FLAG: Feature-based Localization between Air and Ground

Xipeng Wang, Steve Vozar and Edwin Olson

***Abstract*— In GPS-denied environments, robot systems typically revert to navigating with dead-reckoning and relative mapping, accumulating error in their global pose estimate. In this paper, we propose Feature-based Localization between Air and Ground (FLAG), a method for computing global position updates by matching features observed from ground to features in an aerial image. Our method uses stable, descriptorless features associated with vertical structure in the environment around a ground robot in previously unmapped areas, referencing only overhead imagery, without GPS. Multiple-hypothesis data association with a particle filter enables efficient recovery from data association error and odometry uncertainty. We implement a stereo system to demonstrate our vertical feature based global positioning approach in both indoor and outdoor scenarios, and show comparable performance to laser-scan-matching results in both environments.**

I. Introduction

Knowing the *global* position of a robot is critical for coordinating the actions of multiple agents, creating user interfaces for operators, and creating data products that can be more easily shared across agents or over time. Global Positioning System (GPS) sensors are an obvious solution to this problem, but GPS does not work well indoors or in urban canyons, and it can be easily jammed. In some applications, a map can be built ahead of time. That map might either define a canonical global coordinate frame for all the agents that operate within it, or the map may be geo-registered. Once such a map is built, a robot can then obtain global position information by localizing with respect to that map.

For such a map to exist, a robot must have previously operated in that location. That is an unacceptable requirement for many real-world applications such as search and rescue in which robots may be called upon to operate in a novel environment. If GPS is unavailable, how can the robot limit the accumulation of global positioning error that results from operating without a means of global localization?

In this paper, we propose and demonstrate FLAG: Feature-based Localization between Air and Ground, an approach that allows a ground robot to perform global positioning by recognizing landmarks from aerial imagery (see Fig. 1). In particular, we use a stereo camera pair on the robot to detect large vertical features (*e.g.* trees and corners of buildings) in the environment.

In some cases, the vertical features can be directly seen in aerial imagery, but in others, the presence of a feature can only be inferred: the eaves of a roof often obscure the exact location of the corner of the building, for example. In this paper, we hand-label features in the aerial image.

The authors are with the Computer Science and Engineering Department, University of Michigan, Ann Arbor, MI, 48109, USA. {xipengw, svozar, ebolson}@umich.edu

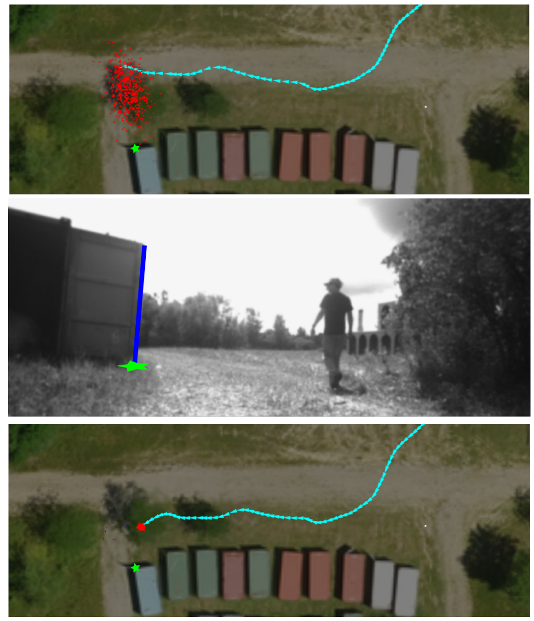

Fig. 1. *Top*: A robot path (cyan), with uncertainty of current robot position indicated by red particles. The green star represents a known global location of a vertical edge feature in the environment. *Middle*: Single image from a stereo camera pair showing the location of a vertical edge feature in the camera frame. *Bottom*: The corrected robot path and updated uncertainty after detection of the vertical edge feature.

The significant challenge is that when observed from above, landmarks may look completely different than when viewed from the ground. The lack of useful appearance data makes data association very difficult. We describe a robust multi-hypothesis data association system based on particle filtering as a way of mitigating this challenge.

Unlike other stereo localization methods based on interest point detectors, FLAG uses descriptorless features that are robust to temporal variations in visual appearance. We describe a stereo processing pipeline that is well-suited to extracting large vertical features. This processing pipeline is one approach to feature extraction for a stereo camera. Feature extraction methods for other sensors such as monocular cameras or LIDAR could be used in the FLAG method, but further discussion is beyond the scope of this paper.

