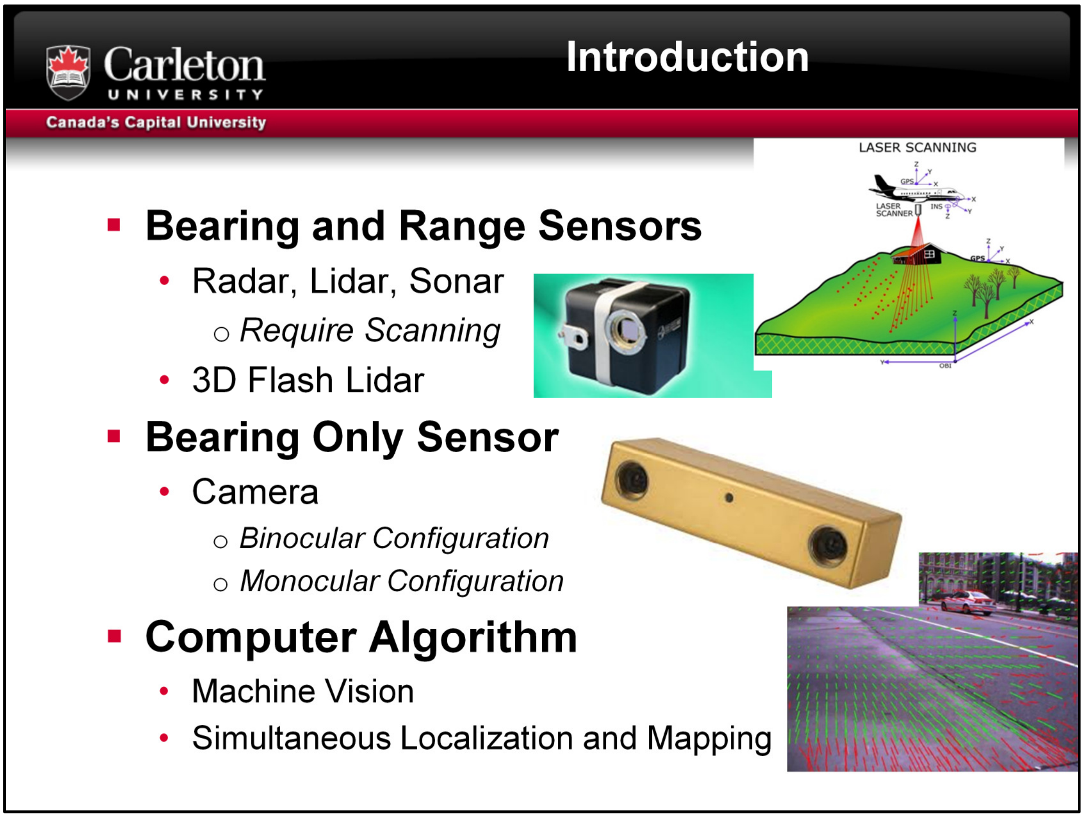
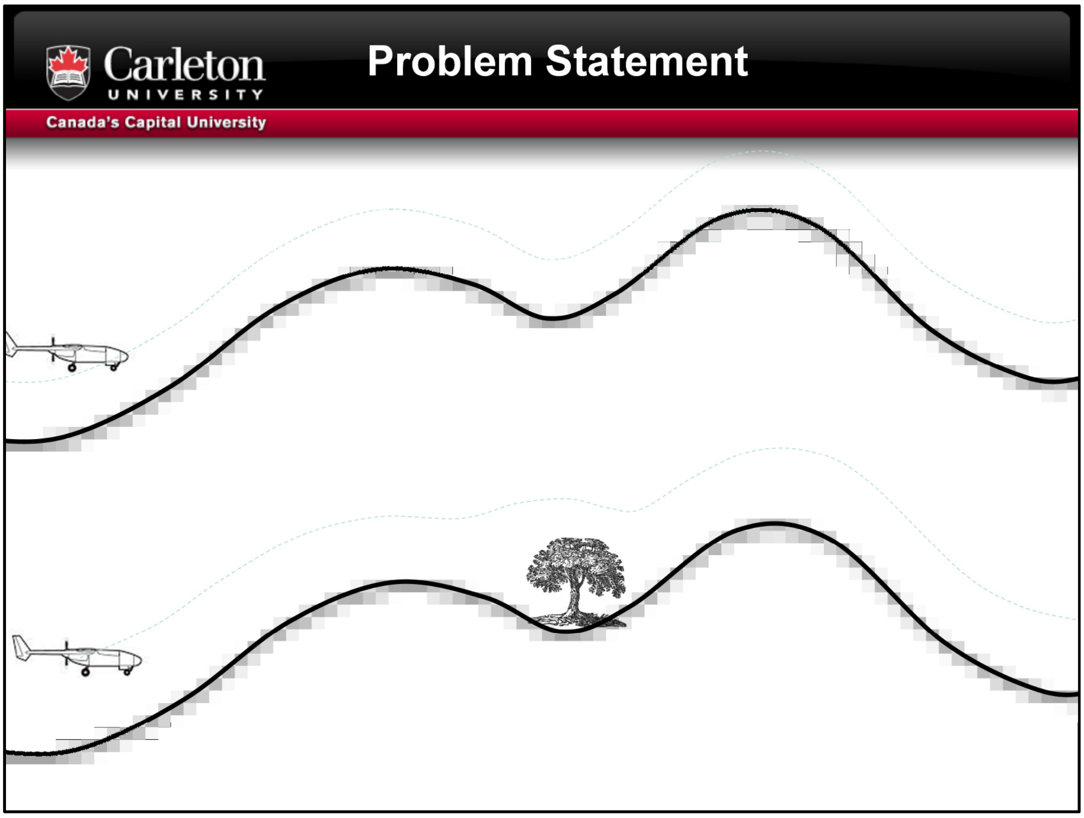
Carleton UNIVERSITY
Canada's Capital University

