

LiDAR/Camera Point & Pixel Aided Autonomous Navigation

Autonomous localization is the process of determining a platform's location without the use of any prior information external to the platform, using only what is available from the environment perceived through sensors. In this article, we describe a technique using a collaborative swarm of UAV's with the goal of assisted navigation in GPS-denied environments using vision-aided Point and Pixel data. We consider two teams of UAVs, operating at different times, and operating collaboratively. The first team, equipped with LiDAR, fly over the unknown area generating a digital terrain map (DTM) or digital surface map (DSM). The second team, equipped with low-end passive-vision sensors, fly over the same area at a later time, using the information generated by the first team, for landmark navigation. The second team operates without GPS using the terrain as a source of localization. To enable this scenario, we present an algorithm for terrain-aided navigation based on Point & Pixel matching to provide environment perception.

Introduction

Autonomous localization is the process of determining a platform's location without the use of any information external to the platform using only

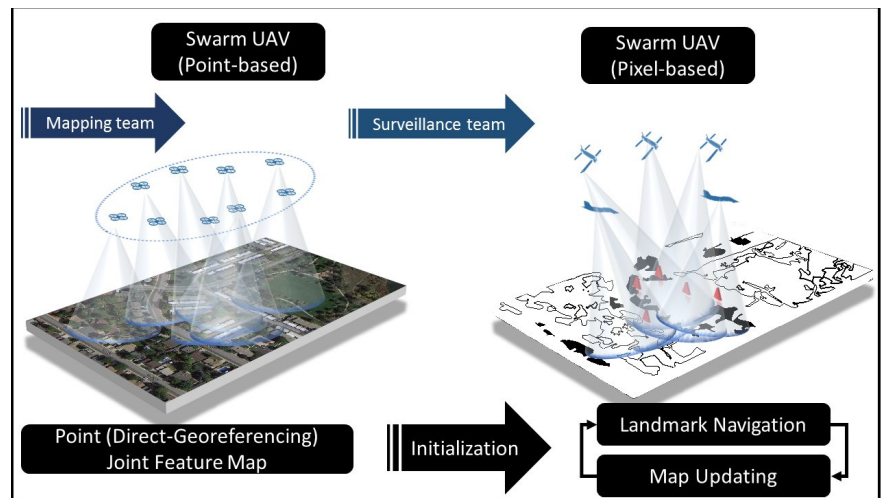


Figure 1: Collaborative navigation vision

what is available from the environment perceived through sensors. In this paper, we describe a technique based on the collaboration of UAV's with the goal of assisted navigation in GPS-denied environments. The capability can be extended to swarm UAV path/mission planning, navigation, and localization and as a mission planning and training tool. We consider two teams of UAV's working at different times in collaborative navigation. First, a small swarm of UAV's, equipped with a low-cost (camera) and high-end sensors (LiDAR), fly over a known/unknown area, when GPS is available. This swarm creates a landmark-map. Later, a larger team

of UAVs, equipped only with low-end electro-optical (EO) sensors (cameras) fly over the same area using the map generated by the first team for landmark navigation. A sketch of our collaborative navigation system based on a swarm of UAVs is shown in **Figure 1**.

In this design, a Mapping team of UAVs flying in a hostile area is responsible for mapping and providing a reference terrain model from which a set of landmark can be identified. Later, a Surveillance team will go to the same area and use the reference terrain model for navigation in GPS-denied conditions. Collaboration is based on perception of the environment with

