State Estimation and Localization for ROV-Based Reactor Pressure Vessel Inspection

Timothy E. Lee and Nathan Michael

Abstract A vision-based extended Kalman filter is proposed to estimate the state of a remotely operated vehicle (ROV) used for inspection of a nuclear reactor pressure vessel. The state estimation framework employs an overhead, pan-tilt-zoom (PTZ) camera as the primary sensing modality. In addition to the camera state, a map of the nuclear reactor vessel is also estimated from a prior. We conduct experiments to validate the framework in terms of accuracy and robustness to environmental image degradation due to speckling and color attenuation. Subscale mockup experiments highlight estimate 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 was robust to environmental image degradation effects.

1 Introduction

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 pan-tilt-zoom (PTZ) camera, fixed with respect to the vessel frame, that is autonomously controlled. To model the reactor vessel, we propose the use of a sparse 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.

A submersible robot that is used to inspect a nuclear reactor pressure vessel is shown in Fig. 1. The robot and vessel are monitored by an external PTZ camera, which is the primary sensing modality of the framework. The key advantage of us-

Timothy E. Lee and Nathan Michael
Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213
e-mail: {timothyelee, nmichael}@cmu.edu

ing using a camera is its zoom capability. High optical resolution images of the scene are still obtained via zoom despite the camera being positioned relatively farther from the reactor, which mitigates the adverse effects of radiation. Indeed, the existence of radiation in this environment restricts the use of others sensors for the framework. Significant radiation exposure excludes the use of localization sensors with sensitive electronics, such as inertial measurement units or depth sensors. Indeed, radiation-sensitive electronics can degrade and fail with exposure of a few krad [21], which is below the expected exposure dosage in this setting. The underwater setting excludes the use of GPS, and water attenuation excludes sensing depth using projective infrared light without the use of an external light source [25]. However, the underwater setting lacks turbidity, so vision-based perception is viable.

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 [15], 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 [14]. An EKF state estimation formulation that includes vision and inertial measurements was found to be successful in underwater navigation of a submersible robot [22]. Another study demonstrated a localization solution for a AUV using acoustic sensors and visual odometry [7]. 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) [4], 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 [11], 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 [18, 20]. 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) [5]; our work differs by estimating both the robot and camera pose with six degrees of freedom.

Fig. 1 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.
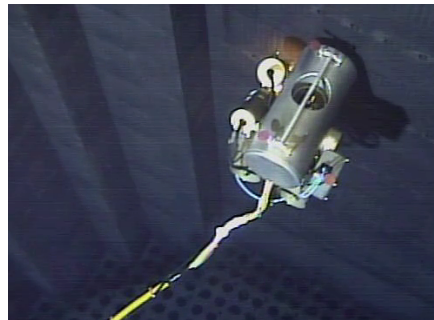
Regarding PTZ cameras, we note that the movement of a small unmanned system with a pan-tilt camera has been estimated using an EKF [8]. This study estimates the projection of the system in the camera image space, not in three dimensions as our framework does. Jain and Neumann [12] employ an EKF to estimate the pose and focal length of a PTZ camera.

2 System Overview

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 illustrates the system and depicts three distinct reference frames:

1. the *body frame* $\{B\}$, located at the robot center of mass;
2. the *external camera frame* $\{E\}$, located at the optical center of the camera; and
3. the inertial *world frame* $\{W\}$, which is the reference frame for the vessel map.

The robot pose, camera pose, and camera focal length are estimated using an EKF. To account for uncertainties in the vessel geometry, the map representation of the vessel is also estimated in the state. The resulting framework is shown in Fig. 3.

