A Multi-Functional Robot For Civil Infrastructure Inspection

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science & Engineering

by

Tuan D. Le

Dr. Hung M. La - Thesis Advisor
May 2017
We recommend that the thesis prepared under our supervision by

**TUAN D LE**

Entitled

*A Multi-Functional Robot For Civil Infrastructure Inspection*

be accepted in partial fulfillment of the requirements for the degree of

**MASTER OF SCIENCE**

Hung M. La, Ph.D., Advisor

George Bebis, Ph.D., Committee Member

Hao Xu, Ph.D., Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

May, 2017
Abstract

This thesis focuses on designing and building a robotic system for civil infrastructure inspections. The robot is equipped with two non-destructive-evaluation (NDE) types of sensors: a ground penetrating (GPR) sensor and two electrical resistivity (ER) sensors. For localization, the robot employs an extended Kalman filter (EKF)-based approach using global positioning system (GPS) signal. Differ from its predecessor, the currently built robot is capable of localizing and navigating in GPS-denied environments by using a visual inertial odometry method with a stereo camera. In addition, the robot can build an elevation map of its surrounding areas using a PrimeSense camera for obstacle avoidance.
Acknowledgments

This thesis research is supported by the National Science Foundation under grant NSF-IIP-ICoprs #1639092, the University of Nevada, Reno, and the INSPIRE University Transportation Center. Financial support for the INSPIRE University Transportation Center was provided by the U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) under grant agreement No. 69A3551747126 and by participating state departments of transportation.

The views, opinions, findings and conclusions reflected in this presentation are the responsibility of the presenter only and do not represent the official policy or position of the USDOT/OST-R, NSF or any State or other entity.

I would like to thank my advisor, Dr. Hung La, who provided advices and supports to help me finish my study.

I would like to express my gratitude to my committee members, Dr. George Bebis and Dr. Hao Xu, for the precious time and effort which you have put in to review this thesis and for providing advice along the way.

I would also like to thank my partners, Nhan Pham and Quynh Tran, for their endless inspirations during my difficult times.

Finally I would like to thank my amazing wife, Iris, who has decidedly sacrificed everything to support me. I could not have done this without her.
Dedication

Dedicated to my parents.
# Table of Contents

1 Thesis Introduction, Contribution, and Content  
   1.1 Introduction ........................................... 1  
   1.2 Contribution ......................................... 2  
   1.3 Content ............................................... 3  

2 Background  
   2.1 Half-cell Potential (HCP) Method .......................... 5  
   2.2 Impact Echo (IE) ......................................... 5  
   2.3 Electrical Resistivity (ER) ................................ 5  
   2.4 Ground Penetrating Radar (GPR) .......................... 6  
   2.5 Related Work and Motivation to Develop A New Inspection System ........................................ 7  
   2.6 Summary ............................................... 10  

3 Mechanical Design, Its Implementation and Sensors  
   3.1 System Design and Implementation .......................... 11  
   3.2 Ground Penetrating Radar and Electric Resistivity ........ 14  
   3.3 Stereo camera and IMU ...................................... 17  
   3.4 PrimeSense camera ....................................... 18  
   3.5 Summary ............................................... 18
## 4 Localization

4.1 GPS-based Localization with Extended Kalman Filter . . . . . . . . . 19  
4.2 Stereo Inertial Odometry . . . . . . . . . . . . . . . . . . . . . . . . . 21  
4.3 Elevation Mapping . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

## 5 Experimental Results

5.1 Field test for nondestructive evaluation sensors . . . . . . . . . . . . . 31  
5.2 Localization with GPS . . . . . . . . . . . . . . . . . . . . . . . . . . 33  
5.3 Visual inertial odometry with stereo camera . . . . . . . . . . . . . . 33  
5.4 Robot deployment and Elevation map inside a garage building . . . . 35

## 6 Conclusions and Future Work

6.1 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42  
6.2 Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
List of Figures

1.1 A multi-function robot for civil infrastructure inspection 3

2.1 Operation of NDE sensors by team members from the Advanced Robotics and Automation (ARA) Lab for bridge deck inspection on the Pleasant Valley Bridge on Highway 580 from Reno, NV toward Carson City, NV. 7

2.2 An overview of RABIT system. (a) Front view. (b) Rear view. (c) Side view. Image courtesy of CAIT, Rutgers University. Image is used with a corresponding author’s consent. 8

3.1 Seekur Jr. as a mobile platform with dimensions in millimeter. 12

3.2 A 3D design of the whole system: A - GPR sensor’s box; B - Mobile platform; C - ER deployment system. 13

3.3 GPR deployment system: design (a) and implementation (b): A - motor; B - gear shaft; C - GPR’s box. 13

3.4 ER deployment system: design (a) and implementation (b): A - motor; B - gear shaft; C - two ER sensors attached to an L-shape bar. 14

3.5 Visual system is mounted in front of the robot. 15

3.6 Working principle of GPR sensor: (a) - Illustration; (b) - Actual data from field test. 16

3.7 A sample of condition map using GPR data from field test. 16
3.8 Electric Resistivity sensor: (a) - Resipod ER; (b) - ER’s working principle

3.9 Stereo inertial odometry: (a) Zed stereo camera; (b) IMU VN-100.

3.10 A PrimeSense Camera from Asus - AsusXtion.

4.1 Overview on the workflow of a feature in the filter state.

5.1 Pleasant Valley Bridge on Highway 580, Reno, Nevada with the surveyed areas marked by yellow lines (image taken from Google Map).

5.2 Condition map of the surveyed areas on Pleasant Valley Bridge. (a) - Northbound part of the bridge; (b) - Southbound part of the bridge.

5.3 (Left sub-figure) Performance of GPS+IMU-EKF-based localization with robot.localization package. The red dots are GPS signal and blue dots are output of EKF localization. (Right-sub figure) Zoom-in/Close-look at one location: the EKF outperforms the GPS alone since it outputs smoother results.

