Video Camera-based Vibration Measurement of Infrastructure

by

Justin Gejune Chen

B.S. Physics
California Institute of Technology, 2009

S.M. Civil and Environmental Engineering
Massachusetts Institute of Technology, 2013

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in the field of Structures and Materials

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

© Massachusetts Institute of Technology 2016. All rights reserved.

Author . . . .
Department of Civil and Environmental Engineering
May 19, 2016

Certified by. . . . .
Oral Büyüköztürk
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by . . . .
Heidi M. Nepf
Donald and Martha Harleman Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Video Camera-based Vibration Measurement of Infrastructure

by

Justin Gejune Chen

Abstract

Functional infrastructure — including transportation, energy, and buildings and other facilities — is key to the economic production of a country and the daily lives of its people. Due to deterioration and potential damage, timely inspections and repairs are necessary to keep infrastructure functioning at full capacity. Visual testing is one of the oldest and most widely used methods for condition assessment; however, this technique is limited since it is based on an inspector's subjective rating rather than on an objective quantitative measure of structural condition.

Quantitative alternatives to visual testing have emerged over time. These include vibration analysis, in which a structure's operational resonant frequencies and mode shapes are measured and compared against a healthy baseline to detect changes. Typically, contact sensors such as accelerometers have been used to measure vibrations. When physical access to a structure is limited or the placement of contact sensors is too time-consuming, new technologies that allow for non-contact measurements can be used. Video cameras, where each pixel is effectively a sensor, can remotely collect a large amount of data from a structure. The challenge is then to interpret these videos into quantitative vibration data.

In this research, newly developed computer vision techniques for analyzing small motions in videos and their application to the vibration measurement and condition assessment of infrastructure are presented. These techniques allow for qualitative visualizations of normally imperceptible motions as a form of enhanced visual testing, and quantitative measurements of the displacements and vibrations of structures as a basis for condition assessment. Computer vision algorithms for processing video are described and the technique is experimentally validated against traditional sensors. The methodology is demonstrated with a series of laboratory measurements of simple representative structures and field measurements of civil infrastructure, including the WWI Memorial Bridge in Portsmouth, New Hampshire.

Thesis Supervisor: Oral Büyüköztürk
Title: Professor of Civil and Environmental Engineering

Acknowledgments

My advisor, Prof. Oral Buyukozturk has been the best mentor I could have ever asked for. He has been supportive through the different paths of research and kept me on the right track with regards to lines of work to pursue. I'm especially thankful for his ability to give the right mix of encouragement and pressure.

Rob Haupt is responsible for setting me on this path, and introducing me to Prof. Buyukozturk. I've learned so much from him, from working at MIT Lincoln Laboratory to the master's thesis and this doctoral work. I hope we get to keep working together on interesting problems with cool technology.

It has been great collaborating with Prof. Bill Freeman on this research project and seeing it develop over the years, and I'm thankful for his enthusiasm in finding interesting and sometimes surprising applications for motion magnification.

My other committee members have also been an invaluable resource over many years of committee meetings. I thank Prof. Jerome Connor and Prof. Nafi Toksoz for their insightful comments, guidance, and suggestions. For all of my committee members, I'm especially grateful for them being able to find time in their busy schedules to always converge on a time to meet.

Much of this research was done in collaboration with Neal Wadhwa and Abe Davis; without them, I suspect none of it would have been completed. I enjoyed working with them and discussing the different potential applications for our research. Even if we end up going different paths with work, I hope we can still find some way to collaborate.

Prof. Frédo Durand and Dr. John Fisher were also extremely helpful in various aspects of research. Others in CSAIL were also great to work with and bounce ideas off of including Michael Rubinstein, Zoran Dzunic, Hossein Mobahi, Katie Bouman, Tianfan Xue, and Donglai Wei.

The research sponsor, Shell, has been instrumental in getting this work off the ground, and are responsible for the great collaboration and original matchmaking between CEE and CSAIL in this project. This includes Chief Scientists Dr. Dirk Smit

and Dr. Sergio Kapusta for their intense interest and help in guiding the research, Dr. Lorna Ortiz-Soto, Dr. Keng Yap, and Dr. Yile Li as project managers, and Dr. Jonathan Kane as the Shell-MIT liaison.

A chance meeting at a conference in Berlin with Prof. Erin Bell at the University of New Hampshire was responsible for many of the best examples and results in this thesis, the Portsmouth bridge study. I am eternally grateful for her invitation and arrangements with the appropriate authorities to make measurements of the bridge. I also thank Travis Smith for joining me in the bridge measurements.

I'd like to thank others that have assisted including Dr. Keng Yap for the NASA 8-bay structure data, Scott Wade for access to the MIT Green building roof, Maryse Vachon for the 3D printed beam, Bing Li and Dr. Stephen Morgan for the granite example, Dr. Kunal Kupwade-Patil for help with the corroded pipe experiments, and Seth Kinderman for access to the MIT central facilities plant.

My group members have been helpful in many small ways that really add up over the course of a Ph.D thesis. They include, originally Chakrapan Tuakta, Denvid Lau, and Tim Emge; in between, Peter Trocha and Young-Jin Cha; and now James Long, Reza Mohammadi Ghazi, Steve Palkovic, Murat Uzun, and Hao Sun. They have helped with everything from an extra hand when doing experiments, to being around the office and making that time pleasant with banter. Also, thanks to Rebecca Fowler for putting up with all my random requests for help.

I'd like to thank my friends from MIT LL, Caltech, the Devens autocross community, and others for keeping me sane under the pressures of work, and my pet cats Freyr, Loki, and Frieda for being fuzzy comfort.

It's hard to believe that we've been together about as long as I've been working on this research topic; Hannah Chiodo, thank you so much for your support and tolerance of the strange lifestyle that a graduate student leads.

My parents Yihwa Liou and Minder Chen have supported and encouraged all my endeavors, academic or otherwise, and I love them dearly for it. Without them, I surely would not have become the scientist/engineer I am now, and I am eternally grateful for them instilling in me a sense of conscientiousness.

Contents

List of Figures

15

17

19

24

26

List of Tables

Chapter 1

Introduction

1.1 Background

The national economy depends on functioning civil infrastructure to operate, which includes structures such as buildings, bridges, roadways, and pipelines. Damage to these structures may occur due to normal in-service usage, aging, environmental deterioration, or specific adverse events such as earthquakes or other abnormal disturbances. It is already the case that much of the nation's infrastructure is in need of desperate repair. One in nine of the nation's bridges is rated as structurally deficient, the average age of the nation's 607,380 bridges is 42 years, and an estimated \$8 billion extra is necessary to eliminate the nation's bridge backlog by 2028 [1].

The most recent well known example of a massive structural failure is the collapse of the I-35W highway bridge over the Mississippi River in Minneapolis, Minnesota as shown in Figure 1-1. Undersized gusset plates in the connections of the bridge from incorrect design, combined with a poor distribution of a much greater than average load due to construction, led to the sudden collapse [66, 46]. The gusset plates were already visibly bowed at a node on the bridge, during a bridge evaluation done in 2003 by URS Corporation, under contract to Minnesota DOT. The evaluation however, did not make note of the bowed condition of the gusset plates other than the pictures included in the inspection report. Reports on the state of corrosion in the gusset plates from 1994-2006 mentioned the presence of rust and section loss, however the

amount was not quantified and no assessment was made to assess the degree or rate of deterioration [66].
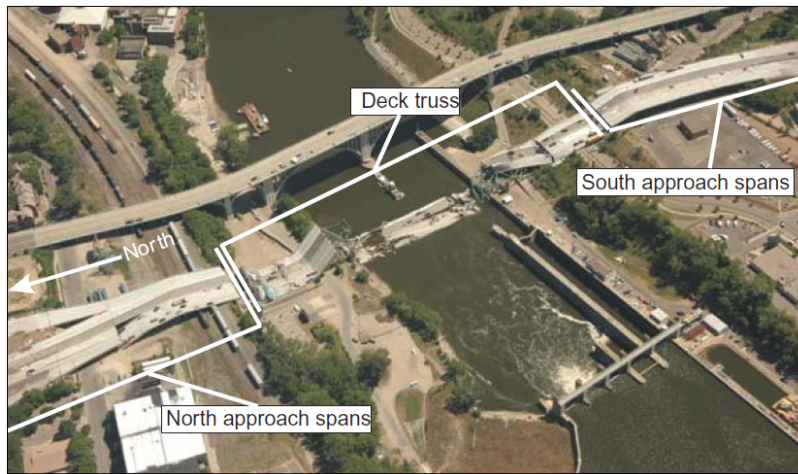
(a)
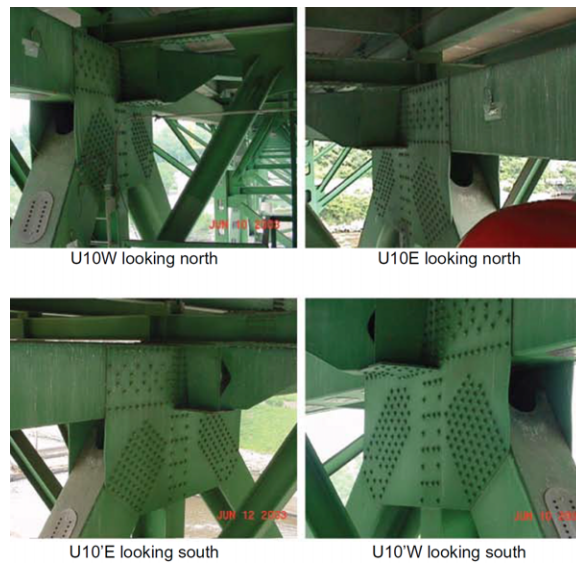
(b)

Figure 1-1: Pictures from the collapse of the I-35W highway bridge over the Mississippi River in Minneapolis, Minnesota showing (a) the bridge after collapse and (b) bowed gusset plates four years before the collapse [66]

Bridges and most other civil infrastructure are for the most part, only visually assessed by inspectors [67, 93]. An inspector will go out to a bridge, have a long checklist of things to inspect in a "hands on" manner, and use their senses of sight, hearing, and touch to assess the condition of the structure, e.g. looking for corrosion,

listening for strange noises, or feeling for excess vibration. Condition ratings for different parts of the structure are simply given on a scale from 1 to 9, where 1 represents imminent failure, 5 is fair, and 9 is excellent condition. Unless more detailed studies are commissioned with quantitative assessments, tracking a bridge's condition over time can become a tedious affair of reading into the inspector's remarks and looking through pictures. A single number for condition of all the girders on a bridge leaves a lot of room for interpretation for the rate of deterioration of a bridge's structural elements, thus quantitative assessment methods are important, especially for tracking a structure's condition over time.

1.1.1 Quantitative Condition Assessment of Structures

For more detailed inspection of structural elements, non-destructive testing (NDT) methods including ultrasound, pulse echo, x-ray, and other methods [83]. Structural health monitoring (SHM) methods using vibration sensors can be used to monitor a structure's condition over time [89, 40, 106, 35]. The goal of SHM is usually to direct inspectors to critical locations at critical times to potential locations of damage, since it can be time consuming to fully inspect a large structure with local NDT methods. Full instrumentation of a structure with typical sensors such as contact accelerometers for SHM can be labor intensive as wiring issues and physical placement can be cumbersome, however systems are usually meant to last for long-term monitoring where the initial time investment pays off. Densely instrumenting a structure can become cost prohibitive and tedious, especially as structural health monitoring algorithms are moving to more statistical and data-driven approaches [90, 61, 36]. Also, when the structure is small compared to the size of an accelerometer, the presence of added mass from accelerometers can affect the result. There are methods for correcting for the effects of accelerometer mass loading, but they are not exact [2, 17]. Non-contact methods of measurement avoid some of these drawbacks and are being researched intensely for the purposes of SHM and NDT.

Non-contact methods typically depend on some form of wave being measured from the structure, in the form of electromagnetic radiation such as radar or light, or

acoustically through sound. Some examples include synthetic aperture radar (SAR) which is an interferometric radar technique that can be used for monitoring bridges or other civil infrastructure [38, 72] or other radar based techniques for imaging of concrete [77, 13, 75, 76, 15, 110, 14, 45]. Laser methods such as laser vibrometry and electronic speckle pattern interferometry have also been studied [11, 37, 21, 22, 91, 96].

Video cameras are an enticing technology for non-contact measurement of structures. They generally measure visible light, and can be easily set up and trained on a structure. They can measure a large scene of interest, and collect high spatial density data, as every pixel collects a time series of intensity values. They can range from precise instruments for high-frequency and high-resolution video to inexpensive units such as those on cell phones or action cameras which can be chosen as necessary for the application. The ubiquity and low cost of video cameras, make them a good candidate for widespread usage in SHM and NDT applications.

1.1.2 Camera Measurement of Vibrations

Video cameras have been used to measure civil infrastructure on a variety of different structures in previous work. Many camera measurements have been made of bridges and the cables on suspension bridges using small lights or paper targets, however this nullifies the ease of instrumentation advantage of the camera, as the targets need to be placed on the structure [101, 57, 53, 28]. Edge detection algorithms have also been used to measure displacements of simple structures [70]. Target-less measurement, a much more ideal use case which does not require any preparation other than the camera itself where virtual sensors are imposed over the video [80], has also been made of bridges and cables [54, 28], and a traffic structure [3]. Regarding long range measurements the resonant frequencies of cables in a cable-stayed bridge were identified from a distance of 850 meters (0.53 miles) with the smallest estimated amplitude being approximately 1/3rd of a pixel [16].

With a single camera, only measurements of in-plane motion can be made, however a stereo camera setup is capable of measuring both in-plane and out-of-plane motion [48]. A newer method of in-plane and out-of-plane measurement with cameras is the

36

time-of-flight camera [55]; however, they currently do not provide enough resolution or speed for typical vibration measurement applications. This thesis focuses on single camera measurements of in-place motions with normal cameras.

Processing of recorded videos to recover quantitative measurements of a structure can be difficult as videos contain an immense amount of data when uncompressed. It is also not necessarily straightforward to interpret brightness changes in pixel values to obtain displacement or motion data of a structure.

Motion Magnification

Recently, new computer vision techniques, collectively called motion magnification, were introduced to magnify imperceptible motions in videos [60, 104, 99, 78, 100, 97]. The most recent motion magnification techniques use a signal processing approach to analyze image motions in a way analogous to an Eulerian framework for fluid flow analysis. They are ideal for visualizing mode shapes because they are capable of detecting small subpixel motions that are present in the videos of vibrating structures. They are able to separate the different modal motions through the use of temporal filtering to produce visualizations of the apparent mode shapes of objects. This works well with an assumption of a stationary camera and a structure that is barely moving. Similar processing was also able to recover sound from a video of an object, being excited by a loudspeaker [31]. These computer vision algorithms allow for different ways of processing videos that have advantages over previous methods, potentially being faster, providing better detectability of smaller motions, and the ability to generate visualizations of the apparent motions.

Research Motivation

Quantitative condition assessment of structures through NDT and SHM requires the accurate and swift collection of data that can be analyzed to determine whether or not a structure's behavior is normal. Contact accelerometers, strain gauges, and other sensors currently used are extremely accurate, but can be labor intensive to use to instrument a structure, especially if dense measurements are necessary. Newer

SHM damage detection algorithms are moving towards data driven approaches that require a massive amount of data not only in time, but also spatially along a structure, which can be difficult to obtain with current measurement techniques. Current camera methods for measuring displacements can provide sufficient spatial resolution and density, however they haven't been widely adopted yet because sensitivity, processing time, and available hardware haven't quite reached acceptable levels; recent research and technological advancements have made good progress. Thus, most current measurement methods are not well matched to the extensive need for data with new SHM algorithms, and more data is better for improved quantitative condition assessment of infrastructure.

1.2   Research Objective

The objective of this research is to develop experimental camera methods and computer vision algorithms for collecting and analyzing videos for the purposes of measuring vibrations. The use of a video camera for vibration measurement represents a unique capability that complements and could replace existing measurement and sensor systems currently deployed for NDT and SHM. Current sensor systems generally rely on contact sensors which only measure the motion locally. The advantage of a single or few cameras at a distance, is thatt hey can collect data from a large portion of the structure, obtaining information from many different locations at once. The camera is also much easier to set up than having to place sensors all over a structure. This is an enormous operational advantage, even if the measurements aren't as accurate. Recent developments in motion magnification and associated computer vision algorithms, advancements in the processing power of computers, and improvements in camera technology make this tractable. Further development of these computer vision algorithms for vibration measurement is a collaborative effort between Prof. Oral Buyukozturk's research group in the Department of Civil and Environmental Engineering and Prof. Bill Freeman's and Prof. Frédo Durand's research groups in CSAIL. The general methodology is as follows. A video camera records a video of a

38

vibrating structure from which displacements and modal characteristics are extracted using computer vision algorithms. The displacements and modal characteristics can be analyzed with system identification, damage detection, or other SHM algorithms for condition assessment of the structure. The system has applications in the monitoring and condition assessment of machinery commonly used in oil and gas, power plants, physical facilities, and other heavy industries, and the monitoring of roadway and railroad bridges, tall buildings, and other forms of civil infrastructure. It has the potential to significantly change the way that these systems are inspected and monitored, and hopefully improve the reliability and condition of civil infrastructure around the world.

1.3   Research Approach

The research approach involves a comprehensive study of the use of a camera and video processing methods as applied to the vibration measurement of civil infrastructure for the purpose of structural health monitoring. We describe the algorithms for processing videos of vibrating objects and structures to extract motion information and the ways in which this motion information can be used to identify the displacements and vibration characteristics of structures. We also describe the use of motion magnification to visualize those small motions and how to produce a plausible simulation of the object under applied forcing. We also discuss a new metric for damage detection in symmetric structures based on a continuous symmetry measure of the resonant mode shapes.

The camera measurement system is experimentally verified with a comparison to traditional vibration measurement methods such as an accelerometer and laser vibrometer. We determine the effects of several parameters on the noise floor of the camera measurement including the local visual contrast in the video, frame rate, distance to the object, lens focal length, and camera motion. For long measurement distances, camera motion can be a significant contributor to the noise floor of the measurement, so we explore methods for measuring and correcting for camera motion.

The proposed method can also be used for sensor fusion of the data from sensors in the field of view of the camera measurement to improve the measurement precision.

A full complement of experiments are conducted demonstrating the camera vibration measurement methodology both in the laboratory and in the field. The laboratory measurements focus on two representative simple structures, the cantilever beam and a circular pipe cross-section. Operational resonant frequencies and mode shapes of these simple structures are determined, and also used for the purposes of damage detection. Damages are simulated by machined defects in a cantilever beam, and induced corrosion damage in a section of steel pipe. We also demonstrate a quasi real-time implementation that is used for measurements of lightweight structures where the attachment of a traditional sensor would alter the dynamics of the structure. Full field displacements are also measured from a granite specimen undergoing a compression test to measure the behavior before, during, and after fracture which involves the interpretation of the displacement measurements to strain values rather than vibration characteristics. Measurements are made in the field of several structures including vibrating machinery and pipes at the MIT physical facilities plant, the antenna tower on top of the MIT's Green Building from approximately 175 meters (574 feet) away, and the tuned mass damper inside the Taipei 101 skyscraper during an earthquake using a video from YouTube.

As a proof of concept for the use of a camera for structural health monitoring of civil infrastructure, we conducted a thorough analysis and measurement of the WWI Memorial Bridge in Portsmouth, New Hampshire, a vertical-lift bridge. Camera measurements are made of the bridge's motions from over 260 feet (80 meters) away and directly cross referenced and confirmed by accelerometer measurements on the bridge. Vibrations and deflections of the bridge as a result of the impact from the lift span and traffic driving over the bridge, are measured. A finite element model of the bridge is created to simulate the response of the bridge due to a lift span impact and a static load at the midspan of the bridge, to compare with the measurements from the real world bridge.

1.4 Thesis Organization

The thesis chapters are organized as follows:

Chapter 1: General background information, research objective, research approach, and thesis organization.

Chapter 2: A new metric for mode shape damage detection based on a continuous symmetry measure. Demonstration on several numerical models including the ASCE benchmark structure, and an experimental study on sections of corroded pipe.

Chapter 3: A description of how to measure vibrations from video, and visualizations created from motion signals extracted from video. This includes an explanation of the full workflow in processing videos to vibration information, from the computer vision algorithms used to process the small motions of objects in videos, to obtaining displacements, and the signal processing and vibration analysis methods to process the displacements into information such as resonant frequencies and mode shapes. This chapter also includes a description of the motion magnification visualization and the use of image-space modal bases for plausible manipulation of objects from the measured motion data.

Chapter 4: Verification of the camera measurement method by comparison to existing methods and characterization of the noise floor of the camera measurement under the effects of different parameters. Demonstration of a method for sensor fusion and motion compensation with external sensors. Discussion of the resulting total noise floor from the camera measurement considering intrinsic noise, measurement parameters, and camera motion.

Chapter 5: Laboratory measurements with representative simple structures such as a cantilever beam and a pipe cross-section, and an example of a measurement of damage. Demonstration of a quasi real-time implementation of

the method. Experimental study on a corroded pipe section, making use of the continuous symmetry measure. Strain measurement of a granite specimen undergoing a compression test.

Chapter 6: Field measurements of machinery in MIT facilities, a long range measurement of the MIT Green Building from the Stata Center, and measurement of the tuned mass damper in the Taipei 101 skyscraper.

Chapter 7: A comprehensive study of the WWI Memorial Bridge in Portsmouth, NH including a field measurement, modeling study, and comparison of the measurement and model.

Chapter 8: Summary of the work presented and discussion of advantages and limitations, the presented applications, and use of the methodology. Conclusions and future work.

The appendices include the following:

Appendix A: Abbreviations and acronyms used in the thesis.

Appendix B: A detailed description of the phase-based motion magnification algorithm, reprinted with permission from [97].

Appendix C: Extra experiments and experimental data.

Appendix D: A catalog of videos included in this thesis.

Chapter 2

A Symmetry Measure for Damage Detection

2.1 Introduction

Structural mode shapes have been commonly used as a feature for structural health monitoring. Methods such as mode shape difference, slope, and curvature are ways to detect damage in beam-like structures using mode shapes and frequency response function data [62, 69, 74]. Direct comparisons of the differences in mode shapes work well in laboratory testing situations; however they may not be particularly informative in the field due to measurement noise or environmental effects. For mode shape curvature in particular, prior knowledge about the structure may not be necessary as long as it is similar to a beam; damage is approximated as a discontinuous change in the EI or flexural rigidity term of the beam equation, which manifests in the second derivative, or curvature of the mode shape. When structures are more complex beyond that of a simple beam, this method is less applicable. Other mode shape methods such as modal strain energy requires a detailed model of the structure which limits the situations in which it can be used [82].

Data-based methods for structural health monitoring entirely eschew models of structures for an approach that uses data from sensors in a presumed healthy state of the structure as a baseline and looks for statistically significant differences in fea-

tures extracted from the measured structural responses for indications of change or damage to the structure [39]. Examples of data-based methods of damage detection include ones based on statistical process control [88] and the one-class support vector machine [61] among others. A key requirement for these statistical methods is that the features are indicative of damage or change in the structure, otherwise the statistical methods won't work. Features can be based on signal characteristics or physics-based information about the structure, and generally a collection of varied features is necessary for the best detection accuracy.

An intrinsic property of many structures, either natural or engineered is symmetry. Usually objects that are symmetric are aesthetically pleasing and also the best design for a structure, as an asymmetry might introduce uneven loading or stresses. Most types of damage to structures occur locally to specific elements of the structure, and generally do not affect its symmetric counterparts elsewhere in the structure. Asymmetric damage in the structure then alters its mode shapes and may cause them to be more asymmetrical. A measurement of how symmetric or asymmetric a structure's mode shapes are, could be a useful indication of an asymmetric or local change in the structure's geometrical or physical properties.

In this chapter, we demonstrate the use of a symmetry metric as a feature for damage detection in structures. The continuous symmetry measure (CSM) was introduced by Zabrodsky et al. as a feature for both the identification of chemical compounds [111], and for image analysis in computer vision [112]. We adapt its use to measure the amount of symmetry present in the mode shapes of a structure, and use it as an indication of damage or change in a structure. It adds to the toolbox of features for detecting damage using mode shapes, specifically in the common case where the structure exhibits some form of symmetry. We explain how to calculate the CSM and demonstrate it on three model structures: a finite element pipe model, the NASA 8-bay truss, and the IASC-ASCE SHM benchmark structure, and also a laboratory experiment on a corroded steel pipe. Then we discuss how the feature can be applied in varying types of structures and its limitations.

2.2 Continuous Symmetry Measure

The CSM is a way to quantify how symmetric a given shape is. The method was initially proposed for rotational symmetry of polygons of N-sides [111], but it can easily be extended for any kind of rotational, translational, or mirror symmetry, and applied to mode shapes of a structure.

2.2.1 Derivation

Given an N-sided polygon that we want to calculate the CSM for cyclic group C_N with rotational symmetry, the methodology is as follows, as shown in Figure 2-1. First, the center of mass c_m of the polygon is found (Figure 2-1b), and the polygon is normalized such that the furthest vertex has a distance of 1. Then a C_N symmetrical version of the original polygon is calculated, by first rotating the vertices n by $-n\frac{2\pi}{N}$ radians, moving them into the same $\frac{2\pi}{N}$ slice (Figure 2-1c) to make a folded shape, averaging their position (Figure 2-1d), and rotating the vertices back. This creates a C_N symmetrical version of the polygon (Figure 2-1e). The mean squared distance of each vertex v from it's corresponding vertex v' in the symmetrical version is found and this value is the continuous symemtry measure as shown in Equation 2.1. It has a range of 0 to 1, with 0 representing perfect symmetry, and 1 perfect asymmetry (see the appendix of [111] for the derivation).
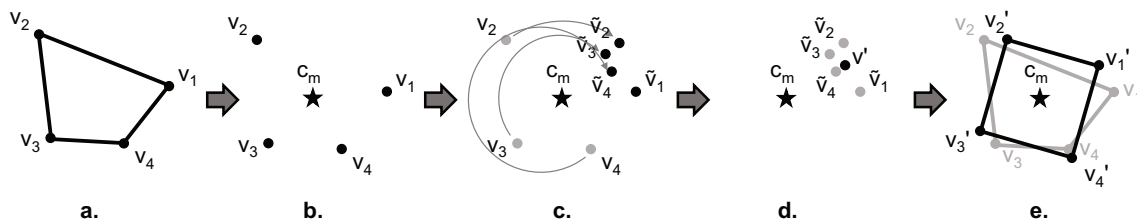
Figure 2-1: Workflow for calculating the CSM: a. The original N-sided polygon. b. Normalizing the size of the polygon and finding the center of mass c_m. c. Rotating vertices by multiples of $\frac{2\pi}{N}$ to so that they are all in the same $\frac{2\pi}{N}$ slice for the folded shape. d. Finding the average v' of the rotated vertices. e. The symmetrized polygon used to calculate the CSM.

$$CSM = \frac{1}{n} \sum_{i=1}^{n} |v_i' - v_i|^2 \tag{2.1}$$

To extend this procedure to mirror symmetry, instead of rotating the points in Figure 2-1c. we can mirror the points across a mirror line, and complete the procedure as usual. Similarly the same procedure can be done for translational symmetry where groups of points are translated, averaged to get a symmerized shape, and then the CSM is calculated. To calculate the CSM of the mode shape of a structure, we only need to use a little imagination to apply the same technique. The analogous procedure is to normalize the mode shape, group the points into the desired symmetry, find the average points, and then calculate the distances from the average points to determine the CSM. The ordering of the points from the structural mode shape should represent some rotational or mirror symmetry of the structure so that useful CSMs of different symmetries can be calculated.

2.2.2   Mode Shape Damage Detection Using the CSM

Intuitively, structures with symmetric geometry, mass, and stiffness have symmetric material properties. Small perturbations in the mass or stiffness matrix result in small perturbations in the resulting resonant frequencies and mode shapes. We can explore this by using matrix perturbation theory which describes the effects of a small perturbation in the eigenvalue problem on the resulting eigenvalues and eigenvectors. A short discussion of the theory follows; the detailed derivation is contained in Wilkinson [102].

Assume we have a simple n degree of freedom system where the mass matrix $\mathbf{M}$ is the identity matrix and the stiffness $\mathbf{K}$ matrix is real symmetric. In this case, we can set up the eigenvalue problem as:

$$\mathbf{M}^{-1}\mathbf{K}x = \mathbf{K}x = \lambda x \tag{2.2}$$

The solution to this eigenvalue problem gives the eigenvalues λ_n, which are the square

of the angular natural frequencies, and eigenvectors x_n.

If we have a small perturbation $\epsilon \mathbf{D}$ on the system matrix representing damage, where $\epsilon << 1$ and $\mathbf{D} \sim O(\mathbf{K})$, we have the perturbed system

$$(\mathbf{K} + \epsilon \mathbf{D})x' = \lambda' x' \tag{2.3}$$

and we can calculate the first-order perturbations of both the eigenvalues and the eigenvectors. For the eigenvalue λ'_i the first order perturbation is

$$\lambda'_i = \lambda_i + \epsilon D_{ii}, \tag{2.4}$$

where D_{ii} are the diagonal elements of the perturbation $\mathbf{D}$. For the eigenvector x'_i, the first order perturbation is given by

$$x'_i = x_i + \epsilon \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{D_{ji} x_j}{(\lambda_i - \lambda_j)}. \tag{2.5}$$

Note how the first order perturbation of the mode shape contaminates one mode shape with contributions from other mode shapes, and thus become only loosely orthogonal. Different mode shapes of structures typically exhibit different kinds of symmetry, like in the case of a simply supported beam, and thus the addition of damage causes the mode shape to become less symmetric for its particular symmetry.

2.3 Example: Finite Element Pipe Model

The ideal structure for demonstrating the CSM is one that exhibits many different types of perfect symmetry. A circular pipe has circumferential mode shapes which correspond to flexural standing waves that go circumferentially around the pipe and have rotational symmetry.

2.3.1   Model Setup

A size 4" schedule 40 steel pipe was modeled in a finite element program as a cross-sectional 2D model. A damaged version of the pipe was also modeled where for a 10 degree section of the pipe, the pipe only had half the thickness. This damage could be representative of pitting or material loss due to corrosion, a local manufacturing defect, or a gouge out of the pipe. This local damage should introduce asymmetry in the typically symmetrical mode shapes of the intact pipe. The mode shapes of the pipe models are determined from a linear perturbation, frequency analysis of the model.

2.3.2   Pipe Model Mode Shapes

The 2 through 6 circumferential mode shapes of the intact pipe are shown in Figure 2-2 and the mode shapes of the damaged pipe are shown in Figure 2-3. The folded shapes of the 2nd, 3rd, and 4th mode shapes for the damaged pipe are shown in Figure 2-4, which show the calculated mean symmetrized shape in red, and the individual folded shapes in blue, which corresponds to 2-1d in the CSM calculation. Not only is the region of damage in the pipe visible as the largest deviation from the mean shape, but it can be seen that the rest of the mode shape also deviates slightly from the mean shape, contributing to a higher CSM value.

(a)   (b)   (c)

(d)   (e)
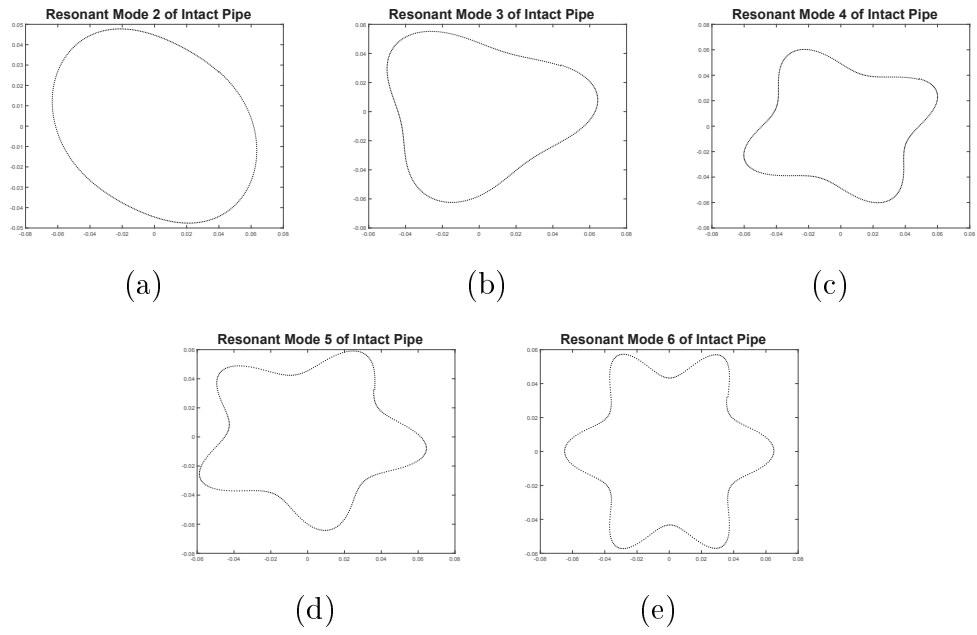
Figure 2-2: Intact pipe model circumferential mode shapes

(a)   (b)   (c)

(d)   (e)
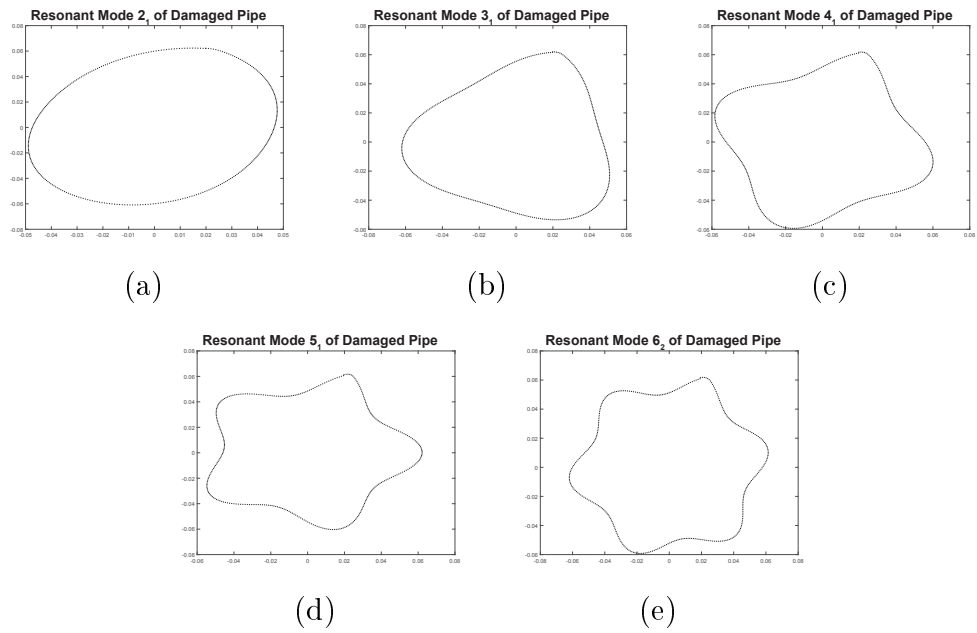
Figure 2-3: Damaged pipe model circumferential mode shapes
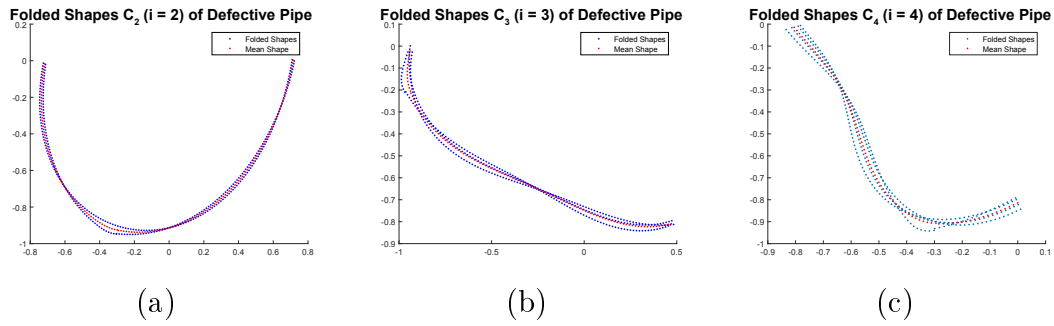
Figure 2-4: Folded mode shapes for defective pipe in CSM calculation: (a) 2nd mode, folded for C_2 (b) 3rd mode, folded for C_3 (c) 4rd mode, folded for C_4

2.3.3 Results: Pipe Model

The CSM values for the relevant rotational symmetries are calculated from the pipe mode shapes. For example, for the 2nd mode, the C_2 CSM value is calculated, and for the 6th mode, the C_2, C_3, and C_6 CSM values are calculated. A summary of the calculated CSM values, multiplied by 10^5 for readability, for the intact pipe are given in Table 2.1, and the CSM values for the damaged pipe and the percent difference from the intact pipe values are given in Table 2.2.

Table 2.1: CSM*100,000 values for the Intact Pipe

Resonant Mode	2	3	4	5	6
Frequency (Hz)	1284.9	3596.5	6799.4	10807	15538

C_N Symmetry		CSM*10^5			
2	1.4573	-	1.4894	-	1.5566
3	-	5.5443	-	-	6.5631
4	-	-	1.8686	-	-
5	-	-	-	1.978	-
6	-	-	-	-	59.741

Table 2.2: CSM*100,000 values for the Damaged Pipe

Mode	2_1	2_2	3_1	3_2	4_1	4_2	5_1	5_2	6_1	6_2	6_3
Fr. (Hz)	1148.2	1148.2	3306.9	3590.2	6375.9	6769.5	10273	10730	14482	15023	15380
% Diff.	10.6%	10.6%	8.05%	0.18%	6.23%	0.44%	4.94%	0.71%	6.80%	3.31%	1.02%

C_N				CSM*10^5 and (% Difference from Intact Pipe)							
2	13.123	1.5779	-	-	24.47	1.7988	-	-	24.379	12.411	4.3846
	(801%)	(8.3%)	-	-	(1543%)	(21%)	-	-	(1466%)	(697%)	(182%)
3	-	-	24.105	4.6593	-	-	-	-	27.333	16.606	8.4888
	-	-	(335%)	(-16%)	-	-	-	-	(316%)	(153%)	(29.3%)
4	-	-	-	-	29.298	2.302	-	-	-	-	-
	-	-	-	-	(1468%)	(23%)	-	-	-	-	-
5	-	-	-	-	-	-	33.805	3.2724	-	-	-
	-	-	-	-	-	-	(1609%)	(65%)	-	-	-
6	-	-	-	-	-	-	-	-	60.833	53.637	56.733
	-	-	-	-	-	-	-	-	(1.83%)	(-10.2%)	(-5%)

Examining the percent differences of the CSM values for the damaged pipe from the intact pipe, reveals that almost all the values show an increase and in most cases a massive increase of over 100%. In the damaged resonant modes, the previously overlapping modes in the intact pipe split into two or three mode shapes with dif-

ferent resonant frequencies which are all shown in Table 2.2. Some of the resonant frequencies of the damaged pipe shift more than others, corresponding to the influence of the defect on the mode shape. Accordingly, the CSM values are also larger, indicating a more asymmetric mode shape when there is a larger shift in the resonant frequencies.

2.4   Example: Corroded Pipe Section Experiment

To complement the finite element model study, an experiment was conducted on sections of steel pipe to validate the CSM feature on a physical model.

2.4.1   Experimental Setup

As a proxy for longer pipe systems common for transporting oil, gas, and water, we used short sections of steel pipes. An intact 6" long pipe section of 6" schedule 10 steel pipe in shown in Figure 2-5(a). Damage was induced by subjecting the pipe sections to varying amounts of impressed current corrosion, placing the pipe in a salt bath with a copper cathode, both connected to a power supply, for either 1 day, 2 days, or 5 days. As shown in Figure 2-5(c) through (e), damage was induced on approximately half the cross section of the pipe, with more significant wall thinning occurring in a smaller sections. The effect of the corrosion was asymmetrical material loss in the pipe wall, similar to that which occurs in real pipe systems.

To measure the operational mode shapes of the pipes, we arranged 12 accelerometers around the pipe as shown in Figure 2-5(b), and used a hammer to excite the circumferential vibrational modes. From the vibration data, we calculated the operational mode shapes for the 2nd and 3rd circumferential modes and from those, the CSM values for C_2 and C_3 respectively. 10 tests were conducted, on each pipe, with the location of the hammer impact moved around between tests. For the CSM values, the highest 2 values from each pipe were dropped as outliers from the average value.
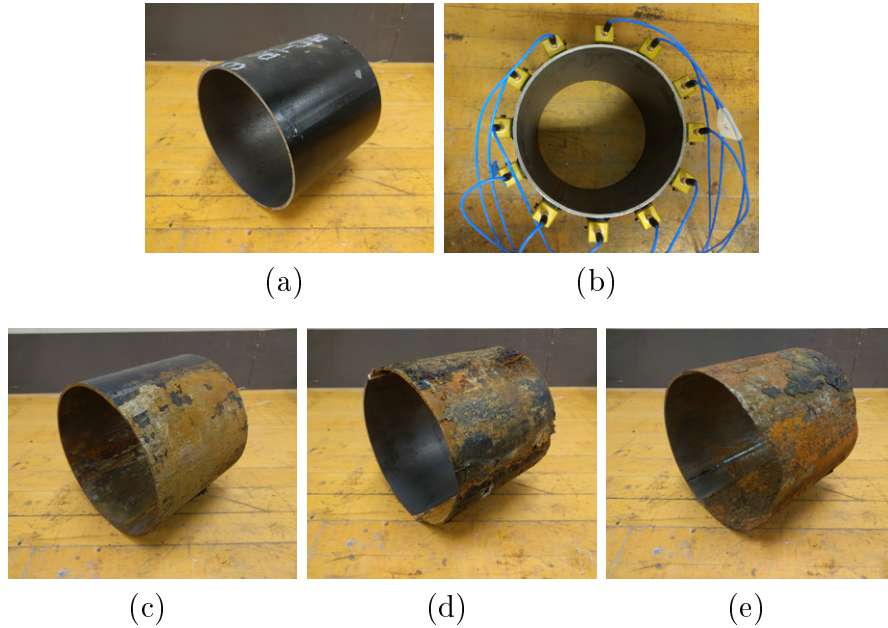
Figure 2-5: Experimental setup for the corroded pipe section study, with (a) intact pipe section, (b) accelerometer experimental setup, and corroded pipes with (c) 1 day, (d) 2 days, and (e) 5 days of impressed current corrosion damage

2.4.2   Results: Corroded Pipe Section Experiment

The average resonant frequencies for the intact and corroded are given in Table 2.3. As expected, the frequencies decrease with progressively greater amounts of damage as the pipe wall becomes thinner. As the pipe gets thinner, the mass decreases proportional to thickness, however the stiffness decreases proportional to thickness cubed, and thus the resonant frequencies decrease. Examples of the mode shapes extracted from representative tests for the intact and the pipe with 5 days of corrosion are shown in Figure 2-6. The normalized mode shapes, ranging in magnitude from -1 to 1 are plotted for visualization on a circle of radius 5. Visually the operational mode shapes for the pipe with 5 days of corrosion look less symmetric when compared to those of the intact pipe, and the CSM values for the individual mode shapes quantitatively reflect this.

The average C_2 and C_3 symmetry CSM values for the 2nd and 3rd resonant modes for the pipes are shown in Figure 2-7. For both modes, the CSM values for the

Table 2.3: Resonant frequencies for the first two operational modes shapes of the intact pipe and pipes with corrosion

Operational Mode	Intact Pipe	1 day	2 days	5 days
2	381.8 Hz	380.1 Hz	371.6 Hz	325.6 Hz
3	966.5 Hz	952.8 Hz	927.3 Hz	847.2 Hz

(a)

(b)
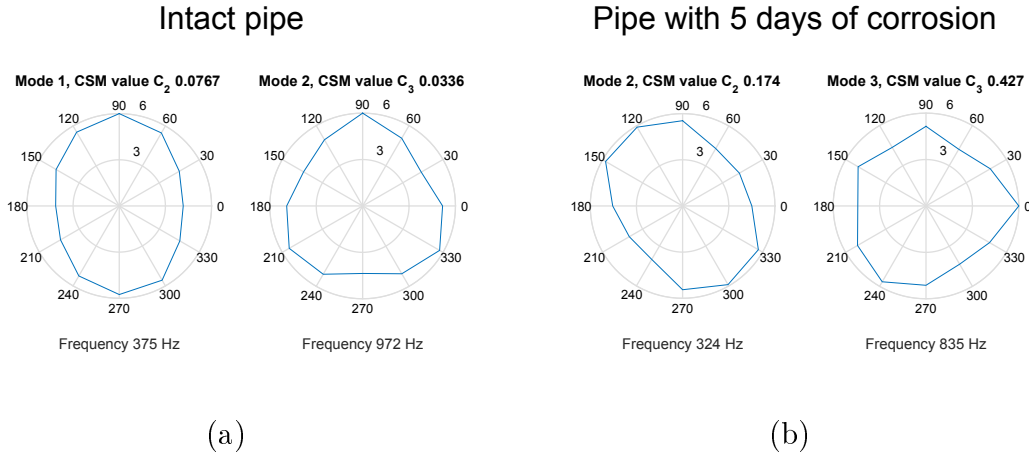
Figure 2-6: Operational mode shapes for the (a) intact pipe and (b) pipe with 5 days of corrosion

damaged pipes are greater when compared to the intact pipe, indicating increasing amounts of asymmetry in the operational mode shapes, characteristic of damage. Even though the corrosion damage is not as localized as the simulated damage in the finite element pipe model, the differences in the CSM values are still clear indicators of damage. A convenient characteristic of the CSM is that even though the mode shapes are rotated around the pipe depending on where the hammer impact occurs, the CSM is agnostic to the orientation of the mode shape itself, and values from multiple tests with different orientations can be directly averaged together.
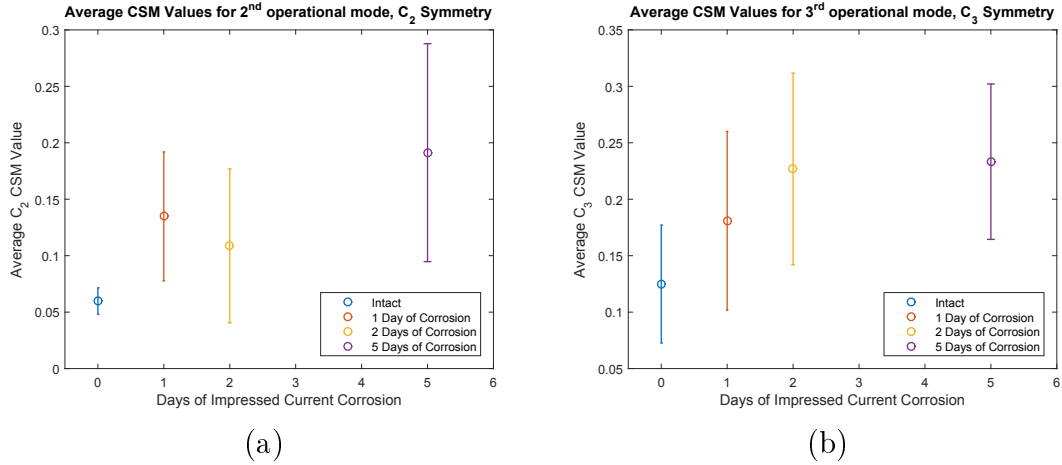
Figure 2-7: CSM values differences for pipes with no corrosion, 1 day, 2 days, and 5 days of corrosion

2.5 Example: NASA 8-bay Model

The NASA 8-bay model is a finite element model of a truss used for testing of damage detection algorithms using modal analysis [52]. The model is a statically indeterminate cantilevered truss. It is a model of a real eight bay cantilevered truss from a scale model of an early Space Station Freedom configuration. Several previous studies have used this structure, and it is thus a known benchmark in the literature for damage detection algorithms and structural health monitoring studies [56, 114, 107]

2.5.1 NASA 8-bay Model Dataset

The data for the model includes the mode shapes for five different resonant modes, shown in Figure 2-8 [107], for the intact model, and eight different damage scenarios. In the damage scenarios, different truss members are removed, i.e. a 100% loss of stiffness, and they are members that are either in the longitudinal, horizontal, or diagonal direction.

The basic assumption for the truss structure is that the motions of the four columns in the mode shape, should be approximately the same and therefore symmetric, when the structure is healthy. Damage might cause the four columns to move
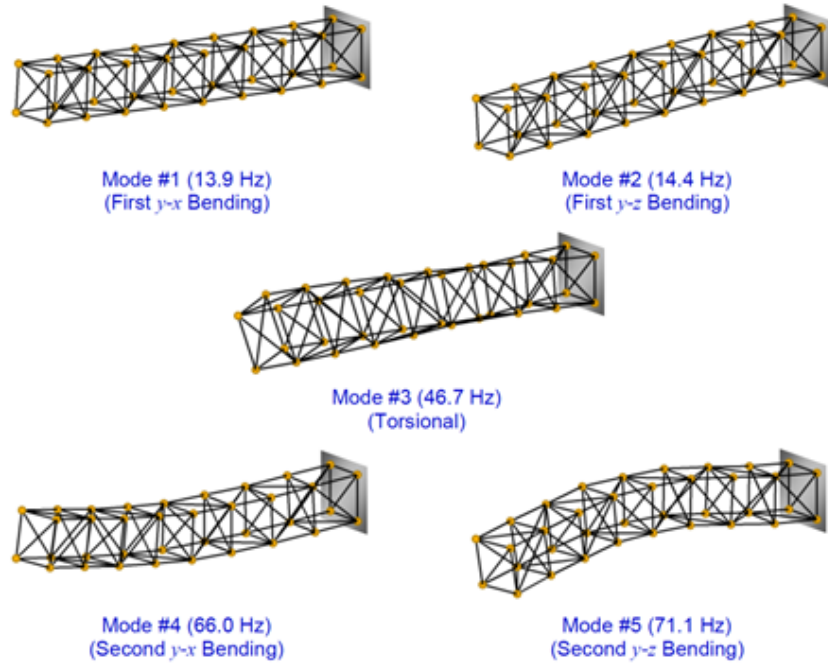
Figure 2-8: NASA 8-bay truss model resonant mode [107]

differently, and have more asymmetric mode shapes which would be reflected in the CSM values. In each damage scenario and for the intact case, the CSM values are calculated for the five resonant modes and averaged to give a single value. The values are then normalized with the value from the intact structure.

2.5.2 Results: NASA 8-bay Model

The results are shown in Figure 2-9, which demonstrate that in cases where diagonal elements are removed, the CSM values increase significantly. In damage cases where either longitudinal or horizontal elements are removed, there is much less change and cannot be distinguished with the CSM values, from the intact structure. The general conclusion from this test is that depending on the structure geometry, it is possible that only some types of damage may influence the CSM values, while others may not.

CSM Values for Different Damage Cases
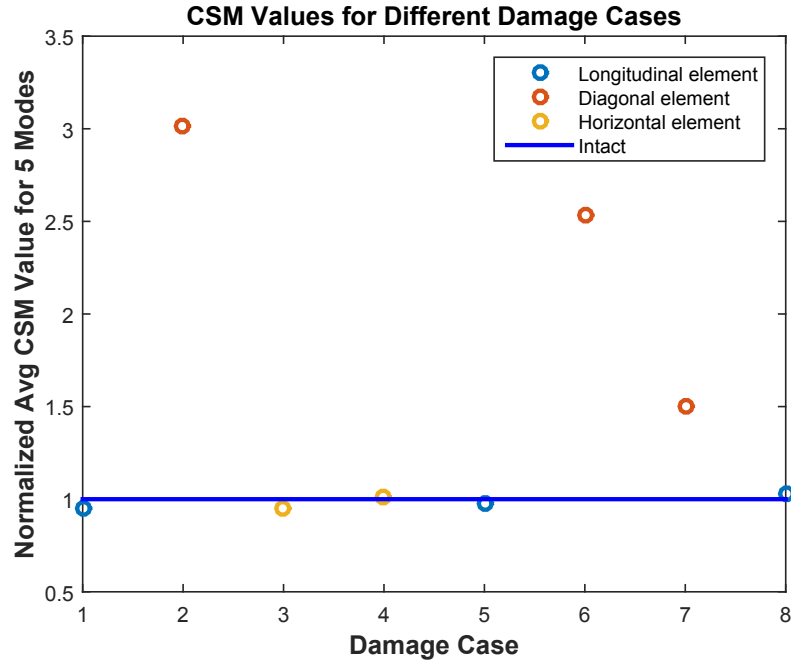
Figure 2-9: NASA 8-bay CSM damage detection results, showing that when diagonal elements are removed the CSM values increase significantly

2.6 Example: ASCE Benchmark Structure

In the case of structures that do not exhibit perfect symmetries, we would like to see if the CSM feature is still applicable. The IASC-ASCE SHM benchmark problem involves a 2 by 2 bay, 4 story, fully braced, moment resisting, steel frame structure, notable for being widely used for demonstrating new SHM algorithms. Conveniently, it has publically available software for generating simulated time series response data for various damage scenarios and excitation cases [50].

2.6.1 Simulation Experimental Setup

The following experimental specifications for the ASCE benchmark structure are use to generate the simulated time response data. Case 5 is used for the time history response simulation, which is a 120DOF model, asymmetric mass distribution, and uses a shaker diagonally oriented on the roof as an excitation. We used the time response data from 16 "sensors" placed on the central columns at each story of the

57

structure, sampled at 1000 Hz for 100 seconds, nominal values as specified by [50]. The damping ratio used was 0.01, noise level was 0, and the force intensity was 150. 8 sensors measured in the X-axis direction on two opposing central columns, and the other 8 sensors measured in the Y-axis direction on the other two central columns. The damage patterns used include the first 6 standard damage patterns, and an additional 8 were created, for a total of 14 damage patterns which are listed in Table 2.4. Damage scenarios 3, 4, 5, 6, 9, 10, 11, and 12 have asymmetrical damage, whereas the other damage cases introduce symmetrical damage. Figure 2-10 shows a diagram of the analytical benchmark structure and additional details about the structural model including the damage pattern resonant frequencies, and damage pattern stiffness losses are listed in the Johnson et al. paper [50].
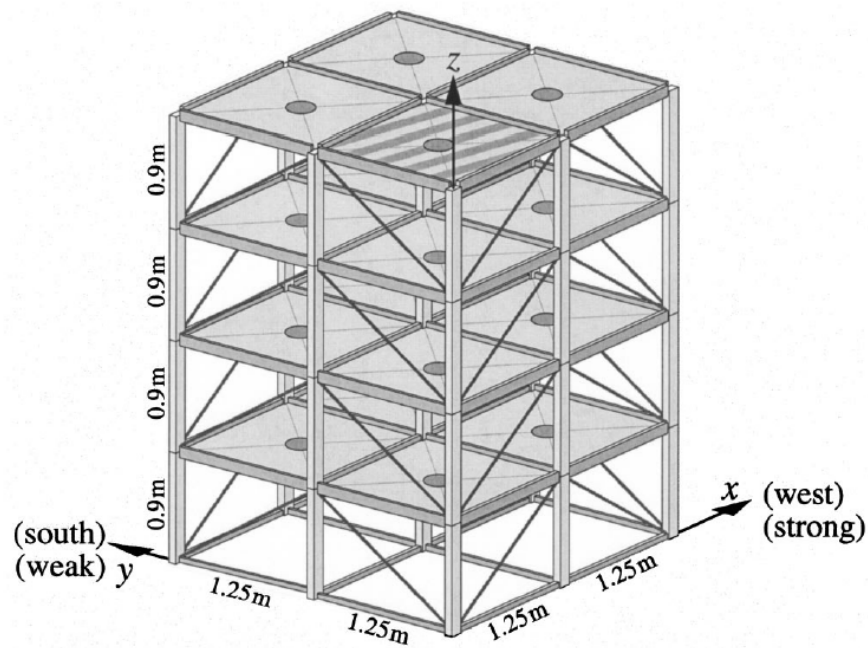
Figure 2-10: Diagram of analytical IASC-ASCE SHM benchmark structure [50]

Table 2.4: Scenario numbers and damage patterns for the ASCE benchmark structure

Scenario #	Damage Pattern (removal of)
0	Intact Structure
1	All first story braces
2	All first and third story braces
3	One first story brace
4	One first and one third story brace
5	Same as 4, also loosening a first story floor beam
6	2/3rds stiffness of one first story brace
7	All second story braces
8	All third story braces
9	One first and one second story brace
10	One first, second, and third story brace
11	One second story brace
12	One third story brace
13	Symmetrically removing two first story braces
14	Symmetrically removing two second story braces

2.6.2 Operational Deflection Shape Extraction and CSM Calculation

Operational deflection shapes (ODS) were extracted from the simulated time series data using frequency domain decomposition (FDD). The first 8 bending modes of the structure, 4 in the X-axis direction and 4 in the Y-axis direction were analyzed to calculate CSM values for the structure. The geometry of the healthy structure is rotationally and mirror symmetric, however the structure has an asymmetrical mass distribution on the roof. The 16 sensors are arranged on the central columns and oriented such that for the bending modes of the structure, for the X-axis modes, two sets of 4 sensors provide information on the ODS on opposite columns, and similarly for the Y-axis modes on the other pair of columns. Given the geometry of the structure, the pairs of columns should move somewhat symmetrically, so we calculate the CSM for the symmetry and bending mode shapes between the pairs of columns for the X-axis sensors from the X-axis bending modes, and similarly using the Y-axis sensors for the Y-axis modes. The feature for detecting damage or change in the structure is thus the mean of the CSM values for the first 8 bending modes.

2.6.3 Results: ASCE benchmark structure

CSM vs. Damage Scenario
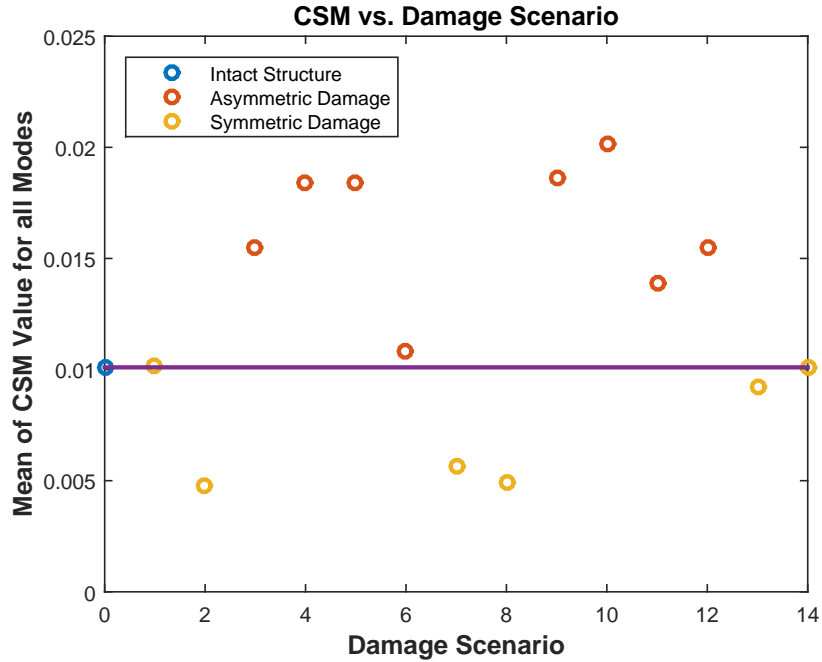
Figure 2-11: Results for CSM Metric for ASCE Benchmark Structure

The mean CSM values for the different damage scenarios are shown in Figure 2-11 where the intact structure has a mean CSM value of 0.0101. Shown in red are the damage scenarios with asymmetric damage and in yellow are the scenarios with symmetric damage. The full tables of extracted ODS frequencies and CSM values are given in Appendix C.1. For the symmetric damage scenarios, the values either stay similar to the intact structure, or actually decrease by a notable amount in the case of scenarios 2, 7, and 8. This decrease is possible because the intact structure is inherently asymmetric due to the mass distribution. For the scenarios with asymmetric damage, the mean CSM values are increased compared to the intact structure, with a greater deviation for more severe damage patterns. The greatest increase is with scenario 10 where single braces are removed from the first, second, and third stories, and the smallest increase is with scenario 6 where just one first story brace has 2/3rds stiffness. More generally, any changes to the structure cause changes in the CSM values of the operational deflection shapes.

60

2.7 Discussion and Conclusions

We have demonstrated the CSM as an indicator of change in a structure by measuring the amount of symmetry present in the mode shapes of a pipe cross section, the NASA 8-bay truss, and the IASC-ASCE SHM benchmark structure. In the case of the pipe cross section for the damage introduced over a 10 degree section of the pipe, the CSM values for the circumferential modes generally increased by a large factor when comparing the damaged and the intact model. This change was also validated experimentally with intact and corroded steel pipes. In the NASA 8-bay truss, the average CSM value for the first 5 modes only increased when diagonal elements were removed, and not for longitudinal or horizontal elements. For the ASCE benchmark structure, mean CSM values for the first 8 bending modes of the structure increased only in the case of asymmetric damage, but sometimes decreased in the case of symmetric damage. In this section we discuss the applicability and the limitations of the CSM feature.

2.7.1 Applicability

Simple structures such as circular pipes with rotational symmetry or simply supported beams with translational symmetry, are the easiest to apply the CSM to, because any asymmetric change in the geometry material properties would cause an asymmetric change in the mode shape, and thus an increase in the CSM value. Structures that are geometrically similar to these simple shapes are good candidates in this respect. For the pipe cross-section this would include cooling towers commonly associated with nuclear power plants, cylindrical storage tanks, and similarly cylindrical objects. Similar to a simply supported beam would be repeating continuous structures such as long highway overpasses or above ground pipelines.

For more complex structures such as buildings or bridges which do not exhibit perfect symmetry, similar to the ASCE benchmark structure, the CSM can still be used, however the CSM is likely to increase or decrease in the case of different types of damage. This makes it less useful as a standalone feature for damage detection,

however it can still be useful in the larger context of a data-based SHM algorithm as the feature is sensitive to changes in mode shapes. Additionally, determining the correct symmetries to use for calculating the CSM values of the mode shapes will take some user thought, as it may not be obvious for all structures. For tall buildings that are somewhat box shaped, an initial suggestion is to measure whether bending mode shapes are symmetrical along the axis orthogonal to the bending direction. For bridges what might be appropriate is measuring the mirror symmetry of the mode shapes across the deck of a bridge to see if both sides experience similar motion.

2.7.2   Limitations

There are limitations with the CSM, especially as structures are more complex and deviate from ideal symmetry. For complex structures, determining reasonable symmetries for calculating CSM values may require some extra analysis of the structure. The ODS measured from the structure also need to be sufficiently spatially dense to be able to make a meaningful CSM calculation. When structures are initially asymmetric to a certain degree, the CSM feature is also less discriminative of damage, as the CSM is only meant to detect asymmetrical damage. This is illustrated by the ASCE benchmark structure example, where in some cases symmetrical damage caused the mean CSM value to decrease. Most damage is expected to be asymmetrical, especially in more complex structures where it is unlikely that damage caused by most typical sources, would simultaneously occur symmetrically in multiple locations in the structure. In general, the CSM seems sensitive to any changes in the structure that impact the mode shapes, and contributes as a useful feature for data-based SHM approaches.