The remainder of this section will address system models and methods: the external camera (Sec. 2.1), the submersible robot (Sec. 2.2), the map representation of the vessel (Sec. 2.3), a method for camera rotation inference via homography-based image registration (Sec. 2.4), and a method for incorporation of landmark projections into the sparse map (Sec. 2.5). The EKF formulation detailed in Sec. 3 will leverage these models and methods to enable ROV state estimation.
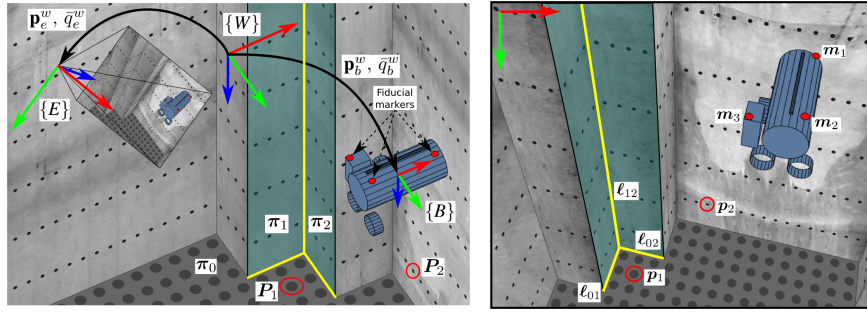

Fig. 2: System representation and landmarks: (left) in three dimensions; (right) in the two dimensional image space of the external camera. The intersections of vessel planes $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_j$ yield Plücker lines that project as lines $\boldsymbol{\ell}_{ij}$. Similarly, three-dimensional points $\boldsymbol{P}_i$ project to two-dimensional points $\boldsymbol{p}_i$ in the image space.
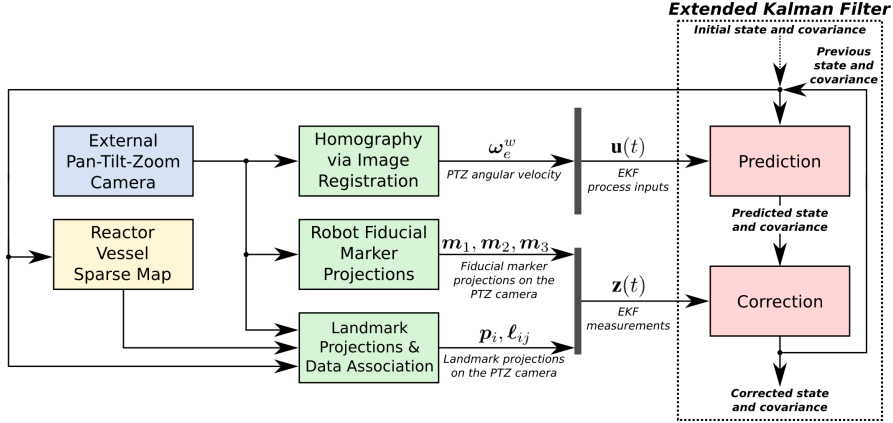
Fig. 3: The system diagram of the state estimation framework.

2.1 External PTZ Camera

A PTZ camera is utilized to monitor the robot and vessel during infrastructure inspection and is mounted to the vessel, external to the robot. The camera is controlled via visual servoing such that the robot is always in view with reasonable magnification. The PTZ images are used for inference of camera rotation (Section 2.4), camera-to-robot localization (Section 2.2), and camera-to-world localization using projections of structural landmarks in the image space (Section 2.5). We assume pinhole projection as the underlying camera model that relates a three-dimensional point in homogeneous coordinates $\widetilde{\boldsymbol{P}} \sim [\boldsymbol{P}^{\mathrm{T}}, 1]^{\mathrm{T}}$ to its image projection in homogeneous coordinates $\widetilde{\boldsymbol{p}} \sim [\boldsymbol{p}^{\mathrm{T}}, 1]^{\mathrm{T}}$, where R and $\mathbf{t}$ represent the transformation from the point frame to the camera frame:

$$\widetilde{\boldsymbol{p}} \sim K[R|\mathbf{t}]\widetilde{\boldsymbol{P}} \qquad (1) \qquad\qquad K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

2.2 Submersible Robot and Fiducial Markers

The submersible robot (Fig. 1) 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 [1] and assigned based on the estimated robot pose.

The position of the markers ($\boldsymbol{M}_i^b$) with respect to the body frame $\{B\}$ is static and known from robot measurements. 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:

$$\widetilde{\boldsymbol{m}}_i \sim K \left[R_w^e \,|\, \mathbf{t}_w^e \right] T_b^w \, \widetilde{\boldsymbol{M}}_i^b \tag{3}$$

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

2.3 Sparse Map from Structural Elements

As shown in Fig. 1, the characteristic geometric appearance of the reactor pressure vessel structure can be described as a series of intersecting planes, with landmarks that exist on these planes. These three-dimensional geometric entities (planes and points) form two types of landmarks in the image space: lines and points. The vessel geometry is specified in the world frame $\{W\}$.

