Visual Localization for UAVs on Mars

Alex Hagiopol

Outline

- Motivation: Autonomy Mission & Problem
- Proposed Solution
- Hardware & Software Implementation
- Results
- Future Work
- References

Motivation

Rover damage caused by terrain on Mars. Images from NASA JPL.

Motivation

Rover damage caused by terrain on Mars. Images from NASA JPL.

Motivation

- Objectives of NASA UAV Missions:
  - On Earth: physical package delivery, data acquisition, and exploration.
  - On Mars: exploration for rover pathfinding. First-ever take-off, flight, and landing on another planet.
- Autonomy necessary on Mars: speed of light communications too slow to pilot a drone.
- Vehicle localization technology needed.
  - No GPS on Mars.
  - Even if there were GPS, would it be good enough?
    - Geology science missions demand the highest quality observations.
    - Centimeter-accuracy localization for measurements?

Proposed Solution

- Goal: Create a fast and accurate localization system for a Mars UAV.
  - Relative to known starting point (e.g. rover itself).
  - Concerned with localization only - not detailed mapmaking.
  - Constraints:
    - Lightweight.
    - Low power.
    - Tight integration schedule*.

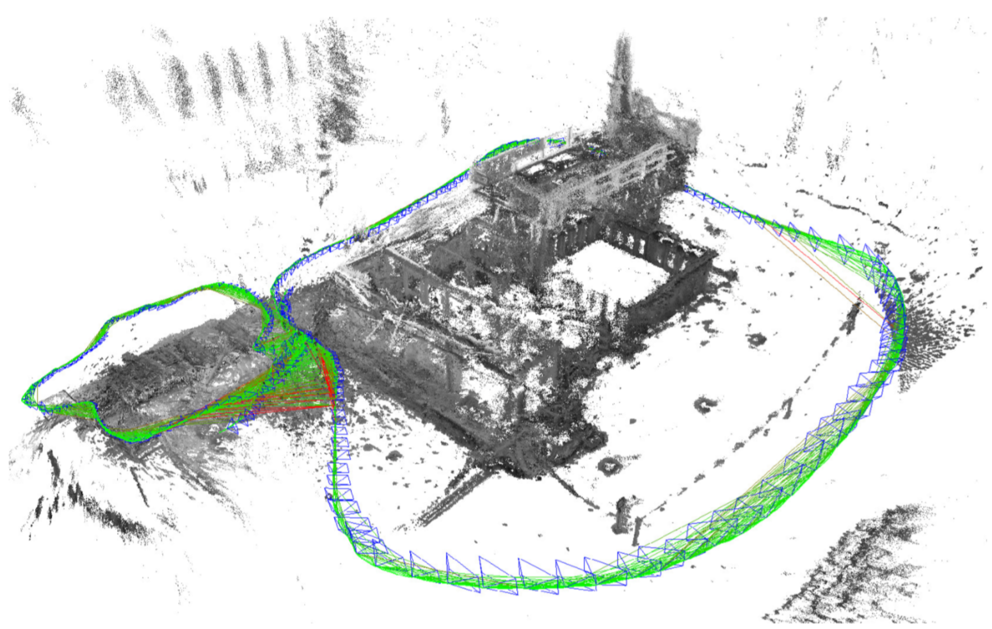
- Proposal: Estimate UAV motion from camera stream using visual odometry.
  - Input = video stream from a camera.
  - Output = estimated camera poses.
  - Algorithmic minimalism: No loop closure, no global bundle adjustment.
  - Hardware minimalism: only a camera and computer are needed.
  - Historical precedent: Spirit & Opportunity rovers applied visual odometry [1].

* At the time of development, different teams were working on different subsystems of the vehicle. The software and hardware infrastructure to fuse IMU data from inner controller subsystem with camera data wasn't yet ready.

[1] M. Maimone, Y. Cheng, and L. Matthies, "Two Years of Visual Odometry on the Mars Exploration Rovers," 2007.

Proposed Solution

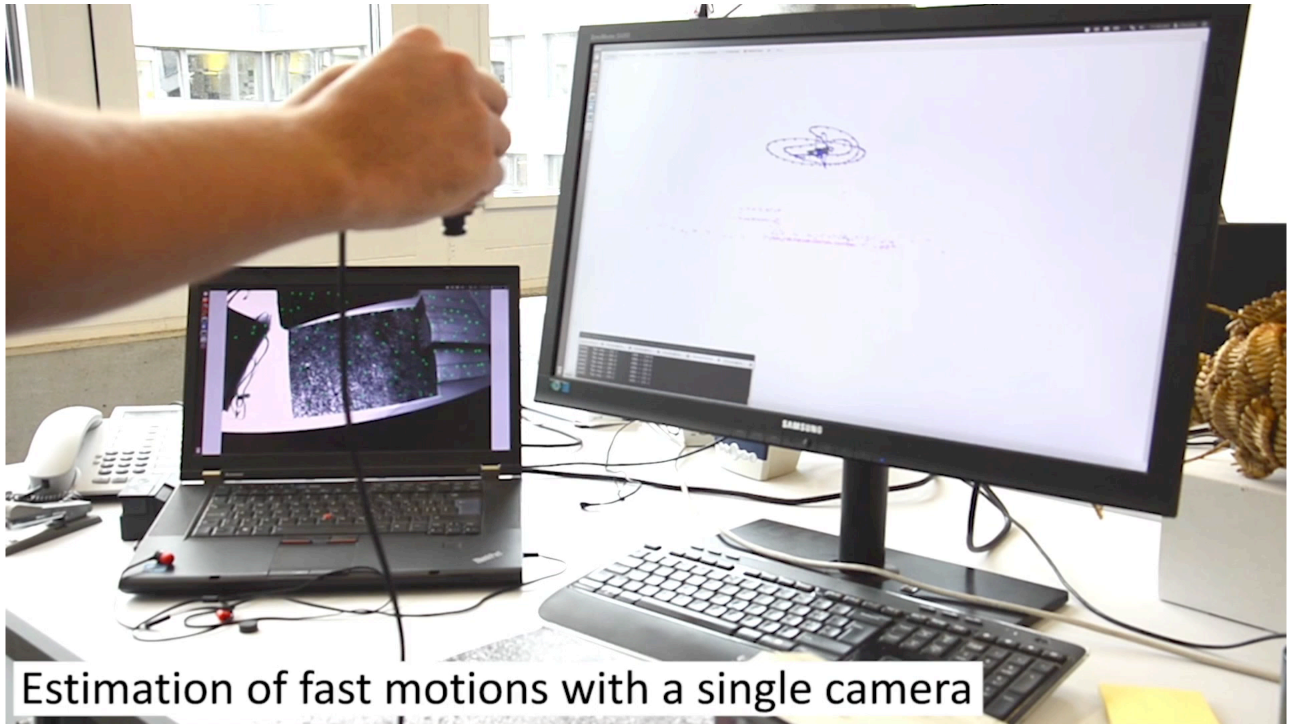
- Candidates:
  - PTAM [3]:
    - Pioneer method in visual SLAM. Based on FAST feature extraction for tracking.
    - Limited in scale of environment: problematic if initialized map leaves view.
  - ORB-SLAM [4]:
    - Great #2 choice, but not as fast. Computes ORB features & descriptors on every frame.
    - Includes loop closure which is useful but maybe not worth speed penalty.
  - LSD-SLAM [5]:
    - Direct method based on optimizing over pixel intensities for tracking.
    - Lower speed & accuracy. Creates dense map which isn't needed:



