Temasek Defence Systems Institute

# Test and Evaluation of An Image-Matching Navigation System for a UAS Operating in a GPS-Denied Environment

Keng Siew Aloysius Han
Advisors: Oleg Yakimenko, Ryan Decker

## Objectives

This thesis studies the use of image-matching techniques to provide positional information in a GPS-denied environment. Navigation systems within unmanned vehicles today are largely reliant on updates from the GPS and, in more capable systems, on the inertial navigation system (INS) as well. Within a GPS-degraded or GPS-denied environment on Earth or other planets, navigational capabilities are degraded significantly because the INS becomes the only source of a vehicle's movement estimate. Numerous unmanned vehicles today can and often are easily equipped with other passive sensors such as video cameras, as these devices have increasingly lower prices and improved sensor resolution. Such alternative sources of information can be used to work out the movement of the vehicle with respect to the operating environment. In the instance of video cameras, vision-based techniques can be harnessed for use as a navigation aid. Specifically, image-matching techniques rely on the stream of data from the cameras and a pre-stored depositoryofgeo-referencedreferenceimagestoestimatethecurrentattitudeandposition of a drone in flight.

This thesis extends the work previously done by Yakimenko and Decker in 2016, by studying the errors associated with the estimated attitude and terrain versus the actual recorded GPS position during data collection flights conducted at King City and Camp Roberts in California. Various parameters that can affect the image-matching navigation algorithm performance are also studied at different altitudes and in two different terrains.

## Research Results

Five major observations from the conducted evaluations are as follows.

1. The Image-Matching approach relies on the feature-richness of both satellite and onboard camera images. To this end, a typical satellite image provides a resolution of 0.5m$^2$ per pixel regardless of the size of the ground footprint. The resolution of on-board camera depends on the field-of-view (FOV, or zoom setting), altitude, and attitude. The best resolution is achieved in a level straight flight at low altitudes with a maximum zoom in. Nevertheless, such a setting results in a very narrow field of view(significant reduction in the number of features that can be used to match those of the satellite image). Specifically, with the TASE-200 sensor used in this research and a field-of-view of 35 degrees (Camp Roberts' flights), a resolution of 0.5m$^2$ per pixel can be achieved only when flying below 400m AGL. Likewise for King City flights, where the videos were taken at field-of-view of 10 degrees, only flights below 1200m can achieve 0.5m$^2$ per pixel resolution.

2. The texture of the Earth's surface has a major role. Specifically, flying over the agricultural area consisting of crop fields(between Greenfield and King City)at low altitudes with a narrow field-of-view results in no features detected in the onboard camera field-of-view. Some features can be detected only when flying in between the crop fields. One way to mitigate this effect might be increasing the field-of-view, but that leads to a decrease in resolution and possible failure to find the matches between two different resolution images. Still, this approach is worth exploring in the future.

3. Onboard camera stabilization (i.e., suppression of vibrations) has a crucial role, as well. In this research two aerial vehicles were used. The same sensor, a TASE-200, had much better stabilization when flying on UAV at 25m/s compared to that of a manned Cessna-206 flying twice as fast.

4. Varying the terrain elevation also contributes to the accuracy of IMMAT navigational solution. That includes are requirement to have a detailed terrain elevation map of the intended area of operations.

5. Aircraft attitude plays a major role, as well. In this research, IMMAT performance was evaluated only for straight level flight. Future evaluation should consider IMMAT performance while turning, climbing and descending. Using a limited set of test data based on a (not high-end) TASE-200 sensor with some vibration isolation problems along with incorrect reporting of pan-tilt information (which was discovered within this research effort and reported to the manufacturer) resulted in an unusually high drop rate. This occurred when there were not enough matching points to construct a projective transformation, which is a basis of the IMMAT approach. Nevertheless, this thesis was able to conduct a detailed assessment of the overall performance of the IMMAT algorithm.

The main conclusion is that when all conditions are met (i.e., at least five matching points are found), the IMMAT algorithm can provide an estimate of an aerial vehicle's position thatisaccuratetowithin50m from its true position (this value correlates with the satellite image resolution), and determine the vehicle's attitude within ±15 degrees for pitch and roll, while finding its yaw angle within just ±2-degree accuracy.

## Potential Future Developments

Optimizing the Number of Reference Frames. This seeks to answer the question on what would be the minimum number of frames required, below which the performance of the image-matching algorithm degrades. One possible idea is to take advantage of the fact that the code-base today is able to plot the viewpoint of the camera onto the aerial view of the area of operations. Using this information, it is possible to work out how far apart the reference frames can be spread out and still contain the viewpoint of the camera.

Creating a Feature-Rich Reference Image Library. During the creation of the RIL, there were instances where the reference images selected had few features. These frames were still included in the RIL to keep the algorithm simple. An immediate extension will be to select reference image frames that have sufficient features, but yet sufficiently spaced out and representative of the nominal trajectory to be covered, immediately reducing drop-rates.

Skip Rate for Incoming Video Stream. This seeks to answer the question of how many frames in an incoming video stream can be skipped to avoid unnecessary processing, but still allow it to provide accurate estimates on the UAV's position.

Investigating Image Feature Extraction Ability of Various Algorithms for Different Terrain Types. The analysis supports the claim that drop-rates are highly associated with the feature extraction capabilities of the image feature extraction algorithm used if the scenes between the reference image and the camera view are indeed overlapping. As the feature extraction algorithm is a component that can be substituted, future work can investigate the use of other feature extraction schemes such as SIFT or BRISK, and which are appropriate for the various terrains.

Managing Drops in Image Matching. On the whole, during the batch processing of the data, high IMMAT drop-rates were observed. Continued work to reduce the drop rates needs to be done to improve the robustness and reliability of the current IMMAT algorithm so that it can function as a viable source for navigational updates.

Improving Terrain Projection. It is assumed that the projected view of the terrain is an adequate approximation of the camera view for the purposes of image-matching. This assumption may be violated should the re-projected view of the planar satellite image differ from the actual perspective view of the physical terrain. In order to study the differences of error in elevation projection, further work needs to be done with a satellite imagery textured digital elevation model of the terrain for in-depth studies.

Using Alternate Video Streams. This thesis was primarily assessing the effectiveness of using the day camera output of a UAV. Some UAVs, however, may also be equipped with IR cameras, which images the environment within a different spectral band. In terrains where the IMMAT algorithm may produce a poor estimation when using a day camera, the output could potentially be substituted with the view from the IR camera, which may reveal features that are otherwise imperceptible in daylight.

Studying the Effect of Actual Camera Field-of-View. Based on the data sets available, the Horizontal Field-of-View of 10.5◦ was used for King City recording and 35.26◦ for Camp Roberts recording. While the drop-rates seen in the King City flights were significantly higher than those for Camp Roberts, it is not possible to conclude whether it was the result of a smaller actual camera FOV or the effect of the terrain in King City that was challenging for the feature extraction algorithm to produce a match. Further studies using more data sets with varying actual camera fields-of-view are required to understand this aspect of the algorithm.

Using High-Fidelity Simulated Urban Environment Fly-By as Reference Images. There are systems that can generate high-fidelity simulations based on the inputs of a fly-through route. While it is difficult to replicate the environment accurately for remote places, it should be possible to get a reasonably accurate model of a 3D urban environment. One pertinent research question is whether the image matching algorithm still is able to provide reasonable estimates of position despite using a simulated scene as a reference.

NUS National University of Singapore