5.4 GPS-based localization with EKF. The robot has been driven along a 34 ft-by14 ft square 10 times. The red plus sign denotes GPS location. The blue dot denotes robot’s location from wheelodometry. The green dot denotes the localization algorithm’s output.

5.5 Static calibration result for left camera.

5.6 Static calibration result for right camera.

5.7 Acceleration error from IMU.

5.8 Angular velocity error from IMU.

5.9 Feature tracking performance from stereo visual inertial sensor.

5.10 The robot is moving inside a garage building.

5.11 Elevation map inside the garage parking. A car appears in front of the robot and the map is being updated accordingly.

5.12 Elevation map is being built when the robot moves along.
Chapter 1

Thesis Introduction, Contribution, and Content

1.1 Introduction

Civil infrastructure is a key element of an economy. In order to sustain an economic growth and social development of a modern society, satisfactory operation performance of civil infrastructure must be guaranteed. For example, a national highway transportation system is one of the most critical foundations of the US economy because it provides crucial nodes to move people and goods in time [1]. The national highway transportation system (NHS) consists of several elements, including roads, bridges, ports, etc. Even though the NHS covers only 5.5% of total nation’s roads, it supports 55% of all vehicle traffic and 97% of truck-borne goods transportation [2]. Obviously, in order to keep pace with this growing usage, roads need to be properly maintained. In addition, there are several other parts of the road system that requires special attentions such as bridges, tunnels, etc. Among them, bridges are crucially
important due to their distinct function: connecting road nodes. Since bridges are constructed using concrete and they are constantly exposed to harsh environments, they are at most vulnerability.

According to [3], the number of deficient bridges in the US is more than 180,000. Those bridges therefore require proper maintenance to avoid any catastrophic accident such as [4]. Inherently, bridge deck inspection for maintenance is labor-intensive work and high cost. However, with the recent growth in technology, there are many interesting research focusing on development new inspecting sensors and techniques [5–8]. These sensors and techniques have their advantages and disadvantages. Yet, they are still being utilized separately and there is lack of an integrated system to provide a complete inspecting process [9,10].

In this thesis, a novel robotic system, which is capable of inspecting multiple types of infrastructure, is presented (Fig. 1.1). The robot is equipped with different types of sensors so that it can perform a complete infrastructure inspection. This thesis presents the design, construction and working capability of this system.

1.2 Contribution

This work provides a mechanical design of a robotic system. Differ than its predecessor [9,11,17], this new system is built on a smaller base, which enables the robot to operate in various environments. The previous robot is solely for bridge deck inspection while the current robot can inspect bridge deck, pavement, parking garage, etc. In addition, the new robot is able to work in indoor environments utilizing its visual inertial sensor.
This thesis is organized as follows. Chapter 2 presents an overview of infrastructure inspection using non-destructive evaluation (NDE) methods. We also discuss incentives to develop a new system for inspections. Chapter 3 provides details about the robot’s mechanical design and its implementation. Chapter 2 presents a brief discussion about NDE sensors that are utilized on the robot and other sensors for robot’s localization capability. Chapter 4 provides details of localization’s algorithms and path planning. In Chapter 5 some preliminary results are presented. Finally, Chapter 6 discusses ideas on future work and concludes the thesis.
Chapter 2

Background

As briefly mentioned in the Introduction, many infrastructures are subjected to harsh environments and constantly under heavy load such as roads, bridges, etc. This makes maintenance processes including inspections, evaluations, and rehabilitations a burden from financial point of view. For example, reinforced concrete structures, such as bridge decks, parking garage’s floors, are prone to several types of deteriorations: corrosion, carbonation, freeze-thaw actions, etc. Affects from these processes are not always visual. Therefore, NDE methods are preferable for inspections due to its simplicity and accuracy. Modern NDE methods for concrete structure inspections have their origins in geophysics. In general, NDE methods exploit the fact that different materials or material’s states response distinctly to external excitations. Hence, NDE methods utilize an approach in which characteristics of an inspected object are revealed by measuring how the object responses to the applied excitation. There are several NDE methods and some selected ones will be discussed here.
2.1 Half-cell Potential (HCP) Method

HCP method is a widely-used method to assess corrosion of steel-reinforced concrete structures by identifying the presence of active corrosive processes [18]. The amount of electro-chemical activity in a path between two points on a structure. If the measurement is less than $-0.35 \text{V}$, then there is a 90% probability of corrosion. If the measurement is higher than $-0.2 \text{V}$, then there is a 90% probability of absence of corrosion [19]. This method is popular for its simplicity of implementation. However, there is not any detailed study of how to use this method for concrete cover depth.

2.2 Impact Echo (IE)

This method uses a mechanical impact such as a hammer to excite and send high-frequency elastic waves into the concrete structure. By evaluating the reflections waves, which are the interactions of the input signals to the subsurface features, one can evaluate various deterioration stages. Due to a significant contrast in rigidity of concrete and air, the impact input source is sufficiently reflected from the bottom of the concrete structure back to its surface. A good and fair condition will be represented by one and two distinct peaks in the response spectrum, respectively. However, a serious condition or being heavily delaminated will always be in the audible frequency range. This method provides very accurate evaluation results but the process is slow and requires manual placement [20].

2.3 Electrical Resistivity (ER)

In a concrete structure, electrical conduction occurs mainly because of electrolytic current flow through the open pore system and the formation of electro-chemical
corrosion cells. In other words, the higher conductive a concrete structure is, the more likely damaged and/or cracked areas take place. It has been observed that a resistivity of less than $5k\Omega \cdot cm$ is a strong indication of high level of corrosion $^{[21]}$. In contrast, the concrete with high resistance ($> 100k\Omega \cdot cm$) show a low probability of being corrosive. When one uses the ER to measure, it is important that the concrete surface has to be prewet and the concrete surface is not coated by any electrical insulating layer. This is the disadvantage of ER sensors in addition to low rate process and readings should be taken in a rather small grid ($0.6 \times by \times 0.6m$ grid) to ensure adequate data quality.

