	202
Figure 6.30 Reconstructed and measured N-S accelerations and displacements of the 96 th floor under Typhoon Maliksi	203
Figure 6.31 Reconstructed and measured E-W accelerations and displacements of the 96 th floor under Typhoon Yagi.....	205
Figure 6.32 Maximum N-S full-field dynamic displacement reconstruction of Kingkey 100 under Typhoon Maliksi	206
Figure 6.33 Maximum E-W full-field dynamic displacement reconstruction of Kingkey 100 under Typhoon Yagi.....	206
Figure 6.34 Maximum N-S full-field displacement reconstruction of Kingkey 100 under Typhoon Maliksi	207
Figure 6.35 Maximum E-W full-field displacement reconstruction of Kingkey 100 under Typhoon Yagi.....	207
Figure 6.36 Ratio of dynamic to total displacements of the top floor under typhoons.	208

LIST OF TABLES

Table 1.1 Measured peak accelerations and estimated dynamic displacements of high-rise buildings under typhoons	3
Table 3.1 Relative errors of the displacements estimated by different methods in Case 1	63
Table 3.2 Relative errors of the displacements estimated by different methods in Case 2	65
Table 3.3 Relative errors of the displacements estimated by different methods in Case 3	67
Table 3.4 Relative errors of the displacements estimated under different apertures	70
Table 5.1 Measured frequencies (Hz) using video clips and accelerometers	130
Table 5.2 Natural frequency measured by the camera, laser sensor, and accelerometer	147
Table 5.3 Relative errors of the different methods on the second floor	148
Table 6.1 Relative errors of the reconstructed quasi-static displacements regarding different singular value numbers	169
Table 6.2 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different singular value numbers	170
Table 6.3 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different measurement point numbers	172
Table 6.4 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different measurement point placements.....	174
Table 6.5 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different excitation frequencies	176
Table 6.6 Relative errors of the reconstructed dynamic displacements regarding different low-frequency noises	184
Table 6.7 Relative errors of the reconstructed dynamic displacements regarding different-level high-frequency noises	185

Table 6.8 Relative errors of the reconstructed dynamic displacements regarding different-level white noises	185
--	-----

LIST OF SYMBOLS

Symbols

x_0	Horizontal coordinate of target center in the reference frame
y_0	Vertical coordinate of target center in the reference frame
x_1	Horizontal coordinate of displaced target center
y_1	Vertical coordinate of displaced target center
$\mathbf{P}(t)$	Coordinate of the target at time t
$\mathbf{I}(\cdot)$	Pixel intensity of the target
$\nabla \mathbf{I}$	Spatial gradient of the pixel intensity
$\mathbf{u}(t)$	Time history of the horizontal displacement
$\mathbf{v}(t)$	Time history of the vertical displacement
$\nabla \mathbf{u}$	Spatial gradient of the horizontal displacement among pixels
$\nabla \mathbf{v}$	Spatial gradient of the vertical displacement among pixels
ϕ	Phase of the target
c	Constant phase value
ϕ_t	Phase deviation in the time domain
ϕ_p	Spatial gradient of the phase
R	Regularization term
λ	Parameters of the regularization term
$\mathbf{R}_I(x, y)$	Real part of the filtered image in the spatial domain

$Z_I(x, y)$	Imaginary part of the filtered image in the spatial domain
$I(x, y)$	Pixel intensity values
$G_{\text{even}}(x, y)$	Real part of the complex filter in the spatial domain
$G_{\text{odd}}(x, y)$	Imaginary part of the complex filter in the spatial domain
w	Spatial support (spatial size) of the complex filter
$\phi(x, y, t)$	Extracted phase values
$D(t)$	Structure displacements
$u(t)$	Horizontal displacements
$v(t)$	Vertical displacements
α_p	Scaling factor
$g(x, y)$	Gabor filter in the spatial domain
f_1	Central frequency of the Gabor filter in the vertical direction
f_2	Central frequency of the Gabor filter in the horizontal direction
j	Imaginary unit;
$w_r(x, y)$	Spatial envelope of the Gabor filter
K	Scaling coefficient unrelated to the phase
σ	Envelope standard deviation perpendicular to the target edge
σ_n	Envelope standard deviation parallel to the target edge
β	Half-magnitude bandwidth of the Gabor filter
f_0	Central frequency of the Gabor filter
b	Pixel value variation induced by illumination changes
a	Integration of the real part filter in the spatial domain

$\mathbf{LG}(f)$	Log-Gabor filter in the frequency domain
f_0	Central frequency of the Log-Gabor filter
σ_r	Standard deviation of the radial envelope in the logarithmic scale
θ	Direction in the polar coordinate
θ_0	Direction of the measured target
σ_θ	Standard deviation of the circumferential envelope
Loc	Localization index
$\mathbf{T}(x, y)$	True image
$\mathbf{n}(x, y)$	Additive noise
$\mathbf{R}_T(x, y)$	Real part of the clean structural feature in the complex domain
$\mathbf{R}_n(x, y)$	Real part of the noise in the complex domain
$\mathbf{Z}_T(x, y)$	Imaginary part of the clean structural feature in the complex domain
$\mathbf{Z}_n(x, y)$	Imaginary part of the noise in the complex domain
$\phi(x, y)$	Phase of the noisy structural feature
$\mathbf{R}_I^m(x, y)$	Real part of the complex signal at level m
$\mathbf{Z}_I^m(x, y)$	Imaginary part of the complex signal at level m
$\mathbf{G}_{\text{even}}^m(x, y)$	Real part of the complex filter (even function) at level m
$\mathbf{G}_{\text{odd}}^m(x, y)$	Imaginary part of the complex filter (odd function) at level m
$\hat{\mathbf{R}}_I^m(x, y)$	Real part of the denoised complex signal at level m
T_m	Noise threshold at level m
$\hat{\mathbf{I}}(x, y)$	Denoised image signal
S_m	Amplitude instability indicator of the complex signal at level m

A_m^1	Level m amplitude of the reference frame
A_m^2	Level m amplitude of the current frame
W	Width of the ROI
H	Height of the ROI
$\text{median}(\cdot)$	Median value operator
a_m	Amplitude value of the content of interest at level m
$\Delta\phi(x, y)$	Noise-induced phase variations
δ	Phase-preserving indicator, smaller values mean more stable phases
C	A cluster of pixels belonging to the measurement region
W	Width of the measurement region
H	Height of the measurement region
ε	Relative error of displacement
Δu	Deviation between the estimated displacement and GT
u	GT of the displacement
$\ \cdot\ $	Euclidean norm
$SNR_{\text{shotnoise}}$	SNR of the shot noise
α_{ISO}	Gain of the PGA
κ	Initial light intake
g_x	Horizontal gradient calculated by Sobel operator
g_y	Vertical gradient calculated by Sobel operator
$LG^m(f, \theta)$	Log-Gabor filter at Level m
f_m	Central frequency at Level m
σ_m	Standard deviation of the radial envelope at Level m

$\phi_l^m(x, y, t)$	Phase at Level m
$A_l^m(x, y, t)$	Amplitude at Level m
j	Imaginary unit
$\text{FFT}(\cdot)$	Fourier transform
$\text{IFFT}(\cdot)$	Inverse Fourier transform
$\odot$	Dot product
$\nabla^2(\cdot)$	Laplacian operator
$\text{div}(\cdot)$	Divergence operator
d_θ	Direction parallel with the target edge
$d_{\bar{\theta}}$	Direction orthogonal to the target edge
$\nabla \phi_l^m$	Spatial phase gradient at Level m
D_m	Amplitude degradation index of Level m at time t
η	A factor deciding to what extent the amplitude value variation is regarded as degradation in a pixel patch with a size of 5×5
$\mathbf{u}_s^b$	Estimated quasi-static displacement after baseline correction
$\mathbf{u}_s^g$	Quasi-static displacement from GNSS
$\mathbf{u}_s^c$	Estimated quasi-static displacement
$E(\mathbf{f})$	Target function of the filter optimization
$\mathbf{S}_f(\mathbf{g}_j)$	Filter result of the j^{th} frame, $\mathbf{S}_f(\mathbf{g}_j) \in \mathbb{R}^{W \times H}$
$\mathbf{g}_j$	Geometric features of the j^{th} frame, $\mathbf{g}_j \in \mathbb{R}^{D \times W \times H}$
$\mathbf{g}_j^d$	d^{th} -channel geometric feature of the j^{th} frame, $\mathbf{g}_j^d \in \mathbb{R}^{W \times H}$
$\mathbf{y}_j$	2D Gaussian label, $\mathbf{y}_j \in \mathbb{R}^{W \times H}$

$\mathbf{f}$	Filters to be optimized, $\mathbf{f} \in \mathbb{R}^{D \times W \times H}$
$\mathbf{f}^d$	d^{th} -channel filter to be optimized, $\mathbf{f}^d \in \mathbb{R}^{W \times H}$
$\otimes$	Convolution operator
D	Number of the channels in the geometric features
$J\{\cdot\}$	Transformer of the discrete sample to a continuous function
T_g	Upper range of the continuous function, set to be 1
N_d	Element number of the d^{th} -channel geometric features
t_g	Variable of the continuous function within $[0, T_g]$
$b_d(\cdot)$	A cubic spline kernel used to interpolate the continuous signal by superpositioning shifted versions of $b_d(\cdot)$
$\mathbf{S}_f\{\cdot\}$	Continuous filter result
$\mathbf{S}_p\{\cdot\}$	Filter result of the compact model
C	Channel number of the compact model
$\mathbf{P}$	Factorized convolution operator, $\mathbf{P} \in \mathbb{R}^{D \times C}$
$p_{d,c}$	Element in the factorized convolution operator
$E(\mathbf{f}, \mathbf{P})$	Target function to obtain the factorized convolution operator
$\hat{\cdot}$	Frequency-domain transformer
$\mathbf{G}_{\text{mix}}(\mathbf{g})$	Gaussian mixture model with L Gaussian components
$\mathbf{G}_l(\mathbf{g}, \boldsymbol{\mu}_l)$	l^{th} Gaussian component with a mean of $\boldsymbol{\mu}_l$ and a standard deviation of unit matrix
$\boldsymbol{\mu}_l$	Mean of the l^{th} Gaussian component, $\boldsymbol{\mu}_l \in \mathbb{R}^{W \times H}$
η_l	Weight of the l^{th} Gaussian component
L	Number of compact samples, $L < m$, which is 50 in this study

γ	Initial learning rate, 0.012 in this study
μ_n	The next sample
μ_k	Existing sample which is most similar to the next sample
η_n	Weight of the next sample, which is the initial learning rate
η_k	Weight of the mixed sample
$\mathbf{m}$	Mass matrix, $\mathbf{m} \in \mathbb{R}^{n \times n}$
$\mathbf{k}$	Stiffness matrix, $\mathbf{k} \in \mathbb{R}^{n \times n}$
$\mathbf{c}$	Damping matrix, $\mathbf{c} \in \mathbb{R}^{n \times n}$
$\mathbf{u}(t)$	Displacement vector, $\mathbf{u} \in \mathbb{R}^{n \times 1}$
$\mathbf{W}(t)$	Load vector, $\mathbf{W} \in \mathbb{R}^{n \times 1}$
$\tilde{\mathbf{W}}(t)$	Equivalent load vector
$\mathbf{D}$	Flexibility matrix
$\mathbf{D}_{m \times n}^+$	Pseudo inverse of the flexibility matrix
$\mathbf{Y}_{n \times 1}$	Error term
$\mathbf{Q}_1$	Left positive definite matrix, $\mathbf{Q}_1 \in \mathbb{R}^{m \times m}$
$\mathbf{Q}_2$	Right positive definite matrix, $\mathbf{Q}_2 \in \mathbb{R}^{n \times n}$
Σ	Non-positive definite singular value matrix, $\Sigma \in \mathbb{R}^{m \times n}$
$\tilde{\mathbf{D}}_{m \times n}$	Truncated flexibility matrix with a rank number of r
$\mathbf{q}_1^r$	r^{th} column of $\mathbf{Q}_1$, $\mathbf{q}_1^r \in \mathbb{R}^{m \times 1}$
$\mathbf{q}_2^r$	r^{th} row of $\mathbf{Q}_2$, $\mathbf{q}_2^r \in \mathbb{R}^{1 \times n}$
σ_r	r^{th} diagonal element of Σ
$\mathbf{x}_k$	State vector, $\mathbf{x}_k \in \mathbb{R}^{2n \times 1}$

$\mathbf{y}_k$	Output vector, $\mathbf{y}_k \in \mathbb{R}^{m \times 1}$
$\mathbf{W}_k$	Input vector, $\mathbf{W}_k \in \mathbb{R}^{n \times 1}$
Δt	Sampling period
$\mathbf{A}$	System matrix, $\mathbf{A} \in \mathbb{R}^{2n \times 2n}$
$\mathbf{B}$	Controllability matrix, $\mathbf{B} \in \mathbb{R}^{2n \times n}$
$\mathbf{C}$	Observability matrix, $\mathbf{C} \in \mathbb{R}^{m \times 2n}$
$\mathbf{D}_I$	Input feedthrough matrix, $\mathbf{D}_I \in \mathbb{R}^{m \times n}$
$\mathbf{w}_k$	System error, $\mathbf{w}_k \in \mathbb{R}^{2n \times 1}$
$\mathbf{v}_k$	Measurement noises, $\mathbf{v}_k \in \mathbb{R}^{m \times 1}$
$\mathbf{Y}_F$	Stacked future output matrix
$\mathbf{Y}_p$	Stacked past output matrix
$\mathbf{P}_i^{\text{ref}}$	Projection of the future output matrix on the past output matrix
$\mathbf{O}_i$	Extended observability matrix
$\hat{\mathbf{X}}_i$	Stacked Kalman state matrix
$\hat{\mathbf{x}}_i$	Kalman state vector
i	Time lag
L	Number of columns of the assembled output matrix
$\mathbf{U}_1$	Left matrix of SVD
$\mathbf{V}_1$	Right matrix of SVD
$\mathbf{S}_1$	Singular value matrix
$\mathbf{V}_{1i}$	Residuals induced by system errors
$\mathbf{V}_{2i}$	Residuals induced by measurement noises

$\mathbf{Q}$	Covariance matrix of noises
$\mathbf{S}$	Covariance matrix of noises
$\mathbf{R}$	Covariance matrix of noises
Ψ	Mode shapes
Λ_d	Natural frequencies in the discrete time domain
$\mathbf{V}$	Observed mode shapes
Λ_c	Natural frequencies in the continuous time domain
$\mathbf{x}_{m,k}$	State vector in the modal coordinate
$\mathbf{K}_m$	Kalman gain in the modal coordinate
$\mathbf{K}$	Kalman gain
$\mathbf{e}_k$	Innovation
$\mathbf{G}$	Covariance between the output and the next time-step state
$\mathbf{R}_e$	Covariance of the innovation
Σ	Covariance of the state
$\mathbf{R}_0$	Covariance of the output
$\hat{\mathbf{y}}_{m,k}^j$	j^{th} modal acceleration at time k
$\mathbf{v}_j$	j^{th} mode shape
$\mathbf{x}_{m,k}^j$	Contribution of j^{th} mode at time k
$\hat{\mathbf{y}}_m^j$	j^{th} modal acceleration
J	Number of modes
$\mathbf{u}_d$	Full-field modal dynamic displacement
$\mathbf{u}_s$	Measured quasi-static displacements
$\mathbf{a}_d$	Measured dynamic accelerations

$\mathbf{u}$	Measured displacements
$\mathbf{a}$	Measured accelerations
χ	Order of the Butterworth filter
f_t	Cutoff frequency of quasi-static responses
$\hat{\mathbf{u}}_s$	Full-field quasi-static displacements, $\hat{\mathbf{u}}_s \in \mathbb{R}^{n \times 1}$
m_d	Number of displacement measurement points
$\hat{\mathbf{u}}_d$	Full-field modal dynamic displacements, $\hat{\mathbf{u}}_d \in \mathbb{R}^{n \times 1}$
$\hat{\mathbf{a}}_m^j$	Full-field j^{th} modal acceleration, $\hat{\mathbf{a}}_m^j \in \mathbb{R}^{n \times 1}$
$\hat{\mathbf{u}}$	Reconstructed full-field displacements, $\hat{\mathbf{u}} \in \mathbb{R}^{n \times 1}$
$V(z)$	Mean wind speed at height z
z_{10}	10-m height
α	Terrain roughness index, 0.22 for category C
z	Height (m)

Abbreviations

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
SHM	Structural Health Monitoring
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
FBG	Fiber Bragg Grating
LVDT	Linear Variable Differential Transformer

LDV	Laser Doppler Vibrometry
CW	Continuous Wave
FMCW	Frequency Modulated Continuous Wave
DIC	Digital Image Correlation
UAV	Unmanned Aerial Vehicle
IOF	Intensity-based Optical Flow
POF	Phase-based Optical Flow
DC	Direct Current
TV	Total Variation
BM3D	Block Matching and 3D
DnCNN	Denoising Convolutional Neural Network
FEM	Finite Element Model
PSD	Power Spectral Density
PSNR	Peak-Signal-to-Noise Ratio
GT	Ground Truth
PGA	Programmable Gain Amplifier
SNR	Signal-to-Noise Ratio
HOG	Histogram of Oriented Gradient
DL	Deep Learning
ROI	Region of Interest
MAC	Modal Assurance Criterion
SVD	Singular Value Decomposition
UAV	Unmanned Aerial Vehicle

CHAPTER 1

INTRODUCTION

1.1 Research background

The rapid development of China's economy and society has been accompanied by substantial construction of large-scale infrastructure, including high-rise buildings and long-span bridges. In 2024 alone, the construction area of buildings in China reached 136.83 billion square meters, while the completed building area amounted to 34.37 billion square meters (Zhao et al., 2025). Notably, super-tall structures exceeding 200 meters in height are more concentrated in the typhoon-prone southeastern coastal regions in China (Chen, 2017; Tang, 2024), as shown in Figure 1.1. On the other hand, the number of bridges in China has grown steadily with the expansion of transportation networks. As illustrated in Figure 1.2, by the end of 2023, China possessed a total of 10,239 super-long-span bridges and 178,000 long-span bridges (Ministry of Transport of the People's Republic of China, 2024).

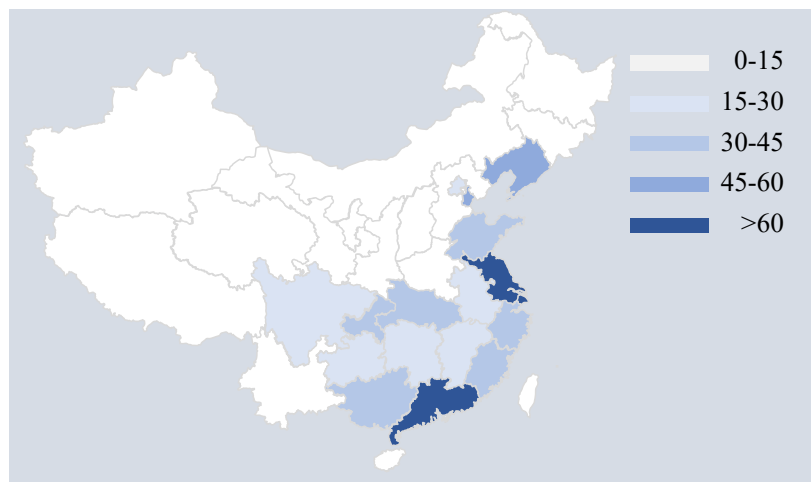


Figure 1.1 Distribution of super-tall buildings in China

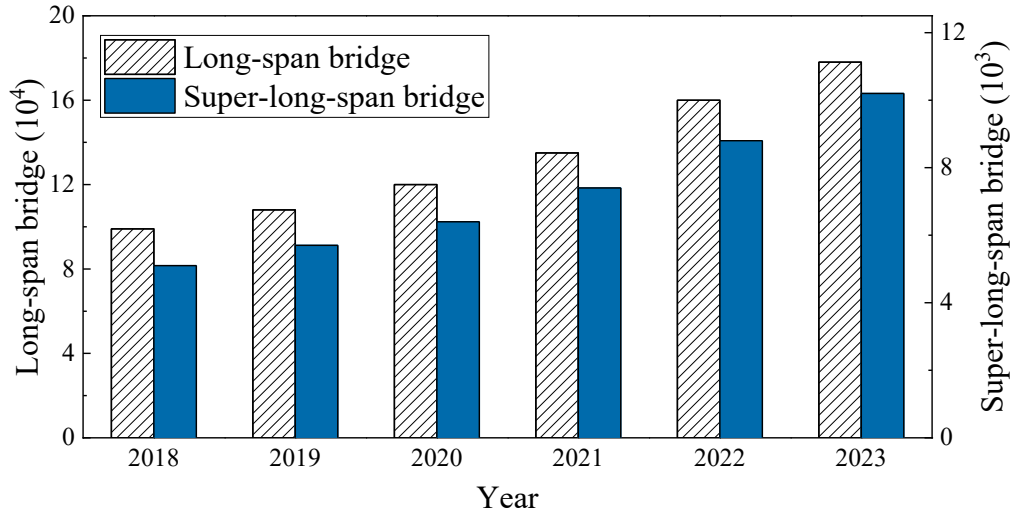


Figure 1.2 The number of bridges in China

The structural performance of large-scale structures is critical to economic security, social security, and life safety. Table 1.1 summarizes the peak accelerations and identified/estimated peak displacements of typical super-tall structures under typhoon conditions. For instance, during Typhoon Mangkhut, the top-floor peak acceleration of the Shenzhen Kingkey 100 was recorded as 0.30 m/s^2 , with an estimated peak displacement of 0.95 m (Liu, 2020). Assuming uniform inter-story drift distribution across all floors, the calculated inter-story drift angle reached 0.00215 , exceeding the elastic limit of 0.002 specified for high-rise structures in China's technical specification for concrete structures of tall buildings (Ministry of Housing and Urban-Rural Development of the People's Republic of China, 2010). Furthermore, even under normal operating conditions, large-scale structures may exhibit significant responses due to the coupling of environmental actions with structural degradation or configuration modifications. A representative case is the vertical vortex-induced vibration of the Humen Bridge in 2020, which experienced vertical displacements of up to 20 cm under moderate winds following alterations to its water-filled barrier arrangement (Zhao and Ge, 2020).

Table 1.1 Measured peak accelerations and estimated dynamic displacements of high-rise buildings under typhoons

High-rise structures	Height (m)	Typhoon	Peak acceleration (m/s ²)	Peak displacement (m)
Shanghai Tower (Wu et al., 2019; Yang et al., 2025)	632	Muifa	0.16	0.23
Ping An Finance Center (He, 2020)	599	Mangkhut	0.25	0.45
Kingkey 100 (Bian, 2019; Liu, 2020)	442	Mangkhut	0.30	0.95
International Finance Centre, Hong Kong (Zhou and Li, 2022)	420	Mangkhut	0.22	0.31
Di Wang Tower (Wu et al., 2007)	384	Rammasun	0.20	0.18
A Tower in Shenzhen (Zhang et al., 2020)	350	Mangkhut	0.14	0.16

Chinese design codes have established requirements for performance indicators related to deformation in large-scale structures, such as the interstory drift ratio for high-rise buildings (Ministry of Housing and Urban-Rural Development of the People's Republic of China, 2010) and the deflection limit for bridges (Ministry of Transport of the People's Republic of China, 2022). However, practical evaluation of the deformation performance of large-scale structures remains challenging, primarily in the following two aspects:

First, obtaining densely distributed measurement data of large-scale structures is a prerequisite for evaluating structural deformation performance (Xu and Xia, 2011; Yu et al., 2021). Contact-based sensors are constrained by the limited number of sensors, making it difficult to achieve high spatial resolution in displacement measurements (Hu

et al., 2020). Non-contact displacement measurement systems (Castellini et al., 2006) have demonstrated considerable potential for directly measuring multi-point displacements on structures, which, however, requires non-contact systems to be installed at significant distances from the target structure, degrading the measurement accuracy of active systems such as laser and radar sensors. The sampling frequency of total stations is insufficient to meet the high-frequency sampling requirements for dynamic displacements in large structures (Yu et al., 2017). Furthermore, these active non-contact measurement techniques typically require sophisticated hardware and complex operation, limiting their practicality. In recent years, rapidly developing vision-based techniques utilize readily available video data to identify structural displacements (Spencer et al., 2019; Xu and Brownjohn, 2018). However, existing vision-based methods struggle to achieve a satisfactory compromise between accuracy and measurement range. Therefore, it is imperative to develop high-precision vision-based methods capable of estimating full-field displacements for large-scale structures. Simultaneously, for scenarios where direct multi-point displacement identification is infeasible, it is necessary to integrate sparse measurement data with structural mechanics models to reconstruct the full-field displacement of large structures.

Second, the complex environments surrounding large-scale structures impose limitations on the full-field displacement estimation performance (Luo et al., 2025; Y. Wang et al., 2023). The actual monitoring environments for such structures are intricate, posing significant challenges for vision-based displacement estimation methods. For example, time-varying illumination conditions and noises in the imaging environment alter the pixel grayscale values of the structure's video frames that are unrelated to its mechanical behavior (Zhang et al., 2017). The presence of closely adjacent extra texture within the structure's imaging plane, or occluding extra texture within the depth of field, hinders the extraction and tracking of structural features (Hur and Roth, 2017). Moreover, mathematical interpolation of discrete pixels in digital images cannot guarantee reliable displacement estimation under these complex environmental conditions. Therefore, it is necessary to develop methods for accurately extracting the geometric features of the target structure from digital images. These methods must separate disturbances caused by temporal illumination variations and spatial interfering objects affecting the structure's

geometric features. Furthermore, establishing a linear relationship between the phase values of discrete pixels at multiple measurement points and the actual structural displacement (Kong et al., 2025) is crucial for enhancing both the accuracy and robustness of digital image-based methods in estimating full-field displacements under complex environments.

Consequently, developing high-precision, easily deployable, and highly robust displacement measurement techniques; establishing methods for estimating and reconstructing the full-field displacement of large-scale structures; and analyzing the deformation and displacement response characteristics of these structures under real operational conditions are critical prerequisites for accurately evaluating their deformation performance.

1.2 Research objectives

This study aims to develop a vision-based motion estimation method with higher precision and robustness, facilitating the full-field displacement measurement of large-scale structures. The research objectives are:

- 1) To develop high-precision phase-based motion estimation methods under complex environments.
 - Investigate the adverse effect of complex environments (the in-plane disturbance, the out-of-the-plane disturbance, and the time-varying illumination condition) on the phase quality;
 - Optimize the complex steerable filter to promote the high-performance phase extraction under complex environments;
 - Evaluate the noise level and denoise the extracted phase under noise with unknown distributions.
- 2) To develop the full-field displacement estimation method for large-scale structures:

- Construct the complex steerable filter bank with multiple spatial sizes to preserve structural displacements with different amplitudes;
 - Design evaluation index to select pixels with reliable phase quality to extract structural displacements at different measurement points;
 - Develop a robust feature detection method to measure multiple-point displacements of large-scale structures with various features.
- 3) To develop the full-field displacement reconstruction method for large-scale structures:
- Reconstruct full-field quasi-static displacements with limited displacement measurements;
 - Reconstruct full-field dynamic displacements with limited acceleration measurements;
 - Fuse limited displacement and acceleration measurements to reconstruct full-field interstory drift ratios of high rises.

1.3 Thesis organization

The thesis is organized into seven chapters, as shown in Figure 1.3. Chapter 1 introduces the research background and objectives and outlines the thesis. Chapter 2 reviews the related literature on displacement estimation under complex environments, full-field displacement estimation, and full-field displacement reconstruction of large-scale structures. Chapter 3 develops the phase-based displacement estimation method under complex environments. Chapter 4 improves the accuracy of phase and displacement estimates using a multi-frame phase-preserving denoiser. Chapter 5 develops a full-field displacement estimation method for large-scale structures based on the self-adaptive complex pyramid and discriminative correlation and applies it to the Tsing Ma Bridge and the Humen Bridge. In Chapter 6, quasi-static displacements and dynamic accelerations are fused to reconstruct full-field displacements of a large-scale structure (Kingkey 100 Building).

In detail, Chapter 3 develops the phase-based displacement estimation method under complex environments. The adverse effects of the large spatial size and the DC component of filters on the performance of the motion estimation method under complex environments are theoretically analyzed. The Log-Gabor filter optimized using the localization index is adopted to extract the local phase under extra textures and varying illuminations, yielding accurate displacements under complex environments.

In Chapter 4, the phase-based displacement estimation method is improved through multi-frame phase-preserving denoising for noisy data with unknown distributions. Multi-scale and multi-orientation quadrature filters transform noisy images into complex pyramids to preserve the phase of the content. An instability indicator is designed to evaluate the noise threshold and to filter noise with unknown distributions in the simulation and experiment examples.

Chapter 5 facilitates the full-field motion estimation of large-scale structures. The multi-resolution complex pyramid is introduced to adaptively capture multi-scale displacements. The amplitude degradation index and phase Laplacian value are proposed to determine reliable pyramid levels and measurement pixels. The discriminative correlation is used to estimate displacements of measurement points with multi-type geometric features. The present methods are applied to estimate the displacements of the Tsing Ma Bridge, Kingkey 100 high-rise building, and Humen Bridge.

The full-field displacement of large-scale structures is reconstructed using limited measurement data in Chapter 6. The phase-based displacement data are used to reconstruct quasi-static displacements via the pseudo-inverse operation. The acceleration data are used to reconstruct modal displacements using stochastic subspace theory. The displacement reconstruction method is validated using the Kingkey 100 high-rise building under typhoons. The results show that the quasi-static and modal displacements are accurately reconstructed with relative errors under 6%.

Chapter 7 concludes the thesis and discusses the future work.

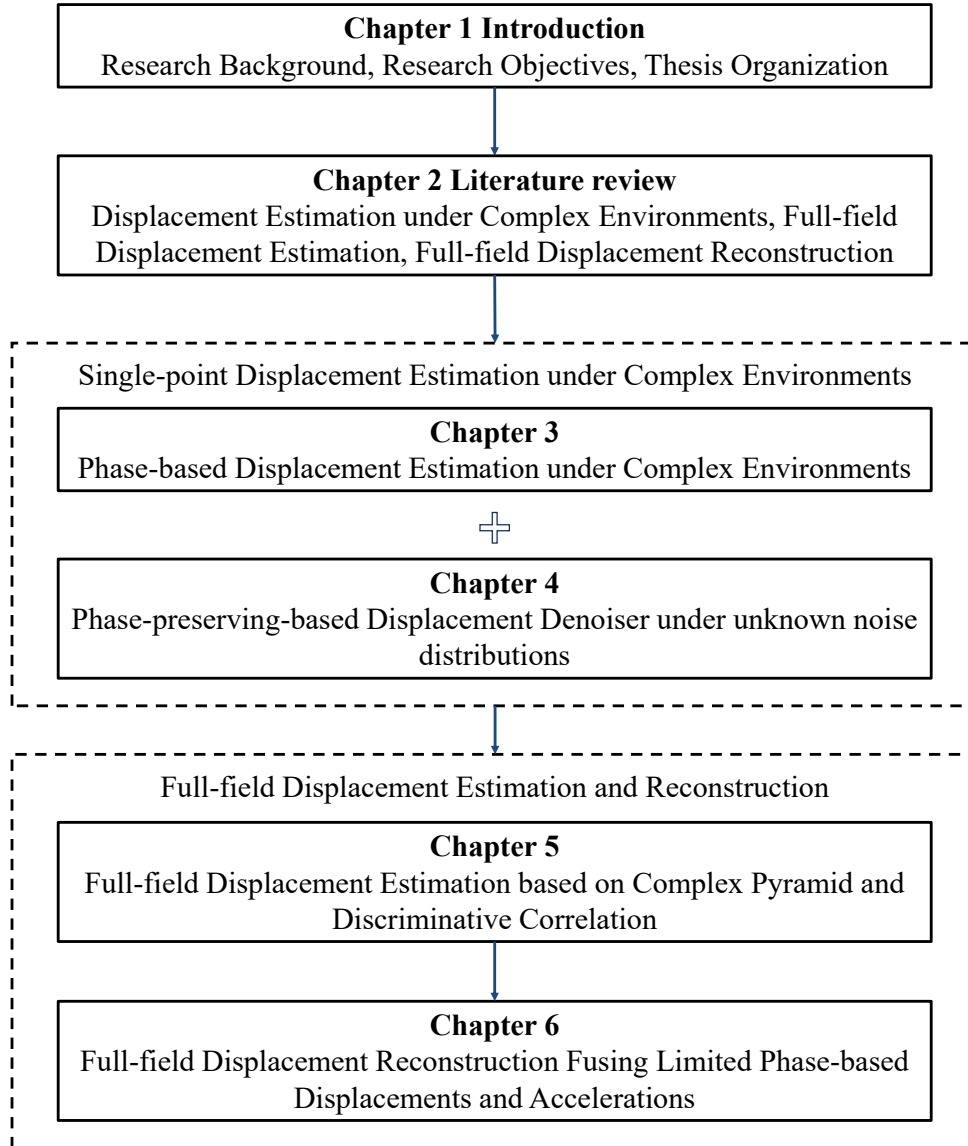


Figure 1.3 Organization of Thesis

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Research on structural displacements has been widely conducted (Ma et al., 2023; Shao et al., 2021; Yang et al., 2025). Nevertheless, measuring displacement in large-scale structures remains challenging.

The environments around large-scale structures adversely affect displacement measurements. First, complex environments, such as extra textures not belonging to the structure, illumination changes, and image noises, can degrade the performance of non-contact displacement estimation methods. Second, the installation of contact sensors and the insufficient measurement range and accuracy of non-contact sensors limit the full-field displacement estimation of large-scale structures. Third, the spatial partition in real-world scenarios impedes non-contact sensors from estimating full-field displacements.

Accordingly, previous studies are categorized into two main areas: displacement estimation under complex environments, full-field displacement estimation and reconstruction of large-scale structures.

2.2 Displacement estimation under complex environments

Displacement measurement methods can be classified as contact or non-contact. Contact methods such as the GNSS, FBG, and LVDT have been widely used in SHM systems.

LDV and vision-based equipment are increasingly being developed to measure structural displacement in a non-contact manner. Depending on the used devices and algorithms, current displacement estimation methods and their applications in complex environments are introduced as follows.

2.2.1 Contact displacement estimation methods

In the field of contact measurement, the most widely used techniques include LVDT displacement sensors, fiber optic displacement sensors, GPS, and displacement derived indirectly from accelerometer integration.

1) LVDT

LVDT displacement sensors offer high precision in measuring displacement, strain, and force (Norton, 1969). These sensors operate by detecting voltage variations between two copper coils to sense core movement. While LVDT technology is now relatively mature, its application is primarily constrained by environmental influences and measurement range limitations. Temperature significantly affects the material properties of the core and coils, thereby altering the sensor's sensitivity. In practical monitoring environments, temperature fluctuations are commonplace; for instance, bridge deck temperatures can vary by up to 20°C between day and night (Xia et al., 2013). A self-compensation method (Ara, 1972) was proposed to isolate the effects of temperature, current, and excitation frequency on the sensor output. However, the fundamental limitation of LVDT displacement sensors lies in their capability to measure only relative deformation between the two ends of the sensor, thus restricting their application to localized, single-point displacement and deformation measurements in structures.

2) FBG

FBG sensors operate similarly to LVDT displacement sensors but are immune to electromagnetic interference due to their use of optical signals. Common FBG sensors typically offer measurement ranges from a few millimeters to several hundred millimeters (Li et al., 2015, 2018; Tao et al., 2016; Xiong et al., 2019). When embedded in fiber-reinforced composites, FBG sensors are protected and can measure the deformation of the composite materials. This technique has been applied to large civil engineering structures such as bridges (Li et al., 2004). Interferometric fiber optic sensors are also widely used in scientific and engineering fields (Huang et al., 2010). However, the measurement range of such interferometers is generally limited to a few millimeters. Similar to LVDT sensors, these fiber optic displacement sensors can only measure localized deformations and are inadequate for capturing displacements relative to distant reference points on a structure.

3) GNSS

The GNSS is an umbrella term for navigation systems, encompassing the GPS, BeiDou Navigation Satellite System, GLONASS, Galileo, Quasi-Zenith Satellite System, and the Indian Regional Navigation Satellite System (Yu et al., 2020). Although satellites enable global signal acquisition by orbiting the Earth, ground-based receivers and base stations must be installed on or near the structure to be measured, making GNSS systems inherently contact-based sensor systems. Four kinematic processing methods can be applied for displacement monitoring: Real-Time Kinematic (RTK), Post-Processing Kinematic (PPK), Network RTK, and Precise Point Positioning (PPP) (Aponte et al., 2009). When monitoring structural displacements, receivers and base stations capture signals from satellites, and structural displacements are derived based on these satellite signals and correction data (Nikitopoulou et al., 2006).

The GPS with a sampling frequency of 10 Hz can achieve real-time monitoring of a 44-story high-rise building and identify its dynamic parameters (Çelebi, 2000). The responses of three high-rise buildings in Chicago were monitored by GNSS, with results consistent with wind tunnel tests and finite element analyses (Kijewski-Correa et al.,

2006). A GNSS system with a 20 Hz sampling frequency was employed to monitor both translational and rotational displacements of the Tianjin 117 super high-rise building (Liu et al., 2017). The Hong Kong Highways Department installed base stations on the Tsing Ma Bridge, Kap Shui Mun Bridge, and Ting Kau Bridge to monitor static influence coefficients at specific structural locations. Displacements of up to 200 mm were reported in the Tsing Ma Bridge over a single day, measured by GPS and caused by temperature variations (Xia et al., 2013).

GNSS performance is constrained by multipath errors, cycle slips, and low sampling frequencies (Yu et al., 2020). To enhance GNSS performance, data integration with total stations has been employed, a technique capable of capturing up to the 36th order natural frequency of a structure (Erdoğan and Güllal, 2013). The multipath error repeats over two sidereal days while the environment around the GPS receiver remains unchanged, which can be utilized to simulate and mitigate 67% of the multipath effect in dynamic displacement monitoring (Dai et al., 2014). However, GNSS measurement points and reference stations require relatively stringent installation conditions to suppress multipath effects, and their application is predominantly suited for measuring displacements at open locations, such as the tops of super high-rise structures.

4) Accelerometers

Accelerometers can also be integrated with other data to enhance displacement measurement performance. The issues of low accuracy and low sampling frequency of GPS sensors can be addressed by incorporating accelerometers (Kim et al., 2014), achieving an accuracy of 2 mm in field measurements. Ma et al. (2021) proposed an FIR filter-based method to fuse strain and acceleration data, improving the performance of displacement estimation through data fusion. Furthermore, accelerometer data fusion techniques can be combined with computer vision or millimeter-wave radar (Ma et al., 2023, 2022), leveraging the high-frequency components of acceleration data to enhance displacement accuracy.

Despite these advancements, the above contact-based measurement methods are still limited by the number of measurement points, making it difficult to obtain high spatial resolution displacement responses for large-scale structures.

2.2.2 Non-contact displacement estimation methods

In the field of contact measurement, the most widely used techniques include LVDT displacement sensors, fiber optic displacement sensors, GPS, and displacement derived indirectly from accelerometer integration.

1) Active non-contact displacement estimation

Active non-contact measurements refer to the method of measuring structures by actively emitting signals such as infrared or microwaves, which possess inherent advantages in complex environments. Typical techniques like total stations and LDV can achieve high precision. Psimoulis and Stiros (2013) used a smart total station to measure a maximum deflection of 6 mm in a railway bridge. Luo et al. (2016) found that when the distance between the total station and the measured surface was approximately 50 m, the accuracy of the total station could reach 1 mm. However, the sampling frequency of total station equipment is insufficient to capture the dynamic displacement time history of structures. Nassif et al. (2005) conducted deflection and vibration measurements on the Doremus Avenue Bridge using LDV and other contact-based sensors, and the results from all sensors were highly consistent. Garg et al. (2019) employed a mobile LDV on an unmanned system to monitor outdoor structures. After angle and linear motion compensation, the root mean square error of the results was 10%.

Radar is another advanced non-contact measurement technique capable of achieving millimeter-level displacement accuracy with a carrier frequency of 2.3 GHz. There are

various types of radar systems, among which those used for structural displacement monitoring are largely developed from CW radar. Doppler interferometric radar and FMCW radar are two primary forms of CW radar (Li et al., 2017). Zhang et al. (2018) successfully obtained the dynamic responses of structures under different loads using a portable radar interferometer on the Ting Kau Bridge and Tsing Ma Bridge in Hong Kong. Guan et al. (2017) installed radar measurement points and signal transponders as reference points on a bridge structure to actively monitor structural displacement.

In complex environments, the aforementioned active non-contact sensors can mitigate the adverse effects of illumination variations caused by time and weather conditions through their active emission-based measurement approach. However, in large-scale structural displacement monitoring scenarios, the extended range of active signal transmission tends to compromise measurement accuracy. Moreover, the high demands for equipment and operational expertise have limited their application in structural health monitoring systems. Consequently, methods utilizing readily available digital image data for structural displacement identification have undergone rapid development in recent years.

2) DIC method

The fundamental principle of DIC is to identify similar pixel regions across successive image frames and compute the coordinate variations of these pixels. While DIC can be employed in laboratory settings to measure local strain in structural components, its application to large-scale structures presents challenges. In such cases, the pixel displacements corresponding to modal vibrations of structural members are often too subtle in digital images. Therefore, achieving adequate accuracy generally requires the use of telephoto lenses and highly distinguishable artificial targets.

The DIC method begins by defining a region of interest in a reference image frame and extracting its feature representation. Subsequently, in the displaced image frame, similarity values are computed between all candidate regions and the originally defined

reference region. The candidate region with the highest similarity is identified as the displaced location of the original region. Finally, the coordinates of the centroid for both regions are extracted, and the displacement value is calculated as follows:

$$\begin{cases} u = x_1 - x_0 \\ v = y_1 - y_0 \end{cases} \quad (2.1)$$

where u — Horizontal displacements;

v — Vertical displacements;

x_0 — Horizontal coordinate of target center in the reference frame;

y_0 — Vertical coordinate of target center in the reference frame;

x_1 — Horizontal coordinate of displaced target center;

y_1 — Vertical coordinate of displaced target center.

Numerous matching criteria, such as cross-correlation and sum of squared differences, have been proposed (Giachetti, 2000; Tong, 2005) to evaluate similarity between regions. To account for variations in illumination across image frames, the normalized cross-correlation criterion can achieve more robust performance (Pan et al., 2007). After assessing inter-region similarity, optimization algorithms such as Newton–Raphson or Levenberg–Marquardt (Schreier and Sutton, 2002; Vendroux and Knauss, 1998) are employed to identify the most similar region and compute coordinate differences.

Based on the above principles, a key limitation of the DIC method is that it identifies displacements with at most pixel-level resolution. This stems primarily from the fact that similarity between regions is computed on a pixel-by-pixel basis. Although various interpolation methods can be used to upsample the original images for sub-pixel displacement identification (Schreier et al., 2000), such approaches rely on mathematical assumptions without a physical foundation, making it difficult to guarantee a linear relationship between pixel value changes and actual displacement.

To facilitate the application of the DIC method, Blaber et al. (2015) released a high-quality open-source DIC toolkit, which employs first-order shape functions to describe

displacement, utilizes B-spline interpolation to enhance displacement resolution, and improves computational efficiency through an inverse compositional Gauss–Newton optimization algorithm. To compare the performance of DIC and strain gauges in laboratory settings, Hoult et al. (2013) conducted a series of tensile tests. They identified and addressed the issue where out-of-plane displacement of the measured object can significantly amplify errors in the identified in-plane displacement. Their results demonstrated that the proposed method could serve as a viable alternative to strain gauges in laboratory experiments.

Ellenberg et al. (2015) mounted a camera on a UAV to measure the deformation of indoor structures using DIC. For real-world structures, McCormick and Lord (2012) applied the DIC method to monitor a highway bridge. They found that DIC performed effectively under both natural ambient lighting and artificial illumination. Xia et al. (2021) deployed a comprehensive structural health monitoring system on a footbridge at the Hong Kong Polytechnic University for research and educational purposes. The system included two cameras using the DIC method to monitor displacements on both sides of the footbridge, clearly capturing the structural displacement responses induced by pedestrian traffic. Ni et al. (2019) explored the application of DIC on super-tall structures by setting up a telephoto camera at the base of the Canton Tower to capture the illuminated logo at its top. The method successfully identified tower-top vibrations with amplitudes of about 1 cm during both clear days and nights, showing good agreement with GNSS measurements.

Although the theoretical framework and commercial equipment for the DIC method are relatively well-established, its underlying principles dictate that subpixel-level displacement relies on mathematically assumed interpolation. Moreover, the method often requires the measured structure to exhibit geometric features with high contrast against the background; otherwise, artificial targets must be manually installed at the measurement points (Hu et al., 2021).

3) IOF method

The optical flow method represents a category of displacement identification techniques capable of achieving subpixel-level resolution. IOF tracks pixels of identical pixel intensity across successive image frames. This approach eliminates the need for artificial markers on the target structure, instead localizing objects through variations in pixel intensity. A fundamental assumption underlying the IOF method is the brightness constancy of the measured structure, which posits that:

$$\frac{d}{dt} I[\mathbf{P}(t), t] = 0 \quad (2.2)$$

where $\mathbf{P}(t)$ ——Coordinate of the target at time t ;

$I(\cdot)$ ——Pixel intensity of the target.

$$\frac{\partial I}{\partial t} + \nabla I \frac{d\mathbf{P}(t)}{dt} = \frac{\partial I}{\partial t} + \nabla I \times [\mathbf{u}(t), \mathbf{v}(t)] = 0 \quad (2.3)$$

where ∇I ——Spatial gradient of the pixel intensity;

$\mathbf{u}(t)$ ——Time history of the horizontal displacement;

$\mathbf{v}(t)$ ——Time history of the vertical displacement.

By solving the above equation, a relationship between pixel intensity and displacement vectors can be established. However, since the displacement vector contains two unknown components in the horizontal and vertical directions, the number of unknowns becomes twice the number of equations, resulting in an underdetermined problem. To resolve the displacement vector, Horn and Schunck (1981) introduced a smoothness constraint term, which assumes that the displacement field varies smoothly across the image plane:

$$E(\mathbf{u}, \mathbf{v}) = \int_{(x,y) \in I} \left\{ \left[\frac{\partial I}{\partial t} + \nabla I * (\mathbf{u}, \mathbf{v}) \right]^2 + (\nabla \mathbf{u}^2 + \nabla \mathbf{v}^2) \right\} dx dy \quad (2.4)$$

where $\nabla \mathbf{u}$ ——Spatial gradient of the horizontal displacement among pixels;

$\nabla \mathbf{v}$ ——Spatial gradient of the vertical displacement among pixels.

The displacement vector solution can be obtained by minimizing the aforementioned

objective function through optimization algorithms. Another widely adopted constraint for solving displacement vectors is the Lucas-Kanade method (Bruhn et al., 2005), which operates under the assumption that pixels within a local neighborhood exhibit identical displacement vectors.

IOF has been widely integrated into open-source libraries such as OpenCV (Sánchez Pérez et al., 2013). The introduction of a total variation-based regularization term has addressed the over-smoothing issue inherent in the Horn-Schunck method's smoothness constraint, resulting in enhanced efficiency and robustness. Yoon et al. (2016) employed a dual-camera setup to capture images and applied the Kanade-Lucas-Tomasi method to measure displacements of a six-story frame model. To improve reliability, distinct geometric features were first extracted from the images using the Harris corner detection method. Compared with results from accelerometer sensors, the proposed method demonstrated strong agreement in transfer functions, mode shapes, and natural frequencies.

In the context of tall building structures, a major challenge lies in the deployment of camera sensors. Weng et al. (2021) utilized an UAV to monitor structural displacements. To compensate for the UAV's ego-motion, a deep learning model was developed to distinguish static and dynamic objects within the image sequences. Experimental results from both laboratory tests and field excitation tests on an actual television tower confirmed that the method can accurately identify structural displacements without being affected by equipment vibration.

To overcome the limitation of conventional cameras in capturing only two-dimensional planar motion, Shao et al. (2021) proposed a binocular vision system for measuring three-dimensional movements. They employed a deep learning model to match identical feature points in the images from both lenses, enabling the identification of structural displacements in different directions from each lens. The method demonstrated excellent performance in laboratory environments.

However, the greatest challenge of IOF lies in the utilization of the raw intensity value, which can be adversely affected by complex environments like shadows and illumination changes, leading to inaccurate displacement results.

4) Deep learning-based displacement estimation

The object tracking in images, known as the optical flow field, also constitutes a significant research topic in the domain of deep learning for computer vision. Dosovitskiy et al. (2015) pioneered the estimation of motion fields for each pixel in images using deep learning models. In their work, they not only developed an effective model comprising two architectures but also released an open-source image dataset for optical flow estimation. They further improved the model's performance by employing a multi-network architecture (Ilg et al., 2017), which for the first time surpassed traditional displacement identification methods in accuracy on benchmark datasets, albeit at the cost of significantly increased computational complexity due to the large number of parameters.

Hui et al. (2018, 2020) utilized multi-resolution sampling techniques (image pyramids) to extract image features, enhancing model performance while reducing computational costs. Shao et al. (2024) predicted three-dimensional mesh deformations of structures from two-dimensional images via deep neural networks, enabling 3D displacement identification using a monocular camera.

Drawing on phase-based motion magnification theory, Oh et al. (2018) proposed a deep learning-based motion amplification model that achieved superior performance and faster processing speeds compared to phase-based methods. Luan et al. (2021) trained a deep learning model using laboratory mechanical experiments to efficiently extract structural displacement fields.

Although deep learning-based methods have demonstrated excellent performance on limited datasets (Baker et al., 2011; Scharstein et al., 2014), their application to real-world structural scenarios still faces two main limitations. First, deep learning models rely on training data, which often differs from actual application environments, thus compromising the reliability of displacement identification. Second, there remains a lack of datasets that adequately account for subpixel-level displacements, particularly for large-scale structures where displacements may manifest at small pixel scales in the image plane.

5) POF method

The brightness constancy assumption in IOF often proves overly restrictive for real-world structural monitoring applications. In response, phase information derived from complex filter transformations has been proposed as a more robust and sensitive representation for displacement compared to pixel intensity (Fleet and Jepson, 1990). Consequently, POF (Ai et al., 2025; Zhang et al., 2024) has gained increasing attention in SHM over the past decade. Unlike IOF, POF operates on the principle that the phase values corresponding to identical structural features remain consistent across successive image frames:

$$\phi[\mathbf{P}(t), t] = c \quad (2.5)$$

where ϕ ——Phase of the target;
 c ——Constant phase value.

Differentiating the phase in the time domain, the displacement can be projected by the phase:

$$\mathbf{u}(t) = \frac{d\mathbf{P}(t)}{dt} = \frac{-\phi}{\|\phi_p\|^2} \phi_p \quad (2.6)$$

where ϕ ——Phase deviation in the time domain;
 ϕ_p ——Spatial gradient of the phase.

As shown in the above equation, the estimation of structural displacement in the POF method directly depends on phase information. Consequently, recent research has focused on improving phase quality under complex environmental conditions and promoting the application of POF. Barron et al. (1992) found that the POF method outperformed other displacement estimation techniques in experiments, further demonstrating the strong correlation between phase and displacement. Subsequently, Gautama and Van Hulle (2002) enhanced the efficiency of POF using spatial filters, achieving satisfactory performance on both synthetic and real-world images. To capture structural displacements of varying amplitudes, they constructed a complex steerable pyramid to identify motions in different orientations and ultimately superimposed the results.

Building upon the concept of POF, Wadhwa et al. (2013) released an open-source, phase-based motion magnification tool, which was applied to identify and visualize the vibration modes of laboratory structures (Guo et al., 2025). Yang et al. (2017) integrated blind source separation algorithms with this phase-based motion magnification tool to accurately extract modal parameters of structures in laboratory settings. They further extended their work by accounting for the removal of rigid body motion effects prior to modal identification (Yang et al., 2020).

To estimate out-of-the-plane motion, Chou et al. (2022) employed a depth camera combined with a phase pyramid to identify out-of-the-plane displacements. Valente et al. (2022) utilized the IOF method to extract displacements from phase-magnified images and demonstrated that the accuracy of the identified displacements was indeed improved through phase magnification.

Utilizing the local phase of large-scale structures is an effective approach for enhancing displacement accuracy, since only the local phase information of the target structure is relevant to its displacement. For instance, the Hilbert transform, which operates on the entire signal, can extract precise phase information only when the displacement is uniform across the whole image (Liu et al., 2022). Therefore, most existing methods employ complex filters such as the Gabor filter as extractors of local phase (Chen et al., 2015).

These filters are confined by finite windows in both the spatial and frequency domains.

Diamond et al. (2017) demonstrated through laboratory experiments that the Gabor filter-based POF method can achieve an accuracy of 1/450 pixel. Miao et al. (2022) introduced a phase nonlinearity index to optimize the parameters of the Gabor filter, improving its adaptability to the geometric features of the target structure. Collier and Dare (2022) validated the performance of the Gabor filter across images with different characteristics. Similarly, Boukerroui et al. (2004) compared the performance of various complex filters and found that the traditional Gabor filter is unsuitable for extracting geometric features of structures in certain scenarios due to its introduction of DC and negative frequency components in the frequency domain.

Besides, the monitoring environments for large-scale real-world structures are highly complex, where interfering objects adjacent to or occluding the target structure, time-varying illumination conditions (Luo et al., 2025), and equipment-induced noise collectively degrade the accuracy of phase extraction and displacement estimation. However, the underlying mechanisms of how such complex environments affect phase remain poorly understood, and the performance of traditional POF under these challenging conditions has not been systematically investigated or validated.

6) Displacement estimation enhancement based on image denoising

For digital image-based displacement estimation methods, the quality of image signals is of critical importance. However, photographs captured in real-world conditions inevitably contain noise, prompting the development of numerous denoising techniques over recent decades. These methods can be categorized into spatial domain approaches, variational-based methods, transform domain-based techniques, and deep learning-based methods (Fan et al., 2019).

Spatial domain methods represent the most common form of denoising, where typical techniques such as windowed mean filtering and Gaussian smoothing can be employed to suppress fluctuations around signal sampling points. Median filtering (Pitas and Venetsanopoulos, 2013) is a nonlinear approach that replaces signal values with the local median of the surrounding area. To better preserve geometric information such as linear edges, the bilateral filter (Tomasi and Manduchi, 1998) incorporates both the distance between the target pixel and its neighbors and gives greater weight to pixels closer to the target during denoising, though this method suffers from relatively low computational efficiency. Overall, spatial filters function as a class of low-pass filters that inevitably cause the loss of high-frequency signal components. However, phase-based displacement estimation for large-scale structures requires the extraction of high-contrast pixel information representing spatial high-frequency variations, such as structural edges.

Variational denoising methods, meanwhile, formulate the denoising process as an optimization problem:

$$\hat{\mathbf{I}} = \arg \min_{\hat{\mathbf{I}}} \frac{1}{2} \|\mathbf{I} - \hat{\mathbf{I}}\|_2^2 + \lambda R(\hat{\mathbf{I}}) \quad (2.7)$$

where $\hat{\mathbf{I}}$ — Denoised image frame;
 R — Regularization term;
 λ — Parameters of the regularization term.

This formulation is grounded in Bayesian theory, where the first term on the right-hand side represents the likelihood estimate of the denoised image, while the second term corresponds to the image prior. The key to this variational approach lies in the selection of an appropriate prior regularizer for the image. The TV regularizer stands as one of the most influential image denoising techniques, operating under the assumption that variations in pixel intensity across the image are gradual (Rudin et al., 1992). While this method can preserve image details to some extent, it still tends to blur these fine features.

In contrast to TV-based methods, non-local regularization incorporates non-local information by considering pixels that are spatially distant yet structurally similar to the

target pixel (Gilboa and Osher, 2009). Non-local variational strategies have been widely adopted in denoising applications. They can be combined with TV regularizers to enhance performance (Sutour et al., 2014) or employ deep learning models to evaluate inter-pixel similarity (Zoran and Weiss, 2011). Furthermore, additional advanced regularization techniques, such as sparse representation and overcomplete dictionaries (Zhang et al., 2012), have been proposed for image reconstruction, though their computational efficiency remains suboptimal.

Transform domain-based denoising methods, primarily represented by wavelet denoising, have evolved from spatial domain techniques, as most transform domain methods employ spatial filters to convert the original signal into a domain where noise can be effectively separated. The core principle of these methods lies in the clear distinction and separation between noise and meaningful image textures in the transform domain. Wavelet denoising (Mallat, 1989) is widely applied to images and other signals. Its strategy involves filtering the original signal with a series of wavelet filters at different spatial resolutions to generate multi-resolution filtered signals. Noise can be distinguished in these signals, and the denoised result is obtained by reconstructing the multi-resolution signals after noise removal. The BM3D Filtering method (Dabov et al., 2009) builds upon wavelet techniques by first searching for similar patches in the signal using a non-local approach and encoding them into 3D arrays. Maggioni et al. (2013) further extended this approach by encoding image signals into 4D arrays, accounting for correlations within and between 3D arrays. After inverse transformation, the denoised 3D arrays were reassembled into the denoised image.

While the aforementioned methods transform image signals into another real-valued domain, phase-based denoising (Kovesi, 1999) converts image signals into the complex domain. This method relies on oriented complex filters (Field, 1987), which are similar in form to wavelet filters but consist of filter pairs with a phase difference to extract the real and imaginary parts of the signal, ultimately computing phase and magnitude. In this approach, noise exhibits significantly smaller magnitude values compared to meaningful information in the image. By eliminating pixels with smaller magnitudes at each

resolution level and reconstructing the denoised multi-resolution images, the denoised result is achieved. Phase-based denoising is particularly suitable for POF displacement estimation, as it ensures that the phase information of the target structure remains undisturbed by the denoising process. However, existing phase-based denoising methods require pre-defined assumptions about the noise distribution, making it difficult to quantify noise thresholds under conditions of unknown noise distribution.

Deep learning-based denoising methods have garnered widespread attention due to their remarkable advantages in enhancing visual quality. In recent years, numerous deep learning models have been continuously proposed to improve performance (Tian et al., 2020). Zhang et al. (2017) introduced the DnCNN model paradigm, which incorporated residual learning and batch normalization for the first time to enhance model performance and efficiency. This method can rapidly remove Gaussian noise from images through parallel computing. However, despite their excellent performance, the application of deep learning methods remains constrained by their limited generalization capability. The primary reason for this limitation lies in the fact that deep learning models learn noise distributions from training datasets, whereas the noise distribution in practical applications remains unknown. For instance, a model trained on data with Gaussian noise at an SNR of 20 would not perform optimally on data with an SNR of 25. To address this issue, Zhang et al. (2018) proposed incorporating noise level as an input during model training, yet this approach still struggles to handle complex non-Gaussian noise distributions. Consequently, the performance of deep learning-based denoising methods cannot be guaranteed when dealing with unknown noise distributions in complex environments.

In summary, existing image denoising techniques primarily focus on enhancing the overall visual quality of images. However, vision-based displacement estimation methods predominantly rely on spatial high-frequency information such as structural edges and fine details within the image. This fundamental discrepancy leads to a situation where current denoising approaches cannot guarantee and may even degrade the performance of vision-based displacement estimation.

2.3 Full-field displacement estimation and reconstruction of large-scale structures

SHM systems acquire structural responses through deployed distributed sensors, yet these measurements remain sparse in nature. For large-scale structures, high spatial-resolution displacements are crucial for performance analysis, necessitating the development of theories and methodologies capable of estimating or reconstructing high spatial-resolution full-field displacements across extensive areas. Two primary approaches, namely, advanced full-field displacement estimation techniques and full-field displacement reconstruction methods, can obtain full-field displacements for large-scale structures. The following section will review relevant research from these two perspectives.

2.3.1 Full-field displacement estimation of large-scale structures

To directly monitor high spatial-resolution full-field responses across large-scale structures over extensive areas, sensors capable of measuring structural responses on a broad scale are essential. For instance, Luo et al. (2021) conducted real-time monitoring during the construction phase of the National Speed Skating Oval using laser displacement transducers connected via wireless networks, achieving high precision. To address the challenge of capturing full-field structural responses, an adaptive partitioned phase demodulation method was proposed to obtain accurate displacements, demonstrating micrometer-level accuracy in laboratory settings (S. Li et al., 2022). However, these non-contact displacement identification methods still require the deployment of contact-based receivers on the structure, thereby yielding sparse measurement points.

Based on the preceding introduction, optical imaging devices represented by cameras can capture spatial positions and geometric information of structures over large areas,

presenting a potentially ideal solution for acquiring high spatial-resolution displacement data. Consequently, numerous recent studies have focused on monitoring full-field displacements of large-scale structures using visual information (Peng et al., 2024).

To measure the full-field structural modes of actual bridges, Wang et al. (2022) proposed a fully non-contact method that uses cameras to progressively measure responses at reference points and moving measurement points, thereby obtaining structural mode shapes. By deploying three cameras, they sequentially identified measurement points at varying distances on both sides of the bridge and correlated the data with images from a camera focused on static reference points near the bridge, enabling the extraction of global bridge modal properties. This method demonstrated satisfactory performance in both scaled models and real bridge applications.

Xu et al. (2018) employed cameras to conduct multi-point measurements on different structural components of a 109-meter-long cable-stayed bridge. For the stay cables, which had insufficient pixel coverage in the image (only 4 pixels in width), edge detection algorithms were applied to extract cable information, followed by computing positional changes of the cables across image frames. For the bridge girder, the geometric features of cable anchorage points were utilized with the zero-mean normalized cross-correlation criterion of the DIC method to perform inter-frame matching. At a distance of 50 meters from the bridge, their proposed multi-point measurement method successfully estimated millimeter-level displacements and captured the first two modal frequencies of the bridge. To further leverage multi-point structural responses for finite element model updating, Martini et al. (2022) used digital images to simultaneously observe vertical deflections and vehicle loads on a laboratory bridge model, and updated the finite element model based on influence lines.

Owing to its high accuracy and robustness, the phase-based method has garnered considerable research interest and practical application in the full-field structural displacement estimation. To enhance the spatial resolution of structural displacement estimation in real-world environments, Zhu et al. (2022) partitioned images of the target

structure into multiple regions based on field-of-view parameters to preprocess the images and obtain high-quality signals. They then employed the POF method to measure displacement signals in each region, denoised the displacements via modal decomposition, and finally amplified the modal components of interest to observe multi-point vibrations of the structure. This approach successfully measured sub-millimeter-level displacements of multiple stay cables on a footbridge from a distance of 30 meters.

By integrating a binocular vision system with a rotating platform, Cui et al. (2025) developed a platform suitable for large-area displacement identification based on the POF method. Combining broadband phase-based motion magnification technology, Luo et al. (2023) used edge tracking techniques to measure cable vibrations, identifying the natural frequencies of stay cables of varying lengths from a distance of 120 meters away from an actual cable-stayed bridge.

Kong et al. (2025) adaptively and accurately extracted pixel displacements using reliability indices and estimated the displacement field in digital images of an air compressor in a laboratory setting. Li et al. (2025) employed Gaussian process regression to match structural measurement points from different perspectives and used a multi-resolution image pyramid-based POF to identify multi-point displacements of a laboratory structure. By integrating the multi-resolution image pyramid POF with compressed sensing, Li et al. (2025) reduced the computational cost associated with identifying high-frequency, full-field displacement time histories using high-speed cameras.

Liu et al. (2024) simplified the computational workflow of POF by treating image grayscale values as the real part and the first-order derivative of the image as the imaginary part, enabling the identification of full-field displacements in a laboratory structure. Zhong et al. (2025) were the first to incorporate multi-task optimization into the optimization of complex filters for POF, achieving displacement field identification for a laboratory-scale spacecraft model. Shan et al. (2024) proposed a POF method that does not require phase unwrapping, successfully estimating full-field displacements in a

laboratory structure. Xie and Cheng (2025) applied a 2D Hilbert transform to compute image phase and performed full-field displacement and mode shape identification for a laboratory structure.

However, the aforementioned vision-based techniques are primarily applicable for estimating relatively large-amplitude displacements at multiple points within a limited scope. For large-scale structures, the considerable filming distance leads to minuscule pixel displacements, while complex external excitations and various structural stiffness result in significant amplitude fluctuations across measurement points. Furthermore, challenging filming environments often compromise the reliability of geometric features at target measurement points. Therefore, further development is required to establish adaptive full-field displacement identification methods that can satisfy the accuracy, measurement range, and adaptability to diverse target geometric characteristics essential for large-scale structural displacement monitoring.

2.3.2 Full-field displacement reconstruction of large-scale structures

In practical monitoring environments, constraints such as structural layout, daily operations, and surrounding conditions may make it difficult to directly estimate spatiotemporally continuous full-field displacements of structures. Given that structural displacement responses are mechanical phenomena induced by external actions, the patterns of structural displacement are inherently related to the structural mechanical model. Leveraging this principle, the displacement at different locations of a structure can be reconstructed based on limited measurements, or the conversion between displacement and other physical response quantities of the structure can be achieved.

For high-rise structures, wind-induced displacements under daily operational conditions are typically on the millimeter scale. However, under extreme weather conditions such as typhoons, displacements at the top increase sharply, posing a risk of the inter-story drift

ratio exceeding the elastic limit of $1/500$. Therefore, obtaining the inter-story drift values at each floor under both serviceability and ultimate limit states is essential for evaluating the deformation safety of high-rise structures. Nevertheless, high-rise structures are characterized by internal spatial partitions between floors and external urban environmental obstructions, while extreme weather conditions may disturb external measurement equipment.

To efficiently obtain displacements of high-rise structures, Hu et al. (2020) proposed a conversion coefficient that transforms the output of inclinometers into structural displacements. This coefficient is defined as the ratio between the first-order mode shape of the structure and its first derivative, and it remains unaffected by the external loads applied to the structure. The performance of this inclinometer-based displacement monitoring method was validated on both a dormitory high-rise structure at Harbin Institute of Technology (Shenzhen) and the super-tall Kingkey 100 building, demonstrating the ability to identify sub-millimeter-level displacements. However, as the method relies on data from inclinometers, it can only provide structural displacements at instrumented floor locations where inclinometers are installed.

Therefore, it is necessary to integrate the mechanical model of the structure and develop the concept of hybrid monitoring-based structural response reconstruction. He et al. (2023) established a simplified spring model for the seismic displacement response of high-rise structures, incorporating acceleration measurements from limited locations and residual displacement monitoring data at the top. The model was then updated using a particle swarm optimization algorithm, allowing extraction of structural responses at specific locations. This method can estimate the response of the entire building using measurement data from only $1/4$ to $1/3$ of the total floors, achieving an error of approximately 5% in laboratory model structures.

From a static perspective, Zhang and Sun (2017) assumed that the bridge response is primarily induced by vehicle loads, which are relatively slow-varying regarding the dimensions of long-span bridges. They established equivalent structural loads by

neglecting damping and inertial forces, and subsequently reconstructed hybrid monitoring-based structural responses using the flexibility matrix derived from various responses in FEM. However, long-span bridges and high-rise structures are also subjected to dynamic loads such as wind and seismic actions, necessitating consideration of dynamic structural responses.

To extend the hybrid monitoring theory to general loading conditions, Li and Sun (2021) formulated a time-discrete state-space representation of the dynamic form of the structural equilibrium equation. They obtained an over-determined control equation and solved for the equivalent load under general loading using generalized inverse operators and truncated singular value decomposition, ultimately reconstructing high-resolution structural responses.

These high-resolution structural response reconstruction methods essentially require estimating full-field responses from limited measurement data. Therefore, the reconstruction performance depends on the type and quantity of known measurements. For displacement response reconstruction of large-scale structures, directly measured displacement time histories and an understanding of displacement response characteristics remain fundamental. However, for large structures with spatial partitions, such as super-tall buildings, while substantial research exists on acceleration response characteristics based on accelerometer data, insufficient means for displacement estimation under extreme weather conditions have hindered a comprehensive study of displacement response characteristics. This limitation consequently restricts the estimation and reconstruction of spatiotemporally continuous full-story displacements and inter-story drift ratios for super-tall structures.

2.4 Summary and challenges

The full-field displacement estimation and reconstruction of large-scale structures under complex environments have been comprehensively reviewed. Contact displacement estimation methods are reliable, but fail to measure multiple points due to their installation requirements. Non-contact methods can measure structures at a long distance, in which the accurate and robust phase-based method, together with the response reconstruction technique, shows good potential in obtaining full-field displacements.