BY SHAHRAM MOAFIPOOR, LYDIA BOCK, JEFFREY A. FAYMAN

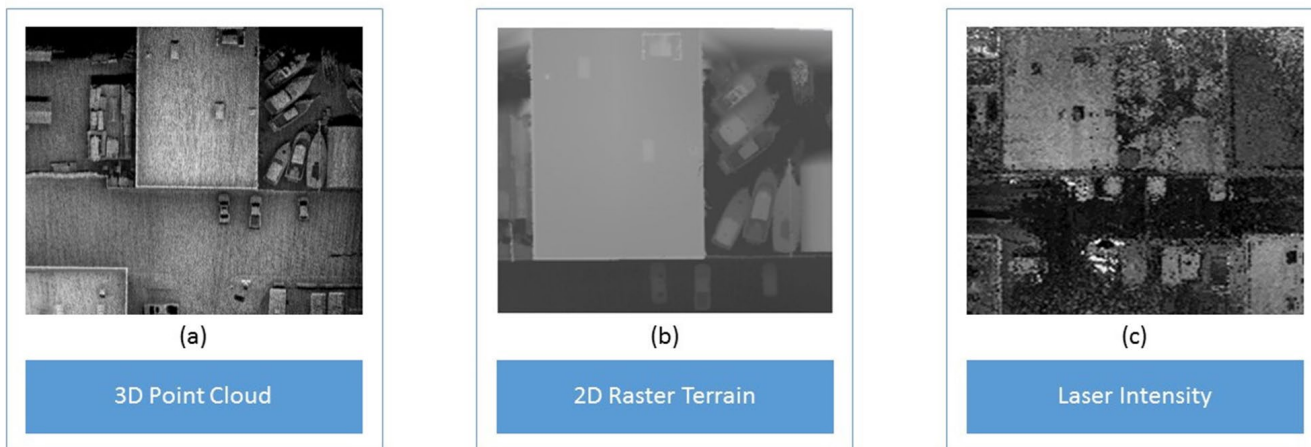


Figure 2: Point & Pixel feature presentation

vision sensory data captured by one swarm when GPS is available and used later by other platforms operating in the same area under GPS-denied conditions. The focus in this paper is on the use of LiDAR for perception by the first swarm of UAV's and EO cameras for perception by the second swarm of UAV's. We present an algorithm for vision-aided navigation based on Point (LiDAR) and Pixel (EO camera) matching to provide environment perception. We refer to this capability as Point & Pixel. In this Point & Pixel matching algorithm, we match camera images (pixels) to point cloud reference 2D range images provided by the LiDAR.

From perspective of data acquisition and sensor operations, there are major differences between the LiDAR and EO cameras. LiDAR is a high-power active sensor enabling direct-georeferencing and encoding 3D pointwise sampling (point clouds) and provides terrain models even in environments containing dense foliage. EO cameras, on the other hand, are low-energy passive sensors that can cover a full area in a snapshot

within multispectral bands and high radiometric pixel information. However, the captured images require post-processing for calibration and encoding. The main advantage of LiDAR-based mobile-mapping is that there is no need for Ground Control Points (GCP) for real-time direct geo-referencing of the point clouds. This is particularly useful when the system operates in hostile environments. Another advantage is availability of multiple laser returns, at a high frequency, enabling penetration through canopy and forest. The sensor can separate off-ground objects from on-ground objects, providing DTM/DSM in areas of dense vegetation or tree canopy. The limitations of LiDAR are related to the classification and identification of objects from point clouds, point density variation with scan angle and the topography, and noise level caused by multi-return beams.

UAV Point & Pixel Mapping

Once the first group of UAV's finishes flying its mission, the map generated will be used as a reference terrain model to aid

navigation of the second group of UAV's operating in GPS-denied conditions. In a navigation system based only on image aiding, the main challenges become real-time image-processing and integrating image features with the inertial system to provide filter updates during GPS-denied periods. In this article, the vision-aided inertial system is supported by the map features. The key in this algorithm is to extract landmark (key-points) from point-based map data and find corresponding features in image-based data. To illustrate the steps used in the Point & Pixel matching algorithm, sample Point & Pixel data, captured by Velodyne VLP-16 laser scanner is shown in **Figure 2**. In **Figure 2a**, a geo-referenced point-based image in which point cloud data are spaced with an irregular point density is shown. To use the geo-referenced point cloud as a terrain aiding map, we utilize several techniques, including (1) Interpolating/extrapolating the irregular point clouds to a regular grid and treating it as an 2D range image as shown in **Figure 2b** in which case each point carries a depth value and the laser intensity; (2) Delaunay

