A COMPUTATIONAL FRAMEWORK FOR NEXT-GENERATION INSPECTION IMAGING

David Lattanzi

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Gregory R. Miller, Chair

Marc O. Eberhard

Duane W. Storti

Program Authorized to Offer Degree:
Civil and Environmental Engineering

University of Washington

Abstract

A COMPUTATIONAL FRAMEWORK FOR NEXT-GENERATION INSPECTION IMAGING

David Lattanzi

Chair of the Supervisory Committee:
Professor Gregory R. Miller
Civil and Environmental Engineering

Civil inspection images are notoriously difficult to interpret for those who were not present during an inspection, and so they have historically had limited value in ongoing condition monitoring applications. Within the last decade the advent of inexpensive digital imaging has compounded this problem by making it possible and common to capture and report numerous images, creating information overload that can easily overwhelm engineers and structure owners. More recently, developments in electronics manufacturing and robotics have made automated structural inspection an emergent technology for civil engineers, which inevitably is leading to even more digital image data that must be handled. This dissertation presents a systematic approach to capturing, processing, and representing inspection image data such that the inherent value of these increasingly large data sets can be realized. Using a data pipeline combining automated image capture, contextualized 3D visualization, and robust computational imaging and regression techniques, the goal is to allow engineers to view inspection images in their 3D spatial context, while aiding them through enhanced damage detection routines, all while helping to minimize field inspection disruptions through the use of robotic imaging. In the development of this pipeline and an associated prototype implementation, several key challenges have

been addressed: (i) automated systems capable of comprehensive field imaging; (ii) 3D reconstruction algorithms which provide *accurate*, photorealistic image interpretations; (iii) robust computer vision algorithms suitable for field applications; and (iv) nonlinear regression models which correlate the relationships between extracted image information and structural performance. The results of prototype testing show that, given due consideration to the inherently large data sets that robots produce, systematic imaging can enable entirely new ways of visualizing and interacting with inspection information.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost I would like to thank my advisor, Dr. Greg Miller, for his guidance throughout my research at the University of Washington. It has been a rare pleasure to work with someone so supportive of academic exploration and innovation, to say nothing of his seemingly bottomless patience. My committee members (Drs. Marc Eberhard, Laura Lowes, Duane Storti, and Jim Pfaendtner) provided invaluable feedback on my work not only with regards to the technical component, but also with respect to my nascent academic career. I owe an enormous debt of gratitude to my family, and in particular my parents Terry and Dan, for their constant support of my (highly nonlinear) career path. Lastly, I must thank my wife Elisabeth. For close to a decade her love and encouragement have allowed me to pursue my research goals, and her quietly persistent faith in me has served as the foundation of my career. To say that I could not have completed my degree without her would be an outrageous understatement.

DEDICATION

For A+B

Chapter 1

DISSERTATION OVERVIEW

Structural inspection and monitoring can be thought of as producing three types of information: observations, images, and measurements (including sensor measurements). In current practice, captured images are decontextualized and underutilized. The purpose of this research has been to rethink how we capture and handle image information, and to explore how a holistic approach to inspection imaging leads to fundamentally new sources of information for inspectors and engineers. The result is a systems-level concept and associated prototype implementation, illustrated in Figure 1.1.

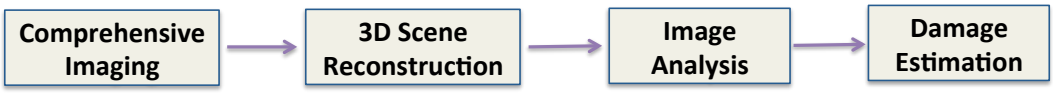


Figure 1.1: A conceptual overview of the image analysis pipeline

The research effort spans the length of this "information pipeline" from image capture to data contextualization and interpretation to damage estimation, exploring the challenges inherent to each step of this new process, and proposing solutions where necessary. The technical challenges to be addressed included: (i) prototyping of an automated system capable of comprehensive field imaging; (ii) development of 3D reconstruction algorithms that provide *accurate*, *scalable*, and *photorealistic* image interpretations; (iii) development of robust computer vision algorithms suitable for field applications; and (iv) the creation of statistical regression models which correlate the relationships between extracted image information and structural performance. Accompanying the overall system concept, prototype solutions to these challenges are provided within this dissertation as self-contained chapters and are the primary contributions of this research. The content of the dissertation is as follows:

Chapter 2

This chapter begins with a discussion of the current state of practice for bridge inspection and assessment, along with the related discipline of structural health monitoring (SHM). A literature review is provided. This is followed by a presentation of the systems-level approach to inspection imaging which forms the primary hypothesis of this research, as well as an overview of the author's contributions to this field.

Chapter 3

Chapter 3 is a self-contained paper on the topic of robotic systems for structural inspection. Structured as a survey paper, the emphasis is on considerations and metrics for developing robots specifically for structural inspections. Any such robot must have sufficient maneuverability to access often complex 3D structures. Relatedly, inspection robots must guarantee, either through human piloting or autonomous path-planning algorithms, that the entire structure is inspected comprehensively. Lastly, proper sensors and instrumentation must be chosen so as to detect structural damage and deterioration. The paper concludes with a discussion of outstanding challenges in the field, as well as emergent technologies.

Chapter 4

Chapter 4 is a self-contained paper which provides a comparative study of two common 3D scene reconstruction techniques — the process of constructing 3D models from 2D images — in the context of automated structural inspection (e.g. via unmanned aerial drones). Structure from Motion and Image Mosaicing are considered. Metrics were developed to compare the two techniques: (i) scalability and interactive renderability; (ii) geometric reconstruction accuracy; (iii) reconstruction artifacts nature and prevalence. A series of model reconstructions using both approaches were developed and compared using these metrics.

Chapter 5

Chapter 5 is a self-contained paper which presents a computer vision algorithm supporting automated surface crack detection in structural concrete. The presented algorithm is designed with the necessary robustness for field implementation in mind, while being capable of operating on the large image sets provided by automated imaging. Specifically, a clustering segmentation and feature extraction scheme that exploits inherent characteristics of fracture images was developed and tested. The approach is shown to perform well on a wide range of highly varied images and provides an improvement in detection accuracy relative to existing techniques.

Chapter 6

Chapter 6 is a self-contained paper which investigates the potential of regression-based techniques utilizing computational computer vision information to estimate post-earthquake damage in structural bridge columns. Images taken during a series of tests on socket column bridge footings were analyzed using crack and spall detection and parameterization algorithms. The numerical parameterizations were then correlated to the known level of maximum drift at the time the image was taken, a key indicator of structural damage. The results indicate that there is a strong relationship between extracted image information and structural damage. However, further work is necessary to develop similar patterns for other structural elements and loading patterns. Recommendations for optimizing future camera placements during testing are also provided.

Chapter 7

Chapter 7 outlines suggested directions for future research. The overarching system-level hypothesis has received proof-of-concept validation by means of the work presented in the preceding chapters. Future work seeks to explore not only how to

extend such a system, but how to integrate it more fully into the larger structural health monitoring paradigm.

Chapter 2

RESEARCH MOTIVATION AND HYPOTHESIS

2.1 Current Bridge Inspection Protocols

Bridges represent a critical component of infrastructure systems, and therefore condition monitoring via periodic inspection has long been a key part of bridge operations and maintenance practice. There are more than 578,000 bridges in the US alone, most of which must be inspected every two years, and so hundreds of millions of dollars per year are spent on the inspection of trillions of dollars in assets. There are also substantial indirect costs associated with required lane closures and related traffic disruptions. Making bridge inspection less costly, less obtrusive, more quantitative, and more effective in regards to the type and quality of data collected thus can lead to significant economic savings and safety improvements. This includes reductions in both the direct and indirect costs of the inspections themselves, the avoidance of unnecessary repairs, the timely implementation of needed repairs, and the opportunity for improved engineering that comes from improved understanding of field performance of designs over time.

Current bridge inspection technology typically requires an inspection team and support equipment to travel to a given bridge to make a series of qualitative observations. Thus, there are a number of personnel, equipment, and travel costs inherent in this approach that scale linearly with the number of bridges needing inspection, the frequency of inspection, the distance between bridges, and the life of the bridges. During a structural inspection, inspectors capture local measurements and close-up images of critical bridge components and damage. However, these photographs and measurements are notoriously difficult to comprehend by anyone other than the inspection team themselves, as they are captured and organized in such a way that they are decontextualized. The images are also presented in a qualitative manner, with no standard methods in place to objectively analyze them. An inspection report which includes the decontextualized data is passed to an engineer who is responsible for assessing deterioration to the structure, considering the temporal context of

the inspection information. The engineer then passes an assessment to managers charged with high-level policy decisions concerning system-wide resource allocations. In general, the movement of data up through this decision-making hierarchy results in huge losses of potentially useful and critical information, and this is equally true in regards to the the transfer of data across time (i.e., from inspection to inspection). Contrasting the effort and expense required to complete a typical bridge inspection with the resulting outcome of passing up the chain a set of coarse-grained, essentially qualitative assessments, it is clear that the cost-to-information ratio is unfavorably high.

The reliability and consistency of the bridge inspection process was studied at length by Phares et al. [2004]. The results of 49 individual inspections were compared against an agreed upon standardized inspection. Condition assessments showed significant variability, with 95% of all element ratings varying by more than 2 condition levels (on a 1-10 scale). There was also substantial variability in both the quality and quantify of field notes taken, as well as no consistency in how photographs were taken.

The high degree of variability in current inspection practices inhibit structure owners and future inspection teams from properly assessing the previous condition of a structure and make life-cycle projections highly ambiguous. Furthermore, the inherently qualitative nature of the process prevents engineers from repeatably and reliably estimating current and future structural performance.

2.2 Current State of Structural Health Monitoring Practices

In order to address the shortcomings of current inspection practice, the most recent decade has seen rapid growth in regards to the use of embedded, fixed sensors for monitoring various kinds of structural systems in a number of contexts, including applications such as damage, deterioration, and health assessments. Referred to as structural health monitoring (SHM), it could ideally lead to unsupervised monitoring

and condition assessment of infrastructure, reducing the need for disruptive and costly human inspections. It could also drastically improve reliability assessments.

There are, in general, two sensing paradigms. Local sensors, of which strain gauges are the most common example, capture material response at what is effectively a single point on the structure. Since damage is not known a priori, local sensors have inherent limitations with regards to SHM. One possible method of overcoming this local sensing limitation is to embed an extremely dense array of local sensors within the structure at the time of construction. However, the lifespan of these sensing systems are not as long as as those of the structures they monitor. The materials and electronics used to create sensing systems (e.g. piezoelectrics, silicon, constantan) experience permanent performance degradation with repeated use and are therefore not suitable for long-term monitoring.

The result is that in most cases global sensors, which measure the *system* response and can be easily replaced during the lifespan of the structure, are the preferred technology for health monitoring. Accelerometers are the most commonly used sensor for this purpose. The basic principle for accelerometer sensing is simple: observe key aspects of the global dynamic response of a structural system over time, and relate changes in response to changes in the condition of the structure itself. The actual application of this concept is quite complex, and much of the activity in this area has focused on the development of sophisticated pattern recognition systems capable of processing and organizing data obtained from an accelerometer array [Cruz and Salgado, 2009].

Modern SHM systems are capable of detecting the fact that some kinds of damage have occurred and, depending on sensor density, can in some instances determine the general location of the damage[Elbehairy, 2007]. However, the definition, quantification, and diagnosis of damage eludes modern systems. The reality is that the engineer is attempting to locate and quantify local damage based on the global system response, a challenging task in direct opposition to St. Venant's principle, which

states that local effects decay rapidly with distance.

2.2.1 Literature Review

No attempt will be made here to review the structural health monitoring field comprehensively—there are a number of representative summary and overview papers that have been published recently. The survey that follows is meant to illustrate the state of practice as it pertains to the current research project.

Worden et al. [2007] discussed the general concepts and limitations of structural health monitoring. They defined SHM as a 4-part process: (i) damage detection; (ii) damage localization; (iii) damage categorization; (iv) damage quantification. With regards to this process, a series of axioms has been established.

- All materials have inherent defects.
- Damage assessment involves a comparison between two system states.
- Damage detection and localization can be done using unsupervised learning, however categorization and quantification can only be performed using supervised learning techniques.
- Sensors cannot measure damage and sensor measurements require intelligent feature extraction for that purpose
- Damage initiation and evolution time and length scales dictate the sensor system's requirements
- Higher sensitivity to damage increases sensitivity to noise
- The frequency range of excitation governs the size of damage that can be detected

Wang and Chan [2009] provided an overview of vibration based SHM techniques. Detection is performed by analyzing the dynamic response of the structure and monitoring damage sensitive parameters. Modal shift analysis is limited to severe damage

events due to noise corruption of the sensed ambient dynamic response. Various methods of dynamic parameter extraction are discussed. They conclude that while SHM methods are capable of damage detection and general localization, quantification of damage or long-term life predictions is not currently feasible.

Fritzen and Kraemer [2009] also provided a discussion of modern SHM, including both time and frequency domain techniques, as well as modal analysis concepts. Limitations due to sensor breakdown, environmental changes, model dependencies, and sparse sensor networks are explained. Case studies in aerospace, civil, and mechanical engineering are used to illustrate the capacity of modern SHM systems.

Cruz and Salgado [2009] evaluated six damage detection techniques, including two case studies. COMAC (coordinate modal assurance criteria), which analyzes modal displacements, has been found to be ineffective for all but the most severe damage. Curvature and strain energy based techniques offer a more sensitive analysis, but an accurate portrayal of structural mode shapes is essential for good results. Wavelet based techniques which operate at varying time and frequency resolutions, as opposed to Fourier techniques that only operate in the frequency domain, have become much more prevalent in recent years. However they are prone to noise contamination and are therefore better suited to severe damage scenarios. Of the two case studies, one was a simulated event and the other a field test. The simulated case study showed that noise levels as low as 1.0% of the signal amplitude dramatically decrease the effectiveness of any vibration based method. Applications for wavelet based damage detection were discussed in Jiang and Adeli [2007]. They used the multi-resolution properties of wavelet signal processing, combined with a feedback neural network to create a damage detection algorithm capable of overcoming some of the noise issues inherent to wavelet decomposition. However, the resulting algorithm still required the analysis of damage data on the level of ambient noise.

Friswell [2007] provided an overview of inverse methods for damage localization and parameterization. Inverse methods attempt to localize detected damage by sys-

tematically altering a highly detailed finite element model of a structure to match in-service vibration information. There are a wide range of methods associated with this technique, however all of them suffer from a dependency on a detailed finite element model of known accuracy. The reliance on models also means that structural damage, which is indicated by relatively small changes in dynamic response, is often on the order of errors due to imprecise modeling.

Sohn and Farrar [2001] developed a time domain analysis technique for damage detection in machinery. Their technique relies neither on accurate finite element modeling nor advanced knowledge of modal shapes. Time history acceleration data is broken into subdivisions and an autoregressive weighting function is used to represent accelerometer readings within these subdivisions. As damage accumulates, residual error between the current system function weights and the weights from a known undamaged condition increases. Larger errors will accumulate from sensors near damaged locations. Damage detection and localization success rates were high. Sensor density is the limiting factor for the system. Their system was implemented for civil structures by Lu et al. [2008] as part of a system designed to provide online wireless structure monitoring. The time domain method was found to be well-suited to embedded computer processing, however the sensitivity of the error function was found to be on the order of low-level white noise errors, mitigating the effectiveness of the approach.

An overview of time-series based methods is provided by Fassois and Sakellariou [2007]. Relevant to the current research project is that, while most time-series methods are capable of damage detection and some degree of localization, damage quantification is not possible without prior knowledge of the behavior of a structure in a damaged state. This knowledge can stem either from previous monitoring or finite element analysis. This is less of a concern for the monitoring of manufactured equipment, for which such data may be available, but is a more serious issue for civil structures, which are constructed on site rather than manufactured, and are difficult

to model with a high degree of accuracy.

A general discussion of the global/local sensing issue is presented by Mal et al. [2007]. The inherent difficulty of using a sparse network of accelerometers to detect local damage is expressed, and the various methods used to extract damage sensitive parameters from global dynamic response are examined. The authors propose the use of both local (in this case, ultrasound) and global sensors to monitor structures. To this end, they have developed a damage detection method based on time-series data which can universally process global and local sensor data.

Lynch and Loh [2006] performed a survey of wireless vibration sensing technologies. Wireless monitoring of structures is often advantageous due to reduced installation costs and fewer maintenance issues compared to wired sensor networks, along with an increased capability for embedment within structures. Furthermore, wireless systems can easily contain on-board capabilities for processing sensor data. This decentralized data management approach leads to reduced sensor power demands due to more limited data transmissions. However, wireless sensors still have power demands beyond what is possible with energy harvesting systems and the majority of research in this area has been directed towards power minimization strategies. Furthermore, these systems are still limited by the inherent gap between local phenomenon and global response.

Several researchers have studied the long-term performance of sensors, an important aspect of embedded "cradle to grave" SHM. Middleton [1986] tested the endurance of strain gauges for fatigue testing purposes. Most strain gauges failed at cycle counts on the order of 10^4, with a maximum on the order of 10^6. Inaudi [2004] tested the durability and reliability of fiber optic strain sensors. Strain gauges were embedded in concrete bridges and monitored over a 7 year span. Failure rates for strain sensors were 1% per year. Cyclic loading was shown to have no effect. Testing by Muhlstein et al. [2001] indicate that MEMS sensors (commonly used for accelerometers) experience nonlinear changes in compliance and resonant frequency

at cycle counts below what is typically experienced during a bridge's lifespan.

Brownjohn [2007] discussed issues specific to civil infrastructure monitoring. He emphasized that the breadth of structural system types means that it is challenging to implement a universal approach to monitoring. Furthermore, it is reiterated that localized damage detection and quantification is still far from being a reality outside of laboratory situations. Most in-service monitoring systems are devoted to tracking known defects or logging stochastic loadings such as wind or heavy truck traffic. He highlights several key problems with civil monitoring: a lack of established cost-benefit analyses to supply to owners, inappropriate instrumentation or over-instrumentation, the resulting data overload and storage problems, sensor communication reliability, compensating for environmental factors, and an overall lack of funding and collaboration with regards to SHM research.

The integration of SHM data into the larger framework of damage prognosis is discussed in Farrar and Lieven [2007]. Damage prognosis is defined as the process of assessing the current state of a structure and predicting future degradation and remaining lifespan. Farrar states that the challenge of prognosis is in integrating sensing systems with computational intelligence techniques that draw upon past experiences as well as analytical system models for the accurate prediction of remaining system lifespans.

Liu et al. [2010] suggested a first step towards condition assessment and damage prognosis. They paired an online damage monitoring system with an offline statistical damage prognosis algorithm populated with data from known fatigue failures. They were able to successfully predict future remaining fatigue life based on the statistical predictions, but their system relied on a database of known damage information as well as preknowledge of the failure mechanism, in this case fatigue.

One well known example of vibration based bridge monitoring was undertaken by Soyoz and Feng [2009]. A three span continuous post-tensioned box girder bridge was instrumented with wireless accelerometers placed at quarter-points along the

span. Monitoring was performed over a 5 year span, beginning 4 years after the bridge was put into service. Frequency domain techniques were utilized in assessing changes to global modal parameters. High variation in the modal parameters (on the order of ±10%) was encountered, due to mass participation from passing vehicles and thermal effects. Modal frequencies were estimated to decrease 5% over the the 5-year monitoring period. A neural network approach was used to correlate the shifts in global modal behavior to more localized stiffness degradation. A 2% decrease in structural stiffness was estimated, but with variations on the order of ±3%.

2.3 Systems Approach to Inspection Imaging

Bridge inspections continue to operate in much the same way as they have for decades, using subjective and highly variable visual methods and a limited information pipeline for communicating inspection information. SHM technologies, while very effective at certain tasks, have fundamental limitations stemming from the need to extract localized structural performance characteristics from inherently global sensor information. The net result is that, despite sweeping changes to computing and sensing technologies in the last two decades, civil infrastructure managers have so far been unable to leverage these technologies for substantive improvements in maintenance practice.

Viewed through a different prism, the limitations of modern bridge maintenance addressed in this research are threefold:

- Engineers disregard massive amounts of potentially valuable and quantitative data in the form of unprocessed and decontextualized images.
- Human factors associated with the tedious, repetitive, and disruptive nature of current inspection practices limit the repeatability and quantification of inspections while adding to their cost.
- Current monitoring methods are incapable of associating inherently local disturbances with global system performance and long-term trends, the ultimate

goals of any monitoring and maintenance practice.

This work presents an alternative, complementary inspection framework which systematizes each step in the processing of inspection imaging (Figure 2.1). This framework uses automated mobile imaging methods that in principle can deliver image data from anywhere (and thus everywhere) on a structure. In conjunction with systematic and automated inspections, 3D scene reconstruction and computer vision methods are used to supplement the traditional qualitative visual inspection process and potentially increase the effectiveness of embedded sensor systems. The resulting computational image data are used not only in a damage *detection* role, but also serves as the foundation of nonlinear regression methods which correlate visually apparent damage information with actual levels of structural performance, allowing for damage *estimation* rooted in image analysis. The key in all these contexts is to have a framework that controls the flow and processing of image information in a manner guaranteeing its comprehensivity and contextualizes it both temporally and spatially, thus enabling robust computational analyses.

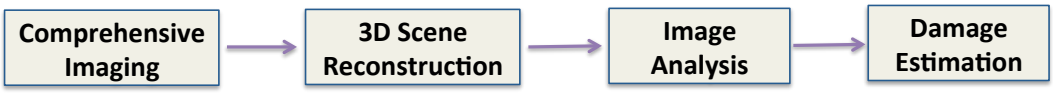


Figure 2.1: The proposed bridge inspection imaging framework

It is useful to consider a future scenario in which the envisioned technologies are fully realized. In this scenario an engineer receives notice that new inspection data have been obtained from a particular bridge via periodic robotic inspection requiring no traffic disturbance, crew travel/safety risk, or equipment rental. The engineer brings up the data in the context of an interactive geometric model of the bridge and queries the system to show all new indications of damage or deterioration since the previous inspection. Each indicated location can then be zoomed in on and examined

with in-context, high-resolution images simulating the experience of viewing the structure in the field, but with optional image enhancement highlighting damage. A second query could then request an update of all previously existing indications of damage or deterioration, followed by zoomed-in animations of the evolution of the damage over time. A third query could request that a particular area of damage be compared against a national database to compare rates of growth relative to similar flaws in similar classes of structures, which could further be used to extrapolate growth rates. A fourth query could ask for correlation between damage growth between inspections and embedded sensor data indicating loading and other environmental history during the period in question. In the case of ambiguous or unclear visual data, additional inspection could be requested using alternative sensing modalities (e.g., ultrasound, infrared, etc.), and so on. Lastly, damage patterns could be used to estimate mechanical deterioration in the structure through regression models constructed from experimental testing.

2.3.1 Contributions

In order to enable this pipeline, several outstanding obstacles are addressed in this dissertation:

- a need for a data-collecting front-end system capable of comprehensive imaging for real bridges in the field (Chapter 3)
- a means of accurately storing and representing inspection information in a photorealistic 3D context (Chapter 4)
- a need for computer vision algorithms capable of robustly handling highly variable field image scenarios (Chapter 5)
- knowledge of the correlations between parameterized image information and sensor data (Chapter 6)

In addition to the development of the overall pipeline, the initial solutions to these obstacles are the specific contributions of this research project, and are the emphasis of the four self-contained papers which make up the body of this dissertation.

Chapter 3

ROBOTIC INSPECTION CONSIDERATIONS: A SURVEY

3.1 Introduction

The rehabilitation and replacement of aging infrastructure, such as buildings and bridges, represents an increasingly critical component of government and private industry budgets, and so condition monitoring via periodic inspection has long been a key part of infrastructure operations and maintenance practice. This paper presents a survey of the state of the art of automated inspection based on mobile platforms (i.e., robots), which while still in a largely nascent phase of development, can offer significant opportunities for improved inspection and monitoring.

The purpose of any inspection regimen is to provide engineers and infrastructure managers with information that helps them assess the need for repairs, as well as to project the remaining in-service life of the structure. During each inspection, the goal is to find and assess signs of structural damage and distress. However, what constitutes damage and how it is observed are not always readily apparent. The deterioration of materials is not always directly observable and often occurs internal to the structure itself. Therefore, most inspections look for changes from some baseline, as-constructed state of the structure. Inspectors search not only for obvious damage such as material cracking and rusting, but also for mechanical deformations and other indicators of distress, such as water intrusion. This leads to three critical criteria which need to be met by any inspection methodology. The inspection must *comprehensively* inspect the structure, observing every possible surface for signs of damage. This observation must be done *repeatably*, over the course of decades, to facilitate long-term life-cycle projections. Lastly, measurement tools must be available to assess the condition of the structure *quantitatively* where possible.

Current inspection methods and standards have been in place for decades. Consider the process of routine bridge inspection. The typical bridge inspection requires an inspection team and support equipment to travel to a given bridge to make a series of qualitative observations. The inspection process is primarily visual and teams typ-

ically spend their time closely examining all of the core components of the structure, taking detailed photographs and logging observations. This element-based assessment is submitted as part of a detailed report to the structure owner along with inspection images and any field measurements. The report is archived and compared against previous inspection reports to provide the basis for a long-term life-cycle maintenance strategy for the bridge. Studies of the reliability and consistency of this inspection process reveal substantial variability in both the quality and quantity of field notes taken, as well as little consistency in how photographs are taken [Phares et al., 2004]. The inherent limitations of current bridge inspection practice extend to the inspection of other structures as well.

Compounding the problem is that the advent of digital imaging has enabled reports to include potentially hundreds of inspection images, many of which are spatially decontextualized, exacerbating information overload and providing limited additional information. This lack of inspection consistency, ineffective image communication, and corresponding data overload inhibit structure owners and future inspection teams from properly assessing the previous condition of a structure and make life-cycle projections unreliable.

Another key limitation to current inspection approaches is the accessibility of infrastructure systems. In many cases, key system components are either completely or partially inaccessible to inspectors. Often, visual access to these components can only be gained through great expense and disruption to the general public. There are also substantial indirect costs associated with required lane closures and related traffic disruptions for transportation structures. It is difficult to quantify the costs associated with either failing to detect critical damage or to err on the side of making unneeded repairs, but they are undoubtedly substantial, as well. It is not surprising that there have been a variety of efforts to make infrastructure monitoring less costly, less obtrusive, more quantitative, and more effective in regards to the type and quality of data collected.

3.1.1 Robotic Infrastructure Inspection

Largely to address the aforementioned limitations of current inspection methods, a new and emerging field of inspection robotics is developing. Robots promise less obtrusive and lower cost inspections, as well as increased objectivity and repeatability for routine inspection tasks. In terms of robot development, the primary challenges of a structural inspection are robust mobility, comprehensive path planning, and damage perception. Perhaps the biggest impediment to widespread implementation of inspection robots is that civil structures are rarely uniform in shape and scale, and generally exist in exposed outdoor environments. Furthermore, structures are not typically designed with considerations for easy inspection access and often require scaling hundreds of vertical feet or navigation around complex 3D obstacles, all of which makes robot mobility in the inspection environment highly challenging. Inspection robots must also be capable of comprehensively inspecting a structure in order to guarantee that its safety for use by the general population. This often requires the development of motion path planning algorithms that guarantee complete structure coverage by the chosen sensor array. Lastly, an appropriate sensor array must be chosen to detect damage, wholly separate from sensors used for navigation and control. Compounding the difficulty is that sensing must sometimes be done in murky, low-light environments which degrades the usefulness of commonly used sensors such as cameras.

3.1.2 Considerations of This Survey

Developing general-purpose robots that are capable of completely tackling all of these challenges consistently in a wide variety of inspection environments is impractical in the short term. Instead, more specialized robots have been developed that are designed to help detect some of these changes or to give inspectors additional performance information from non-visual sensors. The implementations of these robotic

inspection systems and their associated research efforts are disparate due to the interdisciplinary nature of the field. The purpose of this survey is to agglomerate these efforts to represent the state of the modern field and to give researchers a clearer understanding of the outstanding challenges.

A precise definition of what a civil structure is, and therefore exactly what should be included in such a survey, does not exist. This survey will focus on the general class of constructed infrastructure facilities that require routine or special case inspections to assess their long-term viability. Included facilities for the purposes of this survey:

- Buildings
- Bridges
- Tunnels
- Roadways
- Storage Tanks
- Marine Structures

A key distinction to make is between robots that are used to assess structural integrity, as opposed to robots that are designed to *navigate* infrastructure such as those used to find survivors of a disaster. The latter have distinctly different sensing and mobility criteria and are not included in this survey. Pipe inspection, which is a mature field for inspection robotics and implementations, will not be discussed here. The previous survey of pipe inspection robots by Tur and Garthwaite [2010] covers this topic. Transmission line inspection robots were covered by Toussaint et al. [2009] and will also not be discussed here.

This survey is organized into two primary sections. Discussed first are the fundamental mechanisms and methods for mobility, comprehensive inspection, and damage

perception that have been implemented for robotic inspection. The second section presents case study implementations of inspection robots. Some of the presented robots were developed to study fundamental technical challenges for inspection robots, such as mobility in inspection environments, while others represent full implementations being used commercially or in field testing. The survey concludes with a discussion of outstanding challenges and emerging technologies.

3.2 Fundamental Mechanisms and Methods

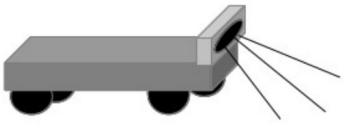
3.2.1 Mobility

The majority of case studies and prototypes for inspections have focused on the problem of robust mobility for structural access. The approaches that engineers have taken are varied but in general the associated locomotive mechanisms can be categorized as ground-based, crawling, climbing, aerial, and marine propulsion (Figure 3.1). The choice of mechanism must balance the trade-off between the need for robust, stable locomotion with the flexibility required for a robot to gain access to all portions of the structure.

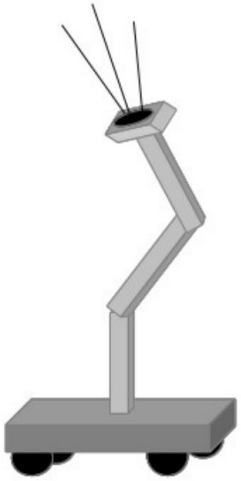
Ground-based Mechanisms

Many successful inspection robots have operated on simple wheeled or continuous tread bases. In highly unstructured environments, as in the case of post-disaster assessment, these may be the preferred option due to the inherent robustness. Such ground-based locomotion is also sensible in scenarios like tunnel inspection. Tunnel inspection typically involves only the inspection of the interior lining of the tunnel itself. This is a tedious and time-consuming task for a human, but it is also a task which poses relatively little difficulty in the way of access to key inspection features. Similarly, ground-based robots are a suitable choice for pavement inspection.