Obstacle Detection Using Monocular Camera for Low Flying Unmanned Aerial Vehicle

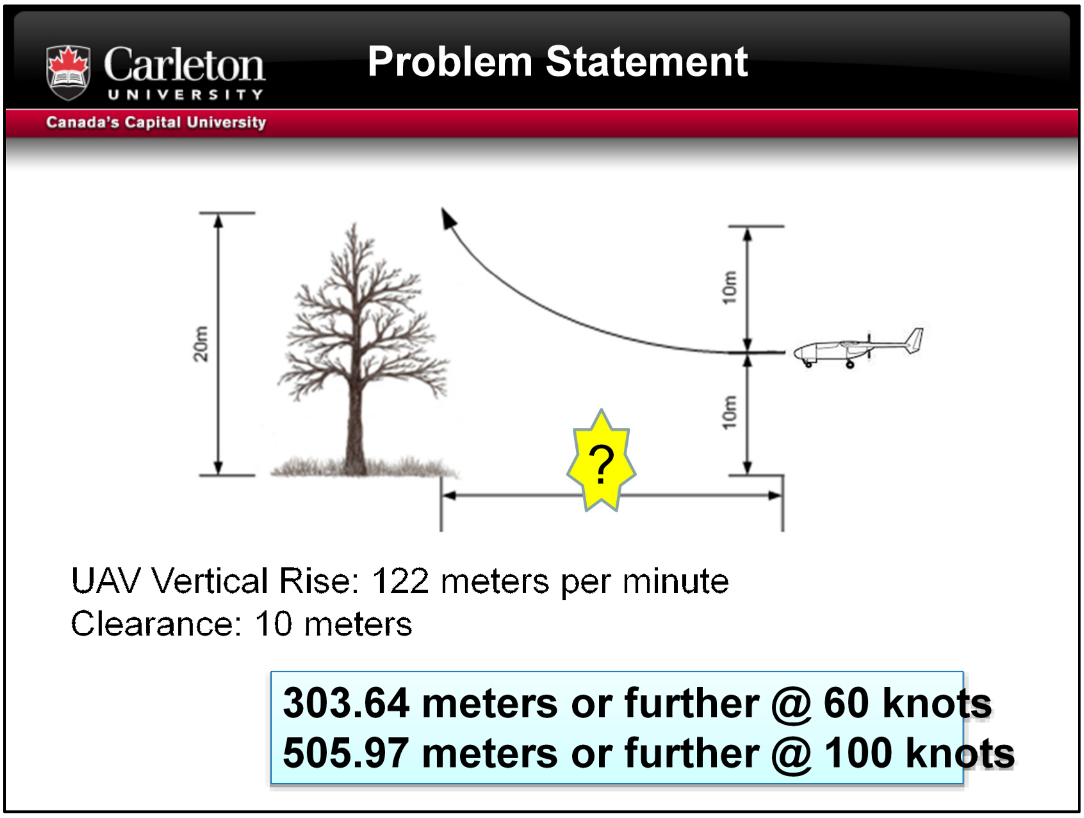
Fan Zhang, B.Sc.
January, 2015



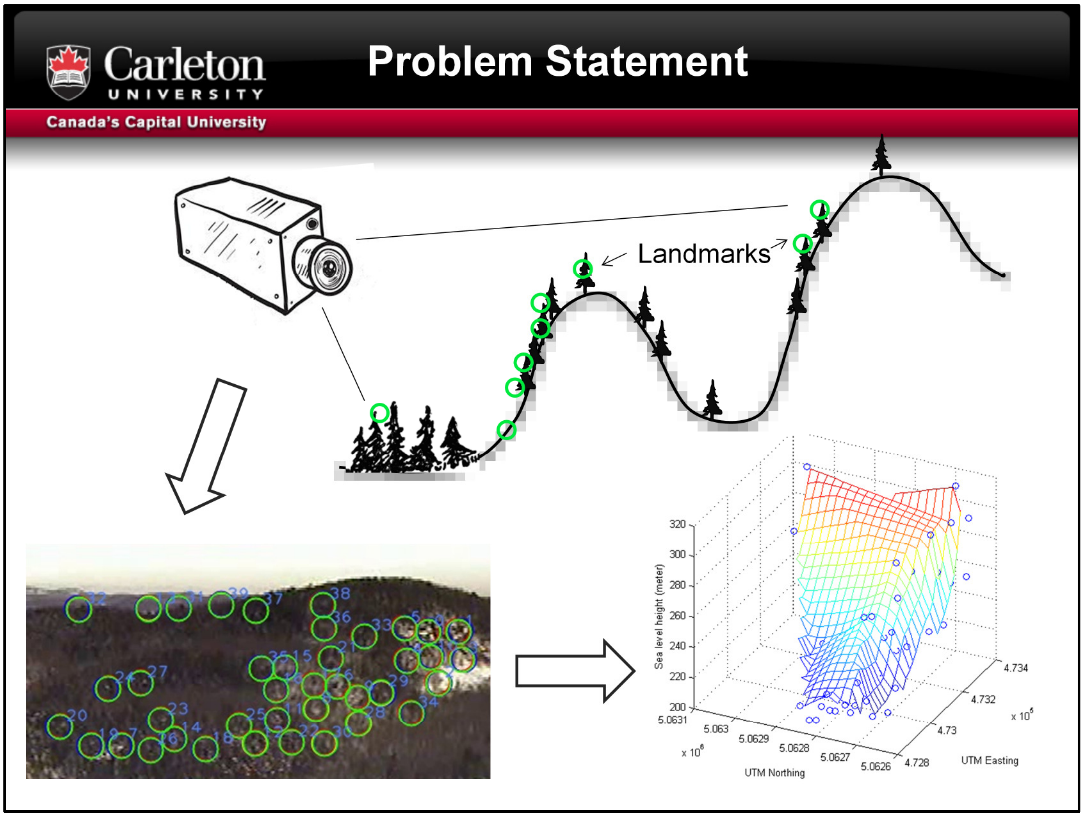
•In order to avoid obstacle, we must know where it is, which is knowing its distance and direction from the vehicel.
•Many sensors are used in obstacle detection application. Some sensors are range and bearing sensor, meaning that they output range and direction measurement of the target, such as radar, lidar, and sonar. To get a 3D depth of the scene, mechanical scanning is required.
•3D flash lidar can obtain a depth measurement of the entire scene simultaneously. But they are very pricy, which limits their use in commercial application.
•Another solution is to use video camera. They are inexpensive, small size, has many options. Because image sensors are bearing only sensor, to measure distance, you need to two view of the same scene, and obtain depth measurement through triangulation. Same way as how human eye works.
•There are two type of configuration, binocular configuration usually has two camera side by side and fixed, like human eye. Monocular configuration use only one camera to capture image sequence, and obtain depth from the image sequence. When no odometry measurement is available, monocular configuration can obtain a structure of the scene, but cannot measure depth. To measure depth, monocular camera must work with odometry measurement.
•Using video camera for depth measurement requires computer algorithm. The machine vision algorithm processes the image sequence to extract information. On top of that, a data fusion framework is used to estimate vehicle trajatory as well as mapping the static scene.