2.8   Chapter Summary

In this chapter we have proposed using the continuous symmetry measure for measuring the symmetry of structural mode shapes as a feature for damage detection in structural health monitoring. We demonstrated the feature on a model with ideal

rotational symmetry, a pipe cross-section, the NASA 8-bay truss, and a model more representative of civil structures, the IASC-ASCE SHM benchmark structure. Additionally we conducted a laboratory experiment with intact and corroded sections of pipe. This experiment and the CSM metric is revisited later in the thesis and applied to camera measurements of a corroded pipe experiment in Section 5.3.4. In the cases of the pipe and the benchmark structure, when asymmetric damage was introduced, the CSM values of the mode shapes increased, indicating increased asymmetry of the mode shapes. In the case of the NASA 8-bay truss, the CSM value only increased when diagonal elements were removed, but not with horizontal or longitudinal elements. The CSM feature is applicable to any structure that exhibits some symmetry known to the end user, albeit easier to apply to simple symmetrical structures like pipes, and more difficult for complex structures like buildings. The feature may be limited in being able to detect all forms of damage, but it is easy to calculate and provides information on an intrinsic physical property of the structure, and thus would work best as one feature of many in a data-based or statistical SHM approach.

Chapter 3

Vibration Measurement and Visualization from Video

This chapter describes how video is processed to measure the motions and vibrations of objects in video and how the extracted information can be used for both quantitative measurement and qualitative visualizations. Quantitatively, displacements can be measured with much higher spatial resolution than with traditional contact sensors, and the signals can be processed with modal analysis techniques to obtain operational resonant frequencies and mode shapes. Qualitatively, visualizations of an object's motion can be created by amplifying the scale of the normally imperceptible motion or by using the measured operational mode shapes to create a plausible interactive simulation.

3.1 Processing Workflow

A general overview of the processing workflow is given in Figure 3-1, and the following sections cover each step of the processing workflow in more detail. The input is a source video of a vibrating structure or object, and the outputs are displacements extracted from the video and a video where small motions are magnified as a visualization. The source video is processed using complex linear filters into an amplitude and phase representation for each frame of the video. The phase signals in time are

representative of motion and can at pixels with good edges or local contrast can be processed into displacement measurements. The phase signals can also be band pass filtered in time, amplified, and recombined with the amplitude signals to generate a motion magnified video. In this video small motions are amplified to be perceptible in a specified frequency band. This frequency band can either be specified a priori, or from a frequency analysis of the displacements determined from the phase signals. Additionally the displacement information can be processed into resonant frequencies and mode shapes. This information can then be used for the purposes of SHM and NDT, and to generate a plausible simulation of the dynamics of the structure or object.
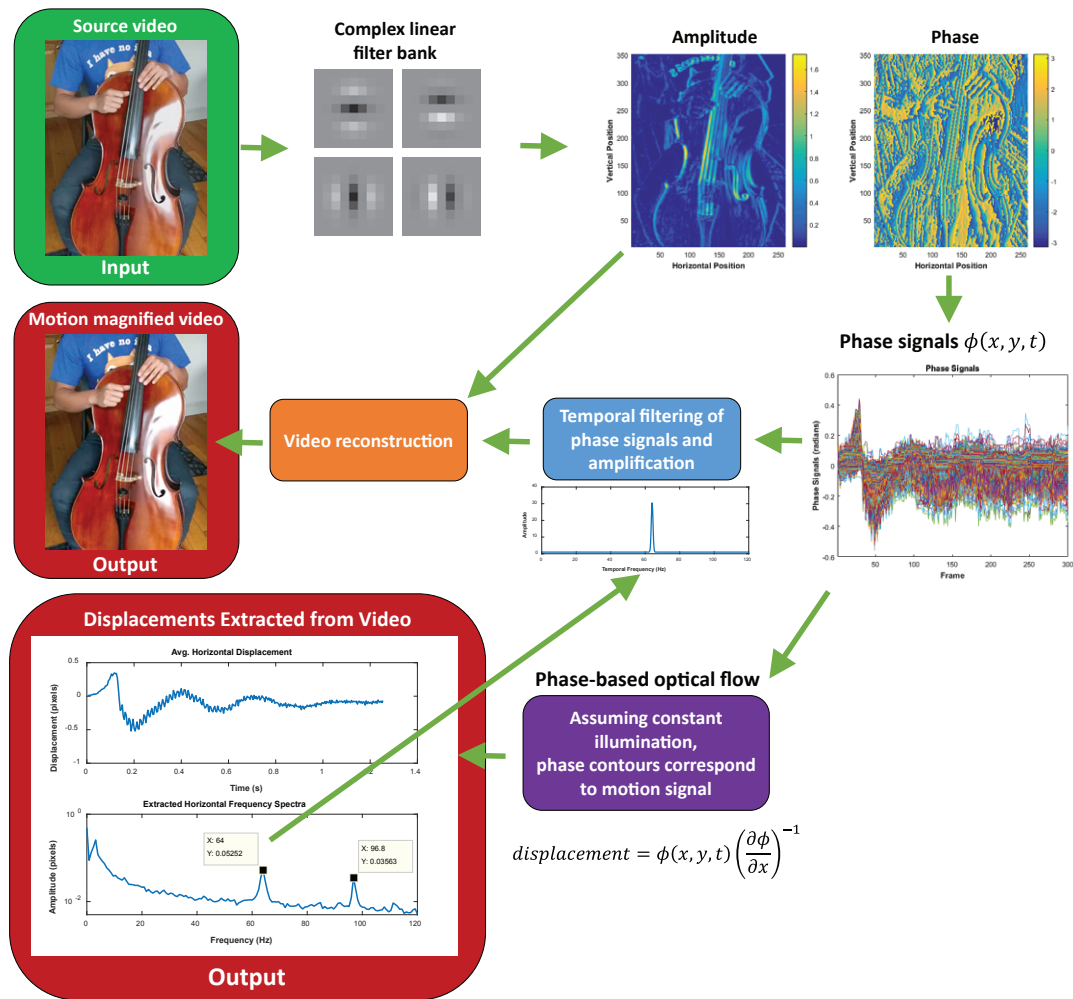
Figure 3-1: A general overview of the processing workflow, from the input of a source video, obtaining motion signals from local phase, and outputs of displacements extracted from the video and a motion magnified video

3.2 Obtaining Motion Signals from Video

In determining the motion signals from a video, there are two general methods which are best described by an analogy to descriptions of fluid flow. Lagrangian descriptions track the fluid as it moves through space and time, which is similar to particle tracking or digital image correlation methods of measuring displacements from video. Eulerian descriptions look at the motion of the fluid as it passes in time, while being at a fixed location, which is the basis for the type of video processing used here to obtain the motion signals from video.

Almost all image processing is done with the same techniques used for time series or typical 1D signal processing, however it is expanded to 2D, as images are effectively 2D signals in the spatial domain. With a series of images in time, or a video, we end up with a 3D signal, however we typically treat the time dimension differently, and separately process the spatial and temporal dimensions. Additionally, images may be not only in greyscale but also in color, however in our processing we typically just look at the brightness information and convert the red, green, and blue channels to a brightness value for each pixel. We can also apply linear filters to 2D images, just as they are applied in the time domain to 1D signals.

3.2.1 Local Phase and Amplitude

For signals, other than the typical time domain (or spatial domain) representation, there is a frequency domain representation that is commonly used to decompose signals i.e. using a Fourier transform. For a specific Fourier series coefficient, there is the amplitude and phase, where the amplitude controls the strength of a 1D sinusoid, and the phase controls its location. A Gabor filter, which is defined by a sinusoidal wave multiplied by a Gaussian function, introduces locality to the Fourier series, in that Fourier coefficients can be calculated for a signal in a local area around a certain sample. Thus we can decompose a video into a local amplitude and local phase representation. Local phase signals have previously been demonstrated to correspond to motion signals in a video, and this is how we will obtain motion signals from video

68

[41, 43]. We use the phase rather than the amplitude, since the phase signal is more robust against intensity noise inherent in camera sensors.

For a video with image brightness specified by $I(x, y, t)$ at spatial location in image-space (x, y) and time t, the local phase and local amplitude in orientation θ at a frame at time t_0 is computed by applying a quadrature filter pair G and H to a frame to get

$$A^\theta(x, y, t_0)e^{i\phi^\theta(x,y,t_0)} = (G^\theta + iH^\theta) \otimes I(x, y, t_0) \tag{3.1}$$

where $A^\theta(x, y, t_0)$ is the local amplitude and $\phi^\theta(x, y, t_0)$ is the local phase. These filter pairs G and H can be any number of filters, such as G_2 and H_2, a Gaussian derivative filter pair described in [42], or the complex steerable filters, approximations of Gabor wavelets, described in [84]. A 9-tap version of the G_2 and H_2 filter pair is shown in Figure 3-2.

(a)   (b)
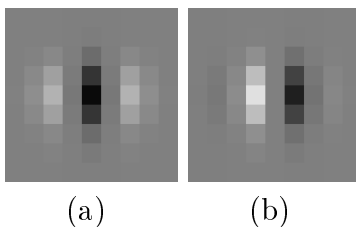
Figure 3-2: G_2 and H_2 quadrature filter pair, 9-tap discrete filters where the grey level represents the value [42]

3.2.2 Motion Signal from Local Phase

It has been demonstrated that constant contours of the local phase through time correspond to the displacement signal assuming no changes in illumination [41, 43]. This can be expressed as

$$\phi^\theta(x, y, t) = c \tag{3.2}$$

for some constant c. Differentiating with respect to time, yields

$$\left(\frac{\partial \phi_\theta(x, y, t)}{\partial x}, \frac{\partial \phi_\theta(x, y, t)}{\partial y}, \frac{\partial \phi_\theta(x, y, t)}{\partial t} \right) \cdot (u, v, 1) = 0 \tag{3.3}$$

where u and v are the velocity in the x and y directions respectively. Because an oriented filter can only measure phase changes in one direction, the temporal phase gradient $\frac{\partial \phi^\theta(x,y,t)}{\partial t}$ only gives information on the component of velocity with the same direction as the phase spatial gradient. Thus, the velocity in units of pixel is

$$u = -\left(\frac{\partial \phi^0(x,y,t)}{\partial x}\right)^{-1} \frac{\partial \phi^0(x,y,t)}{\partial t} \text{ and } v = -\left(\frac{\partial \phi^{\pi/2}(x,y,t)}{\partial y}\right)^{-1} \frac{\partial \phi^{\pi/2}(x,y,t)}{\partial t}$$
$$(3.4)$$

where $\phi^0(x,y,t)$ and $\phi^{\pi/2}(x,y,t)$ are derived by applying horizontal or vertical direction filters, respectively. To get horizontal and vertical displacements d_x and d_y at time t_0 we integrate both sides to get

$$d_x(t_0) = -\left(\frac{\partial \phi^0(x,y,t_0)}{\partial x}\right)^{-1} \left(\phi^0(x,y,t_0) - \phi^0(x,y,0)\right) \text{ and}$$
$$d_y(t_0) = -\left(\frac{\partial \phi^{\pi/2}(x,y,t_0)}{\partial y}\right)^{-1} \left(\phi^{\pi/2}(x,y,t_0) - \phi^{\pi/2}(x,y,0)\right).$$
$$(3.5)$$

which gives the local displacement of the image at that pixel relative to the first frame.

3.2.3 Assumptions

For the local phase signals to be equivalent to the motion signals, a couple of assumptions are necessary. The most important one is constant illumination, otherwise the notion that constant contours of local phase through time correspond to the displacement signal does not hold. A periodic change in the illumination can manifest as an apparent motion signal that appears in the extracted displacements. This happens most commonly when fluorescent lights or lights powered by alternating current illuminate the scene; a common signal is 120 Hz which is the flicker frequency fluorescent lights powered by 60 Hz mains alternating current. Also, lighting changes may have the effect of aliasing the motions of the object to the flicker of the lighting. This can be advantageously used to stroboscopically sample objects vibration at a much higher frequency than the camera's frame rate, as long as the exposure time of the camera

is small enough to properly image the object. When making measurements outdoors, lighting changes due to clouds can also occur, but these are usually on much longer and slower time scales, and occur randomly, so the impact on the measurement is minimal.

Displacements are only valid in regions with sufficient texture, and they are best when the displacement is in the direction that is normal to an edge in an image. Determining the displacement of textureless regions is an open problem in computer vision, as without texture there is insufficient visual contrast to detect motion. In our processing, displacements are only extracted from areas with sufficient texture.

Another important assumption is that the motions in the source video are small. Typically the displacements need to be on the order of one pixel or smaller, but this depends on the size of the filter kernel and whether or not the video is downsampled spatially before being filtered. Since the motion is extracted from the phase signal, if the phase signal goes beyond a change of pi in phase, the method breaks down. The camera itself, is also not allowed to move a significant amount as large apparent motion due to translation or rotation motions of the camera, will also break the method. Small motions of the camera can be accounted for and subracted from the extracted displacement field either by assuming that global motion of the frame is camera motion, or by measuring the motion separately with other sensors. The subject of camera motion will be further discussed in Chapter 4. Pixels where the motion may be too large end up with zero crossings in the $\frac{\partial \phi^0(x,y,t)}{\partial x}$ signal, which being in the denominator causes the displacement calculation in Equation 3.5 to blow up. Thus, pixels with such zero crossings are removed from the extracted displacements.

3.2.4   Additional Pre- and Post-processing

The size of the video can be downsampled by any power of 2 to change the scale of the motion in the video, which also changes the scale of features in the video that the filter acts on. Additionally it has the benefit of speeding up the computation. Since displacements are only valid in regions with sufficient texture, a mask is generated based on the amplitude signal in the first frame after applying the quadrature filter

pair. A threshold is chosen, nominally half of the median of the 30 pixels with the greatest amplitudes, above which pixels are included in the final processing. This threshold can be anything sufficient to remove pixels without sufficient texture and it is shown in Chapter 4 that the noise floor for a pixel is inversely correlated to the amplitude. After the displacements are extracted, pixels with zero crossings in the $\frac{\partial \phi^0(x,y,t)}{\partial x}$ signal are removed because they cause the displacement calculation to blow up when dividing by near-zero.

The result of the procedure is a set of pixels with motion signals corresponding to the displacement of the objects in the video. These displacements can be processed in any number of methods to determine vibrational information of the object measured.

3.3 Processing Motion Signals into Vibration Information

Once a collection of motion signals is obtained, they need to be processed with system identification algorithms to determine the mode shapes or operational deflection shapes of the system. Nominally the operational deflection shapes are determined using an output only modal identification procedure because the input forces are either unidentifiable, or are unknown because they require other experimental equipment to measure. They serve as a good estimate of the mode shapes when the input forcing is broadband and the resonant frequencies are well spaced.

3.3.1 Initial Resonant Frequency Identification

To get an initial idea of the frequencies present in the video, an average frequency spectra is obtained for the video. Either the displacement signals can be averaged first, and then a fast Fourier transform (FFT) is used to get a frequency spectrum, or each of the individual signals can be FFTed and then averaged in the frequency domain to get a single frequency spectrum. The average can be weighted by the amplitude signal of the pixel to put more weight into pixels with good texture. For

particularly long videos, a sliding window can be used to average together many windows of different FFTs.

As not all pixels may belong to objects in the video that are vibrating, especially for pixels with strong edges in the background that may be preferentially weighted, we'd like some way to distinguish pixels with frequency distributions that contain potential vibration peaks, from those that are just the noise floor. We can calculate the skewness of the distribution, of the magnitudes of the frequency spectrum per pixel to only consider pixels with positive skewness above a certain threshold. Positive skewness is characterized by a distribution with a longer or fatter tail on the right side rather than the left, indicative of the possible presence of frequency peaks in the spectrum with the higher amplitudes of frequency peaks forming that larger tail on the right side. This improves the signal to noise ratio by preferentially including only pixels that have the potential to have frequency peaks.

To automatically identify frequencies of interest, a peak finding function can be used [108], which is usually good for obvious resonant frequency peaks in videos with strong vibrational behavior. Frequency peaks can also be chosen by hand, which works better for resonant frequencies with lower magnitudes. In more natural scenes without obviously vibrating objects, the frequency peaks may be wider and shorter and require user intervention to identify. External sources of noise like lighting flicker or electronic noise may cause erroneous frequency peaks that can usually be identified by their unusual sharpness or frequency, uncharacteristic of mechanical vibration behavior.

3.3.2 Modal Identification

To process raw displacement signals into operational resonant frequencies and mode shapes, modal identification techniques are necessary. Two frequency domain techniques are described here, peak picking and frequency domain decomposition.

Peak Picking

Peak picking is one of the most straightforward methods to identify mode shapes or operational deflection shapes from displacement signals in the frequency domain. It involves taking the FFT of the displacement time signals and picking out peaks in the frequency domain that correspond to candidate resonant modes. Cross power spectral densities (PSD) are taken with reference to a single signal to determine the phase difference between the displacement signals at the frequencies of interest. This information combined with the normalized magnitudes of the FFT at the picked frequencies create the operational deflection shape.

Given a displacement signal of $d_{i,j}(t)$ for pixel at location i, j in the video, with FFT $\widehat{d}_{i,j}(f)$ the cross PSD (ignoring scaling factors) referenced to some pixel r is

$$cPSD_{i,j|r}(f) = \overline{\widehat{d_{i,j}(f)}} \circ \widehat{d}_r(f) \tag{3.6}$$

and the phase difference between the pixel at i, j and the reference pixel is the phase angle of the complex value $cPSD_{i,j|r}(f)$. The mode shape $\varphi_f(i, j)$ (not to be confused with the phase) at a specific frequency f_0 is then

$$\varphi_{f_0}(i, j) = \begin{cases} |\widehat{d}_{i,j}(f_0)| & \text{if } |\angle cPSD_{i,j|r}(f_0)| < \frac{\pi}{2} \\ -|\widehat{d}_{i,j}(f_0)| & \text{if } |\angle cPSD_{i,j|r}(f_0)| > \frac{\pi}{2} \end{cases} \tag{3.7}$$

and then it is normalized by the maximum value of $\varphi_{f_0}(i, j)$ such that it ranges from -1 to 1.

When the input forcing frequency spectrum is available, the output displacement frequency spectra are divided by the input forcing frequency spectrum to obtain frequency response functions, that are processed in the same way to obtain mode shapes. The downside to this method is that it does not work well with closely spaced modes and it can be difficult to identify non-physical harmonics in the response [9]. In most cases with coherently vibrating objects, peak picking is more than sufficient to identify the resonant frequencies and operational deflection shapes.

Frequency Domain Decomposition

Frequency domain decomposition (FDD) is a method used for modal identification of output-only systems which addresses some of the drawbacks of the peak picking method [9]. It involves calculating the spectral matrix, a matrix containing the cross PSD between all the output signals. The singular value decomposition (SVD) of the spectral matrix at each frequency bin is taken to decompose the spectral matrix into a set of auto spectral density functions as singular values, where strong resonant peaks in the singular values correspond to possible resonant modes and the singular vectors are the corresponding operational deflection shape. The benefit is that this algorithm performs better than peak picking when signals are noisy, there is better discrimination for closely spaced resonant modes, and the possible capability to identify non-physical harmonics [9, 7]. The downside is that because the FDD depends on the SVD it is much more computationally intensive than peak picking, especially as the number of sensors or pixels increases. FDD can be done separately on each frequency bin, so once frequencies of interest are chosen from the average frequency spectrum for the video, the FDD procedure need only be applied around frequencies of interest in that spectrum.

3.4 Phase-Based Motion Magnification

Motion Magnification is an algorithm for the amplification of small motions in videos. The basic principle behind the algorithm is to obtain a representation for the video in such that signals representing the motions of objects in the video can be time-frequency bandpass filtered, amplified, and reconstructed back into a video where the apparent motion of objects in the video is larger in a certain frequency band. The phase-based motion magnification algorithm [99] decomposes the signal of a video into the local spatial amplitude and phase using a complex-valued steerable pyramid filter bank [84]. The local spatial phase signals are temporally Fourier decomposed into a series of sinusoids representing harmonic motion. The phase signals are then temporally bandpass filtered, amplified, and recombined back to form a motion mag-

nified video. The result is that the video then has magnification of the motion in a specified band of temporal frequencies. Often, for speed and faster processing of videos, we use a similar processing procedure described in [100].

A diagram of this processing workflow is shown in Figure 3-3 and a detailed explanation and derivation of the procedure is given in the reference [99]. Additionally, for the reader's convenience, a description of the phase-based motion magnification algorithm is included in Appendix B, taken from [97] and reproduced with permission. This includes an explanation of the theory behind phase-based motion magnification, and how videos are processed to obtain motion magnified visualizations of normally imperceptible motions.
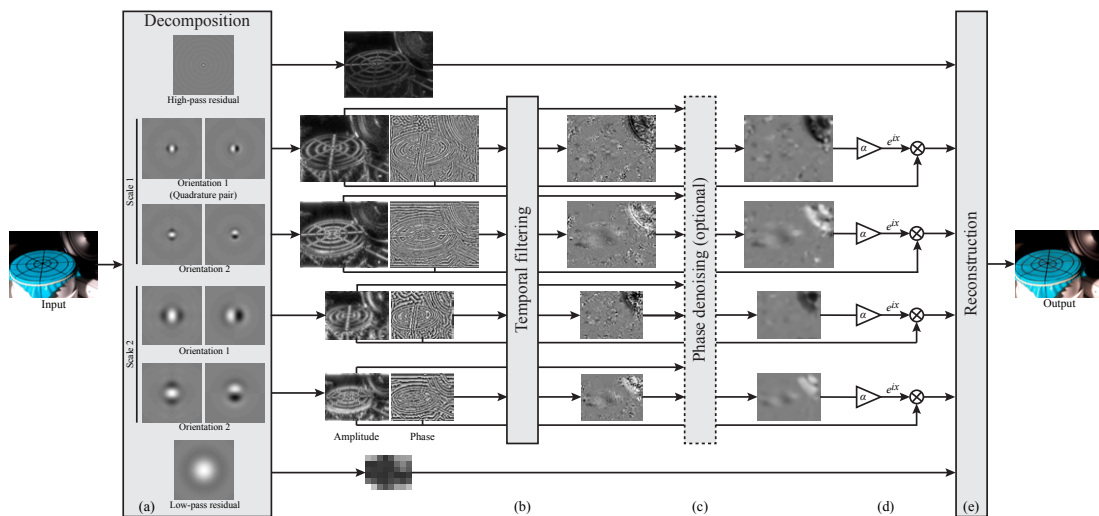
Figure 3-3: Motion Magnification processing workflow: (a) complex steerable pyramid filters decompose the video into amplitude and phase at different scales, (b) the decomposed phases are bandpass filtered in frequency, (c) amplitude-weighted smoothing is applied, (d) the bandpassed phases are amplified or attenuated, and (e) the video is reconstructed [99]

For a vibrating structure, if the band of frequencies is appropriately chosen, the result will be a video of the structure's operational deflection shape at one of its approximate resonant frequencies, assuming the modes are sufficiently spaced. In the case of closely coupled modes, the mode with a larger amplitude will likely visually dominate. An in-depth discussion of the differences and interpretation of operational deflection shapes as opposed to classical mode shapes is given in the reference [49].

This qualitative visualization of a structure's operational vibration modes can be invaluable in the case of more geometrically complex structures.

3.5 Plausible Simulation of Objects in Video

We can use the extracted mode shapes to form a modal basis for simulation of the motion of objects in video. Our approach is based on the same linear modal analysis behind many techniques in physically-based animation. However, unlike most of these techniques, we do not assume any knowledge of object geometry or material properties, and therefore cannot rely on finite element model (FEM) methods to derive a modal basis for simulation. Instead, we observe non-orthogonal projections of an object's vibration modes directly in video. For this we derive a relationship between projected modes and the temporal spectra of optical flow. We then show that, while non-orthogonal, these projections can still be used as a basis to simulate image-space object dynamics. This Section 3.5 of the chapter was previously published and presented in [30].

3.5.1 Modal Analysis

In this section we connect established modal analysis theory to the image-space deformations of an object. Section 3.5.1 reviews some of the relevant theory from linear modal analysis (more detail can be found in [81, 4]). Section 3.5.1 connects this theory to the observed deformations of an object in video and provides a theoretical basis for the algorithms described in Section 3.5.2.

Object Motion

The dynamics of most solid objects under small deformations are well approximated by a finite element model representing a system of masses, dampers, and springs. We assume that objects undergo small deformations in the linear range that do not cause permanent translations or shape alterations. The matrices $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ represent

77

mass, damping, and stiffness relationships between an object's degrees of freedom, and the equation of motion in response to a force $\mathbf{f}(t)$ is given by

$$\mathbf{M\ddot{u}}(t) + \mathbf{C\dot{u}}(t) + \mathbf{Ku}(t) = \mathbf{f}(t), \tag{3.8}$$

where $\mathbf{\ddot{u}}(t)$, $\mathbf{\dot{u}}(t)$, and $\mathbf{u}(t)$ are vectors for acceleration, velocity, and displacement. Assuming sinusoidal solutions to Equation 3.8, the eigenmodes of this system are the orthogonal solutions to the generalized eigenvalue problem given by $\mathbf{K}\phi_i = \omega_i^2 \mathbf{M}\phi_i$. The set of eigenvectors or eigenmodes $\phi_1...\phi_N$ define a modal matrix $\mathbf{\Phi}$ shown in Equation 3.9 which diagonalizes the mass and stiffness matrices into modal masses $\mathbf{m}_i$ and modal stiffnesses $\mathbf{k}_i$.

$$\mathbf{\Phi} = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_N \end{bmatrix} \tag{3.9}$$

$$\mathbf{\Phi}^T\mathbf{M}\mathbf{\Phi} = diag(\mathbf{m}_i) \tag{3.10}$$

$$\mathbf{\Phi}^T\mathbf{K}\mathbf{\Phi} = diag(\mathbf{k}_i) \tag{3.11}$$

The matrix $\mathbf{\Phi}$ defines modal coordinates $\mathbf{q}(t)$ where $\mathbf{u}(t) = \mathbf{\Phi}\mathbf{q}(t)$. In these modal coordinates the equations of motion are decoupled into single degree of freedom systems defined by modal masses $\mathbf{m}_i$, damping $\mathbf{c}_i$, stiffnesses $\mathbf{k}_i$, and forces $\mathbf{f}_i(t) = \phi_i^T\mathbf{f}(t)$. Under the common assumption of Rayleigh damping, modal damping can be expressed by $\mathbf{c}_i = \alpha\mathbf{m}_i + \beta\mathbf{k}_i$ giving the decoupled equation of motion for each mode

$$\mathbf{\ddot{q}}(t) + 2\xi_i\omega_i\mathbf{\dot{q}}(t) + \omega_i^2\mathbf{q} = \frac{\mathbf{f}_i}{\mathbf{m}_i} \tag{3.12}$$

where the undamped natural frequency is $\omega_i = \sqrt{\frac{\mathbf{k}_i}{\mathbf{m}_i}}$, giving the modal damping factor

$$\xi_i = \frac{\mathbf{c}_i}{2\mathbf{m}_i\omega_i} = \frac{1}{2}\left(\frac{\alpha}{\omega_i} + \beta\omega_i\right) \tag{3.13}$$

We can then obtain the unit impulse response for the ith mode by solving Equation

3.12

$$h_i(t) = \left(\frac{e^{-\xi_i\omega_i t}}{\mathbf{m}_i\omega_{di}}\right) sin(\omega_{di}t) \tag{3.14}$$

where the damped natural frequency is $\omega_{di} = \omega_i\sqrt{1 - \xi_i^2}$. Given Equation 3.14 we can construct the response of an object to an arbitrary impulse as the superposition of that object's 1D modal responses.

Taking the Fourier transform of the unit impulse response $h_i(t)$ the product in Equation 3.14 becomes the convolution

$$H_i(\omega) = \left(\frac{1}{\mathbf{m}_i\omega_{di}}\frac{\xi_i\omega_i}{\xi_i^2\omega_i^2 + \omega^2}\right) * \left(\frac{\delta(\omega - \omega_{di}) - \delta(\omega + \omega_{di})}{i}\right) \tag{3.15}$$

which convolves the Fourier transform of the decaying exponential, a Lorentzian distribution; and a pair of delta functions. In other words, the transfer function of a single mode is the convolution of a spike at its resonant frequency and a Lorentzian with a width that depends on modal frequency and damping.

Eigenmodes in Image-Space

In this section we relate deformations observed in video to projections of the mode shapes ϕ_i and show that these projections can be used as a basis for representing image-space dynamics. We first consider the dynamics of a single degree of freedom, which we later relate to the motion of a visible point in video.

An excitation force $\mathbf{f}$ given in modal coordinates can be decomposed into a set of impulses $\mathbf{f}_i = A_i\delta(t)$ where A_i is the amplitude of the impulse at mode ϕ_i. Applying Equation 3.14, the response of the object at one degrees of freedom $u_p(t)$ is given by

$$u_p(t) = \sum_{i=1}^{N} A_i h_i(t)\phi_i(p) \tag{3.16}$$

where $\phi_i(p)$ is the mode shape coefficient of the degree of freedom p of the object for mode i. Using Equations 3.15 and 3.16 we can construct the Fourier transform of Equation 3.16 as

$$U_p(\omega) = \sum_{i=1}^{N} A_i H_i(\omega)\phi_i(p) \tag{3.17}$$

Here we make an assumption that is common in engineering modal analysis [32, 8], but not necessary in FEM-based applications of modal analysis for simulation: that modes are well spaced, or non-overlapping in the frequency domain. Under this assumption, we can represent the frequency response of a single degree of freedom at ω_{di} as

$$U_p(\omega_{di}) = A_i H_i(\omega_{di})\phi_i(p). \tag{3.18}$$

Our next assumption is weak perspective - a common approximation in computer vision, but one that is also not necessary when modes are derived from known models. Using this approximation we align our object's coordinate system with the image plane of an input video, giving us observable degrees of freedom for each pixel's motion in the x and y dimensions of our image. For the purpose of derivation, we represent visibilty across all degrees of freedom with the unknown, binary, diagonal matrix $\mathbf{V}$, which multiplies the visible degrees of freedom in a mode by 1 and all other degrees of freedom by 0. The projection of a mode shape ϕ_i into the image plane is then $\mathbf{V}\phi_i$.

By taking Fourier transforms of all local motions $\mathbf{Vu}$ observed in video we obtain $\mathbf{VU}$, the Fourier spectra for visible degrees of freedom, which, evaluated at resonant frequencies ω_{di}, is

$$\mathbf{VU}(\omega_{di}) = A_i H_i(\omega_{di})\mathbf{V}\phi_i. \tag{3.19}$$

Here, A_i and $H_i(\omega_{di})$ are constant across all degrees of freedom p, meaning that $\mathbf{VU}(\omega_{di}) \propto \mathbf{V}\phi_i$. Therefore we can treat the set of complex ϕ_i', the values of $\mathbf{VU}(\omega_{di})$ measured in video, as a basis for the motion of the object in the image plane.

Assumptions and Limitations

While linear motion is a standard assumption of linear modal analysis that usually applies to the type of small motion we are analyzing, our derivation makes a few key approximations that are not typical of modal analysis applied to simulation:

- *Weak Perspective* - Our analysis assumes that linear motion in 3D space projects to linear motion in the image plane. This can be violated by large motion in the z-plane.

- *Well-spaced modes* - We rely on separation in the frequency domain to decouple independent modes. This can fail in objects with strong symmetries, high damping, or independent moving parts.

- *Broad-Spectrum Forcing* - By using observed modes as a basis for the motion of an object in the image plane, we make an implicit assumption about the ratio of modal masses to observed modal forces. Allowing for an ambiguity of global scale, this assumption is still violated when observed forces are much stronger at some modes than others.

Because we deal with small motion around a rest state, weak perspective is almost guaranteed to be a safe approximation. However, there are many cases where our remaining two assumptions could fail. Fortunately, the consequences of these failures tend to affect the accuracy more than the plausability of simulation. Consider the failure cases of each approxiation. Overlapping modes will cause independent objects to appear coupled in simulation - in other words, the response of an object to one force will incorrectly be an otherwise appropriate response to multiple forces. Similarly, when broad-spectrum forcing is violated, the response of a object to one force will be the appropriate response to a differently scaled, but equally valid set of forces. In both cases, the failure results in inaccurate, but still plausible deformations of the object.

3.5.2 Algorithm

Our algorithms first extracts a volume of candidate vibration modes from an input video. We then provide a user interface for selecting a subset of these candidate modes to use as a basis for simulation.

Extracting Candidate Modes

We measure optical flow in the x and y dimensions of an input video using phase variations of a complex steerable pyramid [85]. This approach has been shown to work well for small motion in several recent works [99, 100, 31], though Lagrangian flow algorithms may be equally well suited to our application. To filter local displacements, we employ the weighted gaussian filtering used in [99]. Local displacements are first given weights proportional to local contrast. The weighted displacements and the weights are both blurred spatially, then the filtered displacements are normalized by their filtered weights. This denoises displacement signals by causing regions with low image contrast, and therefore noisy displacements, to take on the values of nearby regions with high image contrast.

Next we compute the temporal FFT of our filtered displacement signals as in [31]. Each spatial slice of the resulting temporal frequency spectra forms a candidate shape for a possible mode at that frequency.

Note that while our approach combines aspects of [99] and [31], neither of these techniques individually produce mode shapes that are suitable for simulation. [31] employ filtering that introduces nonlinearities into mode shapes, making their recovered modes suitable for visualization, but not synthesis, while [99] visualize modes by magnifying deformations, but do not compute a static representation of their shape.

Mode Selection:

Under ideal conditions, the observed candidate modes ϕ'_ω at each frequency ω would be zero everywhere but at real mode shapes. However, real video contains unintended motion from a variety of sources (e.g., camera shake, noise, moving background). To

distinguish between object deformations and unintended motion from other sources, we first ask users to provide a rough mask of the content they are interested in. We then present them with a graphical interface to help select mode shapes.

Our mode selection interface (shown in Figure 3-4) displays a representative image from the input video, a power spectrum showing the magnitude of motion observed at each frequency, and a visualization of the current selected candidate mode, chosen by the user. The power spectrum shows the average amplitude of unmasked coefficients in each candidate mode shape. In very recent work, [29] showed that resonant modes of an object can be identified as peaks in a similar power spectrum (though the spectra they use are based on the motion signals described in [31]). When a user clicks on the spectrum in our interface, we find the frequency with maximum energy in a small window around the user's mouse, and display the corresponding candidate mode in our shape window. We use the same visualization of candidate mode shapes described in [31] - phases are mapped to hue, and magnitudes are mapped to intensity. Users can select either a set or a range of modes by clicking on different peaks in the power spectrum. This selection process is similar to peak-picking methods that have been used for modal identification of structures in engineering [32]. Informed users are generally able to select a suitable set of mode shapes in less than a minute, though some training to know how to identify 'good' mode shapes is necessary. For a video of mode selection refer to our supplemental material.

Complex Mode Shapes:

Note that the set of mode shape solutions ϕ_i to Equation 3.8 are real-valued, i.e. they only have binary phase relationships. Similarly, the mode shapes derived using FEM in typical simulation applications are also real-valued. In contrast, the mode shapes we recover may have non-binary phases. This can happen for a number of reasons, including noise or a violation of one of our assumptions. We could force mode shapes to be real-valued by projecting them onto their dominant axis in the complex plane, however, we found that allowing non-binary phases actually improves results. Visually, such mode shapes allow for features like traveling waves and partial coupling

that might otherwise require much higher-order modes to represent. By allowing these shapes, we effectively let our representation fit the motion in a video more closely. In this sense, our technique is allowed to behave a little more like methods for exemplar-based motion texture synthesis in situations where motion cannot be explained well with sparse, low-frequency modes.

To ensure that the behavior of our simulation reduces to one using only real mode shapes when observed modes contain only binary phase relationships, we calculate the dominant orientation of each selected mode shapes on the complex plane, and rotate all phases so that this orientation aligns with the real axis.
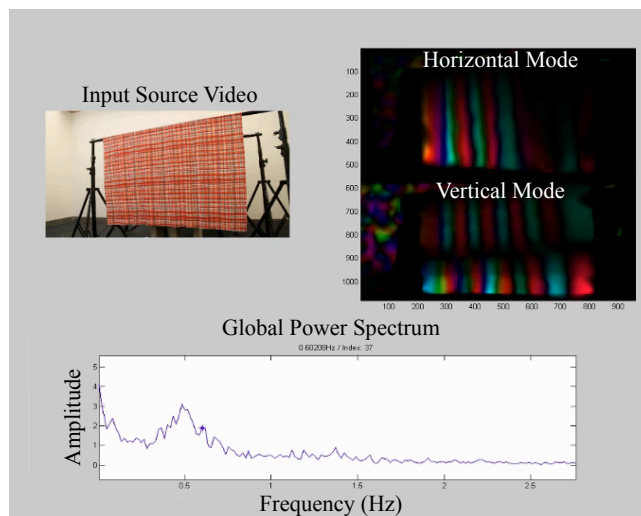
Figure 3-4: To use our mode selection interface, users click on a frequency in the video's motion spectrum (bottom) and are shown a visualization at the corresponding candidate mode shape (right). Using this interface users can select either an individual, or a range of candidate modes to use as a basis for simulation. [30]

Simulation

Our simulation works on the state of an object in modal coordinates. The key components are a way to evolve the state of an object over time, and a way to translate user input into forces, displacements, and velocities.

Given Equation 3.12, we can define a state space model per modal coordinate to simulate the the object over time. We define the state vector $\mathbf{y_i}$ that describes the

system for a single modal coordinate $\mathbf{y_i} = [\varrho_i \ \dot{\varrho}_i]^\intercal$, where ϱ_i and $\dot{\varrho}_i$ are the modal displacement and velocity vectors respectively which relate to the complex modal coordinate by $\mathbf{q_i} = \varrho_i - i\dot{\varrho}_i/\omega_i$. We evolve the state to $\mathbf{y}[n+1]$ given $\mathbf{y}[n]$ and a modal force $\mathbf{f_i}$ using the equation [81]:

$$\mathbf{y}[n+1] = \begin{bmatrix} 1 & h \\ -\omega_i^2 h & 1 - 2\xi_i\omega_i h \end{bmatrix} \mathbf{y}[n] + \begin{bmatrix} 0 \\ h/\mathbf{m}_i \end{bmatrix} \mathbf{f}_i[n], \qquad (3.20)$$

and set h, the amount of time passed in the simulation, to be small enough to ensure that this equation is stable.

3.5.3   User Input

We provide users with modes of interaction that can be divided into two categories: forcing interactions and direct manipulations. Forcing interactions affect state indirectly by changing the force $\mathbf{f}_i$ applied to an object. Direct manipulations translate user input directly into intantaneous state $\mathbf{y}$.

Forcing Interactions: Forcing interactions translate user input into a force to be applied at a specified point. In the simplest forcing interaction, a user clicks at a point $\mathbf{p}$ on the object, and drags their mouse in a direction $\mathbf{d}$. We interpret this as specifying a force $\mathbf{f}$ to be applied at the point $\mathbf{p}$ in the direction $\mathbf{d}$. The scalar modal force $\mathbf{f_i}$ applied to each mode is computed by taking the magnitude of the dot product of $\mathbf{d}$ with the value of that mode shape ϕ_i' at point $\mathbf{p}$:

$$\mathbf{f_i} = \|\mathbf{d}^\intercal \phi_i'(\mathbf{p})\| \alpha \qquad (3.21)$$

where α is used to control the strength of the force, and can be set by the user with a slider. Note that we take the magnitude here because the mode shape ϕ_i' is complex.

Direct Manipulation: Real objects are often found in configurations that are difficult or impossible to achieve through forces applied to one point at a time. However,

fully specifying shaped forces is a difficult user interaction problem. We instead offer a mode of interaction that lets users directly manipulate the position or velocity of a single point. This lets users explore states with greater contributions from higher-order modes that are difficult to achieve without shaped forces. We accomplished this by explicitly setting the state of the object whenever the user's mouse is pressed, and only letting the state evolve once the mouse is released. As with forcing interactions, the user specifies a point $\mathbf{p}$ and direction $\mathbf{d}$ with a mouse. We then compute the magnitude of each modal coordinate in the same way that we computed the magnitude of modal forces before:

$$\|\mathbf{q_i}\| = \|\mathbf{d}^\intercal \phi_i'(\mathbf{p})\|\alpha \tag{3.22}$$

where α is used to control the strength of the manipulation, and can be set by the user with a slider. However, in this case we set the phase of the modal coordinate to maximize either the displacement or velocity of $\mathbf{p}$ in the direction $\mathbf{d}$. This is accomplished by setting the phase $Arg(\mathbf{q_i})$ to

$$\text{Max Displacement:} \quad Arg(\mathbf{q_i}) = -Arg(\mathbf{d}^\intercal \phi_i'(\mathbf{p})) \tag{3.23}$$

$$\text{Max Velocity:} \quad Arg(\mathbf{q_i}) = -Arg(\mathbf{d}^\intercal \phi_i'(\mathbf{p})) + \frac{\pi}{2} \tag{3.24}$$

For objects with real mode shapes, velocity is maximized when displacements are zero, and displacement is maximized when velocities are zero. Intuitively, maximizing displacement lets users 'pull' a point around the screen and see how the the object deforms in response, while maximizing velocity specifies an impulse to be applied when the mouse is released.

Rendering Deformations

We render the object in a given state by warping a single color image, representing the object's rest state, by a displacement field $\mathbf{D}(t)$. $\mathbf{D}(t)$ is calculated as a superposition of mode shapes weighted by their respective modal coordinates:

$$\mathbf{D}(t) = \sum_i^N \mathbf{Re}\{\phi_i' q_i(t)\} \tag{3.25}$$

This can be evaluated efficiently on the GPU by representing each ϕ_i' as an RGBA texture storing two complex numbers per pixel, corresponding to the coupled image-space x and y displacements of ϕ_i'. Each $\phi_i' q_i(t)$ term is computed in a single rendering pass, accumulating $\mathbf{D}_t$ in a framebuffer that can be applied as a displacement map to the color image in a final pass. Our implementation uses depth culling and assigns pixels depth values that are inversely proportional to the magnitude of their displacement, causing parts of the image that move more to occlude parts that move less. This tends to work better than blending pixel values in practice, as objects closer to the camera usually exhibit larger screen space motion due to foreshortening.

Implementation Details

Our mode extraction and selection interface are written in MATLAB. Once modes have been selected, they are exported as 8-bit RGBA TIFF images, and loaded into our simulation software, which is written in C++ and uses Qt, OpenGL, and GLSL.

The slowest part of our algorithm is building a complex steerable pyramid on the input video. Using the MATLAB implementation from [85] this takes less than two minutes on shorter videos like the Wireman, but can take 2-3 hours on longer, or high-speed videos like the Ukulele. The only parameter we set for this is the standard deviation of the gaussian used for filtering local motion signals. Our strategy for setting this parameter is to effectively test out 4 values at once - we pick a standard deviation that is 5-10% of the larger image dimension, filter with this standard deviation at all scales, and use the highest-resolution scale that does not appear noisy. Mode selection can then usually be done in less than a minute, but users may choose to spend more time exploring the recovered spectra with our selection interface.

In the Playground, YoutubeBridge, and ForceTree examples we use inpainting to fill disoccluded parts of the image.

3.5.4 Results

We tested our method on several different examples. Thumbnails showing the rest state of each example can be found in Table 3.2 along with additional details about the corresponding input video.

All of the input videos that we captured were recorded with a tripod. The input video for YoutubeBridge was downloaded from Youtube user KOCEDWindCenter (link).

Our simulations plausibly reproduce the behavior observed in most input videos, and works well with regular cameras operating at 30 frames per second. While higher-frequency modes exist in most objects, their fast temporal dynamics are not usually visible in output videos, just as they are not observable in the input. Our Ukulele example explores the use of a high speed camera to recover modes that are not visible at normal framerates.

Interactive Animations   In each interactive session, an arrow is rendered to indicate where users click and drag. The head of the arrow points to the current mouse location, and the tail of the arrow ends at the displaced point $\mathbf{p}$ where the user initially clicked.
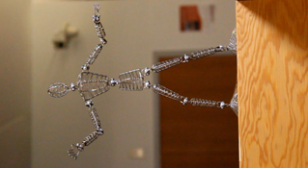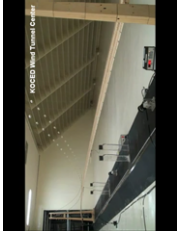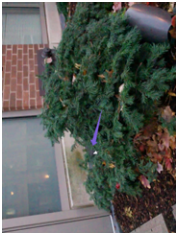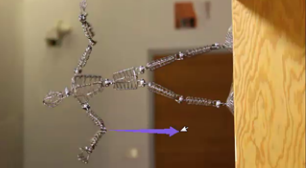
For the most part, these interactive animations are quite compelling. However, in some cases where our non-overlapping modes assumption is violated, independent parts of a scene appear coupled. This effect is subtle in most of our results, so we include an additional failure case designed to violate this assumption in our supplemental material (labeled 'dinos1'). The example shows two dinosaur toys with similar motion spectra resting on the same surface. When a user interacts with one of the toys, this causes some motion in the other toy as well. This problem could be addressed in the future by asking users to provide multiple masks, indicating independent parts of the scene.

Special Effects   A variety of visual effects can be be achieved by specifying forces in different ways. We explore the possibility of using this to create low-cost special

effects. For example, by using forcing interactions and setting $\mathbf{d}$ to be a vector pointing down, we can simulate the effect of increased weight at the point $\mathbf{p}$. For example, we can use this to simulate a small robot rolling along the surface of different objects. When the robot 'lands' on a point $\mathbf{p}$ of the object, we fix the robot to $\mathbf{p}$ by applying the time-varying displacement at $\mathbf{p}$ to the image of the robot at each frame. By moving $\mathbf{p}$ along a trajectory specified in the object rest state, we cause the robot to 'roll' along the object's surface in a way that couples their dynamics.

In another example, ForceTree, we control the force $\mathbf{d}$ applied to branches of a tree so that the branches appear to be controlled by a moving hand elsewhere in the video. In this way, we make it appear as though the leaves of the tree are coupled (or controlled through some supernatural force) by the hand. This is substantially simpler that modeling a synthetic tree and matching its appearence to the filmed scene.

Table 3.1: This table gives a summary of the experimental results. The first row contains the names of all the examples. The middle row contains an image from the input video representing the rest state of the object, and the bottom row is an example of a synthesized deformation.

Example Name							
Bush	Playground	Cloth	Wireman	Ukulele	YoutubeBridge	ForceTree
Input Video Image							
Synthesized Deformation							

Table 3.2: This table gives a summary of the parameters of the experimental results. We give the source, length, framerate, and resolution of the source video. The excitation column describes the type of excitation used to excite the object in the input video where: ambient/wind means natural outdoor excitations mostly due to wind, impulse means that the object or its support was manually tapped, and sound means that a ramp of frequencies was played from 20 Hz to the Nyquist rate of the recorded video. We give the number of mode shapes identifed from the input video local motion spectra that are used to simulate the object response and in the final column, the frequency range of these mode shapes.

Example	Source	Source Length (s)	Framerate (fps)	Resolution	Excitation	Number of Modes	Frequency Range
Bush	SLR	80.18	60	640 × 480	Ambient/Wind	77[†]	1.3 - 4.2 Hz
Playground	SLR	53.85	60	1280 × 720	Impulse	34[†]	0.8 - 22 Hz
Cloth	SLR	59.77	30	1920 × 1080	Ambient/Wind	147[†]	0.3 - 0.8 Hz
Wireman	SLR	5.82	60	720 × 1280	Impulse	6	5 - 20 Hz
Ukulele	High-speed camera	8.87	1400	432 × 576	Sound	13	219 - 670 Hz
YoutubeBridge	Youtube (link)	50	30	640 × 480	Wind	4	0.25 - 11 Hz
ForceTree	SLR	35	60	1280 × 720	Impulse	13	0.6 - 9 Hz

† Range of frequencies selected

3.6 Chapter Summary

This chapter gives a description of how small motions of objects in videos are measured and how they can be used for different purposes. The small motions can be processed into resonant frequencies and mode shapes and used to measure the real displacements of objects. They can be used to make visualizations of an object's motion through motion magnification. They can be also used to create a plausible interactive simulation where forces can be applied to an object and the object's plausible reactions can be visualized.

Chapter 4

Verification and Noise Floor Experiments

This chapter presents a verification of the camera measurement with traditional sensors, a study of the parameters impacting the noise floor of a camera measurement, a comparison of video processing algorithms, a demonstration of motion compensation and sensor fusion techniques, and a discussion on the total noise floor including the effects of different parameters and camera motion.

4.1 Verification with Traditional Sensors

This section describes a laboratory experiment where the camera is measured against an accelerometer and a laser vibrometer. This verification measurement was originally presented in [23, 24].

4.1.1 Experimental Setup

In order to validate the motion magnification algorithm as applied to videos of vibrating structures for the measurement of displacements and operational deflection shapes, an experiment was formulated to compare the results to standard sensors. An accelerometer was mounted on the free end of a steel cantilever beam, and the
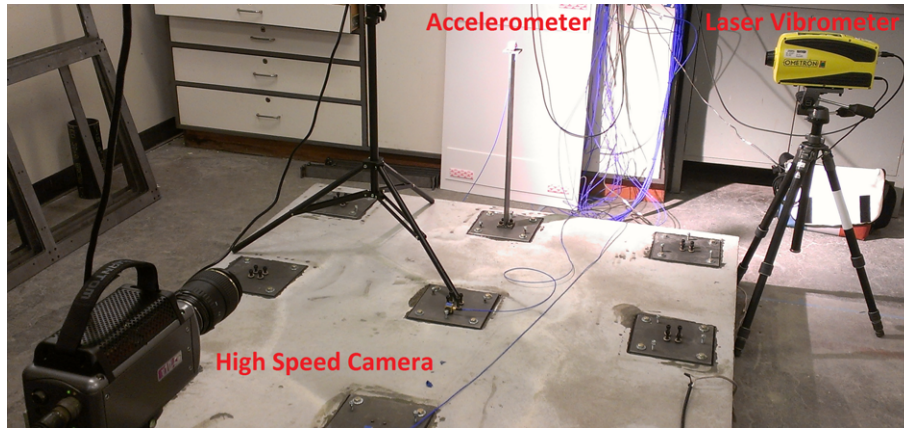
(a)

(b)

Figure 4-1: Verification test: (a) experimental setup and (b) scene from high speed camera video (artifically brightened in text for visibility)

motion of the accelerometer was simultaneously measured by the accelerometer itself, a laser vibrometer, and a high speed camera, as shown in Figure 4-1(a). A screenshot of the video from the camera is shown in Figure 4-1(b), and the resolution of the camera was 480×288, and the frame rate was 5000 frames per second. In the plane of the accelerometer, the video frame was 104 mm wide and 62 mm tall, with 4.615 pixels per millimeter. The cantilever beam was excited with an impact hammer near the base, and the subsequent vibration was measured for comparison. The velocity time series from the laser vibrometer was integrated to obtain displacement to verify against the displacements derived from camera measurements of the optical flow of the accelerometer movement. Time synchronization was not possible between the camera and laser vibrometer data set, so in the data the time series were aligned by hand. A fast Fourier transform (FFT) was performed on the data from the laser vibrometer, accelerometer, and camera derived displacement, and integrated in the

frequency domain to obtain the displacement so that the frequency peaks and noise floors could be directly compared.

4.1.2   Results

The data collected from the verification measurement were the various signals of the accelerometer movement as mounted on the cantilever beam as measured by a laser vibrometer, an accelerometer, and a high speed camera. The displacement from the camera was extracted using the local phase of the measured video as described in the theory and derivation section. The raw signals collected were the displacement time series from the camera, velocity time series from the laser vibrometer, and acceleration time series from the accelerometer. To directly compare the time series results, the laser vibrometer velocity time series was integrated numerically and results are shown in Figure 4-2. The double numerical integration of the acceleration signal in the time domain to obtain displacement did not give satisfactory results due to low-frequency contamination, and is therefore not shown. The issue of low-frequency contamination in a double integrated accelerometer signal is a known issue and motivates direct measurement of displacements when possible [103]. The displacement derived from the camera closely matches the integrated laser vibrometer displacement for the whole 9 seconds of data, however the camera signal does show less detail and has more noise than the laser vibrometer signal. For a quantitative comparison the correlation between the camera and laser vibrometer displacement signals is 99.6%.

The signals were also compared in the frequency domain by using the FFT and then integrating the accelerometer and laser vibrometer signals in the frequency domain to obtain the displacements for each of the measurement methods. This allowed for easier comparison of the noise floor of the displacement derived from the camera with the other measurement methods. The comparison plot is shown in Figure 4-3. Both the accelerometer and laser vibrometer data show eight resonant frequencies above the noise floor from 0 to 2500 Hz, while the camera only shows the first four resonant frequencies of the cantilever beam. The noise floor of the camera for this nine second measurement is 40 nanometers, while the laser vibrometer has a noise floor of
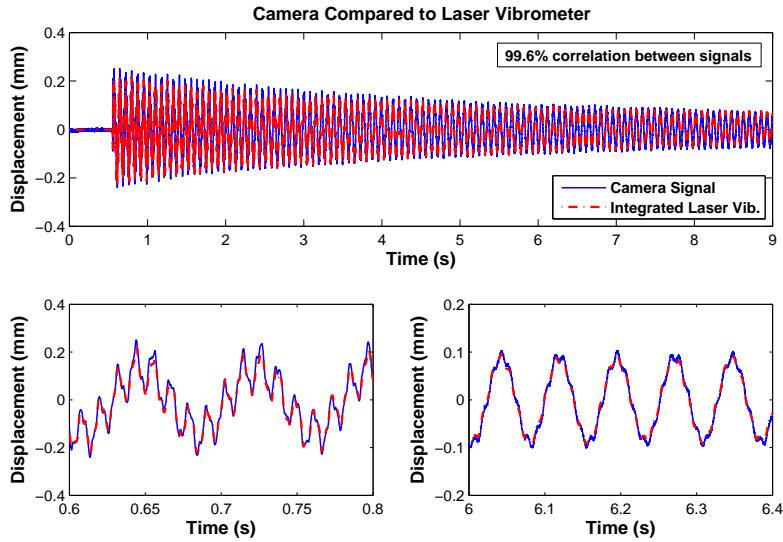
Figure 4-2: Comparison between displacements derived from the camera and the laser vibrometer

0.2 nanometers, and the accelerometer has a noise floor of 0.02 nanometers. Given the conversion factor of 480 pixels for 104 mm, and accounting for the length of the measurement, the noise floor of the camera for this measurement was approximately 5.5×10^{-4} pixels per root Hertz. Possible methods for improving the noise floor of the camera measurement are to increase the number of pixels per millimeter in the frame or to capture more spatial scales with different levels or spatial downsampling of the video. Both increasing the number of pixels and capturing more spatial scales require more video processing time. Increasing the number of pixels may also reduce the achievable frame rate as there is a limit due to the camera hardware. This verification measurement gives us reasonable confidence that the displacement extracted from video is accurate and that the operational deflection shapes of a column can be successfully measured by a high-speed camera.
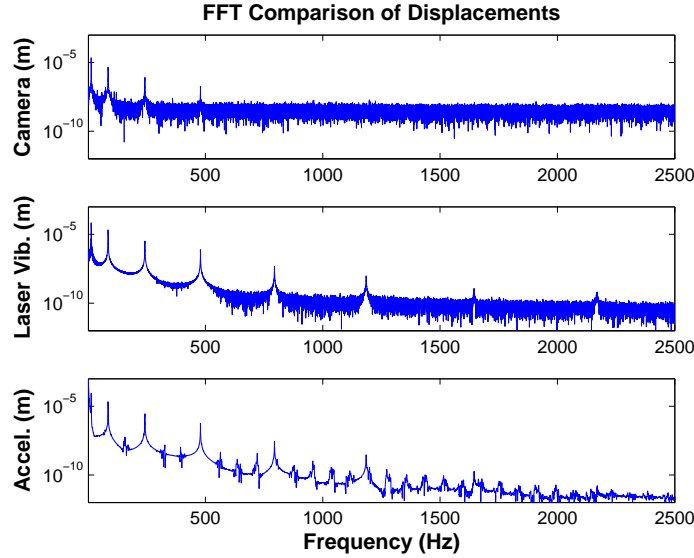
Figure 4-3: Frequency space comparison between displacements derived from the camera, laser vibrometer, and accelerometers

4.2 Noise Floor Parametric Study

The verification measurement showed that the camera measurements were accurate, but with a higher noise floor than an accelerometer and laser vibrometer. To quantify the effects of different parameters on this noise floor, a set of measurements were made with a similar experimental setup under different sets of conditions.

4.2.1 Experimental Setup

The experimental setup, like the verification measurement includes a high-speed camera, laser vibrometer, and accelerometer for measuring the motion of a cantilever beam, as shown in Figure 4-4. The lights used were battery powered, so they did not flicker. There was also a black backdrop to ensure even lighting and reduce the appearance of any shadows. To excite the cantilever beam, a computer controlled shaker was attached to the top which can introduce excitations of any waveform in its frequency response range of approximately 5 Hz to 300 Hz. A detailed characterization of the shaker's response is included in Appendix C.2.

The camera nominally measured the motion of an accelerometer mounted to the
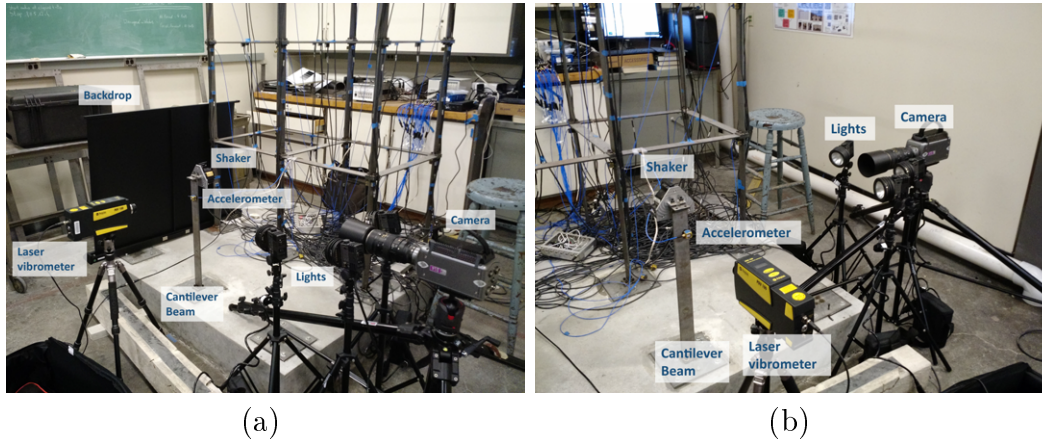
Figure 4-4: Experimental setup from two different angles

cantilever beam, which was covered with a piece of white paper to ensure even lighting of the accelerometer and clean strong edges between it and the background. Measurements were made at 2000 fps with a 490 microsecond exposure time, at a resolution of 576×576 pixels, and 10-bit depth. The 85mm setting on the zoom lens was used on the camera, so the scaling factor was 0.1669 mm per pixel or a magnification ratio of 1:14.5, since the pixel size for the camera is 11.5 μm. A screenshot from a recorded video is shown in Figure 4-5(a) and (b) the pixel mask showing the pixels from which displacement signals were extracted. Only the edges of the accelerometer were used to be consistent between different measurements. To generate a single signal for each camera measurement, the signals from the pixels were averaged together, however they were also analyzed as individual signals from individual pixels. Measurements of the motion were also made with the accelerometer and a laser vibrometer for comparison, both sampling at 6000 Hz.

An example measurement without any excitation from the shaker, from the video shown in Figure 4-5(a) is shown in Figure 4-6. In the frequency domain there is still a large peak in the camera measurement at 74.26 Hz which is attributed to the cooling fan on the high-speed camera. In the spectrum of the accelerometer, peaks corresponding to resonant frequencies of the cantilever beam can be seen, due to very small amounts of motion. In the time domain, the signal of the camera had a much

(a)   (b)
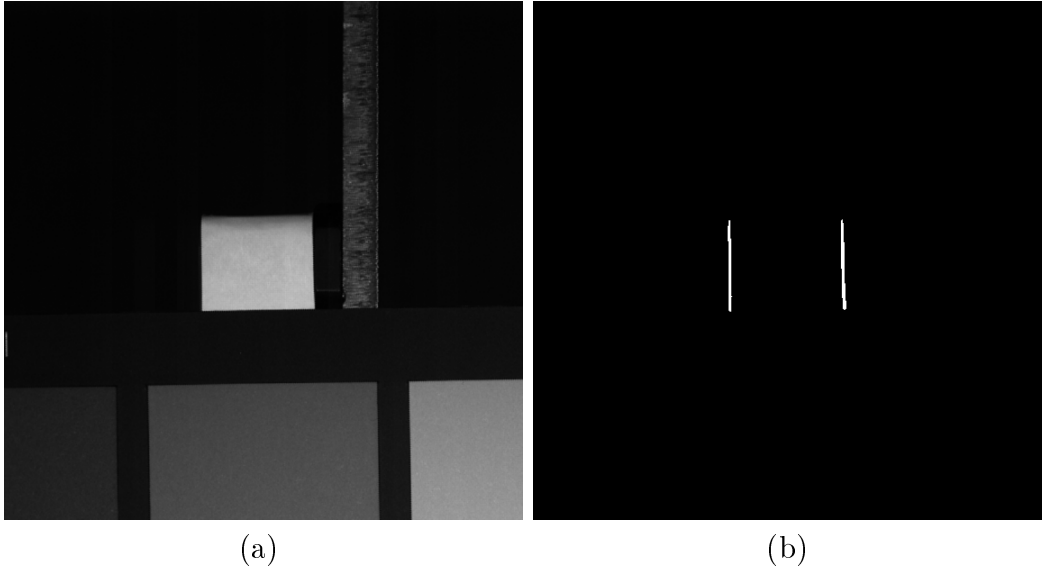
Figure 4-5: (a) Screenshot from recorded video without motion and (b) pixel mask showing pixels from which displacement signals were extracted

larger amplitude than either the laser or the accelerometer, however this corresponds to the vibration from the cooling fan on the camera itself.

