IAC–22,A3,IPB,31,x69374

Robust place recognition with Gaussian Process Gradient Maps for teams of robotic explorers in challenging lunar environments

Riccardo Giubilato
German Aerospace Center (DLR), Germany, riccardo.giubilato@dlr.de

Mallikarjuna Vayugundla
German Aerospace Center (DLR), Germany, mallikarjuna.vayugundla@dlr.de

Cedric Le Gentil
University of Technology Sydney (UTS), cedric.legentil@uts.edu.au

Martin J. Schuster
German Aerospace Center (DLR), Germany, martin.schuster@dlr.de

William McDonald
University of Technology Sydney (UTS), william.t.mcdonald@student.uts.edu.au

Teresa Vidal-Calleja
University of Technology Sydney (UTS), teresa.vidalcalleja@uts.edu.au

Armin Wedler
German Aerospace Center (DLR), Germany, armin.wedler@dlr.de

Rudolph Triebel
German Aerospace Center (DLR), Germany, rudolph.triebel@dlr.de

Teams of mobile robots will play a key role towards future planetary exploration missions. In fact, plans for upcoming lunar exploration, and other extraterrestrial bodies, foresee an extensive usage of robots for the purposes of in-situ analysis, building infrastructure and realizing maps of the environment for its exploitation. To enable prolonged robotic autonomy, however, it is critical for the robotic agents to be able to robustly localize themselves during their motion and, concurrently, to produce maps of the environment. To this end, visual SLAM (Simultaneous Localization and Mapping) techniques have been developed during the years and found successful application in several terrestrial fields, such as autonomous driving, automated construction and agricultural robotics. To this day, autonomous navigation has been demonstrated in various robotic missions to Mars, e.g., from NASA's Mars Exploration Rover (MER) Missions, to NASA's Mars Science Laboratory (Curiosity) and the current Mars2020 Perseverance, thanks to the implementation of Visual Odometry, using cameras to robustly estimate the rover's ego-motion. While VO techniques enable the traversal of large distances from one scientific target to the other, future operations, e.g., for building or maintenance of infrastructure, will require robotic agents to repeatedly visit the same environment. In this case, the ability to re-localize themselves with respect to previously visited places, and therefore the ability to create consistent maps of the environment, is paramount to achieve localization accuracies, that are far above what is achievable from global localization approaches. The planetary environment, however, poses significant challenges to this goal, due to extreme lighting conditions, severe visual aliasing and a lack of uniquely identifiable natural "features". For this reason, we developed an approach for re-localization and place recognition, that relies on Gaussian Processes, to efficiently represent portions of the local terrain elevation, named "GPGMaps" (Gaussian Process Gradient Maps), and to use its gradient in conjunction with traditional visual matching techniques. In this paper, we demonstrate, analyze and report the performances of our SLAM approach, based on GPGMaps, during the 2022 ARCHES (Autonomous Robotic Networks to Help Modern Societies) mission, that took place on the volcanic ash slopes of Mt. Etna, Sicily, a designated planetary analogous environment. The proposed SLAM system has been deployed for real-time usage on a robotic team that includes the LRU (Lightweight Rover Unit), a planetary-like rover with high autonomy, perceptual and locomotion capabilities, to demonstrate enabling technologies for future lunar applications.

I. Introduction

The fast technological development, that the field of mobile robotics experienced within the last decades, creates promising prospects for the application of high-level autonomous robotic skills in several challenging fields of applications, such as search and rescue, mitigation of natural disaster, as well as the exploration of planetary bodies [1, 2, 3].

The autonomous realization of complex tasks, e.g., scouting and exploration of unknown environments, navigation towards targets of scientific importance, or building infrastructure, requires very accurate positioning of the robotic agent with respect to a reference on the ground. Furthermore, the creation of Digital Elevation Models (DEM) with centimeter-level resolution, is paramount for a complete understanding and modeling of the terrain towards a more efficient planning of actions.

Traditional odometry, that makes use of visual perception, as well as global means of localization, can not provide the robots with localization that meets the precision requirements needed to interact with the environment [4, 5, 6] and revisit places. To this end, Simultaneous Localization and Mapping (SLAM) techniques allow the robots to concurrently build a representation of the environment and localize with respect to it. This includes the capability of recognizing previously visited places, and, when revisiting, correcting the pose drift accumulated by the state estimation through the establishment of *loop closures*. This process ultimately leads to an accurate history of poses, but also to a consistent map of the environment.

The principal challenge that is encountered when performing SLAM in unstructured natural environments is related to the ambiguous, repetitive or featureless appearance of the terrain, both from the visual and structural perspectives. This hinders the performances of place recognition as candidate images, or point clouds, are not unambiguously similar to other samples belonging to previous observations [7, 8, 9].

To this end, we developed GPGM-SLAM [10] that, building on top of our modular submap-based multi-robot SLAM [11, 12] system, relies on Gaussian Processes to infer a continuous representation of the local elevation, whose gradient is used to generate Gaussian Process Gradient Maps (*GPGMaps*). We previously demonstrated the accuracy and effectiveness of establishing loop closures with GPGMaps [13] in severely unstructured environments.

With this paper, we present the results obtained during our multi-robot GPGM-SLAM experiment on Mt. Etna, Sicily, during the 2022 ARCHES campaign, and we demonstrate how GPGM-SLAM allows to establish frequent and accurate loop closures, outperforming our baseline system [14].

