# State Estimation and Localization for ROV-Based Reactor Pressure Vessel Inspection Using a Pan-Tilt-Zoom Camera

Timothy Edward Lee
August 25, 2017
CMU-RI-TR-17-48


The Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania

**Thesis Committee:**
Nathan Michael, *Chair*
Stephen Nuske
Eric Westman

*Submitted in partial fulfillment of the requirements for the degree of Master of Science in Robotics.*

**Abstract**

A vision-based state estimation framework is proposed to localize a remotely operated vehicle (ROV) that is used to inspect a nuclear reactor pressure vessel. Using an extended Kalman filter, the framework leverages an external, pan-tilt-zoom (PTZ) camera as the primary sensing modality. In addition to the ROV and camera states, the framework also estimates a map of the reactor lower core from a prior. Both simulation and camera experiments are conducted to establish the correctness of the framework. Camera experiments (consisting of subscale and platform tests) validate the framework in terms of accuracy and robustness to environmental image degradation caused by speckling and color attenuation. Subscale mockup experiments highlight estimation consistency as compared to ground truth despite visually degraded operated conditions. Full-scale platform experiments are conducted using the actual inspection system in a dry setting. In this case, the ROV achieves a lower state uncertainty as compared to subscale mockup evaluation. For both subscale and full-scale experiments, the state uncertainty is robust to environmental image degradation effects. To further increase the efficiency and utility of the framework, we formulate and evaluate an online initialization methodology that bootstraps the framework *in situ* by calibrating the camera from the reactor structure. A Monte Carlo simulation shows that the initialization methodology is not sensitive to camera mounting position and orientation if the camera is placed near the center of the vessel. This framework sets the foundation for improving not only ROV-based inspection through automation, but also visual inspection with the PTZ camera through mosaics and dense reconstructions of the reactor lower core.

# Acknowledgements

To my advisor, Prof. Nathan Michael —

To my colleague, “Master Builder” Curtis Boirum —

To Dr. Stephen Nuske and Eric Westman —

To my fellow RASL and Robotics Institute colleagues —

To my wife, Karin —

To my parents and my sister —

Please accept my deepest gratitude for your support.

“...if we don’t discover what we want from ourselves and what we stand for, we will live passively and unfulfilled.” — Bill Watterson

We gratefully acknowledge support from Westinghouse Electric Company, LLC.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Nuclear reactors require periodic maintenance and inspection to ensure continued and safe operations, and automation of inspection tasks through robotics can minimize the duration of reactor inspection outages in a safe and cost-effective manner. State-of-the-art reactor pressure vessel inspection leverages remotely operated vehicles (ROVs), which are teleoperated, submersible robots. In order to inspect the reactor core, ROVs are subjected to extreme radiation exposure that destroys sensitive, commercial-grade electronics that would otherise enable localization. The lack of viable radiation-hardened sensors motivates the development of novel localization methodologies that overcome these limitations to enable efficient, autonomous inspection campaigns. This work explores the use of an external pan-tilt-zoom (PTZ) camera to address the challenges of sensing and localization in this hazardous, irradiated environment. By decoupling the sensor from the platform, the PTZ camera provides high optical resolution imagery that can be used for ROV localization while limiting radiation exposure, thereby providing a viable path to the automation of reactor core inspection.

In this work, we propose a vision-based state estimation and localization framework to enable submersible robots to conduct inspection of nuclear reactor pressure vessels. The framework is formulated as an extended Kalman filter (EKF) that is robust to sensor degradation and image corruption that may occur due to environmental effects, such as radiation and color attenuation. The proposed framework relies on a PTZ camera, fixed with respect to the vessel frame, that is autonomously controlled. To model the reactor vessel, we propose the use of a planar map that concisely represents the vessel geometry as a series of planes and landmarks. The map is assumed to be known in advance with limited uncertainty arising from differences between engineering blueprints and construction. The map is estimated in the state to enable corrections based on projections of vessel landmarks (points and lines) in the camera image space.

## 1.1 Background Information

The reactor pressure vessel of a nuclear power plant over time becomes embrittled due to neutron irradiation that is a consequence of nuclear fission. Specifically, it is the lower core of the reactor (Fig. 1.1–1.2) that is bombarded with radiation during operation. Embrittlement of the lower core weakens the vessel, leading to cracks, corrosion, and the formation of cavities [4]. These are detrimental to the structural integrity of the vessel. It is therefore motivated to periodically inspect the vessel to ensure its structural integrity for the power plant to continue to operate safely.

In order to inspect the lower core of the reactor, the reactor must be shut down with the fuel rods removed. These inspections may occur during periodic refueling outages, in which the fuel rod arrays are replaced, or in addition to normal refueling

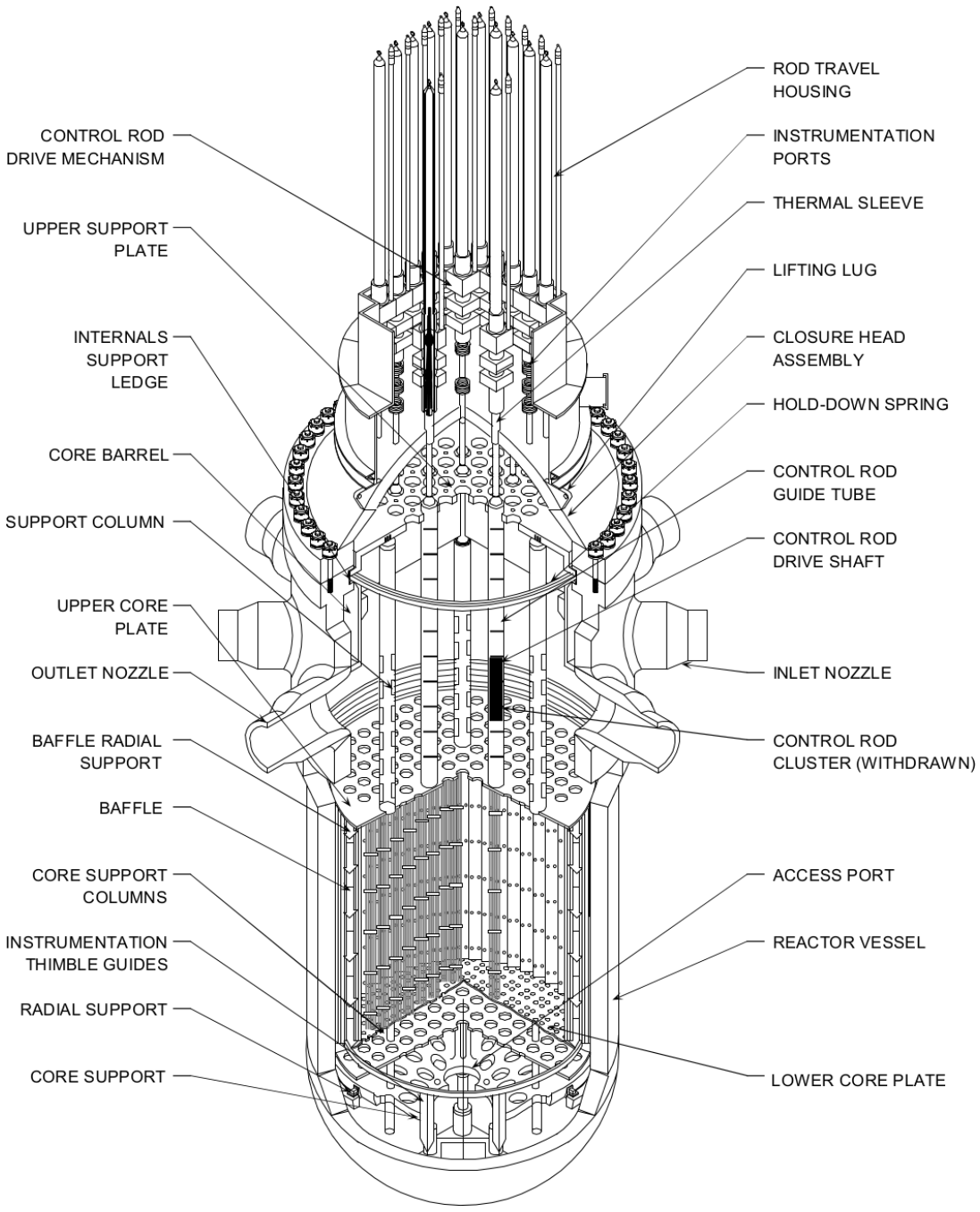



Figure 1.1: Diagram of the class of nuclear reactors that are considered in this work (pressurized water reactors) [1]. Note the shape of the lower core, composed of baffles, that is exposed when the fuel rods are removed from the reactor.

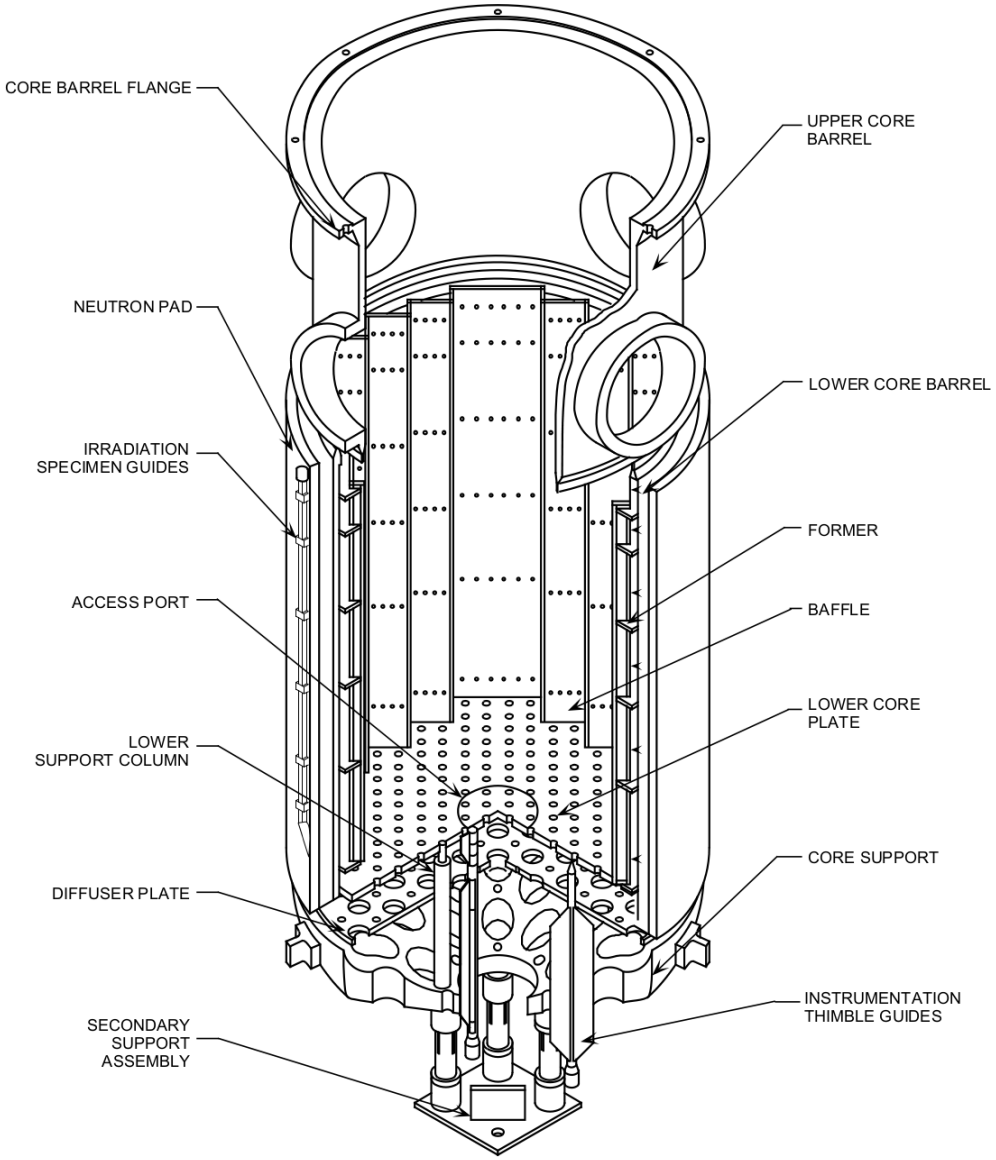



Figure 1.2: Diagram of the lower internals of a PWR nuclear reactor [2]. The lower core barrel, baffles, and core plate are of particular interest for this work.

outages if the damage is warranted. According the U.S. Energy Information Agency, refueling outages typically occur every 18–24 months with an average duration of six weeks [5].

### 1.1.1 Robotic Inspection of Nuclear Reactor Pressure Vessels

Current visual inspection techniques for the lower core of a nuclear reactor include spanning the reactor with a bridge and submerging telescopic cameras, endoscopes, or fibre-scopes that are mounted on long poles into the reactor to detect structural flaws [6, 7]. These existing techniques, while adequate for inspection, have room for improvement in two primary areas: 1) efficiency and 2) reduced radiation dosage to personnel. The inspection processes are slow, contributing to higher costs for nuclear energy while exposing inspection personnel to radiation.

To improve the cost-effectiveness of nuclear energy, it is prudent to consider technologies that can address both of these deficiencies. Current research in technologies for improved inspection of nuclear power plants include methods for perceiving structural components, either with cameras or ultrasonic sensors [8–10]. Additionally, the use of robots for improving the inspection of nuclear reactors is an active area of interest. Robotic inspection could limit radiation dosage to personnel, and inspections could be more efficient through automation. Both of these advantages would yield decreased nuclear energy costs. Robotic technologies may also provide assurance towards the efficacy of existing inspection methods when working side-by-side with inspection personnel. As an example of such a robotic inspection system, consider the submersible robot that is used to inspect a nuclear reactor pressure vessel, shown in Fig. 1.3. Other representative robotic systems used in the nuclear energy domain

include the Stinger robot[1] and the SUSI (Submarine System for Inspection) robot[2].

### 1.1.2 Impact of Radiation on Sensor Modalities

As the advantages of robotic inspection of nuclear reactor are clear, it is therefore necessary in the design of a robot to utilize sensing modalities that enable accurate and robust performance in the presence of radiation.

Specifically, the types of inspections which we seek to automate are those conducted in close proximity to the reactor core barrel, which imposes harsh environmental conditions in which the robot must operate. For these inspections, the reactor is offline with the fuel removed, exposing the lower core of the reactor for inspection. Although no fuel is present during the inspection, the core barrel emits an intense radiation field due to neutron activation of the structural materials. This field, which may reach up to 30 kRad/hr at the lower core baffle plates [11], excludes the use of

[1]GE Hitachi Nuclear Energy Company.
[2]AREVA GmbH.

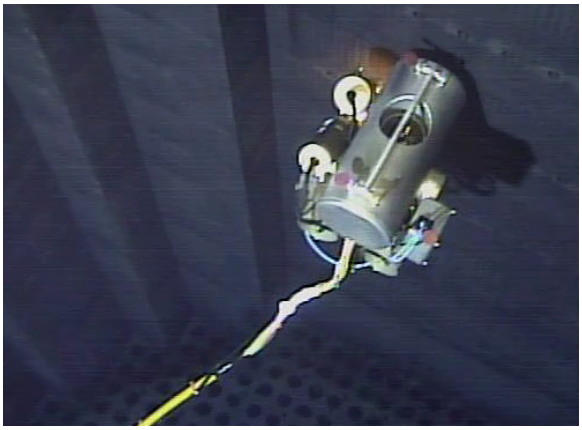


Figure 1.3: This submersible robot is used to inspect reactor pressure vessels. Note the planar structure of the vessel and the geometric features located on the walls and floor. The robot is equipped with three red fiducial markers that are used for pose estimation. A tether is used to transmit robot control signals from the control station.

localization sensors with sensitive electronics. Indeed, radiation-sensitive electronics can degrade and fail with exposure of a few krad [12], well below the expected radiation exposure in this setting. For this reason, the radiation exposure that must be withstood by a robot in this setting restricts the use of all but the most hardened electronic devices and sensors. This generally excludes the use of sensing modalities that would typically be installed on a submersible robot [13], such as depth sensors, Doppler velocity log (DVL), sonar, and inertial measurement units (IMU), as such sensing modalities are not commonly available with sufficient radiation hardening. In addition, the underwater setting precludes the practical use of projective light sensing [14] and GPS.

However, the underwater setting is free of turbidity, so the use of vision-based perception is viable. Specifically, a pan-tilt-zoom (PTZ) camera is especially useful in this setting. A PTZ camera has servos allowing it to rotate in two degrees of freedom (pan, tilt) and has variable zoom. The radiation field from the core falls off by about an order of magnitude for every 16 inches of water, so a PTZ camera would only need to achieve a nominal level of radiation hardening. Furthermore, the geometry of the reactor core is amenable to using this sensing modality, as the camera can always observe the robot without needing to reposition the camera. Lastly, the variable zoom allows for a high optical resolution image no matter the position of the ROV within the vessel. For these reasons, the PTZ camera is the primary sensing modality used to localize the robot. We will now refer to the *inspection system*, which includes both the robot and the PTZ camera.

## 1.2 Previous Work

The use of vision for underwater robots has been studied both in laboratory experiments and in deployed field robots. Although submersible robots can utilize a range of sensing modalities [13], radiation exposure from the nuclear reactor restricts us to considering systems where vision is the primary sensing modality. A visual SLAM formulation with pose-graph optimization was utilized to construct a texture-mapped, three-dimensional model of a ship hull for inspection purposes [15]. An EKF state estimation formulation that includes vision and inertial measurements was found to be successful in underwater navigation of a submersible robot [16]. Another study demonstrated a localization solution for a AUV using acoustic sensors and visual odometry [17]. Our use of structural landmarks is similar to previous work in localizing an underwater robot in a structured environment using only visual perception (an onboard camera) [18], but our study differs in that robot localization is achieved through a fixed, external camera.

Deployed field inspection robots that utilize vision have conducted subsurface bridge inspection [19] and ship hull inspection [20], with sonar imaging as the primary inspection modality in these cases. In the domain of nuclear reactor inspection, the use of cameras and robotics for inspection has been studied [8, 9]. A previous study estimated the $x$- and $y$-position and yaw angle of a submersible robot within a reactor vessel by observing eight LEDs located on the vehicle with an external camera (primarily using a depth sensor for the $z$-position) [21]; our work differs by estimating both the robot and camera pose with six degrees of freedom.

Regarding PTZ cameras, we note that the movement of a small unmanned system with a pan-tilt camera has been estimated using an EKF [22]. This study estimates the projection of the system in the camera image space, not in three dimensions as

our framework does. Jain and Neumann [23] employ an EKF to estimate the pose and focal length of a PTZ camera.

## 1.3 Thesis Problem: ROV Localization

Given the motivation for ROV-based inspection of nuclear reactors, it is therefore critical to formulate and test a localization framework that leverages the sensors that could exist in this harsh environment. State-of-the-art robot localization typically leverages sensors that are mounted on the robot, such as monocular or RGB-D camera and IMU. However, the use of these common sensors are greatly restricted due to radiation exposure. On the other hand, methods for pan-tilt-zoom cameras are typically used for inference within the image space, not in three dimensions as is required for our work. As existing state-of-the-art methods are not satisfactory to address our problem, we therefore propose a state estimation and localization framework that enables ROV localization within the lower core of a nuclear reactor pressure vessel using a pan-tilt-zoom camera. We demonstrate our framework using the ROV example shown in Fig. 1.3.

## 1.4 Thesis Contributions

We identify the following contributions of the work:

1. A state estimation framework that enables three-dimensional pose estimation and localization of the submersible robot with respect to the reactor vessel. To accomplish this, the framework localizes the three-dimensional pose of the pan-tilt-zoom camera and estimates the focal length. The framework also estimates the reactor map, accounting for limited uncertainty in the prior for the reactor

geometry.

2. The planar map representation of the reactor vessel. This representation enables the mapping of the entire class of lower reactor core geometries with a few limited assumptions.

3. Demonstrations of accuracy and robustness of the framework to speckling, a known failure mode that occurs due to the exposure of cameras to radiation. Results will also demonstrate the capability for online initialization.

These contributions demonstrate that we have proposed a system that works not only in principle, but also in the practice of actual inspection of nuclear reactors.

## 1.5 Thesis Outline

This thesis document is outlined as follows. In Chapter 2, we will present the various models and methods that describe the inspection system. Topics include methods for a pan-tilt-zoom (PTZ) camera, such as rotation, zoom, and projection (Sec. 2.2); the ROV description and marker placement (Sec. 2.3); and a geometric representation of the reactor pressure vessel and how its projections arise in the camera image space (Sec. 2.4). These models and methods provide the foundation for the state estimation of the inspection system using an extended Kalman filter, which is described in Chapter 3. This chapter describes not only the framework itself, but a methodology for online initialization (Sec. 3.6), leveraging camera calibration techniques that are amenable to automation. In Chapter 4, we demonstrate the accuracy and robustness of the EKF in simulation and camera experiments (including with the inspection system). In closing, Chapter 5 summarizes the contributions of the work. In this chapter, we also identify how this work sets the foundation for several high-impact research

paths that can transform existing inspection processes, including: 1) automation by enabling closed-loop control of the ROV (Sec. 5.3.1), and 2) dense reconstruction of the reactor pressure vessel through mosaics that are assembled from PTZ camera images (Sec. 5.3.2).

# Chapter 2

# System Models and Methods

The mathematical formulations and representations of the components that compose the inspection system are detailed in this chapter. A model for speckling, which is radiation-induced chromatic image corruption, is also presented. These models are utilized by the state estimation framework (Chapter 3) to localize the system with respect to the reactor vessel.

## 2.1 System Coordinate Frames

The robotic inspection system consists of a submerged PTZ camera that monitors a ROV operating in a reactor pressure vessel. The robot is equipped with three fiducial markers. Figure 2.1 illustrates the system and depicts the three primary reference frames:

1. the *body frame* $\{B\}$, located at the robot center of mass;

2. the *external camera frame* $\{E\}$ at the optical center of the PTZ camera; and

3. the inertial *world frame* $\{W\}$, which is the reference frame for the vessel map.

Note that the world frame $\{W\}$ is defined with $z$-axis pointed downward, towards the bottom of the reactor. This is consistent with typical reference frames for underwater vehicles [24, 25], where positive $z$ refers to increased depth. Although the external camera frame $\{E\}$ will change based on the PTZ camera pan and tilt angle, the mounting of the camera itself is fixed with respect to the world frame $\{W\}$. This fixed orientation is the *scene frame* $\{S\}$, an intermediate frame that is useful in the kinematic chain between the camera frame and the world frame.

## 2.2 Pan-Tilt-Zoom Camera Models

This section outlines the mathematical model for the pan-tilt-zoom (PTZ) camera that is used in the inspection system (Fig. 2.2). The PTZ camera is used to monitor the robot and vessel during inspection. It is mounted external to the robot and is fixed with respect to the world frame $\{W\}$. Specifically, the camera would be mounted at the end of a pole and submerged into the reactor, with the other end of the pole rigidly secured to a bridge than spans the reactor vessel. Submerging the camera requires a waterproof housing, but this is the best option to utilize the images for localization. The alternative — mounting the camera above the water surface of the vessel — is not viable, as it introduces greater optical challenges (imaging an underwater scene from above the water). In order to obtain high-quality footage of the reactor for inspection purposes, we assume the use of a high-quality PTZ camera, such as the Axis V5915 camera. An advantage of using such a camera is that the camera has limited lens distortion.

The camera is controlled via visual servoing to keep the ROV in view. A pro-

portional controller that operates in the image space is used which uses the centroid and distribution of the markers to determine when to pan, tilt, or increase the zoom setting of the camera.

The camera images are used for camera-to-robot localization (using robot marker

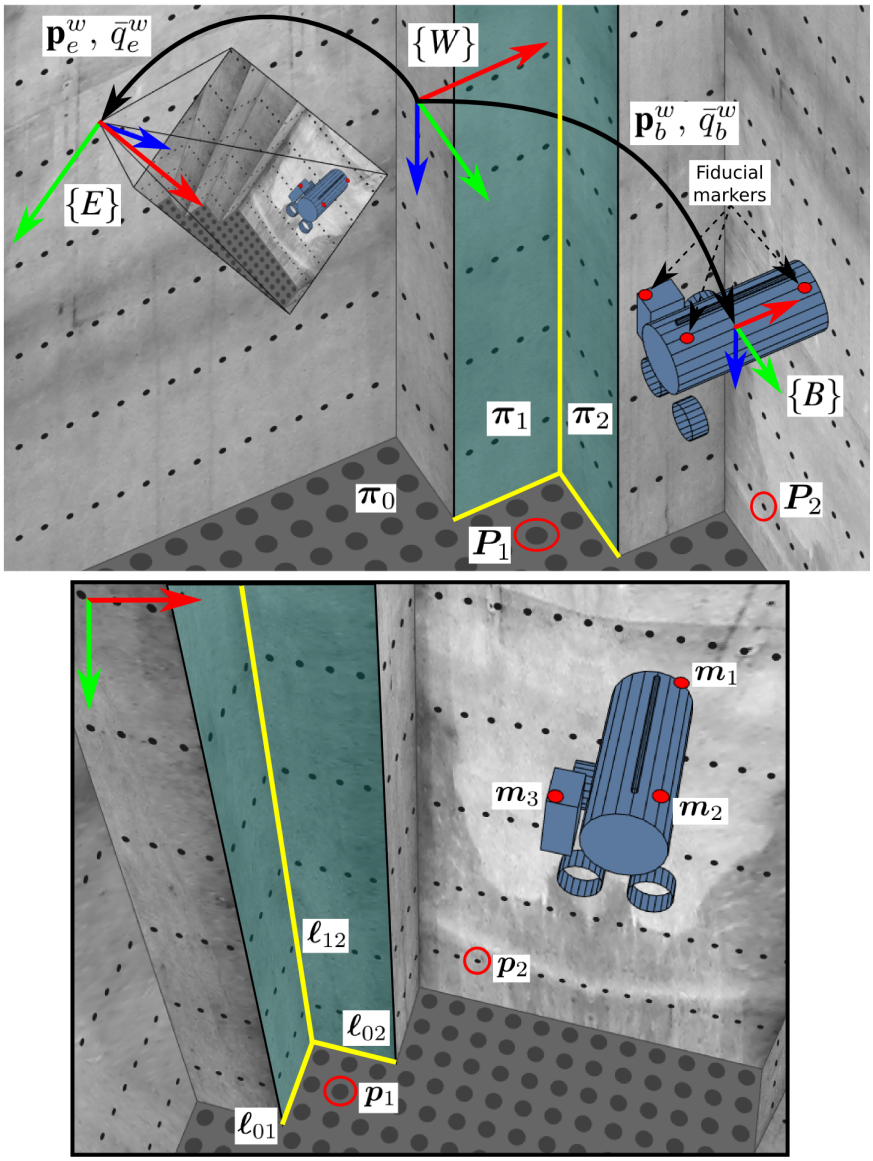
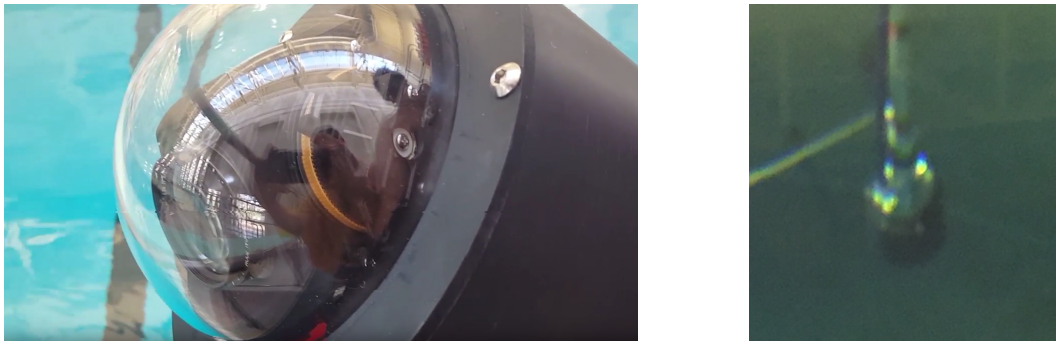



Figure 2.1: System representation and landmarks: (left) in three dimensions; (right) in the two dimensional image space of the PTZ camera. The intersections of vessel planes $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_j$ yield Plücker lines that project as lines $\boldsymbol{\ell}_{ij}$. Similarly, three-dimensional points $\boldsymbol{P}_i$ project to two-dimensional points $\boldsymbol{p}_i$ in the image space.



