

A Bayesian Modeling Framework for Estimation and Ground Segmentation of Cluttered Staircases

Prasanna Sriganesh¹, Burhanuddin Shirose¹, and Matthew Travers¹

Abstract—Autonomous robot navigation in complex environments requires robust perception as well as high-level scene understanding due to perceptual challenges, such as occlusions, and uncertainty introduced by robot movement. For example, a robot climbing a cluttered staircase can misinterpret clutter as a step, misrepresenting the state and compromising safety. This requires robust state estimation methods capable of inferring the underlying structure of the environment even from incomplete sensor data. In this paper, we introduce a novel method for robust state estimation of staircases. To address the challenge of perceiving occluded staircases extending beyond the robot’s field-of-view, our approach combines an infinite-width staircase representation with a finite endpoint state to capture the overall staircase structure. This representation is integrated into a Bayesian inference framework to fuse noisy measurements enabling accurate estimation of staircase location even with partial observations and occlusions. Additionally, we present a segmentation algorithm that works in conjunction with the staircase estimation pipeline to accurately identify clutter-free regions on a staircase. Our method is extensively evaluated on real robots across diverse staircases, demonstrating significant improvements in estimation accuracy and segmentation performance compared to baseline approaches.

Index Terms—Object Detection, Segmentation and Categorization; Probabilistic Inference; Field Robots

I. INTRODUCTION

STAIRCASES, an ubiquitous feature of human-built environments throughout history, have enabled access to different levels within structures. With the increasing availability of legged robot platforms, significant strides have been made in enabling agile locomotion across diverse terrain, including staircases. However, ensuring safe navigation in complex real-world scenarios, such as cluttered or damaged staircases, remains a formidable challenge. This necessitates robots that can not only perceive their environment but also continuously update their knowledge about the environment by inferring the underlying structure.

Accurately perceiving and modeling staircases in such scenarios requires overcoming three key challenges: occlusions, limited field-of-view and sensor noise. Occlusions from objects or clutter can hide parts of the staircase, leading to incomplete estimate of staircase location. Moreover, a robot’s limited field-of-view often restricts perception to a few steps at a time, hindering estimation of the staircase’s overall location

Manuscript received: December, 13, 2024; Accepted February, 22, 2025. This paper was recommended for publication by Editor Giuseppe Loianno upon evaluation of the Associate Editor and Reviewers’ comments.

¹All authors are from The Robotics Institute, Carnegie Mellon University, USA. {pkettava, bshirose, mtravers}@andrew.cmu.edu

Supplementary Video - https://youtu.be/8baHgQ_rGLs

Digital Object Identifier (DOI): see top of this page.

©2026 IEEE

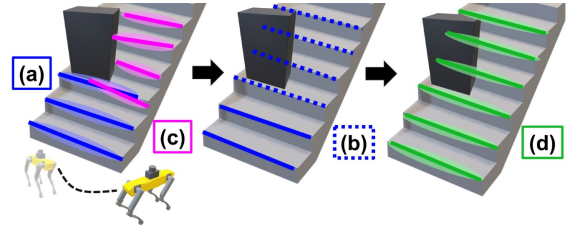


Fig. 1: (a) Prior staircase estimate, $bel({}_{\mathcal{L}}\mathbf{X}_k)$ at time k (solid blue). (b) Predicted staircase location, $p({}_{\mathcal{L}}\hat{\mathbf{X}}_{k+1|k})$ based on prior (dotted blue). (c) New noisy measurement, $p({}_{\mathcal{L}}\mathbf{Z}_{k+1})$ at time $k+1$ (magenta). (d) Filtered staircase estimate $bel({}_{\mathcal{L}}\mathbf{X}_{k+1})$ at time $k+1$ by combining prediction and measurement (green).

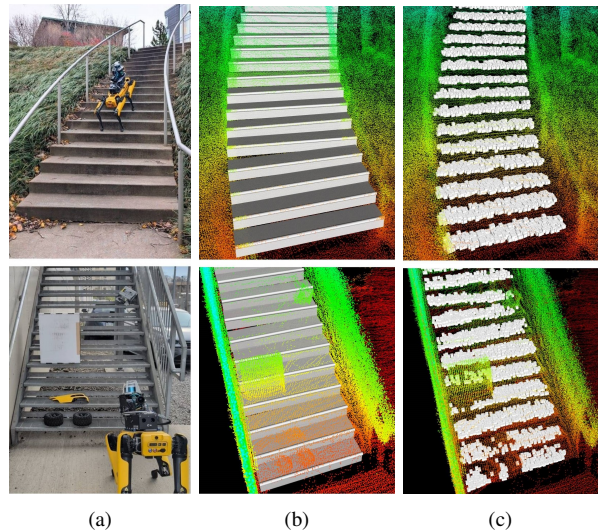


Fig. 2: Results of our proposed staircase modeling and estimation framework. (a) Robot navigating staircase environments (b) Estimated staircase highlighted by white marker. (c) Segmented clutter-free regions on each stair.

and orientation. Robot movement introduces sensor noise, causing errors in measurements of step height, depth, and curvature. These challenges necessitate a robust approach that integrates noisy sensor measurements with prior knowledge for accurate staircase estimation.

In this paper, we propose a novel method for robust staircase estimation employing Bayesian inference, as illustrated in Fig. 1. Our approach leverages prior knowledge about typical staircase geometry obtained from previously observed parts of the staircase to infer the location and geometry of new steps. This prediction is then combined with noisy sensor measurements to generate a maximum a posteriori (MAP) estimate of the staircase over time, effectively filtering out noise and accounting for occlusions. This MAP estimate is fused with 3D point cloud data to enable accurate segmentation of the

stair surfaces, differentiating them from surrounding clutter. The contributions of this letter are:

- **Staircase Modeling:** A novel split state-space model to parameterize and represent large-scale staircases
- **MAP Estimation:** A robust pipeline for estimating staircase location and parameters over time using noisy detections in scenes with occlusion and clutter
- **Stair Surface Segmentation:** An algorithm combining the staircase estimate with the point cloud to segment the staircase ground points in presence of clutter

We implement our proposed method both on real robots and in simulation in diverse environments with large and cluttered staircases (Fig. 2). We achieve real-time performance with 67-89% reduction in estimated staircase parameter errors and 30% reduction in staircase location error compared to baselines

II. RELATED WORK

A. Staircase Detection and Estimation

Researchers have explored various methods for staircase detection using both image-based and point cloud-based approaches. While image-based methods [1] [2] offer computational speed, they are susceptible to environmental factors like lighting conditions, and face challenges in accurately estimating 3D geometry. Point cloud methods leverage the inherent planar geometry of stairs to segment planes and estimate staircase location and parameters. Although RANSAC-based approaches are prevalent [3] [4], they can be non-deterministic and computationally expensive. Perez-Yus et al. used a wearable RGB-D camera to detect staircases [5]. Westfechtel et al. [6] achieved high accuracy in staircase parameter estimation using a graph-based search strategy, but required significant processing time. Qing et al. [7] introduced a measurement correction technique based on estimated stair plane poses, but their approach is sensitive to plane segmentation errors that can be caused by clutter on stairs.

