3D SPACECRAFT RECONSTRUCTION FROM ON-ORBIT IMAGERY

Daniel Crispell[1*] and Scott Richardson[1]; [1]Vision Systems, Inc. Providence, RI, USA.
*daniel.crispell@visionsystemsinc.com

Abstract. *This paper describes a system for generating digital 3D models of spacecraft from 2D imagery collected on-orbit. The 3D model represents the surface location as well as regions of uncertainty geometry due to insufficient information and thus can be used to inform optimal future collection paths. The reconstruction method is demonstrated on real on-orbit imagery of the Intelsat 10-02 commercial communications satellite collected by Northrup Grumman's MEV-2 space vehicle and released publicly via YouTube. Despite low image quality and the lack of available sensor metadata, we demonstrate the feasibility of recovering a 3D representation of the target spacecraft.*

Introduction. Space situational awareness (SSA) is dependent on large scale cataloging and in-depth understanding of select space objects. The 3D geometry of a space object can inform its function, capabilities, and aid in damage assessment. Manual construction of 3D models is feasible but requires a manually intensive and time-consuming process. One of the main challenges of 3D reconstruction of space-based objects is the large variation in observed intensity with respect to lighting conditions and viewpoint. The former is mainly caused by the lack of atmosphere and results in very dark shadow regions and very bright areas under direct illumination from the sun. The former is primarily a result of very specular ("shiny") surfaces typical on visible spectrum images of spacecraft (foil, solar panels, metal). Despite these challenges, recent work [1] has demonstrated the feasibility of generating novel views and extracting depth information of synthetic renderings of satellite models using neural radiance fields (NeRF) [2]. In this work, we demonstrate that a joint system composed of neural radiance fields and a novel volumetric fusion can generate 3D models from real space-based imagery. The resulting representation explicitly represents regions of uncertain geometry, a critical requirement when limited amounts of input viewpoints are available.

Source Data. Image data for the experiment was retrieved from a publicly released promotional video of the Intelsat 10-02 commercial communications satellite based on real footage captured by Northrop Grumman's MEV-2 space vehicle in April 2021 during a successful docking operation [3]. The video includes a good diversity of viewpoints of Intelsat 10-02, making it a good candidate for 3D reconstruction. The video contains "burnt in" metadata indicating range in meters to the target, which is used to constrain the scale of the 3D reconstruction. An example frame is shown in Figure 1.

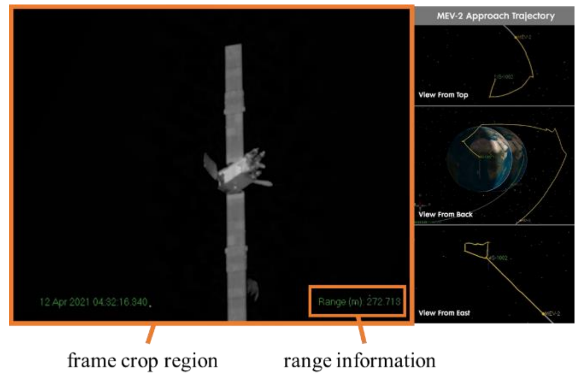

Figure 1. The relevant imagery was cropped from the MEV-2 demonstration video, and approximate range information embedded in the imagery used to determine metric scale.

Camera Estimation. Other than the coarse range information, the video does not include any metadata regarding the relative sensor position, orientation, or intrinsic properties (e.g. focal length), all of which are required for 3D reconstruction. We automatically estimate this information using "structure from motion" (SfM) techniques [4]. The first step in the SfM pipeline is the selection of keypoints in each video frame and matching of the keypoints between frames as shown in Figure 2. Next, the 3D positions of each matched keypoint and the sensor metadata for each image are jointly optimized in a process known as "bundle adjustment".