(a) (b)

Figure 2.2: The PTZ camera used for the inspection system. The camera is contained within a waterproof enclosure that has a spherical housing. (a) Camera close-up. (b) Camera is attached to a pole and submerged. The tether shown here is a preventative measure for testing and is not part of the deployable system.

projections) and camera-to-vessel (world) localization (using vessel landmark projections). Additionally, we require the PTZ camera to also report its state, including pan and tilt angle from some base orientation, referred to as the "scene orientation," and zoom setting. The scene orientation is determined by the camera mounting. The pan and tilt angles are determined from encoders or resolvers that exist on the camera motors that control the pan and tilt angles. The zoom setting provides an indication of the current image magnification as compared to the base zoom level. This corresponds to some zoom level expressed as a multiple of the base zoom (e.g., "2x" zoom). The underlying metric focal length is not observable because the image sensor dimensions are not known, but the variation in zoom setting exposes how the focal length intrinsics change (c.f. Sec. 3.4).

### 2.2.1 Camera Projection Model

We model the PTZ camera as a pinhole projection camera. In this classical projection model, a three-dimensional point $\boldsymbol{P} \in \mathbb{R}^3$ expressed in homogeneous coordinates is mapped from $\mathbb{P}^3 \mapsto \mathbb{P}^2$ to produce a two-dimensional point projection $\boldsymbol{p} \in \mathbb{R}^2$ in

homogeneous coordinates [26]:

$$\widetilde{\boldsymbol{p}} \sim K[\, R \,|\, \mathbf{t} \,]\widetilde{\boldsymbol{P}} \tag{2.1}$$

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.2}$$

where $\widetilde{\boldsymbol{P}} \sim [\, \boldsymbol{P}^{\mathrm{T}}, 1 \,]^{\mathrm{T}}$ is the three-dimensional point in homogeneous coordinates, $\widetilde{\boldsymbol{p}} \sim [\, \boldsymbol{p}^{\mathrm{T}}, 1 \,]^{\mathrm{T}}$ is the two-dimensional point in homogeneous coordinates, and the $\sim$ operator indicates that the transformation is equal up to scale. The rotation matrix $R \in SO(3)$ and vector $\mathbf{t} \in \mathbb{R}^3$ transforms the frame in which $\boldsymbol{P}$ is defined to the camera frame.

The intrinsics matrix $K$ describes the camera projection model and defines the field of view of the camera. The effect of increased zoom causes the focal length components $f_x$ and $f_y$ to increases, resulting in a narrower field of view. In the same manner as Wu and Radke, we assume that the principal point defined by $c_x$ and $c_y$ is constant with respect to zoom [27]. The center of zoom is also assumed to be congruent with the center of projection (the principal point).

In this work, we also consider the projection of three-dimensional lines in the image space. Projecting a three-dimensional line in the scene yields (in general) a two-dimensional line in the image. We use the Plücker representation of three-dimensional lines $\boldsymbol{\mathcal{L}} \in \mathbb{P}^5$, which project to lines in the image space $\boldsymbol{\ell} \in \mathbb{P}^2$:

$$\boldsymbol{\ell} \sim \mathcal{K}\mathcal{D}\boldsymbol{\mathcal{L}} \tag{2.3}$$

The matrix $\mathcal{K} \in \mathbb{R}^{3\times 6}$ is the perspective projection matrix for Plücker lines, and

$\mathcal{D} \in \mathbb{R}^{6\times6}$ is the rigid displacement matrix maps between the Plücker line frame and the camera frame [28]:

$$\mathcal{K} = \begin{bmatrix} \det(K)\, K^{-\mathrm{T}} & 0_{3\times3} \end{bmatrix} \tag{2.4}$$

$$\mathcal{D} = \begin{bmatrix} R & [\mathbf{t}]_{\times} R \\ 0_{3\times3} & R \end{bmatrix} \tag{2.5}$$

In this way, the mapping of points from $\mathbb{P}^3 \mapsto \mathbb{P}^2$ caused by $K$ and $[\, R \,|\, \mathbf{t} \,]$ is analogous to the mapping of lines from $\mathbb{P}^5 \mapsto \mathbb{P}^2$ caused by $\mathcal{K}$ and $\mathcal{D}$.

### 2.2.2 Camera Pose Representation

Unlike a monocular camera that would exist on a robot, a PTZ camera has only two degrees of freedom: pan and tilt. The PTZ camera position with respect to the world $\mathbf{p}_e^w \in \mathbb{R}^3$ is fixed. We also assume that the center of projection is congruent with the camera position.

To consider the action of pan and tilt on the orientation of the camera frame $\{E\}$, it is useful to introduce an additional coordinate frame — the scene frame $\{S\}$ — that is defined as the camera frame orientation with respect to the world frame $\{W\}$ when the pan angle $\theta \in \mathbb{R}$ and tilt angle $\phi \in \mathbb{R}$ are zero (so the scene orientation and camera frame orientation are congruent at zero pan and tilt). The pan and tilt angles are measured with respect to their respective axis: the pan axis $\bar{\Theta} \in \mathbb{R}^3$ and tilt axis $\bar{\Phi} \in \mathbb{R}^3$, which are both unit vectors. The pan and tilt axes are defined with respect to the frame in which they are constant. For our work, the pan axis $\bar{\Theta} = [\, 0,\, 1,\, 0 \,]^{\mathrm{T}}$ is static with respect to the scene frame $\{S\}$. The tilt axis $\bar{\Phi} = [\, 1,\, 0,\, 0 \,]^{\mathrm{T}}$ is static with respect to the camera frame $\{E\}$. Assuming a scene orientation that has zero roll and pitch, a tilt angle of 0° corresponds to the camera $z$-axis aligned within the

$xy$-plane of $\{W\}$. With a tilt angle of $-90°$, the camera $z$-axis is parallel to the $z$-axis of $\{W\}$, viewing the bottom of the reactor.

The rotation matrix $R_e^w$ that describes the orientation of the camera frame $\{E\}$ with respect to the world frame $\{W\}$ can be described through the kinematic chain of successive rotations defined by the scene orientation $R_s$, pan rotation $R_\theta$, and tilt rotation $R_\phi$ as described by Collins and Tsin [29]:

$$R_e^w = R(R_s, \theta, \phi) = R_\phi R_\theta R_s \tag{2.6}$$

Collins and Tsin specify an optional $R_m$ following $R_\phi$ to account for the physical mounting of the camera on the pan-tilt head. For our system, the mounting of the camera with respect to the world is already accounted for by the scene frame $\{S\}$, and so this additional rotation $R_m$ is not necessary.

The rotation matrices $R_\theta$ and $R_\phi$ are constructed from the axis-angle representation and Rodrigues' rotation formula:

$$R_\theta = I_{3\times3} + (\sin\theta)\left[R_s\bar{\Theta}\right]_\times + (1-\cos\theta)\left[R_s\bar{\Theta}\right]_\times^2 \tag{2.7}$$

$$R_\phi = I_{3\times3} + (\sin\phi)\left[R_\theta R_s\bar{\Phi}\right]_\times + (1-\cos\phi)\left[R_\theta R_s\bar{\Phi}\right]_\times^2 \tag{2.8}$$

From the above equations, it is apparent that the camera orientation is coupled between pan and tilt. The tilt axis is defined so that positive tilt yields a positive rotation about the $y$-axis of the external camera frame $\{E\}$. However, the pan axis is defined so that positive pan yields a positive rotation about the $x$-axis of the *scene* frame $\{S\}$. As the tilt angle is decreased, pan rotation tends to induce a smaller rotation about the $x$-axis of $\{E\}$ and a greater rotation about the $z$-axis of $\{E\}$. When the tilt angle is $-90°$, pan rotation yields a pure camera roll — a rotation

about the $z$-axis of $\{E\}$.

Going forward, we estimate the orientation of the camera frame $\{E\}$ with respect to $\{W\}$ using the unit quaternion representation of $\mathrm{SO}(3)$ — $\bar{q}_e^w \in \mathbb{H}$ — by converting to and from $R_e^w$ when needed.

### 2.2.3 Effect of Zoom on Camera Optics in Air and in Water

Given that the inspection system is submerged, it is prudent to consider the effect of underwater optics on the camera image, particularly under camera zoom. We first discuss zoom optics in air.

Consider a different parameterization of camera intrinsics using a singular focal length $f$ with a fixed aspect ratio $\tau$ [30]. In this manner, $f_x = \tau f$ and $f_y = f$. In this model, the effect of zoom is a multiplication factor $s \geq 1$ of the metric focal length $\mathfrak{f}$, where $s = 1$ corresponds to the base metric focal length when the camera is completely zoomed out. The horizontal and vertical field of views ($FOV_H$ and $FOV_V$) are calculated via trigonometry of the scaled metric focal length $\mathfrak{f}s$ and the horizontal and vertical dimensions of the image sensor ($h$ and $v$). Note that increasing $s$ yields a decrease in the field of view.

$$\tan\left(\frac{FOV_H(s)}{2}\right) = \frac{h}{2\mathfrak{f}s} \tag{2.9}$$

$$\tan\left(\frac{FOV_V(s)}{2}\right) = \frac{v}{2\mathfrak{f}s} \tag{2.10}$$

$$f_x(s) = \frac{c_x}{\tan\left(FOV_H(s)\,/\,2\right)} = \frac{2c_x\mathfrak{f}s}{h} \tag{2.11}$$

$$f_y(s) = \frac{c_y}{\tan\left(FOV_V(s)\,/\,2\right)} = \frac{2c_y\mathfrak{f}s}{v} \tag{2.12}$$

$$\tau = \frac{f_x(s)}{f_y(s)} = \frac{c_x v}{c_y h} \tag{2.13}$$

Therefore, in air, the camera aspect ratio $\tau$ is invariant to zoom (assuming a static principal point, as stated previously). However, we will now demonstrate that this outcome does not hold in general for cameras operating underwater, where the camera optics are subject to refractive effects [31].

Consider the same camera mounted directly behind a infinitesimally thin flat-plate housing and submerged in water. From Snell's Law, the amount of refraction depends on the angle of incidence of the ray that is traced from a scene point to the camera. Therefore, the following relationship holds for the horizontal and vertical field of view:

$$\sin\left(FOV_{air}\,/\,2\right) = n_w \sin\left(FOV_{water}\,/\,2\right) \tag{2.14}$$

where $n_w \approx 1.333$ is the refractive index of water.

Repeating the above derivation for this flat-plate housing shows that the aspect ratio is no longer static with respect to zoom while operating underwater:

$$f_{x,\,water}(s) = \frac{c_x}{\tan\left(FOV_{H,\,water}(s)\,/\,2\right)} = \frac{c_x}{h}\sqrt{\left(2\mathfrak{f}s\right)^2 n_w^2 + \left(n_w^2 - 1\right)h^2} \tag{2.15}$$

$$f_{y,\,water}(s) = \frac{c_y}{\tan\left(FOV_{V,\,water}(s)\,/\,2\right)} = \frac{c_y}{v}\sqrt{\left(2\mathfrak{f}s\right)^2 n_w^2 + \left(n_w^2 - 1\right)v^2} \tag{2.16}$$

$$\tau_{water}(s) = \frac{f_{x,\,water}(s)}{f_{y,\,water}(s)} = \frac{c_x v\sqrt{\left(2\mathfrak{f}s\right)^2 n_w^2 + \left(n_w^2 - 1\right)h^2}}{c_y h\sqrt{\left(2\mathfrak{f}s\right)^2 n_w^2 + \left(n_w^2 - 1\right)v^2}} \tag{2.17}$$

where we see if $n_w = 1$, then the optical variation of zoom simplifies to the case of operating in air.

The purpose of the above analysis is to demonstrate that the underwater optics with zoom are non-trivial and inherently coupled with the type of housing that is used. Although the PTZ camera used in the inspection system utilizes a spherical

housing, not a flat-plate housing, it nonetheless stands to reason that the optical model may differ as compared to operation in air.

Developing an optical model for general spherical housing geometries (including possible variation in thickness) is beyond the scope of this work, but it suffices to motivate the rationale for estimating both focal length components in the state estimation and localization framework. Rather than estimating one focal length and an assumed aspect ratio, we allow the focal length components to vary independently to account for the camera underwater optical model to change with zoom. Because the principal point describes the center of projection on the image sensor, we believe that assuming a static principal point for underwater optics is reasonable with the assumption that the center of projection and center of zoom are congruent.

## 2.3 ROV and Fiducial Marker Projection

The submersible robot (Fig. 1.3) is equipped with three fiducial markers to enable pose estimation from the marker projections in the camera image space. These projections ($\boldsymbol{m}_i = [\, u_i,\, v_i \,]^{\mathrm{T}}$, $i = \{1, 2, 3\}$) provide corrections between the external camera and the robot frames. The markers are detected using the K-means clustering algorithm [32]. The markers are associated in one of two ways: 1) based on their Euclidean distance to a known assignment in a previous frame (whereupon the initial assignment could be given manually); 2) automatically using the P3P algorithm and estimate of the camera state, which is used for online initialization of the ROV state (c.f. Section 3.6.2).

The position of the markers with respect to the body frame $\{B\}$ ($\boldsymbol{M}_i^b \in \mathbb{R}^3$, $i = \{1, 2, 3\}$) is static and known within some limited uncertainty $\sigma_M^2 \in \mathbb{R}$ in each marker dimension from measurements of the platform. The marker positions provide the

visual scale that is necessary to infer the three-dimensional pose of the robot from the marker projections. These projections arise in the external camera image space as follows:

$$\widetilde{\boldsymbol{m}}_i \sim K \left[ R_w^e \,|\, \mathbf{t}_w^e \right] T_b^w \, \widetilde{\boldsymbol{M}}_i^b \tag{2.18}$$

In this model, $T_b^w$ is the rigid body transformation matrix that relates points expressed in the body frame $\{B\}$ to the world frame $\{W\}$, calculated from the robot pose estimate ($\mathbf{p}_b^w \in \mathbb{R}^3$, $\bar{q}_b^w \in \mathbb{H}$). Similarly, the extrinsic calibration matrix $[\, R_w^e \,|\, \mathbf{t}_w^e \,]$ is determined from the pose estimate of the external camera ($\mathbf{p}_e^w$, $\bar{q}_e^w$).

## 2.4 Vessel Geometry Representation and Landmark Projection

The geometry of the lower core of a nuclear reactor pressure vessel lends itself well to a "planar map" — a map that is sparsely represented by a series of adjacent planes, that form three-dimensional lines, with landmarks that exist on these planes. We will show how localization with respect to this map is possible with only observations that are projective in nature, i.e., not three-dimensional.

### 2.4.1 Lower Core Geometry

Consider the geometry of the lower core of a nuclear reactor pressure vessel that is shown in Fig. 2.3. The lower core consists of several components that are critical to this work. (Note that Fig. 2.3 shows a fuel assembly mounted to the core plate. For our work, the fuel arrays are removed, leaving no obstruction within the lower core.)

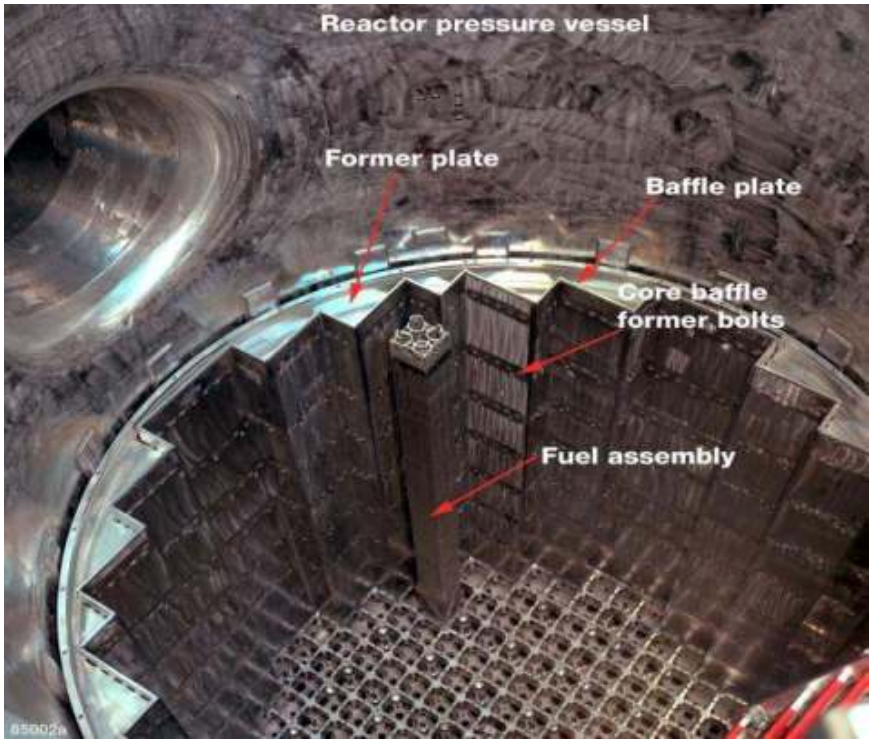



Figure 2.3: A diagram of the primary components of the lower core of a reactor pressure vessel [3]. For our work, all fuel assemblies are removed. Note that the baffle plates can be represented as vertical planes. Additionally, the core plate on the bottom of the core can be represented by a horizontal plane. The core plate contains flow holes, and bolts and other structural elements are mounted to the baffle plates.

First, consider the baffle plates that form the "walls" of the reactor. These plates can be represented by vertical plane segments. A three-dimension line is formed at the intersection of adjacent baffle plates. The positioning of the baffle plates is designed to minimize the flow disturbance of water during normal reactor operation, so this structural design pattern for the lower core is ubiquitous in this nuclear reactor class. Therefore, designing our framework to leverage this geometry is sound and does not limit its use from reactor-to-reactor. Lastly, the baffle plates contain structural elements, such as bolts, that can be used for localization.

The core plate at the floor of the vessel is flat and contains numerous circular extrusions, known as flow holes. The positioning of the fuel rods is tightly controlled.

Similarly, the position and diameter of the flow holes are known to a high degree of precision.

Structurally, little variation exists between lower cores of this type of nuclear reactor. All can be described by adjacent planar segments intersecting with a horizontal plane (the core plate). Some structural variation will occur with the energy generation capacity of the reactor, but such differences will be limited to baffle configuration and the surface area of the core plate. Visually, however, there may be significant variation. Figure 2.4 conveys a sense of variation between reactor cores. Several commonalities emerge: 1) a sparsity of *unique* visual texture exists, suggesting that using feature-based vision methods would be challenging; 2) the flow holes on the core plate are usually well-visible, as are the lines that arise from intersecting baffle plates; and 3) structural elements on the walls are not reliably distinguishable.

### 2.4.2 Planar Representation of Vessel Geometry

Although the visual texture of the lower core may vary, the fundamental structure of core will not. Therefore, we propose a planar representation of the lower core that concisely describes the core geometry. This planar representation thereby enables localization of the inspection system with respect to a map that uses this representation (Section 3.3).

The lower core geometry is defined with respect to the world frame, $\{W\}$. It principally consists of a series of planes, which are the geometric representation of the baffle plates and core plate. Commonly, we refer to the vertical planes as the "walls" of the vessel. Each plane of the vessel $\boldsymbol{\pi} = [\,\bar{n}^{\mathrm{T}},\, d\,]^{\mathrm{T}} \in \mathbb{R}^4$ is defined by a unit normal vector $\bar{n} \in \mathbb{R}^3$ and a distance $d \in \mathbb{R}$. By convention, the normal vector for each wall points outward, and the normal vector for the core plate points in the

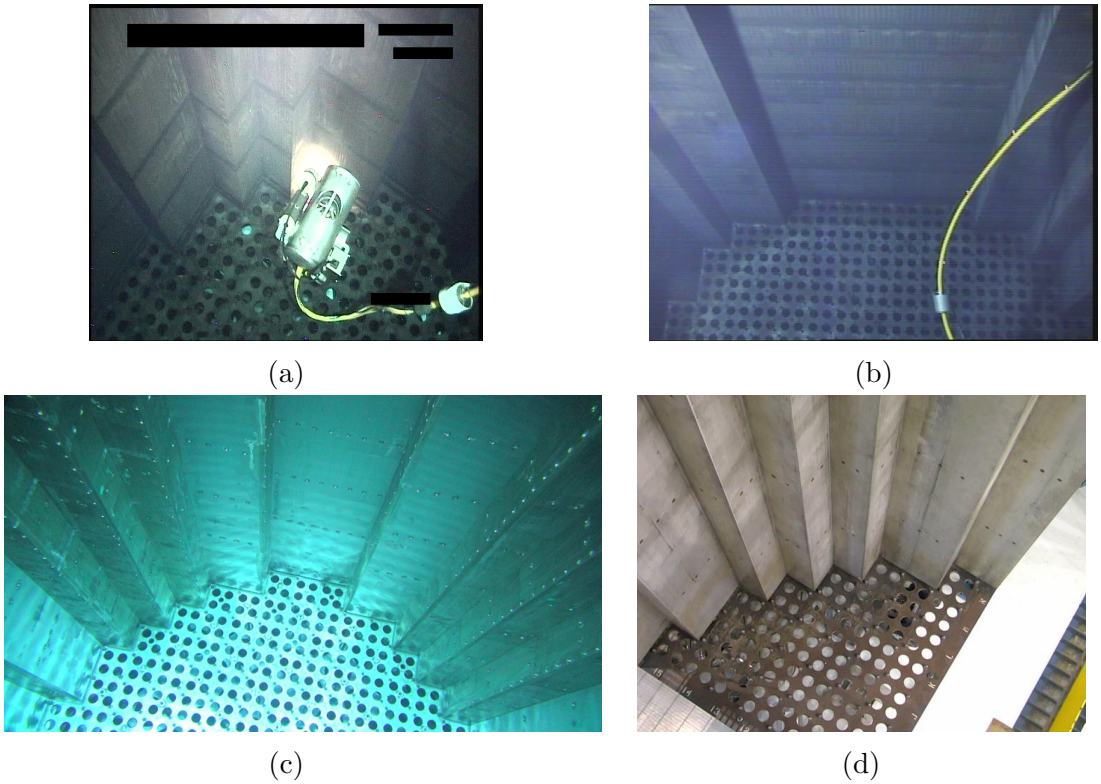


Figure 2.4: Representative examples of the reactor cores and visual texture that the framework must leverage to achieve ROV localization. Despite the differences in texture, these examples share a common structure and thus can all be described by a generalized representation.

positive $z$-direction of $\{W\}$ (i.e., away from the inside of the vessel). The three-dimensional line that arises from the intersection of two adjacent planes, $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_j$, can therefore be represented by the Plücker line $\mathcal{L}_{ij} = \boldsymbol{\pi}_i \wedge \boldsymbol{\pi}_j$, where the "meet" operator $\wedge$ represents the smallest subspace that is formed by the intersection of two subspaces. Note that Plücker lines will occur between the core plate and the baffles plates, at the intersection of baffle plates, and at the vertical edge of the baffle plate (the top of the reactor). These lines along the top rim of the reactor are formed from the intersection of vertical baffle plates and an imaginary plane across the top of the lower core that is parallel to the core plate.

The structural elements that exist within the vessel (e.g., bolts, cavities, and flow

holes) are represented by a three-dimensional point, $\boldsymbol{P}$. These points must exist on a vessel surface, as represented by the constraint $\boldsymbol{\pi} \cdot \widetilde{\boldsymbol{P}} = 0$, where $\widetilde{\boldsymbol{P}} \sim [\, \boldsymbol{P}^{\mathrm{T}},\, 1\,]^{\mathrm{T}}$ is the three-dimensional point in homogeneous coordinates and $\boldsymbol{\pi}$ is the coincident plane. No points exist in the vessel that are not coincident with a planar surface.

### 2.4.3 Vessel Landmark Projections: Points and Lines

The PTZ camera does not provide depth observations, so the projection of vessel points and lines in the image space serve as landmarks used for localization of the inspection system.

