APPLYING NEURAL RADIANCE FIELDS TO ASTEROID SHAPE MODELING

Matthew W. Givens; Advanced Space, 1400 W 122nd Ave, Westminster, CO 80234
matthew.givens@advancedspace.com

Abstract. *Neural Radiance Fields (NeRFs) are a promising new machine learning-driven computer vision technology that may have many future applications in spaceflight problems. This paper provides a short background of the theory and assumptions behind NeRFs and how they can be used for 3D view synthesis and shape reconstruction, then applies a popular NeRF algorithm to the problem of asteroid shape modeling using real image and trajectory data of the asteroid Bennu. Results are compared to the shape model generated by the OSIRIS-REx mission.*

Introduction. Since the publication of the originating paper in 2020,[1] one of the most active new topics in computer vision research has been Neural Radiance Fields (NeRFs), also known as "neural rendering techniques", which are capable of generating *implicit*, highly photo-realistic, novel views of complex, non-Lambertian 3D environments after being trained on a set of 2D images with known poses. Once trained, the model can be sampled at any input pose to yield local color and volume density outputs which are then used to generate a novel image via fast volume rendering techniques. Since then, hundreds of new techniques have been devised to improve on the original formulation and extended it to new use cases and applications including scene relighting, video synthesis, pose estimation, and more.[2]

Neural Radiance Fields. In contrast to classical, Structure from Motion (SfM) techniques for reconstructing a 3D scene from a sequence of 2D images, NeRF algorithms do not encapsulate the 3D geometry of the scene in terms of *explicit* landmarks, control points, or point clouds. Instead, the environment is represented using a field that is approximated by a neural network, or by a plenoxel grid,[3] which describes the color and volume density of every point and every viewing direction in the scene. At a high level, the NeRF can be understood mathematically as

$$F(\mathbf{x}, \theta, \phi) \rightarrow (c, \sigma)$$

which is a mapping between position vector $\mathbf{x}$ and azimuth and elevation angles θ and ϕ respectively (or, equivalently, a 3D unit vector) to a space of color c and volume density σ. The volume density, which captures how much radiance is accumulated by a given ray in space, is constrained to be independent of viewing direction while the color can depend on both the viewing direction and position. This function is approximated using one or more Multi-Layer Perceptrons (MLPs) resulting in an estimated field F_Θ.

To render a novel image R using a NeRF, the pixels of the query image are turned into rays in space parameterized by

$$\mathbf{r} = \mathbf{o} + t_i \mathbf{d}(\theta, \phi)$$

where $\mathbf{o}$ is the camera center. The field can then be sampled distinct distances t_i along direction vector $\mathbf{d}$. The (c_i, σ_i) samples are then composited together into a single color output C via numerical quadrature of the volume rendering equation,

$$C = \sum_i \exp\left(-\sum_{j<i} \sigma_j \delta_j\right)(1 - \exp(-\sigma_i \delta_i))c_i$$

where $\delta = t_i - t_{i-1}$ is the distance between successive points along the ray. During training, the loss

$$L = \sum_{r \in R} \left\| \hat{C} - C_{truth} \right\|_2^2$$

is computed and minimized so that the field output matches the ground truth training images. The position and direction inputs are typically passed through a "positional encoding" in order to better capture higher-frequency information and render higher quality images. The details of this and other aspects of the pipeline vary wildly between extant implementations and variations.

NeRFs for Asteroid Shape Modeling. The problem of turning a sequence of images into a 3D shape model is encountered in spaceflight research in a number of areas, notably Rendezvous and Proximity Operations (RPO) and asteroid shape modeling. While neural rendering techniques are beginning to be explored for the former application,[4] to the knowledge of the author they have never been applied to the latter application up to this point.

Many aerospace authors have pointed out that the asteroid shape modeling problem presents very challenging conditions for vision-based relative navigation and SfM pipelines.[5] Spacecraft safety considerations as well as the extreme low gravity environments of many asteroids result in spacecraft approach and orbital velocities on the order of centimeters per second, making the true parallax between successive images very small. Meanwhile, the asteroid is typically rotating on its axis, causing large changes in lighting conditions and visual feature appearance over significantly shorter time intervals. These factors cause many standard SfM algorithms, typically based on visual feature matching techniques, to fail.