The authors' prior work in [8] presented a fast staircase detection method and introduced a simple matching algorithm to combine multiple detections over time, which averaged individual matched stairs based on proximity. However, this approach proved susceptible to noise, leading to estimation failures due to erroneous matches. Furthermore, the absence of explicit staircase modeling and the inability to leverage existing staircase information limited its capacity to handle incomplete detections arising from the sensor's restricted field of view or occlusions.

B. Ground Segmentation

Extensive research has investigated ground segmentation techniques due to their critical role in ensuring safe robot navigation [9]. Elevation maps are a simple and popular approach, which transform 3D point clouds into 2.5D grid representations for ground point classification [10]. Plane fitting techniques, frequently employing RANSAC, classify points based on their proximity to a fitted plane [11]. These methods encounter difficulties in scenarios with multiple ground planes, such as staircases, as they necessitate multiple plane estimations, which is computationally expensive. While variations

like concentric zone divisions [12] strive to accelerate this process, the inherent complexity of cluttered staircases persists as a challenge.

More sophisticated methods, including Markov Random Fields (MRFs) [13] and Gaussian processes [14], offer enhanced accuracy but demand substantial computational resources, posing challenges for mobile robots with limited compute. Cloth simulation filters [15] present an alternative, yet their performance on stairs is highly dependent on parameter tuning. Furthermore, purely geometric methods struggle to differentiate between small, flat obstacles on stairs and adjacent steps, raising safety concerns. Learning-based methods [16][17] integrate semantic information to improve performance, but their effectiveness remains contingent on the specific scenarios encountered during training.

III. STAIRCASE MODELING

A. Staircase State-Space

We present a split state-space model for representing staircases, enabling a robot to capture and reason about those extending beyond its sensor range. The first component, the 'infinite-line staircase state' ($\mathbf{L}\mathbf{X}$), represents the intersection of each stair's tread and riser as an *infinite line* using polar coordinates. Subsequently, for each stair i in the staircase, its state $\xi^{(i)}$ consists of radius $r^{(i)}$, its angle $\phi^{(i)}$, along with the start and end z-positions ($z_s^{(i)}, z_e^{(i)}$) of each stair, all with respect to the robot's world frame. The infinite-line state is illustrated in Fig. 3a. For a staircase with N steps at time step k , the infinite-line state vector, $\mathbf{L}\mathbf{X}_k$, consists of each stair line in ascending staircase direction.

$$\mathbf{L}\mathbf{X}_k = [\xi_k^{(1)} \dots \xi_k^{(N)}], \xi_k^{(i)} = [r^{(i)} \phi^{(i)} z_s^{(i)} z_e^{(i)}] \quad (1)$$

The second component, the 'staircase endpoint state' ($\mathbf{P}\mathbf{X}$), grounds this representation by incorporating the physical location of each step. $\mathbf{P}\mathbf{X}$ uses cartesian coordinates to define the actual starting (\mathbf{p}_s) and ending points (\mathbf{p}_e) of each stair line with respect to the robot's world frame (Fig. 3). This precise 3D information allows for accurate reconstruction of parts of the staircase that have been observed, while the infinite-line state can help describe potentially unseen parts of the staircase. For a staircase with N steps at time step k , the endpoint state vector, $\mathbf{P}\mathbf{X}_k$, consists of each stair's endpoints in ascending staircase direction.

$$\mathbf{P}\mathbf{X}_k = [\mathbf{p}_s^{(1)} \mathbf{p}_e^{(1)} \dots \mathbf{p}_s^{(N)} \mathbf{p}_e^{(N)}], \quad (2)$$

$$\mathbf{p}_s^{(i)} = [x_s^{(i)} y_s^{(i)} z_s^{(i)}], \mathbf{p}_e^{(i)} = [x_e^{(i)} y_e^{(i)} z_e^{(i)}]$$

As our model represents stairs as edges (lines) it excludes the information of the initial ground plane for ascending staircases. Similarly, for descending staircases, it excludes the final ground plane at the bottom.

B. Parameterized Staircase Model

We model a staircase using six key parameters \mathbf{S}_k . We assume that the parameters remain consistent between any two consecutive steps along the staircase. As depicted in Fig. 3c, the six parameters of our staircase model are:

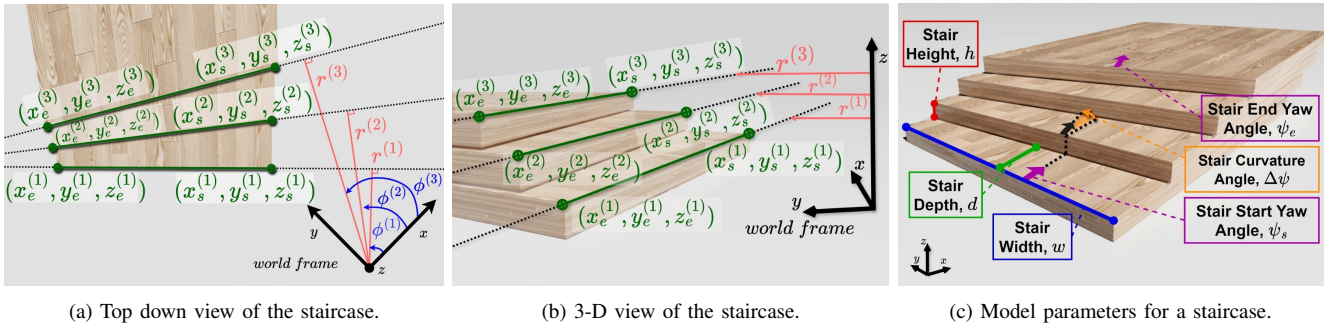


Fig. 3: Staircase annotated with the ‘infinite-line’ and ‘staircase endpoint’ state variables along with its model parameters.

- *Stair Height* (h): The vertical distance between consecutive steps along the entire staircase.
- *Stair Depth* (d): The horizontal depth of each step. In the case of curved staircases, the depth is measured from the center of each stair line, as illustrated in Fig. 3c. This ensures consistent measurement even when the stair edges are not parallel.
- *Stair Width* (w): The overall width of the staircase.
- *Stair Start Yaw Angle* (ψ_s): The orientation of the first stair line with respect to the world frame, defined as the angle perpendicular to the stair line in ascending direction of the staircase.
- *Stair End Yaw Angle* (ψ_e): The orientation of the last stair line with respect to the world frame. This is defined similar to the stair start yaw angle.
- *Stair Curvature Angle* ($\Delta\psi$): The change in angle between consecutive steps, which captures the curvature of the staircase.