Each plane $\boldsymbol{\pi} = [\bar{n}^{\mathrm{T}}, d]^{\mathrm{T}} \in \mathbb{R}^4$ is described by a unit normal vector $\bar{n}^{\mathrm{T}}$ and distance d. The three-dimensional line that arises from the intersection of two adjacent planes, $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_j$, is represented in Plücker coordinates $\mathcal{L}_{ij} = \boldsymbol{\pi}_i \wedge \boldsymbol{\pi}_j$, where $\mathcal{L} \in \mathbb{P}^5$.

The infrastructure contains landmarks, which are engineered structural elements such as relief holes, cavities, or bolts that can be represented as a three-dimensional point, $\boldsymbol{P}$. Specifically, we note the prominence of repeated point elements such as flow holes on the reactor core floor (Fig. 1, *c.f.* Fig. 5) and bolts (*c.f.* Fig. 8). These landmarks exist on a plane, as represented by the constraint $\boldsymbol{\pi} \cdot \widetilde{\boldsymbol{P}} = 0$.

2.4 Homography-Based Inference of Camera Rotation

To infer the change of the external camera rotation, a homography-based methodology is used that leverages image registration. The pixel coordinates of successive images from the external camera are mapped by a homography [23]:

$$\widetilde{\boldsymbol{x}}' = H \widetilde{\boldsymbol{x}} \tag{4}$$

where $\widetilde{\boldsymbol{x}}$ is the homogeneous pixel coordinates, i.e., $\widetilde{\boldsymbol{x}} = [u, v, 1]^{\mathrm{T}}$. Because the external camera does not translate, image pixel displacements do not depend on scene structure [6]. Specifically, this homography H is the infinite homography H_∞ induced by the plane of rotation at infinity [10]. Between frames i and j, this homography has one of two structures, H_{static} and H_{rot}, depending on whether the camera is static or rotating, respectively:

$$H_{static} = I_{3 \times 3} \tag{5}$$

$$H_{rot} = K R_{ij} K^{-1} \tag{6}$$

The homography, H, is calculated via intensity-based image registration between consecutive frames. The resulting camera rotation R_{ij} is then used to drive the process model of the external camera in the state estimation framework.

2.5 *Projections of Structural Elements*

The projections of structural elements that compose the map are observable from the external camera image space (Fig. 2). Landmarks (points and lines) are used for correcting the state estimate by identifying and associating them in the image space. These projections are utilized for localization between the robot and the world, as well as for localization between the external camera and the world.

Landmarks are identified and associated first using feature detection based on geometric shape. Lines are detected using the Canny filter [3] and the probabilistic Hough transform [17]. Points are detected using blob detection and, in the case of circular elements, the Hough transform [26]. 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.

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

$$\widetilde{\boldsymbol{p}}_i \sim K \left[R_w^e \,|\, \mathbf{t}_w^e \right] \widetilde{\boldsymbol{P}}_i^w \tag{7}$$

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

$$\boldsymbol{\ell}_{ij} \sim \mathcal{K} L_w^e \mathcal{L}_{ij}^w \tag{8}$$

The matrix L_w^e is the rigid displacement matrix for lines, and $\mathcal{K}$ is the perspective projection matrix for lines [2].

3 EKF Methodology

We estimate the non-linear state of the system (robot, camera, and map) using a discrete extended Kalman filter (EKF) with quaternions [16] to obtain a recursive state estimate. This state estimation framework leverages the algorithms described in Section 2 (Fig. 3) and requires as priors the position of the fiducial markers on the robot, 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. To initialize the filter, we optimize the initial position of the external camera and focal length to minimize the reprojection error between the

observed and predicted vessel landmarks. Thereafter, the initial pose of the robot may be estimated from the marker projections from constrained optimization.