2.4 Ground Penetrating Radar (GPR)

GPR provides an electromagnetic wave reflection survey. It uses high frequency radar microwaves ($100MHz$ to $3GHz$) to evaluate subsurface features. The radar signal travels through dielectric materials. A portion of the signal’s energy will reflect back to the surface if the signal contacts a different materials (from air to ground, from concrete to steel rebars). The GPR signal can not penetrate through metals, therefore it will be reflected most. By capturing the time of fly of the reflected signals and their amplitudes, one can estimate the subsurface target depth and its condition $^{[22-24]}$. Electrical conductivity, as well as material dielectric properties, affect how a GPR signal travel. For example, GPR will not work in wet environments because radar signals are absorbed by water.
2.5 Related Work and Motivation to Develop A New Inspection System

![Figure 2.1: Operation of NDE sensors by team members from the Advanced Robotics and Automation (ARA) Lab for bridge deck inspection on the Pleasant Valley Bridge on Highway 580 from Reno, NV toward Carson City, NV.](image)

Obviously, each NDE sensor has its advantages and limitations. Moreover, skilled operators are required to use these sensors. For example, in our data collection for field test, there were five operators to conduct data collections. This tedious process is presented in Fig. 2.1. As can be seen, these operators were working in a dangerous environment with continuously flowing traffic next to an inspecting site. Hence, it is desirable to develop an autonomous system with integrated multiple sensor to
exploit complementaries from different sensors to provide a complete and accurate inspecting result and b) replace human operator therefore reducing maintenance cost.

There were some attempts to build an automatic data collection system \[25\]-\[27\], for which, the first attempt was dated back to 1993 [25]. In [25], a portable seismic pavement analyzer (PSPA), utilizing impact echo method, was developed. This system even though was primitive but it inspired a research trend in building advanced systems for bridge deck inspection [28]. In [26\-27], an automatic GPR data collection system HERMES/PERES II was developed. This was an integrated system including a GPR unit, motion control hardware, a calibration and signal conditioning system and a signal processor. The whole system rides on a set of rubber wheels, which are controlled by stepper motors, to move along an inspecting area. The main disadvantage of this system is that it is a simple semi-automatic mechanical system.

![Figure 2.2](image.png)

Figure 2.2: An overview of RABIT system. (a) Front view. (b) Rear view. (c) Side view. Image courtesy of CAIT, Rutgers University. Image is used with a corresponding author’s consent.

Hence, it is desirable to have an integrated system, utilizing not just multiple types of NDE sensors but also other environment-aware sensors as well. To the best of the author’s knowledge, there is only one such system exists [9\-11\-14]. This system was developed at the Center for Advanced Infrastructure and Transportation (CAIT), Rutgers University in 2013 [12]. This system also comprises of various NDE sensors including two real time kinematic (RTK) GPS units, three laser scanners, two GPR units, two acoustic array sensors, four ER sensors, two high-resolution digital cameras
and a panoramic camera \cite{9}. This system is developed for bridge deck inspections. An overview of this system is in Fig.\ref{fig:overview}. This system is developed by a professional team \cite{29}, however, it still has disadvantages.

First, this system is designed for bridge deck inspections specifically. This severely limits its application, since bridges are just a part of a much more complex civil infrastructure system that need to be maintained properly as well. In addition, with its current size the robot would require a spacious area to operate, which also limits its application. Second, for localization, the robot relied on two RTK Novatel GPS units and an inertial measurement unit (IMU). This approach has several drawbacks, such as: the RTK GPS units are expensive and inapplicable in many scenarios: cloudy outdoor, unaccessible to the sky when inspecting multi-level bridges, etc. Even though this is partially compensated by fusing IMU data by using an EKF-based method, good localization result can only pertain for a short distance due to IMU’s drifting and an incorrect robot’s model. And finally, it is unclear how the robot would use laser scanner sensors for its navigation \cite{9,14}.

Motivating by these disadvantages, a new robotic system is proposed in this thesis. By using a vision-based approach, a stereo inertial odometry in combination with GPS and IMU sensor will enhance the system’s versatility. Using a stereo camera for feature tracking, a stereo inertial odometry can work in outdoor and indoor environments, in which, for outdoor environment, whenever GPS signal is available, it will be fed as an external pose to a stereo inertial odometry to maximize the robot’s localization capability. In addition, utilizing pointcloud data from a PrimeSense camera, the robot is able to build an elevation map of its surrounding environment, in which, it can be used for obstacles avoidances. Lastly, with a smaller form design, the new system is capable of operating in a much narrow space, which opens new applications such as indoor storage inspections, parking garage inspections, etc.
2.6 Summary

In this chapter, an overview of existing NDE sensors and inspection methods is presented. Advantages and limitations of each sensor are also discussed. A detailed discussion about related work, its disadvantages is provided, followed by the motivation of this work.
Chapter 3

Mechanical Design, Its Implementation and Sensors

3.1 System Design and Implementation

The Seekur Jr. mobile robot is chosen as a mobile base platform. This is a skid-steering four-wheel-drive robot. Having a smaller form than its sister [1], this mobile platform is capable of operating on narrow bridges as well as pavements, garage parking, storages, etc.

The design goal is to build a system with robust deployment mechanisms for NDE sensors. Deployment mechanisms of NDE sensors should work accurately and fast for data collection. The deployment mechanism should be able to move attached NDE sensors up and down accordingly to data collection process: to collect GPR data, the GPR’s box needs to be in contact with the ground throughout the process. Steps for collecting ER sensors are slightly different. The robot needs to move and stop typically every 2 feet. When the robot stops, a small amount of water is sprayed
Figure 3.1: Seekur Jr. as a mobile platform with dimensions in millimeter.

over the inspecting area to create a conductive area. After that, the deployment mechanism needs to move ER sensors to firmly touch the wet surface to record data. There are also several other sensors that need to be mounted on-board: a stereo camera, a PrimeSense camera, an IMU, GPS unit. It is desirable that with multiple sensors mounted on-board, the mobile platform is still in compact form to ensure efficient movement.