Each of these parameters can be computed directly from the current staircase state ($\mathbf{L}\mathbf{X}_k, \mathbf{P}\mathbf{X}_k$). To minimize the impact of noise, the stair height, depth, width, and curvature angle of the staircase are determined by averaging the corresponding values across all individual steps. Staircases divided by a landing are treated as multiple, separate staircases. The pronounced difference in stair depth at a landing is used to detect these instances. It’s important to distinguish the stair start and end angles (ψ_s, ψ_e) from the infinite-line state angles ($\phi^{(1)}, \phi^{(N)}$). While ψ_s and ψ_e represent the orientation of the perpendicular to the stair line in the ascending direction of the staircase, $\phi^{(i)}$ simply represents the angle of the normal to the line. Hence, based on the staircase’s location in the map ψ_s can be the same as $\phi^{(1)}$ or $\phi^{(1)} + \pi$. The same applies to ψ_e as well.

This parameterized model is crucial for robustly estimating individual stair locations, as it allows us to predict the location of a stair using information from its immediate neighbors. $\xi^{(i,j)}$ denotes the state of stair i predicted given the state of its *adjacent stair* j using the stair parameters, as shown in (3). These equations incorporate several parameters to generalize for diverse staircase configurations. The parameter η controls the direction of prediction, where $\eta = 1$ for predicting a succeeding stair ($j = i - 1$) and $\eta = -1$ for a preceding stair ($j = i + 1$). As detailed before, the orientation of stair j (given by $\psi_s \pm j\Delta\psi$) can differ from the actual angle of the stair line ($\phi^{(j)}$) by π , depending on the staircase’s position and the world frame’s orientation. To account for this, the parameter ρ is introduced. Moreover, since a curved staircase has different

depths at different points on the line, the parameters γ and l account for this difference and correct the stair depth (d) based on distance from the center of the given stair ($c_x^{(j)}, c_y^{(j)}$).

$$\begin{aligned}
 r^{(i,j)} &= r^{(j)} + \eta\rho(d - l \sin(\gamma(\eta\Delta\psi))) & z_s^{(i,j)} &= z_s^{(j)} + \eta h \\
 \phi^{(i,j)} &= \phi^{(j)} + \eta(\Delta\psi) & z_e^{(i,j)} &= z_e^{(j)} + \eta h \\
 \eta &= \begin{cases} 1 & \text{if } j = i - 1 \text{ (predicting next stair),} \\ -1 & \text{if } j = i + 1 \text{ (predicting previous stair)} \end{cases} \\
 \rho &= \begin{cases} 1 & \text{if } (\psi_s + j\eta\Delta\psi - \phi^{(j)}) \approx 0, \\ -1 & \text{if } (\psi_s + j\eta\Delta\psi - \phi^{(j)}) \approx \pi \end{cases} \\
 \gamma &= \begin{cases} 1 & \text{if } \phi^{(i,j)} \in [\phi^{(i)}, \tan^{-1}(c_y^{(j)}, c_x^{(j)})], \\ -1 & \text{otherwise} \end{cases} \\
 l &= \sqrt{(c_x^{(j)} - r^{(j)} \cos(\phi^{(i,j)}))^2 + (c_y^{(j)} - r^{(j)} \sin(\phi^{(i,j)}))^2}, \\
 c_x^{(j)} &= (x_s^{(j)} + x_e^{(j)})/2, & c_y^{(j)} &= (y_s^{(j)} + y_e^{(j)})/2 \quad (3)
 \end{aligned}$$

Similarly, we can predict the endpoint state for stair i given neighboring stair j , denoted as $\mathbf{p}_s^{(i,j)}$ and $\mathbf{p}_e^{(i,j)}$ as shown in (4). The prediction starts from the known endpoint of stair j and extrapolates them based on the direction. The predictions conform to the direction of extrapolation, as it utilizes the starting stair yaw angle (ψ_s) when predicting a succeeding stair and the final stair yaw angle (ψ_e) when predicting a preceding stair. Furthermore, the prediction incorporates variations in step depth (d) for curved staircases by accounting for the curvature ($\Delta\psi$).

$$\begin{aligned}
 x_s^{(i,j)} &= x_s^{(j)} + d_s \cos(\psi_o + \eta(j+1)\Delta\psi) \\
 y_s^{(i,j)} &= y_s^{(j)} + d_s \sin(\psi_o + \eta(j+1)\Delta\psi) \\
 x_e^{(i,j)} &= x_e^{(j)} + d_e \cos(\psi_o + \eta(j+1)\Delta\psi) \\
 y_e^{(i,j)} &= y_e^{(j)} + d_e \sin(\psi_o + \eta(j+1)\Delta\psi) \\
 d_s &= d + \frac{w}{2} \sin(\Delta\psi), & d_e &= d - \frac{w}{2} \sin(\Delta\psi), \\
 \psi_o &= \psi_e, \eta = -1 & \text{if } j = i + 1 \text{ (predict previous stair),} \\
 \psi_o &= \psi_s, \eta = 1 & \text{if } j = i - 1 \text{ (predict next stair)} \quad (4)
 \end{aligned}$$

We can estimate the state of non-adjacent stairs (i.e., stairs where j is not an immediate neighbor of i) by iteratively applying (4), effectively enabling the prediction of any stair along the staircase. This is particularly useful to predict the state of multiple new stairs based on the information from a single observed stair.

C. Staircase Measurements and Initialization

We leverage raw detections from our staircase detection algorithm [8] as measurements ($\mathbf{L}\mathbf{Z}_k, \mathbf{P}\mathbf{Z}_k$). Measurements only capture the portion of the staircase within the robot's current field-of-view and are expressed in the robot's local frame. They are split into two components: measured stair lines ($\mathbf{L}\mathbf{Z}_k$) and measured stair endpoints ($\mathbf{P}\mathbf{Z}_k$). In contrast, the staircase state ($\mathbf{L}\mathbf{X}, \mathbf{P}\mathbf{X}$) provides an estimate of the entire staircase structure in the world coordinate frame. This local frame measurement approach aids in decoupling uncertainty in detection and the robot localization uncertainty. The robot's local frame is defined with its x-y plane parallel to the world x-y plane, simplifying the corresponding robot's pose (Ω_k) to its position (t_r) and yaw orientation (θ_r).