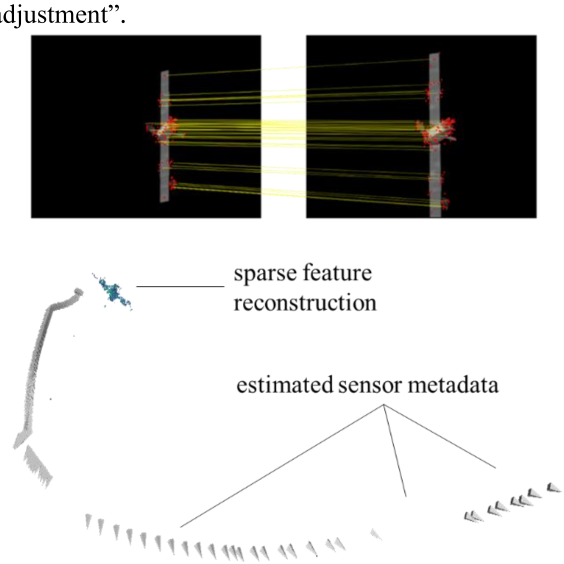

Figure 2. Sparse feature matches (top) are used to estimate sensor metadata and a sparse 3D reconstruction (bottom).

3D Modeling with Uncertainty. Several recent breakthrough 3D reconstruction methods have been published in the computer vision literature under the heading of "Neural Radiance Fields" (NeRF) [2]. These methods leverage a differentiable rendering framework based on a volumetric representation but do not explicitly represent surface probabilities and appearance models. Instead, a neural network is trained to predict surface probability and appearance given a 3D point as input. NeRF methods excel at reproducing images containing complex reflectance functions and have been shown to work well on simulated images of satellites [1]. We observed that baseline NeRF methods suffer considerably when target resolution varies significantly throughout the input imagery, a common feature of "approach" sequences. The recently proposed Mip-NeRF [5] method aims to solve this problem, and indeed produced high-quality results in our experiments. Given an input set of images, Mip-NeRF generates predicted image and depth values given a sensor position as shown in Figure 3.

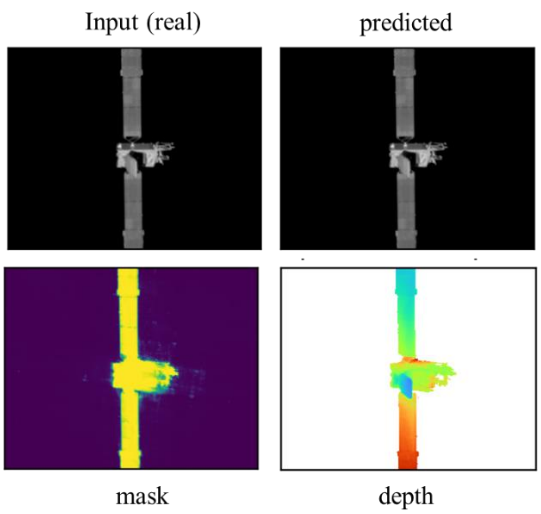

Figure 3. Mip-NeRF reconstructs the input images as well as estimated per-pixel depth values.

Mip-NeRF estimates depth maps for surfaces with good coverage in the input data but is not designed to explicitly predict or represent uncertainty. To mitigate this limitation, we developed a volumetric fusion method that identifies surface regions of high uncertainty. The volumetric fusion algorithm represents an occupancy probability at each point in a regular 3D voxel grid that encompasses the bounds of the target object. Two such occupancy models are simultaneously optimized: a "maximal" and a "minimal" model. The same algorithm is used to optimize both, but a low prior probability is used for the minimal model, and a high prior probability is used for the maximal model. Unobserved regions will tend to converge to "empty" (low occupancy probability) in the minimal model, and "occupied" (high occupancy probability) in the maximal model. Once convergence is reached, surface meshes are extracted for both models using the standard "marching cubes" method. For surfaces with high certainty, the maximal and minimal models will be in close agreement, as the prior has little effect when presented with overwhelming evidence. For uncertain regions, the maximal surface will lie far outside the minimal surface, indicating that more observations are needed. The resulting 3D model is shown below in Figure 4, with model regions containing high uncertainty shown in red.