The challenge of varying lighting conditions is one that prior NeRF authors have approached in the context of "unconstrained photo collections" derived from internet searches. Images taken from different cameras at different times of day with different weather conditions and possible occlusions can result in subpar performance in the original NeRF framework which relied on a "static environment" assumption. To address this drawback, "NeRF-in-the-Wild"[6] added the capability of learning a latent "appearance embedding" for each image that captures photometric and environmental variations between images found "in the wild". This approach effectively decouples static and transient components of the scene into different MLPs and was demonstrated by reconstructing large outdoor scenes. The methodology has since been incorporated into many subsequent NeRF algorithms.

Another challenge of the shape modeling problem is that, as outlined previously, NeRF algorithms do not natively produce an explicit mesh or "shape model" of the environment. They simply create an input-output map that captures the scene implicitly. Converting this output into a more explicit representation such as a mesh or voxel grid is a growing area of NeRF research due to its potential to provide data products useful for applications outside of novel view synthesis. The original NeRF authors proposed the use of the "marching cubes" algorithm for this purpose and more recent authors have taken advantage of signed distance functions (SDFs) for improved surface recovery and rendering speed.[7]

Algorithm Selection. Given the aforementioned requirements on a NeRF algorithm appropriate for the asteroid shape modeling task, we sought an algorithm that included both appearance embedding and high quality mesh generation. NVIDIA's Instant-NGP, released in 2022,[8] provides both of these tools and the added benefit of extremely fast, parallelizable training by leveraging a smaller neural network enabled by a novel multiresolution hash table input encoding. Instant-NGP is also capable of optimizing the pose of the camera given an initial set of estimated poses as well as camera calibration parameters and more. The official version of the code is available through NVIDIA but an open source implementation of Instant-NGP is available in the Nerfstudio project.[9]

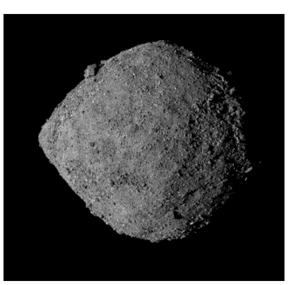


Figure 1. Example MAPCam image from OSIRIS-REx image sequence.

In 2023, NVIDIA also released Neuralangelo[10], an algorithm with the stated goal of creating high-quality shape models from a neural radiance field. These two algorithms were assessed for the asteroid problem in this work.

Bennu Dataset. The OSIRIS-REx spacecraft arrived at the 500-meter wide asteroid Bennu in December of 2018. It carried the OCAMS camera suite,[11] a three-camera package to observe and map the asteroid across varying distances and resolutions over the mission lifetime. The raw data from these cameras is available on the mission's node on the Planetary Data Sciences website. From these sources, we obtained a sequence of 37 images taken by the spacecraft's MAPCam between 00:47:30 UTC and 05:04:57 UTC on December 13, 2018 which provide medium-resolution coverage of the asteroid over a full rotation period. One such image is shown in Figure 1. The MAPCam has a focal length of 125 mm and the image sensor has a pixel pitch of 8.5μm.

Instead of using a software like COLMAP to obtain the spacecraft poses, which is typical in NeRF pipelines but turns out to be problematic in this scenario, we utilized the estimated location and attitude of the spacecraft at the image times from the mission SPICE kernels available online from NASA's Navigation and Ancillary Information Facility (NAIF) website. These resources were used to create a suitable dataset in the IAU Bennu-fixed frame to be processed by Instant-NGP. The resulting image sequence is plotted in Figure 2.

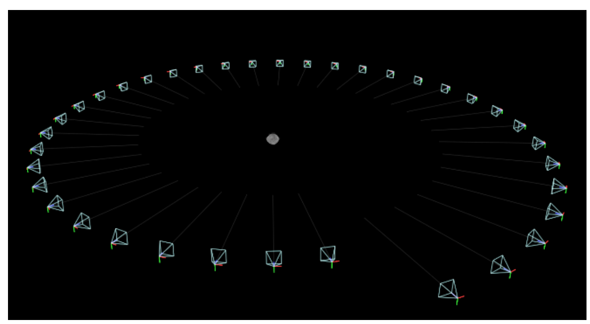
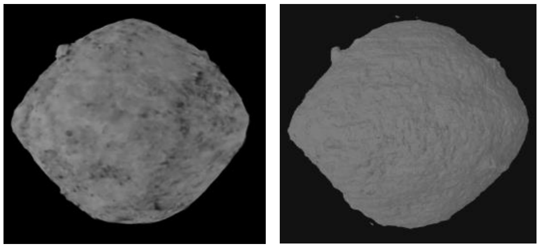