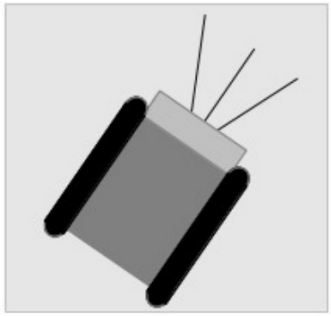
An extension of the ground based locomotion concept has been to affix more



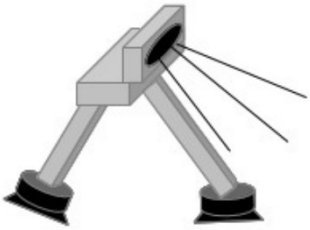
(a) Ground-based



(b) Ground-based with articulator



(c) Vertical crawler



(d) Climbing/walking

Figure 3.1: Mobility schematics for selected inspection robots

sophisticated articulators and end effectors to a wheeled base to provide access to more difficult to reach portions of structures. This combination has proven especially capable in scenarios such as bridge inspection, where complex 3D articulation is a necessity.

Crawling Mechanisms

An adaptation of the ground-based mobility approach is to affix the robot to the structure under inspection.This allows for access to difficult to access portions of the structure, such as high elevation locations, and traversal of vertical surfaces. Such robots are referred to here as "crawlers". Crawlers have been widely implemented for the inspection of steel storage tanks, with magnetic adhesion allowing the crawlers to scale walls. While magnetic continuous treads are commonly used for adhesion to the structure, some implementations are clamped to the structure itself. Crawling is also the most common method for inspecting structural cables [Wang and Xu, 2007]. In most cases, the biggest challenge is in creating crawlers capable of avoiding obstacles, due to their need to be affixed to the structure for support. With regards to the trade-off between stability and flexibility, crawling robots tend to maintain highly accurate motion at the cost of limited flexibility.

Climbing Mechanisms

Instead of being constantly affixed to the structure, climbing robots are able to partially disconnect from the structure. This enables the robot to more easily avoid obstacles and move between individual structural elements. These robots typically attempt to minimize the kinematic degrees of freedom (DOF) of the robot while maximizing potential articulation so as to minimize the weight of the robot [Balaguer et al., 2005]. For most climbing robots, the core problem is how such robots affix themselves to the structure. The chosen adherence mechanism for an inspection robot is often determined by the type of structure being inspected. Traditional

grasping methods, which attempt to grab an arbitrary structure for adhesion, have not been historically popular. Reliable grasping of structures is challenging due to highly variable structural shapes and the large flat surfaces that robots must often traverse. Suction based adhesion has been far more popular. Attachment is gained by creating a suction between the end effectors of the robot and the structural surface. The other core issue is how to carry an often heavy payload reliably, occasionally in an inverted position. The capability to move between planar surfaces (e.g. floor to wall) is also valued in an inspection environment and is a major advantage of climbing robots over crawlers.

Aerial and Marine Propulsion

Entirely different methods of locomotion for inspection are utilized by unmanned aerial vehicles (UAV) and unmanned marine vehicles (UMV). Using propulsion through air or water for locomotion, these systems maximize system flexibility at the expense of state estimation accuracy, as these systems are far more susceptible to environmental control disturbances which disrupt localization and motion planning. Because of these control difficulties, most UAV and UMV systems are remotely operated vehicles (ROV), eschewing autonomy for human-led control and navigation. UAVs include quadcopter drones, rotary helicopters, and fixed-wing systems. UMVs are often classified as unmanned surface vehicle (USV) or unmanned underwater vehicle (UUV) depending on whether or not the robot is capable of operating below the water surface. There are many commercially available UAV and UMV systems, including the Parrot drone [Bristeau et al., 2011] and COMETS [Ollero et al., 2004]. UMVs include the VideoRay [Wang and Clark, 2007] utilized in several of the included case studies.

3.2.2 Comprehensive Path Planning and Autonomy

To date, robot autonomy has not been emphasized generally for structural inspection. Most of the presented robots are tele-operated, in particular those which have been

implemented for field testing. The exception are robots developed to inspect storage tanks. The relatively simple shapes of these structures have facilitated the use of coverage planning algorithms.

In order for autonomous inspection robots to gain acceptance as tools for civil inspection, they must be able to guarantee repeatable and comprehensive inspection through the application of planning algorithms. These path planning algorithms are, in fact, typically coverage algorithms which strive to minimize the time and energy required to guarantee a complete observation. The complex 3D nature of civil structures often requires path planning in 3D, leading potentially to high-dimensional configuration spaces and complex motion paths. Furthermore, the large scale of civil infrastructure often prohibits the use of tethered power systems, so the ability to efficiently inspect every facet of a structure is essential due to battery life limitations.

3.2.3 Sensing and Perception Mechanisms

Perception mechanisms for inspection robots, in addition to those necessary for control and localization, are critical for accurate and reliable inspection of civil structures. The sensors chosen should be able to represent as many of the characteristics of structural distress as possible. This distress can manifest itself as changes to the surface appearance of the structure, such as is the case with steel corrosion or concrete cracking. But it may also exist as excessive deformation of the structure itself or as distress which occurs below the surface of the structure. One example of this kind of damage is corrosion in concrete rebar, which is sometimes only detectable by non-visual methods.

The sensors used in inspection robots run the gamut of available sensing technologies. Often payload capacity and fundamental sensor demands dictate the implemented mechanisms, as is the case with lightweight climbing robots and small UAVs. By far, the most common sensors used by inspection robots are cameras. In most applications, robots are directed simply to capture images of a structure, often in

the form of video. This sort of application is similar to current, human-led visual inspections and so is an intuitive approach for robots, especially tele-operated ROVs. However, there is typically a lack of back-end processing which can result in human inspectors watching hours of inspection footage, looking for flaws. This issue has caused a perceived lack of value for robotic inspection amongst structure owners and managers.

Rangefinding sensors, such as sonar, have seen use in murky marine environments where imaging is difficult. Sonar has been used to detect bridge pier scour, as well as subsurface damage to dams. Infrared rangefinders, and the associated RGB-D cameras, have seen less use due to the highly variable infrared response of civil structure in changing environmental conditions.

Of key interest to structure engineers and managers is the capability of robots to not only capture the surface appearance of structures, but to perceive the internal integrity. There are a wide range of tomographic sensing methods used in robotic inspections, with ultrasound being the most common. However, ultrasound requires direct contact between the transducer and the inspection surface, which makes it impractical for many applications. Magnetic flux-leakage is commonly used to inspect the integrity of steel structures. The detection of flux-leakage corresponds to a loss of integrity of steel structures, with the sensor detecting disruptions in a produced magnetic field due to this leakage. This sensing method has been used to detect damage in steel storage tanks, where imaging is difficult or impossible. Other tomographic sensors which have been used for field structural inspection include ground-penetrating radar, thermographic, and electromagnetic methods.

3.3 Case Studies and Field Implementations

The previous section presented an overview of means and methods for robotic structural inspection. In this section, the individual case studies are presented and discussed. The survey of developed structural inspection robots is subdivided based on

application:

- Bridge and tunnel inspection
- Storage tank inspection
- Post-disaster inspection and assessment
- General inspection (no specific application)
- Miscellaneous applications

The presented case studies utilize a broad range of locomotive mechanisms to gain inspection access. Comprehensive inspection is achieved primarily through the use of a human pilot, with limited robot autonomy. In most cases, CCD cameras are the preferred sensing paradigm, although other sensing methods are explored.

3.3.1 Bridge & Tunnel Inspection Robots

For bridge and tunnel structures, even small flaws can indicate a serious risk to the public. Therefore, routine and comprehensive visual inspections in which all available facets of the structure are observed are mandated by federal and state laws. For human inspectors, this often means sophisticated and expensive climbing equipment and machinery or, in the case of marine bridge pier inspections, highly skilled dive teams. Disruptions to the general public are also a major concern, as is the safety of the public and the inspection team during an inspection. There are readily apparent benefits to using robots in this context, however the complexity of the structures is a significant impediment. Most robots developed for these applications tend to focus on gaining inspection access.

One of the most common and valuable pieces of inspection equipment is the "snooper" truck. A snooper uses a large, articulating arm with an inspection platform

to give inspectors visual access to the underside of bridge. Several researchers have developed robotic inspection systems which use a snooper as a basic platform. Tung et al. [2002] prototyped one such inspector. Their access solution consists of a simple mechanical arm transported by the snooper. They developed an edge-finding based crack detector as a complement to the system. A similar system was developed by Oh et al. [2007]. Their system is transported along a series of rails (connected to the snooper) to simplify controls. Modularity was emphasized to decrease both initial and long term costs of the system. Similarly, Lee et al. [2008] used a rail-guided system(Figure 3.2e). In all cases, damage detection is performed using a CCD camera, which also enables teleoperation capabilities.

Huston et al. [2003] proposed adaptive, mobile sensor arrays capable of moving to areas of detected importance. Complex, highly flexible robot concepts were avoided in favor of low cost and simplicity. Mobility concerns, such as obstacle avoidance, and autonomy were not considered. This led to the development of a robotic prototype which clamps to the bottom of a steel beam and uses a video camera for damage perception (Figure 3.3a).

Choset [2000] developed an inspection robot in response to the extensive rigging and traffic control costs inherent in standard bridge inspections. The robot was of the serpentine arm (proposed 12 degrees of freedom) variety so as to provide optimized manipulation for bridge truss access, and was envisioned to ride on a larger vehicle for maximum mobility. The recognized high motor and stiffness demands for this approach, which requires long arm lengths, required the use of highly efficient arm joints and motor assemblies. A resulting prototype, while only incorporates 2 degrees of freedom, was tested. While the challenges of coverage planning for a 12 DOF serpentine robot were recognized, they were not considered in the final prototype, which is tele-operated(Figure 3.2a). Perception and arm control are performed using a camera as an end-effector.

An unmanned aerial vehichle (UAV) was developed for bridge inspection by Metni

and Hamel [2007]. They believed that a UAV provided the best solution to the challenging access problems of the bridge environment. However, they also recognized that, in order to provide enough flight stability for accurate imaging of damage, that the UAV would have to rely on visual servoing to maintain its position while hovering. This led to the development and testing of a control law that used the 3D homography tensor as feedback. The system was tested on a local bridge and found to be reasonably effective as a feedback mechanism. The prototype is tele-operated, with hovering stability provided by the servoing law(Figure 3.2b).

Mazumdar and Asada [2009] proposed a fleet of smaller robots that would "live" on a bridge, inspecting it constantly. Such a system would provide resolution in time closer to what is provided by a static sensor network. In addressed the mobility challenges, they developed a robot connected via permanent magnetic connection(Figure 3.2c). The robot is able to slide across the bottom of a steel plate, but is fairly limited with regards to the range of motion. No autonomy or damage detection were considered.

One of the most difficult challenges in structural inspection is the assessment of reinforcing steel embedded in concrete. This is particularly true for bridges, where access problems and environmental concerns make corrosion of reinforcement a serious safety issue. Under guidance from the Federal Highway Administration (FHWA), Ghorbanpoor et al. [2000] developed a crawling robot that uses magnetic flux leakage (MFL) to sense the location of damaged or deteriorated reinforcement in a concrete beam. The basic prototype robot clamps laterally to a beam while rolling along the length of the span (Figure 3.2f). The emphasis of this work was on the use of MFL for damage detection, and mobility and comprehensiveness were not design criteria.

Researchers at Oklahoma State University developed a bridge deck inspection robot that looks for cracking in the bridge deck (driving surface)[Lim et al., 2011]. Their robot uses a Pioneer 3-DX robotic wheeled base for mobility, as bridge decks do not pose an access challenge(Figure 3.3c). Damage detection is done primarily via

CCD camera. Much of their work emphasized optimal path planning algorithms for the robot based on optimizing the camera field of view to provide efficient paths for comprehensive inspection.

DeVault [2000] developed a robot to assist underwater bridge pier inspection, which has long posed access problems for human inspectors. Their system is semi-autonomous, with onboard microcontrollers manipulating the system. The primary damage sensing unit is an underwater video camera, although considerations for sonar were made (Figure 3.3b). The resulting robot ia capable of making observations in high turbidity conditions, but field-test results were not published.

Yu et al. [2007] developed an autonomous tunnel inspection robot capable of automatically extracting crack information from tunnel surface images. The robot uses a simple wheeled base for mobility, an appropriate choice for tunnel inspections, which do not pose complex access challenges. A CCD camera is used in tandem with a pan/tilt head to provide comprehensive imaging of the structural surface (Figure 3.2d). Autonomous planning was not implemented.

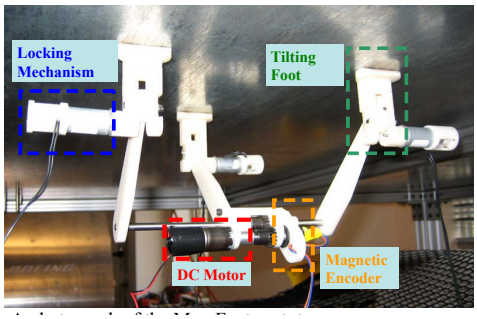
A slightly different robotic bridge inspection robot was developed by Mascarenas et al. [2009]. Instead of using onboard damage sensing, an RC helicopter (Spectra G) was used to wirelessly power and interrogate embedded sensors on a bridge. This approach was taken to solve a longstanding problem with static sensor networks monitoring of structures, that of long-term sensor power supply and management. A Spectra G helicopter was fitted with the communications and powering device, which broadcast power via an electromagnetic wave. No autonomy or onboard perception was used, and one finding of the study was that human piloting of the helicopter was unstable and challenging.

3.3.2 Storage Tank Inspection

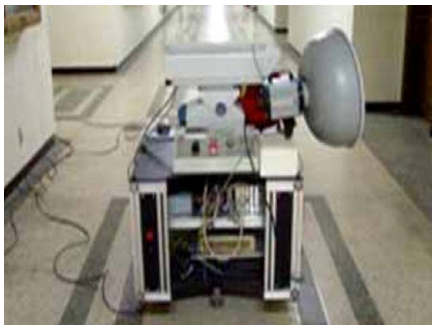
Storage tank inspections are designed to check the integrity of the tank walls from the interior of the tank. In steel storage tanks, structural distress normally takes the form

(a) Serpentine robot by Choset et al

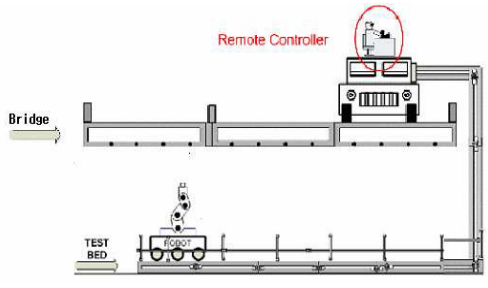
(b) UAV by Metni et al



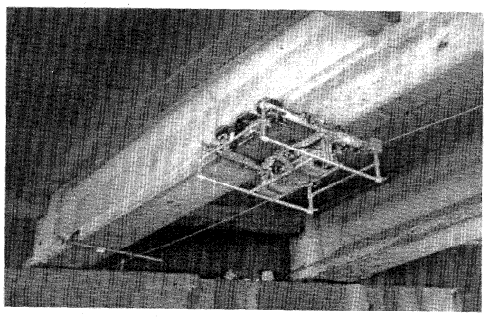
(c) Mag-Foot by Mazdumar et al



(d) Tunnel inspector by Yu et al



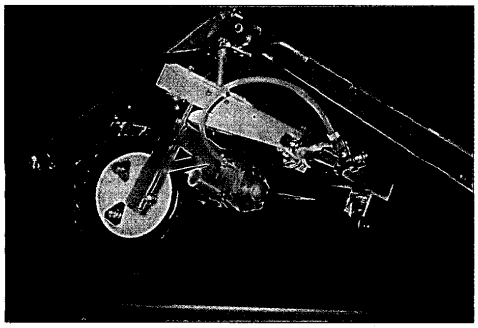
(e) Snooper Based Robot (Lee et al)



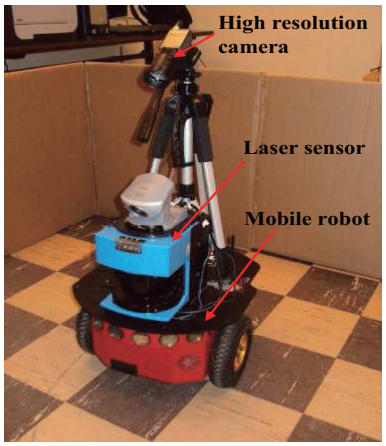
(f) Robot by Ghorbanpoor et al

Figure 3.2: Bridge and tunnel inspection robots

(a) Bridge crawler by Huston et al.



(b) Underwater bridge robot (DeVault et al)



(c) Bridge deck robot (Lim et al)

Figure 3.3: Bridge and tunnel inspection robots, cont'd

of corrosion and cracking of the steel vessel, as well as damage to seams in the welded sections of the tanks. The value of a robot in this scenario is that the tank does not necessarily have to be drained to enable inspection. Compared to bridge inspections, storage tanks are relatively simple shapes. This has enabled more autonomy and effort on comprehensive path planning in development efforts. The primary challenge for robotic tank inspection is detecting damage in murky or dark conditions which can preclude the use of common imaging systems.

The Neptune robot, developed by Schempf et al. [1995], is a mobile wall crawling robot designed to adhere to tank walls using magnetic continuous treads (Figure 3.4a). The robot is, in fact, part of a system which also includes a deployment system and operator console. Autonomy in this setup is minimal, relying on teleoperation for comprehensiveness. Damage detection is performed using a standard CCD camera and ultrasound. The ultrasound sensors are essential for inspections in non-transparent media.

Engineers at the Osaka Gas Company developed an inspection robot to improve the accuracy of their storage tank weld inspection methods [Sogi et al., 2000]. The robot also uses magnetic adhesion and is designed to automatically follow weld seams (Figure 3.4b). Sensing is performed via a combination of ultrasound and high-precision rangefinders, which measure the size and thickness of welds. Their system reduces both the inspection time and the number of inspectors required.

The robot developed by Kalra et al. [2006] was designed to inspect storage tanks from the outside. Similar to the Neptune system, it uses magnetic treads to adhere to the tanks (Figure 3.4c). Much of the development was focused on path planning algorithms to autonomously inspect the tank while avoiding obstacles such as weld joints and minimizing the time required for complete inspection. No sensing, other than a camera used for system verification, is included in the final prototype.

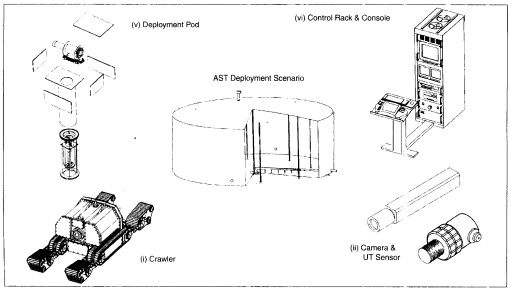
Research by Menegaldo et al. [2009] resulted in a robotic crawler for the inspection of ship hulls and steel offshore marine structures, which have similar inspection

requirements to storage tanks. This robot also uses the magnetic tread concept for inspection access. Ultrasound is used for sensing wall thicknesses. While the crawler is not autonomous, it does perform relative localization via Extended Kalman Filtering of the inertial sensors and motor encoder.

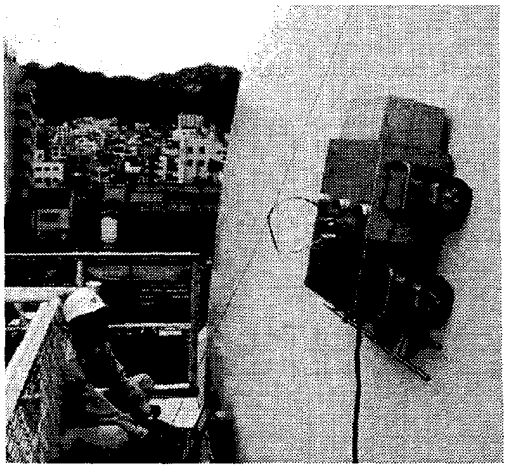
Fury, a robot designed by Marsh et al. [2004] for the Army Corps of Engineers, was designed to test underground storage tanks for leakage. Like other storage inspection robots, Fury uses magnetic wheels to adhere to tank surfaces (Figure 3.4d). It was designed to be small enough to enter underground tanks through access pipes, and uses an articulated chassis to change planes of motion. There is no autonomy, as Fury is tele operated. The onboard inspection sensors use ultrasound to measure wall thickness.

Researchers at ETH Zurich Fischer et al. [2007] adapted the magnetic wall climbing mechanism. Noting that many of the smaller inspection robots used in past research made obstacle avoidance difficult, researchers developed a "mother-child" approach. The larger, "mother", robot is designed to avoid major access obstacles and carry a "child" robot (Figures 3.5a & 3.5b)). The child robot is incapable of obstacle avoidance, but it carries the sensor package and can navigate over thin-walled tanks which likely to be damaged by a larger robot. Both robots use magnetic wheels for mobility. The system is capable of limited autonomy, making decisions about child deployment and obstacle avoidance, but inspection coverage planning was not considered. No inspection sensors were carried by the prototypes.

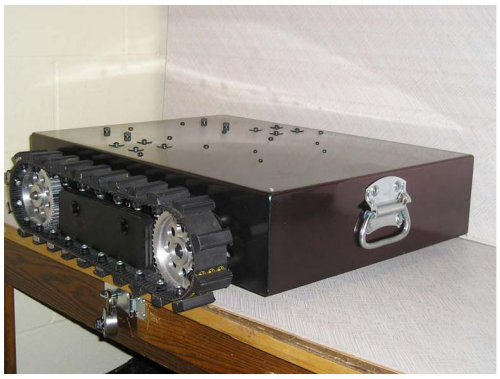
One inspection tank robot using a different locomotive strategy is SURFY [Rosa et al., 2002]. SURFY replaces magnetic wheels with vacuum suction, allowing it flexibility for use on other surfaces and potentially improved mobility (Figure 3.5c). Much of the research effort was devoted to control algorithms for the complex locomotive kinematics. The robot is completely tele-operated. Ultrasonic sensors were used during field testing.



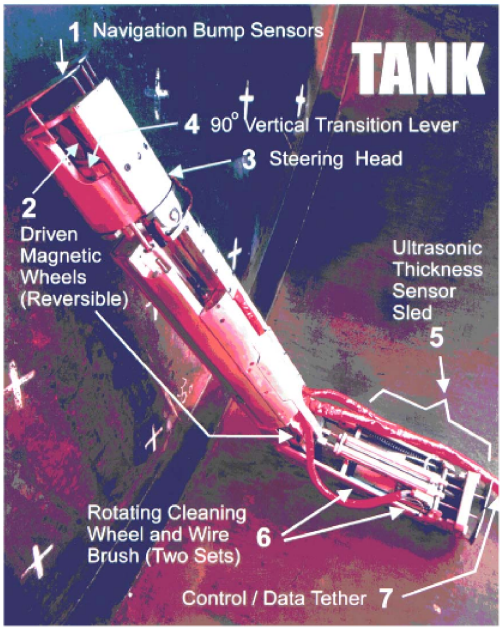
(a) Neptune tank inspection system (Schempf et a)l



(b) Tank inspection robot (Sogi et al)

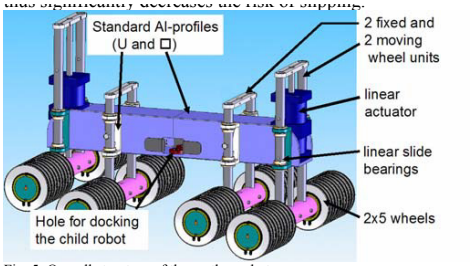


(c) Tank inspection robot by Kalra et al al)

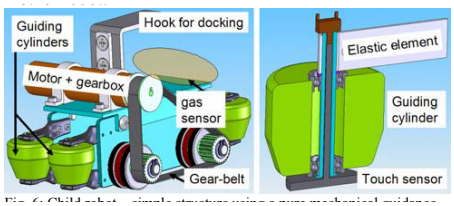


(d) Fury tank inspection robot (Marsh et

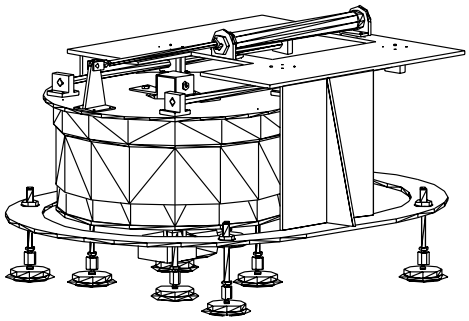
Figure 3.4: Storage tank inspection robots



(a) "Mother" robot design of Fischer et al



(b) "Child" robot design of Fischer et al



(c) Surfy robot by Rosa et al

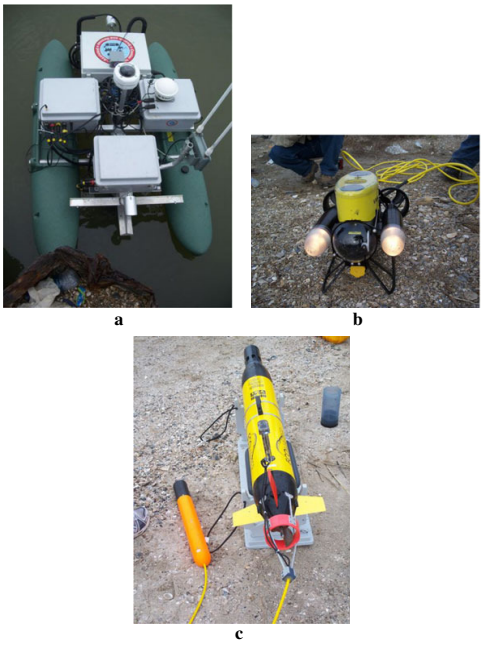
Figure 3.5: Storage tank inspection robots, cont'd

3.3.3 Post-Disaster Inspection Robots

Post-disaster robots have been developed for scenarios where it may not be safe to send human inspectors after events such as earthquakes and hurricanes. It is often difficult to predict how much damage infrastructure has experienced during such a disaster, and keeping structures closed afterwards can have significant impacts to emergency responders and the general public. Robots in these scenarios must often navigate challenging, unstable terrain while providing sensory information for triage structural assessment.

Murphy et al. [2011b] field tested a series of surface and underwater remotely operated vehicles (ROV) for use in post-disaster bridge pier assessment after Hurricane Ike and Hurricane Wilma (Figure 3.6a). They used multiple commercial UMVs, including both subsurface and surface marine vehicles. Subsurface scour inspection was performed using a CCD equipped Video-Ray and an EcoMapper with sonar sensors. Surface investigations utilized a Sea-RAI with both acoustic and CCD sensors. In studies performed after Hurricane Ike, they found that control disturbances due to currents were particularly difficult to handle, sending the ROVs off course. This led to difficulties with 3D registration and reconstruction of image data as well as hazards to the robots themselves. Furthermore, human-ROV interaction was often disorienting due to the limited visual field of the on-board cameras. To address these localization problems, a GPS equipped UAV (Like90 T-Rex helicopter) was incorporated for field studies after Hurricane Wilma. It was used in tandem with the Sea-RAI surface vehicle to aid in vehicle localization by means of a CCD camera observing the UMV [Murphy et al., 2008].

An ongoing series of rubble crawling robots, the Souryu series, is being developed by Arai et al. [2008]. Using durable continuous treads and a segmented chassis, these robots are capable of traversing through rubble or other post-disaster environments (Figure 3.6b). The most recent version, the Souryu-V, uses four cameras. Two are



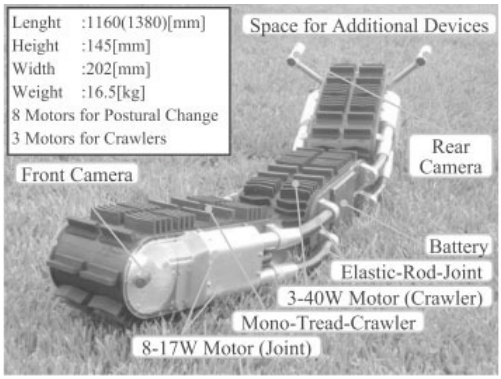
(a) Inspection robots used by Murphy et al

used for search and assessment operations and two are used for teleoperation and control of the robot.

Using a ground-based robot, Torok [2012] developed an assessment system which performs 3D reconstructions of structural elements using Structure from Motion methods [Hartley and Zisserman, 2000] and uses the resulting 3D model to detect anomalies in the structure such as the severe concrete cracking that results from an earthquake. The robot uses a continuous tread system for mobility and a CCD camera for sensing (Figure 3.6c). Robot autonomy was not studied.

3.3.4 Other Applications

The cables of suspension and cable-stayed bridges must be inspected regularly, and are often of key concern for inspection and maintenance teams. Recognizing the



(b) Souryu-V robot (Arai et al)

(c) Damage detection robot by Torok et al

Figure 3.6: Post-disaster inspection robots

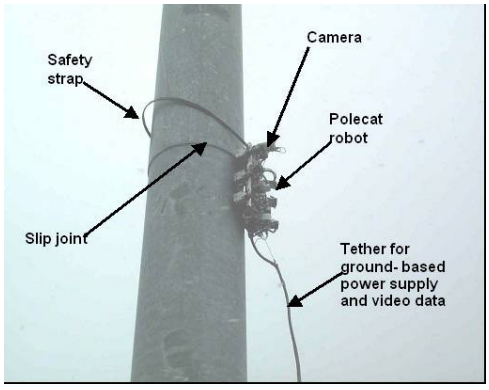
inherent difficulties and disruptions of these inspection, Xu et al. [2011] developed a cable crawling robot to inspect structural cables. The robot chassis is designed to fit a variety of cable diameters and avoid minor obstacles along the cable itself (Figure 3.7a). The robot also includes several safety mechanisms in case of system failure at high elevations, which could jeopardize the robot, the structure, and the traveling public. Two video cameras provide sensing. Autonomy was not implemented.

High-mast highway light poles require disruptive inspections, and the sheer number of required inspections for an owner represent a substantial cost. Polecat, developed by Sheth [2005] was developed as a response to this problem. A crawling robot, Polecat utilizes a series of magnetic wheels to scale light poles, and is capable of avoiding small obstacles due to the articulated suspension (Figure 3.7b). It includes several failsafe features. A video camera provides sensing. The robot is tele-operated.