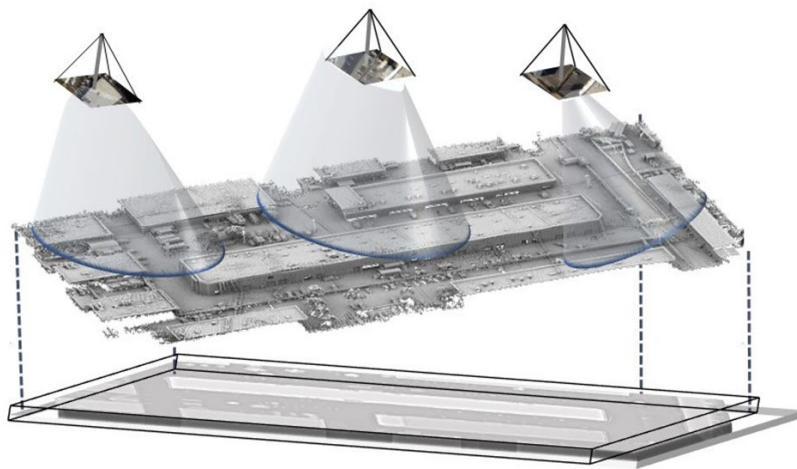


Figure 3: Image matching concept on a raster 2D DTM image

triangulation for terrain modeling (vector based); and (3) use the laser point density as a map **Figure 2c**.

As can be seen, the laser intensity signature is based on the reflectivity of the surface type and cannot be used to represent feature attributes that can be used for feature matching. One advantage of representing the point clouds in the raster form is that they can be treated as images and a variety of Photogrammetry/Computer Vision (CV) image-processing tools and techniques can be applied. Thus, we chose to represent the point-data in a raster 2D image.

Point & Pixel Matching Algorithm

In order to find a match between the camera image, and a raster DTM image, as illustrated in **Figure 3**, we first tried applying a standard feature matching algorithm to the Point & Pixel images. However, the feature-based algorithm failed due to the mixed-scale and radiometric resolution across the DTM

raster image. The approach we used for matching was implemented in two stages; coarse and fine. In the coarse stage, we narrowed the search space in the image with respect to the raster 2D image and estimated the proper image scale and orientation of the pixel-image to the reference terrain image. Once the approximate location of the pixel-image was determined, in the fine stage, a feature matching algorithm was used to find the common features between two

images. These features are used later in a space resection algorithm to estimate the Exterior Orientation Parameters (EOP) of the pixel-images in terms of position and attitude of the projection center. These EOPs are then used as auxiliary data to update the inertial navigation filters in the absence of GPS data.

The algorithm used in the coarse-stage is based on template matching. The template is provided by the pixel image, and the reference image is the DTM raster image. We found that standard template matching works only under ideal conditions. Firstly, the template image should be captured close to the nadir direction, as any over-tilt can cause a mismatch. Additionally, the two images should have a similar radiometric resolution. In our application, the pixel image radiometric resolution is higher than the DTM raster image as the radiometric quantization of a raster DTM image is a function of variations in altitude, while an RGB camera quantization level is usually more than 8-bits. In addition, the spatial resolution of the two images should be similar. The spatial resolution of the raster DTM image is a function of the point density, which directly related

