Fisher Information Based Active Planning for Aerial Photogrammetry

Fisher Information Based Active Planning for Aerial Photogrammetry

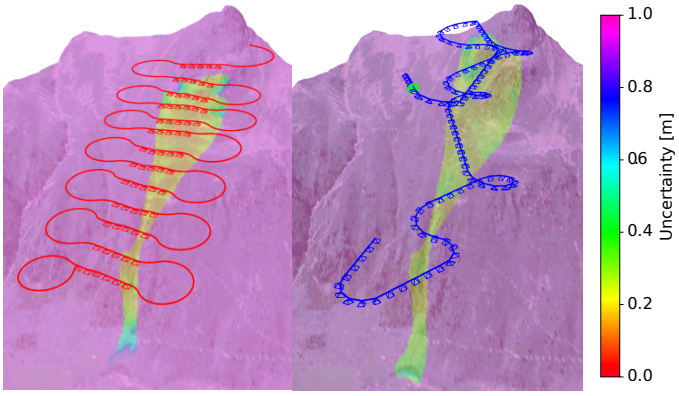
Jaeyoung Lim[1], Nicholas Lawrance[1,2], Florian Achermann[1], Thomas Stastny[1,3]
Rik Bähnemann[1], Roland Siegwart[1]

***Abstract*— Small uncrewed aerial systems (sUASs) are useful tools for 3D reconstruction due to their speed, ease of use, and ability to access high-utility viewpoints. Today, most aerial survey approaches generate a preplanned coverage pattern assuming a planar target region. However, this is inefficient since it results in superfluous overlap and suboptimal viewing angles and does not utilize the entire flight envelope. In this work, we propose active path planning for photogrammetric reconstruction. Our main contribution is a view utility function based on Fisher information approximating the offline reconstruction uncertainty. The metric enables online path planning to make in-flight decisions to collect geometrically informative image data in complex terrain. We evaluate our approach in a photorealistic simulation. A viewpoint selection study shows that our metric leads to faster and more precise reconstruction than state-of-the-art active planning metrics and adapts to different camera resolutions. Comparing our online planning approach to an ordinary fixed-wing aerial survey yields 3.2 × faster coverage of 16 ha undulated terrain without sacrificing precision.**

I. Introduction

Aerial photogrammetry is becoming increasingly important in alpine geoscience. Especially sUASs offer an alternative to costly satellite imagery due to their capability to reach remote places in the mountains on demand. For example, avalanche researchers use sUASs to generate high spatiotemporal resolution maps to measure snow depth to predict hazardous areas [1]. Today, commercial platforms provide autonomous mission planners to collect image data. Most of these rely on simple coverage planning. An operator defines a region of interest (ROI), altitude, and image overlap and the software computes a path with constant lane distance as shown in Fig. 1a.

However, this method oversimplifies the data collection as it assumes a planar ROI. In steep terrain, it will fail to collect sufficient image overlap with constant ground sampling distance (GSD) to compute a complete and precise photogrammetric reconstruction. This issue becomes even more severe for fixed-wing sUAS, which are suited for large area surveys but are susceptible to wind and thus deviation from the assumptions of the preplanned path and prior viewpoints [2]. To overcome this, the user usually reduces the lane distance, which then results in inefficient paths that do not necessarily improve the reconstruction quality.



(a) *Coverage* (b) *Ours*

Fig. 1: Comparison of coverage planning and our proposed active view planner mapping a previously outlined avalanche in Davos, Switzerland, mapped by SLF [3], colored by expected map uncertainty.

In this work, we approach data collection for photogrammetry as an active view planning problem. Here, a utility function defines the quality of a new viewpoint, and the planner iteratively selects the next measurement. As Fig. 1b shows, this paradigm leads to shorter coverage missions and improves view quality over the whole ROI.

Our utility function is an approximation of the offline photogrammetry process. Given a prior surface geometry and the camera frustum, it evaluates the Fisher information for every potential landmark in the scene. Multiple viewpoints observing the same landmark from different angles reduce the landmark uncertainty. Maximizing the Fisher information ultimately improves the offline reconstruction quality.

Our contributions are:

- An online proxy function for predicting offline surface reconstruction quality using viewpoint geometry.
- A resolution adaptive viewpoint planning strategy using the proposed information metric for offline surface reconstruction.
- Demonstration and evaluation of the proposed method with a dynamically constrained aerial vehicle model.

II. Related Work

Traditionally, aerial photogrammetry is approached as a 2D coverage problem solved offline and executed without online adaptation. The paths are computed by Boustrophedon cell decomposition for a polygonal ROI [4]–[6], and the

This work was supported by ETH Research Grant AvalMapper ETH-10 20-1, by NCCR DFab P3 through the Swiss National Science Foundation, and by Amazon Research Award #4808. We would like to thank Yves Bühler and Elisabeth Hafner at the WSL-Institute for Snow and Avalanche Research (SLF) for their expertise and support in avalanche mapping and modelling.

[1] Autonomous Systems Lab, ETH Zürich, Zürich 8092, Switzerland {jalim, acfloria, brik, rsiegwart}@ethz.ch
[2] The Robotics and Autonomous Systems Group, CSIRO Data61, QLD 4069, Australia, nicholas.lawrance@csiro.au
[3] Auterion AG, Zürich 8045, Switzerland thomas.stastny@auterion.com

flight altitude is determined with a constant offset to the varying terrain altitude [6]. However, such patterns result in variable overlaps between images and inconsistent ground sampling distances for non-planar, complex terrain. Further, for dynamically constrained vehicles such as fixed-wing aircraft, it is not trivial to design a feasible trajectory [7], and such paths are not robust against unexpected environmental disturbances, such as wind gusts, leading to incomplete coverage [2]. The following online- and surface-adapting methods can compensate for such modeling errors.