(a)   (b)
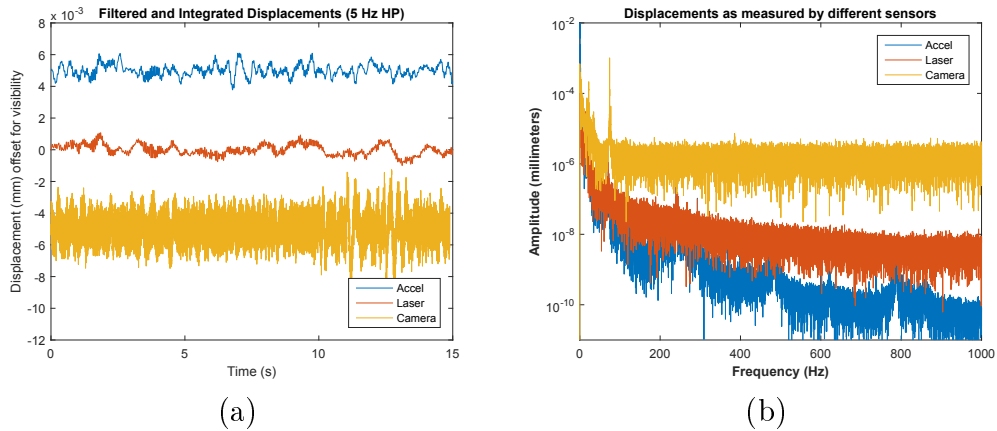
Figure 4-6: Measured displacements from the camera, laser vibrometer, and camera in the (a) time domain and (b) frequency domain

4.2.2 Exposure (Lighting)

The first parameter that was explored that is suspected to have the largest impact on the noise floor is the local contrast of the edge or texture, from which the displacement is being measured. In field measurements this would depend on the visual texture on an object, contrast between a foreground object of interest and the background, and lighting conditions. To change this in a systematic way, we change the exposure time of the camera to reduce the lighting and assume that the black background stays the same brightness, while the white paper on the accelerometer gets darker as it is exposed for shorter times per frame. Exposure times from 12.5 microseconds up to 490 microseconds were considered. Screenshots of the videos from all these measurements and the resulting measured frequency response are included in Appendix C.3.
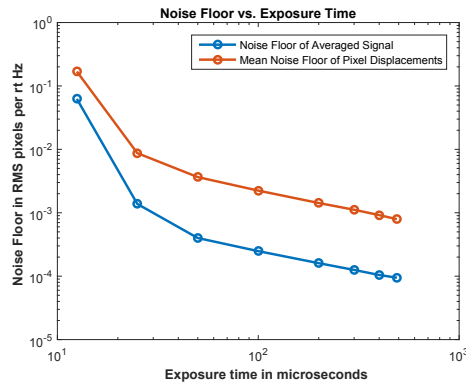
Figure 4-7: Noise floor in pixels per rt Hz for an averaged signal for a whole video and the mean noise floor of individual pixel displacements vs. exposure time as a proxy for lighting conditions

To measure the noise floor, we look at the frequency spectrum of the measured camera displacement of the accelerometer from 100 Hz up to the Nyquist frequency of 1000 Hz. The maximum value of the response in a sliding window of 100 frequency bins is averaged over the interval is considered the noise floor, as a signal above that value should stand out from the noise floor. The noise floor of the average displacement signal for each video and the mean noise floor of individual pixels is shown in Figure 4-7. The individual values for each pixel in the measured videos are plotted against the edge strength, the amplitude of the G2 filter in Figure 4-8(a) and

against the brightness, measured with a 9x9 Gaussian filter in Figure 4-8(b). In both plots, a trend can be seen where the best possible noise floor for any pixel lies on a line defined by the equation $10^{-3.35} * x^{-2/3}$. This equation can be used to quickly determine the likely noise floor from just a single frame of a video, assuming small motions.
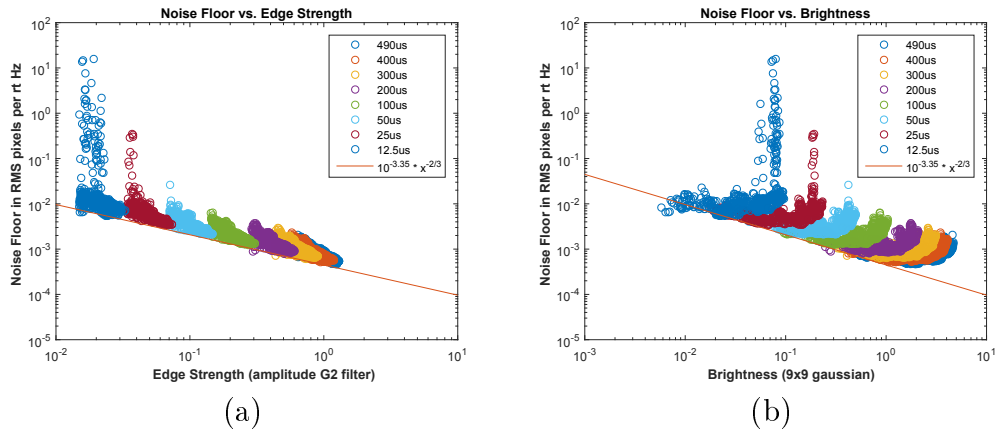
Figure 4-8: Noise floor for individual pixels as a function of (a) edge strength and (b) general brightness

4.2.3   Frame Rate

Another parameter that was varied to see if there was any effect on the noise floor was the frame rate. To change the frame rate while keeping the exposure time constant the video was downsampled by integer amounts. There was no noticeable change in the noise floor, since the amplitudes and amount of noise are the same in the time domain. A summary of the results is shown in Table 4.1. Lower noise floors can be achieved at lower frame rates, because longer exposure times can be used, potentially increasing the amount of contrast between dark and bright objects.

Table 4.1: Relationship between frame rate and noise floor, keeping the brightness constant by downsampling a higher frame rate video

Frame Rate (fps)	Downsampling Factor	Noise Floor of Avg. Displacement (Pixels/rtHz)
2000	n/a	7.9236e-4
1000	2	7.8779e-4
666.7	3	7.8586e-4
500	4	7.8196e-4
400	5	7.8534e-4
333.3	6	7.8310e-4

4.2.4 Lens Focal Length (Distance)

To vary the effective distance from the camera to the object being measured, without changing the experimental setup, the zoom lens used was zoomed out. Measurements were made with the lens set to 85mm, 50mm, and 24mm with screenshots from the measured videos shown in Figure 4-9.

(a)   (b)   (c)
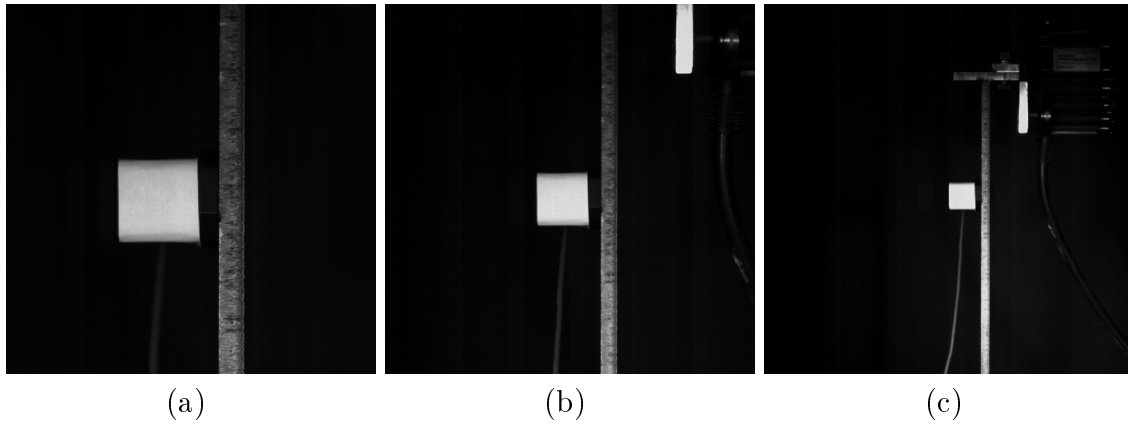
Figure 4-9: The effect of distance by changing the lens zoom, screenshots from videos at (a) 85mm, (b) 50mm, and (c) 24mm; note how the accelerometer looks brighter in the screenshot from the 24mm video

The measurement results and plots of the individual pixel noise floors are shown in Figure 4-10. The noise floor for individual pixels actually changes between the different lens zooms, with the 24mm lens measurement having the best noise floor, while the edge contrast for the individual pixels also changes as well. The lens used on the high-speed camera, is listed as having an f-stop of f/3.5-4.5 which means that

102

at 24mm the lens has an f-stop of f/3.5 while at 85mm the lens has an f-stop of f/4.5, which means that the lens lets in more light at 24mm than at 85mm. The difference in noise floor is thus explained by the change of brightness in the video, rather than anything to do with the change in feature sizes from the lens zoom.
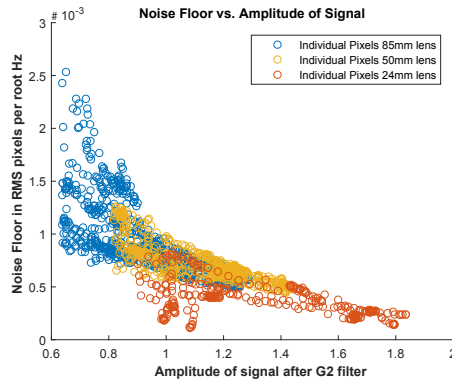
Figure 4-10: Noise floor of individual pixels for different lens zooms showing that the effect on the noise floor is mostly a function of the change in edge strength due to changes in brightness

4.3   Comparison of Video Processing Algorithms

Different video processing algorithms were run on a similar set of videos from the same experimental setup. In this thesis, the method used for determining displacements from video, phase-based optical flow, is compared against an open source package for digital image correlation called Ncorr [5]. The phase-based optical flow algorithm is also run with two different sets of filters, an undercomplete quadrature filter pair, G2 and H2, and an overcomplete filter bank, the complex steerable pyramid (CSP) filters.

4.3.1   Experimental Setup

Two differently sized videos were used to compare video processing algorithms. A full frame of 576×576 and a cropped frame of 301×201 were used, as shown in Figure 4-11 to determine the change in processing time with an increase in the video resolution

to be processed. For this measurement, the shaker was used to harmonically excite the cantilever beam during the video measurement.

(a)   (b)
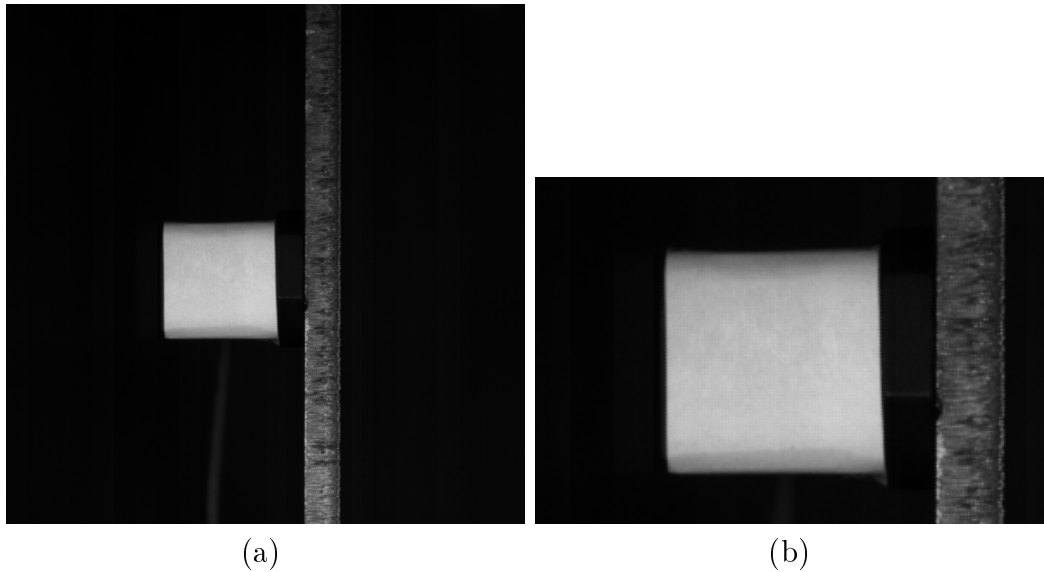
Figure 4-11: Screenshot from the video used to compare different video processing methods with (a) the full frame and (b) a cropped frame

4.3.2   Results

The signals extracted from the video is shown in Figure 4-12, showing them on top of each other for comparison, and vertically separately for clarity. The signals extracted from video with the three different processing methods are extremely similar. The noise floors from this measurement for the different processing methods are shown in Table 4.2, showing improvements for both phase-based optical flow methods, G2 H2 and CSP over Ncorr DIC, although they are all quite similar. The differences between the noise floors likely come down to the differences in the spatial averaging of pixels to obtain the single displacement signal for the cropped video.

The most stark differences between the processing methods are in the processing times for the videos. 100 frames of video Ncorr DIC processing is much slower than the phase-based optical flow methods. The G2 H2 method is the fastest as it uses only 2 linear filters to determine the displacement, and only in the horizontal direction in this
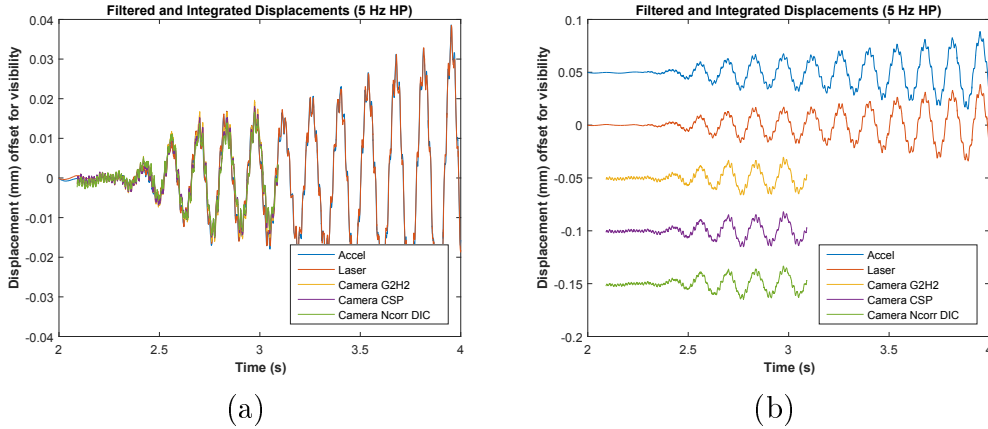
Figure 4-12: Measured displacement signals from the different video processing methods and by integrating the accelerometer and laser vibrometer measurements shown (a) on top of each other and (b) separated vertically for clarity

Table 4.2: Comparison of noise floor for different video processing methods

Method	Noise Floor (pixel/rtHz)
G2 H2	8.0560e-5
CSP	1.0195e-4
Ncorr DIC	1.2521e-4

test. The difference in times between the cropped and the full frame is approximately the same as the increase in the number of pixels for the phase-based optical flow methods, while for Ncorr DIC it is a larger multiple. Since the phase-based optical flow method uses linear filters for the bulk of the processing we expect that the time scales linearly with the number of pixels. The Ncorr DIC algorithm depends on solving an optimization problem and thus scales super linearly with the number of pixels. For very long and large videos the phase-based optical flow method has significant processing time advantages over the Ncorr DIC algorithm.

Table 4.3: Comparison of processing times for 100 frames of video, with different video processing methods and resulting processing time scaling factors

Video Size	301×201 pixels		576×576 pixels		
100 frames	60,501 total pixels		331,776 total pixels		
Method	Elapsed Time (s)	CPU Time (s)	Elapsed Time (s)	CPU Time (s)	Difference
G2 H2	1.31	3.69	6.65	20.3	5.08×
CSP	52.5	142.8	313.8	743.7	5.98×
Ncorr DIC	4,756	4,788	44,373	43,521	9.33×

4.4 Motion Compensation and Sensor Fusion

In Section 4.2 a parametric study of the intrinsic noise floor of the camera was carried out, however in real world measurements, another source of measurement noise is the motion of the camera itself. This is illustrated in the additional signal from the cooling fan vibrating the high speed camera in Figure 4-6. In this section we consider ways to compensate for camera motion using additional sensors to measure the camera motion, and a technique for sensor fusion.

4.4.1 Measurement of Camera Motion

Camera motion is likely to be a problem for outdoor, longer distance measurements where small motions of the camera, especially rotations can become large apparent displacements of objects in the video. The simplest way to measure the camera motion is to use a stationary reference object in the frame of the video and subtract that displacement signal from the signal from a region of interest in the video. This is used later in the thesis in several of the long-range measurements. Another possible way of accounting for camera motion is to measure the motion of the camera itself directly with other sensors.

To see what can be done with the additional information from sensors mounted to the camera, triaxial MEMS accelerometers were attached to the camera in four locations to measure its full 6 degree of freedom (DOF) motion: X (horizontal), Y, and Z (vertical) translation; and roll (Y-axis), pitch (X-axis), and yaw (Z-axis) rotations. The MEMS accelerometers sampled at a rate of 2000 Hz. The total of 12

106

accelerometer measurements are averaged to find the 6 DOF for the camera.

(a)   (b)
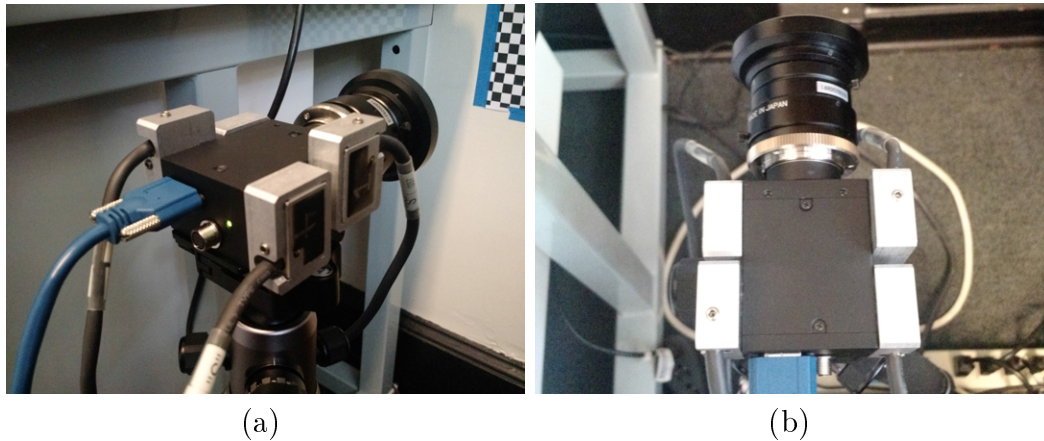
Figure 4-13: Setup for measurement of camera motion, showing wired MEMS sensors attached to camera body

An experimental setup was created in the laboratory to measure camera motions, shown in Figure 4-14. A checkerboard target was used to ensure good displacement signals from many pixels in the frame of the camera. Measurements were made at 100 fps and at a resolution of 480×320. To induce camera motion, the tripod was manually lightly shaken.

The simplest way to compensate for camera motion would be to directly subtract the measurement from the accelerometers after being double integrated from the camera measured displacement. The measured horizontal motion signal from the recorded video and the corresponding accelerometer measurement for Y translation are shown in Figure 4-15(a). A subtraction of the double integrated accelerometer measurement is shown in Figure 4-15(b) which shows less than satisfactory results. The problem is that for the time series signal, any small difference in phase can make the subtraction result in a signal with similar or higher magnitude than the original signal. The subtraction is much better in the frequency domain as the camera and accelerometer measured signals match up quite well as shown in Figure 4-16.

A discussion of the issues with a double integration of an accelerometer to displacement is discussed in references [103] and [86]. The basic issue is that at each time step there is some measurement noise that is amplified by the time integration,

107

(a)   (b)
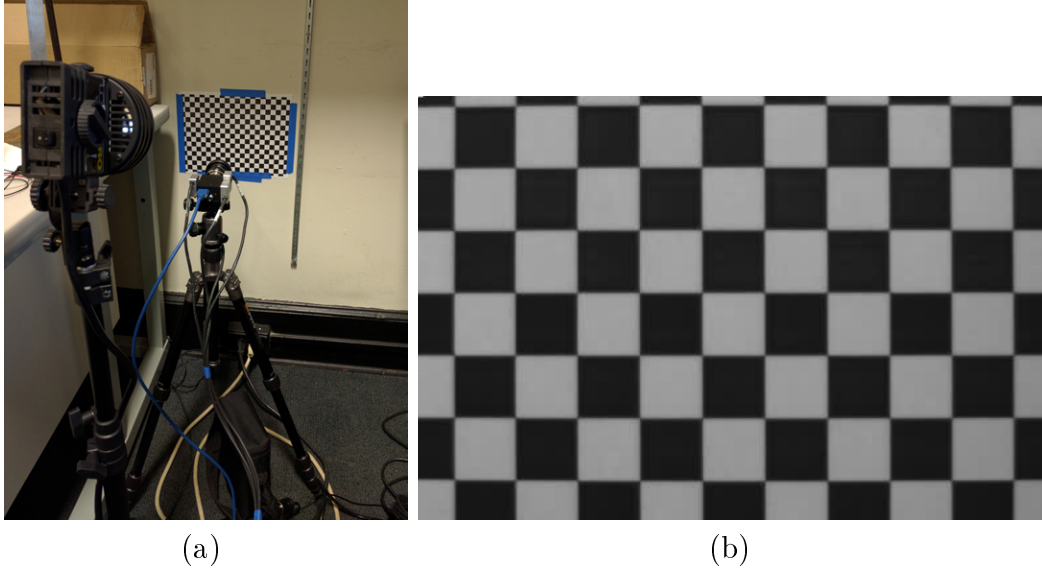
Figure 4-14: Laboratory experimental setup for camera motion test (a) showing setup with camera, light, and checkerboard target, and (b) screenshot from the recorded video

and especially magnified by the double integration. Generally, this issue occurs at low frequencies relative to the Nyquist frequency, which may fall directly in the frequency bands of interest for typical civil infrastructure.

(a)   (b)
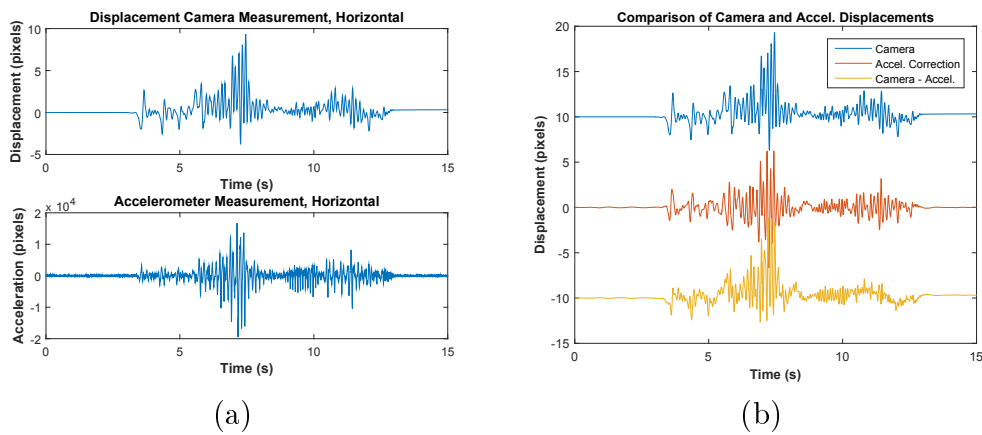
Figure 4-15: (a) Measured horizontal motion signals from the recorded video and accelerometer and (b) displacement signals extracted from the video and derived from the accelerometer measurements with subtraction as motion compensation result
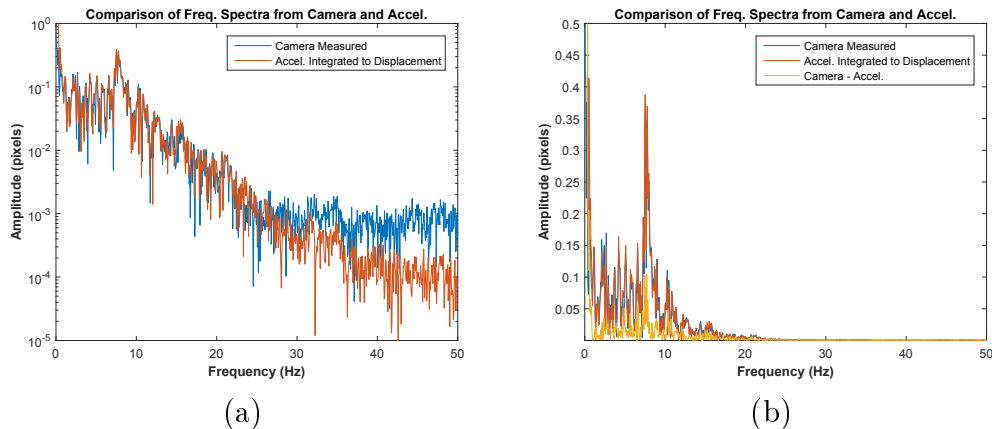
Figure 4-16: (a) Frequency response of measured horizontal motion signals and (b) subtraction as motion compensation in the frequency domain

4.4.2   Kalman Filter Methodology

Since the measurements do not line up well in the time domain due to the inaccuracy in the double integration from acceleration, one possible way to improve the accuracy is to also measure the displacement information of the camera. Because of inherently noisy measurements, accelerometer and camera displacement measurements will not be perfectly matched. A sensor fusion methodology reconciles the two disparate data sources and also accounts for different sampling rates. We can use a Kalman Filter multi-rate methodology for sensor fusion to obtain a better estimate of the camera motion from both an image-space measurement of the camera motion and accelerometer measurement [86].

Kalman filter formulation

A summary of the multi-rate Kalman filter formulation presented in [86] is given here. The state-space formulation of the camera and accelerometer measurements as a discrete process can be modeled as follows:

$$\begin{bmatrix} \dot{x}(k+1) \\ \ddot{x}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & t_a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ \dot{x}(k) \end{bmatrix} + \begin{bmatrix} t_a^2/2 \\ t_a \end{bmatrix} \ddot{x}_m(k) + \begin{bmatrix} t_a^2/2 \\ t_a \end{bmatrix} \eta_a(k) \qquad (4.1)$$

109

$$z(k) = x_m(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ \dot{x}(k) \end{bmatrix} + \eta_d \tag{4.2}$$

(k) with x representing the system displacement, $\dot{x}$ the velocity, and $\ddot{x}$ the accelera-
tion, t_a the time interval for the accelerometer measurement, k as the time step, x_m
and $\ddot{x}_m$ the measured displacement and accelerations, and η_d and η_a the associated
measurement noise for the displacement and acceleration. These equations can be
written compactly (changing $\ddot{x}_m$ to u(k) for the canonical notation) as:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + w(k) \tag{4.3}$$

$$z(k) = \mathbf{H}\mathbf{x}(k) + v(k). \tag{4.4}$$

The covariance matrices of the noise sequences are:

$$\mathbf{Q} = \begin{bmatrix} qt_a^3/3 & qt_a^2/2 \\ qt_a^2/2 & qt_a \end{bmatrix} \tag{4.5}$$

$$\mathbf{R} = \frac{r}{t_a} \tag{4.6}$$

where q is a covariance for the acceleration measurement noise, and r is a covariance
for the displacement measurement noise.

The Kalman filter is an algorithm that produces estimates of unknown variables
of a system from a series of measurements, given some input characteristics about
the process variance and the measurement noise. From the state space formulation
of the measurement, the Kalman filter algorithm is as follows with the time update
step:

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{A}\hat{\mathbf{x}}(k|k) + \mathbf{B}u(k) \tag{4.7}$$

$$\mathbf{P}(k+1|k) = \mathbf{A}\mathbf{P}(k|k)\mathbf{A}^T + \mathbf{Q} \tag{4.8}$$

and the measurement update step:

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)[z(k+1) - \mathbf{H}\hat{\mathbf{x}}(k+1|k)] \tag{4.9}$$

$$\mathbf{P}(k+1|k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}]\mathbf{P}(k+1|k) \tag{4.10}$$

and the Kalman gain $\mathbf{K}(k+1)$ is given by

$$\mathbf{K}(k+1) = \mathbf{P}(k+1|k)\mathbf{H}^T[\mathbf{H}\mathbf{P}(k+1|k)\mathbf{H}^T + \mathbf{R}]^{-1} \tag{4.11}$$

To fuse displacement and acceleration measurements, made at different sampling rates, a multi-rate process can be used [86]. We assume that the displacement sampling interval t_d is an integer multiple M of the acceleration sampling interval t_a, where $t_d/t_a = M$. Only the time update is performed during the accelerometer measurement time steps, and then every displacement measurement interval t_d the full Kalman measurement update is performed.

Toy example

For a toy example, we consider a sinusoidal displacement signal with a frequency sweep, and an additional linear trend. This type of signal would be extremely difficult to perform a double integration on from an acceleration measurement to obtain displacement. We define the displacement signal analytically as:

$$x(t) = 0.2 * t + sin((\pi + 1.5t) * t) \tag{4.12}$$

and we can analytically differentiate to obtain the velocity and acceleration as:

$$\dot{x}(t) = 0.2 + (\pi + 3t) * cos((\pi + 1.5t) * t) \tag{4.13}$$

$$\ddot{x}(t) = 3 * cos((\pi + 1.5t) * t) - (\pi + 3t)^2 * sin((\pi + 1.5t) * t) \tag{4.14}$$

To simulate measurements of this with an accelerometer and camera system we use an acceleration measurement sampling rate of 2000 Hz with RMS Gaussian noise of 10 percent, and a displacement sampling rate of 100 Hz with RMS Gaussian noise of 30 percent. The true analytical signal and the noisy measurements are shown in Figure 4-17, for displacement, velocity, and acceleration.
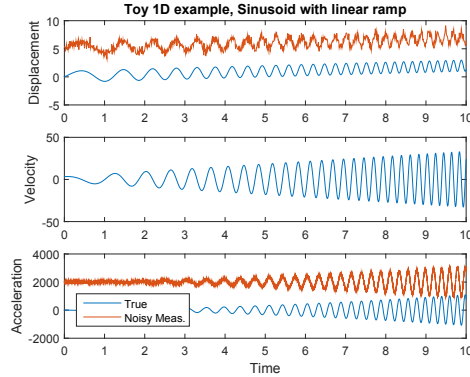
Figure 4-17: Toy example of a ramped sinusoid with an accelerometer sample rate of 2000 Hz and a displacement measurement rate of 100 Hz, showing comparison of true and noisy measurements

We input these noisy measurements into the Kalman filter and process the data for 10 seconds of the time history. Other parameters we use in the Kalman filter are the noise covariances, q and r determined from the 10 percent and 30 percent noise values for the acceleration measurement and Gaussian measurement. The initial value for **P** the error covariance matrix is 50 for both the displacement and measurement. The results from the Kalman filter are shown in Figure 4-18(a) comparing the true analytical displacement and velocity signals with the Kalman filtered signals. In Figure 4-18(b) we plot a comparison of the true displacement, the Kalman filtered displacement, and the displacement signal double integrated straight from the accelerometer measurement. Even with a relatively noisy camera measurement with 30 percent noise, the Kalman filter estimate is much better than the displacement integrated from the accelerometer.

Another toy example where the displacement measurement sampling rate is far lower, at only 10 Hz, with other parameters the same, is shown in Appendix C.4. This represents a situation in which the displacement measurement data might come from a camera with a much larger resolution and lower frame rate or from GPS.

(a)   (b)
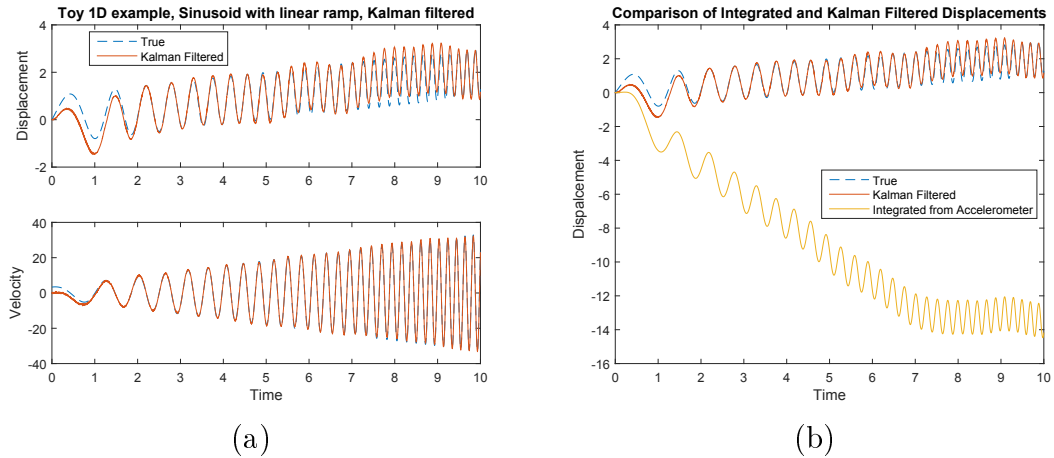
Figure 4-18: Results of the Kalman filter for the toy example of the ramped sinusoid showing (a) comparison between true and Kalman filtered measurements of displacement and velocity and (b) comparison between the Kalman filtered displacement result with integration from the accelerometer measurement without any filtering

Application to Motion Compensation

The multi-rate Kalman filter is applied to the measurement data from Section 4.4.1 and the resulting displacement, velocity, and accelerations are shown in Figure 4-19. For the displacement the camera measurement, Kalman filter displacement output, and the double integrated accelerometer signal are shown. For the velocity, the derivative of the camera measurement, the Kalman filter velocity output, and the integrated accelerometer signal are shown. For the acceleration, the double derivative of the camera measurement, the derivative of the Kalman filter velocity output, and the accelerometer signal are shown. The Kalman filter has the effect of generating a set of consistent displacement and velocity outputs where the displacement signal is very close to the camera measured signal, while the derivative of the velocity output is close to the measured accelerometer signal.

The motion compensation result is shown in Figure 4-20 in both the time domain and the frequency domain, and it is much improved over the straight subtraction, especially in the time domain. In a real-world measurement, the displacement of an object in the video frame that is not of interest and assumed to be stationary can be used as a displacement input to the Kalman filter along with accelerometers

113
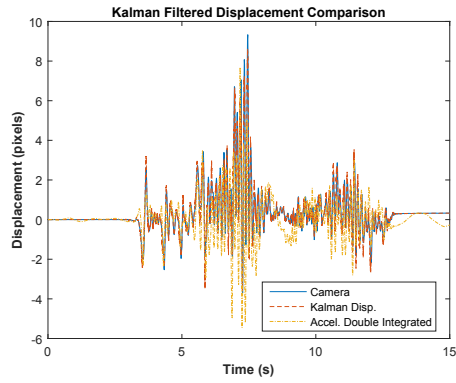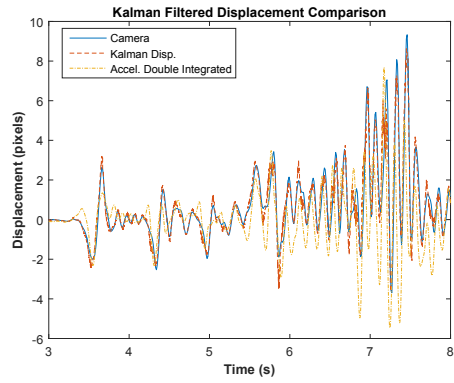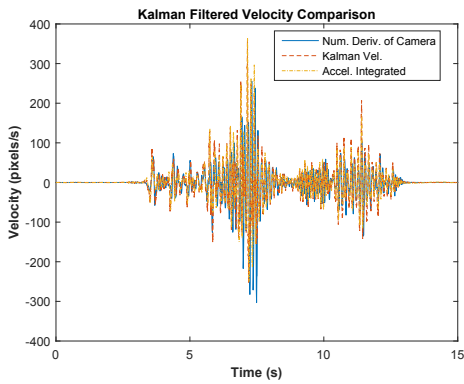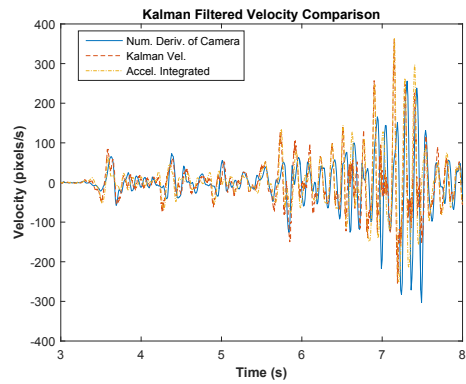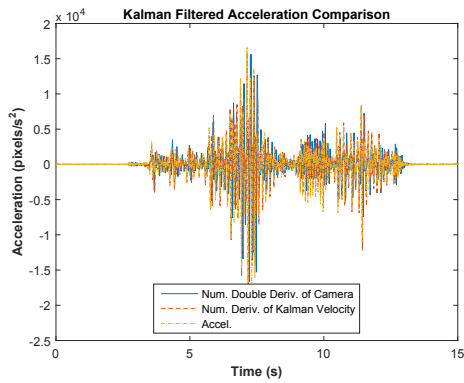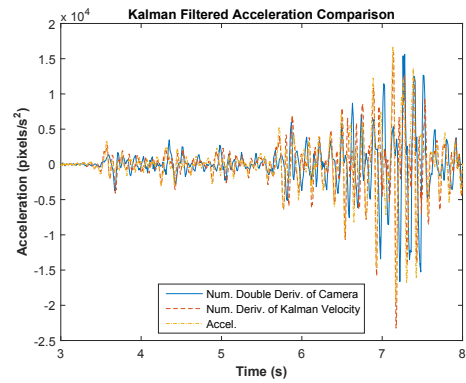
Figure 4-19: Applying the multi-rate Kalman filter to motion compensation showing the measured horizontal motion signals derived from the camera, laser vibrometer, and accelerometer measurements, with (a-b) displacements, (c-d) velocities, and (e-f) accelerations

114

measuring camera motion resulting in a dispalcement signal that can be subtracted from the extracted signal from the motion of an object of interest to correct for camera motion.

(a)
(b)
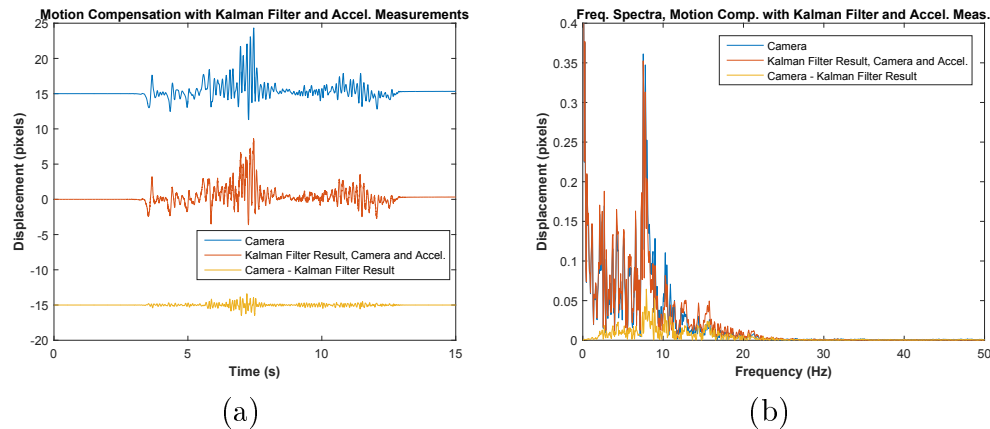
Figure 4-20: Motion compensation result using the Kalman filter derived horizontal displacement in the (a) time domain and (b) frequency domain, both showing much improved results

4.4.3   Sensor Fusion

We can also use the Kalman Filter multi-rate methodology for sensor fusion of accelerometer or other sensors in the frame of the video. It is possible that in some real-world applications, accelerometers or strain gauges might be placed on the structure of interest in a couple of locations, while a camera is used to measure more of the structure at once and give displacements with lower resolution but with more area coverage.

The experimental setup used is similar to the setup in Section 4.2 as shown in Figure 4-21. The measurement of the accelerometer in the frame of the video and the average displacement signal from the pixels corresponding to that accelerometer are used as inputs to the Kalman filter. The motion is also measured by a laser vibrometer for a comparison with both the integrated measurement from the accelerometer, and the result from the Kalman filter.

115
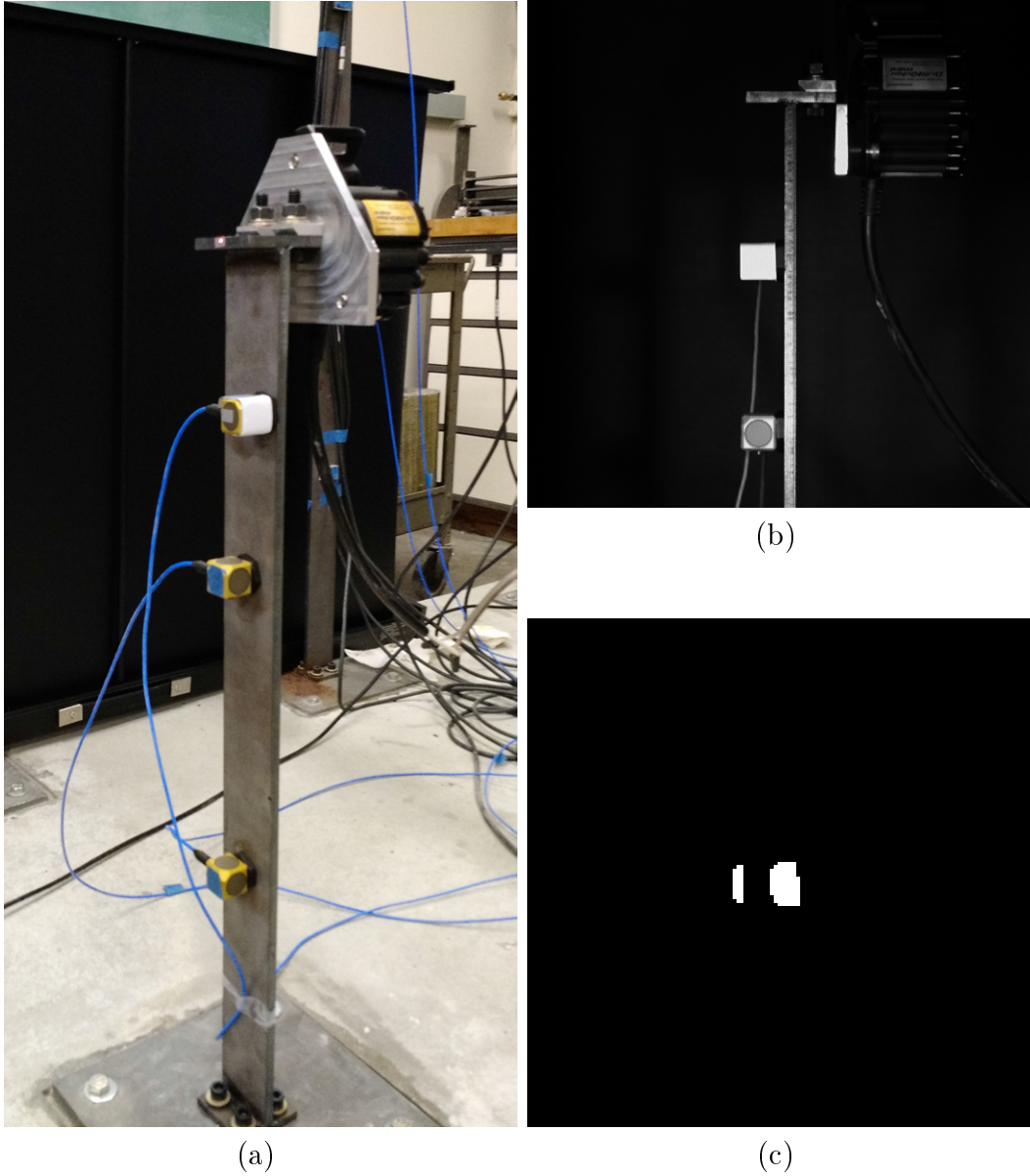
Figure 4-21: Experimental setup for sensor fusion of a single measured accelerometer signal in the frame of a video measurement: (a) a picture of the experimental setup, (b) a frame from the camera video, and (c) the mask showing pixels on the accelerometer with extracted displacements

The results of integration and the multi-rate Kalman filter look quite similar, as shown in Figure 4-22. We can quantitatively compare the results from the Kalman filter and the integration from the accelerometer with the measured laser vibrometer signal by calculating the correlation coefficient. Between the measured laser vibrometer signal and the Kalman filter derived velocity signal, the correlation coefficient is

99.01% while it is slightly less at 98.5% when compared with the signal derived from the accelerometer integrated signal. There is a slight improvement with the Kalman filter methodology when compared to integration, in addition to the fusing of different sampling rate data.
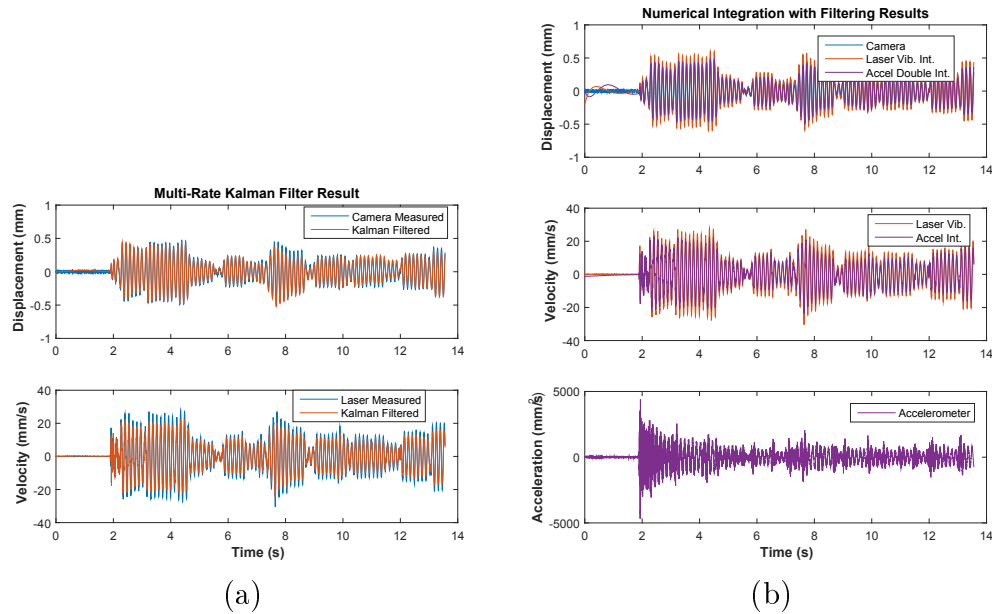
Figure 4-22: Result of data fusion using the multi-rate Kalman filter with (a) comparison between the measured and Kalman filter derived displacements and (b) comparison between measured displacements, velocity, and acceleration from integration of different measurements

4.5   Noise Floor Discussion

From the previous discussion of both the parametric study of the camera noise floor and the effects of camera motion, we can estimate the total noise floor for a typical measurement made outdoors, at different distances.

We can calculate the total noise TN in units of pixels per rt Hz by adding up the camera noise CN in units of pixels per rt Hz to the contribution from camera motion which would be the translational motion TM divided by the pixel size at distance $PSAD$ and the distance d times the rotational motion RM divided by the

pixel size at distance $PSAD$, using a small angle approximation. This is summarized in equation 4.15. Dividing by the $PSAD$ we can obtain the total noise in units of millimeters measured for an object of interest at a certain distance, shown in equation 4.16.

$$TN_{pixel} = CN + \frac{TM}{PSAD} + \frac{d * RM}{PSAD} \tag{4.15}$$

$$TN_{mm} = CN * PSAD + TM + d * RM \tag{4.16}$$

The pixel size at distance, in units of millimeters per pixel can be calculated given the pixel size on the sensor for a camera SPS, the focal length of the lens LFL, and the distance to the object of interest d, using the lens equation and solving for the pixel size as in equation 4.17.

$$PSAD = \frac{SPS * d}{LFL} \tag{4.17}$$

Numbers to input to the equations come from the noise parametric study in Section 4.2 and a later outdoor field measurement in Section 7.2. Camera noise is assumed to be 10^{-4} pixels per rt Hz, translational motion of the camera 10^{-4} mm per rt Hz, and rotational motion of the camera of 5×10^{-5} radians per rt Hz. We can generate plots for the total noise floor including camera motion, assuming a lens focal length of 24mm and a camera pixel size of 5.86 μm, shown in Figure 4-23, in units of both pixels per rt Hz in the video and millimeters per rt Hz physically.

There are two main regimes for the total camera noise that depend on where the line for the noise contribution from translation motion crosses the one for either camera noise or rotation motion, whichever is larger. In the close regime, translational motions of the camera are the dominane source of noise, but in the far regime, either the camera noise or rotational motion will dominate. This is because of the effect of the distance of the object on the effect of the rotational motion, and the camera noise which is effectively controlled by the pixel size at distance.

For long-distance measurements where either the rotational motion or the camera

118
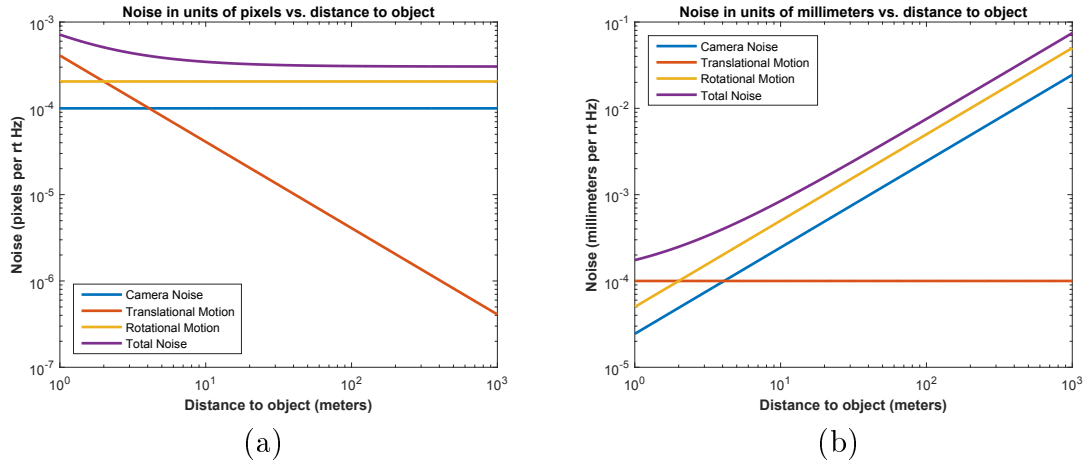
Figure 4-23: Total measurement noise floor plots vs. distance assuming intrinsic camera measurement noise of 10^{-4} pixels per rt Hz, translation motion of 10^{-4} mm per rt Hz, and rotation motion of 5×10^{-5} radians per rt Hz, with a lens focal length of 24mm, in units of (a) pixels and (b) millimeters at the object

noise dominates, we hope to remove the rotational motion actively using the accelerometers on the camera itself and/or a motion compensating gimbal, or passively by using an extremely stable tripod.

We can factor in the effect of the edge strength or local contrast by using the fitted line found in section 4.2, where the noise floor has a $x^{-2/3}$ dependence on the normalized local contrast. This is plotted on a surface shown in Figure 4-24 and as a function of distance.

We can also account for changes in the lens focal length which controls the zoom, and the pixel size at the object being measured. This is shown in Figure 4-25. As the lens focal length increases, the pixel size at distance gets smaller, effectively scaling the contribution from camera noise relative to the contribution from rotational motion. With a large enough focal length lens, the pixel size is small enough to lower the camera noise to the point such that the main contribution is from rotational motion, seen as the flattening of the curves as the focal lengths approach 100mm in the plot.
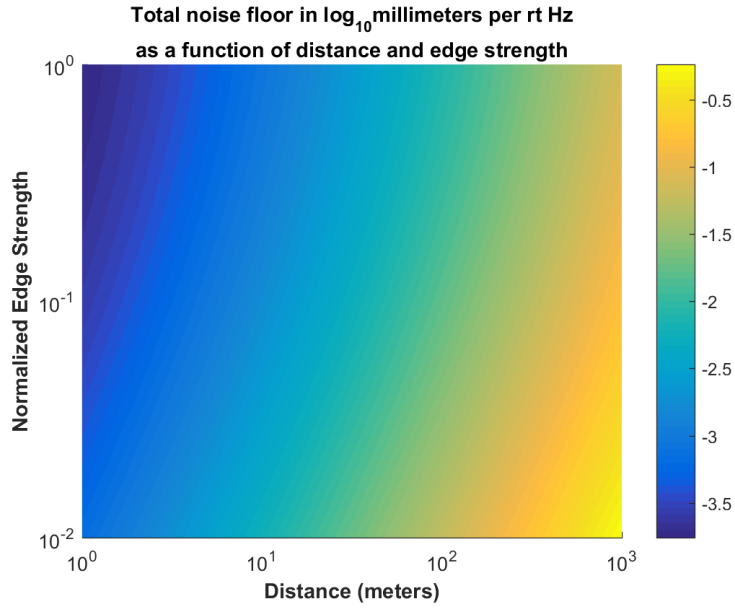
Figure 4-24: Total measurement noise floor in $\log_{10}$ millimeters per rt Hz as a function of distance to the object and normalized edge strength (local contrast)
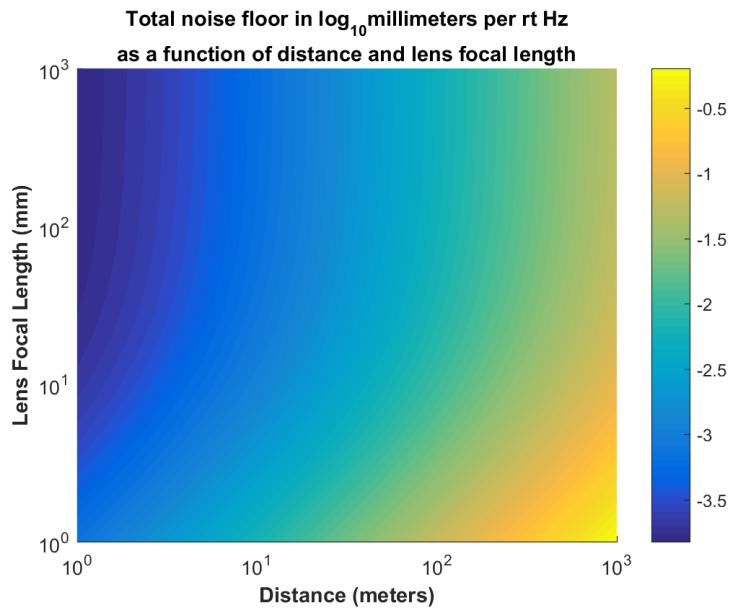
Figure 4-25: Total measurement noise floor in $\log_{10}$ millimeters per rt Hz as a function of distance and lens focal length

4.6 Chapter Summary

This chapter presented a verification of the camera measurement methodology, characterization of the measurement noise floor, comparison of video processing algorithms, study on motion compensation and sensor fusion, and discussion of the noise floor under different parameters and camera motion.

The camera measurement methodology compared well to a laser vibrometer and accelerometer measurement. The most important parameter governing the noise floor of the camera is the edge strength or amount of local contrast, as the higher that is, the better the noise floor. The approximate correlation between the edge strength and the noise floor is $(edgestrength)^{-2/3}$. Frame rate and lens focal length were found to have no effect on the intrinsic noise floor controlling for edge strength or lighting conditions. Phase-based motion processing algorithms used in this thesis are more than an order of magnitude faster than DIC, the current standard for video processing to obtain displacement. Motion compensation can be done using a stationary reference object in the video, accelerometers mounted to the camera itself, or a Kalman filter methodology to fuse the two to obtain a better estimate of the camera motion. The same Kalman filter methodology can also be used for sensor fusion of other sensors in the frame of a recorded video to improve the accuracy of the measured motions. A discussion of the effects of camera motion, lighting, lens focal length, and distance on the noise floor of the camera measurements was done with plots showing the relations between the different parameters and impact on the noise floor of the measurement.

Chapter 5

Laboratory Experiments

Two simple structural configurations, a cantilever beam and the circular cross section of a pipe, were considered as representative structures for a wider range of real world structures. Measurements were made of their resonant frequencies and mode shapes, and studies on damaged configurations were also conducted. A quasi real-time implementation of the full processing workflow is demonstrated on lightweight cantilever beams. Additionally, high-speed video from a test of a granite specimen was studied to see if displacements and strains could be measured with our processing workflow, as a basis for materials testing applications.

5.1   Cantilever Beam

5.1.1   Quantitative ODS Extraction from Camera Video

A measurement of a cantilever beam was made in the laboratory to demonstrate the extraction of operational deflection shapes (ODS) from a video measurement, as presented in [26]. The experimental setup used to test the ODS identification methods involved a high-speed camera, cantilever beam, and sources of illumination as shown in Figure 5-1(a). The column was struck with an impact hammer on the bottom third to induce vibrations in as many different bending modes as possible. Video was recorded at 2000 frames per second with a full frame resolution of 1248×200, with
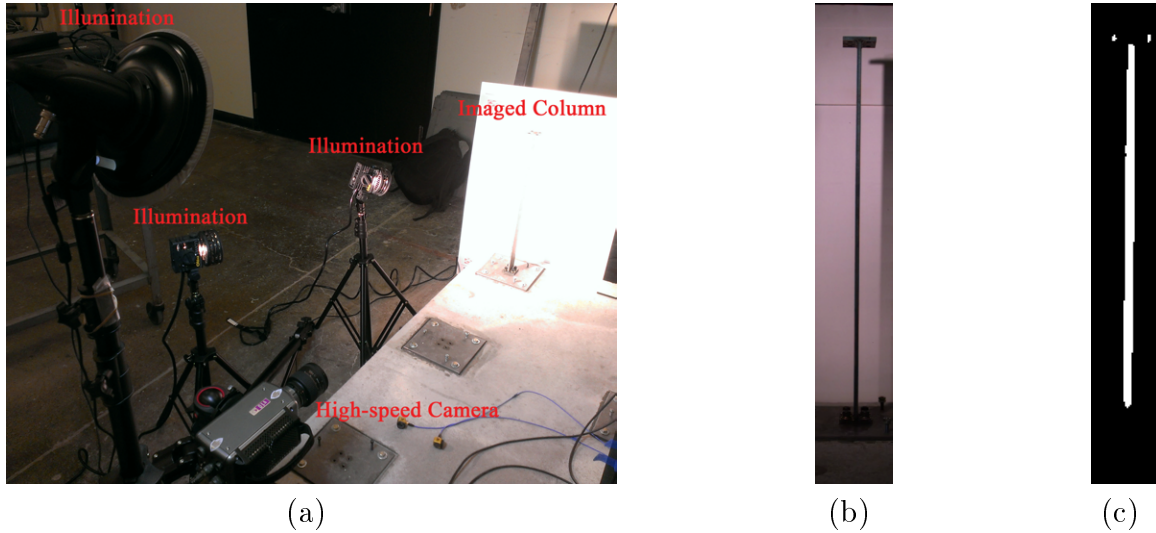
Figure 5-1: Setup for ODS identification of a cantilever beam: (a) a picture of the experimental setup, (b) a frame from the camera video, and (c) the mask showing pixels with extracted displacements

a frame shown in Figure 5-1(b) . The video was downsampled by a factor of two such that after the processing, the valid frame was 304×42 (12768) pixels and from that, edges were strong enough to obtain signals for 1234 pixels in the image. The mask shown in Figure 5-1(c) has in white the pixels where displacement signals were obtained.

From the displacements the ODS were identified using the two methods described in Section 3.3.2, peak picking from FFT spectra, and frequency domain decomposition (FDD). The frequency spectra obtained from the displacement signals are shown in Figure 5-2 with the one used for the peak picking technique in (a) and the singular values given in (b) and (c). The frequencies determined to be candidates for resonant modes are the same for both methods, which makes sense as the FDD method uses the same FFTs of the displacement signals to generate the singular values. The results of the ODS identification are shown in Figure 5-3, with 2D plots showing the normalized magnitude of the ODS at each identified frequency and the phase relationships between the signals.

The resulting identified ODS for both cases end up being the first five bending
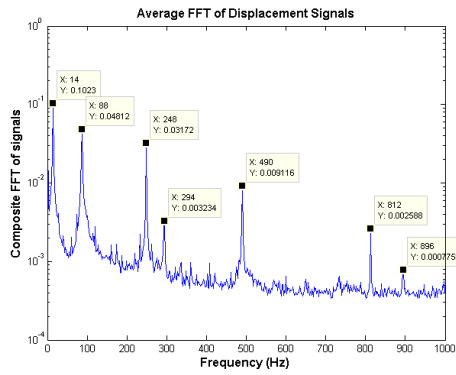
(a) Average of FFT of all displace-
ment signals
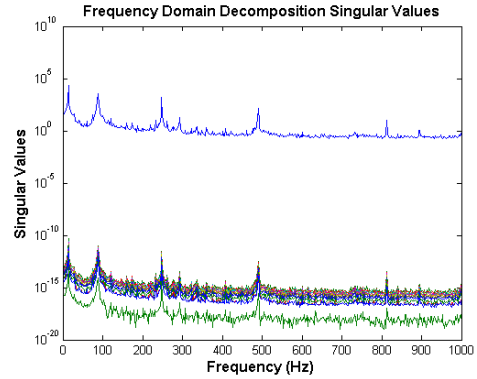
(b) Singular values from FDD
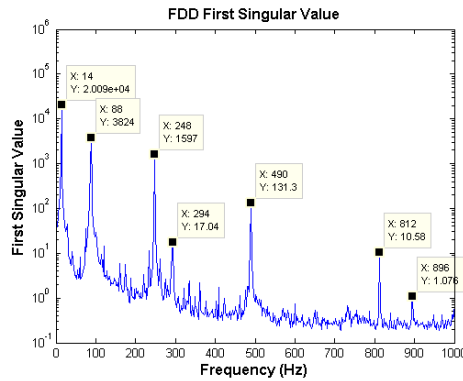
(c) First singular value from FDD

Figure 5-2: Frequency spectra used to determine potential operational resonant
modes, (a) average FFT of displacement signals and selected frequencies of inter-
est, (b) singular values as a result of FDD, and (c) first singular value and selected
frequencies of interest

modes at 14 Hz, 88 Hz, 248 Hz, 490 Hz, and 812 Hz and the first two torsional modes at 294 Hz and 896 Hz. The torsional modes are distinguished from the bending modes by considering the frequencies at which they appear, and by examining the motion magnified videos, shown in `https://youtu.be/sYY5jMRw2Zc`. For the torsional modes, the plate at the top of the beam can clearly be seen rotating in the motion magnified video.