FLAG is a method for localization, not simultaneous localization and mapping (SLAM), and is intended as a replacement or supplement for GPS in areas with poor or nonexistent GPS coverage. The FLAG approach performs occasional global position updates directly in a geo-registered

TABLE I

COMPARISON OF SEVERAL RELATED LOCALIZATION METHODOLOGIES. WHILE MANY VISION-BASED METHODS MEASURE RELATIVE MOTION, FLAG IS A VISUAL METHOD USED FOR GLOBAL REGISTRATION.

System	Feature tracking	Prior map	Global registration
VO	Y	N	N
V-SLAM	Y	N	N
GPS	N	N	Y
FLAG	N	Y	Y

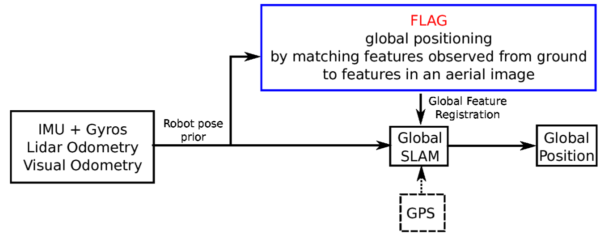

Fig. 2. Example robot localization system architecture showing the role of FLAG as a replacement or augmentation for GPS.

overhead image, and relies upon open-loop odometry or a SLAM framework to maintain local position estimates between global position fixes. Tracking algorithms such as visual odometry [1][2][3], and optimization algorithms such as factor-graph SLAM[4] could be complementary to our global registration approach. A comparison of FLAG with several common localization approaches is shown in Table I, and an overview of how FLAG fits into a typical robot localization architecture is shown in Fig. 2.

The contributions of this paper include:

- An approach to globally localize a ground robot based on featureless descriptors that can be identified in both 2D overhead maps of an area, and from ground-based sensing such as stereo imagery.
- A method to quickly find vertical edge features from stereo imagery of dynamic environments.
- An evaluation of our localization system on both outdoor and indoor real-world datasets demonstrating comparable performance to LIDAR-based localization methods in both environments.

II. RELATED WORK

Our proposed method uses a two-step approach to define and localize to landmarks in the world, using both overhead (orthographic) imagery, and 3D point clouds generated from a stereo camera. In this section, we briefly summarize prior work related to both of these steps: first discussing work in the area of localizing a robot based on satellite imagery, and then prior work related to robot localization based on stereo image features.

A. Orthographic Imagery-Based Localization

Traditionally, mobile robots have relied on offline processing to produce detailed prior maps, and perform localization by matching observed features with landmarks in the map. Recently, some research has focused on localizing a robot using only a satellite image instead of a detailed prior map.

Viswanathan et al. [5] focus on visual matching. In their work, the images collected with a wide-view stereo camera on a robot are warped to obtain a birds-eye view of the ground, and this view is compared to a grid of satellite locations using whole-image descriptors. However, objects in the environment that violate the implicit flat-ground assumption in the warping process, such as buildings, will not match well between the maps. Senlet and Elgammal [6] similarly use satellite images of road maps for localization. However, they do not follow the single-plane assumption and instead build a 3D point cloud of the scene as seen by the stereo camera, and then project these points into a top-view image. However, this method is highly reliant on accurate stereo dense matching. Furthermore, both approaches try to match the stereo view with a satellite image based on image features, which can fail if the visual appearance of the area differs from the satellite map due to different lighting conditions or changes in appearance. In our paper, we directly match edge features to corners in the orthographic image, which is less prone to error due to temporal variations in visual appearance.

B. Features for Stereo Localization

Much of the previous work in stereo localization and mapping has either relied on artificial landmarks [7] in a controlled environment or used algorithms based on local image features such as Scale-Invariant Feature Transform (SIFT) [8] or Speeded Up Robust Features (SURF) [9]. While these local image feature-based methods can achieve accurate localization, they come at a high computational cost [10]. Additionally, the reliability of these features is reduced in the presence of significant light and scale variations that can occur between different viewpoints. Since these systems rely on feature matching across images taken at different poses to estimate robot motion and perform localization, system performance will suffer in the presence of appearance variations [11].