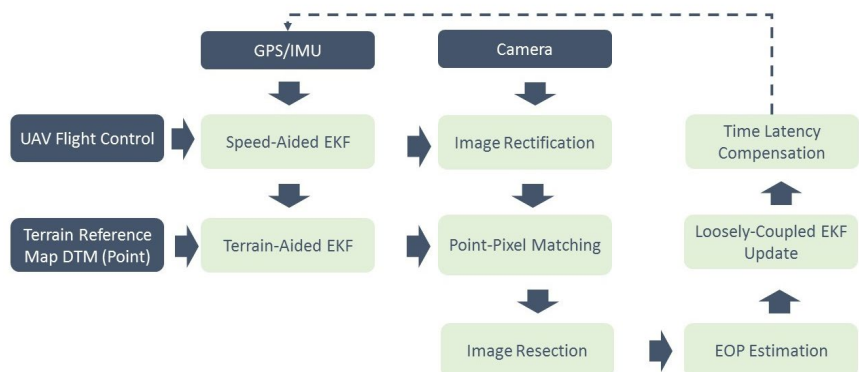


Figure 4: Architecture of the Point & Pixel matching algorithm

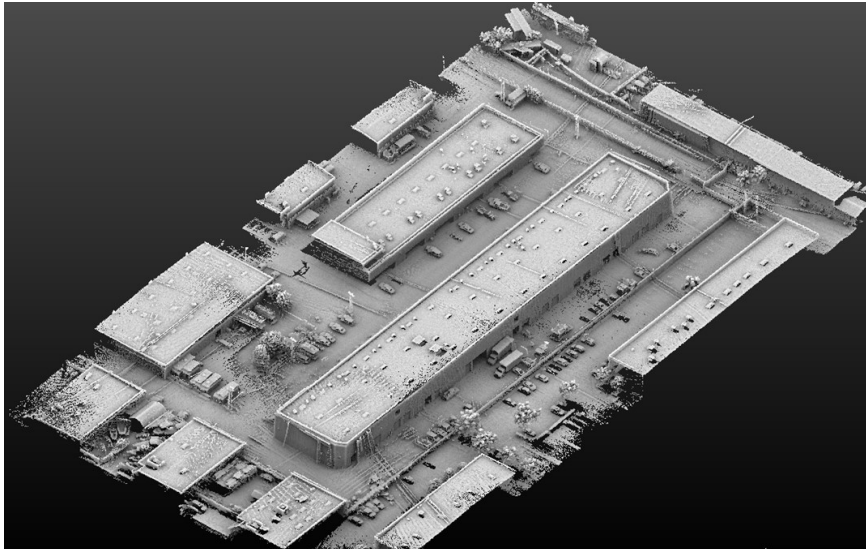


Figure 5: Direct-georeferencing terrain mapping

to the flight altitude (AGL), frequency setting, and platform speed.

Thus, additional quantization and rectification must be applied to the pixel images before they can be used for template matching. After these additional transformations, a Rotation-Scale-Translation (RST) invariant template matching algorithm can be used to estimate the coarse-location of the pixel image with respect to the DTM raster image. The Point & Pixel matching algorithm is illustrated in **Figure 4**. In this architecture, there are four sources of sensory data: (1) terrain reference map DTM generated by the first team of the UAVs, (2) UAV flight control of the second UAV team, (3) the GPS/IMU of the second UAV team with assumption of GPS denied, and (4) the camera mounted on the second team of UAVs.

Experimental Results

To simulate our collaborative navigation application, we collected data from two separate teams of UAVs using Geodetics' Geo-MMS™ Tactical system.

The first was equipped with an autopilot providing inner loop attitude and velocity stabilization control, a Velodyne VLP-16 LiDAR sensor, a MEMS IMU, a radio modem and a dual-frequency RTK GPS sensor. The flight duration was approximately 20 minutes at an altitude of 40m AGL. The FOV of the laser

scanner was 120° with the frequency of 20Hz (1200 RPM). The second group of UAV's were instrumented with an autopilot, a GoPro HERO4 camera, a MEMS IMU and a radio modem. This group of UAVs were flown at 20m AGL. The first team of UAVs, equipped with Geo-MMS LiDAR, were flown over the test area and a DTM was generated. At this stage, the laser point clouds were geo-referenced with the accuracy of $\pm 5\text{cm}$ using RTK. **Figure 5** shows the generated map of the area.