Comparing the two methods the results are almost identical for the five bending modes. They are also similar for the torsional modes, albeit with noisier phase information in the case of the peak picking method. A more detailed comparison of the identified ODS is given in Figure 5-4 where the ODS magnitudes are averaged in horizontal rows to give a single value for each vertical sample which results in a more visually interpretable 1D mode shapes for the column. The ODS are shown in blue for the result from the FFT peak picking method and red for FDD. For the bending modes, the results are almost identical, but for the torsional modes the differences are more clear, especially with the second torsional mode where the peak picking ODS is much noisier than the FDD ODS. The modal assurance criterion (MAC) values comparing the identified ODS are given in Table 5.1 with the bending and torsional modes denoted by either B or T. The MAC values quantitatively match the qualitative observations previously described.
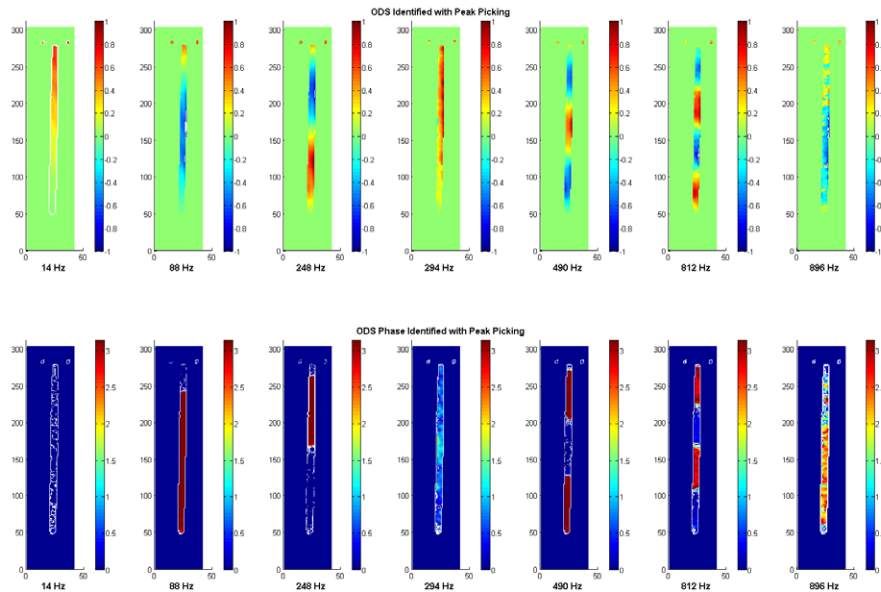
(a) Normalized ODS and phases identified with peak picking
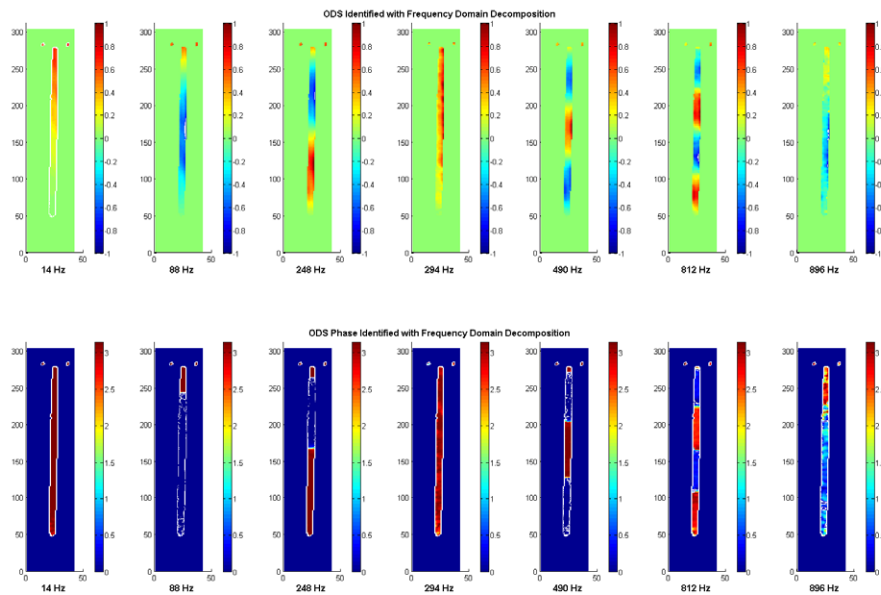
(b) Normalized ODS and phases identified with FDD

Figure 5-3: Identified ODS using the (a) peak picking method and (b) frequency domain decomposition
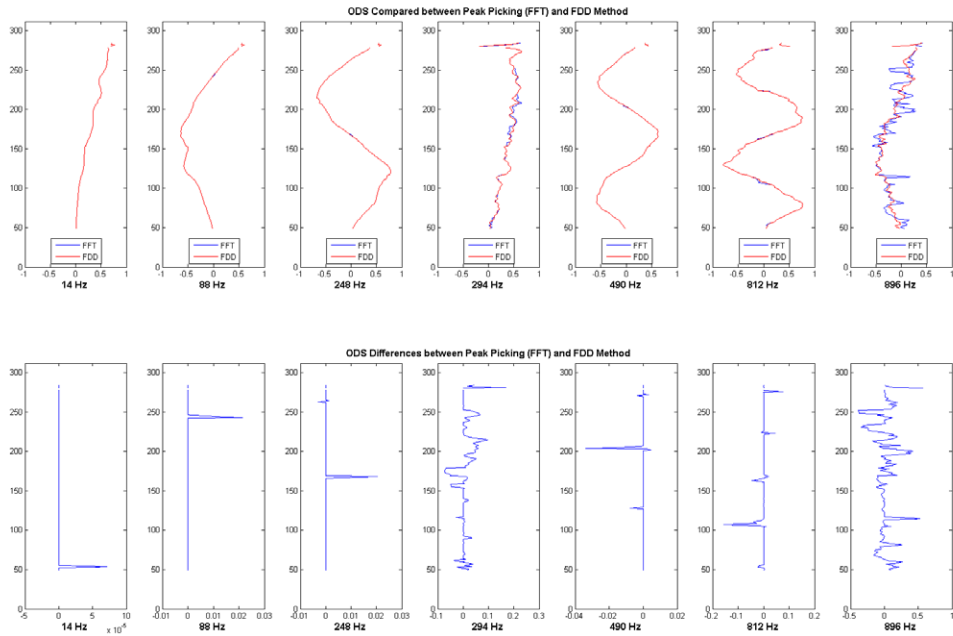
Figure 5-4: Comparison of video identified ODS

Table 5.1: MAC values comparing ODS identified with FDD with peak picking (PP) techniques for the various bending (B) and torsional (T) modes

ODS	PP B1	PP B2	PP B3	PP B4	PP B5	PP T1	PP T2
FDD B1	100.00%	7.75%	4.38%	1.17%	0.00%	84.50%	0.41%
FDD B2	7.78%	100.00%	1.56%	2.53%	1.19%	33.61%	63.21%
FDD B3	4.38%	1.56%	100.00%	0.19%	5.07%	1.38%	20.33%
FDD B4	1.15%	2.57%	0.19%	100.00%	0.09%	1.96%	1.03%
FDD B5	0.00%	1.30%	4.85%	0.12%	99.91%	0.18%	0.54%
FDD T1	83.34%	34.91%	1.07%	1.58%	0.09%	99.52%	13.48%
FDD T2	1.08%	78.09%	25.88%	3.34%	0.10%	14.37%	76.41%

5.1.2 Cantilever Beam Damage Study

Measurements were also made of damaged elements as shown in Figure 5-5(a). One damaged element has a small machined notch approximating a crack, and the other has a large reduction in the cross section. An additional form of damage is the loosening of bolts connecting the column to the base, thus changing the boundary

Table 5.2: Results from measurements of intact and damaged columns

Damage Scenario	Frequency (Hz)	Damping ratio	Element Mass (g)	Est. Equiv. Stiffness (kg/m)
Intact column (IC)	13.1624	0.2974	1690.1	5780
IC 1 bolt loose	13.1275	0.2863	1690.1	5749
IC 2 bolts loose	12.9434	0.2918	1690.1	5589
IC 3 bolts loose	12.8488	0.3039	1690.1	5508
Notched column (NC)	12.7101	0.2692	1676.5	5346
NC 1 bolt loose	12.6904	0.2141	1676.5	5329
NC 2 bolts loose	12.4392	0.2590	1676.5	5121
NC 3 bolts loose	12.336	0.2568	1676.5	5036
Reduced cross-section (RCS)	11.7384	0.2479	1168.2	3177
RCS 1 bolt loose	11.7352	0.2405	1168.2	3176
RCS 2 bolts loose	11.7036	0.2669	1168.2	3159
RCS 3 bolts loose	11.6869	0.3364	1168.2	3150

conditions. A video of the column is recorded as it freely vibrates after being hit by a hammer, with a screenshot shown in Figure 5-5(b). The displacement is extracted from a cropped region of the video near the top of the column, shown in Figure 5-5(c).

To determine the effects of damage, the resonant frequency and damping ratios were measured from the video measured displacements, an example of which is shown in Figure 5-6(a). To measure the resonant frequency the peak was found from the FFT of the displacement. To measure the damping ratio, an exponential function was fitted to the envelope of the decaying free vibration. The envelope was found by taking the Hilbert Transform of the absolute value of the displacement, as shown in Figure 5-6(b). The stiffness is estimated by assuming the cantilever beam acts as a single DOF system from the frequency and a factored mass for the beam.

The results are shown in Figure 5-7 and summarized in Table 5.2. Both forms of damage, the notch and the reduced cross section reduce the frequency of the beam when compared to the intact beam. The measured resonant frequency also decreases as bolts are being loosened, suggesting that the boundary condition becomes less clamped. In this situation, the damping ratio does not follow any obvious trend relative to the bolts loosened or damage in the column.
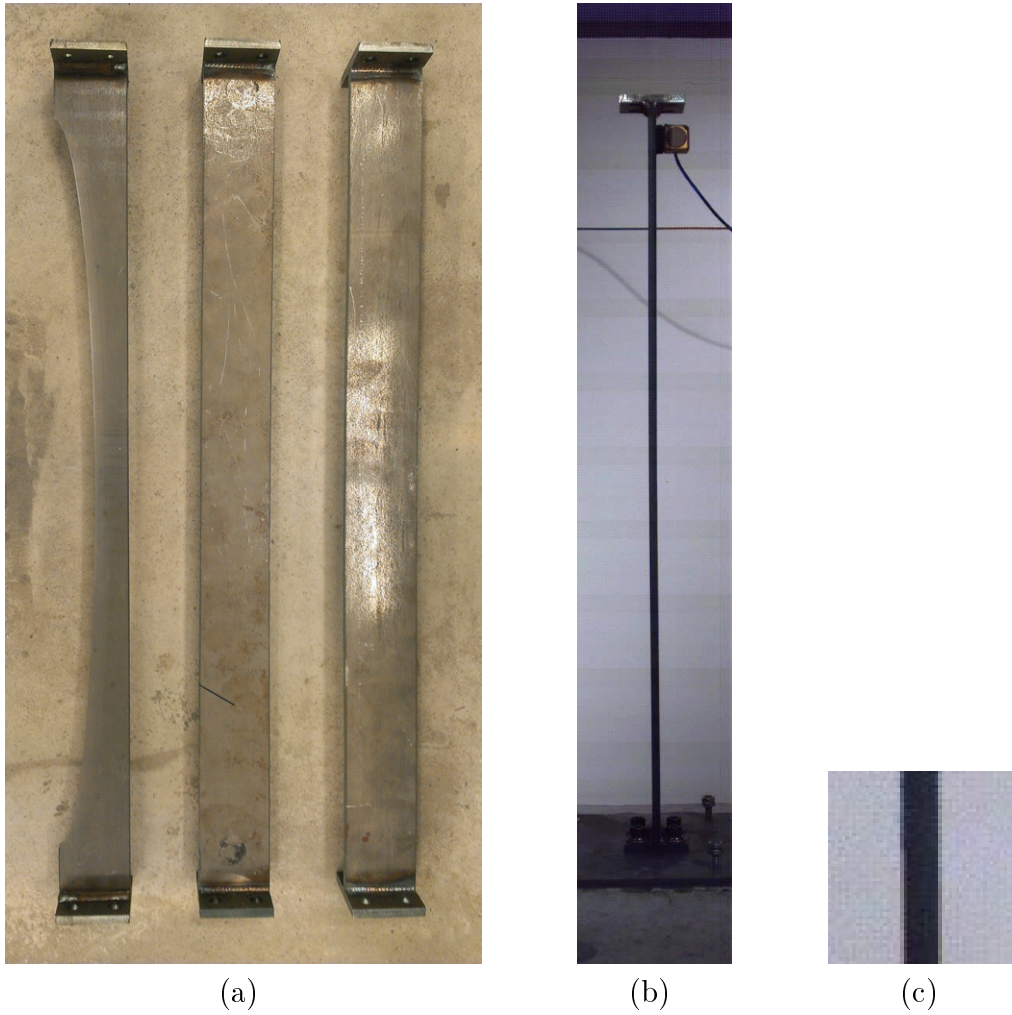
Figure 5-5: (a) Picture of cantilever beam specimens with reduced cross section on the left, machined crack in the center, and intact on the right, (b) screenshot from the recorded video, and (c) cropped frame of the video near the top of the beam from which a displacement signal was extracted
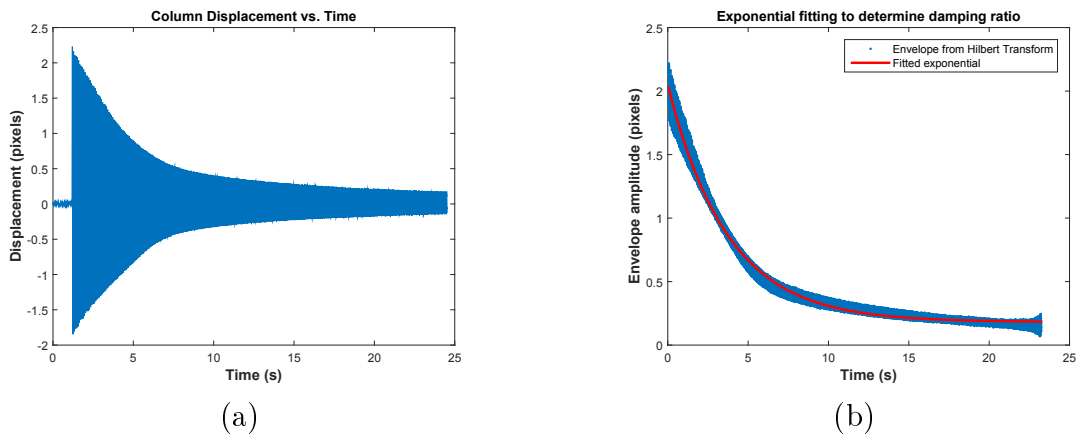
Figure 5-6: (a) Example extracted displacement time series from the cropped video and (b) Hilbert transformed signal to find the envelope with a fitted exponential to measure the damping ratio
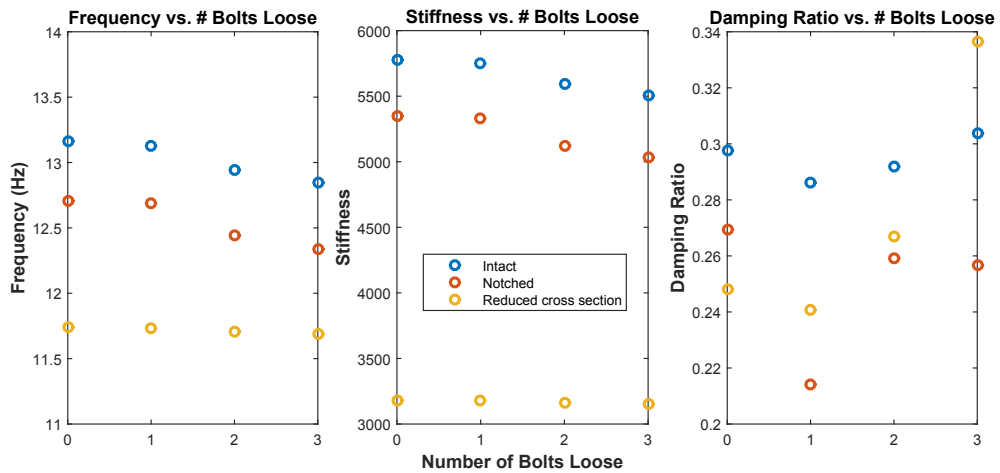
Figure 5-7: Results from measurements of intact and damaged columns with different base connection conditions (number of bolts loose) showing (left) frequency (center) stiffness and (right) damping ratio

5.2 Quasi Real-time Implementation

A desktop experiment was set up where the operational mode shapes and resonant frequencies of a polycarbonate beam were measured using a quasi real-time implementation of the video-camera measurement methodology where many steps are automated. The experimental setup is shown in Figure 5-8(a) where a 2 foot (0.61 m) long and quarter inch (6.35 mm) thick polycarbonate beam to be measured was rigidly attached to the desk via a vise, and the meaurement equipment consists of a video camera connected to a typical laptop computer that processes the recorded video. During the measurement the beam was struck by a hammer which vibrates the beam. The recorded video had a resolution of 236×1888 at 399.2 fps. 1000 frames of the video were recorded during the measurement over a duration of 2.5 seconds. Figure 5-8(b) shows a screenshot of the recorded video.

(a) (b) (c)
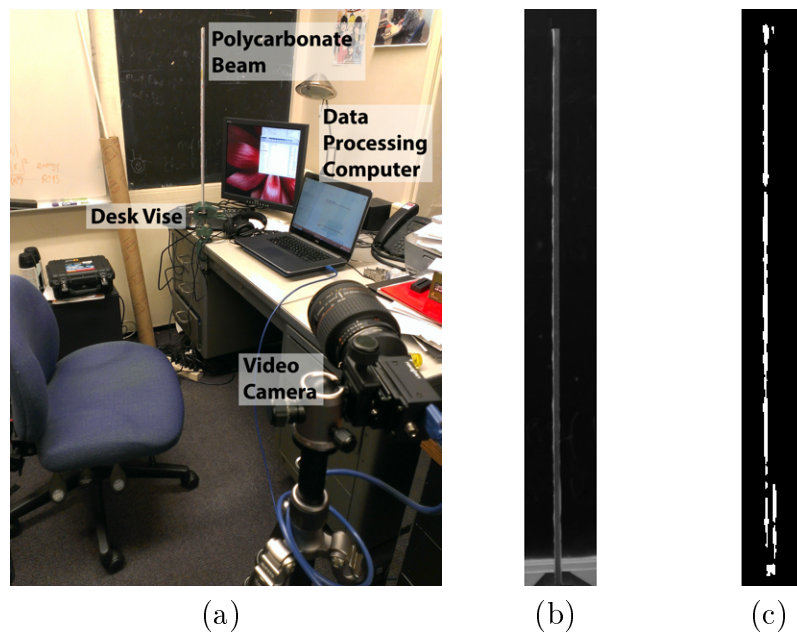
Figure 5-8: Near-real time measurement proof of concept (a) experimental setup, (b) screenshot of recorded video, and (c) pixel mask

Processing of the video took approximately 2 minutes starting from collecting the video to generating the plots showing the operational mode shapes for the automatically selected candidate resonant frequencies. Approximately another 5 minutes were

necessary to generate four motion magnified videos at the resonant frequencies selected by the user. Displacements were calculated for the pixels shown in white in Figure 5-8(c). A composite of all the displacement signals and their frequency spectra are shown in Figure 5-9(a) and 5-9(b) respectively, and the averaged displacement signal and frequency spectrum are shown in Figure 5-10. Four potential resonant frequencies were identified manually, 3.593 Hz, 23.55 Hz, 61.08 Hz, and 135.7 Hz, shown in Figure 5-10(b). Separately, five potential resonant frequencies were automatically identified, 3.593 Hz, 7.185 Hz, 19.96 Hz, 23.55 Hz, and 27.14 Hz. Some of the automatically picked resonant frequencies look like side bands of the first two true resonant frequencies, and thus may be non-physical.

The mode shapes corresponding to the manually and automatically picked resonant frequencies are shown in Figure 5-11 and Figure 5-12 respectively. At the top of each figure is the mode shape amplitude per pixel, the middle are the mode shape phase images, and the bottom shows the equivalent 1 dimensional representation of the mode shape for a column, obtained by averaging horizontal rows of the mode shape. Looking at the mode shapes of the manually picked resonant frequencies shows that they correspond to the first four bending modes of the beam.

(a)   (b)
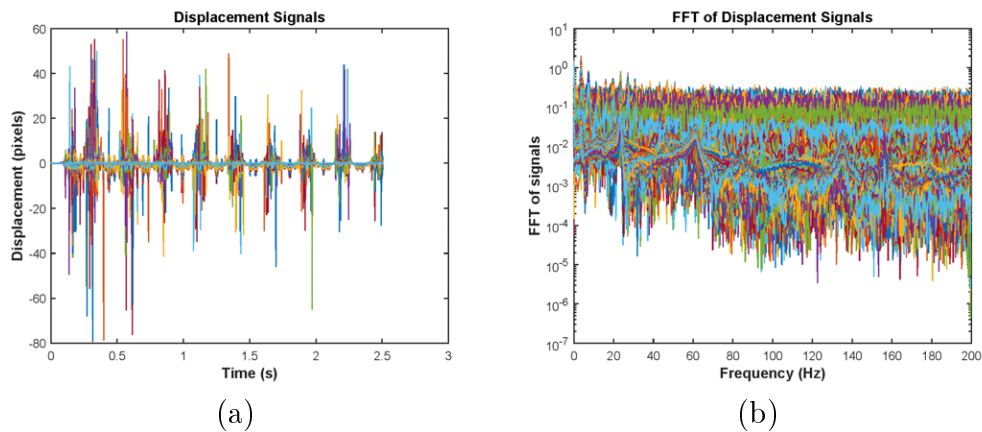
Figure 5-9: (a) Collected displacement time series signals and (b) FFTs for all pixels with measured displacements

133
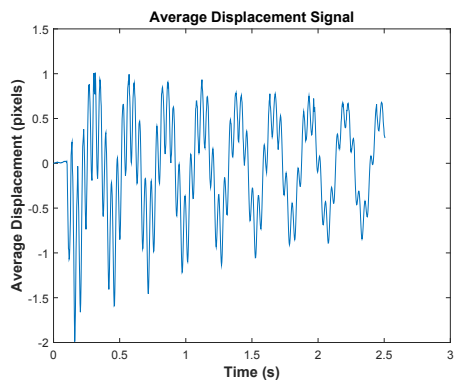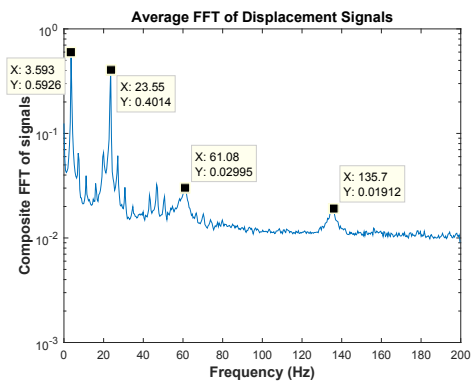
(a)   (b)

Figure 5-10: (a) Average displacement signals and (b) average frequency spectra used for picking suspected resonant frequencies
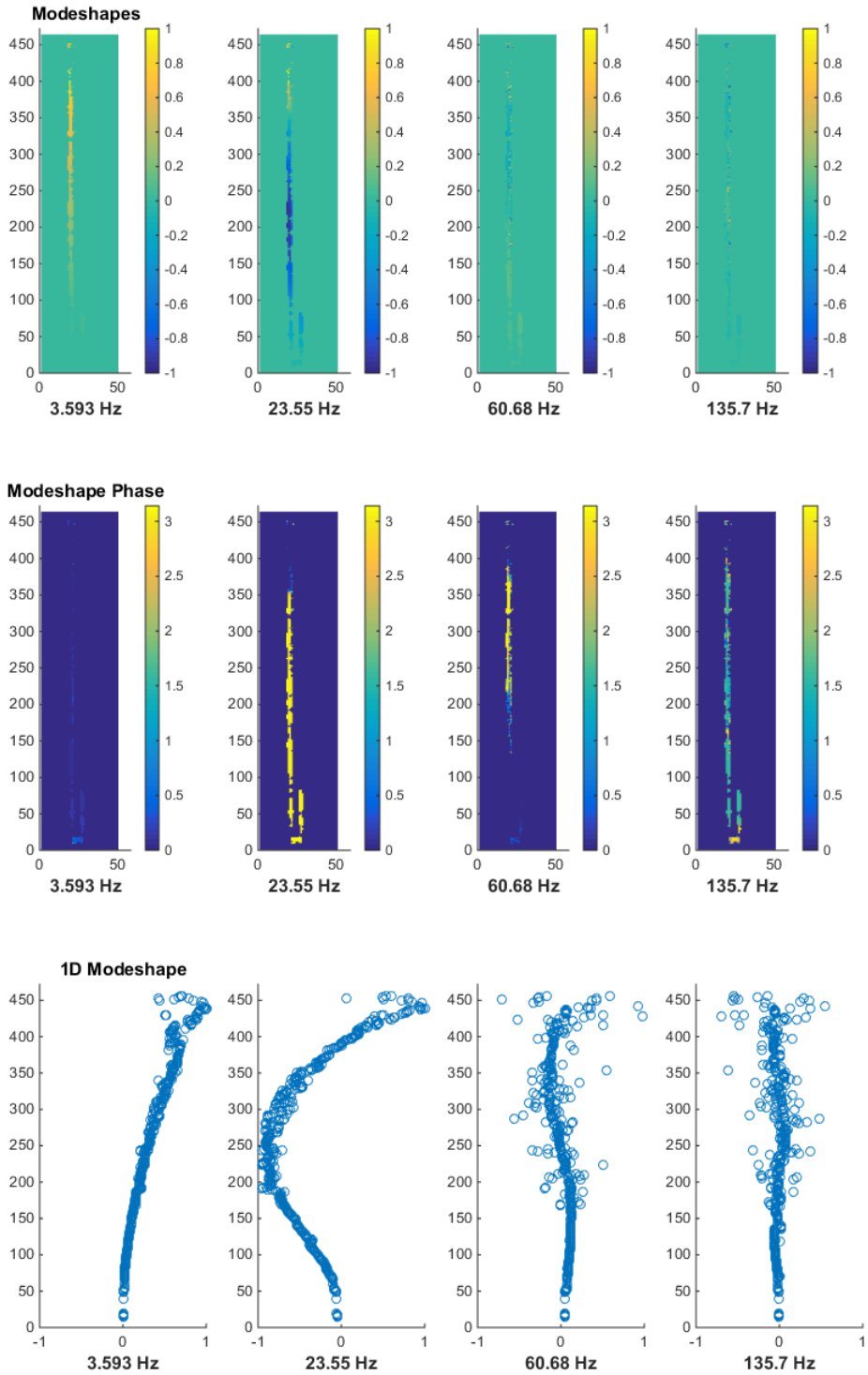
Figure 5-11: Operational mode shapes, phase relationships, and 1D mode shape representations for suspected resonant frequencies picked manually by informed user
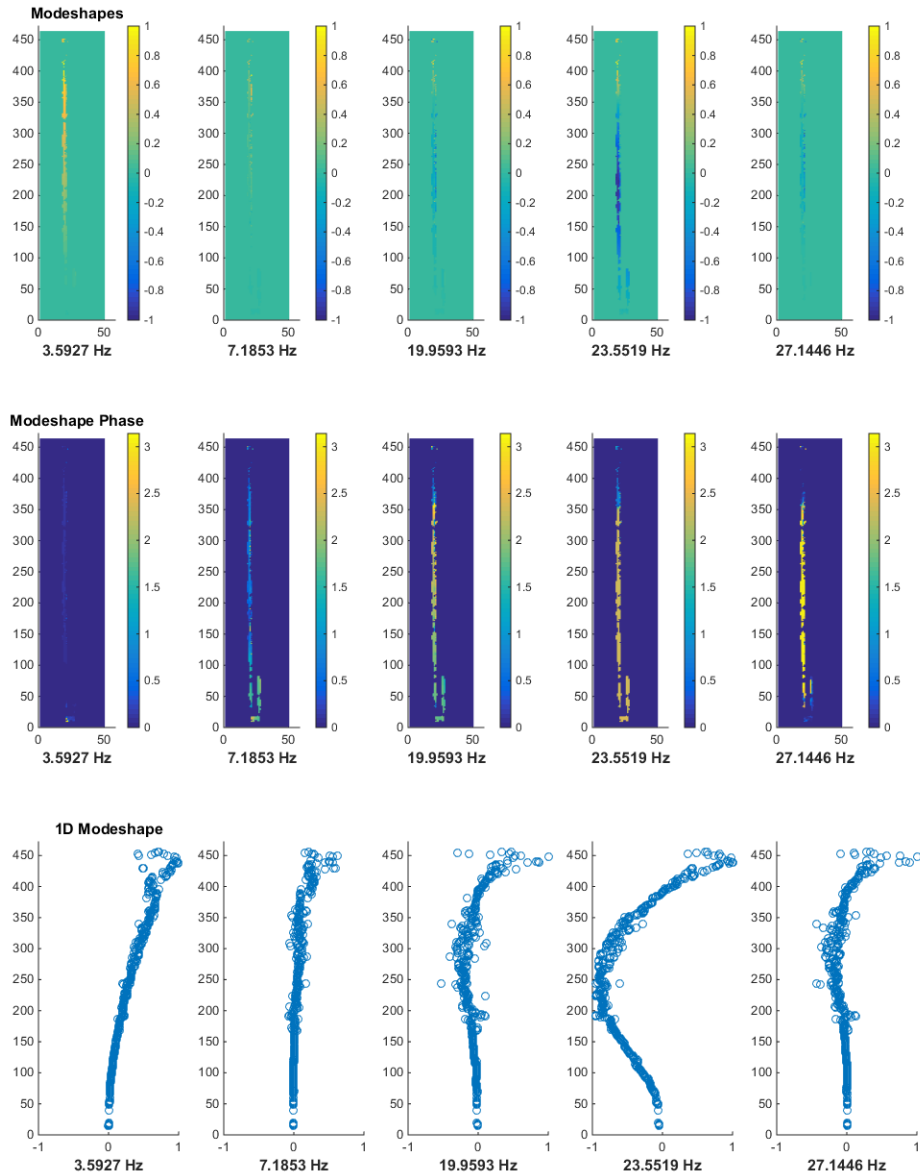
Figure 5-12: Operational mode shapes, phase relationships, and 1D mode shape representations for suspected resonant frequencies picked automatically

5.2.1 Use on 3D Printed Beams

Lightweight structures are a particularly relevant application for non-contact measurement methods because the attachment of traditional sensors such as accelerometers or strain gauges may alter the system dynamics by adding mass or stiffness. As a part of a separate experimental study, we made measurements of a series of lightweight cantilever beams made of different plastic materials to look for differences in the material properties, as described in [95], and presented at [10]. This was quantified by looking at the frequency of the first resonant modes. An example measurement of the Pink specimen, made from VeroMagneta RGD851, a 3d printed material, is presented here. The specimen has dimensions of 210mm × 20mm × 2mm (8.27in × 7.9in × 0.079in) [95]. The desktop experimental setup shown in Figure 5-13(a), and a screenshot of the recorded video shown in Figure 5-13(b). The input video had a resolution of 512 × 1200 pixels, 1000 frames recorded at 161.4 frames per second (fps). As the beam is clamped in the vise, the free span of the beam ends up being 200mm (7.9in). In the frame of the video, the beam is 200mm (7.9in) tall for a height of 790 pixels, thus the pixel size is 3.95mm per pixel (0.156in per pixel). The measurement procedure is the same as the quasi real-time implementation proof of concept measurement, and the video frame is downsampled by a factor of 4 in the processing. The beam is excited at the beginning of the video by a tap from a finger near the base.

<div align="center">(a)　　　　　　　　　　　　　　　(b)</div>
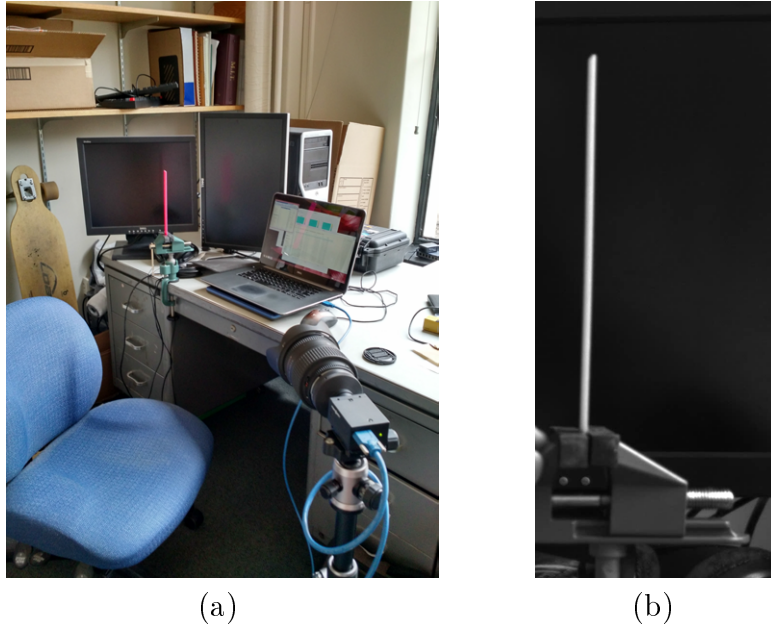
Figure 5-13: (a) Desktop setup for measuring the operational mode shapes of a 3D printed cantilever beam and (b) screenshot from recorded video, 512×1200 pixels

From the recorded video, the first two resonant modes of the beam were measured at 11.95 Hz and 78.94 Hz, as shown in Figure 5-14. The operational mode shapes correspond to the first two bending modes of the beam. The frequency spectrum measured from the beam is shown in Figure 5-15 in orange while measurements made with a contact accelerometer were also made, with the frequency spectra shown in blue and red. This corresponds to the placement of the accelerometer on the beam in two different locations, which shows a significant shift in the frequency of the first resonant mode, due to the fact that the accelerometer mass of 16 grams (0.56oz) is heavier as the beam itself at 10 grams (0.35oz). With only one accelerometer, mode shapes can't be measured without a roving accelerometer or impact hammer, and just by moving the accelerometer or placing one accelerometer the dynamics are changed. For complex lightweight structures where many sensors would be needed to measure the mode shapes, the camera provides much better information as it not only leaves the dynamics of the structure unchanged, but it also provides spatially dense measurements on the structure which can be used to identify the mode shapes.
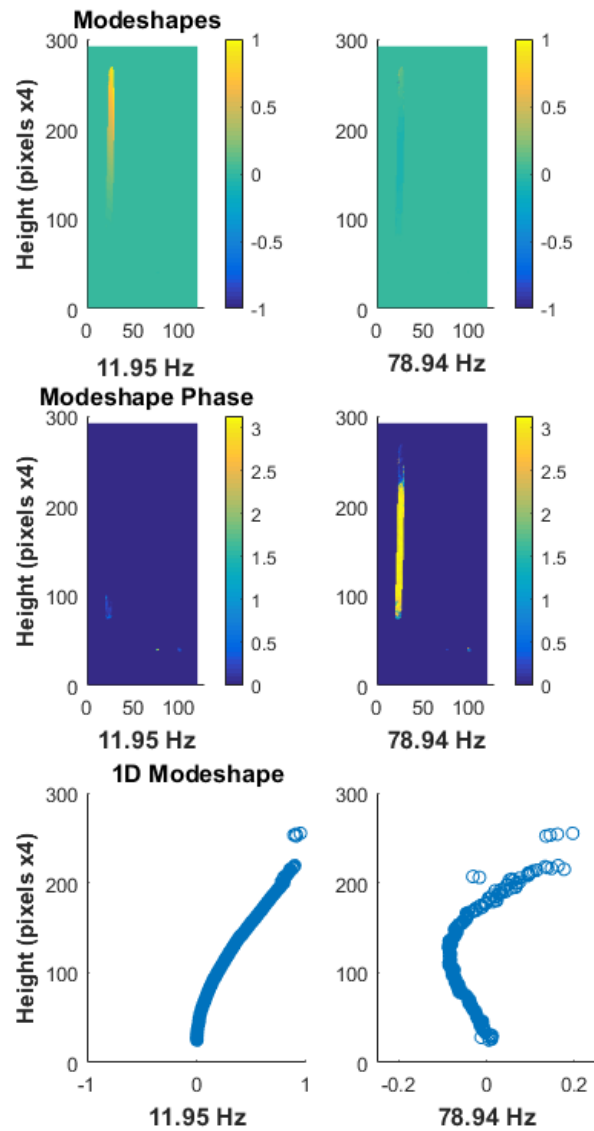
Figure 5-14: Operational mode shapes, phase relationships, and 1D mode shape representations for suspected resonant frequencies
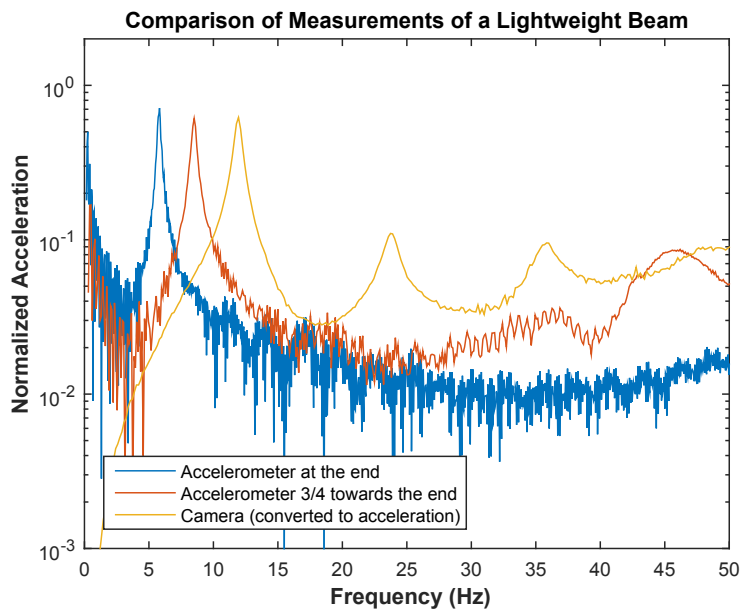
Figure 5-15: Comparison of measurements with an accelerometer in two different positions and a camera of the lightweight beam

5.3   Pipe Cross-section

5.3.1   Initial PVC Pipe Studies

An experiment similar to the instrumented cantilever beam test was performed on a pipe, as published in [23, 24]. A section of 4" schedule 40 PVC pipe was placed on the end of a laboratory bench and held down by hand as shown in Figure 5-16(a). The pipe was struck on the free end by an impact hammer as an excitation and the resulting vibrations of the pipe as viewed in the cross-section were recorded with a high-speed camera as shown in Figure 5-16(b). The camera recorded video at 20,000 fps, with a frame size of 192 by 192 pixels, and in the plane of the pipe the video frame was approximately 147 mm (5.8 in) wide and 147 mm (5.8 in) tall, with 1.306 pixels per millimeter (33.2 pixels per inch).

(a)   (b)
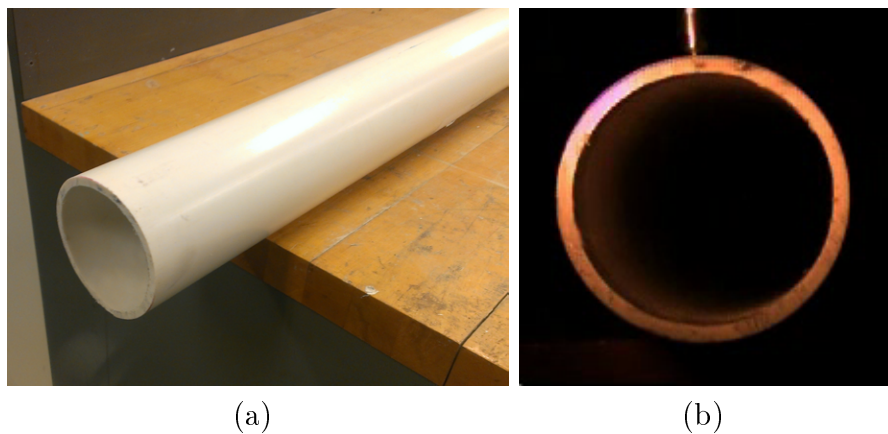
Figure 5-16: Pipe measurement: (a) experimental setup and (b) screenshot from video camera

5.3.2   Visualization of Mode Shapes with Motion Magnification

A displacement signal was extracted from the source video cropped to the bottom edge of the pipe, suggesting frequency bands for the motion magnification algorithm. Operational deflection shapes were extracted from the video using the phase-based

motion magnification algorithm at 480 Hz, 1200 Hz, and 2400 Hz [99]. The frequency bands used were 475 - 485 Hz, 1195 - 1205 Hz, and 2390 - 2410 Hz with amplification factors of 50, 200, and 400 respectively. Screenshots from the motion magnified video are shown in Figure 5-17. These vibrational modes visualized here are the first three circumferential-radial modes of a pipe and match the expected theoretical mode shapes, as described in [6]. The results from an edge detection routine to determine a quantified mode shape from the screenshots are shown in Figure 5-18. A series of screenshots from the motion magnified video in the frequency band of 1195 - 1205 Hz are shown in Figure 5-19.

(a)　　　　　　　　　(b)　　　　　　　　　(c)
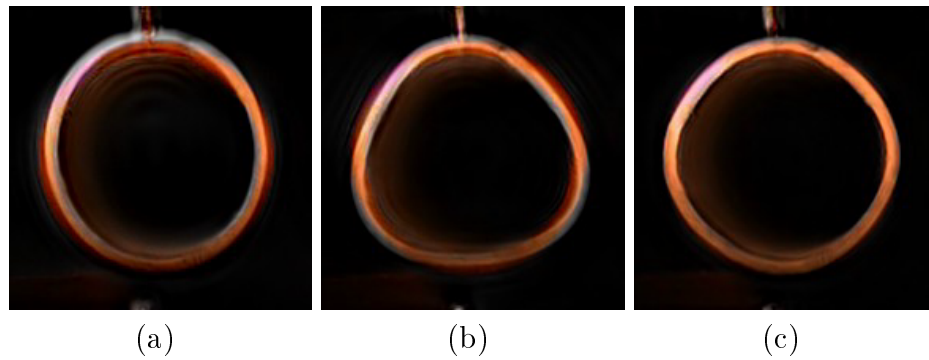
Figure 5-17: Screenshots from motion magnified video of pipe cross-section: (a) 480 Hz, (b) 1200 Hz, and (c) 2400 Hz

(a)　　　　　　　　　(b)　　　　　　　　　(c)
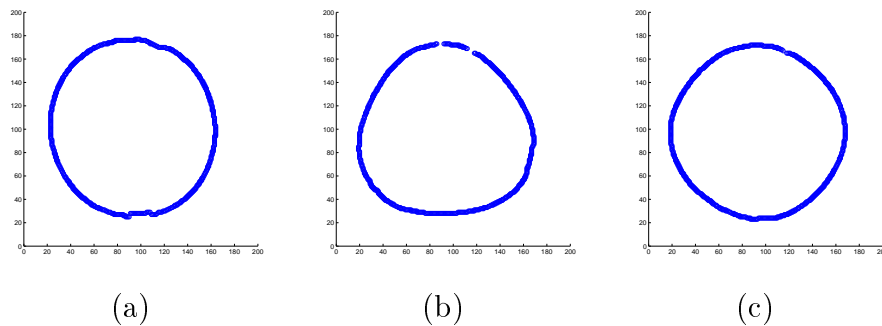
Figure 5-18: Operational deflection shapes derived from edge detection of the screenshots from motion magnified video of the pipe cross-section: (a) 480 Hz, (b) 1200 Hz, and (c) 2400 Hz
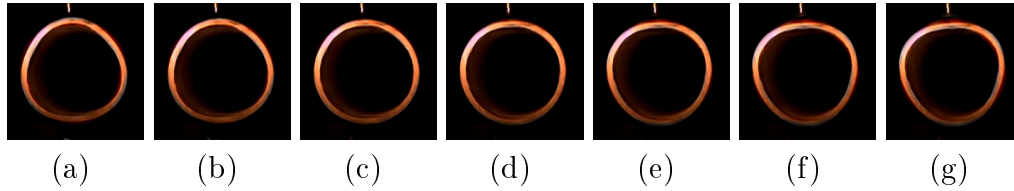
Figure 5-19: Series of frames from motion magnified video of the pipe in the 1195 - 1205 Hz band

5.3.3 Measurement with Non-flickering lights

The measurement of the 4" schedule 40 PVC pipe was repeated with lights that were battery powered so as not to induce any flickering in the video which may have caused the motions of the pipe to alias to multiples of the frequency of the 120 Hz of the previous AC mains powered lights. The experimental setup and screenshot from the video camera are shown in Figure 5-20; the measurement was made at 24,000 fps with a resolution of 192×192. Similarly the pipe was excited by a hammer hit from the top, and the pipe was held against the side of a laboratory bench.
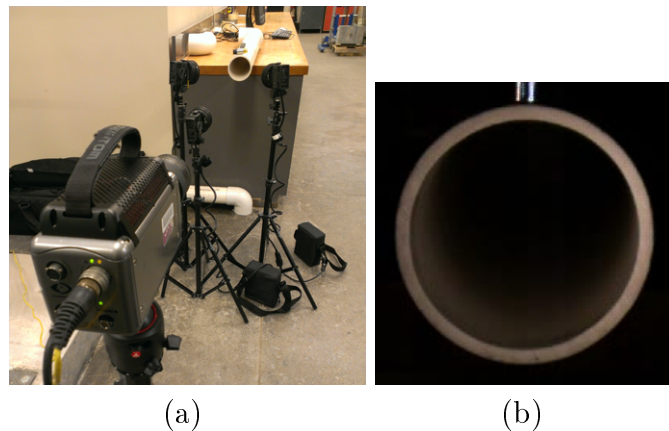
Figure 5-20: Pipe measurement with non flickering lights: (a) experimental setup and (b) screenshot from video camera

Figure 5-21 shows the frequency spectrum of displacement extracted from a cropped region at the bottom of the pipe. The first two resonant frequencies at 466 Hz and 1317 Hz can be seen, however higher frequencies are not visible in the spectrum. Using the known equations for the circumferential-radial resonant frequencies of an

infinite pipe [6], we can make reasonable estimates of the frequencies where the higher resonant modes should be. We then tried motion magnifying the video at these estimated frequencies and visually determined if the motions were coherent enough to represent a mode shape. The first five resonant modes of the pipe were found in the video at frequencies of approximately 466 Hz, 1317 Hz, 2525 Hz, 4100 Hz, and 6000 Hz. Screenshots from the motion magnified videos showing these mode shapes are shown in Figure 5-22.
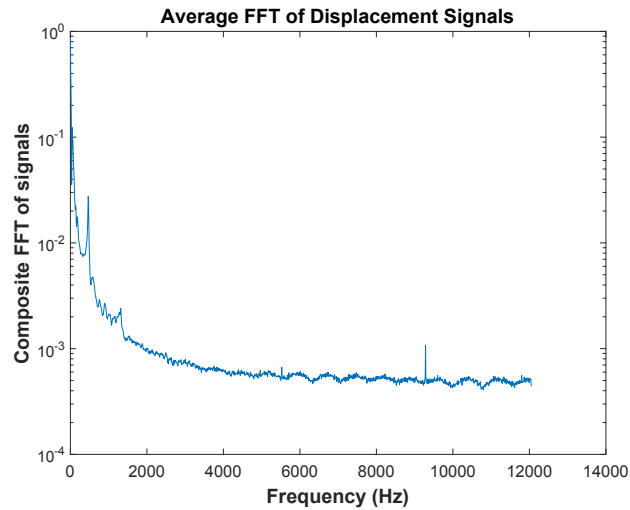
Figure 5-21: Frequency spectrum of vertical displacement extracted from cropped region at the bottom of the pipe
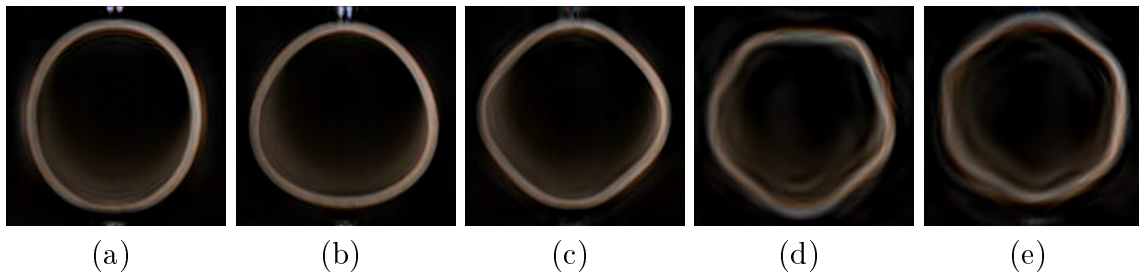
Figure 5-22: Screenshots from motion magnified video from improved measurement of pipe cross-section: (a) 460 - 472 Hz $\times 20$, (b) 1310 - 1325 Hz $\times 100$, (c) 2500 - 2550 Hz $\times 800$, (d) 4000 - 4200 Hz $\times 4000$, and (e) 5900 - 6100 Hz $\times 4000$

5.3.4   Corroded Pipe Experiment

An experimental study was conducted on short sections of steel pipes, damaged in the laboratory to simulate potential damage that longer sections of pipe may experience in the real world. We used the same measurement technique that we demonstrated on the previous study on the PVC pipes to measure the resonant frequencies and motion magnify the mode shapes. To create damage in the pipes we induced corrosion using an impressed current method. The experimental setup is shown in Figure 5-23 where the pipe is placed in a salt bath and connected to a direct current power supply on the positive terminal and a piece of copper is connected to the negative terminal and electrons are stripped from the steel forming rust. Two corroded specimens were produced, corroded for 2 days and 5 days. An intact 6" schedule 10 pipe is seen in Figure 5-24(a) and a corroded pipe is seen in 5-24(b).
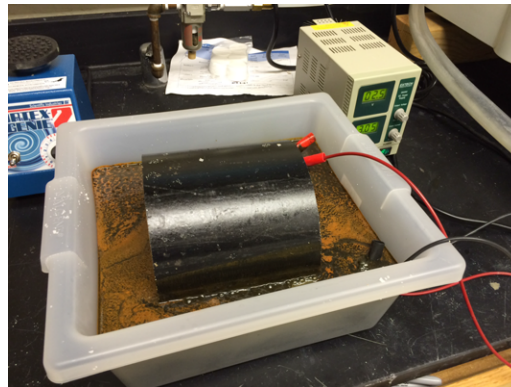
(a)

Figure 5-23: Impressed current corrosion setup for corroding the pipe in the laboratory to simulate damage

The experimental setup for taking video was the same as for the PVC pipes as shown in Figure 5-25. As before, the pipes were rested on the side of a laboratory bench and hit from above with a hammer. A difference in the setup is that a backdrop was added to ensure an evenly dark background. Measurements were made at 5,000 fps with a frame size of 576×512 pixels. Figure 5-26 shows screenshots from videos of the intact, 2 day corroded, and 5 day corroded pipes. Note the large difference in the surface of the pipe, with the intact pipe giving a nice even reflection as it is the

Figure 5-24: (a) An intact 6" Schedule 10 pipe, (b) a corroded 6" Schedule 10 pipe

even cut surface, the 2 day corroded pipe with mild corrosion around the rim, and the 5 day corroded pipe with some parts darker than others and thinner as well due to the mass loss from corrosion.

Figure 5-25: Experimental setup for measuring high-speed videos of the pipe, which rests on the side of a laboratory bench, showing high-speed camera, battery powered lights, pipe section, and backdrop

Frequency spectra were extracted from the three videos for motions in both the horizontal and vertical directions. They are plotted in Figure 5-28 with the intact

(a)   (b)   (c)
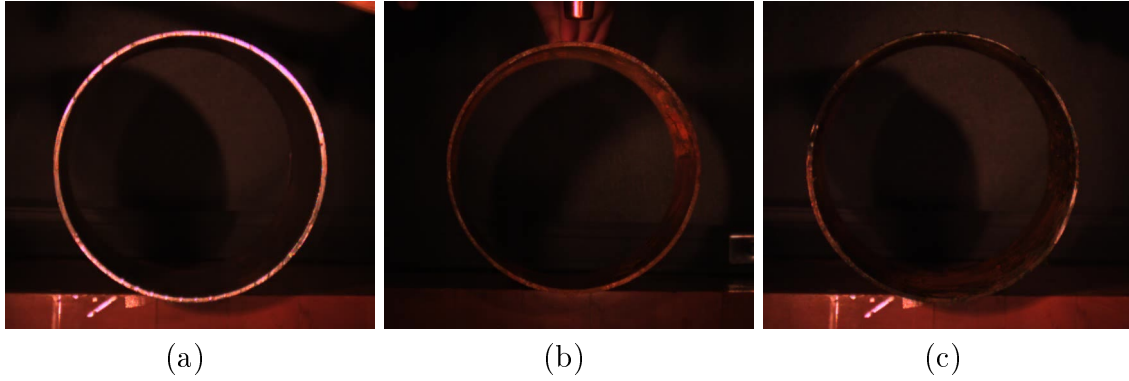
Figure 5-26: Screenshots from recorded video of the (a) intact pipe, (b) 2 day corroded pipe, and (c) 5 day corroded pipe

pipe spectrum on the bottom in blue, the 2 day corroded pipe spectrum in the middle in red, and the 5 day corroded pipe spectrum on the top in orange. The resonant peaks in the spectra are present at approximately 500 Hz for the first resonant mode, 1100 Hz for the second, and 2300 Hz for the third. In general, the resonant frequencies shift lower with increasing damage.

(a)   (b)
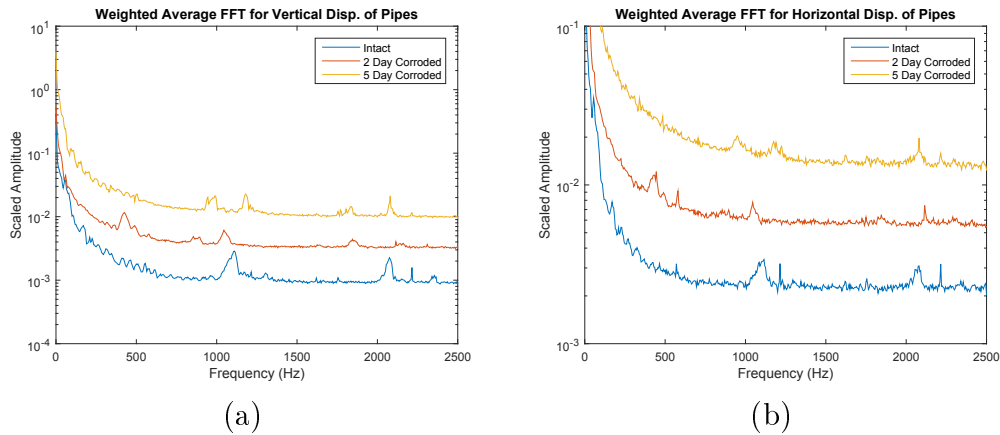
Figure 5-27: Frequency spectra from the video camera measurements of an intact, 2 day corroded, and a 5 day corroded pipe, vertically separated for clarity, (a) for vertical displacement and (b) horizontal displacement

Figure 5-27 shows the figures with superimposed arrows showing these frequency shifts. The loss in thickness of the pipe lowers the resonant frequencies of the pipes, as stiffness is proportional to thickness cubed, and mass is directly proportional to

147

the thickness. As the pipe is corroded, more material and thickness is lost from the pipe, and the frequencies shifts lower. Additionally, in the 5-day corroded pipe, some more resonant peaks appear, suggesting splitting of the resonant modes with damage. This behavior is predicted by the finite element model of the pipe studied in Chapter 2, as shown in Tables 2.1 and 2.2.
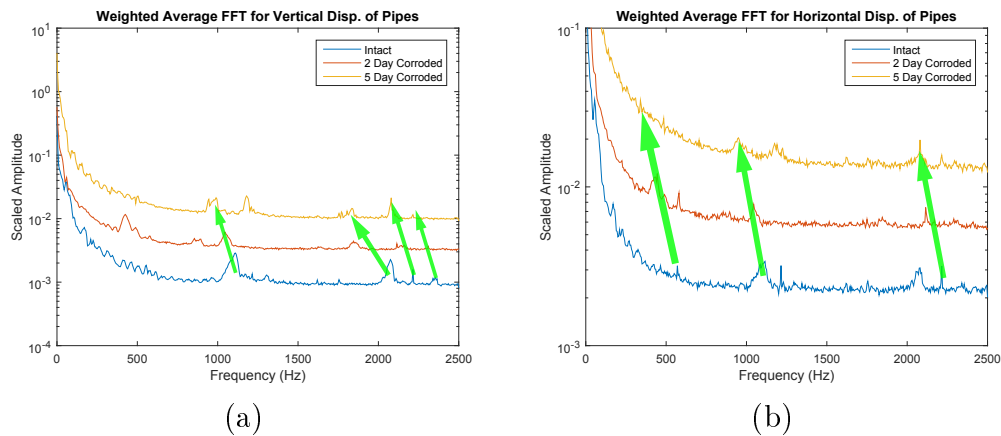
Figure 5-28: Frequency spectra from the video camera measurements of an intact, 2 day corroded, and a 5 day corroded pipe, vertically separated for clarity, (a) for vertical displacement and (b) horizontal displacement

Using the Continuous Symmetry Measure

Instead of extracting the horizontal and vertical displacements, we can use a slightly different processing procedure to extract the displacements in the radial direction from around the pipe. Using these radial displacements, we identified mode shapes, and applied the continuous symmetry measure, as described in Chapter 2, to look at changes in the symmetry of the mode shapes, potentially indicative of damage.

Displacements were extracted from 12 evenly spaced regions of interest defined around the pipe, as seen in Figure 5-29. Horizontal and vertical displacements were extracted from these regions and projected onto the radial direction at the specific location to obtain a radial displacement. FDD was used to process these radial displacements into operational resonant frequencies and mode shapes, with the first singular values for each video shown in Figure 5-30.

148

(a) (b) (c)
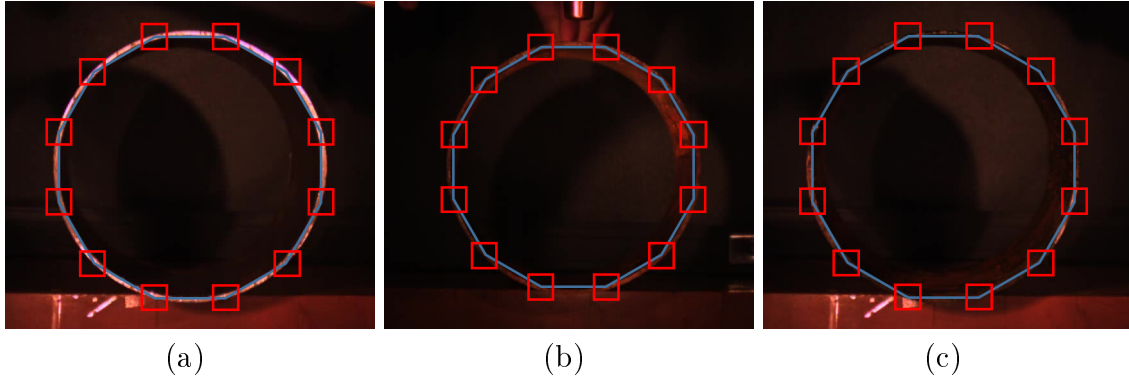
Figure 5-29: Screenshots from videos of the (a) intact pipe, (b) 2-day corroded pipe, and (c) 5-day corroded pipe, showing the locations of 12 virtual sensors in red arranged around the pipe

(a) (b) (c)
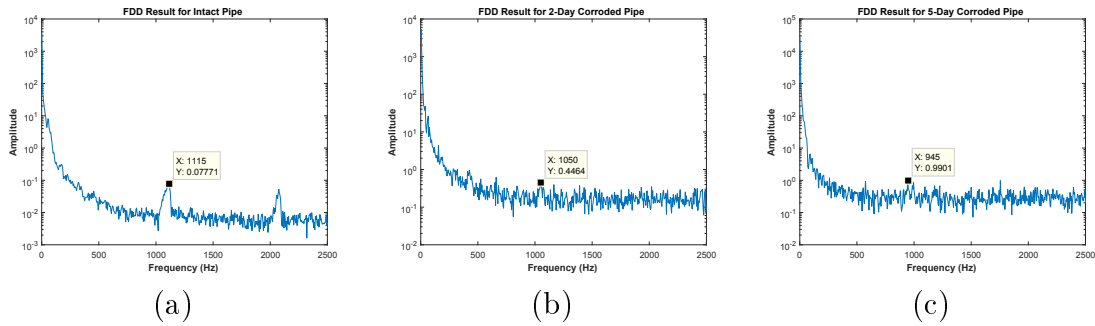
Figure 5-30: The first singular value resulting from frequency domain decomposition, showing frequency spectra extracted from the 12 virtual sensors on the (a) intact pipe, (b) 2-day corroded pipe, and (c) 5-day corroded pipe

In the first singular values from FDD, the only clean modes present in all three tests are the ones at 1115 Hz, 1050 Hz, and 945 Hz, in the intact, 2-day corroded, and 5-day corroded pipes respectively. Note that the frequencies decrease progressively between the intact pipe, the 2-day corroded pipe, and the 5-day corroded pipe, corresponding to increasing amounts of damage.

The extracted operational mode shapes at the selected resonant frequencies are shown in Figure 5-31. These modes belong to the C3 symmetric radial mode shape of the pipe section. The C3 CSM values are increasing from 0.1313 for the intact pipe, 0.1519 for the 2-day corroded pipe, and 0.1568 for the 5-day corroded pipe. Visually
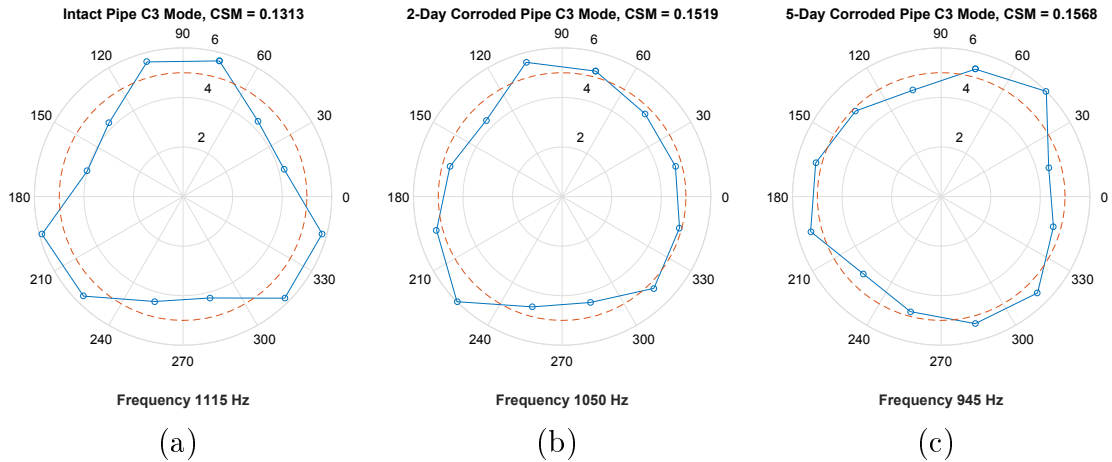
Figure 5-31: C3 symmetric mode shapes extracted from the pipe videos for the (a) intact pipe, (b) 2-day corroded pipe, and (c) 5-day corroded pipe

the pipes get less symmetric, and interestingly enough the nodes of the mode shape for the 5-day corroded pipe have also shifted locations along the pipe, when compared to the intact and 2-day corroded pipe. Close examination of the first singular value as a result of FDD in Figure 5-30(c) shows a second resonant peak at 995 Hz next to the one selected at 945 Hz. The resonant mode has split into two asymmetric modes due to the damage, with nodal positions differently arranged along the pipe, as suggested by the analysis shown in Tables 2.1 and 2.2.

Figure 5-32: A real world example of a pipe as a part of construction in Cambridge, MA

150

It is possible that this kind of measurement could be used as a test for damage in steel or plastic pipes either before they are buried, if they are damaged in transport, or after they have been in use underground and exposed through excavation. Figure 5-32 shows a picture of a construction site in Cambridge, Massachusetts just one block from MIT which shows a pipe that looks conspicuously like the experimental setup.