Three-dimensional points and their projections in the image space are related by

$$\widetilde{\boldsymbol{p}}_i \sim K\, [\, R^e_w \,|\, \mathbf{t}^e_w \,]\, \widetilde{\boldsymbol{P}}^w_i \tag{2.19}$$

For lines, the Plücker line $\boldsymbol{\mathcal{L}}_{ij} \sim \boldsymbol{\pi}_i \wedge \boldsymbol{\pi}_j$ formed from the intersection of adjacent planes defined in the world coordinate frame are projected in the external camera image space as a line $\boldsymbol{\ell}_{ij} \in \mathbb{P}^2$ [28]:

$$\boldsymbol{\mathcal{L}}^w_{ij} \sim \begin{bmatrix} d_i \bar{n}_j - d_j \bar{n}_i \\ \\ \bar{n}_i \times \bar{n}_j \end{bmatrix} \tag{2.20}$$

$$\boldsymbol{\ell}_{ij} \sim \mathcal{K} \mathcal{D}^e_w \boldsymbol{\mathcal{L}}^w_{ij} \tag{2.21}$$

where $\mathcal{D}^e_w$ is defined as the line displacement matrix that maps a Plücker line in the world frame to the camera frame per Eq. 2.5 with $R = R^e_w$ and $\mathbf{t} = \mathbf{t}^e_w$.

In both cases of lines and points, note that the frame of the landmark is the world frame. In the formulation of the map (Section 3.3), we will discuss a minimal representation for planes and points that leverages the assumption that all landmark

points must be coincident with a plane.

## 2.5 Speckling: Radiation-Induced Image Noise

Over time, the PTZ camera electronics will begin to degrade from continual exposure to radiation. This sensor degradation will manifest as speckling, defined as radiation-induced chromatic image noise (Fig. 2.5). Speckling is characterized by the random occurrence of pixel clusters to become activated with a high color intensity, persisting for only one frame.

To model the speckling process for later evaluation of the framework's robustness to speckling, a probabilistic speckling model was quantified using 549 frames from an inspection dataset that exhibited speckling. Table 2.1 describes the model. The number of speckles per frame $n_{frame}$ (normalized by total number of pixels) is modeled as Gaussian distribution. The size of the each speckle in pixels, $s_{size}$, is then chosen from a categorical distribution.

Table 2.1: Speckling model as defined by the number of speckles per frame (normalized by total number of pixels) and the speckle size in pixels.

| Parameter | Value |
|---|---|
| $n_{frame} \sim \mathcal{N}(\mu, \sigma^2)$ | |
| $\mu$ | 6.4655e−5 |
| $\sigma^2$ | 3.0493e−10 |
| $s_{size} \sim Cat(K_i, p_i)$ | |
| $i = \{1, \ldots, 6\}, K_i = i$ | |
| $p_1$ | 0.4034 |
| $p_2$ | 0.4087 |
| $p_3$ | 0.1021 |
| $p_4$ | 0.0542 |
| $p_5$ | 0.0200 |
| $p_6$ | 0.0116 |

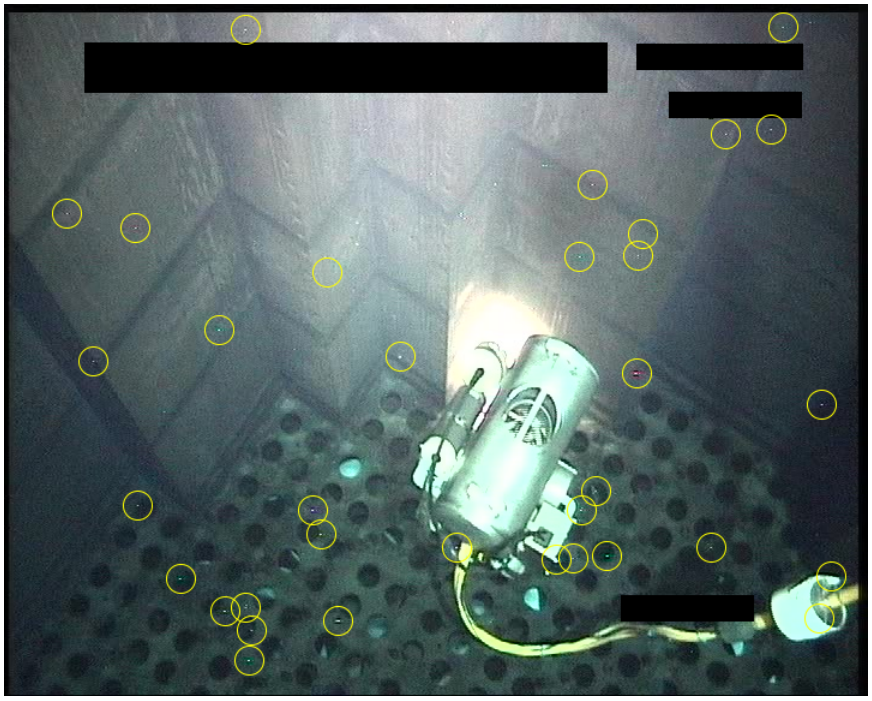


Figure 2.5: Speckling (radiation-induced chromatic image noise) observed during a reactor pressure vessel inspection. Speckle clusters are circled in yellow for ease in viewing.

## 2.6 Summary of Models and Methods

In this chapter, we have established the necessary foundational models and methods to characterize the inspection system. We have presented the models for the PTZ camera and the projection of points and lines; a brief discussion of the effect of zoom on the camera optics; characterization of the ROV and fiducial marker projection; and a planar representation that is viable for the lower core of this class of reactor pressure vessels. With these preliminaries, in the next chapter we will consider how to estimate and localize the system using the camera.

# Chapter 3

# State Estimation and Localization

The underlying state estimation framework that enables ROV localization is described in this chapter. The models and methods described in Chapter 2 are utilized here. In addition to formulating the framework, we also present a method for online initialization of the framework that is achieved through camera self-calibration.

## 3.1 State Estimation and Localization using the Extended Kalman Filter

We estimate the non-linear state of the system (robot, camera, and map) using a discrete extended Kalman filter (EKF) with quaternions [33] to obtain a recursive state estimate. This state estimation framework leverages the methods and models described in Section 2. Figure 3.1 contextualizes the various components of the framework.

The framework requires as priors the position of the fiducial markers on the robot,

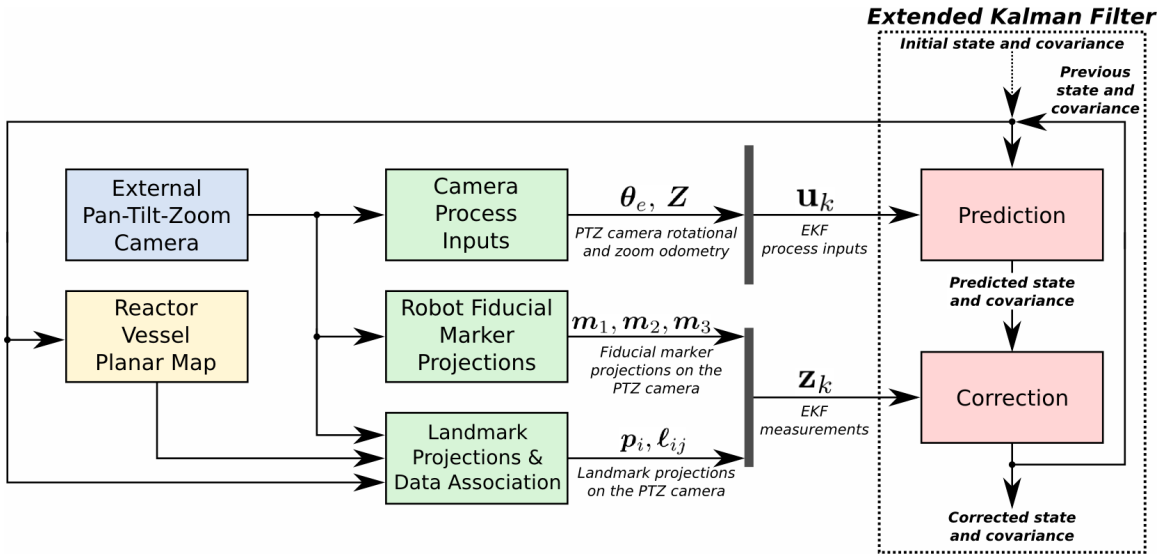



Figure 3.1: The system diagram of the state estimation and localization framework.

the vessel geometry as obtained from blueprint specifications within a limited degree of uncertainty, and the camera radial and tangential distortion parameters, obtained via a previously known calibration. We assume from this point forward that the external camera images are unwarped following a transformation to reverse the lens distortion. The filter is initialized in one of two ways. Manual initialization is done by specifying initial state values that minimize the projection error between the observed and predicted vessel landmarks. As an alternative, an online initialization procedure can estimate the camera parameters from self-calibration from the reactor structure, and the ROV state can be added to the filter state through inference of the camera state and the marker projections.

Of the possible state estimation frameworks, our choice of using an EKF is motivated as follows. Consider that our system has world observations (within some uncertainty), and the environment itself is relatively confined. Pose graphs are highly accurate and especially useful when large-scale loop closures are needed to achieve a globally consistent map [34]. However, because observations of the reactor are known

in the frame that we wish to localize the ROV, the issue of loop closure is not relevant for this application, as camera odometry errors are sufficiently corrected by the observations. For this reason, it is preferred to use a filtering approach because the most recent state is sufficient to describe the temporal evolution of the state, which significantly reduces the computational needs of the system. Between possible filtering options, such as the EKF or the Unscented Kalman Filter (UKF), the UKF has been shown to produce slightly more accurate results by reconstructing the Gaussian belief distribution of the states through a series of sigma points, instead of linearizing the system about a point [34]. However, an advantage of the EKF is that the analytic Jacobian matrices can be used to conduct an observability analysis of the system, providing understanding of which observations are necessary to render the system fully observable. Although we leave the observability analysis for future work, this may provide powerful insights to the filter. Particularly, it may be possible to assess the system observability during operation to lend robustness to the system and detect when the system is no longer completely observable.

## 3.2 System Parameterization

The non-linear state of the inspection system $\mathbf{X}(t)$ is estimated via an extended Kalman filter with inputs $\mathbf{u}(t) \in \mathbb{R}^5$ and a variable number of measurements $\mathbf{z}(t)$, depending on the scene. The system state encompasses the state of the inspection system (robot and external camera) $\mathbf{X}_S(t)$ and the map of the vessel $\mathbf{X}_M$:

$$\mathbf{X}(t) = [\, \mathbf{X}_S(t)^{\mathrm{T}},\, {\mathbf{X}_M}^{\mathrm{T}} \,]^{\mathrm{T}} \tag{3.1}$$

The state of the system is represented by

$$\mathbf{X}_S(t) = [\, \mathbf{p}_b^w(t)^{\mathrm{T}},\, \bar{q}_b^w(t)^{\mathrm{T}},\, \mathbf{p}_e^{w\mathrm{T}},\, \bar{q}_e^w(t)^{\mathrm{T}},\, \boldsymbol{f}_e(t)^{\mathrm{T}} \,]^{\mathrm{T}} \tag{3.2}$$

where $\mathbf{p}_b^w(t) = [\, x_b^w(t),\, y_b^w(t),\, z_b^w(t) \,]^{\mathrm{T}}$ and $\bar{q}_b^w(t)$ are the position and orientation (respectively) of the robot with respect to the world frame, $\mathbf{p}_e^w = [\, x_e^w,\, y_e^w,\, z_e^w \,]^{\mathrm{T}}$ and $\bar{q}_e^w(t)$ are the (static) position and orientation of the external camera with respect to the world, and $\boldsymbol{f}_e(t) = [\, f_x(t),\, f_y(t) \,]^{\mathrm{T}}$ is the vector of focal length components. As described in Section 2.2.3, both focal length components are estimated (instead of a constant aspect ratio $\tau = {}^{f_x}/_{f_y}$) to account for the effect of zoom on underwater optics. Note that we formulate the system and process models (Sec. 3.4) in continuous time to make clear the temporal evolution of the system. In other sections, we use the subscript $k$ to indicate discrete time; the implementation of the framework is in discrete time.

## 3.3 Map Representation and Estimation

Using the representation of reactor geometry presented in Section 2.4, the world structure is expressed as a map in the state estimate $\mathbf{X}_M$ that encompasses the vessel planes and three-dimensional landmarks that exist on these planes.

We propose a minimal representation for utilizing a map of the reactor in the EKF framework by extending the geometric representation described in Sec. 2.4. Assuming that the vessel baffle plates are orthogonal to the core plate, vertical planes are specified by their rotational degree of freedom ($\theta$) about the world $z$-axis and translational degree of freedom ($d$). Therefore, the unit normal for each wall is $\bar{n} = [\, \cos\theta,\, \sin\theta,\, 0 \,]^{\mathrm{T}}$ (so the unit normal points away from the reactor center). For the

core plate ("floor") of the vessel, only the height of the vessel $h$ is needed, for $\bar{n} = [\,0,\,0,\,1\,]^{\mathrm{T}}$ and $d = -h$. Therefore, if the vessel consists of $N$ walls, $2N+1$ parameters are needed for the planar structure of the vessel. An additional "virtual" plane is defined that represents the top rim of the vessel with unit normal $\bar{n} = [\,0,\,0,\,1\,]^{\mathrm{T}}$ and $d = 0$. This plane is necessary to represent the Plücker lines that arise from the top of the reactor baffle plates (where the walls intersect the "top" plane). This representation for planes is summarized as follows:

$$\boldsymbol{\pi}_{wall,\,i} = \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \\ 0 \\ d_i \end{bmatrix} \quad \boldsymbol{\pi}_{floor} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -h \end{bmatrix} \quad \boldsymbol{\pi}_{top} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{3.3}$$

Although landmark points are three-dimensional, they all must exist on a plane. Consider the planar map diagram described by Fig. 3.2. To enforce the coplanarity of a landmark with its associated plane, a point is represented by two translational degrees of freedom within in the plane, $\delta_1$ and $\delta_2$, relative to the point $-\bar{n}d$, which is the point on the plane closest to the origin of the world frame $\{\,W\,\}$. These degrees of freedom represent the two-dimensional position of the landmark in the two-dimensional subspace of $\mathbb{R}^3$ formed by $\boldsymbol{\pi}$. When considering $\bar{n}$ as one axis of an orthonormal basis of the plane, the other two axes are $\bar{v}_1 = [\,-\sin\theta,\,\cos\theta,\,0\,]^{\mathrm{T}}$ and $\bar{v}_2 = [\,0,\,0,\,1\,]^{\mathrm{T}}$ for walls, and $\bar{v}_1 = [\,1,\,0,\,0\,]^{\mathrm{T}}$ and $\bar{v}_2 = [\,0,\,1,\,0\,]^{\mathrm{T}}$ for the floor. Note that this choice of planar orthonormal basis vectors for both the walls and core plate forms a coordinate frame $[\,\bar{v}_1,\,\bar{v}_2,\,\bar{n}\,]$ that is consistent with the right-hand rule.

A landmark's three-dimensional position in the world frame $\{\,W\,\}$ can be recovered

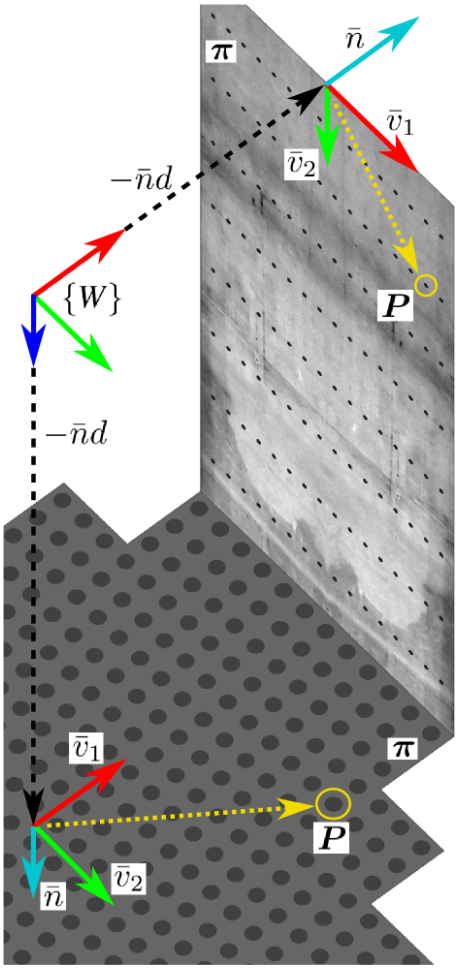



Figure 3.2: Diagram describing the representation of vessel geometry used for the map in terms of planes and points that exist on those planes. All planar coordinate frames adhere to the right-hand rule. The two-dimensional landmark representation in a given plane relative to the local plane coordinate frame origin $-\bar{n}d$ is distance $\delta_1$ along planar basis $\bar{v}_1$ and distance $\delta_2$ along planar basis $\bar{v}_2$.

from its coincident plane and two-dimensional position within this plane:

$$\boldsymbol{P} = \delta_1 \bar{v}_1 + \delta_2 \bar{v}_2 - \bar{n}d \tag{3.4}$$

It follows from Eq. 3.4 that the coplanarity constraint of the landmark, $\boldsymbol{\pi} \cdot \widetilde{\boldsymbol{P}} = 0$, is always satisfied for any choice of $\theta$, $d$, $\delta_1$, or $\delta_2$.

Formulating the planes and landmarks in this manner wholly represents the ge-

ometry with a minimum number of degrees of freedom. With this minimal representation, the map state $\mathbf{X}_M \in \mathbb{R}^{2N+1+2L}$ for $N$ walls and $L$ points is as follows:

$$\mathbf{X}_M = [\,\theta_1,\, d_1,\, \ldots,\, \theta_N,\, d_N,\, h,\, \delta_{1,1},\, \delta_{1,2},\, \ldots,\, \delta_{L,1},\, \delta_{L,2}\,]^{\mathrm{T}} \tag{3.5}$$

In the above formulation, we chose to represent the reactor map in the state in terms of planes. The intersecting Plücker lines could then be inferred from the map as needed. A dual representation could instead estimate the map by its corner points. For an equivalent minimal representation, consider that the corner points on the top rim of the vessel have the same $x$- and $y$-coordinates as those on the reactor core plate. Thus, two parameters for the $x$- and $y$-coordinates of each corner point that defines the two-dimensional "footprint" of the reactor are needed, plus the reactor height. This requires $2C + 1$ parameters, where $C$ is the number of corner points that define the reactor footprint. From these corner points, the reactor wall equations can be constructed. Note that reactor vessels are enclosed, so $N = C$.[1] These representations are duals, and thus either is viable.

## 3.4 Process Models

The system process model characterizes the temporal evolution of the state. For this section, we first describe the process model in continuous time, and then show how it is implemented in discrete time.

[1]For our implementation, this is true even for vessels that aren't enclosed. For example, the subscale mockup utilizes three "virtual walls" to enclose the vessel, so $N = C$ even in this case.

### 3.4.1 Formulation in Continuous Time

The system process input, $\mathbf{u}(t) \in \mathbb{R}^4$, consists of the angular velocity of the external camera with respect to its reference frame, $\boldsymbol{\omega}_e(t) \in \mathbb{R}^3$, and the rate of change of the normalized scaling factor for the focal length components, $\zeta(t) \in \mathbb{R}$. The entire process model expressed in continuous time is

$$\dot{\mathbf{p}}_b^w(t) = 0 \tag{3.6}$$

$$\dot{\bar{q}}_b^w(t) = 0 \tag{3.7}$$

$$\dot{\mathbf{p}}_e^w = 0 \tag{3.8}$$

$$\dot{\bar{q}}_e^w(t) = \tfrac{1}{2}\, Q(\boldsymbol{\omega}_e(t))\, \bar{q}_e^w(t) \tag{3.9}$$

$$\dot{\boldsymbol{f}}_e(t) = \zeta(t)\, \boldsymbol{f}_e(t) \tag{3.10}$$

$$\dot{\mathbf{X}}_M = 0 \tag{3.11}$$

These continuous time update equations are converted into discrete time using Euler discretization. We model the robot's state evolution as being driven forward by a random walk through additive process noise. For the external camera process, the position of the external camera $\mathbf{p}_e^w$ is static. The external camera rotates with an body-referenced angular velocity $\boldsymbol{\omega}_e(t)$. $Q(\boldsymbol{\omega}_e(t))$ is the quaternion kinematic matrix [35] that relates angular velocity in a body-referenced frame and quaternion orientation to quaternion rate. Lastly, the map is static as it is defined in the world frame $\{W\}$.

The process inputs are determined from the inspection camera, which reports a pan angle, tilt angle, and zoom setting. We do not use the angles directly to estimate the state of the camera, as Wu and Radke have shown that in long-term operations, the pan and tilt angles become inaccurate due to various error sources,

such as mechanical errors in the camera motors [27].

For the zoom process, consider that the camera does not provide a metric focal length measurement. Instead, a scalar value representing the camera zoom state is provided that maps to the "times" representation of zoom (e.g., 2x zoom). This scalar is an estimate of the camera's focal length, normalized with respect to the camera's base focal length (when completely zoomed out). Therefore, the rate at which this scalar changes allows for inference of the rate of change in the focal length components. This is formulated as a normalized scaling factor, where $\zeta = 0$ when the camera zoom state is static. Thus, $\zeta > 0$ indicates zoom-in, and $\zeta < 0$ indicates zoom-out.

Note that an alternative to using the PTZ camera state is to infer the process visually via homography-based image registration [36]. However, as the pan, tilt, and zoom settings are reported by the camera, it is straightforward to utilize these instead in the state process model.

### 3.4.2 Implementation in Discrete Time

For implementation in discrete time, we utilize the formulation described by Thrun, Wolfram, and Fox [34]. A non-linear process model $g$ predicts the evolution in the state from the previous state $\mathbf{X}_{k-1}$ to the current state $\mathbf{X}_k$ given a system input $\mathbf{u}_k \in \mathbb{R}^4$ in the current timestep:

$$\mathbf{X}_k = g(\mathbf{u}_k,\, \mathbf{X}_{k-1}) + \mathcal{N}(0, Q_k) \tag{3.12}$$

$$\mathbf{u}_k = \left[\boldsymbol{\theta}_e{}^{\mathrm{T}},\, Z\right]^{\mathrm{T}} \tag{3.13}$$

For our system,

$$g(\mathbf{u},\, \mathbf{X}) = \begin{bmatrix} \mathbf{p}_b^w \\ \bar{q}_b^w \\ \mathbf{p}_e^w \\ || \, \bar{q}_e^w + \frac{1}{2}\, Q(\boldsymbol{\theta}_e)\, \bar{q}_e^w \, || \\ (Z+1)\, \boldsymbol{f}_e \\ \mathbf{X}_M \end{bmatrix} \tag{3.14}$$

where $\boldsymbol{\theta}_e \in \mathbb{R}^3$, $[\boldsymbol{\theta}_e]_\times \in \mathfrak{so}(3)$ represents the rotational odometry arising from the camera rotation over a small timestep. Assuming Euler integration, it can be shown that $\boldsymbol{\theta}_e = \boldsymbol{\omega}_e \Delta t$. The input $\boldsymbol{\theta}_e$ is obtained from successive pan and tilt angles reported by the camera using Eq. 2.6 with a scene orientation of $I_{3\times3}$:

$$R_e = R(I_{3\times3}, \theta_{k-1}, \phi_{k-1})^{-1} R(I_{3\times3}, \theta_k, \phi_k) \tag{3.15}$$

The axis-angle representation of $R_e$ provides $\boldsymbol{\theta}_e$. Note that the scene orientation $R_s$ does not impact the change of camera rotation from the reference frame of the camera and is therefore not needed for calculation of the rotational process input.

For the zoom process, $Z \in \mathbb{R}$ represents the normalized change in focal length components. This is obtained from successive zoom states $z \in \mathbb{R}_{\geq 1}$ reported by the camera:[2]

$$Z = \left( \frac{z_k}{z_{k-1}} - 1 \right) \tag{3.16}$$

[2]Strictly speaking, the zoom state reported by the camera is a rational number $z \in \mathbb{Q}_{\geq 1}$, but the precision is sufficient to represent it as a real number for practical purposes.

where $z$ provides a measure of the current scaling of the focal length from the base focal length (i.e., 2x zoom). It can be shown that $Z = \zeta \Delta t$ for the Euler method of integration, and thus $Z\boldsymbol{f}_e$ can be considered to be "zoom odometry" — the change in focal length components over some $\Delta t$.

With the process inputs, the mean and covariance of the Gaussian state belief distribution following the process update is as follows [34]:

$$\bar{\boldsymbol{\mu}}_k = g(\mathbf{u}_k,\, \boldsymbol{\mu}_{k-1}) \tag{3.17}$$

$$\bar{\Sigma}_k = G_k\, \Sigma_{k-1}\, {G_k}^{\mathrm{T}} + Q_k \tag{3.18}$$

The matrix $G_k$ is used to propagate the state covariance through the non-linear process model, $g$, through linearization. This matrix is obtained by evaluating the Jacobian matrix $G$ with the process input $\mathbf{u}_k$ and state mean $\boldsymbol{\mu}_{k-1}$:

$$G = \frac{\partial g(\mathbf{u},\, \mathbf{X})}{\partial \mathbf{X}} \tag{3.19}$$

$$G_k = G|_{\mathbf{u}_k,\, \boldsymbol{\mu}_{k-1}} \tag{3.20}$$

The process noise matrix $Q_k \in \mathbb{R}^{N \times N}$ accounts for the random process of the ROV state through matrix $W_k \in \mathbb{R}^{N \times N}$, as well as the input noise $M_k \in \mathbb{R}^{4 \times 4}$:

$$Q_k = W_k + V_k\, M_k\, {V_k}^{\mathrm{T}} \tag{3.21}$$

$$W_k = \mathrm{diag}(\sigma_r^2\, I_{3\times 3},\, \sigma_q^2\, I_{4\times 4},\, 0_{N-7\times N-7}) \tag{3.22}$$

where $\sigma_r^2 \in \mathbb{R}$ and $\sigma_q^2 \in \mathbb{R}$ are the position and quaternion variances associated with the ROV random process. The expression diag refers to block diagonalization, which produces a sparse matrix except for the indicated square matrices on the diagonal.