3.1 System Parameterization

The non-linear state of the inspection system $\mathbf{X}(t)$ is estimated via an extended Kalman filter with inputs $\mathbf{u}(t) \in \mathbb{R}^3$ 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 $\mathbf{X}_M$:

$$\mathbf{X}(t) = [\mathbf{X}_S(t)^\mathrm{T}, \mathbf{X}_M{}^\mathrm{T}]^\mathrm{T} \tag{9}$$

3.1.1 Inspection System State

The state of the system is represented by

$$\mathbf{X}_S(t) = [\mathbf{p}_b^w(t)^\mathrm{T}, \bar{q}_b^w(t)^\mathrm{T}, \mathbf{p}_e^{w\mathrm{T}}, \bar{q}_e^w(t)^\mathrm{T}, \boldsymbol{f}_e(t)^\mathrm{T}]^\mathrm{T} \tag{10}$$

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.

3.1.2 Map State

The world structure is represented by a map in the state estimate $\mathbf{X}_M$ that encompasses the vessel planes and three-dimensional landmarks that exist on these planes. Each plane $\boldsymbol{\pi} = [\bar{n}^\mathrm{T}, d]^\mathrm{T}$ is described by the unit normal vector $\bar{n}$ and distance d.

We propose a minimal representation for utilizing a map of the reactor in the EKF framework by extending the geometric representation described in Sec. 2.3. Assuming that the walls of the vessel are orthogonal to the floor, planes are specified by their rotational degree of freedom (θ) about the world z-axis and translational degree of freedom (d). Therefore, the unit normal for each wall is $\bar{n} = [\cos\theta, \sin\theta, 0]^\mathrm{T}$. For the floor of the vessel, only the height of the vessel h is needed in the state, 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.

Although landmark points are three-dimensional, they all must exist on a plane. 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, δ_1 and δ_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 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 $\bar{v}_1 = [-\sin\theta, \cos\theta, 0]^T$ and $\bar{v}_2 = [1, 0, 0]^T$ for walls, and $\bar{v}_1 = [1, 0, 0]^T$ and $\bar{v}_2 = [0, 1, 0]^T$ for the floor.

A landmark's three-dimensional position in the world frame $\{W\}$ can be recovered from its coincident plane and two-dimensional position within this plane:

$$P = -\bar{n}d + \delta_1\bar{v}_1 + \delta_2\bar{v}_2 \tag{11}$$

It follows from Eq. 11 that the coplanarity constraint of the landmark, $\boldsymbol{\pi} \cdot \widetilde{P} = 0$, is always satisfied for any choice of θ, d, δ_1, or δ_2.

With this minimal geometric representation, the map state $\mathbf{X}_M \in \mathbb{R}^{2N+1+2L}$ for N planes and L points is as follows:

$$\mathbf{X}_M = [\theta_1, d_1, \ldots, \theta_N, d_N, h, \delta_{1,1}, \delta_{1,2}, \ldots, \delta_{L,1}, \delta_{L,2}]^T \tag{12}$$

3.2 Process Models

The system process model characterizes the temporal evolution of the state. The process input, $\mathbf{u}(t)$, consists of the angular velocity of the external camera with respect to its reference frame, $\boldsymbol{\omega}_e^w(t)$. The entire process model expressed in continuous time is

$$\dot{\mathbf{p}}_b^w(t) = 0 \tag{13}$$
$$\dot{\bar{q}}_b^w(t) = 0 \tag{14}$$
$$\dot{\mathbf{p}}_e^w = 0 \tag{15}$$
$$\dot{\bar{q}}_e^w(t) = \tfrac{1}{2} Q(\boldsymbol{\omega}_e^w(t)) \bar{q}_e^w(t) \tag{16}$$

$$\dot{f}_e(t) = 0 \tag{17}$$
$$\dot{\boldsymbol{\pi}} = 0 \tag{18}$$
$$\dot{P} = 0 \tag{19}$$

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. For the external camera process, the position of the external camera $\mathbf{p}_e^w$ is static. The external camera rotates with an average angular velocity $\boldsymbol{\omega}_e^w(t)$ determined from the output of the homography (Section 2.4). $Q(\boldsymbol{\omega}(t))$ is the quaternion kinematic matrix [13] 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 $\{W\}$.