For this research, it focus on small to medium size UAV conducting low altitude terrain following flight. [click]
Digital elevation map is usually used to plan the flight path of a UAV. Its resolution is too low for obstacle detection. For example, when in this scenario, the UAV must be able to spot the obstacle, and plan a way around it. [click]
Only static scene is considered at this stage.



Then, how far do we need to detect the obstacle so that the UAV has enough time to avoid it. Based on the specification of the GeoSurv2, and giving about 10m of clearance from the obstacle, the UAV need to detect the obstacle at least 300 to 500 meters away depending on its flying speed.



The solution proposed by the thesis, is to capture image sequence of the scene with a forward looking monocular camera, extract visual feature from the scene, called landmark, and tracked them through the frame. The SLAM algorithm can then use these landmark and the motion measurement of the camera to estimate the trajetory of the camera, as well as generate a sparse map of the scene.

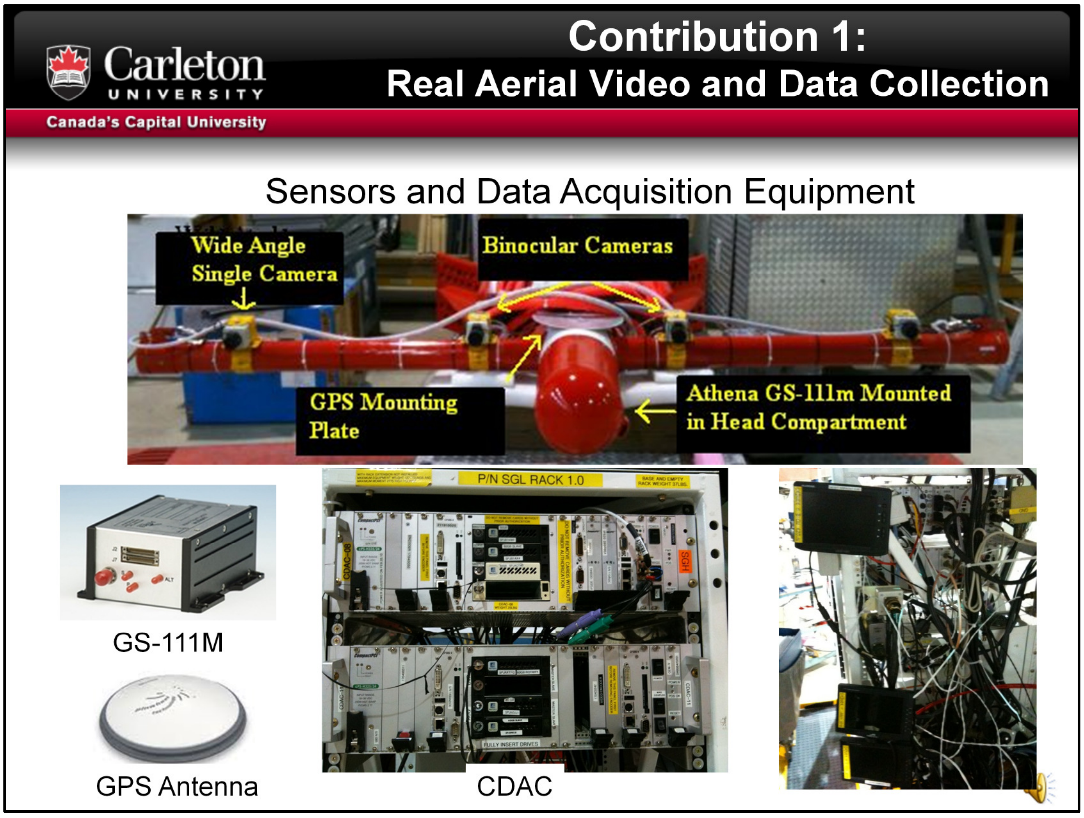
Carleton UNIVERSITY
Canada's Capital University

