Terrain Reconstruction of Glacial Surfaces via Robotic Surveying Techniques

Stephen Williams, Lonnie T. Parker, and Ayanna M. Howard

Abstract—The capability to monitor natural phenomena using mobile sensing is a benefit to the Earth science community given the potentially large impact that humans have on naturally occurring processes. Such phenomena can be readily monitored using networks of mobile sensor nodes that are tasked to regions of interest by scientists. In our work, we hone in on a very specific domain, elevation changes in glacial surfaces, to demonstrate a concept applicable to any spatially distributed phenomena (e.g., temperature or humidity). Our work leverages the sensing of a vision-based odometry system and the design of robotic surveying navigation rules to reconstruct scientific areas of interest, with the goal of monitoring elevation changes in glacial regions. The reconstruction methodology presented makes use of Gaussian process regression to combine sparse visual landmarks extracted from the glacial scenery into a dense topographic map. Further, we introduce a theory behind spatial coverage, in the context of sampling, as achieved by an intelligently navigating agent. Finally, we validate the output from our methodology and provide results that show the reconstructed terrain error complies with acceptable mapping standards found in the scientific community.

The ability to understand the causes and effects of climate change is one of the foremost questions under consideration in the scientific community today. Since the 1970s, scientists have gathered weather-related measurements from around the globe to study this phenomenon, model the major contributing factors, and predict the global ramifications. It has been discovered that the world’s glacial regions are particularly sensitive to changes in climate; the dwindling ice caps are but one sign of this region’s increasing temperatures [1], [2]. Due to the sensitivity of these regions, scientists have focused data gathering efforts towards the poles, setting up networks of automatic weather stations in Greenland and Antarctica [3]. These stations are expensive to install and maintain, yet provide only sparse spatial resolution in these critical areas. To augment the data collection mechanisms available to climate scientists in harsh, glacial terrain, a multi-agent robotic sensor network has been proposed [4]. The network would consist of multiple autonomous robotic rovers equipped with a customizable sensor payload. The scientists would define the region of interest and desired spatial resolution, then task the network to execute the data-gathering mission.

The multi-agent nature of the proposed system poses certain design constraints. Most notably, because the system will consist of many robotic nodes, each node must be inexpensive. This pushes the design away from centimeter accuracy GPS units and military-grade IMUs and towards consumer-grade sensing equipment. Consumer-grade sensing, however, generally does not have the localization accuracy necessary for the positioning of the robotic nodes into the requested sampling topology. Commodity sensors must therefore be augmented with other real-time measurements to create a higher-accuracy localization system.

For efficient traversal to the designated location, each robotic node must plan a safe and efficient path through the terrain. While the algorithms employed by global path planners differ significantly, from the dynamic programming methods of Dijkstra’s algorithm to the random sampling methods of rapidly-exploring random trees [5], all planning strategies require a map on which to plan. Coarse scale maps are generally available from remote sensing technologies. However, at typical resolutions greater than 100 [m] [6], these maps are unable to capture rover-scale terrain structures that could impede travel or affect the accuracy of derived scientific measurements. Additionally, glacial terrain is often dynamic in nature, with snow dunes and exposed ice changing shape and location over time. Examples of these time-varying terrain elements are shown in Figure 1. In order for the planned paths to be useful, the coarse-scale terrain map must be augmented with local-scale features encountered by the robotic platform during the traverse. Ideally, the terrain should be sensed or predicted before the platform encounters these obstacles, allowing new paths to be planned far in advance.

In addition to improving navigational performance, the terrain reconstruction itself can be a key scientific data product. Currently, many remote sensing methods lack the instrumentation necessary to collect important spatial detail of portions of the glaciers of interest to Earth scientists [3]. Sparse automatic weather stations cannot provide the preferred spatial detail. Furthermore, human field expeditions, especially for Arctic surveys in regions where glacial melt...
creates significant hazards, cause safety concerns. Even if the centimeter resolution desired by Earth scientists is obtainable by modern sensing techniques, such as interferometric synthetic aperture radar (InSAR), the specific areas of interest may not always be accessible. For example, naturally occurring occlusions can prevent direct line-of-sight measurements by remote sensing instrumentation while temporally-based constraints, such as weather conditions or satellite orbits, may preclude measurements at the desired time.

The SeaMonster project [7], for example, relies on short-term experiments (10 - 15 days) to measure the transient hydrological processes taking place in Lemon Creek glacier near Juneau, AK. Their work in seasonal snow melt modeling requires registering data collected from static sensor nodes with elevation maps. While time-variant measurements obtained from the static nodes are collected on a semi-daily basis, they must rely on the same, outdated DEM data at each measurement time step, i.e., every 1 to 2 days. A robotic survey system could provide spatially relevant terrain data in synchronization with this sampling frequency, allowing subtle changes in topography to be correlated with other measurement data. Although high-resolution imaging techniques like InSAR are very powerful, conducting repeatable surveys with a robotic surveyor system may prove a more flexible alternative [8]. The value in developing robotic surveying solutions is found, not only in terrain reconstruction, but for other applications requiring the intelligent sampling of geophysical science information. This is particularly true when such information is spatially distributed and not easily accessible by remote sensing technology [4], [9]. Other motivations include mineral prospecting or chemical concentration monitoring in soil [10], [11].