3.3.5 General Inspection Robots

Many prototyped inspection robots were not created for a specific application, but were designed to test and solve specific challenges for inspection. One of the biggest

(a) Cable inspector (Xu et al)



(b) Polecat during inspection (Sheth et al)

Figure 3.7: Robots for miscellaneous inspection purposes

impediments to using robots for structural inspection is that the geometry that robots must traverse can be arbitrarily complex in many cases. The sheer variety of bridge and building types dictates using robots capable of sophisticated 3D mobility. While many robots have been developed to climb walls, far fewer have been designed to navigate the challenging geometry of civil structures. The robots presented here were developed for the express purpose of maneuvering on civil structures.

The climbing robot ROBIN, developed at Vanderbilt University [Pack et al., 1997], was designed to not only climb structures, but to also carry inspection sensors. ROBIN is a 4 DOF robot with enough articulation to transition from from horizontal to vertical surfaces, as well as turn (Figure 3.8a). Adherence to a structure is gained via vacuum suction, which permits adhesion to arbitrarily oriented flat surfaces. As the study was primarily focused on mobility, sensing and autonomy were not considered.

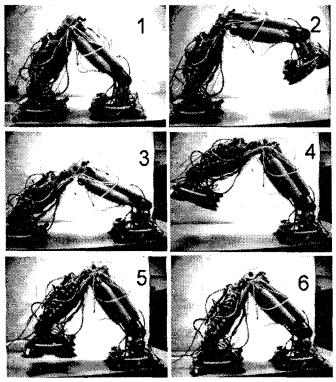
In the development of several robots by Balaguer et al. [2005], the demand for

lightweight systems was paramount. As climbing robots add complexity of motion and additional DOFs, they inevitably grow in weight. The trade-off eventually makes highly articulated robots impractical for field climbing scenarios. This observation led to two series of robots. The ROMA robot is capable of moving in multiple planes as ROBIN does. The first generation ROMA used a somewhat limited grasping mechanism for adherence. The most current version, ROMA II, uses vacuum suction for adherence. ROMA II is a 4 DOF robot, but is capable of sophisticated locomotion (Figure 3.8b). The other robot, MATS, is a 5 DOF robot with high mobility (Figure 3.8c). No perception mechanisms are present on any of the robots.

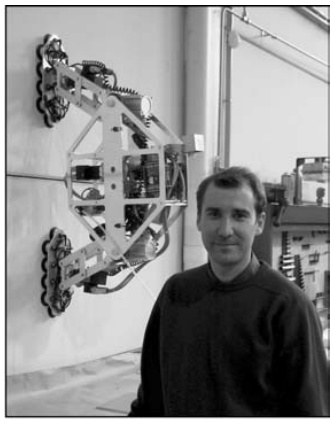
Robug II, developed by Luk et al. [2005] was also designed to navigate in multiple planes. Robug uses four articulated arms with a bifurcated chassis which allows for better mobility in arbitrary environments. Each leg has 3 DOF and uses vacuum suction for adherence (Figure 3.8d). Inspection sensing is performed using ultrasonic sensors on the chassis.

The Alicia series, culminating with the Alicia^3 [Longo and Muscato, 2006], was designed around the criteria of light weight, vertical climbing ability, and obstacle avoidance. The Alicia series began as crawling robots, but the Alicia^3 has an articulated chassis which enables improved obstacle avoidance (Figure 3.8e). Adherence is maintained by vacuum suction, and was successfully tested on concrete surfaces. Infrared proximity sensors are used for obstacle avoidance, but no inspection focused sensors were included.

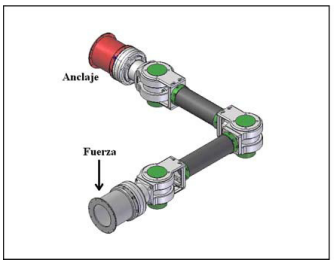
Localization difficulties for UAVs were addressed in Caballero et al. [2005]. GPS is commonly used to localize UAVs in unstructured environments, however GPS accuracy varies with the number of available satellites and can fail entirely. A computer vision based backup system was created which derived relative camera positions using sparse image features and subsequent homography estimations. The method was tested with the COMETS UAV, localizing the UAV with respect to a building (Figure 3.8f).



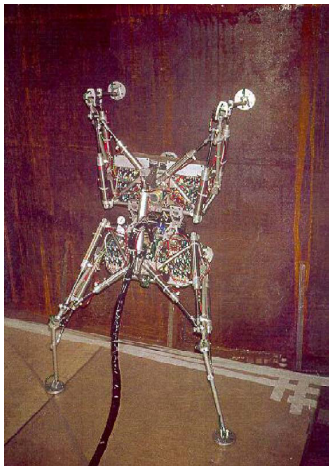
(a) ROBIN climbing robot (Pack et al)



(b) ROMA II climbing robot (Balaguer et al)



(c) MATS climbing robot (Balaguer et al)



(d) ROBUG II climbing robot (Luk et al)

(e) Alicia[3] climbing robot (Luono et al)

Fig. 5 COMETS helicopter.

(f) COMETS rotary UAV used by Caballero et al

Figure 3.8: General inspection robots

3.4 Summary of Robotic Technologies for Infrastructure Inspection

Author(s)	**Robot Name(s)**	**Commercial**	**Sensors**[1]	**Locomotion**[2]	**Application**
Arai	Souryu IV/V	N	RGB	CR	Post-disaster
Balaguer	ROMA, MATS	N	None	CL	General
Cabellero	COMETS		RGB, IR	UAV	General
Choset		N	None	GWA	Bridges
DeVault		N	RGB, O	UMV	Bridges (marine)
Fehlman		N	IR	GB	General
Fischer			None	CR	Storage tanks
Ghorbanpoor		N	MFL	CR	Concrete structures
Huston		N	OS	CR	Bridges
Kalra		N	US	CR	Storage tanks
Lee		N	RGB	GWA	Bridges
Lim	Pioneer 3-DX	N	RGB	GWA	Pavement
Luk	Robug II	N	US	CL	General
Luono	Alicia [3]	N	None	CL	General
Marsh	Fury				Storage tanks
Mazdumar	Mag-Foot	N	None	CL	Bridges
Menegaldo		N	US	CR	Storage tanks
Metni		N	RGB	UAV	Bridges
Murphy	Sea-RAI	Y	RGB	UMV	Post-disaster (Hurr. Ike)
Murphy					
Oh		N	None	GWA	Bridges
Pack	ROBIN	N	None	CL	General
Ridao	Ictineu	N	RGB	UMV	Marine (dams)
Rosa	SURFY		US	CR	Storage tanks
Schempf	Neptune	N	RGB, UW	CR	Storage Tanks
Sheth	Polecat	N	RGB	CR	Lightpoles
Sogi			US, OS	CR	Storage Tanks
Steimle				UMV	Post-disaster (Hurr. Wilma)
Torok		N	RGB	GB	Post-disaster
Tseng	Video-Ray	Y	RGB	GB	Pavement

Tung	N	RGB	GWA	Bridges
Xu	N	RGB	CR	Cables
Yu	N	RGB	GB	Tunnels

[1]RGB, camera; IR, infrared; MFL, magnetic flux leakage; US, ultrasound; SO, sonar; O, other
[2]GB, ground-based; GWA, ground-based with articulator; CR, crawling; CL, climbing; UAV, unmanned aerial; UMV, unmanned marine

3.5 Unsolved Challenges

The field of inspection robotics is still nascent, with many outstanding challenges still unmet. Prototype robot development generally has not been focused on inspection robotics per se, but on exploring particular aspects of the field. Robots that have been implemented have been, more often than not, specialized devices which limit flexibility to gain stability and robustness. From the perspective of an inspector, inspection robots are still limited in their ability to access structures, comprehensively inspect those structures, and to perceive and represent damage consistently and accurately. Fail-safe measures to guarantee the safety of the general public during an inspection also remain outstanding.

The need for flexible locomotive systems imply that UAVs and UMVs, in their respective environments, are a way forward for inspection robotics. Yet the work done by CRASAR [Murphy et al., 2008] highlights why the solution is far from simple. These robots are susceptible to environmental control disturbances, especially if they are operating in GPS denied scenarios. An ongoing challenge for these systems will be the development of localization and control methods which can better leverage sensor data. The authors see computer vision and sensor fusion methods as potential solution techniques, such as those developed by Scherer et al. [2012]. Another important challenge is developing locomotive mechanisms which can operate reliably without risk to the general public (e.g. UAVs inspecting a bridge in traffic).

There is substantial need for more sophisticated autonomy and efficient coverage planning algorithms, a limitation again tied to the challenging and varied inspection environment. On a more fundamental level, questions remain regarding whether ROVs or fully autonomous robots are preferable for field inspection tasks. ROVs allow for operator correction to environmental disturbances, but they tend to mitigate one major benefit of autonomous robots, accurate map building. Furthermore, they provide no guarantee of a comprehensive inspection. Still, for the near future, teleop-

eration is likely the preferred choice in unstructured and unknown environments such as those that exist in post-disaster scenarios.

Finally, the authors have observed three challenges to accurate damage perception. First, while most sensors can detect damage such as cracking or corrosion, it is much more difficult to detect mechanical deformations in a structure as these are often subtle and complex 3D phenomena. The 2D visual representations cameras produce exacerbate this limitation. Second, there is a need to determine acceptable accuracy metrics for robotic inspection measurements and methods for using those measurements to make quantitative structural assessments. Lastly, there are outstanding questions about how to represent robotic inspection information to structure owners. Simply supplying an engineer or owner with video footage and sensor measurements leads to confusion and tedious back-end work such as reviewing hours of close-up footage of structural components, and consequentially to a perceived lack of value for inspection robotics systems. Work must be done to ensure that inspection work is organized in a manner that makes the job of human data review and analysis simpler and more intuitive.

3.6 Emerging Technologies

Robots, and in particular focused robot designs such as those often used for inspection purposes, have shrunk in cost dramatically while expanding in capabilities thanks to the ubiquity of embedded processors and sensing systems. This is largely due to the expansion of the smartphone industry, but also because of the efforts and expansion of the online robotics and electronics enthusiast communities. This affordability and accessibility is already expanding the base of research, as civil engineers are now capable of incorporating robotic technologies in their research.

There is a great deal of current research on the study of RGB-D systems, such as those supplied by PrimeSense [PrimeSense, 2011]. These systems produce fully 3D color point clouds and therefore are potentially capable of providing both intuitive

representations of structural damage and of detecting mechanical deformations in the structure. However, they use infrared depth sensing, which has inherent limitations and noise susceptibility. For consumer entertainment purposes, for which they were originally designed, this is not critical but it can produce significant problems for the construction of accurate 3D environment maps [Huang et al., 2011].

Another area of recent research is in the field of RFID (radio frequency identification) interfacing with embedded sensors [Lynch and Loh, 2006]. Incorporating sensors permanently into civil structures has been the focus of several decades of research in the civil engineering research community. However, these sensors must either be hardwired for power and communications, which limits their durability, or they are limited by the battery demands of wirelessly transmitting sensors. RFID potentially solves this problem by using transmitters which are energized by the receiver itself. If adapted to structural sensing systems, a robotic inspector could potentially energize RFID tags and access structural performance information of a sort that is currently impractical with embedded systems.

3.7 Conclusion

Developing robotic systems for infrastructure inspection requires a broad set of considerations in addition to the common robotics challenges of power, autonomy, and communications. The field environment that inspection robots must navigate is complex and challenging to localize within. It also poses potential safety hazards to the general public in the event of robot system failure. And the degree of sensing accuracy and repeatability necessary for structural inspection can potentially dictate the design of motion control algorithms and payload demands.

Still, the potential advantages of robotic inspection systems outweigh the development challenges. Current inspection practices are expensive and often dangerous for both inspectors and the general public, and robotic systems can mitigate both the costs and the risks. Lastly, if properly employed, inspection robots can also improve

inspection practices by providing more quantitative inspection information than what is typically gained from standard visual inspection methods.

Chapter 4

3D INSPECTION RECONSTRUCTION VIA ROBOTIC IMAGING

4.1 Introduction

This paper presents methods capable of providing a virtualized bridge inspection experience based on robotically captured images embedded in a spatially and temporally contextualized, photorealistic, 3D visualization environment. In particular, different methods of 3D scene reconstruction are studied and compared in regards to suitability for infrastructure inspection applications, considering fundamental issues such as accuracy, artifact behavior, and resolution requirements.

Computing advances have led to an emerging field of inspection robotics [DeVault, 2000, Guo et al., 2009b, Lim et al., 2011], and will likely lead to the increased usage of robots as inspection tools in the near future. There are many decisions to be made regarding the development of a robotic inspection system, as robots can be human piloted or designed to autonomously perform inspections, and a wide range of locomotive mechanisms are available for navigating structures efficiently. In terms of damage sensing, robotic systems often provide massive sets of unstructured and decontextualized digital images as the primary means of damage. In order to maximize the utility of inspection robots, these images must be processed and organized to facilitate their usage.

Robotic inspection systems in current bridge inspection practices most commonly utilize cameras to sense and observe structures during routine inspections. This type of sensing is an intuitive one for inspectors who have for decades performed structural inspections primarily through visual observation [Hartle and Administration, 2004]. However, robotically captured images and videos typically have a narrow field of view and are often high-resolution close-ups of the structure. These images are fundamentally decontextualized, as they do not consider information in the context of the structure the way a human observer in the field intuitively does. This decontextualization has been shown to lead to disorientation for robot operators [Murphy et al., 2008] and devalues robotic inspection as a general practice since engineers must

often parse hours of close-up video footage.

The need for contextualized back-end visualization development to keep pace with front-end robotic capture technologies is apparent. By converting the 2-dimensional images into photorealistic 3-dimensional geometric models of the structure, a robotic inspection regimen can effectively "bring the inspection into the office" and enable inspectors and engineers to engage with robotic imaging in a more intuitive manner. Beyond the direct value of making inspection image data more comprehensible, a properly calibrated 3D visualization environment can provide accurate inspection measurements comparable to those performed in the field, and can enable time-history analyses by accurately overlaying image data from repeated inspections throughout a structure's life-cycle.

In order to create photorealistic virtual bridge inspection environments, the images captured during an inspection must be converted into 3D computer models, a process known as scene reconstruction. There are two primary paradigms for constructing 3D models from the 2D monocular images commonly captured by robots. The first paradigm consists of so-called naïve methods that require no knowledge of the scene in the image or the location of the camera with respect to objects in the image, and typically generate 3D point clouds of the reconstruction. The most commonly used naïve approach today uses a process known as Structure from Motion (SfM) to generate sparse 3D point clouds, followed by a series of associated methods to develop dense 3D point clouds and eventually photorealistic 3D models of a scene [Snavely et al., 2008]. The second scene reconstruction paradigm is referred to herein as Image Mosaicing (IM). With IM, an underlying 3D geometric model is constructed independently of the captured images. Using this known model, the captured 2D images are calibrated and stitched together. They are then texture mapped onto the underlying 3D geometric model.

4.1.1 Prior Research

Scene Reconstruction for Infrastructure Maintenance

Jahanshahi et al. [2011] considered image stitching and scene reconstruction methods for civil inspection purposes. They utilized image stitching methods to detect changes between two stitched scenes, allowing for time-history analyses from 2D inspection images. They did not use their technique to provide 3D reconstructions and visualizations.

3D scene reconstruction via SfM and dense point cloud reconstruction has been applied in a wide variety of computer science applications [Snavely et al., 2008, Pollefeys et al., 2004, Akbarzadeh et al., 2006, Pizarro et al., 2004]. Many of these applications have focused on the reconstruction of single objects or on scene reconstructions that do not consider scale accurate representations of fine details or the scalability of the reconstructions to large models.

Researchers have tested the capabilities of photogrammetric reconstruction methods, including SfM, to provide photorealistic reconstructions for as-built construction verification [Golparvar-Fard et al., 2011, Bhatla et al., 2012]. While these studies tested the reconstruction accuracy of SfM methods, the studies were designed for construction site inspections, as opposed to routine bridge inspections, and so the reconstructions were not evaluated for inspection measurement accuracy or artifact appearance but for their capability as a virtual "as-built" model. Furthermore, the tested methods were considered for use by a human inspector without the restrictions of robotic inspection in mind.

Torok [2012] applied the SfM method to a severely cracked concrete block. Images were captured via a ground-based robot, which were then used to reconstruct the object. The 3D reconstruction was then applied to provide crack detection capabilities based on the reconstructed depth information. This study did not consider scene reconstruction accuracy or the performance of the SfM method on structural materials

other than concrete.

In the work of Dai et al. [2011], a methodology for assessing the suitability of reconstruction algorithms was developed. The proposed method involved: (i) identifying relevant target objects; (ii) reconstructing the target objects; (iii) taking measurements of the objects using both the reconstructed model and the physical test object; (iv) comparing the two sets of objects. The system was not tested in the context of reconstruction accuracy.

Typically, IM is used to reconstruct objects with arbitrary complexity and where photorealism, not accuracy, is the primary metric [Lensch et al., 2000, Jankó et al., 2005]. In the context of civil inspection, Früh et al. [2004] used aerial images captured at an oblique angle, mapping them to a known 3D model of a city. Images were registered to the model by matching line segments in the 3D model to line segments in the images. The method was not designed for bridge inspection, as the level of image detail available from satellite images is inadequate for inspection purposes.

The most detailed assessment of an IM system was presented in Jauregui et al. [2006]. This research analyzed both the accuracy of a commercial IM method, as well as an associated method for generating 3D models of a scene, in the context of routine inspections. The technique required previously defined targets on a structure to aid in the reconstruction. The IM method was found to be suitably accurate for reconstruction, but the accuracy assessment did not consider flaw detection or the nature of reconstruction artifacts.

Robotic Bridge Inspection

Researcher have for several decades explored using robots to perform routine bridge inspections. Such robots have take many forms, ranging from complex articulators attached to vehicles [Choset, 2000, Tung et al., 2002, Lee et al., 2008] to smaller crawling robots [Huston et al., 2003]. Recently, researchers have considered using unmanned aerial vehicles (UAV) [Metni and Hamel, 2007, Caballero et al., 2005]. The

lightweight nature of such robots restricts the allowable sensor payload and limits the application of heavier tomographic or laser-based sensing systems. Instead, UAVs tend to rely on digital camera imaging as the primary means of perception due to the low weight and minimal power consumption of such sensors.

The challenges of processing and representing robotically captured images were explored during a series of post-disaster robotic inspections [Murphy et al., 2011b,a, 2008]. In this work, unmanned aerial and marine robots were used to inspect structures in post-disaster scenarios. One of the key findings of the studies was that the robotically captured imaging is often disorienting and confusing to both the robot operator and the end-user.

4.1.2 Contributions of this work

Photorealistic 3D scene reconstruction methods have been applied in a variety of circumstances. However, these methods have rarely been considered and tested in the context of routine bridge inspection, and never in the context of robotic bridge inspection. To date there has been no study on the fine scale measurement accuracy, reconstruction viability, or resolution demands of these methods with respect to robotic bridge inspection needs.

Given the state of the art of both robotic structural inspection and scene reconstruction techniques, the primary contribution of this research is a comparative study of the two most common methods of monocular scene reconstruction, a SfM-based method and an IM method, and the development of guidelines for their implementation within a robotic bridge inspection framework. The study is designed to lend insight into the suitability of both methods for robotic inspection applications utilizing monocular imaging as the primary inspection sensor.

The SfM-based and IM methods were prototyped and tested in several imaging scenarios to compare their relative merits based on fundamental performance metrics described later in the paper. The methods were first tested on a series of small-scale

structural objects to simplify the reconstruction process and observe the fundamental behaviors of the two methods. This was then followed by a full-scale reconstruction of a bridge using robotic imaging, so as to provide insight into the implementation of a robot in a simulated inspection scenario.

This paper is structured as follows. First a brief overview of the two tested reconstruction techniques, a dense SfM-based method (referred to herein as DSfM), and Image Mosaicing (IM), is presented. The details and results of the performed comparative small-scale tests are then presented, followed by details and observations of the full-scale test. The paper concludes with directions for future research.

4.2 Reconstruction Methodologies

While there are many methods of reconstructing scenes using various sensing technologies, such as LADAR or stereo depth correlations [Golparvar-Fard et al., 2011, Brilakis et al., 2011, Valença et al., 2012], such methods are typically impractical in the context of robotic bridge inspection [Chambers et al., 2011]. Most robotic inspection platforms must maneuver through complex 3D environments and are severely limited by on-board weight, power, and computing restrictions. Thus, single camera sensing is utilized by the majority of inspection robots and this consideration guided the choice of reconstruction methods in this study. DSfM and IM were chosen as they are the most commonly implemented methods for reconstructing photorealistic 3D scenes from monocular 2D images.

4.2.1 Photogrammetric Reconstruction via DSfM

There are many variants of the DSfM approach available, however the general concept is typically the same [Hartley and Zisserman, 2000]. First, salient feature points in unstructured 2D color images are found, $\mathbf{x}_j^i$ for the j^{th} point in the i^{th} image. These points are then matched across multiple images using the descriptions of the features. Using these matches and the intrinsic parameters of the cameras, the spatial

orientations of each camera, represented as a projection matrix $\mathbf{P}^i$, and the 3D points $\mathbf{X}_j$, are derived using geometric epipolar constraints. For each detected feature, the re-projection error of its 3D location onto the relevant 2D image planes is then minimized using a least squares approach, a process known as bundle adjustment. The results of bundle adjustment are *estimated* camera orientations and 3D point locations, $\hat{\mathbf{P}}^i$ and $\hat{\mathbf{X}}_j$ as expressed below:

$$\text{argmin}[\sum_{i,j}(\mathbf{P}^i\mathbf{X}_j - \mathbf{x}_j^i)^2] \rightarrow \hat{\mathbf{P}}^i,\ \hat{\mathbf{X}}_j \tag{4.1}$$

The result of the SfM method is a sparse 3D point cloud, containing only the estimated locations of relevant salient features.

The next step is to create a dense point cloud reconstruction. There are various methods to perform this reconstruction [Furukawa and Ponce, 2010, Pollefeys et al., 2004]. The result is a dense point cloud with oriented surface normals. Finally, in order to provide a realistic, consistent, and accurate visualization this dense point cloud is then employed to create a surface mesh using a Poisson Reconstruction [Kazhdan et al., 2006]. The result is a 3D surface mesh and associated texture map that can be rendered using standard 3D imaging techniques. The complete flowchart for this method can be shown in Figure 4.1.

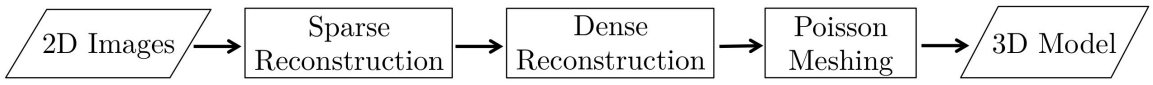


Figure 4.1: DSfM method flowchart

DSfM has two critical limitations with regards to field applications. First, the reliance on salient and robust local features for creation of the point cloud can lead to problems in regions of low texture variations, or where scene lighting or camera

position become too extreme between captured images. The second issue is that, if a cluster of feature points are co-planar, this can lead to degenerate solutions of the SfM equations and unstable depth correspondences for those features [Hartley and Zisserman, 2000]. Something as common as an area of a relatively clean steel plate can give rise to these kinds of issues [Bhatla et al., 2012].

4.2.2 Image Mosaicing Reconstruction

The other major scene reconstruction paradigm is IM. With IM, an underlying 3D geometric model is constructed independently of the captured images. Using this known model and a set of previously calibrated baseline images, the captured 2D images are calibrated and stitched together. They are then texture mapped onto the underlying 3D model (Figure 4.3). A flowchart for the complete IM method is illustrated in Figure 4.2.

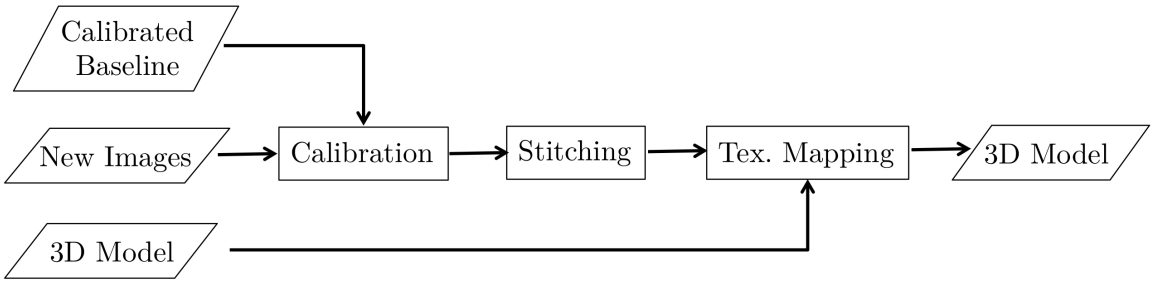
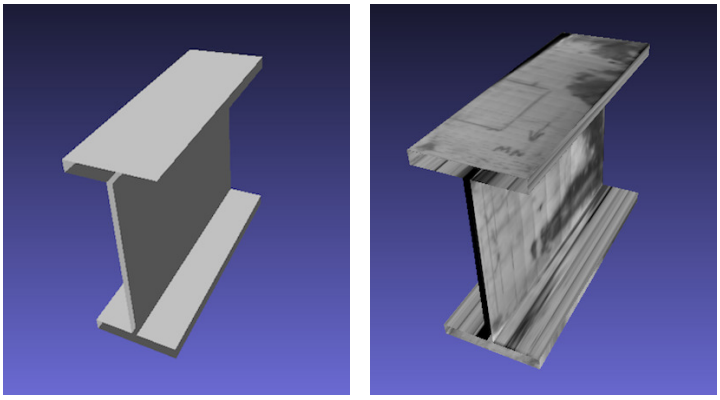


Figure 4.2: IM method flowchart

Image Stitching, and Mapping

After the images are calibrated, they are stitched together as necessary to provide texture maps. UV texture mapping [Wright et al., 2010] is then applied to the resulting stitched images based on the known spatial locations of the images.



(a) Underlying 3D structural model

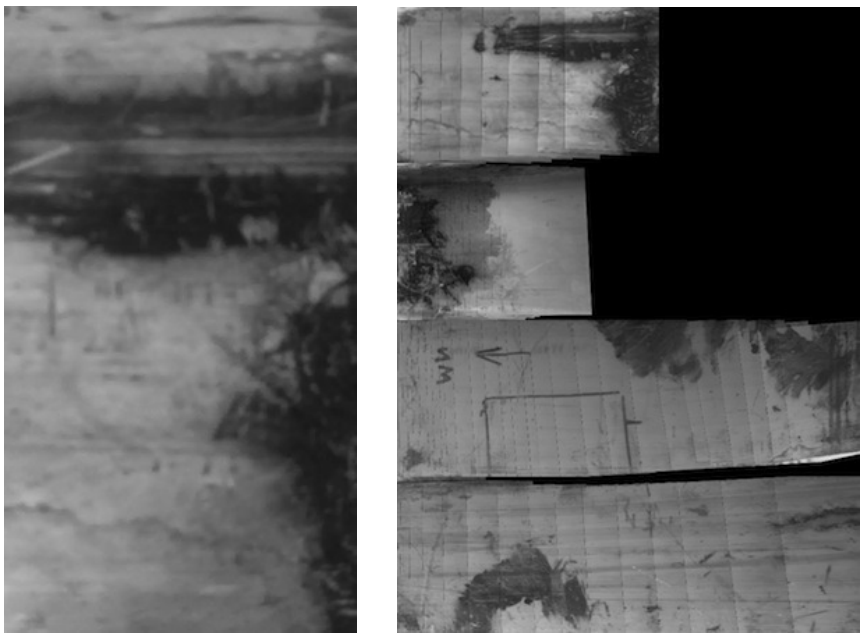
(b) Model with photorealistic texture map.

Figure 4.3: IM texture mapping example

IM methods have several limitations. In order to develop a successful visualization system using this technique, the underlying 3D geometric model must provide an accurate representation of the as-built structure. In many 3D reconstruction problems, the underlying model is not well known, making IM impractical. However, most modern infrastructure projects now include highly detailed 3D computer models of the structures themselves. This makes creating the 3D model relatively simple, and the primary question is how refined to make the model for visualization purposes. However, the 3D model used may have been created for a purpose other than structural inspection, such as architectural renderings, and may not provide a suitable level of model complexity for structural inspection.

4.2.3 Model Calibrations

In either scenario, the model must be calibrated and projected into a global reference frame and to provide perspective corrected images. There are several methods for performing such calibrations, such as methods using checkerboard or known targets



(a) Single captured image (b) Final stitched texture map

Figure 4.4: IM image stitching example

on the structure [Zhang, 2000]. For the work presented here, all models were calibrated manually.

4.3 Preliminary Small Scale Testing

The goals of the preliminary study were to identify any fundamental limitations to the application of either DSfM or IM to bridge inspection, and to explore how material textures and structural complexity affected the different reconstruction methods independent of the complications inherent to full-scale testing. The initial tests were also performed so as to determine key imaging demands such as required reconstruction resolution for the eventual full-scale test.

In order to provide a comparative analysis of the two methodologies presented, a series of small-scale structural objects were comprehensively and systematically photographed. These images were then used to develop 3D reconstructions of the objects, which were imported into the open-source visualization program Meshlab [Cignoni et al., 2008]. MeshLab is a 3D environment that is optimized for rendering performance and provides a suitable comparison platform for working interactively with 3D models. MeshLab also has the functionality to perform scale-accurate 3D measurements and report rendering performance information.

4.3.1 Test Setup

Reconstructed Structural Objects

Three reconstruction targets were chosen: an 11-inch long S8×18 aluminum section, a scale aluminum and acrylic structure, and a tubular steel bridge column with a concrete footing. The S8×18 section was chosen so as to explore the accuracy of the reconstruction methods for single structural components. The section had a variety of pre-existing markings and surface imperfections, which provided useful measurement targets for system accuracy testing. The scale structure was chosen to test each

method's capabilities for more complex 3D structural systems. The steel and concrete bridge column was chosen to test capabilities with regards to variance in structural materials, and to test system accuracy at a larger physical scale. The test subjects are shown in Figure 4.5.

(a) Aluminum Section (b) Scale Structure (c) Column and Footing

Figure 4.5: Test objects for reconstructions

Image Capture

Images were captured in a manner representative of how they might be during a robotic inspection using either of the two proposed approaches. For the DSfM approach, images were captured at roughly 5-10-degree intervals around the test specimens. Images for the IM approach focused on imaging each associated structural face separately. Images were captured using a 12-megapixel camera at an ISO of 400. The equivalent of a 50mm focal length (35mm format, 40° field of view) was used to minimize perspective distortion.