Using edge features for localization and mapping has previously been investigated by several authors. Tomono [12] proposed an edge-based 3D SLAM system that extracts image edges using a Canny detector and finds correspondences between line features in the stereo image pairs. This method uses normalized correlation and dynamic programming (DP) matching [13] that requires significant computational resources. While our paper also uses edges as features, we focus solely on vertical edges, and use a 1D Binary Robust Independent Elementary Feature (BRIEF) descriptor [14] and marriage matching to match edges between stereo images in an efficient manner (see Sec. IV).

These previous systems have only been tested in controlled indoor environments, the structure of which enables the use of simple nearest-neighbor data association. In dynamic outdoor scenes, the motion of grass, trees, and shadows

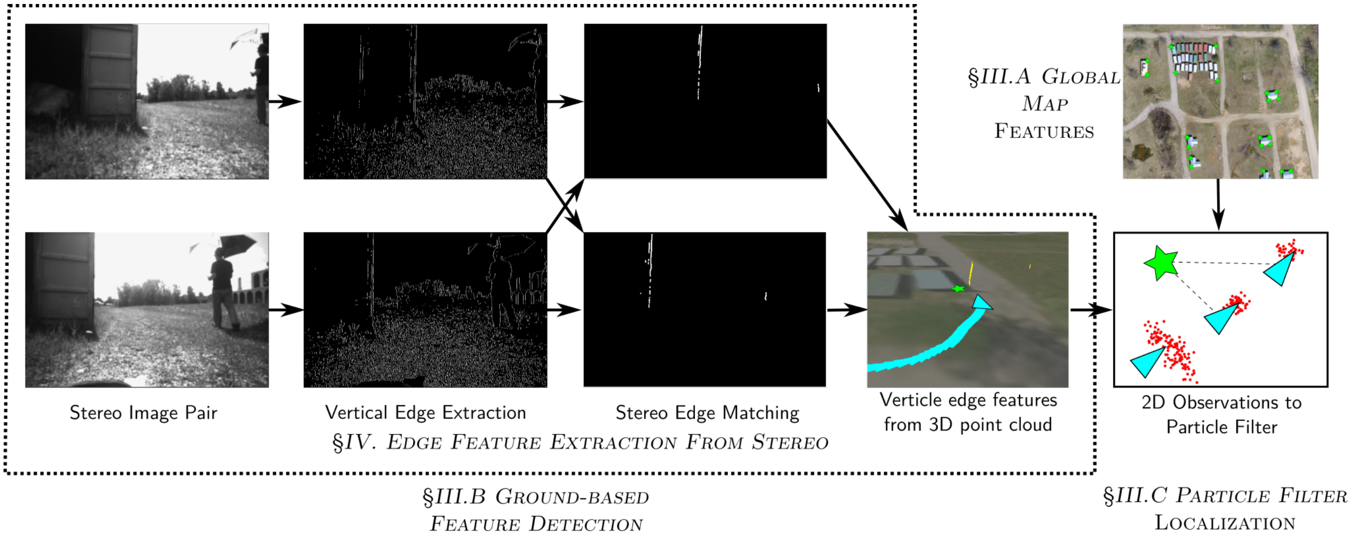

Fig. 3. Schematic overview illustrating the FLAG approach. Ground-based features are compared to the locations of known landmarks, and then used to update the particle filter localization solution. In a particular stereo processing pipeline described in IV, vertical edges detected in both stereo images are extracted, matched and projected into 3D.

results in matching noise and increases the uncertainty of data association, and single hypothesis data association is not feasible. In our paper, a particle filter is used to perform multiple-hypothesis data association to help achieve robustness.

III. Approach

FLAG leverages vertical structures, such as edges of buildings, as robust descriptorless features for localization by matching them to pre-determined locations of vertical structures in an orthographic overhead map. In this paper, we use a stereo camera to detect these edges and project their positions into the 3D world space. Edges are then matched to the known features in the global prior map. A solution for the robot pose is found using a particle filter-based localization algorithm. Figure 3 shows an overview of our approach.