In the following sections, a vision-based simultaneous localization and mapping (SLAM) algorithm is described that was tailored to meet the challenges of using vision systems in low-contrast, glacial environments. As a by-product of calculating the robot’s pose estimate, the SLAM system also estimates the positions of a large number of terrain landmark points. An adaptive terrain reconstruction methodology is proposed that creates a topographic terrain map using these vision-based terrain measurements as input. Additionally, prior terrain knowledge, such as course-scale satellite elevation measurements, can be incorporated into the terrain model in a natural way, further improving the reconstruction quality. This is originally motivated by the need for forward-looking maps for path planning algorithms. However, once the focus changes to “maps as the end product,” different planning mechanisms are needed. Surveying strategies are discussed in the context of sampling methodologies, as well as methods for selecting the surveying path parameters based on prior knowledge of the terrain and common mapping standards. We also introduce the concept of science-centric coverage, a spatially-relevant performance metric, to better evaluate the meaning of collected science information as it relates to the surveying strategies. Both the vision-based terrain sampling methods and the described surveying path planners are validated within a simulation environment created to mimic a glacier field test site.

I. ROBOTIC SENSOR NETWORK

Previous arctic robotics projects, such as Nomad out of CMU [12] and MARVIN from the University of Kansas [13], showcase the ability of the mechanics of a robot to survive the inhospitable climate of glacial environments. However, each of these projects involves the construction of a single, expensive robotic agent. Such an approach is not practical for the development of multi-agent systems, where potentially dozens of robotic agents will be utilized. Three low-cost prototype mobile sensor nodes were constructed as part of this research, enabling data collection and autonomous field testing in analogous arctic terrain.

High terrain mobility is required for testing and proper execution of science missions. While much of the rover’s time will be spent in the central regions of the glacier, the project goal is to construct a system capable of traversing the widest range of expected terrain possible. Typically, the areas of most interest to scientists occur at the extremes of the environment. Collecting data about a forming glacial lake requires descending into the surrounding basin as shown in Figure 2, while investigating the glacier-mountain boundary requires ascending steep slopes.

For these reasons, a snowmobile chassis was selected as the base for the SnoMote prototype robotic mobile sensor [14]. The chassis, based on a 1/8-scale RC model, was heavily modified to incorporate a dual-track design. The modified platform has been equipped with an on-board embedded computer, consumer-grade GPS unit for global localization, and a wide-angle monocular camera for real-time image processing. Only a minimal amount of sensing was incorporated into the rover design to test the extents to which the vision system could supply the situational awareness and terrain assessment needs of the mobile rover.

II. VISION-AUGMENTED LOCALIZATION

Because many agents will be required to perform a data collection task, each agent must cope with low-cost, commodity sensors. However, consumer-grade GPS receivers and IMUs do not have the localization accuracy necessary to position the robotic nodes into the request sampling topology. These
sensors must be augmented with additional information to produce a viable system. In particular, vision is an attractive option. It is the sensing modality relied upon most by humans, and it has been shown effective for both the Mars rovers [15] and DARPA Grand Challenge vehicles [16].

To augment the GPS localization system, a visual odometry system has been implemented, based on SLAM techniques. Vision-based SLAM systems seek to estimate the 3D pose of the camera by tracking the coordinates of visually distinct features in the environment. As the features move in image space, the relative motion is used to update the position of the camera, as well as estimate the 3D location of the features themselves. This requires that image features be reliably extracted and matched within the image stream. However, glaciated environments generally lack these types of distinctive features. The next sections briefly describe the vision system and sensor fusion techniques employed in the vision-augmented localization subsystem.

A. Image Preprocessing

Since standard feature detectors search for pixels exhibiting strong directional gradients, the foreground image gradient must be boosted for these detectors to perform properly in low-contrast glacial environments. Ideally, the image enhancement should be non-uniform, adaptively enhancing the foreground regions while leaving areas of sufficient contrast alone. A contrast-limited adaptive histogram equalization (CLAHE) preprocessing stage has been shown to drastically improve both the detection rate and matching consistency of standard feature detectors when applied to glacial scenes [17].

CLAHE separates the image into different contextual regions. Within each region, a histogram equalization procedure is calculated and stored. To prevent blocking artifacts, the level of equalization performed is interpolated from neighboring regions. Finally, to prevent over-enhancement of local areas, a contrast limit is imposed.

As the goal of the contrast enhancement procedure is to extract features from the snowy foreground, the algorithm performs best if background elements, such as sky and distant mountains, are first removed. However, standard image segmentation algorithms, such as region growing methods or machine learning based approaches, generally use information local to the examined pixel to make segmentation decisions. The properties of glacial images make local examination problematic. Overcast skies, common in glacial environments, often share the same color range as the ground plane snow. To that end, a horizon line detection scheme has been developed that uses multiple visual cues to rank candidate horizon segments, then constructs a horizon line consistent with those cues [18].