Despite the significant progress achieved by these approaches, challenges in obtaining full-field displacements for large-scale structures are shown in Figure 2.1:

- (1) First, complex environments induce variations in pixel values. Under ideal conditions, changes in pixel intensity are solely influenced by the structural displacement. However, in practical complex monitoring environments, three factors, namely, time-varying illumination conditions, displacements of interfering objects adjacent to the measurement points, and image noise generated by the camera's electro-optical sensors, can all perturb the pixel intensity at the measurement points. Therefore, it is necessary to investigate the mechanism of the adverse effects of complex environments on displacement estimation, developing digital image-based displacement estimation and denoising methods suitable for such conditions. These three research factors correspond to the content of Chapters 3 and 4 of this thesis.

Steps of displacement estimation

Challenges for large-scale structures

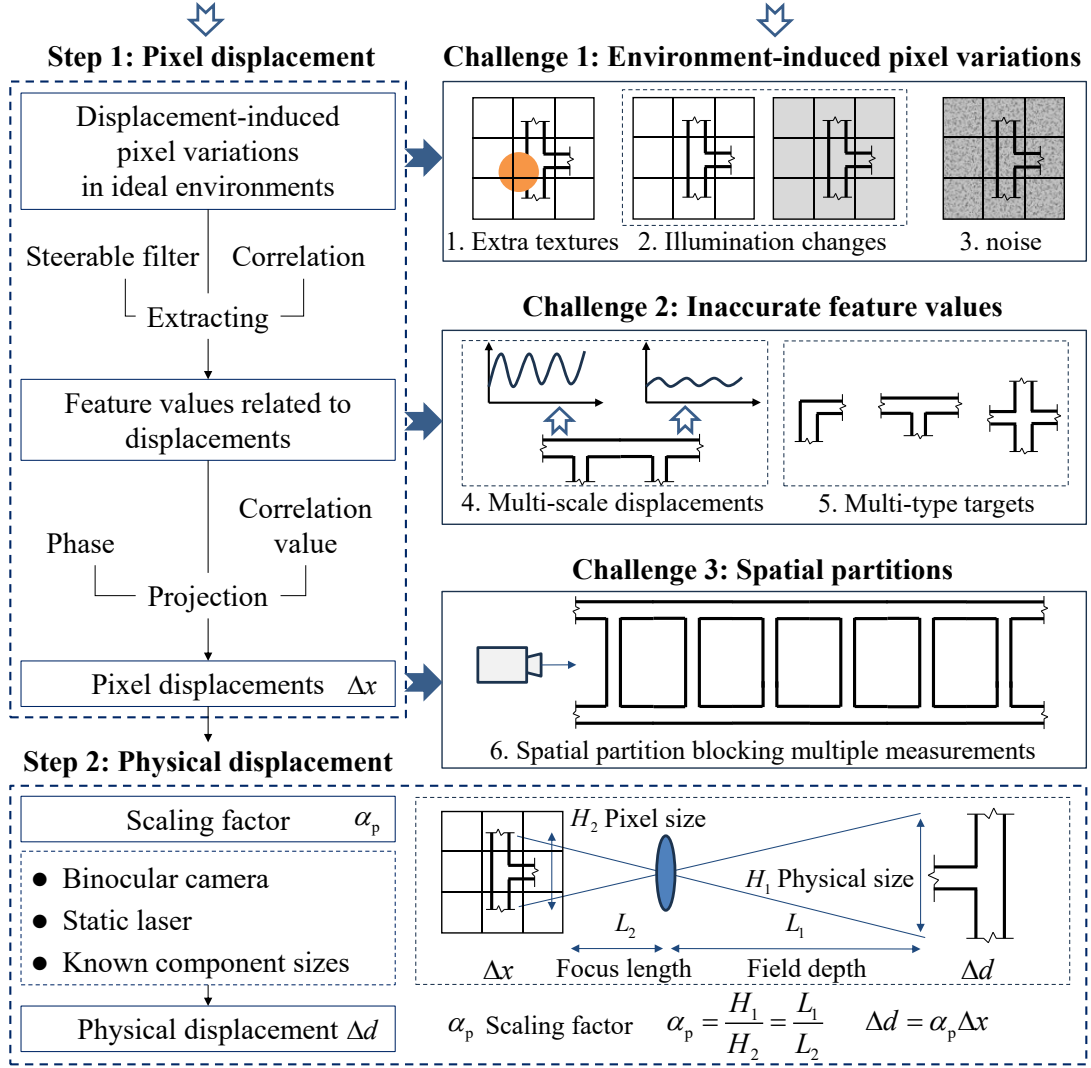


Figure 2.1 Challenges in obtaining full-field displacements

- (2) Second, errors exist in the projection between feature values and displacements in full-field displacement estimation. On one hand, when estimating displacements at multiple measurement points of a large structure within the image plane, the spatiotemporal displacement amplitudes at each point vary according to the stiffness of the specific structural component and the external excitation. This necessitates the development of digital image displacement estimation methods that satisfy both measurement range and accuracy. On the other hand, the geometric shapes of multiple target points differ, requiring the development of feature extraction and tracking methods capable of handling diverse target types to obtain accurate feature values for

each point. These two research aspects correspond to the content of Chapter 5 of this thesis.

- (3) Third, spatial partitions among different measurement points of large-scale structures make it difficult to identify spatiotemporally continuous full-field displacements. In practical large-structure measurement scenarios, physical spatial obstructions may exist between multiple measurement points, such as floor slabs separating different stories in super high-rise structures. Under such spatial partitions, it is challenging to directly estimate the displacements of all measurement points within the digital image plane. Therefore, it is of great significance to develop methods for reconstructing full-field displacements by integrating limited digital image displacement data with measurements from other sensors. This research factor corresponds to the content of Chapter 6 of this thesis.

This study will tackle the above difficulties and explore solutions in the subsequent chapters.

CHAPTER 3

PHASE-BASED DISPLACEMENT ESTIMATION UNDER COMPLEX ENVIRONMENTS

3.1 Introduction

Complex environments impose significant constraints on the accuracy of digital image-based displacement estimation methods. Extracting the phase of geometric features of the target structure from digital images to characterize the spatial position of measurement points has been recognized as an effective approach to enhance the accuracy and robustness of subpixel-level displacement estimation (Chen et al., 2015; Miao et al., 2022). However, the extra textures interfering with targets and varying illumination conditions in complex environments pose three major challenges to the reliable extraction of the target structure's phase:

First, the presence of occluding or adjacent extra textures along the optical axis or within the imaging plane often violates the constant-phase assumption for the target structure. Second, time-varying illumination conditions in the shooting environment cause pixel value fluctuations unrelated to structural mechanical behavior, leading to temporal drift in the extracted phase of the target structure. Third, the coupled effects of interfering objects and illumination variations on phase extraction quality remain insufficiently studied, making it difficult for conventional complex filters to decouple the influence of different environmental factors, thereby limiting the performance of phase-based displacement estimation in complex environments.

To address these challenges, this chapter focuses on phase-based displacement estimation in complex environments using digital images. Specifically, it employs complex filters to extract the spatial phase of geometric features from the target structure and establishes a projection relationship between the temporal phase variations and structural displacements, enabling high-accuracy subpixel-level displacement estimation. Furthermore, the mechanisms of the effects of the extra texture on phase extraction and displacement estimation performance are analyzed. Based on this analysis, a localized phase extraction method is proposed to controls the depth of field and optimizes parameters of the complex filters for accurate phase retrieval. Finally, the chapter investigates the coupled influence of the extra texture and the time-varying illumination on phase extraction and displacement estimation, and establishes a parameter optimization method for Log-Gabor filters based on the localization index, thereby achieving robust phase-based displacement estimation in complex environments.

3.2 Phase and subpixel displacement projection in complex environments

3.2.1 Phase-based subpixel displacement estimation

The phase of the target structure should be first extracted by convolving the digital image with complex filters:

$$\mathbf{R}_I(x, y) = \sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} \mathbf{I}(x + \Delta x, y + \Delta y) \mathbf{G}_{\text{even}}(\Delta x, \Delta y) \quad (3.1)$$

$$\mathbf{Z}_I(x, y) = \sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} \mathbf{I}(x + \Delta x, y + \Delta y) \mathbf{G}_{\text{odd}}(\Delta x, \Delta y) \quad (3.2)$$

where $\mathbf{R}_I(x, y)$ ——Real part of the filtered image in the spatial domain;
 $\mathbf{Z}_I(x, y)$ ——Imaginary part of the filtered image in the spatial domain;
 $\mathbf{I}(x, y)$ ——Pixel intensity values;
 $\mathbf{G}_{\text{even}}(x, y)$ ——Real part of the complex filter in the spatial domain;
 $\mathbf{G}_{\text{odd}}(x, y)$ ——Imaginary part of the complex filter in the spatial domain;
 w ——Spatial support (spatial size) of the complex filter.

The real and imaginary parts of the filtered image are calculated at each time step, yielding the structure phase distribution. The structure displacement can be projected according to the constant phase assumption of the same target. The phase variations and spatial gradient at the selected pixel are used to calculate the structure displacement implicitly:

$$\phi(x, y, t) = \arctan \left[\frac{Z_I(x, y, t)}{R_I(x, y, t)} \right] \quad (3.3)$$

$$\mathbf{D}(t) = \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{v}(t) \end{bmatrix} \alpha_p = \frac{\phi(x_1, y_1, t_1) - \phi(x_1, y_1, t)}{\|\nabla \phi(x_1, y_1, t_1)\|^2} \nabla \phi(x_1, y_1, t_1) \alpha_p \quad (3.4)$$

$$\nabla \phi(x, y, t) = \left[\frac{\partial \phi(x, y, t)}{\partial x} \quad \frac{\partial \phi(x, y, t)}{\partial y} \right]^T \quad (3.5)$$

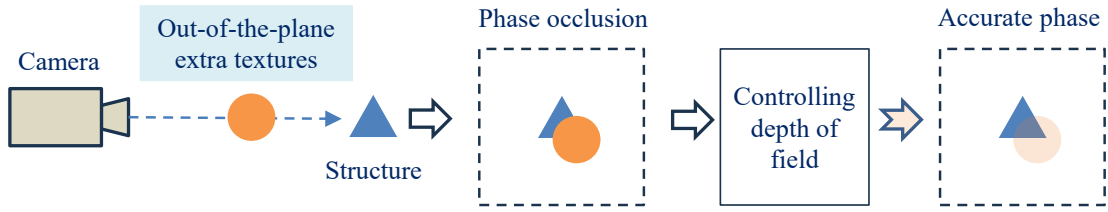
where $\phi(x, y, t)$ — Extracted phase values;
 $\mathbf{D}(t)$ — Structure displacements;
 $\mathbf{u}(t)$ — Horizontal displacements;
 $\mathbf{v}(t)$ — Vertical displacements;
 α_p — Scaling factor;
 t_1 — Reference time;
 x_1 — Reference horizontal coordinate;
 y_1 — Reference vertical coordinate;
 $\nabla \phi(x_1, y_1, t_1)$ — Spatial phase gradient at the reference time.

3.2.2 Mechanisms of complex environment effects on the extracted phase

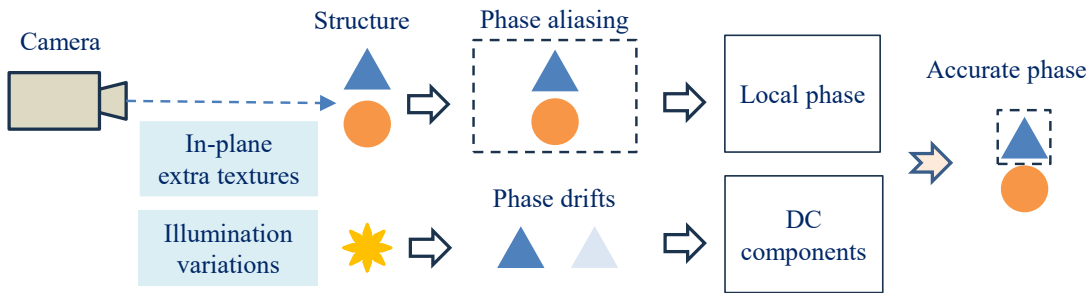
As shown in Equation (3.4), the phase extraction technique is critical to the performance of displacement estimation. Equations (3.1) and (3.2) indicate that the phase calculation for each pixel depends on the intensity value of surrounding pixels. When extra textures are present in the vicinity (including occlusions along the optical axis or adjacent interfering objects within the imaging plane), the extracted phase values incorporate

information from these disturbances and do not fully reflect the spatial position of the target structure.

This study adopts distinct phase extraction strategies for extra textures along the optical axis and those adjacent within the imaging plane, as illustrated in Figure 3.1. In Figure 3.1(a), when an extra texture lies between the camera and the target structure, occluding the optical axis, the aperture is adjusted to control the depth of field. This blurs the extra texture by placing it outside the depth of field, thereby emphasizing the features of the target structure. In Figure 3.1(b), when an extra texture is adjacent to the target structure and cannot be excluded from the depth of field along the optical axis, the spatial size of the complex filters in Equations (3.1) and (3.1) is adjusted to ensure that the surrounding pixels used for phase extraction exclude the extra texture. This requires investigating the parameters controlling the spatial size of the complex filters.



(a) Accurate phase extraction under out-of-the-plane extra textures



(b) Accurate phase extraction under in-plane extra textures and illumination variations

Figure 3.1 Accurate phase extraction under extra textures

In traditional phase-based methods, complex filters with finite spatial envelopes are employed to extract the local phase of the target structure. Among these, the Gabor complex filter stands out as one of the most widely used oriented complex filters for

extracting features from 1D or 2D signals. The theoretical formulation of the Gabor filter incorporates Gaussian-function-defined envelopes in both the spatial and frequency domains, thereby providing localized feature extraction capabilities. The typical theoretical expression of a 2D Gabor complex filter is as follows:

$$\mathbf{g}(x, y) = \mathbf{w}_r(x, y) \left\{ \cos[2\pi(f_1x + f_2y)] + j \sin[2\pi(f_1x + f_2y)] \right\} \quad (3.6)$$

$$\mathbf{w}_r(x, y) = K \exp \left[-\frac{(x \cos \theta + y \sin \theta)^2}{2\sigma^2} - \frac{(-x \sin \theta + y \cos \theta)^2}{2\sigma_n^2} \right] \quad (3.7)$$

$$\theta = \arctan \frac{f_2}{f_1} \quad (3.8)$$

where $\mathbf{g}(x, y)$ — Gabor filter in the spatial domain;
 f_1 — Central frequency of the Gabor filter in the vertical direction;
 f_2 — Central frequency of the Gabor filter in the horizontal direction;
 j — Imaginary unit;
 $\mathbf{w}_r(x, y)$ — Spatial envelope of the Gabor filter;
 K — Scaling coefficient unrelated to the phase;
 σ — Envelope standard deviation perpendicular to the target edge;
 σ_n — Envelope standard deviation parallel to the target edge;
 θ — Direction of the target edge.

The standard deviation of the filter envelope parallel to the edge of the target structure controls the number of pixels included along the direction of the geometric feature edge during phase extraction. A larger standard deviation incorporates more pixels. Since features along the edge direction of the structural geometry remain consistent, this parameter is generally set to a relatively high value to include as many edge-parallel sampling pixels as possible for noise suppression. In the case of curved features, a smaller standard deviation should be used to extract the local feature which is regarded as a straight line. On the other hand, the standard deviation perpendicular to the edge of the target structure controls the number of pixels in the direction normal to the geometric feature edge. Its value determines the spatial extent considered in phase extraction and serves as a key factor influencing phase quality in complex environments.

Typically, complex filters are constructed in the frequency domain for computational purposes. In the frequency domain, the spatial filter size relates to the half-magnitude bandwidth and the center frequency through the following expression:

$$2^\beta = 1 + \frac{2}{2\pi\sigma f - 1} \quad (3.9)$$

$$f_0 = \sqrt{f_1^2 + f_2^2} \quad (3.10)$$

where β — Half-magnitude bandwidth of the Gabor filter;
 f_0 — Central frequency of the Gabor filter.

As evident from the above equation, the spatial size of the complex filter (i.e., the standard deviation of the Gaussian envelope) is determined by the half-magnitude bandwidth and the center frequency in the frequency domain. Both the half-magnitude bandwidth and the center frequency exhibit a negative correlation with the spatial dimensions of the complex filter.

To further illustrate this relationship, Figure 3.2 presents filters with varying half-magnitude bandwidths and center frequencies. The first row displays the frequency domain representations of the filters, while the second row shows their spatial domain counterparts. When the filter's bandwidth increases from 0.6 to 1.6, the major axis of the frequency domain representation elongates, whereas the spatial domain envelope contracts. A similar trend is observed when the center frequency increases from 1/10 Hz to 1/5 Hz.

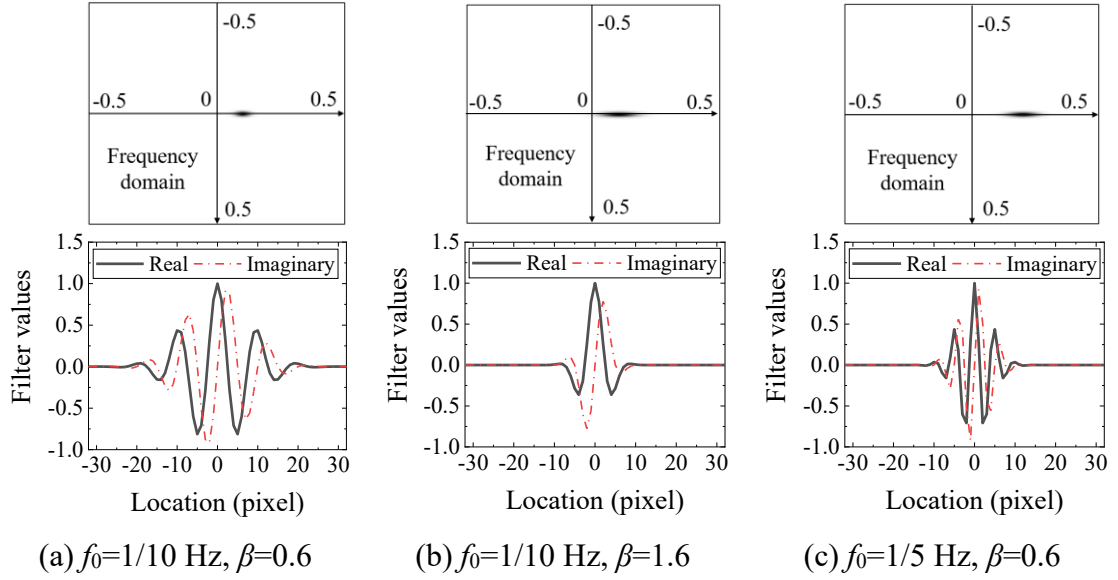


Figure 3.2 Gabor filters with different parameters in the spatial and frequency domains

Therefore, the spatial dimensions of the Gabor filter decrease with increasing center frequency and bandwidth, allowing the adjustment of filter parameters to control the spatial extent of structural feature phase extraction.

When no adjacent extra textures are present within the imaging plane, the spatial size of the Gabor filter should ideally approximate that of the structural feature (Miao et al., 2022). However, when extra textures exist near the target structure, they adversely affect the extracted phase. Under ideal conditions, for a vibrating target structure, the corresponding pixel values in the image signal should remain constant across frames at different time steps:

$$I_{t_1}(x, y) = I_{t_2}(x + u_{t_2-t_1}, y + v_{t_2-t_1}) \quad (3.11)$$

where I_{t_1} — Pixel intensity at t_1 ;
 I_{t_2} — Pixel intensity at t_2 ;
 $u_{t_2-t_1}$ — Horizontal displacement from t_1 to t_2 ;
 $v_{t_2-t_1}$ — Vertical displacement from t_1 to t_2 .

The extracted phase is also invariant:

$$\begin{aligned}
c &= \arctan \left[\frac{\sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} I_{t1}(x + \Delta x, y + \Delta y) \mathbf{G}_{\text{odd}}(\Delta x, \Delta y)}{\sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} I_{t1}(x + \Delta x, y + \Delta y) \mathbf{G}_{\text{even}}(\Delta x, \Delta y)} \right] \\
&= \arctan \left[\frac{\sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} I_{t2}(x + u_{t_2-t_1} + \Delta x, y + v_{t_2-t_1} + \Delta y) \mathbf{G}_{\text{odd}}(\Delta x, \Delta y)}{\sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} I_{t2}(x + u_{t_2-t_1} + \Delta x, y + v_{t_2-t_1} + \Delta y) \mathbf{G}_{\text{even}}(\Delta x, \Delta y)} \right]
\end{aligned} \tag{3.12}$$

where c — Phase value of the structure.

The complex filter will cover both the target structure and the extra texture in the spatial domain when extra textures exist. Consequently, the generated phase information becomes aliased, causing Equation (3.12) to no longer hold and leading to inaccurate displacement estimation results.

To verify this, consider a 1D signal containing a slope structure as the target structure. In Figure 3.3(a), a hyperbolic tangent function is represented by a solid line, while its horizontally shifted version is shown as a dashed line. Only the regions with non-zero gradients in the signal change before and after the shift. A small square wave signal is introduced as an extra texture adjacent to the target structure. As the hyperbolic tangent signal moves at a speed of 0.05 pixels per frame, the square wave remains stationary, differing in motion from the hyperbolic tangent curve.

Applying a Gabor filter with a center frequency of 1/5 Hz and a bandwidth of 0.6 to extract the phase of the target structure, and using Equation (3.4) to estimate displacement, Figure 3.3(b) compares the estimated displacements of the hyperbolic tangent signal when the extra structure is fixed at different distances (indicated as D-20, D-15, D-10, and D-8) from the target structure. It can be observed that when the square wave is positioned further from the hyperbolic tangent signal, the estimated displacement closely matches the GT. However, when the extra texture is closer to the hyperbolic tangent signal, significant errors occur in the estimated displacement. Therefore, when

extra textures are present near the target structure, the spatial scale of the filter must be optimized to avoid such interference.

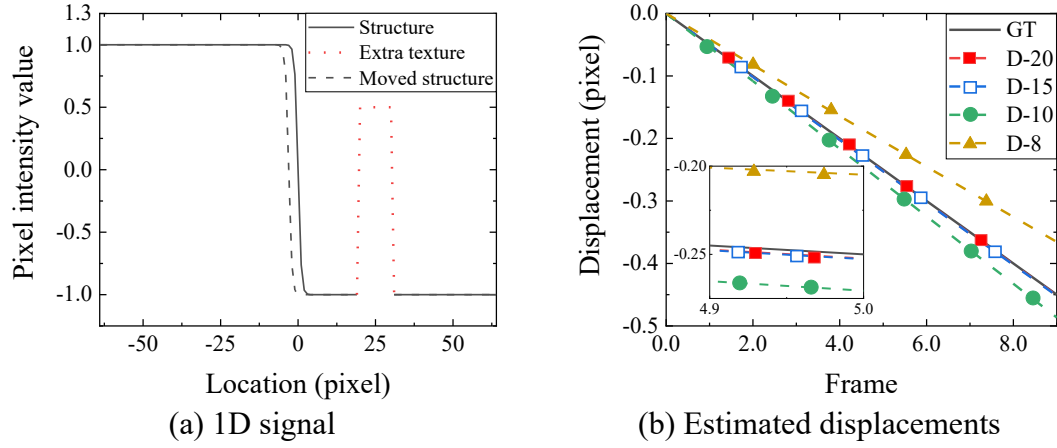


Figure 3.3 Adverse effects of extra textures on estimating 1D signal displacements

According to Equation (3.9), the spatial scale of a typical Gabor filter is determined by its bandwidth and center frequency. Therefore, this study investigates the use of filters with different parameters to optimize displacement estimation performance. A set of signals with extra textures and a set without extra textures are used for comparison, as shown in Figure 3.4. "g-f1/5-b0.8" denotes displacement estimation using a Gabor filter with a center frequency of 1/5 Hz and a bandwidth of 0.8 under extra textures, while "g-f1/5-b0.8-nd" represents displacement estimation using the same Gabor filter parameters in the absence of extra texture interference, and so forth.

The results indicate that in Figure 3.4(a), when the filter center frequency is 1/5 Hz and the bandwidth is 0.8, the presence of the extra texture significantly affects the displacement estimation results, resulting in a relative error of 27% compared to the interference-free case, with a maximum absolute error of approximately 1 pixel. However, when the bandwidth is increased to 1.2, this adverse effect is markedly suppressed, reducing the relative error to 5.56% and the maximum absolute error to about 0.03 pixels. With a further bandwidth increase to 1.6, the relative error in estimated displacement decreases to 0.6%. In Figure 3.4(b), the displacement estimation results

using a Gabor filter with a center frequency of $1/4$ Hz exhibit a trend similar to that observed with the $1/5$ Hz case. Increasing the frequency-domain bandwidth effectively mitigates the adverse impact of in-plane adjacent extra textures on displacement estimation. Moreover, at a bandwidth of 1.2, the displacement estimated with a center frequency of $1/4$ Hz is less affected by in-plane adjacent extra texture compared to the $1/5$ Hz case, yielding a maximum absolute error of approximately 0.005 pixels. In summary, sufficiently high center frequencies and bandwidths enable the Gabor filter to suppress the adverse effects of extra textures on displacement estimation.

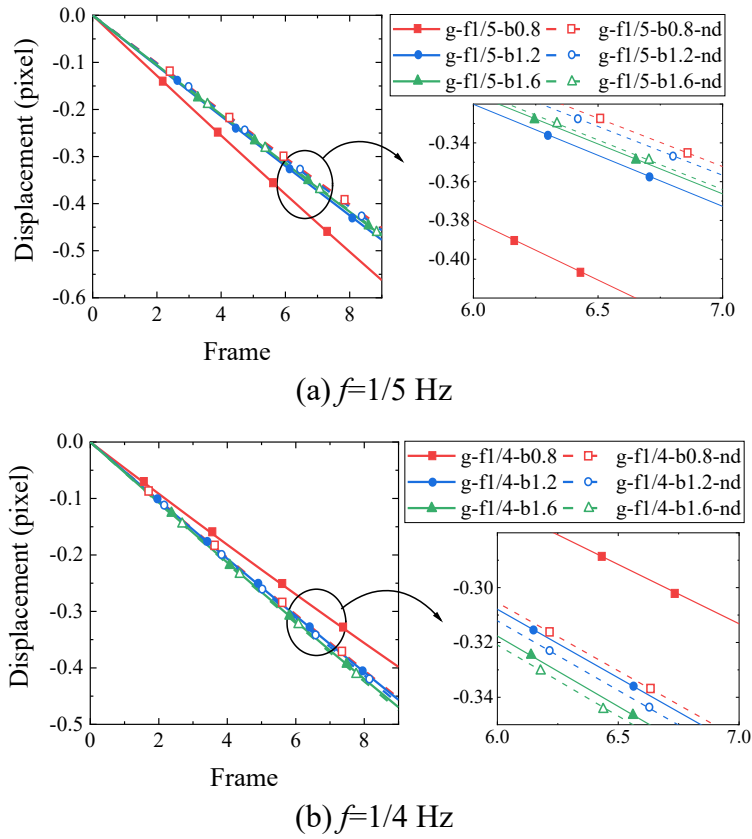


Figure 3.4 Displacements estimated by Gabor filters with different central frequencies

Further examining the spatial dimensions of the aforementioned filters helps clarify how larger bandwidths and higher center frequencies contribute to localized phase-based displacement estimation. Figure 3.5 displays the disturbed signal and the envelope profiles of filters with different parameters. "Envelop-f1/5-b0.8" represents the envelope of a filter with a center frequency of $1/5$ Hz and a half-magnitude bandwidth of 0.8. This

envelope partially overlaps with the extra texture, causing the extracted phase and displacement to incorporate both the movement of the hyperbolic tangent signal over time and the unrelated motion of the extra texture. As a result, the constant phase assumption is violated, leading to inaccurate displacement estimation. In contrast, the envelope of the filter with a center frequency of 1/4 Hz and a bandwidth of 1.6 is notably more compact, avoiding the overlap with the extra texture. Under these conditions, the extracted displacement originates solely from the hyperbolic tangent signal.

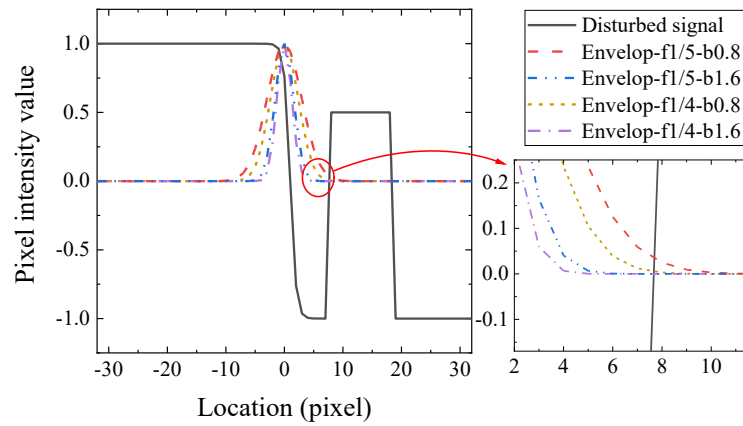


Figure 3.5 Disturbed 1D signal and filter envelopes with different parameters

However, when a filter has a large bandwidth, the uncertainty principle implies that its frequency domain response will include DC components (Boukerroui et al., 2004). This causes low-frequency spatial variations in pixel values, such as those induced by changing illumination conditions, to significantly interfere with displacement estimation results. To simultaneously mitigate the effects of both extra textures and illumination variations in complex environments, this paper proposes a displacement estimation method based on Log-Gabor filters (Log-Gabor-based Motion Estimation, LGME) (Y. Wang et al., 2023).

3.2.3 LGME method in complex environments

The average pixel intensity value of an image signal is determined by illumination conditions and may vary over time. When the frequency domain composition of a filter includes a DC component, such variations in pixel values will cause corresponding changes in the extracted phase values. In this scenario, illumination variations across different image frames are reflected in the real part of the signal after filter convolution:

$$R_I(x, y) = \sum_{\Delta x = -w/2}^{w/2} \sum_{\Delta y = -w/2}^{w/2} [I(x + \Delta x, y + \Delta y) + b] G_{\text{even}}(\Delta x, \Delta y) \quad (3.13)$$

where b — Pixel value variation induced by illumination changes.

When a Gabor filter has a large bandwidth and contains negative frequency and DC components, the influence of illumination variations described by the above equation cannot be neglected, as the spatial-domain integral of the filter's real part is non-zero in such cases. Based on this, the disturbance in the real part of the signal caused by illumination variations can be derived as follows:

$$\sum_{\Delta x = -w/2}^{w/2} \sum_{\Delta y = -w/2}^{w/2} b G_{\text{even}}(\Delta x, \Delta y) = ba \quad (3.14)$$

where a — Integration of the real part filter in the spatial domain.

However, the Gabor filter has to be designed with a large bandwidth to suppress the in-plane extra texture according to Section 3.2.2. Therefore, the Log-Gabor filter (Field, 1987; Xie et al., 2019) without negative and DC components is introduced to estimate displacements:

$$LG(f) = \exp \left\{ - \frac{[\ln(f / f_0)]^2}{2[\ln(\sigma_r / f_0)]^2} \right\} \quad (3.15)$$

where $LG(f)$ — Log-Gabor filter in the frequency domain;

f_0 — Central frequency of the Log-Gabor filter;

σ_r — Standard deviation of the radial envelope in the logarithmic scale.

The bandwidth of the Log-Gabor filter is:

$$\beta = 2 \sqrt{\frac{2}{\ln(2)}} \left| \ln \left(\frac{\sigma_r}{f_0} \right) \right| \quad (3.16)$$

Figure 3.6 compares the frequency domain characteristics of 1D Gabor and Log-Gabor filters. Here, "g-fl/4-b0.8" denotes a Gabor filter with a center frequency of 1/4 Hz and a bandwidth of 0.8, while "lg-fl/4-b0.8" represents a Log-Gabor filter with the same center frequency and bandwidth, and so forth. It can be observed that the Log-Gabor filter does not include negative frequency or DC components even under large bandwidth conditions. In the 2D polar coordinate system, the Log-Gabor filter attenuates in the circumferential direction following a Gaussian function:

$$LG(f, \theta) = \exp \left\{ -\frac{[\ln(f / f_0)]^2}{2[\ln(\sigma_r / f_0)]^2} \right\} \exp \left[\frac{-(\theta - \theta_0)^2}{2\sigma_\theta^2} \right] \quad (3.17)$$

where θ — Direction in the polar coordinate;
 θ_0 — Direction of the measured target;
 σ_θ — Standard deviation of the circumferential envelope.

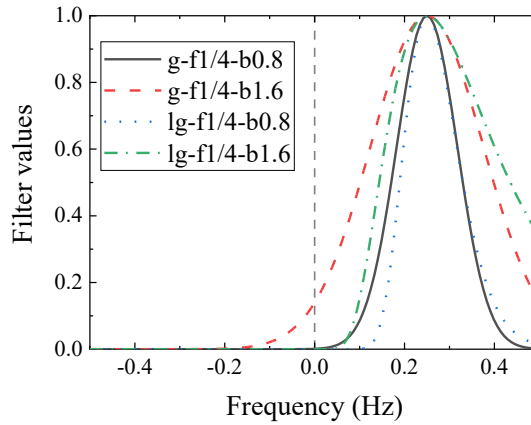


Figure 3.6 Gabor and Log-Gabor filters in the frequency domain

The Log-Gabor filter lacks a theoretical expression in the spatial domain, making it impossible to directly design its spatial dimensions to avoid interference from in-plane extra textures. To address this, this paper proposes a localization index based on magnitude values to characterize the energy of the target structure extracted by the filter relative to the image background. For different filter parameters, a higher localization

index indicates that the extracted structural features exhibit a higher SNR under the current filter settings:

$$Loc = \frac{w^2 \sqrt{\mathbf{R}_I^2(x_0, y_0) + \mathbf{Z}_I^2(x_0, y_0)}}{\sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} \sqrt{\mathbf{R}_I^2(\Delta x, \Delta y) + \mathbf{Z}_I^2(\Delta x, \Delta y)}} \quad (3.18)$$

where Loc ——Localization index.

The procedure of LGME is presented in Figure 3.7.

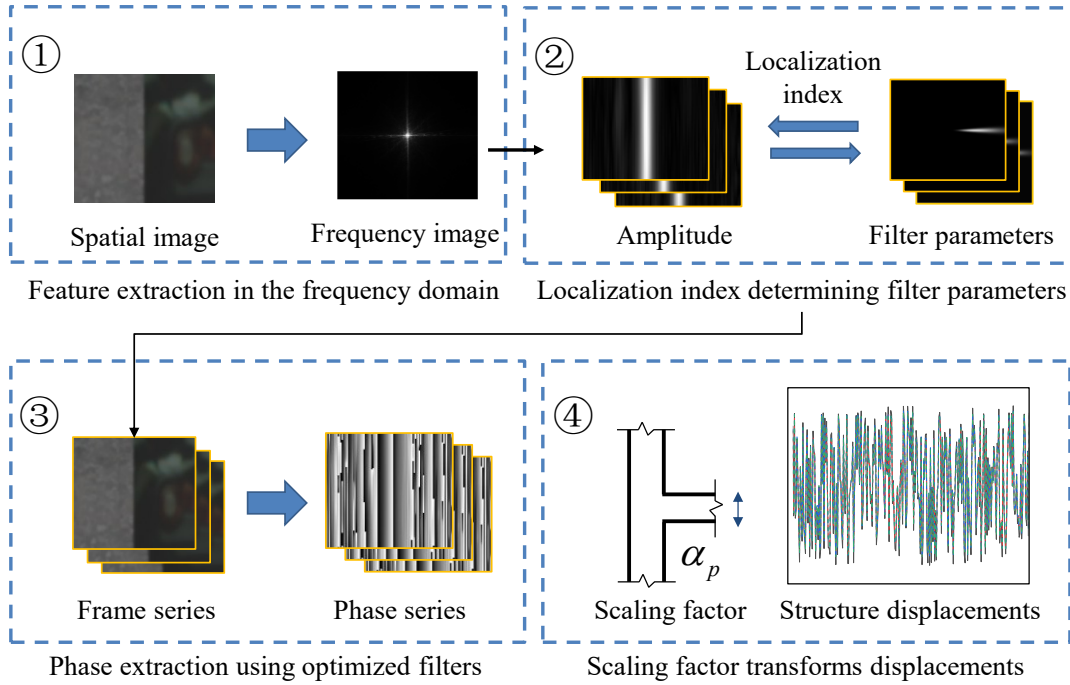


Figure 3.7 Procedure of LGME

First, the image signal is transformed into the frequency domain to improve computational efficiency. Second, the reference image frame is processed using Log-Gabor filters with different parameters to calculate the localization index. Third, the parameter set yielding the maximum localization index is selected as the optimal filter, which is then applied to all image frames for phase extraction. Finally, a projection between phase and displacement is established. The scaling factor between pixels and

physical displacements is calibrated, estimating the displacement time history of the structure.

3.3 Performance of LGME in complex numerical environments

3.3.1 Performance evaluation in complex 1D signals

A hyperbolic tangent function is used as the vibrating target structure signal, while a box function is employed as a stationary extra texture positioned 7 pixels away from the origin. The overall intensity of Signal 1 increases at a rate of 0.02 per frame, whereas Signal 2 experiences a stronger intensity rise of 0.04 per frame to simulate more pronounced illumination variations, as illustrated in Figure 3.8. Both Signal 1 and Signal 2 move horizontally at a speed of 0.1 pixels per frame.

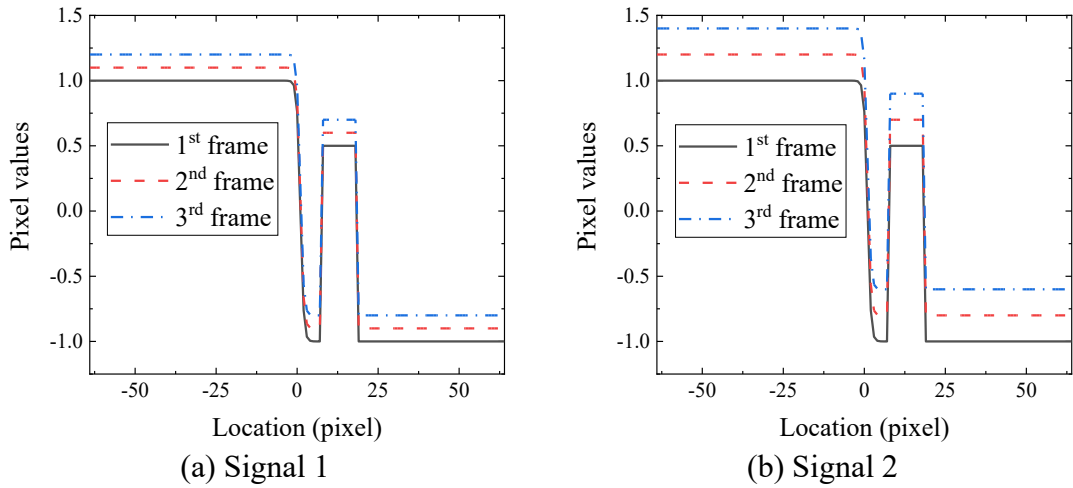


Figure 3.8 1D signals with different solar-induced time-variant pixel values

The reference frame of the signal (identical for Signals 1 and 2) is processed using different filter parameters through Equations (3.1) and (3.2), and the localization index is calculated via Equation (3.18) to optimize the Log-Gabor filter parameters. Figure 3.9 shows the localization index computed for the 1D signal using Log-Gabor filters with

varying parameters. It is evident that the index increases with higher center frequencies and reaches its maximum within the bandwidth range of 1.5 to 3.0. Therefore, two sets of optimal filter parameters, namely, center frequency of 1/4 Hz with bandwidths of 1.8 and 2.2, are selected for displacement estimation.

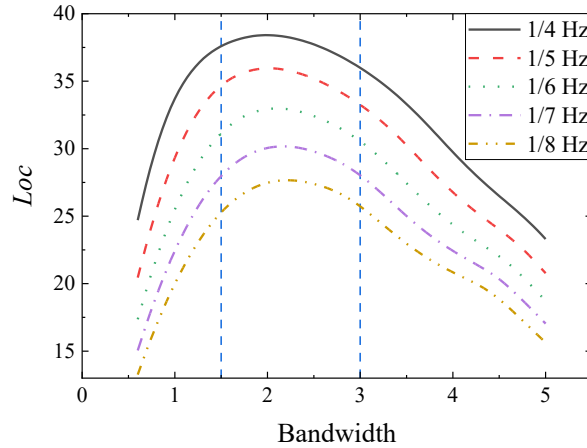
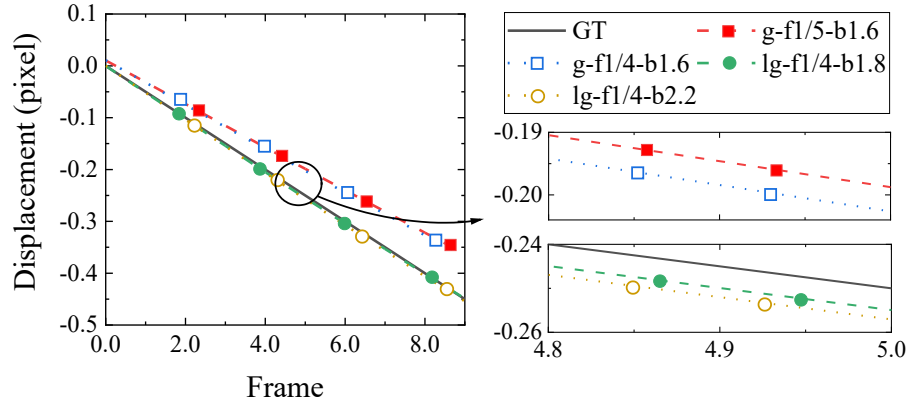
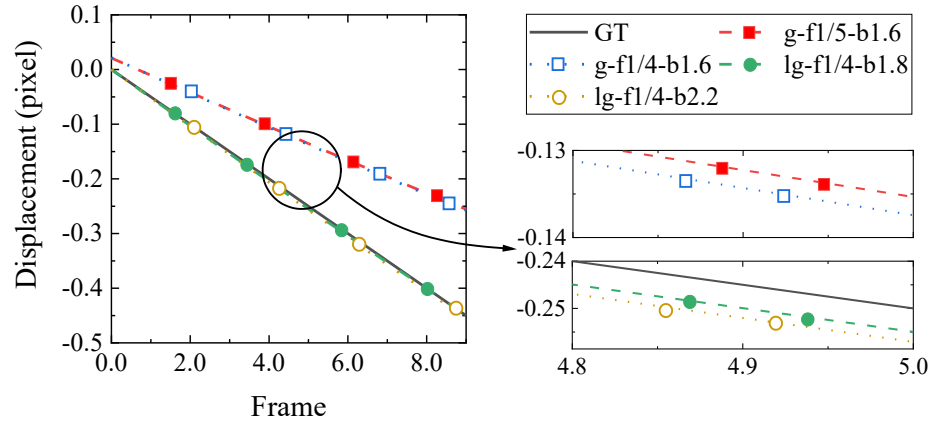


Figure 3.9 Localization indices of different Log-Gabor filters in 1-D signal cases

The displacement of the 1D signals are estimated using the conventional Gabor-based method and the proposed LGME method. For the Gabor filter, a bandwidth of 1.6 is selected to minimize the influence of adjacent in-plane extra texture. The displacement estimation results in Figure 3.10 demonstrate that the displacements obtained by the LGME method align closely with the GT, with a maximum instantaneous error of only 0.005 pixels and a relative error of 1.1%. In contrast, the conventional method estimates displacements with drifts of up to 0.08 pixels for Signal 1 and 0.2 pixels for Signal 2, corresponding to relative errors of 17.8% and 44.4%, respectively. Thus, the proposed LGME method effectively handles complex environments involving both adjacent in-plane extra textures and time-varying illumination conditions.



(a) Estimated displacements of Signal 1



(b) Estimated displacements of Signal 2

Figure 3.10 Displacements estimated by different methods under complex environments

3.3.2 2D signals: in-plane extra textures and illumination variance

A series of image datasets is generated using the rendering software Blender to validate the LGME performance for 2D signals. Each frame in the dataset includes an image background, a moving block structure as the target of interest. An adjacent strip structure serving as an in-plane extra texture, as shown in Figure 3.11. The target block structure measures 0.3 m in length and width, while the extra texture has a width of 0.01 m. The camera is positioned 10 m above the imaging plane with a focus distance of 10 m, a focal length of 200 mm, and an aperture of F/2.8. The motion of the target block structure is applied as velocity vectors at different time steps. In this example, the displacement

magnitude of the target object per frame is less than 0.3 pixels, demonstrating the subpixel displacement estimation capability of the proposed method.

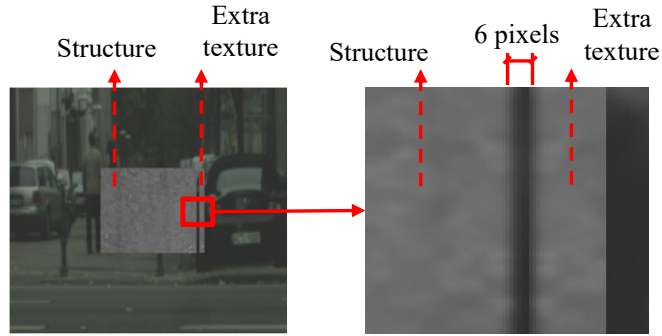


Figure 3.11 Structure to be measured with in-plane extra texture in the image plane

Two datasets of image sequences, each comprising 500 frames with a resolution of 1024×1024 pixels, are rendered. Dataset 1 considers only the interference of strip extra textures, while Dataset 2 incorporates both extra textures and illumination variations. The illumination changes are achieved by progressively increasing solar radiation intensity from the first to the 500th frame, as illustrated in Figure 3.12. In both datasets, the target block structure moves randomly in the horizontal direction by no more than 0.3 pixels per frame, while the background and the extra texture remain stationary. The displacement of the target block structure is identical in both datasets, as shown in Figure 3.13.



Figure 3.12 Dataset 2 with in-plane extra textures and illumination variations

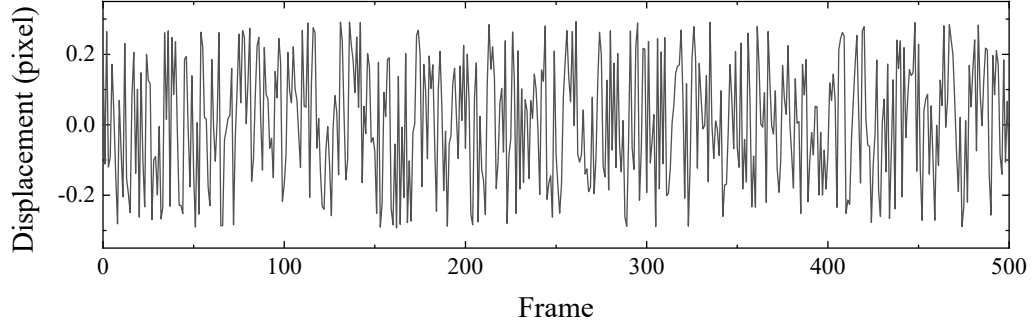


Figure 3.13 Displacements of the structure to be measured

The localization index is calculated to determine the parameters of the Log-Gabor filter. Figure 3.14 depicts the *Loc* values generated by filters with different center frequencies and bandwidths. The maximum values occur within the bandwidth range of 1.5 to 3.5, which is consistent with the 1D case. Based on the *Loc* results, two parameter sets are selected as the optimal filters for displacement identification: a center frequency of 1/5 Hz with a bandwidth of 2.2, and a center frequency of 1/3 Hz with a bandwidth of 2.8.

First, the displacement estimation performance of the proposed LGME method in the presence of adjacent in-plane extra texture is validated. Figure 3.15 displays the estimated displacement using the traditional Gabor method and the LGME method for the first 50 frames of Dataset 1. The Log-Gabor filters optimized with the two sets of parameters based on the *Loc* demonstrate excellent performance in the presence of the in-plane extra texture, achieving a maximum absolute error of only 0.02 pixels between the estimated displacement and the GT. In contrast, the traditional Gabor filter with a small bandwidth of 0.6 produces severely deviated displacement results, significantly underestimating the actual displacement. When the bandwidth of the traditional Gabor filter is increased to 1.2, the displacement results generally align with the GT but still exhibit an absolute error of 0.06 pixels at the lower bound of the displacement time history.

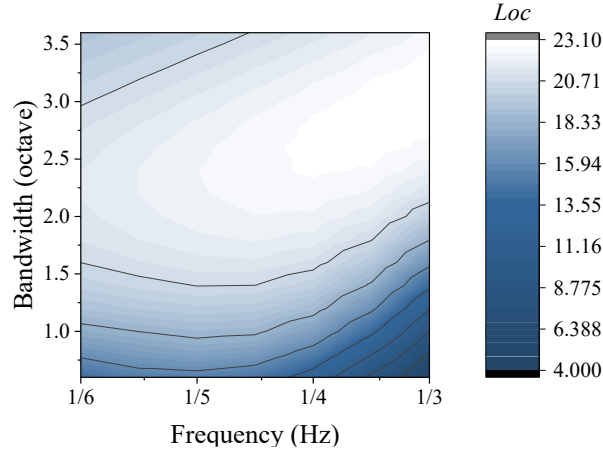


Figure 3.14 *Loc* indexes calculated on the first frame of the 2D dataset

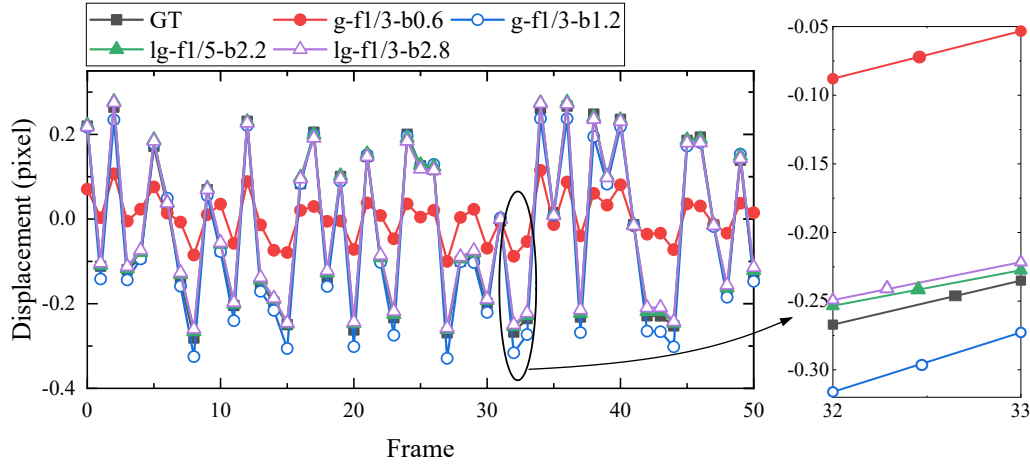


Figure 3.15 Estimated displacements with in-plane extra textures

For the complex environment involving both adjacent in-plane extra texture and time-varying illumination, the displacement estimation performance of traditional Gabor filters with different parameters is first evaluated using Dataset 2. Figure 3.16 shows the displacement results for the last 50 frames, where illumination changes are most pronounced. When the traditional Gabor filter "g-f1/3-b1.2" uses a large bandwidth of 1.2, its frequency domain response covers DC components, resulting in significant displacement drift due to illumination variations (reaching up to 1 pixel, which is three times the actual displacement amplitude). On the other hand, "g-f1/3-b0.6" with a bandwidth of 0.6 avoids DC components, but the larger spatial scale associated with the

smaller bandwidth allows interference from adjacent in-plane extra textures to adversely affect displacement estimation, leading to significantly underestimated displacement amplitudes.

Therefore, while a large bandwidth enables traditional Gabor filters to mitigate interference from adjacent in-plane extra textures in Dataset 1, it causes the traditional method to become susceptible to illumination variations in Dataset 2. In conclusion, traditional Gabor-based methods are unsuitable for complex environments where both adjacent in-plane extra textures and time-varying illumination are present.

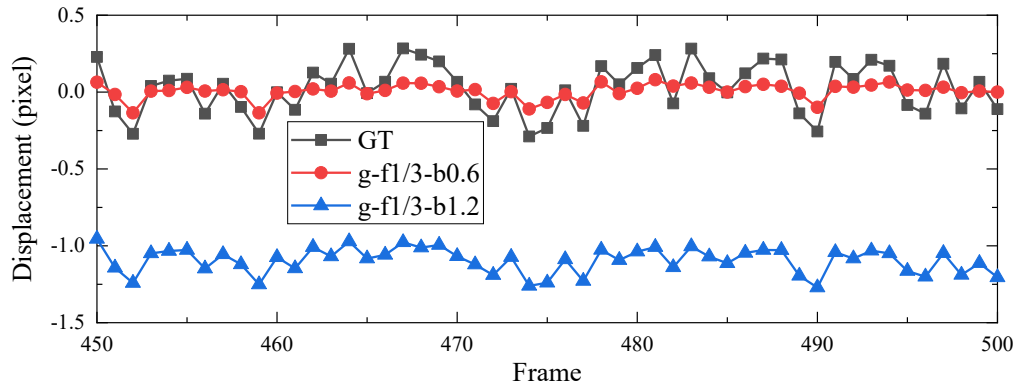


Figure 3.16 Displacements estimated by traditional Gabor filters under in-plane extra textures and illumination variations

The proposed LGME method was applied to Dataset 2, which involves the complex environment of adjacent in-plane extra textures and illumination variations. Figure 3.17 presents the estimated displacement for the last 50 frames, where illumination variation is most significant, along with a zoom-in view of a local segment. As shown, the displacement results estimated by the two sets of optimized Log-Gabor filters almost overlap with the GT, indicating that the filter parameters optimized by the localization index deliver consistently superior performance. Even under maximum illumination variation, the maximum displacement error is only 0.03 pixels, with an overall relative displacement error of 3.4%.

In conclusion, compared to traditional Gabor-based methods, the LGME method determines parameters via the localization index, enabling the extraction of local phase and displacement from the target structure with a compact spatial scale. Moreover, the inherent absence of DC components in the Log-Gabor filter allows accurate displacement estimation in complex environments characterized by both adjacent in-plane extra textures and illumination variations.

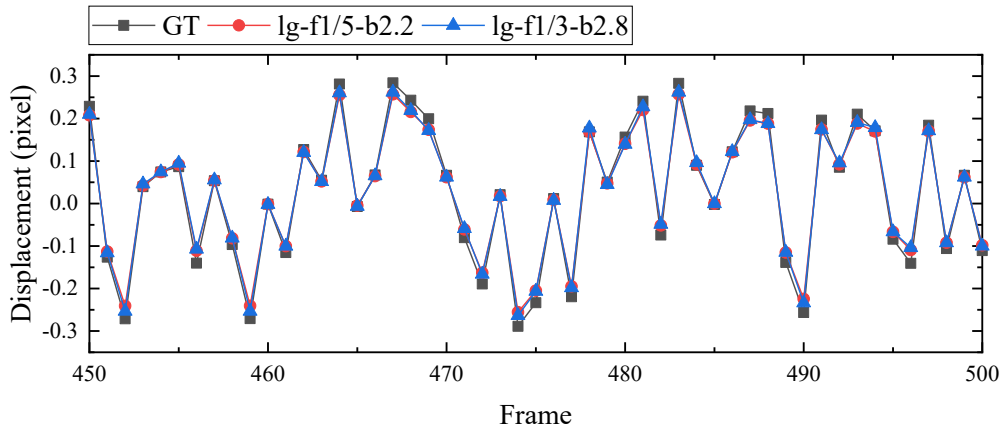


Figure 3.17 Displacements estimated by LGME under in-plane extra textures and illumination variations

3.3.3 2D signals: out-of-the-plane extra texture occlusions

As analyzed in Figure 3.1, in practical structural measurement environments, interference can arise not only from in-plane extra textures but also from out-of-the-plane extra textures along the optical axis, which may occlude the line of sight and prevent the extraction of structural features. This study adjusts the imaging depth of field to blur out-of-the-plane extra textures, thereby improving the accuracy of displacement estimation for the target structure. The depth of field of digital imaging is primarily controlled by two parameters: the lens aperture and the focal length. Since the focal length is typically determined by the required field of view for capturing the region of interest, this study focuses on controlling the depth of field by adjusting the lens aperture size.

Accordingly, a rendering dataset incorporating out-of-the-plane extra textures is built using Blender, enabling control over the aperture value to achieve different depths of field, while the image intensity level remains constant. The structure of the rendering model is shown in Figure 3.18. The target structure is consistent with that in Section 3.3.2: a cube with a side length of 0.3 m. The out-of-the-plane extra texture, located along the optical axis, is a cylinder with a radius of 0.02 m and a depth of 0.1 m. The extra texture is stationary, positioned 8 m above the target structure and offset 0.03 m to the right to occlude the right side of the target structure. The camera is fixed 10 m above the target structure with a focal length of 200 mm. The aperture is set to F/16, F/8, F/5.6, F/2.8, and F/1.4 to control the depth of field. The horizontal displacement of the target structure is the same as that in Figure 3.13.

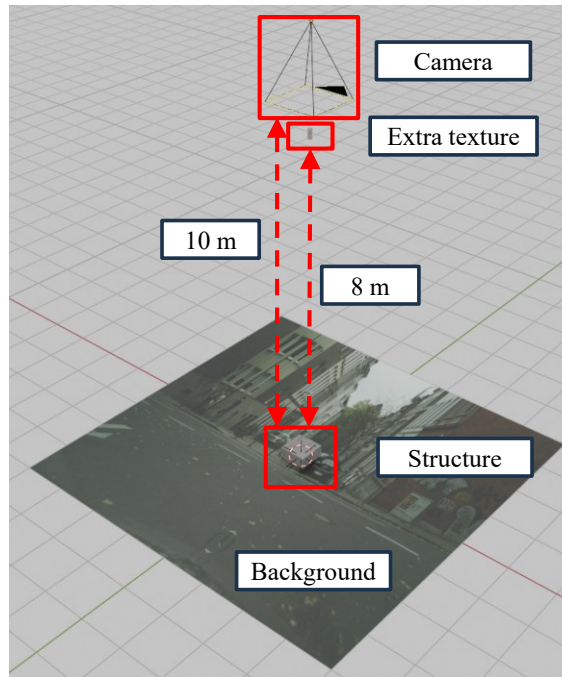


Figure 3.18 Rendered model with extra textures along the light axis

The rendered datasets under different aperture settings are shown in Figure 3.19. At the smallest aperture (F/16), the extra texture, located in front of the target structure, blocks the view of the target. When the aperture is increased to F/8, the reduced depth of field blurs the extra texture to some extent, making the occluded right edge of the target

structure faintly visible. As the aperture is further increased to F/5.6 and F/2.8, the blurring of the extra texture intensifies, and the structural features along the right edge of the target become clearer. At the largest aperture (F/1.4), the extra texture appears as a faint halo in the shallow depth of field, and the right edge of the target structure becomes clearly visible compared to the case with the smallest aperture (F/16).

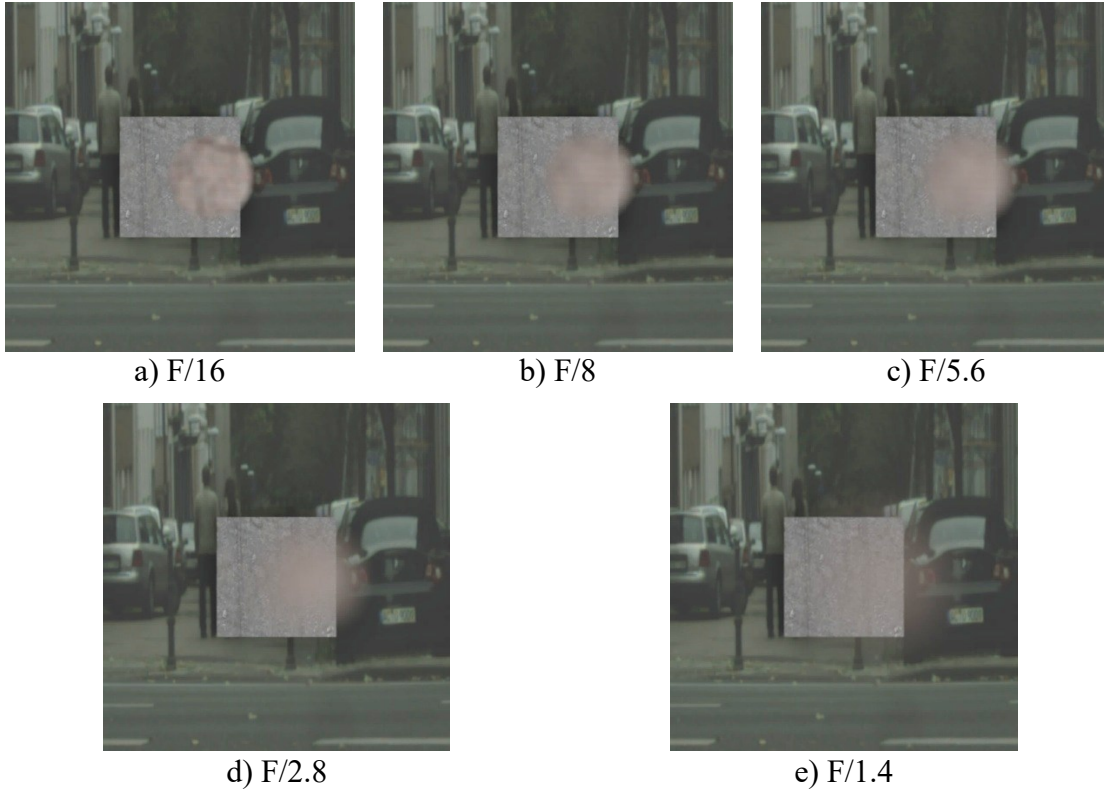


Figure 3.19 Dataset frames under different apertures

The proposed LGME method is employed to estimate the horizontal displacement of the target structure under different aperture settings in the presence of out-of-the-plane extra textures. The displacement measurement point is selected at the midpoint of the right edge of the target block to simulate the most unfavorable scenario of occlusions. The displacement estimation results are shown in Figure 3.20.

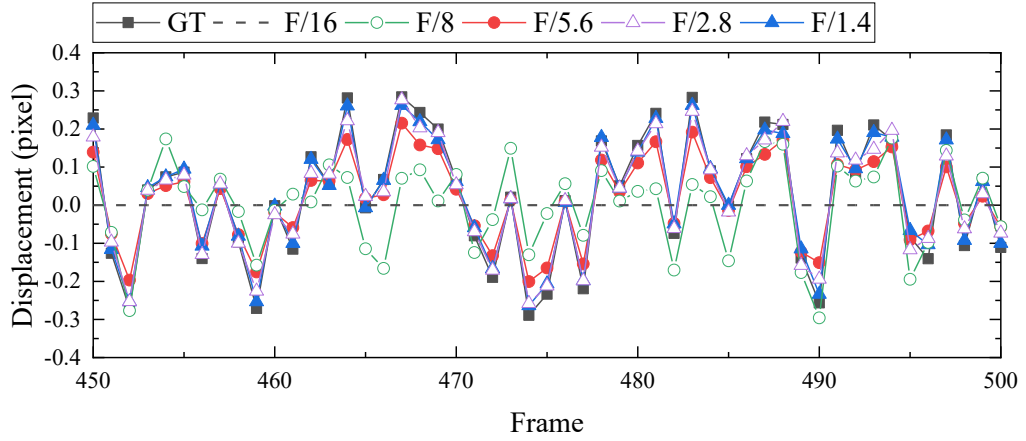


Figure 3.20 Displacements estimated by LGME under extra textures along the light axis

At the smallest aperture of F/16, the phase variation of the target structural feature can not be observed at the measurement point due to occlusion, resulting in a measured displacement of zero throughout. When the aperture is increased to F/8, the phase variation trend of the target structure becomes detectable, but the extracted phase values are perturbed by residual features of the extra texture, leading to significantly underestimated displacement values. With larger apertures (F/2.8 and F/1.4), the extra texture is sufficiently blurred, enabling accurate estimation of the displacement time history of the target structure, achieving a relative error of 2.8%.

In conclusion, when out-of-the-plane extra texture along the optical axis causes occlusion in the measurement environment, using a larger aperture can effectively blur such interference and improve the accuracy of displacement estimation for the target structure.

3.4 Performance of LGME in complex experimental environments

3.4.1 In-plane extra textures and illumination variations

This subsection validates the performance of the LGME method under laboratory conditions involving both in-plane extra textures and time-varying illumination. A five-

story frame structure serves as the target, with an in-plane extra texture placed adjacent to the left column to evaluate the robustness of displacement estimation. Time-varying illumination is simulated by dynamically adjusting the intensity of indoor lighting during observation. Figure 3.21 illustrates the experimental setup. A ZCAM S6G camera equipped with a Canon EF24-105 mm lens was used to capture video at a resolution of 3840×2160 pixels and a frame rate of 29.97 Hz. The lens was configured with a focal length of 24 mm, an aperture of F/4, a shutter speed of $1/60$ s (twice the frame rate), and an ISO of 2000. The camera was positioned approximately 6 m from the target structure. The five-story frame measures $1.8 \text{ m} \times 0.5 \text{ m} \times 0.5 \text{ m}$. The scaling factor between image coordinates and physical coordinates is 0.5 mm/pixel.

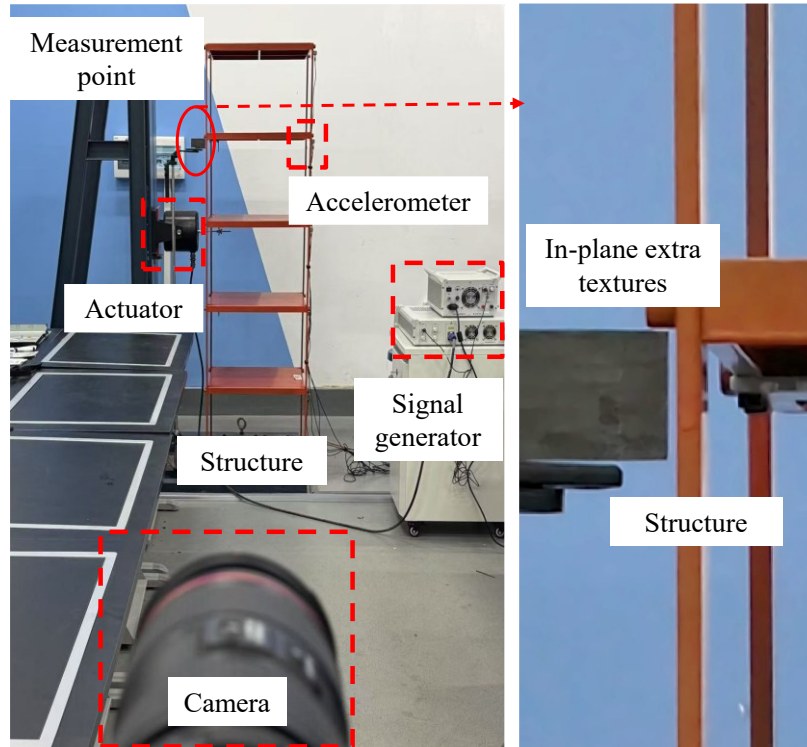


Figure 3.21 Experimental set-up with in-plane extra textures and illumination variations

During the experiment, a 7 Hz harmonic load was applied to the third floor via an actuator. Since low-frequency quasi-static displacements were negligible, a Lance LC0119 accelerometer was installed on the right side of the frame column to record acceleration

at a sampling rate of 30 Hz. The displacement derived from integrating the 1-Hz high-passed acceleration data is used as the benchmark for structural displacement.

An 80 mm $\times$ 40 mm black square mounted on a tripod and positioned to the left of the target frame column is the in-plane extra texture. In the captured video, the gap between the left edge of the target column and the interfering object is approximately 6 pixels, representing an adverse scenario where the complex filter spatially overlaps with the extra texture. During the experiment, only the frame vibrated, while the extra texture fixed on the tripod remained stationary. Displacement is estimated by tracking the left edge of the fourth-story frame column.

Three test cases were designed. In Case 1, only the in-plane adjacent extra texture was considered. The objective of this case is to validate the local feature extraction capability and subpixel displacement estimation performance of the LGME method. Displacement is estimated using both the traditional Gabor filter method and the LGME method, and the results are compared against the benchmark.