Next, a UAV from the second team was flown over the same area at a higher altitude. The second team operates without GPS using the Point & Pixel algorithm described in this paper as a source of localization. As explained earlier, the first step was to make a raster 2D image of the DTM. **Figure 6** shows the raster image generated after a cubic interpolation/extrapolation from the irregular point cloud to a regular grid cell size of 0.1m. The reference image size was 2610 * 1480 pixels with 8-bit depth.

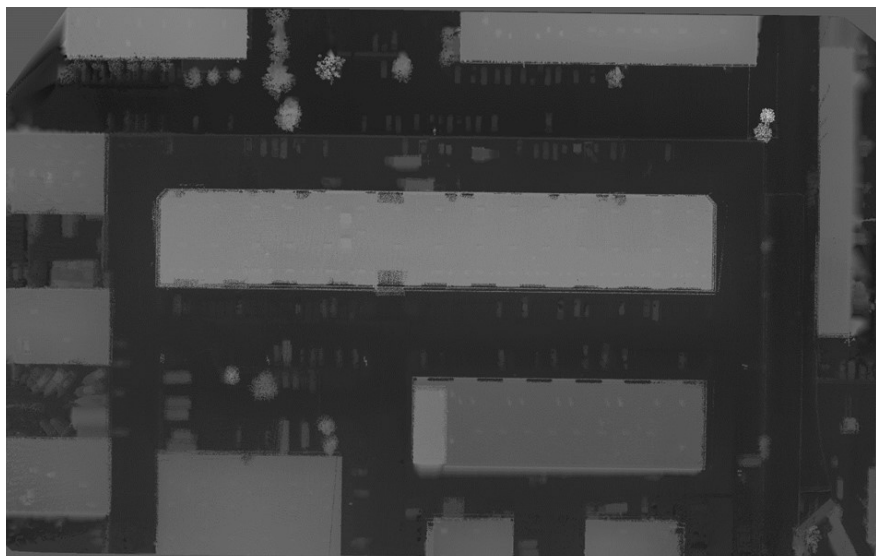


Figure 6: DTM raster image



Figure 7: Original image (a) and single image rectification and intensity normalization (b)

Next, the Point & Pixel matching algorithm was employed. For the purposes of vision-aided navigation, the camera was set to capture images at 1Hz, a sample rate allowing for filter updating in real-time. **Figure 7a** shows an example of the images fed to the Point & Pixel matching algorithm. The camera was installed with a tilt of approximately 30° with respect to the platform body frame. As previously mentioned, the template image should be captured close to the nadir direction, as any over-tilt can cause a mismatch, thus the 30° tilt must be compensated before it can be used for template matching. To accomplish this, an indirect rectification was performed using a central perspective transformation with a fixed focal length. This transformation was a projective transformation with the scale of focal length, as shown in **Figure 7b**. Additionally, intensity normalization was applied to the template image.

Once the image is rectified for tilt, it is used in an RST-invariant template matching algorithm. Due to the unknown scale difference between the template and DTM images, the first image was processed across the whole

range of scale following which the scale factor was narrowed down for faster template matching. **Figure 8** shows the normalized cross correlation of the matching between the rectified image, **Figure 7b**, and the DTM raster image. Once the maximum cross correlation is identified, feature extraction and matching is applied to the restricted DTM raster image and the scaled down image. For camera images, there are a variety of standard methods for feature point extraction and matching, including SIFT and SURF. However, for rasterized DTM image, these methods could not provide a consistent solution. The main reason is related to the low quantitation and resolution of these images, which are

restricted by the laser spacing in object space and quantization restricted by altitude range.