The matrix $V_k \in \mathbb{R}^{N\times 4}$ is used to map the input noise to the state space.

$$V = \frac{\partial g(\mathbf{u},\, \mathbf{X})}{\partial \mathbf{u}} \tag{3.23}$$

$$V_k = V|_{\mathbf{u}_k,\, \boldsymbol{\mu}_{k-1}} \tag{3.24}$$

The process noise expressed in the input space is $M_k$:

$$M_k = \mathrm{diag}(\sigma_a^2 I_{3\times 3},\, \sigma_z^2) \tag{3.25}$$

where $\sigma_a^2 \in \mathbb{R}$ and $\sigma_z^2 \in \mathbb{R}$ are the variances associated with the rotational and zoom process inputs, respectively.

## 3.5 Measurement Models

All system measurements $\mathbf{z}$ consist of projections into the camera image space. In total, there are six types of measurements that can be categorized into two types: 1) $\mathbf{z}_b^e$, relating the robot body frame $\{B\}$ to the camera frame $\{E\}$; and 2) $\mathbf{z}_e^w$, which relates the camera frame $\{E\}$ to the world frame $\{W\}$:

$$\mathbf{z} = [\, \mathbf{z}_b^{e\mathrm{T}},\, \mathbf{z}_e^{w\mathrm{T}} \,]^{\mathrm{T}} \tag{3.26}$$

Measurement predictions $\hat{\mathbf{z}}_k$ are formed from the current state $\mathbf{X}_k$ using a non-linear measurement model $h$ and measurement noise $R_k$ [34]:

$$\hat{\mathbf{z}}_k = h(\mathbf{X}_k) + \mathcal{N}(0, R_k) \tag{3.27}$$

Based on the scene, there are a variable number of measurements available at

each timestep. For this reason, we describe the measurement models individually. The measurements can either be assembled in aggregate for one large observation update, or applied individually as they are processed.

The EKF equations that yield the *a posteriori* state belief distributions are as follows:

$$\Delta \mathbf{z}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k \tag{3.28}$$

$$S_k = H_k \bar{\Sigma}_k {H_k}^{\mathrm{T}} + R_k \tag{3.29}$$

$$K_k = \bar{\Sigma}_k {H_k}^{\mathrm{T}} S_k^{-1} \tag{3.30}$$

$$\boldsymbol{\mu}_k = \bar{\boldsymbol{\mu}}_k + K_k \Delta \mathbf{z}_k \tag{3.31}$$

$$\Sigma_k = \bar{\Sigma} - K_k H_k \bar{\Sigma}_k \tag{3.32}$$

where $\hat{\mathbf{z}}_k = h(\bar{\boldsymbol{\mu}}_k)$.

The Jacobian matrix $H_k$ provides the linearization of the non-linear measurement models at the state mean $\boldsymbol{\mu}_k$:

$$H = \frac{\partial h(\mathbf{X})}{\partial \mathbf{X}} \tag{3.33}$$

$$H_k = H|_{\boldsymbol{\mu}_k} \tag{3.34}$$

With the exception of the fiducial marker model, all models assume a measurement noise of $\sigma_I^2$: $R_{k,\,pt} = \sigma_I^2\, I_{2\times 2}$ for points and $R_{k,\,ln} = \sigma_I^2\, I_{4\times 4}$ for line segments. Thus, we do not detail the measurement noise for these models in their section. However, the measurement noise for the fiducial marker model is described in further detail as it accounts for the uncertainty in the marker positions.

### 3.5.1 Robot-to-Camera Measurements

The body-to-camera measurements, $\mathbf{z}^{e}_{b,\,k} = \mathbf{z}_{m,\,k} \in \mathbb{R}^6$, are determined through robot fiducial marker detection (Section 2.3):

$$\mathbf{z}_{m,\,k} = [\, \boldsymbol{m}_{1,\,k}{}^{\mathrm{T}},\, \boldsymbol{m}_{2,\,k}{}^{\mathrm{T}},\, \boldsymbol{m}_{3,\,k}{}^{\mathrm{T}} \,]^{\mathrm{T}} \tag{3.35}$$

The measurement model is described in Eq. 2.18 that relates the predicted projection of the ROV markers, $\hat{\boldsymbol{m}}_i \in \mathbb{R}^2$, $i = \{1, 2, 3\}$, into the image space:

$$\hat{\mathbf{z}}_{m,\,k} = h_m(\mathbf{X}_k) + \mathcal{N}(0, R_{m,\,k}) \tag{3.36}$$

$$h_m(\mathbf{X}_k) = [\, \hat{\boldsymbol{m}}_{1,\,k}{}^{\mathrm{T}},\, \hat{\boldsymbol{m}}_{2,\,k}{}^{\mathrm{T}},\, \hat{\boldsymbol{m}}_{3,\,k}{}^{\mathrm{T}} \,]^{\mathrm{T}} \tag{3.37}$$

The measurement noise for this model $R_{m,\,k} \in \mathbb{R}^{6\times6}$ accounts for both noise in the image space, but also incorporates the uncertainty of the marker positions in the body frame:

$$R_{m,\,k} = \sigma_I^2 I_{6\times6} + J_k \left(\sigma_M^2 I_{9\times9}\right) J_k{}^{\mathrm{T}} \tag{3.38}$$

where $J_k \in \mathbb{R}^{6\times9}$ is the Jacobian matrix $J$ evaluated at $\boldsymbol{\mu}_k$ that maps the position uncertainty from the marker position space to the image space:

$$J_{ij} = \frac{\partial \mathbf{m}_i}{\partial \mathbf{M}_j^b}, \quad \forall i, j = \{1,\, 2,\, 3\} \tag{3.39}$$

$$J_k = J|_{\boldsymbol{\mu}_k} \tag{3.40}$$

### 3.5.2 World-to-Camera Measurements

Projections of structural elements (Section 2.4.3) provide observations for camera-to-world localization; therefore, robot-to-world localization is possible through the camera from the ROV marker observations. While the number of marker corrections is fixed while the robot is in view, the number of landmark corrections will vary depending on the scene. All measurements assume $\sigma = 3$ noise.

There are five types of world-to-camera observations $\mathbf{z}^{w}_{e,\,k}$, which can be categorized by the underlying geometric projection:

1. **Projection of Points**
    - Projection of floor (core plate) landmarks.
    - Projection of wall (baffle plate) landmarks.
2. **Projection of Lines**
    - Projection of Plücker lines that arise from the intersection of adjacent walls (baffle plates).
    - Projection of Plücker lines that arise from the intersection of a wall (baffle plate) and the floor (core plate).
    - Projection of Plücker lines that arise from the intersection of a wall (baffle plate) and the virtual top plane of the lower core.

### Projection of Landmark Points

The projection of core plate landmarks into the camera as image points is described as follows:

$$\mathbf{z}_{fp,\,k,\,i} = \boldsymbol{p}_{fp,\,k,\,i} \tag{3.41}$$

where $\boldsymbol{p}_{fp,\,k,\,i} \in \mathbb{R}^2$ is the projection of landmark point $\boldsymbol{P}^{w}_{fp,\,i}$ that exists on the core plate (floor). The subscript $i$ here and forward implies that multiple projections may exist at a particular timestep $k$.

Likewise, the projection of a landmark that exists on a wall (baffle plate) arises in the image space as follows:

$$\mathbf{z}_{wp,\,k,\,i} = \boldsymbol{p}_{wp,\,k,\,i} \tag{3.42}$$

where $\boldsymbol{p}_{wp,\,k,\,i} \in \mathbb{R}^2$ is the projection of a wall landmark $\boldsymbol{P}^{w}_{wp,\,i}$.

The measurement model for both cases is described in Eq. 2.19:

$$\hat{\mathbf{z}}_{fp,\,k,\,i} = h_{fp}(\mathbf{X}_k) + \mathcal{N}(0, R_{k,\,pt}) \tag{3.43}$$

$$h_{fp}(\mathbf{X}_k) = \hat{\boldsymbol{p}}_{fp,\,k,\,i} \tag{3.44}$$

$$\hat{\tilde{\boldsymbol{p}}}_{fp,\,k,\,i} \sim \hat{K}\,[\,\hat{R}^{e}_{w}\,|\,\hat{\mathbf{t}}^{e}_{w}\,]\,\hat{\tilde{\boldsymbol{P}}}^{w}_{fp,\,i} \tag{3.45}$$

$$\hat{\mathbf{z}}_{wp,\,k,\,i} = h_{wp}(\mathbf{X}_k) + \mathcal{N}(0, R_{k,\,pt}) \tag{3.46}$$

$$h_{wp}(\mathbf{X}_k) = \hat{\boldsymbol{p}}_{wp,\,k,\,i} \tag{3.47}$$

$$\hat{\tilde{\boldsymbol{p}}}_{wp,\,k,\,i} \sim \hat{K}\,[\,\hat{R}^{e}_{w}\,|\,\hat{\mathbf{t}}^{e}_{w}\,]\,\hat{\tilde{\boldsymbol{P}}}^{w}_{wp,\,i} \tag{3.48}$$

where Eq. 3.4 describes how to form the three-dimensional point from the minimal representation form used in the map and available in the state. For floor landmarks, the reactor height $\hat{h}$ and the planar distances $\hat{\delta}_{fp,i,1}$ and $\hat{\delta}_{fp,i,2}$ are needed to reconstruct $\hat{\boldsymbol{P}}^{w}_{fp,i}$. For wall landmarks, the planar distances $\hat{\delta}_{wp,i,1}$ and $\hat{\delta}_{wp,i,2}$ are needed along with the vertical wall planar representation, $\hat{\theta}_{wp,i}$ and $\hat{d}_{wp,i}$, to reconstruct $\hat{\boldsymbol{P}}^{w}_{wp,i}$. In these equations and those to follow, the hat operator indicates that the parameter is estimated from the framework.

### Projection of Landmark Lines

As previously noted, the projection of a Plücker line forms a line in the image space. However, we do not observe lines directly, but rather, line *segments*, so our measurement models account for this. As shown in Fig. 3.3, each line segment observation yields two points: $\boldsymbol{p}_A \in \mathbb{R}^2$ and $\boldsymbol{p}_B \in \mathbb{R}^2$. The predicted observation is the nearest point on this predicted line to the observed line segment [37]. In this manner, the observed and predicted observations for these class of models is a 4-vector, $[\,\boldsymbol{p}_A{}^{\mathrm{T}},\, \boldsymbol{p}_B{}^{\mathrm{T}}\,]^{\mathrm{T}} = [\,u_A,\, v_A,\, u_B,\, v_B\,]^{\mathrm{T}} \in \mathbb{R}^4$. All predicted observations use a measurement model based on Eq. 2.21.

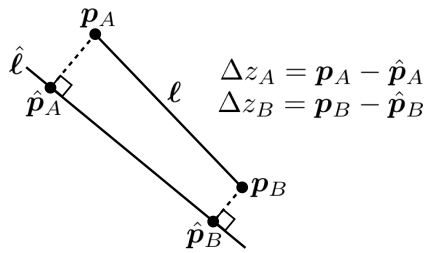



Figure 3.3: Quantifying error between a predicted line $\hat{\boldsymbol{\ell}}$ and an observed line segment $\boldsymbol{\ell}$. Points $\hat{\boldsymbol{p}}_A$ and $\hat{\boldsymbol{p}}_B$ are the closest points on line $\hat{\boldsymbol{\ell}}$ to points $\boldsymbol{p}_A$ and $\boldsymbol{p}_B$, respectively, that define line segment $\boldsymbol{\ell}$.

The observed line segment formed by two adjacent planes is $\mathbf{z}_{wl,\,k,\,i} \in \mathbb{R}^4$:

$$\mathbf{z}_{wl,\,k,\,i} = \begin{bmatrix} \boldsymbol{p}_{wl,\,A,\,k,\,i} \\ \\ \boldsymbol{p}_{wl,\,B,\,k,\,i} \end{bmatrix} \tag{3.49}$$

with measurement model

$$\hat{\mathbf{z}}_{wl,\,k,\,i} = h_{wl}(\mathbf{X}_k) + \mathcal{N}(0, R_{k,\,ln}) \tag{3.50}$$

$$h_{wl}(\mathbf{X}_k) = \begin{bmatrix} \hat{\boldsymbol{p}}_{wl,\,A,\,k,\,i} \\ \\ \hat{\boldsymbol{p}}_{wl,\,B,\,k,\,i} \end{bmatrix} \tag{3.51}$$

$$\hat{\boldsymbol{\ell}} \sim \hat{\mathcal{K}}\hat{\mathcal{D}}^{e}_{w}\hat{\boldsymbol{\mathcal{L}}}^{w}_{wl,\,i} \tag{3.52}$$

$$\hat{\boldsymbol{\mathcal{L}}}^{w}_{wl,\,i} \sim \hat{\boldsymbol{\pi}}_i \wedge \hat{\boldsymbol{\pi}}'_i \tag{3.53}$$

$$\hat{m} = -\frac{\hat{\ell}_x}{\hat{\ell}_y}, \quad \hat{m}_\perp = -\frac{1}{\hat{m}} \tag{3.54}$$

$$\hat{b} = -\frac{1}{\hat{\ell}_y}, \quad \hat{b}_{\perp,\,A} = v_A - m_\perp u_A, \quad \hat{b}_{\perp,\,B} = v_B - m_\perp u_B \tag{3.55}$$

$$\hat{\boldsymbol{p}}_{wl,\,A,\,k,\,i} = \begin{bmatrix} \hat{u}_{wl,\,A,\,k,\,i} \\ \\ \hat{v}_{wl,\,A,\,k,\,i} \end{bmatrix} = \begin{bmatrix} \frac{\hat{b}_{\perp,\,A} - \hat{b}}{\hat{m} - \hat{m}_\perp} \\ \\ \hat{m}\,\hat{u}_{wl,\,A,\,k,\,i} + \hat{b} \end{bmatrix} \tag{3.56}$$

$$\hat{\boldsymbol{p}}_{wl,\,B,\,k,\,i} = \begin{bmatrix} \hat{u}_{wl,\,B,\,k,\,i} \\ \\ \hat{v}_{wl,\,B,\,k,\,i} \end{bmatrix} = \begin{bmatrix} \frac{\hat{b}_{\perp,\,B} - \hat{b}}{\hat{m} - \hat{m}_\perp} \\ \\ \hat{m}\,\hat{u}_{wl,\,B,\,k,\,i} + \hat{b} \end{bmatrix} \tag{3.57}$$

Essentially, the above model determines the predicted Plücker line $\hat{\boldsymbol{\mathcal{L}}}^{w}_{wl,\,i}$ that arises between adjacent walls $\hat{\boldsymbol{\pi}}_i$ and $\hat{\boldsymbol{\pi}}'_i$, then projects it into the image space as $\hat{\boldsymbol{\ell}} = [\,\hat{\ell}_x,\,\hat{\ell}_y\,]^{\mathrm{T}} \in \mathbb{R}^2$. The plane equations can be reconstructed from $\hat{\theta}$, $\hat{d}$, $\hat{\theta}'$, and $\hat{d}'$. Then, a series of intermediate steps are required to calculate the closest points on this predicted line to the observed line segment; this is done using a slope-intercept representation. (Note that an equivalent projective geometry model was attempted,

but it yielded poor results due to numerical instability.)

For the intersection between the floor and a wall, the observed line segment is $\mathbf{z}_{fl,\,k,\,i} \in \mathbb{R}^4$:

$$\mathbf{z}_{fl,\,k,\,i} = \begin{bmatrix} \boldsymbol{p}_{fl,\,A,\,k,\,i} \\ \\ \boldsymbol{p}_{fl,\,B,\,k,\,i} \end{bmatrix} \tag{3.58}$$

with measurement model

$$\hat{\mathbf{z}}_{fl,\,k,\,i} = h_{fl}(\mathbf{X}_k) + \mathcal{N}(0, R_{k,\,ln}) \tag{3.59}$$

$$h_{fl}(\mathbf{X}_k) = \begin{bmatrix} \hat{\boldsymbol{p}}_{fl,\,A,\,k,\,i} \\ \\ \hat{\boldsymbol{p}}_{fl,\,B,\,k,\,i} \end{bmatrix} \tag{3.60}$$

$$\hat{\boldsymbol{\ell}} \sim \hat{\mathcal{K}}\hat{\mathcal{D}}^e_w \hat{\boldsymbol{\mathcal{L}}}^w_{fl,\,i} \tag{3.61}$$

$$\hat{\boldsymbol{\mathcal{L}}}^w_{fl,\,i} \sim \hat{\boldsymbol{\pi}}_i \wedge \hat{\boldsymbol{\pi}}_{floor} \tag{3.62}$$

In this case, $\hat{\theta}$, $\hat{d}$, and $\hat{h}$ are needed from the state to reconstruct the predicted Plücker line. The rest of the measurement model follows from above.

Lastly, for the intersection between the top virtual plane and a wall, the observed line segment is $\mathbf{z}_{tl,\,k,\,i} \in \mathbb{R}^4$:

$$\mathbf{z}_{tl,\,k,\,i} = \begin{bmatrix} \boldsymbol{p}_{tl,\,A,\,k,\,i} \\ \\ \boldsymbol{p}_{tl,\,B,\,k,\,i} \end{bmatrix} \tag{3.63}$$

with measurement model

$$\hat{\mathbf{z}}_{tl,\,k,\,i} = h_{tl}(\mathbf{X}_k) + \mathcal{N}(0, R_{k,\,ln}) \tag{3.64}$$

$$h_{tl}(\mathbf{X}_k) = \begin{bmatrix} \hat{\boldsymbol{p}}_{tl,\,A,\,k,\,i} \\ \\ \hat{\boldsymbol{p}}_{tl,\,B,\,k,\,i} \end{bmatrix} \tag{3.65}$$

$$\hat{\boldsymbol{\ell}} \sim \hat{\mathcal{K}}\hat{\mathcal{D}}^{e}_{w}\hat{\boldsymbol{\mathcal{L}}}^{w}_{tl,\,i} \tag{3.66}$$

$$\hat{\boldsymbol{\mathcal{L}}}^{w}_{tl,\,i} \sim \hat{\boldsymbol{\pi}}_i \wedge \hat{\boldsymbol{\pi}}_{top} \tag{3.67}$$

For this model, only $\hat{\theta}$ and $\hat{d}$ are needed from the state to reconstruct the predicted Plücker line. The rest of the measurement model follows from above.

### 3.5.3 Measurement Acquisition and Association

Vessel landmark projections are identified and associated first using feature detection based on geometric shape. Lines are detected using the Canny filter [38] and the probabilistic Hough transform [39]. Points are detected using blob detection and, in the case of circular elements, the Hough transform [40]. After detection, data association is performed by first projecting map elements into the image space, and comparing them against candidate detected landmarks. The closest detected landmark (within a heuristic threshold) is then associated to a projected landmark.

## 3.6 EKF Initialization

We now propose an automated EKF initialization procedure that leverages the reactor structure to perform self-calibration of the camera. This initialization procedure allows inspection personnel to boot the system with minimum manual effort, further

increasing the system utility and efficiency of inspections by avoiding the time to manually specify initial values for the state parameters.

The online initialization procedure is a two-stage process. First, the camera conducts a series of closed-loop motion sequences to obtain specific measurements of the reactor. These sequences enable self-calibration of the camera intrinsic and extrinsic parameters from the reactor structure. Then, the EKF is initialized with only the camera and map. Then, in the second stage, the state is augmented with the ROV pose. Of course, the filter can be manually initialized with the ROV, camera, and map at once from a known prior estimate of the system.

Priors of the map state are obtained from blueprint specifications of the reactor geometry, which typically are in good agreement with the reactor as-built dimensions.

### 3.6.1 Online Initialization of PTZ State through Self-Calibration from Structure

The full initialization sequence for the camera requires several stages, where calibration parameters are determined stage-by-stage. The entire algorithm is as follows: 1) estimate the camera principal point from a zoom sequence of the core plate; 2) estimate the camera focal length via homography, or utilize a previous estimate; 3) estimate the camera extrinsics from a pan and tilt sweep of the lower core perimeter, yielding the camera position and scene orientation. This procedure can be utilized to bootstrap the EKF from minimal user inputs.

#### Principal Point from Core Plate

We utilize the calibration methodology proposed by Wu and Radke [27] to estimate the camera principal point. Under assumption that the principal point is fixed and

congruent with the center of zoom, landmark correspondences during a pure zoom sequence are related as follows:

$$\frac{u' - c_x}{u - c_x} = \frac{v' - c_y}{v - c_y} = \frac{z'}{z} \tag{3.68}$$

$$c_x (v - v') + c_y (u' - u) = u'v - uv' \tag{3.69}$$

Each correspondence yields one equation. By observing the correspondences of several landmarks during a zoom sequence, a series of equations are constructed to produce an overdetermined system of equations. These equations are linear with $c_x$ and $c_y$, and so can be solved using the Direct Linear Transform (DLT) algorithm to obtain a least-squares estimate of the principal point.

To collect these correspondences, the camera is commanded to -90° so that it points towards the core plate, with the camera zoomed completely out. The camera then conducts a pure zoom sequence: zooming in to several times zoom, and then zooming out to the base zoom level. Figure 3.4 shows the zoom-in portion of this sequence. As the metric position of the landmarks are not required to calibrate the principal point, association of the landmarks with the map is not necessary; only tracking of correspondences through the zoom sequence is needed.

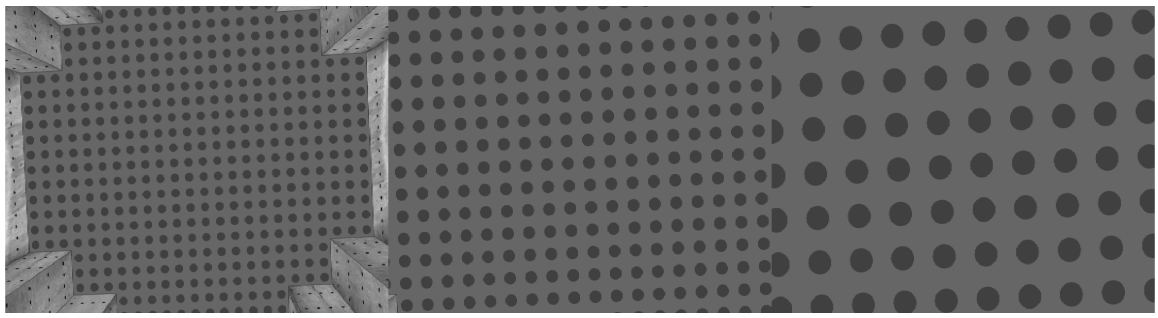


Figure 3.4: Frames from a simulated camera zoom sequence that are used to calibrate the principal point (the zoom-out portion of the sequence is not shown. Left: 1.0x zoom. Middle: 1.5x zoom. Right: 3.0x zoom.

### Focal Length

The standard technique for calibrating the camera focal length from structure is to recreate the image of the absolute conic (IAC) [26]. One such way to do this as suggested by Wu and Radke [27] is to obtain several homographies that relate images at different pan and tilt angles with the base level of zoom. Then, the IAC (and thus the focal length) is obtained as follows:

$$\omega \sim K^{-\mathrm{T}} K^{-1} \tag{3.70}$$

$$H \sim KRK^{-1} \tag{3.71}$$

$$\omega \sim H^{-\mathrm{T}} \omega H^{-1} \tag{3.72}$$

However, calibration through the IAC does not always produce a well-conditioned series of equations [30], and in practice the robustness of this method to image noise and scene conditions is poor. Therefore, for this step, we simply utilize a prior for focal length, as would be obtained from a prior successful run of the state estimate. This is possible because the camera will not change significantly between inspection deployments, and ultimately, only an estimate is needed to initialize the filter. It follows that the first deployment of a camera would require a manual estimate for focal length. Nonetheless, this is still more efficient than requiring a manual initialization each time.

### Extrinsics from Lower Core Perimeter

At this point, the intrinsic parameters of the camera have been calibrated. This final step estimates the extrinsic parameters of the camera, which are the last parameters to be specified before the filter can be initialized (without the ROV). This methodology

first estimates the camera position, then the camera scene orientation, from which the orientation in the world frame can be obtained for any known pan and tilt angle.

To collect the required measurements, the camera is commanded to scan the perimeter of the lower core — see Figs. 3.5 and 3.6. Specifically,

- The camera is commanded to the minimum pan angle and the tilt angle that causes the top rim of the lower core to be vertically centered in the image.
- The camera conducts a pure sweep motion until the maximum pan angle is reached. During this time, the corner points of the reactor structure are identified and recorded, including both the image coordinates and the reported pan and tilt angles.
- The camera conducts a pure downward tilt motion, until the intersecting lines between the baffle plates and the core plate are vertically centered in the image.
- Lastly, a pure sweep motion in the opposite direction as before is commanded, until the minimum pan angle is reached.