Reconstruction Algorithm Implementation

All DSfM reconstructions were performed using Autodesk [2013]. Preliminary testing indicated that 123D-Catch provided the most consistent results compared to reconstructions performed with Arc3D [Tingdahl and Gool, 2011] or Clustering Views for Multi-View Stereo (CMVS) [Furukawa and Ponce, 2010], however the differences were slight. The IM reconstruction algorithm was developed using a combination of the Python programming language, the computer vision package OpenCV [Bradski, 2000], and the OpenGL graphics API. SURF (Speeded Up Robust Features) [Bay et al., 2006] were used for feature detection and matching purposes.

Metrics

The following criteria were used to assess each reconstruction method's utility for bridge inspections:

1. Reconstruction accuracy (i.e., the ability to represent inspection information accurately and repeatably).

2. Graceful failure (failed reconstructions or image artifacts should not provide misleading information).

3. Resolution demands for accurate reconstruction

The reconstruction accuracy was assessed following the work of Dai et al. [2011]. Field measurements were taken using a measuring tape as would be done during a routine bridge inspection [Hartle and Administration, 2004]. As a baseline, the precision of field measurements was estimated to be on the order of $^1/_8$ inch due to the often ambiguous nature of field observations. These measurements were then compared to measurements from the reconstructions, taken in Meshlab. Because the IM method implicitly is based on known global geometries, the measured external

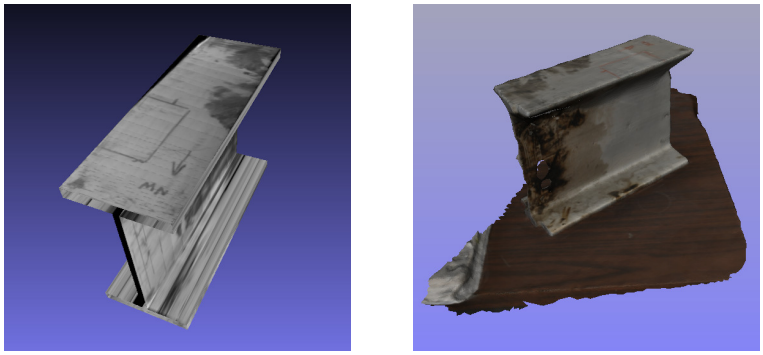
dimensions of the reconstructions were not compared in this analysis. Instead, features such as marking and cracks on each object were measured and compared, referred to as Features A, B & C for each reconstruction. These figures were chosen as representative of the sort of defects or flaws that would require measurement by a field inspection team. Examples of measured features are shown in Figure 4.6.

Figure 4.6: Example features for measurement testing

Regarding image resolution requirements, maximizing a sensor's resolution does not necessarily lead to an optimal reconstruction. Higher resolution images inherently produce more of the salient feature points that both methods rely on, resulting in expanded computational demands for feature matching across multiple images, and in denser meshes in the case of the DSfM method. Both of these expansions reduce the scalability and renderability of reconstructed models. Furthermore, feature points that can only be detected by higher resolution cameras are inherently less robust, particularly when considering the inevitable image degradation that accompanies mechanical vibrations in robotic imaging systems. In this study, the *minimum* resolution required to produce a reconstruction of each object was explored. To determine the resolution demands, the captured images were down-scaled from 12 megapixels until reconstruction failed.

4.3.2 Small Scale Results and Discussion

Reconstructions of the S8×18 aluminum section are shown in Figure 4.7. Reconstructions of the scale structural model are shown in Figure 4.8. For this latter case, DSfM reconstruction performed extremely poorly, despite repeated reconstruction attempts. This failure was due to the low texture of the brushed aluminum components and the highly planar nature of the structure itself, which led to many degenerate DSfM solutions and an unstable solution (Figure 4.8b). Reconstructions of the bridge footing and column are shown in Figure 4.9.



(a) IM reconstruction (b) DSfM reconstruction

Figure 4.7: Reconstructions of aluminum section

Measurement Accuracy

Measurement accuracy comparisons for the aluminum section are shown in Table 4.1. For the aluminum section, both methods provide consistent and scale-accurate results, with discrepancies less than $^1/_4$ of an inch, on the order of the potential accuracy of a field inspection. However, the IM approach failed to capture the full extent of Feature C due to a texture mapping artifact related to inaccuracies in the simplified underlying model of the aluminum section that did not account for a slight coping.

(a) IM reconstruction (b) DSfM reconstruction

Figure 4.8: Reconstructions of scale structure

Measurement accuracy comparisons for the scale structure are shown in Table 4.2. The IM method provided accurate results for all three measurements, with discrepancies less than $^1/_8$ inch. The exception occurred in Feature C, a surface gouge with ambiguous start and end points. The large gaps and inconsistencies in the DSfM reconstruction for this case prevented accurate measurement of any of the features.

Measurement comparisons for the bridge column are shown in Table 4.3. Both methods produced scale accurate measurements, with most measurement discrepancies less than $^1/_8$ of an inch. There is a slight discrepancy in the measurements obtained for Feature B using the IM approach. This feature was a marking on the cylindrical column and the error is due to the approximate modeling of the curved cylinder.

Artifacts and Inaccuracies

In addition to the reconstruction failures illustrated earlier for the DSfM model and the measurement inaccuracies shown for the IM approach, it is important to compare other types of visual artifacts created by each method. Two typical artifacts for the

Feature	Field	IM	DSfM
Feature A (red marker line)	$1^{1}/_{4} \times {}^{1}/_{16}$	$1^{1}/_{4} \times {}^{1}/_{16}$	$1^{3}/_{8} \times {}^{1}/_{16}$
Feature B (red marker line)	$3^{1}/_{8} \times {}^{1}/_{16}$	$3^{1}/_{4} \times {}^{1}/_{16}$	$3^{1}/_{4} \times {}^{1}/_{16}$
Feature C (black smear)	$4^{3}/_{8} \times 1$	$\mathbf{3}^{1}/_{4} \times \mathbf{1}^{a}$	$4^{3}/_{8} \times 1^{1}/_{4}$

[a] Error from inaccurate 3D model

Table 4.1: Measurements on aluminum section (nearest $^{1}/_{16}$ inch)

Feature	Field	IM	DSfM
Feature A (black marking)	$6^{1}/_{4} \times {}^{1}/_{8}$	$6^{1}/_{4} \times {}^{1}/_{8}$	–
Feature B (tape residue)	$1^{3}/_{8} \times 2$	$1^{3}/_{8} \times 2^{1}/_{8}$	–
Feature C (surface gouge)	$\mathbf{15} \times {}^{1}/_{16}{}^{a}$	$\mathbf{14}^{3}/_{8} \times {}^{1}/_{16}{}^{a}$	–

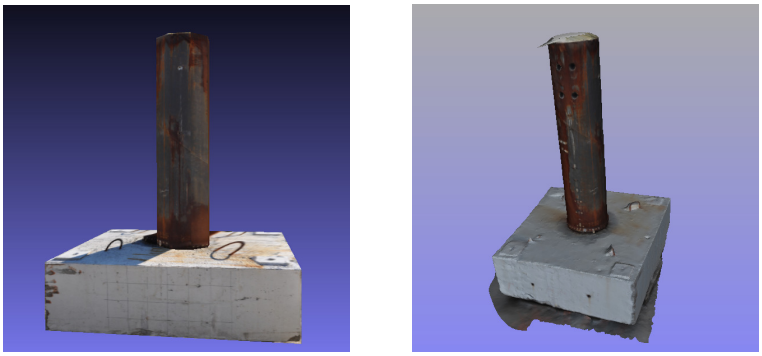
[a] Observational discrepancy

Table 4.2: Measurements on scale structure (nearest $^{1}/_{16}$ inch)

Feature	Field	IM	DSfM
Feature A (black marker line, footing)	$37^{1}/_{2}$ x $^{1}/_{16}$	38 x $^{1}/_{8}$	$37^{3}/_{4}$ x $^{1}/_{8}$
Feature B (tape, column)	$8^{5}/_{8}$ x 1	$\mathbf{7}^{3}/_{4}$ **x** $\mathbf{1}^{a}$	$8^{1}/_{2}$ x 1
Feature C (circular hole, column)	$2^{5}/_{8}$ x $2^{5}/_{8}$	$2^{1}/_{2}$ x $2^{3}/_{8}$	$2^{1}/_{2}$ x $2^{1}/_{2}$

[a] Error from inaccurate 3D model

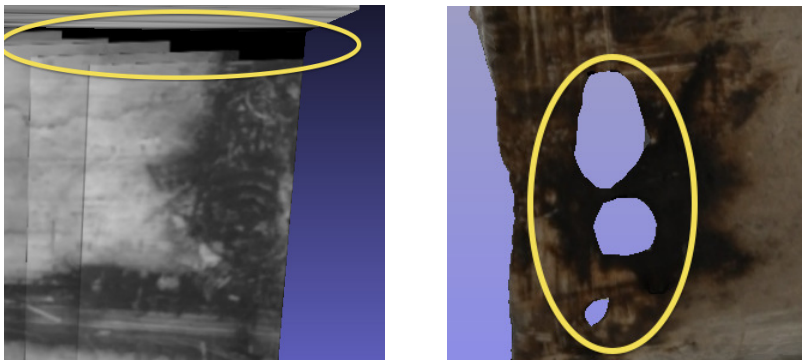
Table 4.3: Measurements on bridge column (nearest $^{1}/_{16}$ inch)



(a) IM reconstruction (b) DSfM reconstruction

Figure 4.9: Reconstructions of bridge column and footing

aluminum section are shown in Figure 4.10a. The IM approach creates blacked out areas where camera motion prevented complete imaging of the section face. The DSfM approach created holes in some surfaces. These holes are due to the inherent difficulties in producing accurate depth correspondences when most nearby salient features all lie on the same plane.

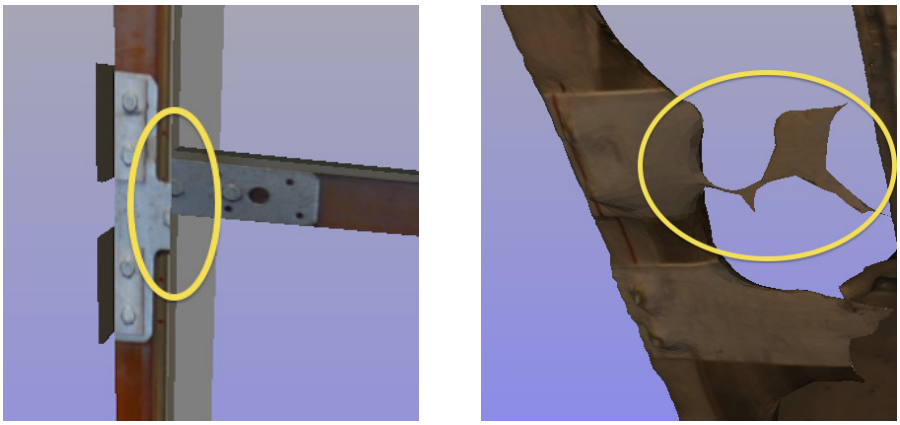


(a) IM artifact (b) DSfM artifact

Figure 4.10: Typical artifacts in aluminum section reconstructions

For the IM reconstruction of the scale structure, some features were accurately

represented and some were not. Inaccuracies again were due to inaccuracies in the underlying geometric model itself. For instance, the height of some columns were incorrectly approximated, which led to misalignment distortions during texture mapping. However, all of the features selected for measurement were represented accurately. The misalignment problem is illustrated in Figure 4.11a, which shows an artifact produced by inaccurate column modeling. The IM reconstruction also introduced artifacts due to another type of modeling error. In the 3D geometric model, all angles between sections were modeled as right angles. However, some of the structure connections were in fact slightly skewed. As mentioned previously, the DSfM reconstruction produced a poor reconstruction due to the lack of texture on the brushed aluminum surfaces and the planar nature of the structure itself. An example of this failure is shown in Figure 4.11b.



(a) IM artifact (b) DSfM artifact

Figure 4.11: Typical artifacts in scale structure reconstructions

For the reconstruction of the bridge column and footing, both methods produced several artifacts. The DSfM model reconstructed a fictitious hole in the bridge column (Figure 4.12b). The DSfM method also inaccurately modeled the footing in regions

of lower surface texture. For the IM reconstruction, poor image stitching due to the curved shape of the column led to artifacts of the type shown in Figure 4.12a. The top face of the column was also not properly imaged and produced a distorted representation. While improved imaging and stitching would mitigate these types of artifacts, they are likely to always be present to some extent. More notable is the inherent failure of the IM method to accurately capture the spalling at one corner of the footing (Figure 4.13). The 3D geometric model was not designed to capture this kind of localized damage, and so the damage is rendered in an unrealistic, distorted fashion. In such cases it would be impractical to include this kind of unpredictable damage geometry in an as-built geometric model, so this kind of distortion of major damage areas would generally be an inherent issue of an IM approach. Because it is neither helped nor hindered by an assumed geometric model, the DSfM reconstruction was much more successful at capturing the true nature of the spalling damage.

(a) IM artifact (b) DSfM artifact

Figure 4.12: Typical artifacts in bridge column reconstructions

Overall, artifacts created by the IM method are typically easy to spot and generally result in distortion of existing features rather than in the introduction of realistic looking features not present in the actual structure. They show up as blocky, blacked

(a) Image of actual damage (b) IM representation (c) DSfM representation

Figure 4.13: Comparison of spalling representations on bridge footing

out, or clearly distorted features (Figure 4.10a). A notable exception is the distorting effect of an inaccurate underlying 3D model. Such inaccuracies may create natural looking but distorted features, as evidenced in the measurement testing. Artifacts created by DSfM reconstructions can be harder to recognize as artifacts. In some cases, they manifest themselves as photorealistic features such as the pseudo-hole in the flat plate in Figure 4.10b. More caution must be used when viewing DSfM reconstructions, in particular if automated computer vision methods are employed as part of the analysis.

Resolution Demands

As an indicative study of the resolution needs of these methods, reconstructions were performed with repeatedly down-sampled versions of the original 12 megapixel captured images until reconstruction methods failed. However, each reconstruction target varied in size relative to the overall frame of the images. The critical factor is not the resolution of the camera, but the pixel density (pixels per inch) of the reconstruction targets in images. In the context of robotic inspection, pixel density demands determine the maximum viewing distance of the robot, relative to the reconstruc-

tion target, so as to produce an accurate reconstruction. Therefore, the results of the small-scale resolution analysis are reported in terms of the approximate required pixel density (pixels per inch).

For the scale structure and the aluminum section, a resolution of about 50 pixels per inch was necessary to create a stable reconstruction using either method. The bridge column required a resolution of approximately 20 pixels per inch. Higher resolution images than those necessary for a stable reconstruction provided more accurate reconstructions with fewer artifacts, but the benefit was often minimal. The higher resolution demands of the scale structure and aluminum section are due to the relatively less textured surfaces of those objects. The results of this analysis are not meant as definitive, as reconstruction capability is dependent on many variables such as sensor sensitivity, lens quality, and lighting conditions. However, the results served as a guideline for the imaging demands of the full-scale robotic test.

4.4 Full Scale Testing

The purpose of the full-scale test was to provide proof-of-concept verification using an unmanned aerial vehicle (UAV) to comprehensively image and reconstruct a local bridge. The bridge used in testing is an in-service reinforced concrete arch bridge that carries pedestrian traffic and spans another pedestrian walkway (Figure 4.14). The bridge is believed to have originally been a rail bridge repurposed for pedestrian use. The bridge was chosen so as to mitigate safety concerns inherent to operating drones in automotive traffic.

The results of small-scale testing yielded two findings that were applied to the full-scale testing. DSfM was shown to produce accurate and highly flexible reconstructions without detailed knowledge of the underlying model and was therefore the only method utilized at full-scale. The lack of an existing underlying geometric model coupled with the relatively complex architectural finish of the bridge ruled out the use of IM. The reconstructions were performed using 123D-Catch, as was done in the

Figure 4.14: Pedestrian bridge chosen for full-scale reconstruction

small-scale study. Also, the maximum viewing distance of the inspection robot was chosen so as to capture images at a pixel density greater than 20 pixels per inch. The 20 PPI criteria was chosen based on the results of the small-scale bridge column results, which had similarly textured concrete surfaces.

The inspection was performed using a human-piloted Parrot AR Drone 2.0 UAV [Bristeau et al., 2011]. Due to the low resolution of the onboard video camera, and a loss of image resolution due to video encoding artifacts, an 11 MP auxiliary camera was used to image the bridge. The auxiliary camera used a rectilinear fisheye lens with a wide horizontal field of view (135°). Automated lens correction was applied to correct the fisheye distortion using Adobe [2013].

The maximum viewing distance (D, inches) of the robot was determined based on the estimated pixel density demand (PPI, pixels per inch) for reconstruction, as well as horizontal sensor density (L_H, pixels) and the horizontal field of view of the lens

Feature	Field	DSfM
Feature A (corner spall)	$7^{1}/_{4} \times 4^{3}/_{8}$	$6^{1}/_{2} \times 4^{1}/_{8}$
Feature B (repair patch)	$3 \times 1^{1}/_{4}$	$2^{3}/_{4} \times 1^{5}/_{8}$
Feature C (repair patch)	$2^{7}/_{8} \times {}^{7}/_{8}$	$3^{1}/_{8} \times 1^{1}/_{16}$

Table 4.4: Measurements on pedestrian bridge (nearest $^{1}/_{16}$ inch)

($HFOV$, degrees). The generalized equation is shown in Equation 4.2. A maximum viewing distance of 3 feet was estimated for the specified UAV and served as a guide for UAV piloting.

$$D = \frac{L_H}{PPI} \cot(\frac{HFOV}{2}) \tag{4.2}$$

4.4.1 Results and Discussion

Reconstructions of the bridge are shown in Figure 4.15. As was done with the small-scale study, a series of tape-measurements were compared against measurements taken from the reconstructions. The measured objects are shown in Figure 4.16. The comparison is shown in Table 4.4. The performance of the method in the field was dependent not just on the choice of lens and sensor resolution, but also on the sensitivity of the digital sensor. It was observed that bright lighting scenarios with heavy shadows and a correspondingly broad range of lighting proved problematic during the reconstruction process. Not only did heavy shadows reduce the apparent surface texture, bright scenes often caused the camera to underexpose the scene and made reconstruction more difficult.

Many of the phenomena observed during small-scale testing were apparent in the full-scale reconstruction. The DSfM method created artifacts of a similar nature

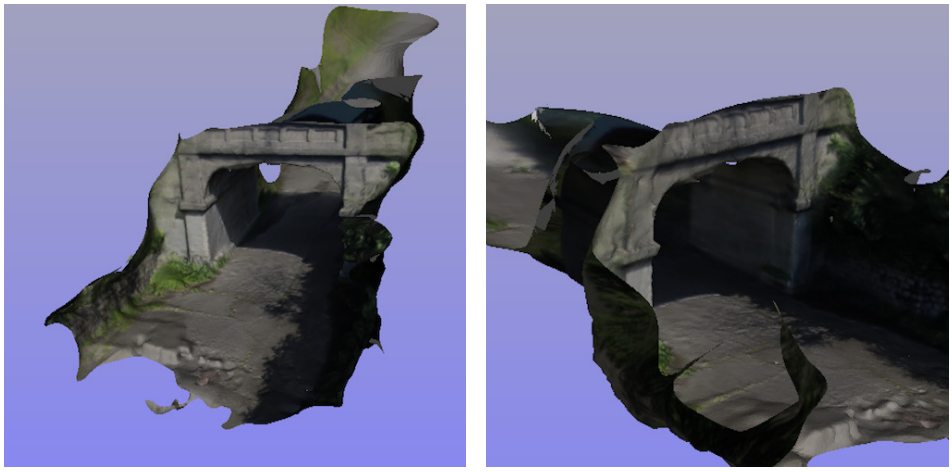


Figure 4.15: DSfM reconstruction of the pedestrian bridge

to those of the small-scale test. In particular, areas of dark shadows during the reconstruction process were problematic due to the low image texture in those regions (Figure 4.17). Several of the relatively large planar surfaces on the bridge distorted the reconstruction as well. While these regions were weathered concrete, and thus highly textured, the co-planar nature of the associated feature points reduced the accuracy of the SfM solution.

4.5 Conclusions

3D scene reconstruction is an area of active and on-going research in the computer science community. While the underlying algorithms that produce reconstructions continue to evolve, the two fundamental reconstruction paradigms presented here likely will not change dramatically for some time. This paper has presented the fundamental capabilities and limitations of these two primary photorealistic reconstruction approaches in realistic but simplified contexts, as well as a full-scale test of the application of the DSfM method to robotic inspection. The results of this study

Figure 4.16: Features used in full-scale measurement testing

indicate that both methods can be used to create 3D reconstructions for photorealistic structural inspection provided attention is paid to the critical characteristics of each approach.

The most readily apparent observation of this study is that the choice of reconstruction paradigm should be based on the intended application. Ultimately, it is a choice between creating a naïve but completely adaptable reconstruction, or mapping image data onto a previously constructed geometric model. For routine inspections in relatively stable and known imaging environments, IM can provide stable and repeatable results with less computational demand due to the efficient representation of the underlying structure. For scenarios without well-known structural maps or that demand substantial reconstruction complexity, methods such as DSfM are the practical choice. It is not difficult to imagine combination approaches, as well, in which global IM-based results are used to drive additional localized reconstructions using DSfM. The spalling damage presented earlier would be a good example where such an approach could be suitable.

Robotic inspection necessitates several additional considerations. Care should be taken to verify the correct exposure parameters for any robotic camera, particularly

Figure 4.17: Artifact on underside of bridge due to low image texture

in scenes with broad dynamic ranges, prior to employing the robot. Motion blur in images is an inevitability for UAVs, and so optimizing on-board cameras to mitigate this effect through proper choice of sensor and lens is essential. The choice of camera lens is critical not just to maximize camera shutter speed, but also to minimize reconstruction errors due to field curvature distortion of the lens. Lastly, the possibility of providing supplemental on-board lighting to minimize dark shadows should be considered.

The results of this paper demonstrate the basic feasibility of generating photo-realistic reconstructed 3D models suitable for enhancing current robotic inspection procedures, but much work remains to be done in terms of broader evaluation and improvement of such methods. There are numerous opportunities for further developments associated with repeated imaging over time, combined machine-learning/data mining and human data analysis, and inclusion of different robotic sensing modalities, all of which are of particular interest for purposes of inspection and monitoring.

Chapter 5

FIELD-ROBUST CRACK DETECTION

(WITH PERMISSION FROM ASCE)

5.1 Introduction

The application of systematic image capture and computer vision techniques holds the promise of significantly changing how infrastructure systems are inspected and maintained. In particular, enabling the automatic location, assessment, quantification, and tracking of visible (or otherwise detectable) damage in appropriately captured inspection images/data can add substantial enhancement to the largely ad hoc, qualitative inspection systems in use currently. A structural inspection vision system could quantify observed damage to a level likely impossible for a human inspector and remove much of the bias associated with an inspector's personal judgements on the severity level of structural damage.

As discussed in detail below, computer vision methods have been successfully applied in infrastructure-oriented situations that provide consistent lighting and structural surface quality, typically with relatively small numbers of images. This paper considers the challenge of addressing highly variable field inspection scenarios and large image sets, which are key issues for broad, practical application of this approach. In particular, concrete crack detection is considered as a damage pattern for computer recognition and analysis based on the common presence of such damage in structures, and based on the computer vision challenges associated with the high variability in concrete surface appearance amongst structures.

5.1.1 Computer Vision For Field Inspection

In order to train a computer to perform automatic image analysis, three core steps are necessary: (i)*segmentation* of the image; (ii) *feature extraction*; and (iii) *recognition* [Gonzalez et al., 2010]. *Segmentation* is the process of isolating only those elements of an image relevant to an analysis, often represented as binary objects for ease of computation and analysis. *Feature extraction* is the process of parameterizing any objects found in the binary image resulting from segmentation and, for the purposes

of this research, creating a feature vector that represents all of the relevant parameters associated with a found object. These first two steps are diagrammed in Figure 5.1. *Recognition* is the process by which a computer analyzes each of these object-specific feature vectors and categorizes them, typically via machine learning algorithms. Once a computer recognizes damage in an image, further image processing and analysis for damage tracking and quantification can begin.

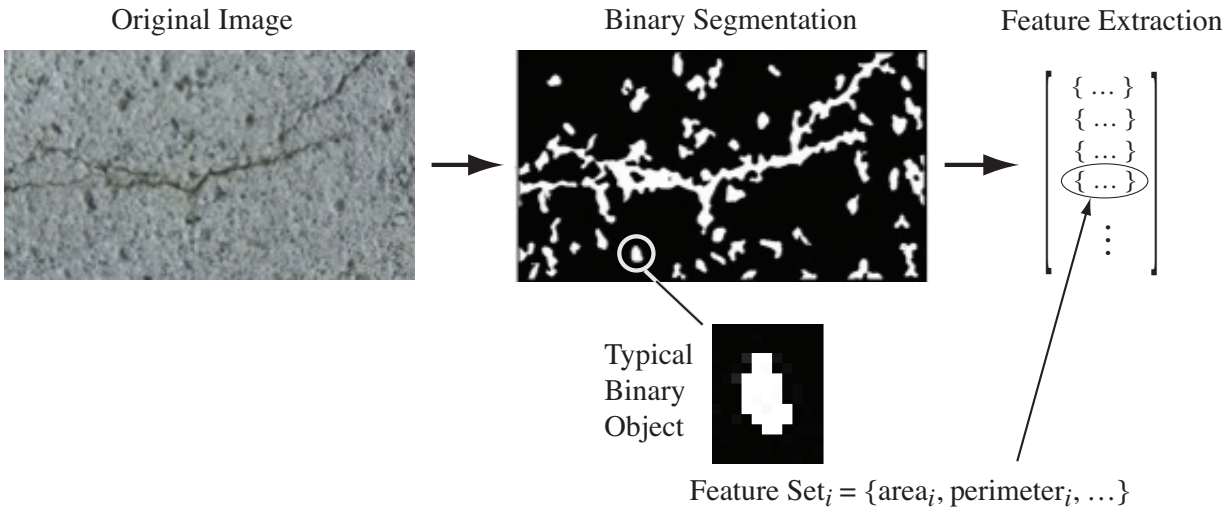


Figure 5.1: Schematic example of segmentation and feature extraction process

The key performance metrics for a field-ready computer vision system are classification accuracy, classifier robustness, and computational cost. To be successful, a trained machine-learning classifier must be able to consistently and accurately analyze inspection images. For general field use, it must also robustly handle a wide range of variations in images, such as the changes in lighting, texture, debris, staining, and contrast that accompany inspection photographs. Similarly, it must operate fast enough to process and analyze large numbers of images within a reasonable time frame for use by an engineer.

5.1.2 Previous Work

Previous crack segmentation research generally has focused on controlled contexts, thereby limiting variability in the images considered. The emphasis has been on the segmentation of cracks from images of concrete surfaces using edge detection schemes. Edge detection schemes generally look for high gradients in pixel values as indications of a physical edge captured in the image. The representation of cracks as edges is intuitive and such methods tend to be very fast in implementation. Edge finding segmentation covers a wide range of related methods, the likes of which were studied and discussed in length in Abdel-Qader et al. [2003]. They determined that edge detection based on Haar wavelet decomposition was the optimal edge detection method. However, the Canny edge detection method [Canny, 1986] was shown to be almost as accurate. The provided analysis did not consider robustness to image variability within the test set. Guo et al. [2009a] combined Canny's method for edge detection with a region growing method to improve segmentation results, but their system required controlled artificial lighting for accuracy. Yu et al. [2007] developed a similar system, but also found that consistent results required controlling the light source in each image. Hutchinson and Chen [2006] addressed the problem of brittle tuning parameters used in edge detection algorithms. Jahanshahi et al. [2009] compared the various available edge detection methods available and their general limitations with respect to "noisy" concrete surfaces.

Another popular segmentation method for crack detection was proposed in Yamaguchi et al. [2008]. Noting the inherent difficulty of using edge detection methods on noisy images, and the low accuracy of morphological methods, they developed a method which uses localized connectivity relationships to "percolate" crack regions in images. This method does not require pixel-to-pixel correspondences to be similar across an image or image set, unlike edge detection methods. More recently, they attempted to speed up their algorithm to scale to larger image sets [Yamaguchi and

Hashimoto, 2010]. However, they noted that the speed of their approach is tightly linked to texturing and noise in the image, with significant slowdowns occurring for highly textured and weathered concrete surfaces. Their system was modified by Zhu et al. [2011], working in tandem with an edge detector for faster segmentation results. However, this system is prone to the noise susceptibility of all edge detection based methods. A pipe inspection robot which adapted the percolation technique was developed, but its performance on field images was indeterminate [Walter et al., 2012].

There has been less work on how to optimize the feature extraction and recognition process, and some of the previously discussed methods use single scalar classifier thresholds [Abdel-Qader et al., 2003, Yu et al., 2007, Walter et al., 2012]. Sinha and Fieguth [2006] developed a feature extraction scheme that utilized size, shape, and orientation descriptors for the parameterization of crack objects in images. However their method is neither scale not orientation invariant. Abdel-Qader et al. [2006] used principal component analysis on sub-images derived from edge detection to detect cracks in localized regions, however the method is dependent on the size of the sub-images and is therefore not invariant to changes in camera perspective and distance. Chae et al. 2001, 2001 developed a fuzzy neural network approach for feature extraction and recognition of cracks in images. However the descriptors used in this work included mostly scale and orientation variant descriptors such as angle, width, and length.

5.1.3 Research Motivation

The methods used in these past studies can be effective provided the variations in both the calibration and field images are limited. However, when presented with more general (i.e., messier) image sets, the performance of these algorithms has been shown to drop significantly. The reality is that almost any crack detection method, and more generally any computer vision method, can be shown to provide a high

degree of accuracy given enough control over the parameters in the image test set. The primary reason for these successes are the inclusion of finely tuned thresholding parameters which filter out information to provide a segmented image. In the case of Canny's method, these thresholding parameters are used to hysteretically filter the edge detection response of the algorithm. With wavelet based methods, it is the high and low pass wavelet coefficients which are filtered. Critically, these thresholding parameters rely on the concept that the localized pixel to pixel correspondences which correlate to cracks remain constant throughout both the image itself and the entire set of images being analyzed. When the relative correspondences which represent cracks within an image change, the thresholding parameters begin to improperly filter results. This algorithm brittleness is aggravated by classifiers which have not properly considered how to robustly describe segmented crack objects and often use only a simple, single, and scale variant description of segmented objects, such as linearity. This leads to computer vision system which have been overtrained on limited data sets.