3.3 Measurement Models

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

$$\mathbf{z}(t) = [\mathbf{z}_b^e(t)^{\mathrm{T}}, \mathbf{z}_e^w(t)^{\mathrm{T}}]^{\mathrm{T}} \tag{20}$$

The body-to-external-camera measurements, $\mathbf{z}_b^e(t)$, are determined through robot fiducial marker detection (Section 2.2):

$$\mathbf{z}_b^e(t) = [\boldsymbol{m}_1(t)^{\mathrm{T}}, \boldsymbol{m}_2(t)^{\mathrm{T}}, \boldsymbol{m}_3(t)^{\mathrm{T}}]^{\mathrm{T}} \tag{21}$$

Projections of structural elements (Section 2.5) provide observations for external-camera-to-world localization and robot-to-world localization. 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.

The predictions for these measurements, $\hat{\mathbf{z}}(t)$, utilize an ideal projective camera model as detailed in Section 2. For points, the correction model is simply the predicted landmark projection in the image space. For lines, we adopt the line error formulation as shown in Fig. 4, which is based on the distance from each point of the detected line to its nearest point on the predicted line [24].

4 Results

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 two different types: 1) camera experiments with a subscale mockup infrastructure system; and 2) platform experiments using the inspection system with a to-scale reactor vessel mockup.

The robustness of the state estimate is assessed against speckling, which is radiation-induced chromatic image noise (Fig. 5). Speckling is characterized by the random occurrence of clusters of pixels to become activated with a high color intensity that persists for only one frame. A probabilistic speckling model was quantified using 549 frames from an inspection dataset with speckling. Table 1 shows the distribution for the number of speckles per frame, n_{frame} (normalized by total number of pixels) and size of the speckle in pixels, s_{size}. All experimental 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.

Fig. 4 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 $\boldsymbol{\ell}$.
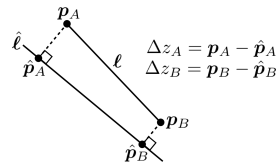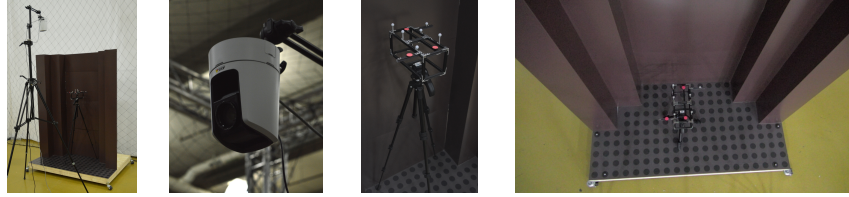
Fig. 6: Subscale mockup system: (far left) experimental setup shown with the subscale mockup of a reactor pressure vessel; (middle left) external camera; (middle right) subscale mockup of the inspection robot; (far right) image of the subscale system from the external camera. Note the complete lack of visual texture on the structure.
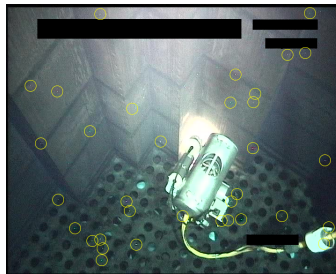

Fig. 5: Speckling (radiation-induced chromatic image noise) observed during a reactor pressure vessel inspection. Speckle clusters are circled in yellow for ease in viewing.

Table 1: Speckling model.

Parameter	Value
$n_{frame} \sim \mathcal{N}(\mu, \sigma^2)$	
μ	6.4655e−5
σ^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

4.1 Camera Experiments with a Subscale Mockup Structure

We perform camera experiments using a subscale mockup system (Fig. 6) 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 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 [9].

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. 7 and Table 2. From ground truth, we calculate that the framework has mean-squared 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. 7, 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 2. 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. 8 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.
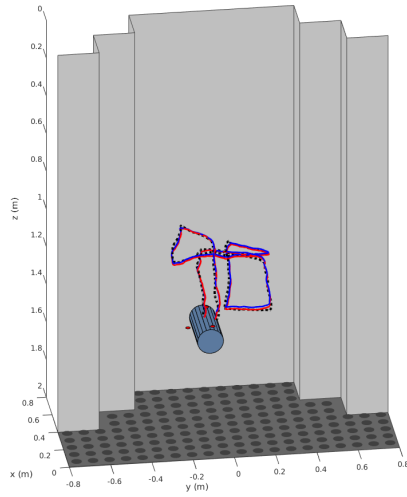

Fig. 7: Path of the robot relative to the subscale mockup structure. Shown are the estimated paths for the clean (red) and degraded (blue) cases. The ground truth path (black) is from motion capture.