The basis of the extrinsic calibration-from-structure methodology is the modeling of three-dimensional triangles that arise from line segments that are formed between two adjacent "corner points" of the structure along the core perimeter, and the (unknown) camera position in the world frame. This is visualized in Fig. 3.7. In vector notation,

$$\mathbf{v}_1 = \boldsymbol{P}_A^w - \mathbf{p}_e^w \tag{3.73}$$

$$\mathbf{v}_2 = \boldsymbol{P}_B^w - \boldsymbol{P}_A^w \tag{3.74}$$

$$\mathbf{v}_3 = \mathbf{p}_e^w - \boldsymbol{P}_B^w \tag{3.75}$$

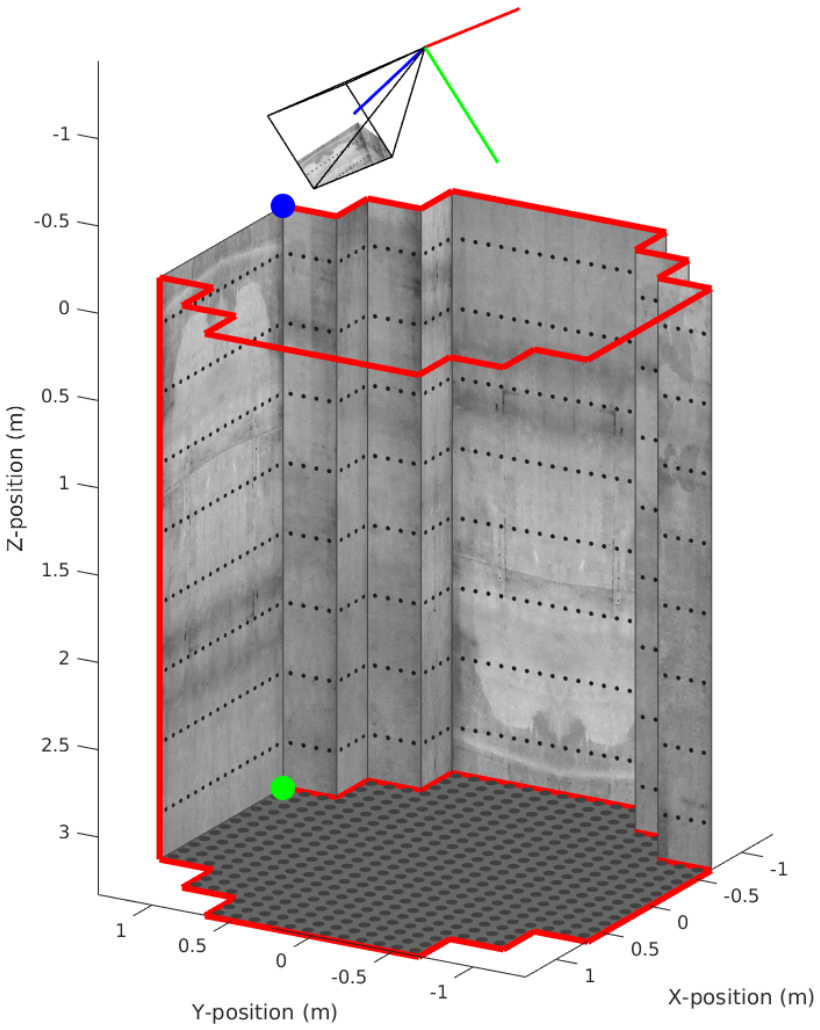



Figure 3.5: The perimeter of the lower core that is viewed by the PTZ camera for online extrinsic calibration. The corners of the vessel are identified and recorded during this sweep. The sequence begins at the blue circle and terminates at the green circle.

where $\boldsymbol{P}_A^w$ and $\boldsymbol{P}_B^w$ are two corner points in the world frame, and $\mathbf{p}_e^w$ is the camera position. It can be shown that $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = 0$, so these vectors form a triangle.

The position of the corner points are known in the world frame, so $\mathbf{v}_2$ is known. Vectors $\mathbf{v}_1$ and $\mathbf{v}_3$ represent rays that are backprojected from the camera center. To solve the triangle, consider that a calibrated camera is essentially a bearing sensor.

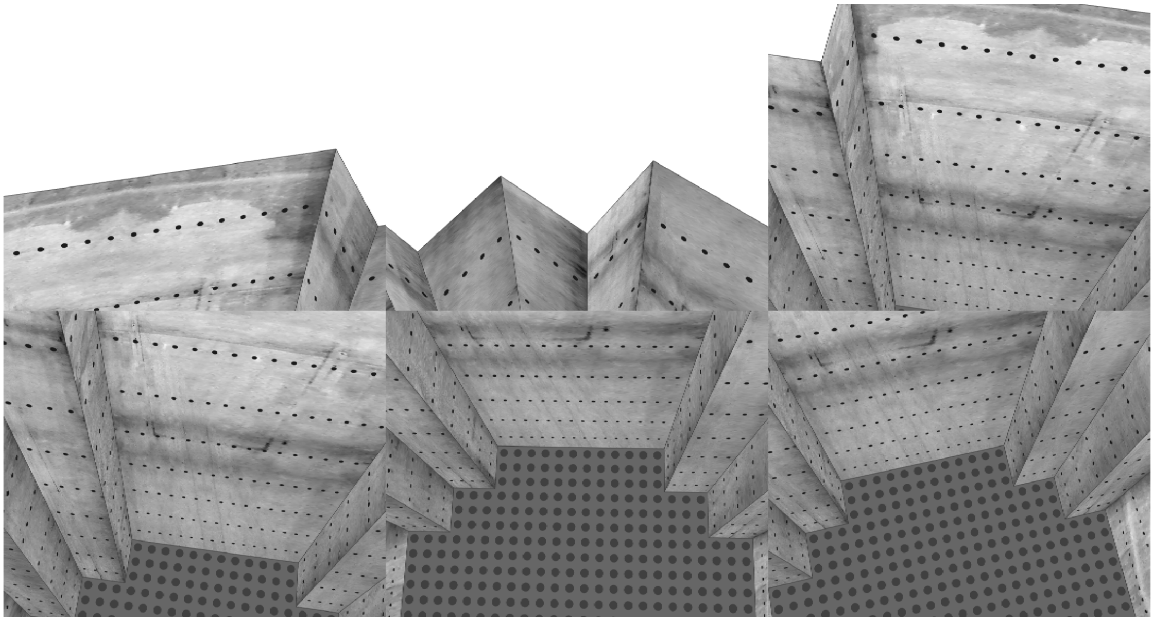


Figure 3.6: Frames from a simulated camera zoom sequence are used to estimate the extrinsic parameters of the PTZ camera. The upper left and lower right images correspond to the camera viewing the blue and green circles of Fig. 3.5, respectively.

Although the scene orientation is unknown (so the direction of the rays in the world frame are unknown), the *relative* difference between these rays — and the angle between them — will be constant between the scene and world frames. Thus, we can provide an additional constraint in the world frame based on the triangle angles.

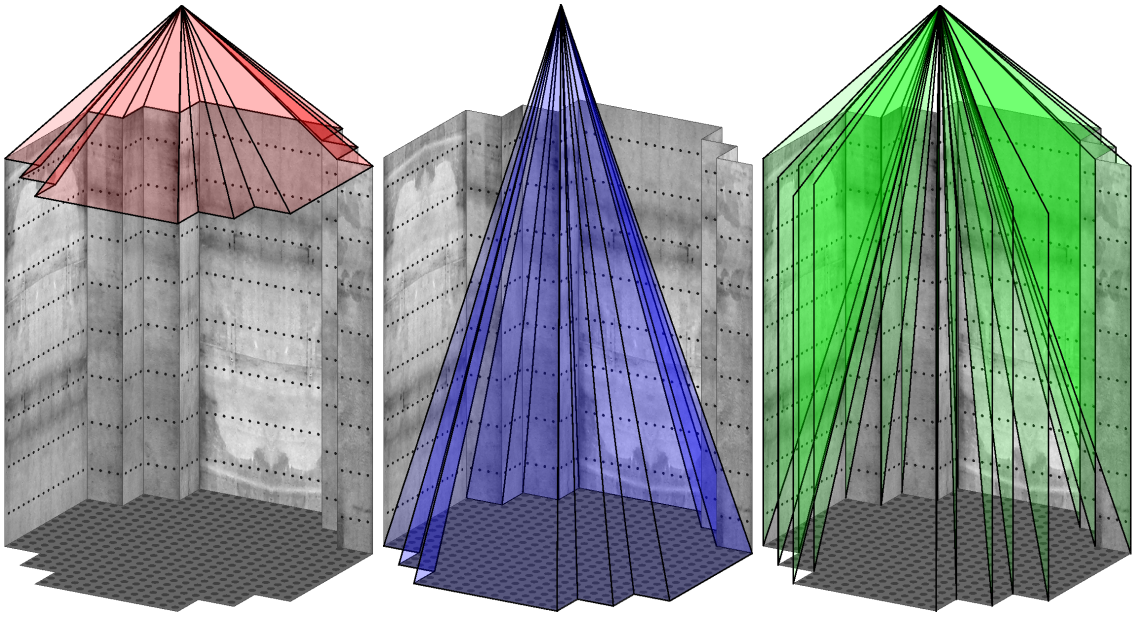
The angles between the triangle sides are as follows:

$$\cos(\theta_{12}) = \frac{-\mathbf{v}_1 \cdot \mathbf{v}_2}{||\mathbf{v}_1||\,||\mathbf{v}_2||} \tag{3.76}$$

$$\cos(\theta_{23}) = \frac{-\mathbf{v}_2 \cdot \mathbf{v}_3}{||\mathbf{v}_2||\,||\mathbf{v}_3||} \tag{3.77}$$

$$\cos(\theta_{13}) = \frac{-\mathbf{v}_1^s \cdot \mathbf{v}_3^s}{||\mathbf{v}_1^s||\,||\mathbf{v}_3^s||} \tag{3.78}$$

$$\mathbf{v}_i^s = R(I_3, \theta_i, \phi_i)\, K^{-1} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \tag{3.79}$$



(a) (b) (c)

Figure 3.7: Adjacent corner points form a series of triangles with the optical center of the camera. The equations specified by these triangles enable the camera position to be estimated. Triangles are formed by corner points (a) along the top of the lower core; (b) along the intersection between the baffle plates and the core plate; and (c) vertically.

where $\mathbf{v}_i^s$ is the backprojected ray for a given corner point in the scene frame that is imaged at image coordinates $[\, u_i,\, v_i \,]^{\mathrm{T}}$ with pan angle $\theta_i$ and tilt angle $\phi_i$.

Two equations can be formed for each triangle that provide constraints for the camera position:

$$||\mathbf{v}_2||^2 = ||\mathbf{v}_1||^2 + ||\mathbf{v}_3||^2 - 2\,||\mathbf{v}_1||\,||\mathbf{v}_3||\cos(\theta_{13}) \tag{3.80}$$

$$\theta_{12} + \theta_{13} + \theta_{23} = \pi \tag{3.81}$$

where the first equation arises from the law of cosines, and the second equation arises from the condition that the interior angles of a triangle must sum to $\pi$ radians. Each adjacent corner point pair detected from the camera sweep footage produces two

equations. Generating equations for these corner point pairs yields an overconstrained nonlinear system of equations. The least-squares solution to this system is obtained from Levenberg-Marquardt nonlinear optimization. An initial position estimate is provided to the optimization solver that assumes the camera is mounted at the center of the $xy$-plane and forms a right triangle with the reactor walls.

Once the camera position in the world frame is known, the position of the corner points in the scene frame can be determined. Then, the scene orientation is found from calculating the best rigid body transformation that maps the corner points from the world frame to the scene frame [41].

As an alternative, one could instead utilize the Efficient PnP (EPnP) algorithm [42], which solves the perspective-n-point problem for cases when the points are coplanar. After associating projections of the core plate landmarks with the map, using EPnP would provide an estimate of the camera pose. However, it was noticed that due to the depth involved, the camera $z$-position is not robustly estimated. For this reason, the core perimeter method is preferable because observations on the camera position from more than one depth will provide a less fragile estimate.

### 3.6.2 Online Initialization of ROV State

In the second stage of the online initialization methodology, the state is augmented with the ROV pose from inference of the fiducial marks and priors of the camera parameters. We note that this methodology obviates the need for marker association.

We now detail the methodology for online ROV initialization. As priors, we require estimates of camera pose and intrinsics, as well as the position of the ROV markers in the body frame. We also require that all three markers are visible in the image, but do not require that the markers have been associated.

First, given three marker projections, we determine the possible permutations for the association of the projections with the markers in the body frame. For three markers, this is $3! = 6$ permutations. Then, for each permutation, we determine the ROV poses which could generate the projections using the P3P algorithm as developed by Knipe, Scaramuzza, and Siegwart [43]. This algorithm produces in closed form four possible poses for which the P3P problem has a solution. We remove poses which are not physically viable (such as when the ROV is not in the camera frustum). Leveraging the assumption that the ROV is typically level (due to the stabilizing buoyancy), we then exclude solutions for which the ROV pose would have a roll or pitch angle greater than a small heuristic value.

The remaining poses are then assessed on their likelihood using one of two metrics: 1) reprojection error or 2) Mahalanobis distance, which is a normalized measure of the distance away from a belief distribution. With reprojection error, we assess the difference between the predicted and actual marker projections for the given association. For Mahalanobis distance, which is the preferred metric to use, we leverage the state belief distribution to assess the likelihood of measurements that would arise from a particular association:

$$d^2 = (\breve{\mathbf{z}}_{m,\,k} - \hat{\mathbf{z}}_k)^{\mathrm{T}} S^{-1} (\breve{\mathbf{z}}_{m,\,k} - \hat{\mathbf{z}}_k) \tag{3.82}$$

where $\breve{\mathbf{z}}_{m,\,k}$ indicates that the ordering of marker observations depends on the association being assessed.

Following the assessments, we retain the best pose that has the best metric value if and only if it is below a heuristic threshold. We only proceed with online ROV initialization if only one valid solution is found. As will be shown in Sec. 4.4.2, these tests produce good initialization cases, suggesting that these pose tests are

well-formulated.

## 3.7 Summary of State Estimation Framework

In this chapter, we have proposed the state estimation framework by which ROV localization is performed. The framework utilizes a discrete EKF that provides a recursive state estimate of the nonlinear system. The state includes the ROV pose, PTZ camera pose and focal length, and the reactor map. The map is defined using a minimum-parameter representation of the lower core geometry.

The EKF models the temporal evolution of the state through a prediction step using a nonlinear process model. The process model predicts the belief state of the system based on inputs from the PTZ camera – specifically, rotational and zoom odometry. The prediction is then corrected through a measurement step. Salient features in the PTZ camera image are extracted, associated, and utilized by several nonlinear measurement models. These features include projections of the ROV markers, reactor lines, and reactor points. The reactor lines and points serve as landmarks to localize the PTZ camera with respect to the world. ROV-to-PTZ observations (projection of fiducial markers) thereby enable ROV-to-world inference – localization of the ROV with respect to the reactor vessel.

To initialize the EKF efficiently during an inspection deployment, as could be done in an automated fashion or with manual inputs that do not require expert knowledge of the EKF, we propose an online initialization methodology. This methodology bootstraps the filter from a series of camera control sequences. The first step initializes the EKF without the ROV. To estimate the initial EKF parameters, the camera conducts a zoom sweep and then a core perimeter sweep. The zoom sweep enables estimates of the principal point, and the extrinsic parameters are estimated from the core perime-

ter sweep. We estimate the focal length from prior knowledge of the camera and prior successful deployments. Once the EKF is online without the ROV, the state is augmented to include the ROV. This ROV initialization algorithm requires prior knowledge of the fiducal marker locations, but does *not* require the marker projections to be associated. The marker association is addressed by assessing the likelihood of the measurements using the Mahalanobis distance for all possible permutations of marker associations. The best association (i.e., yielding measurements that are most likely based on the existing state estimate) is then selected.

This chapter not only proposes a state estimation framework that enables improved inspection through greater efficiency and utility. It also proposes an efficient online initialization procedure that is amenable to being fully automated. The goal of improved reactor inspections is therefore clear in all aspects of the proposed system. In the next chapter, we will present experimental results that demonstrate this framework enables accurate and robust ROV localization.

# Chapter 4

# Results and Discussion

We conduct experiments to demonstrate the correctness, accuracy, and robustness of the state estimation framework. Experimental datasets are representative of actual infrastructure for which the framework was designed. We pursue experiments of three different types: 1) simulation experiments using high-quality synthetic camera datasets to show the correctness of the framework for the possible camera inputs, 2) camera experiments with a subscale mockup infrastructure system, and 3) camera experiments using the inspection system with a to-scale reactor vessel mockup ("platform experiments"). For camera experiments, the robustness of the state estimate is assessed against speckling, which is radiation-induced chromatic image noise as described in Section 2.5. All camera datasets were processed twice: 1) "clean" (no speckling); and 2) "degraded," with artificial speckling and color attenuation to emulate the environmental image effects that are expected when deployed in a nuclear reactor vessel. We also utilize the simulation environment to conduct various analyses to demonstrate the viability of the online initialization methodology.

## 4.1 Simulation Experiments with Synthetic Camera Datasets

We demonstrate the accuracy and uncertainty of the framework in simulation using high-fidelity synthetic camera images for a representative dataset containing realistic ROV motion and PTZ camera commands. The dataset, which is shown visually by Fig. 4.1, consists of the ROV maneuvering to a waypoint within the reactor. The PTZ camera pans, tilts, then zooms in to keep the ROV in view.

Table 4.1 summarizes the relevant parameters used for the simulation, including EKF process and observation noise parameters. Note that these results include the initialization of the ROV pose using the online initialization methodology (Sec. 3.6.2

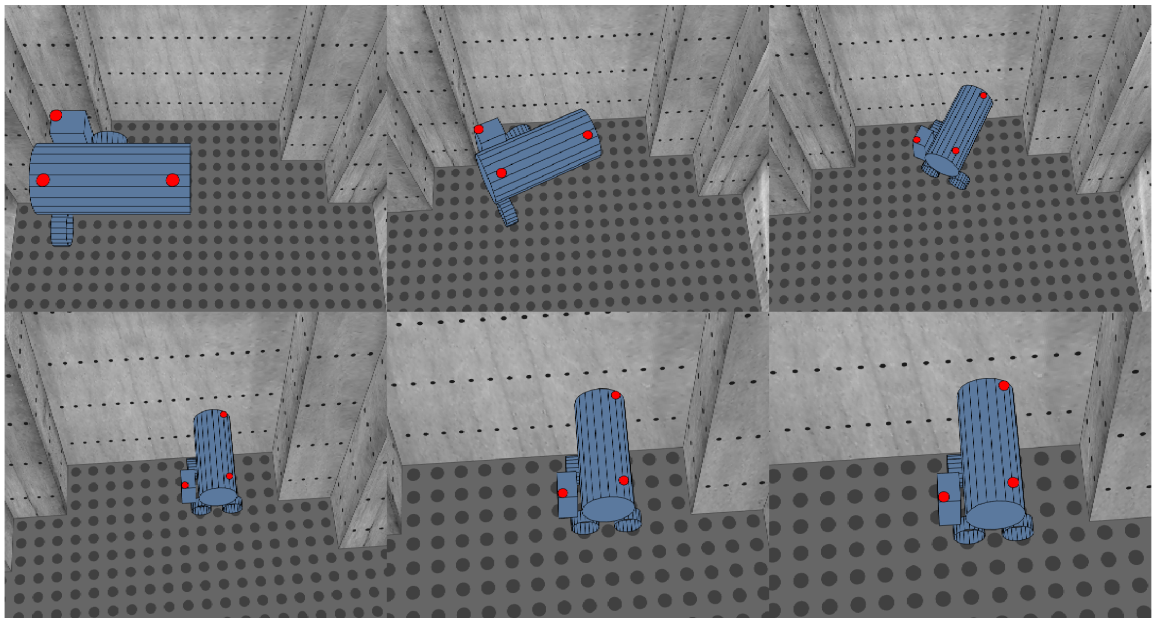


Figure 4.1: Synthetic camera images from a representative dataset. To keep the ROV in view, the PTZ camera pans, tilts, and increases the focal length to approximately 2x zoom. The top-left frame and bottom-right frame correspond to the first and last frames of the sequence, respectively.

Table 4.1: Framework parameters used for simulation results.

| |
|---|
| **EKF initial variance:** |
| ROV position uncertainty: $\sigma^2 = 0.01$ m$^2$ |
| ROV orientation (quaternion) uncertainty: $\sigma^2 = 1\text{e}{-4}$ |
| PTZ camera $x$- and $y$-position uncertainty: $\sigma^2 = 0.01$ m$^2$ |
| PTZ camera $z$-position uncertainty: $\sigma^2 = 6.25\text{e}{-4}$ m$^2$ |
| PTZ camera orientation (quaternion) uncertainty: $\sigma^2 = 1\text{e}{-4}$ |
| PTZ camera focal length uncertainty: $\sigma^2 = 2500$ px$^2$/m$^2$ |
| Vessel height: 1e−12 m$^2$ |
| Vessel baffle plate angles: 1e−12 rad$^2$ |
| Vessel baffle plate distances: 1e−12 m$^2$ |
| Vessel planar landmarks: 1e−12 m$^2$ |
| |
| **EKF process model:** |
| ROV position process noise: $\sigma^2 = 0.0025$ m$^2$ |
| ROV orientation (rotational angle) process noise: $\sigma^2 = 6.25\text{e}{-4}$ rad$^2$ |
| PTZ camera orientation (rotational angle) process noise: $\sigma^2 = 0.01$ rad$^2$ |
| PTZ camera focal length process noise: $\sigma^2 = 0.0025$ px$^2$m$^2$ |
| |
| **EKF measurement model:** |
| Image noise: $\sigma^2 = 2$ px$^2$ |
| ROV marker uncertainty: $\sigma^2 = 1\text{e}{-6}$ m$^2$ |

and c.f. Sec. 4.4.2). Additionally, although the frames in Fig. 4.1 suggest that wall landmarks are viable, we do not utilize these for corrections in this experiment.

The results of the simulation are shown in Fig. 4.2 as time history plots, with the final uncertainty of the system summarized in Table 4.2. Note that we show orientation uncertainty in terms of $ZYX$ Tait-Bryan Euler angles. From Table 4.2, several conclusions can be drawn. First, note the correctness of the filter. Throughout the temporal evolution of the state — ROV motion, PTZ camera motion, and focal length change – the framework produces estimates that are not biased, and no filter divergence is evident. Second, the ROV localization results are sufficiently certain to enable coarse localization of the ROV (within one body length of the robot). Specifically, at the end of the dataset, the $\pm 3\sigma$ uncertainty for ROV localization

within the $xy$-plane is approximately 6 cm, and for $z$-position, the $\pm 3\sigma$ uncertainty is about 10 cm. The camera is well-localized with highly certain estimates. For camera position, this is certain to within 0.7 cm, with most of the uncertainty attributed to $z$-position.

Note that the uncertainty in $x$-position is much less than in $y$-position. This is an outcome of the ROV trajectory and PTZ camera orientation — the camera is pointed along the $y$-axis and the ROV initially moves in the $x$-direction — so this behavior would not hold in general. However, it does indicate that certain dimensions of the ROV position uncertainty will be reduced depending on several factors, including position and orientation of the ROV and the PTZ camera. This suggests that an optimal motion sequence (for both the ROV and the camera) may exist for which the ROV pose uncertainty can be minimized. Furthermore, controlling the ROV and PTZ camera to optimally reduce the system uncertainty could be formulated as an active perception problem. However, this is left for future work. It suffices to show that for these simulation results, the framework is correct with reasonable certainty for representative system inputs and ROV motion.

## 4.2 Camera Experiments with a Subscale Mockup Structure

We perform camera experiments using a subscale mockup system (Fig. 4.3) that is designed to replicate the geometry of a generic reactor pressure vessel on a smaller scale. A frame with markers is used as a mockup for the submersible robot. An Axis V5915 PTZ camera is used for the external camera. We use a VICON motion

Table 4.2: Uncertainty for selected state parameters at conclusion of the simulation dataset.

| Parameter | Uncertainty ($\pm 3\sigma$) |
|---|---|
| Robot, $\{B\}$ | |
| $x_b^w$ | 1.2209 cm |
| $y_b^w$ | 5.8174 cm |
| $z_b^w$ | 9.7990 cm |
| $\theta_b^w$ | 0.0756 rad |
| $\phi_b^w$ | 0.0666 rad |
| $\psi_b^w$ | 0.0269 rad |
| External camera, $\{E\}$ | |
| $x_e^w$ | 0.0772 cm |
| $y_e^w$ | 0.1876 cm |
| $z_e^w$ | 0.6611 cm |
| $\theta_e^w$ | 0.0006 rad |
| $\phi_e^w$ | 0.0008 rad |
| $\psi_e^w$ | 0.0017 rad |

capture system to obtain ground truth pose measurements of the robot. We calibrate the external camera assuming a projective pinhole model and radial and tangential distortion coefficients using the Kalibr calibration toolbox [44].

In this experiment, the robot is translated in a motion that is representative of inspection robot motion. The results of state estimate are shown in Fig. 4.4 and Table 4.3. From ground truth, we calculate that the framework has mean-square error (MSE) in position of under 2.9e$-$4 m$^2$ in $x$, 3.1e$-$4 m$^2$ in $y$, 1.7e$-$3 m$^2$ in $z$. Using $ZYX$ Tait-Bryan Euler angles, the angular position MSE is under 3.8e$-$4 rad$^2$ in roll, 1.3e$-$3 rad$^2$ in pitch, and 6.7e$-$4 rad$^2$ in yaw. Qualitatively, good agreement is observed between the ground truth and estimated paths as shown in Fig. 4.4, and the state estimate is shown to be robust to images degraded by environmental effects.

The uncertainty of the state estimate is also shown in Table 4.3. The $\pm 3\sigma$ uncertainty for the robot within the $xy$-plane is under 4.7 cm and under 10 cm vertically. For the external camera, the lateral $\pm 3\sigma$ uncertainty is 0.5 cm and 4.2 cm vertically.

The relatively higher error and uncertainty in the $z$-direction for both the robot and the external camera is a direct result of the subscale vessel. The subscale vessel by design has no landmarks or visual texture on the walls, which is representative of the most challenging types of reactor vessels for this system. In contrast with this vessel, the to-scale mockup shown in Fig. 4.5 contains landmarks on the walls that improve localization in the $z$-direction. For this reason, we expect that observing wall landmarks or the top edge of the reactor will improve the uncertainty in this dimension. Nonetheless, we note that the error and uncertainty in the $xy$-plane are still suitable for coarse localization of the robot within the vessel.

Table 4.3: Accuracy and uncertainty for subscale mockup testing.

| Parameter | Value (Clean) | Value (Degraded) |
|---|---|---|
| Accuracy (MSE), $m^2$ or $rad^2$ | | |
| $x_b^w$ | 2.7515e−4 | 2.8637e−4 |
| $y_b^w$ | 3.0282e−4 | 3.0809e−4 |
| $z_b^w$ | 1.6192e−3 | 1.6982e−3 |
| $\theta_e^w$ | 3.6473e−4 | 3.7681e−4 |
| $\phi_e^w$ | 1.1025e−3 | 1.2869e−3 |
| $\psi_e^w$ | 6.0344e−4 | 6.7328e−4 |
| | | |
| Uncertainty ($\pm 3\sigma$), m or rad | | |
| Robot, $\{B\}$ | | |
| $x_b^w$ | 0.0435 | 0.0434 |
| $y_b^w$ | 0.0174 | 0.0175 |
| $z_b^w$ | 0.0961 | 0.0960 |
| $\theta_b^w$ | 0.1361 | 0.1358 |
| $\phi_b^w$ | 0.1101 | 0.1118 |
| $\psi_b^w$ | 0.0527 | 0.0520 |
| External camera, $\{E\}$ | | |
| $x_e^w$ | 0.0028 | 0.0037 |
| $y_e^w$ | 0.0021 | 0.0034 |
| $z_e^w$ | 0.0393 | 0.0424 |
| $\theta_e^w$ | 0.0053 | 0.0056 |
| $\phi_e^w$ | 0.0048 | 0.0051 |
| $\psi_e^w$ | 0.0162 | 0.0172 |

## 4.3 Platform Experiments with Inspection System

Next, we demonstrate the capability for estimating the state of the inspection system platform. We perform motion experiments with the inspection system platform hoisted to a crane and translated relative to a (to-scale) reactor vessel quarter mockup (Fig. 4.5).