Segmentation methods which do not rely on consistent localized correspondences, such as percolation methods, avoid the tuning parameter problem and are typically more robust to image variance. This observation has motivated the work presented in this paper, which has focused on developing an algorithm that can perform with improved consistency in the face of significant fluctuations in image features consistent with realistically varying field conditions, all while performing fast enough to be able to efficiently handle the hundreds of images captured during a field inspection. The resulting algorithm, which is optimized to address these concerns, is the primary contribution of this work.

This alternative method uses a custom clustering segmentation algorithm combined with a robust feature descriptor vector that performs well under a wide variety of conditions without localized parameter tuning which is dependent on the imaging context. The effectiveness of this method is compared to commonly implemented

edge detection algorithms using image sets captured from a variety of infrastructure contexts, and it is shown to be more consistent in detection and identification accuracy across varying image scenarios, such as those likely to arise in field inspection settings.

Furthermore, no study to date has attempted to analyze the relative significance of the various types of segmented object descriptors and machine learning methods with respect to crack detection. This analysis is provided as well.

5.2 Algorithm Development

5.2.1 Segmentation

As indicated previously, the first core process in a computer vision system is the segmentation of the image. For the purposes of this system, the intent of the segmentation is to extract binary representations of possible damage from images. These binary representations do not need to be exact representations of the damage in the image, and in most cases they are not. What is critical in this process is that the resulting binary representations can be easily and logically parameterized for eventual recognition by a trained machine learning model. After damage is detected in a processed image, there is an opportunity for further image processing to capture the full extent of the damage.

Method Selection

The criteria for identifying a suitably robust segmentation approach were to find methods that: (i) did not rely on localized pixel correspondences; and (ii) could be computed fast enough to handle large-scale applications, potentially using embedded processors. This method would then be compared against commonly implemented edge detection based systems.

In regards to performance, based on the Federal Highway Administration's bridge

inspection requirements [Hartle and Administration, 2004], it is reasonable to expect that for a medium size bridge at least 500 images per structure would need to be processed by a comprehensive field imaging system. For this number of images, a per-image processing time of 3 minutes each would require a total processing time of about 24 hours. For purposes of the long-term goals of the project driving the present research this was taken as an upper limit of acceptable performance assuming commonly available hardware.

This performance metric eliminated several crack detection methods implemented in previous work. Segmentation based on graph partitioning [Shi and Malik, 2000] and fractal segmentation methods [Chaudhuri and Sarkar, 1995] were not considered in the present work due to the inherently high computational cost associated with them. Color based segmentation methods were not applicable because of the essentially monochrome nature of structural concrete. Morphological methods, which are inherently not shape or scale invariant, were also not implemented.

Segmentation based on the percolation method of Yamaguchi and Hashimoto [2010] was implemented and tested. While computational times for relatively lightly textured concrete images met system performance requirements, the computational time for highly textured images of field concrete increased by an order of magnitude and the method was therefore determined to be unsuitable for this application.

Ultimately, a clustering segmentation method was chosen for further comparative study against commonly implemented edge detection techniques using Canny's method and Haar wavelet filtering. All three of these methods satisfied computational speed requirements, segmenting images in less than 1 minute on the benchmark system (a 2.2GHz multicore Intel laptop, C++, and the optimized image processing system OpenCV [Bradski, 2000]).

Clustering Based Segmentation Via k-means

The clustering method developed is a variant of a k-means clustering algorithm [Hartigan and Wong, 1979]. K-means clustering satisfied the research objectives of not relying on localized pixel correspondences and associated sensitive tuning parameters while meeting computational speed requirements. The method is also inherently scale and orientation invariant, requiring no previous knowledge of the size and shape of image objects. Finally, clustering algorithms can be largely parallelized for optimization. The downsides to k-means clustering are that the method is not inherently invariant to change in image lighting and the method can be highly sensitive to the initial random values chosen [Guo et al., 2009a]. Furthermore, pixel intensities in concrete images tend to follow a Gaussian distribution, making it difficult to distinguish crack pixel values in an image histogram, a common technique for improving the performance of a k-means algorithm [Pena et al., 1999].

Clustering Via k-means*

These shortcomings were addressed by developing a custom variant to the classic k-means algorithm that exploits particular characteristics of the image sets to generate a consistent segmentation. This variant overcomes known deficiencies with the k-means algorithm with regards to crack detection. This approach will be referred to as a k*-means algorithm, indicating its conceptual link to standard k-means clustering techniques.

The key characteristic for this application is the fact that cracks show up as some of the darkest regions in an image. Therefore the cluster with the lowest mean value will likely contain any pixels associated with a crack. The k*-means algorithm exploits this a priori knowledge and allows for a more intelligent initialization of the cluster values. By skewing the initial values of the clusters so that most cluster values are initially set to bright pixel values, dark pixels within cracks are more likely to be

clustered together and more likely to be cleanly separated from the remainder of the image. In particular, skewing via a logarithmic distribution of initial cluster values has been used to produce the results presented below. An algorithm flowchart for the method is illustrated in Figure 5.2.

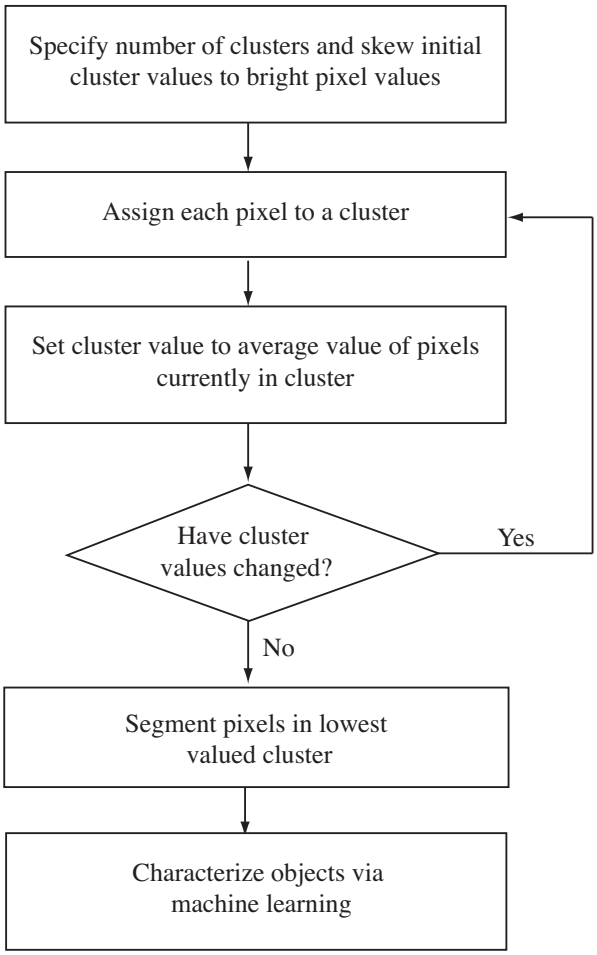


Figure 5.2: Flowchart for the k*-means algorithm.

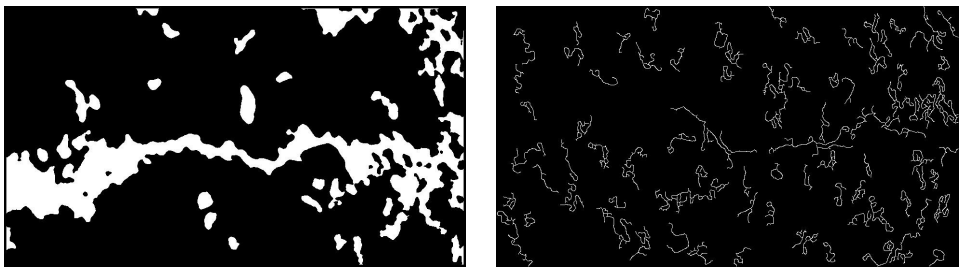
Prior to applying the clustering method, each image can be normalized for lighting variance via standard histogram equalization [Acharya and Ray, 2005] to improve the algorithm's performance. Also, specifying ten initial clusters was empirically determined to provide a consistent result across a series of test images selected to have

a wide variety of image characteristics and noise patterns. The basis for this empirical choice was that ten clusters provided the most consistent segmentation of crack areas without clustering contiguous crack information across disjoint clusters. While this method, like almost all segmentation methods, requires setting various parameters (number of clusters and initial cluster values), these parameters are more robustly set for a wider range of images. The hysteretic thresholding parameters used with Canny's method, and the wavelet coefficients thresholds for wavelet based edge techniques, are related to the relative intensity of the crack pixels with respect to concrete surface pixels. If surface conditions change, as with water staining and concrete weathering, the relative intensity of crack pixels changes and edge detection performance with the original thresholding values degrades accordingly. With a clustering approach, crack pixels will almost always appear as the darkest regions of the image, and the method is not as dependent on local relative pixel intensity for performance. The specified number of clusters and the initial cluster values are globally related parameters which are more general in nature. The downside is that small pockets and defects are often clustered along with cracks when using k*-means. However, these additionally clustered objects can be removed with machine learning techniques. An example of the binary image which results from k*-means is shown in Figure 5.3.

Edge Based Segmentation via Canny's Method and Haar Wavelet Filtering

Canny's method is, like most edge detectors, based on convolving gradient filters with a grayscale image [Canny, 1986]. The resultant edge image is then hysteretically thresholded to produce the edge segmentation. The hysteretic thresholds must be specified by the user and can produce widely varying results depending on the image test set. For the present research the hysteretic thresholds were determined empirically in order to prevent crack edge information from being filtered out of the resulting binary images (see e.g. Hutchinson and Chen [2006] for a discussion of thresholding issues and strategies). An example of the binary image which results from Canny's

method is showing in Figure 5.3.



(a) Result from k*-means (b) Result from Canny's Method

Figure 5.3: Examples of resulting binary images from segmentation

Haar wavelet crack detection was implemented based on the work of Abdel-Qader et al. [2003]. The image is first transformed using the Haar wavelet transform. A high pass filter is then implemented in wavelet space and the resulting coefficients are transformed back into pixel space to produce an edge image. A crack is then classified based on the average intensity of this resultant image. For this research, the filtering and crack classifier thresholds were determined empirically[Hutchinson and Chen, 2006].

In this study, only Canny's method is used in conjunction with a machine learning classifier, as it simplifies the creation of continuous line segments for robust feature extraction relative to wavelet based methods with only a slight decrease in accuracy. Research by Hutchinson and Chen [2006] corroborate this result. The performance of the Haar wavelet approach using edge intensity as a classifier metric is discussed below in the overall testing section.

5.2.2 Feature Extraction

After the image has been converted into a segmented binary image, a connected component analysis is performed to catalogue individual objects in the image [Gonzalez

et al., 2010]. The basic idea is to construct connected sets (blobs) of matching pixels, with each connected set considered to represent an object of possible interest. Once these binary objects in the image have been isolated and identified, they can be parameterized and represented using vectors of parameterized features. As previously discussed, prior research generally has focused on the use of shape and orientation descriptors to describe crack objects in binary images. These descriptors typically include object length, area, perimeter, orientation, aspect ratio, and compactness (also known as roundness). The aspect ratio of an object is defined in terms of distances that are precomputed during the connected component analysis, specifically as the ratio of the major axis length to the minor axis length. Compactness of a region is defined as the square of the perimeter divided by the area [Gonzalez et al., 2010].

A key goal of this work has been to develop a crack detector that is not only robust to variance within the image set, but also maintains scale and orientation invariance. For this reason, all scale variant shape descriptors, including length, area, and perimeter length were excluded from consideration. Object orientation was also excluded. The remaining relevant shape descriptors were aspect ratio and region compactness. Analysis indicated that including compactness descriptors has a small but positive impact on overall results. However, it is not possible to parameterize the compactness of line segments resulting from edge based segmentation, as they are one-dimensional objects and therefore have no area. Thus, in the following analyses, compactness was used for the clustering based system only.

The use of only one or two descriptors (aspect ratio and compactness) leads to a feature vector with only one or two attributes, creating an inadequate descriptor vector for a machine learning algorithm and potentially leading to overtraining. Therefore, a technique was developed to incorporate regional textural information into the object description, as well. Textural descriptors attempt to describe a region by parameterizing its co-occurrence matrix, a representation of pixel intensity variance in a region [Chanda et al., 1985]. The most common descriptors include

correlation, contrast, energy, homogeneity, and entropy [Gonzalez et al., 2010]. All of these descriptors are relatively scale and orientation invariant and were included in the feature descriptor. The resulting feature descriptor is not only scale invariant, but it is semi-universal in that it can be used to describe connected component objects found by clustering as well as line segments found by edge detection techniques, with the minor exception for compactness as noted previously.

The effectiveness of a feature descriptor is measured by its utility in a developed classifier model. Two analysis methods have been used in the present work: (i) the χ^2 test; and (ii) a recursive feature elimination classifier. The reader is referred to Duda et al. [2001] for an explanation of the χ^2 test, which evaluates a class's statistical dependence on a given feature. The recursive feature elimination method functions by building a classifier using all available features and then recursively eliminating the least significant features with regards to classifier accuracy. For this analysis a support vector machine (SVM) classifier was used [Vapnik, 2000].

5.2.3 Object Recognition

The final step in the prototyping process is to construct a machine learning classifier to perform object recognition on extracted image objects. The complete set of parameterized instances, including the class of each object, are used to train the model and verify its accuracy. For this paper, the instances are descriptor vectors for each segmented image object, and the classes have been specified in binary fashion as "crack" or "not a crack". The classes of the training data were manually indicated prior to training, a technique known as *supervised learning*. This introduces some human bias and variability into the overall process, which can result in errors in the test set's "ground truth", especially given the large set of images requiring tagging. Care was taken to minimize such errors, but there is a possibility of some small percentage error in the test set classification.

Tested Algorithms

Within the field of supervised learning there are generally four model types: statistical; instance based; rule based; and generalized linear regression [Witten and Frank, 2005, Duda et al., 2001]. Within each model type, there exists a variety of algorithms and variants, each with its own strengths, weaknesses, and opportunities for tuning. The purpose of this research was not to determine the optimal machine learning algorithm for this system, which could only provide a fleeting answer considering the rapidly evolving landscape of machine learning. Instead, a subgoal was to explore the benefits and drawbacks of the various methodological archetypes to provide a more broadly representative point of reference. Classifiers from each of the four model types (statistical, instance, rules, generalized linear regression) were tested and compared. Adaptive Boosting was utilized where appropriate [Freund and Schapire, 1996].

The statistical classifiers used in this research were a Naive Bayes [Blum and Langley, 1997] classifier and a more sophisticated Bayesian Network [Heckerman et al., 1995]. A K-Nearest Neighbors (KNN) algorithm was used to construct an instance based classifier [Aha et al., 1991]. A pruned decision tree (J48) was used as the representative rule based classifier [Quinlan, 1986]. And a multi-layer perceptron (ANN) was used for the generalized linear regression classifier [Bishop, 1995]. Ten-fold cross-validation [Kohavi, 1995] was used to train and test the classifier models.

A One-Rule (One-R) classifier was also developed to establish baseline performance. As one of the simplest classifiers possible, it represents a comparative benchmark for other methods [Holte, 1993], indicating the value of implementing more complex and computationally expensive machine learning models. The One-R classifier is used as such a benchmark in the later discussion of recognition analysis.

Performance Metrics

To measure classifier accuracy, the simplest metric is the overall accuracy of the classifier (% correct). While this is often a reasonable measure, it does not indicate if the data are biased towards classifying all instances one way or another, nor does it represent the cost of a false positive or false negative, which may be different. Therefore, the true positive rate (% true positive) and true negative rate (% true negative) are also included in the results. In the following analysis, a positive detection is defined as a determination that an extracted object is a crack.

Receiver operating characteristic (ROC) curves can be used to represent overall classifier accuracy in a more general way [Provost and Fawcett, 1997], taking into account the cost of a false positive or negative classification. The true positive rate is plotted against the false positive rate, where the the classifier's discrimination threshold between positive and negative can be shifted to produce differing results (see Figure 5.7 below for an example ROC curve plot). A perfect classifier would plot as a vertical line based at zero false positives and a horizontal line at a true positive rate equal to 1.0, with an area under the curve of 1.0. The area under a non-ideal ROC curve provides a complimentary statistic that can be used to provide a numeric summary of an algorithm's overall level of performance.

The build time for each classifier model is also presented in the recognition analysis discussion as a measure of computational cost.

5.3 Testing and Analysis

Image segmentation and feature extraction were performed using MATLAB [Inc., 2011]. All three portions of the computer vision system were tested using the open source machine learning package WEKA [Hall et al., 2009]. All classifier training and testing data were evenly split between extracted crack and non-crack features. Models were trained and tested using a ten-fold cross-validation process, unless explicitly

noted otherwise. All cost sensitive parameters were set for equivalent costs for both false positives and false negatives.

5.3.1 Image Test Set Development

Prior to vision system testing, the acquisition of an image set is necessary. For many common computer vision problems, standard image sets are available in public repositories for consistent comparative algorithm testing. However, for the selected damage type, concrete cracking under general field conditions, no such standard repository exists. Instead, images were acquired using a hand-held consumer five megapixel digital camera. In order to develop a vision system which improved the robustness of current methods, images were captured which captured the highly variable imaging conditions which occur during inspections. Image variations include regional texture, concrete efflorescence, water staining, crack width, extraneous objects, construction joints, image blurring, and lighting levels.

Image Sets

Three separate image sets were collected, referred to as Sets A, B, and C. Set A contains images of well-controlled concrete surfaces. Some light staining or lighting variations are present but, in general, the quality of the concrete is high. Crack widths are consistent. This set is representative of concrete in a new structure or laboratory setting. Set B contains images of weathered and highly textured concrete surfaces. Scaling and staining are much more prevalent in this image set, and crack widths are more variable. This set is intended to represent deviations from controlled imaging scenarios, as occur frequently during field inspection.

Sets A and B were used in combination for most tests as a means to build the most robust classifier possible. In the image variation tests described below, Set A was used as the training set and Set B as the test set. This simulates a scenario in which a computer vision system has been trained around one type of concrete

surface and then used in an application where crack characteristics varied from the "known" representations. An example: a vision system developed and tuned using images captured from a new bridge and then applied to images from a 60-year-old bridge. An overtrained vision system will inevitably perform worse than a robust system in such a test, providing relative insight into a system's feasibility for field implementation.

Set C images have a mix of variations. Lighting variations are present, as are other small surface defects. Set C was used for the overall systems validation presented near the end of the paper. The image counts for each Test Set and the produced crack and non-crack objects for each segmentation routine are provided in Table 5.1. Note that Set C was used to test crack detection per image (as opposed to per segmented object) and therefore the number of objects is not reported in Table 5.1. Example images for each set are shown in Figures 5.4, 5.5, and 5.6.

	Total	Crack Objects		Non-crack Objects	
	Images	k*-means	Canny	k*-means	Canny
Image Set A	797	904	2486	904	2486
Image Set B	36	38	71	38	71
Image Set C	200	-	-	-	-

Table 5.1: Image test set breakdown

5.3.2 Segmentation Algorithm Testing and Analysis

Comparisons between the Canny edge detection and k*-means clustering segmentation methods were performed using machine learning performance metrics at the

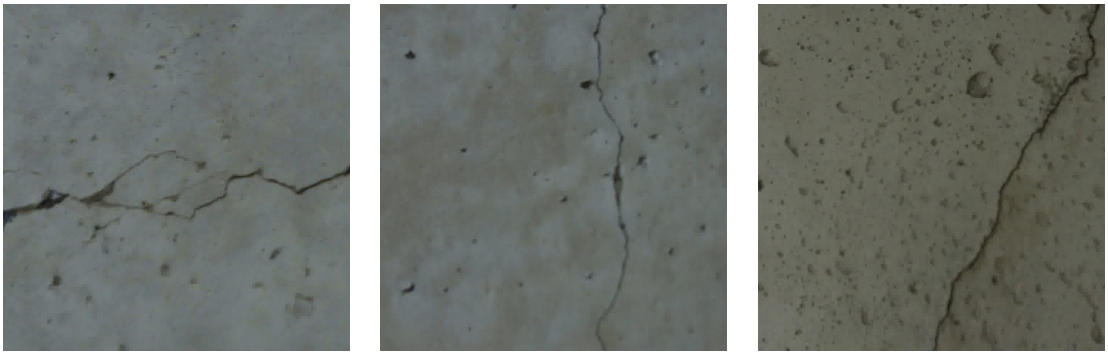


Figure 5.4: Typical image examples from Set A

Figure 5.5: Typical image examples from Set B

object recognition stage of the analysis. For the analysis the results of both segmentation methods were extracted using identical feature descriptors, and identical machine learning methods were trained and tested. The feature descriptor used was the complete vector detailed earlier including the following parameters: aspect ratio, compactness, energy, homogeneity, entropy, correlation, and contrast variance. Compactness was not included for the edge detection methods for the reasons previously discussed. The machine learning classifiers trained to compare segmentation routines were a naive Bayesian statistical classifier, a boosted J48 decision tree using adaptive boosting, and a k-nearest-neighbor instance based classifier (KNN). These three algorithms were chosen not because they were the optimal classifier choices, but

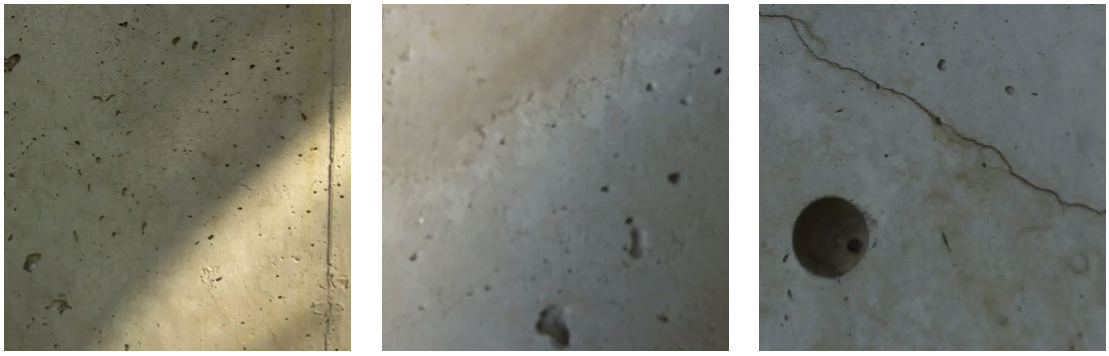


Figure 5.6: Typical image examples from Set C

because they require minimal parameter tuning at the training phase, clarifying the comparison. Support vector machines and neural networks, effective classifiers in this vision system, were not included in comparison of segmentation routines due to the number of tunable parameters. This provides a comparison of how successfully the system is able to extract and recognize the image information that is provided by each segmentation routine. A more general test of the accuracy of such a system relative to a simple classifier is found in the overall system validation section presented near the end of the paper.

Results of the segmentation effectiveness tests are provided in Table 5.2. It can be seen that accurate object recognition was possible for both types of segmented objects, with little to no statistical difference in overall accuracy between the two segmentation object types. The largest variation in performance can be seen for the Naive Bayesian classifier, with k^*-means outperforming the Canny method by a small amount.

Image Variation Testing

In order to further assess the robustness of the two segmentation methods, Test Sets A and B were separated. Set A was used to build and train a classifier model, which

	Naive Bayes		J48		KNN	
	k*-means	Canny	k*-means	Canny	k*-means	Canny
% Correct	89.8	88.7	92.6	92.8	92.4	91.1
% True Pos.	96	91	94	92	93	91
% True Neg.	83	86	91	93	92	91
ROC Area	0.962	0.942	0.973	0.971	0.927	0.911

Table 5.2: Classifier results for two segmentation methods

was then tested on Set B. The purpose of this test was to determine the capability of each computer vision system to handle unfamiliar image scenarios (recall that the Set B images are much noisier than those in Set A). Once again, the same three minimally-parameterized classifiers were considered: Naive Bayes, J48 decision trees, and k-nearest-neighbors. The J48 classifier was boosted using adaptive boosting.

The results for this test are shown in Table 5.3 and the Receiver Operating Characteristic (ROC) curves for the J48 classifier are shown in Figure 5.7 for both the previous low-variation image sets and the current high variance sets. The other classifiers produced similar ROC curves.

	Naive Bayes		J48		KNN	
	k*-means	Canny	k*-means	Canny	k*-means	Canny
% Correct	72.4	87.3	82.9	81.7	86.8	72.5
% True Pos.	100	85	97	70.4	97	59
% True Neg.	45	90	68	93	76	86
ROC Area	0.918	0.913	0.953	0.918	0.868	0.725

Table 5.3: Classifier results for image variance test

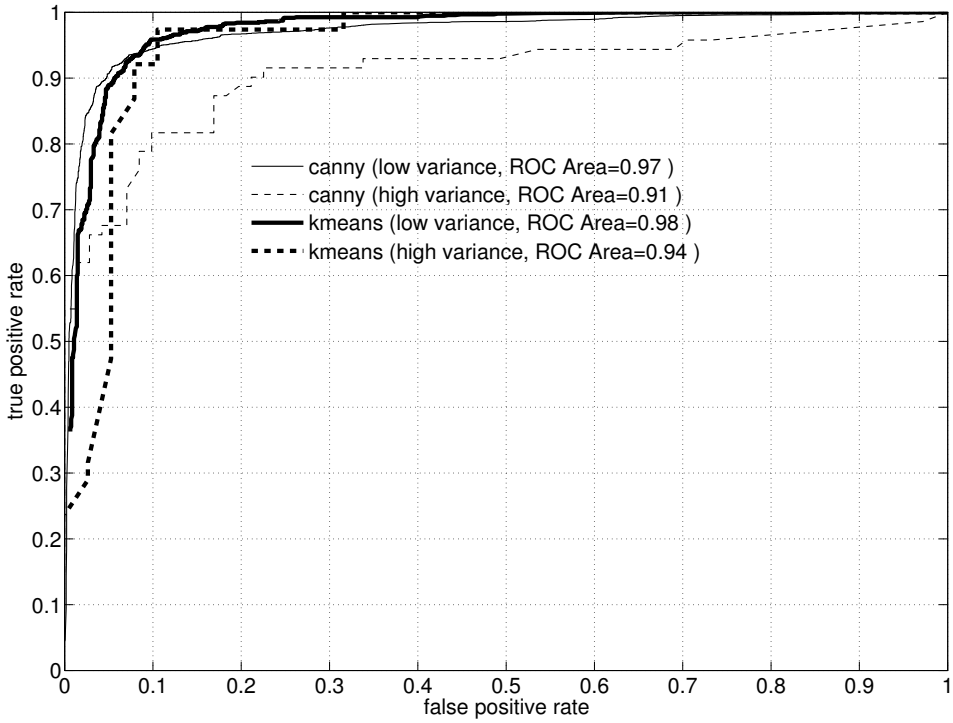


Figure 5.7: ROC plots for low and high variance image sets

In all cases, performance degrades compared to results using the complete image set. However, recognition performed using Canny's method for segmentation is relatively inferior if the primary metrics are the accurate detection of cracks (% True Positive) and the ROC curve area, as opposed to classification of non-cracks. The ROC plot for the J48 classifier (Figure 5.7) is indicative of this issue. Both detection systems perform comparably with the complete image set. However the ROC of the Canny based system degrades significantly with the more varied set of test images. This difference is likely due to the additional tuning parameters required for Canny's method, which contribute to algorithm (and therefore system) brittleness.

5.3.3 Feature Descriptor Testing and Analysis

To better understand the functioning of the feature classifier, the parameters of the feature vector were compared and ranked via a parameter sensitivity analysis performed in Weka. This sensitivity analysis was performed for both the k*-means and Canny segmentation algorithms. Furthermore, the effectiveness of a combined shape and texture based descriptor was compared against vectors using only texture or shape information.

The sensitivity study was performed using the attribute evaluation capabilities of Weka. Two evaluative methods were considered: a ranking provided from a SVM classifier using recursive feature elimination, and a χ^2 score, both of which are presented in Table 5.4. The SVM values represent a simple integer ranking, while the χ^2 values represent the weighted statistical significance of each feature, with values dependent on the total number of feature instances. The most significant feature scores are shown in bold.

These results indicate that not only does the importance of each parameter vary between the k*-means and Canny segmentations, but their importance varies with the chosen metric. What is apparent is that, for both types of segmented features, both textural and shape descriptions provide relevant information for machine learning. Based solely on χ^2 scores, the clustering derived classifier emphasizes a subset of the descriptors (although all parameters are relevant), whereas the Canny derived classifier relies on a more even combination.

This result is further verified by comparing classifier models built with either shape or textural descriptors against a classifier model using the combined information. For the combined classifier, a J48 decision tree was used as the classifier. Other classifiers provided comparable results. In both cases, the classifier model using the combined description outperformed the models using only the textural or shape subset, as evidenced by the ROC plots in Figures 5.8 and 5.9. In no case did the inclusion of all

	k*-Means		Canny	
	SVM	χ^2	SVM	χ^2
aspect ratio (S)	7	482	**7**	**4620**
compactness (S)	2	446	-	-
energy (T)	**8**	868	4	3999
homogeneity (T)	6	80	5	3660
entropy (T)	5	553	6	4073
correlation (T)	4	**923**	1	1090
contrast (T)	3	43	3	3660
variance (T)	1	9	2	4054

Table 5.4: SVM classifier attribute rankings and χ^2 scores. (S) denotes shape descriptors, (T) denotes textural descriptors

feature descriptors actually degrade classifier performance. This can be a common problem with machine learning classifiers, so this is another indication of robustness in the overall framework.

5.3.4 Object Recognition Testing and Analysis

After verification of the segmentation routine and consideration of feature vector sensitivity, a study was performed to determine an optimal classifier type. The optimal classifiers from each machine learning paradigm (determined during preliminary testing) were tested and compared: Bayesian networks (BNet); J48 decision trees; multilayer perceptron (ANN); and k-nearest-neighbors (KNN). Algorithm boosting using adaptive boosting and 10-fold cross-validation were utilized where applicable. A one-rule (1-R) classifier was also developed as a baseline.