Active view planning determines a sequence of viewpoints that maximizes their combined information to generate an optimal scene reconstruction. Determining such an optimal sequence is NP-hard since the informative value of a particular viewpoint depends on all previous observations [8]. Thus, robotic solutions have been approximating the true information-gathering problem.

Next-best-view (NBV) approaches build an online map representation and select informative viewpoints iteratively in a sequential greedy manner. They maximize an information metric related to the expected change of a map representation over a discrete set of candidate views. NBV methods have been used with volumetric maps in the exploration literature, where the information metric is the number of unknown voxels expected to be visible in the view frustum [9]–[12]. Surface-based methods [13]–[15] generate NBV poses based on the improvement of the surface quality of the target scene. While NBV methods are useful for navigation, there is no direct connection to improving the reconstruction quality of an offline structure-from-motion (SFM) or multi-view stereo (MVS) reconstruction process, e.g., improving landmark triangulation.

Some methods which optimize the viewpoints for offline reconstruction take an explore-exploit strategy where the reconstruction runs between flights, and the mesh quality is used to evaluate the utility of views [16]–[19]. These approaches require multiple flights, making it undesirable in our target application where we want to reduce coverage time. Learning-based approaches such as [20], [21] have tried to predict the uncertainty of an offline SFM using learning techniques. An alternative approach is to define a proxy function to evaluate the informativity of camera configurations online [22]–[25]. Roberts et al. [22] introduce a view utility based on a spherical coverage model encouraging viewpoints from varying viewing angles. However, their method does not consider frustum overlap and multi-view baseline in free space, which are essential for offline photogrammetry. [23]–[25] use a heuristically designed 3D reconstruction metric for SFM processes that encourage good camera configurations such as large triangulation angles and overlap. This purely geometric utility function applied in an NBV setting is similar to our proposed approach. However, handcrafted metrics rely on features specific to target scenes [25] and require hand-tuning of various parameters [24], [25]. In this work, we incorporate a probabilistic metric based on Fisher information, which only depends on the uncertainty of the measurement model.

The Fisher information has been used in various active perception problems such as dynamics learning [26] or active SLAM [27], [28]. Fisher information quantifies information based on the sensitivity of a parameter to a measurement. By incorporating Fisher information as a view utility metric, one can predict how much a view contributes to reducing uncertainty [27]–[29]. [27] encodes Fisher information into a volumetric map representation given known landmark positions. Our approach solves the inverse problem, where we assume the vehicle's pose is known and quantify information based on the reduced uncertainty of the landmark position.

III. Fisher Information

Consider a parameter estimation problem, where the goal is to determine the most likely parameters θ given observations of a random variable $\mathbf{z}$ and measurement likelihood density function $p(\mathbf{z}|\theta)$. Fisher information quantifies the amount of information contained in a measurement for such a parameter estimation problem.

The Fisher information matrix is defined as

$$\mathbf{I}_\theta = \left[\frac{\partial}{\partial\theta}\log p(\mathbf{z}|\theta)\right]^T \left[\frac{\partial}{\partial\theta}\log p(\mathbf{z}|\theta)\right]. \tag{1}$$

For a linear Gaussian model with parameterized mean function $\mu(\cdot)$ and measurement covariance $\Sigma(\cdot)$, $\mathbf{z} \sim \mathcal{N}\left(\mu(\theta), \Sigma(\theta)\right)$, the Fisher information matrix can be calculated from the Jacobian $\mathbf{J}_\theta$ and Σ,

$$\mathbf{I}_\theta = \mathbf{J}^T\mathbf{\Sigma}^{-1}\mathbf{J}, \quad \mathbf{J}_\theta = \frac{\partial\mathbf{z}}{\partial\theta}. \tag{2}$$

The Cramér-Rao bounds are defined as the lower bound of the covariance matrix of an unbiased estimator $\hat{\theta}$ and are inversely related to the Fisher information matrix,

$$var(\hat{\theta}) \geq {\mathbf{I}_\theta}^{-1}. \tag{3}$$

Note that the Cramér-Rao bound helps quantify the accumulated information of observations since it does not depend on a specific estimator.

IV. Problem Formulation

To approximate the offline SFM process, we assume that we are given a set of surface landmark positions representing the target geometry. The landmark distribution, and thus the geometric complexity of the terrain, is arbitrary but known. We derive an estimator for every landmark that defines the landmark uncertainty given multiple viewpoints. Minimal Cramér-Rao bounds of all landmarks will lead to a complete and precise reconstruction.

Assume a set of viewpoints $\mathcal{V}$ and surface landmarks $\mathcal{L}$. A single viewpoint $v_k \in \mathcal{V}$, is defined by the camera position $\mathbf{p}_k \in \mathbb{R}^3$ and orientation $R_k \in \mathsf{SO}(3)$. Each viewpoint can contain observations of multiple landmarks. Fig. 2 visualizes the surface and landmark geometry.