We present the results for a 26-second test where the robot was translated vertically. The external camera rotates during this experiment to keep the ROV in view. Figure 4.6 shows the estimated path of the robot for this test. As shown in Table 4.4, the framework estimates the pose of the robot with sufficient uncertainty for coarse localization of the robot within the vessel. Specifically, within the $xy$-plane, the robot uncertainty ($\pm 3\sigma$) is under 1.8 cm; the uncertainty in the $z$-direction is 2.4 cm. For the external camera, the $\pm 3\sigma$ uncertainty is under 0.5 cm within the $xy$-plane and under 1.9 cm vertically, with total rotational uncertainty (in terms of Euler angles) to be approximately 0.01 rad.

We note that the uncertainty estimates were lower overall for the platform experiments as compared to the subscale mockup experiments, due to utilizing wall landmarks on the vessel mockup. Additionally, as in the subscale mockup experiments, we observe that the framework is robust to image degradation effects, with little significant effect on state uncertainty. Although ground truth position data is not available for this experiment, cross-referencing the estimated position against images from a camera installed on-board the robot suggests good agreement between the actual and estimated path.

## 4.4 Online Initialization Results

Having demonstrated the accuracy of the localization and robustness to image degradation from environmental effects, we now establish the correctness of the online initialization methodology detailed in Sec. 3.6. We perform experiments using a ROV and PTZ camera simulation, which can generate high-fidelity synthetic camera

Table 4.4: Uncertainty for platform experiments.

| Parameter | Value (Clean) | Value (Degraded) |
|---|---|---|
| Uncertainty ($\pm 3\sigma$), m or rad | | |
| Robot, $\{B\}$ | | |
| $x_b^w$ | 0.0085 | 0.0084 |
| $y_b^w$ | 0.0161 | 0.0162 |
| $z_b^w$ | 0.0241 | 0.0237 |
| $\theta_b^w$ | 0.0687 | 0.0727 |
| $\phi_b^w$ | 0.0395 | 0.0406 |
| $\psi_b^w$ | 0.0214 | 0.0225 |
| External camera, $\{E\}$ | | |
| $x_e^w$ | 0.0031 | 0.0031 |
| $y_e^w$ | 0.0036 | 0.0036 |
| $z_e^w$ | 0.0166 | 0.0190 |
| $\theta_e^w$ | 0.0035 | 0.0039 |
| $\phi_e^w$ | 0.0033 | 0.0035 |
| $\psi_e^w$ | 0.0092 | 0.0097 |

datasets.

### 4.4.1 Monte Carlo Analysis for Online PTZ Camera Initialization

A Monte Carlo analysis is performed to assess the accuracy of the online initialization methodology for the PTZ camera for a variety of possible parameter values.

For each of 200 runs, the camera parameters were selected from the distributions described in Table 4.5, where $\mathcal{U}$ is a uniform distribution. Figure 4.7 shows the initial position and scene orientation for these 200 cases. The variation in camera position is selected from viable mounting locations for the camera in the reactor. We assume that the camera is roughly level, with neither roll nor pitch angle exceeding 0.1745 rad (10 deg). The location of the reactor corners were corrupted by Gaussian noise with zero mean and variance 1e$-$5 m$^2$ to account for limited uncertainty in the prior

used for the reactor geometry.

Table 4.5: Monte Carlo simulation parameters.

| |
|---|
| Camera intrinsics |
| $r \sim \mathcal{N}(\mu = 0 \text{ px}, \sigma^2 = 50 \text{ px}^2)$ |
| $\theta \sim \mathcal{U}(-\pi \text{ rad}, \pi \text{ rad})$ |
| $c_x = r \cos(\theta) + 360$ |
| $c_y = r \sin(\theta) + 288$ |
| $f_x \sim \mathcal{N}(\mu = 1160 \text{ px/m}, \sigma^2 = 1000 \text{ px}^2/\text{m}^2)$ |
| $f_y \sim \mathcal{N}(\mu = 1215 \text{ px/m}, \sigma^2 = 1000 \text{ px}^2/\text{m}^2)$ |
| |
| Camera extrinsics |
| $r \sim \mathcal{U}(0 \text{ m}, 0.5 \text{ m})$ |
| $\theta \sim \mathcal{U}(-\pi \text{ rad}, \pi \text{ rad})$ |
| $x_e^w = r \cos(\theta)$ |
| $y_e^w = r \sin(\theta)$ |
| $z_e^w \sim \mathcal{U}(-0.75 \text{ m}, -2.0 \text{ m})$ |
| $\theta_e^s \sim \mathcal{U}(-0.1745 \text{ rad}, 0.1745 \text{ rad})$ |
| $\phi_e^s \sim \mathcal{U}(-0.1745 \text{ rad}, 0.1745 \text{ rad})$ |
| $\psi_e^s \sim \mathcal{U}(-\pi \text{ rad}, \pi \text{ rad})$ |
| |
| Reactor corner points |
| Corrupted with Gaussian noise $\sim \mathcal{N}(\mu = 0, \sigma^2 = 1\text{e}{-5} \text{ m}^2)$ |

With the ground truth parameters selected for the camera model, zoom and perimeter sweep sequences were simulated to obtain synthetic measurements, which were corrupted with Gaussian noise. Specifically, we assume that all image measurements are corrupted with zero-mean noise with 2 $\text{px}^2$ variance. Measurement of pan and tilt angles were also corrupted with zero-mean noise with 5.0e$-$5 $\text{rad}^2$ variance. These measurements then were utilized by the online initialization algorithms to estimate the camera parameters. Specifically, the principal point is estimated from the zoom sequence, and the camera position and scene orientation (i.e., orientation with respect to the world when pan and tilt are zero) is estimated from the perimeter sweep sequence. The estimated focal length was chosen as a sample from a Gaussian

distribution with the true focal length as the mean with 20 $\text{px}^2/\text{m}^2$ variance; this step emulates that a reasonable prior is utilized from previous system deployments.

The results of the Monte Carlo simulation are shown in Table 4.6. We show the results in terms of root-mean-square error (RMSE) from the ground truth parameters. From the results, several conclusions about the methodology can be drawn. First, the principal point estimation is highly accurate, yielding an estimate with mean error under one pixel. Second, the extrinsic calibration is also accurate, providing estimates that have approximately 1.6 cm error in the $xy$-plane and 0.6 cm error in $z$-position on average. The relatively higher accuracy for the $z$-position is desirable for initializing the filter, as the measurement models described previously tended to produce relatively greater uncertainty in $z$-position. In terms of $ZYX$ Tait-Bryan Euler angles, the scene orientation provides estimates with approximately 0.3 deg of error on average. The outcome suggests that the methodology is correct, and based on previous runs of the EKF, this accuracy is sufficient to initialize the filter. Additionally, no correlation between either camera height and distance from the centerline and either position or rotational error was evident. This suggests that the cylindrical region of mounting positions for the camera shown in Fig. 4.7 is uniformly viable.

Table 4.6: Monte Carlo simulation results for estimating initial parameters for the PTZ camera from the reactor structure.

| Parameter | RMSE |
|---|---|
| $c_x$ | 0.5512 px |
| $c_y$ | 0.2849 px |
| $x_e^w$ | 0.0109 m |
| $y_e^w$ | 0.0119 m |
| $z_e^w$ | 0.0055 m |
| $\theta_e^s$ | 0.0036 rad |
| $\phi_e^s$ | 0.0037 rad |
| $\psi_e^s$ | 0.0019 rad |

### 4.4.2 Online ROV Initialization

We have established the correctness of the EKF initialization procedure for bootstrapping the framework (in a configuration without the ROV) in situ from specific camera sequences and prior knowledge of the reactor geometry. Now, we demonstrate that the ROV initialization methodology described in Sec. 3.6.2 allows for the ROV pose to be augmented to the state without requiring marker observations to be associated. With this step, the filter can be completely bootstrapped in an online fashion from minimal prior information, and this process can be nearly or fully automated without inspection personnel input.

This analysis is conducted in simulation using the dataset shown in Sec. 4.1. For each timestep of the simulation, we apply the algorithms described in Sec. 3.6.2 using the covariance of the PTZ camera states. We reject any marker association yielding a Mahalanobis distance below 1, and only keep those poses which satisfy the physical and heuristic tests for typical ROV states (i.e., near zero attitude).

As shown in Fig. 4.8, the online ROV initialization is shown to be viable for nearly all timesteps. In fact, a viable ROV pose for initialization was found in 93.73% of the timesteps. No timestep produced more than one viable pose. The root-mean-square error (RMSE) for all viable poses is shown in Table 4.7. The relatively high accuracy of the online ROV initialization suggests it is viable as a component for bootstrapping the entire EKF. Furthermore, this methodology can also assess whether the EKF has diverged during operation.

Table 4.7: ROV online initialization results in terms of RMSE for a representative simulation dataset.

| Parameter | RMSE |
|---|---|
| $x_b^w$ | 0.0058 m |
| $y_b^w$ | 0.0301 m |
| $z_b^w$ | 0.0502 m |
| $\theta_b^w$ | 0.0370 rad |
| $\phi_b^w$ | 0.0337 rad |
| $\psi_b^w$ | 0.0141 rad |

## 4.5 Summary of Results

In this chapter, we have demonstrated through simulation and camera experiments that the state estimation framework is correct and viable for ROV localization. For simulation results, we demonstrated the correctness of the filter for representative ROV motion and possible camera inputs (rotation, zoom). For camera experiments, we show that the framework enables coarse ROV localization for a subscale system and the inspection platform, noting that the prior was successful without the use of landmarks on the baffle plates. Lastly, we demonstrate the accuracy of the online EKF initialization methodology. This procedure is powerful as it enables inspection personnel to bring the system online without needing to manually specify initial state values, which would require technical expertise with the system. Instead, through a series of camera sequences and self-calibration algorithms, the EKF can be bootstrapped first in a camera-and-map configuration from structural observations. Then, the state is augmented with the ROV pose when the fiducial markers are detected in the image space. There is no need to manually specify the association, as the methodology determines the association that maximizes the likelihood of the observations based on a series of heuristic tests.

Ultimately, the purpose of the state estimation framework is to improve the effi-

ciency and utility of reactor inspections. We have demonstrated that this framework achieves that goal. Not only have we designed and evaluated this framework, we have also presented a method to bring the framework online efficiently during inspection deployments with minimal input from personnel.

In the final chapter, we will summarize the work and identify several areas of future research.

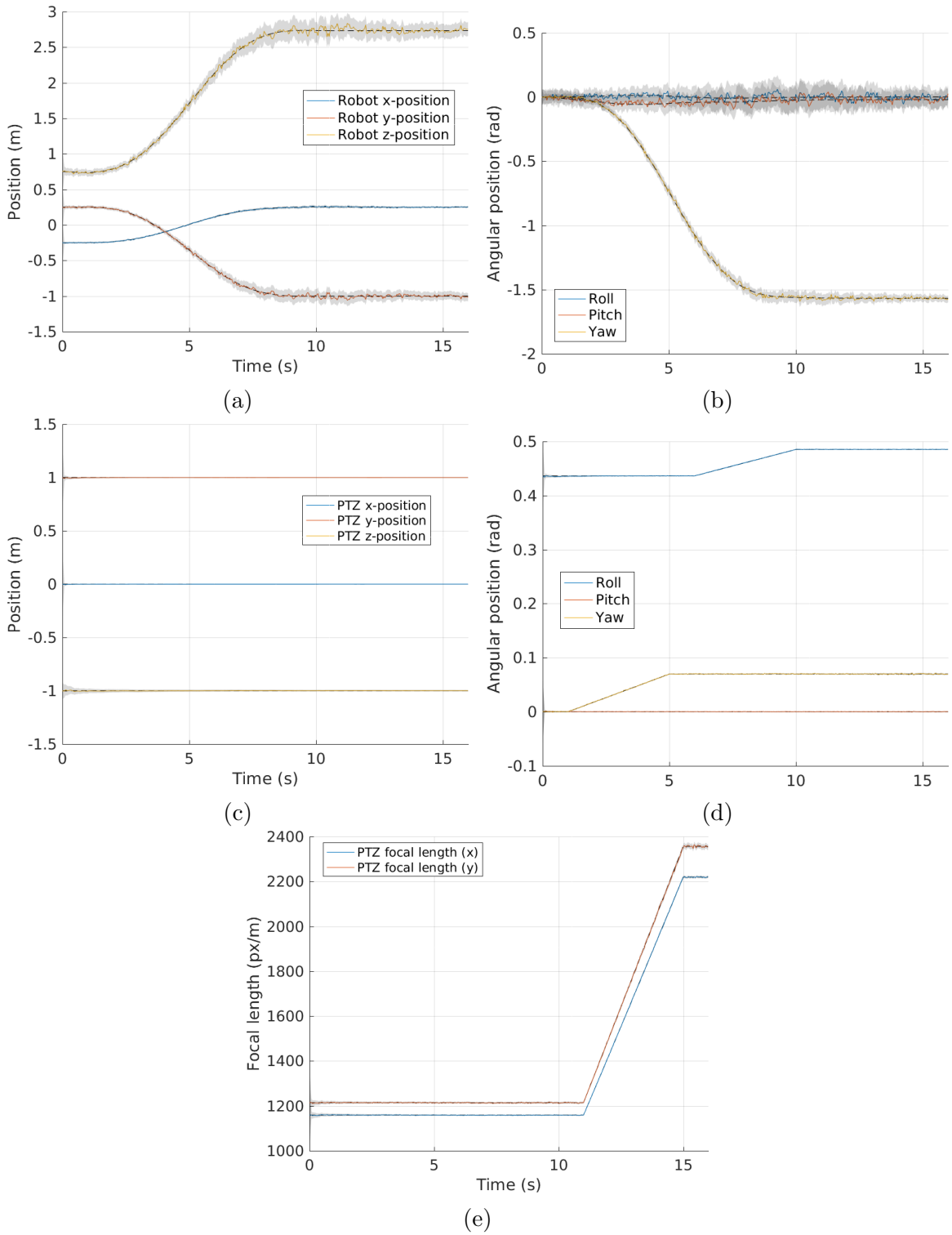



Figure 4.2: Time series simulation results for selected state parameters: (a) ROV position, (b) ROV Euler angle, (c) camera position, (d) camera Euler angle, and (e) camera focal length. The $\pm 3\sigma$ uncertainty is shown for each signal.

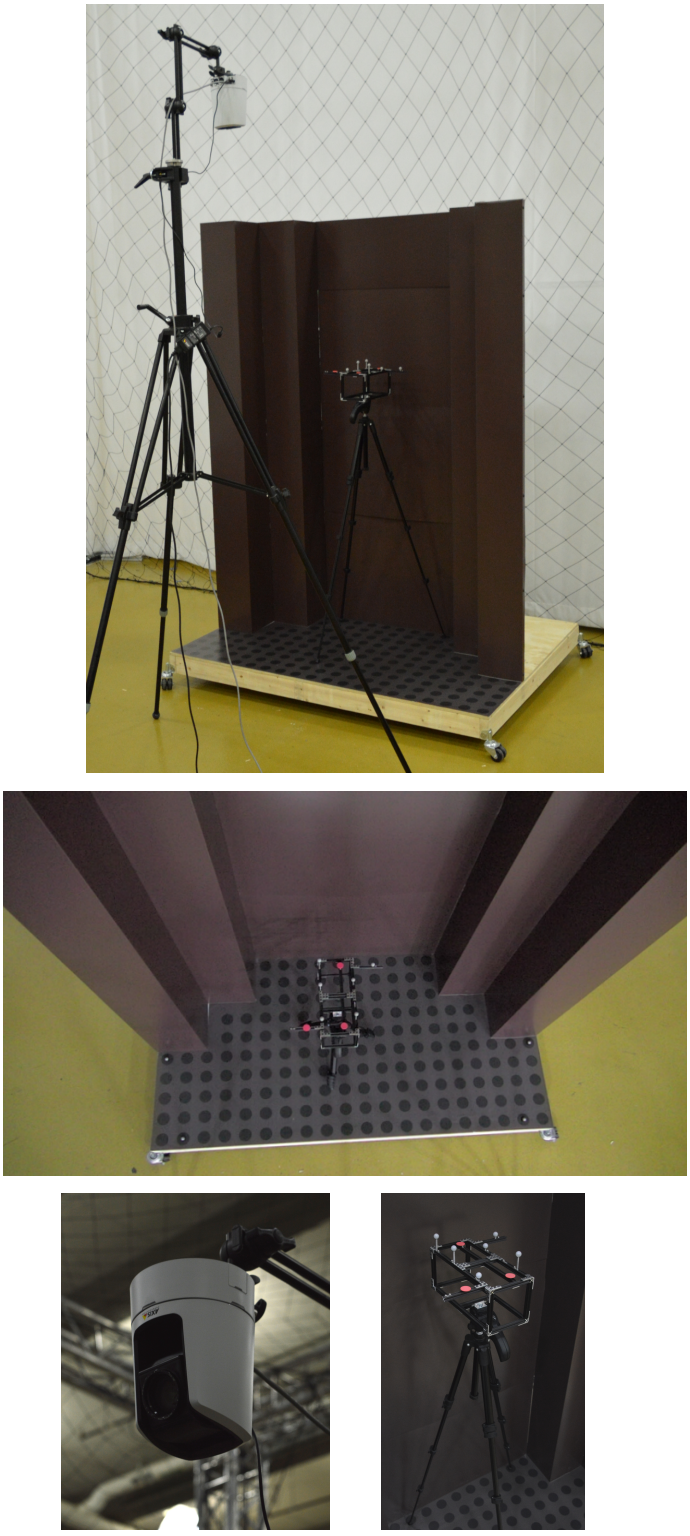


Figure 4.3: Subscale mockup system: (top) experimental setup shown with the subscale mockup of a reactor pressure vessel; (middle) camera image of the subscale system; (bottom left) external camera; (bottom right) subscale mockup of the inspection robot. Note the complete lack of visual texture on the structure.

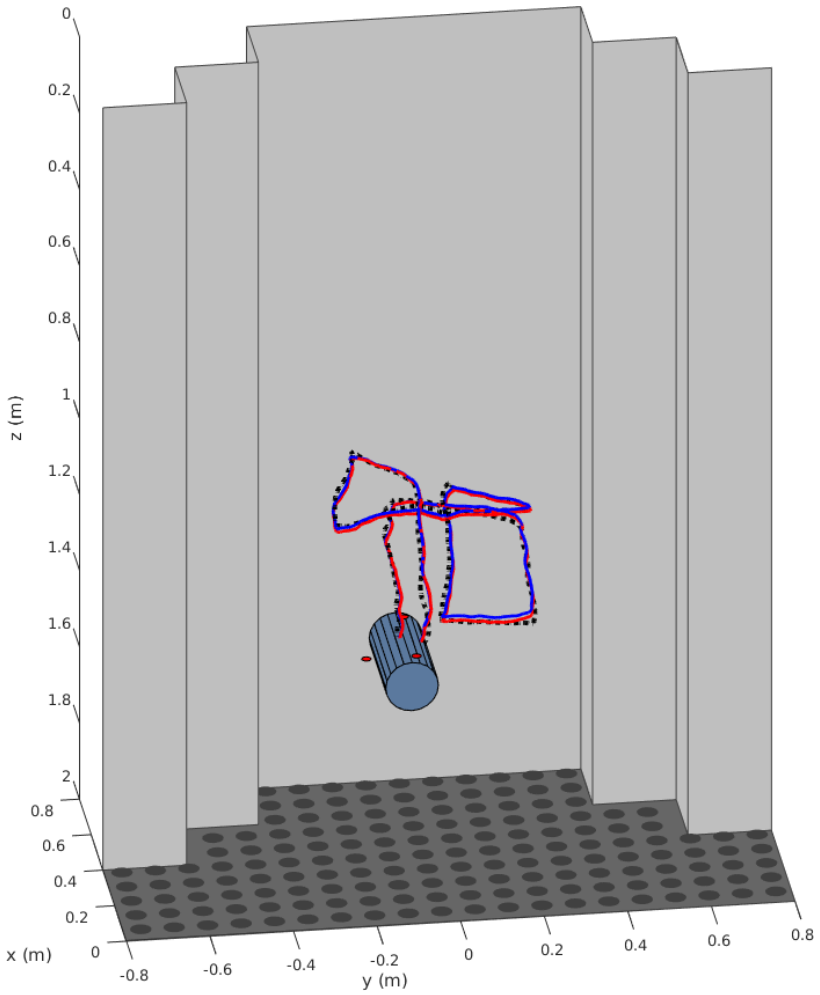



Figure 4.4: Path of the robot relative to the subscale mockup structure. Shown are the estimated paths for the clean (red) and degraded (blue) cases. The ground truth path (black) is from motion capture.

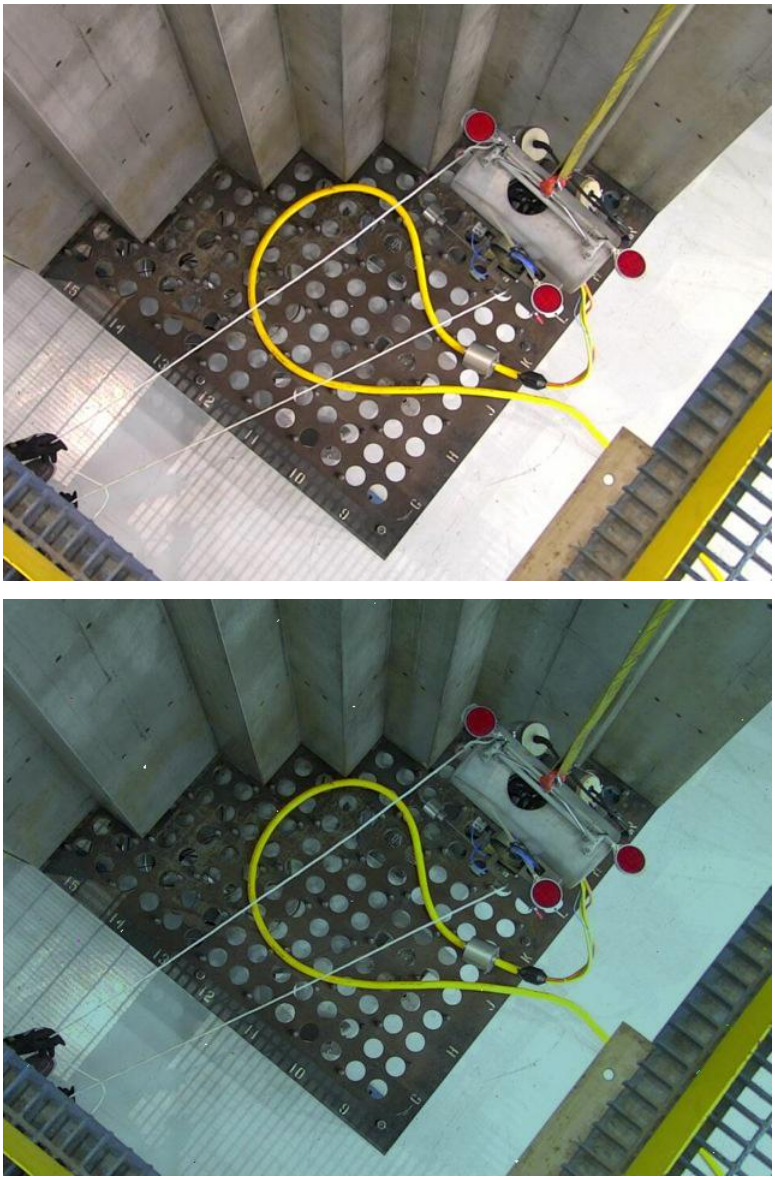


Figure 4.5: External camera images from platform testing (top) clean and (bottom) degraded with artificial environmental image effects (speckling and color attenuation).

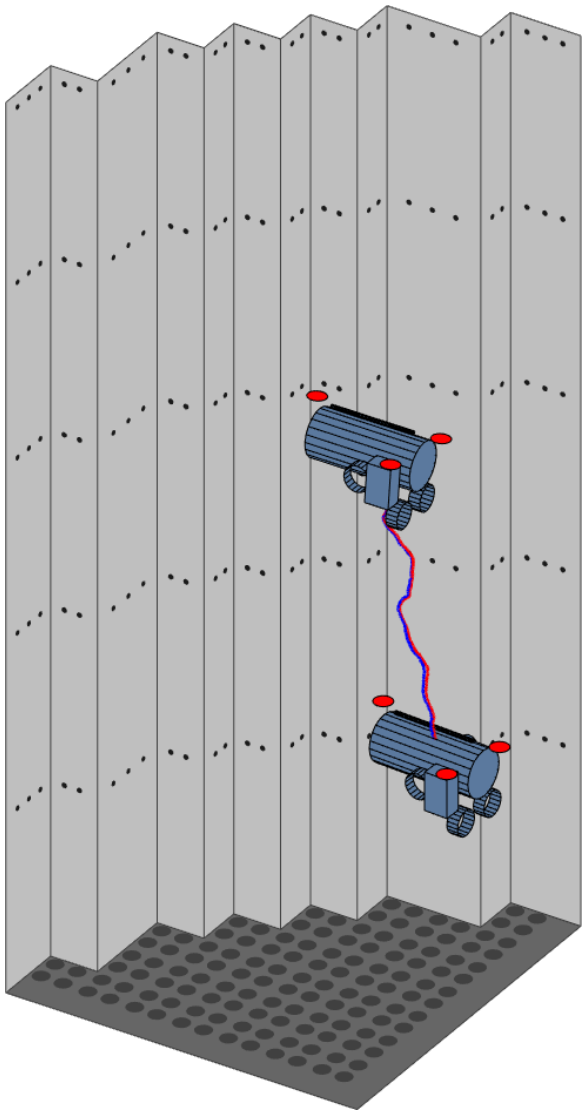


Figure 4.6: Estimated path of the inspection robot relative to a quarter mockup of a reactor vessel for the clean (red) and degraded (blue) cases.