A. Global Map Features

Because point-based image features can change in appearance over the course of a day or season, we seek to use global features that are robust to changes in visual appearance. We also wish to choose features that will persist over long periods of time, so that orthographic imagery need not be recent, and feature maps can be reused for multiple missions. Under normal circumstances, building structures, tree trunks, lamp posts, and other permanent objects are unlikely to change and should be consistently identifiable despite temporal changes in appearance (*e.g.* shadows, snow, dirt), and are thus good feature candidates. One reason LIDAR-based systems are able to achieve high-quality, reusable maps is that they often extract such permanent structures.

While it may be difficult to detect entire buildings with a stereo camera (due to limitations of 3D projection range and field of view), sections of buildings, particularly building edges, are easy to detect. Edges of buildings also have a characteristic appearance in both orthographic overhead views of a scene (corners), and first-person views from the robot (vertical lines). Trees and lamp posts appear as circles from above, and vertical lines from the ground. Therefore, we choose vertical edge features of permanent structures as our ideal global map feature. With this approach, we enable a robot to localize to objects in a scene that have not previously been observed from the robot's perspective.

Given an overhead representation of the robot's environment, such as a satellite image or orthographic aerial photograph (or even a topographic map showing building locations), we can either automatically or manually determine the locations of such vertical edges features in advance of a robot's mission. Note that manual annotation of a prior map can be done with a few clicks of a mouse, and is not necessarily an onerous task – particularly compared to the detailed site analysis that mission preparation may already require. Manual annotation also facilitates feature identification from a wide range of orthographic imagery that could be low-resolution or obscured. Thus, we focus here on maps with hand-annotated features.

B. Ground-based Feature Detection

Vertical edge features are extracted from 3D point clouds of the environment observed by the ground robot. In Section IV, we discuss in detail a method for vertical feature extraction from a stereo camera, however the FLAG method could work with LIDAR, stereo cameras, or any other sensor capable of producing a 3D point cloud.

Given a 3D point cloud of the robot's environment, we project the points into the 2D ground plane, binning points in a 5 cm grid. Long vertical edges will bin to the same grid position, and a simple threshold on number of points per grid position is used to define a vertical edge. The precise threshold depends on the type and resolution of the sensor used to generate the point cloud. For our stereo setup, a minumum threshold of 50 pixels can reliably detect vertical edges at least 1 m high at a viewing distance of 8 m.

C. Particle Filter-Based Localization Solution

In our system, maximum likelihood data association is likely to fail for multiple reasons. First, our global features are descriptorless, which increases the probability of a match, but hinders our ability to reject false matches. Second, outdoor environments are unstructured and prone to noisy feature observations, resulting in false positive feature detections that have no correct association. Third, there may be long periods of dead-reckoned motion without feature observation, due to lack of vertical structures, leading to high uncertainty in robot pose. With a single hypothesis data association method, a false positive data association can corrupt a localization solution beyond repair. We thus rule out linearized Gaussian techniques, such as EKF, UKF, and factor graphs for this process.

Instead, we choose to use a particle filter [15] to account for multiple hypotheses, as shown in Algorithm 1. We represent the posterior of the robot pose distribution by K particles $\chi = \{x_t^{[1]}, \cdots, x_t^{[K]}\}$, and perform data association on a *per-particle* basis using nearest neighbor search. The motion command is denoted u_t, observations are z_t, and importance weights are w_t at time t. This allows us to express uncertainty in the presence of ambiguous data associations, hedging our estimate until more evidence is encountered, and enables recovery from spurious feature matches. For example, our system occasionally detects doorways as vertical features (which are not visible in a satellite image) and misassociates them to annotated building edge features. Having multiple particles allows the system to recover from these situations.

Algorithm 1 Particle Filter for Multi-hypotheses Data Association

procedure INITIALIZATION(χ_0)
    Initialize χ_0 with K particles.

procedure LOCALIZATION(χ_{t-1}, u_t, z_t)
    $\bar{\chi}_t = \chi_t = \emptyset$
    for i = 1 to K **do**
        Sample $x_t^{[i]} \sim p(x_t | u_t, x_{t-1}^{[i]})$
        $w_t^{[i]} = p(z_t | x_t^{[i]})$
        Add $x_t^{[i]}$ into $\bar{\chi}_t$
    for i = 1 to K **do**
        Draw $j \in \{1, \cdots, K\}$ with probability $\propto w_t^{[j]}$
        Add $x_t^{[j]}$ into χ_t
    return χ_t