The localization index used to determine the Log-Gabor parameters is shown in Figure 3.22. The distribution of *Loc* in the experiment is similar to that observed in the two-dimensional simulated signals, indicating that the *Loc* reflects the inherent optimal spatial scale of the Log-Gabor filter for extracting structural feature phase, independent of the observed structure. Based on the *Loc* calculation results, Log-Gabor filter parameters with higher center frequencies and bandwidths between 1.5 and 3.5 exhibit better capability for extracting local structural information. Meanwhile, to avoid periodic phase jumps caused by excessively high center frequencies, the optimal filter parameters are determined as a center frequency of 1/4 Hz and a bandwidth of 2.4.

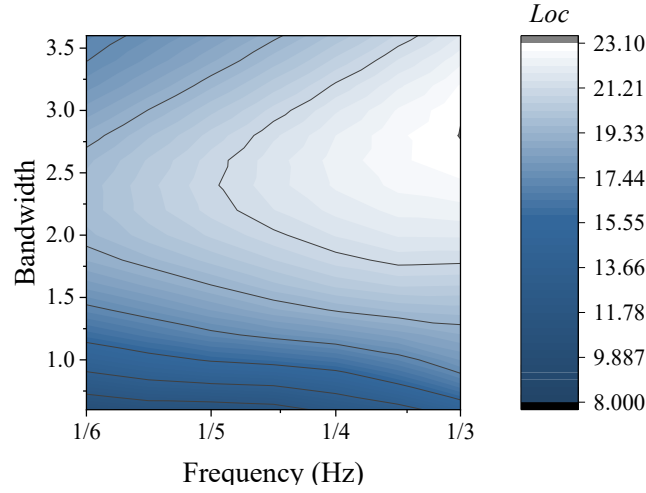


Figure 3.22 *Loc* values calculated on the first frame of the captured video

For comparison, two sets of Gabor filter parameters are selected: one with a larger spatial scale and one with a smaller spatial scale, denoted as "g-f1/5-b0.6" and "g-f1/4-b1.2," respectively. The parameters determined by the *Loc* for LGME are denoted as "lg-f1/4-b2.4." As shown in Figure 3.23, the horizontal displacement amplitude at the measurement point is approximately 0.6 mm, corresponding to 1.2 pixels. It can be observed that all estimation results exhibit displacement trends consistent with the benchmark before and after excitation. However, the displacement amplitudes estimated by the traditional phase-based method using Gabor filters deviate from the benchmark.

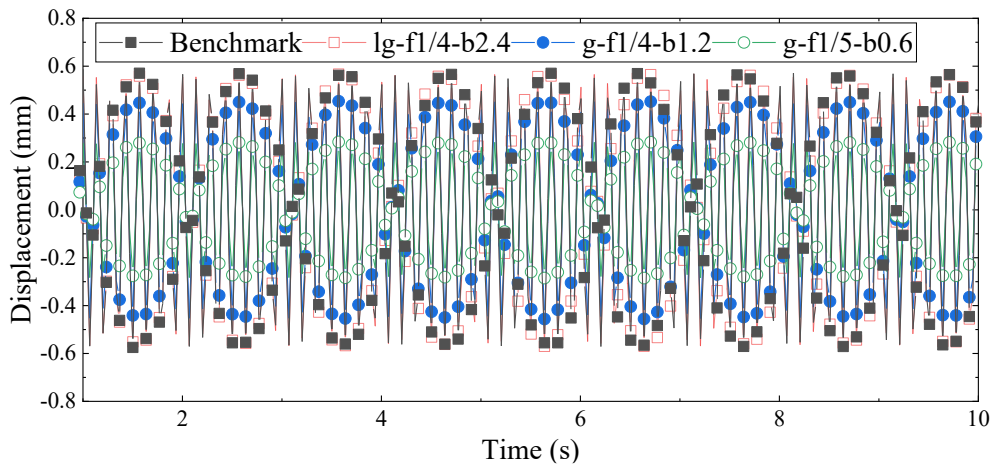


Figure 3.23 Case 1: displacements estimated under neighbor extra texture in the image plane

Figure 3.24 depicts a zoom-in view of the displacement details for Case 1. It can be observed that the displacement estimated by the proposed LGME method aligns closely with the benchmark. In contrast, the traditional Gabor filter with small bandwidth ("g-f1/5-b0.6") yields a displacement amplitude of only 0.3 mm, significantly lower than the benchmark of 0.6 mm. The optimized large-bandwidth Gabor filter ("g-f1/4-b1.2") underperforms compared to the proposed method, consistently underestimating the displacement by approximately 0.1 mm.

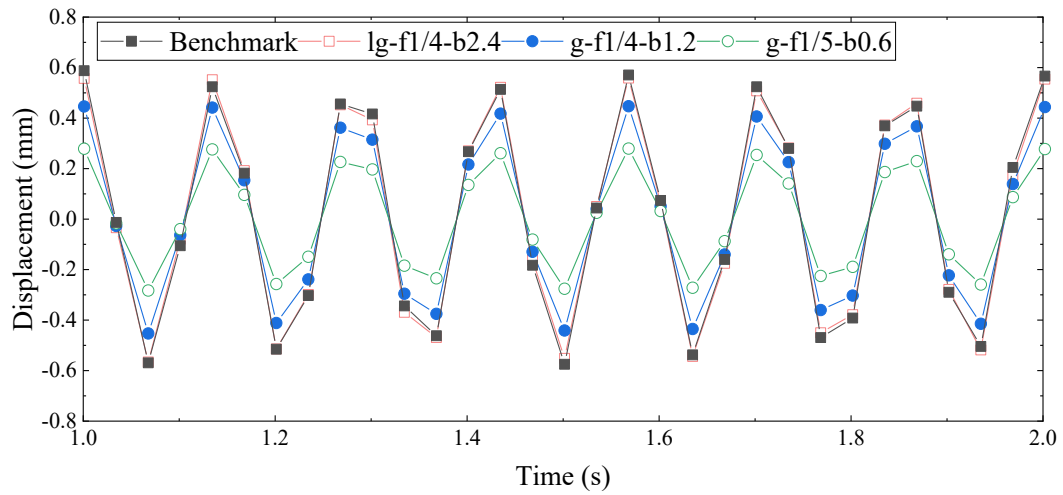


Figure 3.24 Zoom-in of Case 1

Table 3.1 summarizes the relative errors of displacement estimation for different methods in Case 1. The proposed method maintains a relative error below 5%, whereas the traditional methods exhibit relative errors exceeding 10%. These results demonstrate that the LGME method can accurately extract structural displacements when an in-plane extra texture is located as close as 6 pixels from the measurement point.

Table 3.1 Relative errors of the displacements estimated by different methods in Case 1

Methods	lg-f1/4-b2.4	g-f1/5-b0.6	g-f1/4-b1.2
Relative errors	3.48%	36.98%	14.33%

In Case 2, the environmental setup remained identical to Case 1, but the actuator load was reduced to evaluate the performance of the proposed method in estimating subtle displacements under the adverse influence of in-plane adjacent extra textures. The displacement estimation results are shown in Figure 3.25. In this case, the amplitude of the displacement is only 0.02 mm, corresponding to a pixel-scale amplitude of 0.04 pixels. Although minor local fluctuations are observed, the displacements estimated by the proposed method align well with the benchmark. In contrast, the traditional method using Gabor filters exhibited significant deviations from the benchmark.

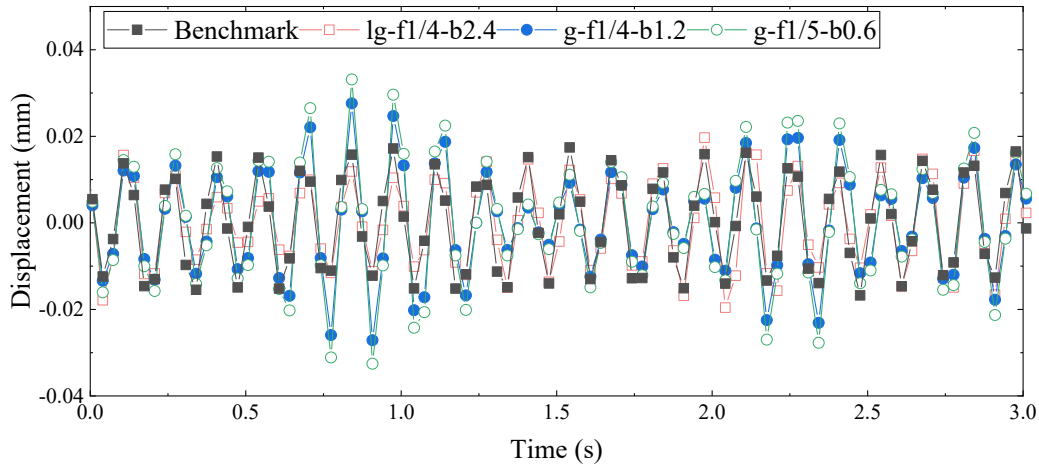


Figure 3.25 Case 2: subtle displacements estimated under neighbor extra texture
in the image plane

Table 3.2 summarizes the relative errors of displacement estimated by each method in Case 2. The traditional method is adversely affected by the influence of in-plane adjacent extra textures in subtle displacements, resulting in relative errors exceeding 60%. In contrast, the proposed method maintained the relative error below 20% even under such extremely subtle displacement conditions.

Table 3.2 Relative errors of the displacements estimated by different methods in Case 2

Methods	lg-fl/4-b2.4	g-fl/5-b0.6	g-fl/4-b1.2
Relative errors	18.35%	74.68%	60.52%

Case 3 combined the influence of time-varying illumination and in-plane adjacent extra textures on displacement estimation performance. Figure 3.26 illustrates the time-varying illumination conditions during the video capture. The indoor lighting intensity was reduced at 5 s and maintained at the lowest level until the end of the recording.

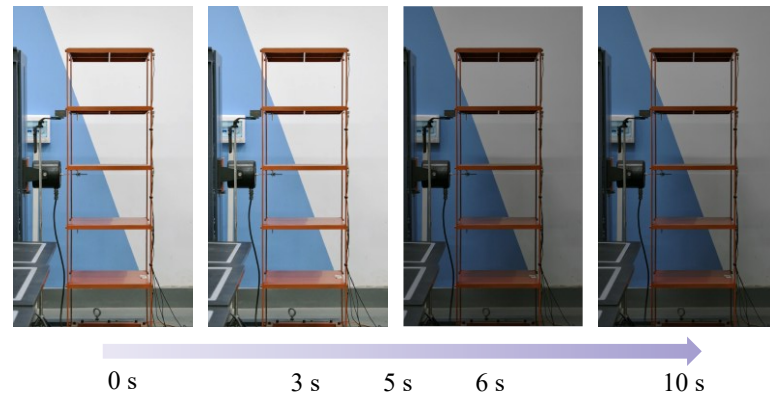


Figure 3.26 Illumination variations in Case 3

Displacements are estimated using the traditional Gabor filter and the LGME method, respectively. The results are shown in Figure 3.27. During the first 5 s with constant illumination, the traditional method with a large bandwidth ("g-fl/4-b1.2") is adversely affected by in-plane adjacent interference, resulting in displacement errors similar to the observations in Case 1. Meanwhile, the traditional method with a small bandwidth ("g-fl/5-b0.6") significantly underestimated the displacement amplitude.

When illumination began to diminish after 5 s, the large-bandwidth traditional method ("g-fl/4-b1.2"), due to its inability to filter out the DC components in the digital images, exhibits notable drift at the upper bound of the displacement curve as a result of illumination changes. The proposed LGME method, using parameters optimized by the

localization index ("lg-fl/4-b2.4"), remains unaffected by the variation in illumination and maintains a displacement trend consistent with the benchmark.

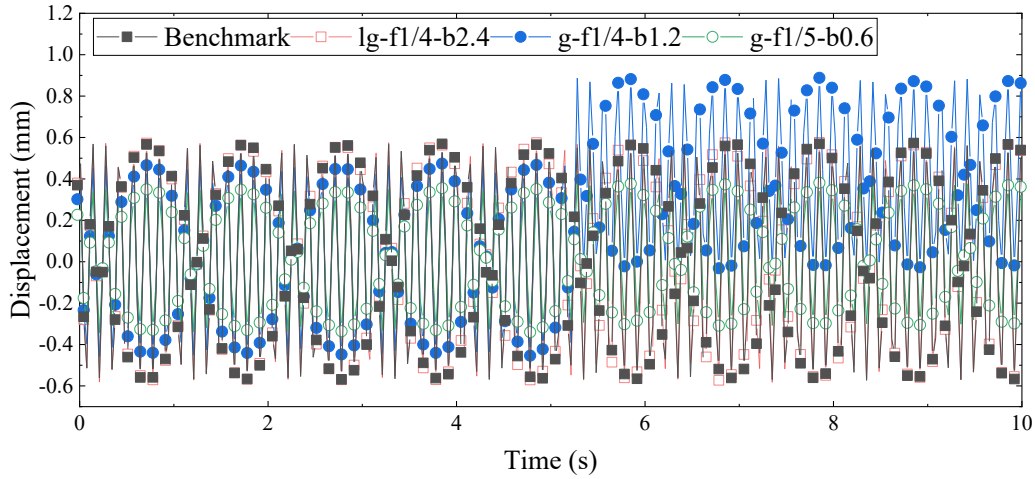
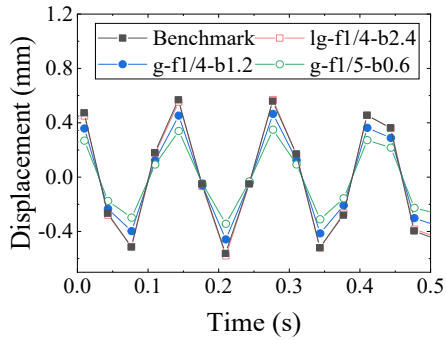
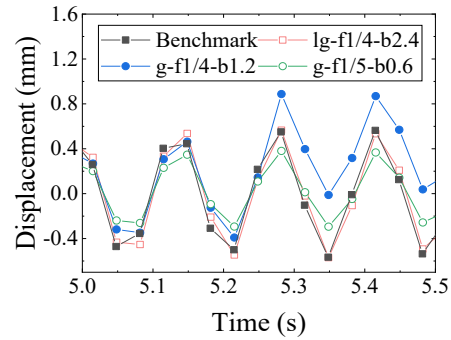


Figure 3.27 Case 3: displacements estimated under neighbor extra texture in the image plane and illumination variations

Figure 3.28 shows the enlarged detail. During the period from 0 to 0.5 s before the illumination change, the displacements estimated by the traditional methods underestimate the amplitude due to the influence of in-plane adjacent extra textures. In the period from 5 to 5.5 seconds after the illumination change, the small-bandwidth traditional method underestimates the displacement amplitude, while the large-bandwidth traditional method exhibits displacement drift caused by the variation in illumination.



(a) Before the illumination variation



(b) After the illumination variation

Figure 3.28 Zoom-in of Case 3

In contrast, the proposed LGME method effectively handles the coupled adverse effects of both in-plane adjacent extra textures and illumination variation, achieving a relative displacement error of only 5.3%, as summarized in Table 3.3.

Table 3.3 Relative errors of the displacements estimated by different methods in Case 3

Methods	lg-f1/4-b2.4	g-f1/5-b0.6	g-f1/4-b1.2
Relative errors	5.30%	41.62%	54.37%

The PSD of the displacement estimated by LGME is compared with that of the accelerometer in Figure 3.29, indicating the consistent natural frequency of 7.16 Hz.

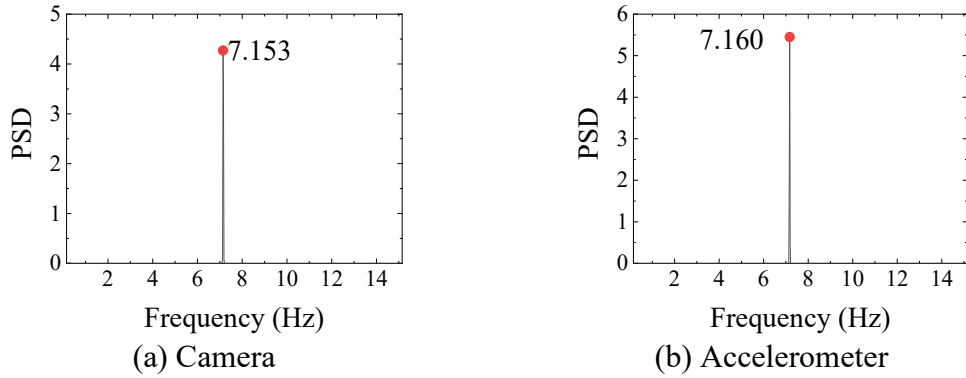


Figure 3.29 PSD of Case 3

3.4.2 Out-of-the-plane extra texture occlusions

This section validates the displacement estimation performance of the proposed LGME method when addressing occlusion caused by out-of-the-plane extra textures along the optical axis. This is achieved by adjusting the camera aperture to blur the extra textures. The out-of-the-plane extra texture, positioned along the optical axis, is an 80 mm $\times$ 40

mm black square fixed on a tripod approximately 0.5 m in front of the lens, as shown in Figure 3.30.



Figure 3.30 Extra texture along the light axis

The camera aperture was set to F/14, F/10, F/8, F/6.3, and F/4 to investigate the effect of varying degrees of blur on displacement estimation performance. To maintain consistent brightness across video frames, the shutter speed was automatically adjusted to control light intake (the shutter speed was always faster than 1/60 s, twice the sampling frequency of 29.97 Hz), while the ISO was fixed to make the same SNR.

With this configuration, the amount of light captured in each digital image of the target structure remained consistent, with the only variation being the depth of field. As shown in Figure 3.31, when a small aperture of F/14 was used, the large depth of field caused the out-of-the-plane extra texture to completely occlude the left column of the target frame. At F/10, the edges of the extra texture were slightly blurred, making the left column faintly visible. As the aperture increased to F/8, the right edge of the extra texture was further blurred, though its partially blurred outline still covered the left column, resulting in low contrast of the geometric features. At the relatively larger apertures of F/6.3 and F/4, the extra texture was sufficiently blurred, allowing clear visibility of the left column.

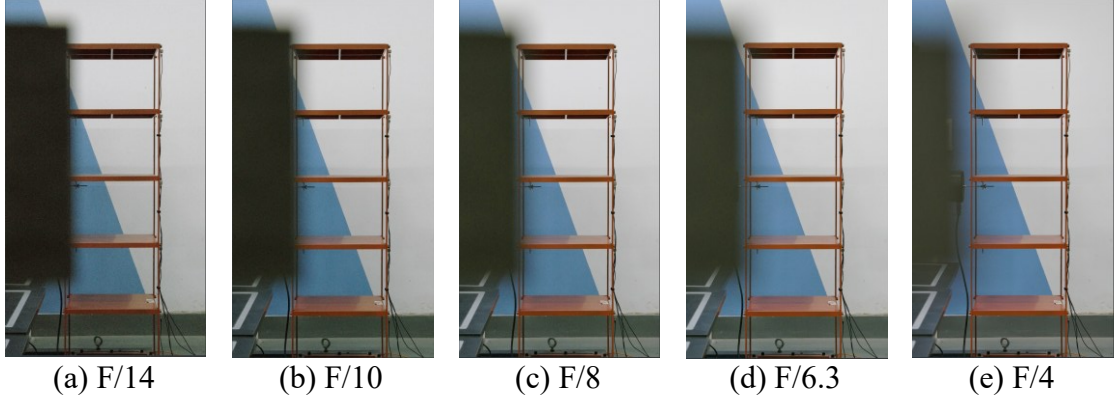


Figure 3.31 Captured structure under different apertures

The actuator excited the frame structure at a frequency of 7 Hz. The proposed LGME method is employed to estimate the horizontal displacement of the fourth-story left column, and the results are compared with the benchmark, as shown in Figure 3.32. When an aperture of F/14 was used, the measurement point was completely occluded, yielding no identifiable displacement. With an aperture of F/10, the blurred edges of the extra texture still covered the measurement point, reducing the contrast of the column's geometric features. This leads to a decrease in the spatial gradient of the phase in Equation (3.5), resulting in a significantly overestimated displacement amplitude of 2.5 mm, far exceeding the benchmark amplitude of 0.6 mm. A similar issue occurs at F/8, though the increased blurring helps constrain the estimated displacement amplitude to below 1 mm.

At an aperture of F/6.3, the estimated displacement amplitude largely aligns with the benchmark, though localized deviations persist. With an aperture of F/4, the extra texture near the measurement point on the left column was sufficiently blurred, enabling the estimation of a displacement amplitude of approximately 0.6 mm. Moreover, the displacement values match closely with the benchmark.

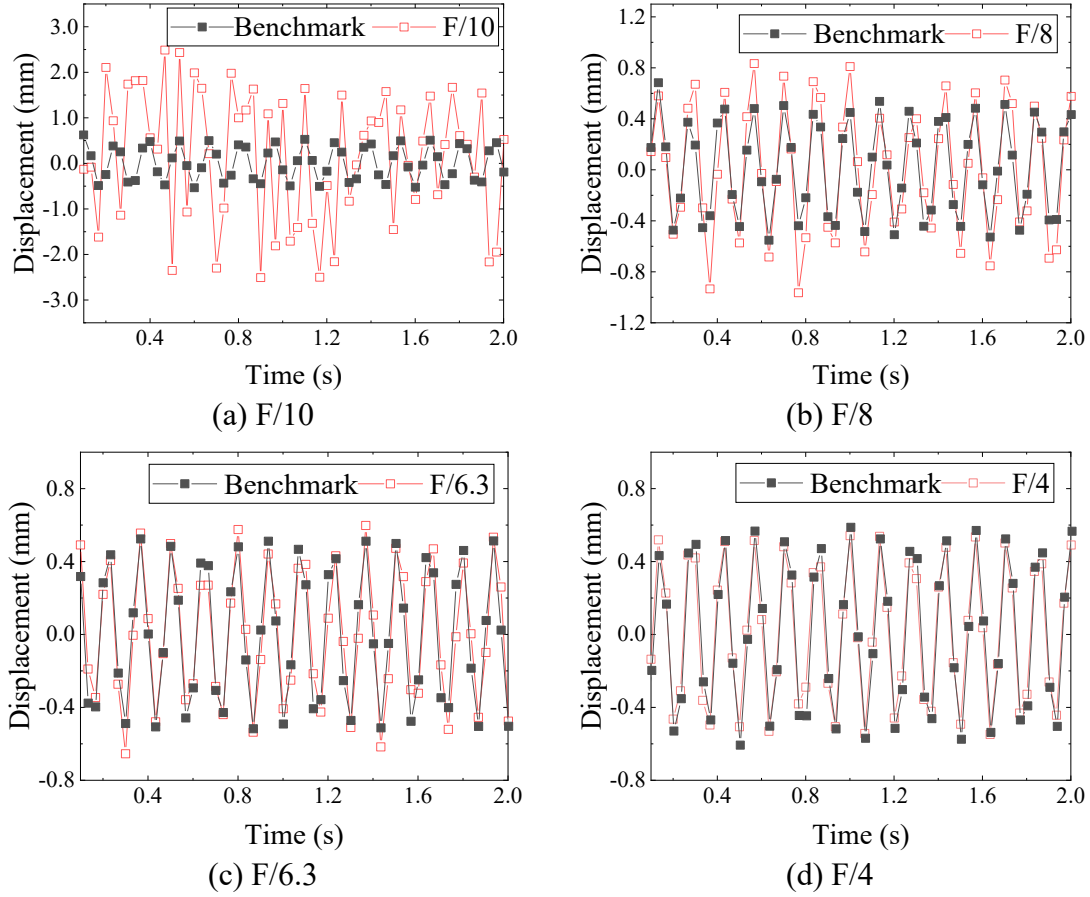


Figure 3.32 Estimated displacements under different apertures

Table 3.4 summarizes the relative errors of estimated displacements under different aperture settings. Compared to the inability to estimate displacement with smaller apertures, using a larger aperture (F/4) to blur the out-of-the-plane extra texture along the optical axis effectively reduces the relative error to 5.74%.

Table 3.4 Relative errors of the displacements estimated under different apertures

Apertures	F/14	F/10	F/8	F/6.3	F/4
Relative errors	/	121.87%	44.92%	20.90%	5.74%

3.5 Summary

This chapter first establishes a phase-displacement projection relationship based on digital image phase, investigating the adverse effects of two factors in complex environments, namely, extra textures and illumination variations, on displacement estimation.

To address extra textures, the mechanism by which they affect the quality of phase extraction from the target structure is analyzed. Local phase extraction strategies are proposed, including depth of field control along the optical axis and optimization of complex filter parameters within the imaging plane.

For the issue of pixel value drift caused by time-varying illumination, which is unrelated to the mechanical behavior of the target structure, the mechanism and limitations of traditional complex filters under multi-factor coupling in complex environments are analyzed. A Log-Gabor complex filter phase extraction method, optimized via a localization index, is proposed to achieve robust phase-based displacement estimation under complex environments.

The performance of the proposed method in complex environments is validated through numerical simulations and laboratory tests.

CHAPTER 4

DISPLACEMENT ESTIMATION FROM NOISY DATA WITH UNKNOWN DISTRIBUTIONS USING PHASE- PRESERVING DENOISING

4.1 Introduction

Chapter 3 addresses the adverse effects of extra textures and illumination variations on phase-based displacement estimation from digital images, while limitations in camera sensor performance or suboptimal settings often introduce image noise as an additional complex factor. Typical digital image sensor noise comprises both pattern noise and random noise, whose distribution characteristics are generally difficult to directly determine. Noise intensity is jointly influenced by the amount of light captured during exposure and the inherent noise generated by the sensor. For instance, shutter speed directly determines exposure time; aperture controls the light intake area; and ISO amplifies the converted electrical signal, which enhances image brightness but also magnifies noise.

Although existing image denoising techniques can effectively improve visual quality, they do not guarantee accurate phase extraction of the target structure for displacement estimation. Therefore, it is necessary to investigate how image denoising algorithms affect phase and displacement, and to develop effective denoising methods that preserve structural phase and enhance displacement accuracy.

To improve displacement estimation accuracy under unknown noise distributions, this chapter examines the impact mechanism of image noise on displacement estimation in

the complex domain. The feature differences between the geometric characteristics of the target structure and noise are analyzed in the multi-frequency complex domain. The multi-frequency pyramid is proposed to denoise structural features and preserve the structural phase. To address the challenge of the unknown spatiotemporal distribution of digital image noise, an instability indicator based on complex-domain magnitudes is introduced. The threshold between unknown noises and structural features across multi-frequency bands in the complex pyramid is quantified, thereby improving displacement estimation accuracy under unknown noises.

4.2 Phase-preserving denoising under unknown noise distributions

4.2.1 Effects of noise on phase-based displacement estimation

The noisy image is regarded as the addition of the true image and a noise signal:

$$\mathbf{I}(x, y) = \mathbf{T}(x, y) + \mathbf{n}(x, y) \quad (4.1)$$

where $\mathbf{T}(x, y)$ ——True image;
 $\mathbf{n}(x, y)$ ——Additive noise;
 $\mathbf{I}(x, y)$ ——Noisy image.

The noise is integrated into the phase after the convolution with complex filters:

$$\mathbf{R}_I(x, y) = \mathbf{R}_T(x, y) + \mathbf{R}_n(x, y) = \sum_{\Delta x = -w/2}^{w/2} \sum_{\Delta y = -w/2}^{w/2} \mathbf{I}(x + \Delta x, y + \Delta y) \mathbf{G}_{\text{even}}(\Delta x, \Delta y) \quad (4.2)$$

$$\mathbf{Z}_I(x, y) = \mathbf{Z}_T(x, y) + \mathbf{Z}_n(x, y) = \sum_{\Delta x = -w/2}^{w/2} \sum_{\Delta y = -w/2}^{w/2} \mathbf{I}(x + \Delta x, y + \Delta y) \mathbf{G}_{\text{odd}}(\Delta x, \Delta y) \quad (4.3)$$

where $\mathbf{R}_I(x, y)$ ——Real part of the noisy structural feature in the complex domain;
 $\mathbf{R}_T(x, y)$ ——Real part of the clean structural feature in the complex domain;
 $\mathbf{R}_n(x, y)$ ——Real part of the noise in the complex domain;
 $\mathbf{Z}_I(x, y)$ ——Imaginary part of the noisy structural feature in the complex domain;

$Z_T(x, y)$ ——Imaginary part of the clean structural feature in the complex domain;

$Z_n(x, y)$ ——Imaginary part of the noise in the complex domain.

The phase of the noisy structural feature is calculated by combining Equations (4.2) and (4.3):

$$\phi(x, y) = \arctan \left[\frac{Z_I(x, y)}{R_I(x, y)} \right] = \arctan \left[\frac{Z_T(x, y) + Z_n(x, y)}{R_T(x, y) + R_n(x, y)} \right] \quad (4.4)$$

where $\phi(x, y)$ ——Phase of the noisy structural feature.

The phase of the same structural feature is identical throughout the time domain in ideal environments.

$$\begin{aligned} \phi(x, y, t_1) &= \phi(x + u_{t_2-t_1}, y + v_{t_2-t_1}, t_2) \\ &= \arctan \left[\frac{Z_T(x, y, t_1)}{R_T(x, y, t_1)} \right] = \arctan \left[\frac{Z_T(x + u_{t_2-t_1}, y + v_{t_2-t_1}, t_2)}{R_T(x + u_{t_2-t_1}, y + v_{t_2-t_1}, t_2)} \right] \end{aligned} \quad (4.5)$$

where $\phi(x, y)$ ——Phase of the clean structural feature;

$u_{t_2-t_1}$ ——Horizontal displacement from time t_1 to t_2 ;

$v_{t_2-t_1}$ ——Vertical displacement from time t_1 to t_2 .

However, the structural phase varies due to the noise variation in complex environments, ruining the constant phase assumption:

$$\phi(x, y, t_1) \neq \phi(x + u_{t_2-t_1}, y + v_{t_2-t_1}, t_2) \quad (4.6)$$

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \frac{\phi(x_1, y_1, t_1) - \phi(x_1, y_1, t)}{\|\nabla \phi(x_1, y_1, t_1)\|^2} \nabla \phi(x_1, y_1, t_1) \quad (4.7)$$

where $u(t)$ ——Horizontal displacement time history;

$v(t)$ ——Vertical displacement time history;

(x_1, y_1) ——Structural feature coordinate at the reference time step;

$\nabla \phi(x_1, y_1, t_1)$ ——Phase spatial gradient calculated at the reference time step.

As evident from the above equation, the adverse effect of noise on displacement estimation is embedded in both the phase values and the phase spatial gradient. When the structural features exhibit low contrast relative to the background (where the phase spatial gradient is small), noise can severely distort the estimated displacement. Equation (4.4) indicates that the impact of noise on the phase is comprehensively reflected in both the real and imaginary components of the structural feature signal.

However, conventional digital image denoising methods operate in the real number domain to remove noise. Moreover, denoising algorithms may also perturb the true image component, potentially further distorting the phase of the target structural features. Therefore, it is of great significance to examine the noise-induced disturbances in both the real and imaginary parts of the structural features, improving displacement estimation accuracy through phase-preserving denoising.

4.2.2 Denoising and phase preservation using the complex pyramid

4.2.2.1 Distinguishing noises using the amplitude

The responses of the denoising method to noises, the target of interest, and other details are different and distinguishable. Figure 4.1 shows an example of a 1D signal consisting of a slope, a step, and Gaussian noise. The step is used to simulate the edge between the target of interest and the background in an image, while the slope and Gaussian noise correspond to the smooth region and image noises, respectively. A Log-Gabor filter with a central frequency of 0.25 Hz and a bandwidth of 0.7 octaves is used to filter the signal. The results show that the step, which is the content of interest, generates significant responses in the real part, the imaginary part, and the amplitude. The phase value has large gradients that can be used in estimating the motion within the step region. Although the noises after Pixel 125 lead to fluctuations in the phase, they can be distinguished because the amplitude is much weaker than that of the step region. The filter rarely responds to other smooth regions, including the slope between Pixels -350 and -125, and

the straight line between Pixels 20 and 120, where the phase value is meaningless because a small value of the real part causes zero-crossing errors in the phase. This example shows that the noises can be apart from the content of interest (i.e., the step region) by considering the amplitude extracted by the quadrature filter.

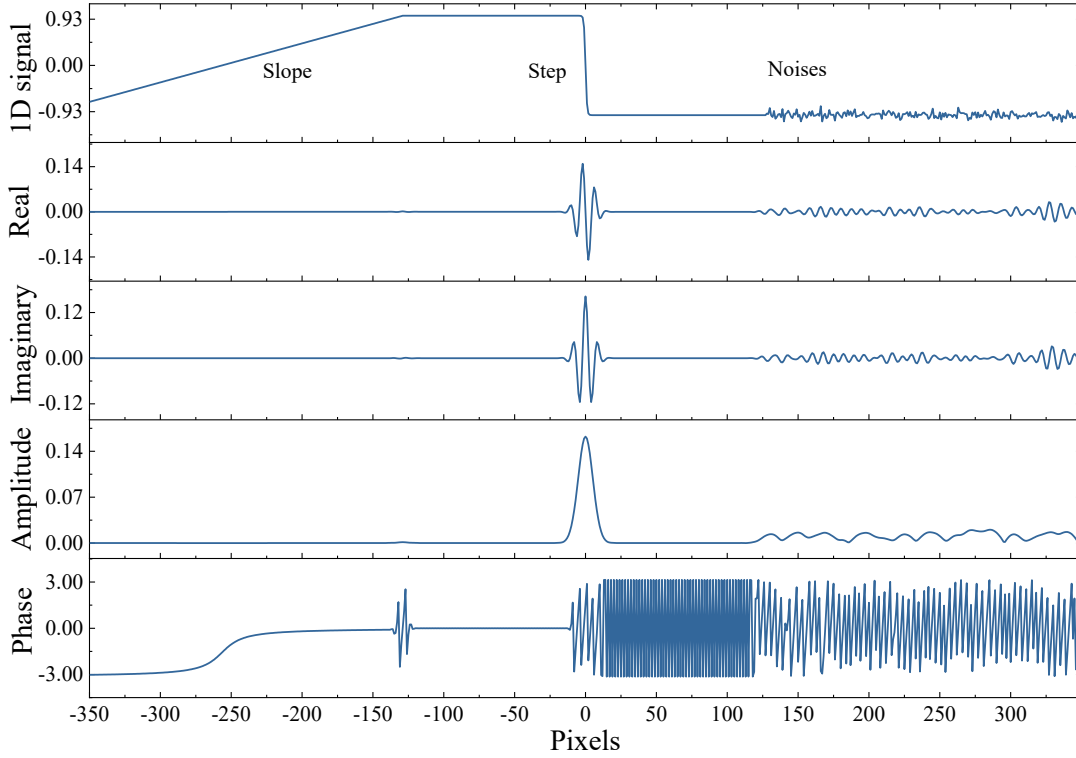


Figure 4.1 1D signal and the filtered results by convolving the log-Gabor filter

However, given a specific quadrature filter, the amplitude of the filtered noise with a high level can be close to that of the filtered interested content. Figure 4.2 demonstrates the amplitude value of a filtered 1D signal with different noise levels. The upper row is the 1D signal consisting of a step and Gaussian noise, whose standard deviation increases from 0.01 to 0.1. The lower row is the amplitude map by filtering the signals with the same log-Gabor filter. The step at the original point can be distinguished in the case of the weak noise with a standard deviation of 0.01. The step is buried by the noise when the noise level has a standard deviation of 0.1. Although the quadrature filter fails to detect the content of interest under the strong noise in this example, the filter size can control

the features extracted within a specific pixel range, which will be analyzed in the next subsection.

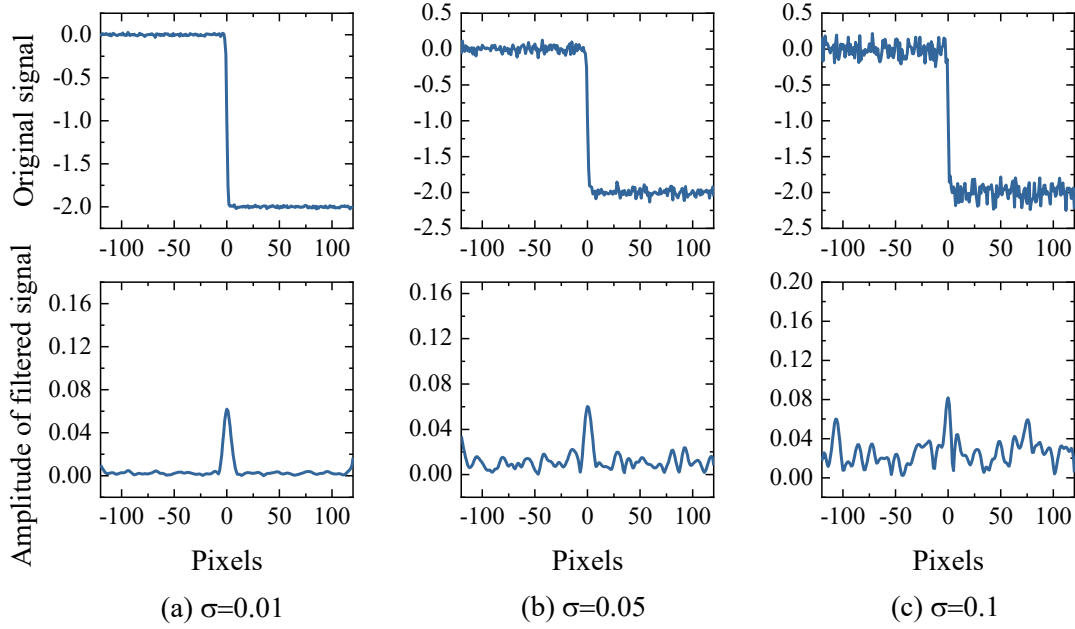


Figure 4.2 1D signal and the amplitude map of the filtered signal

4.2.2.2 Multi-frequency representation using the complex pyramid

Complex quadrature filters, such as the Log-Gabor filter, consist of the even and odd parts in the spatial domain. Both parts fluctuate as a sine function, whose peak value decays to zero at the ends of the filter. The central frequency of the filter controls the filter size, which affects the resolution of the response when convolving the signal with a filter (Miao et al., 2022; Y. Wang et al., 2023). Specifically, features extracted from the signal with a small-sized filter focus on the local variation of the signal, whereas a large filter considers the overall tendency within the range covered by the filter. Therefore, filters with different central frequencies can be used to reflect various features of the same signal in terms of amplitude maps. Figure 4.3 shows the effects of the central frequency on the amplitude of the responses to the step function added by the Gaussian noise with a standard deviation of 0.1 used in Figure 4.2. The decrease in the central frequency of the filter increases the filter size. When the central frequency is 0.5 Hz, the noises and the step function are

messed up in the amplitude map. The step function located around Pixel 0 is emphasized by the filter with a central frequency of 0.184 Hz, whereas the amplitude value of the noise is close to that of the step function. The amplitude map is dominated by the step function until the central frequency of the filter decreases to 0.068 Hz.

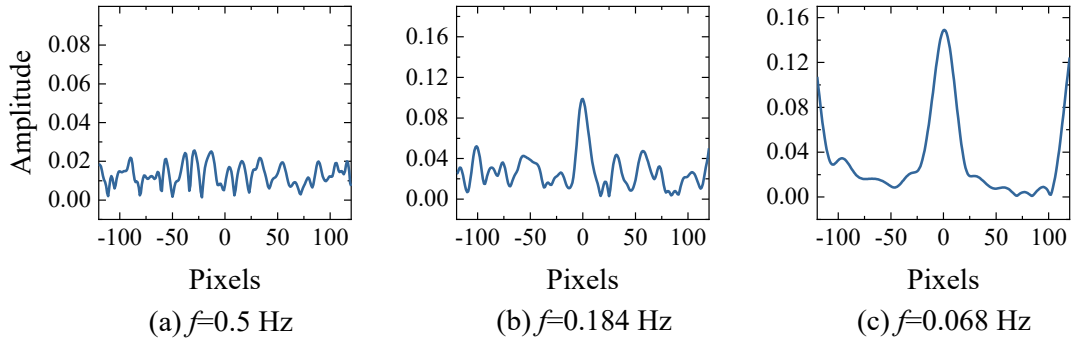


Figure 4.3 Responses by applying filters with different central frequencies

This example demonstrates that the content of interest (i.e., textures like the step function) is underestimated by the quadrature filter with a large central frequency in the amplitude map, in which local noises contribute most to the response. By contrast, the noise can be eliminated by the filter with a small central frequency. However, the reconstruction of the denoised signal within the full frequency band is unavailable with a single filter. For this reason, a class of quadrature filters is applied to the signal and generates a series of complex signals, referred to as a pyramid. As shown in Figure 4.4, the original image signal is filtered by a series of quadrature filters at different central frequencies, from level M to level 0. The multi-resolution amplitude map generated by the filter at the corresponding level is shown in the second row.

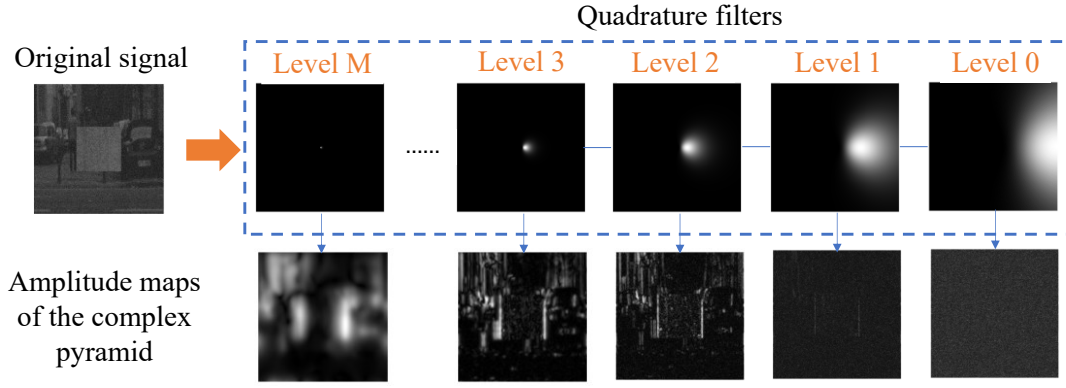


Figure 4.4 Original signal, complex quadrature filter bank, and amplitude map of the complex pyramid

Each level of the pyramid represents the complex signal obtained from the quadrature filter with a specific central frequency at that level, as follows:

$$\mathbf{R}_I^m(x, y) + j\mathbf{Z}_I^m(x, y) = \sum_{\Delta x=-w/2}^{w/2} \sum_{\Delta y=-w/2}^{w/2} \mathbf{I}(x + \Delta x, y + \Delta y) \mathbf{G}_S^m(\Delta x, \Delta y) \quad (4.8)$$

$$\mathbf{G}_S^m(x, y) = \mathbf{G}_{\text{even}}^m(x, y) + j\mathbf{G}_{\text{odd}}^m(x, y) \quad (4.9)$$

where $\mathbf{R}_I^m(x, y)$ — Real part of the complex signal at level m ;
 $\mathbf{Z}_I^m(x, y)$ — Imaginary part of the complex signal at level m ;
 $\mathbf{G}_{\text{even}}^m(x, y)$ — Real part of the complex filter (even function) at level m ;
 $\mathbf{G}_{\text{odd}}^m(x, y)$ — Imaginary part of the complex filter (odd function) at level m .

Each level of the pyramid is denoised by removing the pixel with an amplitude value smaller than the threshold, as follows:

$$\hat{\mathbf{R}}_I^m(x, y) = \begin{cases} 0 & \sqrt{\mathbf{R}_I^m(x, y)^2 + \mathbf{Z}_I^m(x, y)^2} \leq T_m \\ \mathbf{R}_I^m(x, y) & \sqrt{\mathbf{R}_I^m(x, y)^2 + \mathbf{Z}_I^m(x, y)^2} > T_m \end{cases} \quad (4.10)$$

where $\hat{\mathbf{R}}_I^m(x, y)$ — Real part of the denoised complex signal at level m ;
 T_m — Noise threshold at level m .

After denoising, the complex pyramid is shrunk to reconstruct a denoised signal by retaining the real part as follows:

$$\hat{I}(x, y) = \sum_{m=0}^M \hat{R}_l^m(x, y) \quad (4.11)$$

where $\hat{I}(x, y)$ ——Denoised image signal.

Previous analysis indicates that the content of interest and noise in a signal can be distinguished by the amplitude map. The phase of the content of interest is preserved by discarding the noises with small amplitudes in the complex pyramid. However, the threshold depends on the feature of the noises and the responses to the filter used at the specific level, causing difficulty in quantifying the threshold at each level of the complex pyramid. For this reason, an adaptive threshold considering multiple frames is proposed in the next subsection.

4.2.2.3 Denoising based on the amplitude instability among multi-frames

The estimation of the noise at each level of the complex pyramid requires prior knowledge of the probability density function and the spectral distribution of the noise (Kovesi, 1999), which are, however, unknown in practice. Fortunately, multi-frames are used in the phase-based motion estimation. The noises in different frames obey the independent-identical distribution and fluctuate at each pixel over time, whereas the content of interest remains the same in different frames. For this reason, random noises cause instability of the amplitude maps among different frames. Given that the phase-based motion estimation method only focuses on the content of interest, this study proposes to preserve the phase with stable amplitude instead of estimating the noise in the complex pyramid. The instability of amplitude maps among multi-frames at a specific level is evaluated as follows:

$$S_m = \sqrt{\frac{\sum_{W \times H} (A_m^1 - A_m^2)^2}{W \times H \times \text{median}(A_m^1)^2}} \quad (4.12)$$

where S_m ——Amplitude instability indicator of the complex signal at Level m ;
 A_m^1 ——Level m amplitude of the reference frame;
 A_m^2 ——Level m amplitude of the current frame;

W —— Width of the ROI;
 H —— Height of the ROI;
 $\text{median}(\cdot)$ —— Median value operator.

The amplitude instability indicator works as the standard deviation of the amplitude difference between two frames and reaches the maximum value at level 0, where the response only considers the extremely local fluctuation (i.e., the noise). At other levels, the central frequency of the filter decreases (i.e., the filter size increases) in the spatial domain. Therefore, more pixels belonging to the content of interest are covered by the filter, leading to a smaller instability indicator. In the denominator, the median operation is used to introduce the overall amplitude value level of the image, because the amplitude value of the structure edge is much larger than that of other pixels. If the structure edge is counted, the amplitude instability will be underestimated. On the other hand, in the numerator, not only the noise but also the structure edge movement can induce amplitude value variation. This refers to a main limitation of the proposed method, which is that only structures with small displacements (generally smaller than 1 pixel) can be denoised. With small displacements, only the amplitude value fluctuation of the pixels belonging to the structure edge is superimposed in the numerator. Most of the pixels are homogeneous or static textures whose amplitude values only respond to the noise fluctuations.

Thresholds should be estimated at each level of the complex pyramid, in which the amplitude value at the spatial position of the interested content can be directly calculated. The instability indicator is used to determine to what extent the amplitude value is smaller than that of the interested content should be regarded as noise. The responses at level 0 of the complex pyramid are regarded as pure noise because the corresponding central frequency of the filter is the upper bound of the frequency spectrum, and cannot extract the content of interest. The threshold of the noise at other levels can be estimated by comparing the instability indicators between the level concerned and level 0:

$$T_m = \left[1 + \ln \left(\frac{S_m}{S_0} \right) \right] \times a_m \quad (4.13)$$

82

where T_m ——Noise threshold at level m ;
 a_m ——Amplitude value of the content of interest at level m .

The threshold reaches the maximum value at level 0 to discard all pixels and decreases with the decrease in the instability indicator. At each level, the central frequency and the filter size in the frequency domain decrease on a natural exponential scale compared with those at the previous level. The frequency band covered by the filter at the higher level is narrower and generates smaller responses in the amplitude map and a smaller instability indicator. Therefore, the stability indicators at different levels are compared in the natural logarithm scale to avoid underestimating the noises at higher levels of the pyramid.

4.2.3 Multi-frame phase-preserving denoising (MPD)

The procedure of the MPD method (Wang et al., 2024a) is shown in Figure 4.5.

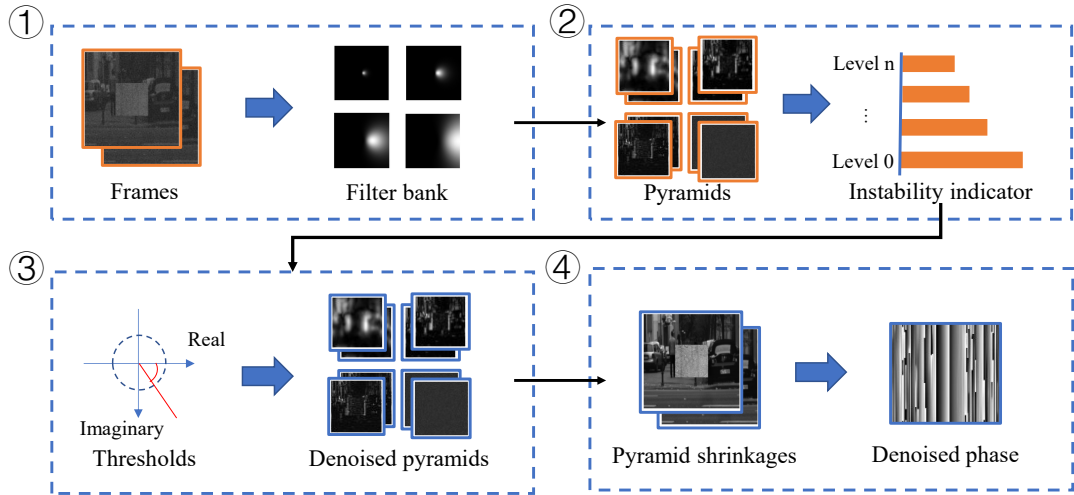


Figure 4.5 Procedure of the MPD

First, two frames used to calculate the motion are transformed by the filter bank (i.e., a series of complex quadrature filters at different directions and scales) to construct complex pyramids. Second, amplitude maps of pyramids between two frames are

compared to generate the instability indicator at each level. Third, thresholds based on the instability indicator are obtained from Equation (4.13) and applied to the complex pyramids to preserve the phase with amplitudes above the threshold at each level. Finally, the denoised pyramids are shrunk using Equation (4.11) to obtain the denoised frames without distorting the phase of interest. Phase-based motion estimation methods (Y. Wang et al., 2023) can then be applied to calculate the denoised phase and motion sequences.

4.3 Numerical examples

In this section, the phase-preserving and motion estimation performance of the MPD method is demonstrated using a 2D numerical example with noise. The dataset in the numerical example is generated by the render engine Blender.

4.3.1 Data generation

Figure 4.6 shows one frame of the dataset with a resolution of 1024×1024 on the left. The content of interest is a concrete block located at the center of the frame. The concrete block can randomly move relative to the background in each frame along the horizontal direction with an amplitude of less than 1 pixel. Three types of noise, namely, Gaussian noise, speckle noise, and local variant noise are applied to the rendered image. The speckle noise is constructed by multiplying the Gaussian noise with the pixel value, whereas the local variant noise obeys different Gaussian distributions at each pixel in the image. The speckle, Gaussian, and location-variant noises are separately applied to the generated 50 images (150 images in total). Figure 4.7 shows the clean image and 50 Gaussian noisy images to be added to the clean one.

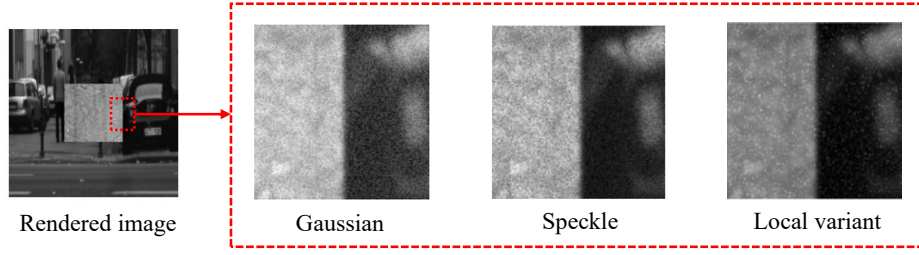


Figure 4.6 Rendered image and different noisy images

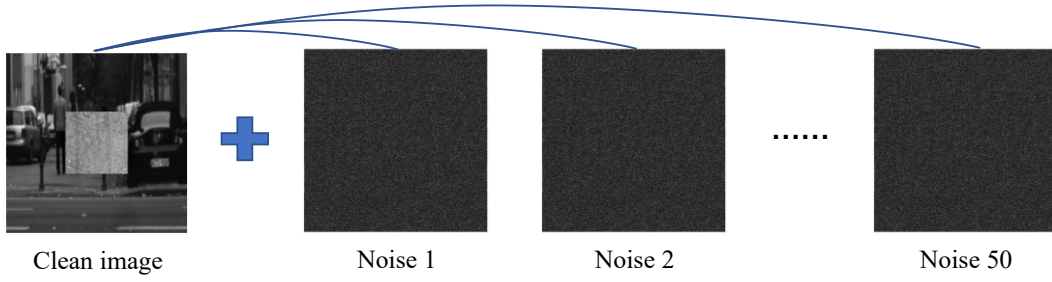


Figure 4.7 Clean and 50 noisy images (Gaussian)

4.3.2 Phase denoising performance

For comparison, all images are denoised by five methods, namely, the proposed MPD methods with six scales and six orientations; the wavelet method (Chang et al., 2000) with Bayes shrink; the non-local mean method (Buades et al., 2005) with a patch size of 5 and a patch distance of 6; the phase method (Kovesi, 1999) with the same configuration with MPD except for the threshold. The state-of-the-art deep learning model (Zhang et al., 2017) was trained by 400 noisy images with various zero-mean Gaussian noises, wherein the standard deviation of the Gaussian noise varies from 0 to 55 when the image intensity values range from 0 to 255. Figure 4.8 shows 64×64 denoised pixels around the measurement point (i.e., the right edge of the concrete block) in the case of the Gaussian noise. The image denoised by the MPD retains the edge to estimate the motion, whereas the noises and unimportant details are removed. Except for the aliasing occurring in the case of the wavelet method, all other denoising methods, especially the deep learning model, achieve satisfactory visual qualities. After denoising the images, the phase of the measurement point is analyzed. For the phase-based motion estimation, evaluating the phase variation between denoised images is beneficial.

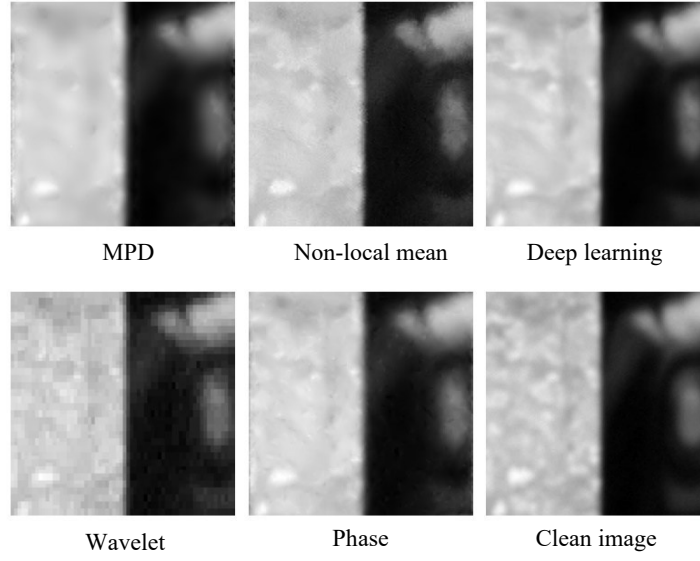


Figure 4.8 Images denoised by different methods (Gaussian noise)

The phase variation due to the noise is studied. The spatial frequency components will be changed due to the denoiser used. However, the traditional Gabor filter focuses on a specific frequency band, introducing manual effects when evaluating the denoising performance of different denoisers. Therefore, the Log-Gabor filter is applied to the denoised images to generate the phase. The Log-Gabor filter has a central frequency of 0.2 Hz and a bandwidth of 2.2 octaves to consider most of the frequency band.

In each noisy image, the noise distribution remains the same, whereas the noise value in each image is independent. The variation of the phase at each pixel between two denoised frames (t_1 and t_2) only comes from the noise and is calculated as follows:

$$\Delta\phi(x, y) = |\phi(x, y, t_1) - \phi(x, y, t_2)| \quad (4.14)$$

where $\Delta\phi(x, y)$ — Noise-induced phase variations.

A perfect denoising method minimizes the difference between $\phi(x, y, t_1)$ and $\phi(x, y, t_2)$, leading to a zero $\Delta\phi(x, y)$. The phase variation between two randomly selected denoised images is demonstrated in Figure 4.9, where phase variations along the measured edge of

the concrete block are exhibited under three types of noises. The result shows that the MPD method has a lower phase variation than other methods at almost all pixels. The phase method performs better than other methods except for the MPD because the images are denoised in the complex pyramid. This example shows that the complex pyramid considering multiple frames effectively eliminates the phase variation due to the noise.

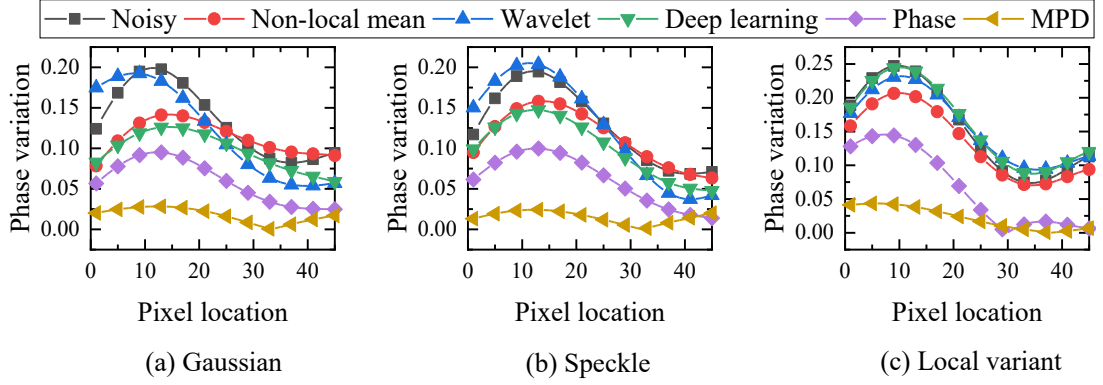


Figure 4.9 Phase variation maps for different noises

The spatial gradient of the phase is engaged in estimating the motion. Therefore, the variations of phase gradients are compared in Figure 4.10. The MPD method leads to lower gradient variations than other methods, especially under the local variant noise whose distribution is complicated. The stability of the phase gradient shows that the phase can be preserved not only at the pixel to be measured but also at the surrounding pixels, leading to a constant linear relationship between the phase and the displacement.

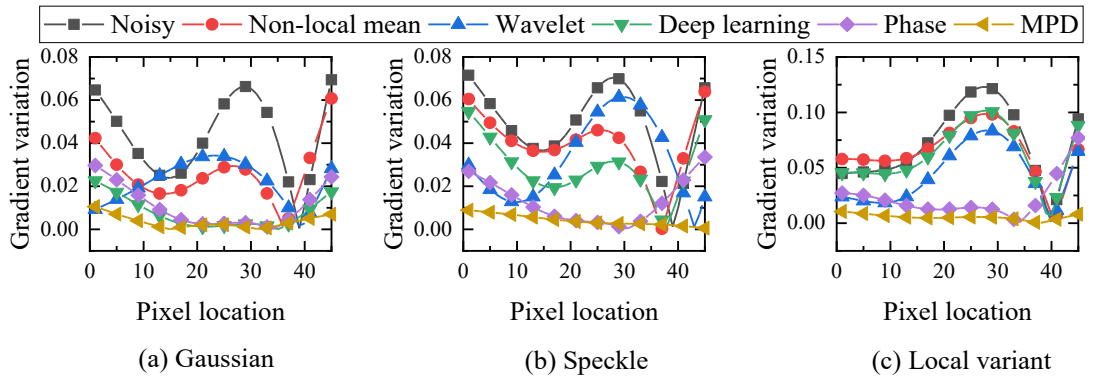


Figure 4.10 Gradient variation maps for different noises

The phase-preserving performance is quantified by the following indicator over all pixels:

$$\delta = \sqrt{\sum_{(x,y) \in C} \frac{\Delta \phi^2(x,y)}{\nabla \phi^2(x,y)WH}} \quad (4.15)$$

Where δ —Phase-preserving indicator, smaller values mean more stable phases;
 C —A cluster of pixels belonging to the measurement region;
 W —Width of the measurement region;
 H —Height of the measurement region.

The phase-preserving performance of each denoising method is shown in Figure 4.11. The proposed MPD method results in a δ value lower than half of the noisy data and outperforms other methods under all types of noises. The wavelet, non-local mean, and deep learning methods do not consider the phase and yield unsatisfactory results. The phase method performs slightly worse than the proposed method in cases of the Gaussian and speckle noises. However, the phase method requires the parameter of the noise distribution in the complex pyramid and does not work effectively under the local variant noise, in which each pixel's noise obeys different distributions. This example demonstrates that the proposed MPD can effectively tackle noises with different distributions.

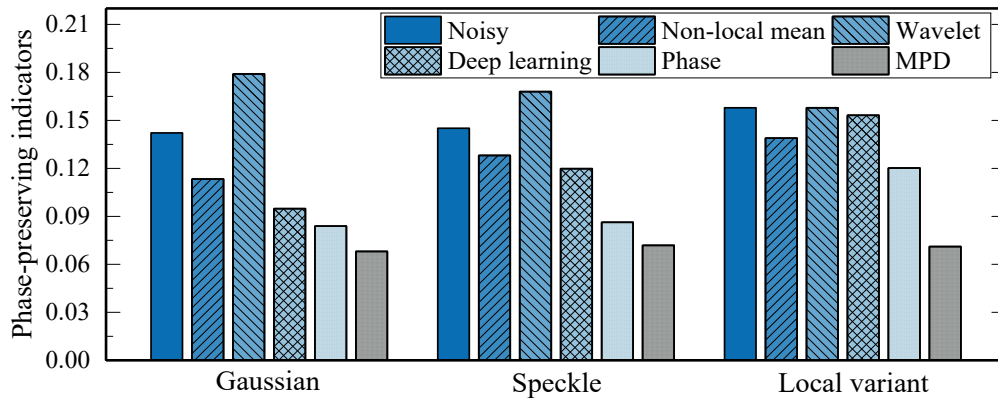


Figure 4.11 Phase preservation performance under different noises

Images with different intensities of the noise are used to evaluate the performance of different denoisers, in which the local variant distribution is selected to consider the complex situation whose noises are unpredictable. The PSNR is used to quantify the noise level. A lower PSNR indicates stronger noise. The phase-preserving performance of the denoising methods is compared in Figure 4.12. The MPD method consistently holds the lowest δ value, which does not exceed 0.12 even under the most significant noise condition with a PSNR of approximately 27. The phase method has twice δ values of the MPD, but, overall, it performs better than other denoisers that do not consider the phase information. The performance of the non-local mean method is similar to that of the phase method, whereas the deep learning and wavelet methods perform ineffectively in all cases.

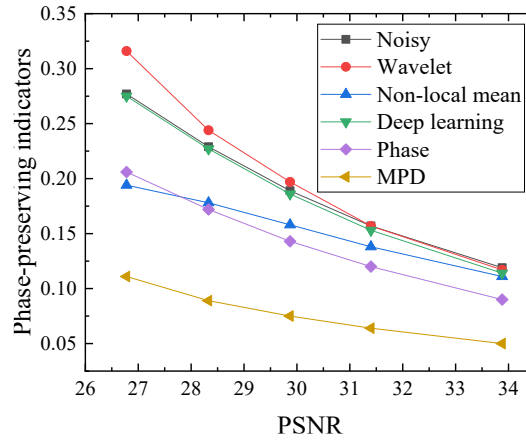


Figure 4.12 Phase preservation performance of denoising methods under different noise levels

The thresholds of the MPD at the first three levels of the pyramid are further extracted and compared with those of the phase method in Figure 4.13. When the PSNR is 33.88, the thresholds of the MPD at levels 0 and 1 exclude most pixels at those levels. This finding is consistent with that shown in Figs. 3 and 4, in which large central frequencies used in lower levels hardly extract the amplitude of the interested content. The phase method yields smaller thresholds because the noise cannot be estimated in the spatial domain. On the other hand, given a larger noise, for example, PSNR of 26.78, the threshold at level 2 generated by the phase method changes slightly, whereas the threshold

of the MPD at this level significantly increases to approximately 0.4. The characteristics of the thresholds show that the proposed MPD method focuses on retaining the content of interest and is sensitive to the noise level.

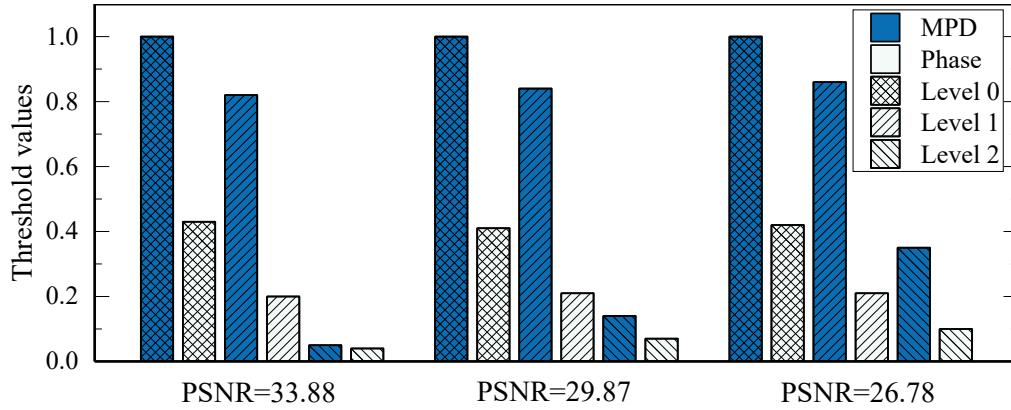


Figure 4.13 Thresholds at the first three levels of the pyramid

4.3.3 Displacement denoising performance

The displacement of the concrete block is estimated. The GT of the motion is accurately extracted from the rendering software. Different from the previous example, the denoising of the motion should distinguish the phase variation caused by the noise apart from that caused by the moving content. Figure 4.14 shows the displacement estimated from 50 images denoised by different methods under the local variant noise. The peak-to-peak displacement is only 0.5 pixels. All denoising methods retain the overall tendency of the motion sequence, wherein some methods cause significant errors at the local peaks of the fluctuations. The MPD method yields the most accurate displacement. A close-up view between Frames 35 and 38 shows that the wavelet, the non-local mean, and the deep learning methods cause the amplitude almost twice that of the GT. The phase and the MPD methods working in the complex pyramid show more accurate results, and the MPD method is more robust due to the threshold strategy considering multiple frames.

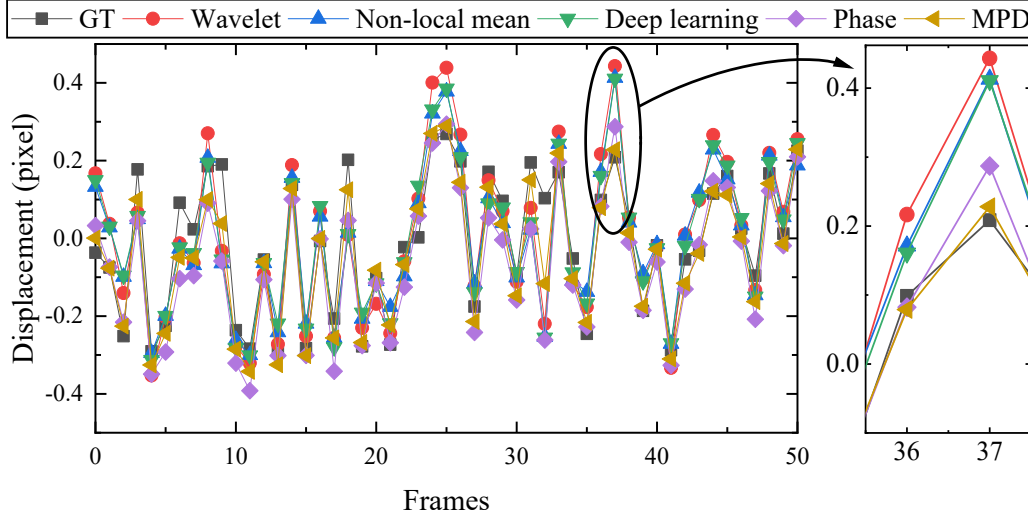


Figure 4.14 Displacement estimated from images denoised by different methods

The relative error of displacement is calculated to quantify the performance of the denoising methods, as follows:

$$\varepsilon = \frac{\|\Delta \mathbf{u}\|}{\|\mathbf{u}\|} \quad (4.16)$$

where ε — Relative error of displacement;

$\Delta \mathbf{u}$ — Deviation between the estimated displacement and GT;

$\mathbf{u}$ — GT of the displacement;

$\|\cdot\|$ — Euclidean norm.