To address those requirements, a simple but highly functional design was developed. In Fig 3.2, an overall system design is presented. The GPR sensor is mounted
Figure 3.2: A 3D design of the whole system: A - GPR sensor’s box; B - Mobile platform; C - ER deployment system.

on the front of the robot while two ER sensors are mounted on the rear. This is to prevent spraying water from ER sensors interfering GPR sensor.

Figure 3.3: GPR deployment system: design (a) and implementation (b): A - motor; B - gear shaft; C - GPR’s box.

In Fig.3.3, a detail of design and implementation of a GPR deployment system is presented. The GPR’s box is attached underneath a moving mechanism. Using a DC motor, the GPR’s box will be lifted up and down accordingly to the current action’s purpose: if the robot is to move without collecting GPR data, the box is lifted off the ground; if the robot is collecting GPR data, then the box is moved down to touch the ground.
A similar deployment system is utilized for ER sensors. In Fig. 3.4, the ER deployment system is presented. Similarly, the ER deployment system operates with the same mechanism of the GPR deployment system: the two ER sensors will be moved to touch the ground whenever ER data is collected.

A GPS unit is mounted at the center of the robot to provide lateral and longitudinal positions. An IMU is attached on top of the stereo camera for stereo inertial odometry. The stereo and IMU set is then placed on top of the PrimeSense Camera and create a camera stack. The stack is mounted on the robot’s front for localization. The camera stack is showed in Fig. 3.5.

### 3.2 Ground Penetrating Radar and Electric Resistivity

A SIR-3000 unit from Geophysical Survey System Inc. (GSSI) is chosen due to its small size, lightweight and multi-functions. Beside performing deep scan on concrete structure for corrosion inspection, it can also perform other geophysical survey tasks. The working principle of GPR sensor is as follows. Extremely short electromagnetic
Figure 3.5: Visual system is mounted in front of the robot.

Pulses are transmitted to an inspecting structure. Reflected from a target, the received signal carries scaled and delayed information of the transmitted one. This effect is known as pulse-echo radar. By recording the time-of-fly and amplitude of the received signal while moving the antenna past the target, the received pulse-echoes form a hyperbolic arc of the target’s surface. In-depth discussion about how GPR works can be found in [30] and references therein. A simple illustration of GPR’s working principle is presented in Fig. 3.6.

For visualization and interpretation, acquisition GPR data is usually presented in a conditional map. The map uses color codes to indicate the condition of inspecting area. In Fig. 3.7, a sample condition map is showed. This condition map uses four colors blue, green, orange and red to indicate good, normal, poor and bad condition, respectively. The use of color code is varied and depended on user’s intention.
Figure 3.6: Working principle of GPR sensor: (a) - Illustration; (b) - Actual data from field test.

Figure 3.7: A sample of condition map using GPR data from field test.

As previously mentioned in Section 2.3 Chapter 2, electrical resistivity data of inspected concrete structure can be used as an evaluation of corrosive level. Two Resipod Electric Resistivity sensors from Proceq Inc. are chosen. The Resipod ER sensor uses a four-electrode Wenner probe. Electrical current runs through two outer electrodes and creates an electrical field, which is measured by the two inner electrodes. The target’s resistivity is then calculated by using the following equation:

\[ \text{Resistivity} = \frac{2\pi a V}{I} \]  

(3.1)

where \( a \) is a distance between two consecutive electrodes.
3.3 Stereo camera and IMU

A Zed stereo camera from Stereo Lab Inc. is used. It consists of two high definition cameras, which can output 1280 × 720 pixel-by-pixel images at 60Hz. The stereo camera has its own Robot Operating System (ROS) driver [31]. An IMU Vn 100 unit from VectorNav Technologies Inc. is chosen to work with the Zed stereo camera. This is an 10-axis MEMS IMU, including 3-axis accelerometer, 3-axis gyroscope, 3-axis magnetometer and a barometric pressure. The IMU had been undergone thermal calibration to accommodate with outdoor inspections. It is placed on top of the Zed stereo camera to create a stereo inertial odometry sensor. The camera IMU calibration process will be detailed in the next Chapter.
3.4 PrimeSense camera

The PrimeSense camera is chosen to produce an elevation map by using its point cloud data. Its working principle is as follows. The PrimeSense camera projects an infrared speckle pattern, which is recaptured by an infrared sensor inside the camera to produce a depth image. The depth image is then correlated to a RGB camera to produce the point cloud data [32]. It is similar to a Microsoft Kinect but this version from Asus is significantly lightweight and has a smaller form.

3.5 Summary

In this Chapter, a detail discussion about the design of the robot is presented. Detail information about sensors including NDE sensors and sensors for localization is provided. Their working principles are also discussed. In the next chapter, a thorough discussion about the robot’s localization will be carried out.
Chapter 4

Localization

Localization and Mapping are the two most fundamental problems for autonomous robots [33]. An autonomous robot must be able to answer two questions “What does my environment look like?” and “Where am I in my environment?”. This is a Chicken-or-Egg problem, which has been studied for decades. Despite significant progresses had been made, the simultaneous localization and mapping or SLAM still attracts many interesting research. In this thesis, two different methods of localization are explored. They are GPS-based localization with Extended Kalman Filter (EKF) and Visual Inertial Odometry. These two methods are chosen due to their robustness and complementarities between them.