Despite significant research into image matching, matching points in a rasterized image remain a challenging problem. Thus, instead of point-based approach, we use feature-based matching based on region description. In the DTM raster image, a region is usually defined by point density, roughness, and elevation variations. In the pixel image, the same region is defined by shadow, texture, and scale. Thus, the region description was generated using a Maximum Stable External Regions (MSER) operator based on region density and region size attributes. Once this

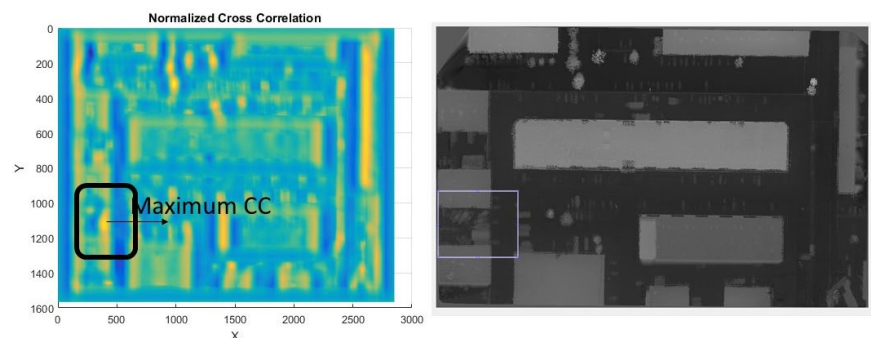


Figure 8: Normalized cross correlation template matching

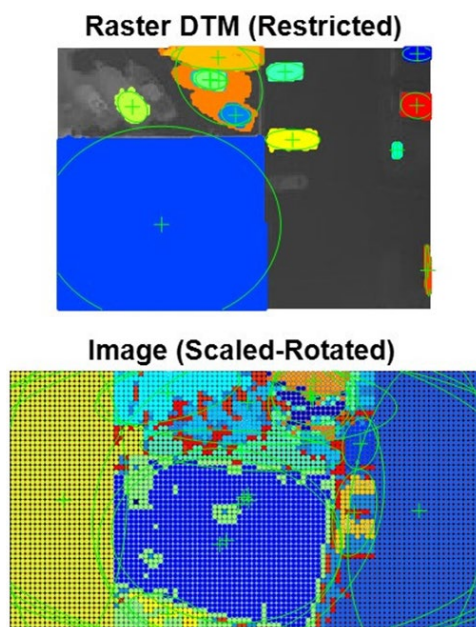


Figure 9: Feature extraction and matching sample case

operator was applied to the two images, a set of corresponding matching features was detected. **Figure 9** shows a sample of image feature matching based on the MSER matching region algorithm.

After removing outliers and detection of the strongest features, the matching features are used in the space resection algorithm, where the EOP of the captured images can be estimated using geo-referenced features in the raster DTM image. Applying this algorithm to a series of captured images along a flight strip can provide position updates that can be fed to the EKF in a loosely-coupled manner. The algorithm was tested on several sample images along the strip, and it was found that the algorithm is sensitive to the texture and the quantization level of the raster DTM image. RST advanced template matching found occurrences of the template regardless of their orientation and scale, but it could not match local

brightness between two Point & Pixel images even after normalizing the intensity of both. One possible approach is to add the camera pixel data to the laser point clouds. In this approach, the DTM raster image will contain the quantization level of the camera image, which can enhance template and feature matching. **Figure 10** shows the results of applying the position updates to the navigation solution based on image EOPs estimation to the collected data.

The “true” solution is shown in blue, while the free-inertial (GPS-denied) solution is shown in red. Note that the red drifts very quickly. Navigation updates from the Pixel-Point Matching algorithm are shown in yellow. This preliminary performance evaluation shows improved performance of the navigation system in GPS-denied conditions based on a collaborative of two UAV teams. A more comprehensive

performance analysis of the Point & Pixel Matching is currently under way for more complicated trajectories and environments, where the terrain signature texture is more homogenous.