A single observation $\mathbf{z}_i^k$ of landmark $l_i \in \mathcal{L}$ from v_k is defined by $\mathbf{p}_k$ and the normalised relative bearing vector $\hat{\mathbf{f}}_i$ from the camera to the landmark position $\mathbf{l}_i \in \mathbb{R}^3$, such that

$$\mathbf{z}_i^k = \begin{pmatrix}\hat{\mathbf{f}}_i^k & \mathbf{p}_k\end{pmatrix}, \quad \text{where } \hat{\mathbf{f}}_i^k = \frac{\mathbf{l}_i - \mathbf{p}_k}{\|\mathbf{l}_i - \mathbf{p}_k\|}. \tag{4}$$

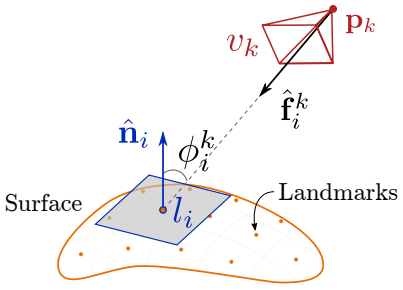


Fig. 2: Surface and viewpoint geometry.

Since a landmark is typically visible from multiple viewpoints, we define the set of viewpoints in which landmark l_i is visible as $\mathbf{v}_i \subset \mathcal{V}$. For each observation z_i^k we define the information matrix $\mathbf{I}_i^k$. Since information is additive, the information matrix of a landmark $\mathbf{I}_i$ visible from multiple viewpoints $\mathbf{v}_i$ can be quantified as the sum of the information matrices of all viewpoints where the landmark is visible, then

$$\mathbf{I}_i = \sum_{v_k \in \mathbf{v}_i} \mathbf{I}_i^k. \tag{5}$$

Our goal is to find an optimal set of viewpoints $\mathcal{V}^\star$ that maximizes the information of all landmarks $\mathcal{L}$ on the surface,

$$\mathcal{V}^\star = \underset{\mathbf{v} \subset \mathcal{V}}{\operatorname{argmin}} \sum_{l_i \in \mathcal{L}} Q_i(\mathbf{v}_i), \tag{6}$$

for the information-based cost function $Q(\cdot)$, defined in Section V. For dynamically constrained views, such as aerial systems, the view sequence must also satisfy the dynamic constraints of the platform, therefore we add an additional constraint to the viewpoint sequence

$$v_{k+1} \in \{v \mid \exists\, u \in \mathcal{U} \text{ s.t. } v = f(v_k, u)\} \tag{7}$$

where $f(\cdot)$ is the dynamic constraint between the viewpoints and $\mathcal{U}$ is the set of feasible control inputs.

V. View Utility

Here we define the novel view utility function that is an online proxy for the quality of the final photogrammetric reconstruction. We consider multiple factors within the view utility, but we are primarily interested in finding a function that can be calculated on future viewpoints and correlates well with final reconstruction quality (coverage and accuracy).

A. Approximating Visibility

We model visibility as a combination of the camera's incidence angle and field of view (FOV). We define priors of these individual components and approximate them with a visibility function $\lambda(v_k, l_i) \in [0, 1]$

$$\lambda(v_k, l_i) = P_{fov}(v_k, l_i) P_{incidence}(v_k, l_i). \tag{8}$$

1) Field of view: In order to approximate if landmark l_i is visible from viewpoint v_k, we use a binary variable that checks whether the bearing vector to the target landmark is inside the camera field of view.

$$P_{fov}(v_k, l_i) = \begin{cases} 1, & \text{if } l_i \text{ is inside FOV of } v_k \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

2) Incident Angle: The incident angle is a visibility constraint of the surface, where the visibility of the landmark improves as the viewing angle becomes less oblique. We follow [30] and define the incident prior as

$$P_{incidence}(v_k, l_i) = \exp\left(-{\phi_i^k}^2 / \left(2\sigma_k^2\right)\right), \tag{10}$$

where $\phi_i^k = \cos^{-1}(\hat{\mathbf{f}}_k^i \cdot (-\hat{\mathbf{n}}_i))$ and σ_k is defined so that $P_{incidence} = 0$ when the surface normal $\hat{\mathbf{n}}_i$ is orthogonal to bearing vector $\hat{\mathbf{f}}_k^i$.

B. Landmark Uncertainty

We are interested in estimating landmark locations from multiple bearing-only observations, so we treat landmark localization as an estimation problem with the parameter vector being the 3D position of the landmark $\theta = \mathbf{l}_i$. Each landmark observation consists of a camera's bearing vector $\hat{\mathbf{f}}_i^k$ and an observation position $\mathbf{p}_k$. We calculate the Fisher information matrix (2) as a function of the Jacobian of the observation relative to the landmark location $\mathbf{J}_\theta$ and the covariance matrix $\mathbf{\Sigma}_z$.

$$\begin{aligned} \mathbf{J}_\theta &= [\frac{\partial \hat{\mathbf{f}}_i^k}{\partial \theta} \quad \frac{\partial \mathbf{p}_k}{\partial \theta}] \\ &= [\frac{1}{n} - \frac{1}{n^3}(\mathbf{l}_i - \mathbf{p}_k)(\mathbf{l}_i - \mathbf{p}_k)^T \quad \mathbf{0}], \\ \text{where } \; n &= \|\mathbf{l}_i - \mathbf{p}_k\|. \end{aligned} \tag{11}$$

The landmark uncertainty is only influenced by the bearing, since the Jacobian is zero for position observations. The covariance matrix consists of bearing uncertainty and position uncertainty. The bearing uncertainty is a function of the angular resolution of the image, which makes our approach adapt to different camera resolutions.