Strong line segments are first extracted from the image. A minimum segment length constraint is enforced to remove the large number of noise-induced edges. A set of heuristic properties, such as segment length and color consistency, are then calculated for each remaining candidate segment. A combined weight is calculated for each candidate segment as the product of the individual weights. The top scoring candidate is selected as a seed segment for the horizon line. A greedy search is then conducted to find additional horizon line segments to connect to the seed segment. Candidate line segments that exhibit weak visual cues serve to reinforce the path of stronger segments, while segments with strong visual cues have the ability to redirect the path of the horizon. Figure 3 shows the results of the preprocessing steps on a typical glacial image acquired during field trials on Mendenhall Glacier, near Juneau, AK.

B. Feature Extraction

With the image suitably enhanced, one of a number of common keypoint detectors, such as Harris or SIFT, can be used to meet the feature detection needs of the visual odometry system [19]. After features have been extracted from the new frame, the detected features must be matched with existing landmarks. As the size of the landmark database grows, the calculation time for the feature association also grows. In this application, where the expected rover path is piecewise straight, the landmarks that are behind the rover are unlikely to ever be viewed again. Using this insight, the database is periodically culled of landmarks that are significantly behind the camera’s image plane. In practice, this limits the number of landmarks that must be actively maintained, while allowing the total number of landmarks used during the traverse to increase without bound. This also means that landmarks are only active for a short time period, reducing problems matching landmarks that were observed under different lighting conditions or at large angle differences. While the small active database size also means that loop closures are unlikely to occur, the inclusion of even low-quality GPS information in the navigation solution prevents the system from drifting over large time periods. Figure 4 shows an example of the resulting features calculated using the SIFT detector on an image from the same Mendenhall Glacier field trial. The detected features tend to track the small-scale surface undulations, visible in the enhanced image as alternating light and dark bands.

C. GPS Fusion

Position drift is one of the fundamental issues when using any incremental localization system, including SLAM. As the system runs, small errors accumulate, resulting in significant localization error over time. In order to remove this drift,
global position information, in the form of low-accuracy GPS data has been fused with vision-based SLAM.

Within this SLAM implementation, the robot state distribution is estimated using a particle filter, which approximates the true robot state distribution from a set of weighted samples. To incorporate the GPS measurements into the state estimate, an additional weighting step is applied to each particle based on the position fix and positional errors reported by the GPS. Particles that naturally evolved near the GPS measurement are reweighted to reflect the positional uncertainty. The positional errors associated with GPS survey data tend to be small and relatively uncorrelated, making them a good fit for GP interpolation. While the use of GPS measurements and high-accuracy laser scan data is common in the GP literature, vision-only reconstructions are rare.

A terrain reconstruction method is presented that uses the sparse landmark position estimates from the localization system as input to a GP interpolation system. Additionally, the GP can incorporate a priori knowledge of the terrain structure through the use of a mean function. The predicted terrain is then pulled away from the mean in response to measurements of the real terrain. This allows the course satellite data to bootstrap the reconstruction system while providing a mechanism for correcting and augmenting this information with measurements on the ground.

A Gaussian Process

A Gaussian process (GP) is a collection of an infinite number of random variables with a jointly Gaussian distribution [22]. This may be interpreted as a distribution over continuous functions, similar to how a Gaussian variable defines a distribution over real values. Instead of sampling a vector in $\mathbb{R}^N$ from the Gaussian variable, a continuous function, $f(\bar{x})$, is drawn from the GP that maps an input vector, $\bar{x} \in \mathbb{R}^N$, to an output value, $y \in \mathbb{R}$. A GP is defined by a mean function, $\mu(\bar{x})$, that describes the mean output value of all possible sample functions evaluated at the input, $\bar{x}$, and a covariance function, $k(f(\bar{x}_i), f(\bar{x}_j))$, that describes the correlation between any pair of output values. The choice of the mean and covariance functions allows prior knowledge of the function’s behavior to be encoded in the GP framework.

To perform interpolation, the GP is conditioned on a set of known measurements [22]. The resulting GP posterior describes only the subset of sample functions that pass through the measurement points. A set of unknown output values, $Y^* = \{y^*_j | j = 1, \ldots, Q\}$, can then be queried, corresponding to a set of known inputs values, $X^* = \{\bar{x}^*_j\}$. The output values are conditioned on the set of known measurements, $Y = \{y_i | i = 1, \ldots, P\}$, corresponding to a second set of known input values, $X = \{\bar{x}_i\}$. The GP posterior mean and covariance satisfying these conditions are shown in Equation (2) and (3) (with a full derivation available in [23]).
\[
p(Y^*|X, Y, X^*) \sim \mathcal{N}(\mu^*, \Sigma^*)
\]
\[
\mu^* = \mu_X + \Sigma_{Y,Y}^{-1} \Sigma_{Y,Y^*}^{-1} (Y - \mu_X)
\]
\[
\Sigma^* = \Sigma_{Y,Y}^{-1} - \Sigma_{Y,Y^*}^{-1} \Sigma_{Y,Y^*}^{-1} S_{11}^{-1} \Sigma_{Y,Y^*}^{-1}
\]
where \( \mu_X \) is a vector of values produced by evaluating the mean function, \( \mu(\cdot) \), over the set, \( S \), and \( \Sigma_{S_1, S_2} \) is a covariance matrix constructed by evaluating the covariance function, \( k(\cdot, \cdot) \), with each pair-wise combination of values from sets \( S_1 \) and \( S_2 \).