5.4 Granite Test

This application involves a high-speed video of a piece of granite with manufactured defects undergoing a test where it is being hydraulically fractured, courtesy of Bing Li. The quantitative values of interest are the displacements and directions of motion at different locations in the specimen, and this application of the camera method is presented in [10]. This is in contrast to many of the previous applications where the measurement is aimed at determining the frequency spectrum of the object in the video. A picture of a similar specimen in the compression testing setup is shown in Figure 5-33(a). The video was recorded at 14,000 fps with a resolution of 640 x 792 pixels over 4828 frames of video; a screenshot of the video which as been brightened for visibility, is shown in Figure 5-33.

The final 1329 frames were processed to determine the displacement at every pixel in the frame with phase-based optical flow described in Section 3.2, with the granite providing sufficient visual texture to properly measure the displacements. To visualize these displacements, they were converted into a video representation with the hue (color) signifying the direction, and the brightness representing the magnitude of the displacement as shown in Figure 5-34(a). For example, red is motion towards the left, teal is to the right, green is upwards, and purple is downwards. To fit within the range of values for a video, the displacements were normalized and the fourth root of the values was taken to magnify smaller values. A frame from the visualization just after fracture occurs is shown in Figure 5-34(b) showing the left side of the video as red, having moved left, and the right side of the video as teal, having moved to the right.
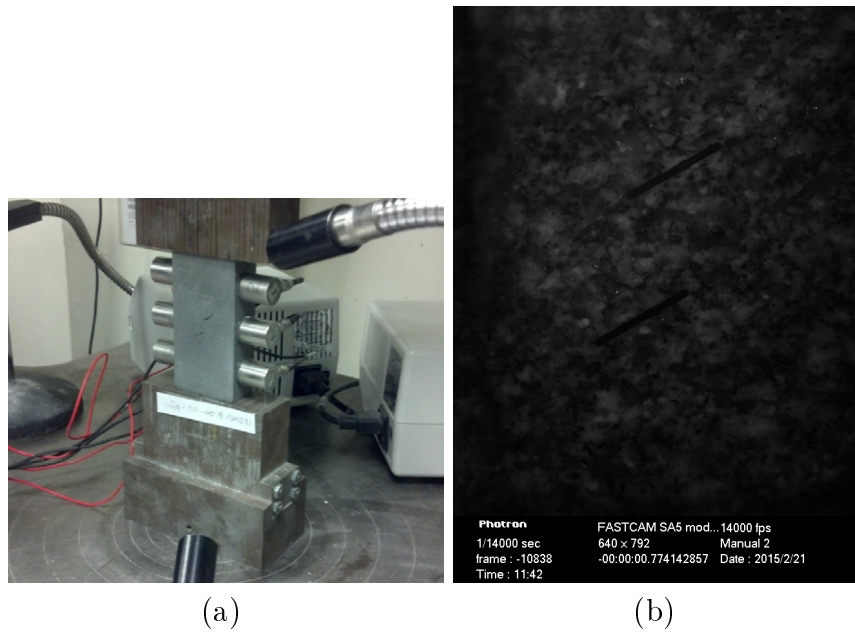
Figure 5-33: (a) Picture of a similar specimen being tested and (b) screenshot from the source video (brightened for visibility)
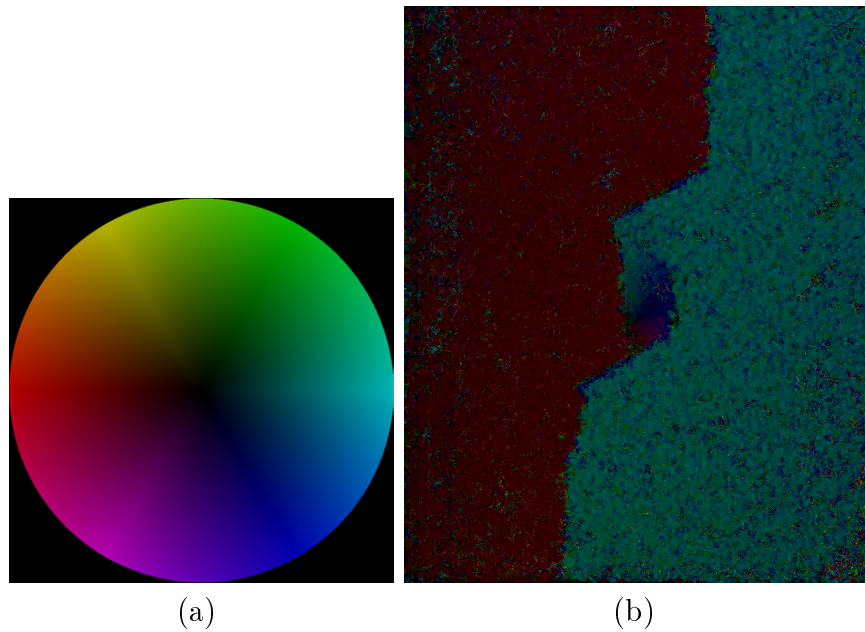
Figure 5-34: Video visualization of displacements from the granite test, (a) Color representation of direction and magnitude of motion, e.g. red for left, teal for right and (b) frame 4100 after cracking, from the generated video visualizing motion
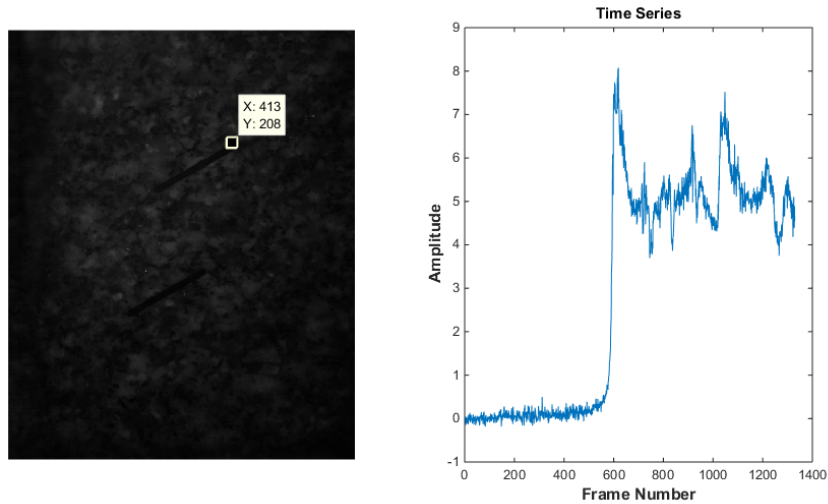
Figure 5-35: (Left) Screenshot of source video with pixel selected and (Right) measured displacement time series of the selected pixel, amplitude in units of pixels

A sample horizontal displacement measured for the pixel located at (413,208) in the video is shown in Figure 5-35 on a screenshot of the video, clearly showing the large amount of movement at fracture, and relaxation afterwards. Using these full field displacement signals we can calculate a strain field for the granite to try and predict the eventual fracture locations. Figure 5-36 shows a set of three frames from the video, 3900, 4050, and 4100. Frame 3900 is from before fracture, 4050 is from a time close to the fracture initiation, and frame 4100 is from after fracture.

Figure 5-37 shows the extracted horizontal displacements at those frames, with blue indicating negative or left motion, and yellow being positive or right. Even at frame 3900, a general trend of the left side of the granite moving left, and the right side moving right can be seen, albeit under a layer of noise. By frame 4050 the motion is clearly visible, and in frame 4100 it's obvious.

Figure 5-38 shows the horizontal strain values calculated from the displacements at the three frames. In frame 3900, a faint line corresponding to the potential fracture location is visible, which gets stronger in frame 4050. This shows the fracture location in frame 4050, before it is visible to the naked eye in the screenshot in Figure 5-36(b). In frame 4100 the fracture locations are clearly visible.

153
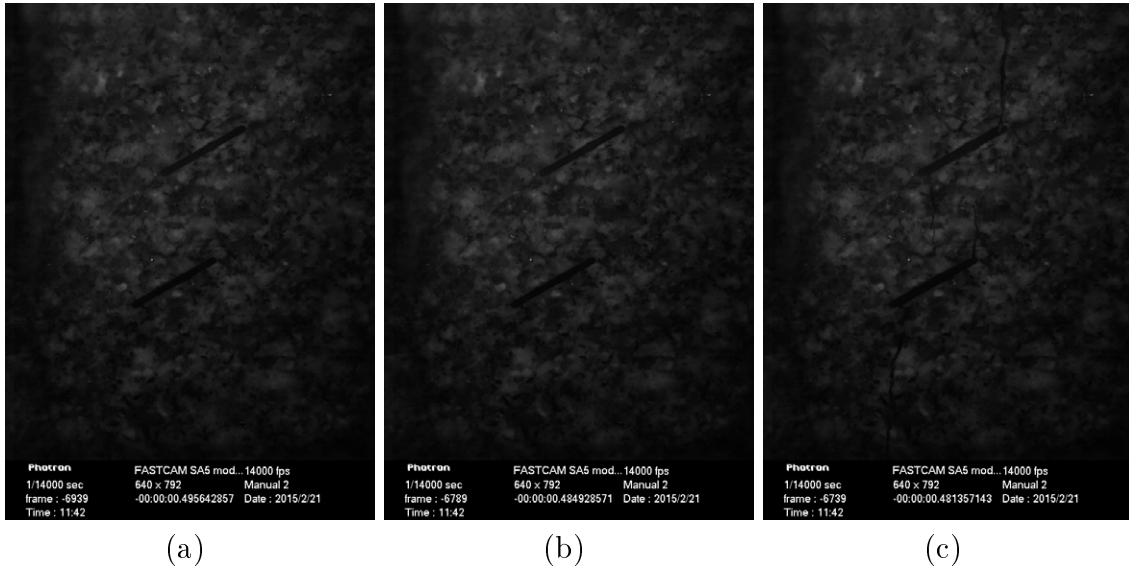
Figure 5-36: Screenshots from the video of the granite test at (a) frame 3900 before fracture, (b) frame 4050 near fracture initiation, and (c) frame 4100 after fracture
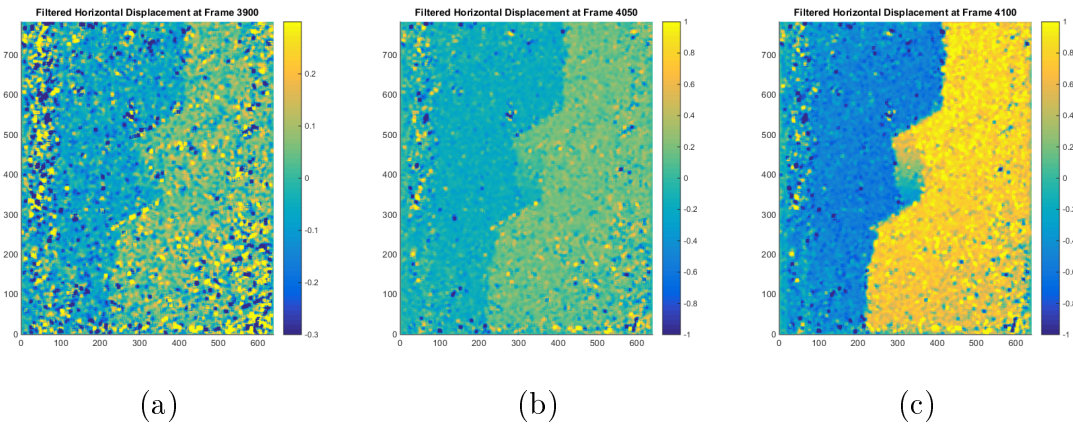
Figure 5-37: Horizontal displacements, with blue as negative or leftwards motion, and yellow as positive or rightwards motion, at (a) frame 3900 and (b) frame 4050 near fracture initiation, and (c) frame 4100 after fracture
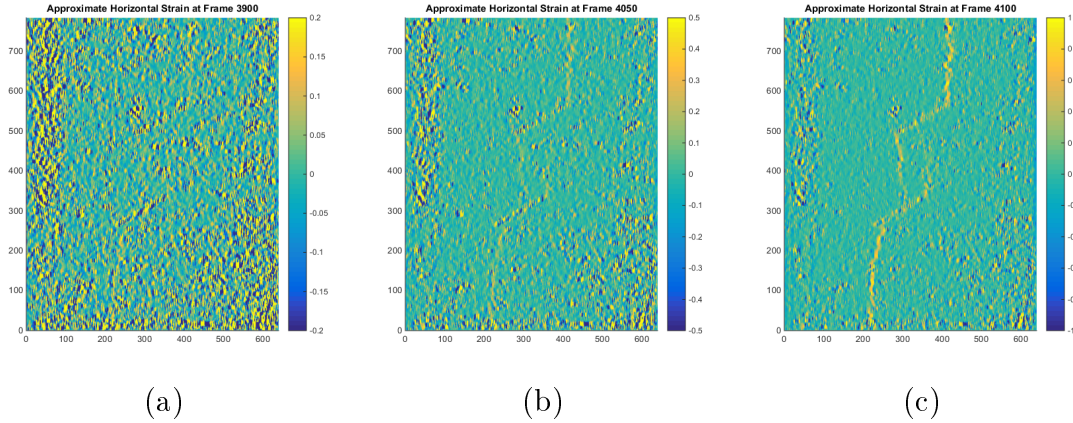
Figure 5-38: Approximate horizontal strain at (a) frame 3900 and (b) frame 4050 before fracture, and (c) frame 4100 after fracture

5.5 Chapter Summary

A series of laboratory measurements were made on two representative simple structures, a cantilever beam and the circular cross section of a pipe. Additionally displacement and strain measurements were made of a granite sample undergoing a compression test.

Measurements were made on a steel cantilever beam to identify the resonant frequencies and mode shapes. Damaged specimens, with a machined crack and a reduced cross section were also measured and distinguished from an intact example by the shift in resonant frequency. A quasi real-time implementation was used to measure lightweight plastic beams, and the resonant frequencies and mode shapes of a 3D printed beam were measured, too lightweight to be measured otherwise by contact accelerometers. These measurements were similar to those made in [29] for estimating the material properties of cantilever rods and fabrics from small motions in video.

Both PVC and steel pipe cross sections were measured by a high-speed camera setup with frame rates of 20,000 fps, 24,000 fps, and 5,000 fps. Damage was induced in the steel pipe cross sections with an impressed current corrosion setup and the lowering of resonant frequencies was correlated with greater amounts of corrosion.

The continuous symmetry measure for detecting damage presented in Chapter 2 is applied with the camera measurement to detect damage in the corroded pipes. The mode shapes became more asymmetric with greater amounts of corrosion damage.

A granite specimen undergoing a loading test was measured with a high speed camera. The varied visual texture on the surface of the specimen allowed for the extraction of full field 2D displacements. The displacements were used to make a movie visualization of the displacements, and were used to calculate the strain in the specimen. The displacements and strain showed the fracture locations in the specimen pre-fracture.

Chapter 6

Field Measurements

This chapter presents three field measurements: machinery and pipes at MIT physical facilities, the MIT Green Building, and one using a YouTube video from the Taipei 101 skyscraper.

6.1 MIT Facilities

6.1.1 Pump

A camera measurement was made of a motor that runs a water pump in the physical facilities of the Massachusetts Institute of Technology (MIT), presented in [10]. Pictures of the experimental setup are shown in Figure 6-1(a) and (b) with a screenshot of the recorded video, 1280 × 720 pixels at 2000 fps, shown in Figure 6-2. Two regions of interest where chosen for extraction of displacements in the video and calculation of the measured frequency spectra. Additionally, a laser vibrometer was used to separately measure the motions of the motor and the ground as a reference for the vibrations in the environment itself, caused by the motor and other machinery in the area.

The frequency spectra extracted from the region of interest in Figure 6-2 are shown in blue in Figure 6-3. The frequency spectra reasonably match for lower frequencies, with those measured for some of the resonant frequencies measured by the laser

(a)   (b)

Figure 6-1: Measurement setup for the motor that drives a chilled water pump showing (a) the camera location relative to the pump and (b) the laser vibrometer positioning

vibrometer, shown in red. There are some issues for a camera measurement in an environment with many sources of vibration as evidenced by the laser vibrometer measurement of the ground. There are many vibrations in the environment that shake the camera, giving the appearance of motion of the object.
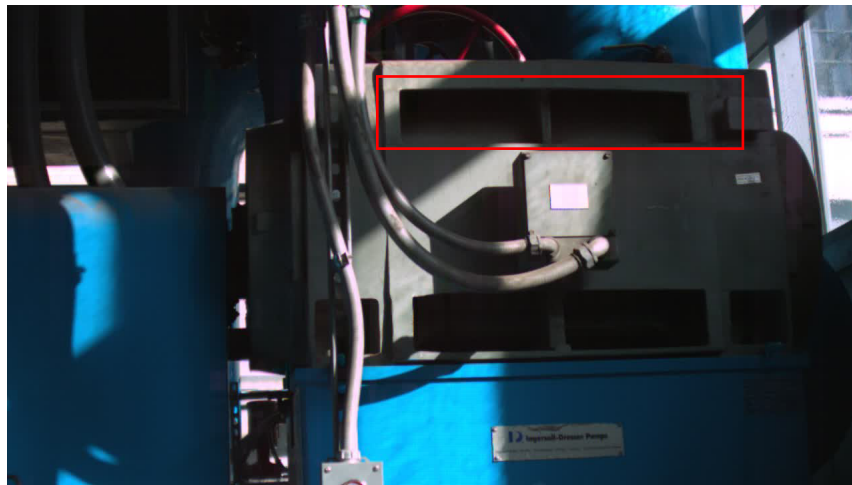
Figure 6-2: Screenshot of the measured pump video with the region of interest outlined in red

Figure 6-3(c) shows frequency spectra extracted from the full video for pixels with sufficient local contrast, in both the horizontal and vertical directions of motion.

158
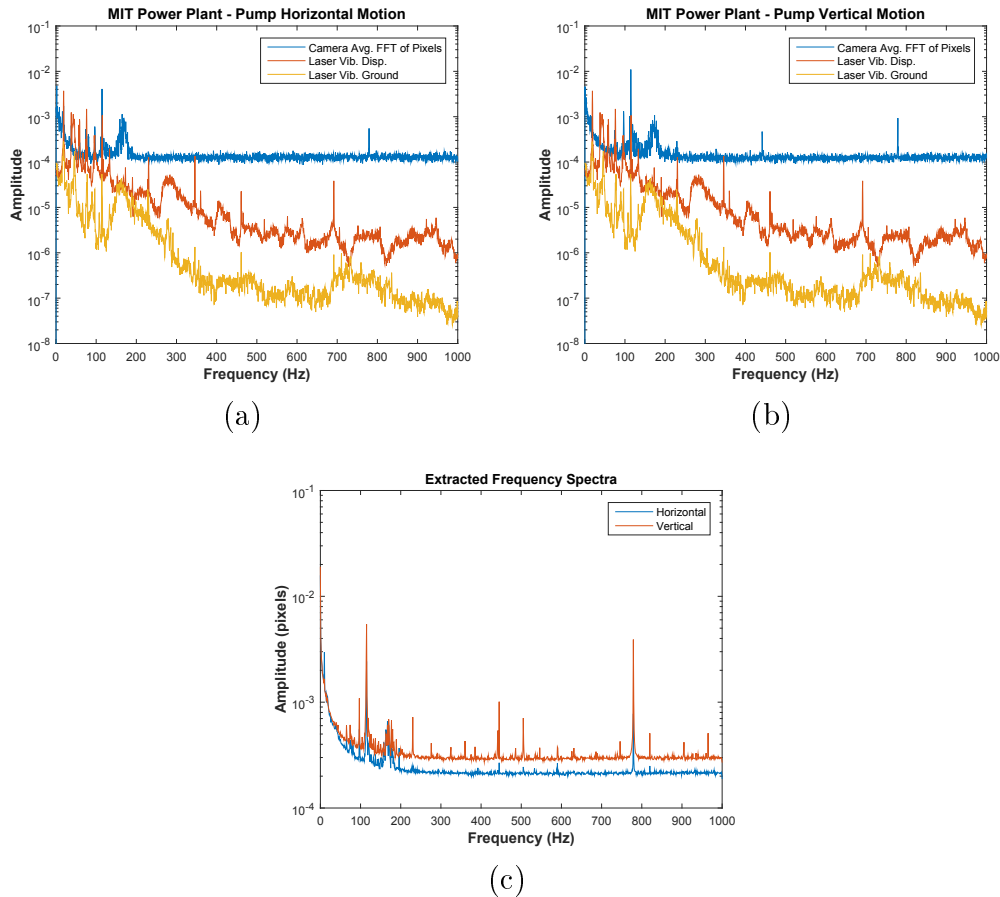
Figure 6-3: Frequency spectra from the laser vibrometer measuring the pump and the ground as reference, and from the video region of interest in the (a) horizontal and (b) vertical direction, and (c) frequency spectra in the horizontal and vertical directions for the full video

These spectra are used to generate pictures of the phase relationships between pixels of their corresponding motion signals at the following frequencies: 97 Hz, 115 Hz, 170 Hz, 230 Hz, 445 Hz, and 779 Hz. The images for horizontal motions are shown in Figure 6-4 and the images for vertical motions are shown in Figure 6-5. The higher frequencies at 230 Hz, 445 Hz, and 779 Hz have noisy phase images that don't give much information, however in the lower modes we can see some information. At 97 Hz and 115 Hz, most of the video is moving together. At 170 Hz, however, we can see that some parts of the pump and connecting wires are moving out of phase with other parts.
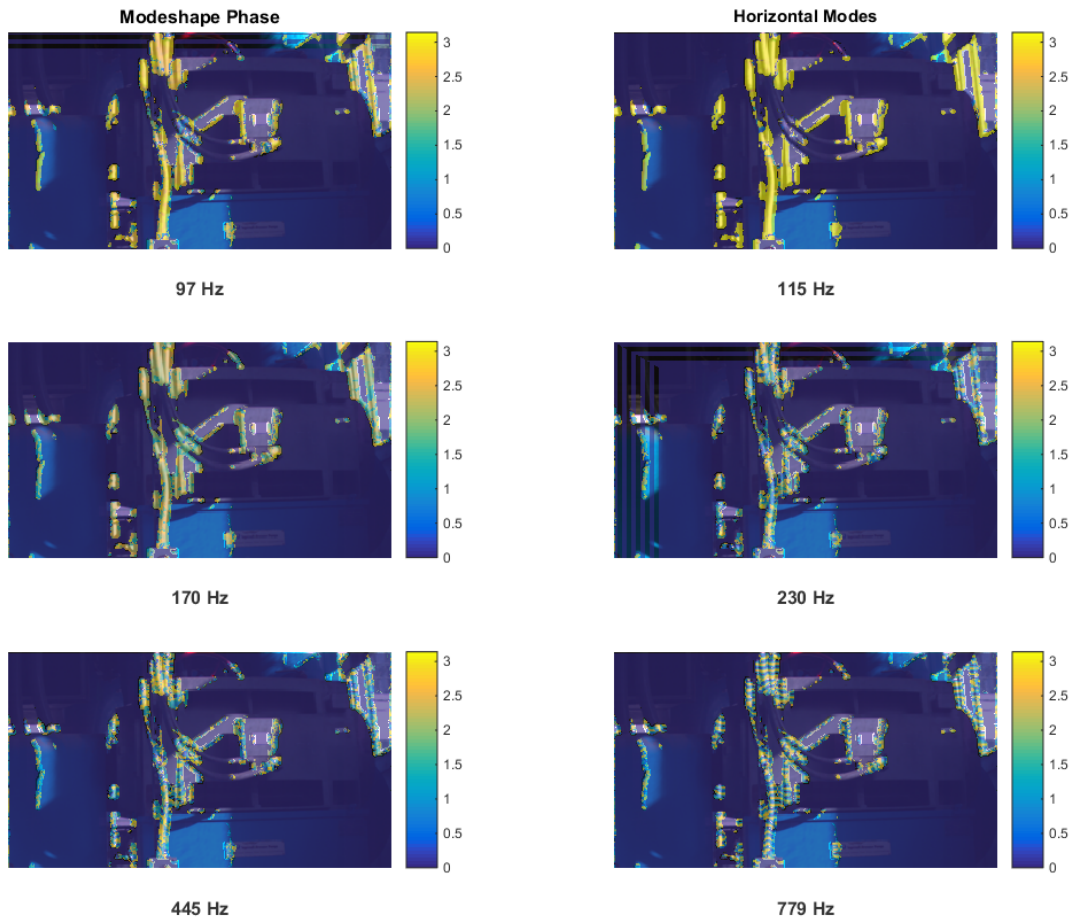
Figure 6-4: Mode shape phase relationships for horizontal displacements in the full pump video
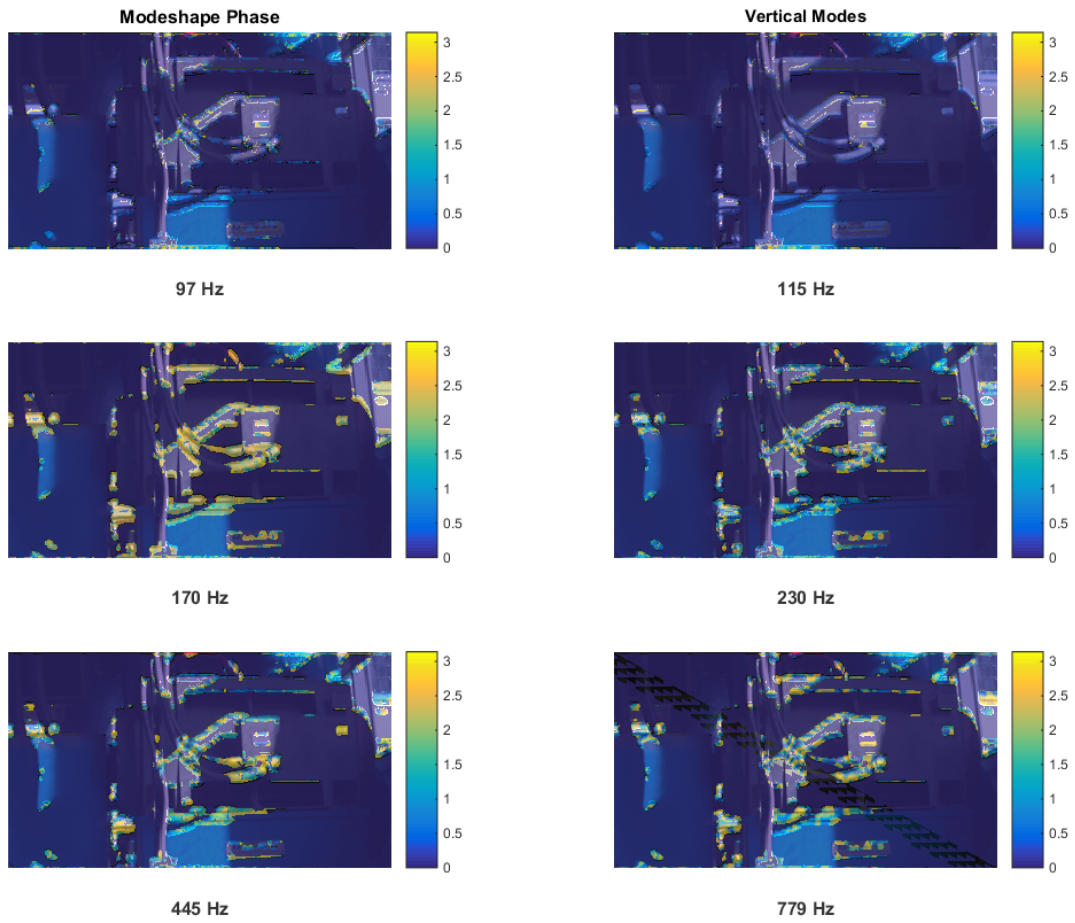
Figure 6-5: Mode shape phase relationships for vertical displacements in the full pump video

6.1.2 Pipes

A group of pipes nearby the pump was also measured. Pictures of the experimental setup are shown in Figure 6-6(a) and (b), showing the camera measurement and laser vibrometer measurement locations respectively. These pipes were connected to smaller pumps that were exciting them and causing them to vibrate. The vibrations were not visible by the naked eye, but could be felt by hand. A screenshot from the video is shown in Figure 6-7 with a region of interest outlined in red.

(a) (b)
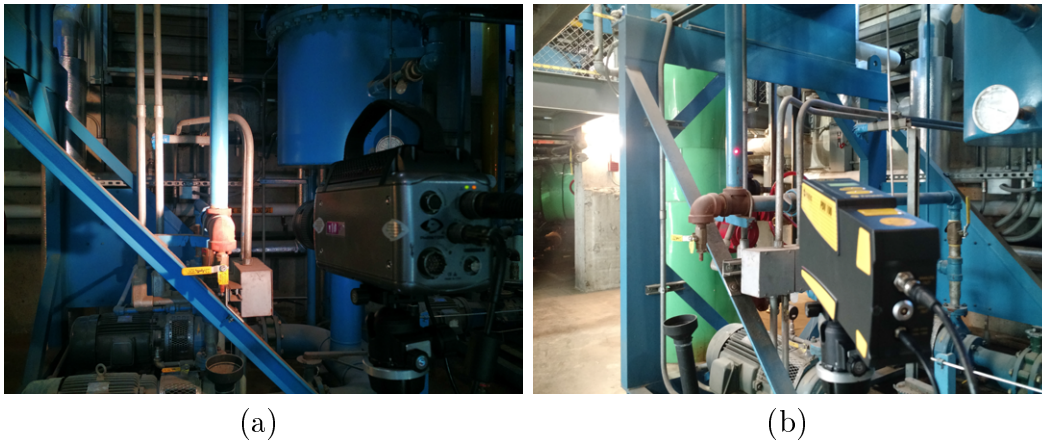
Figure 6-6: Measurement setup for the series of pipes showing (a) the camera location relative to the pipes and (b) the laser vibrometer positioning

The frequency spectrum of motion in the horizontal direction from the region of interest is shown in blue in Figure 6-8(a), with the laser vibrometer measurement shown in red, and the laser vibrometer measurement of the ground as a reference shown in orange. Most of the frequency peaks in the laser vibrometer measurement not seen in the camera measurement, but there are a couple in the lower frequencies below 120 Hz that are visible in both spectra. Figure 6-8(b) shows the vertical and horizontal motion frequency spectra for the full video showing many resonant frequencies. Operational resonant mode shape phase images are plotted at the following frequencies: 28 Hz, 59 Hz, 119 Hz, 230 Hz, 357 Hz, 445 Hz, 505 Hz, and 781 Hz.

The horizontal mode shape phase relationships shown in Figure 6-9 show that some of the pipes move in phase or out of phase with each other at the 28 Hz, 59
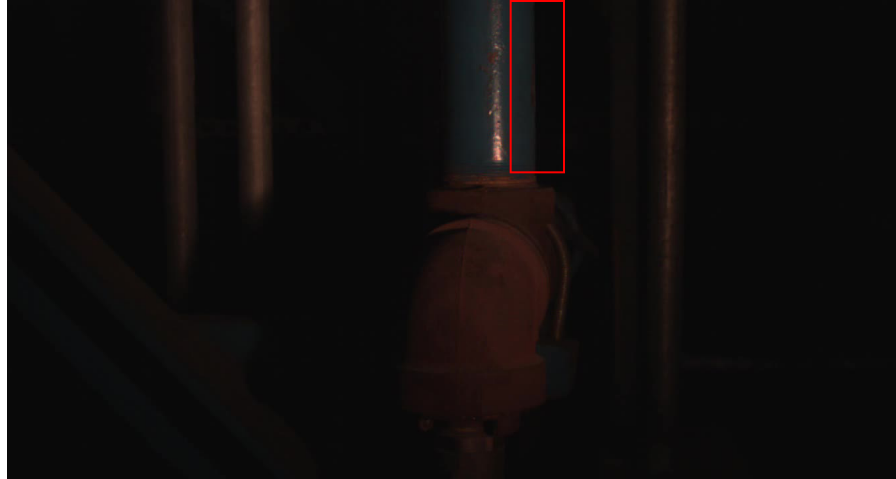
Figure 6-7: Screenshot of the measured facility pipes video with the region of interest outlined in red
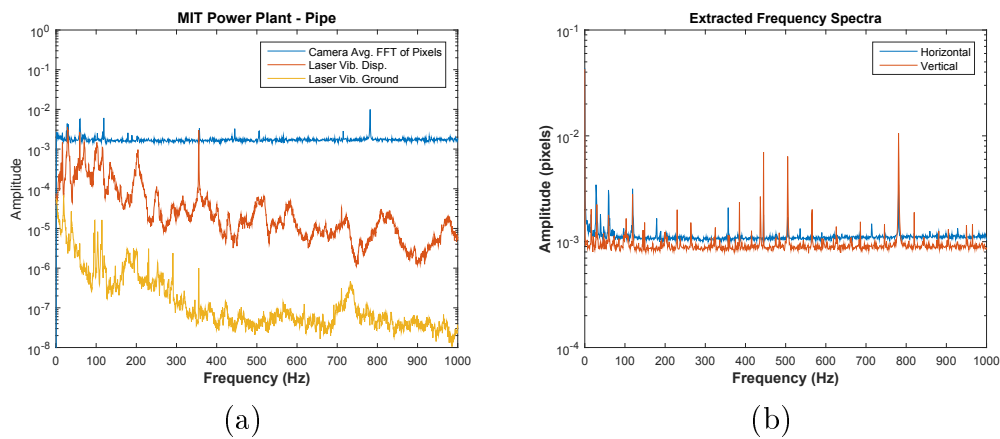
Figure 6-8: Frequency spectra from the laser vibrometer measuring the side of the pipe and the ground as reference, and from the video region of interest in the (a) horizontal direction, and (b) frequency spectra in the horizontal and vertical directions for the full video

Hz, and the 119 Hz frequencies. The 230 Hz, 357 Hz, 445 Hz, 505 Hz, and 781 Hz frequencies are too noisy to draw any conclusions.

The vertical mode shape phase relationships shown in Figure 6-10 are relatively noisy and hard to see anything conclusive, especially since very few of the pixels have good contrast from which displacements were extracted.

163
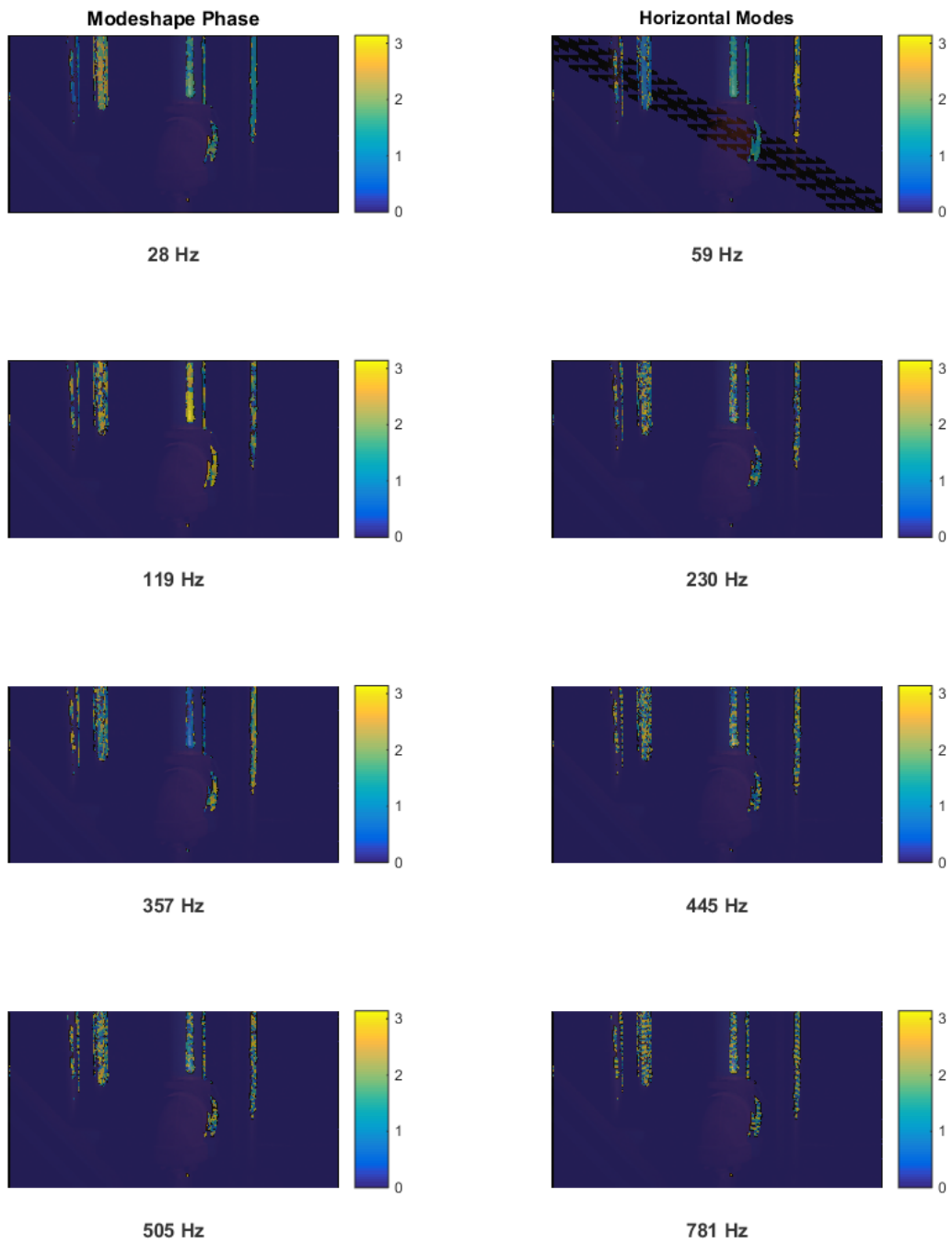
Figure 6-9: Mode shape phase relationships for horizontal displacements in the full pipe video
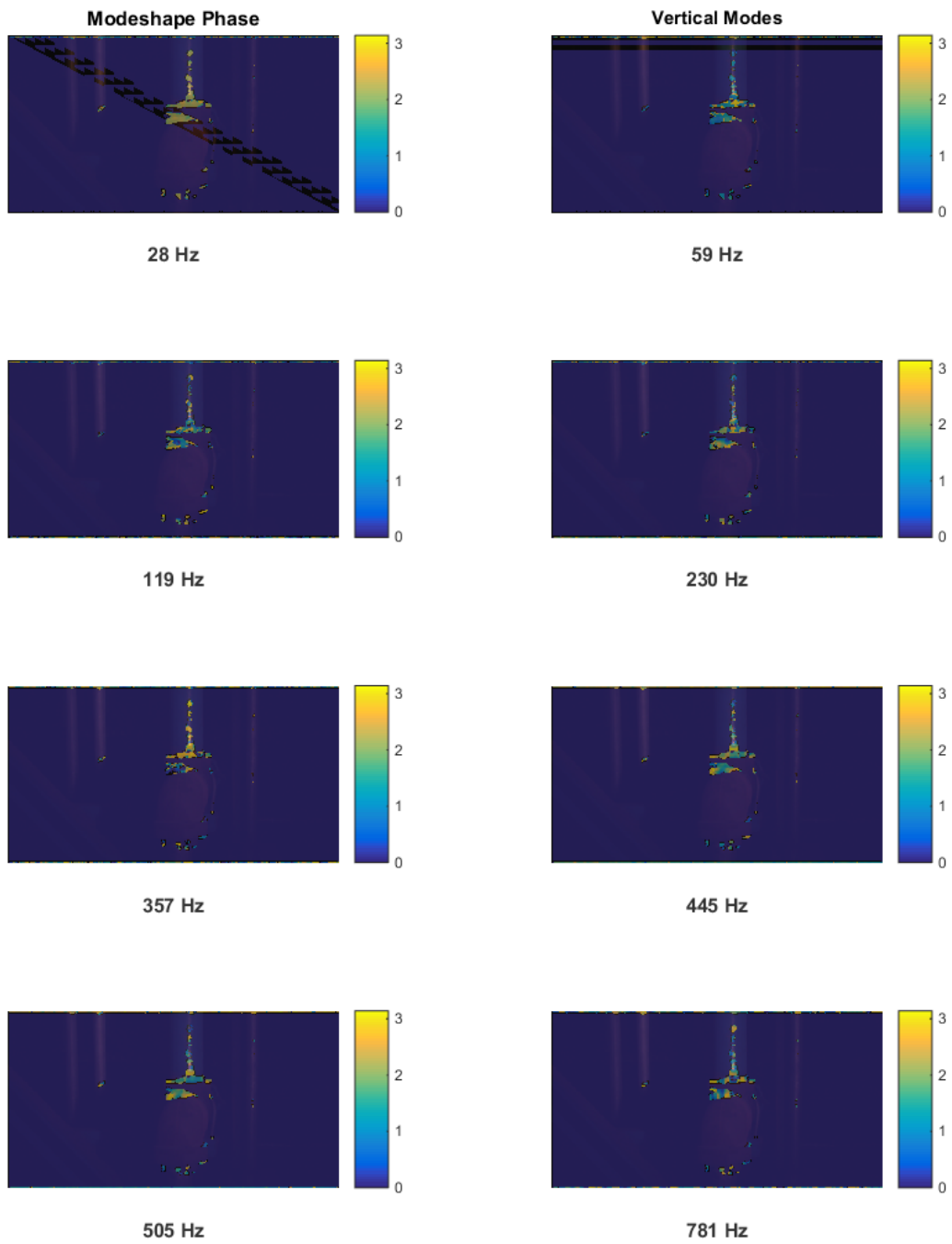
Figure 6-10: Mode shape phase relationships for vertical displacements in the full pipe video

6.2 MIT Green Building

6.2.1 Motivation

A measurement was made from a distance of over 175 m (574 ft) of the Green Building on the campus of the Massachusetts Institute of Technology (MIT) in an outdoor setting with uncontrolled lighting and without any active excitation of the structure. This measurement was presented in [25, 20, 19].

6.2.2 Experimental Setup

The video measurement was made of the Green Building on MIT's campus, a 21 story, 90 m (295 feet) tall reinforced concrete building. The camera was located on a terrace of the Stata Center, another building at MIT, a distance of approximately 175 m (574 ft) away from the Green Building, as shown in a satellite view in Fig. 6-11(a). The view from the measurement location is shown in Fig. 6-11(b) also showing the experimental setup. A Point Grey Grasshopper3 camera was used with a 24-85 mm zoom lens set to 30 mm so as to capture the whole building and the antenna tower on the roof, as seen in a screenshot from the recorded video in Fig. 6-11(c). The resolution of the video was 1200×1920, resulting in a scale factor of 3.65 cm (1.44 in) per pixel at the depth of the structure, determined from the 36 m (118 ft) width of the structure. The video was recorded in greyscale in a raw 8-bit pixel resolution format, which disables any onboard image processing. A 454 second long video at 10 frames per second (fps) was captured. During the measurement itself, the weather was clear with a temperature of 18.9° C and 4.6 m/s winds (66° F and 10 mph) [94].
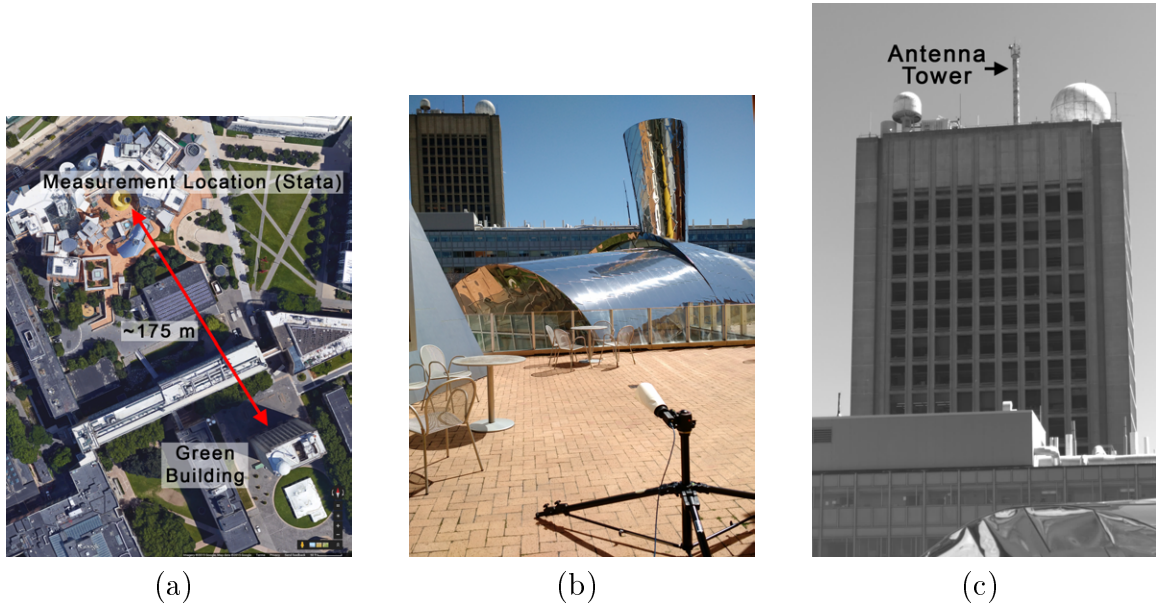
Figure 6-11: Experimental setup for measurement of MIT's Green Building, with (a) satellite view, (b) picture of measurement setup and view of the building, and (c) screenshot from the measured video showing the antenna tower visible on roof

6.2.3 Results

Green Building Initial Processing

The measured video of the Green Building was processed to determine if any frequency peaks indicative of a possible resonance of the building or other structures were captured in the measurement. The video was downsampled by a factor of 2 in each dimension to a size of 600×960. Out of a possible 563,584 pixels, slightly lower than the total number due to the size of the filter kernel with valid convolution results, 1191 pixels with displacements were extracted as shown in white, overlaid over the a video screenshot in Fig. 6-12(a). Of the 454 second data collect, the first 150 seconds is without much camera motion, so the signals were cropped to the first 150 seconds for analysis. After using an FFT to obtain the frequency spectra, they were averaged to obtain an average frequency spectrum for all the signals shown in Fig. 6-12(b). The most prominent resonant peak was at 2.413 Hz, and it was found that the pixels in the video that contributed to this peak corresponded to the antenna

tower located on the roof of the building.

(a)   (b)
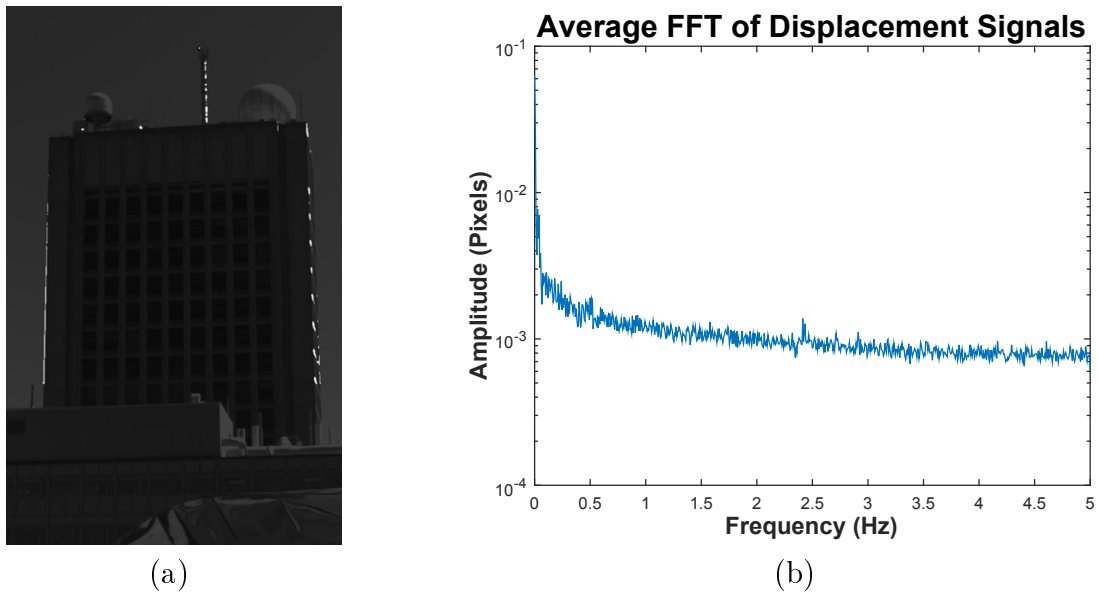
Figure 6-12: Initial data processing results showing (a) screenshot from the video with pixel mask overlaid, bright pixels being those with extracted displacements and (b) Average frequency spectrum of displacement signals extracted from the video

6.2.4 Antenna Tower Processing

In order to better analyze the motion of the antenna tower, the original video of the whole Green Building was cropped to a video containing only the tower, with a resolution of 64×240, shown in Fig. 6-13(a). Before processing, the video was downsampled by a factor of 2 in each dimension to a size of 32×120, and with the filter kernel size a possible 2688 pixels with displacements. 441 out of those 2688 pixels were high contrast pixels with extracted displacements, shown in white in Fig. 6-13(b). The resulting average displacement time series is shown in Fig. 6-14(a). As with the Green Building measurement video, due to camera motion the time series was cropped to the first 150 seconds of the video, which is shown in Fig. 6-14(b).

The jump in displacement of approximately 0.8 pixels, or which occurs from 170s to 177s in the video, is a horizontal motion of the full frame of the video, and thus likely due to camera motion. This motion is present, not only in the cropped video

of the antenna tower, but also in the full video of the Green Building.

(a)   (b)
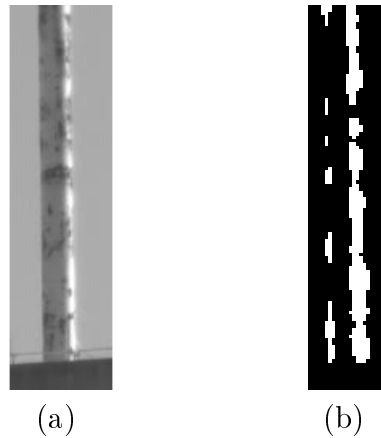
Figure 6-13: (a) Screenshot and (b) pixel mask for the video cropped to the antenna tower

The FFT of the average displacement signal is shown in Fig. 6-14(c) which shows a resonant peak at 2.433 Hz. A peak picking modal analysis method is used to determine the operational deflection shape of the structure at that frequency, and a one dimensional representation of the structure is generated by averaging the amplitude of pixels at the same vertical height. The resulting shape is shown in Fig. 6-15(a) which shows something similar to a coherent first bending mode shape of a cantilever beam, and the phase values are shown in Fig. 6-15(b). For comparison, the mode shape and phase plots for a smaller peak at 0.24 Hz, that is not a suspected resonant mode are shown in Fig. 6-15(c) and (d) respectively, showing a shape that is mostly a constant value through the height of the structure, and phase relationships that are more varied compared to the result at 2.433 Hz. Even though the operational shape at 2.433 Hz is quite noisy, when compared to the non-mode at 0.24 Hz, the trend of the measured shape is plausibly a mode shape and phase relationships are more consistent. The amplitude of the peak is 0.00589 pixels, much smaller than similar previous work has been able to detect. Using the scaling factor of 3.65 cm (1.44 in) per pixel, the amplitude of the resonant peak is calculated to be 0.215 mm (0.0085 in), and given the noise floor of 0.07 mm (0.028 inches), the SNR of this measurement is 3.
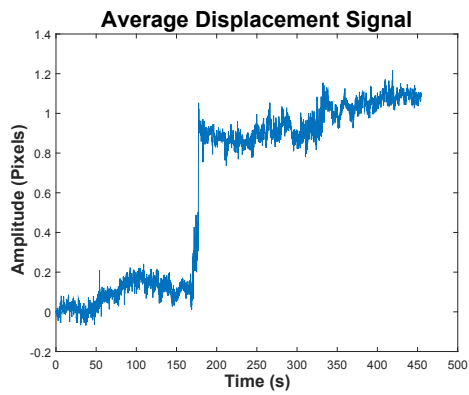
(a)
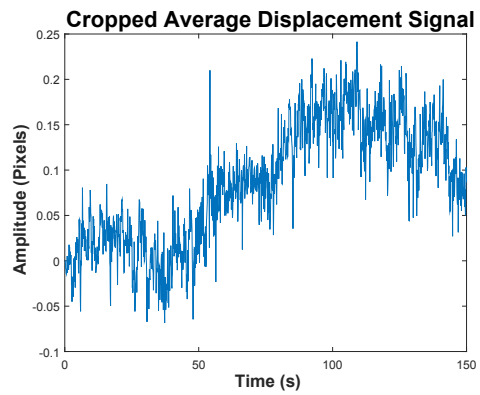
(b)
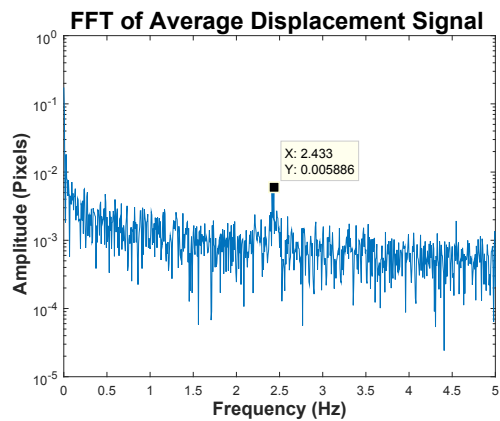
(c)

Figure 6-14: Average displacement signals measured from cropped video of Green Building antenna tower: (a) the full 454 second time series, (b) the cropped first 150s of the time series, and (c) FFT of the cropped time series signal

1D Operational Mode Shape at 2.433 Hz

1D Operational Mode Phase at 2.433 Hz

(a)

(b)

1D Operational Mode Shape at 0.24 Hz

1D Operational Mode Phase at 0.24 Hz

(c)

(d)
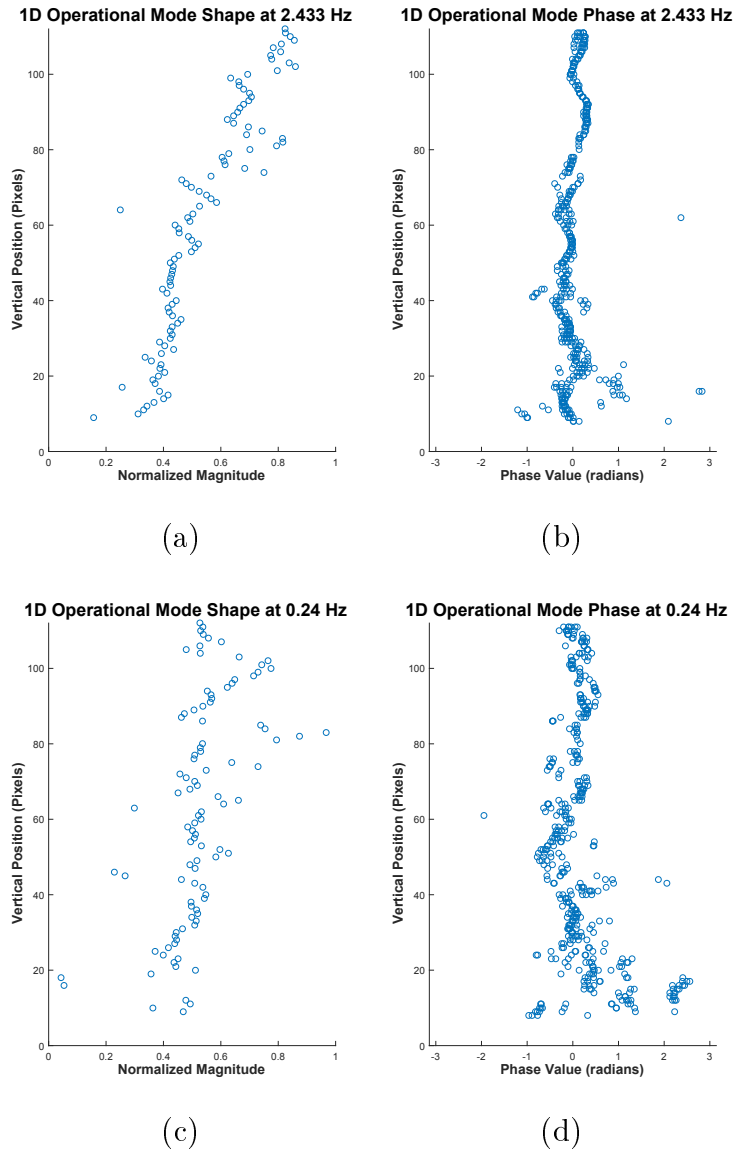
Figure 6-15: Operational modes from the camera measurement of the Green Building antenna tower: (a) operational mode shape and (b) phase at 2.433 Hz; (c) operational mode shape and (d) phase at 0.24 Hz for comparison

6.2.5 Verification

A laser vibrometer was used at close range to measure the frequency response of the antenna tower during a different day with similar weather conditions, clear with a temperature of 28.3° C and 5.7 m/s winds (83° F and 13 mph) [94], to determine the accuracy of the resonant frequency as measured by the video camera. The measurement setup on the roof of the building is shown in Fig. 6-16(a) with the laser vibrometer located about 1 meter (3.3 feet) away from the base of the tower, measuring a point about 1 meter (3.3 feet) high. To discount any potential resonances of the laser vibrometer and tripod system itself, a measurement was also made of the ground next to the antenna tower as a reference. The resulting laser vibrometer measurement from the antenna tower is shown in blue in Fig. 6-16(b) with the reference signal shown in red. The resonant peak that stands out in the antenna tower measurement front the reference spectrum is a peak at 2.474 Hz, which is similar to the 2.433 Hz measured by the camera. The percent difference between the measurements is only 1.7%, within the potential variation in resonant frequencies due to temperature variations [87], 28.3° C (83° F) during the laser vibrometer measurement and 18.9° C (66° F) during the camera measurement.
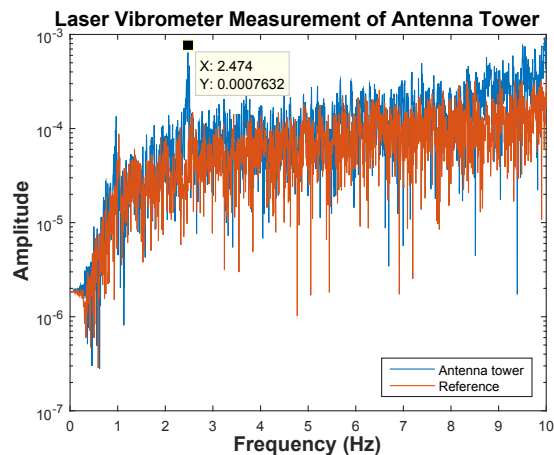
(a)   (b)

Figure 6-16: Laser vibrometer verification of MIT Green Building antenna tower resonant frequency: (a) measurement setup and (b) result with static reference for comparison

Camera Motion Compensation

The full 454 second video was not used in the case of the antenna tower measurement because of camera motion that introduced motion in the measurement displacement signal. Given that there was no measurable response from the Green Building itself, it can be used as a reference to measure the camera motion from the video, and subtract it from the displacement signal of the antenna tower. Fig. 6-17(a) shows the regions of interest used to calculate displacement signals to accomplish this motion compensation. The signals are shown in Fig. 6-17(b) and they include the antenna tower in blue, a slice from the Green Building in red, and a slice from a structure in front of the Green Building, Building 56, in orange. All three signals look quite similar, demonstrating that there is significant camera motion to be corrected.
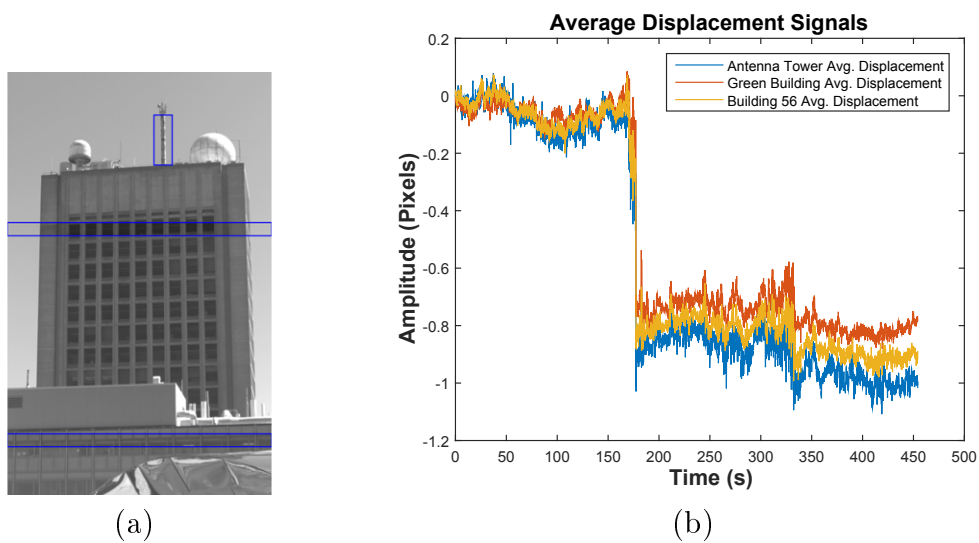
Figure 6-17: Motion compensation signals (a) video regions of interest for motion compensation and (b) average displacement signals of regions of interest
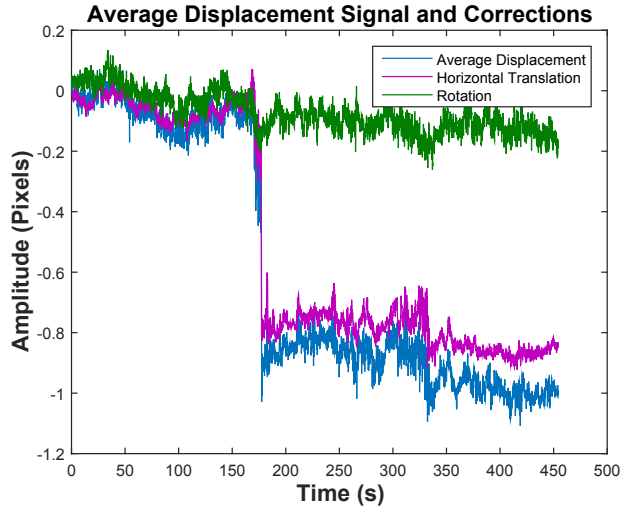
Figure 6-18: Resulting motion compensation correction signals

The two camera motions to be corrected for include both horizontal translation and rotation. Translational motion is calculated from the average of the Building 56 and the Green Building displacement signals, shown in purple in Figure 6-18, while rotational motion of the camera is calculated from the difference of the two signals, shown in green. The resulting corrected displacement signal of the antenna tower is shown in red in Figure 6-19(a), as compared to the original displacement signal shown in blue. Much of the camera motion, especially the large jump at 175 seconds into the signal is removed from the displacement signal. The resulting effects in the fourier domain are shown in Figure 6-19(b). Most of the difference seems to be in the frequencies below 0.2 Hz, which is shown in detail in Figure 6-19(c), where the noise floor is reduced by a factor of 2. Additional improvements in the noise floor can be gained by sacrificing the field of view and zooming in on the structure, to have a lower physical distance per pixel scaling factor.