4.1 GPS-based Localization with Extended Kalman Filter

For outdoor navigation, global position system (GPS) has a long history of extensive research. Interested readers are referred to [34-36] and references therein. For outdoor
navigation, a GPS unit is employed in combination with an extended Kalman filter (EKF) for localization. We implemented a Robot Operating System (ROS) software package robot_localization [37]. Applications of the EKF algorithm have been around for decades [38]. For the developed robot, it was necessary to estimate the 2D pose and velocity of the robot using three sources of information: GPS signal, inertial information from the IMU, and the robot’s wheel odometry. The process is detailed as follows.

\[
x_k = f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1})
\]
\[
w_{k-1} \sim (0, Q)
\]
\[
z_k = h(x_k) + v_k
\]
\[
v_k \sim (0, R_k).
\]

The robot’s state (position and orientation) is \(x_k\) at time \(k\) with a nonlinear state transition \(f\) - a kinematic model derived from Newtonian mechanics, control input \(u\) and zero mean, Gaussian distributed process noise \(w\) with covariance \(Q\). The robot’s pose includes its Cartesian positions, orientations and velocities. The measurement is of the form \(z_k\) with a nonlinear sensor model \(h\) and zero mean, Gaussian distributed measurement noise \(v_k\) with covariance \(R_k\). Then the prediction and correction steps are carried out.

\[
\dot{x}_k = f(x_{k-1})
\]
\[
\dot{P}_k = FP_{k-1}F^T + Q
\]
\[
K = \dot{P}_kH^T(H\dot{P}_kH^T + R)^{-1}
\]
\[
x_k = \hat{x}_k + K(z - \hat{x}_k)
\]
\[
P_k = (I - KH)\dot{P}_k(I - KH)^T + KRK^T
\]
where, $P$ is an estimated error covariance, $F$ is a Jacobian of $f$, $K$ is Kalman gain, and $H$ is an observation matrix. By definition [38], $H$ should be a Jacobian matrix of the observation model function $h$. However, it is desirable to integrate different sensors such as wheel encoders, LIDAR, etc., into the EKF. Therefore, each sensor is assumed to produce measurements that are estimated. Hence, the matrix $H$ is just simply an identity matrix. With this, sensors that do not measure every state variables are supported. In particular, if on $m$ state variables are measured, then $H$ becomes an $m$-by-$12$ matrix, in which nonzero values existing only in the columns of the measure variables. Inherently, it is difficult to estimate process noise covariance $Q$ because it is usually unknown. Typically, there are four general approaches to estimate the process noise covariance $Q$: Bayesian [39], maximum likelihood [40], covariance matching [41] and correlation techniques [42–44]. Bayesian and maximum likelihood methods are computationally expensive and sometimes are impractical. Covariance matching is a method to calculate covariances from the residuals of the state estimation problem. However, it had been shown to give biased estimates of the true covariances. The last method is the most widely applied one, yet, it also had been shown that the conditions for uniqueness of covariances were insufficient [45]. Therefore, in this work, the process noise covariance is tuned manually. This approach has been implemented in [37].

### 4.2 Stereo Inertial Odometry

[46] For an autonomous robot, it is important to reliably estimate the trajectory of its moving body, or so-called “ego-motion”. In application of interests, such as navigation, the robot’s ego motion estimation must have been made precisely with minimal latency. One method to tackle this problem is to use a GPS sensor to
provide a precisely global position. This approach has several drawbacks, which were
discussed in the previous section 4.1. A common practice is to provide aids from
other sensory modalities such as inertial measurement sensors, encoders of wheel
odometry and so on. An inertial device provides information with high rates and
infinitesimal latency. Recent developments in manufacturing microelectromechanical
systems (MEMS) gave way to produce reliable, small size and low-cost inertial devices
[47]. These types of inertial device had been found ideal for a vast amount of robotic
applications. They, however, only provide information of relative motions and suffer
from low frequency drifts. In [48], utilizing Sagnac effect [49], a fiber optics gyroscope
(FOG) device had been developed. Despite the fact that a FOG inertial device
can avoid drifting, its cost is prohibitively expensive and therefore, its applications
are severely limited. Desired as before, some types of aids are required to exploit
advantages of inertial devices. Vision, with its similar properties to inertial sensors
such as: operating at high rate (modern cameras can operate at 60 Hz with 2560×720
pixels output resolution images), providing estimation of semi-local position (relative
to visibility constraints) is a promising candidate [46]. Vision-based navigation or
so-called “visual-odometry” (VO) has been intensively researched. Some of the most
noticeable attempts can be found in the work of [50–56]. In [57], the author provided
a real-time 3D monocular localization. There are also several attempts to couple
vision with inertial model to produce a hybrid visual-inertial system. These attempts
are referred in [58–64]. Those research varied from localization using short-range
sensors [61], to estimation of long-range motion [65] that had been used in spacecraft
applications. One notation is that many of these work used an Earth-centered, Earth-
fixed coordinate system, which is not favorable due to its inflexibility and requires a
priori knowledge of a global position and orientation.

In this thesis, a visual inertial odometry framework, which combines and ex-
tends several previous approaches, is utilized [46, 66, 67]. In this framework, inertial measurements and visual landmarks are combined into a visual-inertial-EKF-SLAM formulation [46, 67]. The overall filter structure is similar to the one that has been used in [46, 67]: to propagate the state of the filter, inertial measurements are used and to perform the update steps, visual information is taken into account. The filter state is also expressed in a fully robot-centric representation. There are three different coordinate frames, which are used in this visual-inertial odometry framework. First, the inertial world coordinate, \( I \), which is served as an absolute reference for both the camera and the IMU. Second IMU fixed coordinate frame, \( B \), whose origin is attached to the center of the IMU body. Third, the camera fixed frame, \( V \), whose origin is attached to the optical center of the camera with the \( z \)-axis aligned with the optical axis of the lens. The filter state is expressed in the following form:

\[
x := (r, v, q, b_f, b_\omega, c, z, \mu_0, ..., \mu_N, \rho_0, ..., \rho_N)
\]  

(4.3)

where \( r \) is the robot-centric position of IMU (expressed in frame \( B \)); \( v \) is the robot-centric velocity of IMU (expressed in frame \( B \)); \( q \) is a measured attitude of IMU (map from \( B \) to \( I \)); \( b_f \) is the additive bias on accelerometer (expressed in frame \( B \)); \( b_\omega \) is the additive bias on gyroscope (expressed in frame \( B \)); \( c \) is the translational part of IMU-camera extrinsic (expressed in frame \( B \)); \( z \) is the rotational part of IMU-camera extrinsic (map from \( B \) to \( V \)); \( \mu_i \) is a bearing vector to feature \( i \) (expressed in \( V \)); \( \rho_i \) is the distance parameter of feature \( i \).