The Fisher information matrix for a single observation is singular since it is not possible to determine the 3D location of a landmark from a single view.Therefore we need to aggregate the Fisher information matrices of multiple views. From (5), the accumulated Fisher information can be found by summing individual matrices from each viewpoint. We modify (5) by including the visibility metric from (8) to estimate the accumulated Fisher information matrix,

$$\hat{\mathbf{I}}_i = \sum_k \lambda(v_k, l_i) \mathbf{I}_i^k \tag{12}$$

Then, the Cramér-Rao bound σ_i for landmark l_i is a lower bound of the variance of an unbiased estimator of the landmark position. It is the inverse of the accumulated Fisher information matrix $\hat{\mathbf{I}}_i$ and estimates the quality of the landmark location prediction. We use the E-optimality from the so-called "alphabet-soup" of experimental design criteria [31] by using the maximum eigenvalue.

$$\sigma_i = \sqrt{max(eig(\hat{\mathbf{I}}_i^{-1}))} \tag{13}$$

Fig. 3 visualizes the spatial distribution of the Cramér-Rao bounds for two viewpoints over a level surface. The bounds predict low reconstruction uncertainty in the center

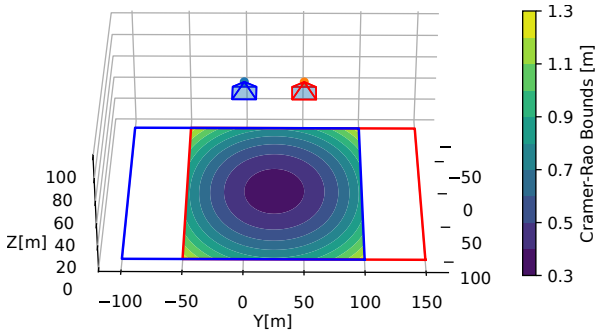


Fig. 3: Spatial distribution of Cramér-Rao bounds (coloured surface) from two nadir viewpoints $v_0 = [0, 0, 100]$ and $v_1 = [0, 50, 100]$ over a level surface. Colored outlines show the respective field of view of each camera.

of the two opposing views and higher uncertainty on the overlap borders where both perspectives are similar, and the incidence angles are oblique. The Cramér-Rao bound is only defined in the overlapping region since the information is singular with only one view.

C. Map Uncertainty

From the uncertainty metric of a single landmark, we define the map uncertainty over the whole surface. It consists of the Cramér-Rao bounds of all $l_i \in \mathcal{L}$ given a viewpoint set $\mathbf{v}$,

$$Q_{\mathcal{L}}(\mathbf{v}) = \sum_{l_i \in \mathcal{L}} min(\sigma_i, \sigma_{max}). \tag{14}$$

σ_{max} is the maximum Cramér-Rao bound. It is an upper limit on the landmark uncertainty, avoiding cost overflow due to ambiguous and weakly observed landmark positions.

VI. Planning

A. View Utility

The view utility describes the reduction of map uncertainty through new viewpoints. Given a candidate viewpoint set $\mathbf{v}'$ (which could also be a single view $\mathbf{v}' = \{v'\}$), the incremental view utility U is the change of map uncertainty,

$$U(\mathbf{v}'|\mathbf{v}) = Q_{\mathcal{L}}(\mathbf{v}) - Q_{\mathcal{L}}(\mathbf{v} \cup \mathbf{v}'). \tag{15}$$

B. View Planning

NBV planning methods try to find the next view that maximizes the view utility.

$$v^* = \underset{v \in \mathcal{V} \setminus \mathbf{v}}{\operatorname{argmax}}\, U(v|\mathbf{v}) \tag{16}$$

While the NBV is the globally optimal viewpoint in terms of view utility, it is often the case that the optimal view is far away or hard to achieve with the dynamic constraints of a mobile robot. Given that it is relatively inexpensive to acquire images on-the-fly, it is better to find the best set of dynamically feasible viewpoints rather than a single viewpoint. Then, from current viewpoint v_k,

$$\begin{aligned} \mathbf{v}^* &= \underset{\mathbf{v}' \subset \mathcal{V} \setminus \mathbf{v}}{\operatorname{argmax}}\, U(\mathbf{v}'|\mathbf{v}) \\ \text{s.t.} \quad \mathbf{v}' &\subset \{v \mid \exists\, u \in \mathcal{U} \text{ s.t. } v = f(v_k, u)\}. \end{aligned} \tag{17}$$

In this paper, we consider the motion constraints of a fixed-wing aerial vehicle. We use a motion primitive planner that generates a set of candidate trajectories, calculates the set of camera poses along each trajectory, and selects the one that maximizes the increase in view utility (15) from the corresponding viewpoint set.

VII. Map Representation

To this point, our viewpoint utility makes no assumptions on the underlying prior surface model, and landmarks can be arbitrarily distributed on the surface. The landmarks the SFM reconstruction uses are typically salient image features that can be robustly associated across multiple images. However, we do not know where the landmarks are expected on the surface without processing the image. In this work, we select a map representation that assumes uniformly distributed proxy landmark locations across the target region and evaluate view utility over these landmarks. Further, we assume access to a surface geometry prior in the form of a digital elevation map (DEM), which is typically required for flight planning purposes. These assumptions suit our target application of terrain mapping, where the surface quality is treated equally.