A GP can also incorporate measurement uncertainty into the reconstruction, if that uncertainty may be modeled by additive independent Gaussian noise. In that case, the measurement covariance matrix, \( \Sigma_{Y,Y} \), is simply augmented by diagonal matrix containing the uncertainty of each measurement.

Finally, while many covariance functions are possible, a common choice is the squared exponential function listed in Equation (4). This covariance function is derived from a Gaussian kernel, exhibits rotation and translation invariance to the inputs, and is infinitely differentiable or infinitely smooth. At this time, the simulation environment consists only of the smooth terrain near the center of a glacial flow, modeled after the terrain encountered during field trails. One field test was conducted in the upper section of the glacier terminus, where the underlying ice is exposed. However, even this terrain is locally smooth. This makes the selection of the standard Gaussian kernel covariance function a natural choice.

\[
k(f(\bar{x}_i), f(\bar{x}_j)) = \alpha \exp \left( -\frac{1}{2} (\bar{x}_i - \bar{x}_j)^T \Gamma (\bar{x}_i - \bar{x}_j) \right)
\]
where \( \Gamma \) is a diagonal matrix of elements \( \frac{1}{\alpha}, \ldots, \frac{1}{\gamma_N} \), and \( \alpha \) is a scaling factor. The variables in the \( N+1 \) dimensional set \( \alpha, \gamma_1, \ldots, \gamma_N \) are known as the hyperparameters for the squared exponential Gaussian process.

To train the hyperparameters, the locations, \( X \), and elevations, \( Y \), of a small segment of the terrain were provided to the GP. The values of the hyperparameters \( \alpha \) and \( \gamma \) were varied over a large range, and the corresponding terrain reconstruction error was calculated from ground truth elevation data. Since the orientation of the world coordinate system should not effect the GP results, the length scales in the two dependent variables are set equal, \( \gamma_x = \gamma_y = \gamma \). The values associated with the lowest reconstruction error, \( \alpha = 10.0 \) and \( \gamma = 315.0 \), were selected for use in the GP terrain reconstruction.

Other research into GPs has shown that non-stationary covariance functions allow the reconstruction to better model abrupt elevation changes [24], [25]. Efforts at the University of Sydney have shown improved terrain modeling performance with GPs that use neural network-inspired covariance functions [26]. If future missions require the reconstruction of more extreme terrain with severe discontinuities, more advanced non-stationary covariance functions can be employed.

### B. Visual Landmarks

The visual SLAM algorithm within the localization system produces a set of 3D point estimates that lie on the terrain surface as a byproduct of the localization process. While these point estimates are superficially analogous to GPS data or laser scan measurements, this data was collected opportunistically while the robot performs a traverse, rather than with the explicit goal of capturing terrain variations. These visual landmarks also cover the terrain only sparsely, with landmarks near the rover’s path occurring far more frequently than landmarks at significant distances. While this may be suboptimal from a terrain sampling standpoint, no additional travel is incurred by the rover to collect this data.

Further, the uncertainty of each SLAM landmark is a joint Gaussian distribution in both the dependent variables, \( (x, y) \), and the independent variable, \( z \). Inclusion of uncertainty in the dependent variables has been studied in the domain of GP regression. One approach conditions the GP on the distribution from which \( \bar{x} \) is drawn, rather than on \( x \) itself. The result of this formulation takes the form of a correction to the standard GP regression results, as shown in Equations (5) - (7) [27]. For this method to be applied, each landmark covariance is converted into a joint distribution on \( x \) and \( y \), and an independent additive noise term on \( z \) by alternately marginalizing out the other set of variables from the joint distribution. Due to the highly directional nature of visual SLAM landmark estimates, removing the dependency of \( x \) and \( y \), even from covariances with a small volume, results in a large elevation uncertainty. For this reason, only those landmark estimates whose depth uncertainty have collapsed to a small region are considered for inclusion in the GP terrain reconstruction.

\[
p(f(\bar{x}^*)|\mu_x^*, \Sigma_{x^*}) = \int p(f(\bar{x}^*)|\bar{x}^*) p(\bar{x}^*) d\bar{x}^*
\]
\[
\approx \mathcal{N}(m(\bar{x}^*), v(\bar{x}^*))
\]
\[
m(\bar{x}^*) = \mu^*
\]
\[
v(\bar{x}^*) = \Sigma^* + \frac{1}{2} \text{Tr} \left( \frac{\partial^2 \Sigma_{x^*}}{\partial \bar{x}^* \partial \bar{x}^*} \bigg|_{\bar{x}^* = \mu^*} \right)
\]
\[
+ \frac{\partial \mu^*}{\partial \bar{x}^*} \bigg|_{\bar{x}^* = \mu^*} \Sigma_{x^*} \frac{\partial \mu^*}{\partial \bar{x}^*} \bigg|_{\bar{x}^* = \mu^*}
\]
where \( p(\bar{x}^*) \sim \mathcal{N}(\mu_{x^*}, \Sigma_{x^*}) \).