$$\mathbf{L}\mathbf{Z}_k = [\zeta_k^{(1)} \dots \zeta_k^{(M)}], \zeta_k^{(i)} = [r_m^{(i)} \phi_m^{(i)} z_{sm}^{(i)} z_{em}^{(i)}], \quad (5)$$

$$\mathbf{P}\mathbf{Z}_k = [\mathbf{p}_{sm}^{(1)} \mathbf{p}_{em}^{(1)} \dots \mathbf{p}_{sm}^{(M)} \mathbf{p}_{em}^{(M)}], \quad (6)$$

$$\mathbf{p}_{sm}^{(i)} = [x_{sm}^{(i)} y_{sm}^{(i)} z_{sm}^{(i)}], \mathbf{p}_{em}^{(i)} = [x_{em}^{(i)} y_{em}^{(i)} z_{em}^{(i)}], \quad (7)$$

$$\Omega_k = [t_r \ \theta_r] = [x_r \ y_r \ z_r \ \theta_r]$$

To enable state estimation, we represent the infinite line state by its belief, $bel(\mathbf{L}\mathbf{X}_k)$ at time k with mean $\mathbf{L}\mathbf{X}_k$ and the covariance $\mathbf{L}\Sigma_k$. Upon initial detection of a staircase, we initialize the infinite-line staircase state ($\mathbf{L}\mathbf{X}_k$) by transforming the measurement to world frame. This transformation, $g(\zeta_k^{(i)}, \Omega_k)$ is shown in (8). We also initialize a state covariance matrix ($\mathbf{L}\Sigma_k$) for the infinite-line state as shown in (9). This covariance incorporates both the inherent noise in the measurements (\mathcal{Q}) and the uncertainty in the estimated robot pose from SLAM (Σ_Ω). We model the measurement noise as zero-mean Gaussian with standard deviations $\sigma_r, \sigma_\phi, \sigma_{z_s}, \sigma_{z_e}$ for each component of the measured stair line ($\mathbf{L}\mathbf{Z}_k$). We also initialize the endpoint state ($\mathbf{P}\mathbf{X}_k$) as shown in (10).

$$\zeta_k^{(i)} = g(\zeta_k^{(i)}, \Omega_k) \quad (8)$$

$$\begin{bmatrix} r^{(i)} \\ \phi^{(i)} \\ z_s^{(i)} \\ z_e^{(i)} \end{bmatrix} = \begin{bmatrix} r_m^{(i)} + x_r \cos(\phi_m^{(i)} + \theta_r) + y_r \sin(\phi_m^{(i)} + \theta_r) \\ \phi_m^{(i)} + \theta_r \\ z_{sm}^{(i)} + z_r \\ z_{em}^{(i)} + z_r \end{bmatrix}$$

$$\mathbf{L}\Sigma_k = \mathbf{G}_z \mathcal{Q} \mathbf{G}_z^T + \mathbf{G}_\Omega \Sigma_\Omega \mathbf{G}_\Omega^T, \quad (9)$$

$$\mathbf{G}_z = \left. \frac{\partial g}{\partial \zeta} \right|_{\zeta_k^{(i)}, \Omega_k} \quad \mathbf{G}_\Omega = \left. \frac{\partial g}{\partial \Omega_k} \right|_{\zeta_k^{(i)}, \Omega_k}$$

$$\mathcal{Q} = \text{Diagonal}(\sigma_r^2, \sigma_\phi^2, \sigma_{z_s}^2, \sigma_{z_e}^2),$$

$$\mathbf{p}_s^{(i)} = \mathbf{R}(\theta_r) \mathbf{p}_{sm}^{(i)} + t_r, \quad \mathbf{p}_e^{(i)} = \mathbf{R}(\theta_r) \mathbf{p}_{em}^{(i)} + t_r \quad (10)$$

$\mathbf{R}(\theta_r)$ is the SO3 rotation matrix with pure yaw θ_r .

It is important to note that, unlike the infinite-line state, we do not maintain a covariance matrix for the stair endpoint state. This design choice is motivated by two factors. Firstly, the weighted line fitting algorithm [18], employed by our staircase detection module [8], inherently handles the noise in polar coordinates during the segmentation process. Secondly, we prioritize refining the infinite-line state via the Extended Kalman Filter (EKF) (explained in Section IV) and subsequently update the endpoints based on this refined state.

IV. STAIRCASE ESTIMATION

In this section we describe our staircase estimation pipeline that utilizes our modeling framework. Our pipeline operates within a Bayesian framework: we begin with a belief about the staircase state at time k , $bel(\mathbf{L}\mathbf{X}_k)$. This belief is then updated at time $k+1$ to $bel(\mathbf{L}\mathbf{X}_{k+1})$ by fusing it with the incoming staircase measurement likelihood, $p(\mathbf{L}\mathbf{Z}_{k+1})$, using the Bayesian update equation:

$$bel(\mathbf{L}\mathbf{X}_{k+1}) = \beta p(\mathbf{L}\mathbf{Z}_{k+1}) \int_{\mathbf{L}\mathbf{X}_k} p(\mathbf{L}\hat{\mathbf{X}}_{k+1|k}) bel(\mathbf{L}\mathbf{X}_k) d\mathbf{L}\mathbf{X}_k$$

Here, β is a normalizing factor and $p(\mathbf{L}\hat{\mathbf{X}}_{k+1|k})$ represents the state transition probability obtained by the process model. Assuming Gaussian noise throughout our system, this Bayesian update could ideally be computed using a standard Kalman Filter [19]. However, as our process model is non-linear, we employ the Extended Kalman Filter (EKF) with a linear approximation of the process model. The overall pipeline is shown in Fig. 4. Further, we utilize the estimated staircase parameters to segment stair surfaces on a cluttered staircase.

A. Stair-to-Stair Matching

After initialization, new measurements undergo data association to match them with the current staircase state. This employs the Mahalanobis distance, which quantifies the similarity between a measurement and the existing state while considering uncertainty. Each stair measurement is classified as either *matched*, *new preceding*, or *new succeeding*. Matched stairs correspond directly to existing stairs (within 3σ) in the state, while new preceding/succeeding stairs indicate those located below/above the current lowest/highest stair respectively.

B. Process Model - Predicting Staircase State

The process model, denoted by $f(\mathbf{L}\mathbf{X}_k, \mathbf{S}_k)$, predicts the next staircase state ($\mathbf{L}\hat{\mathbf{X}}_{k+1|k}, \mathbf{P}\hat{\mathbf{X}}_{k+1|k}$) using the current state ($\mathbf{L}\mathbf{X}_k, \mathbf{P}\mathbf{X}_k$) and staircase parameters (\mathbf{S}_k). This helps track the staircase when new stairs appear or robot motion adds uncertainty. To accommodate new stairs, the state vector

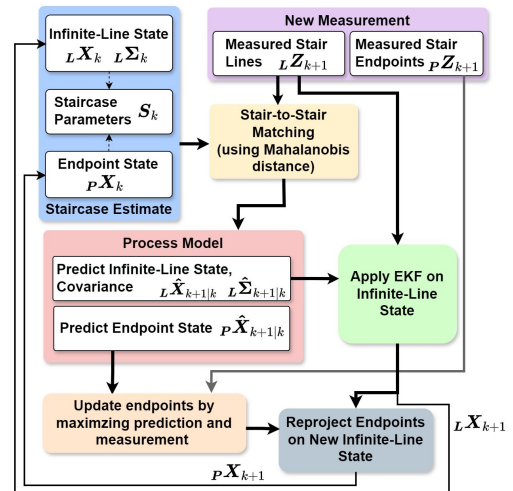


Fig. 4: Block diagram depicting the overall estimation pipeline.

dynamically resizes, placing ‘new preceding stairs’ at the beginning and ‘new succeeding stairs’ at the end of the initial N stairs. Thus, for u new preceding and v new succeeding stairs, the stairs 1 to u become the new preceding stairs, and stairs $u+N+1$ to $u+N+v$ become the new succeeding stairs. The process model slightly differs based on this classification.