[3] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," 2007.
[4] R. Mur-Artal, J. Montiel, J. Tardos, "ORB-SLAM: a Versatile and Accurate Monocular SLAM System," 2015.
[5] J. Engel, T. Schops, D. Cremers, "LSD-SLAM: Large Scale Direct Monocular SLAM," 2014.

Proposed Solution

- Integrate Semi-Direct Monocular Visual Odometry (SVO) [2] on vehicle.
    - Extremely fast: >300Hz on i7 CPU.
    - Accuracy comparable to competing methods.
    - Monocular: save weight of stereo rig.
    - Not SLAM: skip expensive steps for speed.
        - No loop closure & large scale bundle adjustment. No per-frame feature extraction.
        - No descriptor computation and matching for triangulation.
    - Frame-to-frame transformations estimated using photometric error minimization.
    - 3D points estimated using Bayesian estimation: "depth filtering."



[2] C. Forster, M. Pizzoli, and D. Scaramuzza, "SVO: Fast Semi-Direct Monocular Visual Odometry," 2014.

Proposed Solution

- SVO[2] in depth:
  - Initialization:
    - Given first 2 views, assume scene is planar, estimate homography, initialize map with 3D points.
  - Motion Estimation:
    - For each new frame, optimize transformation w.r.t previous frame using difference of reprojected patches as error.
    - Optimize 2D position of patches in current frame, 3D position of visible map points, and camera pose using same error metric.
  - Mapping:
    - If a frame is far enough away from other keyframes, it becomes a keyframe.
      - Compute FAST[6] feature points (unknown depths).
      - Initialize depth filters for each point.
    - If not a keyframe…
      - Update other depth filters.
      - Add depth filtered 3D points to map when converged.

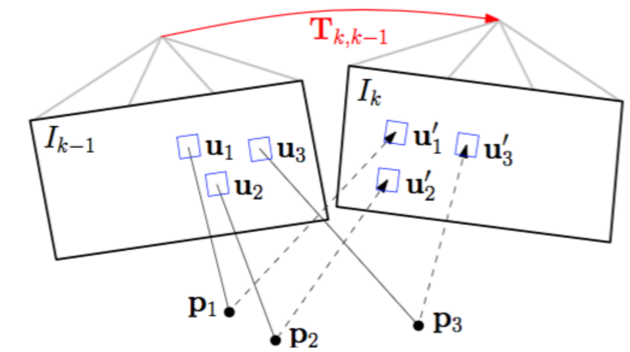
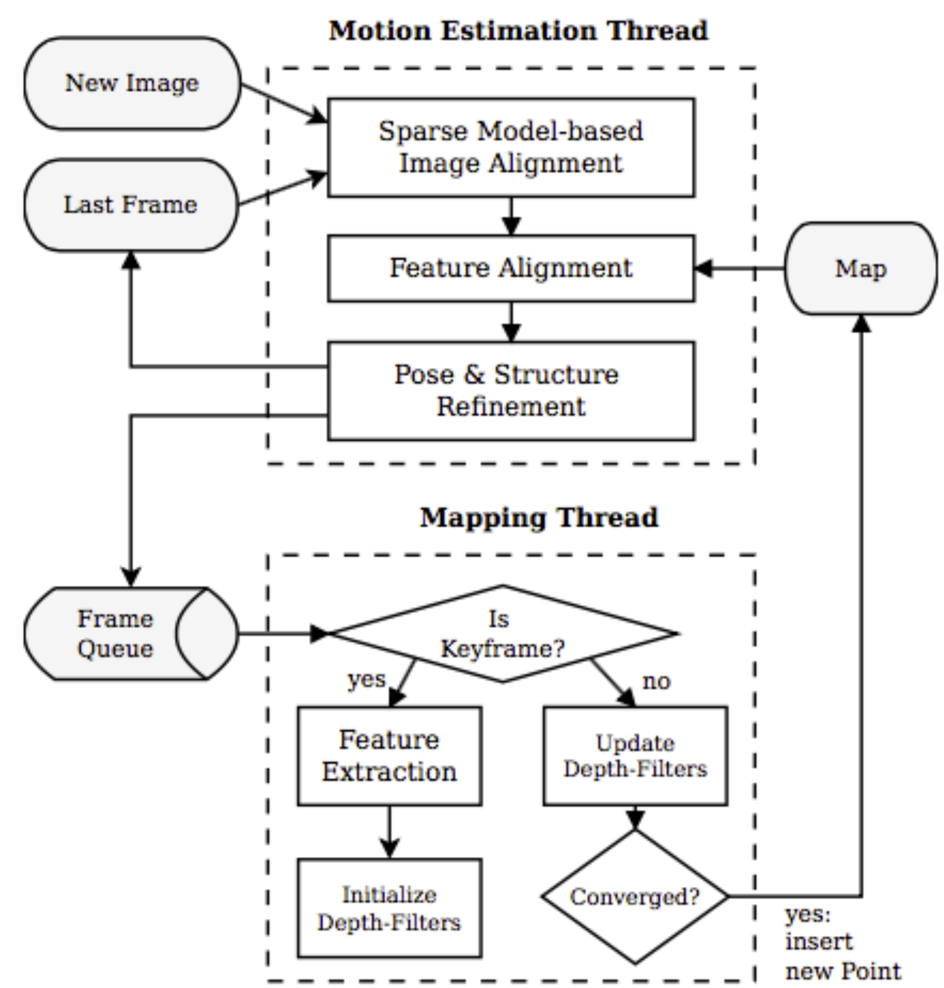