### C. Satellite Elevation Data

DEM’s produced by satellite missions such as the Shuttle Radar Topography Mission (SRTM) or ICESat average the terrain elevation over a large area compared with the size of the rover. While this information cannot capture the local-scale hazards faced by the robotic sensor node, it can serve as an indication of large-scale terrain variations. Within the GP framework, a mean function, \( \mu(\bar{x}) \), is specified. This is typically set to a constant value, calculated from the mean of all the observation values. The GP then models the terrain deviation from the mean. However, if a better terrain elevation expectation is available, this can be easily incorporated into the GP. In particular, while it is referred to as a mean function, it need not be written in analytical form. It simply must provide
terrain elevations at the measurement locations, $X$, and the query locations, $X^*$. This can easily be accommodated using simpler interpolation methods on the satellite data, or even from an online mapping services such as U.S. Geological Survey (http://www.usgs.gov) or Google Earth.

**D. Simulation Examples**

The prototype robotic network has been fielded at several test sites on Mendenhall Glacier near Juneau, AK. However, performing numerical evaluation of the vision system is difficult from the field trial data. An accurate terrain map of the test site locations is unavailable, making assessment of the terrain reconstruction problematic. In order to perform comprehensive numerical analysis of the vision system results, a 3D robotic simulation was developed. This simulation system, which uses Gazebo [29] as its base, has been extended to provide a visually faithful environment including realistic large scale terrain, local scale hazards, and background imagery. Figure 5 shows a visual comparison of the simulated terrain and the real terrain from which it was developed [30]. As the simulation can provide true robot pose information and operates with a known terrain topology, it is an ideal testing platform for localization and terrain reconstruction algorithms.

To test the terrain reconstruction system, one of the Mendenhall Glacier field trial sites was reconstructed within the simulation environment. The control commands from one of the field trial traverses were sent to the simulated rover while the augmented localization algorithm executed against the simulated camera images. During the traverse, any SLAM landmark whose final 1-$\sigma$ covariance ellipse was no larger than 0.5 [m] was logged to an external database. During the simulation trial, approximately 50,000 surface landmarks were sufficiently localized, a vast majority of which occurred very near the rover’s path. Because of the proximity of these landmarks, the information they provide is largely redundant. To reduce the number of measurements that must be processed by the GP, only those landmarks that were initialized more than 5 [m] from the rover’s position are used within the reconstruction. This reduces the set to approximately 5,000 landmarks disbursed over the 600 [m] x 600 [m] simulation site.

To compare the performance of the terrain reconstruction system, three different methods have been tested. The first uses a simple linear triangular mesh interpolation method. The Delaunay triangulation [31] is formed from the SLAM landmark positions. Any query point that falls within the triangulation is estimated using the plane formed by the triangle’s vertices. Because query points must fall within a Delaunay triangle to be estimated, this method only produces terrain estimates within the convex hull of the input measurements. Also, there is no obvious mechanism for incorporating measurement uncertainty or a priori information into a triangular mesh model.

The second reconstruction incorporates the sparse visual SLAM landmark data into a Gaussian Process model. Unlike the triangular mesh interpolation scheme, the GP model is valid over all of $\mathbb{R}^2$. The mean function used within this reconstruction is a constant derived from the mean elevation of all landmarks used within the reconstruction. The final reconstruction is based on a GP model using the sparse SLAM landmarks as measurements, but also incorporates a non-constant mean function. Raw elevations were extracted from the best available SRTM data products for the simulated test site. These elevation values were used to generate a triangular mesh terrain model capable of interpolating the elevation at any point within the test environment. While the simulated environment is derived from over 200,000 unique elevations, the raw DEM contained only 64 values. The resulting terrain reconstructions of the simulation environment after the completion of the pre-planned path are shown in Figure 6.

Perhaps the most striking aspect of the three reconstructions is the limited data provided by the triangular mesh. Only 26% of the terrain could be reconstructed after the traverse was completed. In contrast, both GP reconstructions were able to predict the elevation of the entire terrain based on the local observations, even terrain sections that were located behind the rover over the entire traverse. The landmark-only GP reconstruction is able to capture the basic structure of the terrain from the limited data provided, though significant reconstruction errors exist at the terrain boundaries. The landmark plus satellite data GP reconstruction is able to make use of the provided large-scale terrain structure, drastically reducing reconstruction errors at large distances while still adapting locally to the measured environment.