Figure 2. Camera pose trajectory in Bennu body-fixed frame

Results. Initial results of Instant-NGP on the Bennu dataset are promising. With an appearance embedding of the same size as the image sequence length, the algorithm successfully recovers the shape of the asteroid in 3D after training for less than five minutes on a Windows laptop with an RTX4000. Novel views are generated at 30 frames per second. One of these, as well as a shape model derived from the field, are shown in Figure 3. Due to the changing lighting conditions across the image sequence, the lighting in the synthesized image is not representative of a particular lighting condition or sun direction. Instead, it is a synthesis of the lit geometry in all the images, with each having an estimated lighting direction. The shape model in Figure 3b shows that the 3D geometry of the scene is being accurately captured by the field, albeit with some artifacts present. It is likely that these results can be improved with further hyperparameter tuning and/or additional input images.



a) b)

Figure 3. a) A novel image of asteroid Bennu and b) a derived shape model generated by NVIDIA's Instant-NGP after 3 minutes of training on a laptop.

We assessed the accuracy of the reconstructed shape by comparing it to one of the shape models generated by the OSIRIS-REx mission. This is done efficiently by building a KD-tree of the vertices of the reference shape model and then computing the average Cartesian position difference to the k-nearest neighbors to each vertex of the NeRF shape model, where in this case k=3. These error points are then parameterized by latitude and longitude and then binned so that a 2D contour plot of the position errors can be generated. After scaling the result to real-world units, we obtain Figure 4. Note that the error is computed as truth minus NeRF. Given that the asteroid is approximately 500 meters across, the errors are generally on the order of 2% or less with a median value of -1.049 meters. The camera is about 10 kilometers from the surface at any given time. There does not seem to be a general trend in these errors in latitude or longitude but pockets of high negative errors are evident across the map.

Similarly, Figure 6 and 7 showcase the analogous results derived using Neuralangelo. Training for Neuralangelo took significantly longer than Instant-NGP, on the order of 18 hours on a Linux machine with a single Tesla T4. However, the visual appearance of the shape model is much more detailed than the shape derived from Instant-NGP, and the errors in Figure 7 are appreciably lower. The median value of the errors from Neuralangelo is -0.4554 meters and fewer concentrated regions of significantly negative errors are visible.

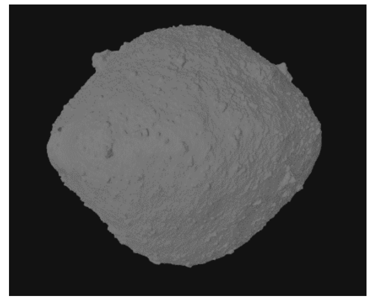


Figure 6. Image of shape model generated by NVIDIA's Neuralangelo.

Conclusion. Neural radiance fields and related deep learning technologies are a rapidly expanding field that deserves the attention of aerospace vision professionals. Using some available algorithms, we have shown that these algorithms, in their current state, are capable of providing good performance on the challenging asteroid shape modeling task with a limited set of images. While the computational requirements for these algorithms is likely higher than current onboard computing resources permit, opportunities for lowering these requirements will likely manifest in future NeRF algorithms in parallel with forthcoming improvements in available onboard computing hardware.

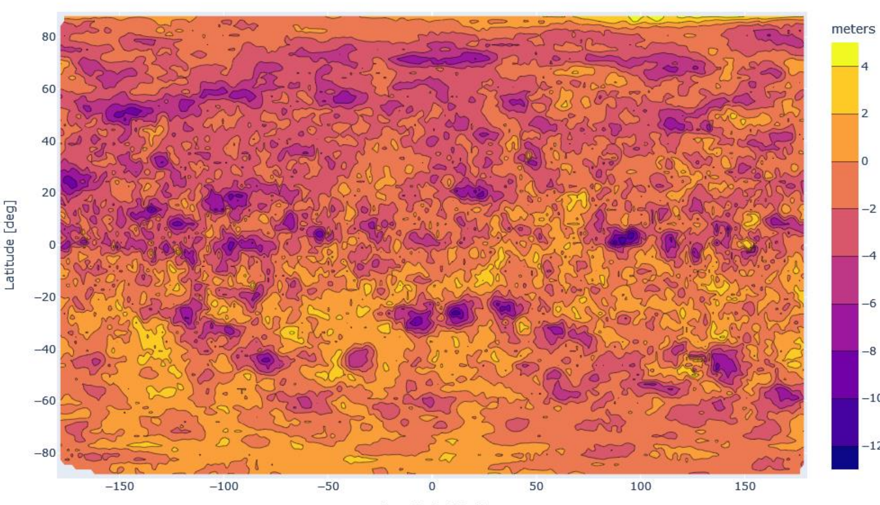


Figure 5. Shape errors computed with Instant-NGP with respect to Bennu_v20_200k.obj from OSIRIS-REx