Contribution 1: Real Aerial Video and Data Collection

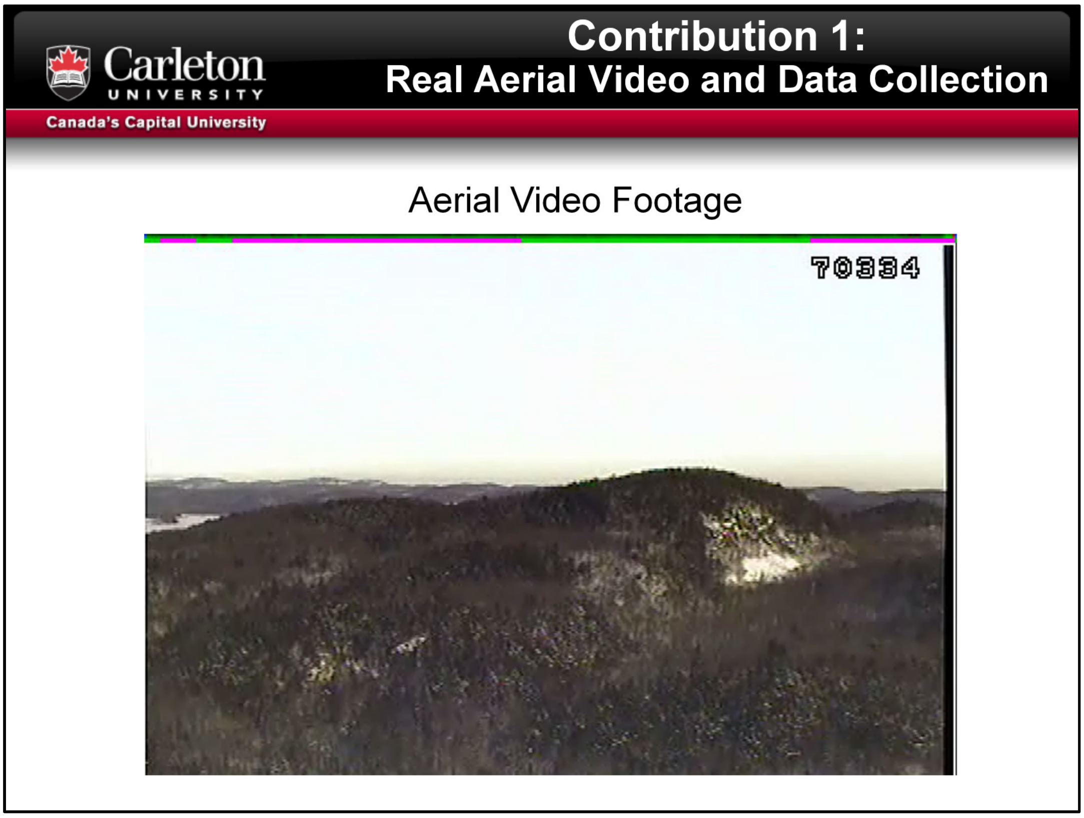
Video of the Test Flight

The first contribution of this thesis is that I have collected real aerial video and navigation data. Here is a video of the test flight.

The test flight site is in the mountain area north of Gatineau. A simulated unmanned aircraft system (SUAS) was fabricated at SGL, and was used to carry all sensors. The SUAS was towed by a helicopter via a tow rope of 33 meters long. The helicopter flew a planned path at 100m above ground. The SUAS approximately at 60-70 meters above ground.



Sensors mounted on the SUAS included one wide angle CCD camera with 6 mm focal length capturing monocular video at 30 fps, one GPS antenna mounted on the plate here, and one INS/GPS navigation unit Athena GS-111m mounted in the head compartment of the SUAS. There were also two narrow angle cameras install for capturing binocular video. These are installed to another research. Analog videos were sent to the helicopter through BNC cables. Navigation data were sent through RS485 signaling. All the videos and data are recorded with CDAC installed in the helicopter. There were three monitors installed in the helicopter so that the operator can monitor those video.



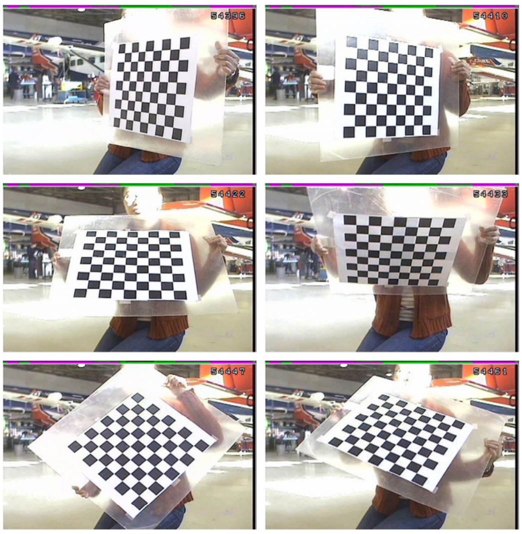
The video were captured by the cameras was digitized to 720x480 resolution images and time-stamped with GPS second on the image screen for post-flight synchronization with the inertial measurements.

Contribution 1: Real Aerial Video and Data Collection

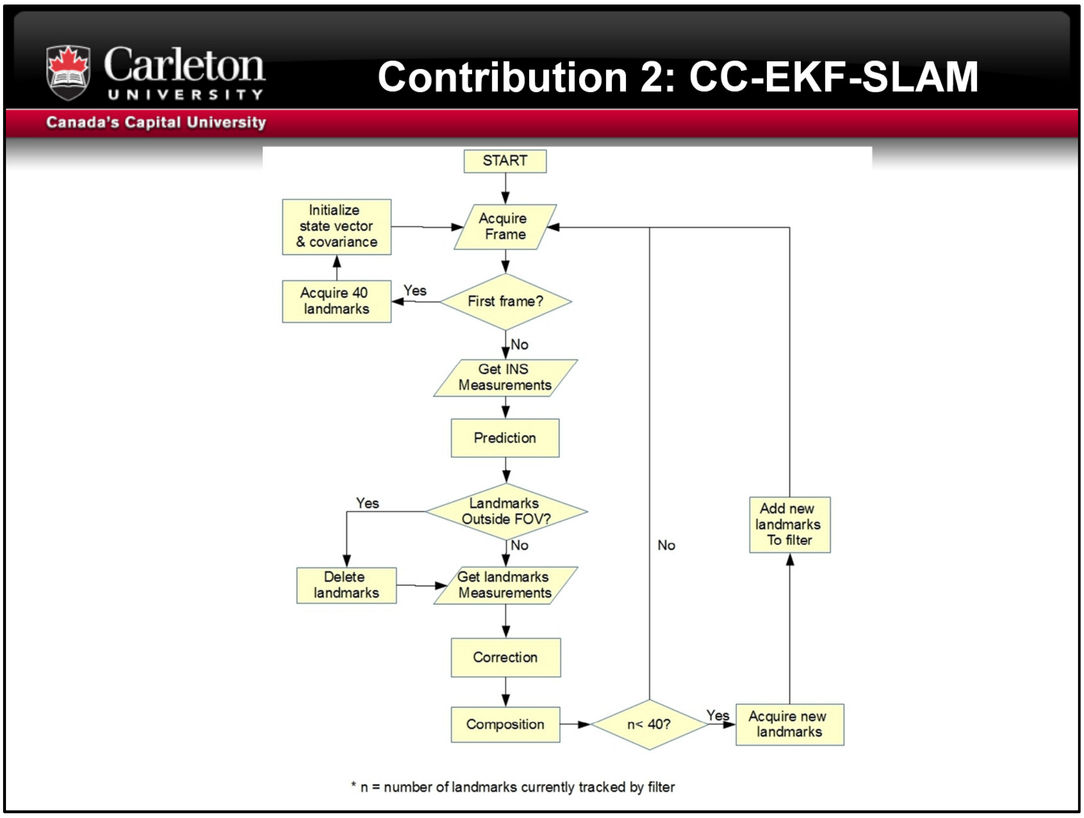
Carleton University
Canada's Capital University