**IV. ROBOTIC SURVEYING**

In the previous section, a method for generating a terrain reconstruction was presented, motivated by the need to provide an accurate map for path planning algorithms when traveling in unknown or dynamic environments. However, the terrain reconstruction itself may be viewed as valuable scientific information. To ensure the reconstruction results in a usable data product, the terrain model should comply with accepted topographic mapping standards. The GP reconstruction method discussed previously showed results based on terrain points gathered opportunistically. In order to ensure compliance with mapping requirements, we require survey paths designed to meet maximum reconstruction error requirements. In traditional robot surveying projects, navigation patterns typically follow a lawnmower structure [32], [33]. In the scientific realm, however, sample locations are planned to capture the distribution of environmental phenomena characteristics [34].
Many successful surveying techniques \cite{32}, \cite{36} focus on A. Robot Navigation by the photogrammetric and cartographic professions. In addition, our methods must sample methodologies that properly cover the space of changes points is indicated by a dashed line. The rover’s path is shown as a solid black line, while the convex hull of landmark

![Triangular Mesh Reconstruction](image1)

![Triangular Mesh Error](image2)

![SLAM Reconstruction](image3)

![SLAM Error](image4)

![SLAM+Satellite Reconstruction](image5)

![SLAM+Satellite Error](image6)

![Ground Truth](image7)

![Axonometric View](image8)

Fig. 6. Terrain reconstructions using data from an example traverse. The sensed information can be used to influence navigation decisions within smooth and continuous terrains. The differentiability condition of the terrain \cite{38} enables the use of gradient-based control rules to guide navigation. For example, the sensed terrain slope between sequential sample locations is leveraged to create a navigation policy that potentially yields a more spatially diverse set of observations.

We restrict the application of this policy to environments best approximated as $C^K$ continuous, where $K \geq 1$. This requires that at least the first derivative of the environment surface must exist and be continuous. Based on preliminary experiments with simulated terrains, sample sets that exhibit lateral, spatial diversity are more likely to reduce reconstruction error than clustered sample sets for these types of undulating terrains. The specific heuristic applied in this work is a policy defined by a switching mechanism, alternating between gradient-ascent and gradient-descent rules for changing direction based on locally-sensed features within the terrain, as described by Equation (8).

\[
  f_{s+1}(q) = \begin{cases} 
    \min(\nabla f_s(q)) & \text{if } \eta = 0 \\
    \max(\nabla f_s(q)) & \text{if } \eta = 1 
  \end{cases}
\]  

In Equation (8), the function $f_{s+1}(q)$ represents the next state or location of the navigating agent, $f_s(q)$ represents the agent’s current state, $s$ is the sample index, and $q$ is the specific location of the agent. The switching feature is represented by $\eta$, a random binary value resampled whenever the agent reaches a local extrema in the terrain or when the agent reaches a lateral boundary distance from the reference swath. For example, if the agent is navigating according to reference swaths that extend east to west, then local north-south boundaries are set as a function of swath width. As the terrain characteristics at a given state are unknown, $\eta$ is used to guide the agent towards different features within the terrain, increasing the lateral spatial diversity of collected samples.

As such, we need to employ navigation patterns based on sampling methodologies that properly cover the space of changes in environmental characteristics. In addition, our methods must be validated based on actual mapping requirements set forth by the photogrammetric and cartographic professions.

A. Robot Navigation

In robotics, the coverage problem is typically defined to maximize the total area covered by a robotic system \cite{35}. Many successful surveying techniques \cite{32}, \cite{36} focus on performing a raster scan (i.e., lawnmower) by designating evenly distributed, linear traverses across an area of interest. By designating swath widths, this type of navigation pattern enables the system to retrieve an even distribution of samples within the shortest distance possible. Unfortunately, this static approach is usually implemented for the purposes of search and thus does not adapt to environmental phenomena measured in situ during the traverse. In the scientific community, however, coverage is defined to properly measure the space of environmental phenomena \cite{34}. As such, our objectives are to define a navigation pattern that sufficiently samples the space of environmental phenomena and quantify performance with useful metrics as it relates to collecting scientific information. To achieve the first objective, we augment the lawnmower navigation pattern using a method called piecewise continuous \cite{37}. The sampling path deviates around a linear reference swath based on sensed phenomena within the terrain according to a locally-applied heuristic. This method empirically tests how sensed information can be used to influence navigation decisions within smooth and continuous terrains. The differentiability condition of the terrain \cite{38} enables the use of gradient-based control rules to guide navigation. For example, the sensed terrain slope between sequential sample locations is leveraged to create a navigation policy that potentially yields a more spatially diverse set of observations.