Camera Noise Floor vs. Expected Building Motion

In the previously presented analysis of the data there is no evidence of the motion of the Green Building itself. Previous work has measured the resonant frequencies of the Green Building in the East-West direction, 0.68 Hz for the first bending mode

174
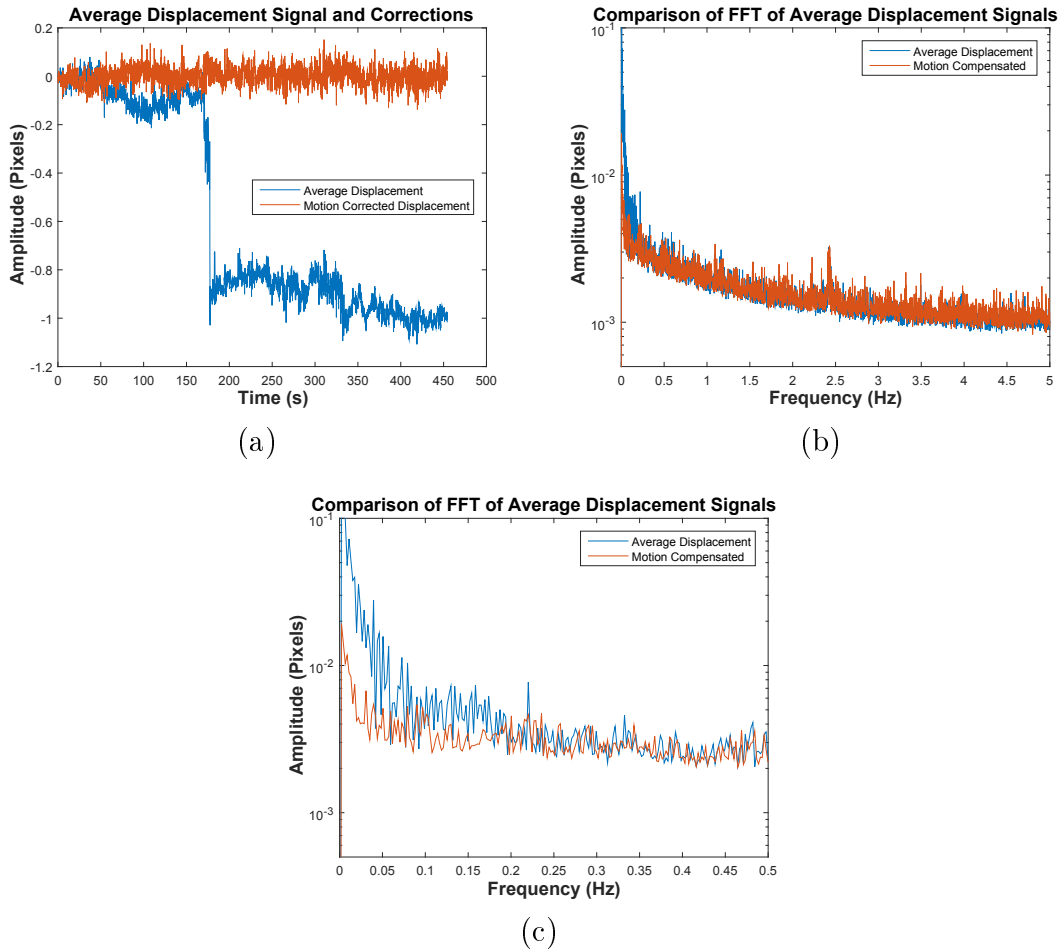
Figure 6-19: (a) Motion compensated antenna tower signal comparison, (b) frequency spectrum comparison, and (c) 0 to 0.2 Hz range of the frequency spectrum comparison in detail

and 2.49 Hz for the second [18]. The most likely explanation is that the motion of the building is under the noise floor of the camera measurement. In Figure 6-12(b) the noise floor in the region of the expected first resonant frequency is approximately 0.003 pixels, which corresponds to 0.1 mm (0.004 in) or 1×10^{-4} meters of displacement. Figure 6-20 shows accelerometer data from the roof of the Green Building, in the E-W direction, on two different days, one with typical spring weather on April 15th, 2013, and the other during a wind advisory and severe thunderstorm watch with winds of 8.9 m/s (20 mph) from the SSW on October 7th, 2013. The displacement of the structure on the calmer spring day in Figure 6-20(a) shows a displacement of

175

the structure at 0.68 Hz of only 5×10^{-6} meters or 0.005 mm (0.0002 in), which is about 1/20th of that of the noise floor of the camera measurement. Thus, it is likely that on the day of the camera measurement, the motion of the Green Building was under the noise floor. On the windy day in Figure 6-20(b), the building motion is an order of magnitude larger at 1×10^{-4} meters or 0.1 mm (0.004 in), which would be measurable by the camera. One potential issue however, is that the windy conditions would exacerbate the problem of camera motion, if the measurement were to be made outdoors.

(a)   (b)
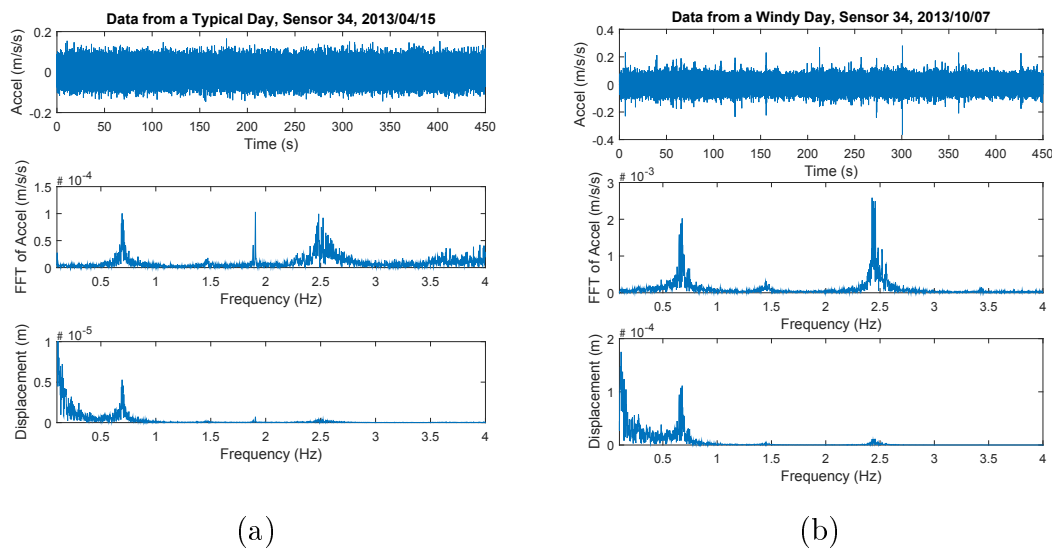
Figure 6-20: Green building accelerometer data from two days with different weather, (a) typical spring weather, (b) windy day with a wind advisory

Video Downsampling Comparison

A comparison was conducted to determine the effect of spatial downsampling on the noise floor and signal to noise ratio (SNR) of the resulting measured frequency spectrum. The previous processing was all done with the video spatially downsampled by 1 level, or reduced by a factor of 2 in each dimension; a comparison is made to no downsampling called level 0, or downsampling twice, or by a factor of 4 in each dimension of the video. Fig. 6-21(a) shows the mean noise floor for individual

176

pixels between 3 Hz and 5 Hz vs. the edge strength of the pixel, for different levels of downsampling for all pixels with valid displacements in the video. More spatial downsampling lowers the noise floor of each individual pixel, however there are fewer pixels in total. Also this shows that the noise floor gets much worse for pixels with insufficient edge strength, which is the rationale behind using a threshold to select pixels good texture or edges.
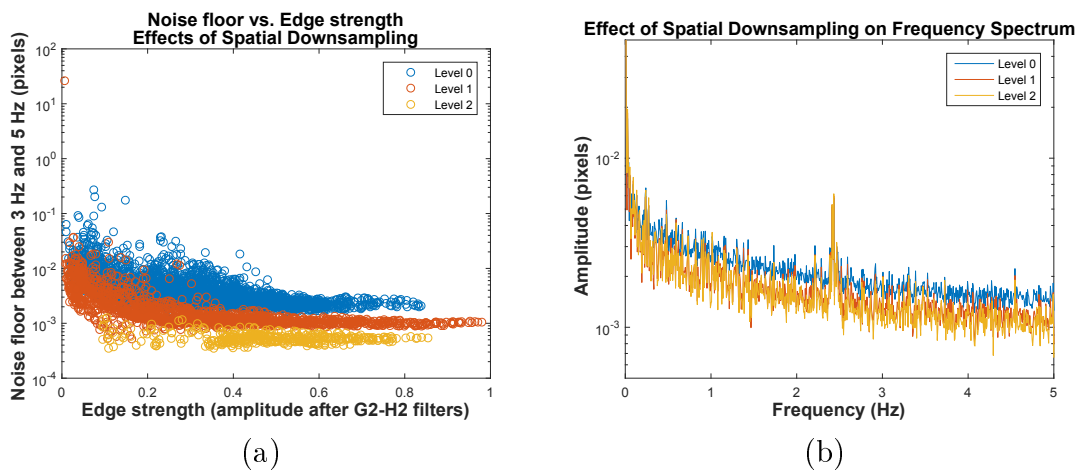
Figure 6-21: Comparison of the effects of video spatial downsampling on (a) the noise floor of individual pixels and (b) the measured frequency spectrum

Fig. 6-21(b) shows the differences in the measured average frequency spectrum, and downsampling noticably reduces the noise floor of the measurement while preserving the amplitude of the vibration peak. Table 6.1 summarizes the results for the different levels of downsampling. In this case, the SNR is calculated from the average frequency spectrum with pixels above the previously specified threshold for edge strength. With 2 levels of downsampling, not only is the processing time reduced by half, as there are far fewer pixels to process, but the SNR is improved as well. However, this is at the expense of less spatial resolution, as the frame size ends up being approximately one sixteenth of the size of the original video.

177

Table 6.1: Comparison of different spatial downsampling levels

Level	Processing time	Frame size	Valid pixels	SNR
0	64.5 seconds	56×232	3908	3.03
1	40.7 seconds	24×112	1729	3.98
2	28.2 seconds	8×52	416	4.54

Discussion

The results present the various processing steps of the measured Green Building video from the initial processing of the full video, to specific processing of the antenna tower, and compensation for camera motion. Each step is necessary to narrow down the focus of what is actually being measured by the camera. In the initial measurement, the camera measures the relative motion between the camera itself and any objects in the frame. The antenna tower was determined to be an object of interest, and as camera motion was quite significant, other portions of the video were used as references to subtract the camera motions from the displacement signal from the antenna tower. The resonant frequency of the antenna tower measured by the camera was successfully verified against the laser vibrometer measurement made from the roof. The signal from the antenna tower also contains the motion of the Green Building, as the tower is directly and rigidly mounted to the roof. However, there is no evidence of the Green Building resonant frequencies which are likely under the noise floor of the current measurement as the resonant frequencies of the Green Building in the East-West direction, 0.68 Hz for the first bending mode and 2.49 Hz for the second, are not present in the displacement signals [18].

The measurement of the antenna tower in this situation represents a proof of concept for remote vibration measurement of civil infrastructure, as a basis for condition assessment. There are similar antenna towers on many structures and they are subject to rain and high winds among other environmental effects that can cause deterioration such as corrosion and metal fatigue. By tracking the resonant frequency and other properties of the structure's displacement signal over time, any changes in those properties could be indicative of potential damage that would warrant further inspection with more traditional and definitive NDT methods. Possible methods

for monitoring changes in the displacement signal include statistical pattern recognition techniques [90], one-class machine learning methods [61], analysis of non-linear features [64], or other damage detection algorithms. Without the need for physical access to instrument a structure, cameras can more easily collect data from structures that might otherwise be difficult or time consuming to instrument. Continuous monitoring of structures with cameras seems tractable in the near future as the camera hardware involved is relatively inexpensive, since civil infrastructure tends to vibrate at relatively low frequencies accessible to normal consumer cameras.

The main improvement that needs to be made to be able to obtain meaningful displacement signals from many different types of civil infrastructure is a lower noise floor. Much of the previous work with camera-based measurements require multiple pixel displacements, and only a few methods are able to access the 1/5th of a pixel range of sub-pixel measurements. This technique gives a noise floor of less than 1/500th of a pixel or 0.07 mm, which in this measurement is sufficient to measure the motion of the antenna tower, but not enough to measure the motion of the Green Building itself. The Çelebi et al. study [18] found that ambient motion of the Green Building had accelerations with an amplitude of 0.02 cm/s^2 in the East-West direction, which at the first bending mode corresponds to displacements of 0.01 mm, below the noise floor of our measurement. A more powerful zoom lens would give a lower physical distance per pixel ratio, and lower the physical scale of the noise floor at the expense of a smaller field of view. Longer measurement windows would also improve the SNR, achievable as a part of a continuous monitoring scheme where a video camera might be able to collect valid data during any daylight hours. Because structural motions may not persist throughout the full measurement and coherent averaging may end up average out the vibration signal, incoherent averaging over a sliding window for a very long measurement time may be necessary. Another significant contributor to the noise floor is camera motion. For low frequencies less than 1 Hz and long-term measurements it is possible that background ground motions in an urban environment may be a source of seismic noise that causes camera motion [44], that cannot be averaged out. Instead, camera motion could be reduced by

correcting the collected data with accurate measurements of the camera's motion, possibly exogenous to the video itself. There is work in computer vision that uses gyroscopes and accelerometers to correct for camera shake in deblurring images [51], which suggests that similar methods could be used for videos as well, while preserving the subtle motions present in objects in the video.

6.3   YouTube Video: Taipei 101 Earthquake

In this application, we demonstrate the approach on a video from YouTube `https://youtu.be/LA-ZcEWCOIc` not recorded by the researchers, for measurement of civil infrastructure, presented in [10]. The example video is from a fixed camera within the observation deck of the Taipei 101 tower in Taipei, Taiwan during a M6.4 earthquake which occurred at a distance of 120.85 km (75 mi) south east, on 2015/04/20 at 01:42:58 UTC [109].

The Taipei 101 tower is a 101 story and 508 m (1667 ft) tall building in Taipei, Taiwan. It was completed in 2004, and when completed was the world's tallest building [73]. In Taiwan, there are potentially very large forces due to seismic and typhoon events, and the structure makes use of tune mass dampers (TMD) to withstand these forces [47]. The main TMD for the building spans 5 floors of the structure and is a pendulum consisting of a 660,000 kg steel ball and viscous dampers. The TMD is equal to the 0.24 % of the total building mass, and is located with its equator 1 meter (3.3 feet) above level 88 [73].

The source video starts at 01:43:01 UTC, 3 seconds after the earthquake starts, and continues for 2 minutes. It has half the typical framerate of 15fps. The Taipei 101 tower is unique in that the structure's main TMD, is plainly visible in the indoor observation desk, and also this video. A screenshot of the video is shown in Figure 6-22(a). The displacement of the TMD is extracted from this video sequence, from the cropped video of size 640 x 130 pixels, shown in Figure 6-22(b). The mask for pixels with satisfactory displacement information is shown in Figure 6-22(c).

The displacement extracted from the TMD in the video is shown in Figure 6-23,

(a)
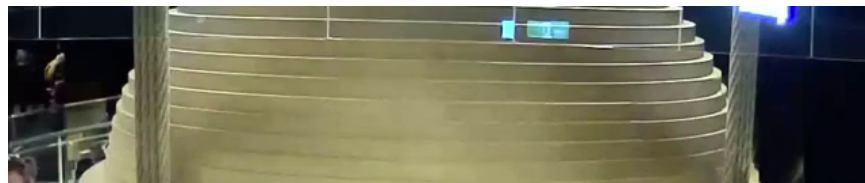
(b)
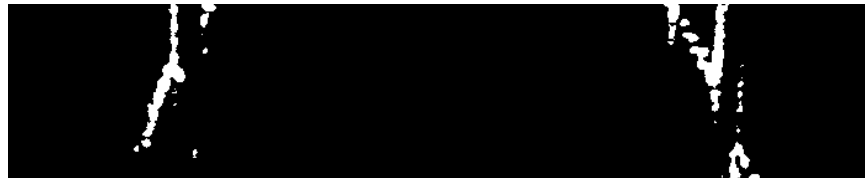
(c)

Figure 6-22: (a) Screenshot from source video, (b) video cropped to tuned mass damper, and (c) pixels with valid displacements

showing both the time series and the frequency spectrum. At 43 seconds after the video starts, there is clearly a large amount of motion, also visible in the video, that likely corresponds to the arrival of S-waves from the earthquake. At 28 seconds, there is a much smaller disturbance not visible in the video, but revealed by the processing which corresponds to the arrival of P-waves from the earthquake. This behavior is standard for earthquakes as the wave speed of P-waves is faster than those of S-waves propagating in the Earth. Using the distance and the timing of the vibrations seen in the video, the P-wave speed is calculated to be 4.32 km/s and the S-wave speed is 2.81 km/s. These numbers for the wave speeds are plausible for the Taipei basin [58].
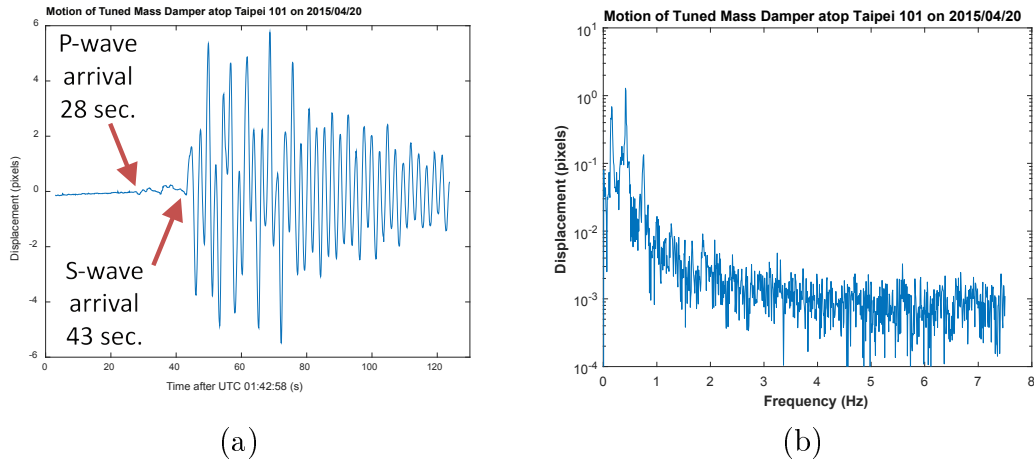
Figure 6-23: Extracted motion of Taipei 101 tuned mass damper from video showing (a) displacement time series and (b) frequency spectrum

A log-log plot of the frequency spectrum of the TMD motion is shown in Figure 6-24(a), which better shows the lower frequencies. The three dominant modes in the spectrum are at 0.1575 Hz, 0.4144 Hz, and 0.7541 Hz. The 0.1575 Hz mode is close to the expected first natural vibration period of 6.8 seconds, or a frequency of 0.147 Hz [47]. Compared to the response of the building as a result of the 2011 Tohoku earthquake [27], the frequencies match up well to the equivalent modes at 0.15 Hz, 0.43 Hz, and 0.78 Hz. During the 2011 Tohoku earthquake, the response was measured by a seismometer in the N-S direction on the 90th floor of the building, and the frequency spectrum is shown in Figure 6-24(b).
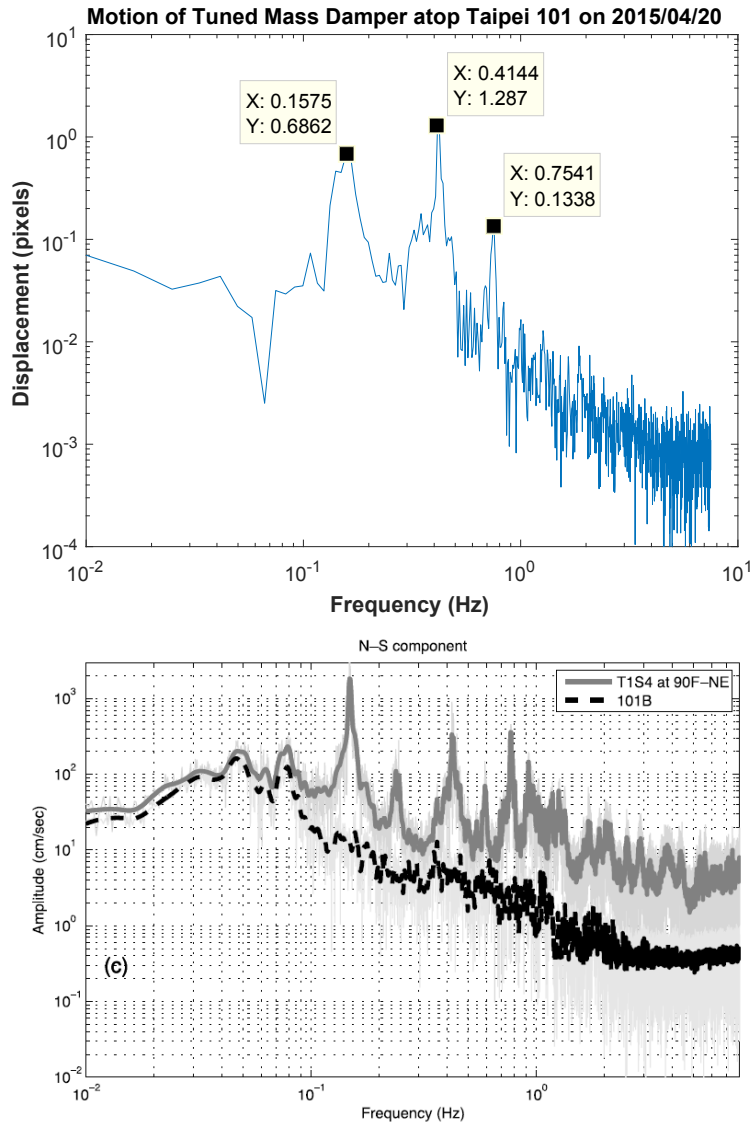
Figure 6-24: (a) Log plot of extracted motion from the Taipei 101 tuned mass damper as compared to (b) spectrum of building motion measured by a seismometer on the 90th floor of the Taipei 101 building (in grey) as a result of the 2011 Tohoku earthquake [27]

6.4 Chapter Summary

In this chapter, field measurements were made of pumps and pipes at the MIT physical facilities plant, the MIT Green Building, and using a video from YouTube of the TMD in the Taipei 101 building.

Measurements were made of vibrating pumps and pipes at the MIT physical fa-

cilities and showed that vibration frequency spectra could be extracted from video in that environment that compared well to a laser vibrometer measurement. The measurement environment is quite challenging as the pump being measured vibrates much of the ground nearby.

The ambient vibration response of the antenna tower atop MIT's Green Building was measured with a camera system from a distance of over 175 m (574 ft). The resonant frequency of 2.433 Hz measured by the camera agreed within 1.7% of the frequency measured by a laser vibrometer from close range. The amplitude of the motion detected was 0.215 mm (0.0085 in) or 1/170th of a pixel, and the noise floor of the measurement was 0.07 mm (0.028 in) or 1/500th of a pixel for an SNR of 3. Camera motion was compensated for by using static reference objects in the same video scene, and noise in the range of 0 to 0.2 Hz was reduced by a factor of 2. The effects of video downsampling were shown in that it improved the SNR at the expense of spatial resolution.

The video from YouTube shows the tuned mass damper in the visitor center at the top of the Taipei 101 skyscraper during an earthquake, and a displacement signal was measured from the mass showing the p-waves and s-waves from the earthquake. The wave speeds from the timing of the measured waves were reasonable for the Taipei basin. The resonant frequencies from the TMD response match those observed from the building in literature.

Chapter 7

Portsmouth Bridge Study

7.1   Background and Motivation

The WWI Memorial Bridge goes over the Piscataqua River from Portsmouth, New Hampshire to Badger's Island in Kittery, Maine with a total length 1201 ft (366.1 m). It is a vertical-lift truss bridge with a clearance of 130 ft (39.6 m), finished in 2013 to replace the previous vertical-lift memorial bridge originally completed in 1923 which had to be permanently closed in 2011 for safety and structural concerns. Pictures of the bridge are shown in Figure 7-1, specifically showing the central lift span.

A research team from the University of New Hampshire (UNH) headed by Professor Erin Bell plans on instrumenting the bridge with a SHM system consisting of wired accelerometers and strain gauges, and a tidal turbine for power, as a benchmark structure to test both model-based and non-model-based condition assessment algorithms. Similarly, this camera measurement study of the bridge is meant to be a proof of concept of the camera-vibration measurement system of civil infrastructure. The goal is to determine the kind of information that can be extracted from the video measurements, including deflections and vibrational behavior. This information could then be used for monitoring and condition assessment, as a basis for application of the system to other bridges worldwide.

The lift motions provide consistent excitations that vibrate the bridge, which simplifies the task of SHM. They remove some of the problems associated with excitation

variability, and significantly excite the bridge so that it can be measured without waiting for environmental or other external conditions to induce sufficient motion in the bridge. The lift mechanism can only be operated safely in normal environmental conditions; it is unsafe to do a lift in excess winds. Currently, the bridge operator looks at a windsock to see if the wind is blowing too strongly to make a safe lift, corresponding to an approximate wind speed. This critical wind speed is currently determined by the structural design limitations of the bridge, not necessarily on any empirical testing on the built structure. Thus, it would be useful to use a vibration measurement system to determine the effect of wind on the motion of the lift towers, and also to measure the effects of each lift on the trusses itself. The lifts also introduce unique stresses and loads to the bridge that could cause it to deteriorate differently than from traffic loading over time.

(a)

(b)

(c)

Figure 7-1: Pictures of the WWI Memorial Bridge in Portsmouth, NH (a) memorial dedication plaque, (b) center span in a lowered state, and (c) center span in a lifted state

7.2 Camera Measurements

7.2.1 Experimental Setup

A pilot study was conducted with the goal of demonstrating that a video camera could measure the vibration characteristics of the bridge and be used in support of a wider SHM system for decision making about the serviceability and maintenance schedule for the bridge. A camera was set up approximately 260 feet (80 meters) away in a nearby park, with the satellite view as shown in Figure 7-2(a) and the measurement setup shown in Figure 7-2(b).

The camera measurement system consists of a Point Grey camera, a heavy tripod,
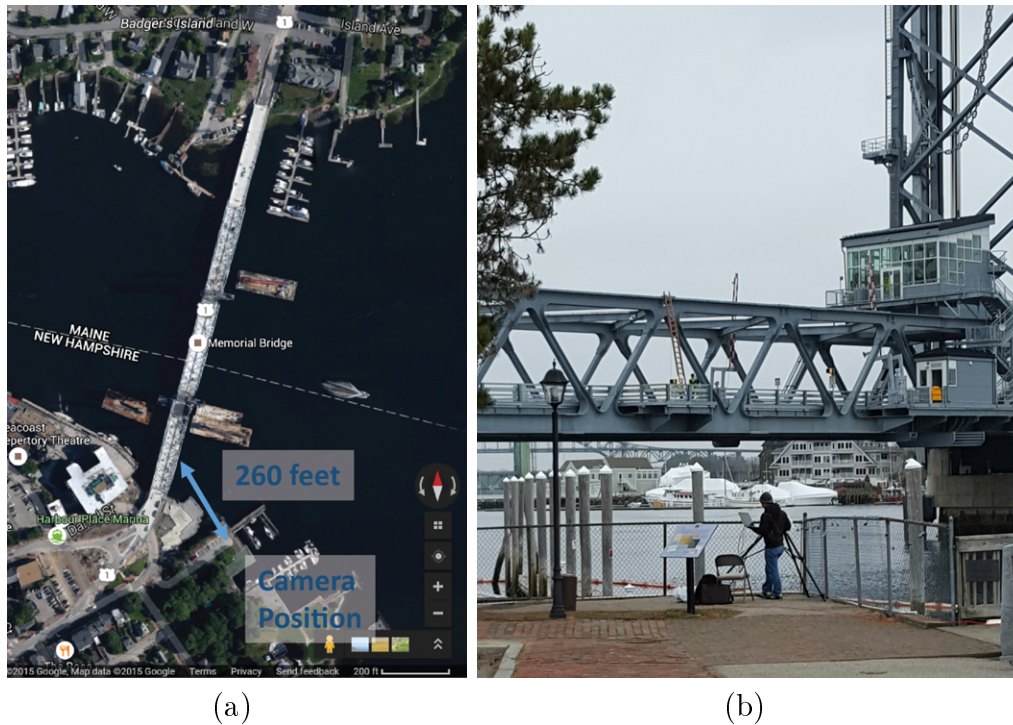
Figure 7-2: (a) Satellite view of the camera measurement location relative to the NH span of the bridge and (b) picture of both the camera measurement setup and view of the NH span of the bridge

a MEMS accelerometer system to measure the camera motions, and a laptop to collect the data. The whole system is portable and can be run with a set of 12V batteries and an AC inverter for portability, as shown in Figure 7-3.

Measurements were made on two different dates, October 8, 2015 and December 14, 2015. The measurement on October 8 was a quick measurement to see if a lift would produce motions that could be seen in the measured video, while the December 14 measurement was made concurrently with accelerometers placed on the bridge in the frame of the video.

Figure 7-3: Picture of the camera measurement setup

7.2.2 October 8, 2015 Measurement

The video was collected at a frame rate of 30 fps, with a resolution of 1184 × 700 pixels. The view from the camera is shown in Figure 7-4(a), and outlined in red is the field of view of the camera, shown in Figure 7-4(b), a screen shot from the camera video. The video was recorded for a full 4 minutes during a lift event, when the center section of the bridge is lifted to allow marine traffic to pass under, and lowered back into place to allow road traffic to resume. As the bridge is lowered back into place, the bridge is excited and vibrates more than with typical car traffic. The video also includes a segment where cars are traveling over the bridge, and also vibrating the bridge. In Figure 7-4(b) the section outlined in red, part of the bottom girder of the bridge, was cropped and processed to avoid erroneous motion signals from cars or pedestrians on the bridge.

The results are shown in Figure 7-5(a) where we can see that three resonant frequencies are identified from the signal extracted from the video. The resonant frequencies are 1.71 Hz, 2.46 Hz, and 8.43 Hz. Figure 7-5(b) shows a temporary accelerometer measurement on the bridge, from a different day, on August 27, 2015, which shows recovered resonant frequencies of 1.713 Hz and 2.47 Hz for the bridge. The small differences in the frequencies are explained by the likely different tempera-

(a)   (b)

Figure 7-4: (a) Picture showing the view from the camera measurement location with the video area outlined in red and (b) screenshot of the recorded video with a region of interest for the bottom chord of the bridge outlined in red

tures on the different days of the camera and accelerometer measurement. They are nonetheless quite similar and show that the frequencies the camera is measuring from 260 feet (80 meters) away is the actual motion of the bridge. We could then use this information for a decision making SHM system, and determine how the denser set of measurements can augment and provide different information from a sensor system on the bridge itself.

(a)   (b)
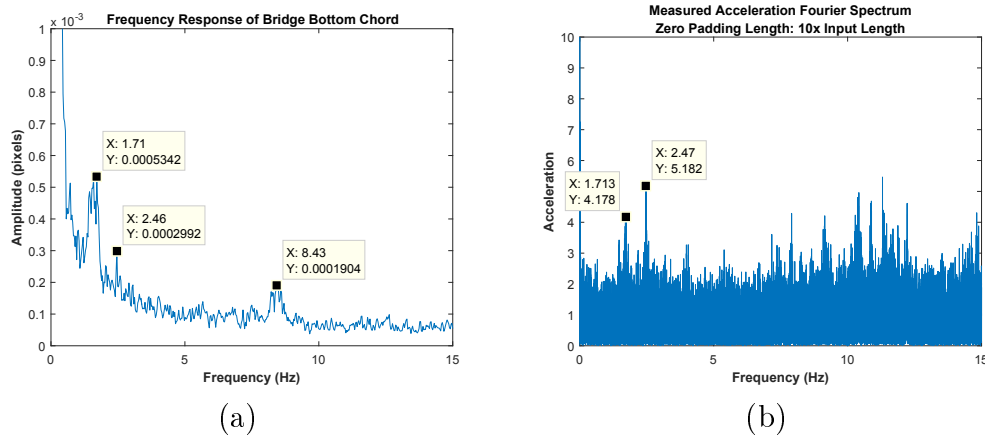
Figure 7-5: (a) Frequency spectrum of vertical motion of the bridge bottom chord as measured by the camera, (b) accelerometer measurement of the bridge at midspan on the sidewalk on August 27, 2015

Figure 7-6 shows measurements from the MEMS accelerometers mounted to the camera to monitor any camera motions. The frequency spectra for both the camera

pitch motions and the vertical translations are smooth, showing that the resonant peaks seen in the camera measurements aren't caused by camera motion and can be attributed to motion of the bridge in the video.
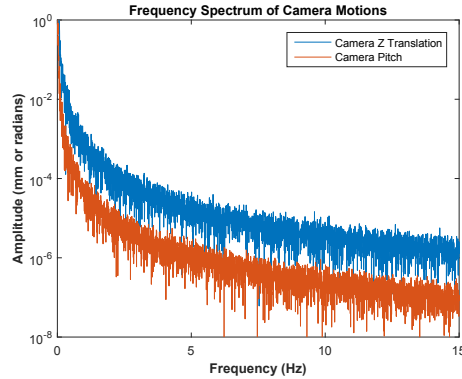
Figure 7-6: Frequency spectra from the camera mounted accelerometers measuring translation in the vertical (Z) direction and pitch

The source video over a small window during the lift span impact can be seen at URL: `https://youtu.be/hfYgbKEszqA`. The motion magnified video over the frequency range of the first resonant mode, 1.65 - 1.75 Hz, with a magnification factor of 400 can be seen at `https://youtu.be/Gd4CdFvYce4`. The motion seen in this video is a general vertical motion of the whole span in the frame of the video. For the second resonant mode, the motion magnified video for 2.4 - 2.6 Hz with a magnification factor of 400 can be seen at `https://youtu.be/-ioY2J9CrWc`. The motion seen is a torsional mode and the video shows a rotation or torsional motion of the bridge.

A full frame video of 500 frames around the lift span impact was processed for the October 8, 2015 measurement. The downsampled video frame and pixel masks are shown in Figure 7-7, the extracted average displacement signals shown in Figure 7-8, and the resulting modeshape phase images for three resonant modes at 1.44 Hz, 1.74 Hz, and 2.46 Hz, are shown in Figure 7-9. Specifically in Figure 7-9(e) and (f), the images for the the 2.46 Hz mode, we can clearly see that this mode is torsional as the far side and near side of the bridge move out of phase with each other.
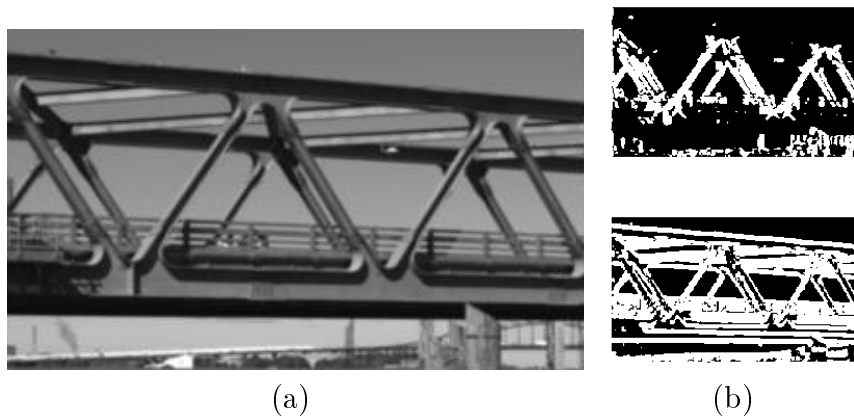
Figure 7-7: Processing of the full frame video for the October 8, 2015 measurement (a) Downsampled frame and (b) top: horizontal and bottom: vertical masks for pixels with valid displacements
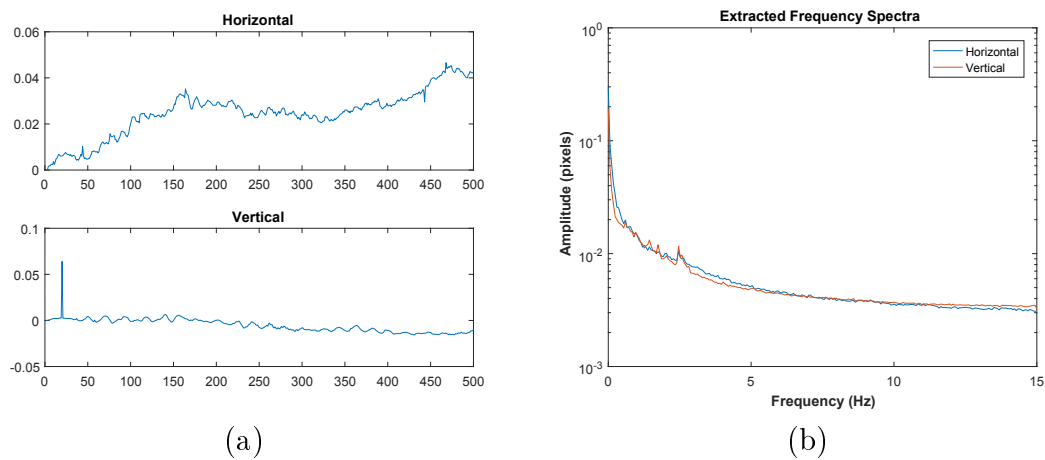
Figure 7-8: Average displacement signals for the full frame video for the October 8, 2015 measurement for horizontal and vertical directions in the (a) time domain and (b) frequency domain
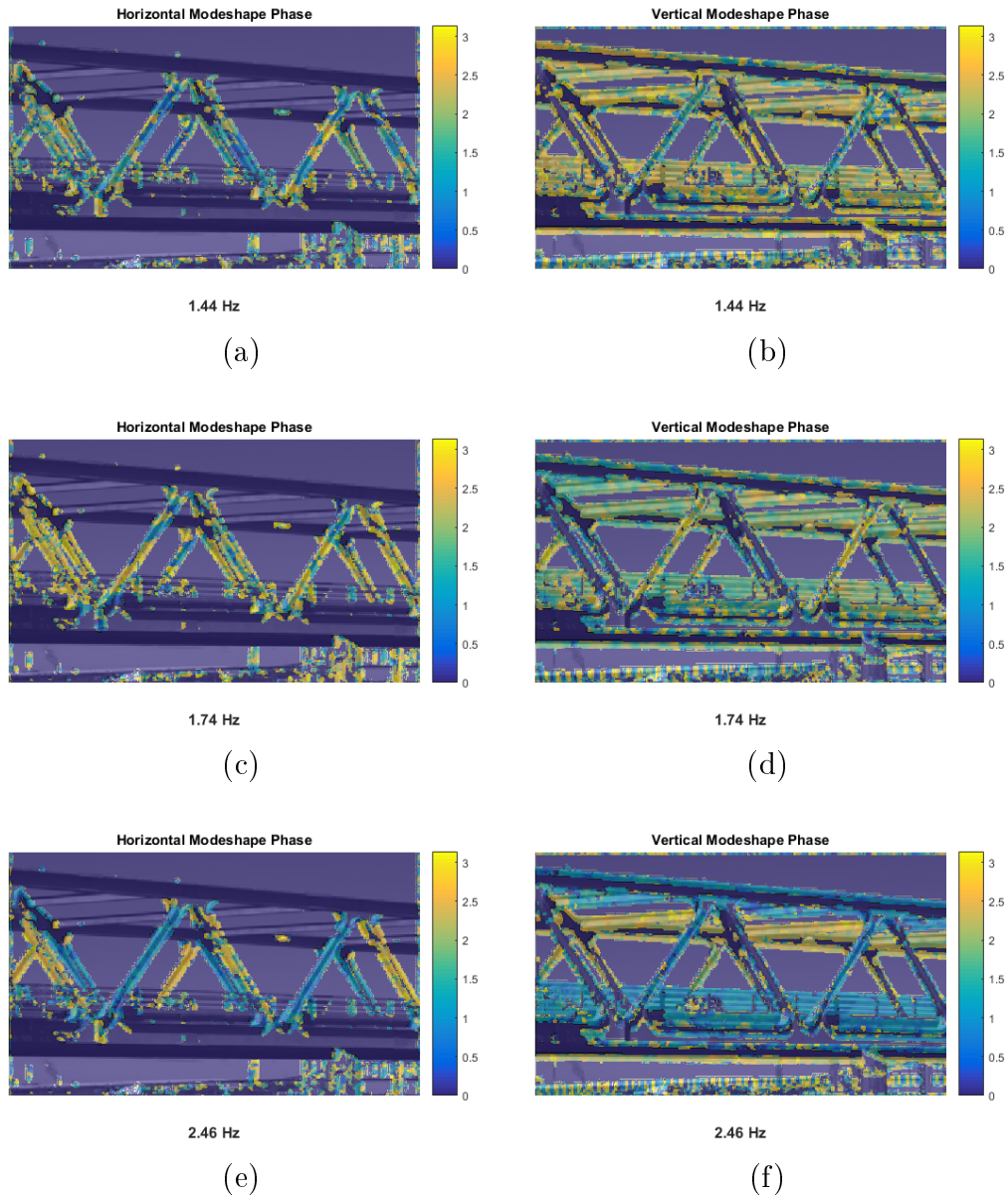
Figure 7-9: Modeshape phase images for the December 12, 2015 measurement, horizontal and vertical displacement directions for (a-b) 1.44 Hz, (c-d) 1.74 Hz, (e-f) 2.46 Hz

7.2.3   December 14, 2015 Measurement

The December 14, 2015 measurement was made on a day with weather much more miserable than the October 8, 2015 measurement. In general the weather was colder, and there was a light drizzle. A picture of the view from the camera location is shown in Figure 7-10(a), and the camera is covered with a plastic bag to protect it from the drizzle. During this measurement, two accelerometers and two strain gauges were placed on the bridge to concurrently measure the response due to lift events. Figure 7-10(b) shows the locations of the two accelerometers, A2192 and A2193, and the two strain gauges, B4821 and B4824, on a screenshot directly from the recorded video. Time synchronization was accomplished approximately by using a GPS application on cell phones carried by the experimenters at both locations, as related to the system time on the camera measurement system laptop, and the accelerometer data acquisition laptop. Small manual adjustments were made to get the two time series to line up exactly.

(a)   (b)

Figure 7-10: (a) Picture of the view from the camera for the December 14, 2015 measurement with (b) screenshot from the recorded video with comments as to the locations of the accelerometers and strain gauges on the bridge

Lift impact response

A lift was conducted at approximately 10:30am for the purposes of making a measurement of the bridge's response. Figure 7-11 shows the resulting vertical direction

measurements from the accelerometer A2192 as compared with the video measurement in a region of interest cropped around the location of the sensor. In these plots the accelerometer measurement is double integrated to displacement. We use the fact that the web of the top chord of the bridge is 24 inches (0.61 meters) tall, corresponding to 30 pixels in the frame of the video, to convert the measured pixel displacements to units of inches. For the comparison with the A2193 accelerometer shown in Figure 7-12, the scaling factor is 36 inches (0.91 meters) to 44 pixels. In both these comparisons, the resonant frequencies present in the frequency spectra, 1.267 Hz, 1.68 Hz, and 2.52 Hz line up right on top of each for both the accelerometer and the camera. The time series measurement also compare quite well. This measurement shows that the camera measures the same information as an accelerometer from a distance of 260 feet (80 meters).

(a)   (b)
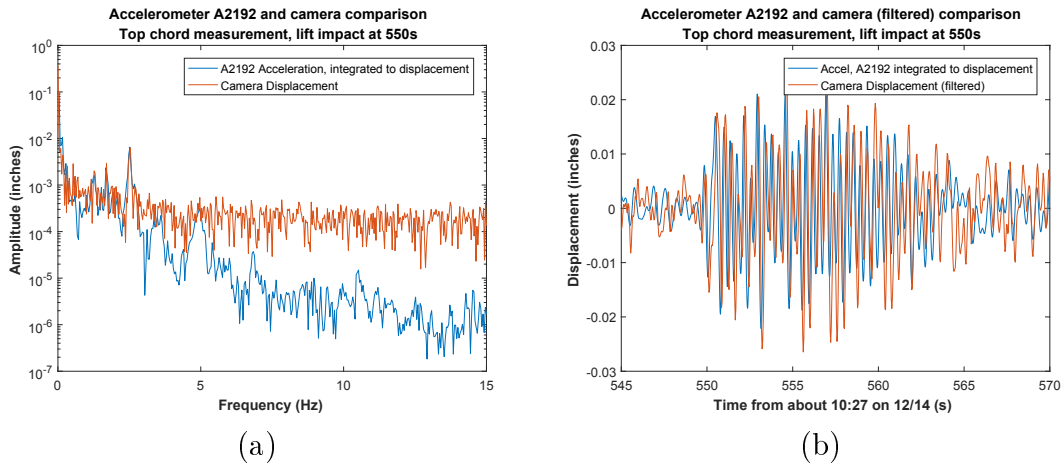
Figure 7-11: Comparison of measurements from the camera in the region around the A2192 accelerometer, and a measurement from the A2192 accelerometer integrated to displacement in the (a) time domain and (b) frequency domain

A full frame video of 1000 frames around the lift span impact was processed for the December 14, 2015 measurement. The downsampled video frame and pixel masks are shown in Figure 7-13, the extracted average displacement signals shown in Figure 7-14, and the resulting modeshape phase images for three resonant modes at 1.26 Hz, 1.71 Hz, and 2.52 Hz, are shown in Figure 7-15.

(a)   (b)
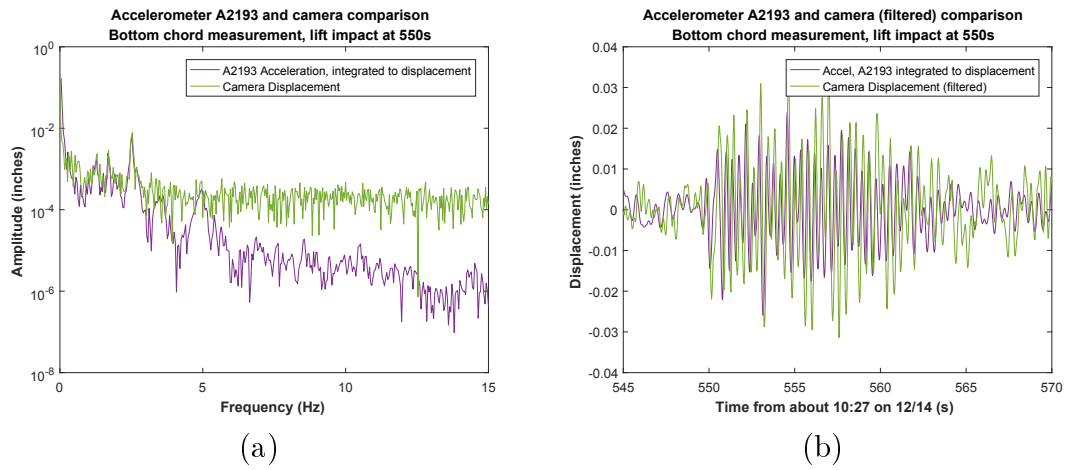
Figure 7-12: Comparison of measurements from the camera in the region around the A2193 accelerometer, and a measurement from the A2193 accelerometer integrated to displacement in the (a) time domain and (b) frequency domain

(a)   (b)
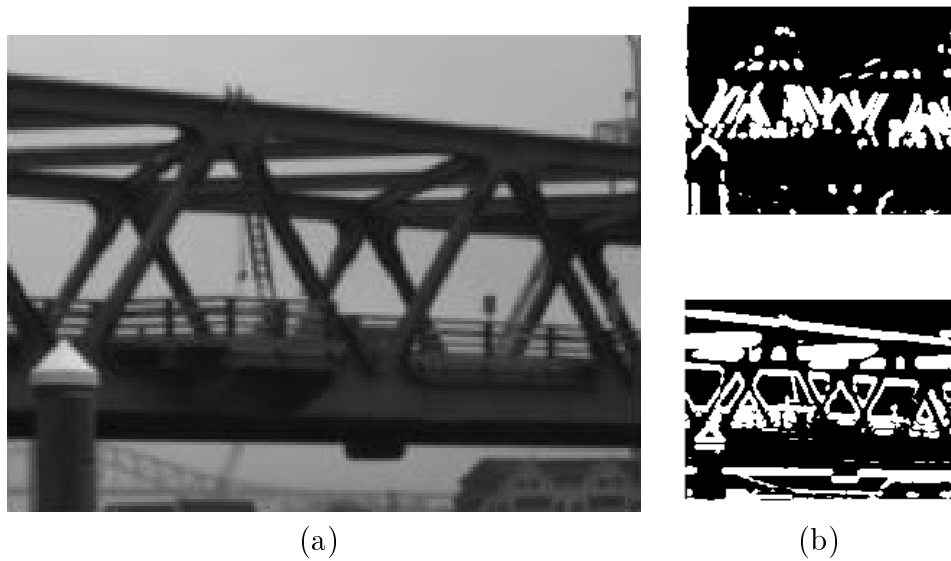
Figure 7-13: Processing of the full frame video for the December 12, 2015 measurement (a) Downsampled frame and (b) top: horizontal and bottom: vertical masks for pixels with valid displacements

(a)   (b)
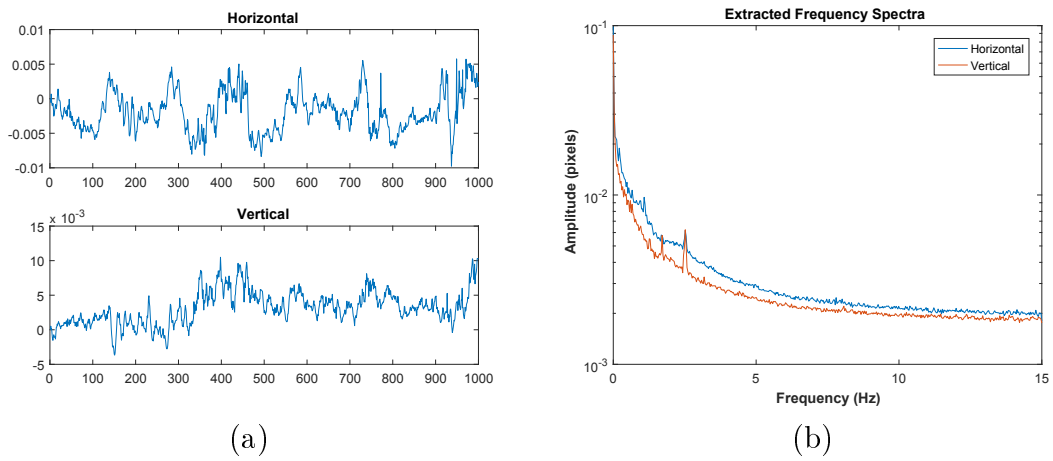
Figure 7-14: Average displacement signals for the full frame video for the December 12, 2015 measurement for horizontal and vertical directions in the (a) time domain and (b) frequency domain
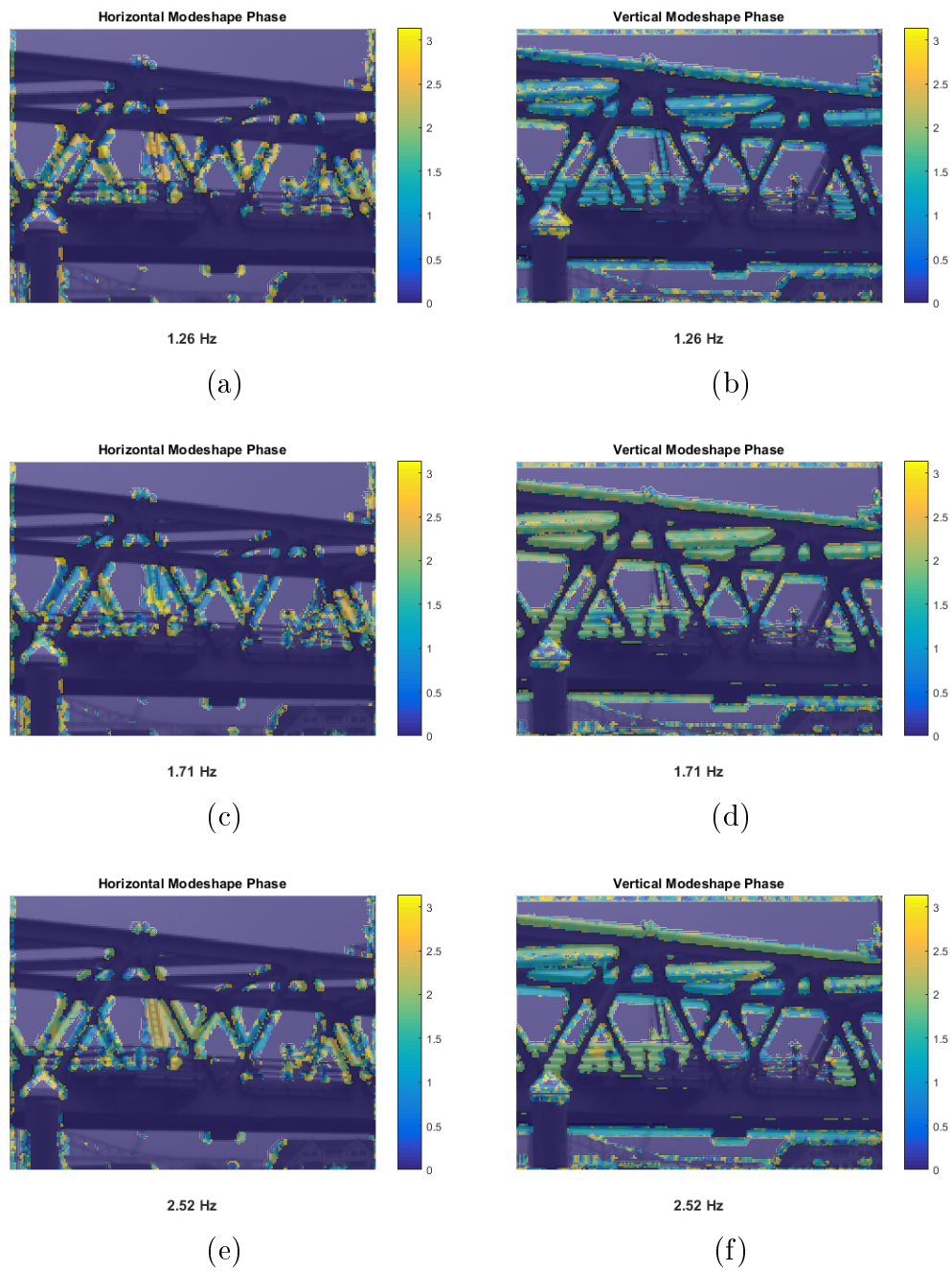
Figure 7-15: Modeshape phase images for the December 12, 2015 measurement, horizontal and vertical displacement directions for (a-b) 1.26 Hz, (c-d) 1.71 Hz, (e-f) 2.52 Hz

Traffic induced loading response

Before the lift at 10:30 am, there was normal traffic flowing across the bridge. A video recorded starting at approximately 10:21:03 am, begins with a concrete mixer truck at midspan going northbound. Later in the video other cars and smaller pick-up trucks also cross the bridge. To measure the displacement of the bridge, a region of interest is made on the bottom chord at the midspan of the bridge, and another region of interest is selected below that of a stationary object in the background as a reference for the purpose of removing camera motion. A screenshot from the start of the video with the regions of interest outlined in red is shown in Figure 7-16.

Figure 7-16: Screenshot from the start of the 10:21:03 am video showing a concrete mixer truck and regions of interest on the bottom chord and an object in the background

Vertical displacements were extracted from these regions of interest and are shown in blue for the signal from the bottom chord and red for the stationary reference as shown in Figure 7-17. To compensate for motion of the camera, the stationary reference signal was subtracted from the signal from the bottom chord to the signal in yellow.

We can offset the motion compensated by 0.208 pixels to give the displacement in pixels of the bridge from an estimated unloaded state. This amount is determined by approximately lining up the highest vertical displacements seen with a zero line. To translate from units of pixels to inches displacement at the bridge, we use the fact
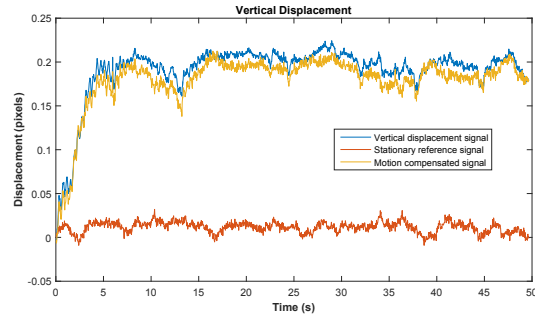
Figure 7-17: Vertical displacements extracted from the video measuring traffic induced loading with a signal from the bottom chord in blue, signal from a background stationary object in red, and the motion compensated signal from the bottom chord in yellow

that the web of the bottom chord is 44 pixels in the video and 36 inches (0.91 meters) tall. The resulting measured vertical displacement of the bottom chord of the bridge is shown in Figure 7-18.
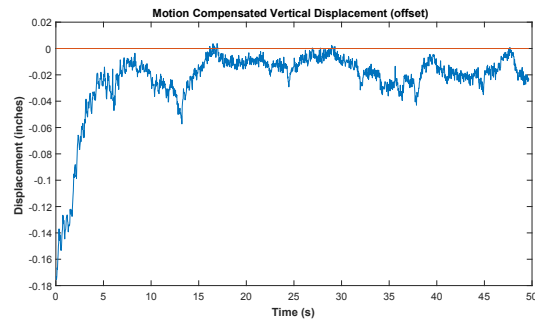
Figure 7-18: Vertical displacement of the bottom chord in inches in blue, approximately offset to represent the displacement from an unloaded state, shown in red

7.3   Finite Element Modeling

This section details the creation of a finite element model for the Portsmouth bridge for comparison with the resonant frequencies and mode shapes measured by the camera from the bridge.

7.3.1 Model Construction

The plans for the bridge were provided by the UNH team and are the drawings for the structure as built [115]. The goal was to create a model of the bridge to work out the dynamics as the lift span is lowered and vibrates the bridge at the end of being lowered. From looking at the plans, the central lift span moves on rails on the two fixed spans, and are otherwise not structurally connected. Since the south New Hampshire fixed span is the one measured by the camera, we choose to model only the south fixed truss and lift tower. The relevant drawings for the lift tower and the truss are shown in Figures 7-19 and 7-20.
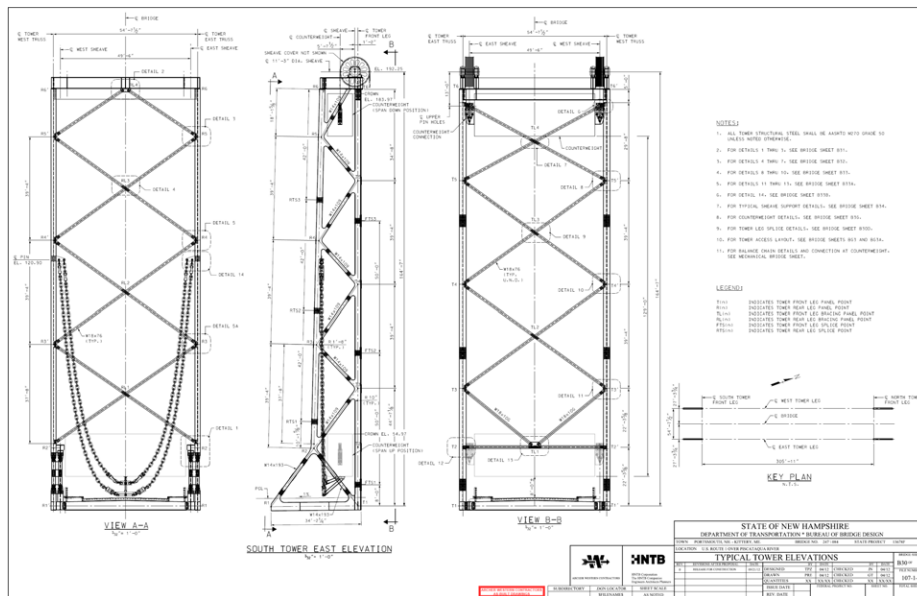
Figure 7-19: Plans for the south tower of the bridge [115]

The majority of the elements used in the bridge are I beams of various different sections, so for simplicity we choose to model the bridge using beam elements. The only non I beam element is the portal beam at the end of the fixed truss, and this is approximated with a rectangular section beam with the same dimensions. The I beams are made from AASHTO M270 Grade 50 structural steel which is assumed to have a Young's Modulus of 29,700 ksi, a Poisson's ratio of 0.29, and a density of 0.284 lb/in^3.

201
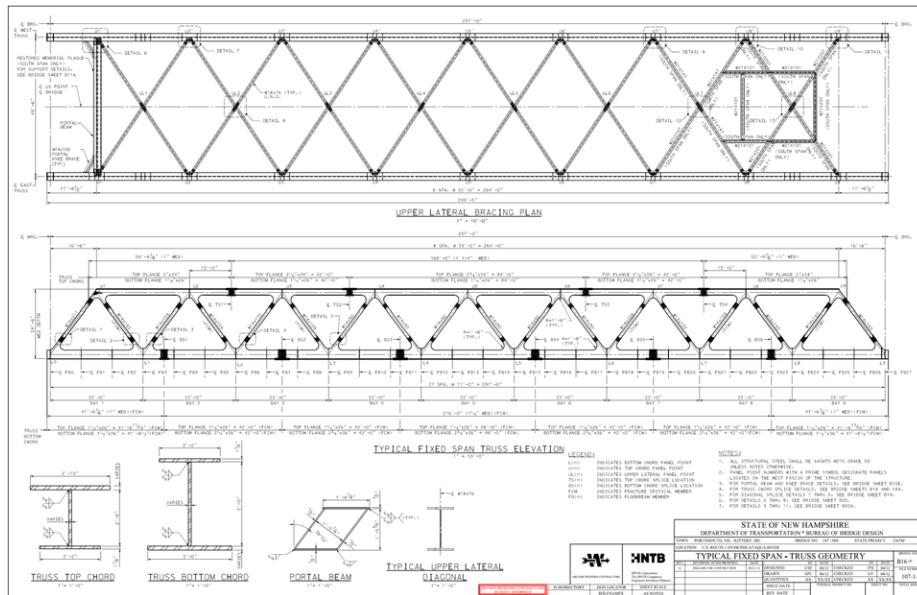
Figure 7-20: Plans for the fixed span of the truss [115]

Figure 7-21 shows a cross section of the lift span at the south tower. This both shows the dimensions of the concrete and asphalt bridge deck and how the lift span could potentially provide an impulse to the fixed span. There is a centering device for the lift span which centers the span as it is lowered. If there is slight misalignment, the lift span may end up impacting the support pier as the centering device is engaged and the lift span is lowered into position.

Assumptions and Simplifications

The concrete and asphalt deck is modeled as a rectangular cross-section beam, of 11.75" thick and width of 32', with the material properties of concrete, a density of 0.0868 lb/in^3, Young's Modulus of 4000 ksi, and Poisson's ratio of 0.2. This beam approximating the bridge deck is connected to the floor beams that span between the trusses of the lift span. This may result in some local modes, however the general bridge behavior should remain well approximated.