The distance parameter \( \rho_i \) used in the mapping \( d_i = d(\rho_i) \) of the generic parametrization for the distance \( d_i \) of a feature \( i \). Using the method that is proposed in [68], rotations and unit vectors are parametrized in the form of \( q, z \in SO(3) \) and \( \mu_i \in S^2 \), respectively. More over, a \( \boxminus \)-operator is defined to compute the difference between
two unit vectors within a 2D linear subspace. The state propagation is described in these differential equations:

\[
\begin{align*}
\dot{r} &= -\hat{\omega}_{\text{skew}} r + v + w_r \\
\dot{v} &= -\hat{\omega}_{\text{skew}} v + \hat{f} + q^{-1}(g) \\
\dot{q} &= -q(\hat{\omega}) \\
\dot{b}_f &= w_{bf} \\
\dot{b}_w &= w_{bw} \\
\dot{b}_w &= w_{bw} \\
\dot{c} &= w_c \\
\dot{z} &= w_z
\end{align*}
\]

\[(4.4)\]

\[
\dot{\mu}_i = N^T(\mu_i)\hat{\omega}_V - \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} N^T(\mu_i) \frac{\dot{v}_V}{d(\rho_i)} + w_{\mu,i}
\]

\[
\dot{\rho}_i = -\mu_i^T \dot{v}_V / d'(\rho_i) + w_{\rho,i}
\]

where \(\hat{f}\) and \(\hat{\omega}\) are acceleration measurement and rotational rate measurement, respectively; in which, the subscript \(\text{skew}\) denotes the skew symmetric matrix of a vector; \(N^T(\mu)\) linearly projects a 3D vector onto the 2D tangent space around the bearing vector \(\mu\), with the bias corrected and noise affected IMU measurements:

\[
\begin{align*}
\hat{f} &= \tilde{f} - b_f - w_f \\
\hat{\omega} &= \tilde{\omega} - b_\omega - w_\omega
\end{align*}
\]

\[(4.5)\]

and the camera linear velocity and rotational rate are defined as follows:

\[
\begin{align*}
\dot{v}_V &= z(v + \hat{\omega}_{\text{skew}} c) \\
\dot{\omega}_V &= z(\hat{\omega})
\end{align*}
\]

\[(4.6)\]
In addition, $g$ is the gravity vector in the world coordinate frame $\mathcal{I}$. $w$ with subscripts are white Gaussian noise processes. Following $[68]$, with $\boxplus$-operator, the set of equations $4.4$ are transformed into a set of discrete prediction equations which are used during the prediction of the filter state.

The filter update steps are carried out as follows. First assuming that the intrinsic calibration of the camera is known, then the projection of a bearing $\mu$ to the corresponding pixel coordinate $p = \pi(\mu)$. The update step is performed for every captured image. A 2D linear constraint for each feature $i$, $b_i(\pi(\hat{\mu}_i))$, is derived. The feature $i$ is predicted to be visible in the current frame with bearing vector $\hat{\mu}_i$. The innovation term within the Kalman update uses this constraint, which represents the intensity errors associated with a specific feature, as follows:

$$
y_i = b_i(\pi(\hat{\mu}_i)) + n_i
$$

$$
H_i = A_i(\pi(\hat{\mu}_i)) \frac{d\pi}{d\mu}(\hat{\mu}_i)
$$

(4.7)

where $H_i$ is the Jacobian and $n_i$ is additive discrete Gaussian pixel intensity noise. The standard EKF update is performed for all visible features. Furthermore, a simple Mahalanobis based outliers detection is implemented. The process will reject unsuitable measurements by comparing the obtained innovation with the predicted innovation covariance with a predefined threshold.

The overall work-flow is presented in Fig.4.1.

$[66]$ For each captured image and a given bearing vector $\mu$, patches $P_l$ of $8 \times 8$ pixels are extracted for each image level $l$ at the corresponding pixel coordinate $p = \pi(\mu)$. With a described multilevel patch, features tracking is performed more robustly. In addition, new features are detected using a standard fast corner detector. An adapted Shi-Tomasi score for selecting new features is utilized to add new features to the state. The adapted Shi-Tomashi score considers the combined Hessian on multiple image
Figure 4.1: Overview on the workflow of a feature in the filter state.

level. This brings an advantage, which is that a high score means a high accuracy of the corresponding multilevel patch feature. One important part of using this visual-inertial odometry approach is a calibration process. With a poor camera-IMU extrinsic calibration, the EKF will diverge in a very short amount of time. The calibration process will be detailed in Chapter 5.

4.3 Elevation Mapping

As previously discussed in Section 4.2, a local map of the surrounding environment of a robot is preferable due to its flexibility and inexpensive computation. The inspection robot might be performing inspections in different working environments. Hence, it is not always possible to obtain a global map using robot’s absolute position measurements from GPS. A local map for navigation therefore is desirable. In this work, elevation mapping techniques are of interests because of two reasons. First, the robot moves on the grounds, therefore it is beneficial to represent the robot’s map as a
two dimensional surface. Second, with equipped camera sensors, previous work on elevation maps, such as [69–71], can be applied. In this work, an elevation mapping technique described in [72] is utilized.

In pioneering work of elevation mapping [69, 73], a grid map, in which each cell represents the height of the terrain, is built by matching corresponding transformation between several scans. In [70], the authors proposed a method to fuse range measurements information into cells. The update process is performed based on previously stored data and measurements uncertainty. However, this method requires absolute position measurements, which is not always accessible. Other work such as [71, 74] proposed methods to build local elevation maps by incorporating robot’s pose estimations. In [72], the elevation map is tightly coupled to the robot’s motion, i.e. a robot-centric approach.

