3D Ground Truth For Simulating Rover Imagery 3D Vision Testing

Gerhard Paar[1], Christoph Traxler[2], Piluca Caballo Perucha[1], Arnold Bauer[1], Manfred Klopschitz[1], Laura Fritz[2], Rebecca Nowak[2], and Jorge Ocón Alonso[3]

[1]JOANNEUM RESEARCH Forschungsgesellschaft mbH, DIGITAL - MVA, Graz, Austria (gerhard.paar@joanneum.at)

[2]VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, Vienna, Austria (traxler@vrvis.at)

[3]GMV Aerospace and Defence, Madrid, Spain (jocon@gmv.com)

1   Introduction & Scope

3D vision (mapping, localization, navigation, science target recognition etc.) using Planetary Rover imaging requires high-level test assets, including a "ground truth" against which the processing results (rover locations, 3D maps) can be compared. Whilst end-to-end simulation for functional testing is realized by visualization of a drone-based DTM (Digital Terrain Model), the accuracy and robustness of vision-based navigation and 3D mapping can only be verified by high-fidelity data sets. The approach followed in the EU Horizon-2020 Project ADE [3] used a terrestrial-captured image data set for high- and medium resolution (2mm / 3dm grid size) DTM generation of a representative Mars-analog environment, followed by batch rendering to be presented to the respective 3D vision components (Visual Odometry – VO, and stereovision-based point cloud generation):

- Capturing terrestrial & drone-based images for photogrammetric reconstruction using ground control points (GCPs)
- COTS (Commercial-Off-The-Shelf) compilation of 3D textured models in different resolutions using Structure-from-Motion (SfM)
- Fusion of the gained textured point clouds in the visualization component PRo3D, and batch rendering of simulated stereo images at poses along a Rover trajectory
- Using these images to validate / evaluate vision-based navigation and mapping frameworks.

Figure 1: clockwise from top left: Ground Control Point; dGPS measurement of GCP; Example of an input image; registered camera poses in CapturingReality tool

2   Ground Truth 3D Data Assembly

Several images (6000x4000 pixels, Figure 1 bottom-left) on a test site at the Canary Islands were taken with a SONY ILCE-6500 digital camera. For DTM geo-referencing, 4 points were dGPS-measured (Figure 1, top). The reconstruction tool used was CapturingReality, using the following workflow:

- Import (a total of 1754) images and Ground Control Points (GCPs).
- Geo-registration: Identified GCPs on the images.
- Image alignment/registration. Find matching points, calculate camera poses and internal geometry, and a sparse 3D point cloud (Figure 1, bottom-right).
- Model generation for a complete dense 3D mesh in high quality.
- Texture mapping on the dense 3D model.
- Ortho image and DSM computation (49948 x 30768 pixels) with spatial resolution of 2mm in epsg:32628-WGS 84/UTM zone 28N coordinate system.

A similar technique was used for an image sequence captured by a drone flown by ESA in medium height, gaining a DTM with 3cm resolution. Without GCPs, a co-registration was possible within the following visualization component (see next section). See the various 3D data sets on Figure 2.

Figure 2: clockwise from top left: Drone DTM / Ortho (diameter ~800m) / Fusion of textured 5mm and 2mm DTM from terrestrial image sequence (~50m, 3D view) / 2mm resolution DTM from terrestrial image sequence (~20m) / respective 2mm Ortho image.

3   Simulation with PRo3D

PRo3D [1] allows planetary scientists fluent navigation through a detailed geospatial environment with a visual experience close to field work. It offers much of the required functionality to generate large volumes of reference images with a fidelity close to field captures and a clear relation to the geometry of the terrain and the capturing position. The fidelity is achieved by a high resolution of the DTMs and its image textures. Combination of textured point clouds (OPCs – Ordered Point Clouds) in various resolutions is easily possible by superimposing very high resolution data sets onto global medium / low resolution data sets (Figure 3). PRo3D's batch rendering enables mass-production of such ground truth data by defining many different viewpoints from which to render different types of images, including camera animations between such viewpoints.
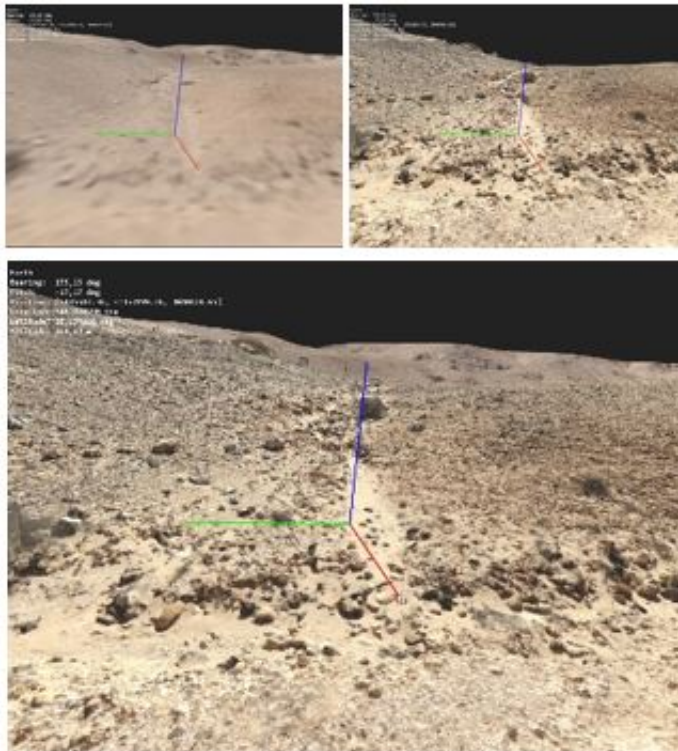
Figure 3: Rendered reference images. Top left: Using UAV 3cm DTM/Ortho only. Top right: 2mm DTM/Ortho from terrestrial images only. Bottom: UAV 3cm and terrestrial 2mm DTM/Ortho combined.