Figure 4.15 shows the relative errors of the results by different denoising methods. The proposed MPD method outperforms other denoisers in all cases. In the Gaussian and speckle noises, the phase method performs effectively because the parameters of the noise distribution can be roughly estimated in the spatial domain, whereas the phase method cannot tackle complex noises, such as the local variant noise. The wavelet method does not work and yields a result worse than or close to that of the noisy data, while the non-local mean method only mitigates the adverse effect of the Gaussian noise. The deep learning model trained by the specific dataset fails when the training set does not contain the noise type. This example demonstrates that the MPD method can suppress the noise and retain the phase of the interested content regardless of the noise type involved.

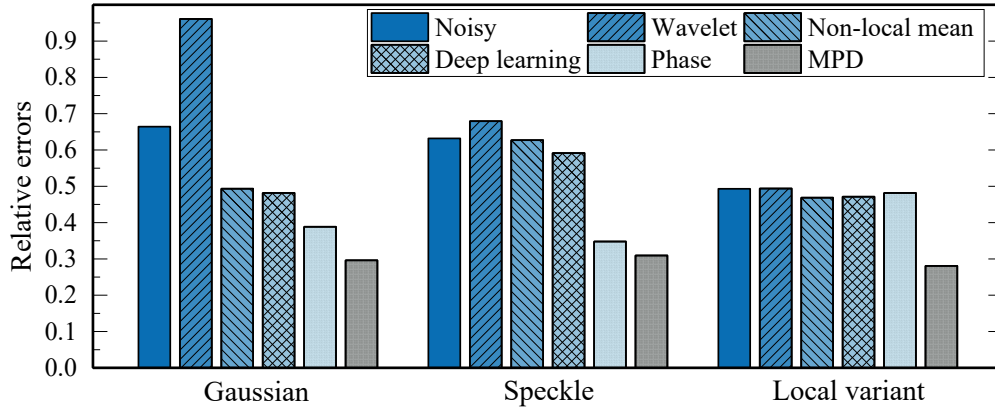


Figure 4.15 Relative errors of estimated motions

4.4 Experimental studies

The MPD method is applied to a steel cantilever beam tested in the laboratory under different lighting and noise conditions.

4.4.1 Factors affecting the noise and experimental setup

This section presents experimental studies on a five-story frame structure to evaluate the performance improvement of the phase-based method using denoising algorithms under different camera settings. Accordingly, parameters affecting noise intensity in experimental conditions must be determined.

In video acquisition, image noise distribution consists of pattern noise and random noise (B. Li et al., 2022; Lukas et al., 2006) generated during the sensor's photoelectric conversion process. This is collectively influenced by exposure time (light intake), recording temperature, and bit depth. Among these, light intake is strongly correlated with shot noise: higher light intake increases the absolute intensity of shot noise but reduces its relative strength compared to valid image information, thereby improving the SNR.

Recording temperature affects circuit temperature, leading to thermal noise due to electron motion. Bit depth refers to the resolution of pixel intensity values (dynamic range), with quantization noise arising below the resolution limit. Among these three noise sources, shot noise contributes the majority of noise intensity and is the primary target for adjusting the SNR during acquisition.

Generally, ISO sensitivity, aperture, and shutter speed are considered the main adjustable parameters affecting shot noise in video recording. Aperture and shutter speed influence total light intake in the spatial domain (light intake area) and temporal domain (light intake duration), respectively. Regarding aperture, as analyzed in Chapter 3, phase-based displacement estimation requires the use of the largest possible aperture to effectively control the depth of field, thereby mitigating adverse effects from out-of-the-plane extra texture. For shutter speed in video recording, the sampling uncertainty principle dictates that it should generally be set to twice the video frame rate and at least twice the vibration frequency of the target structure. Therefore, this study does not treat aperture and shutter speed as controllable parameters for video noise.

ISO sensitivity is widely recognized as a key parameter controlling the perceived strength of shot noise in images and videos. Using a higher ISO essentially increases the gain of the PGA, enhancing photoelectric sensitivity and amplifying the number of output electrons from valid signals to suppress noise transmitted to the analog-to-digital converter. Assuming shot noise caused by photon randomness follows a Poisson distribution, the SNR can be expressed as the ratio of the mean to the standard deviation of the number of photons converted to electrons at a given pixel:

$$SNR_{\text{shotnoise}} = \frac{\kappa}{\sqrt{\kappa}} \sqrt{\alpha_{\text{ISO}}} = \sqrt{\kappa \alpha_{\text{ISO}}} \quad (4.17)$$

where $SNR_{\text{shotnoise}}$ — SNR of the shot noise;

α_{ISO} — Gain of the PGA;

κ — Initial light intake.

The above equation indicates that, under constant light intake conditions, increasing the ISO sensitivity amplifies the electron count, leading to a growth in the SNR of the resulting digital image signal at a rate proportional to the square root of the amplification factor. It should be noted that in photography and videography, low ISO settings are often used to suppress shot noise. This requires coordinated adjustment of aperture and shutter speed to ensure higher initial light intake in Equation (4.17), thereby achieving a higher initial SNR and enabling properly exposed images or videos without increasing ISO. When the initial light intake is fixed, ISO sensitivity is positively correlated with SNR.

Therefore, this study maintains constant aperture and shutter speed while varying the ISO sensitivity to capture videos with different levels of shot noise, thereby investigating the impact of the proposed MPD method on the accuracy of phase-based displacement estimation from digital images.

In the experimental setup, indoor lighting intensity was reduced to achieve a lower initial SNR, as illustrated in Figure 4.16. The video was captured using a ZCAM S6G camera with a Canon EF24-105 mm lens, at a resolution of 3840×2160 pixels, a frame rate of 29.97 Hz, a focal length of 24 mm, an aperture of F/4, and a shutter speed of $1/60$ s (twice the frame rate). The camera was positioned approximately 6 m from the target structure. The five-story frame measures $1.8 \text{ m} \times 0.5 \text{ m} \times 0.5 \text{ m}$, and the conversion factor between image coordinates and physical coordinates is calibrated as 0.5 mm/pixel. The ISO sensitivity was adjusted across test scenarios to amplify the converted electrical signal, thereby influencing the SNR of the digital images of the target structure.

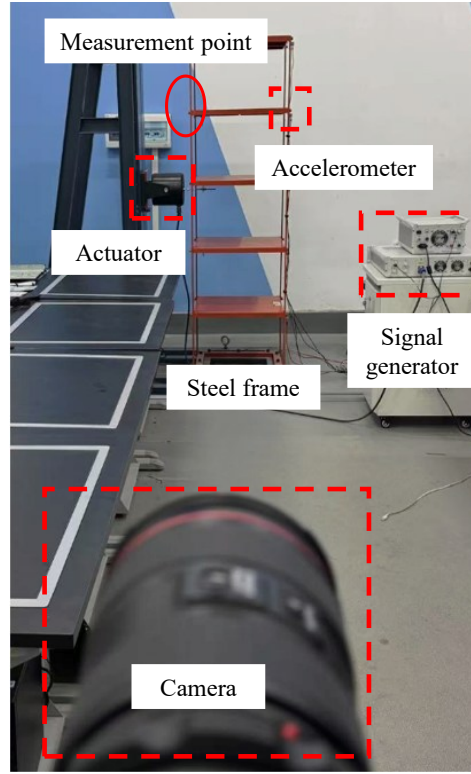


Figure 4.16 Experimental setup

4.4.2 Displacement denoising under different ISO levels

By setting the camera ISO to 800, 1600, and 3200, the displacement denoising performance of the proposed method is evaluated under different SNRs. In all ISO scenarios, the actuator excited the target structure at 7 Hz with consistent amplitude. Displacement derived from integrated acceleration data served as the benchmark. The original videos are denoised using the proposed MPD method, wavelet-based denoising, non-local means denoising, deep learning-based denoising, and phase-based denoising. Displacement at the fourth story is estimated by the proposed LGME method.

The relative errors of the estimated displacement at ISO 800 are compared in Figure 4.17. Under insufficient light intake and ISO 800, the low SNR of the digital images leads to a relative error of 0.31 in the directly estimated displacement. Wavelet-based denoising, which does not account for denoising in the complex domain of the digital image, perturbs

the phase and increases the relative error compared with the no-denoising case. Both non-local means and deep learning-based denoising provide some improvement, but the relative errors remain above 0.2. Phase-based denoising performs better than the other methods. However, its effectiveness is still limited due to the reliance on a Gaussian noise distribution assumption for threshold determination. In contrast, the proposed MPD method directly denoises the phase of structural features in the complex domain and estimates the noise threshold based on the instability of magnitude values, reducing the relative displacement error to approximately 0.1.

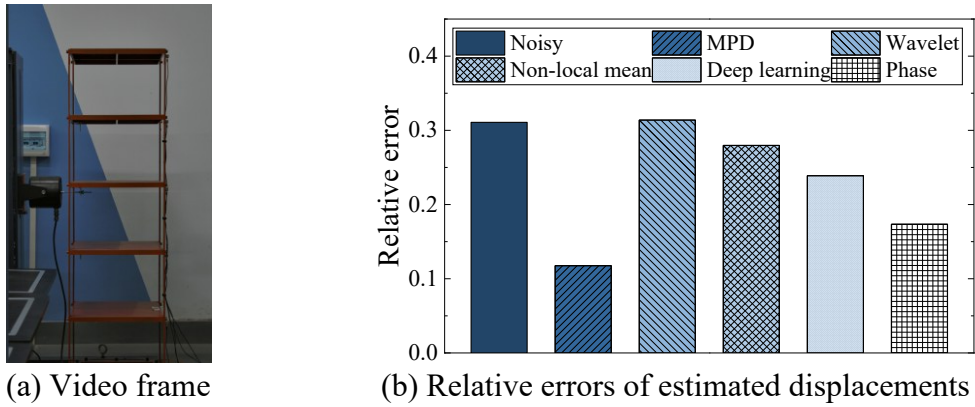


Figure 4.17 Relative errors of estimated displacements under ISO 800

Figure 4.18 displays the estimated displacement after denoising by various methods at ISO 800. Wavelet-based denoising perturbs the phase of the target structure, leading to significant fluctuations in the displacement results. The maximum absolute error reaches 0.12 mm at 0.14 s, exceeding the actual displacement amplitude. Deep learning and non-local means methods perform better than wavelet denoising, with displacement results generally aligning with the benchmark after denoising. However, localized peak deviations can still reach approximately 0.1 mm from the true values. Phase-based denoising outperforms the above methods but still fails to accurately estimate the noise threshold in the complex domain, resulting in displacements deviating from the benchmark by around 0.06 mm. In contrast, the proposed MPD method produces displacement results that closely match the benchmark amplitude of 0.1 mm, with the maximum absolute error controlled within about 0.04 mm, significantly outperforming

other denoising approaches. Therefore, in low-light environments, the application of the proposed MPD method effectively enhances displacement estimation accuracy.

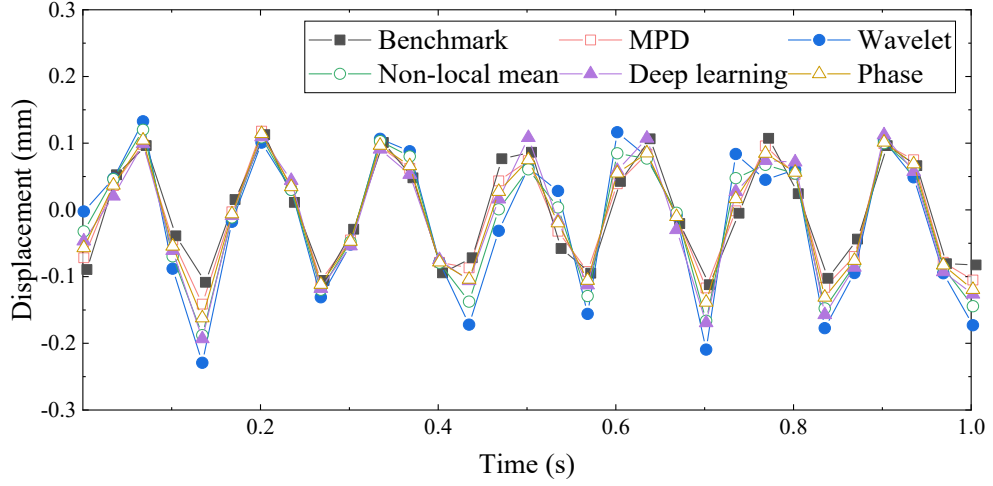
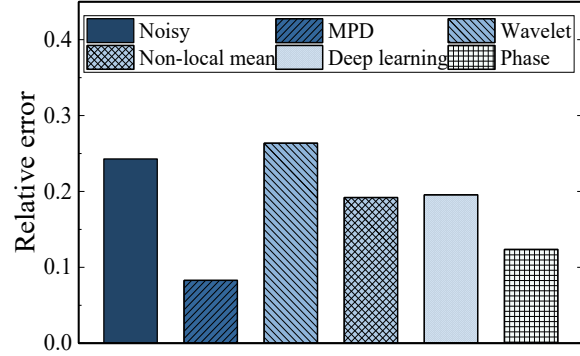


Figure 4.18 Estimated displacements from videos denoised by different methods under ISO 800

With the light intake held constant, increasing the ISO from 800 to 1600 results in the relative displacement errors shown in Figure 4.19. As the ISO amplifies the photoelectric signal, the SNR of the video images improves according to Equation (4.17). Additionally, the higher image brightness enhances the contrast of the geometric features of the target structure, increasing the phase gradient value and reducing its susceptibility to noise disturbance. Consequently, without denoising, the relative error of the estimated displacement is approximately 0.24, which is better than the 0.30 observed at ISO 800. In this scenario, wavelet-based denoising still fails to improve displacement estimation. Non-local means and deep learning-based denoising show limited effectiveness, performing worse than the phase-based method, which denoises directly in the complex domain. However, the proposed MPD method, which determines the noise threshold using the complex-domain instability indicator, outperforms all other denoising approaches, reducing the relative displacement error to around 0.08.



(a) Video frame



(b) Relative errors of estimated displacements

Figure 4.19 Relative errors of estimated displacements under ISO 1600

The estimated displacements at ISO 1600 are shown in Figure 4.20. The level of noise-induced fluctuation in the displacement results is lower compared to the case at ISO 800. However, localized deviations remain noticeable. Wavelet denoising, non-local means, and deep learning-based methods still exhibit deviations ranging from 0.5 mm to 0.1 mm at time points such as 0.24 s and 0.95 s. In contrast, the displacement results processed by the proposed MPD method closely align with the benchmark.

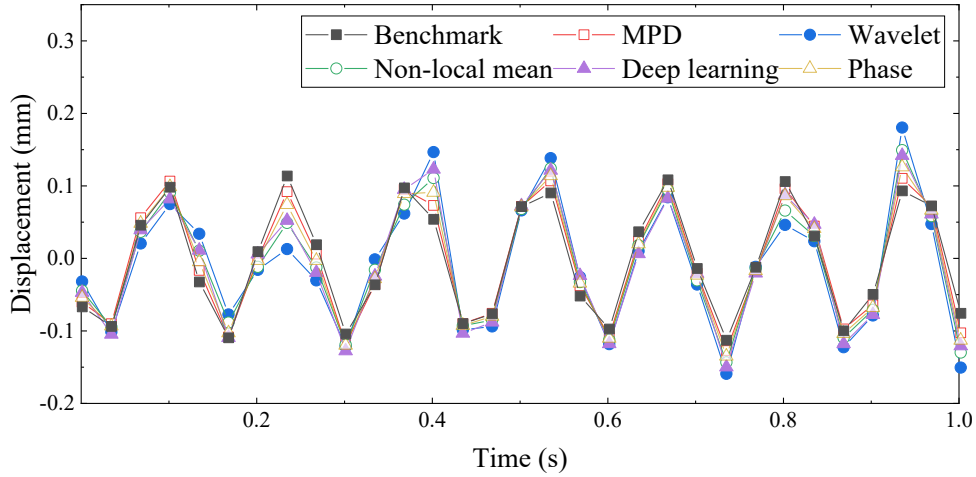


Figure 4.20 Estimated displacements from videos denoised by different methods under ISO 1600

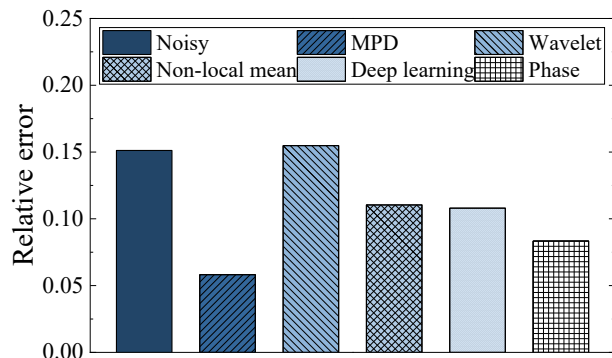
Further increasing the ISO from 1600 to 3200 enhances the contrast of the geometric features of the target structure, leading to a larger spatial phase gradient and an improved

SNR of the structural features relative to the noise. As shown in Figure 4.21, the higher ISO improves the SNR of the original video signal, reducing the relative error of the estimated displacement from 0.30 at ISO 800 to 0.15 without denoising. However, the error remains significant, necessitating further denoising.

Among the various denoising methods applied, wavelet-based denoising shows no substantial improvement. Non-local means and deep learning-based methods reduce the relative displacement error to approximately 0.11. Phase-based denoising, which operates in the complex domain, lowers the relative error to 0.08. In contrast, the proposed MPD method achieves a relative displacement error of only 0.05 under low-light conditions.



(a) Video frame



(b) Relative errors of estimated displacements

Figure 4.21 Relative errors of estimated displacements under ISO 3200

Figure 4.22 shows the displacement denoised by each denoising method. It can be observed that the wavelet-based method causes the peak displacement to exceed the benchmark by 0.05 mm (equivalent to half of the displacement amplitude). The results from other denoising methods also exhibit varying degrees of noise-induced fluctuations. Among them, the proposed MPD method yields displacement results that are closer to the benchmark than those of the other methods.

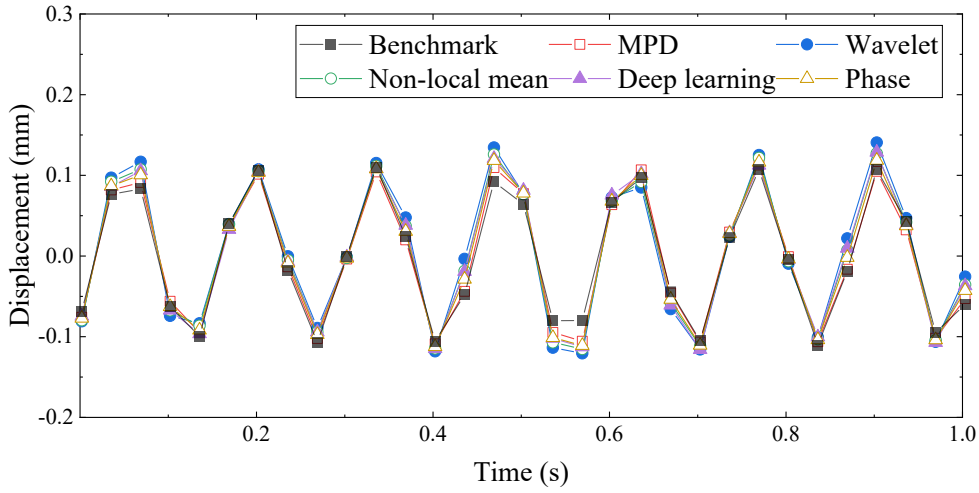


Figure 4.22 Estimated displacements from videos denoised by different methods under ISO 3200

4.5 Summary

A novel phase-preserving denoising method adopting the complex pyramid is proposed to enhance the accuracy of the phase-based motion estimation, in which the responses of the quadrature filter at each level of the pyramid are used to distinguish the noise apart from the content of interest in the complex domain. As the motion is estimated from the image sequence, an instability indicator among multiple frames is designed to determine the threshold of the noise in the complex pyramid with the unknown noise distribution.

A numerical example is designed to compare the performance of the proposed MPD and various image-denoising methods under different noise types. The results show that the classical denoising and the deep learning methods can generate denoised images with good visual qualities. However, the phase-preserving performance is unsatisfactory even under the typical Gaussian noise. The MPD method with the threshold considering the instability among multiple frames preserves the phase of interest and removes the noise, leading to an accurate phase-preserving and motion estimation performance under different noise types. An experimental study on a steel frame with different noise conditions has also been conducted. The results show that the proposed method

effectively decreases the relative error of the estimated displacement and alleviates the local fluctuations. In all cases, the proposed MPD method outperforms the classical and deep learning methods.

CHAPTER 5

FULL-FIELD DISPLACEMENT ESTIMATION BASED ON COMPLEX PYRAMID AND DISCRIMINATIVE CORRELATION

5.1 Introduction

The phase-based displacement estimation and denoising methods proposed in Chapters 3 and 4 of this study effectively treat each pixel in a digital image as a non-contact displacement sensor, thereby establishing a foundation for estimating full-field displacements of structures. However, practical full-field displacement estimation for large-scale structures requires further investigation into two key factors: adaptive estimation of displacements with varying amplitudes and adaptive estimation for multiple geometric types of targets.

First, adaptive estimation of displacements with varying amplitudes. In phase-based displacement estimation methods, the linear projection between phase and displacement holds only within a limited phase variation range (Kong et al., 2025; Miao et al., 2023). Although employing complex filters with larger spatial scales can compress the phase variation range and extend the linear relationship to larger displacement amplitudes, thereby increasing the measurable range under external excitations (e.g., typhoons, earthquakes, traffic loads) and differences in stiffness among structural components. This variability poses challenges for adaptively balancing displacement range and accuracy in phase-based methods.

Second, adaptive estimation for multiple types of targets. The phase extraction of structural features relies on linear geometric features (e.g., edges of bridge towers, beams, or columns) and requires that these linear features are not parallel to the direction of displacement. In real-world large-scale structures, where imaging conditions are difficult to control, effective linear features may not always be available. Thus, it is necessary to develop multi-type target displacement estimation methods capable of flexibly tracking diverse geometric features.

To address these challenges, this chapter aims to achieve adaptive full-field displacement estimation for large-scale structures under varying amplitudes and multi-type targets. Specifically, the phase extraction quality in multi-scale displacement amplitudes is analyzed and improved. A multi-resolution complex pyramid is constructed to observe structural displacements across different displacement amplitudes. A displacement amplitude tracking method based on the amplitude degradation index in the multi-resolution complex pyramid is proposed. Additionally, a linear phase extraction method based on phase spatial gradient divergence is developed, enabling adaptive full-field displacement estimation balancing both range and accuracy.

To handle multi-type geometric targets at different measurement points in real large-scale structures, discriminative correlation filters are introduced. By integrating continuous convolution operators, filter compression, and sample dimension reduction, an adaptive displacement estimation method for multiple target types in large-scale structures is established. Ultimately, this approach achieves simultaneous estimation of full-field displacements and mode shapes for real large-scale structures, accommodating both varying displacement amplitudes and multiple target geometries.

5.2 Multi-scale displacement estimation based on multi-resolution complex pyramid

5.2.1 Phase extraction of multi-scale displacements

The complex filter pairs with a spatial phase difference of $\pi/2$ are applied to extract the phase according to Chapter 3. The phase variation in the time domain and the spatial phase gradient are calculated to estimate the displacement:

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \frac{\phi(x_1, y_1, t_1) - \phi(x_1, y_1, t)}{\|\nabla \phi(x_1, y_1, t_1)\|^2} \nabla \phi(x_1, y_1, t_1) \quad (5.1)$$

$$\nabla \phi(x_1, y_1, t_1) = \begin{bmatrix} g_x & g_y \end{bmatrix}^T \quad (5.2)$$

where t_1 ——Reference time step;

ϕ ——Phase;

$\nabla \phi$ ——Spatial gradient of the phase;

g_x ——Horizontal gradient calculated by Sobel operator;

g_y ——Vertical gradient calculated by Sobel operator.

Equations (5.1) and (5.2) rely on two assumptions: first, phase variations are continuous in the time domain; second, the spatial phase gradient should be constant, projecting the phase and the displacement linearly.

However, the phase is computed using the arctangent function in Equation (5.1). When the displacement of the target structure is large, the phase may undergo abrupt jumps, necessitating the use of complex filters with larger spatial scales to compress the phase variation range. Although phase unwrapping (Collier and Dare, 2022; Shan et al., 2024) can convert the phase into a continuous variable, it does not guarantee the consistency of the spatial gradient of the phase, thereby adversely affecting the linear projection between phases and displacements. When the displacement of the target structure is small, the temporal variation of the phase at the pixel coordinates of the measurement point is also minor. In such cases, the linear relationship between phase and displacement is easily disturbed by complex environmental factors. Therefore, it is necessary to use the local optimization coefficient proposed in Chapter 3 to determine the filter parameters and extract the local phase at a smaller spatial scale. In summary, for full-field displacement estimation in large-scale structures, displacement amplitudes can range from sub-millimeter to decimeter levels, depending on structural stiffness and external excitation.

Constructing a single complex filter can hardly satisfy both the accuracy and amplitude requirements for displacement estimation.

To address multi-scale displacement estimation, the controlling factors of the spatial scale of the Log-Gabor complex filter in Equation (5.3) and its impact on the displacement estimation range are further investigated. Since the Log-Gabor complex filter lacks a theoretical expression in the spatial domain, its spatial scale is analyzed based on its frequency-domain representation. According to the uncertainty principle, a reduction in the frequency-domain scale of the filter corresponds to an increase in its spatial-domain scale. In the 1D case, the frequency-domain scale of the Log-Gabor complex filter is determined by the standard deviation and the center frequency. A higher center frequency directly amplifies the frequency components of the complex filter, thereby reducing its spatial-domain envelope:

$$\mathbf{LG}(f) = \exp \left\{ -\frac{[\ln(f/f_0)]^2}{2[\ln(\sigma_r/f_0)]^2} \right\} \quad (5.3)$$

where $\mathbf{LG}(f)$ ——1D Log-Gabor filter in the frequency domain;

f_0 ——Central frequency;

σ_r ——Standard deviation of the radial envelope in the logarithmic scale.

Therefore, the structure to be measured may exceed the envelope of the filter when a large motion happens. For example, a 1D step signal moving horizontally and the amplitude value of the signals filtered by a complex quadrature filter are shown in Figure 5.1. A higher amplitude value at the measurement point indicates that the filter emphasizes more on the structure, and vice versa. The measurement point is Pixel 0, at which the observed amplitude reaches the maximum when the 1D signal has no movement. However, the amplitude value at Pixel 0 decreases to 0.870 and 0.586 as the signal moves 1 and 2 pixels, respectively, indicating that the signal cannot be captured by the filter centered at Pixel 0 when a relatively large movement occurs.

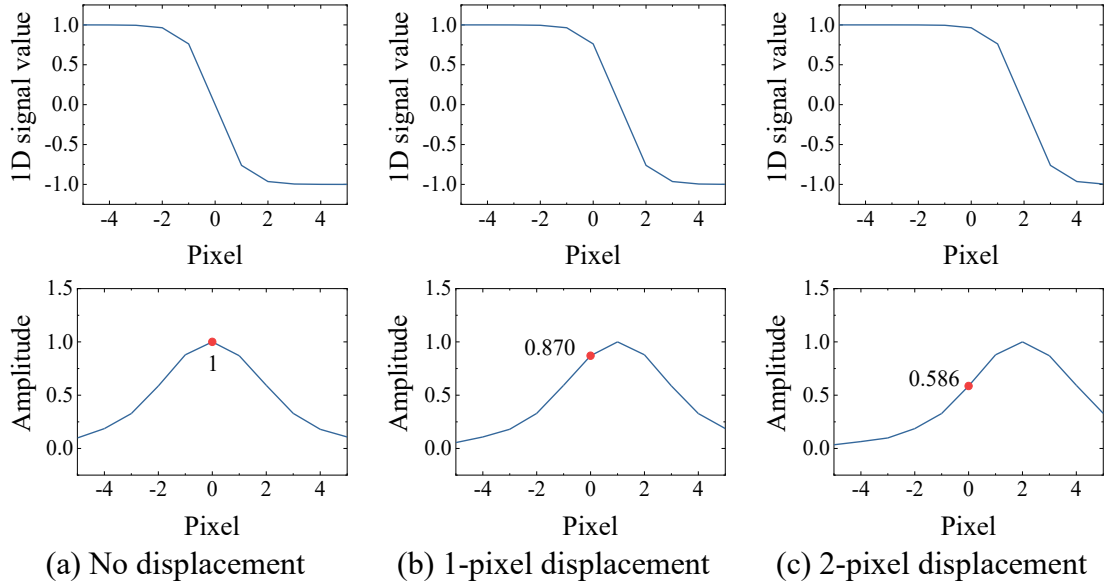


Figure 5.1 Moving 1D signals and corresponding amplitudes

An effective way to tackle the large motion is by employing a filter with a broader spatial size. Figure 5.2 shows the filtered results of the 1D signal with a movement of 2 pixels when a filter with different central frequencies is applied. The normalized amplitude values at Pixel 0 steadily increase from 0.586 to 0.831 when the central frequency decreases from $1/3$ to $1/12$ Hz. However, a filter of small central frequency indicates a large spatial size, adversely affecting the precision of the phase due to extra textures. This phenomenon is explained in the next subsection.

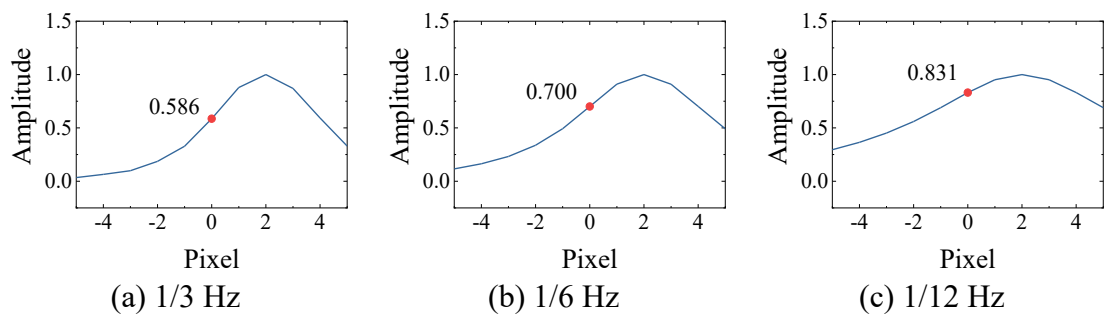


Figure 5.2 Amplitude values of the moving 1D signal filtered by different filters

However, filters with larger spatial dimensions make the extracted structural phase more susceptible to adverse effects from complex environments, thereby compromising estimation accuracy under small-displacement conditions. In real large-scale structures, displacement amplitudes vary over time due to differences in local stiffness and external excitations. Designing a single set of filter parameters cannot adequately address such varying displacement amplitudes.

Therefore, to estimate displacement responses across different amplitude scales in large-scale structures, a family of complex filters with varying spatial dimensions is designed. These filters are adaptively selected based on the changing displacement amplitudes, enabling adaptive estimation of multi-scale displacements.

5.2.2 Multi-resolution complex pyramid based on complex filters

Based on the preceding analysis, a single filter is insufficient to maintain the linear phase-displacement relationship across multi-scale displacements. Therefore, this study proposes constructing a multi-resolution complex pyramid, which incorporates a series of complex filter pairs with different center frequencies. Each level of the pyramid extracts displacements corresponding to a specific amplitude range of the structure. The Log-Gabor filter is adopted as the basis of the multi-resolution complex pyramid:

$$LG^m(f, \theta) = \exp \left\{ -\frac{[\ln(f/f_m)]^2}{2[\ln(\sigma_m/f_m)]^2} \right\} \exp \left[\frac{-(\theta - \theta_0)^2}{2\sigma_\theta^2} \right] \quad (5.4)$$

$$f_m = 2^{-m} f_0 \quad (5.5)$$

$$\sigma_m = 2^{-m} \sigma_r \quad (5.6)$$

where $LG^m(f, \theta)$ ——Log-Gabor filter at Level m ;

θ_0 ——Direction of the measured target;

σ_θ ——Standard deviation of the circumferential envelope;

f_0 ——Central frequency determined by Equation (3.18);

σ_r ——Standard deviation of the radial envelope;

f_m ——Central frequency at Level m ;

σ_m ——Standard deviation of the radial envelope at Level m .

The multi-resolution phases and amplitudes are extracted by the filters determined by Equation (4.4):

$$\phi_l^m(x, y, t) = \arctan \left[\frac{\mathbf{Z}_l^m(x, y, t)}{\mathbf{R}_l^m(x, y, t)} \right] \quad (5.7)$$

$$A_l^m(x, y, t) = \sqrt{\mathbf{Z}_l^m(x, y, t)^2 + \mathbf{R}_l^m(x, y, t)^2} \quad (5.8)$$

$$j\mathbf{Z}_l^m(x, y, t) + \mathbf{R}_l^m(x, y, t) = \text{IFFT} \left\{ \text{FFT} [\mathbf{I}(x, y, t)] \odot \mathbf{LG}^m(f, \theta) \right\} \quad (5.9)$$

where $\phi_l^m(x, y, t)$ ——Phase at Level m ;
 $A_l^m(x, y, t)$ ——Amplitude at Level m ;
 $\mathbf{I}(x, y, t)$ ——Image pixel value;
 j ——Imaginary unit;
 $\text{FFT}(\cdot)$ ——Fourier transform;
 $\text{IFFT}(\cdot)$ ——Inverse Fourier transform;
 $\odot$ ——Dot product;
 $\mathbf{Z}_l^m(x, y, t)$ ——Imaginary part at Level m ;
 $\mathbf{R}_l^m(x, y, t)$ ——Real part at Level m .

Equations (4.7) and (4.8) are the multi-resolution complex pyramid, embedding multi-scale geometrical features of the target. The higher level in the pyramid extracts a larger spatial size, deriving a larger measurement range. In contrast, the lower level in the pyramid extracts local information, which has the advantages of sensing subtle and local vibration of structures under complex environments as introduced in Chapter 3.

5.2.3 Multi-resolution complex pyramid-based displacement estimation

Two challenges must be addressed to adaptively extract displacements from the multi-resolution complex pyramid based on the actual vibration displacement amplitudes of large-scale structures. First, digital images of real large-scale structures contain massive

pixels, making it difficult to select pixels with high-quality phase information. Second, since the displacement amplitudes of large-scale structures are unknown before measurement, a method is needed to determine these amplitudes, enabling the selection of the appropriate pyramid level and extraction of corresponding phase information at each time step.

To address these challenges, this paper proposes two metrics: the divergence of the spatial phase gradient (the spatial Laplacian of the phase) and the amplitude degradation index. These metrics enable the adaptive extraction and integration of displacement time histories for large-scale structures from the multi-resolution complex pyramid.

5.2.3.1 Linear phase extraction based on the spatial Laplacian of the phase

The linear relationship between the phase and displacement is the basis of the phase-based motion estimation method. Although the structure to be measured is emphasized by the localization index in image signals, the phases of most pixels are not strictly linear in the spatial domain owing to the limited spatial support of the filter, leading to inaccurate motion estimation results. The phase of pixels must be analyzed to ensure the phase extracted from each level in the pyramid has good linearity.

The linearity of the phase regarding the displacement in the time domain cannot be directly calculated because the displacement is unknown. Fortunately, the variation of the spatial positions of a structure is recorded in terms of intensity change in the spatial domain. On this basis, the linearity of the phase can be evaluated by the phase change in the spatial domain, that is, the phase among adjacent pixels linearly changing along the pixel location indicates a linear relationship with the displacement. The linear change of the phase in the spatial domain is quantified by the Laplacian operator at each level in the pyramid:

$$\nabla^2 \phi^m = \text{div}(\nabla \phi^m) = \text{div} \left[\frac{\partial \phi^m}{\partial d_\theta}, \frac{\partial \phi^m}{\partial d_{\bar{\theta}}} \right] = \frac{\partial^2 \phi^m}{\partial d_\theta^2} + \frac{\partial^2 \phi^m}{\partial d_{\bar{\theta}}^2} \quad (5.10)$$

where $\nabla^2(\cdot)$ ——Laplacian operator;
 $\text{div}(\cdot)$ ——Divergence operator;
 d_θ ——Direction parallel with the target edge;
 $d_{\bar{\theta}}$ ——Direction orthogonal to the target edge;
 $\nabla \phi^m$ ——Spatial phase gradient at Level m .

Generally, the phase variation along the structural edge is negligible; thereby, the phase variation in the direction orthogonal to the target edge dominates the Laplacian value. In this case, Equation (5.10) can be simplified to a 1D condition:

$$\nabla^2 \phi^m = \frac{\partial^2 \phi^m}{\partial d_{\bar{\theta}}^2} \quad (5.11)$$

The smallest spatial Laplacian value of the phase means the linearity between the phase variation and the displacement of a pixel is the best among adjacent pixels. This value will be selected to extract the linear phase at this level.

5.2.3.2 Level selection based on the amplitude degradation index

Another problem is adaptively selecting the pyramid level according to the scale of the structural vibration. Multi-scale motions are contained in the complex pyramid. Lower levels such as Level 0 offer high resolution but have a narrow measurement range. Therefore, ambient-induced weak vibrations should be extracted from lower levels, whereas vehicle-induced large vibrations can be extracted by higher-level filters.

The vibration condition is preliminarily analyzed to match the suitable level in the complex pyramid. Based on the previous observation, the amplitude reaches a large value at the pixel representing the structure. In contrast, the amplitude decreases at a pixel from

which the structure to be measured is far. For this reason, an amplitude degradation index is designed to measure the scale of the vibration and select the pyramid level at each frame:

$$D_m = A_I^m(x_0, y_0, t) - \eta A_I^m(x_0, y_0, t_0) \quad (5.12)$$

$$\eta = \frac{\min_{x \in [x_0-2, x_0+2], y \in [y_0-2, y_0+2]} A_I^m(x, y, t_0)}{\max_{x \in [x_0-2, x_0+2], y \in [y_0-2, y_0+2]} A_I^m(x, y, t_0)} \quad (5.13)$$

where D_m —— Amplitude degradation index of Level m at time t ;

η —— A factor deciding to what extent the amplitude value variation is regarded as degradation in a pixel patch with a size of 5×5 .

A positive D_m indicates that the vibration does not exceed the measurement of Level m , whereas a negative D_m means the vibration is large and should be measured at the next level. At each time step, the level selection starts at Level 0 with the smallest measurement range. The motion is extracted at Level 0 if D_0 is positive. Otherwise, the motion estimation is conducted at the next level. The amplitude degradation index is repeatedly calculated until a positive degradation index occurs at a certain level. The motion signal is then extracted.

The procedure of the multi-scale displacement estimation using the multi-resolution complex pyramid (Wang et al., 2024b) is introduced in Figure 5.3. First, a series of Log-Gabor filters is designed using the localization index at a reference frame. Second, multi-scale filters are applied to the frame and generate a complex pyramid, in which multi-scale motions are contained at different levels of the pyramid. Third, linear-phase pixel and pyramid levels are adaptively selected using the spatial Laplacian of the phase and the amplitude degradation index, respectively. Finally, the full-field motion is estimated by repeating the above steps over the frame sequence.

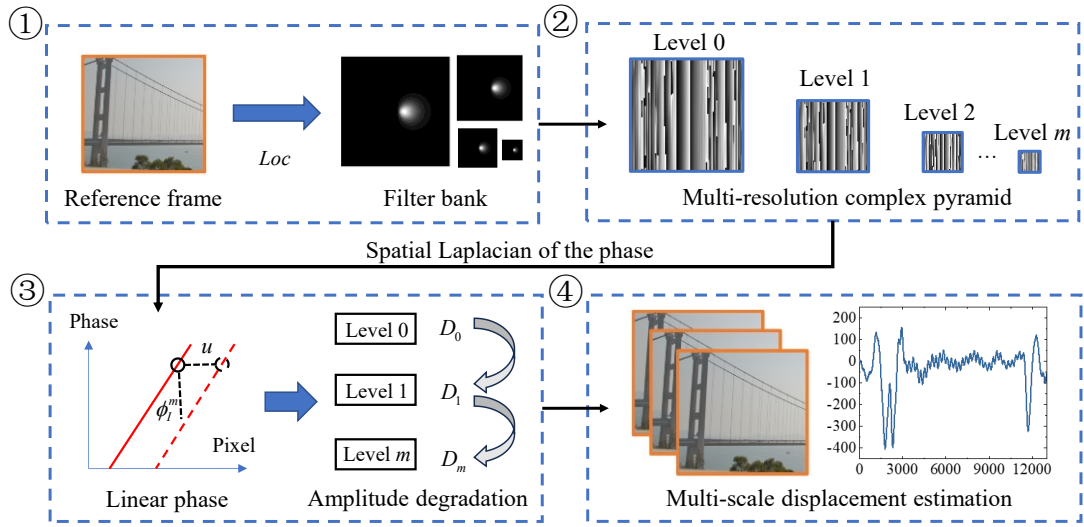


Figure 5.3 Procedure of the multi-scale motion estimation based on the multi-resolution complex pyramid

5.2.4 Experimental study

The multi-scale motion estimation performance is first validated on a two-story steel frame in the laboratory.

5.2.4.1 Experimental setup

Figure 5.4 shows the experimental setup. A two-story steel frame is 1000 mm high and 600 mm wide. The width of the beams and columns is 50 mm, and the thickness of the steel components is 5 mm. The steel frame was fixed on a stable steel girder. The steel frame was excited by hitting the left side of the frame. Two Dytran 3055D3 accelerometers (500 mV/g sensitivity, 10 g range) were installed at the right column (Points 1 and 2) to measure the horizontal acceleration at a sampling rate of 2000 Hz. The measured acceleration data were processed by high-pass filters and then double-integrated to obtain the displacement as benchmark without motion drifts. The video data with a resolution of 4096×1728 pixels was recorded by a ZCAM S6-G camera with a

Canon EF24-105 mm lens at a sampling rate of 59.94 Hz. The focal length and aperture were 24 mm and F/4, respectively.

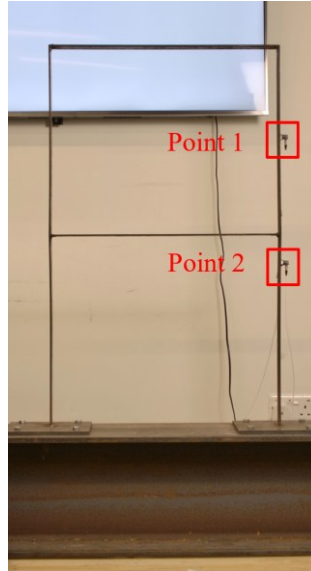


Figure 5.4 Experimental setup

5.2.4.2 Filter optimization

The localization index is first calculated to determine the filter parameters. Figure 5.5 shows the localization indices of Points 1 and 2 by applying Log-Gabor filters with different parameters to the first frame. The localization index of Point 1 varies from 13.25 to 26.90 and reaches the maximum value when the central frequency is over 1/6 Hz and the bandwidth is between 1.5 and 4.0 octaves. Point 2 has a larger localization index than Point 1 due to the stronger lighting condition. Nevertheless, the range of the optimal filter parameters at Point 2 is similar to that of Point 1, that is, central frequencies higher than 1/6 Hz and bandwidths between 1.5 and 3.5 octaves. Considering an extremely high central frequency will hasten the period change of the estimated phase and lead to more sudden jumps of phase angles, the central frequency of 1/5 Hz is selected for both points. The best bandwidths for Points 1 and 2 are 2.4 and 2.2, respectively.

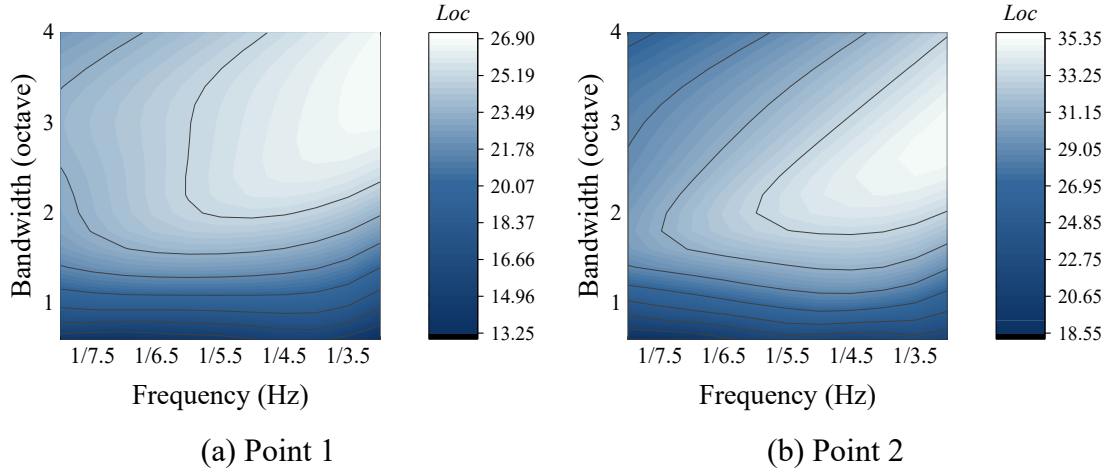


Figure 5.5 *Loc* values of Points 1 and 2

5.2.4.3 Linear measurement pixels selection

With the optimized Log-Gabor filters, the amplitude and phase at each pixel around Points 1 and 2 are calculated. The displacements are estimated by the phase variation. However, the accuracy of the estimated displacement strongly depends on the phase quality at these pixels. Therefore, pixels with a high amplitude and a small spatial Laplacian are selected to extract the structural motion.

Figure 5.6 demonstrates the image, amplitude, and spatial Laplacian of the phase at Point 1. At Level 0, stable measurement pixels are selected within a 100×100 pixel patch around Point 1, as shown in Figure 5.6(a), in which a horizontal 1D slice across the pixel patch center is extracted. Three peaks, corresponding to the left and right edges of the steel plate and the edge of the accelerometer foundation, can be found in the amplitude map of Point 1 at Level 0. Within a 5×5 pixel patch centered at the maximum amplitude value, the pixel with the lowest spatial Laplacian (0.012) is indicated by a dashed line at Pixel 48 on the left edge of the steel plate. The 100×100 pixels patch is then downsampled to 50×50 pixels at Level 1, wherein Pixel 24 has a small phase Laplacian and a high amplitude. Figure 5.6(c) shows that the amplitude map at Level 2 (25×25 pixels) has only two peaks compared with other levels because the downsampling makes textures (i.e.,

edges of steel components and the accelerometer) overlap. At Level 2, the lowest spatial Laplacian occurs at Pixel 19, which is the boundary between the steel component and the accelerometer. In this case, the steel components and the accelerometer at the same height share the same motion pattern. Therefore, the motion of Pixel 19 at Level 2, Pixel 24 at Level 1, and Pixel 48 at Level 0, indicated by yellow dashed lines in Figure 5.6, can be integrated into the multi-scale displacement.

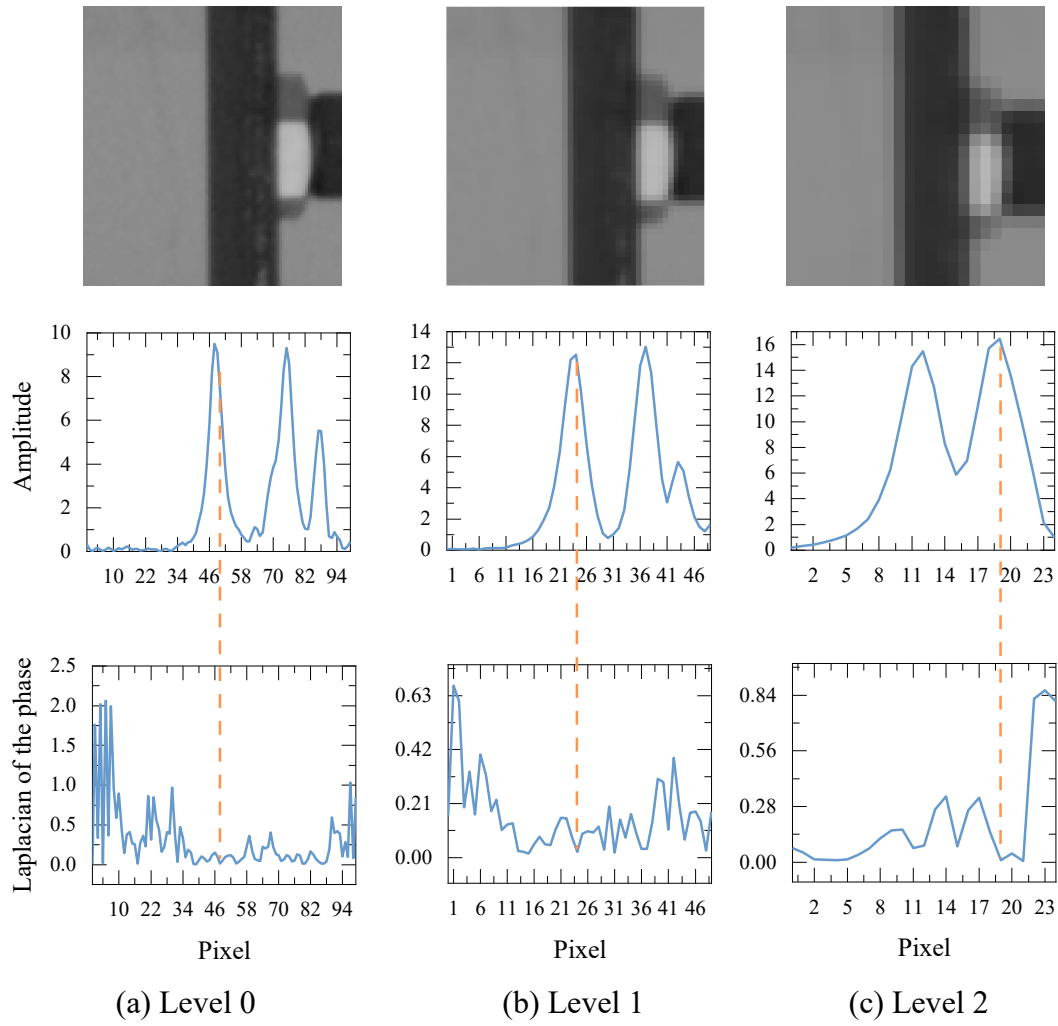


Figure 5.6 Measurement pixels of Point 1

5.2.4.4 Displacement results

The multi-scale displacement can be estimated after integrating the displacements at different levels. Figure 5.7 compares the displacements from the camera and accelerometer at Points 1 and 2 during 100 s. Differences occur at the beginning of the excitation because the double integration of the acceleration with a sudden change is inaccurate. After the initial period, the estimated displacement agrees with the accelerometer, even though the displacement diminishes to 0.05 mm at Point 1 and 0.03 mm at Point 2. The results demonstrate that the multi-resolution complex pyramid method can estimate multi-scale motions with vibration amplitudes varying 30 times.

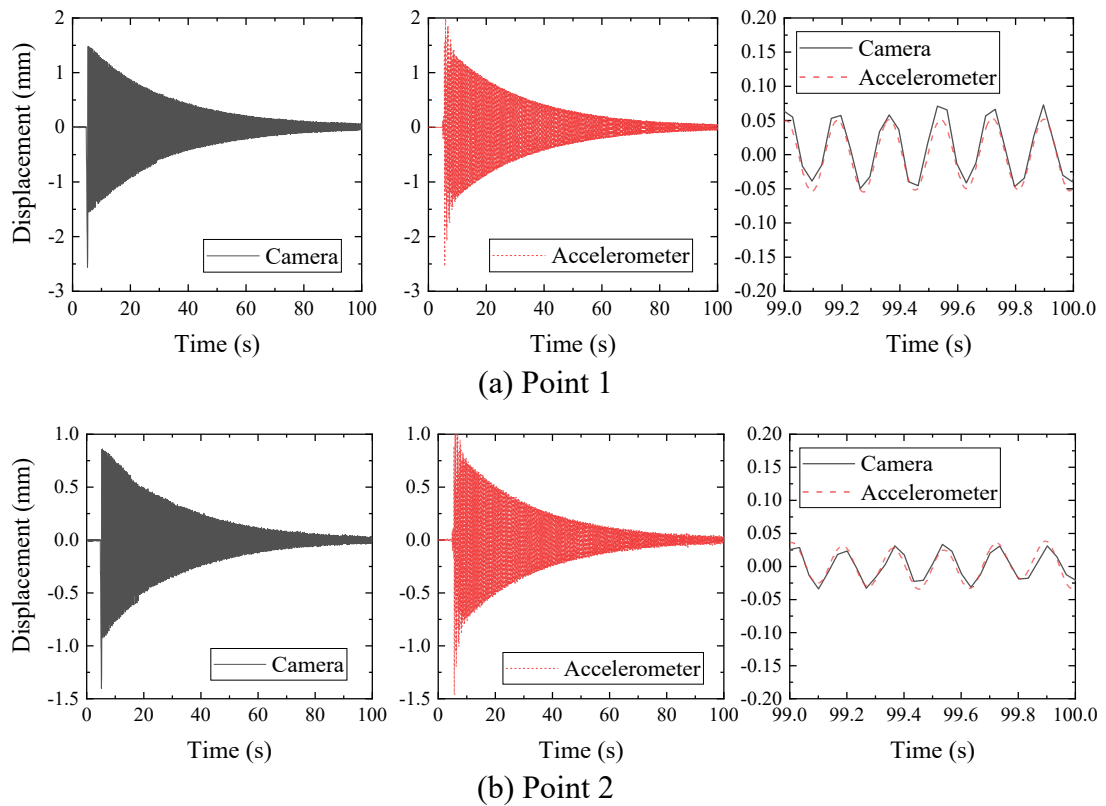


Figure 5.7 Displacements from the camera and accelerometers

The PSD results of the displacements at Point 2 are calculated using Welch's method (Hanning window) and compared in Figure 5.8. The window lengths for the camera and accelerometer data are 1,800 and 60,000 samples, resulting in the same frequency resolution of 0.033 Hz. 50% overlapping is used in the calculation. The modal frequencies

estimated by the videos and accelerometers are close. The phase-based motion estimation method offers a notable advantage: only one camera can visualize structural mode shapes.

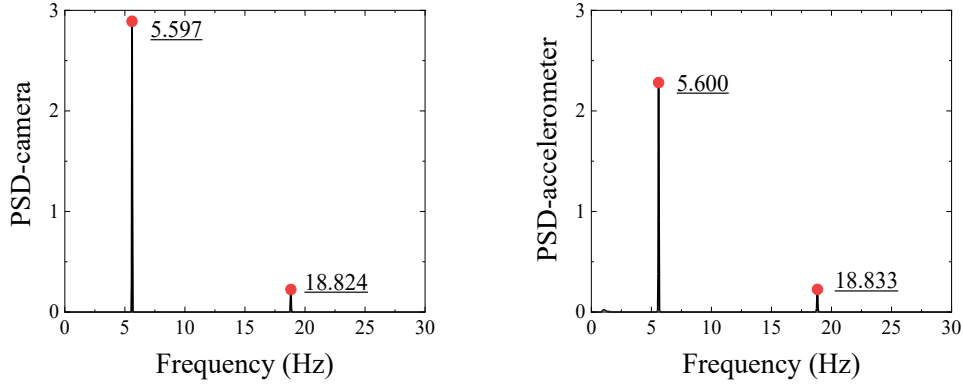


Figure 5.8 PSD of the displacement estimated from the video clip and accelerometer

The proposed technique estimates the full-field horizontal motions of the entire frame. After bandpassing, the displacements within the frequency bands 5–6 Hz and 18–19 Hz (the first two mode shapes) are obtained and demonstrated in Figure 5.9. The mode shape magnitude at the frame foundation is close to 0, then steadily increases along the height of the frame column in the first mode. For the second mode, two peak mode shape values are observed at the beam locations.

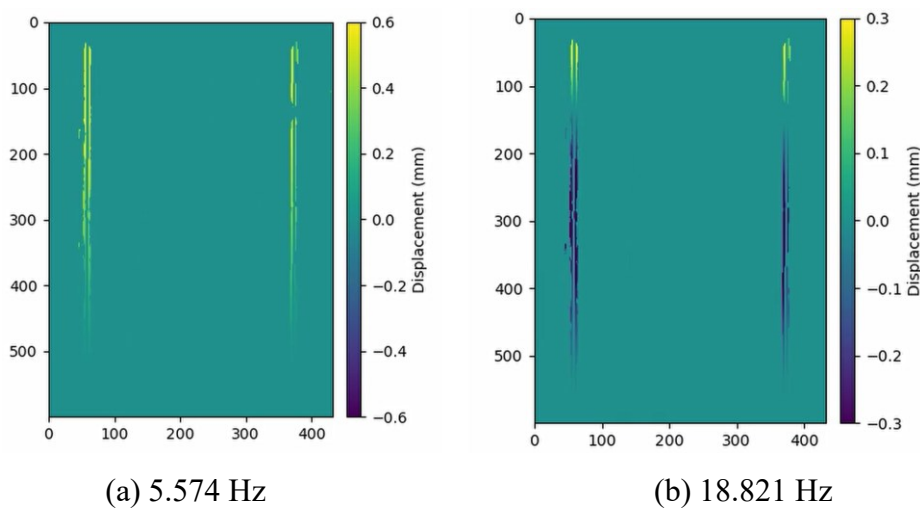


Figure 5.9 Mode shapes estimated from the video clip

Two mode shapes are compared with the FEM model in Figure 5.10, indicating that the estimated mode shapes match well with the FEM. The absolute mode values at the bottom of the vision-based results are smaller than the FEM model. This may be because the vibration at the bottom is too small.

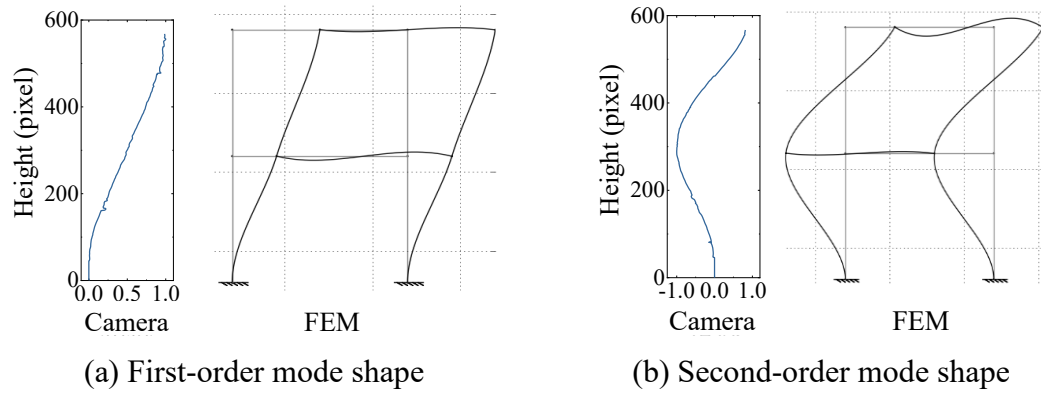


Figure 5.10 Comparison between the estimated mode shape and the FEM result

5.3 In-field study 1: Full-field displacement estimation of the Tsing Ma Bridge

5.3.1 Measurement environment

The Tsing Ma Bridge has a main span of 1377 m between two towers. It is subjected to multi-scale vibrations because of traffic loads and ambient excitations. The Tsing Ma Bridge has been equipped with a long-term SHM system, including GPS stations (Xu and Xia, 2011), making it an ideal testbed to validate the proposed method. Figure 5.11 shows the field measurement environment. A ZCAM S6G camera with a Canon EF24-105 mm lens was placed at the Lantau Link Viewing Platform to shoot the bridge's half span on the side of Tsing Yi. The camera is over 600 m from the bridge tower and over 1000 m from the middle span. The lens aperture was set to F/8 and F/4 to evaluate the displacement estimation performance under different field depths. Although the large camera aperture makes the sight depth smaller, the very short focus length of 24 mm is

used to enlarge the sight depth. Moreover, the focus point is at around the center pixel of the image. As a result, the textures at the mid span and the side span are clear.



Figure 5.11 Measurement environment of the Tsing Ma Bridge

Figure 5.12 shows one video frame captured by the camera with a focal length of 24 mm. The aperture value was set to F/8 to consider out-of-the-plane extra texture in the field depth. The videos with a resolution of 6064×2560 pixels were taken with scaling factors of 72.22 mm/pixel and 127.38 mm/pixel at the bridge tower and the mid-span, respectively. Two GPS stations (accuracy better than $1\text{ cm} \pm 2\text{ ppm}$ at a sampling frequency of 10 Hz) installed at the tower top and the mid-span are used to validate the results at Points 2 and 5 from the video. The GPS reference station is on the roof of a storage building near the bridge.



Figure 5.12 Measurement points

5.3.2 Filter optimization

The localization index is used to determine the filters' parameters. Figure 5.13 indicates the localization index of Points 2 and 5 for demonstration. The overall localization index is smaller than that in the laboratory study because the real environment is more complex than the indoor condition. Specifically, the optimal filter bandwidth is located between 1.5 and 4 octaves when the central frequency is above 1/5 Hz at Point 2. The localization index at Point 5 further decreases compared with Point 2 because the longer shooting distance makes the structural texture extraction more difficult. Nevertheless, Point 5 has optimal bandwidths ranging from 1.0 to 3.5 octaves, similar to Point 2. As a result, the bandwidths at Points 2 and 5 are selected as 2.2 and 2.0 octaves, respectively. A fairly large central frequency of 1/5 Hz is used to achieve a good localization capability and fewer phase wraps.

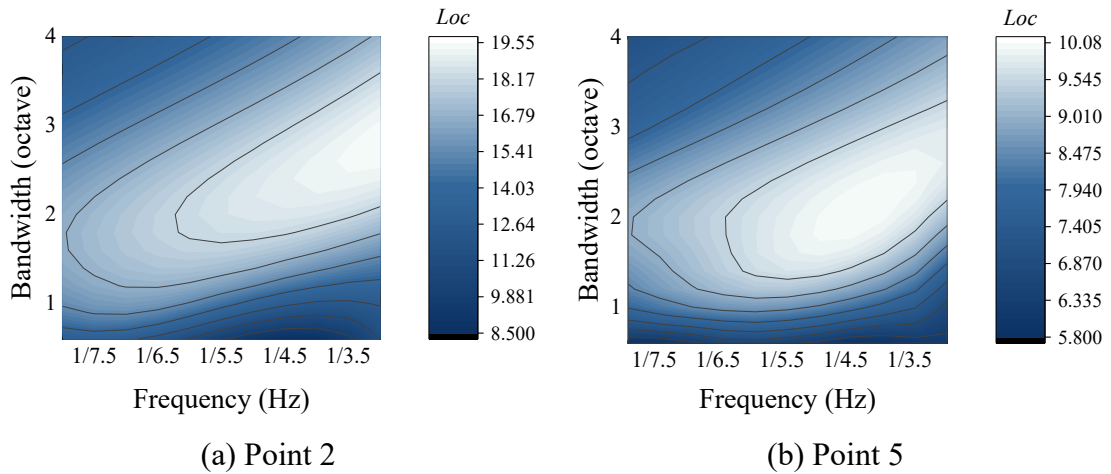


Figure 5.13 Localization indices at Points 2 and 5 of Tsing Ma Bridge

5.3.3 Linear measurement pixel selection

The blurred texture region between the structure and the background, as shown in Figure 5.14, causes difficulty in selecting pixels with accurate phases of the structure. Therefore,

the amplitude maps and the spatial Laplacian of the phase in the complex pyramid should be analyzed. At Point 2, the Log-Gabor filter at the direction of the tower edge is constructed under the instruction of the localization index, and the amplitude and Laplacian of the phase are obtained and shown in Figure 5.15. A 100×100 pixel patch at the center of Point 2 (the left edge of the tower) in Figure 5.14 is extracted to estimate the motion. Stable pixels will be selected among the horizontal 1D slice in the horizontal direction. The pixel coordinate of the pixel patch center is $(50, 50)$. The pixel patch is filtered at Level 0, in which the result of pixels on the 1D horizontal slice is demonstrated in Figure 5.15(a). A significant peak indicating the tower edge is located near Pixel 48, corresponding to a local minimum Laplacian of the phase of 0.024. Therefore, Pixel $(48, 50)$ is selected as the measurement pixel at Level 0. Similarly, linear measurement points at other downsampled levels are determined by the local minimum phase fluctuations within a 5×5 pixel patch centered at the pixel with the largest amplitude value. As a result, Pixels $(24, 25)$ and $(12, 12)$ are selected to extract the motion at Levels 1 (size of 50×50 pixels) and 2 (size of 25×25 pixels), respectively. The pixels selected are indicated by yellow dashed lines in Figure 5.15.

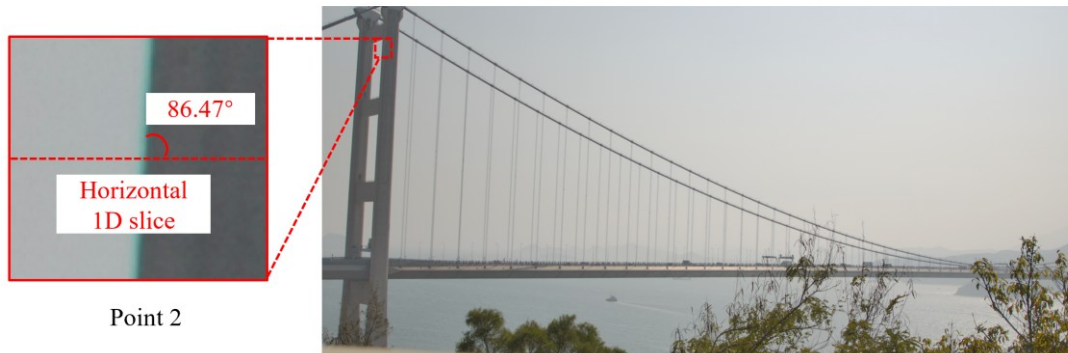


Figure 5.14 Textures at Point 2

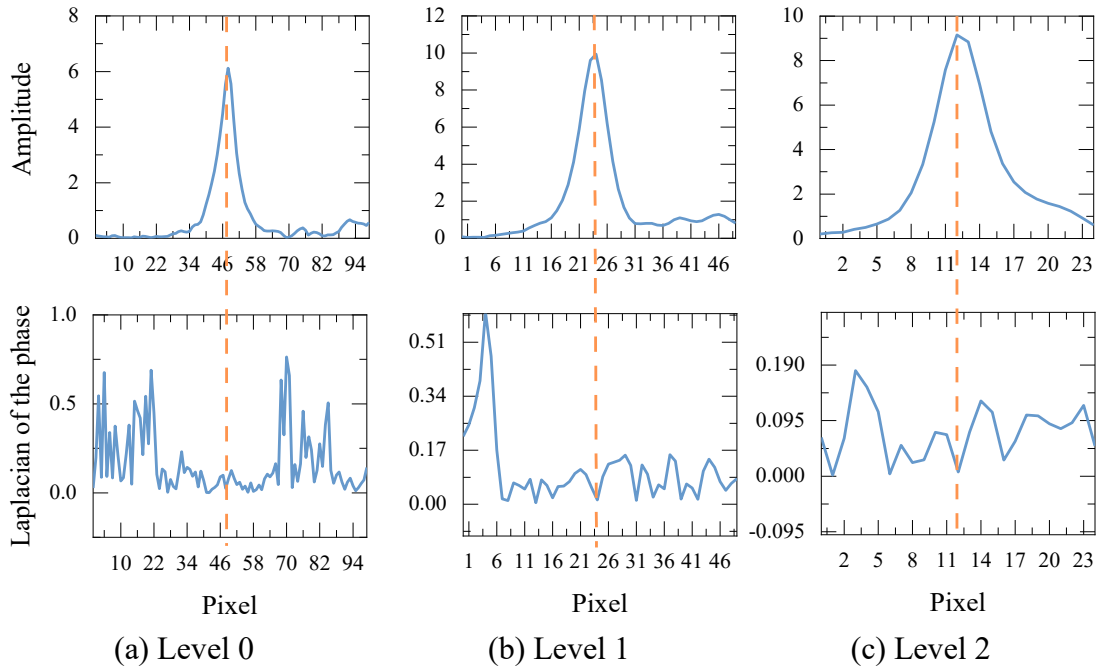


Figure 5.15 Pixel selection with the linear phase at measurement point 2

Point 5 is at the bottom of the mid-span girder, which is 91.03° to the vertical, as shown in Figure 5.16. The filter optimized by the localization index is applied, generating a complex pyramid with three levels. Figure 5.17 shows the vertical 1D slice passing the center of the 100×100 pixel patch at Point 5 (the bottom of the girder) in Figure 5.16 to select stable pixels in the vertical direction. The pixel coordinate of the pixel patch center is (50,50). The amplitude value at Point 5 is smaller than that at Point 2 because a longer shooting distance blurs the former's texture, further increasing the Laplacian of the phase at Point 5. The linear measurement pixels are determined as Pixels (50,52), (25,26), and (12,13) at Levels 0, 1, and 2, respectively.