Table 2: Accuracy and uncertainty.

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
θ_e^w	3.6473e−4	3.7681e−4
ϕ_e^w	1.1025e−3	1.2869e−3
ψ_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
θ_b^w	0.1361	0.1358
ϕ_b^w	0.1101	0.1118
ψ_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
θ_e^w	0.0053	0.0056
ϕ_e^w	0.0048	0.0051
ψ_e^w	0.0162	0.0172

4.2 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. 8).

We present the results for a 26-second test where the robot was translated vertically by approximately 1.42 m. The external camera rotates during this experiment to keep the ROV in view. Figure 9 shows the estimated path of the robot for this test. As shown in Table 3, the state estimation framework estimates the pose of the

robot with sufficient uncertainty for coarse localization of the robot within the vessel. Specifically, we note that within the *xy*-plane the robot uncertainty ($\pm 3\sigma$) is under 1.8 cm and 2.4 cm vertically. 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.
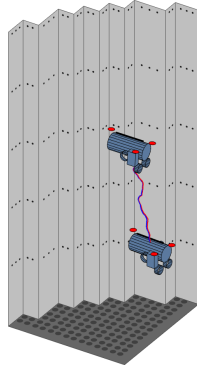

Fig. 9: Estimated path of the inspection robot relative to a quarter mockup of a reactor vessel for the clean (red) and degraded (blue) cases.

Table 3: Uncertainty.

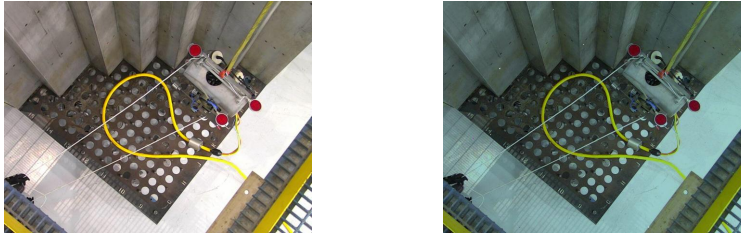| Parameter	Value (Clean)	Value (Degraded)
Uncertainty ($\pm 3\sigma$), m or rad		
Robot, $\{B\}$		
x_b^w	0.0085	0.0084
y_b^w	0.0161	0.0162
z_b^w	0.0241	0.0237
θ_b^w	0.0687	0.0727
ϕ_b^w	0.0395	0.0406
ψ_b^w	0.0214	0.0225
External camera, $\{E\}$		
x_e^w	0.0031	0.0031
y_e^w	0.0036	0.0036
z_e^w	0.0166	0.0190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

Fig. 8: External camera images from platform testing (left) clean and (right) degraded with artificial environmental image effects (speckling and color attenuation).

5 Conclusion and Future Work

In this work, we have proposed an state estimation and localization framework designed for coarse localization of a submersible robot within a nuclear reactor pressure vessel that primarily utilizes a PTZ camera. We have proposed a map representation for reactor pressure vessels that models the struture as a series of orthogonal planes, with structural points of interest that exist on the 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. The rotational motion of the camera is inferred from intensity-based homography.

We have shown that the proposed framework is suitable for coarse localization by conducting two types of experiments. First, we confirmed the accuracy of the filter for localizing the robot 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. We verified that our framework is robust to the environmental image effects (speckling and color attenuation) that are expected to degrade the system sensing when operating in the reactor vessel.

Our current work has shown the capability of this framework for state estimation and localization to enable ROV-based inspections of nuclear reactor vessels. For future work, we will investigate how image-based registration could also be used to model variations in zoom setting of the PTZ camera. Additionally, we will pursue an automated EKF initialization procedure that bootstraps the filter with minimal effort from inspection personnel. We will also pursue testing of our framework in real-time during an inspection of a reactor pressure vessel.