Matched Stairs. For stairs that matched with existing stairs in the state estimate, the process model predicts the state by incorporating information from neighboring stairs. Specifically, the predicted state of each matched stair i is updated by averaging three components: its current state, the prediction from the preceding stair ($i-1$), and the prediction from the succeeding stair ($i+1$). For the stairs at the edges of the staircase (the lower-most and top-most existing stairs), only two predictions are averaged, as they only have one neighboring stair. Using (3) and (4), the predicted infinite-line state ($\hat{\xi}_{k+1|k}^{(i)}$) and the predicted endpoint state ($\hat{\mathbf{p}}_s^{(i)}, \hat{\mathbf{p}}_e^{(i)}$) for a matched stair i can be computed as,

$$\begin{aligned}\hat{\xi}_{k+1|k}^{(i)} &= \frac{1}{3} \left(\xi_k^{(i)} + \xi_k^{(i,i-1)} + \xi_k^{(i,i+1)} \right), \\ \hat{\mathbf{p}}_s^{(i)} &= \frac{1}{3} \left(\mathbf{p}_s^{(i)} + \mathbf{p}_s^{(i,i-1)} + \mathbf{p}_s^{(i,i+1)} \right), \\ \hat{\mathbf{p}}_e^{(i)} &= \frac{1}{3} \left(\mathbf{p}_e^{(i)} + \mathbf{p}_e^{(i,i-1)} + \mathbf{p}_e^{(i,i+1)} \right), \\ \text{where } i &\in [u+1 \dots u+N].\end{aligned}\quad (11)$$

New Preceding/New Succeeding Stairs. When new stairs are detected, the process model iteratively predicts their states using (3) and (4). For preceding stairs, we extrapolate from the lower-most existing stair (current index is $u+1$) to predict the closest preceding stair and progress downwards, and can be computed as:

$$\begin{aligned}\hat{\xi}_{k+1|k}^{(i)} &= \xi_k^{(i,u+1)}, \quad \hat{\mathbf{p}}_s^{(i)} = \mathbf{p}_s^{(i,u+1)}, \quad \hat{\mathbf{p}}_e^{(i)} = \mathbf{p}_e^{(i,u+1)}, \\ \text{where } i &\in [u \ u-1 \dots 1].\end{aligned}\quad (12)$$

For succeeding stairs, we extrapolate from the top-most existing stair (current index is $u+N$) to predict the closest succeeding stair and progress upwards, using:

$$\begin{aligned}\hat{\xi}_{k+1|k}^{(i)} &= \xi_k^{(i,u+N)}, \quad \hat{\mathbf{p}}_s^{(i)} = \mathbf{p}_s^{(i,u+N)}, \quad \hat{\mathbf{p}}_e^{(i)} = \mathbf{p}_e^{(i,u+N)}, \\ \text{where } i &\in [u+N+1 \dots u+N+v].\end{aligned}\quad (13)$$

Covariance Prediction. The predicted covariance matrix ($\mathbf{\Sigma}_{k+1|k}$) not only reflects the uncertainty propagated from the current state but also incorporates the uncertainty associated with the staircase parameters themselves. This is achieved through the \mathcal{R} matrix, which models the uncertainty in the staircase parameters ($\sigma_h, \sigma_d, \sigma_w, \sigma_{\psi_s}, \sigma_{\psi_n}, \sigma_{\Delta\psi}$). Crucially \mathcal{R} can be user-selected, providing a mechanism to adjust the trust in the process model. Higher values in \mathcal{R} indicate less trust, causing the filter to rely more on incoming measurements.

$$\begin{aligned}\mathbf{\Sigma}_{k+1|k} &= \mathbf{F}_X \mathbf{\Sigma}_k \mathbf{F}_X^T + \mathbf{F}_S \mathcal{R} \mathbf{F}_S^T, \\ \mathbf{F}_X &= \left. \frac{\partial f}{\partial \mathbf{L}\mathbf{X}} \right|_{\mathbf{L}\mathbf{X}_k, \mathbf{S}_k} \quad \text{Jacobian of prediction} \\ &\quad \text{w.r.t infinite-line state,} \\ \mathbf{F}_S &= \left. \frac{\partial f}{\partial \mathbf{S}} \right|_{\mathbf{L}\mathbf{X}_k, \mathbf{S}_k} \quad \text{Jacobian of prediction} \\ &\quad \text{w.r.t staircase parameters.}\end{aligned}\quad (14)$$

C. Updating Staircase Estimate

The staircase state is updated in a two-step process using new measurements. First, the predicted infinite-line state ($\mathbf{L}\hat{\mathbf{X}}_{k+1|k}$) is refined through an Extended Kalman Filter (EKF). This involves transforming the predicted state to the robot’s local frame and incorporating the new measurement data ($\mathbf{L}\mathbf{Z}_{k+1}$) and its uncertainty (\mathcal{Q}). This step produces a maximum a posteriori estimate of the infinite-line staircase state $\mathbf{L}\mathbf{X}_{k+1}$. If a stair has no match in the new measurement, EKF is not applied and its state remains unchanged.

Next, the endpoint state is updated through a maximization process that refines the endpoints of each stair to maximize its stair width. For each stair i , this process considers both the predicted endpoints from the process model ($\hat{\mathbf{p}}_s^{(i)}, \hat{\mathbf{p}}_e^{(i)}$) and the measured endpoints ($\mathbf{p}_{sm}^{(i)}, \mathbf{p}_{em}^{(i)}$), selecting the two endpoints that maximize the XY Euclidean distance. This ensures that the endpoint state is updated upon observing previously unseen stair segments. Finally, these refined endpoints are re-projected onto the updated infinite-line state ($\mathbf{L}\mathbf{X}_{k+1}$) to ensure consistency between the two representations. This re-projection effectively refines the endpoint estimates by incorporating the updated stair line location obtained from EKF. The resulting staircase estimate integrates the refined infinite-line and endpoint states, providing an accurate and robust representation of the staircase structure.

D. Stair Surface Segmentation

Once the staircase state and the staircase parameters are updated, we can utilize this information to segment stair surfaces from cluttered scenes on each stair. For each individual stair, a 3D crop-box is defined. The length and width of the crop-box is computed using the estimated endpoints and staircase depth, whereas the height of the crop-box is derived by the uncertainty in the estimated z-positions ($z_s^{(i)}, z_e^{(i)}$) of the stair. This effectively isolates the point cloud region corresponding to each stair, as visualized in Fig. 5b.