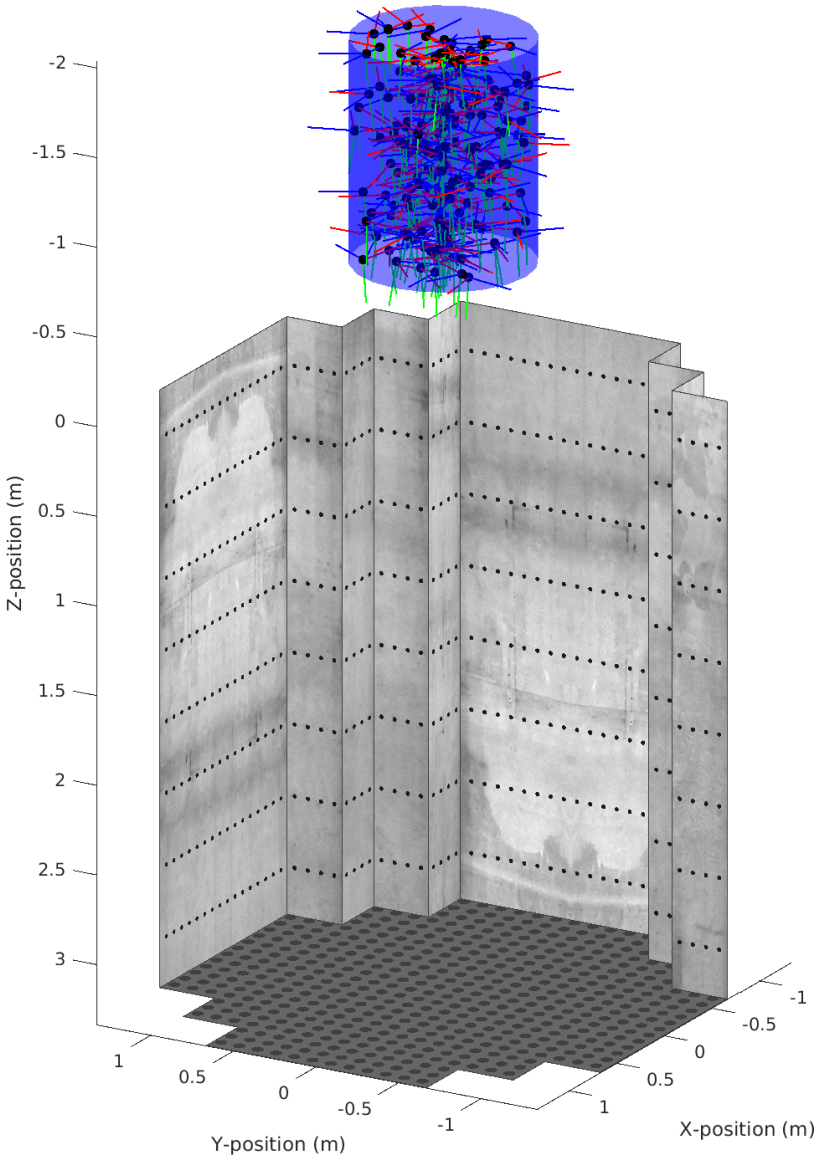
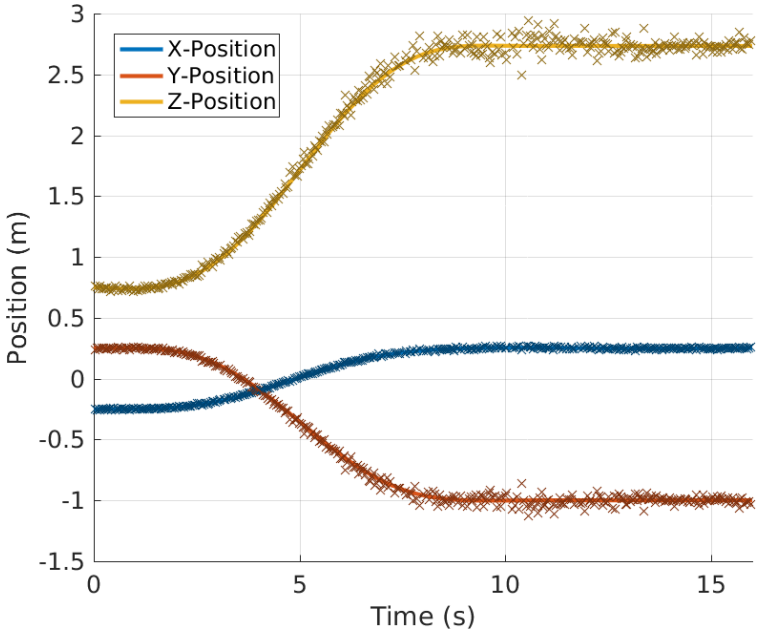
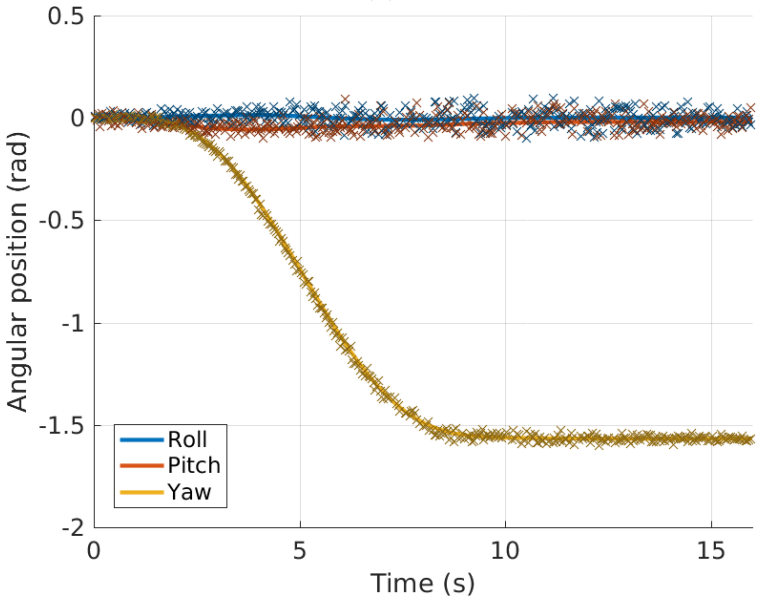



Figure 4.7: The 200 camera instances of the Monte Carlo simulation fall within the blue cylinder, which represents possible mounting positions for the camera during an inspection deployment. The position of each instance is visualized as a black circle, with the scene orientation shown as a frame triad.




(a)




(b)

Figure 4.8: Time series results for ROV online initialization. Each mark represents a viable ROV pose that was determined at that timestep.

# Chapter 5

# Conclusion

## 5.1 Summary and Contributions

In this work, we have proposed a state estimation and localization framework that primarily utilizes a PTZ camera to localize a submersible robot within a nuclear reactor pressure vessel. We have proposed a map representation for reactor pressure vessels that models the structure as a series of adjacent planes, with structural points of interest that exist on these planes. The intersection of vessel planes project to lines in the camera image space. These lines, as well as the points of interest on the planes, serve as landmarks for correcting the state estimate.

We have shown that the proposed framework is suitable for coarse localization by conducting several types of experiments. First, we showed that the framework is correct in a simulation experiment for realistic ROV motion and the possible camera inputs of pan, tilt, and zoom. Then, we assessed the framework performance by conducting camera experiments. We confirmed the accuracy of the filter for localizing a subscale ROV mockup with respect to a challenging vessel with no visual texture

of wall landmarks. Second, we validated the framework using the actual inspection system, showing that the estimated path uncertainty ($\pm 3\sigma$) is under 1.8 cm in the $xy$-plane and 2.4 cm vertically. For the camera, the position uncertainty ($\pm 3\sigma$) is under 0.5 cm in the $xy$-plane and under 1.9 cm vertically, with total rotational uncertainty ($\pm 3\sigma$) of about 0.01 rad. For these camera experiments, we verified that our framework is robust to the environmental effects (speckling and color attenuation) that are expected to degrade the camera images when operating in the reactor vessel.

Lastly, we have detailed an online intialization methodology for the framework. The methodology performs self-calibration of the camera from the reactor structure through a zoom sequence and a perimeter sweep sequence. These sequences are sufficient to estimate the PTZ camera parameters to bring the filter online. Then, the state is augmented with the ROV when the fiducial markers are evident. The ROV initialization component also accounts for the unknown association of markers. This methodology was demonstrated to be viable in simulation results. First, through a Monte Carlo simulation, a variety of possible mounting options for the camera still yielded estimates of sufficient accuracy. Then, for a simulation dataset, the ROV initialization methodology inferred a viable initial pose in nearly all cases. It can be seen that this end-to-end initialization is quite powerful, as it bootstraps the framework with minimum input from inspection personnel. In this way, the online initialization is a critical component of the overall system that seeks to improve the efficiency and utility of ROV-based nuclear reactor inspection.

## 5.2 Improving Localization Performance

As this framework is ultimately intended to enable localization for ROV-based inspection, it is prudent to consider what improvements to the system may yield better

performance in terms of reduced uncertainty. While the existing framework is proven to be sufficient for coarse ROV localization, the level of uncertainty in the system may constrain the types of inspections that may occur in an automated fashion (e.g., the system will need to be upgraded in some capacity to achieve sub-centimeter localization capabilities). To this end, we consider the limiting factors in the system, and address possible options for further reducing system uncertainty.

### 5.2.1 Limiting Factors

We have shown in camera experiments that the framework is robust to environmental image degradation, which consists of speckling and color attenuation. However, as with all vision-only systems, scene lighting is a critical factor in the performance of the system. The camera that is utilized for this system has a relatively large aperture, which provides some robustness to low lighting conditions (at the cost of increased image noise). The limiting factor for lighting is whether the scene illumination is inadequate, and the image observations cannot be detected. In this case, insufficient system observations will certainly lead to either degraded performance or filter divergence. As the environment in which the robot operates is confined, it is recommended to control the scene illumination with sufficient underwater lighting to address this limiting factor.

Additionally, we consider the fact that marker occlusions may occur due to the ROV tether or other inspection equipment. Compensating for marker occlusion is possible through two strategies: 1) utilize a variant of the marker measurement model based on the number of markers that are apparent (e.g., two markers, if one marker is occluded), instead of not processing these observations at all; and 2) remove the ROV from the state if marker occlusion causes a significant degradation in the pose

estimate to no longer be useful, e.g., an extended duration of total marker occlusion. Regarding this second option, the ROV can be added back to the state with the online ROV initialization methodology when the markers are apparent again.

### 5.2.2 System Upgrades

System performance may be improved by upgrades to either the platform or the system algorithms. For the ROV, mounting additional fiducial markers is relatively simple and should yield improved performance. However, improvements may be subject to diminishing returns. The dimension with most uncertainty is the $z$-direction, and viewing the ROV primarily from above (as is currently done with the PTZ camera) will not change that. Instead, it would be prudent to integrate measurements that could be obtained from a sensor on the robot. Although the inspection platform has significant sensor constraints, one such viable sensor is a radiation-hardened, vacuum tube-based analog camera. As it lacks transistors and a CCD image sensor, it can withstand the severe radiation field near the baffle plates. However, the image quality is poor and prone to analog corruption, so this upgrade would not be trivial. However, the advantages are numerous: obtaining measurements from the robot body frame may yield significant increases in localization performance, as now direct ROV-to-world observations would be available. Additionally, this would obviate issues with marker occlusion, as the on-board camera could be used to correct the ROV pose if the PTZ camera lost visibility of the markers.

For software improvements, we may wish to change the ROV motion model. Presently, we model the ROV motion as a random (Brownian) process. However, the use of a constant velocity model [45] may better represent the relatively slow dynamics of the robot, at least in linear position, at the cost of requiring more states.

The possibility of moving the underlying framework to an Unscented Kalman Filter (UKF) or pose graph may yield better performance, but at the tradeoff in computational resources. Both of these frameworks have been shown to provide better performance for nonlinear systems [34]. Here, the EKF is useful as it is amenable to conducting an observability analysis, which we will do for future work.

## 5.3 Future Work

The state estimation and localization framework outlined in this thesis sets the foundation for increased efficiency and utility for conducting inspection of reactor pressure vessels. Specifically, we outline two primary directions for future work: 1) ROV navigation from closed-loop control of ROV position and heading, and 2) generating high-quality mosaics and dense reconstructions of reactor pressure vessels to open completely new avenues for improved inspection without a ROV. Additionally, we will consider implementation of the items addressed in Sec. 5.2, especially conducting an observability analysis of the EKF to better understand exactly which measurements are necessary to render the state fully observable.

### 5.3.1 ROV Navigation

As localization is a fundamental component of mobile robotics, automation of ROV-based inspection of reactor pressure vessels using the proposed framework is now possible. Specifically, we will seek to design time-efficient trajectories and control algorithms that enable the ROV to navigate within the vessel in a closed-loop manner using the proposed state estimation and localization framework.

As an example, we demonstrate how navigation is possible in simulation. As shown in Fig. 5.1, we calculate a trajectory from a start state to a goal state, which is

defined by a position and heading. The trajectory generation seeks to minimize the linear and heading acceleration of the ROV over a specified time horizon. A closed-loop feedback controller using proportional and derivative terms is used to minimize the error between this trajectory and the ROV position by determining the proper ROV thruster inputs. Note that we do not consider control of roll and pitch angle for ROV navigation. The four thrusters on the ROV provide lateral, longitudinal, vertical, and yaw motion. The buoyancy of the platform is designed to keep the ROV level, and so roll and pitch angles are not directly controllable.

The various forces acting on the ROV (such as hydrodynamic [viscous fluid damping], buoyancy, gravitational, and thruster forces) yield the linear and rotational acceleration of the ROV based on a dynamics model within a small discretization of time. Using Euler's method, these accelerations are integrated to model the six degree-of-freedom motion of the ROV [46].

The example of ROV navigation shown in Fig. 5.1 is performed in simulation, but we will now begin to demonstrate navigation with the inspection system. In this way, blending the proposed state estimation framework with trajectory and control algorithms will yield ROV navigation capabilities that will improve the efficiency of reactor pressure vessel inspections by minimizing the time required for the ROV to traverse the vessel.

### 5.3.2 Dense Reconstruction of Reactor Pressure Vessels via Mosaics

We now demonstrate how our state estimation and localization framework can be utilized to generate high-quality, dense reconstructions of reactor pressure vessels from PTZ camera images and priors of the camera intrinsics and orientations.

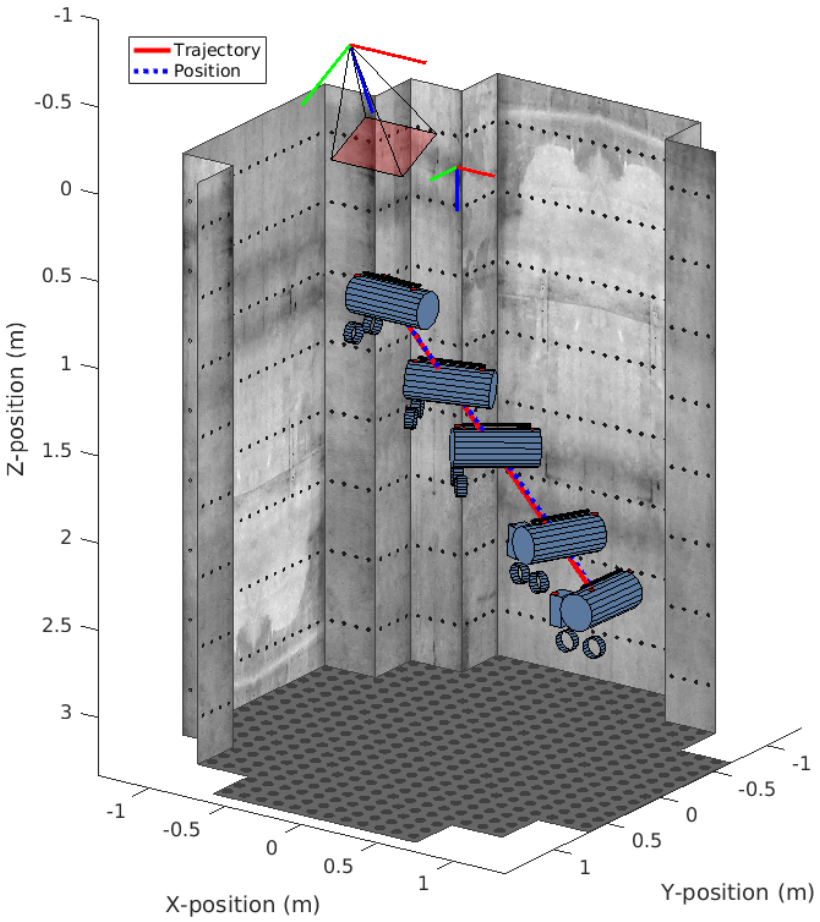



Figure 5.1: An example of how trajectory generation and closed-loop control can enable ROV navigation. Trajectories are generated to guide the ROV to a goal state. A PD controller determines the thruster inputs to control the ROV such that the error between the actual and commanded position and heading is minimized.

### Two-Dimensional Mosaics

Generation of image mosaics (also referred to as panoramas) is possible using a pan-tilt-zoom camera. Because the camera (and therefore the center of projection) does not translate, image pixel displacements do not depend on scene structure [29]. In this way, the pixel coordinates of images $i$ and $j$ from a PTZ camera are always

mapped by a homography $H : \mathbb{P}^2 \mapsto \mathbb{P}^2$ [26]:

$$\widetilde{\boldsymbol{x}}_j = H\,\widetilde{\boldsymbol{x}}_i \tag{5.1}$$

where $\widetilde{\boldsymbol{x}}$ is the homogeneous pixel coordinates, i.e., $\widetilde{\boldsymbol{x}} = [\,u,\, v,\, 1\,]^{\mathrm{T}}$.

Specifically, this homography is the infinite homography $H_\infty$ induced by the plane of rotation at infinity [26]. Camera frames $i$ and $j$ are mapped by the homography $H$, defined by the camera intrinsics $K$ (assumed to be constant in this analysis) and relative rotation $R_{rel}$ of the camera between frames $i$ and $j$ [47]:

$$H = K\,R_{rel}\,K^{-1} \tag{5.2}$$

To construct the mosaic, the homographies are determined between each frame and a common base frame. The base frame is selected to minimize the overall size of the mosaic, which is equivalent to minimizing the warping that occurs from mapping images from the individual frame to the base frame. Generally, it was found that selecting a base frame in the median of the image sequence leads to mosaics that are visually superior.

To demonstrate this methodology, we utilize a dataset that was collected with the inspection system camera. The camera performed a sweep of a lower core mockup that was submerged, pausing at specified pan and tilt angles (Fig. 5.2). The intrinsics were manually estimated to maximize the alignment between adjacent images of the mosaic. The relative rotation $R_{rel}$ was calculated from the reported pan $\theta$ and tilt $\phi$

angles, using a manually estimated $R_{scene}$:

$$R_{rel} = R_j^w \left(R_i^w\right)^{-1} \tag{5.3}$$

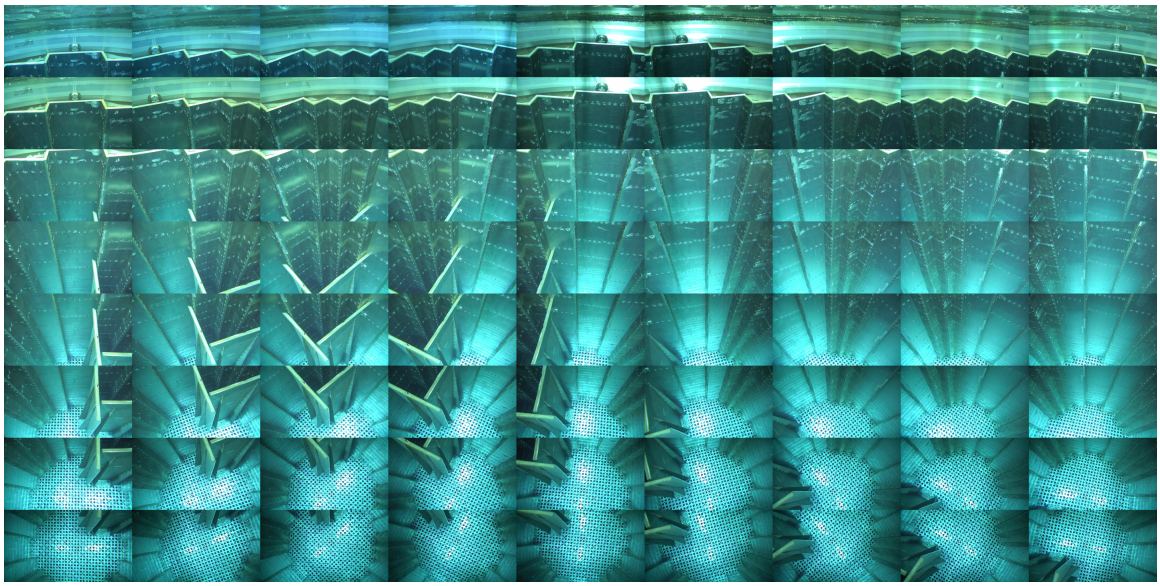
$$R_i^w = R(R_{scene}, \theta_i, \phi_i) \tag{5.4}$$

$$R_j^w = R(R_{scene}, \theta_j, \phi_j) \tag{5.5}$$

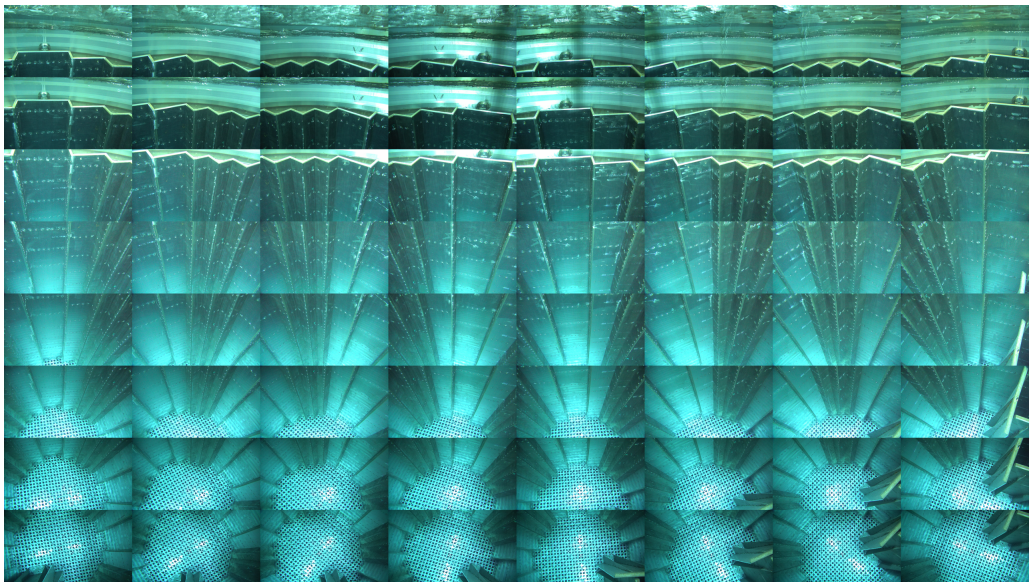
This methodology was utilized to construct various mosaics to model specific areas of the reactor: baffle plates (Fig. 5.3), the core plate (Fig. 5.4), and an overhead view of the entire vessel (Fig. 5.5). Although the mosaics have not been color adjusted (via blending of individual images into the mosaic), and some scene misalignment between individual images may be present due to manual estimation of parameters, these result suffice as a proof-of-concept to show how high-quality mosaics can be constructed for visual assessment and inspection. Furthermore, one could utilize the state estimation framework proposed in this thesis to obtain a better estimate of camera intrinsics and orientations for use in mosaic generation, which would lead to mosaics with better intra-image alignment.

### 5.3.3 Three-Dimensional Dense Reconstruction

The methodology for generating reactor mosaics can be extended to reconstruct the reactor geometry in a high-quality, three dimensional model. To build a dense reconstruction, several mosaics of different vessel surfaces (similar to Figs. 5.3 and 5.4) were assembled, and three-dimensional homographies were used to sample these mosaics to collect color information in order to generate a dense point cloud of the vessel.



(a)



(b)

Figure 5.2: Still images from a dataset where the inspection system camera conducted a visual sweep of a submerged reactor mockup. Rows correspond to decreasing tilt angle in descending order, from $-20°$ to $-90°$ in increments of $-10°$. Columns correspond to increasing pan angle from left to right in increments of 21.25°: (a) $-170°$ to 0°; (b) 21.25° to 170°.