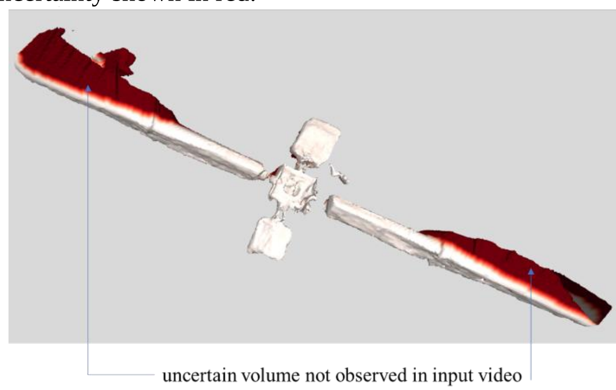

Figure 4. The 3D model reconstructed from the MEV-2 video contains regions of uncertainty (shown in red) in regions not directly observed in the input video sequence.

Conclusion. Despite the challenges presented by the low-quality video, the MEV-2 experiment provides a valuable demonstration of the proposed 3D reconstruction technique operating successfully on real imagery captured in space. Future work will expand on this system by integrating "best next view" estimation for minimizing uncertainty in active sensing scenarios.

References.

[1] A. Mergy, G. Lecuyer, D. Derksen, and D. Izzo, "Vision-based Neural Scene Representations for Spacecraft," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Nashville, TN, Jun. 2021.

[2] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis," 2020.

[3] "Northrop Grumman and Intelsat Make History with Docking of Second Mission Extension Vehicle to Extend Life of Satellite," https://news.northropgrumman.com/news/releases/northrop-grumman-and-intelsat-make-history-with-docking-of-second-mission-extension-vehicle-to-extend-life-of-satellite

[4] J. L. Schönberger and J.-M. Frahm, "Structure-from-Motion Revisited," presented at the Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[5] J. T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, and P. P. Srinivasan, "Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields," *arXiv:2103.13415 [cs]*, May 2021.

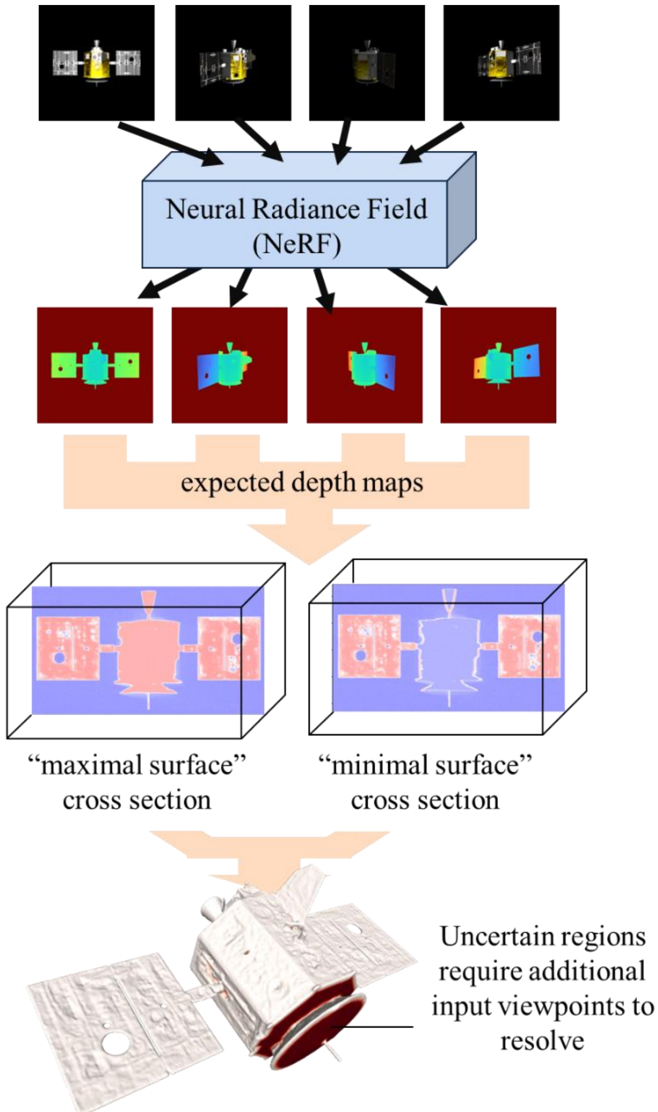

Figure 5. 3D Reconstruction Method Overview: The proposed method generates a 3D model with explicit geometric uncertainty using NeRF-generated depth maps followed by optimization of a pair of volumetric occupancy probability models.