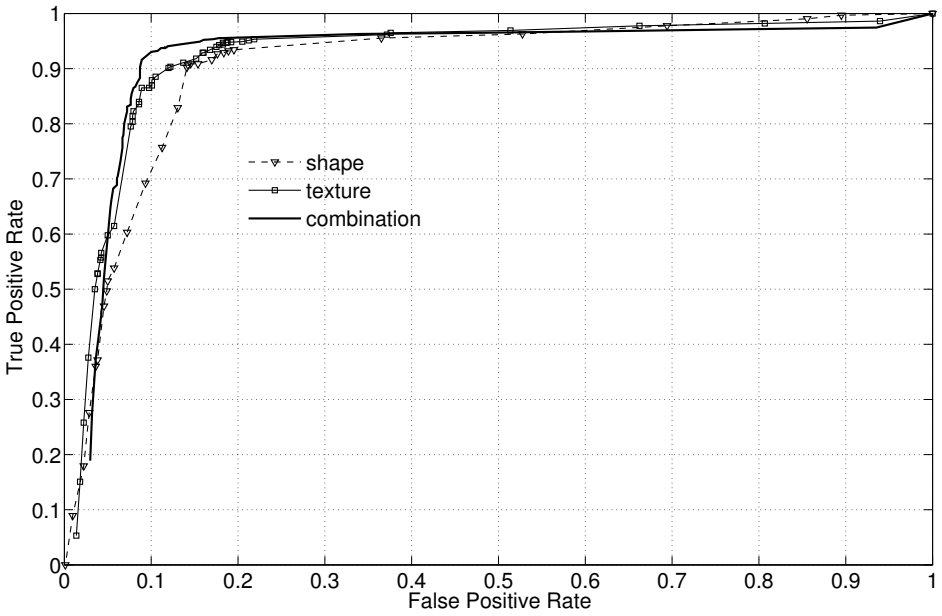


Figure 5.8: ROC comparison plot for k*-means segmentation

The results of this analysis are summarized in Tables 5.5 and 5.6, with the best outcomes shown in bold. These tables also contain the computer times required to generate the results using each algorithm, which are useful for comparing relative computational costs.

Once again, the results show that the four algorithms considered are comparable in performance, and all of them outperform the 1-R classifier, although with significantly increased computational costs. In terms of accuracy, there is little to distinguish between classifier methods, and a paired t-test indicated that there is essentially no statistical distinction in performance between these classifiers.

5.3.5 Overall System Validation

As a final test to verify the benefits of using a more complex machine learning approach compared to a system that uses a single classifier metric, the two fully devel-

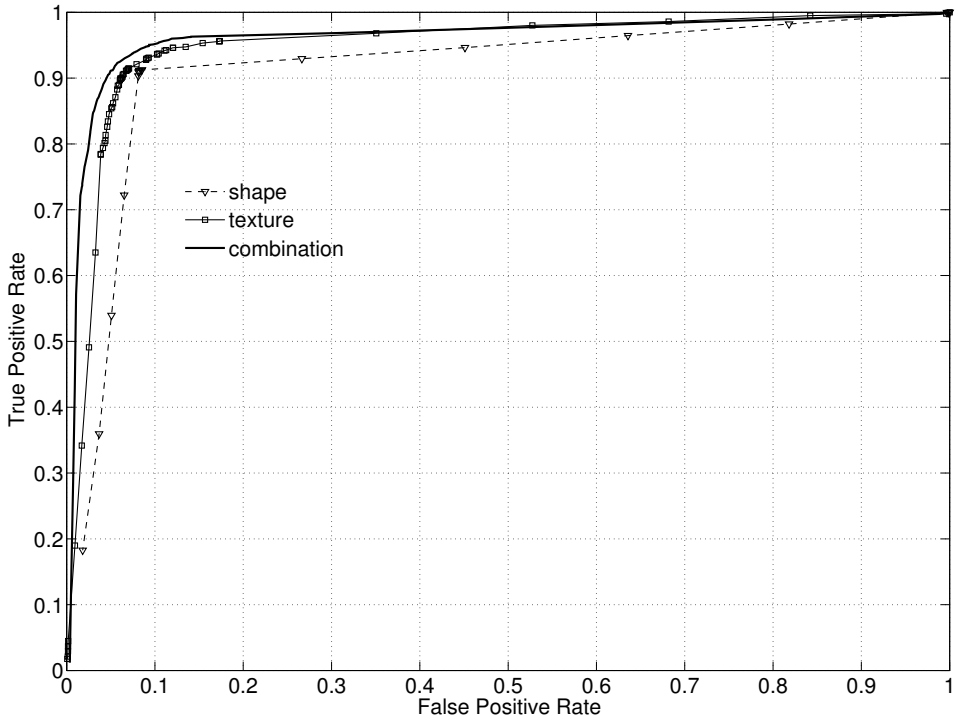


Figure 5.9: ROC comparison plot for Canny segmentation

oped and trained vision systems were run on a separate image test set of 100 images with cracks and 100 images without cracks. The established Haar wavelet method [Abdel-Qader et al., 2003], which uses average pixel intensity in an image as a classifier threshold, was the representative system using a single metric classifier. The test images, as described earlier, included lighting and surface quality variations. Based on its previously observed good accuracy and relatively low computational cost a J48 decision tree utilizing adaptive boosting was used as the classifier for the k*-means and Canny results. The outcomes of this analysis are shown in Table 5.7.

Both machine learning based systems performed comparably well and it is apparent that both systems are weighted to produce false positive results far more often

	1-R	BNet	J48	ANN	KNN
% Correct	86.6	91.5	**92.7**	92.6	92.4
% True Pos.	89	92	**94**	93	93
% True Neg.	84	91	91	92	92
ROC Area	0.866	0.960	**0.973**	0.961	0.927
Build Time	0.09 s	**0.43 s**	0.74 s	18.4 s	0.46 s

Table 5.5: Classifier results for k*-means clustering-based segmentation

	1-R	BNet	J48	ANN	KNN
% Correct	91.0	92.9	92.7	**93.3**	91.1
% True Pos.	91	92	**94**	93	91
% True Neg.	91	94	91	94	91
ROC Area	0.91	0.971	**0.973**	0.965	0.911
Build Time	0.01 s	1.44 s	0.73 s	26.5 s	**0.33 s**

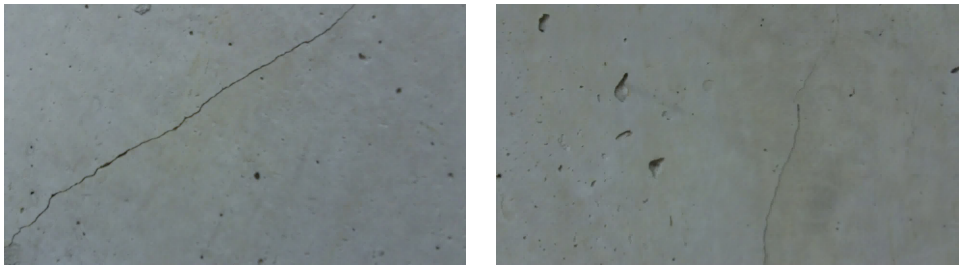
Table 5.6: Classifier results for Canny edge-based segmentation

than false negatives. It is instructive to observe the missed positive images for each case, which are shown in Figure 5.10. The Canny based system failed due to the faintness of the hairline crack combined with the staining around the crack edge. This likely led to a weakened edge response and the resulting edge was thresholded out. The miss for the k*-means system is less apparent. While difficult to perceive in the image, the crack had several gaps in it once the image was median filtered. This led to smaller objects that were not recognized as one continuous crack. In general, this last test serves as verification that using a machine learning approach for crack classification and detection improves detection accuracy rates compared to basic discriminators such as represented by the Haar results in Table 5.7. The high false

	k*-Means	Canny	Haar
% Correct	86.5	87.5	63.0
% True Pos.	99	99	62
% True Neg.	74	76	64

Table 5.7: System test results for k*-means, Canny, and Haar wavelet based systems

positive rate can be reduced using a cost-sensitive analysis in any implementation.



(a) k*-means method failure (b) Canny's method failure

Figure 5.10: Missed positive images for k*-means and Canny based vision systems

5.4 Summary and Conclusions

A computer vision methodology for automatically detecting cracks in concrete images has been developed and tested. The system is capable of accurate classification that is more robust to the image variations inherent in field inspections than established edge-based approaches. The complete system is computationally fast enough to handle large image sets, potentially in embedded applications.

Three segmentation methods were considered. An edge based approach using Canny's method and a clustering approach using a custom variant of the k-means algorithm (k*-means) were combined with a machine learning classifier. While both

segmentation methods perform well on uniform test sets, clustering segmentation is more robust for highly variable image sets and is therefore better suited as a segmentation method for general field image analysis. Clustering methods can also be more optimally parallelized for large-scale applications. Both machine learning approaches significantly outperformed a Haar wavelet based approach using a simple crack discriminator.

The combined shape and textural feature descriptor used in this paper is essentially invariant to crack size and orientation, and is an improvement over descriptors using only shape or texture parameters. One advantage of the k^*-means based system is the availability of shape descriptors that are not useable with the line segments resulting from Canny's method.

In general, the present results show that most machine learning methods will provide comparable performance with regards to accuracy, in particular if meta routines such as adaptive boosting are used. Conveniently, this means the choice of classifier can be made based on execution speed and ease of implementation within a larger framework. The most likely source of error in the results stems from possible mistakes during the manual tagging and registration of training and testing data.

Continuing work in this area is being done to expand the ability of the computer vision system to distinguish between different damage types. The detection system is also being implemented within a larger vision system which will be capable of not only detecting damage, but also of quantifying and cataloguing damage.

In addition to the technical approaches presented in this paper to improve automated damage characterization, the case is also made for the need for broadly available image databases to enable further research to move forward on common ground. Computer vision systems demand a large and varied set of images for optimal model performance, but the handful of sample images typically capable of being presented in a paper are inadequate for reproducing results, evaluating and comparing outcomes using alternative algorithms and approaches, and being able to apply

statistically valid analyses. This clearly limits the rate at which this field of research can develop. The authors propose the development of an open database for both submitting new images and accessing images used in prior research efforts. As a first step in this effort, all images used for the development of this paper can be obtained by contacting the authors.

Chapter 6

IMAGE ANALYSIS FOR POST-EARTHQUAKE DRIFT ESTIMATION

6.1 Introduction

This paper investigates the feasibility of using computer vision for quantitative structural condition assessment, motivated by the need for a better and more complete understanding regarding the estimation of damage in structural components. In particular, it considers the fundamental question as to the degree correlations between images of bridge columns and maximum column displacements exist, in the context of visually observable damage and an associated displacement history.

Current methods of structural inspection and assessment are limited in their ability to quantitatively estimate the extent of damage and deterioration that structural components have experienced. These limitations are manifested not only during routine inspections of structures, where qualitative estimates of damage have been shown to be highly variable and susceptible to inspector bias [Phares et al., 2004], but also during post-disaster triage assessment of critical infrastructure.

In such a triage scenario, rapid assessments must be made regarding the structural integrity of infrastructure to enable critical rescue and transport operations. Often, these assessments must be made without the time or training required to perform complex analyses of structures, and are therefore performed via simple visual observation. This process is normally not quantitative and exhibits many of the same variances as routine infrastructure inspections. However, inspectors typically take digital images of structures under inspection, providing a potential source of quantitative information.

Digital images can provide a powerful complement to qualitative inspection observations. Digital images are, in essence, complex sensor data. These two, and sometimes three, dimensional images can contain a great deal of information regarding the extent of damage and deterioration of infrastructure components. Extracting and making use of this information is an ongoing challenge that extends well beyond the civil engineering community.

Yet, image analysis provides only simple descriptions of damage states such as concrete cracking, metal corrosion, or concrete efflorescence. These descriptions do not inherently provide correlation to structural performance and so, on their own, processed digital images are not capable of enabling quantitative damage estimations. However, were the correlation between structural damage and image data known, it would be possible to use images as a means of generating damage estimates via supervised machine learning techniques.

The goals of this study were threefold. The primary purpose was to determine if there was in fact a strong, statistically robust relationship between bridge column performance under lateral load and visually observable column damage. The analysis was also designed to determine the limits of applicability for derived regression models with regards to other structural elements of similar size but varying design. Lastly, the work was performed to provide recommendations for camera setups of future experimental testing programs so as to maximize the utility of captured and synchronized images. The results of the study indicate that there is a robust relationship between parameterized image information and quantifiable structural performance, and that future experimental testing programs can be optimized to enable these correlations for other structural components.

6.1.1 Prior Work

There is a large body of work on using static sensors for damage detection and estimation [Wang and Chan, 2009] in structures. While finite element model updating and vibration analysis can successfully both detect and analyze certain kinds of damage, the majority of structures are not instrumented with static sensor networks, limiting the utility of such techniques for applications such as post-earthquake triage assessment.

To date there have been no studies that have explored correlations between objects in images and structure displacements. Prior efforts have focused on detecting

damage, assigning a prescribed damage level, or estimating permanent deformations from images, but not on numerically estimating the mechanical integrity of structural elements from images.

German et al. [2012] developed a method for detecting and parameterizing spalling in images. Their approach used features of the image covariance matrix which relate to spalling. These spall regions were then parameterized via template matching of exposed rebar in the images. Zhu et al. [2011] developed a technique for parameterizing concrete cracking in images. The percolation method of crack segmentation developed by Yamaguchi and Hashimoto [2010] was used to provide a crack map that was then described numerically and associated with prescribed generalized damage states. The method was used on images taken after the January 2012 earthquake in Haiti.

Digital image correlation (DIC) has been widely applied to crack detection in images [Destrebecq et al., 2011]. Via a comparison of two images, it is possible to detect and measure deformations and cracking in concrete. These techniques, however, require camera placement to be the same between images. Furthermore, to date there has been no effort in correlating DIC detected cracks to underlying structural causes.

There have been a series of research efforts using augmented reality (AR) systems to aid in damage detection post-earthquake [Kamat and El-Tawil, 2007, Dai et al., 2011]. In this work, the position of a field inspector was determined via photogrammetric analysis of a known model of the structure. This known, undamaged, model of the structure was then overlaid to the AR system to aid in determining if permanent inter-story drift had occurred. However, assessments of permanent drift do not necessarily capture the extent of structural damage due to scenarios such as the re-centering of compromised structural elements due to global structural effects.

Carreno et al. [2010] developed a technique for damage estimation using machine learning. Their system used a fuzzy artificial neural network and a wide range of

heterogenous inputs. These inputs included damage to structural elements and non-structural elements, site conditions, and issues such as material variability. The system was capable of quantifying damage, but did not link observations to the displacement history of the tested structures, and was only able to categorize structures based on prescribed damage states.

OBrien et al. [2011] also developed and tested a post-earthquake triage assessment tool. Based on broad, easy to obtain structural measurements such as building floor plan dimensions and areas of lateral load resisting elements, a vulnerability index was constructed. This index was designed to enable building inspectors to estimate the likelihood of significant structural damage without complex analysis. However, the developed index did not numerically relate observations to quantitative displacement histories.

6.1.2 Research Objectives and Approach

The research presented in this paper explores the relationship between parameterized image information and maximum experienced structural displacements. Such estimations are more closely related to the mechanical performance and integrity of structural elements relative to techniques which estimate residual displacements or associate observed damage with previously defined, and more qualitative, damage states.

For this research, concrete bridge columns with socketed footing connections were studied. Using laboratory testing specimens originally tested as part of a Pacific Earthquake Engineering Research (PEER) project [Haraldsson et al., 2012, Davis et al., 2012], known levels of maximum lateral displacement were recorded synchronously with digital images of the columns from various angles and perspectives. The images of the columns were captured after the column had returned to a zero-displacement configuration. These images were analyzed, parameterized, and then correlated with the known maximum displacements by constructing regression mod-

els via machine learning.

6.2 Methodology

In order to build a machine learning model using synchronized experimental testing and image data, images must first be reduced to vectorized numeric descriptors of relevant objects in the images themselves. Once vectorized, the complete set of descriptors for all images can be correlated to the known maximum drift levels via any number of learning methods. In this work, images were first segmented to produce sub-images of only the relevant visible damage. These segmentations were then numerically described and parameterized using non-dimensional descriptors such as normalized size and location. These first two steps are diagrammed in Figure 6.1. Finally, regression analysis was performed via a suite of supervised learning algorithms. The resulting regression models were tested for accuracy and potential overfitting.

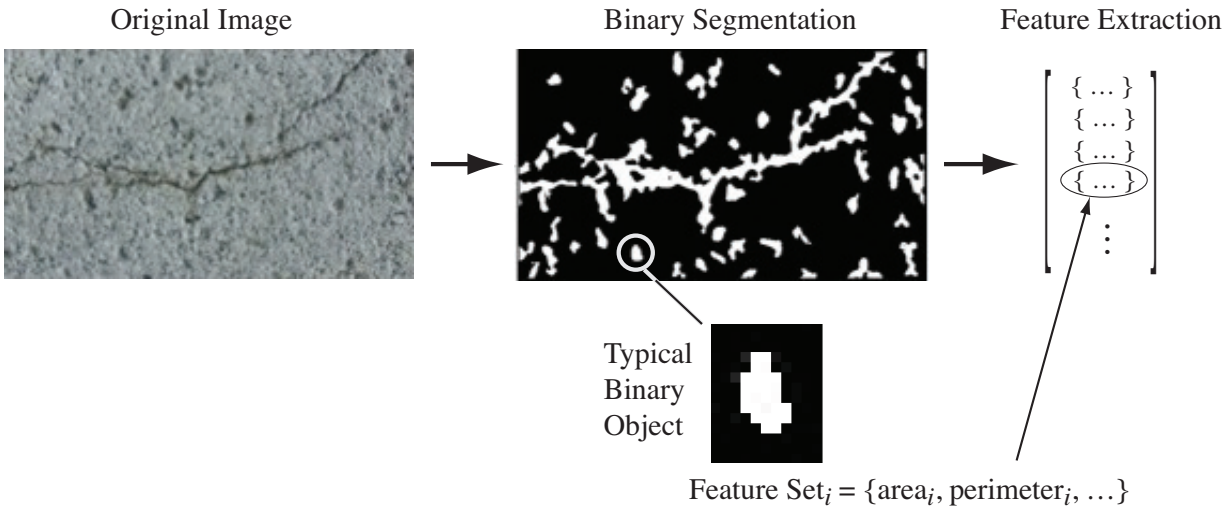


Figure 6.1: Schematic example of segmentation and feature extraction process

6.2.1 Image Segmentation

In order to parameterize the images, relevant objects within the image such as cracks or spalls were first isolated, a process known as image segmentation [Gonzalez et al., 2010]. The cracks that occurred during testing were marked using a red or blue marker. As a preliminary study, the image segmentation process was aided by these additional manual marking of cracks during laboratory observation and this distinction was used to develop a repeatable crack segmentation routine with minimal complexity. This simplification was performed in order to minimize potential variables in the study. The images were segmented twice, once for crack patterns and again for concrete spall patterns.

For the first segmentation, the RGB images were projected into the CIE-LAB color space [Fairchild, 2005]. CIE-LAB isolates the red and blue channels in an image, which helps reduce the impacts of color temperature variances. The A and B channels of the image were thresholded using simple binary thresholds to develop binary images of the cracks, which were then combined. Morphological dilation was performed to connect small discontinuities in the crack patterns and reduce variance caused by crack widths [Serra, 1982]. Connected component analysis was then performed [Gonzalez et al., 2010], discarding small objects that corresponded to image noise. The result of the segmentations were binary geometric skeletons of the crack patterns.

For the second segmentation, the initial image of the uncracked specimen was compared to each subsequent image to provide a differential image. Both images were projected into the CIE-LAB color space. Only the luminance channel, which provides a grayscale image, was used. The differential of the two images was converted into a binary image using simple thresholding. Again, morphological dilation and connected components analyses were performed. The resulting images were binary representations of only the spalls patterns.

It is important to note that this research was not a study of image segmenta-

tion methods, and the simple method used is suitable for laboratory specimens with marked crack patterns only. In the field, issues such as crack closing make segmentation more challenging. There are many methods for segmenting objects from images and it is an active field of research in the computer vision community. However, given current research it is reasonable to assume that an accurate and robust segmentation is possible in a field environment, particularly if an inspector distinctly marks observed cracks.

6.2.2 Feature Extraction

The two binary images, one of the crack patterns and one of the spall patterns, were then reduced to numeric descriptions of the objects in the images. Normalized descriptions of the images were used to reduce variances due to changes in camera perspective between test cameras. The emphasis was on performing bulk parameterization of the images, as opposed to describing individual binary objects in the images, in order to reduce sensitivity to small segmentation variances.

The parameterizations were designed to describe the formation of a plastic hinge in the tested columns, as well as to describe damage in the columns after hinging had occurred and concrete cracking had subsequently stabilized. For the segmented images of the crack patterns, the area of crack pixels in both the binary image and the luminance channel image were included. The centroid of segmented crack objects was also included. All dimensional descriptors were normalized to reduce scale variances. Normalization was performed by leveraging the known dimensions of the column and corresponding pixel size to express all parameters as a ratio of the column dimensions. For the segmented images of the spalls, three descriptors were used. The first was a binary flag indicating whether or not a spall had been detected. The other two descriptors used were the area of the spall and the centroid of the spall. Both the area and centroid descriptors were normalized to reduce scale variances.

6.2.3 Machine Learning

The parameterizations of the images were compiled into a data set that also included the known level of maximum experienced column drift corresponding to each image. This data set was then used to construct a series of regression models via machine learning. These models were designed to predict the maximum experienced drift in the column, as opposed to residual drift, using only the parameterizations of the images.

6.3 Testing and Analysis

The test setup included both the experimental testing of the columns and the numerical analysis of the synchronized images.

6.3.1 Structural Test Setup

A series of 1:2 scale structural tests were performed in which 4 columns were monotonically and laterally loaded to prescribed displacements, specified as a ratio of the column height to the lateral displacement, or percent drift. After each load cycle, the columns were brought back to a neutral "at rest" position. Images of the columns were captured in this at rest position from various camera angles, resulting in 13 camera sequences from the 4 tests (Figure 6.2). The reader is referred to Haraldsson et al. [2012] and Davis et al. [2012] for more details concerning the testing. These four tests are referred to as SF-1, SF-2, SF-3, and PT-SF in the following discussions. The orientation of the cameras during each test are illustrated in Figure 6.3 and key structural parameters for each test are shown in Table 6.1. Several differences between the tests are of note. The first is that the differences between SF-1 and SF-2 involve detailing of the confinement reinforcement, which is not indicated in the table. The second key difference is the thinner footing used during SF-3. The thinner footing led to a punching shear failure in the footing at high drifts, as opposed to the bar

buckling observed during the other tests. Lastly, while the design of the column is similar, PT-SF also included significant prestressing force.

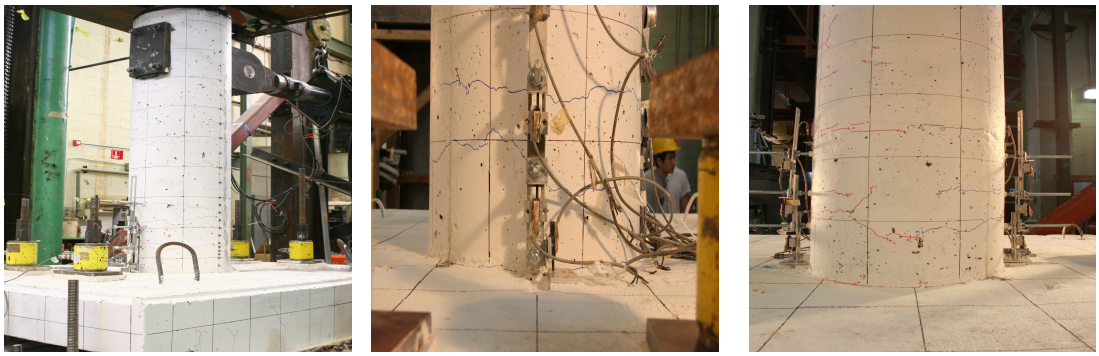


Figure 6.2: Sample images from experimental tests

6.3.2 Computer Vision Test Setup

Image segmentation and parameterization were performed in Inc. [2011], using the Image Processing Toolbox. All images were analyzed with a standardized resolution of about 50 pixels per inch. Regression analysis was performed using the WEKA machine learning software suite [Hall et al., 2009]. A series of linear and nonlinear regression models were constructed. M5P regression trees [Quinlan, 1992] provided the highest degree of accuracy, as is discussed in Section 6.4.1, and were used throughout the analyses unless noted otherwise.

6.3.3 Analysis Procedure

As a first step in the analysis of the image-based regression models, a study was performed in order to assess the need for a nonlinear regression model, as opposed to simple linear regression. A data set of SF-1, SF-2, and SF-3 was used for training and testing via 10-fold cross-validation [Kohavi, 1995].

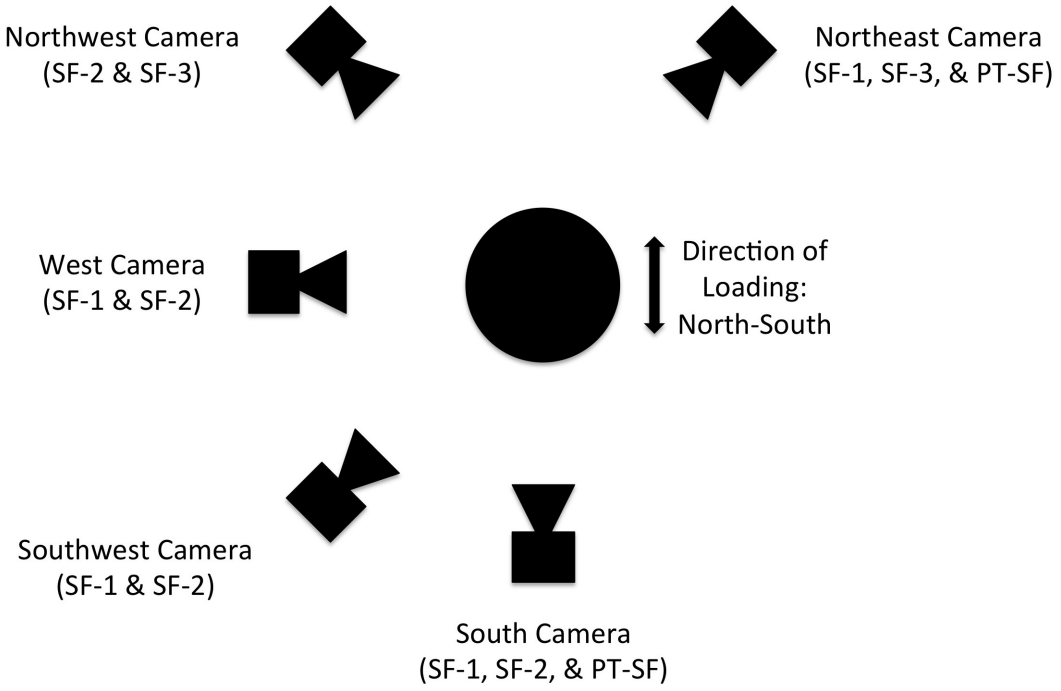


Figure 6.3: Orientation of cameras during column testing

Once a suitable regression model was chosen, the next step was to determine the need for various image descriptors. In particular, the necessity of modeling visually observable spalls was studied. For this study, each of the data sets from SF-1, SF-2, and SF-3 were trained and tested separately via 10-fold cross-validation, both with and without spall descriptors. Additional preliminary tests were performed to determine the need for the other included parameters.

In addition to standard regression analyses, it was vital that the constructed data set be tested for overfitting. Overfitting refers to the construction of statistical models that exhibit a high degree of prediction accuracy on a limited data set, but perform poorly on more varied image sets [Witten and Frank, 2005]. While there is no guaranteed way to prevent or detect overfitting, there are several commonly used techniques generally considered to be effective in minimizing the likelihood of

Specimen	Column Diameter (in)	Column Height (in)	Footing Depth (in)	Reinforcement Ratio (%)	Axial Load (kip)
SF-1	20	60	22.5	1.12%	159
SF-2	20	60	22.5	1.12%	159
SF-3	20	60	10	1.12%	159
PT-SF	20	60	22.5	1.0%[a]	159 (+PS)

[a] Total reinforcement ratio. Mild steel ratio was 0.36%

Table 6.1: Details of experimental tests

overfit. The most common is to construct the regression model using one data set and test the model's accuracy on data from an entirely different scenario. Such tests illustrate the robustness of a regression model with regards to new information. For the purposes of this research, regression models trained on data from SF-1 were tested on data from SF-2 and SF-3. A model trained on test SF-2 was tested on SF-1 and SF-3 as well. The differences in camera positioning, lighting, and column detailing, as well as variances in column fabrication and material behavior, provided significant differentiation between tests.

After selecting a descriptor set and classifier, as well as checking against model overfitting, regression models built using the complete data set were trained and tested via 10-fold cross-validation. The accuracy of this complete data set model was examined for each individual column test in regards to predicting maximum experienced drift based on image data. The impact of camera variables such as camera position and lighting were explored by analyzing the regression models' accuracy on a per-camera basis.

Lastly, the PT-SF column, which included prestressing, was tested using the re-

gression model built from the other column tests. This was done in order to explore whether or not the visual model could be extended to socket columns with a significantly different internal structural design. The expectation was that the differences in concrete cracking due to additional prestressing would reduce the accuracy of the regression model.

Performance Metrics

The metrics associated with analyses of the the type described above are related to the performance of the regression models. Key regression accuracy metrics include the correlation coefficient between the predicted and known drift levels, the mean absolute error, and the root mean squared error. The correlation coefficient provides a relative measure of model fit, or the scatter in the results. The other metrics provide an estimate of the overall accuracy of a regression model [Witten and Frank, 2005].

Scatter plots which indicate the correlation between the true and predicated maximum drifts are provided for all analyses in the next section. Each data point in the scatter plots represents the predicted drift for one test image and displacement. A perfect prediction would fall on a 45 degree line in the plots. Such a line is included in all plots as a visual aid, labeled as the "ideal trend" line.

6.4 Results

6.4.1 Regression Model Selection

Regression models were built using a series of linear and nonlinear algorithms using a data set compiled from SF-1, SF-2, & SF-3. Initial testing indicated that M5P regression trees provided the most accurate and reliable results. A comparison between the performance of linear regression and M5P is shown in Figures 6.4 and 6.5 and Table 6.2. Both methods indicated a correlation between parameterized image information and structural performance, however the M5P algorithm allowed for an arbitrarily

	R	Mean Abs. Err.	RMS	Rel. Err.	Root Rel. Sq. Err.
Linear	0.91	0.93	1.23	40.0%	42.3%
Linear (low drift)	**0.73**	0.73	0.91	**45.4%**	52.1%
M5P	0.94	0.66	0.99	28.5%	33.0%
M5P (low drift)	**0.90**	0.35	0.49	**21.6%**	27.6%

Table 6.2: Results for linear regression and M5P tree models (complete data set)

complex nonlinear model capable of better results, particularly at low drifts where spalling is not observable. The linear regression model optimizes overall accuracy by heavily weighting spall parameters that are unavailable at lower drifts. The result is a model that predicts an almost constant maximum drift if spalling is not observed.