Camera Calibration

Parameter	Result
f_x	887.6 pixels
f_y	805.7 pixels
c_x	381.8 pixels
c_y	293.7 pixels
k_1	-0.102
k_2	-0.535
p_1	1.15e-003
p_2	8.40e-003



A camera calibration was done after the flight. The calibration algorithm requires different view of the calibration target. The results are listed here



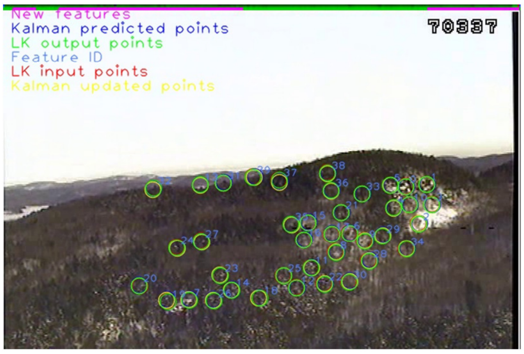
2nd contribution is the implementation of the CC-EKF-SLAM algorithm.

Carleton UNIVERSITY

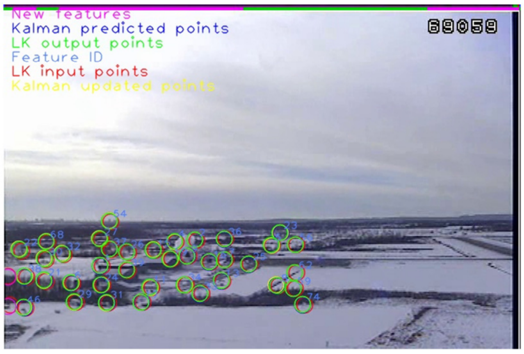
Canada's Capital University

Contribution 3: Aerial Data Processed by CC-EKF-SLAM