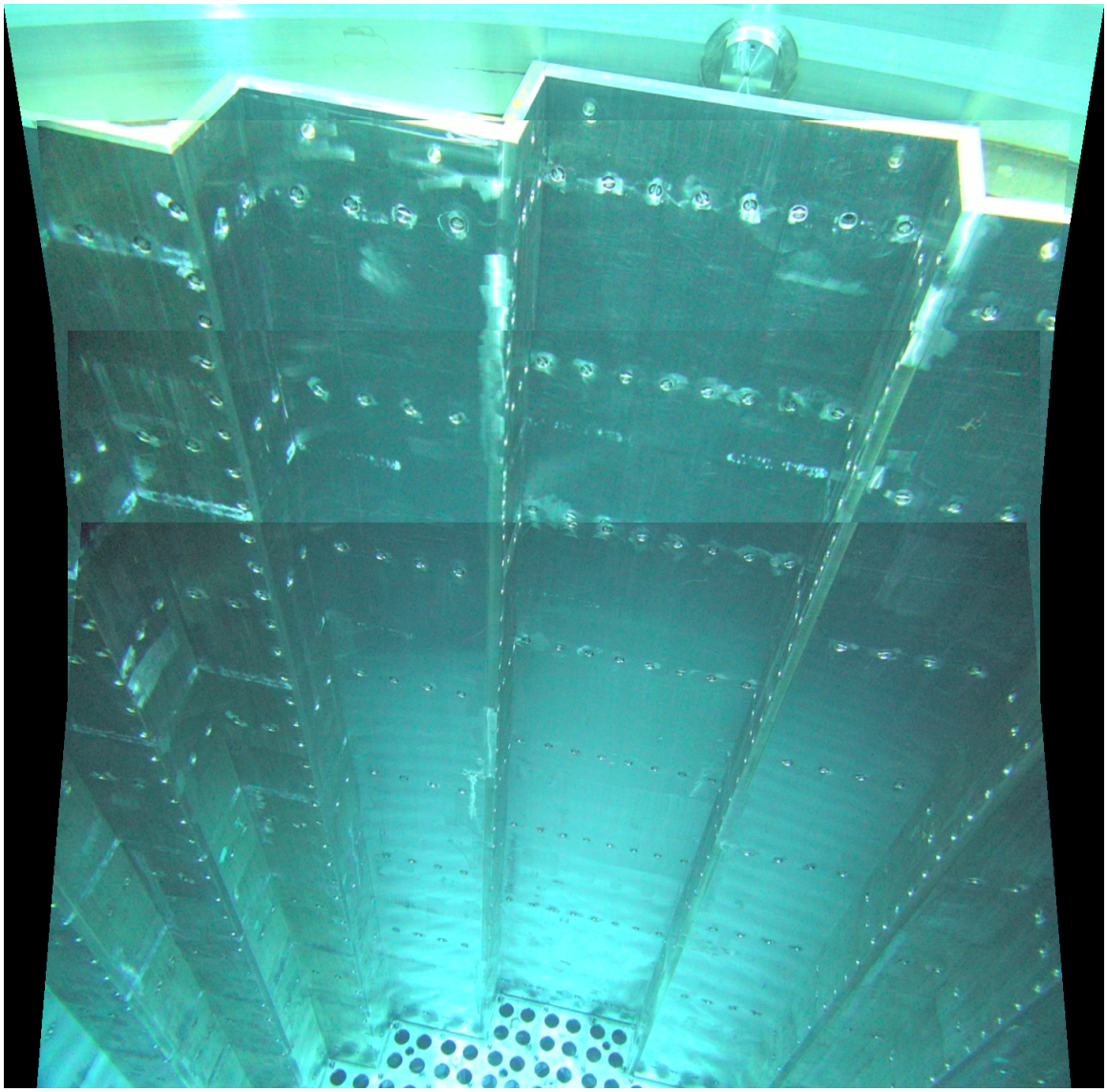


Figure 5.3: A mosaic of several baffle plates of a mockup reactor.

For each planar surface, a three-dimensional homography $H_{3d} : \mathbb{P}^3 \mapsto \mathbb{P}^3$ was used to map from the mosaic space to three dimensions:

$$\widetilde{\boldsymbol{X}}' = H_{3d}\,\widetilde{\boldsymbol{X}} \tag{5.6}$$

where $\widetilde{\boldsymbol{X}}$ is the homogeneous pixel coordinates on the mosaic plane defined by $z = 0$,

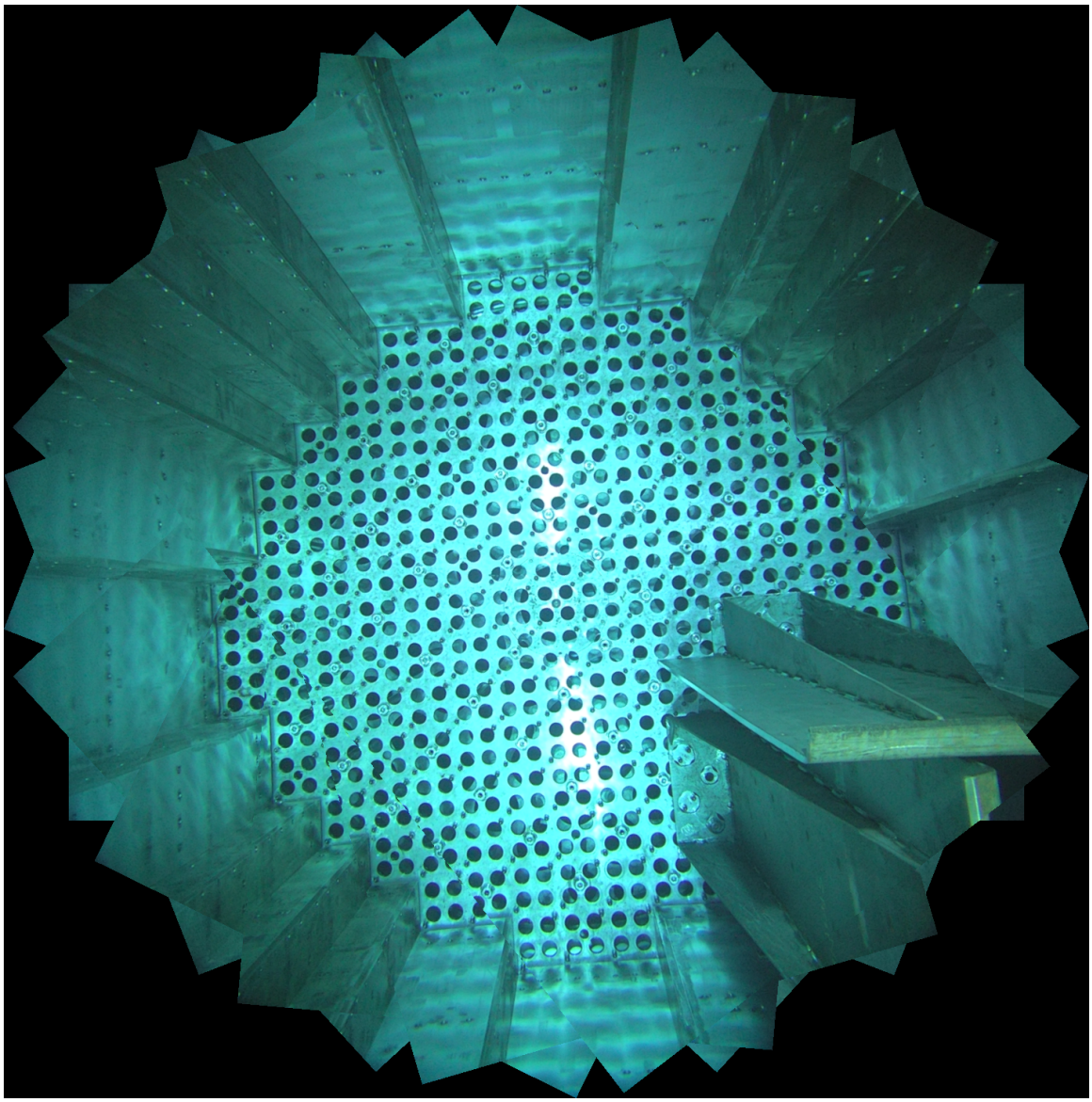


Figure 5.4: A mosaic of the core plate of a mockup reactor. The obstruction that is shown is not typical in operational reactors.

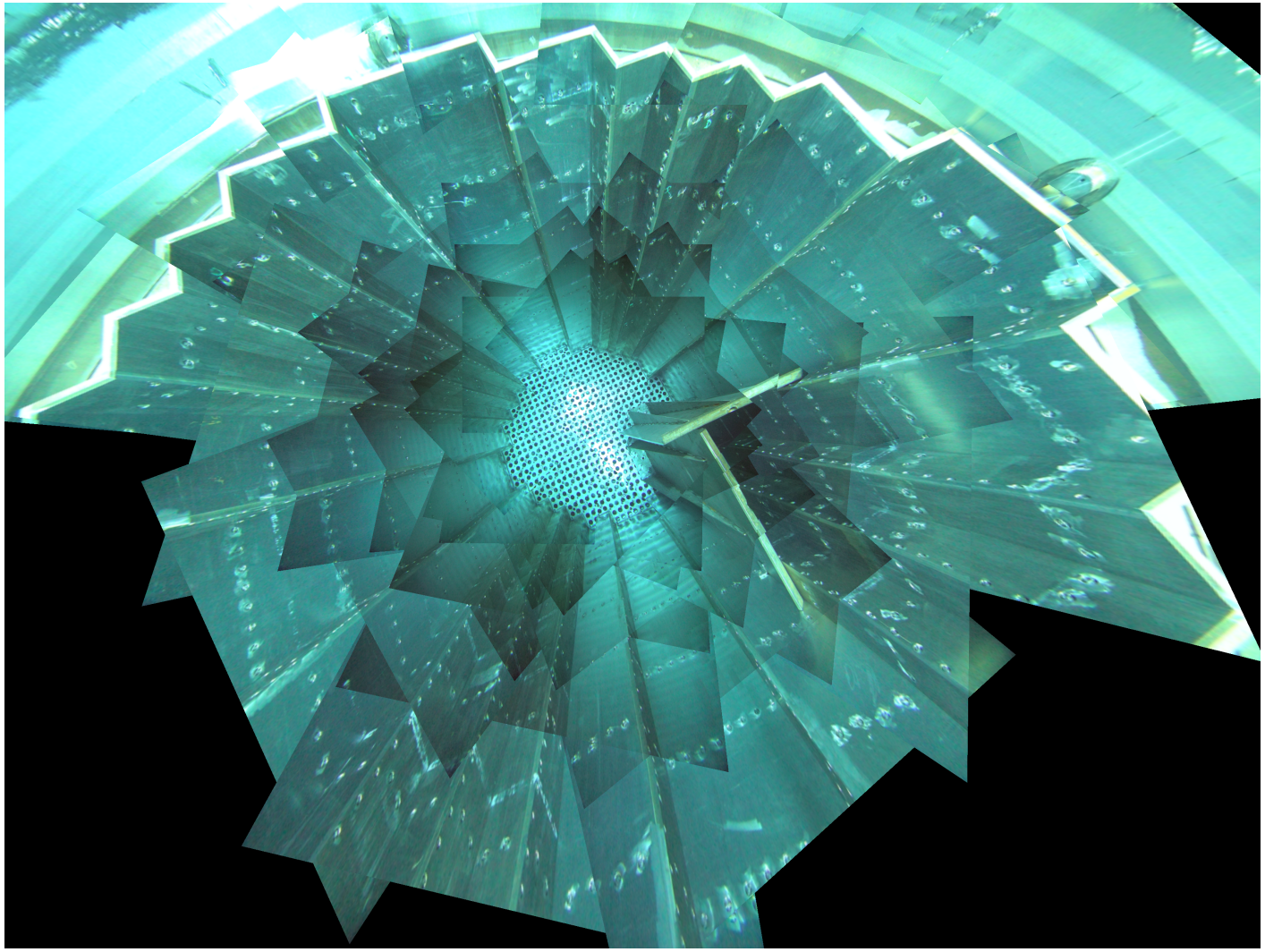


Figure 5.5: A mosaic of the entire mockup reactor. The warping that exists in this mosaic occurs is relatively larger than the other mosaics because the field of view is significantly larger.

i.e., $\widetilde{\boldsymbol{X}} = [\, u,\, v,\, 0,\, 1\,]^{\mathrm{T}}$, and $\widetilde{\boldsymbol{X}'}$ is the mapped point in three dimensions in homogeneous coordinates, i.e., $\widetilde{\boldsymbol{X}'} = [\, x,\, y,\, z,\, 1\,]^{\mathrm{T}}$.

Using the direct linear transform (DLT) algorithm [26], the homography $H_{3d}$ was computed from the correspondences between four corners of a baffle plate that appear in the mosaic and the associated three-dimensional coordinates. (Note that, with only four correspondences, the resulting homography is not unique. However, the mapping is nonetheless suitable for this methodology.) The same process was conducted for the core plate.

In order to produce a dense point cloud that was uniformly distributed in three dimensions, sample points were first generated from three-dimensional discretization of each planar surface. These sample points were then mapped to the mosaic image space using the homography, where the color information was then selected. This process produces a dense, colorized point cloud for each planar surface. Repeating this process for each surface yields the dense reconstruction of the mockup reactor shown in Fig. 5.6. In this way, dense reconstructions of reactor pressure vessels can be utilized for visual assessment and inspection of structural integrity in a similar manner as the vessel mosaics. Although the dense reconstruction contains the same visual fidelity, it also represents it in a three-dimensional form, which may provide better utility for inspection purposes.

The results of the mosaic and dense reconstruction work are compelling. The technologies developed by this work can be leveraged to impact multiple inspection processes, both ROV-based and without a ROV (i.e., visual inspection using mosaics and dense reconstructions). Indeed, the goal of developing new technologies to transform existing inspection processes has been achieved, and we look forward to their deployment in actual inspection campaigns.

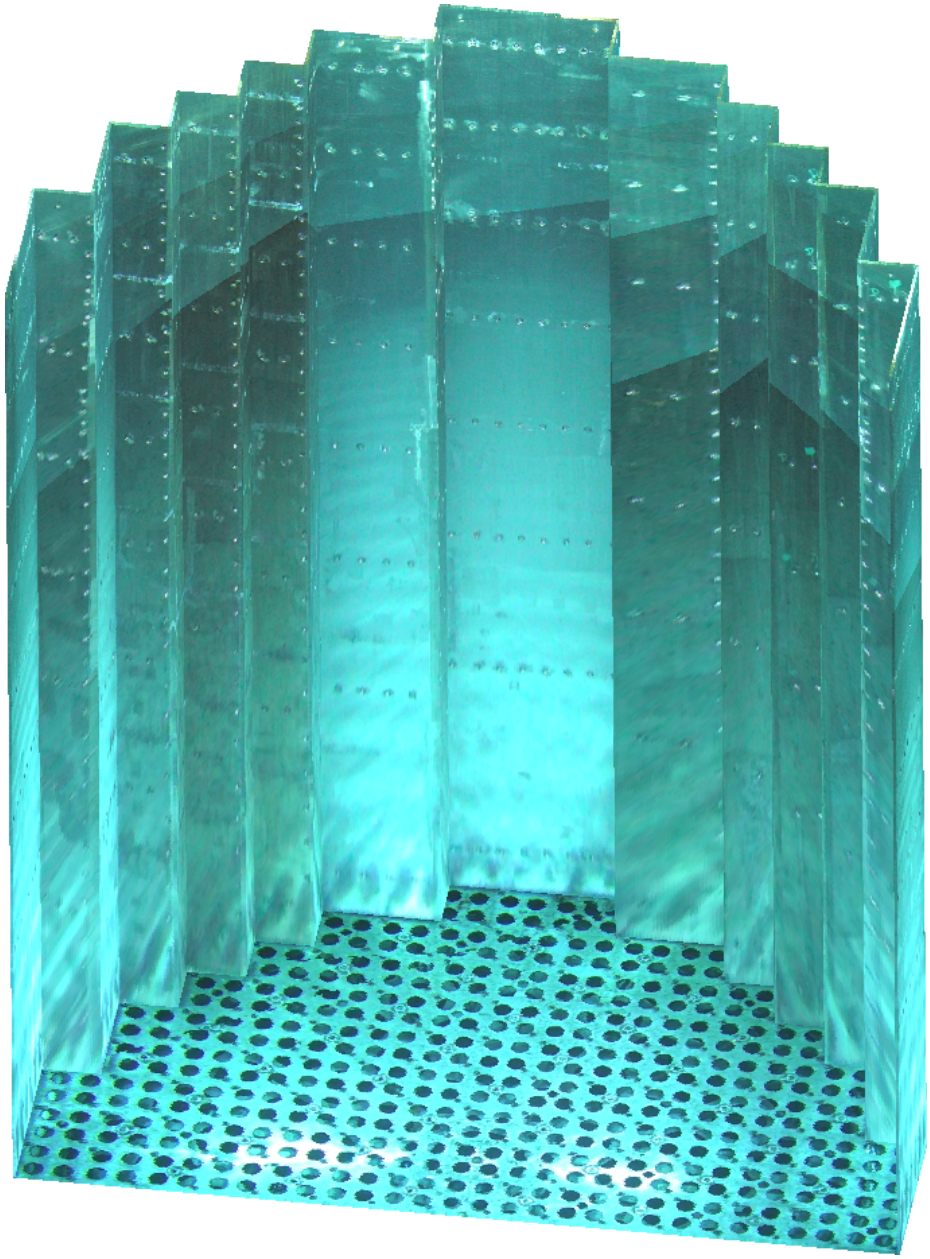


Figure 5.6: A dense reconstruction of a reactor mockup, as generated by mapping the underlying mosaics to three dimensional surfaces.

# Bibliography

[1] U.S. Nuclear Regulatory Commission Technical Training Center, "Reactor Concepts Manual, Chapter 4: Pressurized Water Reactor (PWR) Systems."

[2] U.S. Nuclear Regulatory Commission Human Resources Training and Development, "Westinghouse Technology Systems Manual, Section 3.1: Reactor Vessel and Internals."

[3] Westinghouse Electric Company, LLC, "Your Trusted Reactor Internals Aging Partner." http://www.westinghousenuclear.com/About/News/Features/View/Your-Trusted-Reactor-Internals-Aging-Partner. Accessed: 2017-07-23.

[4] U.S. Nuclear Regulatory Commission, "Backgrounder: Reactor Pressure Vessel Issues." http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/prv.html, February 2016. Accessed: 2017-08-20.

[5] U.S. Energy Information Administration, "Today in Energy (September 28, 2015): U.S. Nuclear Outages Were Less Than 3% of Capacity This Summer." https://www.eia.gov/todayinenergy/detail.php?id=23112. Accessed: 2016-01-18.

[6] S. Cumblidge, M. Anderson, S. Doctor, F. Simonen, and A. Elliot, "A Study of Remote Visual Methods to Detect Cracking in Reactor Components," tech. rep., NUREG/CR-6943, PNNL-16472, US Nuclear Regulatory Commission, Washington, DC, 2007.

[7] R. Alcala, "Application of Non-Destructive Testing and In-service Inspection to Research Reactors," tech. rep., IAEA-TECDOC-1263. International Atomic Energy Agency (IAEA), Vienna, Austria, 2001.

[8] A. Mazumdar, M. Lozano, A. Fittery, and H. H. Asada, "A Compact, Maneuverable, Underwater Robot for Direct Inspection of Nuclear Power Piping Systems," in *Proc. of the IEEE Intl. Conf. on Robot. and Autom.*, pp. 2818–2823, 2012.

[9] M. Odakura, Y. Kometani, M. Koike, M. Tooma, and Y. Nagashima, "Advanced Inspection Technologies for Nuclear Power Plants," *Hitachi Review*, vol. 58, no. 2, pp. 82–87, 2009.

[10] F. Dirauf, B. Gohlke, and E. Fischer, "Innovative Robotics and Ultrasonic Technology at the Examination of Reactor Pressure Vessels in BWR and PWR Nuclear Power Stations," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 42, no. 9, pp. 590–593, 2000.

[11] T. E. Lee, N. Michael, and L. J. Petrosky, "Autonomous Inspection of Nuclear Reactor Pressure Vessels via a Remotely Operated Vehicle," in *American Nuclear Society Decommissioning and Remote Systems (D&RS)*, Pittsburgh, PA, July 31 – August 4, 2016.

[12] H. R. Shea, "Effects of Radiation on MEMS," in *Proc. of SPIE*, vol. 7928, pp. 79280E–1–79280E–13, 2011.

[13] J. C. Kinsey, R. M. Eustice, and L. L. Whitcomb, "A Survey of Underwater Vehicle Navigation: Recent Advances and New Challenges," in *IFAC Conf. of Manoeuvering and Ctrl. of Marine Craft*, vol. 88, 2006.

[14] C.-L. Tsui, D. Schipf, K.-R. Lin, J. Leang, F.-J. Hsieh, and W.-C. Wang, "Using a Time of Flight Method for Underwater 3-Dimensional Depth Measurements and Point Cloud Imaging," in *IEEE OCEANS Conf.*, pp. 1–6, 2014.

[15] A. Kim and R. Eustice, "Pose-graph Visual SLAM with Geometric Model Selection for Autonomous Underwater Ship Hull Inspection," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, pp. 1559–1565, 2009.

[16] F. Shkurti, I. Rekleitis, M. Scaccia, and G. Dudek, "State Estimation of an Underwater Robot using Visual and Inertial Information," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, pp. 5054–5060, 2011.

[17] P. Corke, C. Detweiler, M. Dunbabin, M. Hamilton, D. Rus, and I. Vasilescu, "Experiments with Underwater Robot Localization and Tracking," in *Proc. of the IEEE Intl. Conf. on Robot. and Autom.*, pp. 4556–4561, 2007.

[18] M. Carreras, P. Ridao, R. García, and T. Nicosevici, "Vision-based Localization of an Underwater Robot in a Structured Environment," in *Proc. of the IEEE Intl. Conf. on Robot. and Autom.*, vol. 1, pp. 971–976, 2003.

[19] R. R. Murphy, E. Steimle, M. Hall, M. Lindemuth, D. Trejo, S. Hurlebaus, Z. Medina-Cetina, and D. Slocum, "Robot-Assisted Bridge Inspection," *J. of Intell. & Robot. Syst.*, vol. 64, no. 1, pp. 77–95, 2011.

[20] F. S. Hover, R. M. Eustice, A. Kim, B. Englot, H. Johannsson, M. Kaess, and J. J. Leonard, "Advanced Perception, Navigation and Planning for Autonomous

In-Water Ship Hull Inspection," *The International Journal of Robotics Research*, vol. 31, no. 12, pp. 1445–1464, 2012.

[21] B.-H. Cho, S.-H. Byun, C.-H. Shin, J.-B. Yang, S.-I. Song, and J.-M. Oh, "KeproVt: Underwater Robotic System for Visual Inspection of Nuclear Reactor Internals," *Nuclear Engineering and Design*, vol. 231, no. 3, pp. 327–335, 2004.

[22] D. D. Doyle, A. L. Jennings, and J. T. Black, "Optical Flow Background Estimation for Real-Time Pan/Tilt Camera Object Tracking," *Measurement*, vol. 48, pp. 195–207, 2014.

[23] S. Jain and U. Neumann, "Real-Time Camera Pose and Focal Length Estimation," in *IEEE Intl. Conf. on Pattern Recognition*, vol. 1, pp. 551–555, 2006.

[24] J. Yuh, "Modeling and Control of Underwater Robotic Vehicles," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 20, no. 6, pp. 1475–1483, 1990.

[25] J. Yuh, "Design and Control of Autonomous Underwater Robots: A Survey," *Autonomous Robots*, vol. 8, no. 1, pp. 7–24, 2000.

[26] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[27] Z. Wu and R. J. Radke, "Keeping a Pan-Tilt-Zoom Camera Calibrated," *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. 35, no. 8, pp. 1994–2007, 2013.

[28] A. Bartoli and P. Sturm, "The 3D Line Motion Matrix and Alignment of Line Reconstructions," in *International Journal of Computer Vision*, vol. 57, pp. 159–178, 2004.

[29] R. T. Collins and Y. Tsin, "Calibration of an Outdoor Active Camera System," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 1999.

[30] P. F. Sturm and S. J. Maybank, "On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*, vol. 1, pp. 432–437, 1999.

[31] L. Traffelet, T. Eppenberger, A. Millane, T. Schneider, and R. Siegwart, "Target-based Calibration of Underwater Camera Housing Parameters," in *International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pp. 201–206, 2016.

[32] D. Arthur and S. Vassilvitskii, "K-Means++: The Advantages of Careful Seeding," in *Proc. of the ACM-SIAM Symposium on Discrete Algorithms*, pp. 1027–1035, 2007.

[33] J. J. LaViola, "A Comparison of Unscented and Extended Kalman Filtering for Estimating Quaternion Motion," in *Proceedings of the 2003 American Control Conference*, vol. 3, pp. 2435–2440, 2003.

[34] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT press, 2005.

[35] J. Kelly and G. S. Sukhatme, "Visual-Inertial Sensor Fusion: Localization, Mapping and Sensor-to-Sensor Self-Calibration," *The Intl. Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.

[36] T. E. Lee and N. Michael, "State Estimation and Localization for ROV-Based Reactor Pressure Vessel Inspection," in *Field and Service Robotics 2017*, Zürich, Switzerland, September 12–15, 2017 (to appear).

[37] J. Sola, T. Vidal-Calleja, J. Civera, and J. M. M. Montiel, "Impact of Landmark Parametrization on Monocular EKF-SLAM with Points and Lines," *International Journal of Computer Vision*, vol. 97, no. 3, pp. 339–368, 2012.

[38] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 679–698, 1986.

[39] J. Matas, C. Galambos, and J. Kittler, "Robust Detection of Lines Using the Progressive Probabilistic Hough Transform," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 119–137, 2000.

[40] H. Yuen, J. Princen, J. Illingworth, and J. Kittler, "Comparative Study of Hough Transform Methods for Circle Finding," *Image and Vision Computing*, vol. 8, no. 1, pp. 71–77, 1990.

[41] O. Sorkine, "Least-Squares Rigid Motion Using SVD," *Technical Notes*, vol. 120, no. 3, p. 52, 2009.

[42] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An Accurate O(n) Solution to the PnP Problem," *International Journal of Computer Vision*, vol. 81, no. 2, pp. 155–166, 2009.

[43] L. Kneip, D. Scaramuzza, and R. Siegwart, "A Novel Parametrization of the Perspective-Three-Point Problem for a Direct Computation of Absolute Camera Position and Orientation," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 2969–2976, 2011.

[44] P. Furgale, J. Rehder, and R. Siegwart, "Unified Temporal and Spatial Calibration for Multi-Sensor Systems," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, pp. 1280–1286, 2013.

[45] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software*. John Wiley & Sons, 2004.

[46] T. E. Lee, "Modeling State Estimation, Dynamics, and Control of a Submersible Robot." May 2016.

[47] S. N. Sinha and M. Pollefeys, "Pan–Tilt–Zoom Camera Calibration and High-Resolution Mosaic Generation," *Computer Vision and Image Understanding*, vol. 103, no. 3, pp. 170–183, 2006.