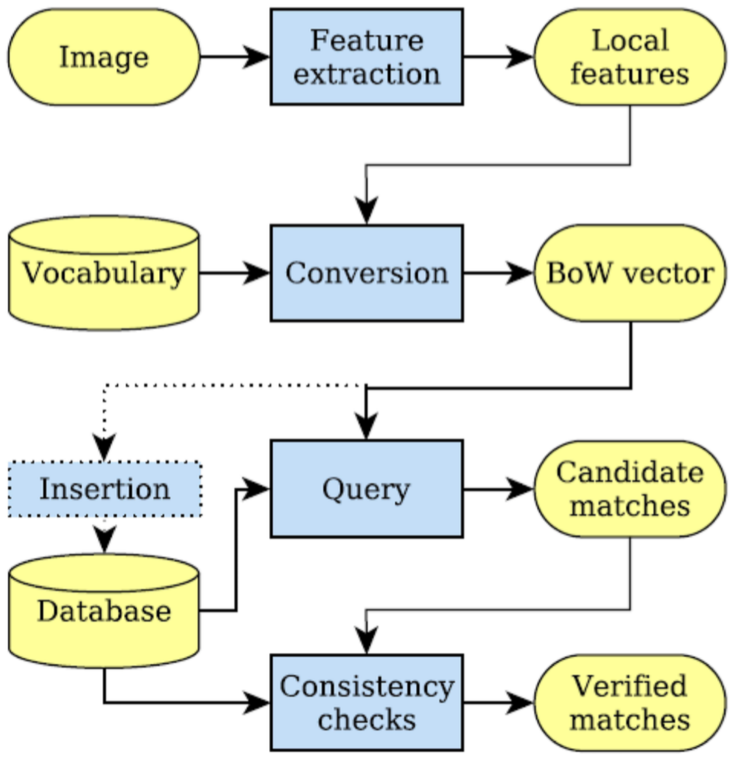
Fig. 2: Changing the relative pose $\mathbf{T}_{k,k-1}$ between the current and the previous frame implicitly moves the position of the reprojected points in the new image $\mathbf{u}'_i$. Sparse image alignment seeks to find $\mathbf{T}_{k,k-1}$ that minimizes the photometric difference between image patches corresponding to the same 3D point (blue squares). Note, in all figures, the parameters to optimize are drawn in red and the optimization cost is highlighted in blue.



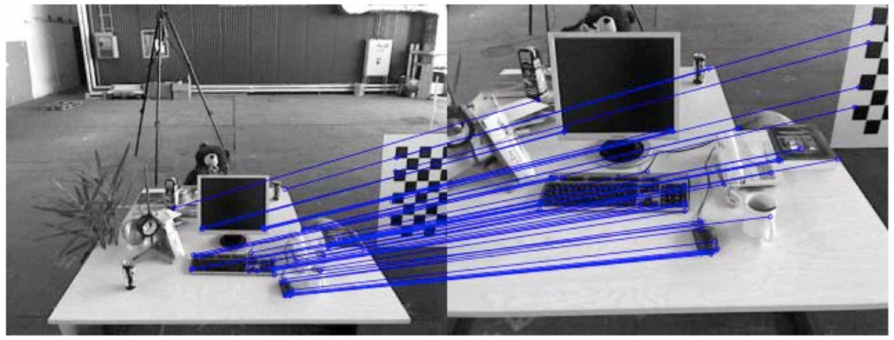
[6] E. Rosten and T. Drummond, "Machine Learning for High-Speed Corner Detection," 2006.

Proposed Solution

- Problem: tracking loss.
  - Occlusions
  - Low texture
  - Very fast or pure rotational motions.
- Solution: borrow relocalization from ORB SLAM [4, 8]
  - Create offline visual vocabulary tree [9].
  - In mapping thread, convert each keyframe to a bag of words (BoW) vector [11].
  - If tracking is lost - in tracking thread convert each frame to a BoW vector and query most similar keyframe.
  - Attempt to reinitialize SVO using current frame and closest keyframe! Again, make the same assumption of planar scene.
- Costs:
  - Features, descriptors, and BoW vectors [7, 10] must be computed for each keyframe.
- Benefits:
  - Allows system to recover from otherwise fatal tracking loss.