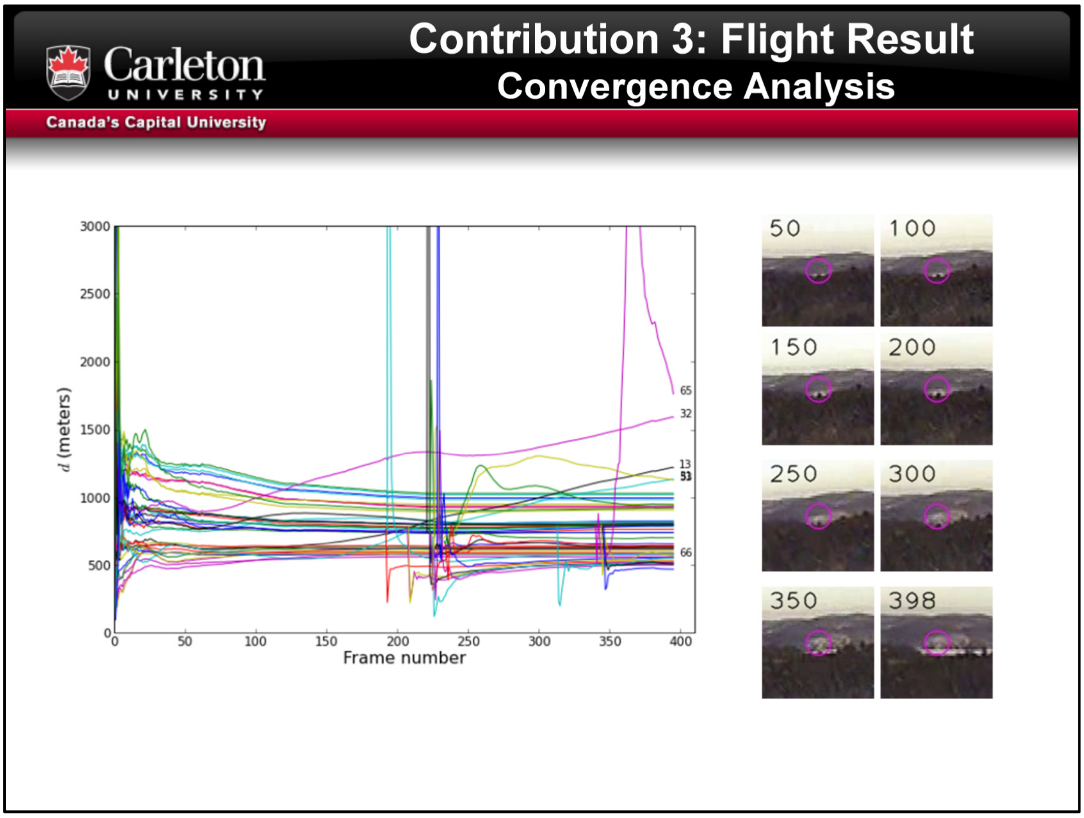
Natural Scene



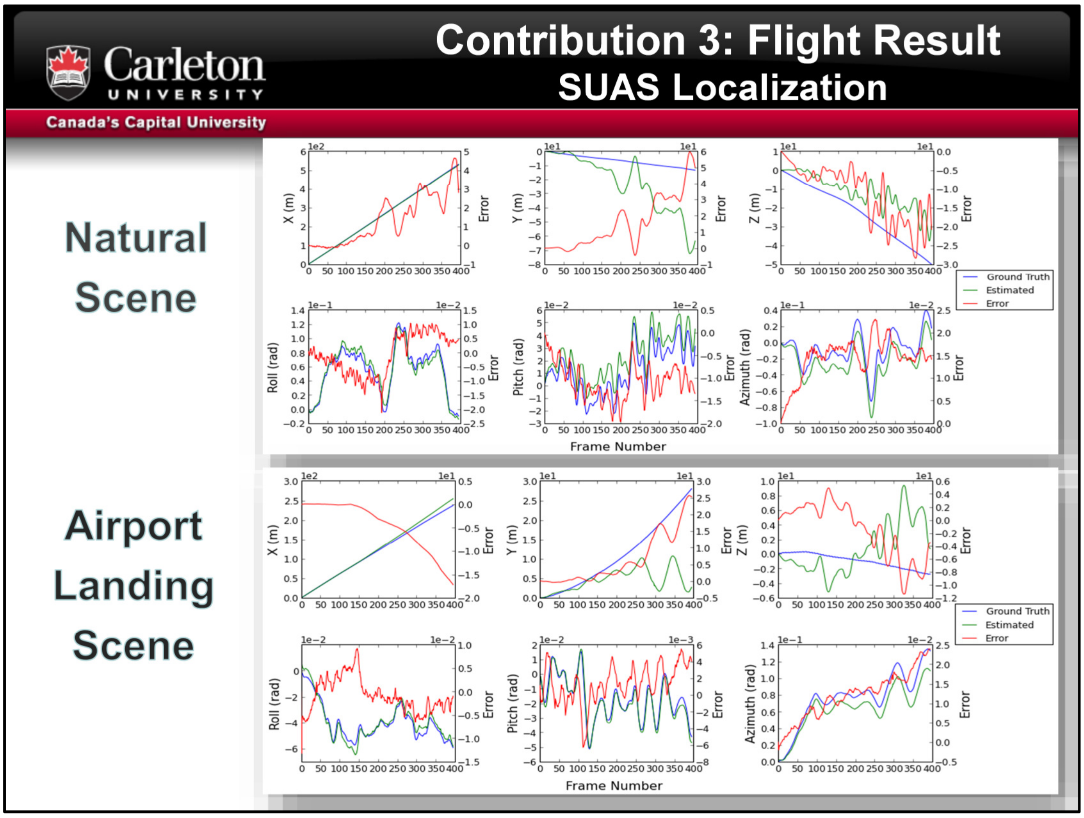
Airport Landing Scene



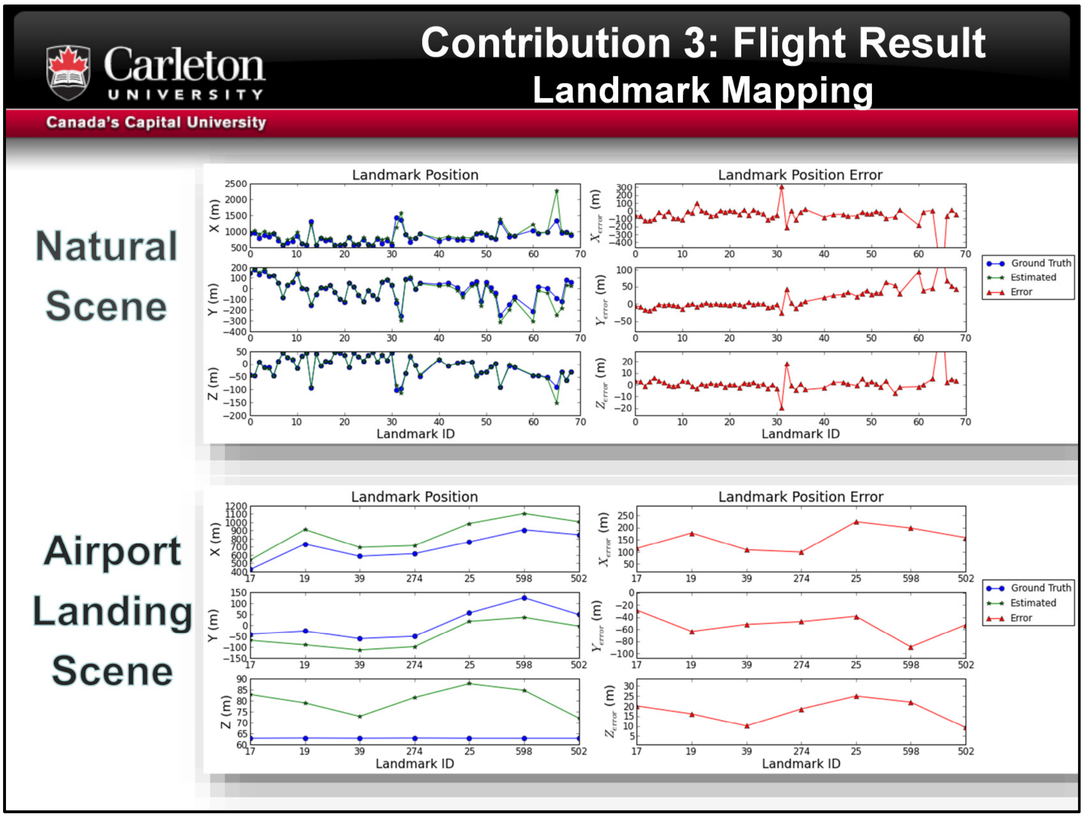
Two pieces of video were processed by the CC-EKF-SLAM algorithm. Each piece is 400 frames long.



First the convergence behavior is analyzed. Most landmarks converged very quickly. Some landmarked that's drifting away are those located on the hill top. Because the small error allowed in visual tracking algorithm, these visual feature changed slightly frame to frame, and eventually became something different.