The counterweight mass is accounted for by making the girder at the top of the tower heavy enough to match the weight of the counterweight, as it is fully supported by the east and west sheave as seen in Figure 7-19. This effectively applies the mass

202
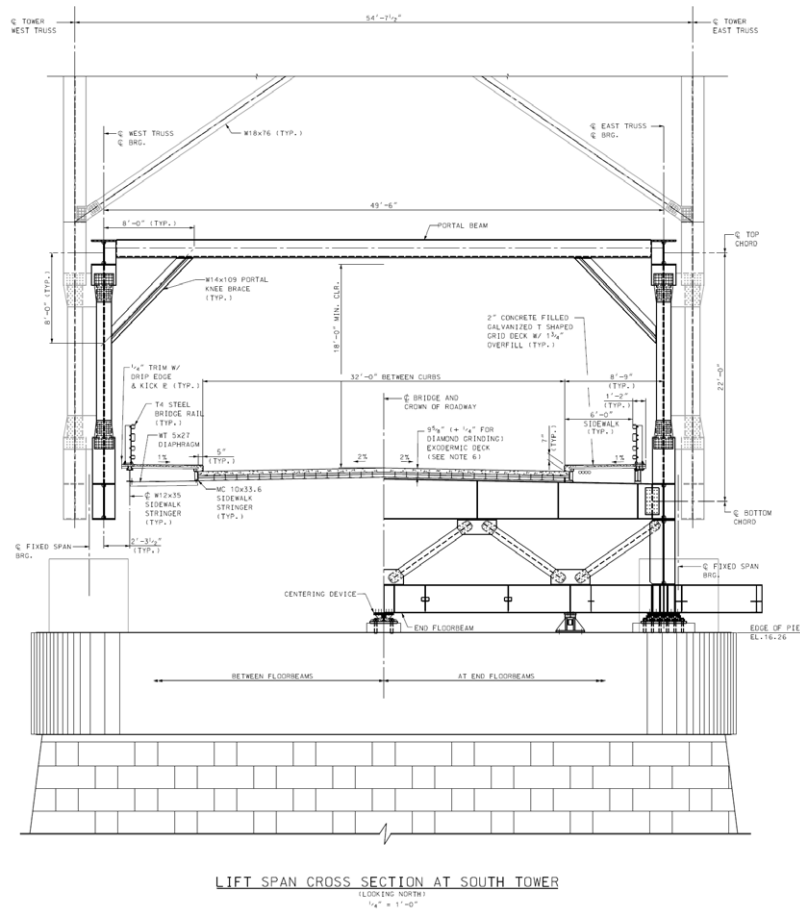
Figure 7-21: Cross section of the bridge at the south tower [115]

to those nodes in the model, however it neither accounts for the swinging of the counterweight on it's cables, nor the balance chains, which are assumed to be small effects.

As seen in the plans in Figures 7-19 and 7-20, and in the pictures in Figure 7-22, this bridge has connections of a special design where the joint is a monolithic element as part of the bottom or top chord, and the typical gusset plates which connect the truss elements are moved onto the elements. The gusset plates connect both the web and the flange parts of the I beams on both sides, and thus they transmit moment. To account for this connection, we use an approximate model where elements in the bridge as assumed to connect at a single point, rather than modeling the full joint as it would be far too computationally intensive as the joint would need to be fully

203

modeled using solid elements. We assume a frame behavior at all the joints in the bridge which transmits moments between the beam elements.

<center>(a)</center>

<center>(b)</center>

Figure 7-22: Pictures of the special connections on the bridge with a monolithic joint and the gusset plates moved onto the elements with pictures from the (a) bridge interior showing diagonal supports between the trusses and (b) bridge exterior

The connection between the south tower and the fixed truss span is shown, highlighted in red in Figure 7-23 and in the plans in Figure 7-21. They are tied together by the gusset plates that connect the doubled elements for that triangle of the truss shared by both the fixed truss and tower. In the model we approximate this with a tie condition between the corresponding nodes in the model so that they are constrained to translate and rotate together.

There are details in the plans such as the bridge tender house, stairways, sidewalks, and others that are neglected for modeling simplicity. The main elements of the lift tower, fixed truss span, bridge deck, and counterweight are included in the model.

The boundary conditions for the bridge are pinned at the two nodes for the lift tower and fixed truss side of the bridge, and a roller at the two nodes on the other end, as the bearings in the plans show a fixed and an expansion bearing at these locations respectively. Damping is not included in the model as that would require much more model tuning to get that value correct due to the multiple sources of damping in the actual bridge (e.g. connection, material, support, etc.).

Figure 7-23: Picture of the south tower and fixed truss with the truss-tower connection outlined in red showing gusset plates tying the two together

Model

The resulting model uses 1790 linear beam elements, with 1637 nodes and a total of 9776 unrestrained degrees of freedom, shown in Figure 7-24. A high density element mesh was used to ensure that the model would not have any simulator errors, as memory or computational problem was not expected to be an issue with linear beam elements. This model will be used as a basis for comparison with the camera measurements by conducting a time history analysis to simulate the impact from the lift span, and an eigenvalue frequency analysis to determine the mode shapes of the resonant frequencies that participate in the response from the time history analysis. This model is derived directly only from the information from the acquired plans and pictures of the bridge. The bridge behavior from the camera or accelerometer

measurements of the real-world bridge is not used to tweak the model in any way.
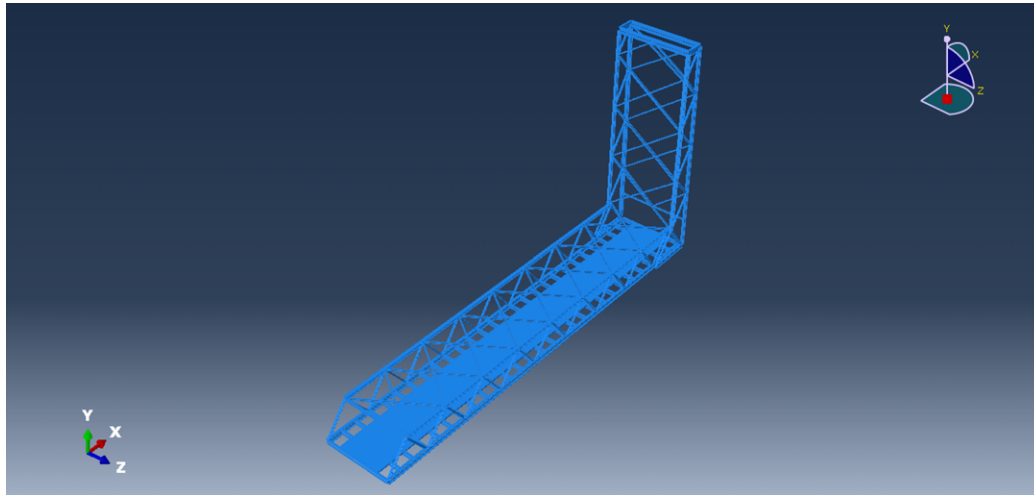
(a)
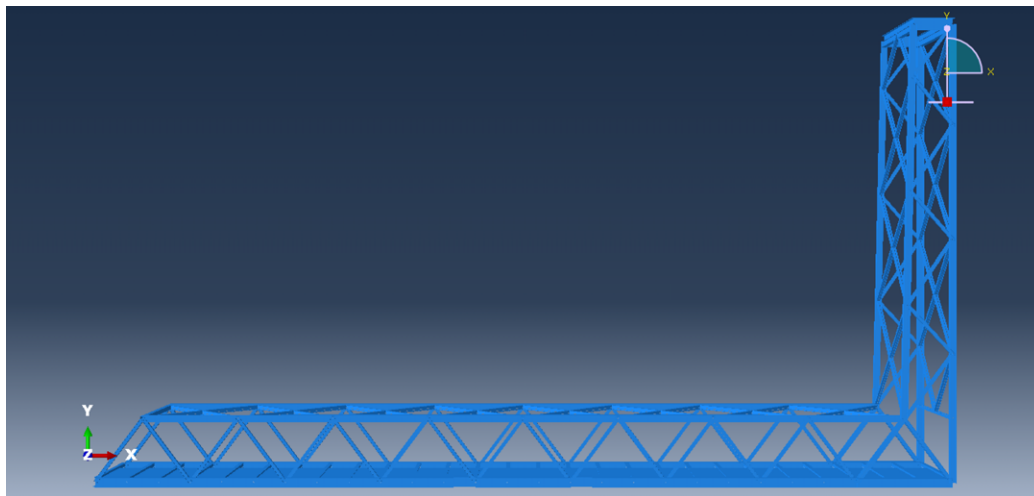
(b)

Figure 7-24: Modeled bridge from a (a) front isometric view and (b) side view

7.3.2 Time History Analysis

To simulate the impact of the lift span on the support holding the tower and truss side of the fixed span of the bridge, the boundary condition is displaced vertically downwards by 1 inch (25.4 mm) for 0.1 seconds at the beginning of the simulation. An implicit time history analysis was carried out over 20 seconds with a fixed time step of 0.01 seconds. Since we expect the majority of the response to be under 10

206

Hz, as has been measured from the bridge, the time step of 0.01 seconds, a factor of 10 shorter than the period of 0.1 seconds for 10 Hz, is sufficient. Figure 7-25(a) shows the first frame of the time history, showing the displacement of the support at the tower end. Figure 7-26(b) shows the 59th frame 0.59 seconds into the analysis showing the vibrations having propagated through the bridge and tower.
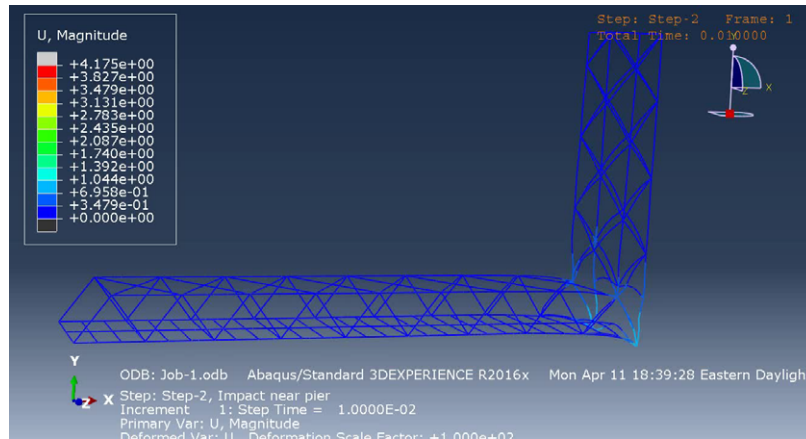
(a)
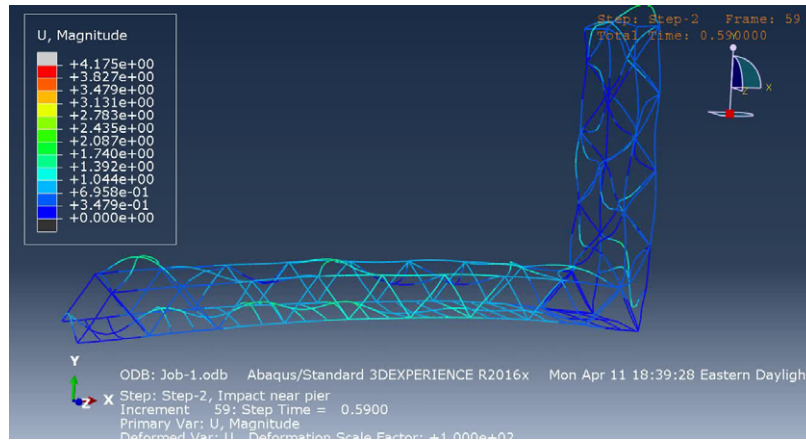
(b)

Figure 7-25: Frames from the time history analysis of the bridge showing the displacement magnitude over a deformed model of the bridge with a scale factor of 100 at (a) the first frame showing the impact at the beginning of the time history and (b) the 59th frame or 0.59 seconds into the analysis

To match something similar to the camera measurements, the vertical displacement of the midspan node on the top chord is extracted from the time history analysis. Figure 7-26 shows the time history and frequency response of that node. The domi-

nant response at that node is at 1.6 Hz. The first four modes with significant response are 1.25 Hz, 1.6 Hz, 1.9 Hz, and 2.9 Hz. In order to see the mode shapes corresponding to these frequencies, an eigenvalue frequency analysis is carried out.

(a)   (b)
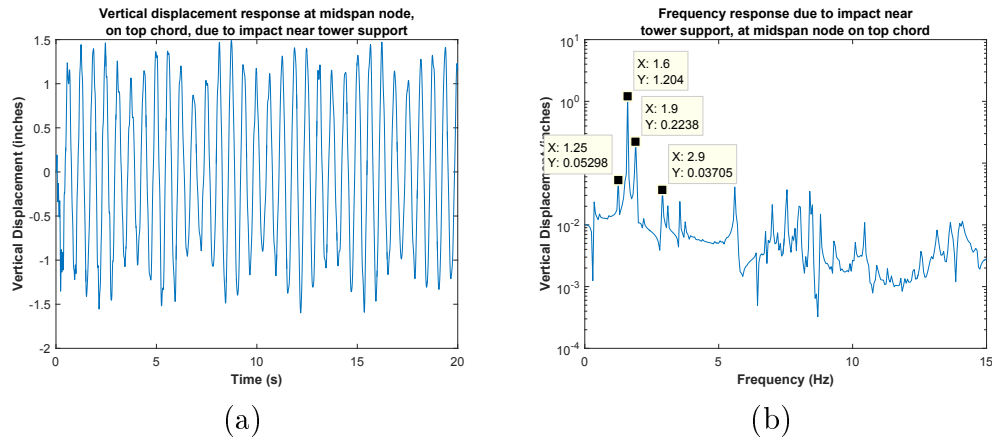
Figure 7-26: Vertical displacement of the bridge at the midspan node on the top chord for the time history analysis, (a) time series and (b) frequency spectrum

The same time history analysis was also carried out on a model with damping with all other parameters being equal, with the resulting video shown at `https://youtu.be/lqLzNJjR-gM`. Rayleigh damping was added to the model with the parameters of $\alpha = 0$ and $\beta = 0.001$ for the steel, and $\alpha = 0$ and $\beta = 0.003$ for the concrete deck. This results in approximately 0.4% damping for the structure at 1 Hz. The resulting displacements for the midspan node on the top chord is shown in Figure 7-27.
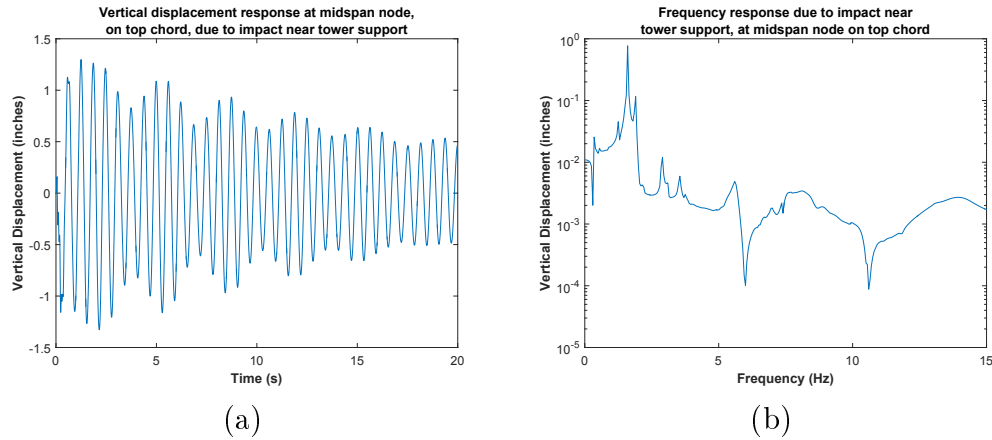
Figure 7-27: Vertical displacement of the bridge at the midspan node on the top chord for the time history analysis with damping, (a) time series and (b) frequency spectrum

7.3.3 Eigenvalue Frequency Analysis

The first 30 modes from the eigenvalue frequency analysis for the bridge model are given in Table 7.1. The first four modes with significant response in the time history analysis correspond to mode 7 at 1.248 Hz, mode 10 at 1.598 Hz, mode 12 at 1.890 Hz, and mode 20 at 2.892 Hz. The mode shapes for these four modes are shown in Figure 7-28. Modes 7, 10, and 12 involve some form of tower motion and a first bending mode of the truss fixed span. Mode 20 is a torsional mode of the bridge, where the truss fixed span is rotating and torsioning about its longitudinal axis.

Table 7.1: Frequencies of the first 30 modes from the eigenvalue frequency analysis for the bridge model

Mode	Frequency (Hz)	Mode	Frequency (Hz)	Mode	Frequency (Hz)
1	0.327	11	1.704	21	3.111
2	0.377	12	1.890	22	3.558
3	0.533	13	1.949	23	3.738
4	0.745	14	1.983	24	3.756
5	0.880	15	2.147	25	3.925
6	1.094	16	2.188	26	4.031
7	1.248	17	2.292	27	4.140
8	1.416	18	2.380	28	4.257
9	1.524	19	2.833	29	4.413
10	1.598	20	2.892	30	4.781
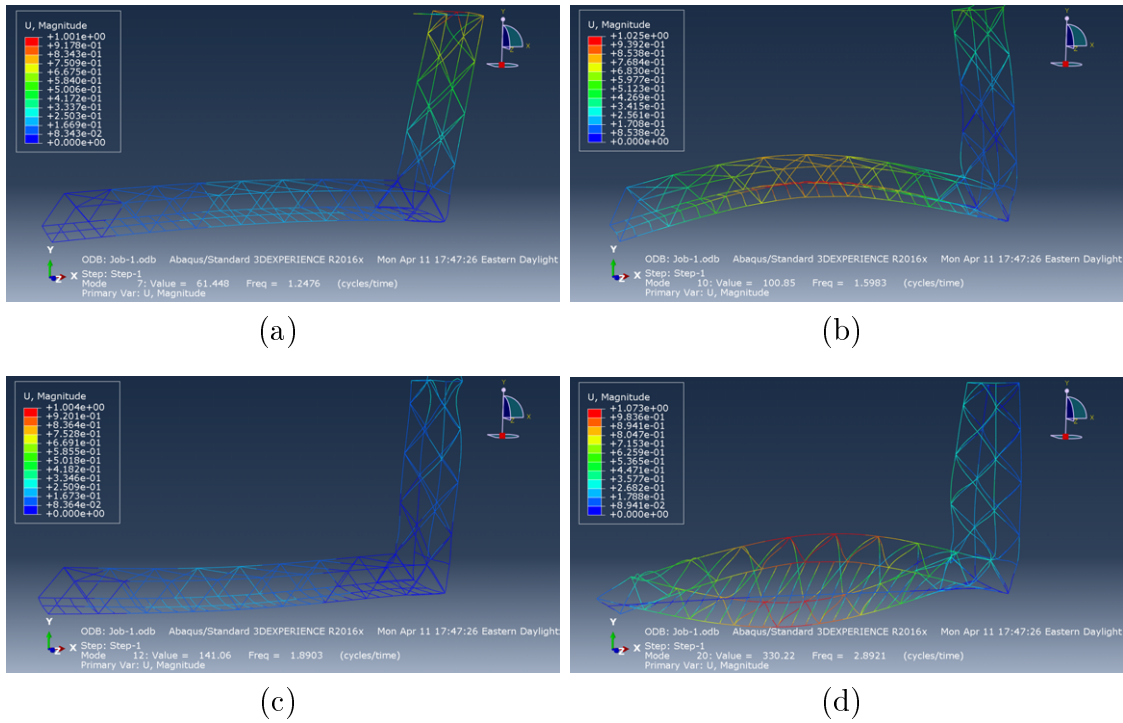
Figure 7-28: Mode shapes for the bridge model at the first four frequencies with significant response in the time history analysis, (a) mode 7, 1.2476 Hz, (b) mode 10, 1.5983 Hz, (c) mode 12, 1.8903 Hz, and (d) mode 20, 2.8921 Hz (torsion)

7.3.4   Static Load Analysis

A static load analysis was also conducted to determine the approximate deflection of the bridge at midspan due to traffic driving across the bridge. Point loads of 250 lbf (1112 N) were applied to four nodes at the center on the northbound side of the fixed truss for a total loading of 1000 lbf (4448 N), as seen in Figure 7-29. This approximates a 1000 lb (454 kg) load on the northbound lane of the bridge at midspan.

The result from the static load analysis is shown in Figure 7-30, showing the vertical displacements on a deformed model. The average displacement at the two nodes on the centerline of the bridge was -0.00535 inches (-0.136 mm), and the average displacement for the two nodes on the near side of the bridge was -0.00506 inches (-0.129 mm). We can use this information to correlate displacements measured in the real world bridge to approximate the imposed loads.
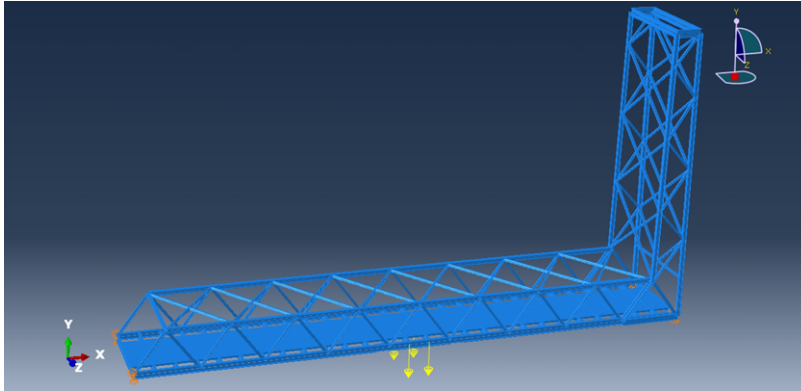
Figure 7-29: Load applied at four central nodes on the northbound side, total loading of 1000 lbf
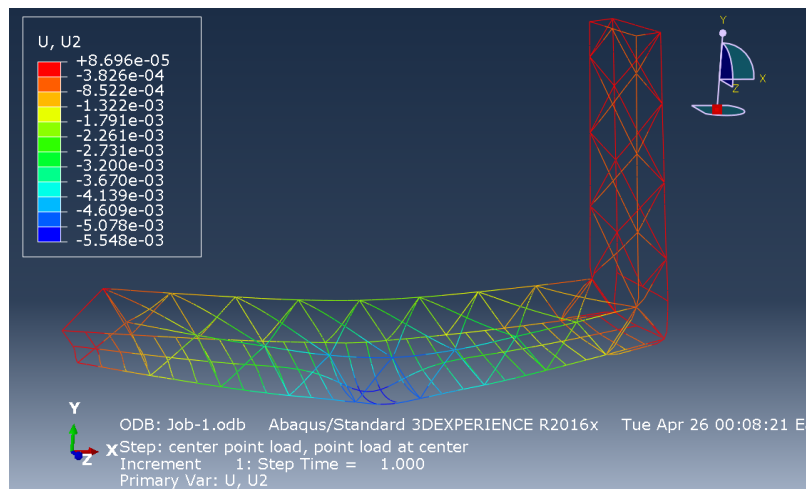
Figure 7-30: Static load analysis result showing vertical displacement on the deformed model

7.4 Comparison and Analysis of Modeling and Measurements

7.4.1 Lift Impact Response Comparison

To compare the results from the finite element model and camera measurements of the bridge, we examine the responses to a lift span impact. Figure 7-31 shows the frequency spectra for the finite element model and both camera measurements. The frequency peaks almost directly compare between the model and the December 14,

2015 measurement, 1.25 Hz and 1.267 Hz, 1.6 Hz and 1.7 Hz, and 2.9 Hz and 2.533 Hz, respectively. In the October 8, 2015 measurement, the fundamental mode of vibration is a wide peak from 1.23 Hz to 1.9 Hz, which corresponds to the series of peaks in the model response at 1.25 Hz, 1.6 Hz, and 1.9 Hz. Frequency wise, the difference between the frequencies of the torsional modes is the largest between the real bridge and the model, however note that the mode was only constructed from the plans and pictures of the bridge, and has not been massaged to match any of the measurements.
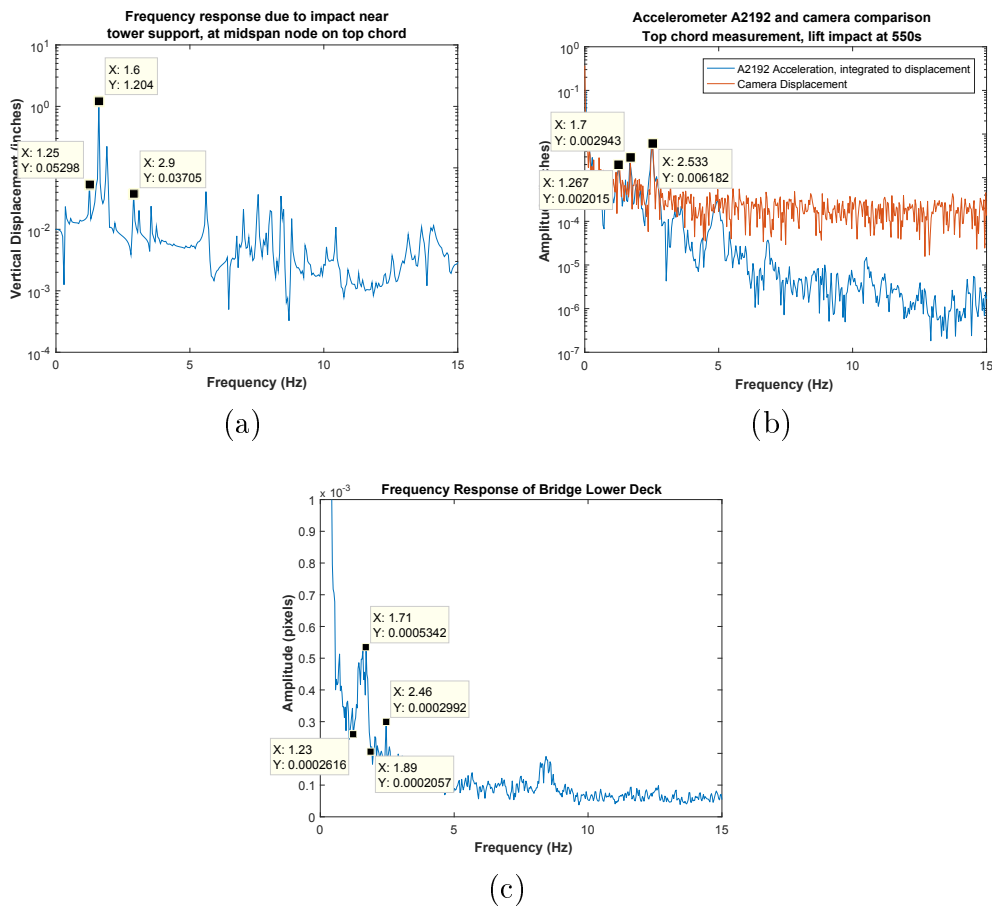
Figure 7-31: Comparison of frequency spectra from the (a) finite element model, (b) top node of the Portsmouth bridge, December 14, 2015 camera measurement, and (c) bottom chord of the Portsmouth bridge, October 8, 2015 camera measurement

Qualitatively we can look at the time domain responses for the model at the midspan node on the top chord and compare it to the December 14, 2015 measure-

ment by the camera at the location of the A2192 accelerometer, which is in the same location. The two responses are shown side by side in Figure 7-32, and they look similar, noting the lack of damping in the finite element model. The peak to peak displacement of approximately 0.03 pixels in the December 14, 2015 measurement corresponds to approximately 0.025 inches (0.635 mm) of displacement. In the simulation the support is displaced by 1 inch (25.4 mm), and the node moves with an peak to peak displacement of 3 inches (76.2 mm). Since the analysis is completely linear, we can scale the values down and estimate that a support displacement of 0.0083 inches (0.21 mm) would cause a peak to peak displacement of 0.025 inches (0.635 mm). This may suggest that the impact of the lift span effectively displaces the support of the bridge by that amount to cause the vibration in the fixed truss.

(a)   (b)
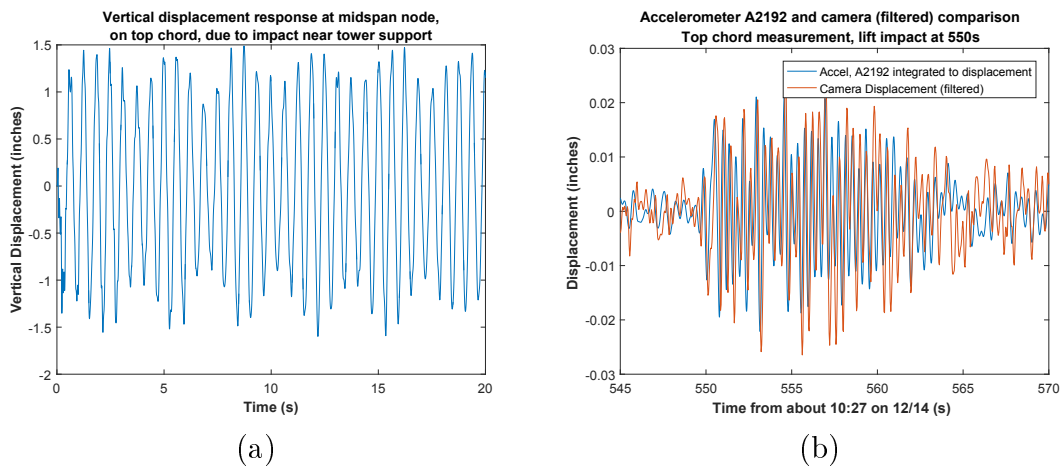
Figure 7-32: Comparison of responses due to a lift span impact for the (a) finite element model (b) as measured on December 14, 2015 by both a camera from long-range and an accelerometer on the bridge

7.4.2 Mode Shape Comparison

We can compare representations for the torsional mode shape at 2.5 Hz in the real bridge, and 2.9 Hz in the finite element model. The representations look very similar and the animation of the model matches the apparent motion in the motion magnified video of the bridge at the frequency of the torsional mode. The animation of the model

213

at 1.6 Hz where the fixed truss has a vertical bending mode also matches the motion magnified video of the bridge at 1.7 Hz. Comparison videos are given in Appendix D.5.
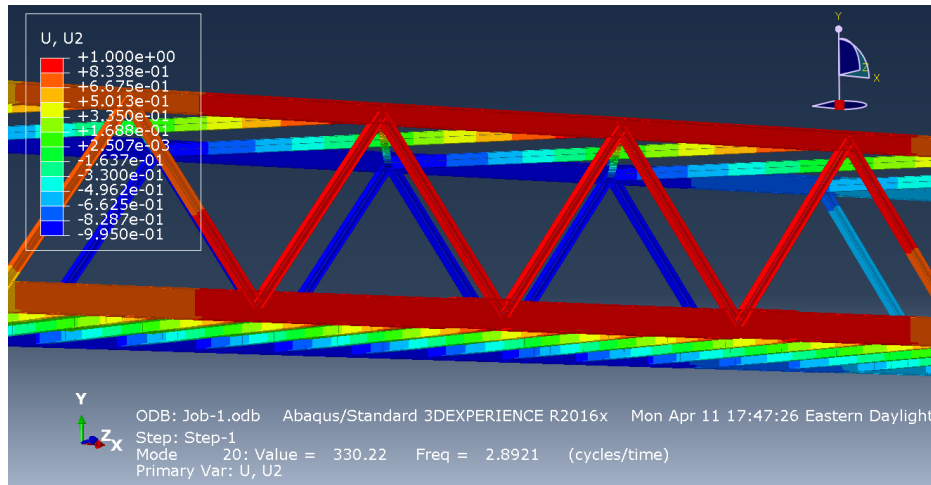
(a)
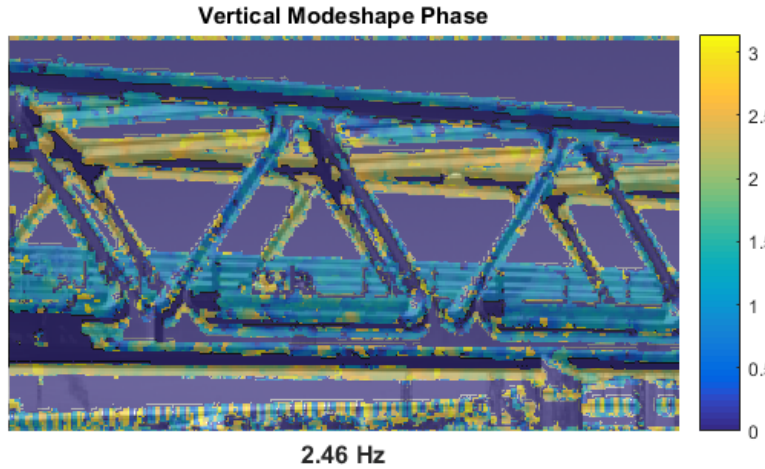
(b)

Figure 7-33: Comparison of representations of the modeshape for the torsional mode (a) image of the vertical direction modeshape phase at 2.46 Hz from the October 8, 2015 measurement, and (b) vertical component of the modeshape at 2.89 Hz for mode 20 of the finite element model

7.4.3 Traffic Induced Loading Measurement

The measurement of the vertical displacement of the bridge under traffic induced loading can be combined with the static load analysis from the model to estimate the weight of the traffic load on the bridge. The static load analysis tells us that a 1000 lbf (4448 N) at the midspan of the bridge on the northbound side will displace the bridge at the center vertically by approximately 0.005 inches (0.127 mm). Given that for a simple beam with a concentrated load at the center, the maximum displacement at the center will vary directly proportionally with the applied load, we make the assumption that we can estimate the load on the bridge by taking the displacement measured by the camera and scaling that by the 1000 lbf (4448 N) per 0.005 inches (0.127 mm) of displacement. We combine the raw video and the data from the measurement scaled to give an estimated load at the midspan of the bridge to see the loadings of certain vehicles as they drive across the bridge. Figures 7-34 through 7-38 show vehicles and estimated loads at different times in the video. The measurement gives plausible loads on the bridge for the weights of the vehicles seen in the video.
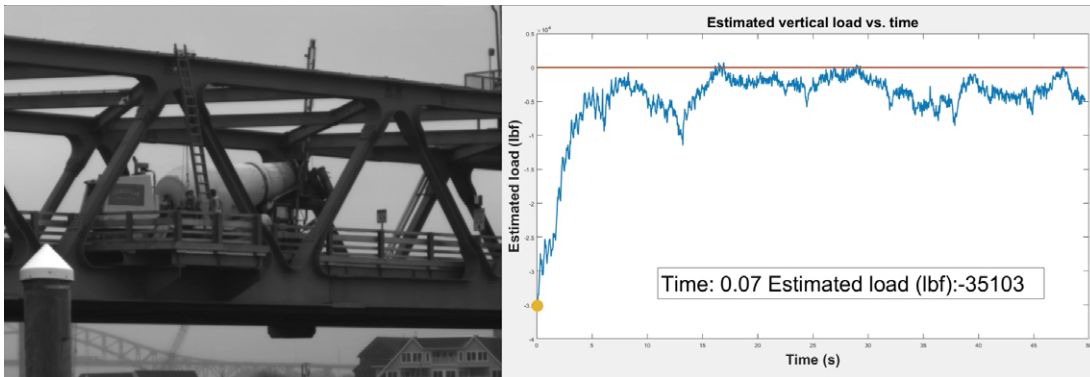
Figure 7-34: Concrete truck at 0.03s in the video and estimated load of 35,103 lbf (156,146 N)
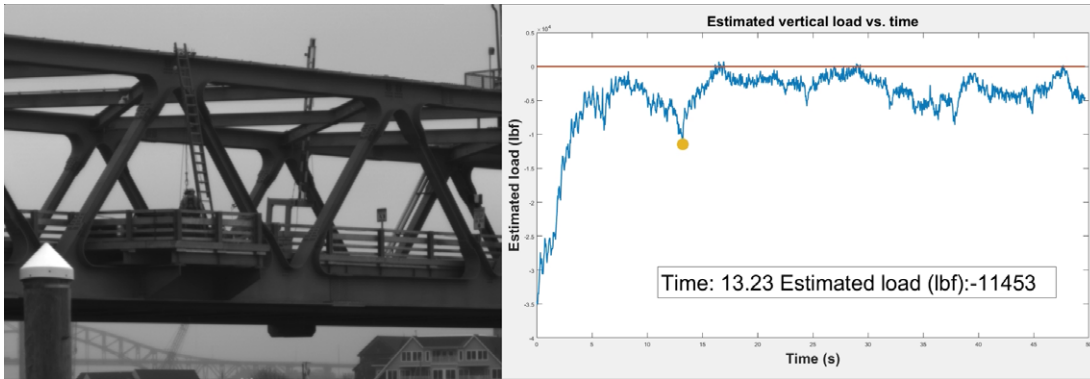
Figure 7-35: Heavy work truck at 13.23s in the video and estimated load of 11,453 lbf (50,945 N)
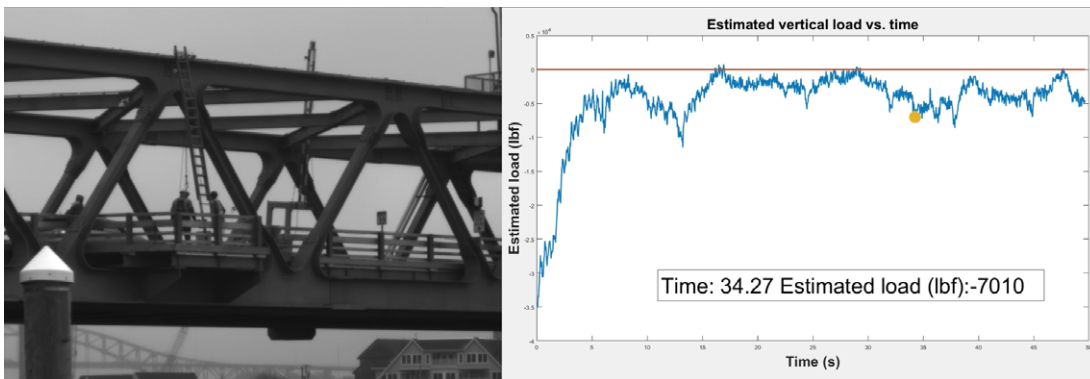
Figure 7-36: Mazda CX-5 (or similar) at 34.27s in the video and estimated load of 7,010 lbf (31,182 N)
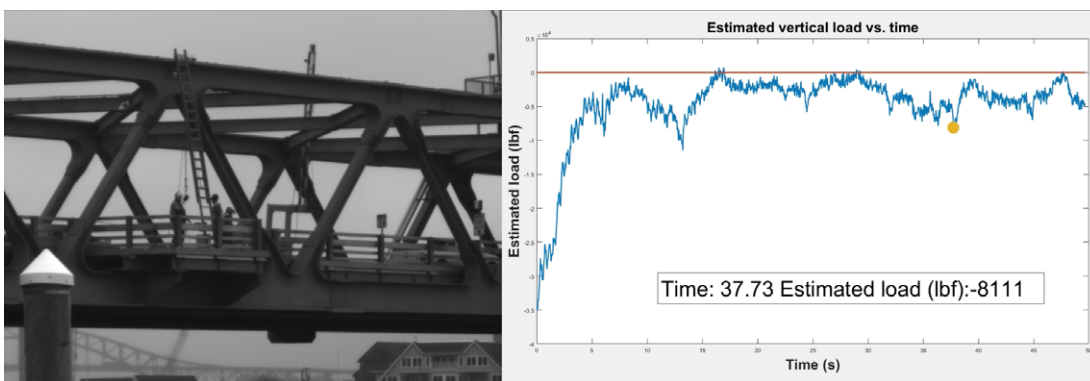
Figure 7-37: Sedan and truck nearby at 37.73s in the video and estimated equivalent load of 8,111 lbf (36,080 N)
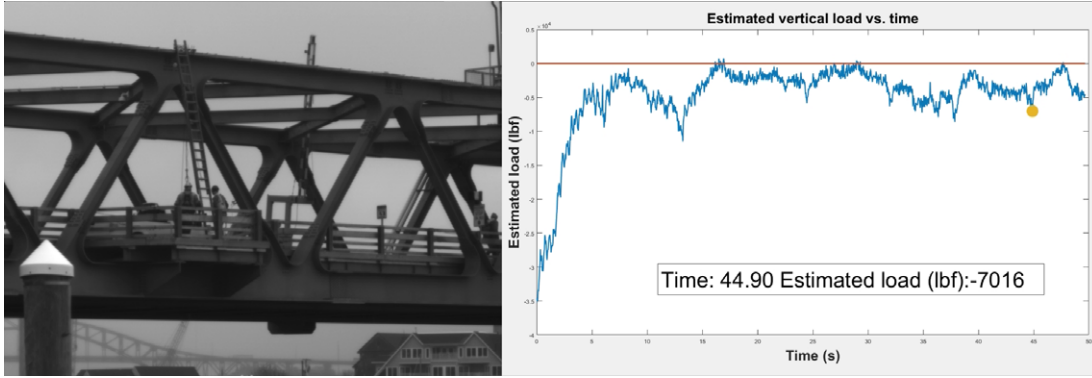
Figure 7-38: Pickup truck at 44.90s in the video and estimated load of 7,016 lbf (31,209 N)

7.5 Chapter Summary

In this chapter we have done a comprehensive study of the WWI Memorial Bridge in Portsmouth, NH involving field measurements with a camera, a finite element model constructed directly from the plans, and a comparison and synthesis of information from both the measurement and model data. The vertical-lift bridge provides an excitation at the end of every lift that makes it easier to measure the operational resonant frequencies, which were identified at 1.7 Hz for a bending mode and 2.5 Hz for a torsional mode from the camera measurements. The camera measurement agree well with an accelerometer measurement on the bridge. The finite element model has resonant frequencies at 1.6 Hz and 2.9 Hz that directly correlate to the bending mode and torsional mode, that are seen in a simulated lift span impact in the model. Additionally a measurement was made of the deflection of the bridge due to traffic induced loading, and using a static load analysis in the finite element model, the weights of vehicles on the bridge can be estimated from the amount of deflection.

Chapter 8

Summary and Discussion, Conclusions, and Future Work

8.1 Summary and Discussion

8.1.1 Summary

This thesis described a methodology for using a video camera to measure the vibrations of civil infrastructure. We first presented a new metric for damage detection based on the symmetry of measured mode shapes, which was demonstrated on several simulated models including the ASCE benchmark structure, and an experimental study on corroded pipes. A full description was given of the algorithms for measuring imperceptibly small motions from video, processing them into displacements and vibration information, and generating visualizations of either the motions in the video directly or using the information to make a plausible simulation of user-specified manipulation of the object.

The measurement system and methodology is fully verified and compared well against co-located and simultaneous accelerometer and laser vibrometer measurements, albeit with a higher noise floor. The noise floor was fully characterized under the effects of different parameters and the local contrast, lens focal length, distance to the target, and camera motion were found to have significant effects. To lessen the

effects of camera motion, a multi-rate Kalman filter was used to fuse accelerometer measurements of camera motion with camera video displacement measurements to provide a better estimate of the motions of the camera. The same multi-rate Kalman filter was also demonstrated for data fusion of an accelerometer within the video frame of a camera measuring displacement.

Experimental studies were conducted on steel cantilever beams and pipe cross sections where operational resonant frequencies and mode shapes were measured, and damage caused by machined defects and corrosion respectively, were detected from shifts in resonant frequency in the camera measurements. A quasi real-time implementation was demonstrated on measurement of a lightweight 3D printed cantilever beam, where contact accelerometer measurements had previously failed due to mass loading effects. Displacements and strains were measured from a granite specimen undergoing a compression test and the fracture locations could be estimated from the extracted information in the video, pre-fracture of the specimen.

Field measurements were made on a series of different structures. Resonant frequencies were measured from a pump and pipe system at the MIT physical facilities, consistent with laser vibrometer measurements. The antenna tower on the top of the MIT Green Building was measured with the camera system from over 175 meters (574 feet) away, and the resonant frequency observed matched one measured by a laser vibrometer on the roof. Earthquake motions of the tuned mass damper in the Taipei 101 skyscraper were identified from a video from YouTube showing the arrival of the P and S waves, with travel times and wave speeds consistent with the region. Frequencies measured from the video of the TMD matched those found in literature for the skyscraper. The full study of the WWI Memorial Bridge in Portsmouth, New Hampshire included measurements of the bridge's response from 260 feet (80 meters) away due to the lift span impact and traffic loading. It also included the construction of a finite element model of the bridge that exhibited similar response with a simulated lift span impact with similar resonant frequencies and mode shapes. Information from a static load analysis of the finite element model was also used to estimate the traffic loading on the bridge using the displacements as measured by the

camera. This study also showed a direct correlation between the camera measurement and simultaneous accelerometer measurement on the bridge, verifying that the long range camera measurement was accurate.

The following sections discuss the advantages and limitations of the camera-based measurement methodology when compared to current methods, and the lessons learned from the presented experiments and applications.

8.1.2   Advantages of the Methodology

Non-contact measurement

The camera is a non-contact measurement method that does not require the placement of physical sensors on a structure. For lightweight structures, this is particularly important as the attachment of sensors could change the dynamics of the system by adding mass. For historically or aesthetically important structures there is also less impact, as a camera measurement system can replace a contact measurement system or reduce the number of contact sensors necessary. The external placement of wiring and sensors may be undesirable.

Area of coverage and high spatial density measurements

The area of coverage of the camera is superior to point sensors such as accelerometers, strain gauges, and laser vibrometers, as a single camera can potentially measure many signals at once as every pixel in a video is potentially a sensor. Virtual sensors can be composed from video over a structure and provide measurements with high spatial density. With data-driven algorithms for SHM requiring more signals to work with to identify potential damage, changes, and disturbances to a structure, the camera methodology is well positioned to easily provide many measurements on a single structure. Even though camera measurements may be less precise, the area of coverage and spatial density of the measurements may make up for it.

221

Cost of measurement hardware and set up

If many measurements at different locations are necessary, it is possible that only a couple of cameras are needed to fully measure the motions of a structure. To match the number of measurements from a couple of cameras, potentially one or two orders of magnitude more of contact sensors are necessary. Since civil infrastructure tends to vibrate at relatively low frequencies, relatively inexpensive cameras can be used. Additionally, set up of a couple cameras is much quicker than instrumenting a structure with many contact sensors, saving labor costs as well. These factors combined, mean that the camera measurement system is likely to be less expensive than a traditional instrumentation setup.

8.1.3 Limitations of the Methodology

Noise floor

The main issue the the camera methodology, is that the noise floor is higher when compared to contact accelerometers, by up to several orders of magnitude depending on the frequency range. Particularly at higher frequencies, structures will vibrate with smaller amplitudes, which are more difficult to measure as displacements. It is possible that multiple camera setups and sensor fusion techniques with contact sensors can improve the noise floor. Measurements over longer time frames with averaging, as long as camera motion can be corrected, may also allow for lower noise floors.

Stationary camera requirement

Currently we require that the camera be somewhat stationary so that only small motions, less than a couple of pixels, are present in the video. This currently precludes the method from being used from moving platforms such as cars, drones, or planes. With measurement of the camera motion using external sensors such as accelerometers and gyroscopes, and from global motion in the image, it is possible that with sufficient motion compensation algorithms, measurements from moving platforms can be made in the future.

Flickering lighting or changing background conditions

Flickering lighting can introduce periodic apparent motion into a video when there is none. When unintentional, it can also alias vibrations to different frequencies. Changing background conditions can also introduce apparent motion, such as clouds moving behind a building. Atmospheric turbulence could also introduce extra apparent motions that are otherwise not well characterized. Flickering lighting can be avoided as the frequencies are well known, as they are typically related to the frequency of AC power. In indoor measurements where fluorescent lighting is common, additional DC lights can be brought in to illuminate the object under test. For outdoor measurements, flickering lighting typically isn't an issue with the sun as a lighting source, however changing background or weather conditions may play a role. To alleviate these effects an image matting procedure could be used [59]. Atmospheric turbulence is another issue that can become increasingly difficult to deal with as measurement distances become large. The effects could possibly be measured and negated by using multiple camera setups that look at a structure through different columns of air, as in [105].

8.1.4 Discussion of Experiments and Applications

The applications and experiments represent a wide variety of different applications in civil infrastructure. The Portsmouth bridge study (Chapter 7) is a proof of concept for vibration and deflection assessment of bridges through a camera measurement. The measurement of displacements from the lift span impacts and traffic induced loading, can be combined with information from the FEA model to determine stresses in the structure. With the stress levels and counting the number of loading cycles, the approximate remaining fatigue life of the bridge can be determined by referencing S-N curves [113]. Measurements of the vibrational behavior of structures can also be used to update FEA models as a form of SHM [65, 92]. The measurement of the corroded pipe (Section 5.3.4) and damaged cantilever beams (Section 5.1.2) represent laboratory proof of concepts of the method for damage detection, where the object

223

under test is intentionally excited and a measurement of the response is made and compared to a healthy baseline. The motion magnification visualization (Section 3.4 and Appendix B) and plausible simulation of object dynamics (Section 3.5) are useful as a quick way to understand how structures vibrate and behave. In the case of operational resonant frequency and mode shape identification on the cantilever beam (Section 5.1), motion magnification was key to identifying the torsional modes of the beam as the visualization clearly showed a torsional motion in the beam. It could also be useful in showing how machinery vibrates (Section 6.1) which can assist in preventing destructive vibrations. In the situation of heavy machinery however, the camera is not as well suited for monitoring because of insufficient noise floor in an inherently vibrationally noisy environment. The study on the Taipei 101 tower (Section 6.3) illustrated that useful displacement and vibration information could be extracted from ubiquitous security videos, even when not controlled by the experimenter and taken from the wild (e.g. YouTube). The noise floor and other studies completed in Chapter 4 allow for determination of the necessary parameters to make successful measurements of a structure given expected displacements, highlighted in the Green Building measurement (Section 6.2.5).

8.2   Conclusions

This research involved a multi-disciplinary effort combining work from computer vision, signal processing, structural dynamics, modal analysis, and experimental design. New computer vision methods including motion magnification were applied to the measurement of vibrations in structures and information in a way that is a massive improvement over previous camera measurement methods. The newly developed methods are potentially faster by over an order of magnitude than current camera measurement methods such as DIC, opening up the opportunity for real-time measurements. It does not require the placement of patterns or lit targets on the object to be measured, as long as there is sufficient local contrast on the object itself. The displacements that the new method can measure on the order of $1/1000^{th}$ to

224

$1/100^{th}$ of a pixel, depending on the experimental conditions represents a significant improvement over current methods.

8.2.1 Methodology Applicability and Usage

Given certain advantages and limitations of the camera-based methodology, and based on the applications presented in this thesis we can make judgments on the situations in which camera measurements would be particularly useful and/or advantageous over contact sensors. Long span bridges and tall buildings are prime candidates as they are likely difficult to instrument with contact sensors. The best solution for those structures might be a hybrid system, where contact sensors are placed in structurally critical and easy to access locations, and cameras are used to measure locations that are difficult to access and to double up on critical locations. Sensor fusion and damage detection algorithms could then make use of the disparate data sources to better characterize the structure, than any sensor system alone. Just as a diverse set of damage detection and SHM algorithms can improve system performance [12], so could a diverse set of measurement methodologies. The spatial density of measurements of the camera can make up for fewer contact sensor measurements, and the better sensitivity of accelerometers or strain gauges can augment the camera measurements.

In a laboratory or testing environment, any structures or objects that require especially dense measurements that would be overwhelmed by the placement of contact sensors, a couple of cameras could be used instead. Measurements from a variety of angles would provide the necessary displacement and vibration information to replace most accelerometer measurements for modal testing or model validation. In these situations, the non-contact nature is important as sensors would provide additional mass loading, possibly changing the dynamics of the structure [2, 17]. Motion magnification also provides a visualization of the structure behavior that can be used to quickly understand the behavior of objects and structures, which can be difficult if the number of contact sensors is limited. For vibration testing of equipment this may be particularly useful.

In infrastructure and industrial environments, there already exist many cameras either for security or monitoring purposes. These cameras are good candidates for quantitative monitoring of structures or machinery, as they record the same scene for days on end and no additional installation of equipment is necessary. Virtual sensors could be placed in the video, and displacements processed from those locations on a structure. Then, over time a baseline of the structure's behavior can be built up with SHM algorithms, and when any change were to occur, operators could be alerted to a change in the behavior. The most likely first applications for the methodology will be critical structures where cameras are already installed, and only the video stream and the addition of data processing is necessary to start monitoring the structure.

8.3   Future Work

The current study is primarily focused on vibration measurements of machinery and large infrastructure, and there are potentially other applications to explore with the same methods and techniques. The earlier motion magnification publications [104, 99], included applications in health care, for measuring a person's respiration and pulse rate. The methods developed in this thesis could allow for different applications with a measurement setup that allows for longer-distance measurements, and more detailed analysis of a person's motions. Other than health care and structural applications, once this methodology is well understood, developed, and circulated amongst many different scientific and engineering communities, hopefully researchers find other uses for it to solve their problems involving distinctive small motions.

For measurements of especially large infrastructure even though vibration characteristics are important to have, a series of static displacements due to external loading over time may be sufficient information for SHM purposes, with much less data needing to be collected. A technique involving small deviations from assumed geometries or local variations in pictures rather than videos may be able to reveal displacement or deviation information could also be used for structural measurements [98, 34]. This has potential advantages in that less data needs to be stored and processed, however

instead of collecting full field displacements and vibration measurements, deviations from local ideal geometries are measured. It may also be more straight forward to collect information from structures such as curvature as a result of bending or moments, rather than trying to infer that from dynamic displacement measurements.

Measurements of material characteristics are also possible with the same processing procedures described in Section 3.2. The measurement of the granite specimen in Section 5.4 is a proof of concept of the methodology for strain measurement for macroscale materials testing. There has always been great interest in the micro and nanoscale characterization of materials qualitatively through optical microscopy, scanning electron microscopy, and other imaging methods. Chemical imaging methods bring another dimension of information to the pictures. Raman spectroscopy measures the energies of vibrational and other modes in molecules which give a specific spectra that can be used for identification of specific chemical compounds. Recently it has been used for mapping of collagen fibril orientations in human osteonal bone [79]. Another chemical imaging technique, multi-channel energy dispersive spectroscopy (EDS) elemental mapping and diffraction-limited confocal Raman spectroscopy was used with backscattered scanning electron microscopy to create maps of elemental and chemical contributions on images on a sea urchin tooth [63]. Interestingly, Raman spectroscopy also has the capability of measuring local stresses inherent in the material, which causes a shift in the Raman spectra due to changes in the lattice and bond structure in chemicals [33, 71]. The video processing methods can likely be applied to images from electron microscopy and/or optical microscopy to measure micro and nano-scale strains and displacements of specimens over time due to changes in loading or environment. Combined with continuum level measurements of stresses and strains from a load cell and load frame, local measurements of stresses with Raman spectroscopy, and a measurement of chemical composition using Raman spectroscopy and elemental distribution in a material using EDS, this could provide extremely insightful information about the local behavior of a material under different loading conditions and correlate that to the local chemical and elemental composition. This so-called stress-strain microscopy connected to chemical imaging could

227

provide new insights into material behavior in informing the design of materials from the micro-scale up.

Appendix A

Abbreviations and Acronyms

1D   One dimensional

2D   Two dimensional

3D   Three dimensional

AASHTO American Association of State Highway and Transportation Officials

AC   Alternating current

ASCE American Society of Civil Engineers

CSAIL Computer Science and Artificial Intelligence Laboratory

CSM Continuous symmetry measure

DC   Direct current

DIC   Digital image correlation

DOF Degree of freedom

DOT Department of Transportation

EDS Energy dispersive spectroscopy

FDD Frequency domain decomposition

FEA	Finite element analysis
FEM	Finite element modeling
FFT	Fast Fourier transform
fps	Frames per second
GLSL	OpenGL Shading Language
GPS	Global positioning system
GPU	Graphics processing unit
IASC	International Association of Structural Control
IID	Independent and identically distributed
MA	Massachusetts
MAC	Modal assurance criterion
MEMS	Microelectromechanical system
MIT	Massachusetts Institute of Technology
NASA	National Aeronautics and Space Administration
NH	New Hampshire
NDT	Non-destructive testing
ODS	Operational deflection shapes
PSD	Power spectral density
PVC	Polyvinyl chloride
RGBA	Red green blue alpha
ROI	Region of interest

SAR   Synthetic aperture radar

SHM   Structural health monitoring

SNR   Signal to noise ratio

SVD   Singular value decomposition

TIFF   Tagged Image File Format

UNH   University of New Hampshire

UTC   Coordinated Universal Time

WWI   World War I

Appendix B

Phase-Based Motion Magnification

This section is reprinted in this Appendix with permission from [97], and reproduced here for the convenience of the reader. It describes phase-based motion processing for motion magnification.

In the case of videos that are global translations of a frame over time, there is a representation that is exactly what we want: the Fourier series. Its basis functions are complex-valued sinusoids that, by the Fourier shift theorem, can be translated exactly by shifting their phase (Figure B-1a,c). However, using the Fourier basis would limit us to only being able to handle the same translation across the entire frame, precluding the amplification of complex spatially-varying motions. To handle such motions, we instead use spatially-local complex sinusoids implemented by a wavelet-like representation called the complex steerable pyramid [85, 84]. This is the same processing workflow, with a different set of filters as in Section 3.2. This representation decomposes images into a sum of complex wavelets corresponding to different scales, orientations and positions. Each wavelet has a notion of local amplitude and local phase, similar to the amplitude and phase of a complex sinusoid (Figure B-2a). The key to our new approach is to perform the same 1D temporal signal processing and amplification described earlier on the local phase of each wavelet, which directly corresponds to local motion as we discuss below.
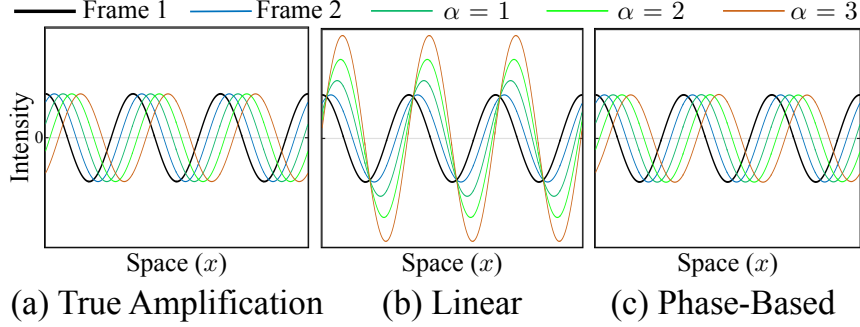
Figure B-1: Phase-based motion magnification is perfect for Fourier basis functions (sinusoids). In these plots, the initial displacement is $\delta(t) = 1$.

B.1 Simplified global case

To provide intuition for what phase is and how it can be used to magnify motion, we work through a simplified example in which one dimensional translation of an image is magnified using the phase of global Fourier basis coefficients (Fig. B-1.) This derivation is closely related to the derivation of the Fourier Shift Theorem [68].

Specifically, again let image intensity $I(x, t)$ be given by $f(x - \delta(t))$ where $\delta(0) = 0$. The profile $f(x)$ can be decomposed into a sum of complex coefficients times sinusoids using the Fourier transform

$$f(x) = \sum_{\omega} A_\omega e^{i\phi_\omega} e^{-i\omega x}. \tag{B.1}$$

Because the frames of I are translations of f, their Fourier transform is given by a phase shift by $\omega\delta(t)$:

$$I(x, t) = \sum_{\omega} A_\omega e^{i\phi_\omega} e^{-i\omega(x - \delta(t))} = \sum_{\omega} A_\omega e^{i(\phi_\omega + \omega\delta(t))} e^{-i\omega x}. \tag{B.2}$$

where the phase of these coefficients becomes $\phi_\omega + \omega\delta(t)$. If we subtract the phase at time 0 from the phase at time t, we get the phase difference

$$\omega\delta(t), \tag{B.3}$$

which is proportional to the translation. Amplifying this phase difference by a factor

α and using it to shift the Fourier coefficients of $I(x, t)$ yields

$$\sum_\omega A_\omega e^{i\phi_\omega + (1+\alpha)\omega\delta(t)} e^{-i\omega x} = f(x - (1+\alpha)\delta(t)), \tag{B.4}$$

a new image sequence in which the translations have been *exactly* magnified.

Phase-based magnification works perfectly in this case because the motions are global and because the transform breaks the image into a representation consisting of exact sinusoids (formally, the Fourier transform diagonalizes the translation operator.) In most cases, however, the motions are not global, but local. This is why we break the image into local sine waves using the complex steerable pyramid.

B.2   Complex Steerable Pyramid

The complex steerable pyramid [85, 84] is a complex, overcomplete linear transform. It decomposes a single channel image $I(x, y)$ into a set of coefficients that correspond to basis functions that are simultaneously localized in position (x, y), spatial scale (r) and orientation (θ). The image can be reconstructed by multiplying the coefficients by the basis functions and summing.

The transform is best-described by its self-similar basis functions. Each one is a translation, dilation or rotation of another. So, it is sufficient to look at just one, a one dimensional version of which is shown in Fig. B-2a. It resembles an oriented complex sinusoid windowed by a Gaussian envelope. The complex sinusoid provides locality in frequency while the windowing provides locality in space. Each basis function is complex, consisting of a real, even-symmetric part (cosine) and an imaginary, odd-symmetric part (sine). This gives rise to a notion of local amplitude and phase as opposed to the global amplitude and phase of Fourier basis functions.

The basis functions are chosen so that when the pyramid is built and collapsed without modification, the reconstructed image I_R perfectly matches the input image I. This requirement imposes conditions on the basis functions [84], which we derive now. Because the basis functions corresponding to a single scale r and orientation θ are
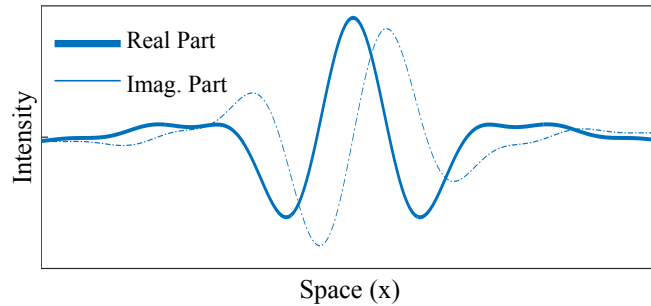
(a) Single Complex Basis Function
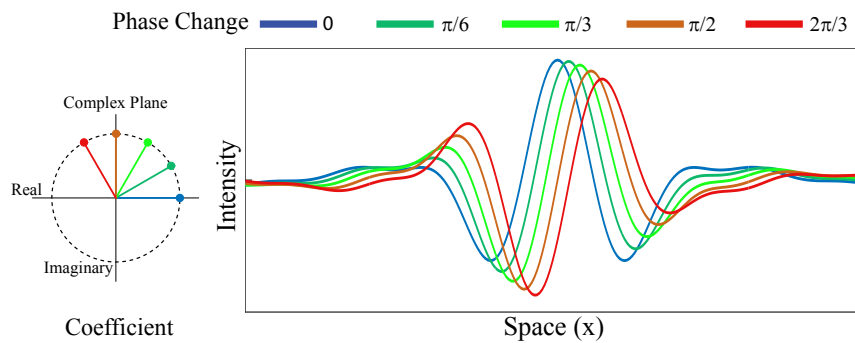
(b) Varying Phase of Coefficient

Figure B-2: Increasing the phase of complex steerable pyramid coefficients results in approximate local motion of the basis functions. A complex steerable pyramid basis function (a) is multiplied by several complex coefficients of constant amplitude and increasing phase to produce the real part of a new basis function that is approximately translating (b).

translated copies of one another, the complex steerable pyramid can be implemented by convolving the image $I(x, y)$ with a basis function from each level $\Psi_{r,\theta}$:

$$L_{r,\theta}(x, y) = I(x, y) * \Psi_{r,\theta}. \tag{B.5}$$

These are the coefficients of the complex steerable pyramid representation. To reconstruct the image, we convolve them with the basis functions and then sum over the scales and orientations

$$\sum_{r,\theta} L_{r,\theta} * \Psi_{r,\theta} = I_R(x, y), \tag{B.6}$$

where $I_R(x, y)$ is the reconstructed image. To ensure that $I_R = I$, we need

$$I_R(x, y) = \sum_{r,\theta} L_{r,\theta} * \Psi_{r,\theta} = I(x, y) * \left(\sum_{r,\theta} \Psi_{r,\theta} * \Psi_{r,\theta} \right). \tag{B.7}$$

If we take the Fourier transform of both sides of this equation, we get

$$\hat{I}_R = \hat{I} \sum_{r,\theta} \left(\hat{\Psi}_{r,\theta} \right)^2 \Rightarrow \sum_{r,\theta} \left(\hat{\Psi}_{r,\theta} \right)^2 = 1 \tag{B.8}$$

as the necessary and sufficient condition for perfect reconstruction of the complex steerable pyramid.