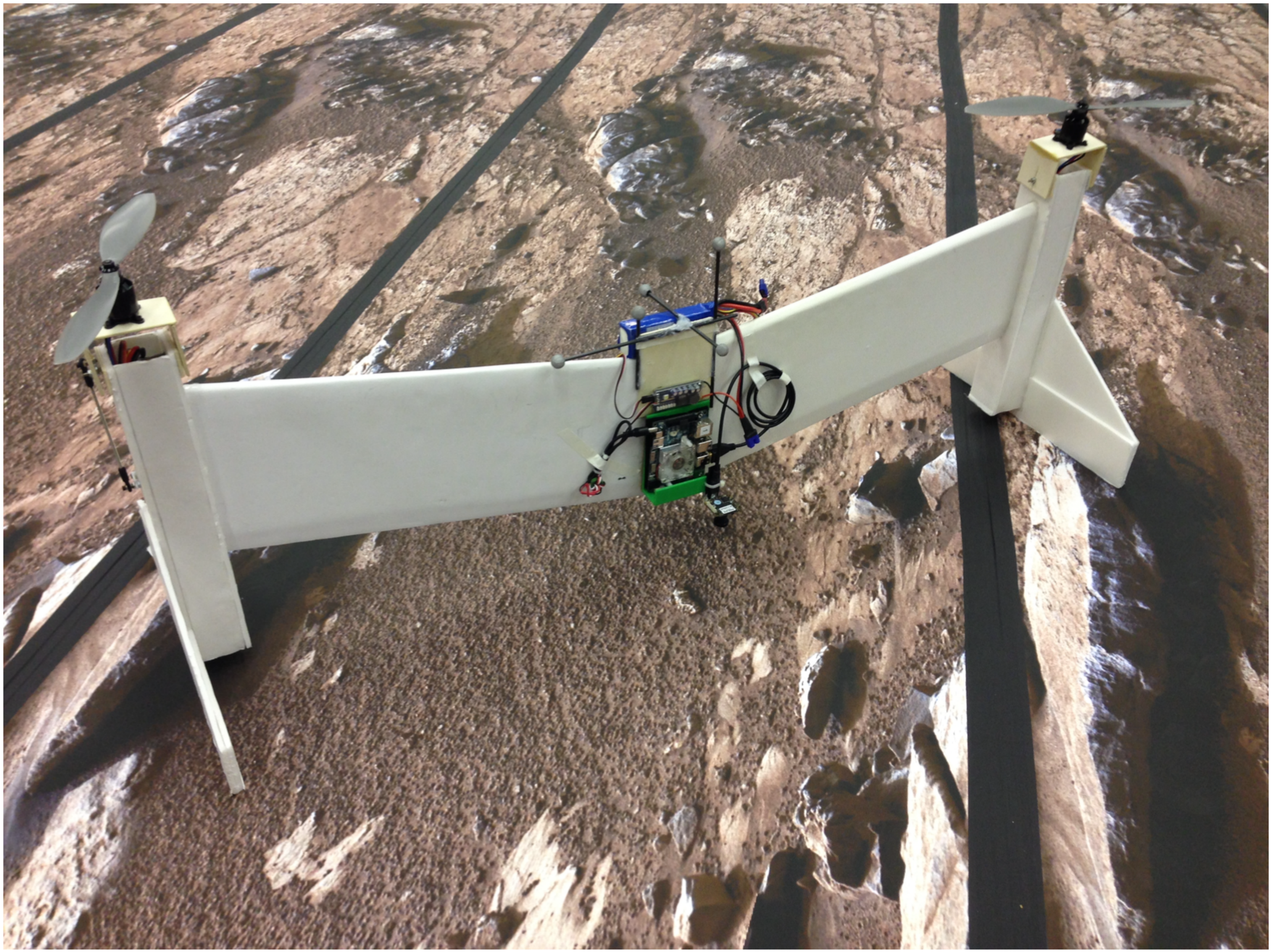
Bag of Words process depiction [10].



Bag of Words applied to loop closure [10].

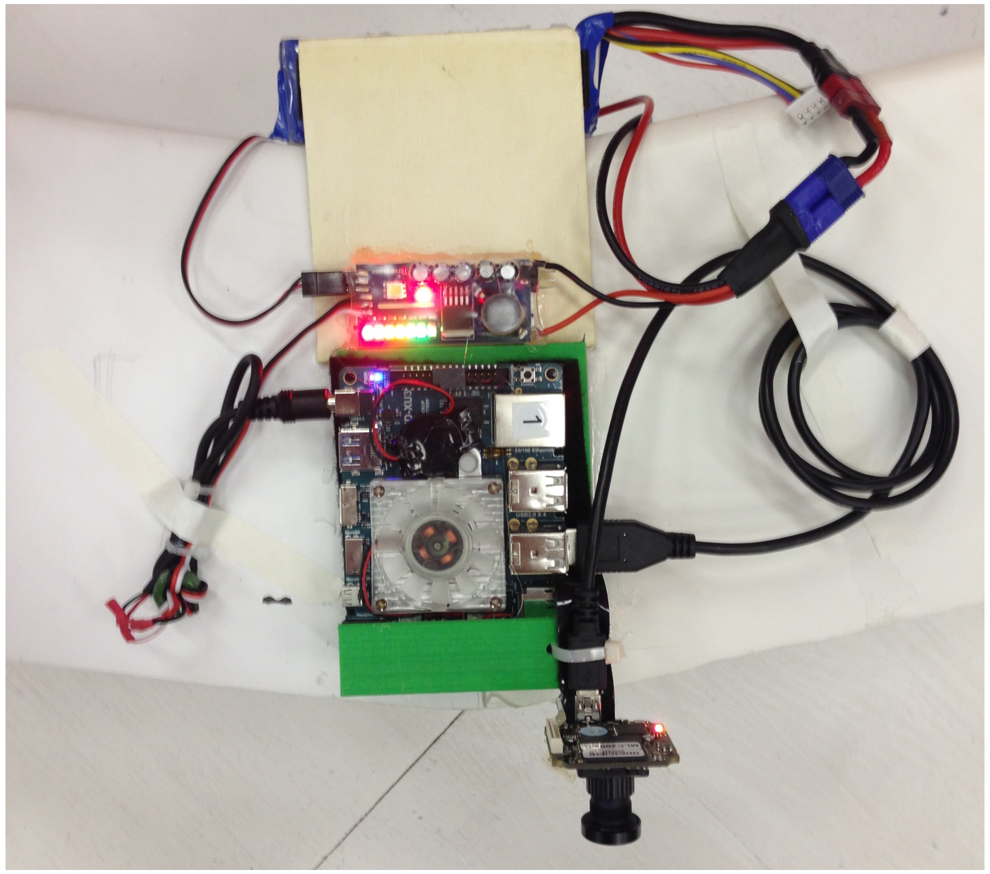
[7] M. Calendar et al, "BRIEF: Binary Robust Independent Elementary Features," 2010.
[8] R. Mur-Artal and J. Tardos, "Fast Relocalization and Loop Closing in Keyframe-Based SLAM," 2014.
[9] D. Galvez-Lopez and J. Tardos, "Real-Time Loop Detection with Bags of Binary Words," 2011.
[10] E. Rublee et al, "ORB: An efficient alternative to SIFT or SURF," 2011.
[11] J. Tardos "Feature Based Visual SLAM," 2016.