For points within each crop-box, the algorithm further refines the segmentation by fitting a plane through the points using RANSAC. Crucially, this plane fitting incorporates a

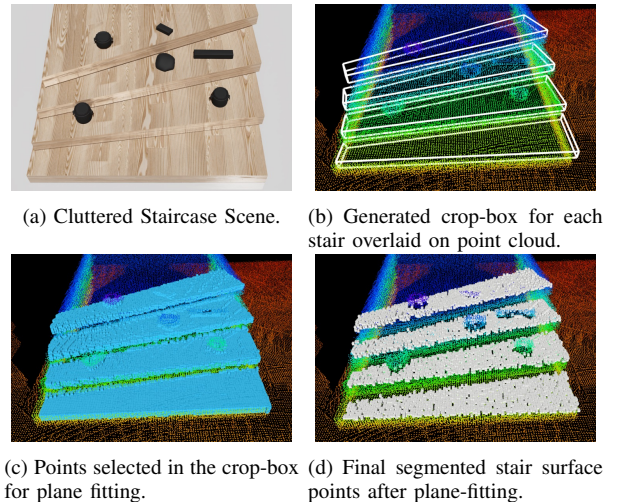


Fig. 5: Stair surface segmentation process for a cluttered staircase.

constraint ensuring that the identified plane remains parallel to the ground plane. This constraint leverages the inherent geometric property of stairs, where the tread surface is typically parallel to the ground. The inliers to this fitted plane are then classified as the stair tread surface, effectively separating them from any remaining clutter or outliers within the crop-box. This approach, illustrated in Fig. 5, enables robust segmentation of stair surfaces even in cluttered environments. As our staircase state excludes the initial ground plane, it is not segmented in our algorithm. Incorporating this ground plane is left as future work.

V. EXPERIMENTS AND RESULTS

A. Experimental Setup

1) *System Overview*: Experiments were conducted both in simulation (using Isaac Sim) and real-world environments (using a Boston Dynamics Spot robot). The Spot robot was equipped with an NVIDIA Jetson AGX Orin, a Velodyne VLP-16 LiDAR for localization and mapping (LiPO [20]). To enhance pointcloud density, points from the Spot’s onboard RealSense sensors were registered to the LiDAR-based SLAM map. In Isaac Sim, a Velodyne VLP-128 LiDAR mounted on a simulated Spot robot generated high-fidelity point cloud data. Data collection involved manually teleoperating Spot to navigate around the staircase, ensuring complete capture by the onboard sensors. Manual teleoperation was crucial because the challenging nature of cluttered staircases hindered Spot’s autonomous navigation.

2) *Dataset*: A comprehensive dataset of 23 distinct staircase examples (17 real-world, 6 simulation) was curated to evaluate the proposed system, encompassing a diverse range of configurations (size, length, curvature, clutter). Specifically, 10 examples featured staircases without any clutter, while the remaining 13 included various types of clutter, such as large boxes obstructing staircase visibility and smaller items on the steps, representing different levels of clutter. It included staircases with widths ranging between 1 and 10 meters and step counts ranging from 4 to 20. These included both ascending and descending orientations as well as curved and open-rise staircases.

3) *Baselines*: To evaluate our approach, we compare it against two staircase estimation baselines adapted from prior state-of-the-art [8] and one segmentation baseline. The simple averaging (AVG) merging algorithm estimates the staircase state by averaging the current estimate of each stair’s endpoints with corresponding new measurements. The maximizing algorithm (MAX) is similar to the averaging algorithm, except it selects the endpoints that maximize the Euclidean distance between the start and end points, effectively maximizing the estimated stair width. For segmentation, we use a Cloth Simulation Filter [15] as a baseline. This method simulates a cloth draped over the estimated staircase region, conforming to the underlying geometry to approximate stair surfaces.

4) *Error Metrics*: Ground truth data for each staircase was meticulously generated through manual segmentation, involving precise annotation of staircase locations and parameters (height, depth, width, curvature), along with the removal of

debris points from the point clouds. We assess staircase estimation accuracy by evaluating errors in two aspects: *Staircase Parameter RMSE*, which quantifies the root-mean-square error between estimated and ground truth staircase parameters; and *Stair Location RMSE*, which evaluates individual stair location accuracy using a point-to-line distance metric, calculating the distance between estimated endpoints of each stair and the corresponding ground truth stair line, along with the stair orientation error.

B. Results

Figure 6 highlights the robustness of our proposed staircase estimation method in challenging real-world scenarios, featuring a curved staircase, and a long cluttered staircase with objects obstructing the steps. In the first example, the inferior performance of the baseline approaches can be noticed in the top portion of the stairs, where the estimated steps vary in height and fail to accurately capture the consistent rise of the actual staircase. Our approach, on the other hand, accurately captures the geometry of the curved staircase, including the consistent step heights throughout its entirety. In contrast to the curved staircase, the second scenario presents a long staircase with significant clutter that leads to noisy measurements. This noise causes significant issues for the baseline approaches: the averaging method completely misses an entire step in its estimation, while the maximizing approach produces a severely distorted and inaccurate estimate. Conversely, our method, which employs Mahalanobis distance for model matching, effectively accounts for this uncertainty, leading to a complete and accurate estimation. Our approach is able to estimate the full extent of the staircase, even in the presence of a large occlusion (red cone) that obscures a portion of the stairs. This accurate staircase estimation enables the precise stair segmentation shown in Fig. 6f, where our proposed algorithm classifies points belonging to the stair surfaces.

As shown in TABLE I, our proposed method demonstrates significantly improved accuracy in estimating both the param-

	R/S	AVG [8]	MAX	EKF (Ours)
Stair Height RMSE (cm)	R	1.8 ± 0.8	2.3 ± 0.9	0.6 ± 0.3
	S	1.1 ± 0.5	1.0 ± 0.7	0.3 ± 0.1
Stair Depth RMSE (cm)	R	2.0 ± 1.3	3.8 ± 2.1	1.3 ± 0.8
	S	3.6 ± 4.0	3.5 ± 0.9	2.0 ± 1.5
Stair Width RMSE (cm)	R	113.7 ± 112.5	29.9 ± 18.8	12.0 ± 18.7
	S	424.0 ± 341.8	30.7 ± 23.7	11.8 ± 8.8
Stair Curvature RMSE (cm)	R	2.0 ± 1.1	3.3 ± 2.8	1.5 ± 0.8
	S	11.5 ± 26.1	1.4 ± 1.7	0.8 ± 0.7
Stair Location XY RMSE (cm)	R	5.0 ± 3.4	6.8 ± 5.5	3.6 ± 1.5
	S	4.0 ± 2.5	3.8 ± 1.8	2.9 ± 0.9
Stair Location Z RMSE (cm)	R	3.3 ± 1.7	3.4 ± 2.1	2.3 ± 1.2
	S	1.4 ± 0.7	1.2 ± 0.6	1.0 ± 0.9
Stair Orientation RMSE (deg)	R	1.7 ± 0.8	2.6 ± 1.9	1.5 ± 0.7
	S	0.9 ± 1.0	1.0 ± 1.1	0.7 ± 0.6

TABLE I: Comparison of staircase estimation RMSE. **R** indicates real-world scenarios, and **S** indicates simulation.