For the accuracy of SUAS localization, the estimated poses generally agree with the ground truth. The position on the Y and Z axis sees more drift, and the position on the X axis has very good accuracy. At the same time, it can be observed there is correlation between the position estimate and the rotational motion of the SUAS. The position error on Y are correlated to rotation on Z, and position error on Z is correlated to the rotation on Y.



•Natual scene, Average distance on the X axis is 1000meters, with error within +/- 100meter. A good indicator that the algorithm is able to map object around 1000m range.

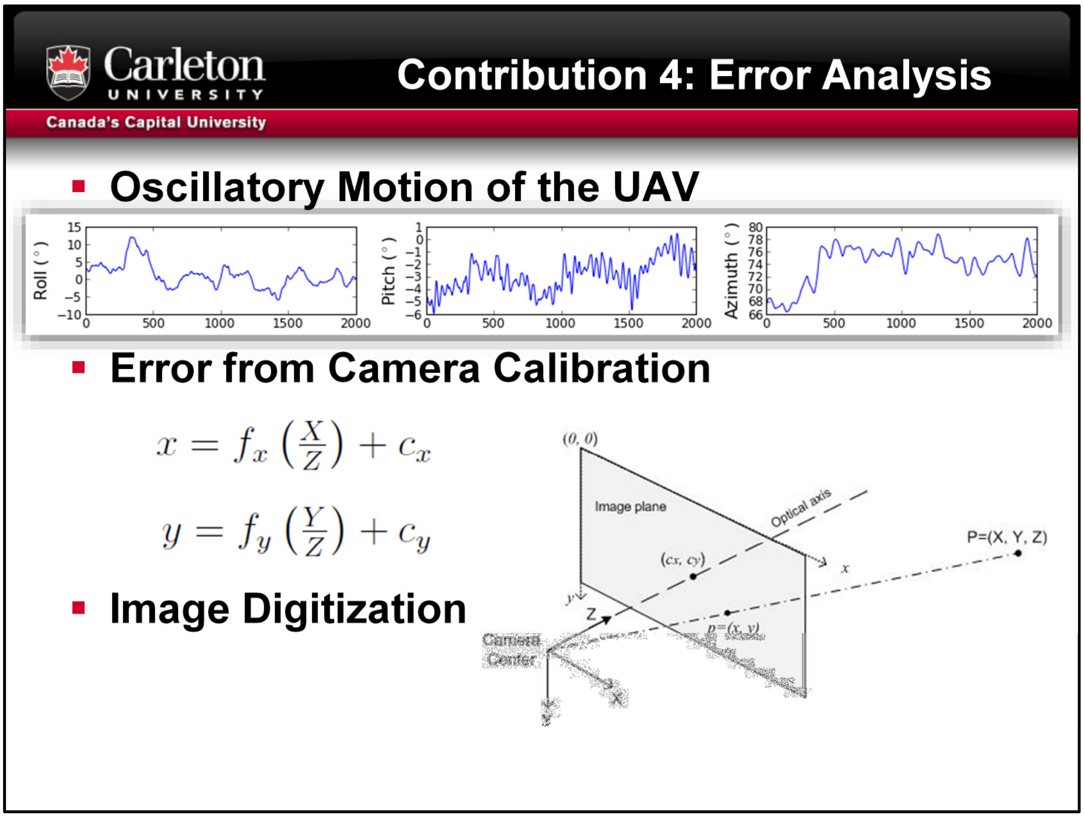
•Landmark with ID bigger than 40 has offsets

•In airport landing Landmark at the corner of the image plane has offset error

Carleton UNIVERSITY

Canada's Capital University

Contribution 3: Flight Result
Consistency Analysis

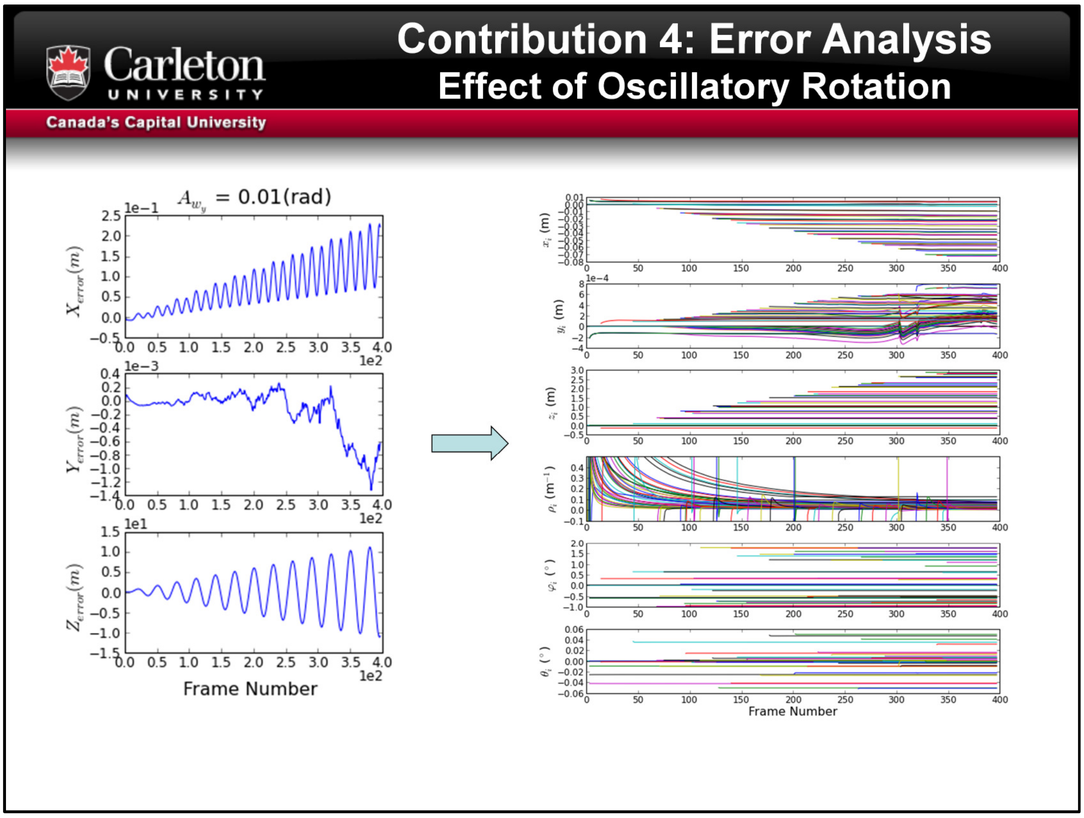


The 4th contribution is that I analyzed the performance of the algorithm under a number of scenario.