We restrict the application of this policy to environments best approximated as $C^K$ continuous, where $K \geq 1$. This requires that at least the first derivative of the environment surface must exist and be continuous. Based on preliminary experiments with simulated terrains, sample sets that exhibit lateral, spatial diversity are more likely to reduce reconstruction error than clustered sample sets for these types of undulating terrains. The specific heuristic applied in this work is a policy defined by a switching mechanism, alternating between gradient-ascent and gradient-descent rules for changing direction based on locally-sensed features within the terrain, as described by Equation (8).

\[
  f_{s+1}(q) = \begin{cases} 
    \min(\nabla f_s(q)) & \text{if } \eta = 0 \\
    \max(\nabla f_s(q)) & \text{if } \eta = 1 
  \end{cases}
\]  

In Equation (8), the function $f_{s+1}(q)$ represents the next state or location of the navigating agent, $f_s(q)$ represents the agent’s current state, $s$ is the sample index, and $q$ is the specific location of the agent. The switching feature is represented by $\eta$, a random binary value resampled whenever the agent reaches a local extrema in the terrain or when the agent reaches a lateral boundary distance from the reference swath. For example, if the agent is navigating according to reference swaths that extend east to west, then local north-south boundaries are set as a function of swath width. As the terrain characteristics at a given state are unknown, $\eta$ is used to guide the agent towards different features within the terrain, increasing the lateral spatial diversity of collected samples. If no significant gradient is experienced during navigation, the policy simplifies to traditional lawnmower navigation. In contrast, employing a random sampling scheme will achieve the desired spatial diversity \cite{39}, but it will also incur large costs in terms of distance traversed. The piecewise continuous navigation strategy generates a policy that correlates to sampling strategies found in the sampling and surveying literature \cite{39}, \cite{40} while also considering the practicality and cost of navigating from one sample location to another. Figure 7 illustrates the traditional sampling patterns viewed in the context of navigation waypoints, while Figure 8 provides example trajectories for the two sampling strategies.
Fig. 7. Traditional sampling patterns viewed in the context of navigation waypoints. While gridded sampling offers the advantage of resource-efficient navigation (LEFT), significant information is lost. Alternatively, random sampling achieves large-scale coverage, yet incurs large resource costs as a function of required distance traversed.

(a) Traditional Lawnmower (b) Piecewise Continuous

Fig. 8. Example of executed rover paths illustrating the spatial coverage differences of robotic surveying navigation options.

B. Science-centric Coverage

Typical navigation work in the robotics community defines coverage as the ratio of the total Euclidean distance traveled by a robotic agent, $D_S$, to some maximum distance, $D_T$. Measuring coverage in this way places attention on the agent and its performance rather than the search space and the quality of samples collected during navigation. If, instead, the search space is discretized into a finite set of possible sample locations, a more useful definition of coverage for science sampling can be generated.

We define a coverage metric relative to the cumulative sum of distances from all possible sample locations to the center of the area of interest. Percent science-centric coverage (SCC) is the ratio of $T_M$, the sum of relative distances between actual samples, $(x_m, y_m)$, and a reference location within the search space, $(X_{ref}, Y_{ref})$, to $T_S$, the sum of relative distances between all possible samples, $(x_s, y_s)$, and that same reference location. This is detailed in Equation (9).

$$ \% \text{ SCC} = \frac{T_M}{T_S} = \frac{\sum_{m=1}^{M} \left\| (x_m, y_m) - (X_{ref}, Y_{ref}) \right\|}{\sum_{s=1}^{S} \left\| (x_s, y_s) - (X_{ref}, Y_{ref}) \right\|} \tag{9} $$

For each level of desired coverage designated by a scientist, we assess a measure of percent SCC, attaching with it the success of the sample set. Specifically, the performance of the samples collected are measured in the form of root mean squared (RMS) error.

By using the representation that science-centric coverage provides, we define distances relative to a specific point, $(X_{ref}, Y_{ref})$. This relationship prioritizes the importance of the samples collected and provides each sample with a relative weight. For our experiments, we presume an approximately rectangular, discretized search space in $\mathbb{R}^2$.

C. Mapping Accuracy Standards

Lastly, we must ensure that the map regeneration we produce meets a predefined error maximum set forth by those interested in the science product. Although desired map accuracies can vary, we refer to the accepted accuracy standards employed by professionals in the cartographic and photogrammetry fields [41]. In the case of map elevation, the American society for photogrammetry and remote sensing (ASPRS) standard quantifies vertical RMS error specifications that dictate how a mapping product may be classified.

APRS standardizes the accuracy of map data when represented as a 2D contour plot, with each contour line representing a specific elevation. When seeking a map product in the form of a 2D contour map with contour separation equal to $P$, the average map error estimated must be no greater than $P/6$ or one-sixth the contour separation. Thus, a desired contour separation of 3 [ft] (0.91 [m]) requires an average error no greater than 0.5 [ft] (0.15 [m]). According to ASPRS, there exist three distinct classes of map accuracies based on contour line separations ranging from 1 [ft] to 9 [ft]. The specific maximum allowable error varies depending on the needs of the scientist, but once a value is agreed upon, it provides a benchmark for validating our results. Figure 9 shows the evaluation of these navigation patterns against typical terrains, such as the one shown in Figure 6g, relative to accepted mapping standards.