We approximate the terrain surface $\mathcal{L}$ as a 2.5D regular grid map $\mathcal{M} = \bigcup_{\forall i} c_i$ [32] where each cell,

$$c_i = [\mathbf{x}_i, h_i, \hat{\mathbf{n}}_i, \sigma_i, \mathbf{I}_i, A_i]. \tag{18}$$

For each cell c_i in the grid, $\mathbf{x}_i$ is the x, y position of the cell, h_c is the altitude of the cell, $\hat{\mathbf{n}}_i$ is the surface normal at the center of the cell, and A_i is the cell area. We assume that a landmark l_i is located at the center of each cell, and therefore we optimise our utility function over these landmarks.

As we are discretizing the surface, the absolute value of the map uncertainty depends on the grid resolution. To address this, we normalize the map uncertainty (14) by area,

$$Q_{\mathcal{M}}(\mathbf{v}) = \frac{\sum_{c_i \in \mathcal{M}} \min(\sigma_i, \sigma_{max})\, A_i}{\sum_{c_i \in \mathcal{M}} A_i}. \tag{19}$$

VIII. Evaluation Setup

A. Simulation Setup

Obtaining ground truth terrain models is challenging in the real world and is essentially the primary goal of this work. To evaluate the proposed approach, we used the Unreal Engine-based photorealistic simulation framework Airsim [33] in a mountainous environment. We considered two different target areas, one inside a *valley* ($382\,\mathrm{m} \times 412\,\mathrm{m}$) and the other one on a *slope* ($400\,\mathrm{m} \times 600\,\mathrm{m}$). The images were captured at a resolution of 1440×1080 pixels and used to reconstruct the surface with COLMAP [30], [34], an SFM, and an MVS pipeline.

B. Evaluation

1) Sampling and registration: The Delaunay meshes resulting from the reconstruction are projected onto a 2.5D map with the same resolution ($1\,\mathrm{m}$) as the ground truth. The mesh is registered through an iterative closest point (ICP) on the viewpoint positions provided from the simulation after the SFM pipeline has finished.

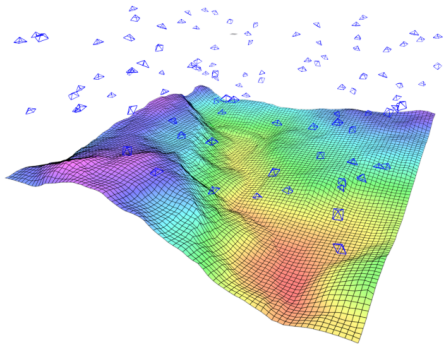


Fig. 4: Visualization of terrain and a set of 100 candidate views for NBV planning

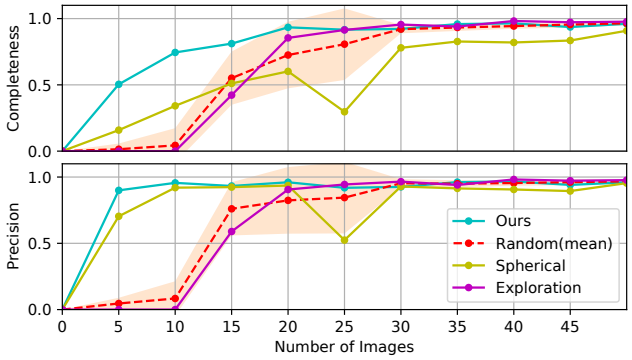


Fig. 5: Incremental reconstruction quality for different NBV metrics. *Ours*, *Spherical* and *Exploration* are deterministic. *Random* shows standard deviation.

2) Metrics: We evaluate reconstruction precision and completeness as in MVS benchmarks [35], [36]. Precision represents how much of the resampled reconstructed mesh $\mathcal{R}$ is within the error threshold e_{max} of the ground truth mesh $\mathcal{G}$,

$$P(e_{max}) = \frac{1}{|\mathcal{R}|} \sum_{i \in \mathcal{R}} [e_{g \to r} < e_{max}]. \tag{20}$$

Completeness shows how much the reconstruction covers the target surface within the error threshold e_{max} from the ground truth map $\mathcal{G}$.

$$C(e_{max}) = \frac{1}{|\mathcal{G}|} \sum_{i \in \mathcal{G}} [e_{r \to g} < e_{max}] \tag{21}$$

Precision and completeness complement each other. For example, if a small area has a very precise reconstruction, the precision metric will be high but the completeness metric will be low.

IX. Next-Best-View Planning

A. Setup

We compare our metric to the spherical coverage metric of [22], an exploration planner [9], and randomly selected views from a finite view set $\mathcal{V}$ as described in Section VI-B. We populate $\mathcal{V}$ with 100 randomly sampled viewpoints, with positions sampled uniformly above the target region, altitude from $U(50\,\mathrm{m}, 150\,\mathrm{m})$, and orientations from $U(-45°, 45°)$ as visualized in Fig. 4. Therefore, different view utility metrics differ on how quickly the reconstruction becomes complete.

For the spherical coverage metric of [22], we used a reference distance $t_0 = 100\,\mathrm{m}$ and half distance (the additional distance at which the utility halves) of $t_{\mathrm{half}} = 30\,\mathrm{m}$. The view utility for the exploration planner is the number of unobserved landmarks (map cells) on the surface. For the purpose of reconstruction, we count a map cell *observed*, if it has two or more viewpoints.