IV. Edge Feature Extraction from Stereo

The approach outlined in Section III is agnostic to the source of the point cloud. In this section, we propose a novel and efficient method to detect vertical edge features

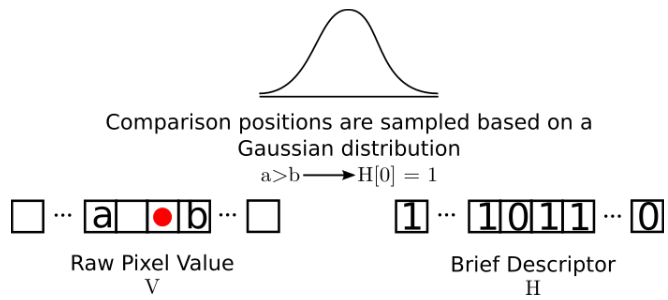

Fig. 4. 1D BRIEF Descriptor. Each edge point is associated with a local 64-bit BRIEF descriptor evaluated using 15 surrounding pixels in the rectified image. For each point shown as a red dot in one rectified image, the Hamming distance of its descriptor is compared to the descriptor of every other filtered point along the epipolar constraint in the other rectified image. Points are matched to one another if they are mutually most compatible.

from a ground-based stereo camera. This stereo method is the basis for the implementation and evaluation described in Section V.

A. Vertical Edge Detection

Given a stereo image pair, we use a Sobel filter to detect edges in each image. Because we are only interested in detecting vertical edges, we only calculate the gradients along the horizontal direction. Assuming the robot is level, we use the filter

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{1}$$

Eight-neighbors connection clustering [16] is used to find pixels belonging to segments of vertical lines. For each edge pixel, a 64-bit BRIEF descriptor [14] is generated to describe the point. We first extract a 15-pixel patch vector around the edge points, as shown in Fig. 4. Then we sample 64 pairs of comparison locations based on a Gaussian distribution. In each pair, if the pixel value of the first location is greater than that of second location, the corresponding bit in the BRIEF descriptor is set to 1. Otherwise, it is set to 0.

We can leverage the horizontal epipolar constraint [17] to quickly evaluate potential feature matches between the stereo images. For every point in the right image, we declare all filtered points in the left image that lie on the corresponding epipolar line as candidate matches. This process is repeated for all points in the left image. The Hamming distance [18] between all candidate pairs' BRIEF descriptors denotes their compatibility for matching, and a marriage procedure is used to match points. This approach is conservative in that it only considers pairs that are *most* compatible with each other. Making the system less sensitive reduces the density of detected landmarks, and can make data association easier. Some pairs that do not lie on a vertical edge may be matched, but they will be culled in the next step.

Note that simple BRIEF descriptors are sufficient since the matching process is facilitated by spatially constrained

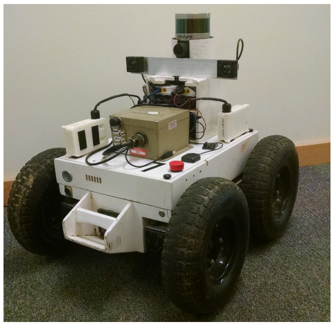

Fig. 5. April MAGIC2 robot platform used for experimental tests. The 30 cm baseline stereo camera pair is located beneath the standard monocular camera and LIDAR sensor.

viewpoints and synchronized images. Additionally, while a BRIEF descriptor is used to match vertical edge pixels in corresponding stereo images, the resulting features in the ground plane used for localization are descriptorless.

B. 3D projection and clustering

A point cloud consisting of all the matched edge points from the stereo image pair is generated based on least square triangulation [17].

The marriage matching procedure described above leaves some outliers in the 3D point cloud that do not belong to any vertical edge. These are removed using density-based spatial clustering (DBSCAN) [19]. Given a set of points in space, DBSCAN counts the number of nearby neighbors of each point, grouping points that are closely packed together, and rejecting outlier points that lie in low-density regions. For our hardware setup (see Sec. V-A), using DBSCAN to remove points with fewer than 5 neighbors within a radius of 10 cm has produced good results.