V. RESULTS

We evaluate the performance of our vision system and surveying methods based on the successful reconstruction of our environment. As a test, the system has been tasked to create a Class 3 elevation map. In this context, a reconstruction is considered successful if the final elevation model meets the minimum criteria for the map type selected. This implies a maximum terrain reconstruction error of 0.46 [m]. The area to be surveyed is a simulation of a 600 [m] x 600 [m] field test site on Mendenhall Glacier. Using satellite elevation data, the maximum terrain variation over this area is found...
to be approximately 25 [m]. However, this is merely an estimate of the terrain variation, as each satellite measurement is actually an average elevation over a large area. Using this maximum terrain variation estimate, a set of random terrains were simulated numerically, using the procedure outlined in Section IV. Figure 10 shows the average reconstruction error of the random environments when surveyed by different path planning approaches. The maximum error requirement for the Class 3 map is superimposed on the results. From this graph, the minimum number of surveying swaths for each algorithm may be extracted (four swaths for traditional lawnmower and two swaths for piecewise continuous).

The simulation system described in Section III has again been employed to validate the surveying path predictions. The simulated rover is placed at a starting point within the simulation, which is assumed to be a known location. The rover is then tasked to drive through a series waypoints calculated from the selected surveying pattern. During these traverses, the vision-augmented localization system maps the location of any visually distinct texture points encountered. These 3D surface estimates are used as inputs to the Gaussian Process (GP) terrain model. These surface points are analogous to GPS survey information, with the exception that the sampled locations are controlled by the visual surface appearance rather than a planned sampling scheme. In the case of the Piecewise Linear survey algorithm, the path actually adapts in response to the surface conditions. These decisions are based on the pose estimate of the rover, as intermediate terrain reconstructions are not available to the rover during the surveying process. Figure 11 shows the executed rover paths for each surveying strategy, and an indication of the spacial distribution of the visual landmarks.

Finally, a terrain reconstruction is performed for each surveying algorithm using the GP framework described in Section III (Figure 12). Based upon our simulated prediction of the number of swaths required to achieve the maximum ASPRS mapping standard error for a Class 3 map (0.4572 [m]), our system achieves RMS error of 0.2827 [m] and 0.3323 [m], when navigating according to the traditional Lawnmower pattern and Piecewise Continuous navigation respectively.

The outcome of our testing highlights two salient aspects of our work. The first point is the performance of our vision system as a sensor to generate useful science information for terrain reconstruction despite its inherent error-prone measurements. The information extracted from Figure 10 used to dictate the minimum number of traverses to achieve maximum error for each navigation strategy was obtained presuming a perfect sensor. In accordance with this prediction from simulation, we are pleased with the RMS error obtained with visual SLAM sensing. Yielding a difference between predicted and actual of 0.0358 [m] (12%) when navigating according to the traditional Lawnmower pattern and 0.0733 [m] (22%) when adhering to the Piecewise Continuous navigation path, both absolute error values meet the desired Class 3 maximum error limit.

The second noticeable item is the importance of the terrain’s spatial complexity when selecting one navigation strategy over another. The simple downward-slope feature of our testing terrain reduces the need for spatially diverse paths, where as a test area more analogous to the one shown in Figure 8 typically demands more flexibility in changing navigation directions given the increased presence of hills and valleys.

Fig. 10. Average reconstruction error annotated with maximum error requirement for Class 3 map for each navigation pattern.

Fig. 11. The executed rover paths and an indication of the spacial distribution of the visual landmark for (a) the lawnmower surveying strategy and (b) the piecewise continuous surveying strategy.

Fig. 12. Terrain reconstructions and reconstruction errors using data from (a-b) the lawnmower surveying strategy and (c-d) the piecewise continuous surveying strategy.
VI. CONCLUSIONS

In this paper, we have discussed a methodology for terrain reconstruction of glacier environments based on observed phenomena during a robot traverse. The principle take-away from our work highlights the significance of augmenting intelligent navigation schemes with environmentally-relevant sensing capabilities to comply with desired scientific objectives. Future work will involve applying our approach to observing alternative phenomena, i.e., soil moisture, as well as deploying multiple agents in this field. These field campaigns range from accurately monitoring chemical plumes to providing timely spatial characterization of radiation distribution across an area. We believe that by coupling robotics with science-based objectives such as these, major life-preserving opportunities could continue to be addressed by the robotics community.

ACKNOWLEDGMENTS

This work was supported by the National Aeronautics and Space Administration under the Earth Science and Technology Office, Applied Information Systems Technology Program. The authors would also like to express their gratitude to Dr. Matt Heavner, Research Scientist, Los Alamos National Lab (formerly Associate Professor of Physics, University of Alaska Southeast), for providing logistical support and his expertise in glacial field work.

REFERENCES


Stephen Williams School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: swilliams8@gatech.edu.

Lonnie T. Parker School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: lonnie@gatech.edu.

Ayanna M. Howard School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: ayanna.howard@ece.gatech.edu.