4   Testing the approach using a customer application

The demonstrated approach allows to verify navigation and reconstruction algorithms with simulated imagery. The DTMs and navigation poses resulting from processing this imagery can be directly and accurately compared with the original DTM & image poses from which the images were rendered. Thus, deviations between the ground truth and the resulting reconstruction as well as image poses can be more efficiently investigated.

To test an application, 101 stereo image pairs along a straight 10m path were rendered and fed into ORB-SLAM [6] (https://github.com/Phylliida/orbslam-windows) VO with satisfactory results (Figure 4).
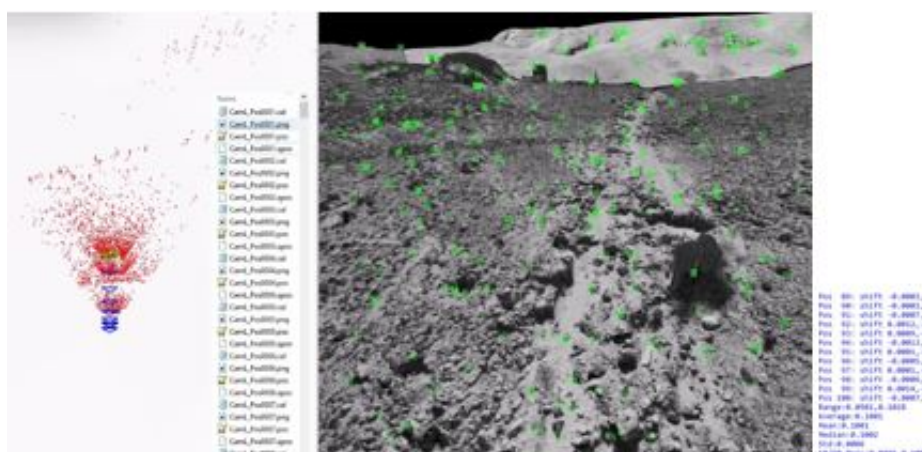

Figure 4: VO Accuracy assessment. Middle: Simulated image with VO feature points. Left: Sparse reconstructed point cloud. Right: Output of the process with camera coordinates gained, for comparison with ground truth.

5   Conclusion

We present a high-fidelity method to generate realistic rendered ground truth images for validating planetary 3D vision mechanisms (VO / stereo mapping / SLAM), based on terrestrial capturing of planetary analog environment by means of the PRo3D viewer's batch render capability. The approach is currently in use for preparation of ExoMars PanCam 3D Vision, and HERA 3D Vision [4], [5].

Future work can include albedo maps in connection with a shader and artificial illumination, the usage of models that support ambiguities on Cartesian DTMs (e.g. overhangs, caves), adding different types of noise to simulate real conditions, and the augmentation of the landscapes by additional (real captured) rocks [2]. On the application side, adding noise to image content and poses, longer trajectories, multiple locations for mapping, and the combination / data fusion with other simulated sensors (LIDAR, INS) can be envisaged.

References

[1] Barnes R., Gupta S., Traxler C., Hesina G., Ortner T., Paar G., Huber B., Juhart K., Fritz L., Nauschnegg B., Muller J.P., Tao Y. and Bauer A. Geological analysis of Martian rover-derived Digital Outcrop Models using the 3D visualisation tool, Planetary Robotics 3D Viewer - PRo3D. In Planetary Mapping: Methods, Tools for Scientific Analysis and Exploration, Volume 5, Issue 7, pp 285-307, July 2018.

[2] Paar, G., Traxler, C., Sidla, O., Bechtold, A., Koeberl, C. "Science Autonomy Training by Visual Simulation". *International Symposium on Artificial Intelligence, Robotics and Automation in Space (I-SAIRAS), 2020.*

[3] Ocón, J., et al. "ADE: Autonomous DEcision making in very long traverses." *International Symposium on Artificial Intelligence, Robotics and Automation in Space (I-SAIRAS), 2020*.

[4] Maria del Pilar Caballo Perucha, Rebecca Nowak, et. Al. Didymos images´ simulation for 3D reconstruction within HERA project. Abstract and oral presentation at Europlanet Science Congress 2020. Virtual Meeting. 21st September – 9th October 2020. Space Missions to Small Bodies: Planetary Defense

[5] Maria del Pilar Caballo Perucha, Rebecca Nowak, et. Al. Imaging Simulations of HERA Didymos Approach. Mission & Campaign Design. Abstract and E-poster at 7th IAA Planetary Defense Conference. Online Event. 26th- 30th April 2021.

[6] Mur-Artal, Raul, and Juan D. Tardós. "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras." *IEEE Transactions on Robotics*5 (2017): 1255-1262.

Acknowledgments