References

1. Arthur, D., Vassilvitskii, S.: K-means++: The advantages of careful seeding. In: Proc. of the ACM-SIAM Symposium on Discrete Algorithms, pp. 1027–1035 (2007)
2. Bartoli, A., Sturm, P.: The 3d line motion matrix and alignment of line reconstructions. In: International Journal of Computer Vision, vol. 57, pp. 159–178 (2004)
3. Canny, J.: A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence (6), 679–698 (1986)
4. Carreras, M., Ridao, P., García, R., Nicosevici, T.: 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)
5. Cho, B.H., Byun, S.H., Shin, C.H., Yang, J.B., Song, S.I., Oh, J.M.: Keprovt: Underwater robotic system for visual inspection of nuclear reactor internals. Nuclear engineering and design **231**(3), 327–335 (2004)
6. Collins, R.T., Tsin, Y.: Calibration of an outdoor active camera system. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1 (1999)

7. Corke, P., Detweiler, C., Dunbabin, M., Hamilton, M., Rus, D., Vasilescu, I.: Experiments with underwater robot localization and tracking. In: Proc. of the IEEE Intl. Conf. on Robot. and Autom., pp. 4556–4561 (2007)
8. Doyle, D.D., Jennings, A.L., Black, J.T.: Optical flow background estimation for real-time pan/tilt camera object tracking. Measurement **48**, 195–207 (2014)
9. Furgale, P., Rehder, J., Siegwart, R.: 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)
10. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2003)
11. Hover, F.S., Eustice, R.M., Kim, A., Englot, B., Johannsson, H., Kaess, M., Leonard, J.J.: Advanced perception, navigation and planning for autonomous in-water ship hull inspection. The International Journal of Robotics Research **31**(12), 1445–1464 (2012)
12. Jain, S., Neumann, U.: Real-time camera pose and focal length estimation. In: IEEE Intl. Conf. on Pattern Recognition, vol. 1, pp. 551–555 (2006)
13. Kelly, J., Sukhatme, G.S.: Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. The Intl. Journal of Robotics Research **30**(1), 56–79 (2011)
14. Kim, A., Eustice, R.: 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)
15. Kinsey, J.C., Eustice, R.M., Whitcomb, L.L.: A survey of underwater vehicle navigation: Recent advances and new challenges. In: IFAC Conf. of Manoeuvering and Ctrl. of Marine Craft, vol. 88 (2006)
16. LaViola, J.J.: A comparison of unscented and extended kalman filtering for estimating quaternion motion. In: American Control Conference, 2003. Proceedings of the 2003, vol. 3, pp. 2435–2440. IEEE (2003)
17. Matas, J., Galambos, C., Kittler, J.: Robust detection of lines using the progressive probabilistic hough transform. Computer Vision and Image Understanding **78**(1), 119–137 (2000)
18. Mazumdar, A., Lozano, M., Fittery, A., Asada, H.H.: 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)
19. Murphy, R.R., Steimle, E., Hall, M., Lindemuth, M., Trejo, D., Hurlebaus, S., Medina-Cetina, Z., Slocum, D.: Robot-assisted bridge inspection. J. of Intell. & Robot. Syst. **64**(1), 77–95 (2011)
20. Odakura, M., Kometani, Y., Koike, M., Tooma, M., Nagashima, Y.: Advanced inspection technologies for nuclear power plants. Hitachi Review **58**(2), 82–87 (2009)
21. Shea, H.R.: Effects of radiation on mems. In: Proc. of SPIE, vol. 7928, pp. 79,280E–1–79,280E–13 (2011). DOI 10.1117/12.876968
22. Shkurti, F., Rekleitis, I., Scaccia, M., Dudek, G.: 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)
23. Sinha, S.N., Pollefeys, M.: Pan–tilt–zoom camera calibration and high-resolution mosaic generation. Computer Vision and Image Understanding **103**(3), 170–183 (2006)
24. Sola, J., Vidal-Calleja, T., Civera, J., Montiel, J.M.M.: Impact of landmark parametrization on monocular ekf-slam with points and lines. International Journal of Computer Vision **97**(3), 339–368 (2012)
25. Tsui, C.L., Schipf, D., Lin, K.R., Leang, J., Hsieh, F.J., Wang, W.C.: Using a time of flight method for underwater 3-dimensional depth measurements and point cloud imaging. In: IEEE OCEANS Conf., pp. 1–6 (2014)
26. Yuen, H., Princen, J., Illingworth, J., Kittler, J.: Comparative study of hough transform methods for circle finding. Image and vision computing **8**(1), 71–77 (1990)