There are three coordinate frames: the inertial frame $I$, which is attached to stationary terrain; the sensor frame $S$, which is attached to a range sensor and the map frame $M$. The transformation between frame $I$ and $S$ is $r_{IS}$ and $C_{IS}$, which are translation and rotation respectively. Using a stereo visual odometry from Section 4.2 the robot’s state estimation is obtained and described by the pose covariance matrix $\Sigma_P$. The map frame $M$ is obtained from the sensor frame $S$ by rotational and translational transformations $C_{SM}$, $r_{SM}$ respectively. To express the elevation map in the robot-centric view, the map frame $M$ and the inertial frame $I$ are aligned by $z$-axis. The yaw angle $\psi$ between frames $I$ and $M$ and the yaw angle between frames $I$ and $S$ are equal.

The height measurements are expressed in pointcloud type data. This allows a map cell $(x, y)$ updated with a new corresponding height measurement $\tilde{p}$. Measurement $\tilde{p}$ is Gaussian with mean $p$ and variance $\sigma_p^2$. To transform a point to the map frame $M$ from the sensor frame $S$ with corresponding height measurement $p$, the
following formula is employed:

$$p = P(C_{SM}^T M(q))_{SRSP} - M_{SM}$$  \hspace{1cm} (4.8)$$

where \( q \) is the unit quaternion between map frame and sensor frame \( C_{SM} \). \( P = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \) maps the 3D measurement to the scalar height measurement \( p \) in the map frame \( M \). The Jacobians of the sensor measurement \( J_S \) and sensor frame rotation \( J_q \) are defined as follows:

$$J_S = \frac{\partial p}{\partial S_{SRSP}} = PC_{SM}^T(q)$$

$$J_q = \frac{\partial p}{\partial q} = PC_{SM}^T S_{SRSP}^\times$$ \hspace{1cm} (4.9)$$

where the superscript \( \times \) denotes a skew symmetric matrix of a vector. The error propagation for the variance \( \sigma_p^2 \) is given as:

$$\sigma_p^2 = J_S \Sigma_S J_S^T + J_q \Sigma_{P,q} J_q^T$$ \hspace{1cm} (4.10)$$

where \( \Sigma_S \) denotes the covariance matrix of the range sensor model and \( \Sigma_{P,q} \) denotes the covariance matrix of the sensor rotation. Using Kalman filter as in [70], the estimation of the elevation map \( (\hat{h}, \sigma_h^2) \) is updated with new height measurements \( (p, \sigma_p^2) \) as follows:

$$\hat{h}^+ = \frac{\sigma_p^2 \hat{h}^- + \sigma_h^2 - \hat{p}}{\sigma_p^2 + \sigma_h^2}$$

$$\sigma_h^2 = \frac{\sigma_h^2 - \sigma_p^2}{\sigma_h^2 + \sigma_p^2}$$ \hspace{1cm} (4.11)$$

where the \(-\) and \(+\) superscripts denote estimation before and after an update, re-
spectively.

Since the elevation map is robot-centric, it is desirable to estimate the terrain in the moving map frame $M$. When the pose of the robot changes, the height information $\hat{h}$ and its variance $\hat{\sigma}^2_h$ of each map cell are updated accordingly. This is computationally expensive and not always needed. To avoid this, instead of computing the variance $\hat{\sigma}^2_h$ of each map cell, the variance in the $x-$ and $y-$ horizontal directions, $\hat{\sigma}^2_x$ and $\hat{\sigma}^2_y$, are evaluated. Consider the mapping from a fixed point in the inertial frame, $r_{IP}$, to its corresponding position in the elevation map $r_{SM}$:

$$
M_r^k_{MP} = C_{SM}^T s^k_{SP} - M_r^k_{SM}
$$

where $k$ is the time step. The Jacobians of the sensor frame translation and rotation, $J_r$ and $J_q$, are derived as follows:

$$
J_r = \frac{\partial M_r^k_{MP}}{\partial r^k_{IS}} = -C_{SM}^T C_{IS}^T (q^k)
$$

$$
J_q = \frac{\partial M_r^k_{MP}}{\partial q^k} = -C_{SM}^T C_{IS}^T (q^k)(r_{IP} - r_{IS})^X
$$

The map for time $k$ with variance $\hat{\sigma}^2_x$ and $\hat{\sigma}^2_y$ is updated as follows:

$$
\begin{bmatrix}
\hat{\sigma}^2_x \\
\hat{\sigma}^2_y \\
\hat{\sigma}^2_h
\end{bmatrix}
= \begin{bmatrix}
\hat{\sigma}^2_x \\
\hat{\sigma}^2_y \\
\hat{\sigma}^2_h
\end{bmatrix} + \text{diag} \left(J_r (\Sigma_{r,r}^k - \Sigma_{r,r}^{k-1}) J_r^T + J_q (\Sigma_{q,q}^k - \Sigma_{q,q}^{k-1}) J_q^T\right)
$$

The process of transforming the elevation map from $(\hat{h}, \hat{\sigma}^2_h, \hat{\sigma}^2_x, \hat{\sigma}^2_y)$ representation to
\((\hat{h}, \hat{\sigma}_h^2)\) is described as follows:

\[
\hat{h} = \frac{\sum_n w_n \hat{h}_n}{\sum_n w_n}
\]

\[
\hat{\sigma}_h^2 = \frac{\sum_n w_n (\hat{\sigma}_{h,n}^2 + \hat{h}_n^2)}{\sum_n w_n} - \hat{h}^2
\] (4.15)

where the weight \(w_n\) for a cell \(n\) is derived as:

\[
w_n = \left( \Phi_x(d_x + \frac{r}{2}) - \Phi_x(d_x - \frac{r}{2}) \right) \left( \Phi_y(d_y + \frac{r}{2}) - \Phi_y(d_y - \frac{r}{2}) \right)
\] (4.16)

where \(\Phi_x\) and \(\Phi_y\) are the cumulative normal distribution with covariance \(\sigma_x, \sigma_y\) respectively; \(d_x\) and \(d_y\) are the distance of cell \(n\) to the cell being updated; \(r\) is the length of the cell side.
Chapter 5