B. Results

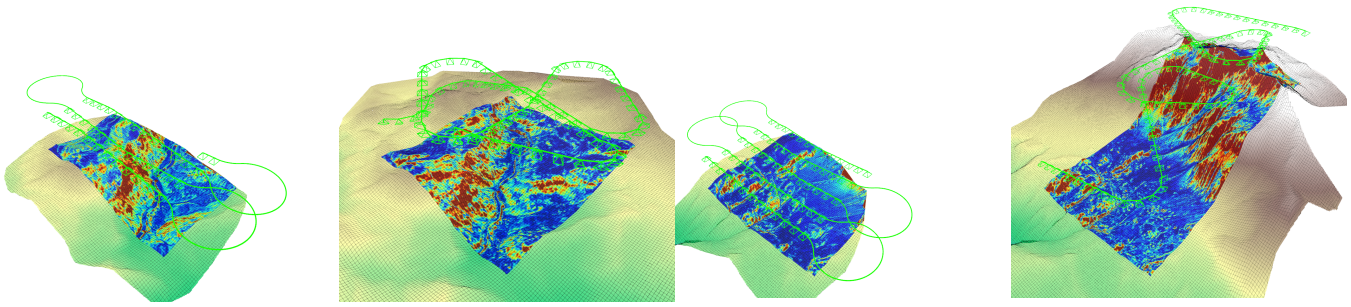
Fig. 5 shows the progress of completeness and precision of the reconstruction result as the viewpoints are selected sequentially with different metrics. Our proposed viewpoint metric achieves the fastest completion with immediate high precision. This shows that our utility indeed optimizes both coverage and precision. The utility prioritizes finding two viewpoints with favorable triangulation angles for all landmarks to minimize Eq. (14). Given the consistently high precision of our approach, even an early mission abortion will deliver a useful reconstruction. All other metrics achieve slower progress on reconstruction results. The spherical metric [22] also ensures high precision but not necessarily good overlap; thus, reconstruction is not guaranteed, as the drop at 25 images indicates. The exploration method performs most inadequately as expected. It does not optimize image overlap.

Since our proposed information metric explicitly includes the bearing uncertainty (see Sec. V-B), our metric automatically adapts to different image resolutions. To demonstrate this, we repeat the NBV experiments with lower resolution images. Fig. 7 shows that *Ours* reduces the map coverage rate with lower resolution images to maintain high precision. Our utility selects viewpoints that are closer to the surface and have more overlap. *Exploration* and *Spherical* metrics do not change the image selection with different image resolutions, resulting in poor reconstruction qualities with lower resolution images. This even yields to situations where the low-resolution reconstruction fails at 10 and 20 images due to a lack of feature overlap.

X. Dynamically Constrained View Planning

A. Setup

The NBV selection study demonstrated that the proposed view utility metric selects informative viewpoints. We also show that this information metric accelerates photogrammetric aerial surveys for dynamically constrained aerial vehicles. In this section, we consider a fixed-wing aircraft modeled as a *Dubins Airplane* [37], where the vehicle is flies at $15\,\mathrm{m\,s^{-1}}$, with a minimum turn radius of $60\,\mathrm{m}$ and climb rates $\{-3, 0, 3\}\,\mathrm{m\,s^{-1}}$. A receding horizon motion primitive planner with $10\,\mathrm{s}$ segments and a planning depth of 3 segments with 7 candidate maneuvers is used to generate candidate view sets where viewpoints are sampled at a fixed rate ($0.5\,\mathrm{Hz}$) along the path. The vehicle executes the first segment of the best candidate sequence, and the planning is repeated. One planning cycle (1715 evaluated views) takes $4.30 \pm 4.00\,\mathrm{s}$ on a desktop computer running



(a) *Coverage*, *Valley*, $t = 113s$ (b) *Ours*, *Valley*, $t = 120s$ (c) *Coverage*, *Slope*, $t = 224s$ (d) *Ours*, *Slope*, $t = 200s$

Fig. 6: Qualitative comparison of a reconstruction snapshot at completion of our planner over two terrains. The colored patch shows the height error over the ROI from dark blue $0\,\mathrm{m}$ to dark red $>1\,\mathrm{m}$. The proposed method yields near-complete coverage and precise reconstruction while *Coverage* is not even half way finished.

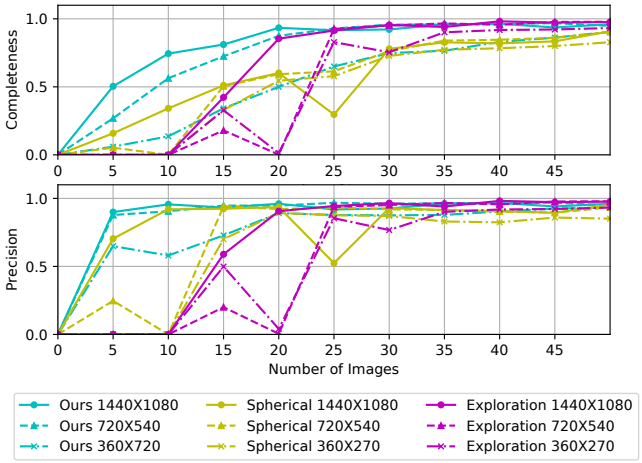


Fig. 7: Reconstruction quality comparison with different resolution images.