B.3   Local Phase Shift is Local Translation

The link between local phase shift and local translation has been studied before in papers exploring phase-based optical flow [41, 43]. Here, we demonstrate how local phase shift approximates local translation for a single basis function in a manner similar to the global phase-shift theorem of Fourier basis functions. We model a basis function as a Gaussian window multiplied by a complex sinusoid

$$e^{\frac{-x^2}{(2\sigma^2)}} e^{-i\omega x}, \tag{B.9}$$

237

where σ is the standard deviation of the Gaussian envelope and ω is the frequency of the complex sinusoid. In the complex steerable pyramid, the ratio between σ and ω is fixed because the basis functions are self-similar. Low frequency wavelets have larger windows.

Changing the *phase* of the basis element by multiplying it by a complex coefficient $e^{i\phi}$ results in

$$e^{\frac{-x^2}{(2\sigma^2)}} e^{-i\omega x} \times e^{i\phi} = e^{\frac{-x^2}{(2\sigma^2)}} e^{-i\omega(x-\phi/\omega)}. \tag{B.10}$$

The complex sinusoid under the window is translated, which is approximately a translation of the whole basis function by $\frac{\phi}{\omega}$ (Fig. B-2b).

Conversely, the phase difference between two translated basis elements is proportional to translation. Specifically, suppose we have a basis element and its translation by δ:

$$e^{\frac{-x^2}{(2\sigma^2)}} e^{-i\omega x}, \ e^{\frac{-(x-\delta)^2}{(2\sigma^2)}} e^{-i\omega(x-\delta)}. \tag{B.11}$$

The local phase of each element only depends on the argument to the complex exponential and is $-\omega x$ in the first case and $-\omega(x-\delta)$ in the second. The phase difference is then $\omega\delta$, which is directly proportional to the translation. Local phase shift can be used both to analyze tiny translations and synthesize larger ones.
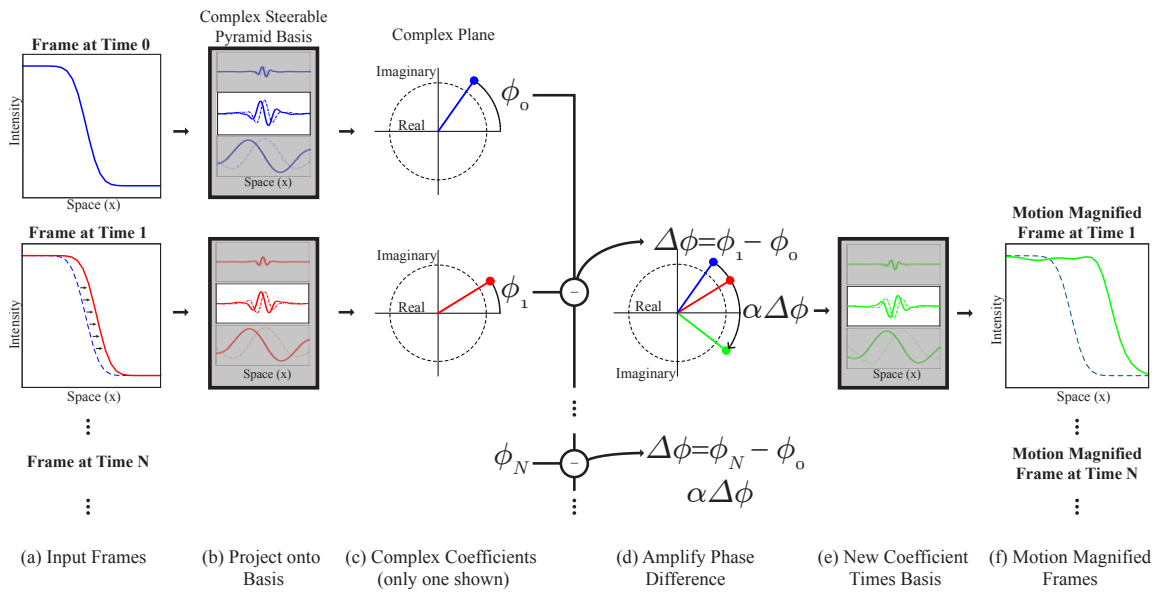
Figure B-3: Using the local phase of complex steerable pyramid coefficients to amplify the motion of a moving step edge. Two frames from a video of a subtly translating step edge (a) are transformed to the complex steerable pyramid representation by projecting onto basis functions (b). The phase between the resulting complex coefficients (c) is computed and amplified (d). Only the coefficient corresponding to exactly one location and scale is shown; this processing is done to every coefficient. The new coefficients are used to shift the basis functions (e) and a reconstructed video is produced in which the motion between the two step edges is evident.

B.4 Phase-based Method

The observation that local phase differences can be used to manipulate local motions motivates our pipeline. We take an image sequence, project each frame onto the complex steerable pyramid basis and then independently amplify the phase difference between *all* corresponding basis elements. This is identical to the linear amplification pipeline except that we have changed the representation from intensities to local spatial phases.

To illustrate the pipeline, consider again an image sequence $I(x, t)$, in which the frame at time 0 is $f(x)$ and the frames at time t are translations $f(x - \delta(t))$ (Fig. B-3a). In our first step, we project each frame onto the complex steerable pyramid basis (Fig. B-3b), which results in a complex coefficient for every scale r, orientation θ and spatial location x, y and time t. Because the coefficients are complex, they can be expressed in terms of amplitude $A_{r,\theta}$ and phase $\phi_{r,\theta}$ as

$$A_{r,\theta}(x, y, t)e^{i\phi_{r,\theta}(x,y,t)}. \tag{B.12}$$

In Fig. B-3c, we show a coefficient at a specific location, scale and orientation in the complex plane at time 0 and at time 1.

Because the the two frames are slight translations of each other, every coefficient has a slight phase difference. This is illustrated in Fig. B-3c, in which the coefficients have roughly the same amplitude but different phases. The next step in the basic version of our processing is therefore to take the phase difference between the coefficients in the video and that of a reference frame, in this case the frame at time 0:

$$\Delta\phi_{r,\theta}(x, y, t) = \phi_{r,\theta}(x, y, t) - \phi_{r,\theta}(x, y, 0). \tag{B.13}$$

This phase difference is then amplified by a factor α as shown in Fig. B-3d. This amplification yields a new set of coefficients for each frame, in which the amplitudes are the same, but the phase differences are larger. We can reconstruct the new frames using these coefficients by multiplying them by the basis functions (Fig. B-3e)
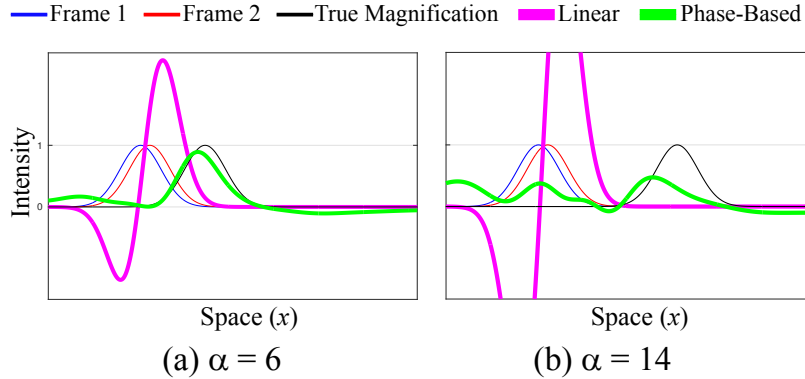
Figure B-4: For general non-periodic structures, the phase-based method can support amplification factors of around four times as high as the linear method and does not suffer from intensity clipping artifacts (a). For large amplification, the different frequency bands break up due to the higher frequency bands having smaller windows (b).

and then summing the real part to get new frames, in which the translations—and therefore the motions in the video—are revealed.

Amplifying phase differences rather than pixel intensity differences has two main advantages: (a) it can support larger amplification factors, and (b) noise amplitude does not get amplified. In Fig. B-4, we show the two different methods being used to amplify the motions of a Gaussian bump. Amplifying raw pixel differences results in overshoot and undershoot causing the signal to appear as pure white or pure black. In contrast, amplifying phase differences allows us to push the Gaussian bump much farther. At very high amplification levels, the different spatial scales of the bump break apart because the high frequency components cannot be pushed as far as the lower frequency components.

In Figure B-5, we show the effect of both methods on a video, which consists of independent and identically distributed (IID) Gaussian noise. Unlike the linear method that increases the power of the noise, the phase based method preserves the noise level preventing objectionable artifacts from making their way into the motion magnified output. For these reasons, we found that amplifying phase differences rather than pixel differences is a better approach for magnifying small motions.

(a) Input   (b) Linear   (c) Phase-based
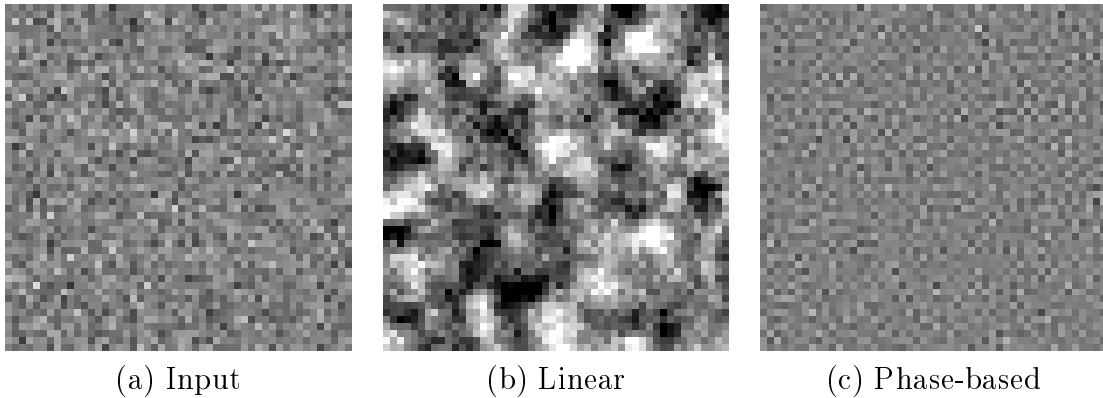
Figure B-5: Comparison between linear and phase-based Eulerian motion magnification in handling noise. (a) A frame in a sequence of IID noise. In both (b) and (c), the motion is amplified by a factor of 50, where (b) amplifies changes linearly, while (c) uses the phase-based approach.

(a) Source   (b) Linear [Wu et al. 2012]   (c) Phase-based (this paper)
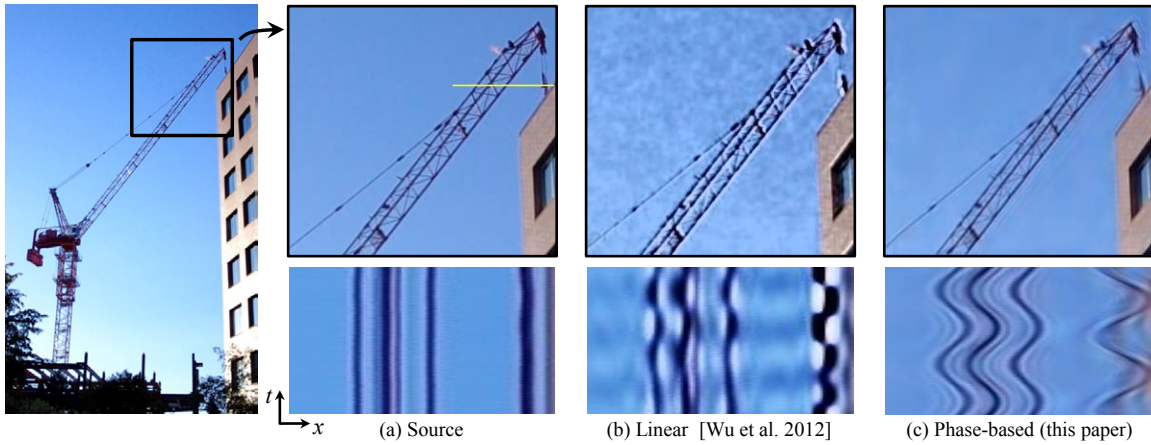
Figure B-6: Motion magnification of a crane imperceptibly swaying in the wind. (a) Top: a zoom-in onto a patch in the original sequence (*crane*) shown on the left. Bottom: a spatiotemporal XT slice of the video along the profile marked on the zoomed-in patch. (b-c) Linear [104] and phase-based motion magnification [99] results, respectively, shown for the corresponding patch and spatiotemporal slice as in (a). The previous, linear method visualizes the crane's motion, but amplifies both signal and noise and introduces artifacts for higher spatial frequencies and larger motions, shown by the clipped intensities (bright pixels) in (b). In comparison, our new phase-based method supports larger magnification factors with significantly fewer artifacts and less noise (c). The full sequences are available in the supplemental video.

242

Appendix C

Additional Experiments

C.1 ASCE Benchmark Structure Identified Modes and CSM Values

Table C.1: Identified Operational Deflection Shape Frequencies for ASCE Benchmark Structure

Damage Scenario	Resonant Mode # and Direction							
	1-Y	1-X	2-Y	2-X	3-Y	3-X	4-Y	4-X
0	8.46	8.98	23.15	25.52	36.38	41.84	47.04	56.69
1	5.47	7.2	19.09	22.38	34.96	40.03	46.51	56.13
2	4.9	6.58	12.27	17.81	34.54	40.03	37.95	48.9
3	8.02	8.98	22.18	25.52	35.93	41.84	46.51	56.69
4	8.02	8.81	22.18	24.66	35.93	41.55	46.51	54.65
5	7.91	8.81	22.18	24.66	35.93	41.55	46.51	54.65
6	8.35	8.98	22.9	25.52	36.38	41.84	46.77	56.69
7	5.52	6.83	22.9	25.52	27.7	35.93	44.36	53.14
8	6.47	7.65	17.4	21.49	36.38	41.55	38.83	50.19
9	7.58	8.98	22.18	25.52	33.76	41.84	45.43	56.69
10	7.43	8.98	20.84	25.52	33.76	41.84	43.58	56.69
11	8.08	8.98	23.15	25.52	34.54	41.84	45.43	56.69
12	8.23	8.98	22.18	25.52	36.38	41.84	45.11	56.69
13	7.49	8.98	21.25	25.52	35.25	41.84	46.51	56.69
14	7.49	8.98	23.15	25.52	32.41	41.84	45.42	56.69

Table C.2: CSM Values for ASCE Benchmark Structure

| Damage Scenario | Resonant Mode # and Direction | | | | | | | | Mean Value |
	1-Y	1-X	2-Y	2-X	3-Y	3-X	4-Y	4-X	
0	0.00182	0.00181	0.00541	0.00513	0.01343	0.03322	0.00979	0.01019	0.01010
1	0.00135	0.00375	0.00480	0.00520	0.03642	0.01031	0.00610	0.01348	0.01018
2	0.00158	0.00484	0.00592	0.00782	0.00325	0.00491	0.00597	0.00404	0.00479
3	0.01445	0.00300	0.01167	0.00561	0.04280	0.02754	0.00793	0.01089	0.01549
4	0.01499	0.00898	0.01140	0.02683	0.05233	0.00768	0.00825	0.01680	0.01841
5	0.01449	0.00904	0.01142	0.02684	0.05251	0.00769	0.00821	0.01680	0.01838
6	0.00274	0.00231	0.00534	0.00531	0.02374	0.03097	0.00590	0.01043	0.01084
7	0.00185	0.00399	0.00330	0.00431	0.00433	0.01088	0.00803	0.00860	0.00566
8	0.00194	0.00359	0.00467	0.00355	0.00741	0.00749	0.00541	0.00513	0.00490
9	0.02608	0.00379	0.01412	0.00590	0.04127	0.02763	0.01788	0.01241	0.01863
10	0.03043	0.00400	0.02466	0.00721	0.03869	0.01720	0.02668	0.01251	0.02017
11	0.01308	0.00340	0.00502	0.00534	0.02505	0.03354	0.01428	0.01163	0.01392
12	0.00520	0.00299	0.01923	0.00695	0.02425	0.01686	0.03808	0.01020	0.01547
13	0.00137	0.00176	0.00562	0.00556	0.01332	0.02108	0.01330	0.01158	0.00920
14	0.00154	0.00327	0.00653	0.00517	0.00739	0.03178	0.01047	0.01455	0.01009

C.2   Shaker Characterization

The shaker used in Section 4.2 was characterized in a measurement with a laser vibrometer. The shaker was directly bolted to the concrete reaction mass sitting directly on the floor slab of the laboratory, oriented vertically, and run at a low enough amplitude such that the piston would not hit the outside casing of the shaker and induce extraneous impacts. The actual motion of the mass piston inside the shaker was measured with a laser vibrometer while a frequency sweep from 0 to 300 Hz was inputted to the shaker from the computer. The measured motion of the mass inside the shaker in both the time domain and frequency domain are shown in Figure C-1. To convert the output to force, the frequency response is multiplied by $2\pi f$ and the frequency response of the shaker in relative units of force is shown in Figure C-2. The shaker has reasonable output from 10 Hz up to 200 Hz, which a peak at about 72 Hz, and a dip at around 20 Hz.
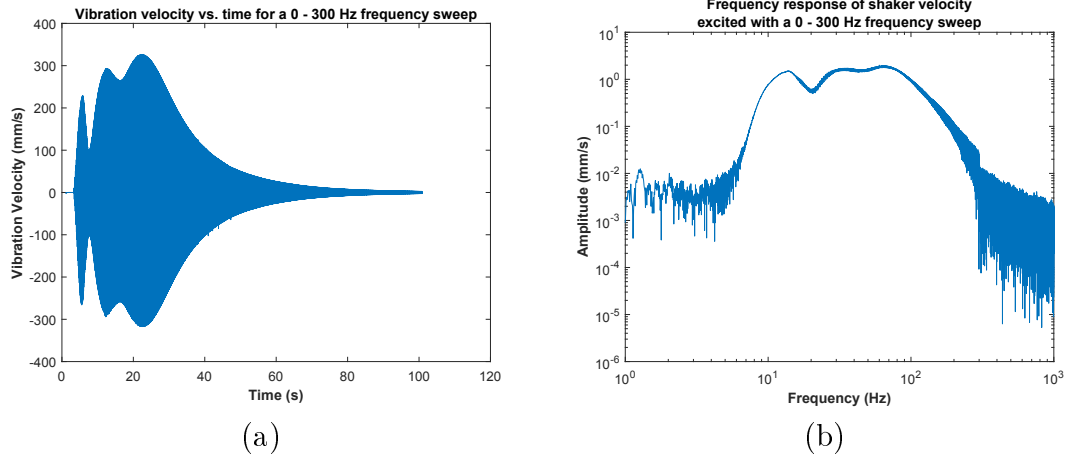
Figure C-1: Measured response of the mass piston in the shaker with a laser vibrometer due to an input frequency sweep from 0 to 300 Hz over 100 seconds in the (a) time domain and (b) frequency domain
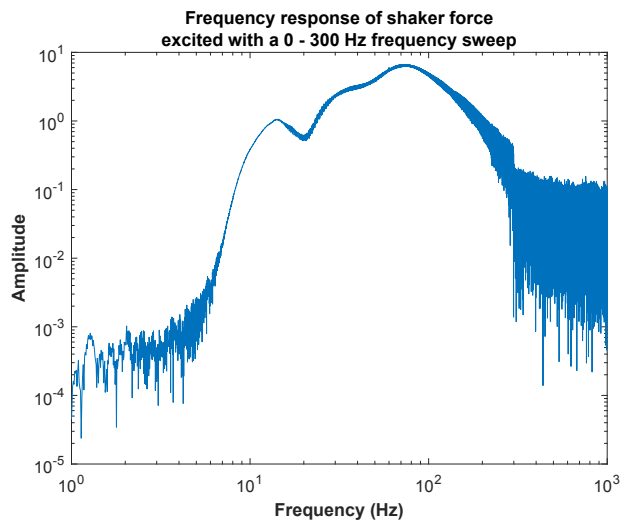
Figure C-2: Frequency response in relative units of force for the shaker, as driven from 0 to 300 Hz

C.3 Noise Floor Parametric Study Measurements for Exposure Time

(a)   (b)
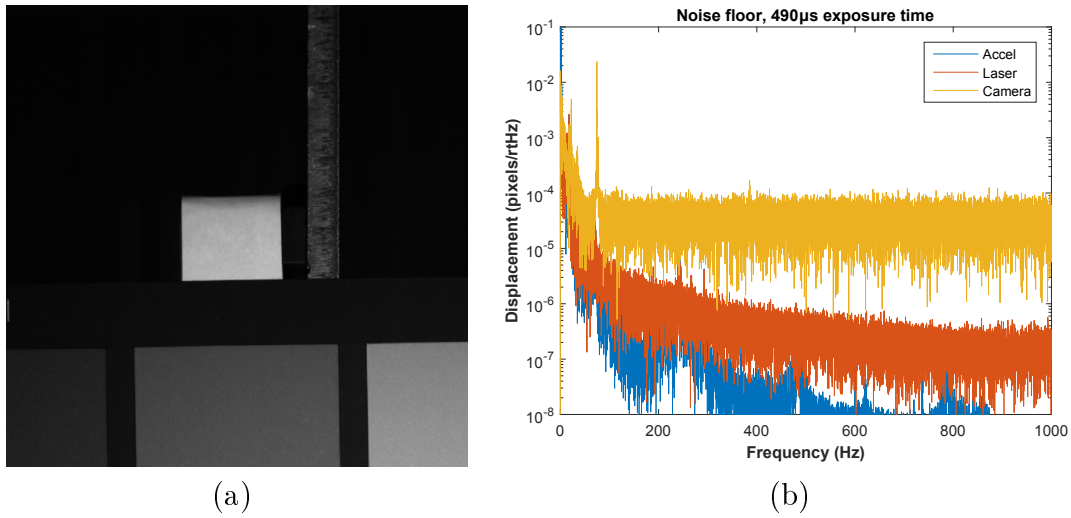
Figure C-3: Exposure time of 490μs, (a) screenshot from video and (b) displacement frequency responses

(a)   (b)
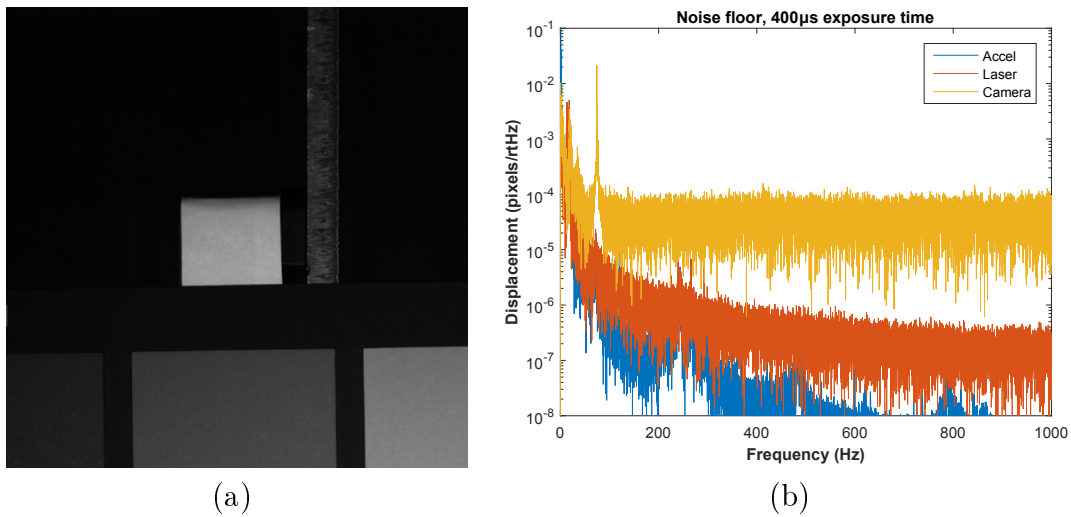
Figure C-4: Exposure time of 400μs, (a) screenshot from video and (b) displacement frequency responses
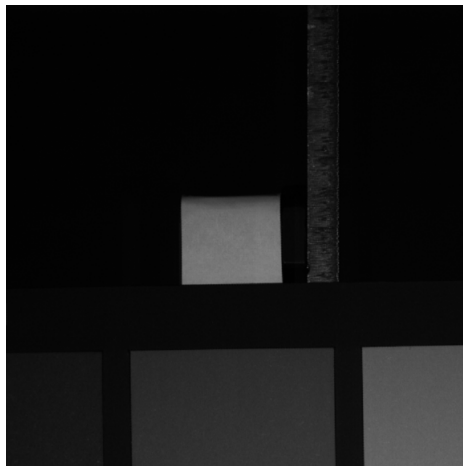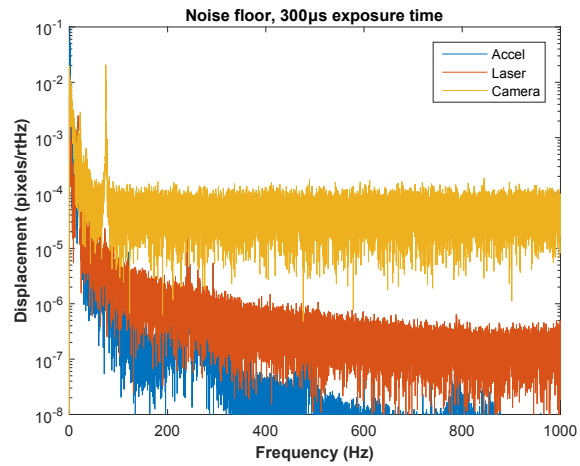
Figure C-5: Exposure time of 300μs, (a) screenshot from video and (b) displacement frequency responses
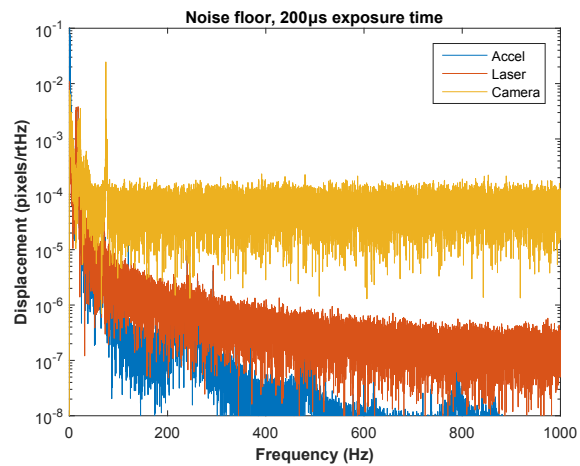
Figure C-6: Exposure time of 200μs, (a) screenshot from video and (b) displacement frequency responses
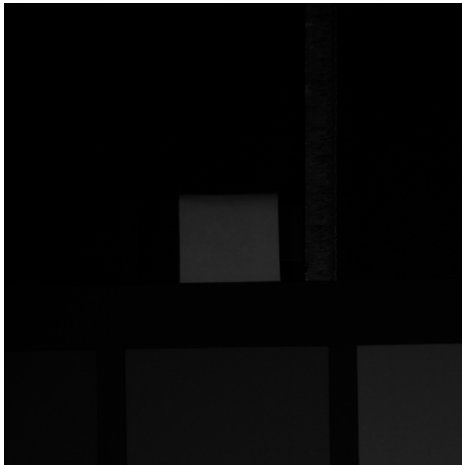
(a)
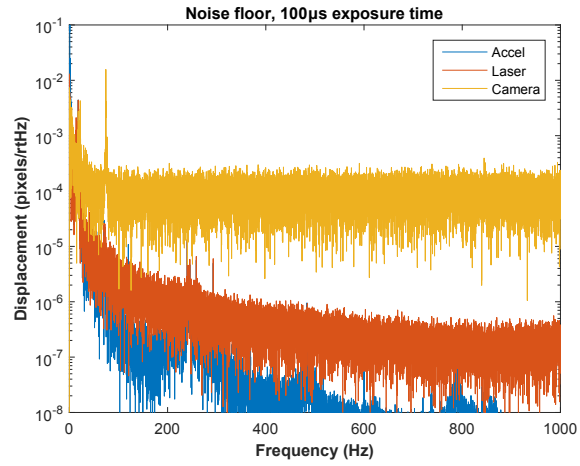
(b)

Figure C-7: Exposure time of 100μs, (a) screenshot from video and (b) displacement frequency responses

(a)
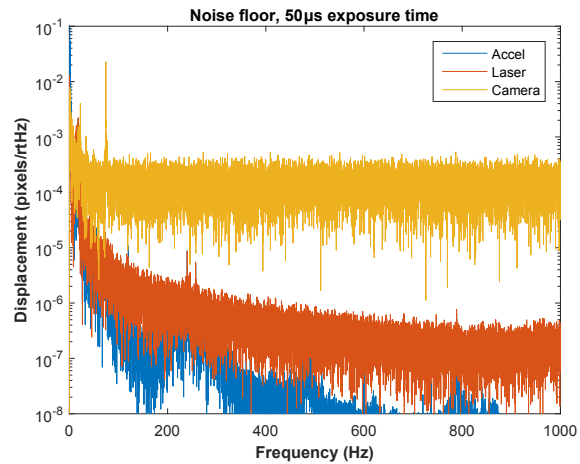
(b)

Figure C-8: Exposure time of 50μs, (a) screenshot from video and (b) displacement frequency responses
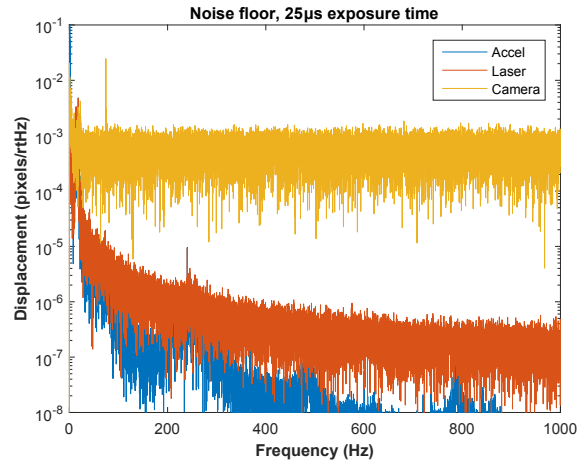
Figure C-9: Exposure time of 25μs, (a) screenshot from video and (b) displacement frequency responses
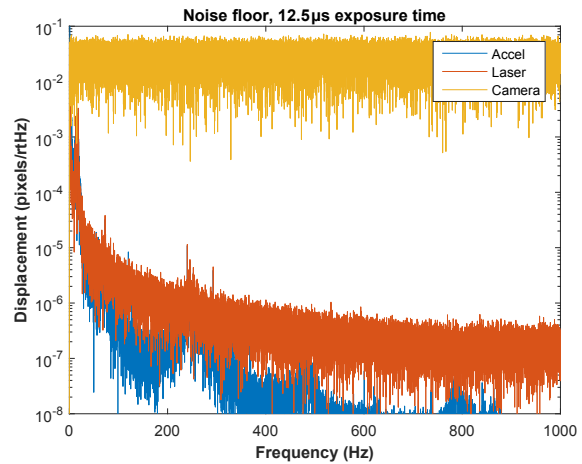
Figure C-10: Exposure time of 12.5μs, (a) screenshot from video and (b) displacement frequency responses

C.4 Kalman Filter Toy Example
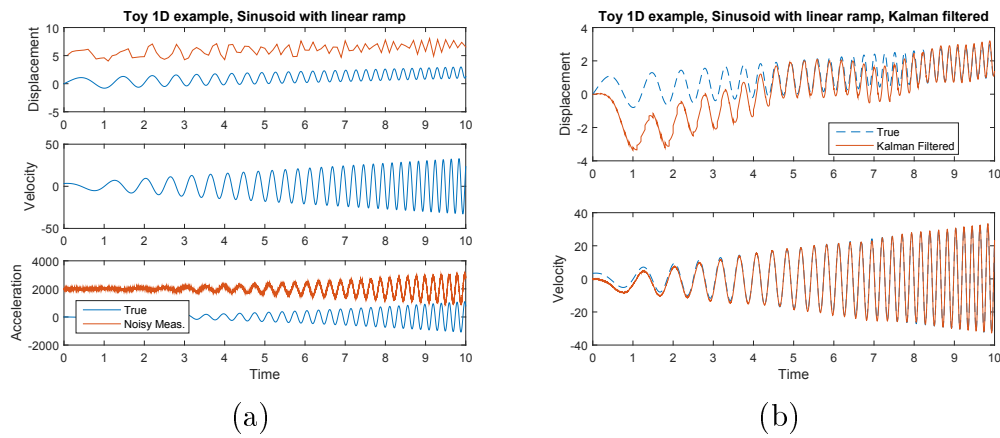
Figure C-11: Toy example of a ramped sinusoid with an accelerometer sample rate of 2000 Hz and a displacement measurement rate of 10 Hz, showing (a) comparison of true and noisy measurements, (b) results of the Kalman filter

Appendix D

Video Catalog

In an effort to catalog the videos corresponding to the work in this thesis, many videos are on YouTube as a repository. A catalog of the videos with links and a short description is included below. The playlist containing videos referenced in this thesis is `goo.gl/g1vORF` or

`youtube.com/playlist?list=PLI3mJDDULMzk0DaqHYkUM7eLzGIiTSqCS`

D.1 Chapter 3: Vibration Measurement and Visualization from Video

Cello Example Video `https://youtu.be/PfhGjbQbGMk`
Source video, and motion magnified videos of strings being plucked on a cello, Figure 3-1

D.2 Chapter 4: Verification and Noise Floor Experiments

Camera Motion Study `https://youtu.be/m8BWYb3wZpI`
Camera motion study input video, Figure 4-14(b)

D.3   Chapter 5: Laboratory Experiments

Cantilever Beam   https://youtu.be/sYY5jMRw2Zc

Cantilever beam as tested in the laboratory showing motion magnified bending and torsional modes, Figure 5-1

Polycarbonate Beam   https://youtu.be/U2eIouSO2yM

Measurement of a polycarbonate cantilever beam with the quasi real-time implementation, Figure 5-11

3D Printed Beam   https://youtu.be/iMsMH3W5SiM

Source video for 3D printed beam [95], Figure 5-13(b)

PVC Pipe   https://youtu.be/LZug9o6EzDU

Initial PVC pipe measurement [24], Figure 5-17

PVC Pipe, DC Lights   https://youtu.be/Iq4OYSj0f_E

PVC pipe measurement with DC lights, Figure 5-22

Intact Steel Pipe   https://youtu.be/RXrw-FhBGSU

Motion magnified videos of intact schedule 10 6" steel pipe, Figure 5-26(a)

Corroded Pipes   https://youtu.be/4LVZOBmThkI

Motion magnified video comparison of the C3 mode for the intact, 2-day, and 5-day corroded pipes, Figure 5-31

Granite Visualization   https://youtu.be/a7UFy_gLbpk

Visualization of the granite specimen hydraulic fracture test, Figure 5-34

D.4   Chapter 6: Field Measurements

MIT Facilities Pump   `https://youtu.be/04WJCpaieTY`
Motion magnified videos of a pump at the MIT facilities plant, Figure 6-2

MIT Facilities Pipes   `https://youtu.be/o3rYwTAHzv4`
Motion magnified videos of pipes at the MIT facilities plant, Figure 6-7

Taipei 101 TMD   `https://youtu.be/LA-ZcEWCOIc`
`https://youtu.be/wbVipGAPJOO`
Taipei 101 interior visitor's center during 4/20/2015 M6.4 earthquake, Figure 6-22

D.5   Chapter 7: Portsmouth Bridge Study

10/8/2015 Measurement   `https://youtu.be/hfYgbKEszqA`
The Portsmouth, NH Bridge source video from the 10/8/2015 measurement, Figure 7-4

10/8/2015 Measurement   `https://youtu.be/Gd4CdFvYce4`
The Portsmouth, NH Bridge on 10/8/2015, motion magnified from 1.65 - 1.75 Hz, ×400, Figure 7-4

10/8/2015 Measurement   `https://youtu.be/-ioY2J9CrWc`
The Portsmouth, NH Bridge on 10/8/2015, motion magnified from 2.4 - 2.6 Hz, ×400, Figure 7-4

12/14/2015 Measurement   `https://youtu.be/DY3BjKsq52g`
The Portsmouth, NH Bridge on 12/14/2015, source video from lift span impact, Figure 7-10

12/14/2015 Measurement	`https://youtu.be/yknHkMY9Q1A` The Portsmouth, NH Bridge on 12/14/2015, video from lift span impact, motion magnified 1.6 - 1.8 Hz ×400, Figure 7-10
12/14/2015 Measurement	`https://youtu.be/UnUrcKJl4Pc` The Portsmouth, NH Bridge on 12/14/2015, video from lift span impact, motion magnified 2.4 - 2.6 Hz ×400, Figure 7-10
Traffic Induced Loading	`https://youtu.be/ukoTU3duglA` The Portsmouth, NH Bridge on 12/14/2015, video from traffic induced loading showing estimated load, Figure 7-16 and 7-34
Bridge Time History FEA	`https://youtu.be/3m2HAJRgp0U` Bridge FEA model, time history analysis, Figure 7-25
Bridge Time History FEA	`https://youtu.be/lqLzNJjR-gM` Bridge FEA model, time history analysis, with damping, Figure 7-27
Bridge FEA 1.248 Hz Mode	`https://youtu.be/B4r0eHGcQww` Bridge FEA model, Mode 7 at 1.248 Hz, Figure 7-28(a)
Bridge FEA 1.598 Hz Mode	`https://youtu.be/6jGFgh4V1lY` Bridge FEA model, Mode 10 at 1.598 Hz, Figure 7-28(b)
Bridge FEA 1.890 Hz Mode	`https://youtu.be/p0S9_rXKdJw` Bridge FEA model, Mode 12 at 1.890 Hz, Figure 7-28(c)

Bridge FEA 2.892 Hz Mode   `https://youtu.be/K4Ram34H1nw`

Bridge FEA model, Mode 20 at 2.892 Hz, Figure 7-28(d)

Bridge FEA 1.598 Hz Mode   `https://youtu.be/-2iLagsmVps`

Bridge FEA model, Mode 10 at 1.598 Hz, to compare with measured motion magnified video, Section 7.4.2

Bridge FEA 2.892 Hz Mode   `https://youtu.be/3GeUs_8T6Kg`

Bridge FEA model, Mode 20 at 2.892 Hz, to compare with measured motion magnified video, Section 7.4.2

References

[1] ASCE. 2013 report card for america's infrastructure.

[2] ASHORY, M. Correction of mass-loading effects of transducers and suspension effects in modal testing. In *Proceedings of the 16th International Modal Analysis Conference* (1998), vol. 3243, p. 815.

[3] BARTILSON, D. T., WIEGHAUS, K. T., AND HURLEBAUS, S. Target-less computer vision for traffic signal structure vibration studies. *Mechanical Systems and Signal Processing 60* (2015), 571–582.

[4] BATHE, K.-J. *Finite element procedures.* Klaus-Jurgen Bathe, 2006.

[5] BLABER, J., ADAIR, B., AND ANTONIOU, A. Ncorr: open-source 2d digital image correlation matlab software. *Experimental Mechanics 55*, 6 (2015), 1105–1122.

[6] BLEVINS, R. D. *Formulas for natural frequency and mode shape.* Van Nostrand Reinhold New York, 1979.

[7] BRINCKER, R., ANDERSEN, P., AND MØLLER, N. An indicator for separation of structural and harmonic modes in output-only modal testing. In *International Modal Analysis Conference (IMAC XVIII), San Antonio, TX, Society for Engineering Mechanics, Bethel, CT* (2000), pp. 1649–1654.

[8] BRINCKER, R., VENTURA, C., AND ANDERSEN, P. Why output-only modal testing is a desirable tool for a wide range of practical applications. In *Proc. Of the International Modal Analysis Conference (IMAC) XXI, paper* (2003), vol. 265.

[9] BRINCKER, R., ZHANG, L., AND ANDERSEN, P. Modal identification of output-only systems using frequency domain decomposition. *Smart materials and structures 10*, 3 (2001), 441.

[10] BUYUKOZTURK, O., CHEN, J. G., WADHWA, N., DAVIS, A., DURAND, F., AND FREEMAN, W. T. Smaller than the eye can see: Vibration analysis with video cameras. In *19th World Conference on Non-Destructive Testing* (2016).

[11] Büyüköztürk, O., Haupt, R., Tuakta, C., and Chen, J. Remote detection of debonding in FRP-strengthened concrete structures using acoustic-laser technique. In *Nondestructive Testing of Materials and Structures*. Springer, 2013, pp. 19–24.

[12] Buyukozturk, O., Long, J., Ghazi, R. M., Cha, Y.-J., Chen, J., and Smit, D. Structural health monitoring: A quest towards the use of combined approaches. In *EWSHM-7th European Workshop on Structural Health Monitoring* (2014).

[13] Büyüköztürk, O., and Rhim, H. C. Radar imaging of concrete specimens for non-destructive testing. *Construction and Building Materials 11*, 3 (1997), 195–198.

[14] Büyüköztürk, O., and Yu, T.-Y. Far-field radar ndt technique for detecting gfrp debonding from concrete. *Construction and Building Materials 23*, 4 (2009), 1678–1689.

[15] Büyüköztürk, O., Yu, T.-Y., and Ortega, J. A. A methodology for determining complex permittivity of construction materials based on transmission-only coherent, wide-bandwidth free-space measurements. *Cement and Concrete Composites 28*, 4 (2006), 349–359.

[16] Caetano, E., Silva, S., and Bateira, J. A vision system for vibration monitoring of civil engineering structures. *Experimental Techniques 35*, 4 (2011), 74–82.

[17] Cakar, O., and Sanliturk, K. Elimination of transducer mass loading effects from frequency response functions. *Mechanical systems and signal processing 19*, 1 (2005), 87–104.

[18] Çelebi, M., Toksöz, N., and Büyüköztürk, O. Rocking behavior of an instrumented unique building on the mit campus identified from ambient shaking data. *Earthquake Spectra 30*, 2 (2014), 705–720.

[19] Chen, J. G., Davis, A., Wadhwa, N., Durand, F., Freeman, W. T., and Buyukozturk, O. Video camera-based vibration measurement for civil infrastructure applications. *Journal of Infrastructure Systems*. in review.

[20] Chen, J. G., Davis, A., Wadhwa, N., Durand, F., Freeman, W. T., and Buyukozturk, O. Video camera-based vibration measurement for condition assessment of civil infrastructure. In *International Symposium Non-Destructive Testing in Civil Engineering (NDT-CE) 2015* (2015).

[21] Chen, J. G., Haupt, R. W., and Buyukozturk, O. Acoustic-laser vibrometry technique for the noncontact detection of discontinuities in fiber reinforced polymerretrofitted concrete. *Materials evaluation 72*, 10 (2014), 1305–1313.

[22] CHEN, J. G., HAUPT, R. W., AND BUYUKOZTURK, O. Operational and defect parameters concerning the acoustic-laser vibrometry method for FRP-reinforced concrete. *NDT & E International 71* (2015), 43–53.

[23] CHEN, J. G., WADHWA, N., CHA, Y.-J., DURAND, F., FREEMAN, W. T., AND BUYUKOZTURK, O. Structural modal identification through high speed camera video: Motion magnification. In *Topics in Modal Analysis I, Volume 7.* Springer, 2014, pp. 191–197.

[24] CHEN, J. G., WADHWA, N., CHA, Y.-J., DURAND, F., FREEMAN, W. T., AND BUYUKOZTURK, O. Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration 345* (2015), 58–71.

[25] CHEN, J. G., WADHWA, N., DAVIS, A., DURAND, F., FREEMAN, W. T., AND BUYUKOZTURK, O. Long distance video camera measurements of structures. In *10th International Workshop on Structural Health Monitoring* (2015).

[26] CHEN, J. G., WADHWA, N., DURAND, F., FREEMAN, W. T., AND BUYUKOZTURK, O. Developments with motion magnification for structural modal identification through camera video. In *Dynamics of Civil Structures, Volume 2.* Springer, 2015, pp. 49–57.

[27] CHEN, K.-C., WANG, J.-H., HUANG, B.-S., LIU, C.-C., AND HUANG, W.-G. Vibrations of the taipei 101 skyscraper caused by the 2011 tohoku earthquake, japan. *Earth, planets and space 64*, 12 (2012), 1277–1286.

[28] CIGADA, A., MAZZOLENI, P., ZAPPA, E., ET AL. Vibration monitoring of multiple bridge points by means of a unique vision-based measuring system. *Experimental Mechanics 54*, 2 (2014), 255–271.

[29] DAVIS*, A., BOUMAN*, K. L., CHEN, J. G., RUBINSTEIN, M., DURAND, F., AND FREEMAN, W. T. Visual vibrometry: Estimating material properties from small motions in video. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on* (2015), IEEE, pp. 5335–5343. *Joint first author.

[30] DAVIS, A., CHEN, J. G., AND DURAND, F. Image-space modal bases for plausible manipulation of objects in video. *ACM Transactions on Graphics (TOG) 34*, 6 (2015), 239.

[31] DAVIS, A., RUBINSTEIN, M., WADHWA, N., MYSORE, G., DURAND, F., AND FREEMAN, W. T. The visual microphone: Passive recovery of sound from video. *ACM Transactions on Graphics (Proc. SIGGRAPH) 33*, 4 (2014), 79:1–79:10.

[32] DE ROECK, G., PEETERS, B., AND REN, W.-X. Benchmark study on system identification through ambient vibration measurements. In *Proceedings of IMAC-XVIII, the 18th International Modal Analysis Conference, San Antonio, Texas* (2000), pp. 1106–1112.

[33] DE WOLF, I. Micro-raman spectroscopy to study local mechanical stress in silicon integrated circuits. *Semiconductor Science and Technology 11*, 2 (1996), 139.

[34] DEKEL, T., MICHAELI, T., IRANI, M., AND FREEMAN, W. T. Revealing and modifying non-local variations in a single image. *ACM Transactions on Graphics (TOG) 34*, 6 (2015), 227.

[35] DOEBLING, S. W., FARRAR, C. R., PRIME, M. B., ET AL. A summary review of vibration-based damage identification methods. *Shock and vibration digest 30*, 2 (1998), 91–105.

[36] DZUNIC, Z., CHEN, J. G., MOBAHI, H., BUYUKOZTURK, O., AND FISHER III, J. W. A bayesian state-space approach for damage detection and classification. In *Dynamics of Civil Structures, Volume 2*. Springer, 2015, pp. 171–183.

[37] EMGE, T., AND BUYUKOZTURK, O. Remote nondestructive testing of composite-steel interface by acoustic laser vibrometry. *Materials evaluation 70*, 12 (2012), 1401–1410.

[38] FARRAR, C. R., DARLING, T. W., MIGLIORI, A., AND BAKER, W. E. Microwave interferometers for non-contact vibration measurements on large structures. *Mechanical Systems and Signal Processing 13*, 2 (1999), 241–253.

[39] FARRAR, C. R., AND LIEVEN, N. A. Damage prognosis: the future of structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 365*, 1851 (2007), 623–632.

[40] FARRAR, C. R., AND WORDEN, K. An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 365*, 1851 (2007), 303–315.

[41] FLEET, D. J., AND JEPSON, A. D. Computation of component image velocity from local phase information. *Int. J. Comput. Vision 5*, 1 (Sept. 1990), 77–104.

[42] FREEMAN, W. T., AND ADELSON, E. H. The design and use of steerable filters. *IEEE Transactions on Pattern analysis and machine intelligence 13*, 9 (1991), 891–906.

[43] GAUTAMA, T., AND VAN HULLE, M. A phase-based approach to the estimation of the optical flow field using spatial filtering. *Neural Networks, IEEE Transactions on 13*, 5 (sep 2002), 1127 – 1136.

[44] GROOS, J., AND RITTER, J. Time domain classification and quantification of seismic noise in an urban environment. *Geophysical Journal International 179*, 2 (2009), 1213–1231.

[45] GÜNEŞ, O., AND BÜYÜKÖZTÜRK, O. Microwave imaging of plain and reinforced concrete for NDT using backpropagation algorithm. In *Nondestructive Testing of Materials and Structures*. Springer Netherlands, 2013, pp. 703–709.

[46] HAO, S. I-35w bridge collapse. *Journal of Bridge Engineering 15*, 5 (2009), 608–614.

[47] HASKETT, T., BREUKELMAN, B., ROBINSON, J., AND KOTTELENBERG, J. Tuned mass dampers under excessive structural excitation. *Motioneering Inc.* (2013). Guelph, Ontario, Canada N1K 1B8.

[48] HELFRICK, M. N., NIEZRECKI, C., AVITABILE, P., AND SCHMIDT, T. 3d digital image correlation methods for full-field vibration measurement. *Mechanical Systems and Signal Processing 25*, 3 (2011), 917–927.

[49] HERMANS, L., AND VAN DER AUWERAER, H. Modal testing and analysis of structures under operational conditions: industrial applications. *Mechanical Systems and Signal Processing 13*, 2 (1999), 193–216.

[50] JOHNSON, E., LAM, H., KATAFYGIOTIS, L., AND BECK, J. Phase i iasc-asce structural health monitoring benchmark problem using simulated data. *Journal of Engineering Mechanics 130*, 1 (2004), 3–15.

[51] JOSHI, N., KANG, S. B., ZITNICK, C. L., AND SZELISKI, R. Image deblurring using inertial measurement sensors. In *ACM Transactions on Graphics (TOG)* (2010), vol. 29, ACM, p. 30.

[52] KASHANGAKI, T. A.-L., SMITH, S. W., AND LIM, T. W. Underlying modal data issues for detecting damage in truss structures. In *Proceedings of the 33rd AIAA/ASME/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (1992), pp. 1437–1446.

[53] KIM, S.-W., AND KIM, N.-S. Multi-point displacement response measurement of civil infrastructures using digital image processing. *Procedia Engineering 14* (2011), 195–203.

[54] KIM, S.-W., AND KIM, N.-S. Dynamic characteristics of suspension bridge hanger cables using digital image processing. *NDT & E International 59* (2013), 25–33.

[55] KOLB, A., BARTH, E., KOCH, R., AND LARSEN, R. Time-of-flight sensors in computer graphics. In *Proc. Eurographics (State-of-the-Art Report)* (2009), pp. 119–134.

[56] LARSON, C. B., ZIMMERMAN, D. C., AND MAREK, E. L. A comparison of modal test planning techniques: excitation and sensor placement using the nasa 8-bay truss. In *Proceedings of the 12th International Modal Analysis* (1994), vol. 2251, p. 205.

[57] LEE, J.-J., AND SHINOZUKA, M. Real-time displacement measurement of a flexible bridge using digital image processing techniques. *Experimental mechanics 46*, 1 (2006), 105–114.

[58] LEE, S.-J., CHEN, H.-W., LIU, Q., KOMATITSCH, D., HUANG, B.-S., AND TROMP, J. Three-dimensional simulations of seismic-wave propagation in the taipei basin with realistic topography based upon the spectral-element method. *Bulletin of the Seismological Society of America 98*, 1 (2008), 253–264.

[59] LEVIN, A., LISCHINSKI, D., AND WEISS, Y. A closed-form solution to natural image matting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 30*, 2 (2008), 228–242.

[60] LIU, C., TORRALBA, A., FREEMAN, W. T., DURAND, F., AND ADELSON, E. H. Motion magnification. *ACM Trans. Graph. 24* (Jul 2005), 519–526.

[61] LONG, J., AND BUYUKOZTURK, O. *Automated structural damage detection using one-class machine learning.* Springer, 2014.

[62] MAIA, N., SILVA, J., ALMAS, E., AND SAMPAIO, R. Damage detection in structures: from mode shape to frequency response function methods. *Mechanical systems and signal processing 17*, 3 (2003), 489–498.

[63] MASIC, A., AND WEAVER, J. C. Large area sub-micron chemical imaging of magnesium in sea urchin teeth. *Journal of structural biology 189*, 3 (2015), 269–275.

[64] MOHAMMADI GHAZI, R., AND BÜYÜKÖZTÜRK, O. Damage detection with small data set using energy-based nonlinear features. *Structural Control and Health Monitoring* (2015). DOI: 10.1002/stc.1774.

[65] MOTTERSHEAD, J., AND FRISWELL, M. Model updating in structural dynamics: a survey. *Journal of sound and vibration 167*, 2 (1993), 347–375.

[66] NATIONAL TRANSPORTATION SAFETY BOARD. Collapse of i-35w highway bridge, minneapolis, minnesota, august 1, 2007. Highway Accident Report NTSB/HAR-08/03, Washington, DC.

[67] OF STATE HIGHWAY, A. A., ON BRIDGES, T. O. S., AND STRUCTURES. *The Manual for Bridge Evaluation.* American Association of State Highway and Transportation Officials, 2011.

[68] OPPENHEIM, A. V., AND SCHAFER, R. W. Discrete-time signal processing. *Prentice Hall, New York* (2010).

[69] PANDEY, A., BISWAS, M., AND SAMMAN, M. Damage detection from changes in curvature mode shapes. *Journal of sound and vibration 145*, 2 (1991), 321–332.

[70] Patsias, S., and Staszewskiy, W. Damage detection using optical measurements and wavelets. *Structural Health Monitoring 1*, 1 (2002), 5–22.

[71] Pezzotti, G., and Sakakura, S. Study of the toughening mechanisms in bone and biomimetic hydroxyapatite materials using raman microprobe spectroscopy. *Journal of Biomedical Materials Research Part A 65*, 2 (2003), 229–236.

[72] Pieraccini, M., Fratini, M., Parrini, F., and Atzeni, C. Dynamic monitoring of bridges using a high-speed coherent radar. *IEEE Transactions on Geoscience and Remote Sensing 44*, 11 (2006), 3284.

[73] Poon, D. C., Shieh, S.-s., Joseph, L., and Chang, C. Structural design of taipei 101, the world's tallest building. In *Proceedings of the CTBUH 2004 Seoul Conference, Seoul, Korea* (2004), pp. 271–278.

[74] Ratcliffe, C. P. A frequency and curvature based experimental method for locating damage in structures. *Journal of vibration and acoustics 122*, 3 (2000), 324–329.

[75] Rhim, H. C., and Buyukozturk, O. Electromagnetic properties of concrete at microwave frequency range. *Materials Journal 95*, 3 (1998), 262–271.

[76] Rhim, H. C., and Büyüköztürk, O. Wideband microwave imaging of concrete for nondestructive testing. *Journal of Structural Engineering 126*, 12 (2000), 1451–1457.

[77] Rhim, H. C., Buyukozturk, O., and Blejer, D. J. Remote radar imaging of concrete slabs with and without a rebar. *Materials evaluation 53*, 2 (1995), 295–299.

[78] Rubinstein, M. *Analysis and Visualization of Temporal Variations in Video.* PhD thesis, Massachusetts Institute of Technology, Feb 2014.

[79] Schrof, S., Varga, P., Galvis, L., Raum, K., and Masic, A. 3d raman mapping of the collagen fibril orientation in human osteonal lamellae. *Journal of structural biology 187*, 3 (2014), 266–275.

[80] Schumacher, T., and Shariati, A. Monitoring of structures and mechanical systems using virtual visual sensors for video analysis: Fundamental concept and proof of feasibility. *Sensors 13*, 12 (2013), 16551–16564.

[81] Shabana, A. A. *Theory of vibration*, vol. 2. Springer, 1991.

[82] Shi, Z., Law, S., and Zhang, L. Structural damage localization from modal strain energy change. *Journal of Sound and Vibration 218*, 5 (1998), 825–844.

[83] Shull, P. J. *Nondestructive evaluation: theory, techniques, and applications.* CRC press, 2002.

[84] SIMONCELLI, E. P., AND FREEMAN, W. T. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *Proceedings of the 1995 International Conference on Image Processing (Vol. 3)-Volume 3 - Volume 3* (Washington, DC, USA, 1995), ICIP '95, IEEE Computer Society, pp. 444–447.

[85] SIMONCELLI, E. P., FREEMAN, W. T., ADELSON, E. H., AND HEEGER, D. J. Shiftable multi-scale transforms. *IEEE Trans. Info. Theory 2*, 38 (1992), 587–607.

[86] SMYTH, A., AND WU, M. Multi-rate kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring. *Mechanical Systems and Signal Processing 21*, 2 (2007), 706–723.

[87] SOHN, H. Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 365*, 1851 (2007), 539–560.

[88] SOHN, H., CZARNECKI, J. A., AND FARRAR, C. R. Structural health monitoring using statistical process control. *Journal of Structural Engineering 126*, 11 (2000), 1356–1363.

[89] SOHN, H., FARRAR, C. R., HEMEZ, F. M., SHUNK, D. D., STINEMATES, D. W., NADLER, B. R., AND CZARNECKI, J. J. A review of structural health monitoring literature: 1996-2001, 2004.

[90] SOHN, H., FARRAR, C. R., HUNTER, N. F., AND WORDEN, K. Structural health monitoring using statistical pattern recognition techniques. *Journal of dynamic systems, measurement, and control 123*, 4 (2001), 706–711.

[91] STANBRIDGE, A., EWINS, D., AND KHAN, A. Modal testing using impact excitation and a scanning ldv. *Shock and Vibration 7*, 2 (2000), 91–100.

[92] SUN, H., AND BÜYÜKÖZTÜRK, O. Probabilistic updating of building models using incomplete modal data. *Mechanical Systems and Signal Processing* (2016).

[93] SWEENEY, R. A., AND UNSWORTH, J. F. Bridge inspection practice: Two different north american railways. *Journal of Bridge Engineering 15*, 4 (2010), 439–444.

[94] THE WEATHER COMPANY, LLC. Weather History & Data Archive | Weather Underground, 2016. http://www.wunderground.com/history/.

[95] VACHON, M. *Dynamic Response of 3D Printed Beams with Damping Layers.* PhD thesis, Massachusetts Institute of Technology, May 2015.

[96] VAN DER AUWERAER, H., STEINBICHLER, H., VANLANDUIT, S., HABERSTOK, C., FREYMANN, R., STORER, D., AND LINET, V. Application of stroboscopic and pulsed-laser electronic speckle pattern interferometry (espi) to modal analysis problems. *Measurement Science and Technology 13*, 4 (2002), 451.

[97] WADHWA, N. *Revealing and Analyzing Imperceptible Deviations in Images and Videos*. PhD thesis, Massachusetts Institute of Technology, Feb 2016.

[98] WADHWA, N., DEKEL, T., WEI, D., DURAND, F., AND FREEMAN, W. T. Deviation magnification: revealing departures from ideal geometries. *ACM Transactions on Graphics (TOG) 34*, 6 (2015), 226.

[99] WADHWA, N., RUBINSTEIN, M., DURAND, F., AND FREEMAN, W. T. Phase-based video motion processing. *ACM Transactions on Graphics (TOG) 32*, 4 (2013), 80.

[100] WADHWA, N., RUBINSTEIN, M., DURAND, F., AND FREEMAN, W. T. Riesz pyramid for fast phase-based video magnification. In *Computational Photography (ICCP), 2014 IEEE International Conference on* (2014), IEEE.

[101] WAHBEH, A. M., CAFFREY, J. P., AND MASRI, S. F. A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Materials and Structures 12*, 5 (2003), 785.

[102] WILKINSON, J. H., WILKINSON, J. H., AND WILKINSON, J. H. *The algebraic eigenvalue problem*, vol. 87. Clarendon Press Oxford, 1965.

[103] WORDEN, K. Data processing and experiment design for the restoring force surface method, part i: integration and differentiation of measured time data. *Mechanical Systems and Signal Processing 4*, 4 (1990), 295–319.

[104] WU, H.-Y., RUBINSTEIN, M., SHIH, E., GUTTAG, J., DURAND, F., AND FREEMAN, W. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph. (Proc. SIGGRAPH) 31* (Aug 2012).

[105] XUE, T., RUBINSTEIN, M., WADHWA, N., LEVIN, A., DURAND, F., AND FREEMAN, W. T. Refraction wiggles for measuring fluid depth and velocity from video. In *Computer Vision–ECCV 2014*. Springer, 2014, pp. 767–782.

[106] YAN, Y., CHENG, L., WU, Z., AND YAM, L. Development in vibration-based structural damage detection technique. *Mechanical Systems and Signal Processing 21*, 5 (2007), 2198–2211.

[107] YAP, K. C. Efficient and robust solution of inverse structural dynamic problems for vibration health monitoring. In *Topics in Modal Analysis, Volume 7*. Springer, 2014, pp. 155–166.

[108] YODER, N. Peakfinder. http://www.mathworks.com/matlabcentral/fileexchange/25500-peakfinder-x0–sel–thresh–extrema–includeendpoints–interpolate-.

[109] YOUTUBE VIDEO, 2015. https://youtu.be/LA-ZcEWCOIc.

[110] YU, T.-Y., AND BÜYÜKÖZTÜRK, O. A far-field airborne radar ndt technique for detecting debonding in gfrp–retrofitted concrete structures. *NDT & E International 41*, 1 (2008), 10–24.

[111] ZABRODSKY, H., PELEG, S., AND AVNIR, D. Continuous symmetry measures. *Journal of the American Chemical Society 114*, 20 (1992), 7843–7851.

[112] ZABRODSKY, H., PELEG, S., AND AVNIR, D. Symmetry as a continuous feature. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 17*, 12 (1995), 1154–1166.

[113] ZHOU, Y. E. Assessment of bridge remaining fatigue life through field strain measurement. *Journal of Bridge Engineering 11*, 6 (2006), 737–744.

[114] ZIMMERMAN, D. C., AND KAOUK, M. Structural damage detection using a minimum rank update theory. *Journal of Vibration and Acoustics 116*, 2 (1994), 222–231.

[115] ZOLI, T. P. Memorial bridge structural plans, October 2012.