V. Experimental Evaluation

We evaluated our method with outdoor and indoor real-world datasets in order to demonstrate its performance and compare it to state-of-the-art laser-scanmatch-based SLAM localization solutions. The following subsections detail the setup and results of the evaluation.

A. Robot Platform

Experiments were performed using the MAGIC2 robot platform developed by the April Laboratory shown in Fig. 5. Relevant sensors onboard the robot include wheel encoders and a fiber optic gyro (KVH DSP-1715) for robot odometry, a stereo camera pair (2× Point Grey Chameleon) with a 30 cm stereo baseline, and a 3D LIDAR sensor (Velodyne VLP-16), for performance benchmarking purposes. The stereo cameras were calibrated using AprilCal [20], and image capture was synchronized via physical triggering.

B. Outdoor Evaluation

An outdoor dataset was collected at the Camp Atterbury training facility, near Edinburgh, Indiana. An orthographic map of the 9-acre (3.6-hectare) test area was generated using

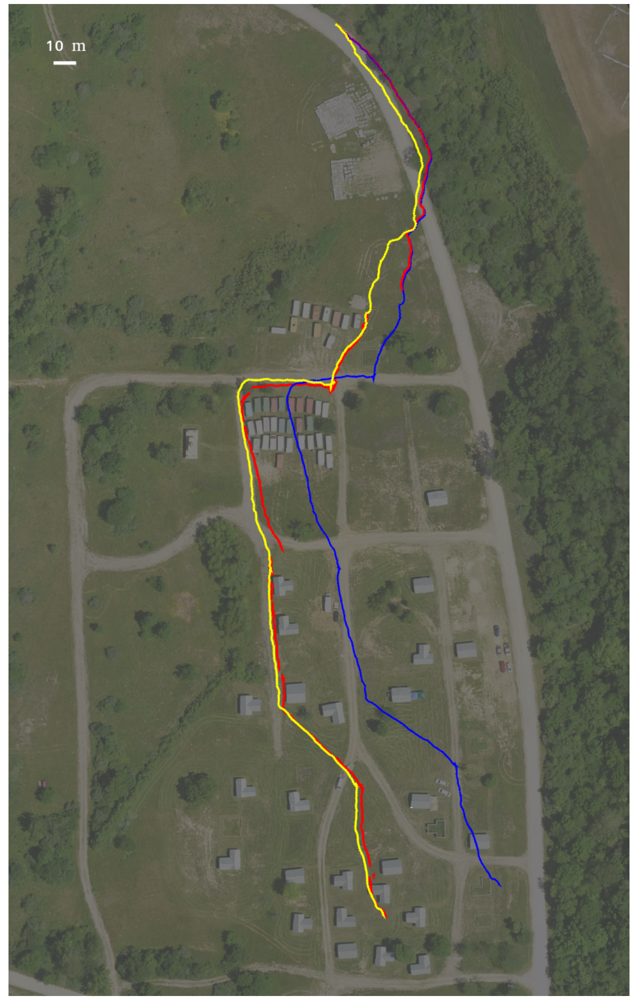

Fig. 6. Comparison of dead-reckoning (blue), laser-scan-based localization (yellow), and FLAG (red) methods over a 0.5 km outdoor course through fields and areas with structure. Note that a vertical edge feature is not detected for the first section of the course, so the FLAG solution is identical to the dead-reckoning solution. Because it does not contain global features, the laser-scan based localization was manually-aligned to the map in a post-processing step, making this the best-case LIDAR solution. Discontinuities in the stereo localization solution are due to feature observations after long periods of dead-reckoning and are shown purposely to demonstrate the recovery abilities of the particle filter.

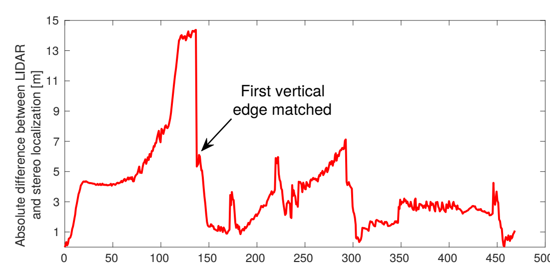

Fig. 7. Absolute difference between robot pose as determined by the benchmark 2D laser-scan-based SLAM algorithm and FLAG over the course of a 0.5 km path. The difference increases with increasing time during which the only updates are from odometry, and sharply decreases when the stereo camera observes a global feature. The difference is consistently low in dense areas of permanent structure. Once a vertical edge factor is found (at 140 m), the difference does not exceed 8 m.

aerial imagery captured from an unmanned aerial vehicle flown over the test site prior to testing. The orthographic overhead imagery had a resolution of 0.27 m per pixel, similar to commercially and freely available satellite imagery. Vertical edge landmark positions were hand-annotated into the prior map at select locations of building corners. Note that an exhaustive labeling of landmarks is not required, as the system uses dead-reckoning, relying on sparse vertical edge features for infrequent global fixes.