6.4.2 Spall Modeling Analysis

Preliminary models only described crack patterns in images and neglected spall information entirely. However those models did not perform well at higher drift levels (>4%) due to the stabilization of crack patterns after the formation of a plastic hinge in the columns. A comparison of regression models both with and without spall descriptors are shown in Figures 6.6, 6.7, & 6.8 and Table 6.3. In all three cases, the inclusion of spall information both reduces scatter in the data and improves the accuracy of the overall trend. The improvement is most noticeable at higher drift levels, as expected. The outliers for SF-1 at 7% drift (Figure 6.6) are due to a disturbance of the test specimen that created a false spall pattern high on the column, leading to an overestimated drift prediction. The noticeable outliers for SF-3 (Figure 6.8) can be attributed to the ultimate failure mechanism of the column. For SF-3, the footing thickness was significantly less than for the other tests and resulted in a punching shear failure. This failure caused the entire column to shift downwards

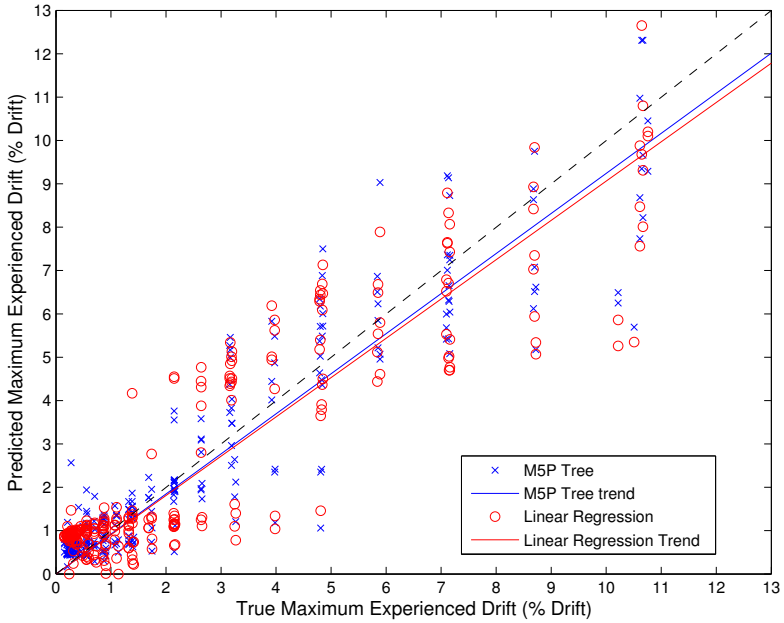


Figure 6.4: Comparison of M5P tree and linear regression based models using complete training set

substantially, lowering the centroid of both the cracking and spalling patterns and causing an under-estimation of the maximum drift.

6.4.3 Overfitting Analysis

The results of the overfitting analyses are shown in Figures 6.10, 6.12, & 6.11 and Table 6.4. For models tested and trained on SF-1 and SF-2, error increases, but the correlations remain strong and indicate that overfitting is unlikely. Performance of models tested on SF-3 perform worse, predominantly under-predicting the maximum drift that the column had experienced.

This difference was due to relatively severe lighting changes during the SF-3 test (Figure 6.9). Light sources for cameras during tests SF-1 and SF-2 remained much more consistent throughout testing. The variance inhibited the consistent segmen-

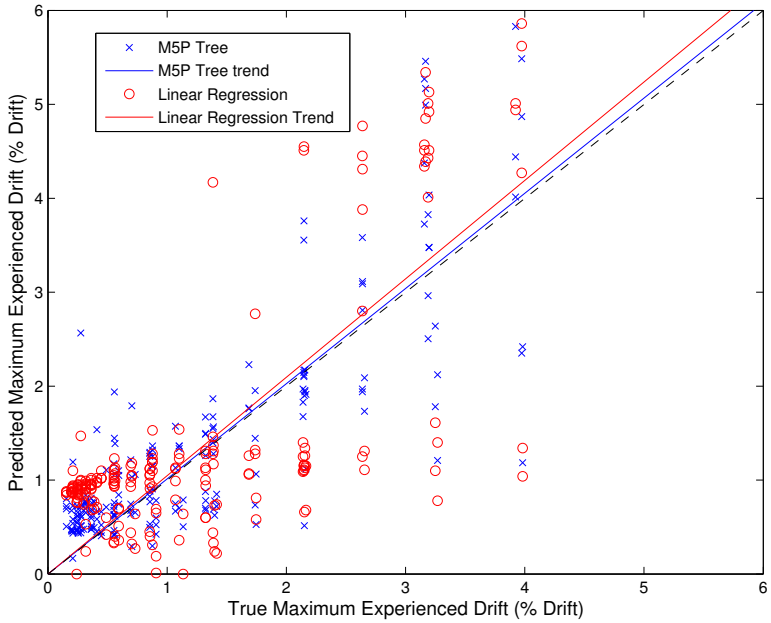


Figure 6.5: Comparison of M5P tree and linear regression based models at low drift level

tation of cracking in images from SF-3 and the resulting crack parameterizations indicated less cracking in the column than was actually present, leading to the underestimation of column drifts, despite a correlation coefficient above 0.8 (Table 6.4).

6.4.4 Complete Model Analysis

After initial model verification, the accuracy of the regression model built using the complete data set was tested. The model was constructed and tested using 10-fold cross-validation to minimize overfit. The results for the complete model, as well as the results broken out for each individual training set, are shown in Figures 6.14, 6.15, 6.16, & 6.17 and Table 6.5. Prediction accuracy is relatively consistent for SF-1 and SF-2. Accuracy is lower for predictions on data from SF-3, for the reasons discussed earlier.

	R	Mean Abs. Err.	RMS	Rel. Err.	Root Rel. Sq. Err.
SF-1 with spall	0.97	0.47	0.69	19.7%	22.9%
SF-1 no spall	0.91	0.79	1.26	32.9%	41.9%
SF-2 with spall	0.98	0.40	0.58	18.0%	20.5%
SF-2 no spall	0.85	0.83	1.49	37.3%	52.5%
SF-3 with spall	0.90	0.75	1.28	30.8%	43.0%
SF-3 no spall	0.74	1.45	1.96	60.1%	65.7%

Table 6.3: Results for M5P tree regression models considering spall descriptors

	R	Mean Abs. Err.	RMS	Rel. Err.	Root Rel. Sq. Err.
SF-1 on SF-2	0.89	0.89	1.40	39.7%	50.0%
SF-1 on SF-3	0.83	1.75	2.46	75.9%	83.8%
SF-2 on SF-1	0.92	0.95	1.50	40.9%	50.1%
SF-2 on SF-3	0.81	2.46	3.12	108.0%	105.7%

Table 6.4: Results of overfit analysis

	R	Mean Abs. Err.	RMS	Rel. Err.	Root Rel. Sq. Err.
Complete Data	0.93	0.75	1.08	32.4%	36.9%
SF-1 breakout	0.98	0.49	0.66	20.7%	22.2%
SF-2 breakout	0.96	0.56	0.87	24.9%	30.9%
SF-3 breakout	0.93	0.95	1.40	41.2%	47.8%

Table 6.5: Performance of complete data set (includes breakouts for individual column tests)

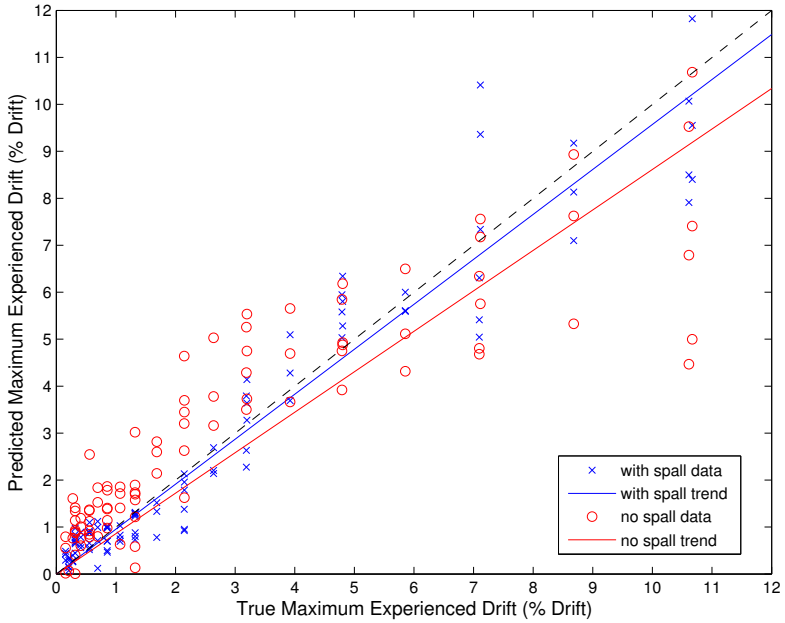


Figure 6.6: Performance of SF-1 trained model with and without spall information

6.4.5 Camera Analysis

It is instructive to study the performance of the complete regression model on a per camera basis, and these results are shown in Figures 6.18, 6.19, & 6.20 and Table 6.6. The prediction accuracy for SF-1 and SF-2 data sets remains high *regardless of camera orientation*. Prediction accuracy results from SF-3 indicate that models built using the northwest camera provided less accurate predictions than the northeast camera. This was due to the problematic lighting variances experienced by the northwest camera.

6.4.6 Pretensioned Column Analysis

As a final test of the limits of the model, the complete regression model constructed from SF-1, SF-2, and SF-3 was tested on the images from pretensioned socket column PT-CB (Figure 6.21 and Table 6.7). Both error and correlation outcomes are

	R	Mean Abs. Err.	RMS	Rel. Err.	Root Rel. Sq. Err.
SF-1 Northeast	0.98	0.46	0.57	19.0%	18.2%
SF-1 South	0.99	0.49	0.74	20.1%	23.7%
SF-1 Southwest	0.97	0.60	0.76	24.5%	24.2%
SF-1 West	0.42	0.34	0.44	17.0%	21.5%
SF-2 Northwest	0.98	0.55	0.76	22.8%	24.2%
SF-2 South	0.98	0.74	1.07	30.3%	34.3%
SF-2 Southwest	0.93	0.57	0.95	28.3%	39.2%
SF-2 West	0.97	0.20	0.22	10.0%	10.9%
SF-3 Northeast	0.93	0.81	1.11	36.9%	40.2%
SF-3 Northwest	0.96	1.08	1.63	45.1%	52.9%

Table 6.6: Performance of complete data set per camera

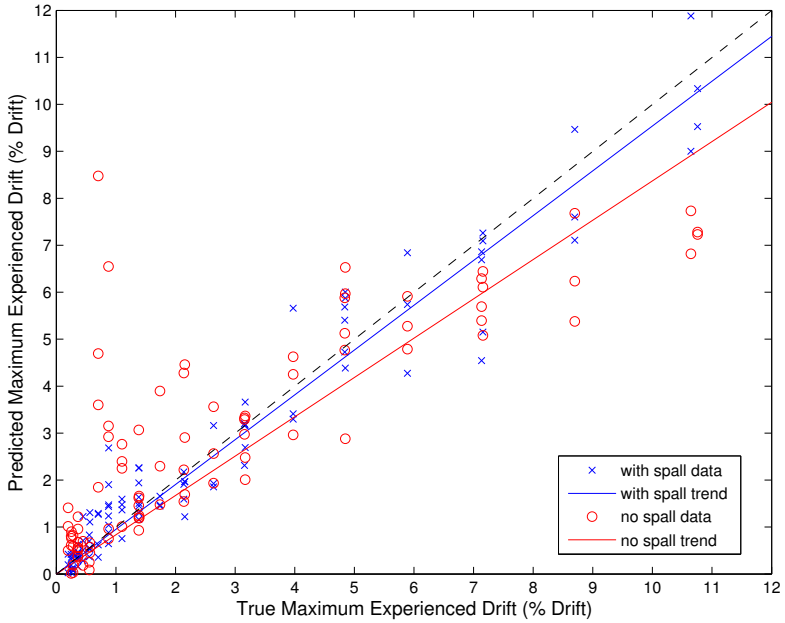


Figure 6.7: Performance of SF-2 trained model with and without spall information

significantly worse for this test than for any other analysis performed in this work. This is an expected result, as the pretensioned column exhibited an entirely different cracking pattern and failure mechanism.

The inability of the regression model to predict the maximum displacement of the pretensioned column does not negate the value of the overall approach. Instead, it serves as an example as to the limitations of applying any single regression model to a wide variety of structural elements. In a fully realized implementation of this computer vision approach, regression models for various structural elements would be necessary, as well as models which analyzed cracking patterns from other potential loading patterns, using a hierarchical pattern recognition methodology.

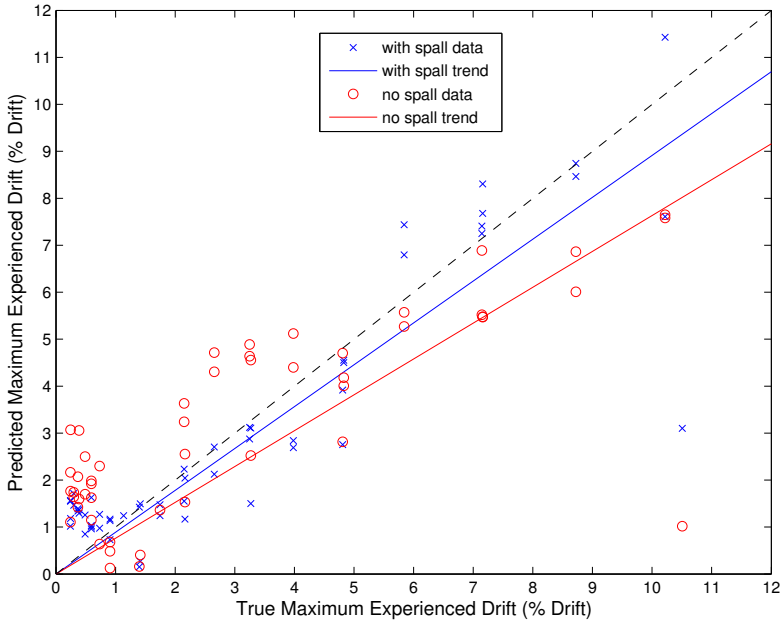


Figure 6.8: Performance of SF-3 trained model with and without spall information

	R	Mean Abs. Err.	RMS	Rel. Abs. Err.	Root Rel. Sq. Err.
PT-SF	0.34	2.83	4.37	107.3%	130.9%

Table 6.7: Performance of regression model on pretensioned column

6.5 Conclusions

A series of regression models capable of predicting the maximum experienced lateral displacement of a socket column have been developed based solely on numeric information extracted from images of the bridge columns via computer vision techniques. Analysis of the regression models indicates that there is a strong correlation between parameterized image data and structure displacement histories for the tested bridge columns. This correlation is robust to changes in camera position and column detail-

Figure 6.9: Example of lighting variance during test SF-3

ing, but severe lighting variances or significant structural changes, such as additional column pretensioning, inhibit model prediction accuracy.

Changes in lighting during the test were determined to be a significant source of segmentation error with respect to the images of test SF-3. In order to maximize the utility of images taken during future experiments, care should be taken to provide consistent lighting throughout testing. Furthermore, analysis of experimental images can be optimized if cracks are marked using distinct colors such as was done for the tests used herein. A marking scheme which distinguishes between cracks observable in the "at-rest" condition from cracks only observable at maximum displacement would also aid further model development and calibration.

Given the correlation and predictive accuracy of the models in this study, image-based regression models have the potential to be used as triage assessment tools in order to better determine the mechanical integrity of bridge columns post-earthquake. However, the limitations of the constructed models with respect to the pre-tensioned column highlight the substantial work that must be completed before such assessment

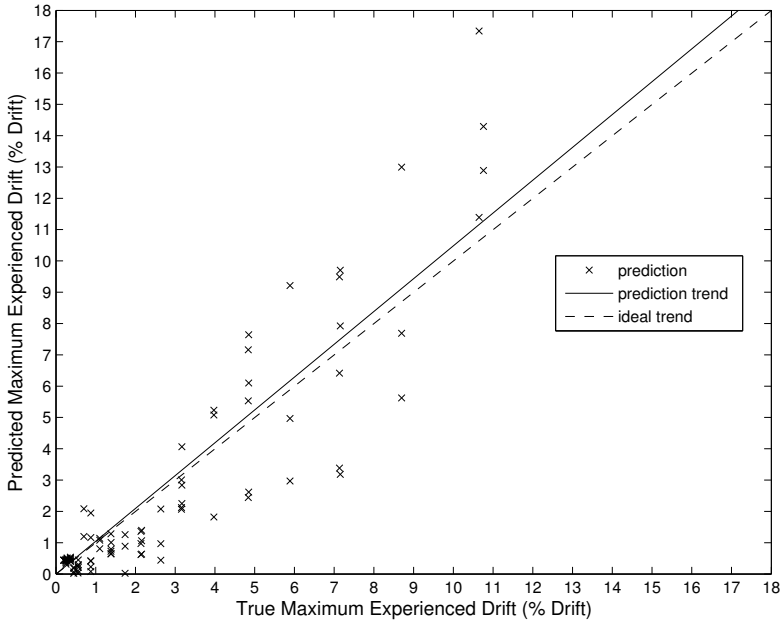


Figure 6.10: Overfit check: correlation for training on SF-1 and tested on SF-2

tools can be realized. In order to achieve a fully comprehensive field implementation, image-based regression models must be developed for other column shapes and designs, as well as for other loading patterns such as torsional loadings. There are many other practical issues to be addressed as well. Ultimately, a suite of regression models should be constructed, as well as a computer vision methodology for discriminating between potential damage patterns a column could experience. Both the number of models necessary to construct a comprehensive assessment tools and the limits of applicability or any one model are open research questions.

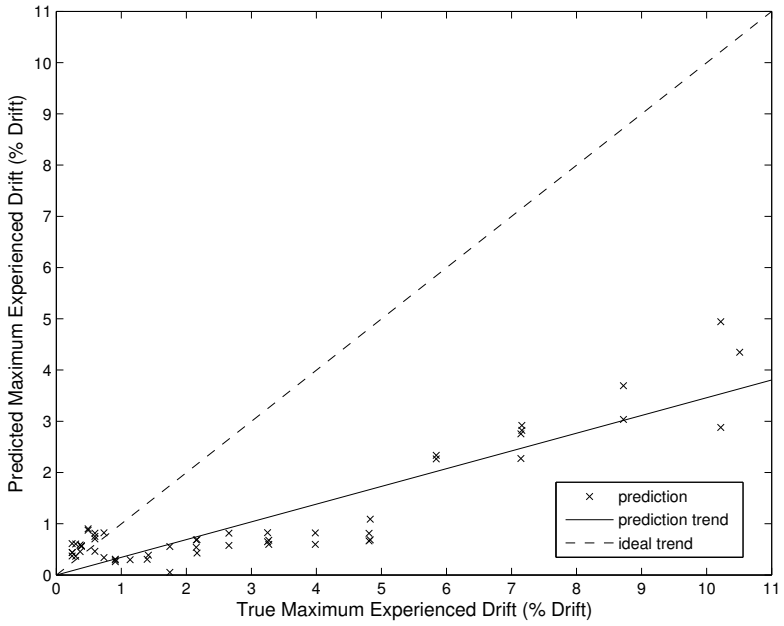


Figure 6.11: Overfit check: correlation for training on SF-1 and tested on SF-3

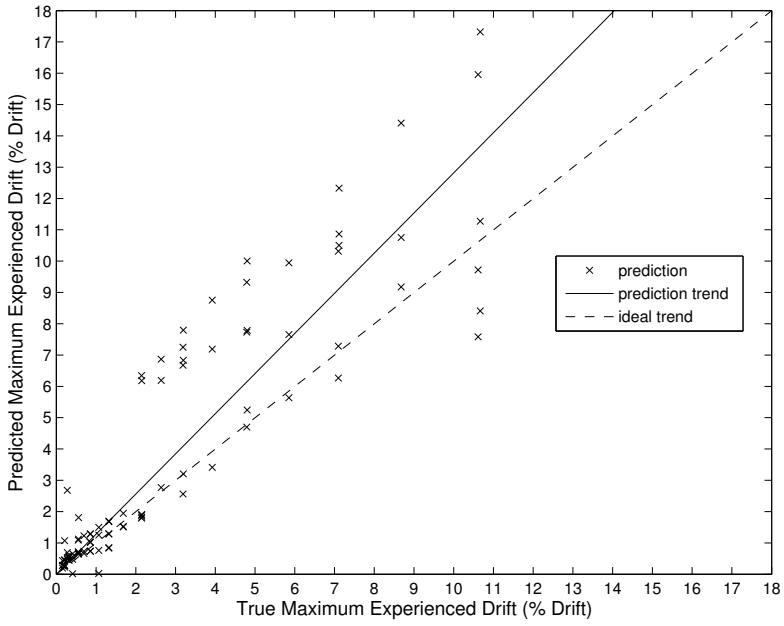


Figure 6.12: Overfit check: correlation for training on SF-2 and tested on SF-1

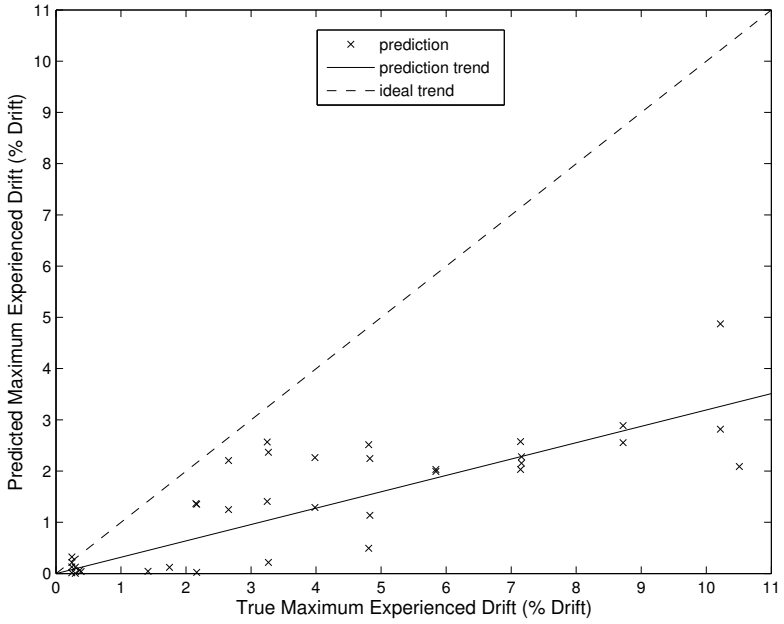


Figure 6.13: Overfit check: correlation for training on SF-2 and tested on SF-3

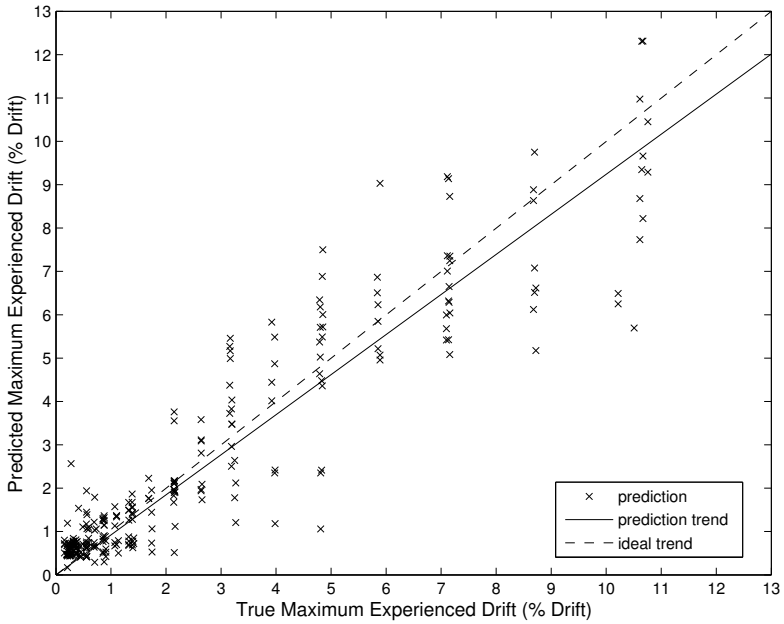


Figure 6.14: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation)

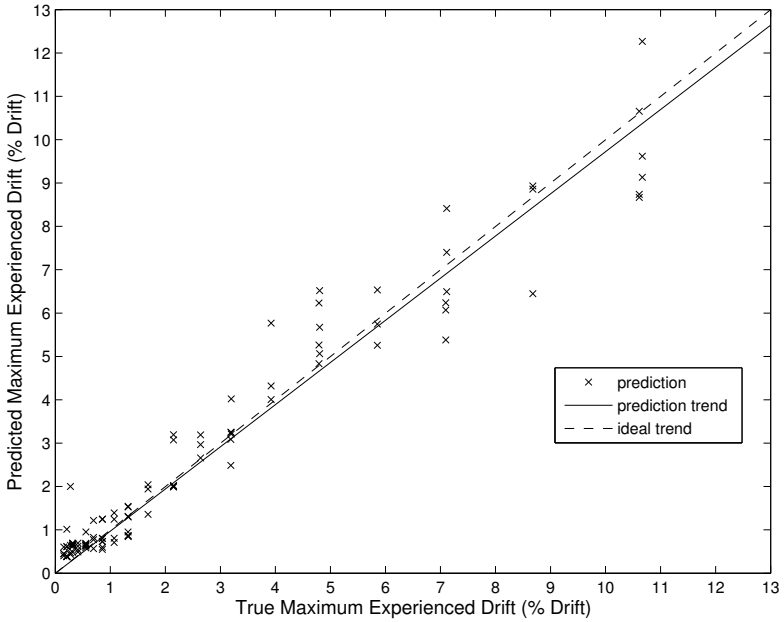


Figure 6.15: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation): SF-1 breakout

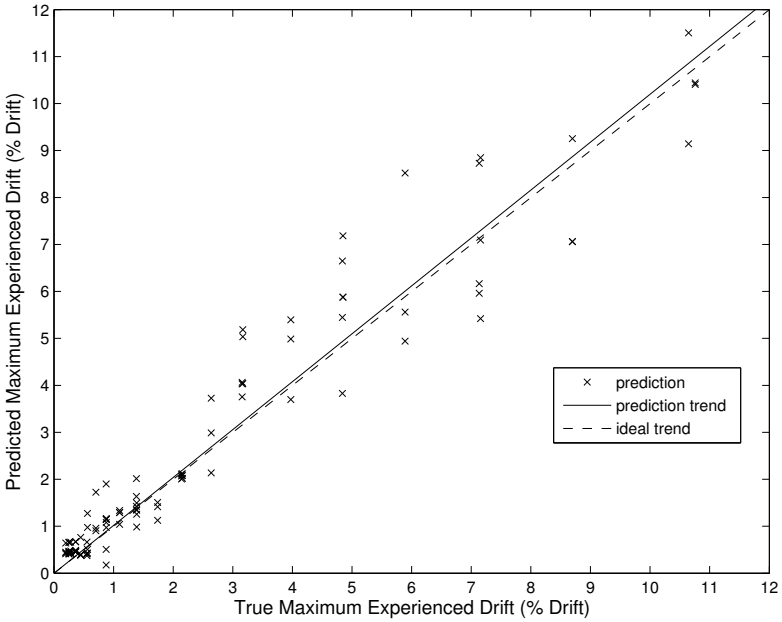


Figure 6.16: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation): SF-2 breakout

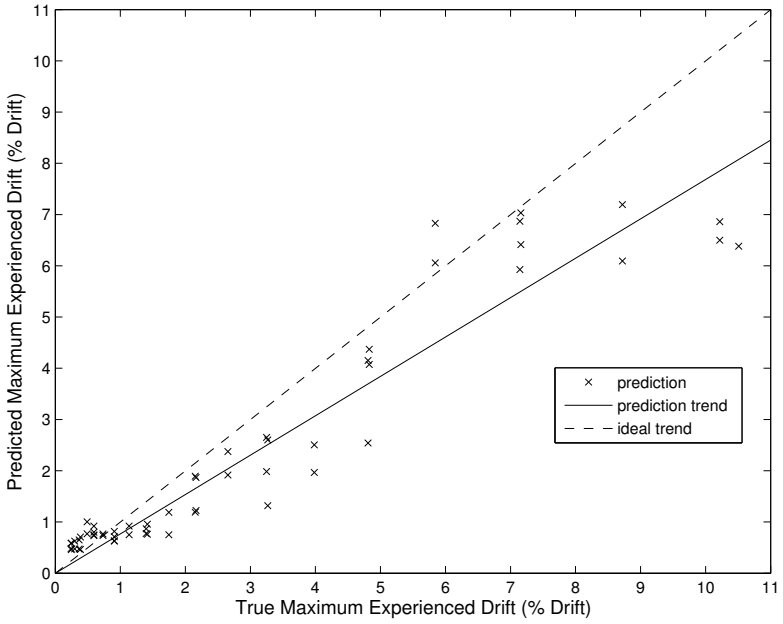


Figure 6.17: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation): SF-3 breakout

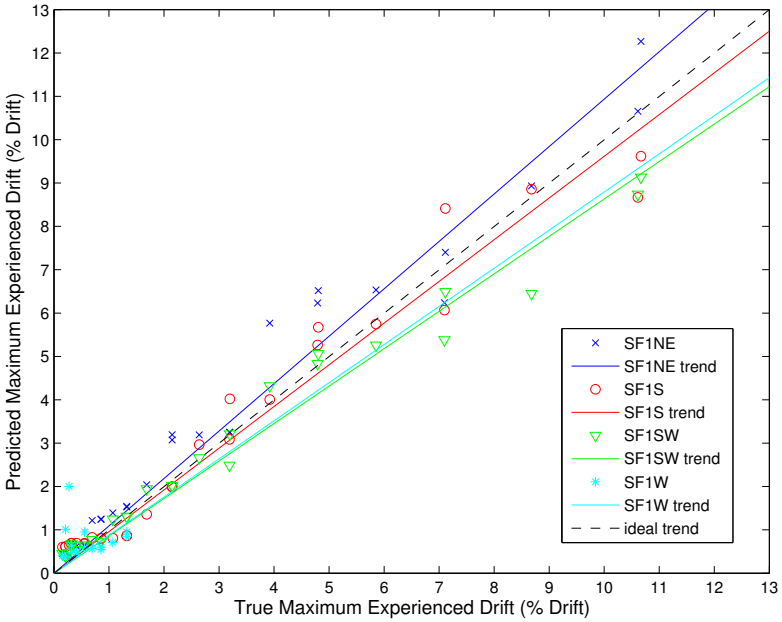


Figure 6.18: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation): SF-1 breakout by camera