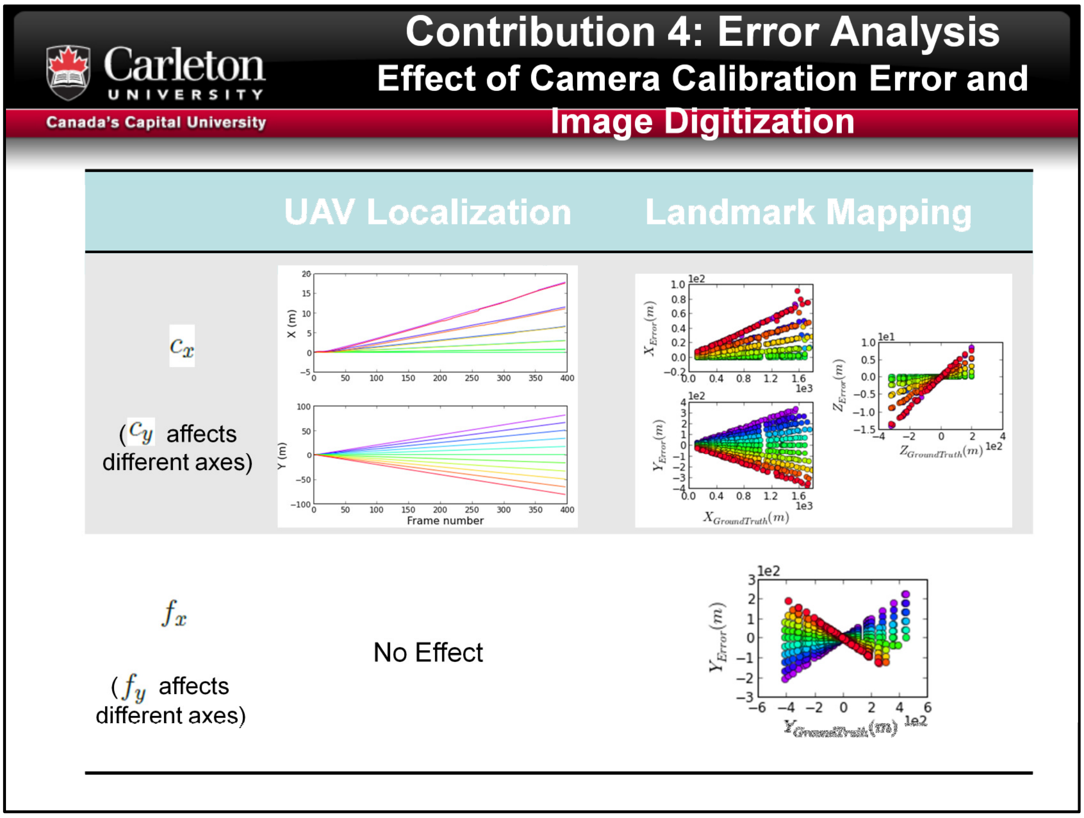
1st scenario is oscillatory motion of the UAV. I found from the flight data that the UAV is under a lot of oscillatory motion, especially the rotations. Here is a piece of rotation recording of the SUAS. We can see that the Y and Z axis has lots of oscillatory rotations.

2nd scenario is error in camera calibration. I found that when choosing different number and different view of the input images, the output of the calibration program is different. Therefore, the calibration results do have error in them, and it would be good to know how much impact they have on the accuracy of the algorithm.

3rd scenario is the image digitization. For the machine vision algorithm to process the recorded video, they must be digitized first. There are lots of choice for image sensor at various resolution. This analysis will help us deciding how much resolution we need for the targeted distance.



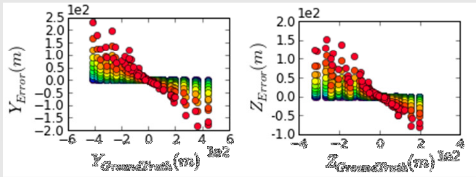
The simulation confirmed that CC-EKF-SLAM is very sensitive to the oscillatory rotation. This slides shows result of UAV experiencing rotation around Y. The estimated UAV position are showing oscillatory and diverging error on both X and Z. As a result of this self localization error, the landmark mapping also have offset error for landmark not initialized on the first frame



The next two slides summarized the effect of camera calibration error. For error in Cx, and Cy, which is the coordinate where optical axis intercept the image plane, it affect both UAV localization and landmark mapping, and its effect can be modelled by 1st order polynomial function. For the scaling factor fx,and fy, it doesn't affect UAV localization, only the landmark mapping.

Carleton UNIVERSITY
Canada's Capital University

Contribution 4: Error Analysis
Effect of Camera Calibration Error and Image Digitization

	UAV Localization	Landmarks Mapping
Lens Distortion	Diverging Error	
Image Digitization	1080 x 1440 or higher	1080 x 1440 or higher

For lens distortion, it caused diverging error in UAV localization, as well as error in landmark mapping. The further the landmark locate from the optical center, the more error it suffer. At last for sensor resolution, in order to achieve a good accuracy for distant subject, sensor resolution should be 1080x1440 or higher.

Carleton UNIVERSITY
Canada's Capital University

Future Work

- Add lens distortion model
- Increase sensor resolution
- Add landmark quality checking function
- Research on map joining algorithm
- Increase accuracy by syncing to GPS
- Investigate on the sensitivity problem to oscillatory rotation