Figure 5.16 Textures at Point 5

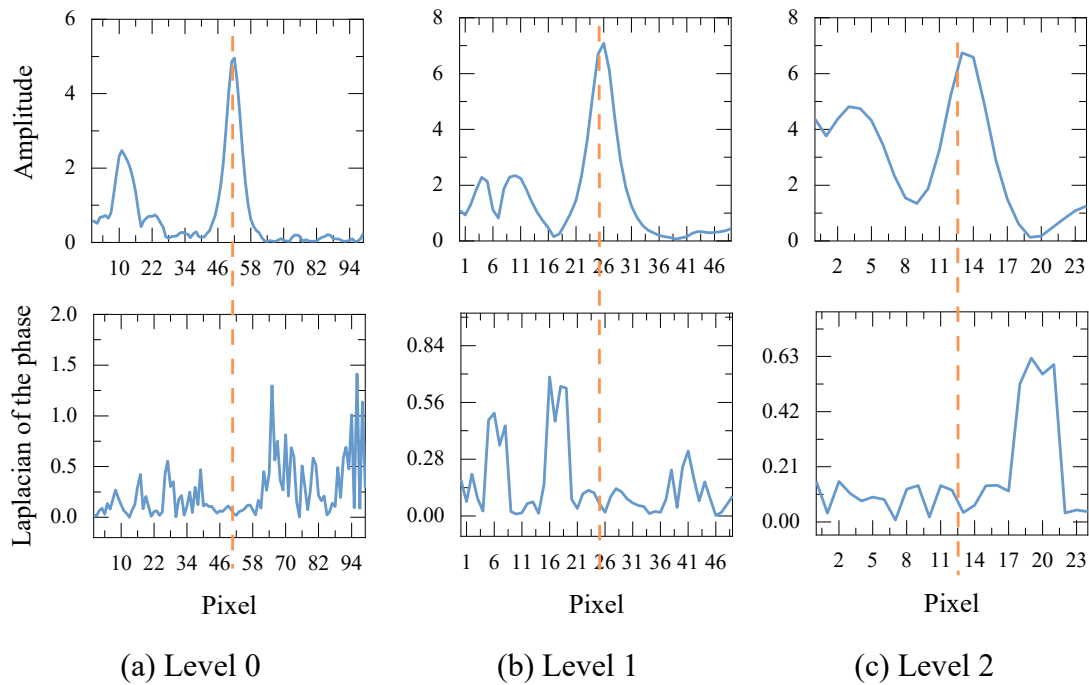


Figure 5.17 Pixel selection with the linear phase at measurement point 5

5.3.4 Displacement results

In Figure 5.18, the displacements at Points 1, 3, and 4 are also estimated to show the performance of the proposed method in estimating multi-scale motions. The displacement of Point 1 at the top of the bridge tower is similar to Point 2 with a dynamic vibration amplitude under 15 mm and a traffic-induced horizontal deflection of around 22 mm. The

maximum deflection at Point 3 is approximately 112 mm, which is smaller than Points 4 and 5 because Point 3 is close to the tower. Point 4 at the quarter girder has similar structural responses to Point 5. Besides, the traffic condition can be estimated by analyzing displacements at multiple measurement points. For instance, Point 3 has a significant deflection first at Sample 3608, followed by peaks of Point 4 at Sample 3724 and Point 5 at Sample 3991, indicating that a train was passing from Point 3 to Point 5. The traveling time from Point 3 to Point 5 is 12.77 s, and the distance is 512 m. Therefore, the train was running from Kowloon to the Hong Kong International Airport at a speed of 40.09 m/s. Similarly, Samples from 4671 to 5072 indicate a train running from the Airport to Kowloon at a speed of 38.26 m/s.

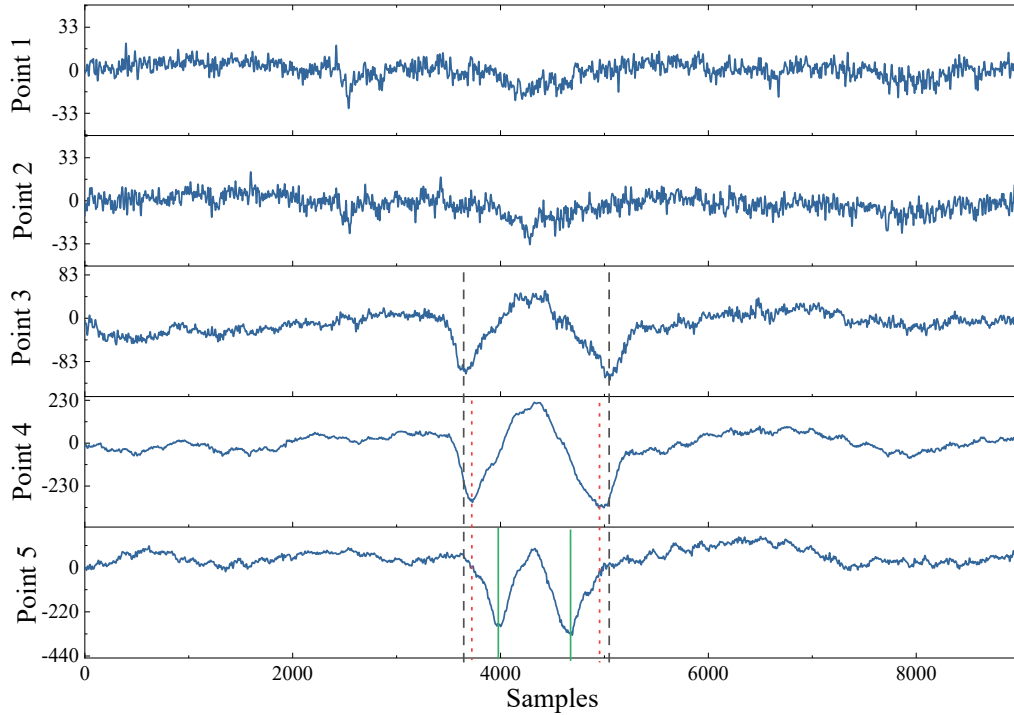


Figure 5.18 Displacements (mm) from the video clip (Points 1–5)

The displacements at Points 2 and 5 are estimated. The results are low-passed with an upper frequency of 2 Hz to suppress the self-vibration of the camera. The structural responses are not distorted because the natural frequencies of the bridge are much lower than 2 Hz. Figure 5.19 compares the displacements from the video clip and the GPS at Point 2. The noise floor of local fluctuations of the displacement estimated by the multi-

resolution complex pyramid method is from -7 mm to 7 mm, which is higher than the GPS data. Nevertheless, the tendency of the structural vibration can be revealed by the camera from a long distance. The results are consistent with the GPS data.

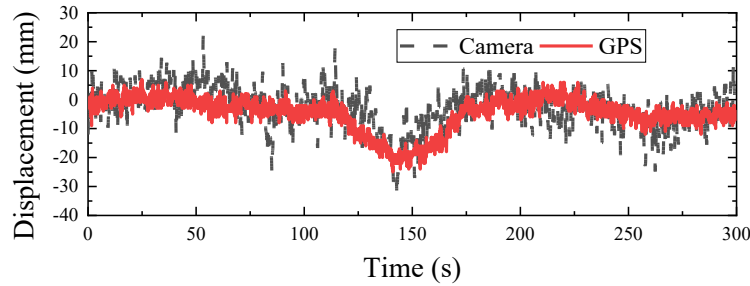


Figure 5.19 Displacements from the video clip and the GPS (Point 2)

At Point 5, the deflection of 400 mm can be observed by the camera and GPS (Figure 5.20). The large deflections may be due to the trains. The displacements from the video clip and the GPS match well. The small differences lie in the different measurement points. As shown in Figure 5.16, Point 5 is on the left of the GPS, which is blocked by plants. The distance of around 80 m between the GPS and Point 5 makes the monitored responses different.

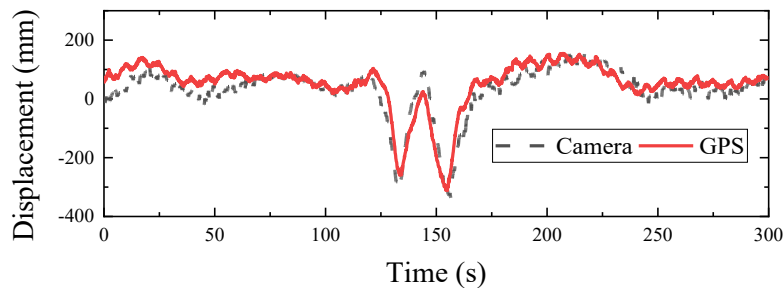


Figure 5.20 Displacements from the video clip and the GPS (Point 5)

The lens aperture was increased to F/4 to suppress the adverse effects of out-of-the-plane extra textures (plants) at the mid span. The ISO was decreased to 1200 to avoid overexposure. The climate was complex during the testing. Figure 5.21 depicts that the

environment was sunny at the beginning and cloudy in the middle, finally turning rainy with weak illumination.

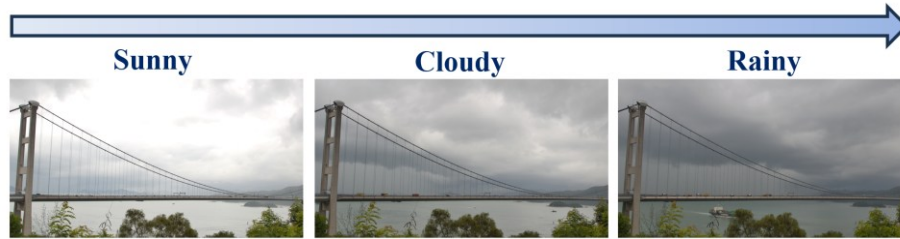


Figure 5.21 Video of Tsing Ma Bridge captured by F/4 aperture under climate change

The mid-span displacement is estimated by the multi-resolution complex pyramid and traditional Gabor phase-based method, and compared with the GPS result in Figure 5.22. The traditional method uses a Gabor filter with a central frequency of 1/4 Hz and a bandwidth of 1.2 to extract the local displacement of the mid span. The estimated train-induced deflection cannot meet the GPS result due to the nonlinear phase. Besides, the baseline has a significant drift under climate change. In contrast, the proposed method yields accurate multi-scale displacements with a noise floor of around 7.05 mm (0.05 pixels), agreeing well with the GPS result.

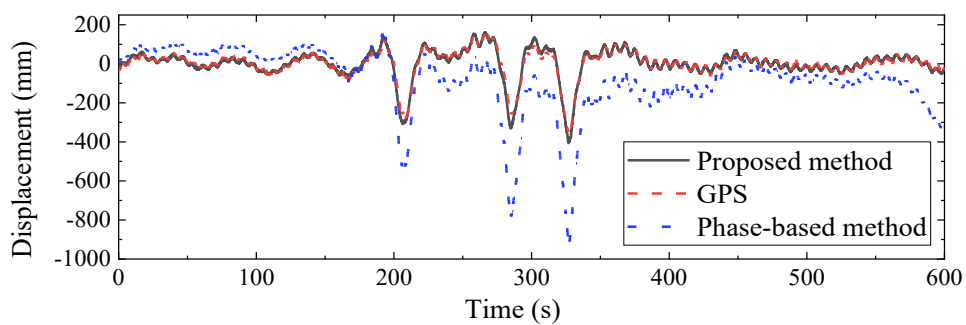


Figure 5.22 Displacements estimated from the proposed method, the traditional phase-based method, and the GPS at the mid span under climate changes

5.3.5 Full-field displacement results

To obtain the full-field displacement of the Tsing Ma Bridge, displacements of eight points on the main cable (C1–C8) and nine points on the girder (S1–S9), as shown in Figure 5.23, are measured.

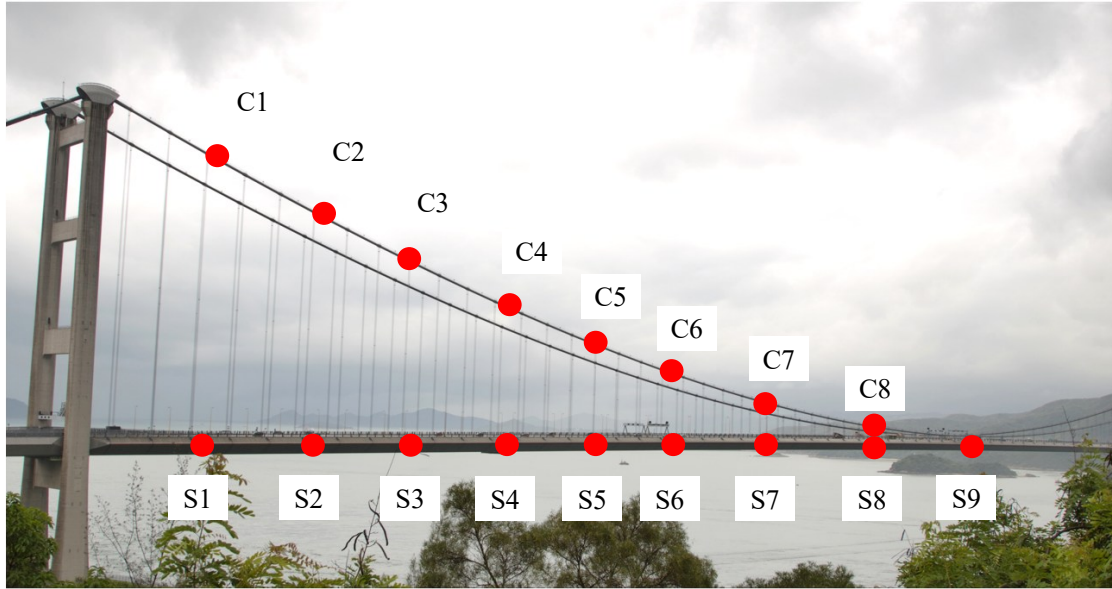


Figure 5.23 Full-field measurement points

The full-field displacements are estimated by the multi-resolution complex pyramid method. Vertical displacements of the main cables and girder (C1, C5, C8, S1, S5, and S8) are demonstrated in Figure 5.24. The displacements of main cables are projected to the vertical direction since the total displacement is vertical. The main cable has a similar displacement to the girder in the same location because they are connected through suspenders. Nevertheless, the traffic-induced deflections of the main cable are slightly smaller than the girder because the suspenders deform due to the tension force. Moreover, the difference becomes smaller at the mid-span because the steel wire is shorter than that close to the tower.

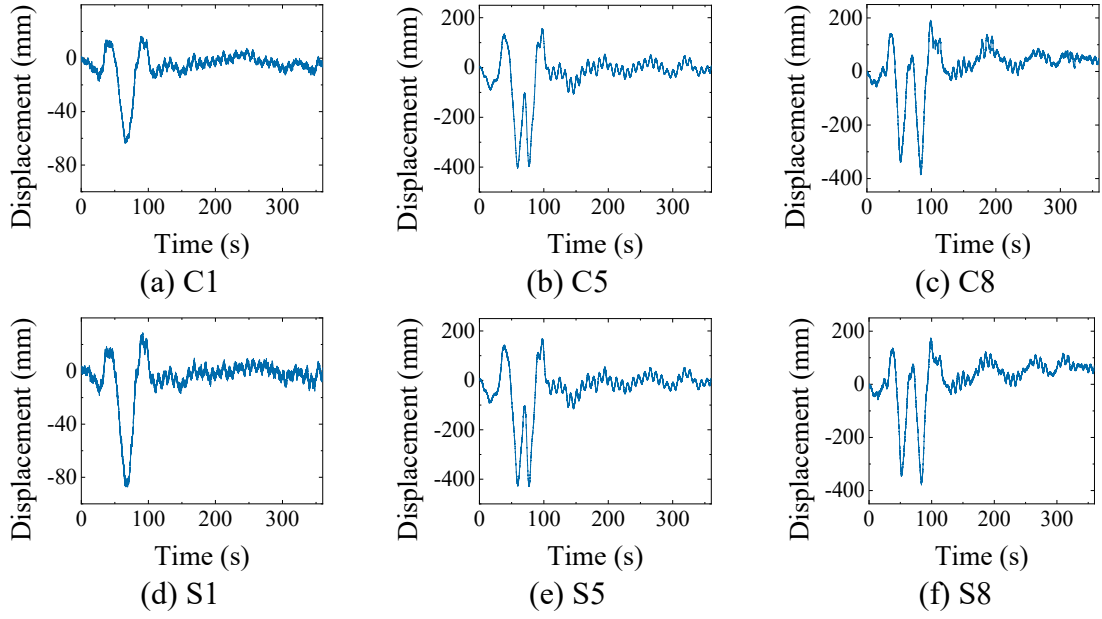
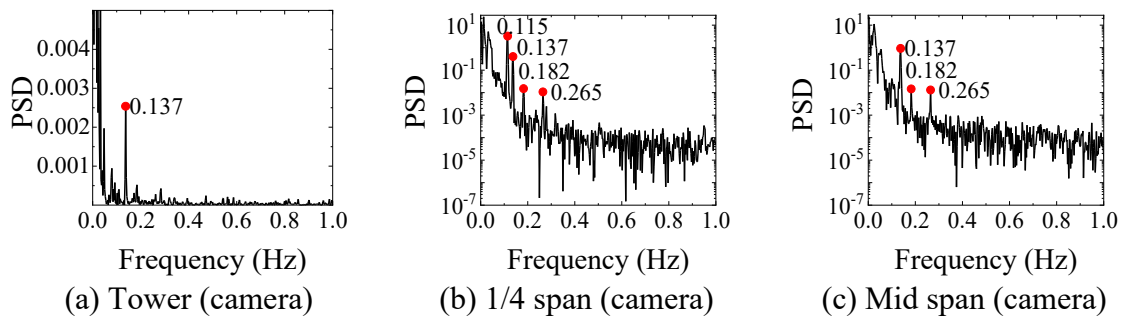


Figure 5.24 Displacements of the main cable and girder

PSDs of the displacements estimated by the video clip and the GPS are compared in Figure 5.25, in which the window lengths (Hanning window) for the camera and the GPS data are 12,000 and 4,000 samples, leading to the same frequency resolution of 0.0025 Hz. 50% overlapping is adopted in both cases. A natural frequency of 0.137 Hz can be identified from both data at the bridge tower. In Figure 5.25(c), the second bending mode of 0.137 Hz, the third bending mode of 0.182 Hz, and a mode of 0.265 Hz can be identified from the video clip. However, only the second bending mode of 0.137 Hz can be determined from the GPS. Moreover, the first bending mode of 0.115 Hz and the other three modes can be revealed from the video clip at the quarter girder, where the GPS data are unavailable.



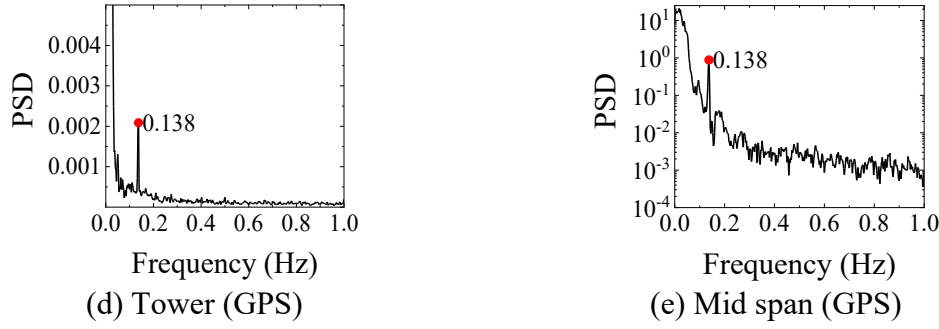


Figure 5.25 PSD of the displacements from the video clip and GPS stations

Table 5.1 compares the first four frequencies of the Tsing Ma Bridge reported by (Xu and Xia, 2011) and the present technique. The mode near 0.265 Hz is identified as the first torsion mode by 13 Honeywell Q-Flex QA700 accelerometers (1.23 mA/g sensitivity, 30 g range) employed in the SHM system. The good agreement verifies the accuracy of the present study.

Table 5.1 Measured frequencies (Hz) using video clips and accelerometers

Mode	First bending	Second bending	Third bending	First torsion
Accelerometers	0.113	0.139	0.184	0.267
Video clips	0.115	0.137	0.182	0.265

Furthermore, the estimated displacements are filtered within the frequency bands, namely, the first bending mode of 0.11 to 0.12 Hz and the second bending mode of 0.13 to 0.14 Hz. The mode shapes can then be visualized by plotting the filtered displacements with a high spatial resolution in Figure 5.26. The mode shapes from the video clip and the FEM match well. The first bending mode has a mode shape value of nearly 0 at the mid-span of the girder. Therefore, the GPS cannot measure the first bending mode in Figure 5.25, whereas the full-field response estimated by the multi-resolution complex pyramid method reveals high spatial-resolution mode shapes.

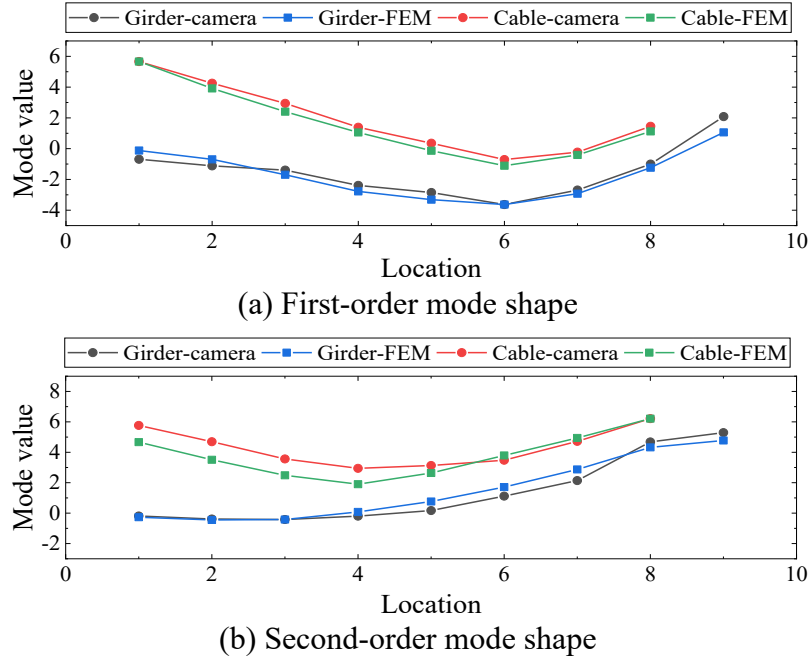


Figure 5.26 Mode shapes from the video clip and the FEM

5.4 In-field study 2: Multi-scale displacement estimation of the Kingkey 100 high-rise structure

5.4.1 Measurement environment

The multi-resolution complex pyramid is applied to estimate the multi-scale horizontal displacement of the Kingkey 100 high-rise structure under Typhoons Maliksi and Yagi. The camera is inevitably vibrated by the typhoon when installed outside. Therefore, the camera was installed on the 96th floor to measure the relative displacement between the 96th floor and the static target near the ground, as shown in Figure 5.27. The horizontal distances of the selected target to measure the N-S and E-W displacements are 412 m and 502 m, respectively. A ZCAM-S6G camera, with a Sigma 150-600 mm F5-6.3 DG OS HSM lens, was used to capture 5760×3240 -resolution videos with a sampling frequency of 25 Hz, at a focal length of 600 mm, an aperture of F/6.3, an ISO of 1600, and a shutter speed of 1/60 s. The scaling factors of the N-S and E-W displacements are calculated as 3.04 mm/pixel and 3.70 mm/pixel, respectively.

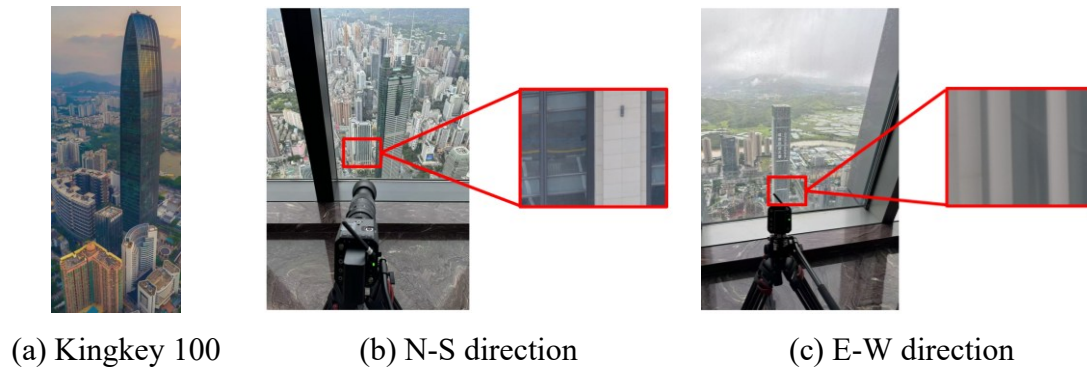


Figure 5.27 Measurement setup of the Kingkey 100 high-rise building

Typhoon Maliksi moved from south to north on June 1st, aligning with the weak-axis direction of the Kingkey 100 structure; Typhoon Yagi moved from east to west on September 6th, aligning with the strong-axis direction of the Kingkey 100 structure. The displacement of the Kingkey 100 structure under typhoon loads exhibits varying amplitude levels due to fluctuations in external excitation. Furthermore, heavy rainfall and cloudy conditions during the typhoons caused significant variations in illumination during video capture. Figure 5.28 shows the minute-by-minute illumination variations in videos recorded on the 96th floor of Kingkey 100 during Typhoons Maliksi and Yagi. It can be observed that the videos exhibited grayscale shifts due to the heavy rainfall.

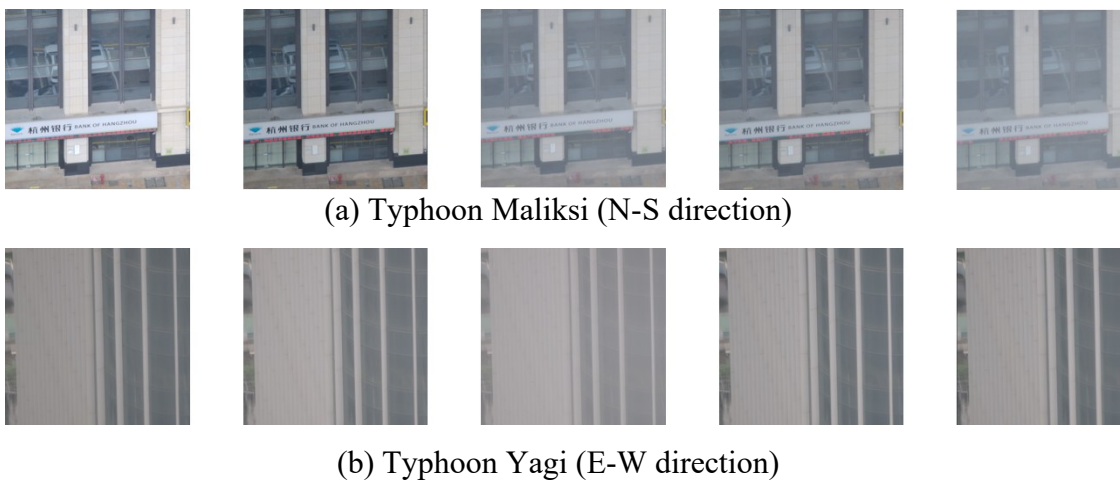


Figure 5.28 Illumination variations (per minute) of videos captured on the 96th floor of the Kingkey 100

5.4.2 Torsional mode-induced axis rotation and baseline correction

The proposed multi-resolution complex pyramid method for varying amplitudes is applied to monitor the horizontal displacement on the 96th floor of the Kingkey 100 structure. It should be noted that the displacement measurement at the 96th floor of the Kingkey 100 structure relies on the camera's movement driven by the floor displacement. When the structural response includes torsional modes, the rotation of the camera's optical axis can amplify and distort the observed torsional displacement, with the degree of distortion depending on the measurement distance. As shown in Figure 5.29, the N-S displacement measured using the proposed method exhibits periodic vibrations.

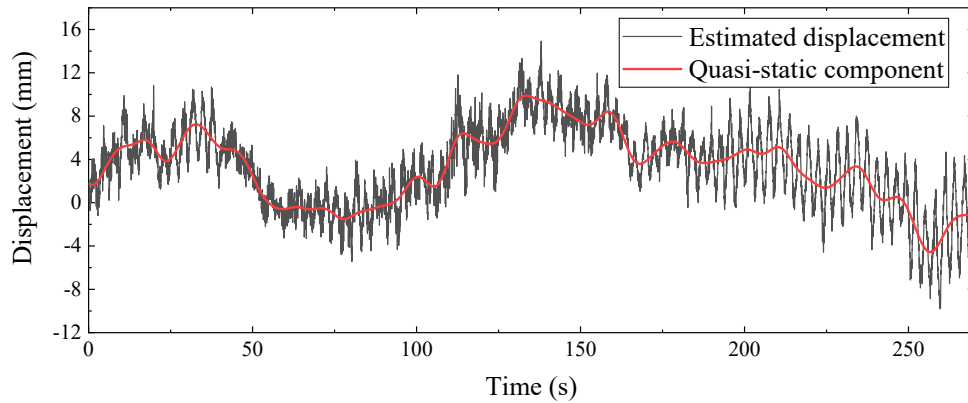


Figure 5.29 Original horizontal displacement estimated by the phase-based method on the 96th floor under Typhoon Maliksi

Frequency domain analysis in Figure 5.30 reveals that the energy corresponding to the first torsional mode of the structure (0.338 Hz) in the measured displacement was amplified due to the deflection of the camera's optical axis. Therefore, a low-pass filter with a cutoff frequency of 0.10 Hz (below the fundamental frequency of the structure) is applied to the displacement time series obtained from the camera. This filtering removes the adverse effects caused by optical axis rotation and relatively high-frequency

environmental vibrations, retaining only the low-frequency quasi-static displacement of the structure under typhoon loading.

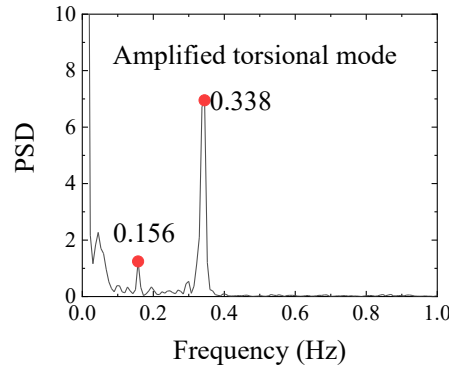


Figure 5.30 PSD of the horizontal displacement estimated by the phase-based method on the 96th floor under Typhoon Maliksi

Besides, the displacement of the reference frame is assumed to be zero in the displacement estimation. The initial deflection due to the external loads should be introduced through the baseline correction:

$$\text{mean}(\mathbf{u}_s^b) = \text{mean}(\mathbf{u}_s^g) \frac{\max(\mathbf{u}_s^c) - \min(\mathbf{u}_s^c)}{\max(\mathbf{u}_s^g) - \min(\mathbf{u}_s^g)} \quad (5.14)$$

where $\mathbf{u}_s^b$ — Estimated quasi-static displacement after baseline correction;
 $\mathbf{u}_s^g$ — Quasi-static displacement from GNSS;
 $\mathbf{u}_s^c$ — Estimated quasi-static displacement.

5.4.3 Displacement results

After baseline correction, the low-frequency quasi-static component of the N-S displacement estimated from the digital image phase on the 96th floor is compared with the quasi-static N-S displacement identified by the GNSS system at the top of the structure, as shown in Figure 5.31. Under typhoon loading, the horizontal displacement at the top of the Kingkey 100 structure is generally slightly larger than that on the 96th floor, with the low-frequency quasi-static components following the same trend. The

cross-correlation coefficient between the displacement time histories at these two floors is 0.9248. Moreover, even during the most significant illumination variations in the 3rd and 4th minutes of Typhoon Maliksi, the displacement estimated by the camera on the 96th floor remains consistent with the GNSS-measured displacement at the top. The maximum and minimum values of the low-frequency quasi-static N-S displacement identified by GNSS at the top are 21.96 mm and 5.38 mm, respectively, while those estimated by the digital image phase method on the 96th floor are 20.67 mm and 4.69 mm. These results indicate that under the low-frequency quasi-static wind component of Typhoon Maliksi moving from south to north, the Kingkey 100 structure experienced an overall quasi-static drift toward the north.

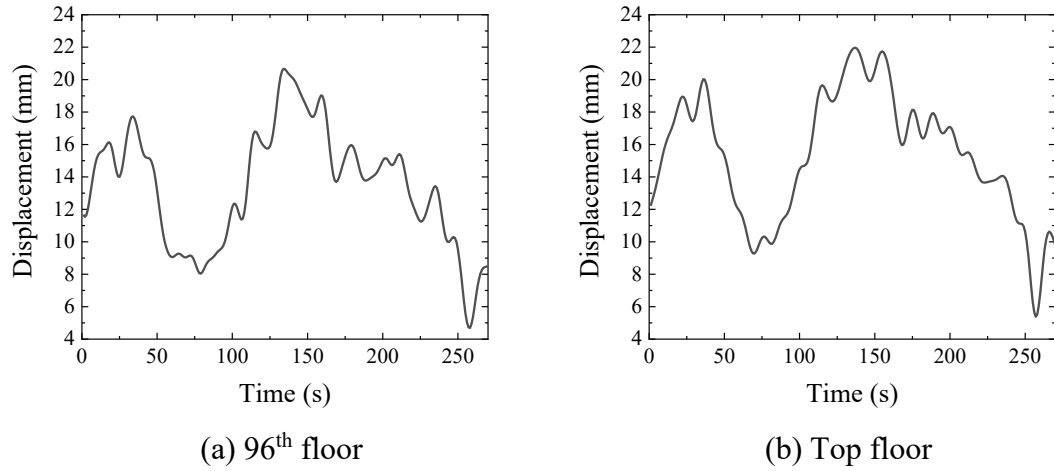


Figure 5.31 Quasi-static N-S displacements estimated by the phase-based method (96th floor) and the GNSS (top floor) under Typhoon Maliksi

The low-frequency quasi-static component of the E-W displacement after baseline correction on the 96th floor is compared with the quasi-static E-W displacement identified by the GNSS system at the top of the structure, as shown in Figure 5.32. The maximum and minimum absolute values of the top-level displacement are 21.90 mm and 12.20 mm, respectively, while those of the 96th-floor displacement are 20.55 mm and 10.89 mm. The cross-correlation coefficient between the E-W displacements estimated at the top and 96th floor is 0.9036. Moreover, even when the image turned grayish-white due to heavy

rainfall around the 3rd minute, the displacement estimated by the camera remains consistent with the GNSS-measured displacement.

In summary, the proposed displacement estimation method for multi-scale amplitudes accurately captures the horizontal displacements of the high-rise structure under typhoon conditions, despite displacement amplitudes varying by a factor of 4 and being coupled with significant illumination changes.

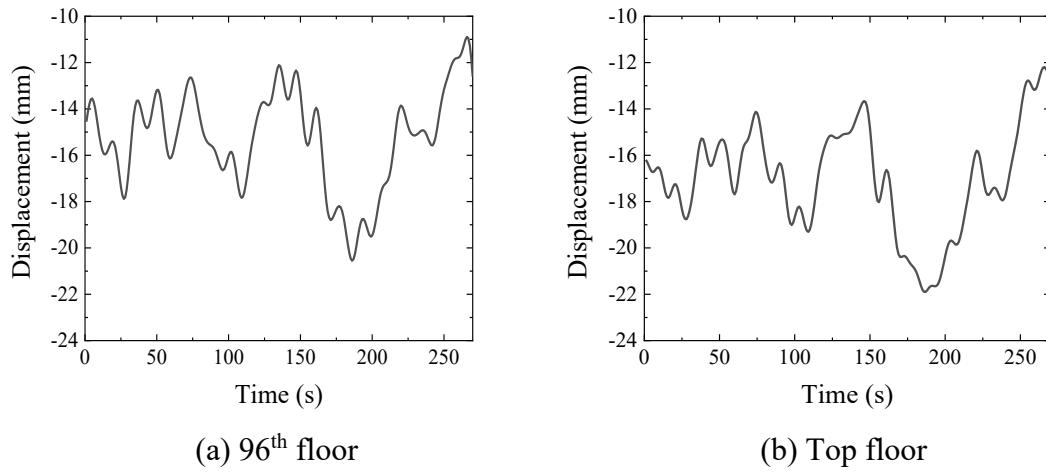


Figure 5.32 Quasi-static E-W displacements estimated by the phase-based method (96th floor) and the GNSS (top floor) under Typhoon Yagi

5.5 Multi-type target displacement estimation using discriminative correlation

5.5.1 Multi-type target detection based on discriminative correlation

The multi-resolution complex pyramid displacement estimation method for multiple amplitudes relies on the quality of the extracted phase, which must be derived from structural vibrations perpendicular to linear geometric features (e.g., straight edges of structural members). This limitation restricts its applicability to multi-type targets (such

as corners, T-junctions, and other geometric features) in large-scale structures for displacement estimation along different directions.

To address scenarios involving nonlinear structural components or vibrations parallel to linear features, this paper proposes a discriminative correlation-based approach for multi-type target displacement estimation in large-scale structures. The discriminative correlation filter adaptively optimizes the filter based on the geometric characteristics of the selected structural component:

$$E(\mathbf{f}) = \sum_{j=1}^m \left\| \mathbf{S}_f(\mathbf{g}_j) - \mathbf{y}_j \right\|^2 + \sum_{d=1}^D \left\| \mathbf{f}^d \right\|^2 \quad (5.15)$$

$$\mathbf{S}_f(\mathbf{g}_j) = \sum_{d=1}^D \mathbf{g}_j^d \otimes \mathbf{f}^d \quad (5.16)$$

where $E(\mathbf{f})$ ——Target function of the filter optimization;

$\mathbf{S}_f(\mathbf{g}_j)$ ——Filter result of the j^{th} frame, $\mathbf{S}_f(\mathbf{g}_j) \in \mathbb{R}^{W \times H}$;

$\mathbf{g}_j$ ——Geometric features of the j^{th} frame, $\mathbf{g}_j \in \mathbb{R}^{D \times W \times H}$;

$\mathbf{g}_j^d$ —— d^{th} -channel geometric feature of the j^{th} frame, $\mathbf{g}_j^d \in \mathbb{R}^{W \times H}$;

$\mathbf{y}_j$ ——2D Gaussian label, $\mathbf{y}_j \in \mathbb{R}^{W \times H}$;

$\mathbf{f}$ ——Filters to be optimized, $\mathbf{f} \in \mathbb{R}^{D \times W \times H}$;

$\mathbf{f}^d$ —— d^{th} -channel filter to be optimized, $\mathbf{f}^d \in \mathbb{R}^{W \times H}$;

$\otimes$ ——Convolution operator;

m ——Number of the training samples;

D ——Number of the channels in the geometric features;

$W \times H$ ——Image frame size.

Among these, the 2D Gaussian function centered on the target needs to be manually selected in the initial frame. In subsequent frames, it is determined by the recognition results of that frame. Its expression in the spatial domain is as follows:

$$\mathbf{y}_j(x, y) = \frac{2\pi}{WH} e^{-\frac{2\pi^2(x-x_0)^2}{W^2} - \frac{2\pi^2(y-y_0)^2}{H^2}} \quad (5.17)$$

where x_0 — Horizontal coordinate of the target center;
 y_0 — Vertical coordinate of the target center.

Figure 5.33 shows the basic process of implementing discriminative correlation. Instead of using raw pixel values, such as the DIC, or hand-crafted filters, such as the phase-based method, discriminative correlation first learns the filter from multiple features, including the raw pixel value, the HOG, and the pre-trained DL feature imagenet-vgg-m-2048 (Chatfield et al., 2014) in this study. The learned filter is used to distinguish the target of interest from the background by filtering the target to generate a 2D Gaussian label. Finally, the peak value of the 2D Gaussian function is identified as the target position:

$$(x_0, y_0) = \arg \max_{(x,y) \in [0,W] \times [0,H]} \mathcal{S}_f(x, y) \quad (5.18)$$

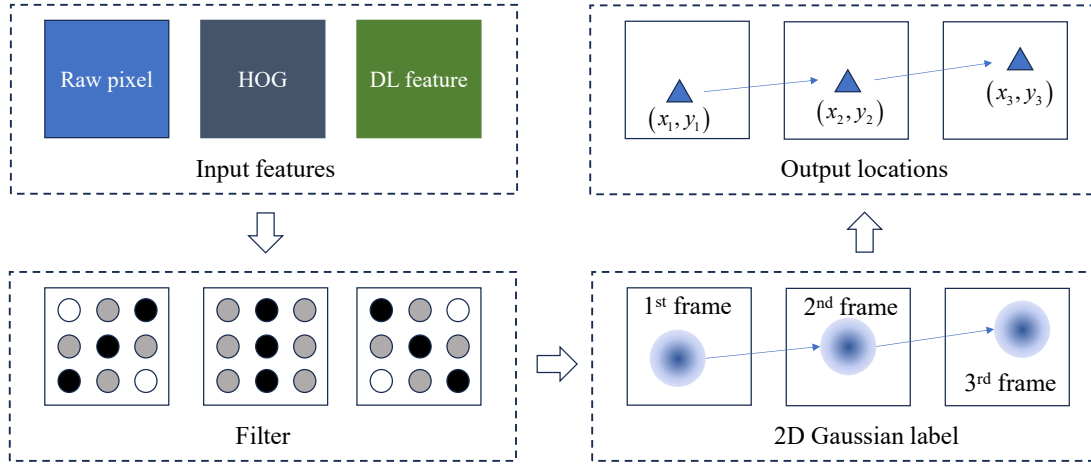


Figure 5.33 Basic process of implementing discriminative correlation

The filter optimization is conducted in a discrete manner according to Equations (5.16) and (5.17) and cannot achieve subpixel accuracy. Therefore, the geometric features are transformed into a continuous function (Danelljan et al., 2016):

$$J\{\mathbf{g}^d\}(t_g) = \sum_{n=0}^{N_d-1} \mathbf{g}^d(n) \mathcal{P}_d\left(t_g - \frac{T_g}{N_d}n\right) \quad (5.19)$$

where $J\{\cdot\}$ — Transformer of the discrete sample to a continuous function;

- T_g ——Upper range of the continuous function, set to be 1;
- N_d ——Element number of the d^{th} -channel geometric features;
- t_g ——Variable of the continuous function within $[0, T_g]$;
- $b_d(\cdot)$ ——A cubic spline kernel used to interpolate the continuous signal by superpositioning shifted versions of $b_d(\cdot)$.

Equation (5.16) is revised as:

$$\mathbf{S}_f\{\mathbf{g}\} = \sum_{d=1}^D \mathbf{f}^d \otimes J\{\mathbf{g}^d\} \quad (5.20)$$

where $\mathbf{S}_f\{\cdot\}$ ——Continuous filter result.

Unlike Equation (5.16), the above expression allows convolution of the filter with coordinates from any continuous domain, generating a continuous convolution response for the target. This study adopts a two-stage target displacement estimation strategy. First, a grid search is performed using Equation (5.18) to locate the pixel-wise coordinates corresponding to the maximum convolution response. Second, Equation (5.20) is applied to search for the maximum response in the continuous domain around the integer-pixel coordinates, achieving subpixel-level accuracy. Based on the above, Equation (5.15) is also reformulated as the objective function for continuously optimizing the filter:

$$E(\mathbf{f}) = \sum_{j=1}^m \|\mathbf{S}_f\{\mathbf{g}_j\} - \mathbf{y}_j\|^2 + \sum_{d=1}^D \|\mathbf{f}^d\|^2 \quad (5.21)$$

However, numerous parameters are required to be optimized in the continuous function, raising the necessity of compacting the parameter number (model).

5.5.2 Compact model

A factorized convolution operator is introduced to evaluate the weights of different filters learned from different features based on the useful information they contribute. The factorized convolution operator (Danelljan et al., 2017) compresses the filter dimension number from D to C , as follows:

$$\mathbf{S}_p\{\mathbf{g}\} = \sum_{d=1}^D \left(\sum_{c=1}^C p_{d,c} \mathbf{f}^c \right) \otimes J\{\mathbf{g}^d\} = \mathbf{P}\mathbf{f} \otimes J\{\mathbf{g}\} \quad (5.22)$$

where $\mathbf{S}_p\{\cdot\}$ — Filter result of the compact model;
 C — Channel number of the compact model;
 $\mathbf{P}$ — Factorized convolution operator, $\mathbf{P} \in \mathbb{R}^{D \times C}$;
 $p_{d,c}$ — Element in the factorized convolution operator.

The factorized convolution operator linearly combines C filters to estimate the original-dimension filter with D channels. For example, 514 layers of features, namely, the raw pixel value, the histogram of oriented gradient (only consider the direction of the target edge), and the 512 deep learning feature layers, are used, requiring 514 filters with numerous parameters to be learned (Figure 5.34). The 514×66 -size factorized convolution operator is adopted to ensure that only 66 filters can be linearly combined to generate 514 filters.

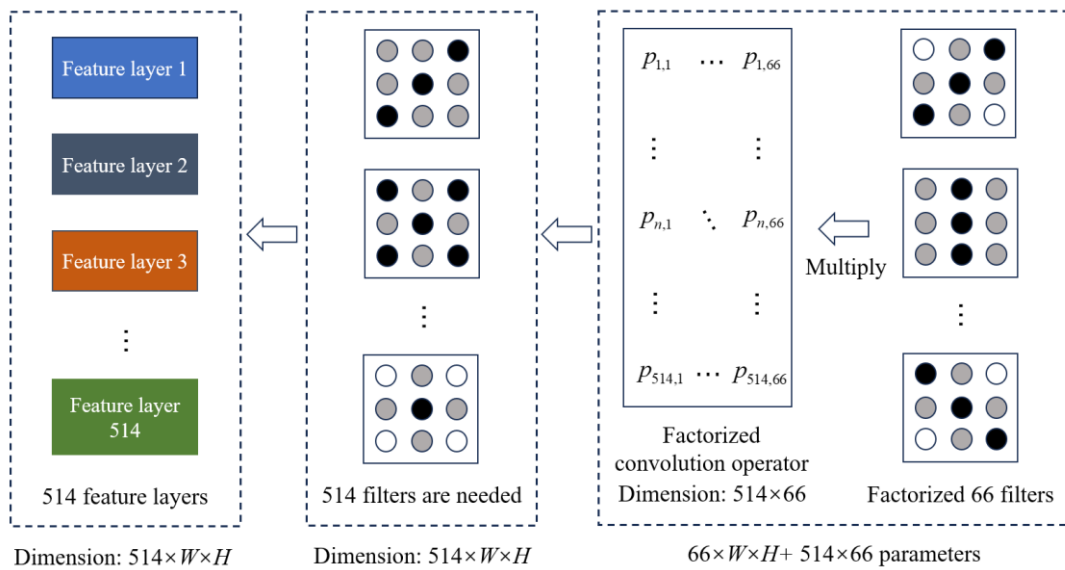


Figure 5.34 Model parameter compacting using the factorized convolution operator

The factorized convolution operator quantifies the importance of each of the C filters to assemble the original filter with D channels. The contribution of each filter rarely changes because the geometric texture of the ROI is constant during monitoring. Therefore, the matrix $\mathbf{P}$ is learned only from the initial image frame:

$$E(\mathbf{f}, \mathbf{P}) = \left\| \hat{J}(\mathbf{g})^T \mathbf{P} \hat{\mathbf{f}} - \mathbf{y}_1 \right\|_2^2 + \sum_{c=1}^C \left\| \hat{\mathbf{f}}^c \right\|_2^2 + \left\| \mathbf{P} \right\|_2^2 \quad (5.23)$$

where $E(\mathbf{f}, \mathbf{P})$ ——Target function to obtain the factorized convolution operator;

$\hat{\cdot}$ ——Frequency-domain transformer.

5.5.3 Compact samples

In a typical discriminative correlation, the samples used to learn the filter are stored at every time step, resulting in redundancy and poor diversity. A Gaussian mixture model (Danelljan et al., 2017) is introduced to compact the training samples:

$$\mathbf{G}_{\text{mix}}(\mathbf{g}) = \sum_{l=1}^L \eta_l \mathbf{G}_l(\mathbf{g}, \boldsymbol{\mu}_l) \quad (5.24)$$

where $\mathbf{G}_{\text{mix}}(\mathbf{g})$ ——Gaussian mixture model with L Gaussian components;

$\mathbf{G}_l(\mathbf{g}, \boldsymbol{\mu}_l)$ —— l^{th} Gaussian component with a mean of $\boldsymbol{\mu}_l$ and a standard deviation of unit matrix;

$\boldsymbol{\mu}_l$ ——Mean of the l^{th} Gaussian component, $\boldsymbol{\mu}_l \in \mathbb{R}^{W \times H}$;

η_l ——Weight of the l^{th} Gaussian component;

L ——Number of compact samples, $L < m$, which is 50 in this study.

The next sample is integrated into the Gaussian mixture model if the sample number is smaller than $L+1$:

$$\boldsymbol{\mu}_l = \mathbf{g} \quad (5.25)$$

$$\eta_l = \gamma \quad (5.26)$$

where γ ——Initial learning rate, 0.012 in this study.

The Gaussian mixture model should be simplified after adopting the new frame. Therefore, the two components with the smallest Euclidean distance are merged into one component, whose weight and parameters are determined by Equations (5.28) and (5.29). The updating process of the Gaussian mixture model is depicted in Figure 5.35. A new sample is added to the training sample when forwarding to the next frame, increasing the size of the training set from L to $L+1$. The Euclidean distances between different samples in the training set are calculated to find the sample pair with the shortest distance. A short distance indicates that the samples are similar, making them redundant for training filters. Therefore, the sample pair with the shortest distance is merged to generate a new sample.

$$k = \arg \min \|\mu_n - \mu_k\|_2 \quad (5.27)$$

$$\mu_k = \frac{\eta_k \mu_k + \eta_n \mu_n}{\mu_k + \mu_n} \quad (5.28)$$

$$\eta_k = \eta_k + \eta_n \quad (5.29)$$

where

μ_n — The next sample;

μ_k — Existing sample which is most similar to the next sample;

η_n — Weight of the next sample, which is the initial learning rate;

η_k — Weight of the mixed sample.

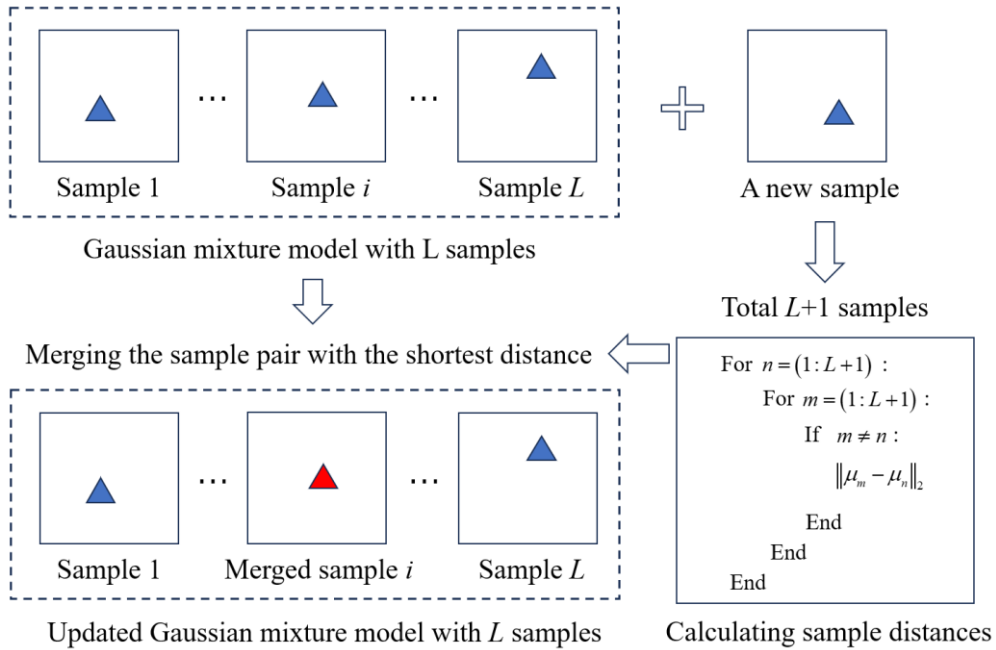


Figure 5.35 Updating process of the Gaussian mixture model

The Gaussian mixture model is updated for every frame. Nevertheless, the size remains L . The filter is efficiently learned by combining a compact model of Equation (5.22):

$$E(\mathbf{f}) = \sum_{l=1}^L \eta_l \left\| \mathbf{S}_p \{ \boldsymbol{\mu}_l \} - \mathbf{y}_l \right\|_2^2 + \sum_{d=1}^D \left\| \sum_{c=1}^C p_{d,c} \mathbf{f}^c \right\|_2^2 \quad (5.30)$$

The proposed full-field motion estimation method (Teng et al., 2025) is shown in Figure 5.36. In Phase 1, the ROI to be detected is selected in the first frame. A Gaussian function centered at the ROI is generated as a label. The ROI is transformed into a continuous domain. The feature dimension is reduced by minimizing Equation (5.23) at the first frame in the continuous domain, based on which the filter is learned through Equation (5.30). In Phase 2, the filter learned in Phase 1 is applied to the i^{th} frame, in which the target position is identified at the coordinate with the maximum response. A scaling factor is calculated on the basis of the geometrical relationship between the pixel and the real-world coordinate systems, estimating the structural motion at the i^{th} frame. In the last phase, the label is updated at the $(i+1)^{\text{th}}$ frame. The filter should be learned as in Phase 1 every five frames to suppress the adverse effects of the target deformation, because the target tracking should consider the pose of the target. Although the large-scale structure rarely changes the geometrical envelope, a relatively large displacement can cause local deformation, such as the slope of the column in the ROI. In the large displacement case, the filter should be updated to locate the deformed target. The filter is updated every five frames, which is significantly higher than the mode frequency of large-scale structures. Finally, the motion of each frame is estimated as in Phase 2.

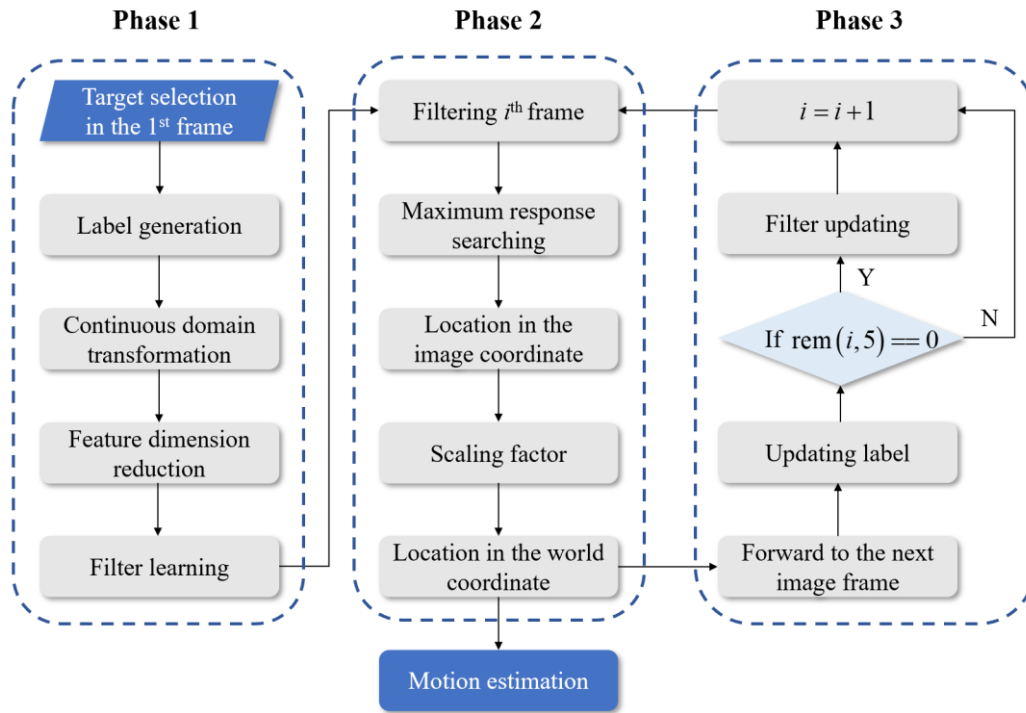


Figure 5.36 Procedure of the motion estimation for multi-type targets based on the discriminative correlation

5.5.4 Experimental study on the four-story steel frame

A four-story frame model test was conducted in the laboratory to verify the performance of the proposed multi-type target displacement estimation method.

5.5.4.1 Experimental setup

Figure 5.37 shows the experimental setup. The four-story steel frame is 3200 mm high and 1400 mm wide. The steel component has a depth of 300 mm and a thickness of 8 mm. The steel frame fixed on the ground was excited by applying a horizontal initial deflection of 2.5 cm on the frame top. The structural responses were measured using three types of sensors, namely, a camera, accelerometers, and laser displacement sensors. A Canon 70D camera (Figure 5.38(a)) was placed 2.5 m away from the frame to shoot

videos with a resolution of 1920×1080 at a sampling rate of 30 Hz, resulting in a scaling factor of 1.608 mm/pixel. Micro-Epsilon laser displacement sensors and 941B accelerometers (Figure 5.38(b) and (c)) were installed to measure the displacement and acceleration at each floor. The sampling rates of the accelerometers and laser displacement sensors were set to 1000 and 30 Hz, respectively. The laser result is used as the benchmark to verify the proposed displacement estimation method.

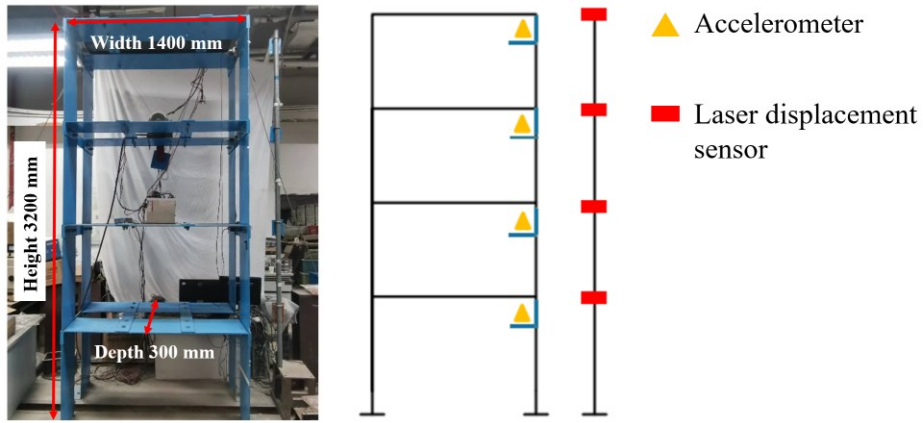


Figure 5.37 Experimental setup

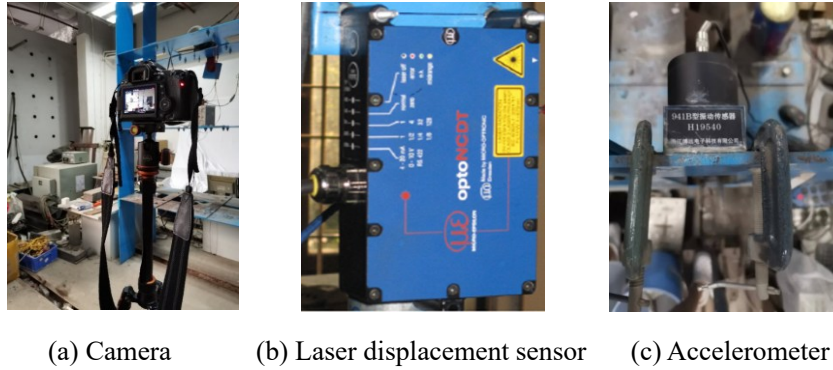


Figure 5.38 Sensors used to measure structural responses

5.5.4.2 Displacement estimation

The displacements of each floor during the 40-s free attenuation are estimated using the proposed method and compared with the laser displacement sensor results in Figure 5.39.

The displacements measured by the camera and laser displacement sensor on the fourth floor match well when the structural vibration decays from 25 mm to 7.5 mm over time (Figure 5.39(a)). The initial vibration amplitudes of other floors are 6 mm on the first floor, where the minimum displacement amplitude during the last 5 s is lower than 2 mm. At all measurement points, the camera results are smooth and almost overlap with the laser displacement sensor results, achieving a submillimeter resolution.

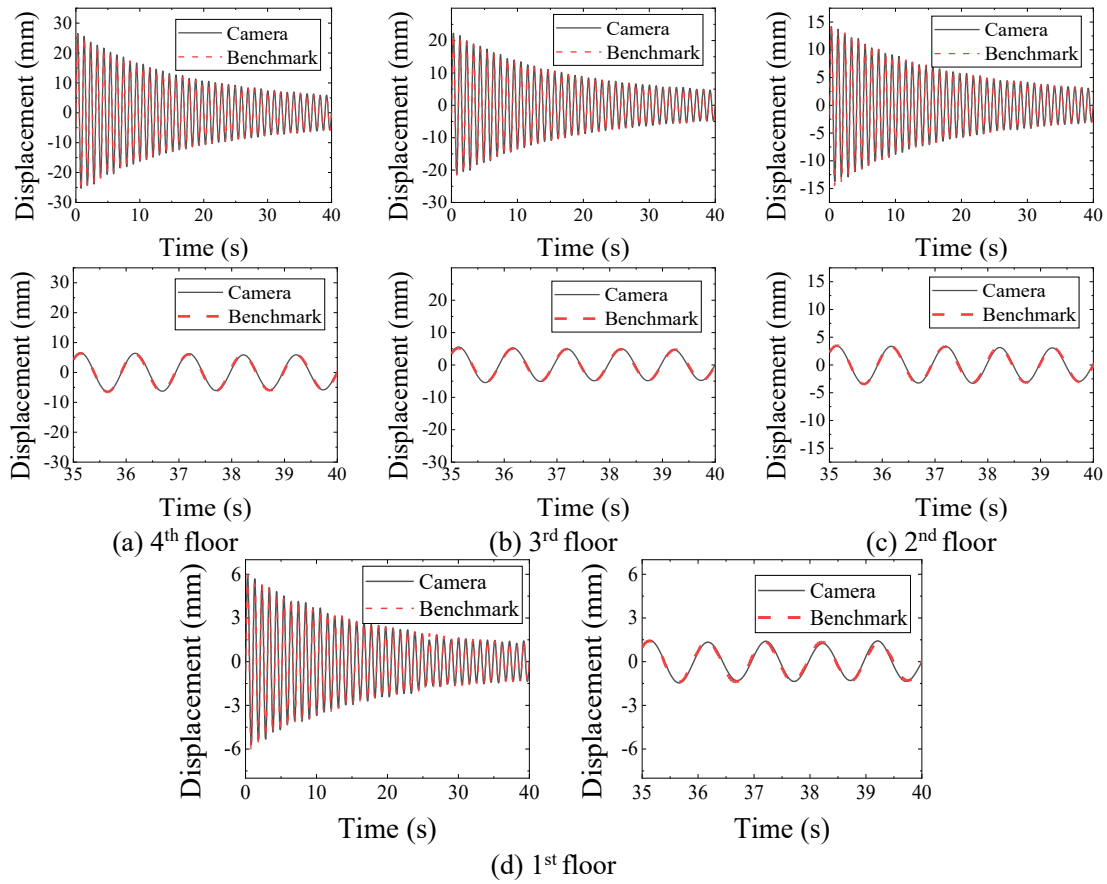


Figure 5.39 Measured displacements on each floor

The PSD results of the displacements estimated by the present method are shown in Figure 5.40. The first-order frequency (0.9799 Hz) dominates the vibration of each floor because the frame is activated by an initial deflection applied to the frame top. The natural frequencies of the four floors are identified and averaged. The natural frequencies from the camera, accelerometer, and laser displacement sensor and compared in Table 5.2. The good agreement confirms the accuracy of the present approach.

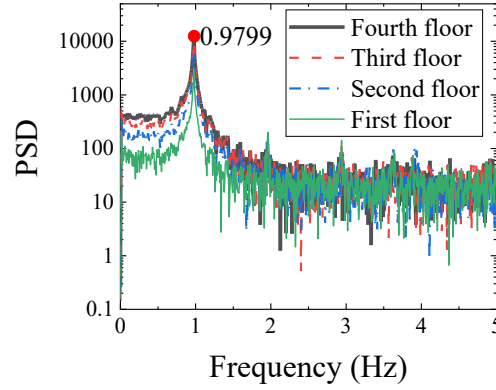


Figure 5.40 PSD of the displacements estimated from the video

Table 5.2 Natural frequency measured by the camera, laser sensor, and accelerometer

Sensor	Camera	Laser	Accelerometer
First-order frequency (Hz)	0.9799	0.9785	0.9796

5.5.4.3 Displacement estimation performance comparison regarding different methods

The second-floor displacement, with amplitudes varying from 2 pixels to 7 pixels, is estimated by the DIC (Blaber et al., 2015), the phase-based method (Collier and Dare, 2022), and the proposed method for comparison. The DIC starts by making an initial guess using normalized cross-correlation among raw pixel values to determine the integer location and refine the result to achieve subpixel accuracy. The subset radius of the DIC is set to the default value (20 pixels). The phase-based method uses the Gabor filter (0.25 Hz central frequency and 1.2 octaves bandwidth) to extract the phase, which estimates the displacement at a subpixel level. The ROI size is 64×64 pixels for all methods. The estimated displacements are compared across different displacement amplitudes and methods in Figure 5.41. The phase-based method efficiently performs when the displacement amplitudes are 2.80 and 2.16 pixels in Figure 5.41(c) and (d). However, the displacements with amplitudes higher than 4 pixels are significantly underestimated by the phase-based method due to the phase nonlinearity for large motions, which is also reported by relevant studies (Miao et al., 2023; Wang et al., 2024b; Weng et al., 2025).

Although the DIC result roughly matches the laser sensor, a maximum drift of around 1.2 mm is observed because the raw pixel value is insufficient to locate the target when similar pixels along the vertical direction of the column exist. The proposed method extracts the steel frame column from the background using the discriminative filter and multiple features, estimating accurate displacements with amplitudes varying from 2 pixels to 7 pixels. The relative errors of the different methods compared with the benchmark are shown in Table 5.3. The present method with a relative error below 2% outperforms the DIC and phase-based methods.