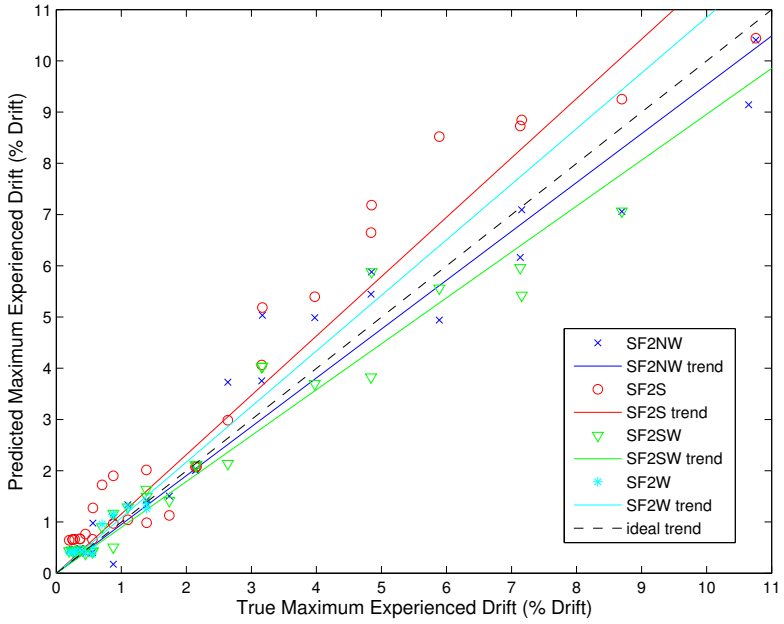


Figure 6.19: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation): SF-2 breakout by camera

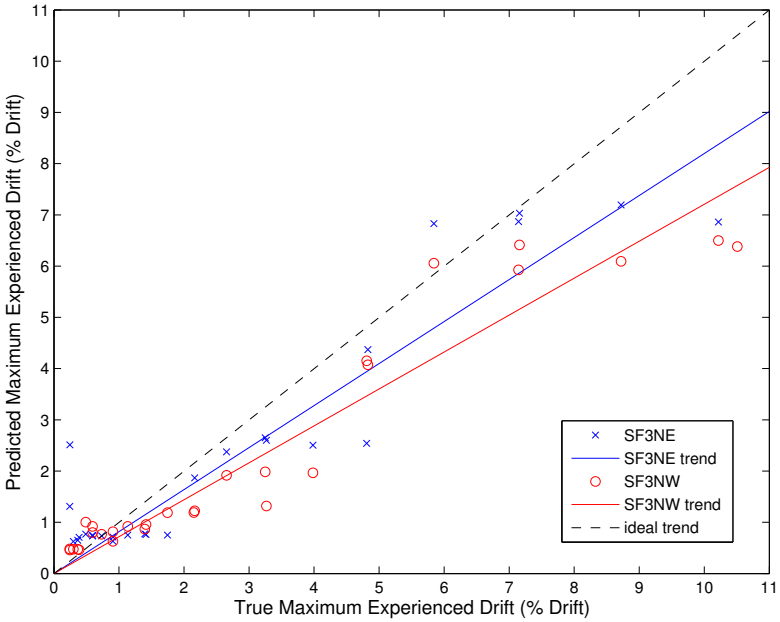


Figure 6.20: Correlation for training and testing on SF-1, SF-2, & SF-3 (10-fold cross-validation): SF-3 breakout by camera

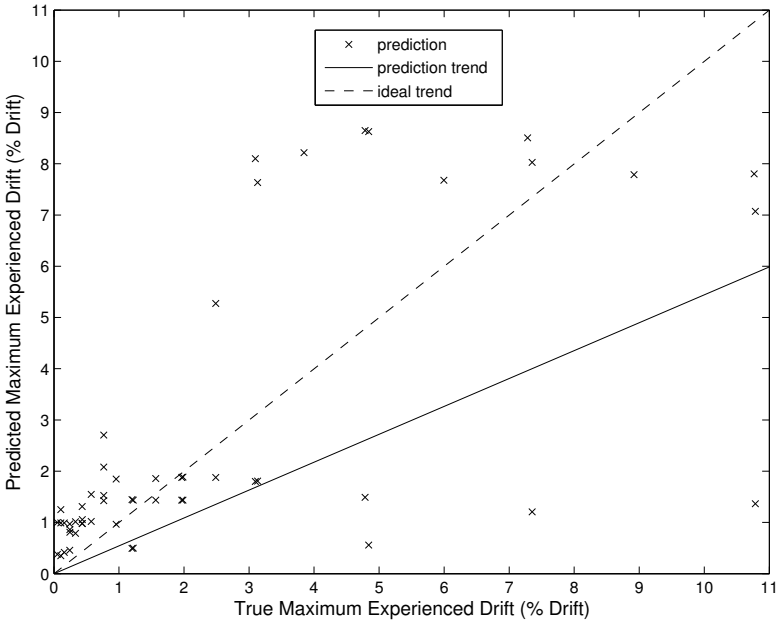


Figure 6.21: Correlations for training on SF-1, SF-2, & SF3, testing on PT-SF

Chapter 7

CONCLUSIONS & AVENUES FOR FUTURE RESEARCH

7.1 Conclusion

A new, systematic, methodology for capturing and analyzing inspection images has been presented and prototyped. This system processes images captured via unmanned aerial drones capable of comprehensive imaging. These images are reconstructed into an accurate and scalable 3D geometric model of the inspected structure, providing spatial context to the high-resolution 2D images. Structural damage in these 2D images is then detected via field-robust image segmentation and trained machine learning pattern classifiers. The detected damage in the images is then parameterized for input into nonlinear regression models that estimate mechanical damage.

This research program strove to prototype the system, addressing key challenges at each phase to enable the application of each process within the larger framework. However, the primary goal of this research has not been the optimization of any one process. Instead, the goal has been to show the feasibility of this systematic approach to inspection and to illustrate the possibilities for new damage models and inspection techniques that are a direct result of this approach. Specifically, by incorporating a holistic methodology to inspection imaging, the following are achieved:

- Minimization of the costs and impacts of inspection through automation
- A contextualized environment that allows for consistent transmission of inspection data
- Improvements in how inspectors and engineers interact with inspection data
- New, quantitative models of the deterioration of structures

The various required processes in this systematized inspection framework are each complex and challenging research fields in their own right and potential for future research exists in all phases of the imaging pipeline, extending it and expanding it to incorporate information on a broader scale.

7.2 Future work in inspection robotics

Based on an analysis of the robots currently used for structural inspection, there is an immediate need to develop robots capable of performing inspections more autonomously in a field environment. Most implementations, including the one used for this research, require human piloting. Such ROVs are more capable of avoiding obstacles and are quicker to implement than autonomous robots. However, they provide no guarantee as to inspection coverage and tend to be highly inefficient with respect to inspection speed. One fundamental reason that autonomy has not be implemented for structural inspections is that the problem of mapping and exploring an unknown and dynamically changing field environment, often referred to as the Simultaneous Localization and Mapping (SLAM) problem, is difficult for lightweight flying robots. These UAVs are highly susceptible to environmental disturbances such as wind gusts, and inherent weight restrictions inhibit the use of on-board computers capable of on-line pose recalibration.

The 3D reconstructions that are a result of the systematic inspection methodology provide a potential solution. These 3D models include robust salient image features, which are used to produce the reconstructions. These features can also be used as "waypoints", potentially enabling a robot to localize itself in the environment. Furthermore, such a known underlying 3D model can facilitate autonomous path planning that guarantees comprehensive structural inspection. The challenges in both case are developing the initial model of the structure and utilizing the information in a manner that allows for on-line robotic decision making.

7.3 Future work in 3D scene reconstructions

There are two primary avenues for future work in 3D scene reconstruction, in the context of structural inspection. The first is the development of methods that provide 3D reconstructions using underwater imaging and sensing modalities. This has

been done in the simpler context of mapping the ocean floor, but more sophisticated 3D reconstructions have not been performed. In conjunction, how to connect 3D reconstructions that utilize different sensing modalities is an open research question.

More specific to the field of structural inspection, there is a need for work that relates 3D reconstructions to the deformations and mechanics of structures. Often, structural damage can only be perceived as a deformation of the structure (e.g. cross-frame buckling). Repeated imaging and reconstruction of a structure can capture these changes, but in order to relate changes to deformations and potentially to strain fields within the structure requires utilizing the reconstruction information in new ways. Relatedly, there is potential to use reconstruction point clouds to perform finite element model updating.

7.4 Future work in damage detection

The majority of potential future research lies in extensions of the developed computer vision methodologies. There is a need for research into damage detection and parameterization methods that can detect and numerically describe other forms of damage such as corrosion and concrete efflorescence. Damage and defect detection could also be improved by developing computer vision techniques that not only detect defects, but can match them against a database of defects in similar structures to provide a more complete picture of the causes and impacts of damage.

7.5 Future work in image-based damage estimation

The regression models developed herein are only applicable within the narrow range of socket columns subjected to a uniaxial lateral load. There is substantial need to build similar regression models for other structural elements and loading patterns. Furthermore, techniques for determining the limits of applicability for developed regression models are needed. Lastly, there is potential to extend these regression models by incorporating additional numerical information, such as is provided by accelerome-

ters and other global response sensors. If achieved, such an image-sensor fusion could address long standing challenges in the structural health monitoring community with regards to combining localized damage descriptions with changes in the global structural response.

BIBLIOGRAPHY

I. Abdel-Qader, O. Abudayyeh, and M. Kelly. Analysis of edge-detection techniques for crack identification in bridges. *Journal of Computing in Civil Engineering*, 17:255, 2003.

I. Abdel-Qader, S. Pashaie-Rad, O. Abudayyeh, and S. Yehia. Pca-based algorithm for unsupervised bridge crack detection. *Advances in Engineering Software*, 37(12):771–778, 2006.

T. Acharya and A. K. Ray. *Image processing: Principles and applications*. Wiley-Interscience, Hoboken NJ, 2005.

Adobe. Lightroom version 4.4, 2013.

David W. Aha, Dennis Kibler, and Marc K. Albert. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.

A. Akbarzadeh, J. M. Frahm, P. Mordohai, B. Clipp, C. Engels, D. Gallup, P. Merrell, M. Phelps, S. Sinha, and B. Talton. Towards urban 3d reconstruction from video. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on*, pages 1–8. IEEE, 2006.

M. Arai, Y. Tanaka, S. Hirose, H. Kuwahara, and S. Tsukui. Development of souryu-iv and souryu-v: Serially connected crawler vehicles for in-rubble searching operations. *Journal of Field Robotics*, 25(1-2):31–65, 2008.

Autodesk. 123d-catch, 2013.

C. Balaguer, A. Gimenez, and A. Jardon. Climbing robots mobility for inspection and maintenance of 3d complex environments. *Autonomous Robots*, 18(2): 157–169, 2005.

H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. *Computer Vision-ECCV 2006*, pages 404–417, 2006.

Ankit Bhatla, Soo Young Choe, Oscar Fierro, and Fernanda Leite. Evaluation of accuracy of as-built 3d modeling from photos taken by handheld digital cameras. *Automation in Construction*, 28(0):116 – 127, 2012.

C. M. Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.

Avrim L Blum and Pat Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1-2):245 – 271, 1997.

G. Bradski. The opencv library. *Dr.Dobb's Journal of Software Tools*, 2000.

Ioannis Brilakis, Habib Fathi, and Abbas Rashidi. Progressive 3d reconstruction of infrastructure with videogrammetry. *Automation in Construction*, 20(7):884 – 895, 2011.

Pierre-Jean Bristeau, Francois Callou, David Vissiare, and Nicolas Petit. The navigation and control technology inside the ar. drone micro uav. In *World Congress*, volume 18, pages 1477–1484, 2011.

JMW Brownjohn. Structural health monitoring of civil infrastructure. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):589, 2007.

F. Caballero, L. Merino, J. Ferruz, and A. Ollero. A visual odometer without 3d reconstruction for aerial vehicles. applications to building inspection. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 4673–4678. Ieee, 2005.

J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 8(6):679–698, 1986.

Martha L. Carreno, Omar D. Cardona, and Alex H. Barbat. Computational tool for post-earthquake evaluation of damage in buildings. *Earthquake Spectra*, 26(1):63–86, 2010.

M. Chae and D. Abraham. Neuro-fuzzy approaches for sanitary sewer pipeline condition assessment. *Journal of Computing in Civil Engineering*, 15(1):4–14, 2001.

MJ Chae. *Automated interpretation and assessment of sewer pipeline*. Purdue University, 2001.

Andrew D Chambers, Supreeth Achar, Stephen T. Nuske, Joern Rehder, Bernd Manfred Kitt, Lyle J. Chamberlain, Justin Haines, Sebastian Scherer, and Sanjiv Singh. Perception for a river mapping robot. In *Proceedings of the*

2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '11), September" 2011.

B. Chanda, BB Chaudhuri, and D. Dutta Majumder. On image enhancement and threshold selection using the graylevel co-occurence matrix. *Pattern Recognition Letters*, 3(4):243–251, 1985.

BB Chaudhuri and N. Sarkar. Texture segmentation using fractal dimension. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(1):72–77, 1995.

H. Choset. Bridge inspection with serpentine robots. Technical Report For Highway IDEA Project 56, Transportation Research Board, 2000.

P. Cignoni, M. Corsini, and G. Ranzuglia. Meshlab: an open-source 3d mesh processing system. *Ercim news*, 73:45–46, 2008.

P. J. S. Cruz and R. Salgado. Performance of vibration-based damage detection methods in bridges. *Computer-Aided Civil and Infrastructure Engineering*, 24 (1):62–79, 2009.

Fei Dai, Suyang Dong, Vineet R. Kamat, and Ming Lu. Photogrammetry assisted measurement of interstory drift for rapid post-disaster building damage reconnaissance. *Journal of Nondestructive Evaluation*, 30(3):201–212, 2011.

P. M. Davis, T. M. Janes, M. O. Eberhard, and J. F. Stanton. Unbonded pre-tensioned columns for bridges in seismic regions. *PEER*, 04, 2012.

J. F Destrebecq, E. Toussaint, and E. Ferrier. Analysis of cracks and deformations in a full scale reinforced concrete beam using a digital image correlation technique. *Experimental Mechanics*, 51(6):879–890, 2011.

J. E. DeVault. Robotic system for underwater inspection of bridge piers. *IEEE Instrumentation & Measurement Magazine*, 3(3):32–37, 2000.

R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*, volume 2. Wiley New York, 2001.

H. Elbehairy. *Bridge management system with integrated life cycle cost optimization*. PhD thesis, University of Waterloo, 2007.

Mark D. Fairchild. *Color appearance models*. Wiley, 2005.

C. R. Farrar and N. A. J. Lieven. Damage prognosis: the future of structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):623, 2007.

S. D. Fassois and J. S. Sakellariou. Time-series methods for fault detection and identification in vibrating structures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851): 411, 2007.

W. Fischer, F. Tache, and R. Siegwart. Inspection system for very thin and fragile surfaces, based on a pair of wall climbing robots with magnetic wheels. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 1216–1221. IEEE, 2007.

Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. In *Machine Learning-International Workshop Then Conference*, pages 148–156. Citeseer, 1996.

M. I. Friswell. Damage identification using inverse methods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):393, 2007.

C. P. Fritzen and P. Kraemer. Self-diagnosis of smart structures based on dynamical properties. *Mechanical Systems and Signal Processing*, 23(6):1830–1845, 2009.

C. Früh, R. Sammon, and A. Zakhor. Automated texture mapping of 3d city models with oblique aerial imagery. In *3D Data Processing, Visualization and Transmission, Proceedings. 2nd International Symposium on*, pages 396–403. IEEE, 2004.

Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(8):1362–1376, 2010.

Stephanie German, Ioannis Brilakis, and Reginald DesRoches. Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments. *Advanced Engineering Informatics*, 2012.

A. Ghorbanpoor, R. Borchelt, M. Edwards, and E. A. Salam. Magnetic-based nde of prestressed and post-tensioned concrete members-the mfl system. Technical report, Federal Highway Administration, 2000.

Mani Golparvar-Fard, Jeffrey Bohn, Jochen Teizer, Silvio Savarese, and Feniosky Pea-Mora. Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in Construction*, 20(8):1143 – 1155, 2011.

R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital image processing using MATLAB*. Pearson, 2010.

W. Guo, L. Soibelman, and JH Garrett Jr. Visual pattern recognition supporting defect reporting and condition assessment of wastewater collection systems. *Journal of Computing in Civil Engineering*, 23:160, 2009a.

W. Guo, L. Soibelman, and JH Garrett Jr. Automated defect detection for sewer pipeline inspection and condition assessment. *Automation in Construction*, 18 (5):587–596, 2009b.

M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

O. S. Haraldsson, T. M. Janes, M. O. Eberhard, and J. F. Stanton. Seismic resistance of socket connection between footing and precast column. *Journal of Bridge Engineering*, 10(1061), 2012.

J. A. Hartigan and M. A. Wong. Algorithm as136: A k-means clustering algorithm. *Journal of the Royal Statistical Society.Series C (Applied Statistics)*, 28 (1):100–108, 1979.

R. A. Hartle and Federal Highway Administration. *Bridge Inspector's Reference Manual*. United States Department of Transportation, 2004.

R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000.

D. Heckerman, D. Geiger, and D. M. Chickering. Learning bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20(3): 197–243, 1995.

R. C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11(1):63–90, 1993.

A. S. Huang, A. Bachrach, P. Henry, M. Krainin, D. Maturana, D. Fox, and N. Roy. Visual odometry and mapping for autonomous flight using an rgb-d camera. In *Int. Symposium on Robotics Research (ISRR),(Flagstaff, Arizona, USA)*, 2011.

D. Huston, B. Esser, J. Miller, and X. Wang. Robotic and mobile sensor systems for structural health monitoring. In *Proceedings of the 4th International Workshop on Structural Health Monitoring, Stanford, CA, September*, page 1517, 2003.

T. C. Hutchinson and Z. Q. Chen. Improved image analysis for evaluating concrete damage. *Journal of Computing in Civil Engineering*, 20(3):210–216, 2006.

D. Inaudi. Testing performance and reliability of fiber optic sensing system for long-term monitoring. In *Proceedings of SPIE*, volume 5502, pages 552–555, 2004.

MathWorks Inc. Matlab version 7.12.0 (r2011a), 2011.

M. R. Jahanshahi, J. S. Kelly, S. F. Masri, and G. S. Sukhatme. A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. *Structure and Infrastructure Engineering*, 5(6):455–486, 2009.

M. R. Jahanshahi, S. F. Masri, and G. S. Sukhatme. Multi-image stitching and scene reconstruction for evaluating defect evolution in structures. *Structural Health Monitoring*, 10(6):643–657, 2011.

Z. Jankó, G. Kós, and D. Chetverikov. Creating entirely textured 3d models of real objects using surface flattening. *Machine Graphics and Vision*, 14(4):379, 2005.

David V. Jauregui, Yuan Tian, and Ruinian Jiang. Photogrammetry applications in routine bridge inspection and historic bridge documentation. *Transportation Research Record: Journal of the Transportation Research Board*, 1958 (1):24–32, 2006.

X. Jiang and H. Adeli. Pseudospectra, music, and dynamic wavelet neural network for damage detection of highrise buildings. *International Journal for Numerical Methods in Engineering*, 71(5):606–629, 2007.

L. P. Kalra, J. Guf, and M. Meng. A wall climbing robot for oil tank inspection. In *Robotics and Biomimetics, 2006. ROBIO'06. IEEE International Conference on*, pages 1523–1528. Ieee, 2006.

Vineet R. Kamat and Sherif El-Tawil. Evaluation of augmented reality for rapid assessment of earthquake-induced building damage. *Journal of Computing in Civil Engineering*, 21(5):303–310, 2007.

M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, pages 61–70. Eurographics Association, 2006.

R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint conference on artificial intelligence*, volume 14, pages 1137–1145. Citeseer, 1995.

J. H. Lee, J. M. Lee, H. J. Kim, and Y. S. Moon. Machine vision system for automatic inspection of bridges. In *Image and Signal Processing, 2008. CISP'08. Congress on*, volume 3, pages 363–366. IEEE, 2008.

H. P. A. Lensch, W. Heidrich, and H. P. Seidel. Automated texture registration and stitching for real world models. In *Computer Graphics and Applications, Proceedings. The Eighth Pacific Conference on*, pages 317–452. IEEE, 2000.

R. S. Lim, H. M. La, Z. Shan, and W. Sheng. Developing a crack inspection robot for bridge maintenance. In *Robotics and Automation (ICRA), IEEE International Conference on*, pages 6288–6293. IEEE, 2011.

Y. Liu, S. Mohanty, and A. Chattopadhyay. Condition based structural health monitoring and prognosis of composite structures under uniaxial and biaxial loading. *Journal of Nondestructive Evaluation*, 29(3):181–188, 2010.

D. Longo and G. Muscato. The alicia3 climbing robot: a three-module robot for automatic wall inspection. *Robotics & Automation Magazine, IEEE*, 13(1): 42–50, 2006.

K. Lu, C. H. Loh, Y. S. Yang, J. P. Lynch, and KH Law. Real-time structural damage detection using wireless sensing and monitoring system. *Smart Structures and Systems*, 4(6):759–777, 2008.

B. L. Luk, D. S. Cooke, S. Galt, A. A. Collie, and S. Chen. Intelligent legged climbing service robot for remote maintenance applications in hazardous environments. *Robotics and Autonomous Systems*, 53(2):142–152, 2005.

J. P. Lynch and K. J. Loh. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock and Vibration Digest*, 38(2): 91–130, 2006.

A. Mal, S. Banerjee, and F. Ricci. An automated damage identification technique based on vibration and wave propagation data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365 (1851):479, 2007.

Charles P Marsh, Amer Siddique, Brian Temple, Vincent M Hock, and Frank Robb. Fury: Robotic in-situ inspection/condition assessment system for underground storage tanks. Technical report, DTIC Document, 2004.

David Mascarenas, Eric Flynn, Charles Farrar, Gyuhae Park, and Michael Todd. A mobile host approach for wireless powering and interrogation of structural health monitoring sensor networks. *Sensors Journal, IEEE*, 9(12):1719–1726, 2009.

A. Mazumdar and HH Asada. Mag-foot: A steel bridge inspection robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009. IROS 2009*, pages 1691–1696, 2009.

L. L. Menegaldo, G. Ferreira, M. F. Santos, and R. S. Guerato. Development and navigation of a mobile robot for floating production storage and offloading ship hull inspection. *Industrial Electronics, IEEE Transactions on*, 56(9):3717–3722, 2009.

N. Metni and T. Hamel. A uav for bridge inspection: Visual servoing control law with orientation limits. *Automation in Construction*, 17(1):3–10, 2007.

JC Middleton. Strain gauge performance during fatigue testing. *Strain*, 22(3): 135–140, 1986.

CL Muhlstein, SB Brown, and RO Ritchie. High-cycle fatigue and durability of polycrystalline silicon thin films in ambient air. *Sensors and Actuators A: Physical*, 94(3):177–188, 2001.

R. R. Murphy, E. Steimle, C. Griffin, C. Cullins, M. Hall, and K. Pratt. Cooperative use of unmanned sea surface and micro aerial vehicles at hurricane wilma. *Journal of Field Robotics*, 25(3):164–180, 2008.

R. R. Murphy, K. L. Dreger, S. Newsome, J. Rodocker, E. Steimle, T. Kimura, K. Makabe, F. Matsuno, S. Tadokoro, and K. Kon. Use of remotely operated marine vehicles at minamisanriku and rikuzentakata japan for disaster recovery. In *Safety, Security, and Rescue Robotics (SSRR), 2011 IEEE International Symposium on*, pages 19–25. IEEE, 2011a.

R. R. Murphy, E. Steimle, M. Hall, M. Lindemuth, D. Trejo, S. Hurlebaus, Z. Medina-Cetina, and D. Slocum. Robot-assisted bridge inspection. *Journal of Intelligent & Robotic Systems*, 64(1):77–95, 2011b.

J. K. Oh, A. Y. Lee, S. M. Oh, Y. Choi, B. J. Yi, and H. W. Yang. Design and control of bridge inspection robot system. In *International Conference on Mechatronics and Automation, 2007. ICMA 2007*, pages 3634–3639, 2007.

Anibal Ollero, Joaquin Ferruz, Fernando Caballero, Sebastian Hurtado, and Luis Merino. Motion compensation and object detection for autonomous helicopter visual navigation in the comets system. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, volume 1, pages 19–24. IEEE, 2004.

Patrick OBrien, Marc Eberhard, Olafur Haraldsson, Ayhan Irfanoglu, David Lattanzi, Steven Lauer, and Santiago Pujol. Measures of the seismic vulnerability of reinforced concrete buildings in haiti. *Journal Information*, 27(S1), 2011.

R. T. Pack, J. L. Christopher Jr, and K. Kawamura. A rubbertuator-based structure-climbing inspection robot. In *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, volume 3, pages 1869–1874 vol. 3. Ieee, 1997.

J. M. Pena, J. A. Lozano, and P. Larranaga. An empirical comparison of four initialization methods for the k-means algorithm. *Pattern Recognition Letters*, 20(10):1027–1040, 1999.

B. Phares, G. Washer, D. Rolander, B. Graybeal, and M. Moore. Routine highway bridge inspection condition documentation accuracy and reliability. *Journal of Bridge Engineering*, 9:403, 2004.

O. Pizarro, R. Eustice, and H. Singh. Large area 3d reconstructions from underwater surveys. In *OCEANS '04. MTTS/IEEE TECHNO-OCEAN'04*, volume 2, pages 678–687. IEEE, 2004.

M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual modeling with a hand-held camera. *International Journal of Computer Vision*, 59(3):207–232, 2004.

PrimeSense, 2011. URL `http://www.primesense.com/`.

F. Provost and T. Fawcett. Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions. In *Proceedings of the third international conference on knowledge discovery and data mining*, pages 43–48. AAAI Press, 1997.

J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1(1):81–106, 1986.

J. R. Quinlan. Learning with continuous classes. In *Proceedings of the 5th Australian joint Conference on Artificial Intelligence*, volume 92, pages 343–348. Singapore, 1992.

G. La Rosa, M. Messina, G. Muscato, and R. Sinatra. A low-cost lightweight climbing robot for the inspection of vertical surfaces. *Mechatronics*, 12(1):71–96, 2002.

H. Schempf, B. Chemel, and N. Everett. Neptune: above-ground storage tank inspection robot system. *Robotics & Automation Magazine, IEEE*, 2(2):9–15, 1995.

Sebastian Scherer, Joern Rehder, Supreeth Achar, Hugh Cover, Andrew Chambers, Stephen Nuske, and Sanjiv Singh. River mapping from a flying robot: state estimation, river detection, and obstacle mapping. *Autonomous Robots*, 33(1-2):189–214, 2012.

Jean Serra. *Image analysis and mathematical morphology*. London: Academic Press, 1982.

Pradip N Sheth. Use of robotics for nondestructive inspection of steel highway bridges and structures. Technical report, Transportation Research Board, 2005.

J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.

S. K. Sinha and P. W. Fieguth. Segmentation of buried concrete pipe images. *Automation in Construction*, 15(1):47–57, 2006.

N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *International Journal of Computer Vision*, 80(2):189–210, 2008.

T. Sogi, Y. Kawaguchi, H. Morisaki, K. Ohkawa, N. Kai, and H. Hayakawa. Inspection robot for spherical storage tanks. In *Industrial Electronics Society, 2000. IECON 2000. 26th Annual Conference of the IEEE*, volume 1, pages 393–398. IEEE, 2000.

H. Sohn and C. R. Farrar. Damage diagnosis using time series analysis of vibration signals. *Smart Materials and Structures*, 10:446, 2001.

S. Soyoz and M. Q. Feng. Long-term monitoring and identification of bridge structural parameters. *Computer-Aided Civil and Infrastructure Engineering*, 24(2):82–92, 2009.

D. Tingdahl and L. Van Gool. A public system for image based 3d model generation. *Computer Vision/Computer Graphics Collaboration Techniques*, pages 262–273, 2011.

M. M. Torok. Autonomous sample collection using image-based 3d reconstructions. Master's thesis, Virginia Polytechnic Institute and State University, Virginia, 2012.

K. Toussaint, N. Pouliot, and S. Montambault. Transmission line maintenance robots capable of crossing obstacles: State of the art review and challenges ahead. *Journal of Field Robotics*, 26(5):477–499, 2009.

P. C. Tung, Y. R. Hwang, and M. C. Wu. The development of a mobile manipulator imaging system for bridge crack inspection. *Automation in Construction*, 11(6):717–729, 2002.

J. M. Mirats Tur and W. Garthwaite. Robotic devices for water main in-pipe inspection: A survey. *Journal of Field Robotics*, 27(4):491–508, 2010.

J. Valença, E.N.B.S. Júlio, and H.J. Araújo. Applications of photogrammetry to structural assessment. *Experimental Techniques*, 36(5):71–81, 2012.

V. N. Vapnik. *The nature of statistical learning theory*. Springer Verlag, 2000.

C. Walter, J. Saenz, N. Elkmann, H. Althoff, S. Kutzner, and T. Stuerze. Design considerations of robotic system for cleaning and inspection of large-diameter sewers. *Journal of Field Robotics*, 29(1):186–214, 2012.

L. Wang and T. H. T. Chan. Review of vibration-based damage detection and condition assessment of bridge structures using structural health monitoring. In *QUT Conference Proceedings*, 2009.

Wei Wang and Christopher M. Clark. Modeling and simulation of the videoray pro iii underwater vehicle. In *OCEANS 2006-Asia Pacific*, pages 1–7. IEEE, 2007.

X. Wang and F. Xu. Design and experiments on cable inspection robot. In *Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE*, pages 2870–2875. IEEE, 2007.

I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

K. Worden, C. R. Farrar, G. Manson, and G. Park. The fundamental axioms of structural health monitoring. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 463(2082):1639, 2007.

R. S. Wright, N. Haemel, G. Sellers, and B. Lipchak. *OpenGL SuperBible: comprehensive tutorial and reference*. Addison-Wesley Professional, 2010.

Fengyu Xu, Xingsong Wang, and Lei Wang. Cable inspection robot for cable-stayed bridges: Design, analysis, and application. *Journal of Field Robotics*, 28 (3):441–459, 2011.

T. Yamaguchi and S. Hashimoto. Fast crack detection method for large-size concrete surface images using percolation-based image processing. *Machine Vision and Applications*, 21(5):797–809, 2010.

T. Yamaguchi, S. Nakamura, R. Saegusa, and S. Hashimoto. Image-based crack detection for real concrete surfaces. *IEEJ Transactions on Electrical and Electronic Engineering*, 3(1):128–135, 2008.

S. N. Yu, J. H. Jang, and C. S. Han. Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel. *Automation in Construction*, 16(3):255–261, 2007.

Z. Zhang. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330–1334, 2000.

Z. Zhu, S. German, and I. Brilakis. Visual retrieval of concrete crack properties for automated post-earthquake structural safety evaluation. *Automation in Construction*, 20(7):874–883, 2011.