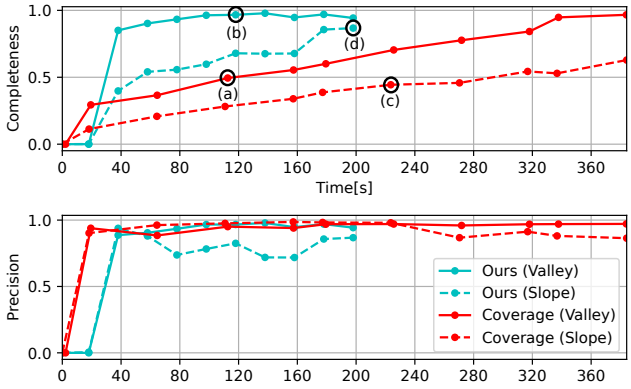


Fig. 8: Reconstruction quality over time for dynamically-constrained viewpoint planning up to 100 viewpoints. The circles mark the time stamps of the terrain snapshots in Fig. 6.

on a $2.9\,\mathrm{GHz}$ Intel core i7-10700 CPU. For the baseline method, we consider a coverage path with $75\,\%$ overlap, a standard configuration for aerial photogrammetry, and the same kinematic constraints. The approaches are evaluated on both *valley* and *sloped* terrain.

B. Results

Fig. 8 shows the reconstruction progress over the first 100 survey images for every 10 images. Our approach has a short time to completion, while *Coverage* only gains linearly. For the *valley* terrain, we reach completeness with high precision after $120\,\mathrm{s}$, which is $3.2\times$ faster than *Coverage* which finishes after $384\,\mathrm{s}$. Fig. 6 shows a snapshot of the reconstruction at the time when *Ours* completed reconstruction and *Coverage* is still running. The *Coverage* approach creates many redundant views while our planner almost immediately covers the entire map with dynamic maneuvers. The quick completion is mainly due to the metric taking advantage of oblique views while *Coverage* orients the camera nadir. Fig. 6(D) shows an interesting behavior where our planner climbs up the terrain to find a balance between covering a large area and creating a precise reconstruction.

XI. Conclusions and Future Work

In this work, we proposed an active mapping strategy for offline reconstruction that uses an online Fisher information-based view utility metric. The probabilistic method only requires a prior landmark distribution and camera model to predict the landmark reconstruction quality from multiple camera views. Our aerial mapping target application shows that the metric is a good predictor of an offline photogrammetry pipeline. It selects highly informative viewpoints with effective feature baselines and terrain coverage leading to precise and complete reconstruction. Aerial survey strategies based on the proposed metric and receding horizon planning result in precise reconstruction much faster than exhaustive coverage surveys.

Currently, the proposed method only considers the geometric viewpoint configuration. Future work will consider landmark quality estimation based on the texture of the surface. Further, while a (snow-free) prior terrain model is available for the current target application, this may not always be the case, and future work could evaluate the view utility sensitivity to poor prior models or online reconstructions in unknown terrain.

REFERENCES

[1] Y. Bühler, M. S. Adams, R. Bosch, and A. Stoffel, "Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): Potential and limitations," *Cryosphere*, vol. 10, no. 3, pp. 1075–1088, 2016.

[2] M. Coombes, W. H. Chen, and C. Liu, "Boustrophedon coverage path planning for UAV aerial surveys in wind," in *International Conference on Unmanned Aircraft Systems, ICUAS*, 2017, pp. 1563–1571.

[3] WSL Institute for Snow and Avalanche Research SLF. (2022) Avalanche mapping in davos. [Online]. Available: https://www.slf.ch/en/avalanches/avalanche-warning/avalanche-mapping.html

[4] H. Choset, "Coverage for robotics–a survey of recent results," *Annals of mathematics and artificial intelligence*, vol. 31, no. 1, pp. 113–126, 2001.

[5] H. Choset and P. Pignon, "Coverage path planning: The boustrophedon cellular decomposition," in *Field and Service Robotics*, A. Zelinsky, Ed. Springer London, 1998, pp. 203–209.

[6] R. Bähnemann, N. Lawrance, J. J. Chung, M. Pantic, R. Siegwart, and J. Nieto, "Revisiting boustrophedon coverage path planning as a generalized traveling salesman problem," in *Field and Service Robotics*, G. Ishigami and K. Yoshida, Eds., 2021, pp. 277–290.

[7] A. Xu, C. Viriyasuthee, and I. Rekleitis, "Optimal complete terrain coverage using an unmanned aerial vehicle," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 2513–2519.

[8] L. Heng, A. Gotovos, A. Krause, and M. Pollefeys, "Efficient visual exploration and coverage with a micro aerial vehicle in unknown environments," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 1071–1078.

[9] J. Delmerico, S. Isler, R. Sabzevari, and D. Scaramuzza, "A comparison of volumetric information gain metrics for active 3D object reconstruction," *Autonomous Robots*, vol. 42, no. 2, pp. 197–208, 2018.

[10] H. Oleynikova, Z. Taylor, R. Siegwart, and J. Nieto, "Safe local exploration for replanning in cluttered unknown environments for microaerial vehicles," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1474–1481, 2018.

[11] L. Schmid, M. Pantic, R. Khanna, L. Ott, R. Siegwart, and J. Nieto, "An efficient sampling-based method for online informative path planning in unknown environments," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1500–1507, 2020.

[12] M. Selin, M. Tiger, D. Duberg, F. Heintz, and P. Jensfelt, "Efficient autonomous exploration planning of large-scale 3-d environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1699–1706, 2019.

