Self-Improving Structure-from-Motion in Orbit

Multiview Onboard Computational Imager Satellite (MOCI Sat) – University Nanosat Program 9 – AFRL

Jacob C. Conley, Dr. David Cotten
jake@jakeconley.com, dcotte1@uga.edu
Small Satellite Research Laboratory, Center for Geospatial Research, University of Georgia

Overview

The Multiview Onboard Computational Imager (MOCI) Sat is a satellite under development by the Small Satellite Research Laboratory at the University of Georgia. The satellite, equipped with two cameras, captures images of the Earth's surface at a point as it passes over it, and from these images it can reconstruct a Digital Elevation Model (DEM) of the surface features, or a 3-D mesh of the surface or any features above it, by a process known generally Structure-from-Motion (SfM).
The many variables involved, such as the number of photos taken, the angle between them, and the offset from which the satellite's trajectory deviates from the target's position can greatly affect the result. As well, though the process is mathematically rigorous, some approximations must be made in order to make the computations feasible to be performed on board the satellite.
MOCI employs a neural network to optimize the SfM parameters as described above. Each parameter has an associated input and output node, with additional inputs describing the target in further detail. There is one hidden layer. This deep learning helps to make the most out of each mission and ensures that the results of each mission continue to improve.

The MOCI Pipeline

Figure 1 describes the pipeline MOCI uses to process images.. After the image set is captured, features are detected and tracked in each image using the Tomasi-Lucas-Kanade feature detection and tracking algorithm on a Field Programmable Gate Array (FPGA) specifically optimized for this task. Afterwards, camera parameters (specifically, position and orientation) are calculated for each image; this is what is technically known as the Structure-from-Motion step. At this point enough information is known to be able to compute reconstruction of the scene; this is known as the Multi-View Stereo (MVS) problem. MOCI uses the Clustering Multi-View Stereo (CMVS) package on a Jetson TX1 (TX1) to perform the sparse reconstruction, and builds upon this reconstruction with the Patch-Based Multi-View Stereo (PMVS) software package to create the dense reconstruction. After the dense point cloud is generated, a mesh is generated using the Poisson mesh reconstruction algorithm and the mesh can thereafter be converted into a DEM.

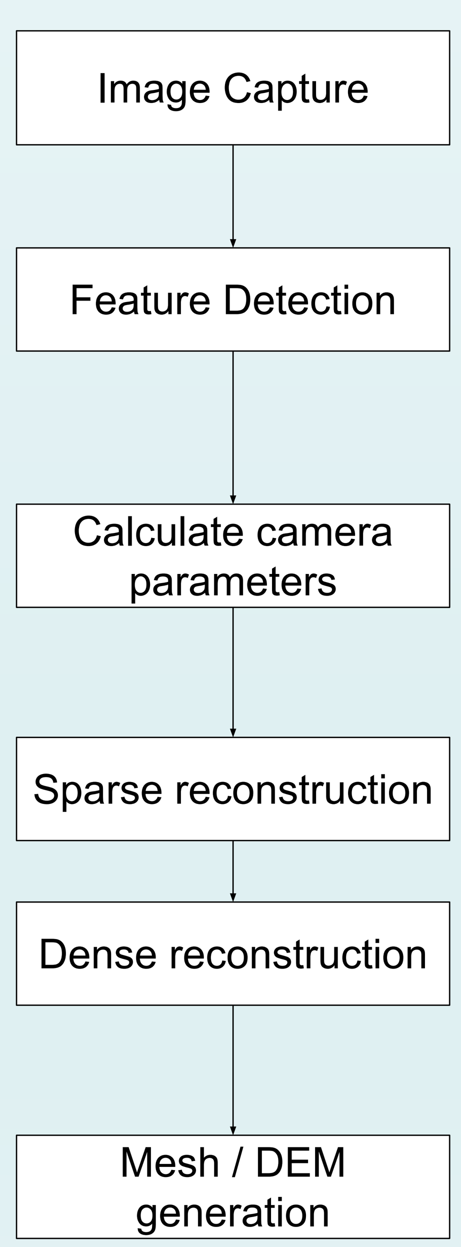
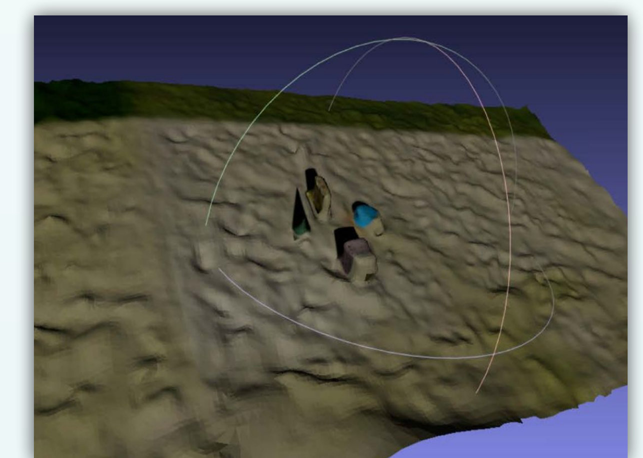