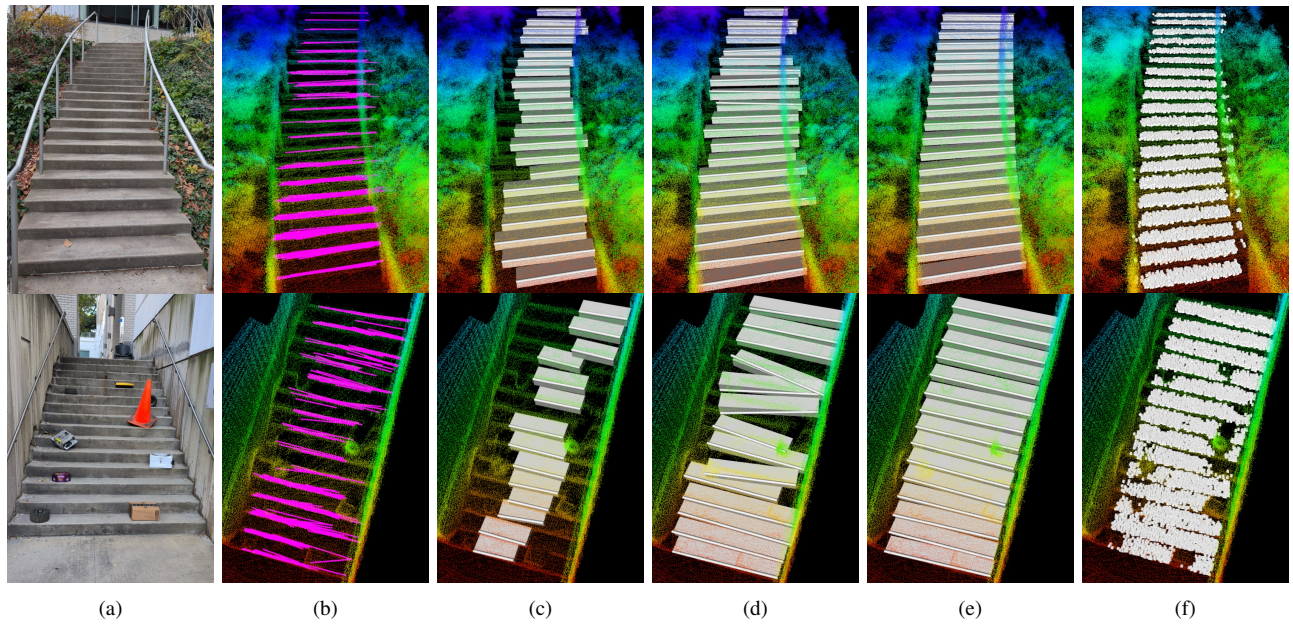


Fig. 6: Staircase estimation results for 2 different examples (a) Original staircase scenarios with curved staircase (top) and long cluttered staircase (bottom) (b) All input measurements (magenta) overlaid on the point cloud (c) Estimated staircase by using averaging algorithm [8] on the input measurements. (d) Estimated staircase by using maximizing algorithm on the input measurements (e) Estimated staircase by using our proposed framework on the input measurements (f) Points classified as stair surface using our segmentation algorithm

eters and location of staircases across all metrics in simulation and in real-world. For real-world staircases, we achieve a 89% reduction in width error and 67% reduction in height error compared to the averaging approach [8]. Furthermore, our method achieves approximately a 30% reduction in RMSE for stair location (XY and Z). While the maximizing algorithm addresses some limitations of this prior work, our approach consistently yields the most accurate results. It's worth noting that in simulation, the width error for the averaging method is considerably higher due to the presence of exceptionally wide (10m) staircases. Our superior performance is attributed to our method's ability to effectively incorporate information from the entire staircase structure, which helps to handle noise and incomplete measurements.

TABLE II presents the accuracy, precision, and recall of our proposed method for stair surface segmentation, compared against the Cloth Simulation Filter (CSF) [15]. Our method consistently outperforms CSF across all metrics. This improvement stems from our method's ability to leverage information about the staircase structure, enabling more accurate differentiation between stair surface and flat debris. In contrast, CSF is purely geometric and often misclassifies such debris as ground. While CSF can be tuned to improve performance, we observed inconsistent results across different cluttered staircases with the same parameters. Our approach consistently provided accurate segmentation across all scenarios.

To verify real-time performance, we evaluated our C++ implementation on an NVIDIA Jetson AGX Orin. We analyzed

Metric	Cloth Simulation Filter (CSF) [15]	Proposed Stair Segmentation Method
Accuracy (%)	86.36	93.13
Precision (%)	91.19	97.35
Recall (%)	94.43	95.56

TABLE II: Comparison of Stair Surface Segmentation Metrics

the computational performance of both our EKF pipeline, model prediction and matching algorithm, as well as the stair surface segmentation. Fig. 7 illustrates the performance of the EKF (shown in red) and the model prediction/matching (shown in blue) components, plotting the total computation time as a function of the number of stairs in the staircase (ranging from 3 to 20). Both the prediction and matching times increase with the number of stairs, with the EKF component exhibiting a steeper rise. Fig. 8 illustrates the time required for stair surface segmentation, with times normalized for a point cloud of 250k points to account for varying point cloud

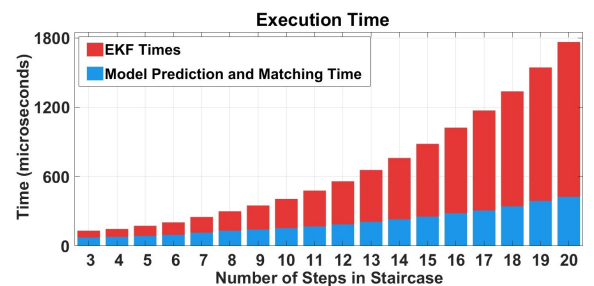


Fig. 7: Plot of execution times vs number of steps for the proposed estimation pipeline

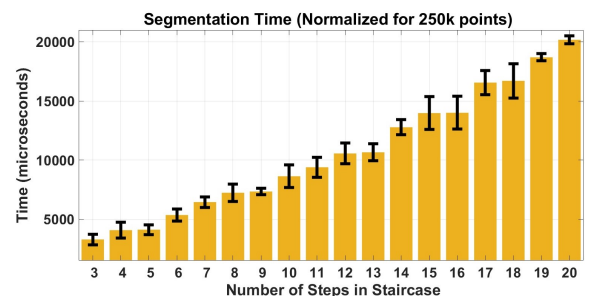


Fig. 8: Plot of execution times vs number of steps for the stair surface segmentation algorithm. Error bars indicate standard deviation.

densities across different staircases. Similar to the EKF, the computation time for segmentation shows an upward trend as the number of steps increases, reflecting the increased time required for a larger number of points. Despite these trends, the absolute segmentation times remain manageable, even for staircases with 20 steps. Crucially, the total time taken by both the estimation and segmentation methods, even for large staircases, remains under 30ms. This is well within 50ms budget to achieve our real-time execution target rate of 20Hz, confirming the feasibility of real-time operation.