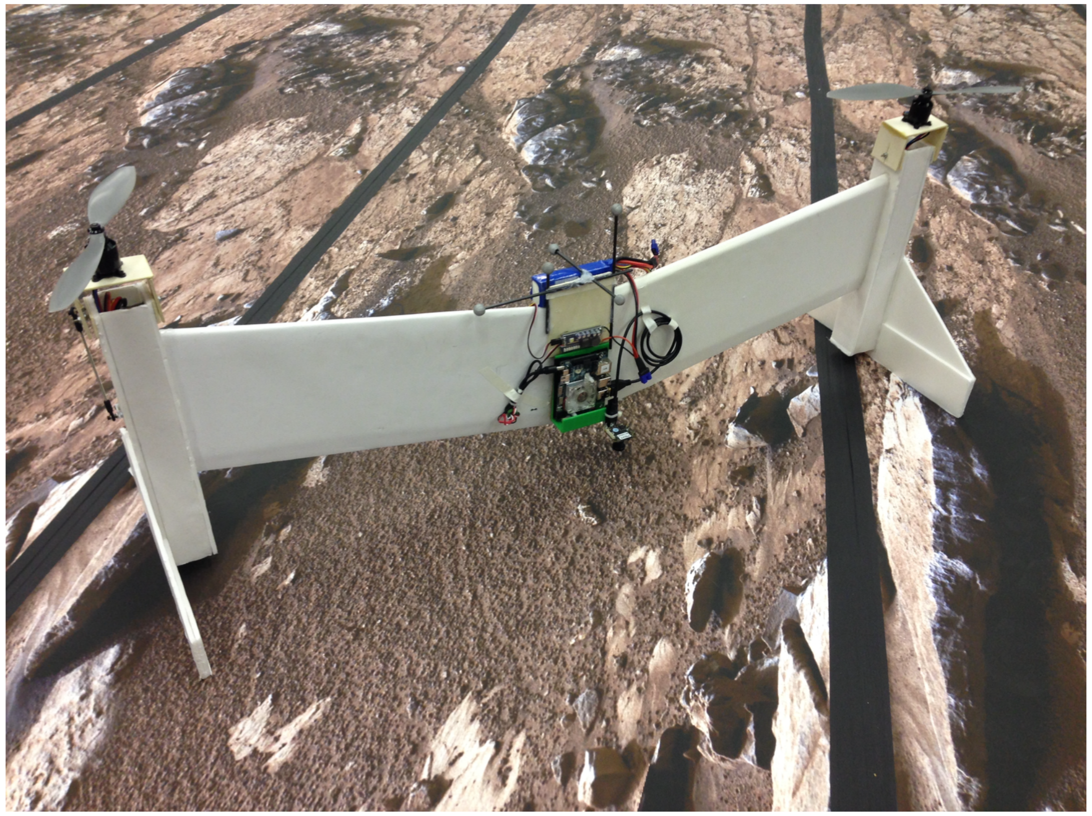
Implementation



Implementation

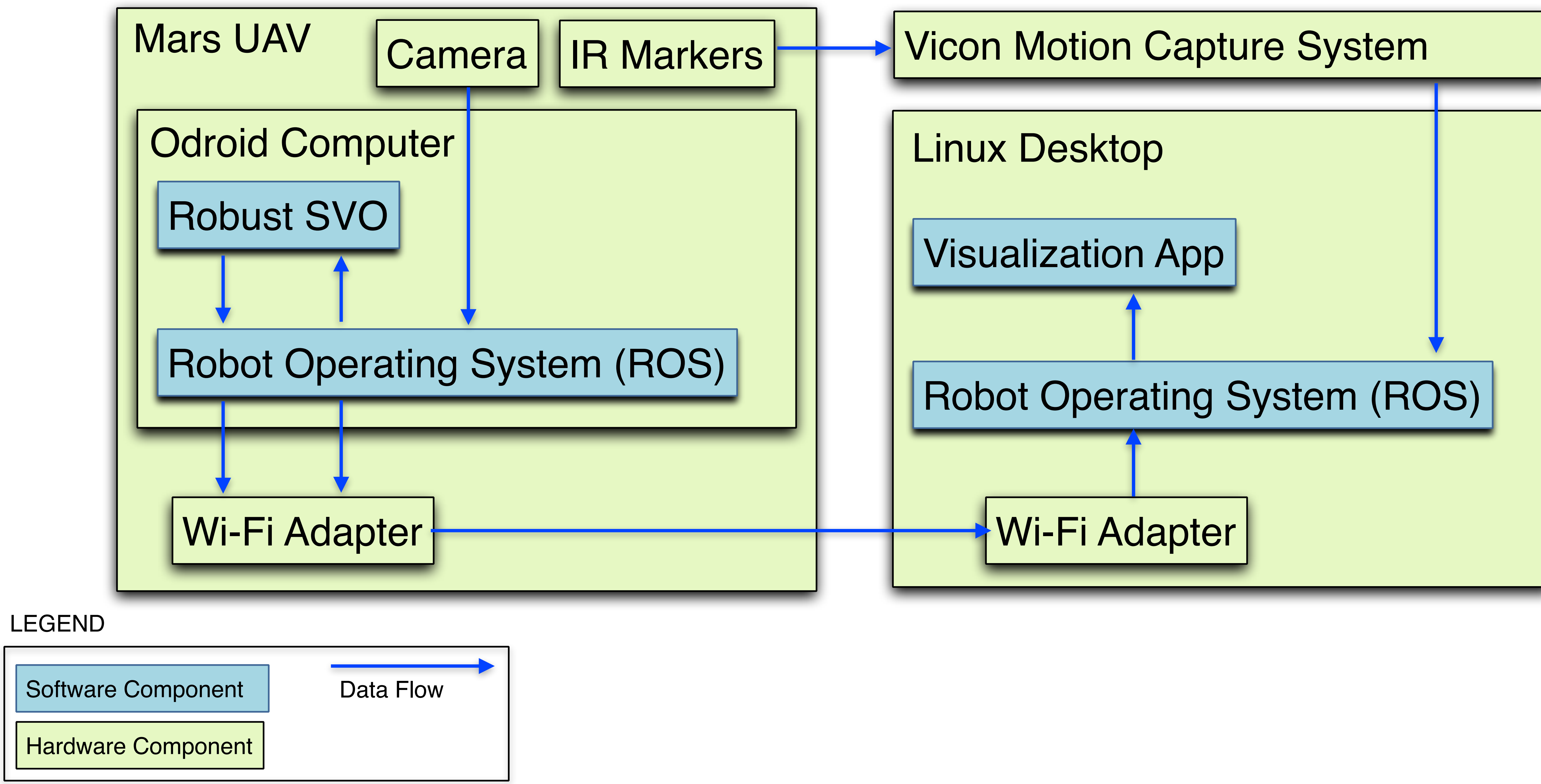
- Hardware
  - Servo-adjustable rotors
  - IR Markers
  - Lithium Polymer Battery
  - ARM 8-Core 2GB RAM Computer
  - 70 fps Global Shutter Camera
  - Wi-Fi Adapter
  - Simulated Mars surface.