The robot was manually driven from the north end of the test area to the south end, on roads and grass through areas including clusters of buildings over a course of about 0.5 km. Figure 6 shows the path of the robot as determined by odometry-based dead-reckoning (blue), a factor graph-based SLAM solution with 2D laser-scan-matching factors [21] (yellow), and our proposed FLAG method (red). Note that the posterior LIDAR-derived solution was post-processed to align it to several surveyed GPS points in order to improve its utility as a "gold standard" (near-ground truth) result.

In contrast, our stereo localization was run in real-time with no manual alignment other than a coarse initial pose estimate. This is evident in the first portion of the path, where a misalignment in initial heading causes the dead-reckoning estimate to deviate from the true path. Our stereo-matching method is able to recover from this initial heading error upon observation of the first vertical feature, despite tens of meters of featureless open-loop driving. We purposely show discontinuities in the stereo-matching solution to illustrate localization recovery.

Figure 6 shows the path estimated by FLAG is similar to that determined by the 2D laser-scan-matching algorithm when vertical edge features are present, and is significantly better than dead-reckoning alone over long distances.

The absolute difference between the laser-scan-based SLAM solution and the proposed stereo edge feature-matching scenario over the robot path is shown in Fig. 7. Once the initial global feature is detected, the absolute difference is never more than 8 m, despite long stretches of featureless driving, and has a mean difference of 2.8 m.

C. Indoor Evaluation

To demonstrate FLAG's versatility, an indoor dataset was collected in an office building setting. Instead of searching for vertical edge features in a satellite image, vertical edge landmark positions were determined and hand-labeled from a simple floor plan map.

The robot was driven in a 160 m loop through the halls of the BBB Building at the University of Michigan. Figure 8 indicates that localization with our stereo vertical edge matching approach shows clear improvements over the dead-reckoning localization solution, resulting similar performance to the baseline scan-matching approach. Figure 9 shows the absolute difference between the baseline and FLAG approaches, indicating that our localization method closely tracks the localization solution determined from the full scan-matching based factor graph solution.

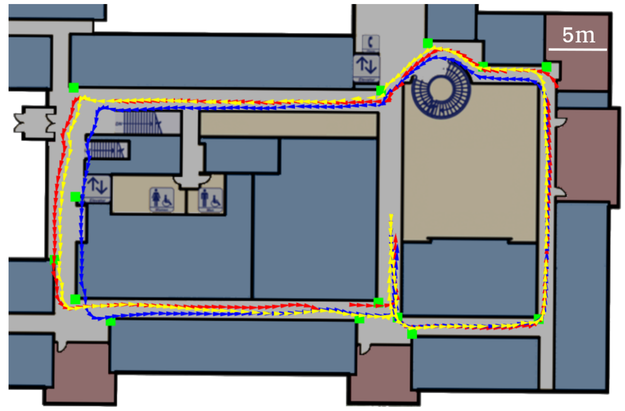

Fig. 8. Comparison of dead-reckoning (blue), laser-scan-based localization (yellow), and our FLAG (red) methods over a 160 m counter-clockwise looped path in an indoor environment. Green squares are the global features, hand-annotated into a simple floor plan map of the building. Note that not every possible vertical edge needs to be in the global map. Imagery has a resolution of 0.10 m per pixel, but could have been much coarser.