[13] R. Border, J. D. Gammell, and P. Newman, "Surface edge explorer (SEE): Planning next best views directly from 3D observations," in *IEEE International Conference on Robotics and Automation*, 2018, pp. 6116–6123.

[14] G. A. Hollinger, B. Englot, F. Hover, U. Mitra, and G. S. Sukhatme, "Uncertainty-driven view planning for underwater inspection," in *IEEE International Conference on Robotics and Automation*, 2012, pp. 4884–4891.

[15] A. Bircher, K. Alexis, M. Burri, P. Oettershagen, S. Omari, T. Mantel, and R. Siegwart, "Structural inspection path planning via iterative viewpoint resampling with application to aerial robotics," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 6423–6430.

[16] B. Hepp, M. Nießner, and O. Hilliges, "Plan3D: Viewpoint and trajectory optimization for aerial multi-view stereo reconstruction," *ACM Transactions on Graphics*, vol. 38, no. 1, pp. 1–17, 2018.

[17] C. Peng and V. Isler, "Adaptive view planning for aerial 3D reconstruction," in *IEEE International Conference on Robotics and Automation*, 2019, pp. 2981–2987.

[18] S. Song, D. Kim, and S. Jo, "Active 3d modeling via online multi-view stereo," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5284–5291, ISSN: 2577-087X.

[19] S. Arce, C. A. Vernon, J. Hammond, V. Newell, J. Janson, K. W. Franke, and J. D. Hedengren, "Automated 3d reconstruction using optimized view-planning algorithms for iterative development of structure-from-motion models," *Remote Sensing*, vol. 12, no. 13, 2020.

[20] C. Mostegel, M. Rumpler, F. Fraundorfer, and H. Bischof, "UAV-based autonomous image acquisition with multi-view stereo quality assurance by confidence prediction," June 2016.

[21] B. Hepp, D. Dey, S. N. Sinha, A. Kapoor, N. Joshi, and O. Hilliges, "Learn-to-score: Efficient 3d scene exploration by predicting view utility," vol. 11219 LNCS, pp. 455–472, ISBN: 9783030012663 _eprint: 1806.10354.

[22] M. Roberts, S. Shah, D. Dey, A. Truong, S. Sinha, A. Kapoor, P. Hanrahan, and N. Joshi, "Submodular trajectory optimization for aerial 3D scanning," in *IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 5334–5343.

[23] C. Peng and V. Isler, "View selection with geometric uncertainty modeling," in *Proceedings of Robotics: Science and Systems*, Pittsburgh, Pennsylvania, June 2018.

[24] N. Smith, N. Moehrle, M. Goesele, and W. Heidrich, "Aerial path planning for urban scene reconstruction: A continuous optimization method and benchmark," accepted: 2018-10-08T13:26:27Z Publisher: Association for Computing Machinery (ACM). [Online]. Available: https://repository.kaust.edu.sa/handle/10754/628907

[25] Y. Liu, R. Cui, K. Xie, M. Gong, and H. Huang, "Aerial path planning for online real-time exploration and offline high-quality reconstruction of large-scale urban scenes," *ACM Trans. Graph.*, vol. 40, no. 6, dec 2021. [Online]. Available: https://doi.org/10.1145/3478513.3480491

[26] I. Abraham and T. D. Murphey, "Active learning of dynamics for data-driven control using koopman operators," *IEEE Transactions on Robotics*, vol. 35, no. 5, pp. 1071–1083, 2019.

[27] Z. Zhang and D. Scaramuzza, "Fisher information field: an efficient and differentiable map for perception-aware planning," 2020, [arXiv preprint arXiv:2008.03324, 2020].

[28] A. Kim and R. M. Eustice, "Active visual SLAM for robotic area coverage: Theory and experiment," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 457–475, 2015.

[29] S. Ponda, R. Kolacinski, and E. Frazzoli, "Trajectory optimization for target localization using small unmanned aerial vehicles," in *AIAA Guidance, Navigation, and Control Conference*, 2009, AIAA 2009-6015.

[30] J. L. Schönberger, E. Zheng, J.-M. Frahm, and M. Pollefeys, "Pixelwise view selection for unstructured multi-view stereo," in *European Conference on Computer Vision*, 2016, pp. 501–518.

[31] G. E. Flaspohler, "Statistical models and decision making for robotic scientific information gathering," Master's thesis, MIT, 2018.

[32] P. Fankhauser and M. Hutter, "A universal grid map library: Implementation and use case for rough terrain navigation," in *Robot Operating System (ROS): The Complete Reference*, A. Koubaa, Ed. Springer International Publishing, 2016, vol. 1, pp. 99–120.

[33] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "AirSim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Robotics*, M. Hutter and R. Siegwart, Eds. Springer International Publishing, 2018, pp. 621–635.

[34] J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 4104–4113.

[35] A. Knapitsch, J. Park, Q.-Y. Zhou, and V. Koltun, "Tanks and temples: benchmarking large-scale scene reconstruction," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–13, 2017.

[36] T. Schöps, J. L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, and A. Geiger, "A multi-view stereo benchmark with high-resolution images and multi-camera videos," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3260–3269.

[37] H. Chitsaz and S. M. LaValle, "Time-optimal paths for a Dubins airplane," in *46th IEEE Conference on Decision and Control*, 2007, pp. 2379–2384.