Figure 1

Figure 2

Figure 5



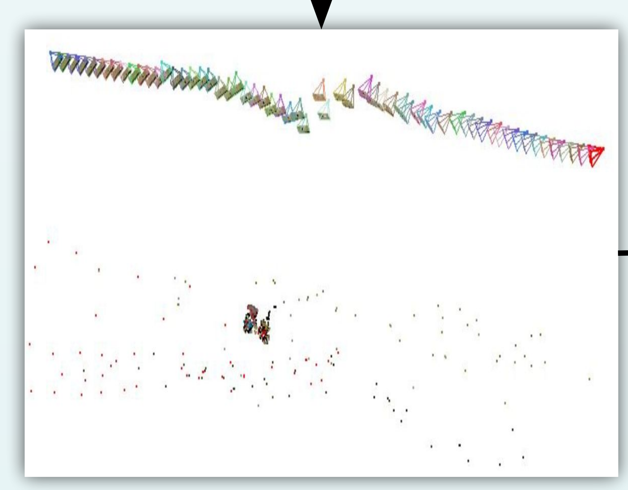


Figure 3

Figure 4

The above is a proof-of-concept simulation from early 2017. A test scene (Figure 2) was rendered to scale and images were rendered at each point MOCI would take an image. Figure 3 shows the sparse reconstruction output by CMVS, Figure 4 shows the dense reconstruction output by PMVS, and Figure 5 shows the final reconstructed mesh.

Parameters to Optimize

The number of images input and the angle between them has been shown to wildly vary the results of the SfM process. As well, many calculations in the SfM and MVS processes involve approximations, which generally involve tunable sensitivity parameters. One such approximation is for the Tomasi feature detection, which involves determining if the eigenvalues of the below expression are very large or very small:

$$E(u,v) = \begin{bmatrix} u & v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \quad \text{, where M = } \quad \sum_{x,y} W(x,y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

However, computing eigenvalues is extremely expensive and impractical on the scale of MOCI's mission. Harris and Stephens suggested[1] approximating it with the single expression below:

det(M) - k(trace(M))2

Where k is an adjustable parameter. This parameter is a perfect candidate for the neural network to optimize as it bears no physical meaning but nonetheless impacts the reconstruction process greatly.

The Neural Network

Artificial Neural Networks (ANNs) emulate machine learning by approximating linear functions. It consists of "neurons", arranged in layers, as seen in Figure 6. MOCI's ANN loosely resembles the ANN described in Figure 6 as it has one input layer, one middle "hidden" layer, and one output layer. Each neuron has or can be given a value which determines how the next layer is processed. These values change as the satellite "learns" through a process known as backpropogation.

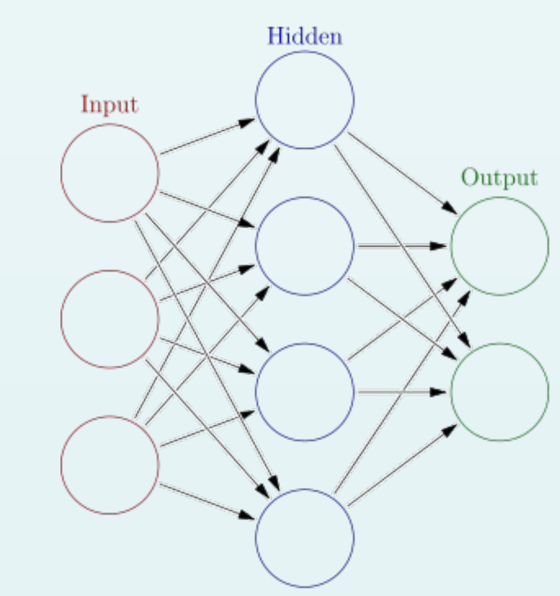


Figure 6

References and Acknowledgements

1) Harris, C., Stephens, M., "A Combined Corner and Edge Detector", 1988
2) Lucas, D., Kanade, T., "Optical Navigation by the Method of Differences", Proc. of the Ninth International Joint Conference on Artificial Intelligence, 1985

For more information about CubeSats, the Small Satellite Research Laboratory (SSRL), our missions, visit www.smallsat.uga.edu.