Table 5.3 Relative errors of the different methods on the second floor

Method	Proposed method	DIC	Phase-based method
Relative error	1.97%	4.88%	39.98%

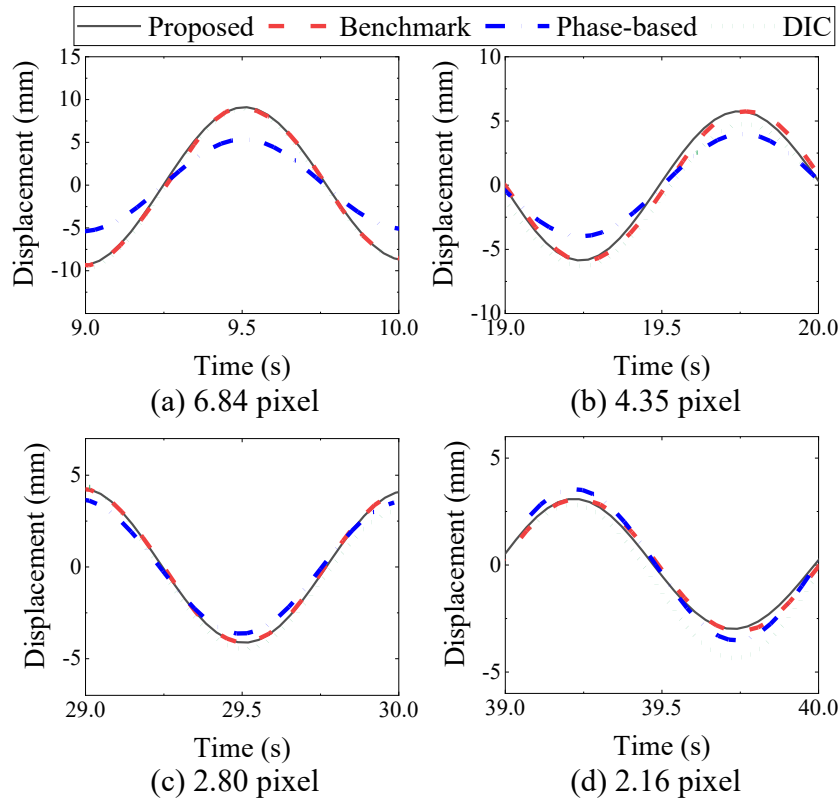


Figure 5.41 Displacements estimated by different methods under different amplitudes

5.6 In-field study 3: Multi-type target displacement estimation of the Humen Bridge

The 888 m-long main span Humen Bridge suffered unexpected vortex-excited resonance on May 5, 2020. Full-field vibration data were needed to evaluate the structural performance. Accordingly, the surveillance video (Figure 5.42) captured by a camera installed on the bridge was used to estimate the vertical full-field motion. The video was taken with a resolution of 1920×1080 pixels at a frequency of 30 Hz.



Figure 5.42 Frame of the surveillance video

The light poles distributed along the bridge are selected as the measurement points. Three reasons support the use of light poles as targets to measure the vertical displacements of the bridge deck. First, each light pole is rigidly connected to the deck at specific locations, ensuring that they have the same vertical displacement. Second, the light pole has distinguishable textures along the vertical direction, allowing them to be detected by the proposed method. Third, the spatial size of the light pole is known, which helps in calculating the scaling factor between the pixel and the real-world coordinates. Figure 5.43 marks 28 measurement points (14 for each side) on two sides of the bridge. The scaling factors range from 21.08 mm/pixel to 54.01 mm/pixel, given that the height of the light pole is 8.5 m.

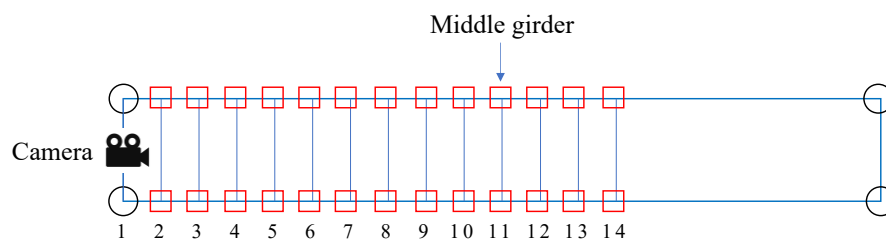
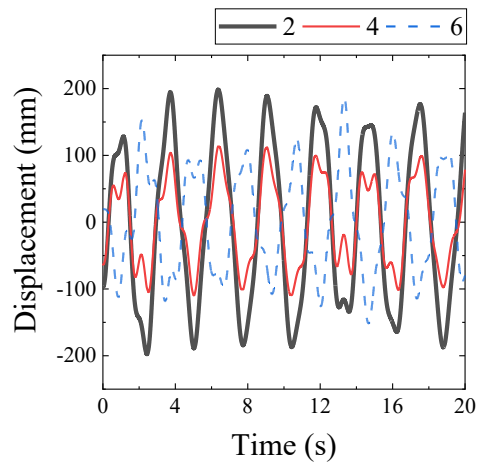
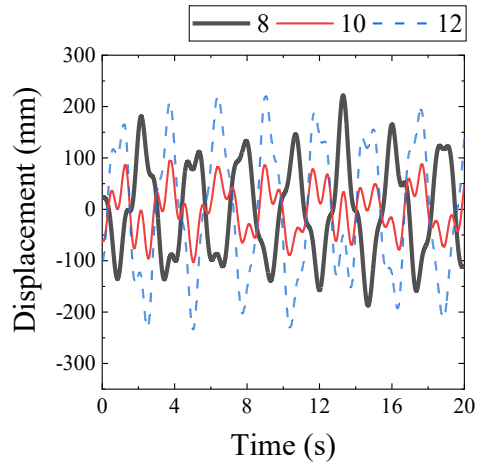


Figure 5.43 Measurement points of the bridge

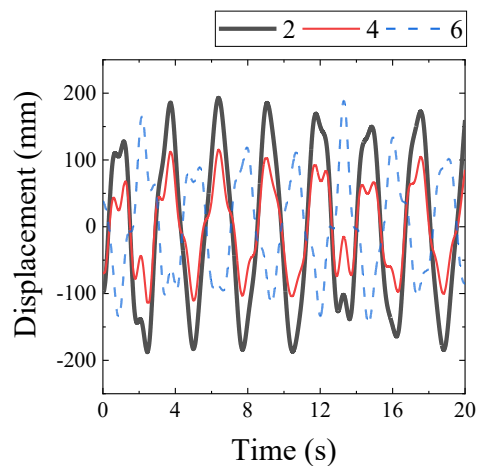
The full-field displacements of different measurement points are estimated from the surveillance video. Figure 5.44 shows the measurement points with even indices on the left and right sides of the bridge. The maximum amplitude of the vibration is around 200 mm at Points 2, 8, and 12. The displacements of all points show a dominated period of approximately 3 s. Points 2, 4, 10, and 12 share the same phase in the time domain, while Points 6 and 8 have a 1.5 s phase shift in displacement, indicating that the vibration directions are different. The displacements at the left and right sides are roughly the same because few torsional mode components exist during the vortex-excited resonance.



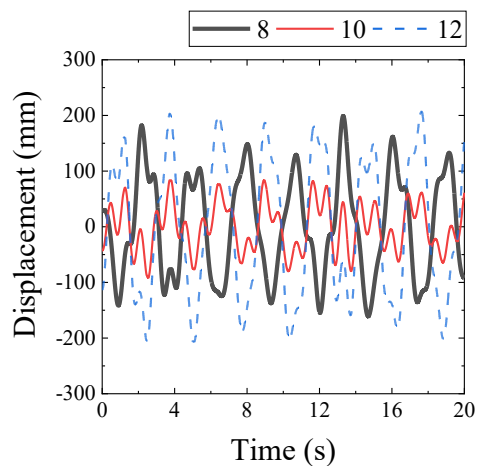
(a) Left side (Points 2, 4, and 6)



(b) Left side (Points 8, 10, and 12)



(c) Right side (Points 2, 4, and 6)



(d) Right side (Points 8, 10, and 12)

Figure 5.44 Estimated displacements of the bridge

The PSD of the displacements is calculated and shown in Figure 5.45, where two modes of the mode frequency (0.3662 and 1.1719 Hz) can be identified. The high-resolution mode shape is estimated by introducing the SSI-COV (Hu et al., 2022). Accelerometers are used to provide the benchmark of the mode shape. The mode shapes calculated from the accelerations are also identified by using the SSI-COV method. The accelerometers (8 for each side) are uniformly distributed along the longitude direction of the bridge deck, which is sparse and cannot be directly aligned with the vision-based measurement points (i.e., locations of light poles). Consequently, the sparse mode shapes are interpolated and normalized to generate high-resolution mode shapes using the spline interpolation method. Figure 5.46 compares the normalized mode shapes obtained from the video and the accelerometers at the frequency of 0.3662 Hz, indicating a good agreement between the video and accelerometer results. The mode values at Points 2, 7, and 12 are larger than those at other positions, which is consistent with the displacement results in Fig. 26 because the vibration is dominated by the frequency of 0.3662 Hz. Meanwhile, the mode values at Points 5 and 9 are close to zero, which explains the vibration direction differences among different points in Figure 5.44. The MAC values of the left and right girders of the bridge are 0.9725 and 0.9792, respectively.

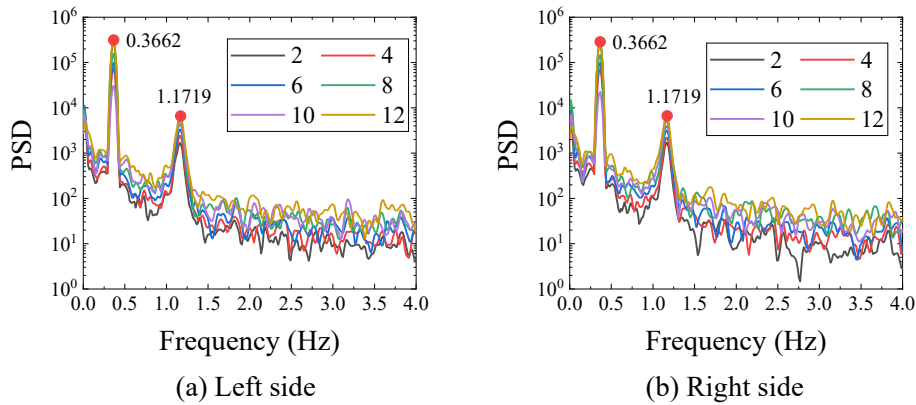
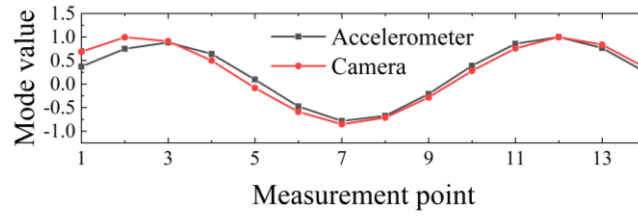
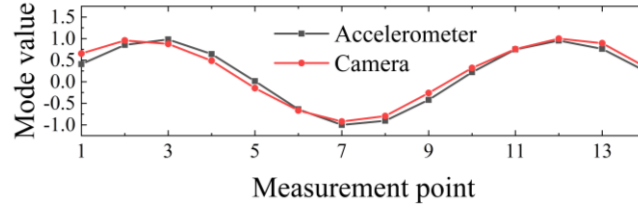


Figure 5.45 PSD of the estimated displacements



(a) Mode shape magnitudes of the left girder



(b) Mode shape magnitudes of the right girder

Figure 5.46 Normalized mode shape magnitudes estimated from the video and the accelerometers (0.3662 Hz)

5.7 Summary

This chapter addresses two critical factors in multi-point displacement estimation for large-scale structures: multi-scale displacement amplitudes and multi-type targets. It extends the digital image phase-based displacement estimation methods developed in Chapters 3 and 4 to achieve adaptive full-field displacement monitoring for large-scale structures.

To tackle the challenge of estimating full-field displacements with varying amplitudes across time and space in large-scale structures, a multi-resolution complex pyramid based on complex quadrature filters is established. This framework enables adaptive multi-point displacement estimation that balances both measurement range and accuracy. For the challenge of displacement estimation for multi-type targets in large-scale structures, an adaptive method using discriminative correlation filters is proposed to extract displacements from diverse structural features.

The proposed adaptive full-field displacement estimation method, integrating multi-resolution complex pyramid and discriminative filtering, has been validated through experimental tests and field measurements on real large-scale structures.

CHAPTER 6

FULL-FIELD DISPLACEMENT RECONSTRUCTION BY FUSING LIMITED DIGITAL IMAGE PHASE AND ACCELERATION MEASUREMENTS

6.1 Introduction

The adaptive full-field displacement estimation method for large-scale structures proposed in Chapter 5 of this study is applicable when all measurement points can be captured by cameras. However, in real-world displacement monitoring scenarios for large-scale structures, spatial partitions (e.g., floor slabs of high rises) may prevent direct observation of multi-point displacements using methods such as cameras, LDV, or GNSS (Yi et al., 2013). Therefore, it is necessary to integrate limited phase-based displacement data with measurements from other sensors to reconstruct full-field displacements of large-scale structures.

Response reconstruction methods, which combine structural mechanical models with monitoring data from limited measurement points, have been widely used to reconstruct responses across all degrees of freedom in large-scale structural models (Zhu et al., 2013, 2023a, 2023b). However, external loads such as typhoons acting on large-scale structures contain significant low-frequency components. Low-frequency accelerations measured by accelerometers are often subtle and susceptible to local environmental interference, leading to drifts in the low-frequency part of displacement reconstruction. Meanwhile, in displacement measurement, GNSS lacks sufficient dynamic accuracy at high frequencies, while non-contact methods such as LDV and digital image phase-based displacement estimation exhibit relatively high-frequency noise due to environmental factors. These issues indicate that reliable data from different sensors are distributed across different

frequency bands, making it difficult to determine the reliability of each sensor's data in response reconstruction. Therefore, it is essential to fuse the reliable frequency bands of different sensors to reconstruct full-field displacements of large-scale structures.

To address the challenge of obtaining full-field displacements in large-scale structures with spatial partitions, this chapter establishes a full-field displacement reconstruction method that integrates limited phase-based displacement data, acceleration measurements, and structural mechanical models. Based on limited phase-based displacement data, a quasi-static displacement reconstruction method is proposed using pseudo-inverse operation and truncated singular value decomposition of the flexibility matrix, enabling the reconstruction of quasi-static displacements below the structural fundamental frequency. Based on limited acceleration data, a modal dynamic displacement reconstruction method is proposed that uses stochastic subspace identification to extract modal dynamic displacement components at and above the structural fundamental frequency. Finally, by fusing limited phase and acceleration data, full-field displacement reconstruction for large-scale structures is achieved.

6.2 Full-field displacement reconstruction theory fusing limited phase-based displacements and accelerations

6.2.1 Pseudo-inverse-based quasi-static displacement reconstruction

Under typhoon loads, large-scale structures exhibit non-negligible low-frequency quasi-static displacement responses due to the low-frequency components of the wind. Yi et al. (2018) reported that a 260-meter-high building in Hong Kong experienced low-frequency quasi-static displacements of up to 5 cm during a typhoon, which was five times the peak amplitude of its fundamental-frequency vibrations. However, the accelerations corresponding to these low-frequency quasi-static displacements are extremely small and difficult to measure directly with accelerometers.

The phase-based displacement estimation method proposed in this study, along with GNSS systems, can capture displacements that include low-frequency quasi-static components. However, accurately measuring relatively high-frequency modal dynamic displacements remains challenging due to extreme environmental interference and the limited dynamic accuracy of GNSS. Therefore, the reconstruction of low-frequency quasi-static and modal dynamic displacements is separated for large-scale structures. To achieve this, the dynamic equilibrium equation of a large-scale structure with n degrees of freedom is first established as follows:

$$\mathbf{m}\ddot{\mathbf{u}}(t) + \mathbf{c}\dot{\mathbf{u}}(t) + \mathbf{k}\mathbf{u}(t) = \mathbf{W}(t) \quad (6.1)$$

where $\mathbf{m}$ — Mass matrix, $\mathbf{m} \in \mathbb{R}^{n \times n}$;
 $\mathbf{k}$ — Stiffness matrix, $\mathbf{k} \in \mathbb{R}^{n \times n}$;
 $\mathbf{c}$ — Damping matrix, $\mathbf{c} \in \mathbb{R}^{n \times n}$;
 $\mathbf{u}(t)$ — Displacement vector, $\mathbf{u} \in \mathbb{R}^{n \times 1}$;
 $\mathbf{W}(t)$ — Load vector, $\mathbf{W} \in \mathbb{R}^{n \times 1}$.

The displacement vector is separated from the velocity and acceleration:

$$\mathbf{u}(t) = \mathbf{D}[\mathbf{W}(t) - \mathbf{m}\ddot{\mathbf{u}}(t) - \mathbf{c}\dot{\mathbf{u}}(t)] = \mathbf{D}\tilde{\mathbf{W}}(t) \quad (6.2)$$

where $\tilde{\mathbf{W}}(t)$ — Equivalent load vector;
 $\mathbf{D}$ — Flexibility matrix.

In Equation (6.2), the load acting on the structure is integrated with inertial and damping forces into an equivalent load. When only the low-frequency quasi-static response of the structure is considered, the acceleration and velocity of the structure are negligible. Under this assumption, the equivalent load can be regarded as representative of the load at each degree of freedom of the structure (Choi et al., 2017; Li et al., 2023), that is:

$$\mathbf{u}_{n \times 1}(t) = \mathbf{D}_{n \times n} \mathbf{W}_{n \times 1}(t) = \mathbf{D}_{n \times n} \tilde{\mathbf{W}}_{n \times 1}(t) \quad (6.3)$$

The degree of freedom n is generally larger than the number of sensors m . The elements without measurement in the displacement vector are eliminated in Equation (6.4):

$$\mathbf{u}_{m \times 1}(t) = \mathbf{D}_{m \times n} \tilde{\mathbf{W}}_{n \times 1}(t) \quad (6.4)$$

To solve the above equation, it is necessary to address the underdetermined problem of estimating the full-field loads from limited measurement points. By introducing the pseudo-inverse operator, the general solution for the full-field loads is constructed as follows (Li, 2021):

$$\tilde{\mathbf{W}}_{n \times 1}(t) = \mathbf{D}_{m \times n}^+ \mathbf{u}_{m \times 1}(t) + (\mathbf{I} - \mathbf{D}_{m \times n}^+ \mathbf{D}_{m \times n}) \mathbf{Y}_{n \times 1} \quad (6.5)$$

where $\mathbf{D}_{m \times n}^+$ — Pseudo inverse of the flexibility matrix;
 $\mathbf{Y}_{n \times 1}$ — Error term.

The equivalent loads calculated via Equation (6.5) satisfy the constraints imposed by the measured displacements. When the number of measurement points equals the number of degrees of freedom, the error term in Equation (6.5) becomes zero. Since the flexibility matrix at the measurement degrees of freedom is non-positive definite, its pseudo inverse relies on the SVD of the non-positive definite flexibility matrix:

$$\mathbf{D}_{m \times n} = \mathbf{Q}_1 \mathbf{\Sigma} \mathbf{Q}_2^T \quad (6.6)$$

$$\mathbf{D}_{m \times n}^+ = \mathbf{Q}_2 \mathbf{\Sigma}^+ \mathbf{Q}_1^T \quad (6.7)$$

where $\mathbf{Q}_1$ — Left positive definite matrix, $\mathbf{Q}_1 \in \mathbb{R}^{m \times m}$;
 $\mathbf{Q}_2$ — Right positive definite matrix, $\mathbf{Q}_2 \in \mathbb{R}^{n \times n}$;
 $\mathbf{\Sigma}$ — Non-positive definite singular value matrix, $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$.

The singular values of the structural flexibility matrix and its decomposition matrices are strongly correlated with the natural frequencies and mode shapes of the structure (Bart Peeters, 2000). Therefore, the physical interpretation of the pseudo inverse is to consider the contributions of the principal deformation modes truncated to the r^{th} eigenvalue ($r \leq m$). When computing the full-field equivalent loads, it is constrained by the displacement responses observed at the measurement points as follows:

$$\tilde{\mathbf{D}}_{m \times n} = \mathbf{q}_1^1 \sigma_1 \mathbf{q}_2^1 + \mathbf{q}_1^2 \sigma_2 \mathbf{q}_2^2 + \cdots + \mathbf{q}_1^r \sigma_r \mathbf{q}_2^r \quad (6.8)$$

where $\tilde{\mathbf{D}}_{m \times n}$ — Truncated flexibility matrix with a rank number of r ;

$\mathbf{q}_1^r$ —— r^{th} column of $\mathbf{Q}_1$, $\mathbf{q}_1^r \in \mathbb{R}^{m \times 1}$;
 $\mathbf{q}_2^r$ —— r^{th} row of $\mathbf{Q}_2$, $\mathbf{q}_2^r \in \mathbb{R}^{1 \times n}$;
 σ_r —— r^{th} diagonal element of Σ .

The truncated flexibility matrix in Equation (6.8) is introduced into Equation (6.4), deriving the full-field quasi-static displacements according to the limited measured data.

$$\mathbf{u}_{n \times 1}(t) = \mathbf{D}_{n \times n} \tilde{\mathbf{D}}_{m \times n}^+ \mathbf{u}_{m \times 1}(t) \quad (6.9)$$

6.2.2 Modal decomposition-based dynamic modal displacement reconstruction

The quasi-static displacement is reconstructed by the pseudo inverse, in which the modal dynamic displacement with a frequency band higher than the structure's fundamental frequency is neglected. Therefore, the modal dynamic displacement is modal decomposed and reconstructed by introducing the state-space equation:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{W}_k + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{W}_k + \mathbf{v}_k \end{cases} \quad (6.10)$$

$$\mathbf{x}_k = [\mathbf{u}(k\Delta t), \dot{\mathbf{u}}(k\Delta t)]^T \quad (6.11)$$

where $\mathbf{x}_k$ —— State vector, $\mathbf{x}_k \in \mathbb{R}^{2n \times 1}$;
 $\mathbf{y}_k$ —— Output vector, $\mathbf{y}_k \in \mathbb{R}^{m \times 1}$;
 $\mathbf{W}_k$ —— Input vector, $\mathbf{W}_k \in \mathbb{R}^{n \times 1}$;
 $\mathbf{u}(\cdot)$ —— Displacements;
 k —— Time step;
 Δt —— Sampling period;
 $\mathbf{A}$ —— System matrix, $\mathbf{A} \in \mathbb{R}^{2n \times 2n}$;
 $\mathbf{B}$ —— Controllability matrix, $\mathbf{B} \in \mathbb{R}^{2n \times n}$;
 $\mathbf{C}$ —— Observability matrix, $\mathbf{C} \in \mathbb{R}^{m \times 2n}$;
 $\mathbf{D}_I$ —— Input feedthrough matrix, $\mathbf{D}_I \in \mathbb{R}^{m \times n}$;
 $\mathbf{w}_k$ —— System error, $\mathbf{w}_k \in \mathbb{R}^{2n \times 1}$;
 $\mathbf{v}_k$ —— Measurement noises, $\mathbf{v}_k \in \mathbb{R}^{m \times 1}$.

In Equation (6.10), the eigenvalues and eigenvectors of the state matrix contain information about the natural frequencies and mode shapes of the structure, respectively. To solve for the state matrix, the effects of the unknown input vector and noise must be eliminated. Since the external loads acting on the structure are difficult to measure directly, they are assumed to be white noise and merged into the noise term. Using past outputs as an instrumental matrix, a data-driven stochastic subspace model is established as follows:

$$\mathbf{P}_i^{\text{ref}} = \mathbf{O}_i \hat{\mathbf{X}}_i = \mathbf{Y}_F \mathbf{Y}_P^T (\mathbf{Y}_P \mathbf{Y}_P^T)^{-1} \mathbf{Y}_P \quad (6.12)$$

$$\mathbf{O}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} \quad (6.13)$$

$$\hat{\mathbf{X}}_i = [\hat{\mathbf{x}}_i \quad \hat{\mathbf{x}}_{i+1} \quad \cdots \quad \hat{\mathbf{x}}_{i+L-1}] \quad (6.14)$$

$$\mathbf{Y}_F = \frac{1}{\sqrt{L}} \begin{bmatrix} \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{i+L-1} \\ \mathbf{y}_{i+1} & \mathbf{y}_{i+2} & \cdots & \mathbf{y}_{i+L} \\ \vdots & & & \\ \mathbf{y}_{2i-1} & \mathbf{y}_{2i} & \cdots & \mathbf{y}_{2i+L-2} \end{bmatrix} \quad (6.15)$$

$$\mathbf{Y}_P = \frac{1}{\sqrt{L}} \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{L-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_L \\ \vdots & & & \\ \mathbf{y}_{i-1} & \mathbf{y}_i & \cdots & \mathbf{y}_{i+L-2} \end{bmatrix} \quad (6.16)$$

where $\mathbf{Y}_F$ — Stacked future output matrix;
 $\mathbf{Y}_P$ — Stacked past output matrix;
 $\mathbf{P}_i^{\text{ref}}$ — Projection of the future output matrix on the past output matrix;
 $\mathbf{O}_i$ — Extended observability matrix;
 $\hat{\mathbf{X}}_i$ — Stacked Kalman state matrix;
 $\hat{\mathbf{x}}_i$ — Kalman state vector;
 i — Time lag;
 L — Number of columns of the assembled output matrix.

It can be observed that the extended observability matrix in Equation (6.13) contains the system matrix and the observability matrix, both related to modal parameters. By performing SVD on the projection of the stacked future output matrix onto the past output matrix, the extended observability matrix and the stacked Kalman state matrix can be obtained:

$$\mathbf{P}_i^{\text{ref}} = \mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^T \quad (6.17)$$

$$\mathbf{O}_i = \mathbf{U}_1 \mathbf{S}_1^{1/2} \mathbf{I} \quad (6.18)$$

$$\hat{\mathbf{X}}_i = \mathbf{O}_i^+ \mathbf{P}_i^{\text{ref}} \quad (6.19)$$

where $\mathbf{U}_1$ ——Left matrix of SVD;
 $\mathbf{V}_1$ ——Right matrix of SVD;
 $\mathbf{S}_1$ ——Singular value matrix.

To solve for the system matrix and the observability matrix, an additional row representing the next time step is appended to the stacked past output matrix, while the first row is removed from the stacked future output matrix to obtain the stacked Kalman state matrix at time $i+1$:

$$\hat{\mathbf{X}}_{i+1} = \mathbf{O}_{i-1}^+ \mathbf{P}_{i-1}^{\text{ref}} \quad (6.20)$$

$$\mathbf{O}_{i-1} = \mathbf{O}_i [1:(i-1),:] \quad (6.21)$$

$$\mathbf{P}_{i-1}^{\text{ref}} = \mathbf{Y}_{\text{F2}} \mathbf{Y}_{\text{P2}}^T (\mathbf{Y}_{\text{P2}} \mathbf{Y}_{\text{P2}}^T)^{-1} \mathbf{Y}_{\text{P2}} \quad (6.22)$$

$$\mathbf{Y}_{\text{F2}} = \frac{1}{\sqrt{L}} \begin{bmatrix} \mathbf{y}_{i+1} & \mathbf{y}_{i+2} & \mathbf{y}_{i+L} \\ \mathbf{y}_{i+2} & \mathbf{y}_{i+3} & \mathbf{y}_{i+L+1} \\ \mathbf{y}_{2i-1} & \mathbf{y}_{2i} & \mathbf{y}_{2i+L-2} \end{bmatrix} \quad (6.23)$$

$$\mathbf{Y}_{\text{P2}} = \frac{1}{\sqrt{L}} \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{L-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_L \\ \vdots & & & \\ \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{i+L-1} \end{bmatrix} \quad (6.24)$$

By combining the output matrix at time i with the stacked Kalman state matrix at time $i+1$, and considering that neither is correlated with the noise, the system matrix and the

observability matrix can be solved via least squares from Equation (6.25). The covariance of the noise is then calculated from the least squares residuals:

$$\begin{bmatrix} A \\ C \end{bmatrix} = \begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} \hat{X}_i^+ \quad (6.25)$$

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} = \begin{bmatrix} V_{1i} \\ V_{2i} \end{bmatrix} \begin{bmatrix} V_{1i}^T & V_{2i}^T \end{bmatrix} \quad (6.26)$$

$$Y_{i|i} = [y_i \quad y_{i+1} \quad \cdots \quad y_{i+L-1}] \quad (6.27)$$

where V_{1i} — Residuals induced by system errors;
 V_{2i} — Residuals induced by measurement noises;
 Q — Covariance matrix of noises;
 S — Covariance matrix of noises;
 R — Covariance matrix of noises.

The mode shapes and natural frequencies can be extracted from the system matrix and observability matrix:

$$A = \Psi \Lambda_d \Psi^{-1} \quad (6.28)$$

$$V = C \Psi \quad (6.29)$$

$$\Lambda_d = e^{\Lambda_c \Delta t} \quad (6.30)$$

where Ψ — Mode shapes;
 Λ_d — Natural frequencies in the discrete time domain;
 V — Observed mode shapes;
 Λ_c — Natural frequencies in the continuous time domain.

To decompose the dynamic responses obtained from accelerometers into modal components and integrate them into displacements, it is necessary to compute the contribution of each modal component. Accordingly, a forward innovation model in the modal coordinate is introduced as follows:

$$\begin{cases} x_{m,k+1} = (\Lambda_d - K_m V) x_{m,k} + K_m y_k \\ e_k = y_k - V x_{m,k} \end{cases} \quad (6.31)$$

$$\Psi^{-1} x_k = x_{m,k} \quad (6.32)$$

$$\Psi^{-1} K = K_m \quad (6.33)$$

where $\mathbf{x}_{m,k}$ — State vector in the modal coordinate;
 $\mathbf{K}_m$ — Kalman gain in the modal coordinate;
 $\mathbf{K}$ — Kalman gain;
 $\mathbf{e}_k$ — Innovation.

The Riccati equation is used to solve the Kalman gain:

$$\mathbf{P} = \mathbf{A}\mathbf{P}\mathbf{A}^T + (\mathbf{G} - \mathbf{A}\mathbf{P}\mathbf{C}^T)(\mathbf{R}_0 - \mathbf{C}\mathbf{P}\mathbf{C}^T)^{-1}(\mathbf{G} - \mathbf{A}\mathbf{P}\mathbf{C}^T)^T \quad (6.34)$$

$$\mathbf{K} = (\mathbf{G} - \mathbf{A}\mathbf{P}\mathbf{C}^T)\mathbf{R}_e^{-1} \quad (6.35)$$

$$\mathbf{R}_e = \mathbf{R}_0 - \mathbf{C}\mathbf{P}\mathbf{C}^T \quad (6.36)$$

$$\mathbf{R}_0 = \mathbf{C}\mathbf{\Sigma}\mathbf{C}^T + \mathbf{R} \quad (6.37)$$

$$\mathbf{G} = \mathbf{A}\mathbf{\Sigma}\mathbf{C}^T + \mathbf{S} \quad (6.38)$$

$$\mathbf{\Sigma} = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^T + \mathbf{Q} \quad (6.39)$$

where $\mathbf{P}$ — Positive definite solution for Riccati;
 $\mathbf{G}$ — Covariance between the output and the next time-step state;
 $\mathbf{R}_e$ — Covariance of the innovation;
 $\mathbf{\Sigma}$ — Covariance of the state;
 $\mathbf{R}_0$ — Covariance of the output.

Each element in the modal state vector of Equation (6.31) indicates the contribution of the corresponding mode. The full-field modal dynamic acceleration can be reconstructed by multiplying the mode shape vector by the modal state vector:

$$\hat{\mathbf{y}}_{m,k}^j = \mathbf{v}_j \mathbf{x}_{m,k}^j \quad (6.40)$$

where $\hat{\mathbf{y}}_{m,k}^j$ — j^{th} modal acceleration at time k ;
 $\mathbf{v}_j$ — j^{th} mode shape;
 $\mathbf{x}_{m,k}^j$ — Contribution of j^{th} mode at time k .

The full-field modal acceleration can be integrated in the frequency domain to reconstruct the full-field modal dynamic displacement:

$$\mathbf{u}_d = \sum_j^J \text{IFFT} \left[-\frac{\text{FFT}(\hat{\mathbf{y}}_m^j)}{f^2} \right] \quad (6.41)$$

where $\hat{\mathbf{y}}_m^j$ — j^{th} modal acceleration;
 IFFT($\cdot$) — Inverse Fourier transformer;
 J — Number of modes;
 $\mathbf{u}_d$ — Full-field modal dynamic displacement;
 FFT($\cdot$) — Fourier transformer.

6.2.3 Full-field displacement reconstruction fusing limited phases and accelerations

Under external actions such as typhoons, the displacement response of large-scale structures comprises not only modal dynamic components but also non-negligible low-frequency quasi-static displacements. Therefore, the low-frequency information of digital image phase-based displacements and the high-frequency information of accelerometer data are fused to reconstruct the full-field displacement. This allows the separate reconstruction of multi-point quasi-static displacements and modal dynamic displacements for large-scale structures. By superimposing these two components, the full-field displacement reconstruction of large-scale structures is achieved.

The dynamic and quasi-static components of directly measurable accelerations and displacements are separately extracted. Taking typhoon loading as an example, the quasi-static displacement below the fundamental frequency induced by low-frequency and mean wind components is obtained by applying a low-pass filter to displacements measured by GNSS or digital image phase-based methods, thereby removing modal dynamic responses. On the other hand, the modal dynamic response at and above the fundamental frequency caused by fluctuating wind components is extracted by applying a high-pass filter to measured accelerations:

$$\mathbf{u}_s = \text{IFFT} \left\{ \text{FFT}(\mathbf{u}) \odot \left[1 / \left(1 + \frac{f^{2\chi}}{f_t^{2\chi}} \right) \right] \right\} \quad (6.42)$$

$$\mathbf{a}_d = \mathbf{a} - \text{IFFT} \left\{ \text{FFT}(\mathbf{a}) \odot \left[1 / \left(1 + \frac{f^{2\chi}}{f_t^{2\chi}} \right) \right] \right\} \quad (6.43)$$

where $\mathbf{u}_s$ — Measured quasi-static displacements;
 $\mathbf{a}_d$ — Measured dynamic accelerations;
 $\mathbf{u}$ — Measured displacements;
 $\mathbf{a}$ — Measured accelerations;
 $\odot$ — Dot product;
 χ — Order of the Butterworth filter (set to be 7);
 f_t — Cutoff frequency of quasi-static responses.

The measured quasi-static displacements and dynamic accelerations are introduced into Equations (6.9) and (6.41), respectively, to reconstruct full-field quasi-static and dynamic displacements:

$$\hat{\mathbf{u}}_s = \mathbf{D}_{n \times n} \tilde{\mathbf{D}}_{m_d \times n}^+ \mathbf{u}_s \quad (6.44)$$

$$\hat{\mathbf{u}}_d = \sum_j \text{IFFT} \left[- \frac{\text{FFT}(\hat{\mathbf{a}}_m^j)}{f^2} \right] \quad (6.45)$$

where $\hat{\mathbf{u}}_s$ — Full-field quasi-static displacements, $\hat{\mathbf{u}}_s \in \mathbb{R}^{n \times 1}$;
 m_d — Number of displacement measurement points;
 $\hat{\mathbf{u}}_d$ — Full-field modal dynamic displacements, $\hat{\mathbf{u}}_d \in \mathbb{R}^{n \times 1}$;
 $\hat{\mathbf{a}}_m^j$ — Full-field j^{th} modal acceleration, $\hat{\mathbf{a}}_m^j \in \mathbb{R}^{n \times 1}$.

The full-field displacement is reconstructed by superpositioning the quasi-static and modal dynamic displacements:

$$\hat{\mathbf{u}} = \hat{\mathbf{u}}_s + \hat{\mathbf{u}}_d \quad (6.46)$$

where $\hat{\mathbf{u}}$ — Reconstructed full-field displacements, $\hat{\mathbf{u}} \in \mathbb{R}^{n \times 1}$.

In summary, Figure 6.1 outlines the data-fusion-based full-field displacement reconstruction (DFDR) procedure. The process begins by applying a low-pass filter with a cutoff frequency below the fundamental frequency of the structure to the limited phase-based displacement data. The quasi-static component of the displacement at the

measurement points is extracted. Concurrently, the acceleration data are high-pass filtered to obtain the dynamic component. The filtered measurement data are then input into Equations (6.44) and (6.45) to estimate the quasi-static and modal dynamic components of the full-field displacements using the pseudo-inverse method and the modal decomposition, respectively. Finally, the estimated quasi-static and modal dynamic displacement components are superimposed to reconstruct the full-field displacement of the large-scale structure. For high-rise structures, this approach further enables the calculation of full-field inter-story drift ratios across all floors.

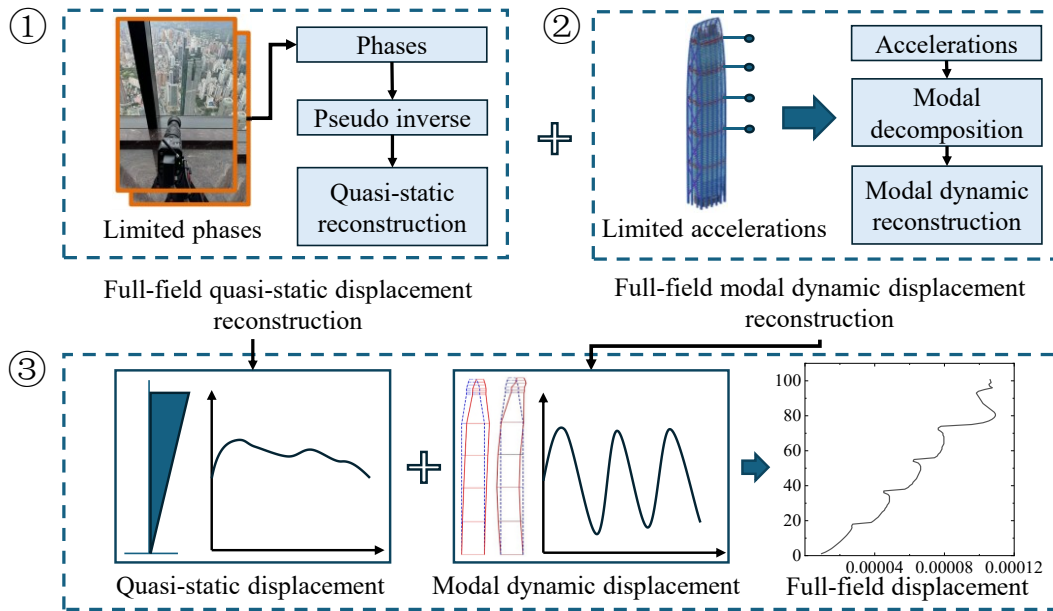


Figure 6.1 Procedure of DFDR

6.3 Performance of DFDR under different numerical cases

For the full-field quasi-static displacement reconstruction in Equation (6.9), the primary influencing factors are the truncation number of singular values in the pseudo inverse of the flexibility matrix and the locations of the measurement points. On the other hand, the elimination of inertial and damping forces in Equation (6.2) relies on the quasi-static response assumption. The influence of the excitation frequency of loads acting on large-scale structures on the reconstruction accuracy of quasi-static displacements requires

further analysis. Therefore, it is necessary to discuss the impact of the number of truncated singular values and measurement point locations on the reconstruction results of structural quasi-static displacements. Besides, for full-field modal dynamic displacement reconstruction, the effects of different input excitations and sensor noise on the performance of full-field displacement reconstruction must be investigated.

6.3.1 Effects of the number of truncated singular values on quasi-static displacement reconstruction

This study investigates the impact of singular value truncation on the reconstruction performance of displacement and inter-story drift ratios for each floor of the high-rise structure under low-frequency and mean wind loads, based on the FEM model of the Kingkey 100 skyscraper (Bian, 2019). The analysis is conducted by applying mean wind loads to the FEM model to obtain the displacements and interstory drift ratios of each floor. The variation of the mean wind load along the building height is determined using an exponential law, with the mean wind speed at 10 m height taken as 25 m/s, corresponding to a 50-year return period:

$$V(z) = V(z_{10}) \left(\frac{z}{z_{10}} \right)^{\alpha} \quad (6.47)$$

where $V(z)$ — Mean wind speed at height z ;
 z_{10} — 10-m height;
 α — Terrain roughness index, 0.22 for category C;
 z — Height (m).

37th, 55th, 74th, 92nd, 96th, 98th, and top floors are used as known response measurement points. Equation (6.9) is applied to reconstruct the quasi-static displacements of all floors. Singular value truncation numbers of the flexibility matrix were set to 7, 5, 3, 2, and 1. The reconstructed results are compared with those computed by the FEM model to verify the accuracy of quasi-static displacement reconstruction. Figure 6.2 presents the reconstructed full-field displacements under different singular value truncation numbers

of the flexibility matrix, compared with FEM. Under the mean wind load with a 50-year return period, the horizontal displacement at the top of the Kingkey 100 skyscraper reaches 0.45 m. Horizontal displacements across floors exhibit an accelerating growth trend along the building height, indicating that interstory displacements in upper floors are larger than those in lower floors. In most cases, the reconstructed quasi-static horizontal displacements under low-frequency and mean wind loads show good agreement with the finite element results. When the singular value truncation number is 1, the horizontal displacements in lower floors are relatively smaller compared to other truncation cases, with minor deviations from the FEM.

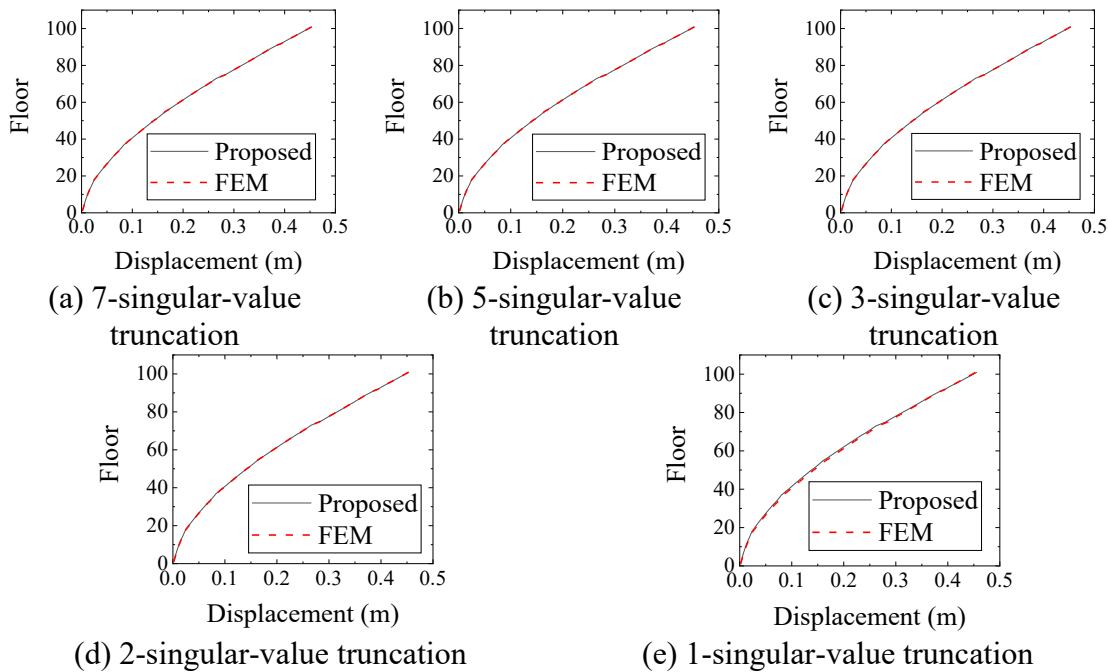


Figure 6.2 Displacements comparison between the reconstructed and FEM results under different numbers of singular values

Table 6.1 lists the relative errors of the reconstructed quasi-static displacements under different singular value truncation numbers. When the truncation number is 2 or higher, the relative errors between the reconstructed full-story displacements and the FEM results are all well below 1%. With a truncation number of 1, the relative error of full-field displacement reconstruction increases notably but remains at only 1.48%. This is attributed to the dominance of a single quasi-static deformation mode under low-

frequency and mean wind loads. Therefore, for the Kingkey 100 skyscraper with 7 displacement measurement points, a singular value truncation number of 2 or higher is sufficient to achieve accurate quasi-static displacement reconstruction.

Table 6.1 Relative errors of the reconstructed quasi-static displacements regarding different singular value numbers

Singular value number	1	2	3	5	7
Relative errors (%)	1.48	0.13	0.02	0.01	0.01

Based on the horizontal displacements and floor heights of each level, the interstory drift ratios are calculated. Figure 6.3 presents a comparison between the reconstructed interstory drift ratios of the Kingkey 100 skyscraper under low-frequency and mean wind loads and the corresponding FEM results. When the singular value truncation number of the structural flexibility matrix ranges from 2 to 7, the reconstructed interstory drift ratios across all floors align closely with the FEM results, with relative errors remaining below 1% (as shown in Table 6.2). However, with a truncation number of 1, the reconstructed interstory drift ratios below the 70th floor are underestimated. To satisfy the displacement constraints at the known measurement points (74th, 92nd, 96th, 98th floors, and the roof) in Equation (6.9), the interstory drift ratios from the 70th floor upward are significantly overestimated to increase the horizontal displacements in these upper levels. Consequently, while the reconstructed full-story horizontal displacements remain consistent with the FEM results due to the constraints from known measurement points, the reconstructed interstory drift ratios exhibit relative errors exceeding 10% when the truncation number is 1.

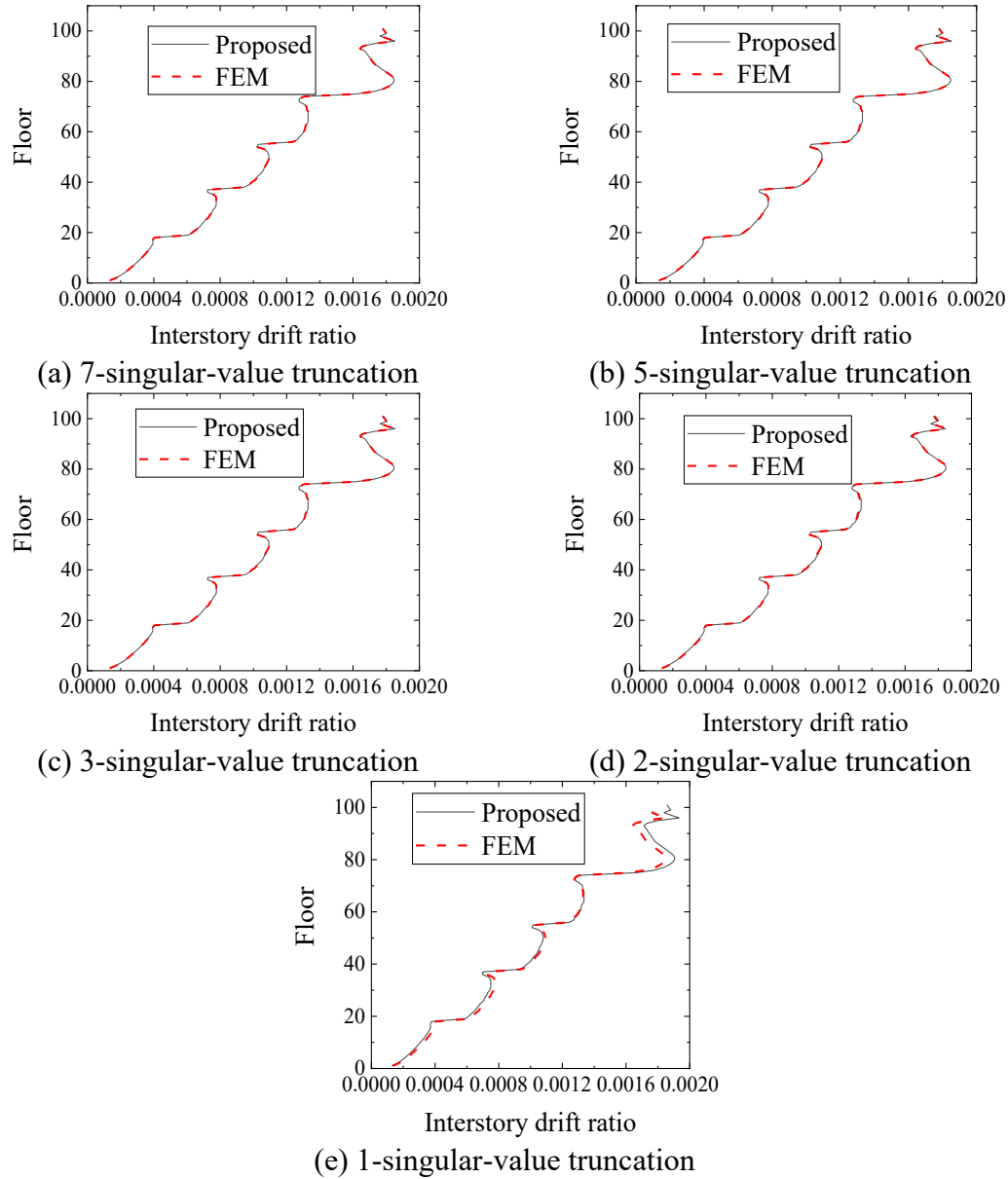


Figure 6.3 Reconstructed interstory drift ratios under different numbers of singular values

Table 6.2 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different singular value numbers

Singular value number	1	2	3	5	7
Relative errors (%)	10.26	0.34	0.09	0.05	0.04

6.3.2 Effects of the measurement points on quasi-static displacement reconstruction

To analyze the influence of the number and arrangement of measurement points on the quasi-static displacement reconstruction of high-rise structures and to determine the optimal placement of digital image phase-based displacement measurement points for the Kingkey 100 skyscraper, Equation (6.9) is applied with different known measurement point configurations to reconstruct the full-story displacements of the structure. The analysis cases with varying numbers of displacement measurement points are illustrated in Figure 6.4.

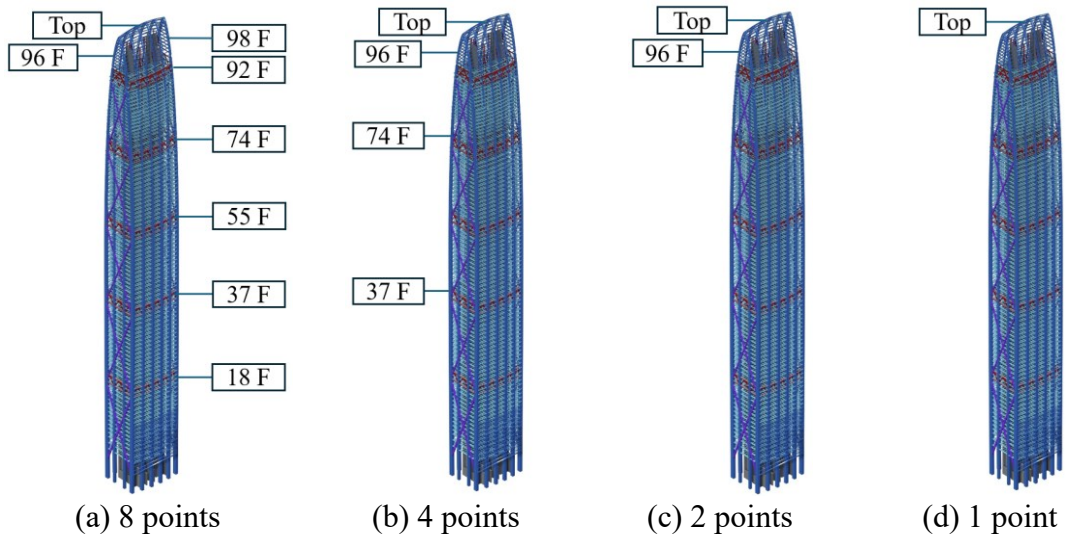


Figure 6.4 Cases with different numbers of measurement points

Displacement measurement points are arranged on both the outrigger trusses and the steel frames, with configurations of 8, 4, 2, and 1 points, to reconstruct the quasi-static displacements under low-frequency and mean wind loads. The singular value truncation number of the structural flexibility matrix is consistently set to 2. The calculated interstory drift ratios for each floor are shown in Figure 6.5. When 8, 4, or 2 measurement points are used, the reconstructed interstory drift ratios across all floors show excellent agreement with the FEM results, with relative errors below 0.5% (as indicated in Table 6.3). However, when only the top-floor displacement measurement point of the Kingkey 100 skyscraper is retained, Equation (6.9) is constrained solely by the single observation at the top. Due to insufficient information on displacement deformation in the mid-level

floors, the reconstructed interstory drift ratios below the 70th floor are underestimated, while those above the 70th floor are overestimated, resulting in a relative error of 10.58% for the interstory drift ratios.

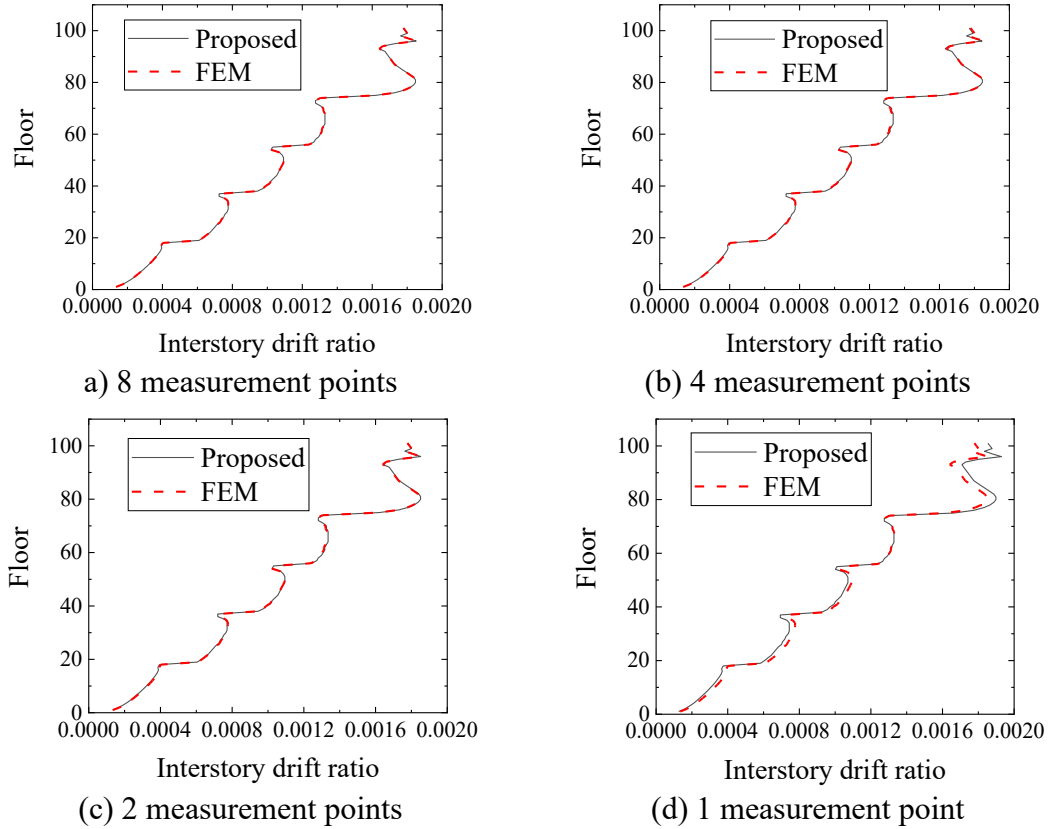


Figure 6.5 Interstory drift ratios comparison between the reconstructed and FEM results under different numbers of measurement points

Table 6.3 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different measurement point numbers

Measurement point number	8	4	2	1
Relative errors (%)	0.27	0.31	0.34	10.58

When two displacement measurement points are retained, the influence of their placement at different floors on the reconstruction accuracy of inter-story drift ratios is investigated. The measurement point configurations are shown in Figure 6.6. In Case 1, the

displacement points are located at the lower floors (18th and 37th floors). As a result, the reconstructed displacements of the upper floors lack constraints from observed data. Figure 6.7 shows that the interstory drift ratios above the 50th floor deviate significantly from the FEM reference values, leading to a relative error of 8.73% (Table 6.4). In Case 2, the measurement points are placed at mid-level floors (55th and 74th floors). However, the reconstructed displacements above the 80th floor remain insufficiently constrained. The relative error between the reconstructed full-story interstory drift ratios and the FEM reference values is 1.68%, which is an improvement over Case 1. In Cases 3 and 4, the displacement points are positioned at upper floors (92nd and 98th floors in Case 3; 96th and top floors in Case 4). It can be observed that when two distinct upper-floor displacement measurements are available to constrain the reconstruction, the resulting interstory drift ratios align closely with the FEM results, with relative errors controlled below 0.5%.

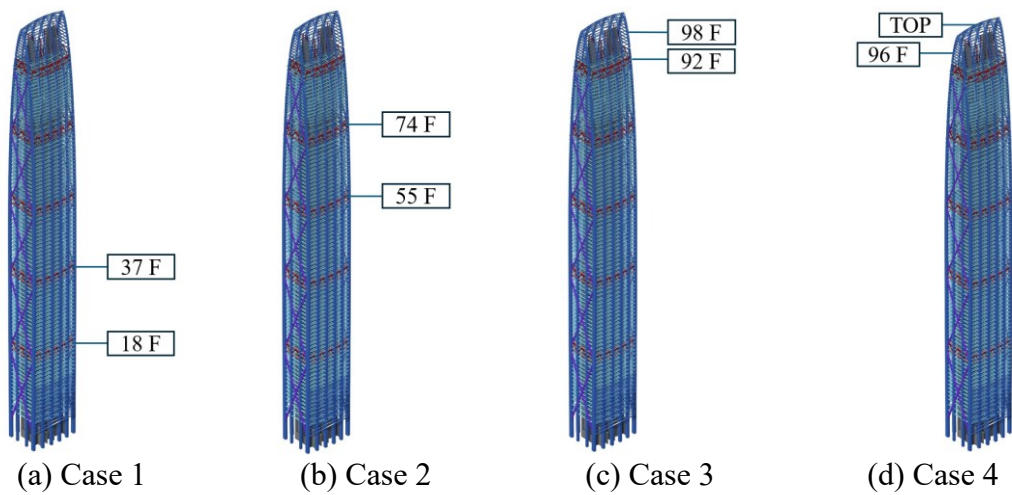
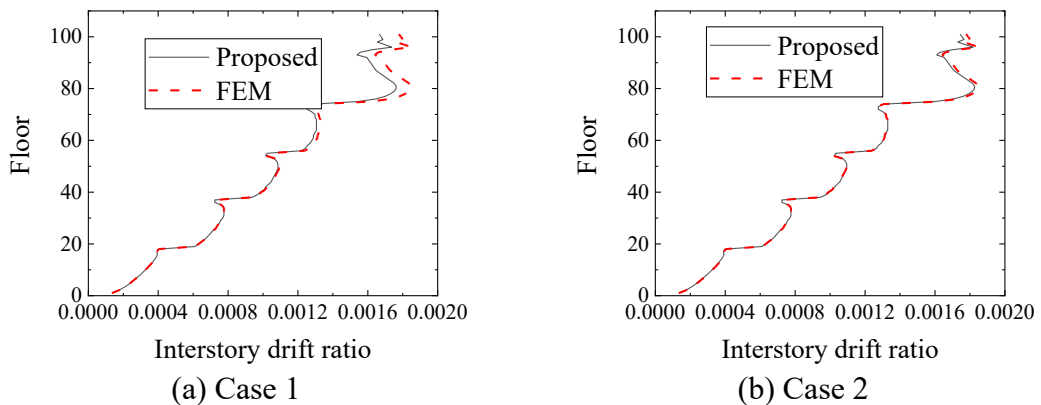
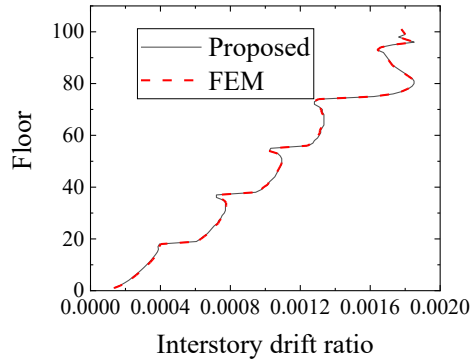
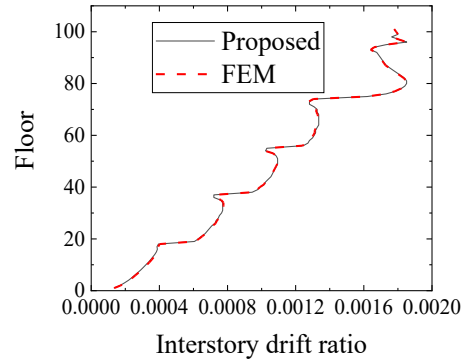


Figure 6.6 Cases with different measurement point placements





(c) Case 3



(d) Case 4

Figure 6.7 Interstory drift ratios comparison between the reconstructed and FEM results under different measurement point placements

Table 6.4 Relative errors of the reconstructed quasi-static interstory drift ratios regarding different measurement point placements

Measurement point placement	Case 1	Case 2	Case 3	Case 4
Relative errors (%)	8.73	1.68	0.36	0.34

6.3.3 Effects of the excitation frequency on quasi-static displacement reconstruction

The assumption of considering only the elastic forces of the structure serves as the prerequisite for estimating full-field displacements using the singular value-truncated flexibility matrix and the constraints from limited displacement measurements. However, the mean wind load of a typhoon is a statistical value based on observed wind speeds; actual low-frequency wind components are not static loads but exhibit slow temporal fluctuations, thus acting as quasi-static loads. To analyze the influence of different load excitation frequencies on the static displacement reconstructed via Equation (6.9), the 50-year return period mean wind load is multiplied by sinusoidal functions of 0.01 Hz, 0.05 Hz, 0.10 Hz, and 0.15 Hz to simulate external loads with different frequencies. These frequencies are selected to be below or near the fundamental frequency of the structure. Using a singular value truncation number of 2 for the structural flexibility matrix and known displacement measurement points at the 96th floor and the roof, the interstory drift ratios across all floors are reconstructed under each excitation frequency. The maximum

interstory drift ratios are then compared with those computed by the Newmark Beta method, as shown in Figure 6.8.

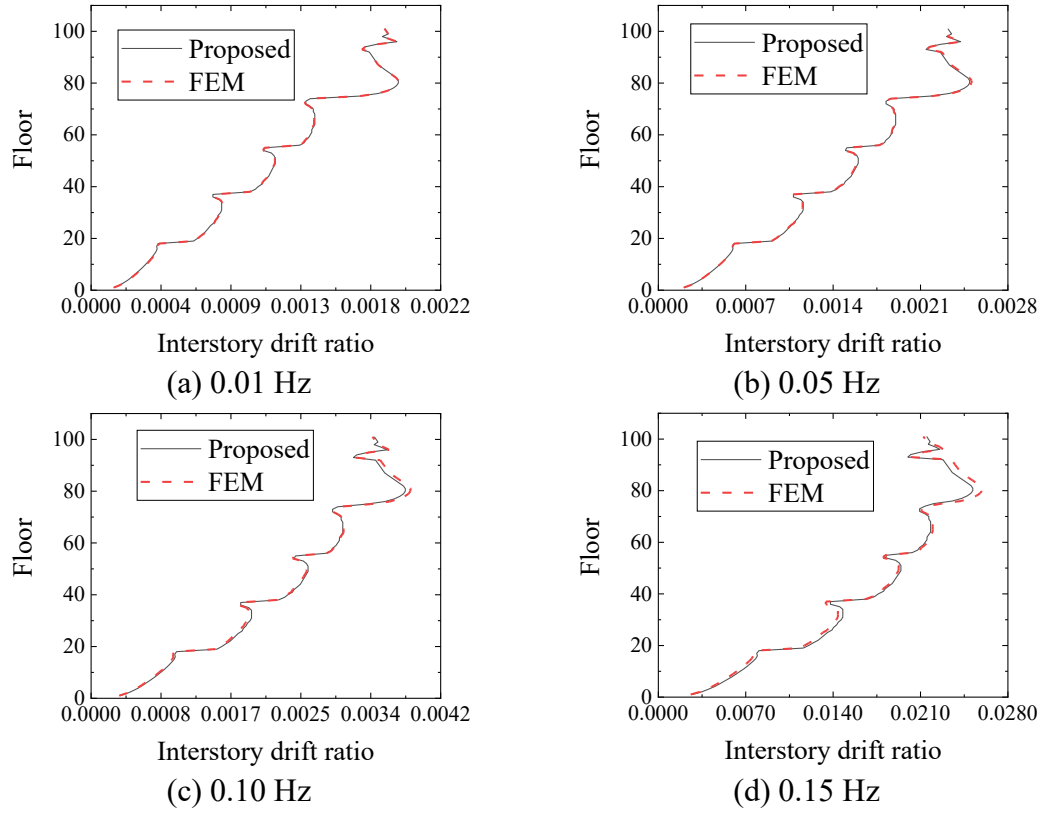


Figure 6.8 Interstory drift ratios comparison between the reconstructed and FEM results under different excitation frequencies

When the excitation frequency is below 0.10 Hz, the amplitude of the reconstructed interstory drift ratios shows good agreement with the FEM results, with relative errors remaining below 1% (as presented in Table 6.5). However, when the excitation frequency reaches 0.15 Hz, which is close to the first bending frequency of the Kingkey 100 structure, the inertial and damping forces of the structure become non-negligible. As a result, the equivalent horizontal load in Equation (6.2) cannot be adequately substituted for the external load, leading to a relative error of 8.90% in the reconstructed full-story interstory drift ratios.

Table 6.5 Relative errors of the reconstructed quasi-static interstory drift ratios
regarding different excitation frequencies

Excitation frequency	0.01 Hz	0.05 Hz	0.10 Hz	0.15 Hz
Relative errors (%)	0.34	0.35	0.95	8.90

6.3.4 Effects of the excitation type on modal dynamic displacement reconstruction

Since the external excitation in Equation (6.10) is assumed as the white noise merged into the noise term, it is necessary to investigate the influence of excitation types with different spectral characteristics on dynamic displacement reconstruction. Therefore, this study uses the white noise and the fluctuating wind as inputs, respectively, based on the FEM model of the Kingkey 100 structure. The dynamic displacements of each floor are reconstructed and compared with the FEM results to validate the effectiveness of the proposed modal dynamic displacement reconstruction method. Based on the previous section, it is concluded that low-frequency displacements below $2/3$ of the structural fundamental frequency (0.1 Hz) can be obtained through the proposed quasi-static displacement reconstruction method. Therefore, this part focuses on the modal dynamic displacements induced by excitations above 0.1 Hz. The 37th, 55th, 74th, 92nd, 96th, 98th, and top floors are selected as known acceleration response measurement points.

Figure 6.9 presents the reconstructed acceleration results of the first four weak-axis bending modal components at the 80th floor under white noise excitation, compared with the FEM results. All acceleration values are normalized by the peak acceleration at the 80th floor from the FEM analysis. It can be observed that the peak acceleration of the first mode at the 80th floor is smaller than that of the second mode, and contributions from higher-order modes are also present, resulting in slightly lower reconstructed accelerations compared to the FEM results.