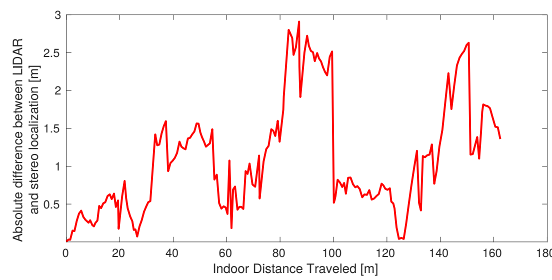

Fig. 9. Absolute difference between robot poses as determined by the benchmark 2D laser-scan-based SLAM algorithm and the FLAG method over a 160 m looped path in an indoor environment. As in the outdoor experiment, the difference increases with increasing time since the last global feature, and is consistently low in areas where many vertical edge features are present. The mean difference is 1.1 m, and never exceeds 3 m.

VI. Conclusion

In this paper, we proposed Feature-based Localization between Air and Ground (FLAG), a method that enables a ground robot to localize itself globally by identifying landmarks visible in overhead orthographic imagery. While landmarks may look very different to a robot than when observed from above, FLAG uses a stereo vision-processing pipeline to quickly identify large vertical features (such as trees, poles, and building corners) on the ground, which can then be matched to known or inferred structural reference points in an aerial image. A multi-hypothesis data association technique is used to mitigate the difficulty of matching features between two different viewpoints (air and ground). We have demonstrated that our system has comparable performance to LIDAR-based scan-matching localization in indoor and outdoor environments, showing its feasibility for real-time global localization without LIDAR or GPS.

Acknowledgements

This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA). The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the

official views or policies of the Department of Defense or the U.S. Government.

The authors thank Carl Kershaw and Rob Goeddel for insightful advice and valuable discussions. Additionally, Victoria Edwards, Justin Tesmer, and John Wang are acknowledged for their assistance with data collection.

REFERENCES

[1] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry for ground vehicle applications," *Journal of Field Robotics*, 2006.

[2] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in *Intelligent Robots and Systems (IROS), IEEE International Conference on*, 2008.

[3] K. Konolige, M. Agrawal, and J. Sola, "Large-scale visual odometry for rough terrain," in *Robotics research*, 2010.

[4] G. Grisetti, R. Kummerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based $slam$," *IEEE Intelligent Transportation Systems Magazine*, 2010.

[5] A. Viswanathan, B. R. Pires, and D. Huber, "Vision based robot localization by ground to satellite matching in GPS-denied situations," in *Intelligent Robots and Systems (IROS), IEEE International Conference on*, 2014.

[6] T. Senlet and A. Elgammal, "A framework for global vehicle localization using stereo images and satellite and road maps," in *Computer Vision Workshops (ICCV Workshops), IEEE International Conference on*, 2011.

[7] S. Takezawa, D. C. Herath, and G. Dissanayake, "SLAM in indoor environments with stereo vision," in *Intelligent Robots and Systems (IROS), IEEE International Conference on*, 2004.

[8] S. Se, D. Lowe, and J. Little, "Vision-based mobile robot localization and mapping using scale-invariant features," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2001.

[9] H. Badino, D. Huber, and T. Kanade, "Real-time topometric localization," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2012.

[10] O. Miksik and K. Mikolajczyk, "Evaluation of local detectors and descriptors for fast feature matching," in *Pattern Recognition (ICPR), 21st International Conference on*, 2012.

[11] W. Churchill and P. Newman, "Practice makes perfect? Managing and leveraging visual experiences for lifelong navigation," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2012.

[12] M. Tomono, "Robust 3d SLAM with a stereo camera based on an edge-point ICP algorithm," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2009.

[13] Y. Ohta and T. Kanade, "Stereo by intra-and inter-scanline search using dynamic programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1985.

[14] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in *European Conference on Computer Vision (ECCV)*, 2010.

[15] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte carlo localization for mobile robots," in *Robotics and Automation (ICRA), IEEE International Conference on*, 1999.

[16] A. Datta and S. K. Parui, "A robust parallel thinning algorithm for binary images," *Pattern Recognition*, 1994.

[17] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2003.

[18] R. W. Hamming, "Error detecting and error correcting codes," *Bell Labs Technical Journal*, 1950.

[19] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Knowledge Discovery and Data Mining*, 1996.

[20] A. Richardson, J. Strom, and E. Olson, "AprilCal: Assisted and repeatable camera calibration," in *Intelligent Robots and Systems (IROS), IEEE International Conference on*, 2013.

[21] E. Olson, "Real-time correlative scan matching," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2009.