VI. CONCLUSION

In this paper we presented a novel approach to staircase estimation in cluttered real-world environments. We introduced a split staircase state-space model, and a Bayesian inference framework utilizing this model for robust state estimation. Our method effectively addresses challenges posed by noise, clutter, and occlusions, leading to significant improvements in accuracy compared to existing approaches. We demonstrate the robustness and effectiveness of our approach through extensive experiments on staircases captured both on real robots and in simulation, showcasing its ability to accurately estimate staircase parameters and location even in complex scenarios. Moreover, our method achieves real-time performance, making it suitable for deployment on robotic systems operating in dynamic environments.

While our method has proven effective for a wide range of staircases, it does have limitations due to the assumption of consistent staircase parameters. This can lead to inaccuracies when encountering irregular staircases with varying step depths or heights. To address this, we plan to explore adaptive filtering approaches to improve estimation accuracy by dynamically changing the \mathcal{R} matrix based on the variations in the computed staircase parameters. We also plan to investigate the feasibility of running an ensemble of EKFs in parallel with varying \mathcal{R} matrices, and selecting the estimate that minimizes the Mahalanobis distance between the residual and innovation.

In the future, we aim to utilize the segmented stair surface regions for quadruped motion planning, enabling safe navigation on cluttered staircases. Additionally, we will investigate approaches to enable fusion across multiple robots using methods such as Covariance Intersection for deployment on a multi-robot system [21] in multi-staircase environments.

REFERENCES

- [1] S. Murakami, M. Shimakawa, K. Kivota, and T. Kato, "Study on stairs detection using rgb-depth images," in *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)*. IEEE, 2014, pp. 1186–1191.
- [2] M. Ilyas, A. K. Lakshmanan, A. V. Le, and M. R. Elara, "Staircase recognition and localization using convolutional neural network (cnn) for cleaning robot application," *Mathematics*, vol. 11, no. 18, p. 3964, 2023.
- [3] J. A. Sánchez-Rojas, J. A. Arias-Aguilar, H. Takemura, and A. E. Petrilli-Barceló, "Staircase detection, characterization and approach pipeline for search and rescue robots," *Applied Sciences*, vol. 11, no. 22, p. 10736, 2021.
- [4] J. Fourre, V. Vauchey, Y. Dupuis, and X. Savatier, "Autonomous rgbd-based industrial staircase localization from tracked robots," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10 691–10 696.
- [5] A. Perez-Yus, D. Gutiérrez-Gómez, G. Lopez-Nicolas, and J. Guerrero, "Stairs detection with odometry-aided traversal from a wearable rgbd camera," *Computer Vision and Image Understanding*, vol. 154, pp. 192–205, 2017.
- [6] T. Westfechtel, K. Ohno, B. Mertsching, R. Hamada, D. Nickchen, S. Kojima, and S. Tadokoro, "Robust stairway-detection and localization method for mobile robots using a graph-based model and competing initializations," *The International Journal of Robotics Research*, vol. 37, no. 12, pp. 1463–1483, 2018.
- [7] C. Qing, R. Zeng, X. Wu, Y. Shi, and G. Ma, "An onboard framework for staircases modeling based on point clouds," *arXiv preprint arXiv:2405.01918*, 2024.
- [8] P. Sriganesh, N. Bagree, B. Vundurthy, and M. Travers, "Fast staircase detection and estimation using 3d point clouds with multi-detection merging for heterogeneous robots," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 9253–9259.
- [9] T. Gomes, D. Matias, A. Campos, L. Cunha, and R. Roriz, "A survey on ground segmentation methods for automotive lidar sensors," *Sensors*, vol. 23, no. 2, p. 601, 2023.
- [10] A. Asvadi, P. Peixoto, and U. Nunes, "Detection and tracking of moving objects using 2.5 d motion grids," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 788–793.
- [11] P. Narksri, E. Takeuchi, Y. Ninomiya, Y. Morales, N. Akai, and N. Kawaguchi, "A slope-robust cascaded ground segmentation in 3d point cloud for autonomous vehicles," in *2018 21st International Conference on intelligent transportation systems (ITSC)*. IEEE, 2018, pp. 497–504.
- [12] H. Lim, M. Oh, and H. Myung, "Patchwork: Concentric zone-based region-wise ground segmentation with ground likelihood estimation using a 3d lidar sensor," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6458–6465, 2021.
- [13] W. Huang, H. Liang, L. Lin, Z. Wang, S. Wang, B. Yu, and R. Niu, "A fast point cloud ground segmentation approach based on coarse-to-fine markov random field," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7841–7854, 2021.
- [14] T. Chen, B. Dai, R. Wang, and D. Liu, "Gaussian-process-based real-time ground segmentation for autonomous land vehicles," *Journal of Intelligent & Robotic Systems*, vol. 76, pp. 563–582, 2014.
- [15] W. Zhang, J. Qi, P. Wan, H. Wang, D. Xie, X. Wang, and G. Yan, "An easy-to-use airborne lidar data filtering method based on cloth simulation," *Remote sensing*, vol. 8, no. 6, p. 501, 2016.
- [16] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss, "Rangenet++: Fast and accurate lidar semantic segmentation," in *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2019, pp. 4213–4220.
- [17] D. He, F. Abid, Y.-M. Kim, and J.-H. Kim, "Sectorgsnet: Sector learning for efficient ground segmentation of outdoor lidar point clouds," *IEEE Access*, vol. 10, pp. 11 938–11 946, 2022.
- [18] S. T. Pfister, S. I. Roumeliotis, and J. W. Burdick, "Weighted line fitting algorithms for mobile robot map building and efficient data representation," in *2003 IEEE International Conference on Robotics and Automation*, vol. 1. IEEE, 2003, pp. 1304–1311.
- [19] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.
- [20] D. Mick, T. Pool, M. S. Nagaraju, M. Kaess, H. Choset, and M. Travers, "Lipo: Lidar inertial odometry for icp comparison," *arXiv preprint arXiv:2410.08097*, 2024.
- [21] P. Sriganesh, J. Maier, A. Johnson, B. Shirose, R. Chandrasekar, C. Noren, J. Spisak, R. Darnley, B. Vundurthy, and M. Travers, "Modular, resilient, and scalable system design approaches - lessons learned in the years after DARPA subterranean challenge," in *IEEE ICRA Workshop on Field Robotics*, 2024.