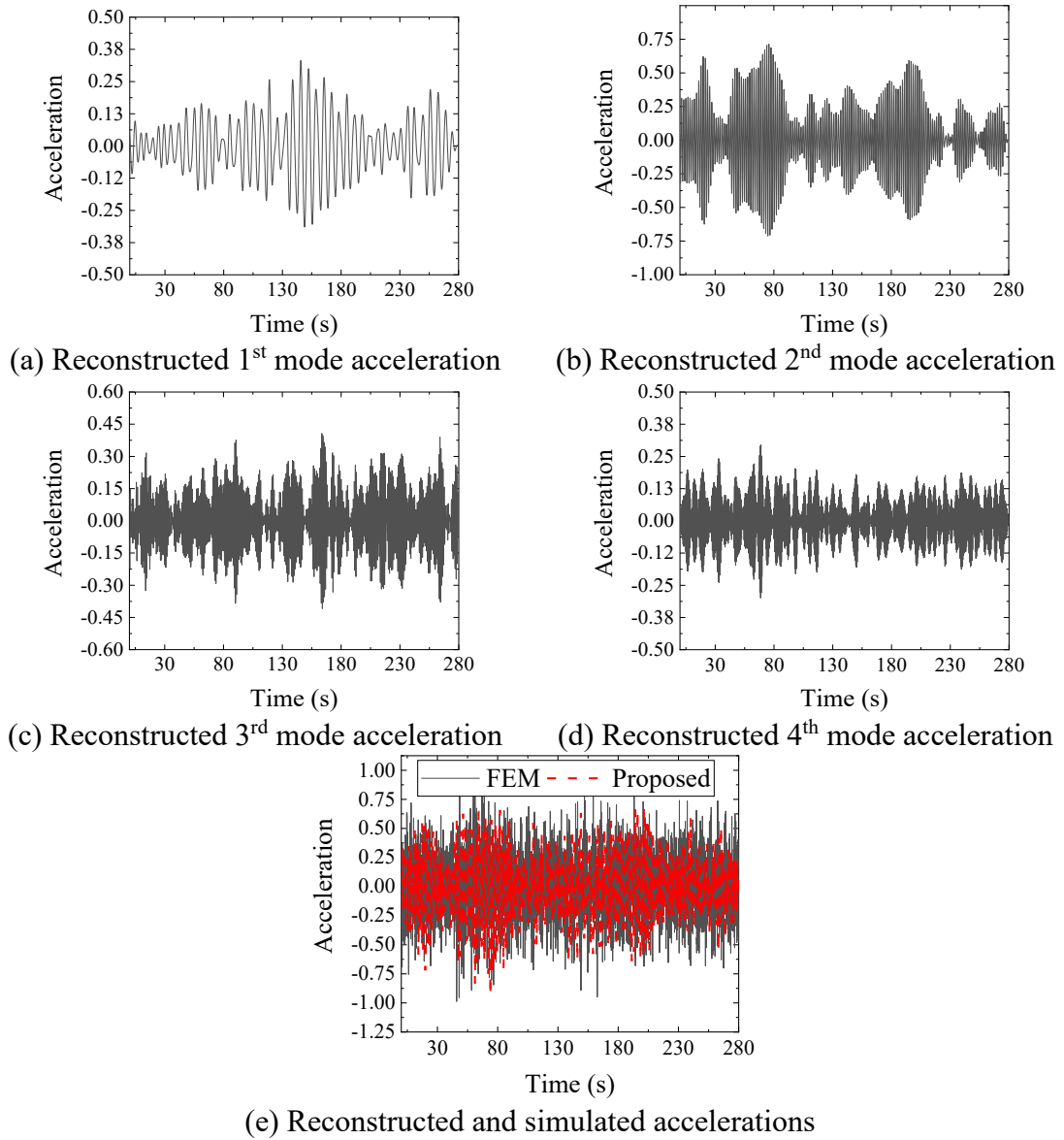


Figure 6.9 Reconstructed and simulated accelerations of the 80th floor under white noise inputs

The reconstructed modal accelerations at the 80th floor of the structure are integrated and superimposed, and the resulting displacements are compared with the FEM displacement results, as shown in Figure 6.10. Compared to acceleration responses, displacement responses are less sensitive to higher-order modes because displacements in the frequency domain are obtained by dividing acceleration by the square of frequency. In the reconstructed displacement response at the 80th floor of the Kingkey 100 structure, the first-mode displacement dominates the response, while the contributions from the second,

third, and fourth modes are significantly reduced. The peak displacement of the fourth mode accounts for only about 0.6% of the total displacement. The displacements reconstructed from each mode are superimposed and compared with the FEM results. The displacements from the first four reconstructed modes align well with the FEM displacements, with a relative error of 5.05%.

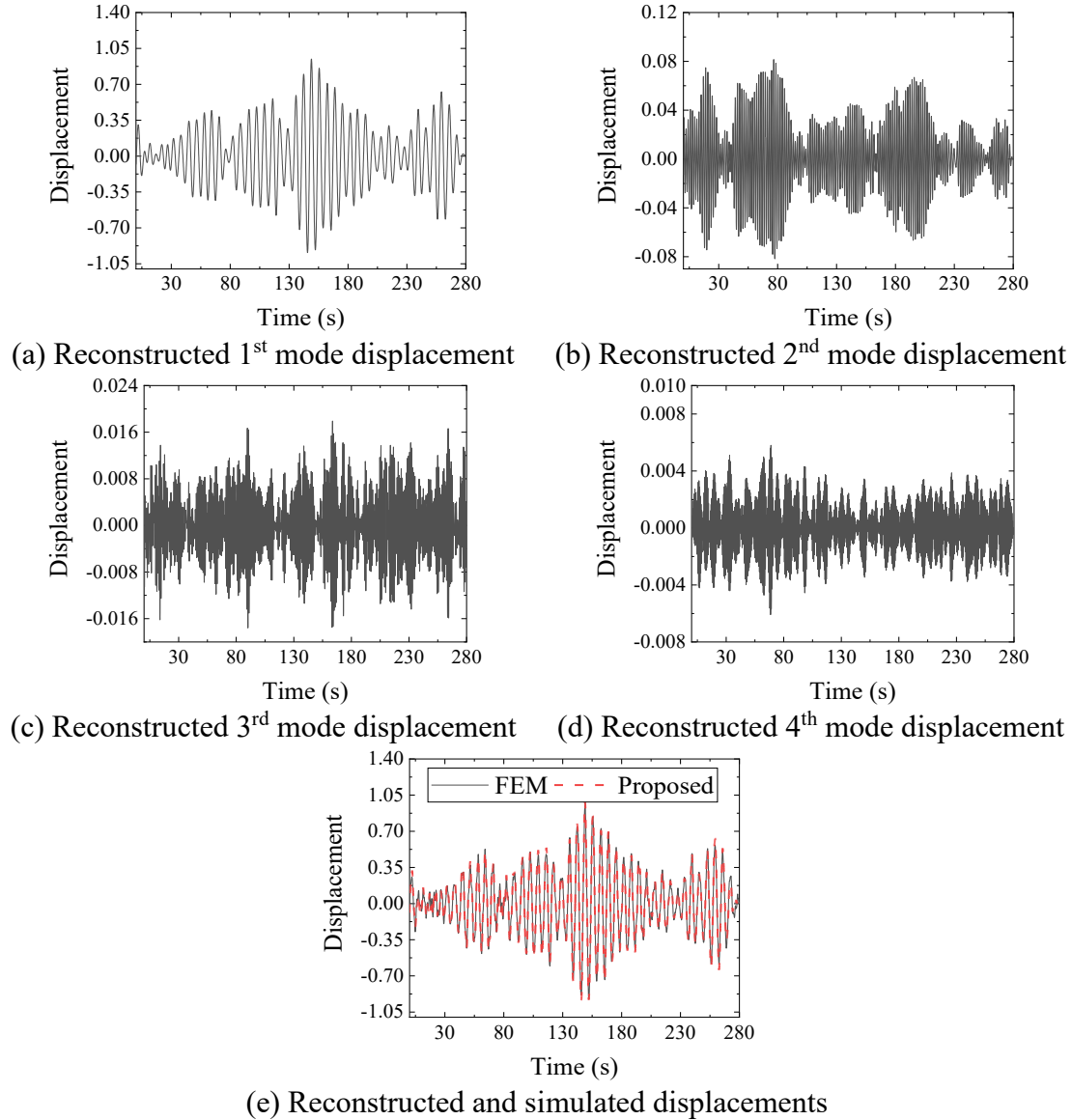
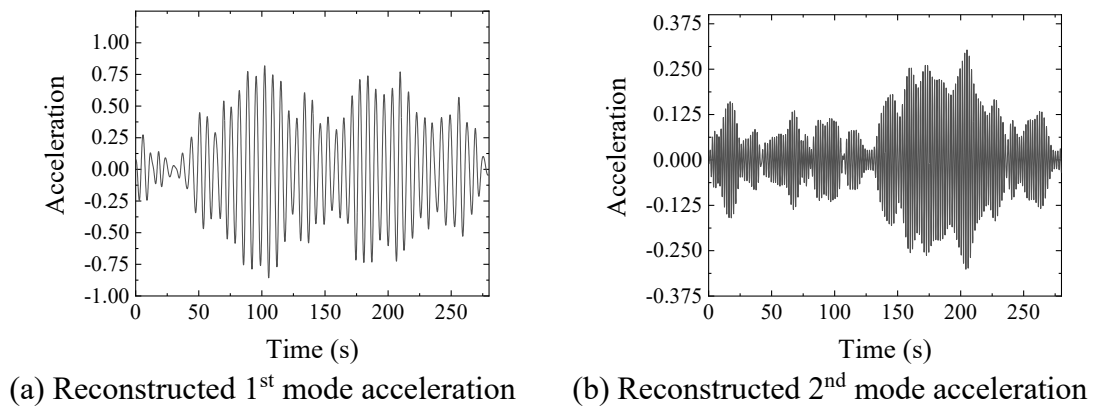


Figure 6.10 Reconstructed and simulated displacements of the 80th floor under white noise

To investigate the dynamic response reconstruction performance under typhoon loads, the fluctuating wind spectrum is used to simulate the dynamic effects of the typhoon. The measured acceleration responses above 0.10 Hz are employed to reconstruct the dynamic responses of all floors. The fluctuating wind is described by the Von Karman spectrum (Fu and Li, 2015; G. Wang et al., 2023). The wind speed spatial correlation regarding different heights is simulated by the Cholesky decomposition (Zhao and Huang, 2020). The fundamental frequency of the Kingkey 100 skyscraper is 0.156 Hz; therefore, responses above 0.10 Hz are treated as dynamic responses for reconstruction. Moreover, since the displacement response of supertall structures under typhoon loads is dominated by low-order modal frequencies, the first four bending modes in the weak-axis direction are considered in Equation (6.41).

Figure 6.11 compares the reconstructed modal accelerations and their superposition at the 80th floor of the structure with the finite element results, normalized by the peak acceleration at the 80th floor from the FEM. It can be observed that under fluctuating wind loads, the acceleration response of the high-rise structure is dominated by the first-mode response, while the second-mode response is also significant, approximately one-third of the first-mode amplitude. The third- and fourth-mode responses are considerably smaller than those of the lower-order modes.



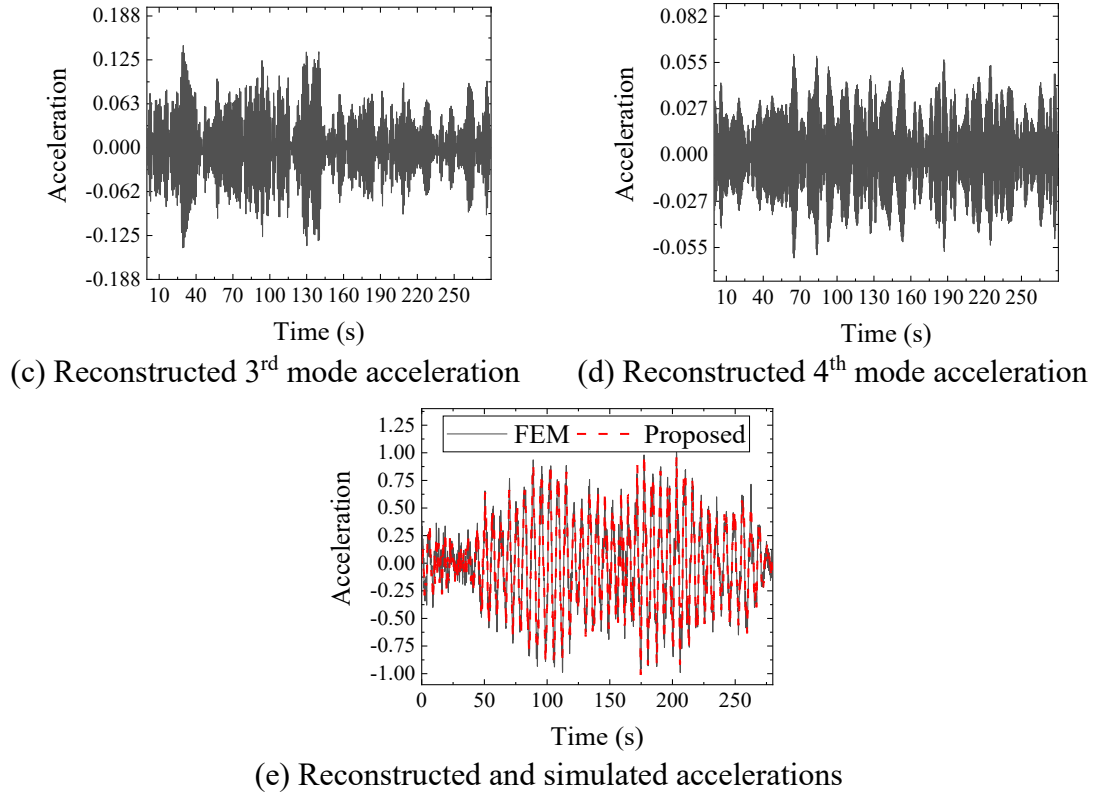


Figure 6.11 Reconstructed and simulated accelerations of the 80th floor under fluctuating wind

After superimposing the accelerations of the first four modes, the reconstructed acceleration at the 80th floor of the structure shows good agreement with the FEM results, with a relative error of about 6.68%. This indicates that under typhoon-induced fluctuating wind loads, the first four modal responses are sufficient to effectively reconstruct the acceleration responses across all floors of the high-rise structure.

The modal dynamic displacement response under fluctuating wind loads is obtained by integrating and superimposing the reconstructed modal acceleration responses. The reconstructed displacement is compared with the FEM displacement results in Figure 6.12, where both sets of displacement values are normalized by the peak displacement at the 80th floor from the FEM. Under typhoon-induced fluctuating wind loads, the displacement of the high-rise structure is dominated by the first-mode response, accounting for approximately 97% of the peak total displacement. The contributions from the second, third, and fourth modes to the displacement response are minimal.

Superimposing the displacements of the first four modes yields a reconstructed displacement response that aligns closely with the FEM displacement results, with a relative error of 3.73%. In summary, under typhoon fluctuating wind loads, the dynamic displacement reconstruction method considering the first four modes can effectively estimate the dynamic displacement responses across all floors, including those without installed sensors.

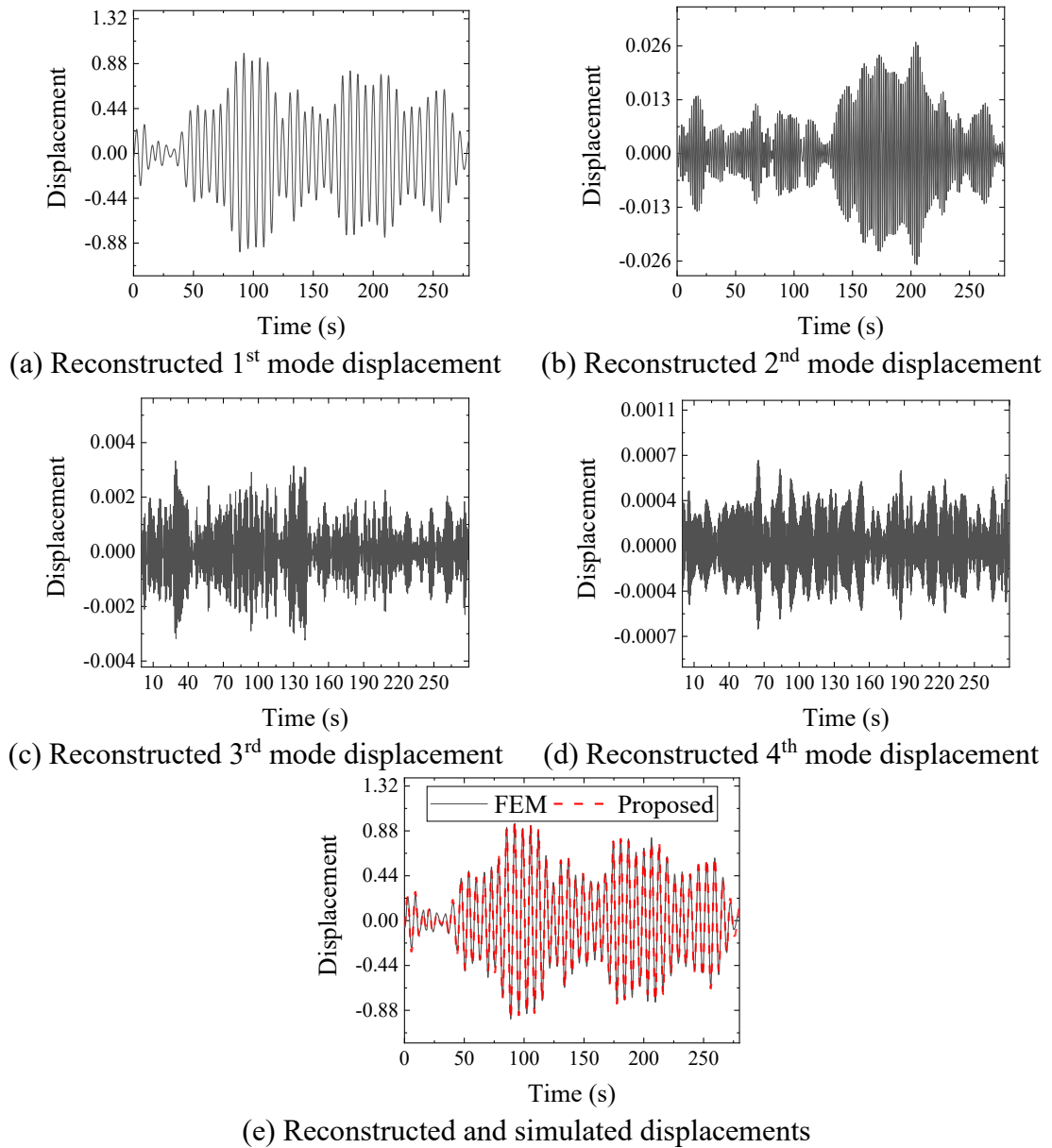


Figure 6.12 Reconstructed and simulated displacements of the 80th floor under fluctuating winds

6.3.5 Effects of the noise on modal dynamic displacement reconstruction

In practical SHM systems for large-scale structures, the environment and the sensor performance introduce unknown noise. Therefore, it is necessary to analyze the impact of noise on the dynamic displacement reconstruction performance for high-rise structures. For this purpose, the 37th, 55th, 74th, 92nd, 96th, 98th, and top floors of the Kingkey 100 structure are selected as known acceleration response measurement points. Different levels of low-frequency, high-frequency, and white noise are superimposed onto the accelerations to investigate the influence of noise intensity and distribution on the displacement reconstruction of unmeasured floors. Under fluctuating wind loads, the displacement response of the Kingkey 100 high-rise structure is dominated by the first four bending modes, with natural frequencies below 2 Hz. Therefore, low-frequency noise is generated by applying a low-pass filter below 1 Hz to white noise, while high-frequency noise is obtained by applying a high-pass filter above 1 Hz to white noise.

The proposed dynamic displacement reconstruction method is applied to estimate the dynamic displacements of unmeasured floors under different levels of low-frequency noise, and the results are normalized by the peak displacement from the FEM analysis. Noise intensity is defined as the ratio of the variance of the noise to the variance of the FEM displacement time history. Noise with intensity ranging from 0.02 to 0.60 is applied to the measured acceleration. Figure 6.13 compares the reconstructed displacements at the 80th floor under different low-frequency noise intensities with the FEM results, with the corresponding relative errors listed in Table 6.6. It can be observed that when the low-frequency noise variance to displacement variance ratio is 0.02 or 0.05, the reconstructed dynamic displacements at the 80th floor align closely with the FEM results, with relative errors below 5%. As the low-frequency noise level increases to 0.40, discrepancies between the reconstructed and the FEM displacement emerge during periods of smaller displacement amplitudes, resulting in an overall relative error of 8.22%. At a low-frequency noise level of 0.60, the noise amplitude approaches the displacement amplitude, further perturbing the reconstructed displacement. Nevertheless, the overall

characteristics of the reconstructed displacement remain consistent with the FEM results, with the relative error controlled at 11.66%. In summary, low-frequency noise has a limited impact on the performance of dynamic displacement reconstruction. Even under an extreme noise level of 0.60, reasonably accurate dynamic displacements can still be reconstructed.

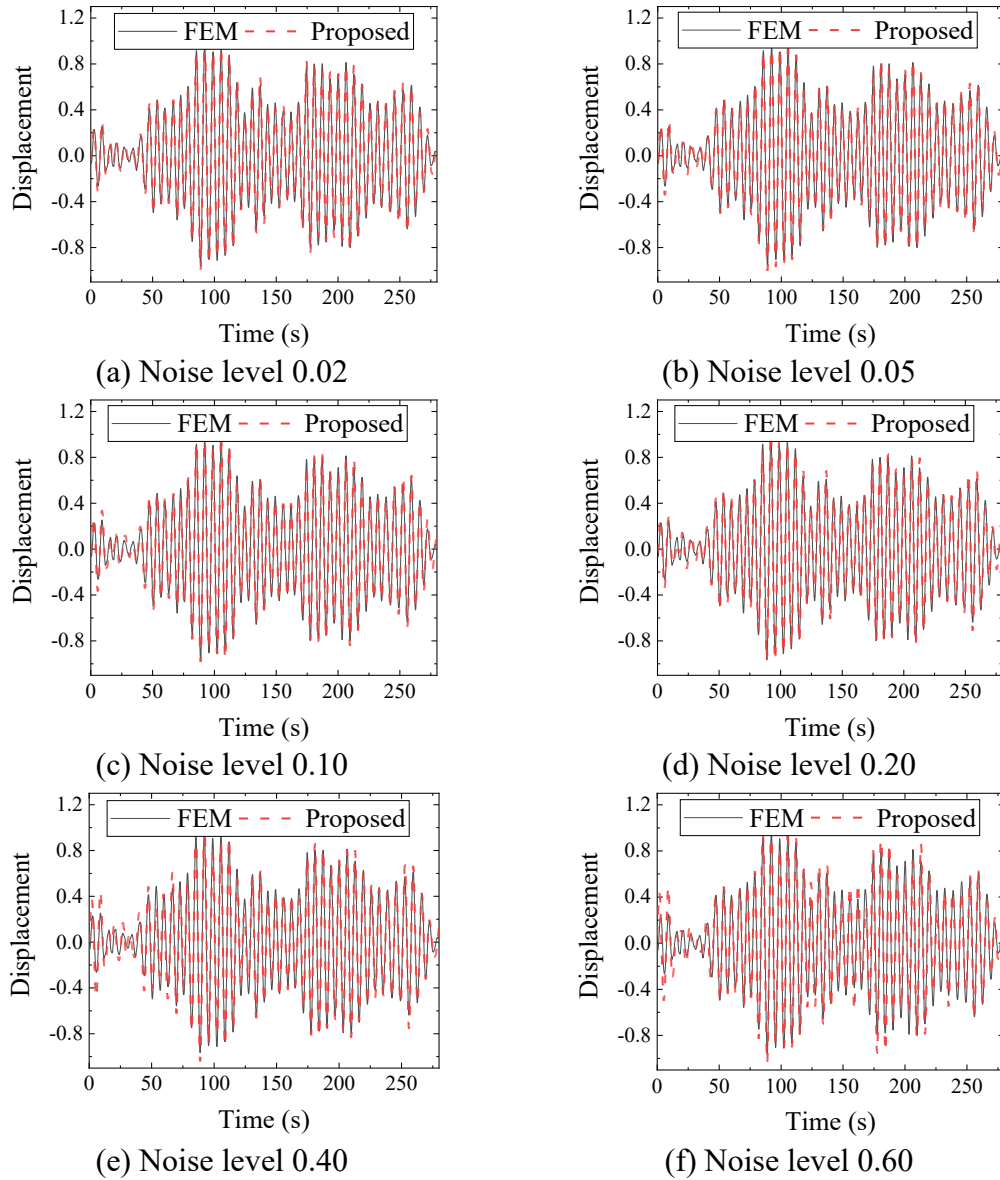
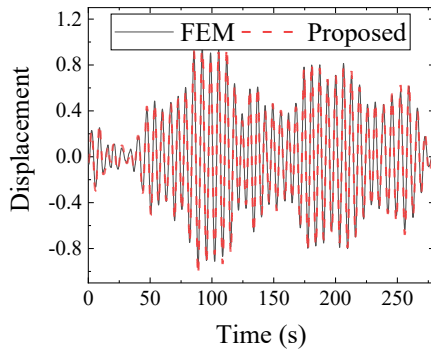


Figure 6.13 Reconstructed and simulated displacements of the 80th floor under different-level low-frequency noises

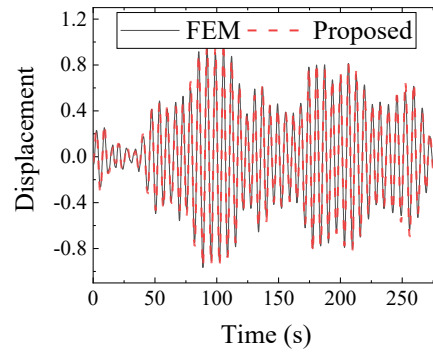
Table 6.6 Relative errors of the reconstructed dynamic displacements regarding different low-frequency noises

Noise level	0.02	0.05	0.10	0.20	0.40	0.60
Relative errors (%)	4.93	4.96	5.34	6.31	8.22	11.66

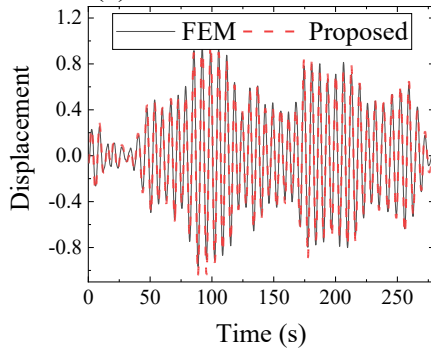
Figure 6.14 compares the reconstructed dynamic displacement at the 80th floor with the FEM results under the influence of high-frequency noise above 1 Hz. As analyzed earlier, under fluctuating wind loads, the displacement response of the high-rise structure is primarily dominated by the first mode. Since the first two bending modal frequencies of the tall structure are both below 1 Hz, the impact of high-frequency noise on dynamic displacement reconstruction is minimal across all tested intensity levels. Table 6.7 indicates that even when the variance of high-frequency noise reaches 0.60 times the variance of the displacement time history, the relative error of the reconstructed dynamic displacement remains approximately 5%.



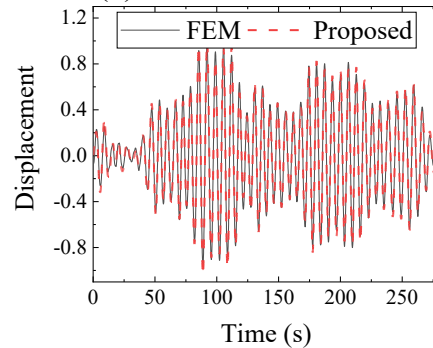
(a) Noise level 0.02



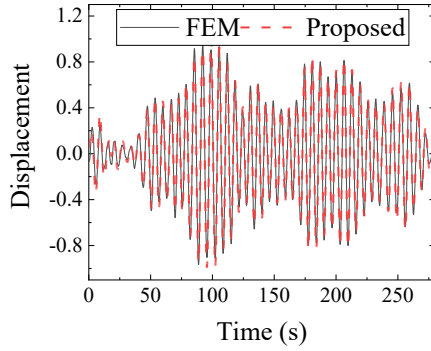
(b) Noise level 0.05



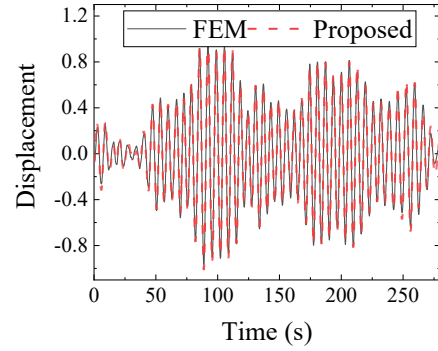
(c) Noise level 0.10



(d) Noise level 0.20



(e) Noise level 0.40



(f) Noise level 0.60

Figure 6.14 Reconstructed and simulated displacements of the 80th floor under different-level high-frequency noises

Table 6.7 Relative errors of the reconstructed dynamic displacements regarding different-level high-frequency noises

Noise level	0.02	0.05	0.10	0.20	0.40	0.60
Relative errors (%)	3.75	3.79	3.98	4.98	5.14	5.32

Under white noise disturbance, the reconstructed displacement is compared with the FEM results in Figure 6.15. Similar to the low-frequency noise case, when the white noise variance is 0.20 or less relative to the displacement variance, the reconstructed dynamic displacement at the 80th floor aligns well with the FEM results in both periodic characteristics and amplitude, with the relative error controlled around 6% (as shown in Table 6.8). When the white noise intensity increases to 0.40 and 0.60, the relative errors between the reconstructed displacement and the FEM values reach 9.56% and 12.31%, respectively.

Table 6.8 Relative errors of the reconstructed dynamic displacements regarding different-level white noises

Noise level	0.02	0.05	0.10	0.20	0.40	0.60
Relative errors (%)	5.01	5.54	6.23	6.62	9.56	12.31

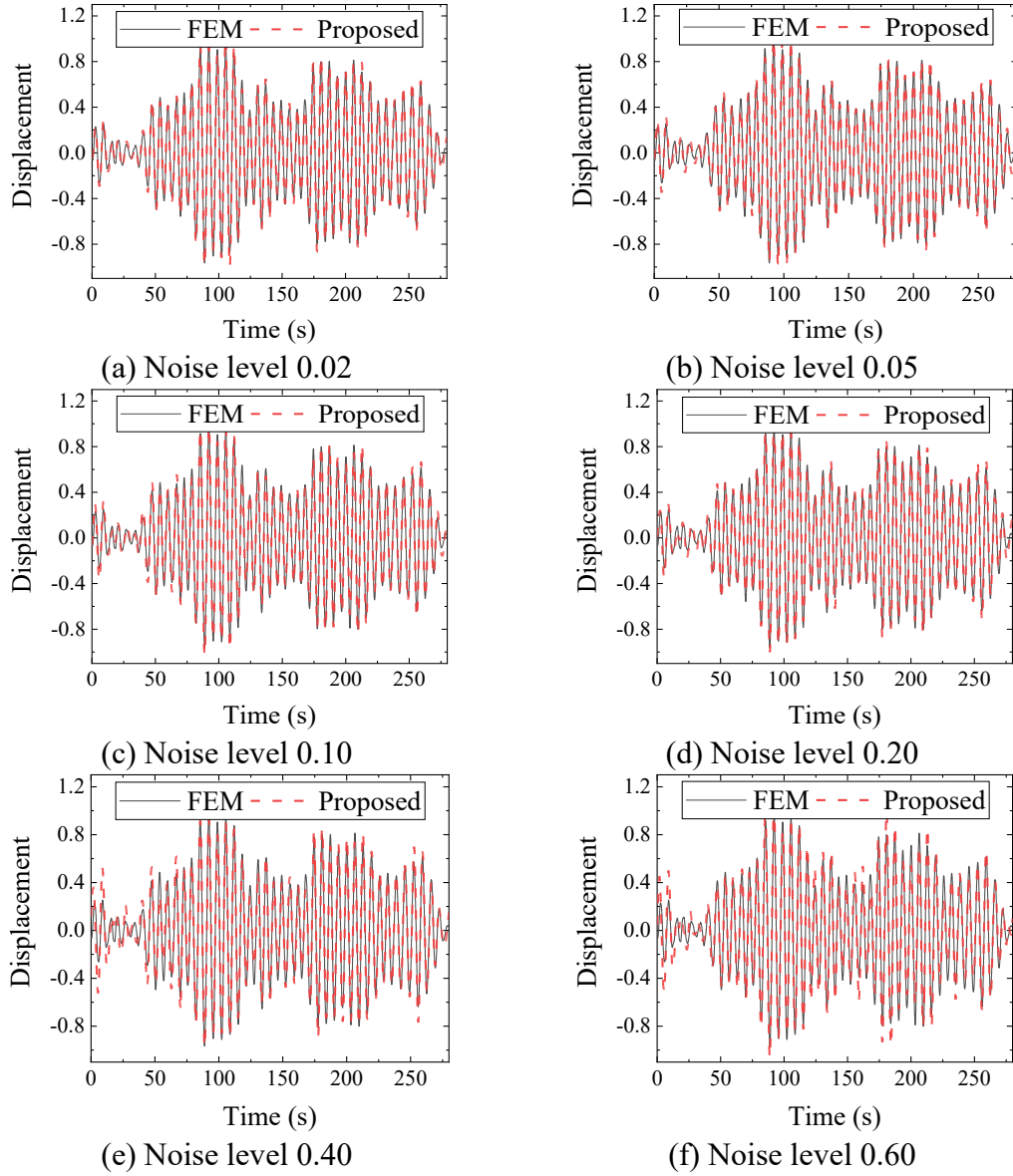


Figure 6.15 Reconstructed and simulated displacements of the 80th floor under different-level white noises

In summary, since the modal dynamic displacements of large-scale structures under typhoon loads are primarily dominated by low-order mode vibrations, low-frequency noise below 1 Hz and white noise containing low-frequency components have a limited impact on the displacement reconstruction results. Even under extreme noise intensity (noise variance up to 0.60 times the variance of the structural displacement), the reconstructed displacement have a relative error of 12%. In practical monitoring environments for large-scale structures, the main environmental disturbances affecting

accelerometer installations typically arise from sources such as pedestrian movement and mechanical operations (e.g., air conditioning systems), whose vibration frequencies are generally much higher than the fundamental frequency of large structures. Therefore, under real monitoring conditions, the proposed displacement reconstruction method exhibits strong robustness.

6.4 In-field study: Full-field displacement reconstruction of the Kingkey 100

This study takes the Kingkey 100 high-rise structure in Shenzhen as an example to estimate, reconstruct, and analyze the displacement response characteristics of high-rise structures under typhoon loads. By fusing limited phase-based displacement measurements and distributed acceleration measurements from the monitoring system, full-field displacements and interstory drift ratios are reconstructed. The Kingkey 100 structure is located in the Luohu District of Shenzhen, with a total height of 441.8 m. It comprises 100 above-ground floors (including the top steel-frame roof) and 4 underground levels, as shown in Figure 6.16. The ground floor measures 57 m $\times$ 46 m and gradually reduces along the height. At the 18th, 37th, 55th, 74th, and 92nd floors, five outrigger trusses are used to constrain the core tube and the frame. The cross support is set between each two outrigger trusses along the N-S direction to enhance the stiffness.

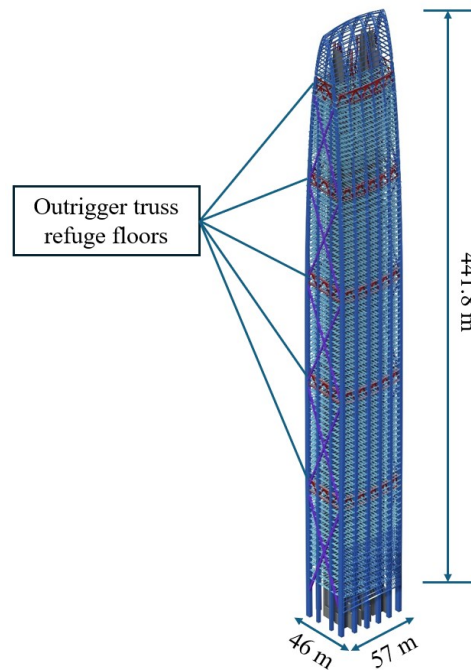


Figure 6.16 Kingkey 100 high-rise structure

6.4.1 Response characteristic analysis under typhoons

6.4.1.1 Distributed SHM system of the Kingkey 100

Hu et al. (2022) established a long-term SHM system on the Kingkey 100 Structure. This system employs wireless transmission to overcome the constraints posed by spatial partitions. Figure 6.17 illustrates the SHM system and the accelerometers, which operate at a sampling frequency of 20 Hz.



(a) Data acquisition system



(b) Accelerometers (991B)

Figure 6.17 Wireless distributed SHM system and sensors

A Zhonghaida MS302 GNSS rover is employed at the top of the structure to directly measure structural displacements for validation of the proposed method. The rover offers static measurement accuracy of 2.5 mm and dynamic measurement accuracy of 8 mm, with a sampling frequency of 1 Hz. A static reference station is set up on the roof of the Guiyuan Primary School building, approximately 200 m away from the Kingkey 100. During displacement processing, only satellite signals above a 15-degree elevation mask angle are utilized, and the daily periodic characteristics of multipath effects are considered (Li et al., 2014) to reduce their interference with displacement measurements. The GNSS displacement measurement point on the roof of the Kingkey 100 and the reference station setup are illustrated in Figure 6.18.

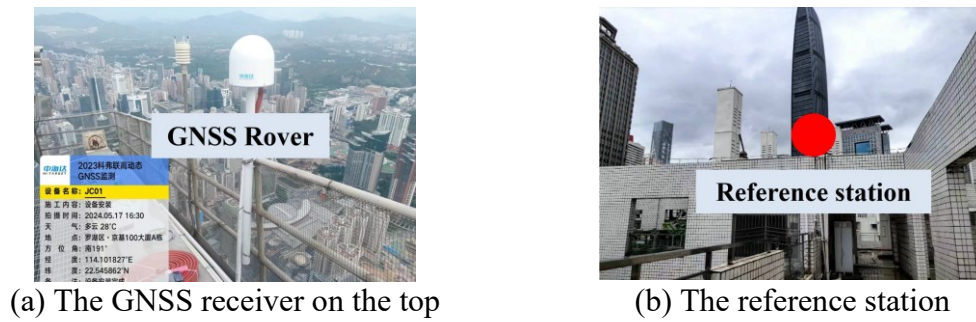


Figure 6.18 GNSS displacement sensors

The measurement point layout for the SHM system in the Kingkey 100 is shown in Figure 6.19. Specifically, accelerometers oriented in the N-S and E-W directions are installed on the outrigger truss refuge floors at the 37th, 55th, 74th, and 92nd floors. The 96th floor houses the hotel lobby at the top of the Kingkey 100, and accelerometers in the N-S and E-W directions are mounted on the steel frame levels 3, 5, and 7 above this floor. GNSS displacement sensors for the N-S and E-W directions are installed on the roof of the Kingkey 100.

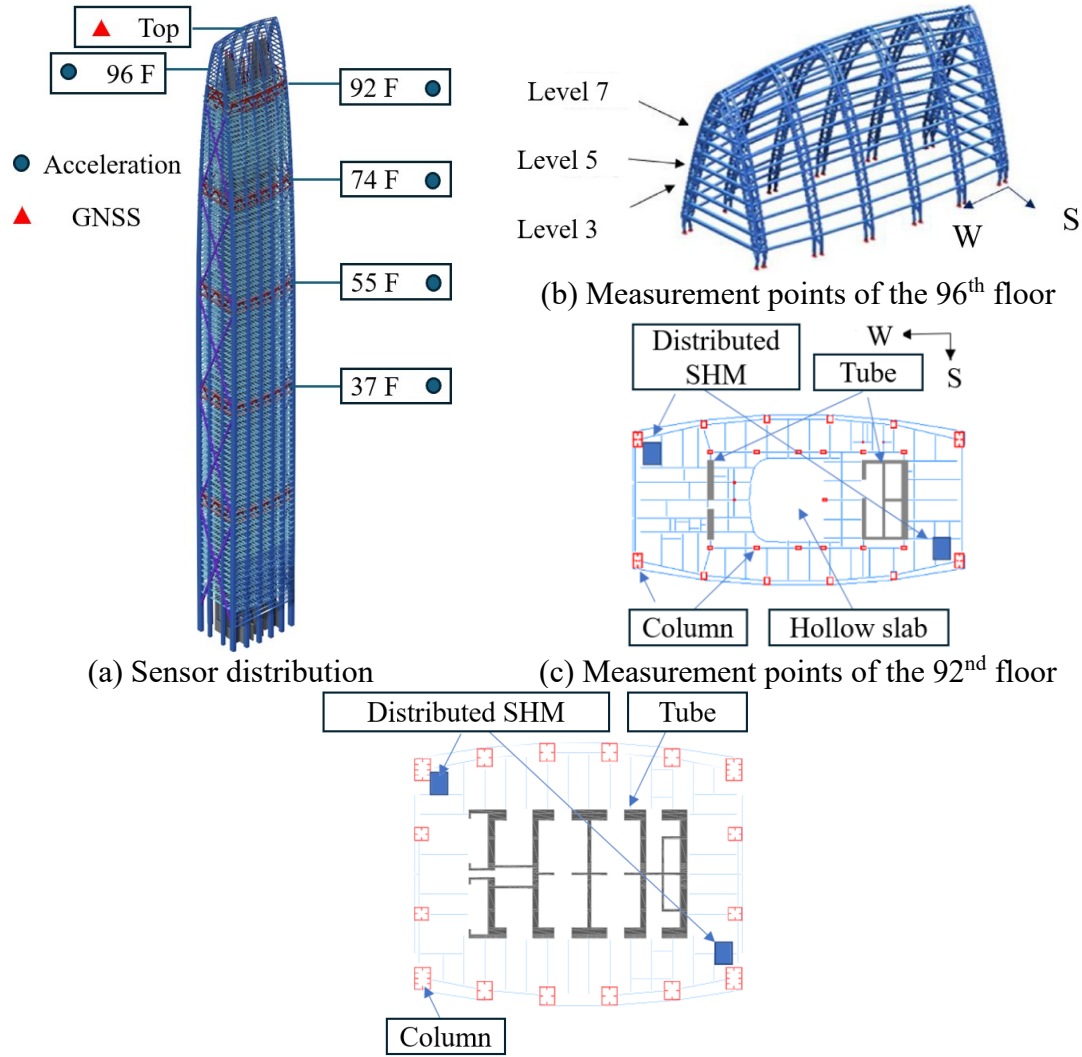


Figure 6.19 Measurement point layout

6.4.1.2 Phase-based displacement measurement points of the Kingkey 100

Displacements of the Kingkey 100 under typhoon loads are monitored using the proposed phase-based method. The 80th floor of the Kingkey 100 is located between the two high-rise outrigger trusses, where the constraints from the outrigger trusses are relatively weak, leading to comparatively larger interstory deformations. The 96th floor houses the hotel lobby area, from which the upper steel frame levels of the Kingkey 100 can be observed. Therefore, monitoring equipment and cameras are installed in the stairwell on the 80th floor and in the hotel lobby on the 96th floor to monitor interstory deformations. At the

measurement point on the 80th floor, the relative deformation between the ceiling and slab is observed. This was achieved by using a surveillance camera mounted under the ceiling of the stairwell to capture a laser beam marker projected from the floor slab, as shown in Figure 6.20(a) and Figure 6.20(b). The layout of the measurement point is illustrated in Figure 6.21, and the results are compared and validated with LDV measurements.



(a) Surveillance camera



(b) Laser beam target

Figure 6.20 The surveillance camera and laser beam target on the 80th floor

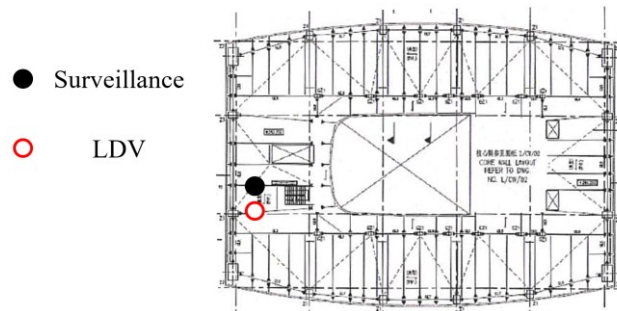


Figure 6.21 The interstory deformation measurement point on the 80th floor

In the hotel lobby on the 96th floor, no floor slabs are obstructing the view of the steel frame levels above this floor. A camera was set up here to monitor the relative displacement of the steel frame at a height of 10 m above the 96th floor relative to the floor slab of the hotel lobby, with simultaneous LDV monitoring for validation, as shown in Figure 6.22. Since both the camera and the LDV equipment mounted on the floor slab of the 96th floor rotate along with the slab's rotation, the relative displacement observed at this level does not account for the cumulative horizontal displacement induced by bending deformations of the lower floors; it only reflects the shear deformation. Considering that the displacement identified by the camera lies within the image plane,

while the displacement identified by the LDV is parallel to its optical axis, the camera and LDV are positioned at the E-W and N-S edges of the 96th floor, respectively.

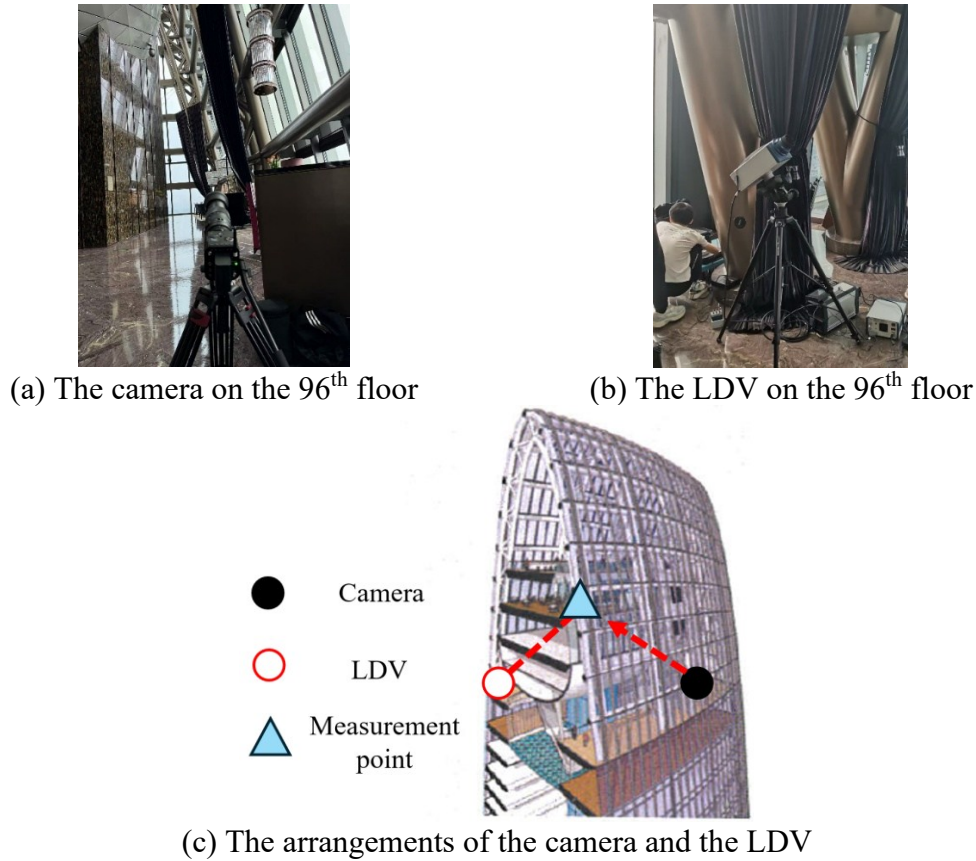


Figure 6.22 The interstory deformation of the steel frame on the 96th floor

6.4.1.3 Response characteristics under Typhoon Maliksi

The displacement and acceleration monitoring and response characteristic analysis under Typhoon Maliksi are conducted based on the aforementioned structural health monitoring system of the Kingkey 100 structure. At the 80th-floor stairwell, the surveillance video has a resolution of 3840×2160 pixels and a frame rate of 25 Hz. The scaling factor for the digital image at the 80th floor is determined as 0.2599 mm/pixel by calculating the ratio of the physical dimension of the laser marker width (7.2 cm) to its pixel width in the surveillance video (277 pixels), as shown in Figure 6.23. At the 96th-floor hotel lobby, a ZCAM-S6G camera with a Sigma 150-600 mm F5-6.3 DG OS HSM lens was used to

measure the relative displacement of the steel frame at a height of 10 m above the floor slab. The camera was configured with a focal length of 600 mm, a video resolution of 5760×3240 pixels, a frame rate of 25 Hz, an aperture of F/6.3, and an ISO of 1250. According to the Nyquist sampling theorem, the shutter speed in video recording is generally set to at least twice the frame rate. To suppress motion blur of the structural target and considering that the natural frequencies and environmental excitations of supertall structures are far below 25 Hz, the shutter speed was set to 1/100 s. The scaling factor for the digital image at the 96th floor is calculated as 0.144 mm/pixel by taking the ratio of the steel column diameter (30 cm) to its corresponding pixel width in the video (2083 pixels), as illustrated in Figure 6.23. Since the steel column is inclined at 50.43° from the vertical direction, the total horizontal displacement is obtained by dividing by the cosine of this angle.



(a) The frame captured on the 80th floor (b) The frame captured on the 96th floor

Figure 6.23 Frames captured on the 80th and 96th floors

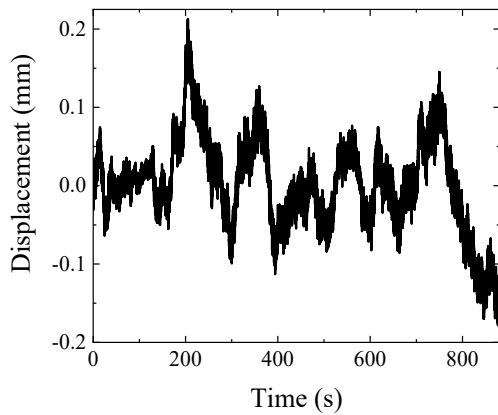
The proposed digital image phase-based method is used to estimate the floor displacements at the 80th-floor stairwell and the 96th-floor steel frame at 13:40 and 12:40, respectively. A Polytech RSV-150 LDV displacement sensor (displacement resolution: 0.38 nm, sampling frequency: 25 Hz) was employed as a benchmark to validate the accuracy of the proposed digital image phase-based displacement estimation. The comparison results are shown in Figure 6.24.

The results indicate that the displacements estimated from the surveillance video on the 80th floor align with the low-frequency quasi-static components measured by the LDV,

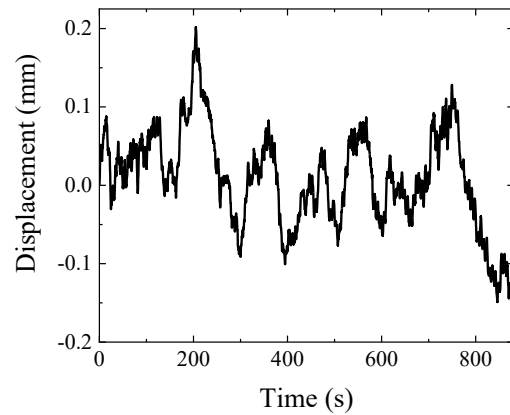
with displacement amplitudes ranging from -0.2 mm to 0.2 mm. Similarly, the vision-based estimation results for the steel frame on the 96th floor correspond to the low-frequency quasi-static components measured by the LDV, with displacements fluctuating between -0.25 mm and 0.35 mm. The displacement amplitudes measured on the 96th floor are larger than those on the 80th floor, which is attributed to the measurement height of 10 m on the 96th floor, compared to the 3.5 m story height on the 80th floor.

It should be noted that the 96th-floor location is a hotel lobby, where environmental excitations from pedestrian and trolley movements are more significant than in the 80th-floor stairwell. Additionally, the LDV and camera on the 96th floor were installed on different adjacent sides of the structure, meaning the environmental excitations affecting the LDV and camera on this floor differed and were generally stronger than those on the 80th floor. Furthermore, considering that the temperature did not change significantly during the measurement period, no low-frequency quasi-static effects on structural displacements due to temperature variations are expected.

In summary, the consistency between the displacement results obtained from the LDV and the camera confirms that the measured displacements reflect the mechanical behavior of the Kingkey 100 structure under typhoon loads and are not significantly influenced by environmental vibrations or thermal disturbances.



(a) Video results on the 80th floor at 13:40



(b) LDV results on the 80th floor at 13:40

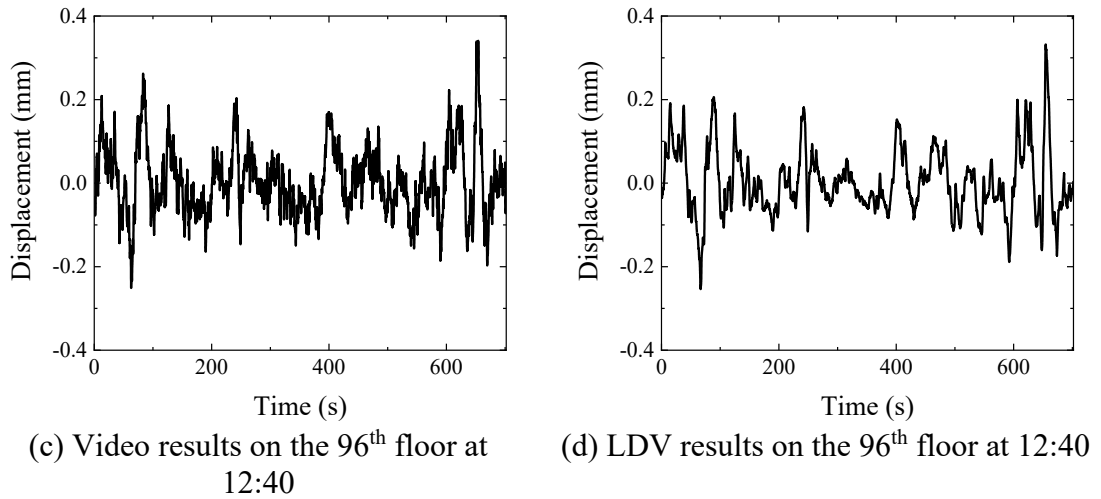


Figure 6.24 Comparison between the estimated and LDV-measured displacements under Typhoon Maliksi

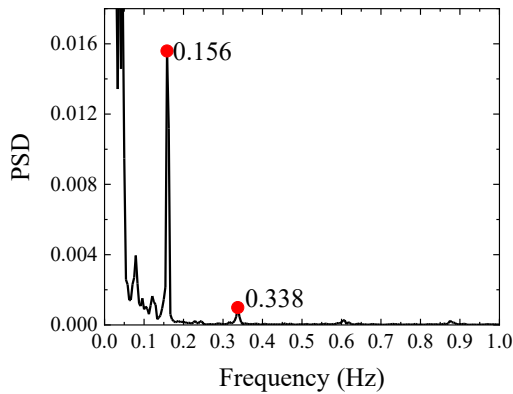
In terms of the periodic characteristics of the displacements, it is observed that both the 80th-floor stairwell and the 96th-floor steel frame exhibit distinct long-period fluctuations in interstory displacements under typhoon loads, with periods exceeding 100 s. This is significantly longer than the natural period of 6.41 s corresponding to the fundamental frequency (0.156 Hz) of the Kingkey 100 structure.

To further investigate the frequency-domain characteristics of the displacement responses, the PSD of the estimated displacements is calculated and compared with results from GNSS and accelerometers, as shown in Figure 6.25. The frequency-domain signals reveal that the displacements measured by cameras, surveillance systems, LDV, and GNSS on the Kingkey 100 structure under Typhoon Maliksi all exhibit distinct features corresponding to the first bending frequency.

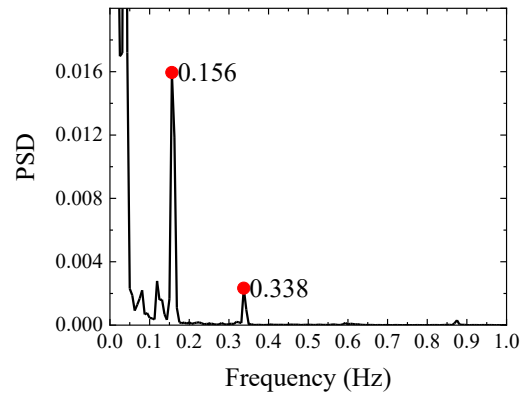
It should be noted that the PSD results from all displacement sensors in the Kingkey 100 show significantly low-frequency components, which correspond to the long-period displacement fluctuations observed in the time-domain results. As analyzed earlier, the proposed phase-based displacement estimation method employs Log-Gabor complex filters to remove DC components, and temperature variations were controlled by the pre-

operation. Therefore, these long-period displacement fluctuations do not originate from high-frequency environmental excitations or temperature variations. Moreover, the consistency of long-period fluctuations measured by different devices in Figure 6.24 collectively indicates that these fluctuations result from the low-frequency quasi-static components of the typhoon.

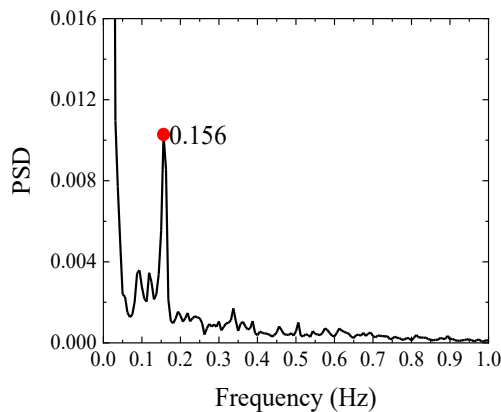
In the accelerometer results, the first bending frequency is the dominant component. Additionally, the first torsional frequency, the second bending frequency, and the second torsional frequency are also prominent. However, low-frequency components below the fundamental frequency are almost absent. This high-frequency dynamic behavior contrasts with the characteristics observed in the displacement results.



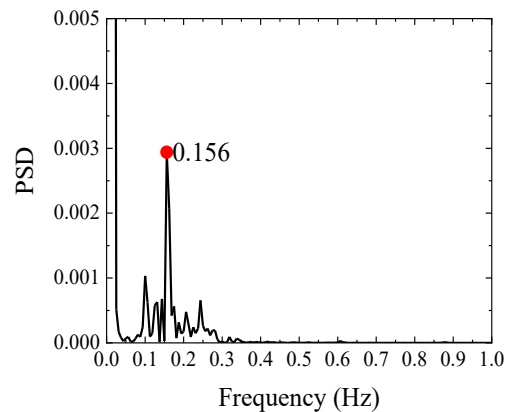
(a) PSD of video results on the 80th floor



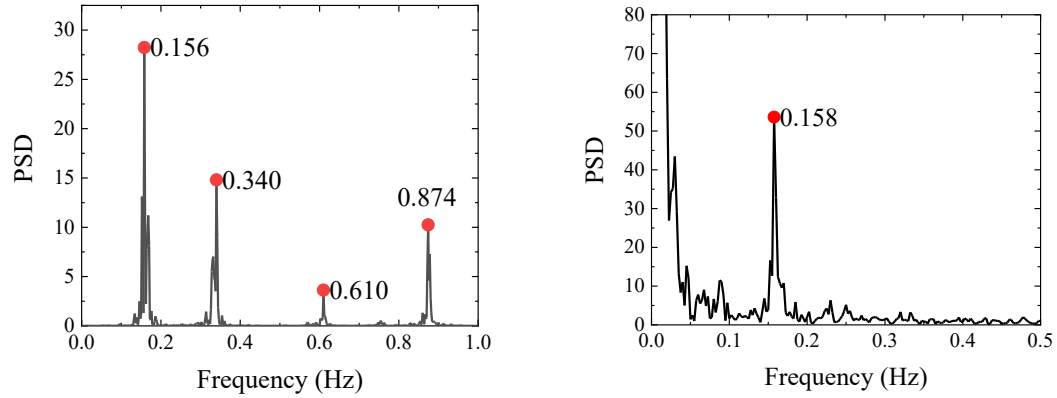
(b) PSD of LDV results on the 80th floor



(c) PSD of video results on the 96th floor



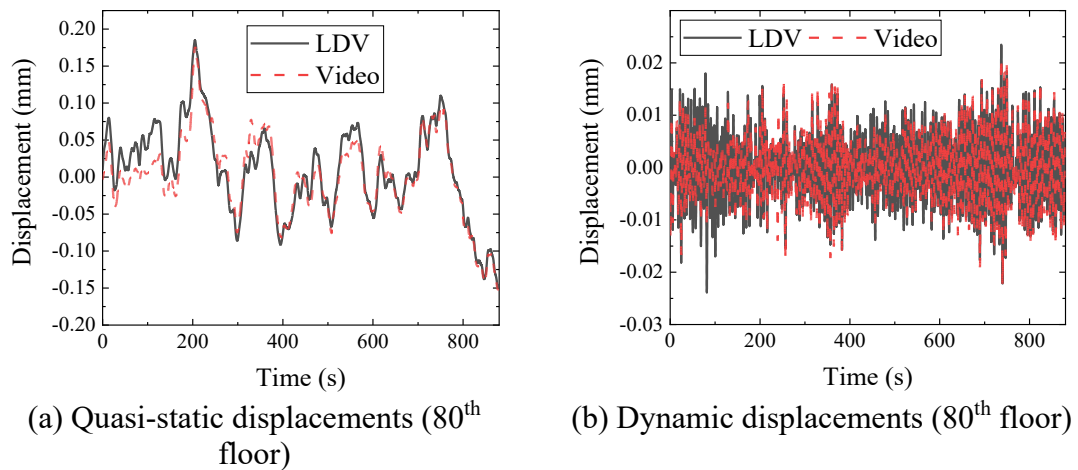
(d) PSD of LDV results on the 96th floor



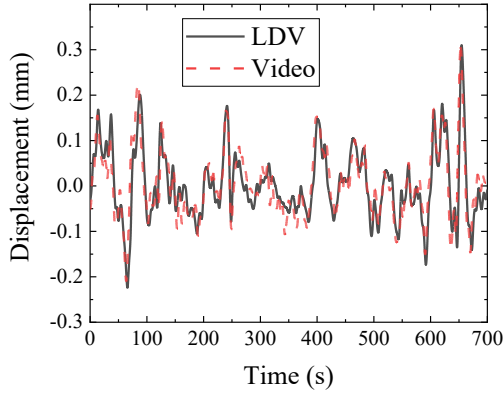
(e) PSD of the acceleration on the 96th floor (f) PSD of GNSS results on the top floor

Figure 6.25 PSD of measured displacements and accelerations under Typhoon Maliksi

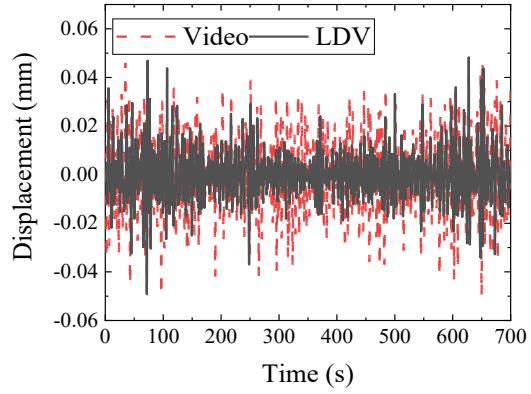
To analyze the amplitudes of low-frequency and relatively high-frequency components in the wind-induced displacement responses of the supertall structure, a 7th-order Butterworth filter is applied in the frequency domain to separate the low-frequency and high-frequency components of the displacements measured from video and LDV data, with a cutoff frequency of 0.10 Hz. The low-pass filtered results are classified as quasi-static displacements, while the difference between the original displacement responses and the low-pass filtered results represents the relatively high-frequency dynamic displacement components, as illustrated in Figure 6.26.



(a) Quasi-static displacements (80th floor) (b) Dynamic displacements (80th floor)



(c) Quasi-static displacements (96th floor)



(d) Dynamic displacements (96th floor)

Figure 6.26 Quasi-static and dynamic displacements measured by the video and the LDV

The quasi-static displacement results from both video and LDV measurements at the 80th-floor stairwell and the 96th floor are in good agreement, with correlation coefficients of 0.9531 and 0.9194, respectively. Since the measurement environment in the 80th-floor stairwell is more stable than that in the 96th-floor hotel lobby, the dynamic displacement results from video and LDV at the 80th floor show better consistency compared to those at the 96th floor.

At the 80th floor, the quasi-static displacement amplitude is 0.182 mm, and the dynamic displacement amplitude is 0.025 mm. The quasi-static component of the displacement is 7.28 times larger than the dynamic component. At the 96th floor, the quasi-static displacement amplitude is 0.310 mm, which is 6.07 times larger than the dynamic displacement amplitude of 0.051 mm.

In summary, under Typhoon Maliksi, the displacement response of the Kingkey 100 supertall structure is primarily characterized by dynamic vibrations superimposed on a quasi-static drift, with the quasi-static drift being the dominant component. This should be taken into account when calculating and evaluating interstory drift ratios for structural assessment under typhoon loads.

6.4.2 Full-field displacement reconstruction based on DFDR

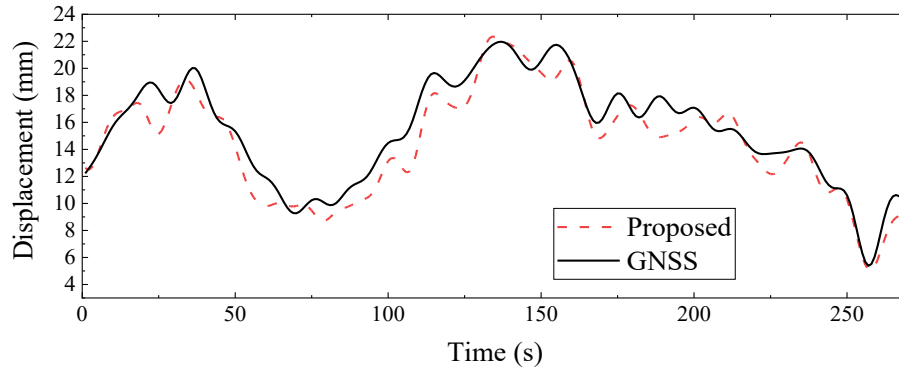
For the Kingkey 100 supertall structure, this study reconstructs the full-field horizontal displacements and interstory drift ratios under Typhoon Maliksi on June 1, 2024, and Typhoon Yagi on September 6, 2024. The reconstruction fuses limited phase-based displacement data and acceleration measurements from distributed sensors.

6.4.2.1 Quasi-static displacement reconstruction under typhoons

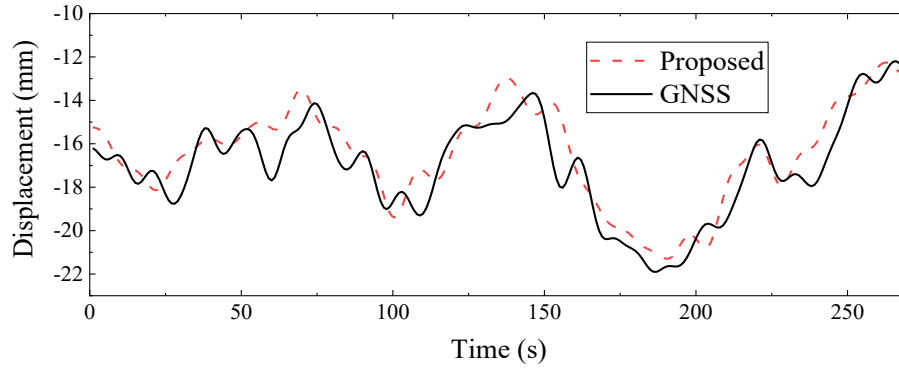
The quasi-static displacements induced by low-frequency components of the typhoon on the supertall structure are reconstructed. According to the simulation analysis in Section 6.3, when displacement measurement points are placed at the top and the 96th floor of the Kingkey 100 supertall structure, the quasi-static horizontal displacements below two-thirds of the fundamental frequency can be accurately estimated for all floors. Therefore, in field measurements, the horizontal displacement at the 96th floor and the top of the Kingkey 100 structure, obtained from digital image phase-based methods and GNSS as described in the case study of Section 5.4, are used to reconstruct the full-field displacements and interstory drift ratios across all floors. Based on the measured horizontal displacements at the top and the 96th floor of the Kingkey 100 supertall structure, after extracting the quasi-static components below 0.10 Hz from the displacement, Equation (6.44) is applied to reconstruct the quasi-static horizontal displacements below 0.10 Hz for all floors.

First, a single displacement measurement point at the 96th floor is used to reconstruct the top-floor quasi-static displacement, which is verified by the GNSS. Figure 6.27 compares the reconstructed and measured quasi-static N-S and E-W displacements at the top floor under Typhoons Maliksi and Yagi, with relative errors of 4.98% and 4.80%, respectively. These results demonstrate that the proposed quasi-static displacement reconstruction method for large-scale structures can effectively estimate the horizontal displacements of

other floors even when only a single upper-floor displacement measurement point is available.



(a) N-S quasi-static displacements under Typhoon Maliksi



(b) E-W quasi-static displacements under Typhoon Yagi

Figure 6.27 Reconstructed and measured quasi-static displacements of Kingkey 100 top floor under typhoons

According to the analysis in Section 6.3.2, two or more displacement measurement points can accurately reconstruct the interstory drift ratios of all floors. Therefore, the displacements from both the top and the 96th floor of the Kingkey 100 structure are used to reconstruct the quasi-static interstory drift ratios of all floors.

Figure 6.28 presents the maximum quasi-static displacement and the corresponding quasi-static interstory drift ratios in the N-S direction reconstructed from the monitoring data of Typhoon Maliksi. Under the low-frequency quasi-static components of Typhoon Maliksi, the horizontal displacements of all floors exhibit significant bending deformation

characteristics. Moreover, the maximum N-S interstory drift ratio of the Kingkey 100 supertall structure does not exceed 0.0001, which is far below the code-specified elastic interstory drift ratio limit of 0.002. Due to slab openings in the core tube for guest room functions on and around the 80th floor, the weakened horizontal components result in relatively larger increases in interstory drift ratios compared to other standard floors. Among all floors, the interstory drift ratio at the 80th floor (located between the two upper outrigger trusses) reaches the spatial maximum of 0.0000917. Furthermore, the interstory drift ratios near the outrigger truss floors exhibit distinct local minima in the spatial distribution, confirming the enhancement of outrigger trusses on the lateral stiffness of the supertall structure.