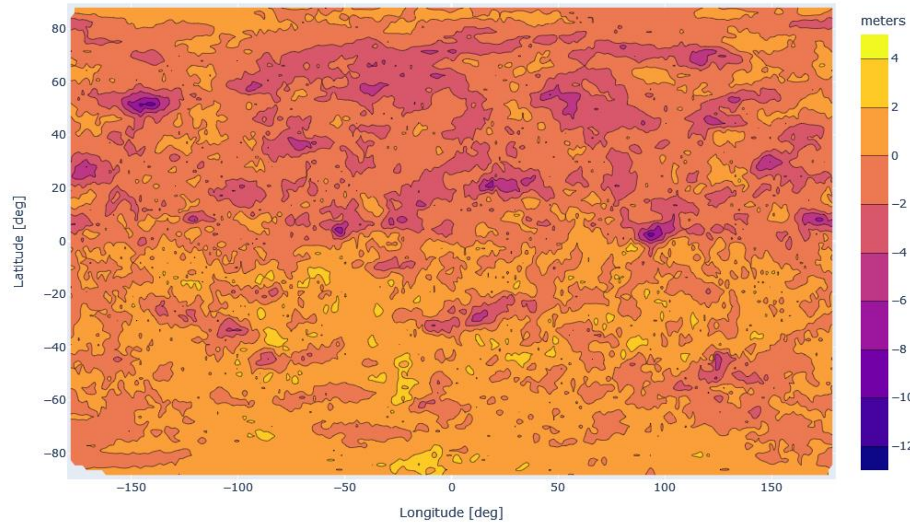


Figure 7. Shape errors computed with Neuralangelo with respect to Bennu_v20_200k.obj from OSIRIS-REx

References.

[1] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99-106.

[2] Gao, K., Gao, Y., He, H., Lu, D., Xu, L., & Li, J. (2022). NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review. ArXiv, 10.48550/arXiv.2210.00379

[3] Fridovich-Keil, S., Yu, A., Tancik, M., Chen, Q., Recht, B., & Kanazawa, A. (2022). Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5501-5510). 10.1109/CVPR52688.2022.00542.

[4] Mergy, A., Lecuyer, G., Derksen, D., & Izzo, D. (2021). Vision-based neural scene representations for spacecraft. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2002-2011).

[5] Villa, J. McMahon, J., & Nesnas, I. (2022). Robust Landmark and Hazard Detection on Small Body Surfaces Using Shadow Imagery. In *2022 AAS Astrodynamics Specialist Conference.*

[6] Martin-Brualla, R., Radwan, N., Sajjadi, M. S., Barron, J. T., Dosovitskiy, A., & Duckworth, D. (2021). Nerf in the wild: Neural radiance fields for unconstrained photo collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7210-7219).

[7] Yariv, L., Hedman, P., Reiser, C., Verbin, D., Srinivasan, P. P., Szeliski, R., & Mildenhall, B. (2023, July). Bakedsdf: Meshing neural sdfs for real-time view synthesis. In *ACM SIGGRAPH 2023 Conference Proceedings* (pp. 1-9).

[8] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)*, *41*(4), 1-15.

[9] Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Kerr, J., Wang, T., Kristoffersen, A., Austin, J., Salahi, K., Ahuja, A., McAllister D., & Kanazawa, A. (2023) Nerfstudio: A Modular Framework for Neural Radiance Field Development. In *ACM SIGGRAPH 2023 Conference Proceedings*.

[10] Li, Z., Müller, T., Evans, A., Taylor, R. H., Unberath, M., Liu, M. Y., & Lin, C. H. (2023). Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8456-8465).

[11] Edmundson, K. L., Becker, K. J., Becker, T. L., Bennett, C. A., DellaGiustina, D. N., Golish, D. R., ... & Lauretta, D. S. (2020). Photogrammetric processing of Osiris-Rex images of asteroid (101955) Bennu. *IsPRs Annals of the Photogrammetry, Remote sensing and spatial Information sciences*, *3*, 587-594.