Conclusion And Future Work

In this paper, we presented an algorithm for UAV navigation based on the collaboration of a swarm of UAVs that provide accurate terrain landmark mapping for use in terrain-aided navigation. We developed an algorithm to find matches between images captured by the camera and the point-based map data. The algorithm is called Point & Pixel Matching, which estimates the transformation from each captured image to the reference terrain. The algorithm was implemented in two steps and tested on several different images with different texture attributes show that performance is limited due to

continued on page 47

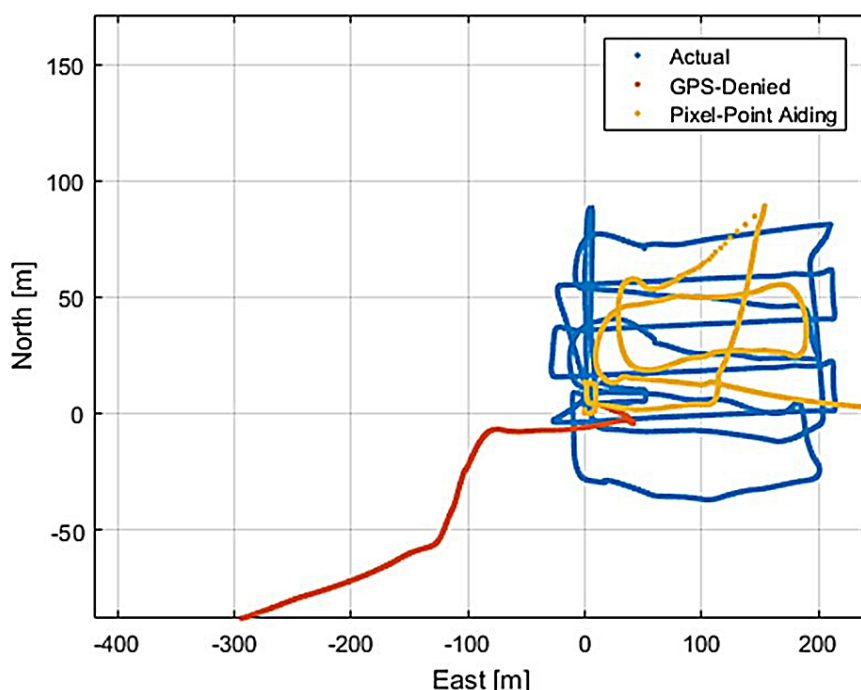


Figure 10: Navigation solution aided by Point & Pixel matching

Moafipoor, continued from page 41

low resolution and quantization levels in the point-based image. One possible approach to resolving this issue is to add the camera pixel information to the laser point clouds by the first team. In this approach, the DTM raster image can be represented with a quantization level closer to the camera-based image; and thus, the matching algorithm can be implemented by point-to-pixel and vice versa. In addition, new features can be added into the map, while also improving the accuracy of the prior terrain information. In future flights, a camera will be added to the first set of the UAVs and the alignment between the laser and the camera will be developed. Another future effort is to merge the two steps of building the reference map and navigation in the GPS denied condition to a single step using a mono-SLAM technique for real-time mapping and using the map for navigation in GPS denied conditions. ■

Dr. Shahram Moafipoor is a senior navigation scientist, focusing on new sensor technologies, sensor-fusion architectures, application software, embedded firmware, and sensor interoperability in GPS and GPS-denied environments. Dr. Moafipoor's work includes image-based navigation, LiDAR-based navigation, relative/collaborative navigation, and personal navigation systems.

Dr. Lydia Bock is the President and Chief Executive Officer (CEO) of Geodetics Inc. Dr. Bock has 35+ years of industry experience including electronics, semiconductors, telecommunications, in the commercial and the defense industries. Dr. Bock holds a Ph.D. from the Massachusetts Institute of Technology.

Dr. Jeff Fayman serves as Vice President of Business and Product Development at Geodetics. Dr. Fayman had many years of experience developing custom software solutions in the fields of Robotics, Computer Vision, Computer Graphics and Navigation. Dr. Fayman holds a B.A. in Business Administration and a M.Sc. in Computer Science, both from San Diego State University. He holds a Ph.D. in Computer Science from the Technion—Israel Institute of Technology.