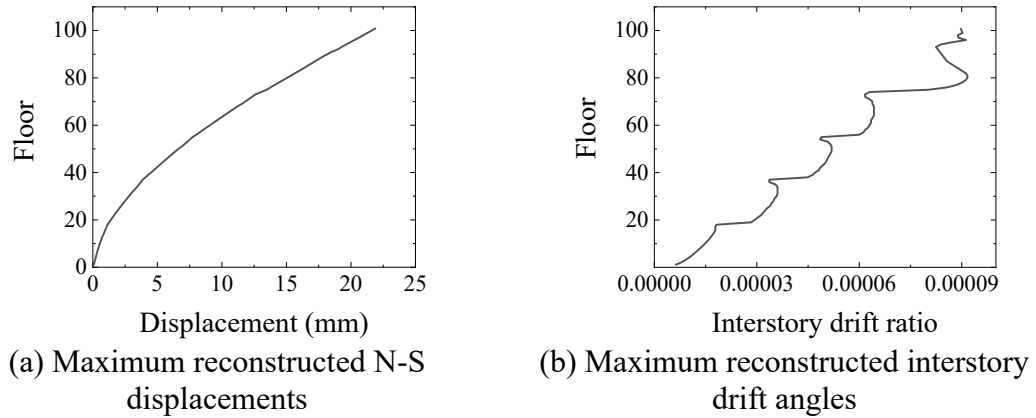


Figure 6.28 Maximum N-S full-field quasi-static horizontal displacement reconstruction of Kingkey 100 under Typhoon Maliksi

Figure 6.29 shows the maximum quasi-static displacement in the time domain and the corresponding quasi-static interstory drift ratios in the E-W direction reconstructed from the monitoring data of Typhoon Yagi. The spatial distribution of E-W interstory drift ratios similarly displays a convergence pattern near the floors with outrigger trusses, with the maximum value of 0.0000860 occurring at the 80th floor, again well below the elastic limit of 0.002.

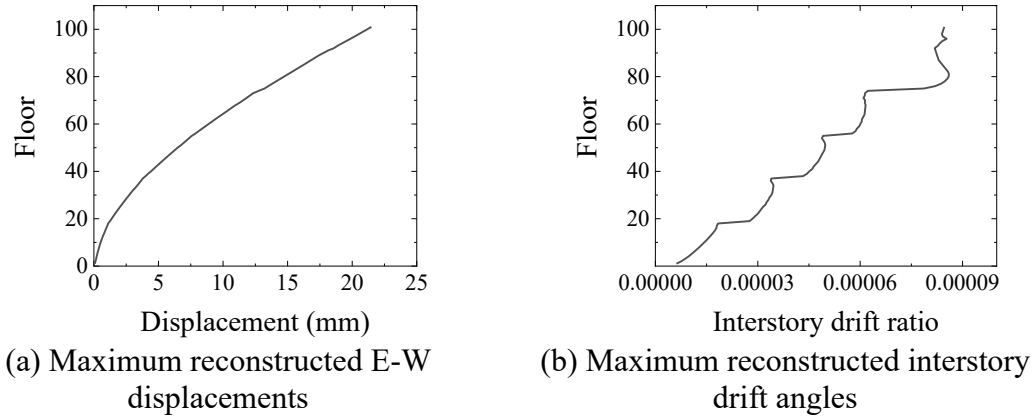


Figure 6.29 Maximum E-W full-field quasi-static horizontal displacement reconstruction of Kingkey 100 under Typhoon Yagi

6.4.2.2 Modal dynamic displacement reconstruction under typhoons

Accelerometers installed on the 37th, 55th, 74th, 92nd, and 96th floors of the Kingkey 100 supertall structure are utilized to identify the modal contributions under typhoon loads and reconstruct the modal dynamic displacements of all floors.

The accelerometers are located at the corners of the floor plane, measuring accelerations with torsional modal components. However, this study focuses on the average interstory drift ratios of each floor, so only the contributions of bending modes are considered in the dynamic displacement reconstruction. First, the 37th, 55th, 74th, and 92nd floors are used as known measurement points to reconstruct the acceleration and dynamic displacement at the 96th floor, and the results are compared with measured values to validate the performance of modal dynamic displacement reconstruction, as shown in Figure 6.30.

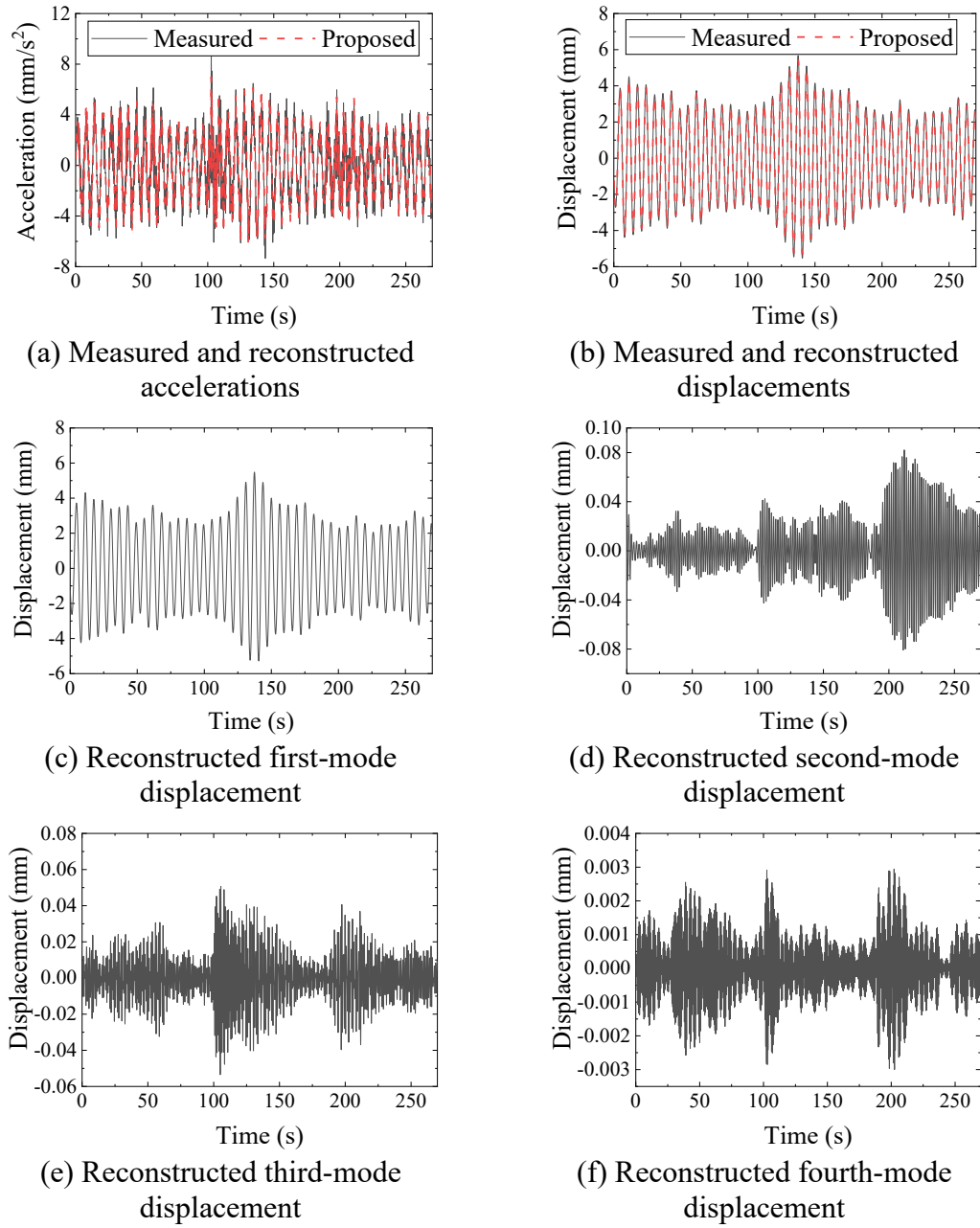
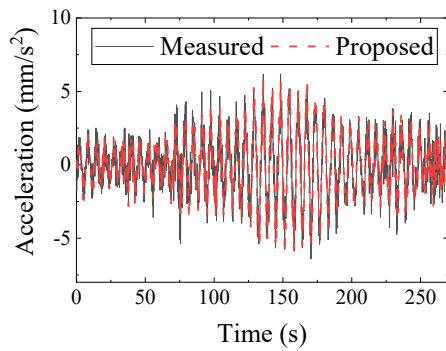


Figure 6.30 Reconstructed and measured N-S accelerations and displacements of the 96th floor under Typhoon Maliksi

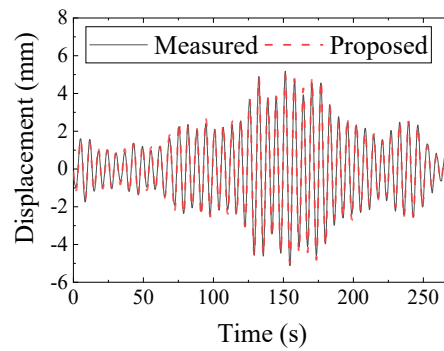
In Figure 6.30(a), the N-S acceleration at the 96th floor reconstructed by other floors aligns with the measured amplitude, but the reconstructed acceleration neglects higher-order modes, resulting in a relative error of 9.70% compared to the measured values. The N-S dynamic displacement at the 96th floor is reconstructed using Equation (6.45) and compared with the dynamic displacement obtained by integrating the acceleration

components above 0.10 Hz measured at the 96th floor, as shown in Figure 6.30(b). The reconstructed N-S dynamic displacement at the 96th floor matches the measured values, with a relative error of 4.40%. Under Typhoon Maliksi, the amplitude of the N-S dynamic displacement at the 96th floor is approximately 5.42 mm, with the majority contributed by the first-mode displacement. The second and higher-mode displacements are less than 2% of the first-mode amplitude. Therefore, considering the first four modes is sufficient to accurately reconstruct the dynamic displacements.

For Typhoon Yagi, which primarily acts in the E-W direction, the proposed method reconstructs the E-W acceleration and dynamic displacement at the 96th floor, as shown in Figure 6.31. The acceleration at the 96th floor reconstructed from the first four modes follows the same trend as the measured values, but the neglect of higher-order modes leads to a relative error of 11.42%. Since the dynamic displacement under typhoon loads is predominantly characterized by the first mode in the frequency domain, the E-W dynamic displacement at the 96th floor reconstructed from the first four modes aligns well with the dynamic displacement obtained by integrating the acceleration components above 0.10 Hz measured at the 96th floor, with a relative error of 5.05%.



(a) Measured and reconstructed accelerations



(b) Measured and reconstructed displacements

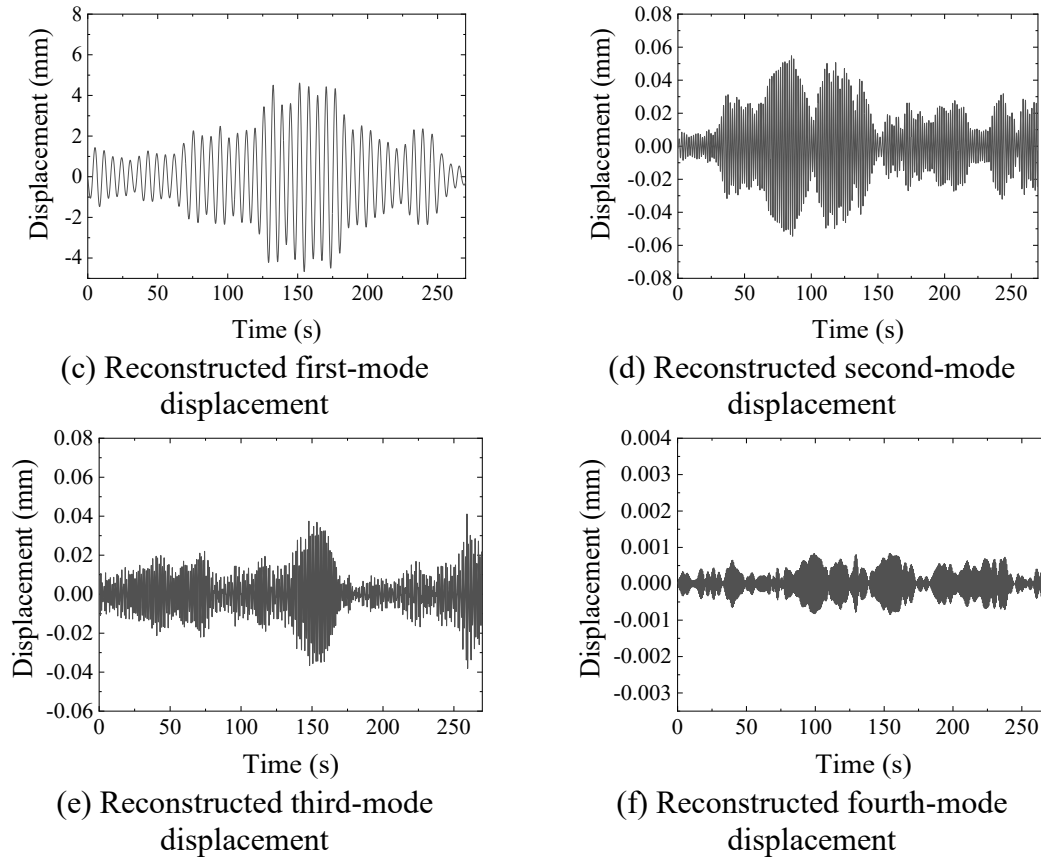


Figure 6.31 Reconstructed and measured E-W accelerations and displacements of the 96th floor under Typhoon Yagi

The temporarily maximum horizontal displacements and interstory drift ratios in the N-S and E-W directions under typhoon loads are presented in Figure 6.32 and Figure 6.33, respectively. The maximum N-S roof displacement of the Kingkey 100 supertall structure is 5.76 mm under the fluctuating wind of Typhoon Maliksi. The maximum interstory drift ratio occurs at the 80th floor, with a value of 0.0000185. Under the fluctuating wind components of Typhoon Yagi, the maximum E-W roof displacement is 4.71 mm. The maximum interstory drift ratio at the corresponding moment is also located at the 80th floor, with a value of 0.0000147.

Moreover, the reconstructed modal dynamic displacements under fluctuating wind loads are smaller than the contributions of the corresponding low-frequency quasi-static

displacements. This indicates that the low-frequency quasi-static displacements of the structure under typhoon loads cannot be neglected.

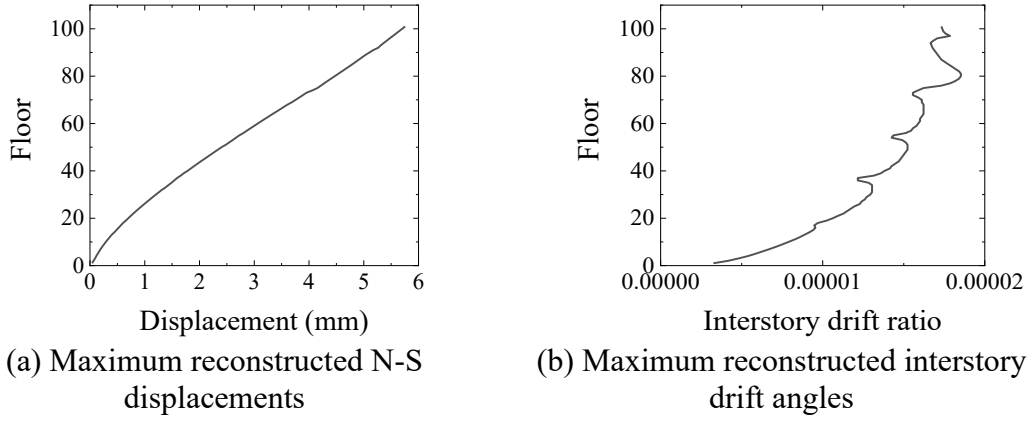


Figure 6.32 Maximum N-S full-field dynamic displacement reconstruction of Kingkey 100 under Typhoon Maliksi

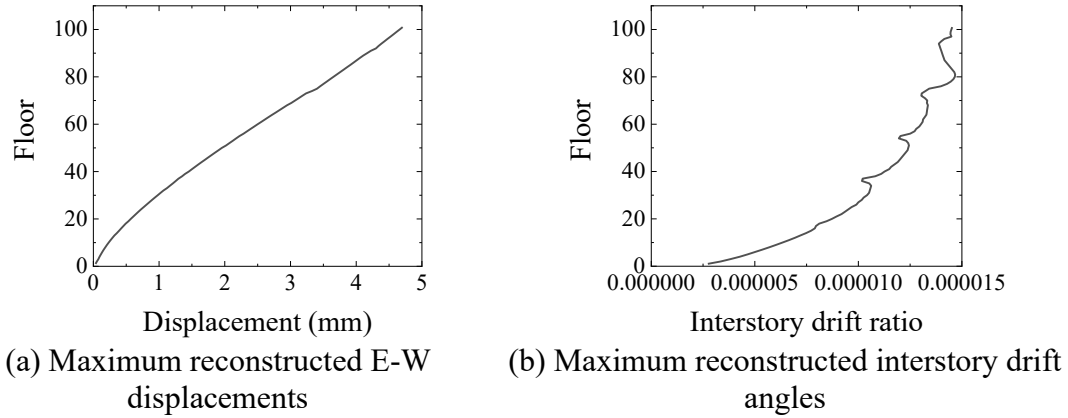


Figure 6.33 Maximum E-W full-field dynamic displacement reconstruction of Kingkey 100 under Typhoon Yagi

6.4.2.3 Full-field displacement and interstory drift angle reconstruction under typhoons

The full-field low-frequency quasi-static horizontal displacements and modal dynamic horizontal displacements of the Kingkey 100 supertall structure are substituted into Equation (6.46) to superimpose and reconstruct the full-field displacement and interstory

drift ratios under typhoon loads. The maximum values in the time domain of the N-S and E-W displacements and interstory drift ratios for all floors under Typhoon Maliksi and Typhoon Yagi are shown in Figure 6.34 and Figure 6.35, respectively. The N-S roof displacement is 27.39 mm, and the maximum interstory drift ratio across all floors in the spatial domain occurs at the 80th floor (located between the two upper outrigger truss levels) with a value of 0.000109, which is far below the elastic limit of 0.002. The E-W roof displacement is 24.77 mm, and the corresponding maximum interstory drift ratio in the spatial domain, 0.0000957, also occurs at the 80th floor. Furthermore, the outrigger trusses significantly constrain the N-S and E-W interstory drift ratios at the 18th, 37th, 55th, 74th, and 92nd floors.

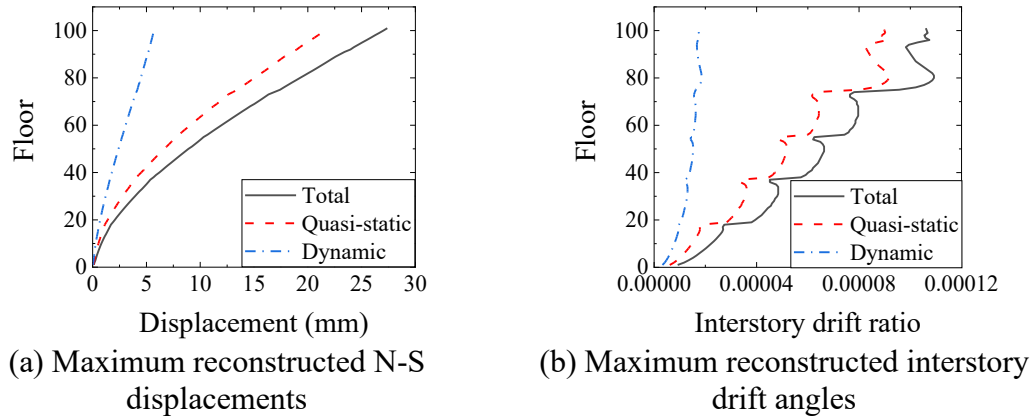


Figure 6.34 Maximum N-S full-field displacement reconstruction of Kingkey 100 under Typhoon Maliksi

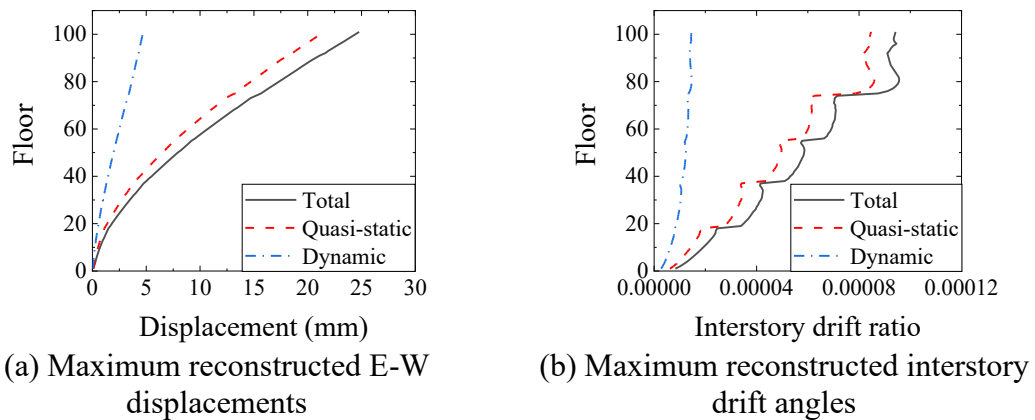


Figure 6.35 Maximum E-W full-field displacement reconstruction of Kingkey 100 under Typhoon Yagi

The low-frequency quasi-static and dynamic components of the horizontal displacement time history are compared to analyze the response characteristics of the structure under typhoons. Figure 6.36 shows the ratio of the absolute dynamic displacement to the absolute total displacement at the roof level. It can be observed that under Typhoons Maliksi and Yagi, the average ratios of dynamic displacement amplitude to total displacement amplitude are 0.1608 and 0.0904, respectively. Moreover, the proportion of the dynamic component in the total horizontal displacement consistently remains below 0.6.

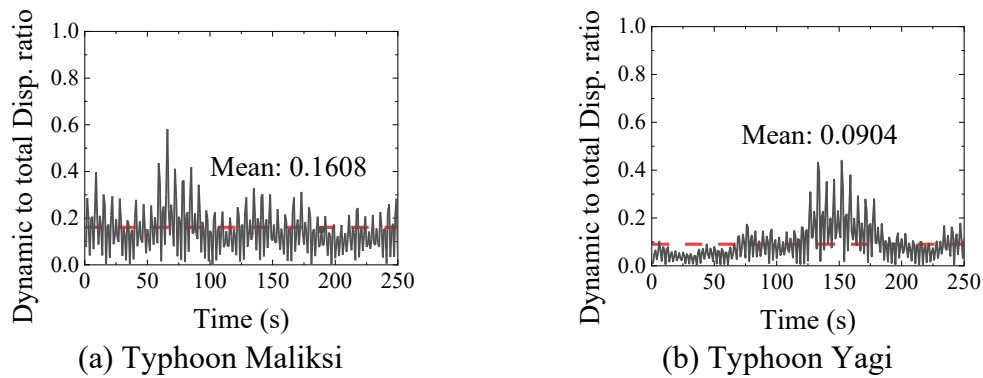


Figure 6.36 Ratio of dynamic to total displacements of the top floor under typhoons

In summary, under typhoon loads, the quasi-static displacement induced by low-frequency wind components contributes significantly to the total horizontal displacements and interstory drift ratios of the supertall structure.

6.5 Summary

This chapter addresses the challenge of obtaining spatiotemporally continuous full-field displacements for large-scale structures in the presence of spatial partitions. By extracting the low-frequency quasi-static components from phase-based displacements and the modal dynamic components from acceleration data, and integrating them with the

structural mechanical model, a full-field displacement reconstruction method based on fusion of limited phase-based displacement and acceleration data is proposed for large-scale structures.

For the low-frequency quasi-static displacements of large-scale structures, a quasi-static displacement reconstruction theory based on the pseudo-inverse method is introduced. For the modal dynamic displacements, a modal dynamic displacement reconstruction method based on stochastic subspace identification is proposed. The quasi-static and modal dynamic displacements are then superimposed to reconstruct the spatiotemporally continuous full-field displacements across all degrees of freedom. This approach enables the reconstruction of full-field interstory drift ratios for the Kingkey 100 supertall structure under Typhoons Maliksi and Yagi. The reconstructed maximum full-field interstory drift ratios in the N-S and E-W directions are 0.000109 and 0.0000957, respectively, both far below the elastic limit of 0.002. Moreover, the proportion of the dynamic component in these ratios consistently remains below 60%, indicating the significant quasi-static displacement of supertall structures under typhoons.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

This study has developed digital image phase-based techniques for full-field displacement estimation and reconstruction of large-scale structures through integrating theoretical analysis, numerical simulations, laboratory experiments, and field measurements. First, a method for accurate local phase extraction of target structures under extra textures, illumination variations, and image noises is proposed to accurately estimate displacement in complex environments. Second, the multi-resolution complex pyramid and discriminative correlation are developed to adaptively estimate the full-field displacements. This approach is applied to real large-scale structures to estimate their full-field displacements. Finally, a full-field displacement reconstruction method is established for large-scale structures with spatial partitions by integrating limited phase-based displacement and acceleration data. This approach achieves the full-field displacement reconstruction for the high-rise structure under typhoons. Based on the above research, the main conclusions can be drawn as follows:

- (1) A digital image phase-displacement projection relationship has been established to deal with the effects of complex environments (e.g., extra textures and illumination changes) on displacement estimation. The out-of-the-plane and in-plane extra textures are suppressed by the large aperture and small filter envelope, respectively. A Log-Gabor complex filter phase extraction method, optimized via a local optimization coefficient, is developed, achieving phase-based displacement estimation in complex environments. The accuracy of the developed method is validated through numerical

simulations and laboratory tests, in which the resolution of the developed method is better than 0.04 pixels.

- (2) The mechanism of image noise on displacement results is analyzed in the complex domain to address the adverse effects of unknown image noise in displacement estimation. An enhanced phase-based displacement estimation framework is then introduced, incorporating a phase denoising strategy to improve accuracy in digital image phase analysis. Specifically, a multi-frequency phase-preserving denoising technique tailored to structural features is developed, along with a noise threshold quantification method that leverages amplitude instability in the complex domain. In the numerical and experimental examples across a range of noise distributions, the displacement accuracy of the multi-frame phase-preserving denoising approach significantly outperforms existing methods.
- (3) For full-field displacement estimation with multi-scale amplitudes in large-scale structures, the quality of phase extraction under different displacement amplitudes is investigated. A multi-resolution complex pyramid is constructed, enabling adaptive displacement estimation that balances measurement range and accuracy. For multi-type target displacement estimation in large-scale structures, the discriminative correlation is introduced to extract geometric features of diverse targets. The proposed adaptive full-field displacement estimation method, integrating multi-resolution complex pyramid and discriminative correlation, is validated through experimental tests and field measurements on real large-scale structures such as the Tsing Ma Bridge, Kingkey 100 high-rise structure, and Humen Bridge.
- (4) The limited phases and accelerations are fused to obtain the full-field displacement of large-scale structures with spatial partitions. The low-frequency quasi-static and modal dynamic displacements are reconstructed by the pseudo-inverse and stochastic subspace identification, respectively, by which the full-field displacement is superimposed. The in-field study reveals that the displacement response of supertall structures under typhoon loads is a combination of quasi-static displacements, driven by low-frequency and mean wind effects, and modal dynamic displacements induced by turbulent wind fluctuations. Applying the developed method enables the

reconstruction of full-field displacements and inter-story drift ratios for the Kingkey 100 super-tall structure.

7.2 Future Work

Despite the contributions in this thesis, obtaining full-field displacement of large-scale structures can be further developed in the following aspects:

- (1) The performance of the digital image phase-based displacement estimation method proposed in this study is affected by vibrations of the imaging device itself. When the frequency-domain characteristics of such vibrations are unknown, it is necessary to integrate hardware and algorithms capable of detecting and mitigating the adverse effects of device vibrations.
- (2) Under varying environmental or operating temperatures, minor deformations of lenses and variations of the air refractive index lead to pixel-level thermal drift in static images. This makes vision-based displacement monitoring of large-scale structures challenging. Further development of temperature control hardware, such as semiconductor cooling systems, is required to stabilize device temperature.
- (3) The displacement estimation method for large-scale structures proposed in this study is only applicable to in-plane displacements. Suitable imaging perspectives must be selected for digital image phase-based displacement estimation. Multi-viewpoint feature matching algorithms could be developed to utilize multiple cameras for estimating the spatial coordinates of measurement points.
- (4) For application scenarios where it is difficult to find suitable sites for installing fixed cameras, there is a need to develop equipment for displacement measurement using UAV photogrammetry, along with algorithms that can remove the interference of UAV vibrations on displacement estimation results. Based on the UAV photogrammetry system, the 3D displacement can be amplified and visualized by operating the phase.

- (5) The present techniques work well for daytime only. Nighttime displacement estimation is required for real-time continuous monitoring. Therefore, artificial lighting sources and the feature extraction method can be utilized to track the displacement at night.
- (6) The phase-based method only works for displacement not parallel to the structural feature. Artificial target systems should be developed to improve the method so that displacement along different directions can be estimated.

REFERENCES

- Ai, S., Dong, C., Mei, Q., 2025. The optimal scale level of complex steerable pyramid for phase-based motion estimation under different motion ranges and target sizes. *Measurement* 253, 117440. <https://doi.org/10.1016/j.measurement.2025.117440>
- Aponte, J., Meng, X., Hill, C., Moore, T., Burbidge, M., Dodson, A., 2009. Quality assessment of a network-based RTK GPS service in the UK. *Journal of Applied Geodesy* 3, 25–34. <https://doi.org/10.1515/JAG.2009.003>
- Ara, K., 1972. A differential transformer with temperature- and excitation-independent output. *IEEE Trans. Instrum. Meas.* 21, 249–255. <https://doi.org/10.1109/TIM.1972.4314011>
- Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M.J., Szeliski, R., 2011. A database and evaluation methodology for optical flow. *Int. J. Comput. Vision* 92, 1–31. <https://doi.org/10/crp46s>
- Barron, J.L., Fleet, D.J., Beauchemin, S.S., Burkitt, T.A., 1992. Performance of optical flow techniques, in: *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Presented at the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Comput. Soc. Press, Champaign, IL, USA, pp. 236–242. <https://doi.org/10/fmhfpf>
- Bart Peeters, 2000. System identification and damage detection in civil engineering (Doctoral dissertation). Katholieke Universiteit Leuven, Leuven.
- Bian, X., 2019. Study on dynamic characteristics of high-rise structure based on distributed synchronous acquisition. Harbin Institute of Technology, Harbin.

- Blaber, J., Adair, B., Antoniou, A., 2015. Ncorr: open-source 2D digital image correlation matlab software. *Exp. Mech.* 55, 1105–1122. <https://doi.org/10.1007/s11340-015-0009-1>
- Boukerroui, D., Noble, J.A., Brady, M., 2004. On the choice of band-pass quadrature filters. *J. Math. Imaging Vision* 21, 53–80. <https://doi.org/10.1023/B:JMIV.0000026557.50965.09>
- Bruhn, A., Weickert, J., Schnörr, C., 2005. Lucas-kanade meets horn-schunck: combining local and global optic flow methods. *Int. J. Comput. Vision* 61, 211–231. <https://doi.org/10.1023/B:VISI.0000045324.43199.43>
- Buades, A., Coll, B., Morel, J.M., 2005. A review of image denoising algorithms, with a new one. *Multiscale Model. Simul.* 4, 490–530. <https://doi.org/10.1137/040616024>
- Castellini, P., Martarelli, M., Tomasini, E.P., 2006. Laser doppler vibrometry: development of advanced solutions answering to technology's needs. *Mech. Syst. Signal Process.*, Special Issue: Laser Doppler Vibrometry 20, 1265–1285. <https://doi.org/10.1016/j.ymssp.2005.11.015>
- Çelebi, M., 2000. GPS in dynamic monitoring of long-period structures. *Soil Dyn. Earthquake Eng.* 20, 477–483. [https://doi.org/10.1016/S0267-7261\(00\)00094-4](https://doi.org/10.1016/S0267-7261(00)00094-4)
- Chang, S.G., Yu, B., Vetterli, M., 2000. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Process.* 9, 1532–1546. <https://doi.org/10.1109/83.862633>
- Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A., 2014. Return of the devil in the details: delving deep into convolutional nets. *Arxiv Prepr. Arxiv, Computer Vision and Pattern Recognition* 1405.3531. <https://doi.org/10.48550/arXiv.1405.3531>

- Chen, J.G., Wadhwa, N., Cha, Y.-J., Durand, F., Freeman, W.T., Buyukozturk, O., 2015. Modal identification of simple structures with high-speed video using motion magnification. *J. Sound Vib.* 345, 58–71. <https://doi.org/10/ghtnn9>
- Chen, T., 2017. Preliminary study on the development trend of high-rise building. China Architecture Design&Research Group, Beijing.
- Choi, J., Lee, K., Lee, J., Kang, Y., 2017. Evaluation of quasi-static responses using displacement data from a limited number of points on a structure. *Int. J. Steel Struct.* 17, 1211–1224. <https://doi.org/10.1007/s13296-017-9027-4>
- Chou, J.-Y., Chang, C.-M., Spencer Jr, B.F., 2022. Out-of-plane modal property extraction based on multi-level image pyramid reconstruction using stereophotogrammetry. *Mech. Syst. Signal Process.* 169, 108786. <https://doi.org/10/gn4g87>
- Collier, S., Dare, T., 2022. Accuracy of phase-based optical flow for vibration extraction. *J. Sound Vib.* 117112. <https://doi.org/10.1016/j.jsv.2022.117112>
- Cui, D., Wang, W., He, Y., Zhang, Yan, Zhang, X., Zhao, Y., Zhang, Yukun, 2025. Research on wide-area displacement monitoring based on rotating platform and binocular vision. *Measurement* 253, 117463. <https://doi.org/10.1016/j.measurement.2025.117463>
- Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K., 2009. BM3D image denoising with shape-adaptive principal component analysis. Presented at the SPARS 09 Signal Processing with Adaptive Sparse Structured Representations.
- Dai, W., Huang, D., Cai, C., 2014. Multipath mitigation via component analysis methods for GPS dynamic deformation monitoring. *GPS Solutions* 18, 417–428. <https://doi.org/10.1007/s10291-013-0341-9>

- Danelljan, M., Bhat, G., Shahbaz Khan, F., Felsberg, M., 2017. ECO: efficient convolution operators for tracking. Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6638–6646.
- Danelljan, M., Robinson, A., Shahbaz Khan, F., Felsberg, M., 2016. Beyond correlation filters: learning continuous convolution operators for visual tracking, in: Computer Vision – ECCV 2016, Lecture Notes in Computer Science. Presented at the European Conference on Computer Vision (ECCV), Springer International Publishing, Cham, pp. 472–488. https://doi.org/10.1007/978-3-319-46454-1_29
- Diamond, D.H., Heyns, P.S., Oberholster, A.J., 2017. Accuracy evaluation of sub-pixel structural vibration measurements through optical flow analysis of a video sequence. *Measurement* 95, 166–172. <https://doi.org/10/gmc9mr>
- Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., Smagt, P. van der, Cremers, D., Brox, T., 2015. FlowNet: learning optical flow with convolutional networks, in: 2015 IEEE International Conference on Computer Vision (ICCV). Presented at the 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, Santiago, pp. 2758–2766. <https://doi.org/10/gfrm2v>
- Ellenberg, A., Branco, L., Krick, A., Bartoli, I., Kontsos, A., 2015. Use of unmanned aerial vehicle for quantitative infrastructure evaluation. *J. Infrastruct. Syst.* 21, 4014054. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000246](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000246)
- Erdoğan, H., Güllal, E., 2013. Ambient vibration measurements of the bosphorus suspension bridge by total station and GPS. *Exp. Tech.* 37, 16–23. <https://doi.org/10.1111/j.1747-1567.2011.00723.x>
- Fan, L., Zhang, F., Fan, H., Zhang, C., 2019. Brief review of image denoising techniques. *Visual Comput. Ind., Biomed., Art* 2, 7. <https://doi.org/10.1186/s42492-019-0016-7>

- Field, D.J., 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Amer. A* 4, 2379. <https://doi.org/10.1364/JOSAA.4.002379>
- Fleet, D.J., Jepson, A.D., 1990. Computation of component image velocity from local phase information. *Int. J. Comput. Vision* 5, 77–104. <https://doi.org/10/d8hwq3>
- Fu, X., Li, H.-N., 2015. Comparative analysis for wind and rain-induced responses of structures with normal wind and typhoon spectra. *Zhendong Yu Chongji/j. Vib. Shock* 34, 7–10 and 22. <https://doi.org/10.13465/j.cnki.jvs.2015.11.002>
- Garg, P., Moreu, F., Ozdagli, A., Taha, M.R., Mascareñas, D., 2019. Noncontact dynamic displacement measurement of structures using a moving laser doppler vibrometer. *J. Bridge Eng.* 24, 4019089. <https://doi.org/10/gkkxrm>
- Gautama, T., Van Hulle, M.A., 2002. A phase-based approach to the estimation of the optical flow field using spatial filtering. *IEEE Trans. Neural Netw.* 13, 1127–1136. <https://doi.org/10/cgk6cv>
- Giachetti, A., 2000. Matching techniques to compute image motion. *Image Vision Comput.* 18, 247–260. [https://doi.org/10.1016/S0262-8856\(99\)00018-9](https://doi.org/10.1016/S0262-8856(99)00018-9)
- Gilboa, G., Osher, S., 2009. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* 7, 1005–1028. <https://doi.org/10.1137/070698592>
- Guan, S., Rice, J.A., Li, C., Li, Y., Wang, G., 2017. Structural displacement measurements using DC coupled radar with active transponder. *Struct. Control Health Monit.* 24, e1909. <https://doi.org/10.1002/stc.1909>
- Guo, L., Guo, W., Chen, D., Duan, B., Shi, Z., 2025. Structural vibration measurement based on improved phase-based motion magnification and deep learning. *Mech. Syst. Signal Process.* 224, 111945. <https://doi.org/10.1016/j.ymssp.2024.111945>

- He, Y., 2020. Study of typhoon-induced dynamic responses of a super tall building based on structural health monitoring system. Huazhong University of Science & Technology, Wuhan.
- He, Z.-Z., Zhang, L.-X., Gao, H.-G., Wang, H.-S., Pan, P., 2023. Estimation of the displacement time history of high-rise building structures using limited measurement data and structural information. *Mech. Syst. Signal Process.* 202, 110716. <https://doi.org/10.1016/j.ymssp.2023.110716>
- Horn, B.K.P., Schunck, B.G., 1981. Determining optical flow. *Artif. Intell.* 17, 185–203. [https://doi.org/10.1016/0004-3702\(81\)90024-2](https://doi.org/10.1016/0004-3702(81)90024-2)
- Hoult, N.A., Andy Take, W., Lee, C., Dutton, M., 2013. Experimental accuracy of two dimensional strain measurements using digital image correlation. *Eng. Struct.* 46, 718–726. <https://doi.org/10.1016/j.engstruct.2012.08.018>
- Hu, W.-H., Xu, Z.-M., Bian, X.-H., Tang, D.-H., Lu, W., Li, C., Teng, J., Cunha, Á., 2022. Operational modal analysis and continuous dynamic monitoring of high-rise building based on wireless distributed synchronized data acquisition system. *Struct. Control Health Monit.* 29, e3063. <https://doi.org/10.1002/stc.3063>
- Hu, W.-H., Xu, Z.-M., Liu, M.-Y., Tang, D.-H., Lu, W., Li, Z.-H., Teng, J., Han, X.-H., Said, S., Rohrmann, R.G., 2020. Estimation of the lateral dynamic displacement of high-rise buildings under wind load based on fusion of a remote sensing vibrometer and an inclinometer. *Remote Sens.* 12, 1120. <https://doi.org/10.3390/rs12071120>
- Hu, Y.-D., Xia, Q., Hou, R., Xia, Y., Yan, J., 2021. Computer vision-based displacement measurement with m-sequence target. *Smart Struct. Syst. Int. J.* 27, 537–546.
- Huang, Y., Wei, T., Zhou, Z., Zhang, Y., Chen, G., Xiao, H., 2010. An extrinsic fabry–perot interferometer-based large strain sensor with high resolution. *Meas. Sci. Technol.* 21, 105308. <https://doi.org/10.1088/0957-0233/21/10/105308>

- Hui, T.-W., Loy, C.C., 2020. LiteFlowNet3: resolving correspondence ambiguity for more accurate optical flow estimation, in: *Computer Vision – ECCV 2020, Lecture Notes in Computer Science*. Presented at the European Conference on Computer Vision (ECCV), Springer International Publishing, Cham, pp. 169–184. https://doi.org/10.1007/978-3-030-58565-5_11
- Hui, T.-W., Tang, X., Loy, C.C., 2018. LiteFlowNet: a lightweight convolutional neural network for optical flow estimation, in: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Presented at the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Salt Lake City, UT, USA, pp. 8981–8989. <https://doi.org/10/gf9zhc>
- Hur, J., Roth, S., 2017. MirrorFlow: exploiting symmetries in joint optical flow and occlusion estimation, in: *2017 IEEE International Conference on Computer Vision (ICCV)*. Presented at the 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, Venice, pp. 312–321. <https://doi.org/10/gfvjq4>
- Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., Brox, T., 2017. FlowNet 2.0: evolution of optical flow estimation with deep networks, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI, pp. 1647–1655. <https://doi.org/10/gf2j38>
- Kijewski-Correa, T., Kilpatrick, J., Kareem, A., Kwon, D.-K., Bashor, R., Kochly, M., Young, B.S., Abdelrazaq, A., Galsworthy, J., Isyumov, N., Morrish, D., Sinn, R.C., Baker, W.F., 2006. Validating wind-induced response of tall buildings: synopsis of the chicago full-scale monitoring program. *J. Struct. Eng.* 132, 1509–1523. [https://doi.org/10.1061/\(ASCE\)0733-9445\(2006\)132:10\(1509\)](https://doi.org/10.1061/(ASCE)0733-9445(2006)132:10(1509))
- Kim, J., Kim, K., Sohn, H., 2014. Autonomous dynamic displacement estimation from data fusion of acceleration and intermittent displacement measurements. *Mech. Syst. Signal Process.* 42, 194–205. <https://doi.org/10/f5n7jq>

- Kong, Y., Lee, S., Miao, Y., Nam, H., Young Jeon, J., Park, G., 2025. Phase nonlinearity–weighted optical flow for enhanced full-field displacement measurement and vibration imaging. *Mech. Syst. Signal Process.* 223, 111933. <https://doi.org/10.1016/j.ymssp.2024.111933>
- Kovesi, P., 1999. Phase preserving denoising of images. *Signal* 4.
- Li, B., Chan, P.H., Baris, G., Higgins, M.D., Donzella, V., 2022. Analysis of automotive camera sensor noise factors and impact on object detection. *IEEE Sensors J.* 22, 22210–22219. <https://doi.org/10.1109/JSEN.2022.3211406>
- Li, C., Peng, Z., Huang, T.-Y., Fan, T., Wang, F.-K., Horng, T.-S., Muñoz-Ferreras, J.-M., Gómez-García, R., Ran, L., Lin, J., 2017. A review on recent progress of portable short-range noncontact microwave radar systems. *IEEE Trans. Microw. Theory Techn.* 65, 1692–1706. <https://doi.org/10.1109/TMTT.2017.2650911>
- Li, G., Zhang, Y., Xu, Z., Qin, M., He, M., Meng, A., 2025. Marker-free full-field approach for structural 3D vibration response measurement based on gaussian process regression and pyramidal optical flow. *J. Sound Vib.* 119133. <https://doi.org/10.1016/j.jsv.2025.119133>
- Li, H.-N., Li, D.-S., Song, G.-B., 2004. Recent applications of fiber optic sensors to health monitoring in civil engineering. *Eng. Struct.* 26, 1647–1657. <https://doi.org/10.1016/j.engstruct.2004.05.018>
- Li, J., Neumann, H., Ramalingam, R., 2015. Design, fabrication, and testing of fiber bragg grating sensors for cryogenic long-range displacement measurement. *Cryogenics* 68, 36–43. <https://doi.org/10.1016/j.cryogenics.2015.02.002>
- Li, Q.S., Zhi, L.H., Yi, J., To, A., Xie, J., 2014. Monitoring of typhoon effects on a super-tall building in hong kong: monitoring of typhoon effects. *Struct. Control Health Monit.* 21, 926–949. <https://doi.org/10/f5w6st>

- Li, R., Tan, Y., Li, T., Liu, H., Huang, B., Song, H., Zhou, Z., Cai, L., 2018. A temperature-insensitive FBG displacement sensor with a 10-nanometer-grade resolution. *IEICE Electron. Express* 15, 20180694–20180694. <https://doi.org/10.1587/elex.15.20180694>
- Li, S., Xiong, Y., Shen, X., Peng, Z., 2022. Multi-scale and full-field vibration measurement via millimetre-wave sensing. *Mech. Syst. Signal Process.* 177, 109178. <https://doi.org/10.1016/j.ymssp.2022.109178>
- Li, X., Zhang, Y., Wen, G., Zhang, X., Wu, L., 2025. High-frequency sub-pixel vibration measurement based on compressed sensing. *IEEE Sensors J.* 1–1. <https://doi.org/10.1109/JSEN.2025.3556247>
- Li, Y., 2021. Dynamic hybrid monitoring theory and approach for bridges. Tongji University, Shanghai.
- Li, Y., Tian, W., Xia, Y., Sun, L., 2023. Principal force pattern and impulse response mode for structural equivalent force estimation and full-field response reconstruction. *Mech. Syst. Signal Process.* 200, 110660. <https://doi.org/10.1016/j.ymssp.2023.110660>
- Liu, G., Li, M.Z., Mao, Z., Yang, Q.S., 2022. Structural motion estimation via hilbert transform enhanced phase-based video processing. *Mech. Syst. Signal Process.* 166, 108418. <https://doi.org/10/gm4tt5>
- Liu, M., 2020. Research on estimation method of horizontal displacement of high-rise buildings under wind loads: based on the fusion of laser vibrometer and inclinometer. Harbin Institute of Technology, Harbin.
- Liu, S., Yu, L., Niu, W., Wang, J., Zhong, Z., Huang, J., Shan, M., 2024. Fast and accurate visual vibration measurement via derivative-enhanced phase-based optical flow. *Mech. Syst. Signal Process.* 209, 111089. <https://doi.org/10.1016/j.ymssp.2023.111089>

- Liu, T., Yang, B., Zhang, Q., 2017. Health monitoring system developed for tianjin 117 high-rise building. *J. Aerosp. Eng.* 30, B4016004. [https://doi.org/10.1061/\(ASCE\)AS.1943-5525.0000602](https://doi.org/10.1061/(ASCE)AS.1943-5525.0000602)
- Luan, L., Zheng, J., Wang, M.L., Yang, Y., Rizzo, P., Sun, H., 2021. Extracting full-field subpixel structural displacements from videos via deep learning. *J. Sound Vib.* 505, 116142. <https://doi.org/10/gkzcx8>
- Lukas, J., Fridrich, J., Goljan, M., 2006. Digital camera identification from sensor pattern noise. *IEEE Trans. Inf. Forensics Security* 1, 205–214. <https://doi.org/10.1109/TIFS.2006.873602>
- Luo, K., Kong, X., Wang, X., Jiang, T., Frøseth, G.T., Rønnquist, A., 2023. Cable vibration measurement based on broad-band phase-based motion magnification and line tracking algorithm. *Mech. Syst. Signal Process.* 200, 110575. <https://doi.org/10.1016/j.ymssp.2023.110575>
- Luo, Y., Chen, J., Xi, W., Zhao, P., Qiao, X., Deng, X., Liu, Q., 2016. Analysis of tunnel displacement accuracy with total station. *Measurement* 83, 29–37. <https://doi.org/10.1016/j.measurement.2016.01.025>
- Luo, Y., Chen, Y., Wan, H.-P., Yu, F., Shen, Y., 2021. Development of laser-based displacement monitoring system and its application to large-scale spatial structures. *J. Civ. Struct. Health Monit.* 11, 381–395. <https://doi.org/10.1007/s13349-020-00459-4>
- Luo, Z., Merainani, B., Baltazart, V., Zhang, Q., Döhler, M., 2025. Subpixel motion estimation for video-based target-free vibration monitoring under complex environmental conditions. *Mech. Syst. Signal Process.* 226, 112342. <https://doi.org/10.1016/j.ymssp.2025.112342>
- Ma, Z., Choi, J., Liu, P., Sohn, H., 2022. Structural displacement estimation by fusing vision camera and accelerometer using hybrid computer vision algorithm and

- adaptive multi-rate kalman filter. *Autom. Constr.* 140, 104338. <https://doi.org/10.1016/j.autcon.2022.104338>
- Ma, Z., Choi, J., Yang, L., Sohn, H., 2023. Structural displacement estimation using accelerometer and FMCW millimeter wave radar. *Mech. Syst. Signal Process.* 182, 109582. <https://doi.org/10.1016/j.ymssp.2022.109582>
- Ma, Z., Chung, J., Liu, P., Sohn, H., 2021. Bridge displacement estimation by fusing accelerometer and strain gauge measurements. *Struct. Control Health Monit.* 28, e2733. <https://doi.org/10.1002/stc.2733>
- Maggioni, M., Katkovnik, V., Egiazarian, K., Foi, A., 2013. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Trans. Image Process.* 22, 119–133. <https://doi.org/10.1109/TIP.2012.2210725>
- Mallat, S.G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 674–693. <https://doi.org/10.1109/34.192463>
- Martini, A., Tronci, E.M., Feng, M.Q., Leung, R.Y., 2022. A computer vision-based method for bridge model updating using displacement influence lines. *Eng. Struct.* 259, 114129. <https://doi.org/10.1016/j.engstruct.2022.114129>
- McCormick, N., Lord, J., 2012. Digital image correlation for structural measurements. *Proc. Inst. Civ. Eng. Civ. Eng.* 165, 185–190. <https://doi.org/10.1680/cien.11.00040>
- Miao, Y., Jeon, J.Y., Kong, Y., Park, G., 2022. Phase-based displacement measurement on a straight edge using an optimal complex gabor filter. *Mech. Syst. Signal Process.* 164, 108224. <https://doi.org/10/gmf6wd>
- Miao, Y., Kong, Y., Nam, H., Lee, S., Park, G., 2023. Phase-based vibration imaging for structural dynamics applications: marker-free full-field displacement

- measurements with confidence measures. *Mech. Syst. Signal Process.* 198, 110418. <https://doi.org/10.1016/j.ymssp.2023.110418>
- Ministry of Housing and Urban-Rural Development of the People's Republic of China, 2010. Technical specification for concrete structures of tall building.
- Ministry of Transport of the People's Republic of China, 2024. 2023 statistical bulletin on transport industry development. *China Transportation News* 2.
- Ministry of Transport of the People's Republic of China, 2022. Technical specifications for structural monitoring of highway bridges.
- Nassif, H.H., Gindy, M., Davis, J., 2005. Comparison of laser doppler vibrometer with contact sensors for monitoring bridge deflection and vibration. *NDT & E Int., Structural Faults and Repair* 38, 213–218. <https://doi.org/10.1016/j.ndteint.2004.06.012>
- Ni, Y.-Q., Wang, Y.-W., Liao, W.-Y., Chen, W.-H., 2019. A vision-based system for long-distance remote monitoring of dynamic displacement: experimental verification on a supertall structure. *Smart Struct. Syst.* 24, 769–781. <https://doi.org/10/gkkxpr>
- Nickitopoulou, A., Protopsalti, K., Stiros, S., 2006. Monitoring dynamic and quasi-static deformations of large flexible engineering structures with GPS: accuracy, limitations and promises. *Eng. Struct.* 28, 1471–1482. <https://doi.org/10.1016/j.engstruct.2006.02.001>
- Norton, H.B.N., 1969. Handbook of transducers for electronic measuring systems. Prentice Hall, New Jersey.
- Oh, T.-H., Jaroensri, R., Kim, C., Elgharib, M., Durand, F., Freeman, W.T., Matusik, W., 2018. Learning-based video motion magnification, in: *Computer Vision – ECCV 2018, Lecture Notes in Computer Science*. Presented at the European Conference

- on Computer Vision (ECCV), Springer International Publishing, Cham, pp. 663–679. https://doi.org/10.1007/978-3-030-01225-0_39
- Pan, B., Xie, H., Guo, Z., Hua, T., 2007. Full-field strain measurement using a two-dimensional savitzky-golay digital differentiator in digital image correlation. *Opt. Eng.* 46, 33601. <https://doi.org/10.1117/1.2714926>
- Peng, Z., Li, J., Hao, H., Zhong, Y., 2024. Smart structural health monitoring using computer vision and edge computing. *Eng. Struct.* 319, 118809. <https://doi.org/10.1016/j.engstruct.2024.118809>
- Pitas, I., Venetsanopoulos, A.N., 2013. Nonlinear digital filters: principles and applications. Springer Science & Business Media.
- Psimoulis, P.A., Stiros, S.C., 2013. Measuring deflections of a short-span railway bridge using a robotic total station. *J. Bridge Eng.* 18, 182–185. <https://doi.org/10/gkxrf>
- Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. *Physica D* 60, 259–268. [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F)
- Sánchez Pérez, J., Meinhardt-Llopis, E., Facciolo, G., 2013. TV-L1 optical flow estimation. *Image Process. Line* 3, 137–150. <https://doi.org/10/gcpx2x>
- Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., Westling, P., 2014. High-resolution stereo datasets with subpixel-accurate ground truth, in: *Pattern Recognition, Lecture Notes in Computer Science*. Presented at the German Conference on Pattern Recognition, Springer International Publishing, Cham, pp. 31–42. https://doi.org/10.1007/978-3-319-11752-2_3
- Schreier, H.W., Braasch, J.R., Sutton, M.A., 2000. Systematic errors in digital image correlation caused by intensity interpolation. *Opt. Eng.* 39, 2915–2921. <https://doi.org/10.1117/1.1314593>

- Schreier, H.W., Sutton, M.A., 2002. Systematic errors in digital image correlation due to undermatched subset shape functions. *Exp. Mech.* 42, 303–310. <https://doi.org/10.1007/BF02410987>
- Shan, M., Liu, S., Xiong, X., Liu, Y., Niu, W., Yu, L., Zhong, Z., 2024. Camera-based vibration sensor using time-directional division-based motion estimation without unwrapping. *IEEE Sensors J.* 1–1. <https://doi.org/10.1109/JSEN.2024.3404996>
- Shao, Y., Li, L., Li, J., An, S., Hao, H., 2021. Computer vision based target-free 3D vibration displacement measurement of structures. *Eng. Struct.* 246, 113040. <https://doi.org/10/gn9mhc>
- Shao, Y., Li, L., Li, J., Li, Q., An, S., Hao, H., 2024. 3D displacement measurement using a single-camera and mesh deformation neural network. *Eng. Struct.* 318, 118767. <https://doi.org/10.1016/j.engstruct.2024.118767>
- Spencer, B.F., Hoskere, V., Narazaki, Y., 2019. Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering* 5, 199–222. <https://doi.org/10/gfz2s7>
- Sutour, C., Deledalle, C.-A., Aujol, J.-F., 2014. Adaptive regularization of the NL-means: application to image and video denoising. *IEEE Trans. Image Process.* 23, 3506–3521. <https://doi.org/10.1109/TIP.2014.2329448>
- Tang, D., 2024. Wind-induced time-variant dynamic characteristics tracking and variation mechanism of super high-rise structures. Harbin Institute of Technology, Harbin.
- Tao, S., Dong, X., Lai, B., 2016. Temperature-insensitive fiber bragg grating displacement sensor based on a thin-wall ring. *Opt. Commun.* 372, 44–48. <https://doi.org/10.1016/j.optcom.2016.03.092>

- Teng, J., Wang, Y., Xia, Y., Hu, W., 2025. Robust discriminative correlation-based full-field motion estimation of large-scale structures using a single video camera. *Eng. Struct.* 334, 120224. <https://doi.org/10.1016/j.engstruct.2025.120224>
- Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., Lin, C.-W., 2020. Deep learning on image denoising: an overview. *Neural Networks* 131, 251–275. <https://doi.org/10.1016/j.neunet.2020.07.025>
- Tomasi, C., Manduchi, R., 1998. Bilateral filtering for gray and color images, in: *Sixth International Conference on Computer Vision (IEEE Cat. No.98ch36271)*. Presented at the Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), pp. 839–846. <https://doi.org/10.1109/ICCV.1998.710815>
- Tong, W., 2005. An evaluation of digital image correlation criteria for strain mapping applications. *Strain* 41, 167–175. <https://doi.org/10.1111/j.1475-1305.2005.00227.x>
- Valente, N.A., do Cabo, C.T., Mao, Z., Niezrecki, C., 2022. Quantification of phase-based magnified motion using image enhancement and optical flow techniques. *Measurement* 189, 110508. <https://doi.org/10.1016/j.measurement.2021.110508>
- Vendroux, G., Knauss, W.G., 1998. Submicron deformation field measurements: part 2. Improved digital image correlation. *Exp. Mech.* 38, 86–92. <https://doi.org/10.1007/BF02321649>
- Wadhwa, N., Rubinstein, M., Durand, F., Freeman, W.T., 2013. Phase-based video motion processing. *ACM Trans. Graphics* 32, 80:1-80:10. <https://doi.org/10/gbdg5f>
- Wang, G., Zhang, F., Feng, Z., 2023. Practical algorithm for dynamic response factor of tall buildings based on bending-shearing beam model and von karman's wind speed spectrum. *J. Tongji Univ.* 51. <https://doi.org/10.11908/j.issn.0253-374x.21563>

- Wang, M., Kei Ao, W., Bownjohn, J., Xu, F., 2022. Completely non-contact modal testing of full-scale bridge in challenging conditions using vision sensing systems. *Eng. Struct.* 272, 114994. <https://doi.org/10.1016/j.engstruct.2022.114994>
- Wang, Y., Hu, W., Teng, J., Xia, Y., 2024a. Motion estimation from noisy data with unknown distributions using multi-frame phase-preserving denoising. *Mech. Syst. Signal Process.* 206, 110924. <https://doi.org/10.1016/j.ymssp.2023.110924>
- Wang, Y., Hu, W., Teng, J., Xia, Y., 2024b. Full-field displacement measurement of long-span bridges using one camera and robust self-adaptive complex pyramid. *Mech. Syst. Signal Process.* 215, 111451. <https://doi.org/10.1016/j.ymssp.2024.111451>
- Wang, Y., Hu, W., Teng, J., Xia, Y., 2023. Phase-based motion estimation in complex environments using the illumination-invariant log-gabor filter. *Mech. Syst. Signal Process.* 186, 109847. <https://doi.org/10.1016/j.ymssp.2022.109847>
- Weng, Y., Quek, S.-T., Ker-Wei Yeoh, J., 2025. Robust vision-based sub-pixel level displacement measurement using a complementary strategy. *Mech. Syst. Signal Process.* 223, 111898. <https://doi.org/10.1016/j.ymssp.2024.111898>
- Weng, Y., Shan, J., Lu, Z., Lu, X., Spencer Jr., B.F., 2021. Homography-based structural displacement measurement for large structures using unmanned aerial vehicles. *Comput.-Aided Civ. Infrastruct. Eng.* 36, 1114–1128. <https://doi.org/10/gk7w9n>
- Wu, J., Xu, H., Zhang, Q., 2019. Dynamic analysis of parameters of shanghai tower in normal wind and typhoon. *J. Tongji Univ.* 47, 6.
- Wu, J.R., Liu, P.F., Li, Q.S., 2007. Effects of amplitude-dependent damping and time constant on wind-induced responses of super tall building. *Comput. Struct.* 85, 1165–1176. <https://doi.org/10.1016/j.compstruc.2007.01.012>
- Xia, Q., Wu, W., Li, F., Xia, Y., Ding, X., Lam, W.H.K., Chung, W., Xu, Y., 2021. System design and demonstration of performance monitoring of a butterfly-

- shaped arch footbridge. *Struct. Control Health Monit.* 28, e2738. <https://doi.org/10/gmp9vg>
- Xia, Y., Chen, B., Zhou, X., Xu, Y., 2013. Field monitoring and numerical analysis of Tsing Ma suspension bridge temperature behavior. *Struct. Control Health Monit.* 20, 560–575. <https://doi.org/10.1002/stc.515>
- Xie, K., Cheng, L., 2025. A novel two-dimensional hilbert transform method for structural vibration measurement and modal analysis. *Mech. Syst. Signal Process.* 225, 112324. <https://doi.org/10.1016/j.ymssp.2025.112324>
- Xie, X., Zhang, Y., Ling, X., Wang, X., 2019. A novel extended phase correlation algorithm based on log-gabor filtering for multimodal remote sensing image registration. *Int. J. Remote Sens.* 40, 5429–5453. <https://doi.org/10.1080/01431161.2019.1579941>
- Xiong, L., Guo, Y., Jiang, G., Jiang, L., Zhou, X., 2019. Fiber bragg grating displacement sensor with high measurement accuracy for crack monitoring. *IEEE Sensors J.* 19, 10506–10512. <https://doi.org/10.1109/JSEN.2019.2930761>
- Xu, Y., Brownjohn, J., Kong, D., 2018. A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge. *Struct. Control Health Monit.* 25, e2155. <https://doi.org/10.1002/stc.2155>
- Xu, Y., Brownjohn, J.M.W., 2018. Review of machine-vision based methodologies for displacement measurement in civil structures. *J. Civ. Struct. Health Monit.* 8, 91–110. <https://doi.org/10/gg3w3r>
- Xu, Y.L., Xia, Y., 2011. *Structural health monitoring of long-span suspension bridges.* CRC Press, Florida.
- Yang, M., Wu, J., Zhang, Q., 2025. GNSS and accelerometer data fusion by variational bayesian adaptive multi-rate kalman filtering for dynamic displacement

- estimation of super high-rise buildings. *Eng. Struct.* 325, 119396. <https://doi.org/10.1016/j.engstruct.2024.119396>
- Yang, Y., Dorn, C., Farrar, C., Mascareñas, D., 2020. Blind, simultaneous identification of full-field vibration modes and large rigid-body motion of output-only structures from digital video measurements. *Eng. Struct.* 207, 110183. <https://doi.org/10/gh2rcp>
- Yang, Y., Dorn, C., Mancini, T., Talken, Z., Kenyon, G., Farrar, C., Mascareñas, D., 2017. Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. *Mech. Syst. Signal Process.* 85, 567–590. <https://doi.org/10/gnpkpt>
- Yi, J., Zhang, J.W., Li, Q.S., 2013. Dynamic characteristics and wind-induced responses of a super-tall building during typhoons. *J. Wind Eng. Ind. Aerodyn.* 121, 116–130. <https://doi.org/10/f5f9f4>
- Yi, Z., Kuang, C., Wang, Y., Yu, W., Cai, C., Dai, W., 2018. Combination of high- and low-rate GPS receivers for monitoring wind-induced response of tall buildings. *Sens.* 18, 4100. <https://doi.org/10/gkkxp8>
- Yoon, H., Elanwar, H., Choi, H., Golparvar-Fard, M., Spencer Jr., B.F., 2016. Target-free approach for vision-based structural system identification using consumer-grade cameras. *Struct. Control Health Monit.* 23, 1405–1416. <https://doi.org/10/f9bjpr>
- Yu, J., Meng, X., Yan, B., Xu, B., Fan, Q., Xie, Y., 2020. Global navigation satellite system-based positioning technology for structural health monitoring: a review. *Struct. Control Health Monit.* 27. <https://doi.org/10/gkkxq4>
- Yu, J., Zhu, P., Xu, B., Meng, X., 2017. Experimental assessment of high sampling-rate robotic total station for monitoring bridge dynamic responses. *Measurement* 104, 60–69. <https://doi.org/10/gkkxqv>

- Yu, L., Xiong, C., Chen, W., Gao, Y., Ye, Z., Shi, Q., 2021. A combined algorithm for denoising GNSS-RTK positioning solutions with application to displacement monitoring of a super-high-rise building. *Meas. Sci. Technol.* 32, 115121. <https://doi.org/10.1088/1361-6501/ac0b6d>
- Zhang, B., Ding, X., Werner, C., Tan, K., Zhang, Bin, Jiang, M., Zhao, J., Xu, Y., 2018. Dynamic displacement monitoring of long-span bridges with a microwave radar interferometer. *ISPRS J. Photogramm. Remote Sens.* 138, 252–264. <https://doi.org/10.1016/j.isprsjprs.2018.02.020>
- Zhang, K., Gao, X., Tao, D., Li, X., 2012. Multi-scale dictionary for single image super-resolution, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition. Presented at the 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1114–1121. <https://doi.org/10.1109/CVPR.2012.6247791>
- Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* 26, 3142–3155. <https://doi.org/10.1109/TIP.2017.2662206>
- Zhang, K., Zuo, W., Zhang, L., 2018. FFDNet: toward a fast and flexible solution for CNN-based image denoising. *IEEE Trans. Image Process.* 27, 4608–4622. <https://doi.org/10.1109/TIP.2018.2839891>
- Zhang, Lele, Hu, X., Xie, Z., Shi, B., Zhang, Li, Wang, R., 2020. Field measurement study on time-varying characteristics of modal parameters of super high-rise buildings during super typhoon. *J. Wind Eng. Ind. Aerodyn.* 200, 104139. <https://doi.org/10.1016/j.jweia.2020.104139>
- Zhang, W., 2017. Research of key theories and methodologies on bridge hybrid monitoring. Tongji University, Shanghai.

- Zhang, W., Li, H., Wang, J., Hong, Y., Meng, G., 2025. Full-field phase-based vibration measurement and visualisation using many knowledge transfer-assisted optimal log-gabor filters. *Mech. Syst. Signal Process.* 225, 112256. <https://doi.org/10.1016/j.ymssp.2024.112256>
- Zhang, W., Wang, X., Li, H., Zhou, J., Luo, Z., Meng, G., 2024. Vibration measurement from an adaptive phase-based motion estimation using parameter optimised log-gabor filter. *Measurement* 224, 113812. <https://doi.org/10.1016/j.measurement.2023.113812>
- Zhao, F., Wang, Y., Jin, L., Li, X., 2025. Statistics analysis of China's construction industry development in 2024. *J. Eng. Manag.* 39, 1–6. <https://doi.org/10.13991/j.cnki.jem.2025.02.001>
- Zhao, L., Ge, Y., 2020. Emergency measures for vortex-induced vibration of Humen bridge. *Mater. Res.* 9, 25–28.
- Zhao, N., Huang, G., 2020. Wind velocity field simulation based on enhanced closed-form solution of cholesky decomposition. *J. Eng. Mech.* 146, 4019128. [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0001712](https://doi.org/10.1061/(ASCE)EM.1943-7889.0001712)
- Zhou, K., Li, Q.-S., 2022. Investigation of time-varying structural dynamic properties of high-rise buildings under typhoon conditions. *J. Build. Eng.* 46, 103796. <https://doi.org/10.1016/j.jobbe.2021.103796>
- Zhu, Q., Cui, D., Zhang, Q., Du, Y., 2022. A robust structural vibration recognition system based on computer vision. *J. Sound Vib.* 117321. <https://doi.org/10.1016/j.jsv.2022.117321>
- Zhu, S., Zhang, X.-H., Xu, Y.-L., Zhan, S., 2013. Multi-type sensor placement for multi-scale response reconstruction. *Adv. Struct. Eng.* 16, 1779–1797. <https://doi.org/10.1260/1369-4332.16.10.1779>

- Zhu, Z., Lu, J., Zhu, S., 2023a. Multi-rate kalman filtering for structural dynamic response reconstruction by fusing multi-type sensor data with different sampling frequencies. *Eng. Struct.* 293, 116573. <https://doi.org/10.1016/j.engstruct.2023.116573>
- Zhu, Z., Zhu, S., Wang, Y.-W., Ni, Y.-Q., 2023b. Structural dynamic response reconstruction with multi-type sensors, unknown input, and rank deficient feedthrough matrix. *Mech. Syst. Signal Process.* 187, 109935. <https://doi.org/10.1016/j.ymssp.2022.109935>
- Zoran, D., Weiss, Y., 2011. From learning models of natural image patches to whole image restoration, in: 2011 International Conference on Computer Vision. Presented at the 2011 International Conference on Computer Vision, pp. 479–486. <https://doi.org/10.1109/ICCV.2011.6126278>