Experimental Results

5.1 Field test for nondestructive evaluation sensors

For GPR and ER sensors, a field test has been conducted. These two sensors were used to collect data manually of the Pleasant Valley Bridge on Highway 580, Reno, Nevada, USA. The survey areas are the two slow lanes of the bridge. The data collected by the GPR sensor was processed by the RADAN™ software. This software is accompanied with the GPR sensor kit. The software identifies the rebars’ locations inside the concrete surveyed areas. This information is utilized to build the condition map of the bridge, as presented in Fig. 5.2. Judging from the condition map, it is safe to evaluate that the bridge is still in good condition. There are four color codes to indicate the deteriorate level, ranging from blue (good condition) to green (fairly good condition), to orange (bad condition) and to red (severe condition). Unfortunately, recent condition assessments of the bridge are not publicly available, it is hard to compare and evaluate GPR sensor’s working quality. Even though ER sen-
Figure 5.1: Pleasant Valley Bridge on Highway 580, Reno, Nevada with the surveyed areas marked by yellow lines (image taken from Google Map).

Figure 5.2: Condition map of the surveyed areas on Pleasant Valley Bridge. (a) - Northbound part of the bridge; (b) - Southbound part of the bridge.
sors were also deployed (Fig. 2.1), the concrete bridge deck surface was protected by coating overlays. Hence, data from ER sensors were inapplicable to bridge condition evaluation.

5.2 Localization with GPS

The robot was manually driven to collect wheelodometry and GPS data. The localization algorithm discussed in Section 4.1 is then run on these data. The results are presented in following figures.

Figure 5.3: (Left sub-figure) Performance of GPS+IMU-EKF-based localization with robot localization package. The red dots are GPS signal and blue dots are output of EKF localization. (Right-sub figure) Zoom-in/Close-look at one location: the EKF outperforms the GPS alone since it outputs smoother results.

5.3 Visual inertial odometry with stereo camera

The camera-IMU calibration process was carried out followed the descriptions in [75, 77]. In this work, the Kalibr package [78] was utilized to automate the calibration
Figure 5.4: GPS-based localization with EKF. The robot has been driven along a 34 ft-by-14 ft square 10 times. The red plus sign denotes GPS location. The blue dot denotes robot’s location from wheelodometry. The green dot denotes the localization algorithm’s output.

The calibration results are presented as follows. The extrinsic transformation process.
from left camera frame to IMU frame is:

\[
\begin{bmatrix}
0.99994075 & 0.00259196 & 0.0105729 & -0.07692214 \\
-0.01063558 & 0.02549736 & 0.99961831 & 0.01076284 \\
0.00232139 & -0.99967153 & 0.02552342 & -0.02766903 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (5.1)

The extrinsic transformation from right camera frame to IMU frame is:

\[
\begin{bmatrix}
0.99999025 & 0.00278813 & 0.00342473 & 0.04309024 \\
-0.00349755 & 0.02654717 & 0.99964144 & 0.01090429 \\
0.002692621 & -0.99964367 & 0.02655666 & -0.02746205 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (5.2)

The extrinsic transformation from left camera to right camera is:

\[
\begin{bmatrix}
0.99997445 & -0.0019256 & 0.0071454 & -0.12001128 \\
0.00018505 & 0.99999943 & 0.00105219 & -0.00013146 \\
-0.0071456 & -0.00105084 & 0.99997392 & -0.00055791 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (5.3)

5.4 Robot deployment and Elevation map inside a garage building

The robot has been deployed inside a garage building. Using AsusXtion RGBD camera, the robot uses algorithms discussed in Section 4.3 to build an elevation map. The results are presented below.
Figure 5.5: Static calibration result for left camera.
Figure 5.6: Static calibration result for right camera.
Figure 5.7: Acceleration error from IMU.
Figure 5.8: Angular velocity error from IMU.

Figure 5.9: Feature tracking performance from stereo visual inertial sensor.
Figure 5.10: The robot is moving inside a garage building.

Figure 5.11: Elevation map inside the garage parking. A car appears in front of the robot and the map is being updated accordingly.
Figure 5.12: Elevation map is being built when the robot moves along.
Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis presented a robotic system for civil infrastructure inspection. The system was designed and implemented with focusing on multiple sensors fusion, including nondestructive evaluation sensors, GPS, cameras, IMU. Detailed discussion about each sensor was provided. Several techniques for localization were provided with emphasizing on complementaries among sensors. Some experimental results were presented.

6.2 Future Work

Future work may include exhaustive field tests to evaluate the robot performance. It is desirable to combine information from NDE sensors with navigation sensors to generate intelligent inspection plans. One possible approach is: the robot first relies on visual data to perform inspection. While reading data from the GPR sensor, the
robot might decide to deploy ER sensor at some specific locations, where GPR shows bad data reading, which may indicate high level of deterioration. This approach might reduce the total inspection time by performing deep inspection only where it is needed. Additionally, several inspection robot could be employed to form a multi-robot inspection network to quickly cover any vast area \[79,80\].
Bibliography


[54] M. Nuttin, T. Tuytelaars, and L. Van Gool, “Omnidirectional vision based topo-

2006 IEEE Computer Society Conference on Computer Vision and Pattern

[56] G. Klein and D. Murray, “Parallel tracking and mapping for small ar

[57] A. J. Davison, “Real-time simultaneous localisation and mapping with a single

[58] E. D. Dickmanns and B. D. Mysliwetz, “Recursive 3-d road and relative ego-state
recognition,” *IEEE Transactions on pattern analysis and machine intelligence*,

[59] G. Qian, R. Chellappa, and Q. Zheng, “Robust structure from motion estimation

navigation with image-based motion estimation,” in *Robotics and Automation,
2002. Proceedings. ICRA’02. IEEE International Conference on*, vol. 4. IEEE,

measurements,” in *Applications of Computer Vision, 2002.(WACV 2002). Pro-