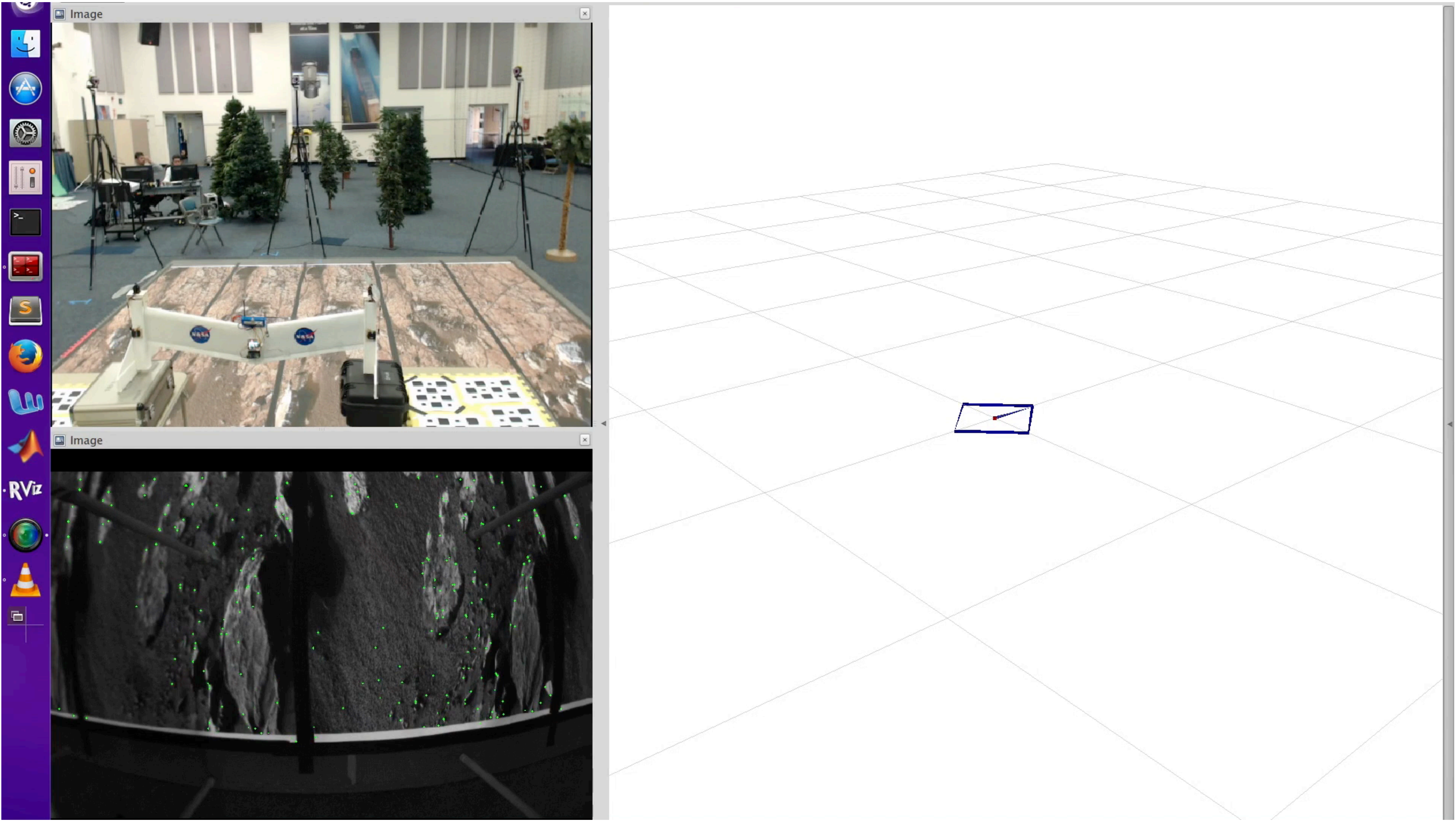


Implementation

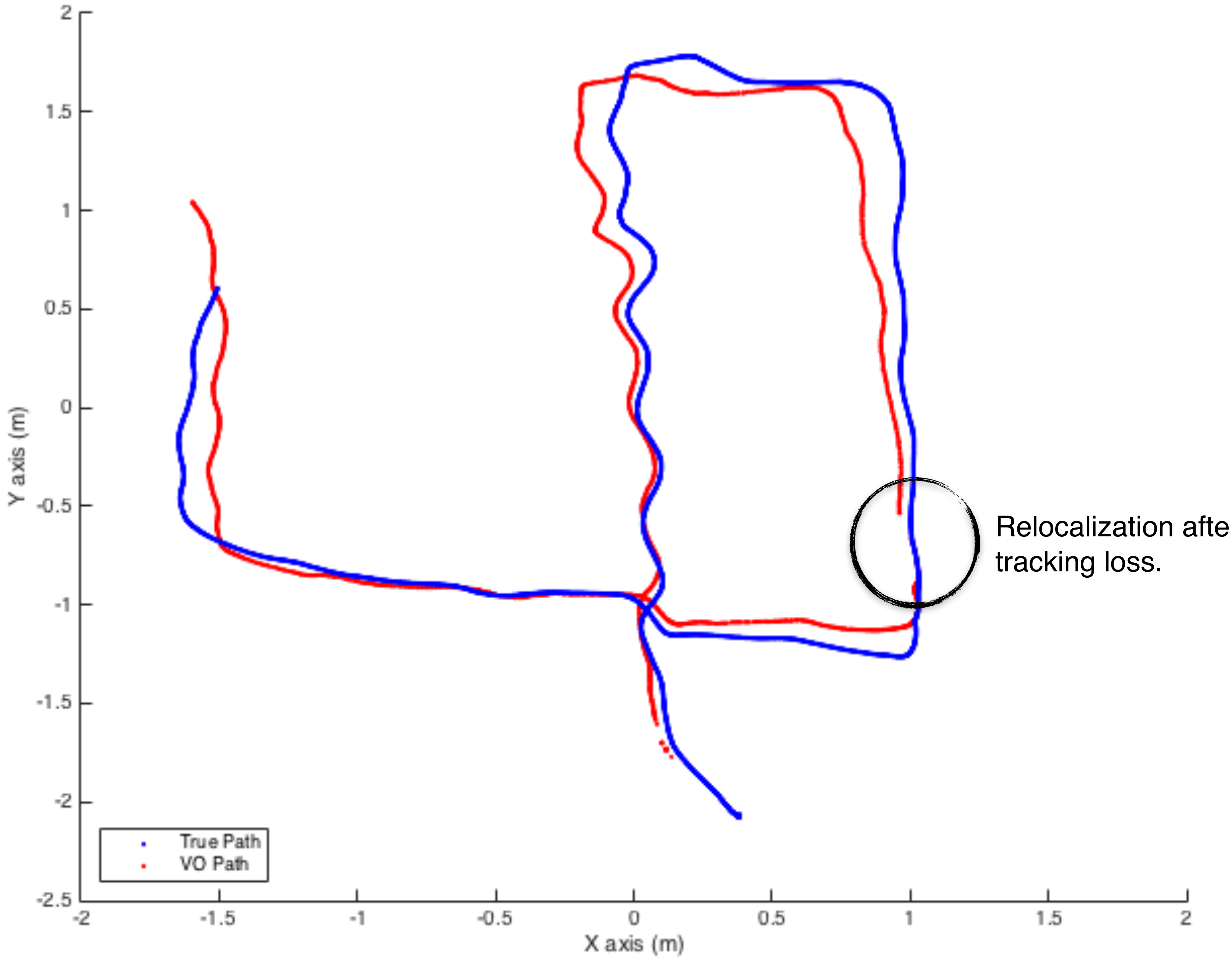
- Mars Flyer Prototype: Software
  - Ubuntu 14.04 running on vehicle.
  - Robot Operating System (ROS) for communications plumbing and live visualization.
  - Enhanced SVO with automatic relocalization.



Results

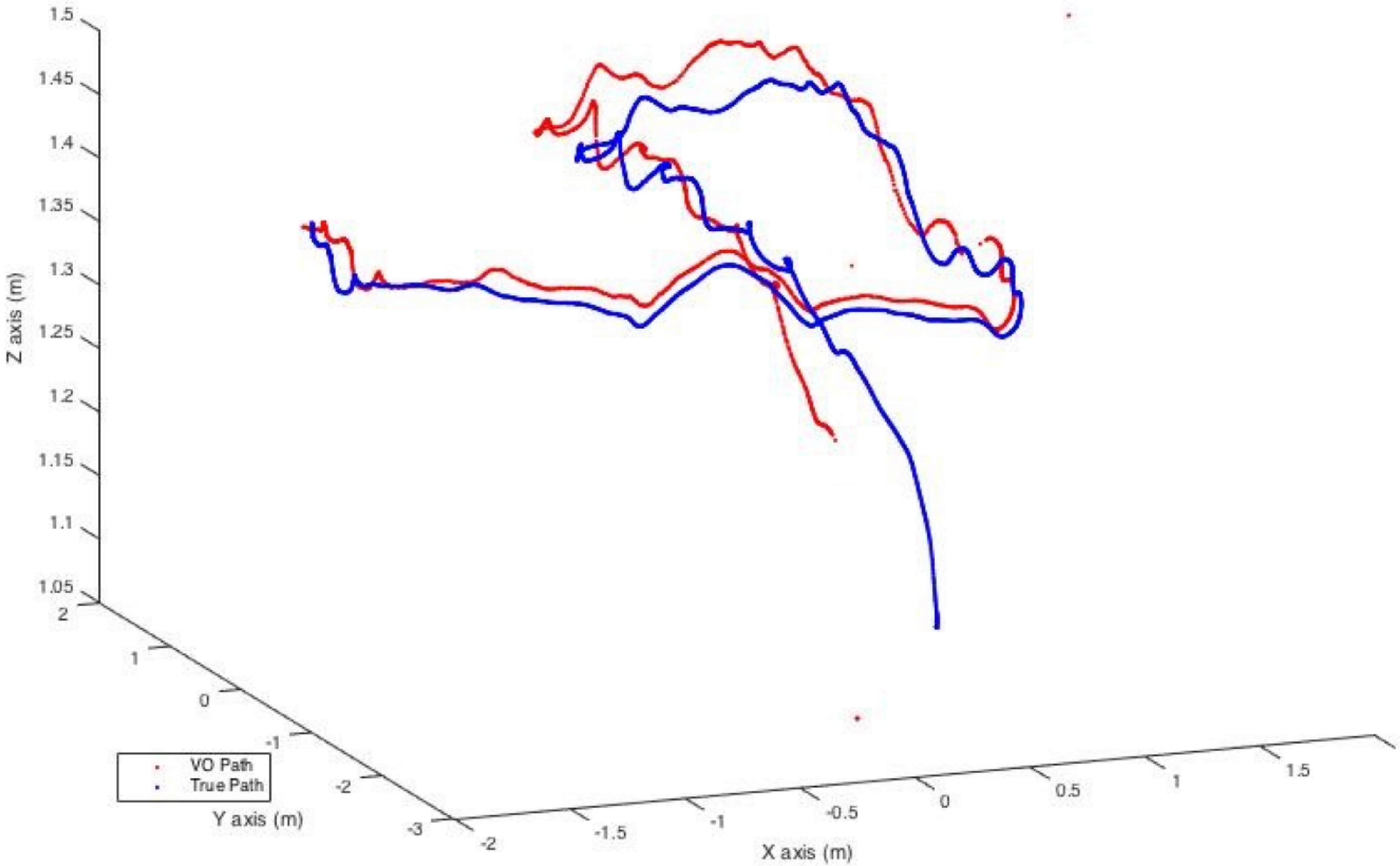


Results



True Path
VO Path
X axis (m)
Y axis (m)
Relocalization after tracking loss.

Results



1.5
1.45
1.4
1.35
1.3
1.25
1.2
1.15
1.1
1.05
Z axis (m)
2
1
0
-1
-2
-3
Y axis (m)
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
X axis (m)
VO Path
True Path

Summary & Future Work

- Results Summary:
  - Accurate: Centimeter accuracy compared to motion capture ground truth.
  - Fast: 70 Hz framerate - limited by the camera, not compute platform.
  - Robust: Survives tracking loss with relocalization.
- Future work:
  - Perform more experiments for much deeper performance evaluation.
  - Optimize software: remove some redundant work in relocalization integration.
  - Integrate IMU information into pose estimator. Lots of literature and existing software on this subject e.g. [12, 13].

[12] C. Forster et al., "IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation," 2015.
[13] S. Lynen et al., "A Robust and Modular Multi-Sensor Fusion Approach Applied to MAV Navigation," 2013.

References

[1] M. Maimone, Y. Cheng, and L. Matthies, "**Two Years of Visual Odometry on the Mars Exploration Rovers,**" In *Journal of Field Robotics,* 2007.

[2] C. Forster, M. Pizzoli, and D. Scaramuzza, "**SVO: Fast Semi-Direct Monocular Visual Odometry,**" In IEEE Intl. Conf. on Robotics and Automation (ICRA), 2014.

[3] J. Engel, T. Schops, D. Cremers, "**LSD-SLAM: Large Scale Direct Monocular SLAM,**" In 13th European Conference on Computer Vision (ECCV), 2014.

[4] R. Mur-Artal, J. Montiel, J. Tardos, "**ORB-SLAM: a Versatile and Accurate Monocular SLAM System,**" In IEEE Transactions on Robotics, Volume 31, Issue 5, 2015.

[5] G. Klein and D. Murray, "**Parallel Tracking and Mapping for Small AR Workspaces,**" In IEEE and ACM Symposium on Mixed and Augmented Reality (ISMAR), 2007.

[6] E. Rosten and T. Drummond, "**Machine Learning for High-Speed Corner Detection,**" In European Conference on Computer Vision (ECCV), 2006.

[7] M. Calendar et al., "**BRIEF: Binary Robust Independent Elementary Features,**" In European Conference on Computer Vision (ECV), 2010.

[8] R. Mur-Artal and J. Tardos, "**Fast Relocalization and Loop Closing in Keyframe-Based SLAM,**" In IEEE International Conference on Robotics and Automation (ICRA), 2014.

[9] D. Galvez-Lopez and J. Tardos, "**Real-Time Loop Detection with Bags of Binary Words,**" In IEEE International Conference on Intelligent Robots and Systems (IROS), 2011.

[10] E. Rublee et al., "**ORB: An efficient alternative to SIFT or SURF,**" In IEEE International Conference on Computer Vision (ICCV), 2011.

[11] J. Tardos "**Feature Based Visual SLAM,**" Tutorial presented at IEEE International Conference on Robotics and Automation, 2016.

[12] C. Forster et al., "**IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation,**" in IEEE Robotics: Science and Systems (RSS), 2015.

[13] S. Lynen et al., "**A Robust and Modular Multi-Sensor Fusion Approach Applied to MAV Navigation,**" in IEEE International Conference on Intelligent Robots and Systems (IROS), 2013.

Acknowledgements

- This work was funded by NASA Langley Research Center in Hampton, VA and started in NASA's Autonomy Incubator group.
- Key people to thank are research engineers Loc Tran and James Neilan for their tireless mentorship and deep expertise in computer vision, Danette Allen for her leadership of the team and for creating the opportunities for ambitious projects within NASA, and Christian Forster for his illuminating advice about the real-world deployment of SVO.