Fig. 1: Impressions of the environment on Mt. Etna. **Top**: View of the test environment, seen from the camp site. **Center**: The LRU rover during autonomous navigation tests. **Bottom**: Lander and the LRU rover.

II. The ARCHES Campaign

The ARCHES (Autonomous Robotic Networks to Help Modern Societies) project by the Helmholtz Association joined institutes from the deep sea and space domains with the aim of developing technologies to let networked robots act autonomously in teams in order to reach high-level objectives such as collecting measurements or deploying infrastructure in challenging and inaccessible terrains. Within the space domain, the campaign took place on the volcanic slopes of Mt. Etna, Sicily, between the 13th of June and the 9th of July 2022, where it was demon-

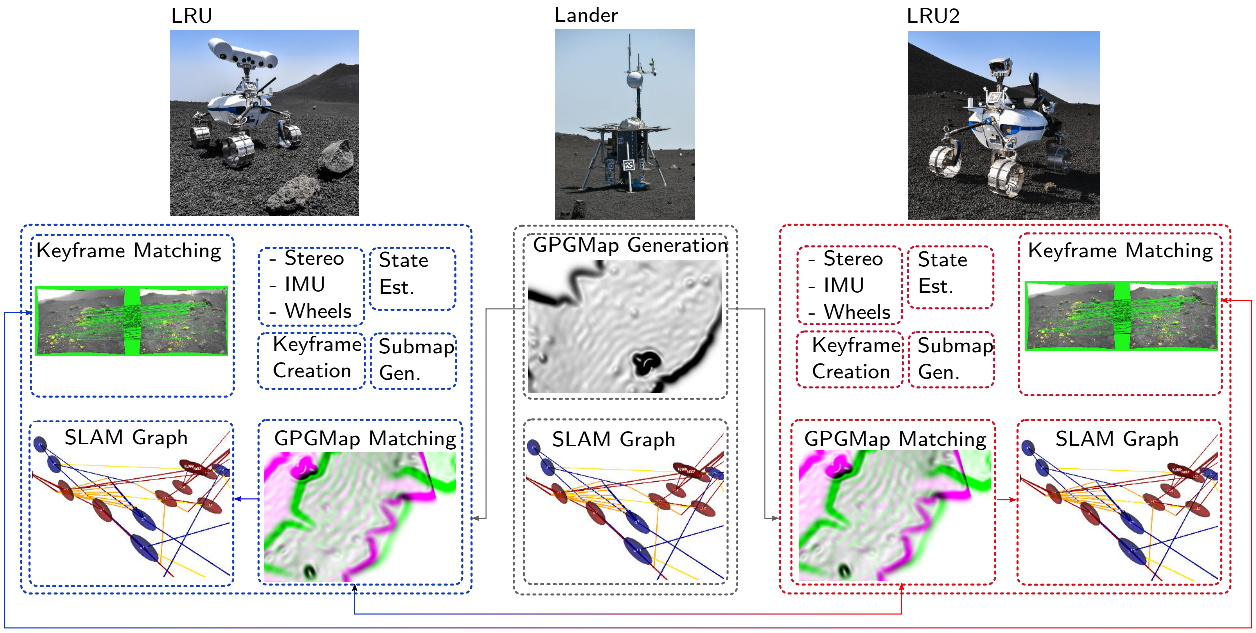

Fig. 2: Short overview of the components of our multi-robot SLAM system that are specifically related to the GPGM-SLAM experiments. Each robot performs state estimation, creates visual keyframes and submaps. The Lander computes the GPGMap representation of each incoming submap and dispatches them back to the robots. LRU and LRU2 utilize and share GPGMaps and keyframes to establish *loop closures* and therefore add pose constraints into the SLAM graph.

strated how a team of heterogeneous robots could act autonomously to deploy infrastructure and take scientifically-relevant samples from the terrain [15].

III. GPGM-SLAM

GPGM-SLAM is a submap-based distributed SLAM system for teams of multiple robots that are equipped with depth sensors as their primary sensory input for the purpose of mapping. Local state estimation is implemented as a Local Reference Filter where IMU measurements, Visual Odometry and Wheel Odometry are fused together to obtain accurate local pose estimates. Local reference frames determine the origin of each *submap*, within which depth measurements from stereo cameras are used to accumulate a rigid point cloud. The origin of each submap is afterwards included in the SLAM graph and subsequently optimized following robot detections or submap matches. In GPGM-SLAM, submaps are converted to GPGMaps, where a continuous model of the elevation is inferred through Gaussian Processes, and its gradients in the spatial domain are matched relying on traditional visual techniques.

More specifically, each GPGMap inherits a point cloud from its parent submap. This set of points constitutes a sparse training set for a Gaussian Process to infer the elevation from a continuous model. The standard formulation for a Gaussian Process has a computational complexity of $\mathcal{O}(n^3)$ for the first inference, where n is the number of training points. In [10], we introduce a variation of the SKI (Structured Kernel Interpolation) [16], that we denote SKI-D, where a fast inference approach, based on inducing points, is used in conjunction with linear operators on the output to obtain elevation gradients in a fast and computationally efficient manner, such that it can be used for real-time operations.

Fig. 2 shows a schematic overview of the implementation of GPGM-SLAM in the context of our multi-robot system during the ARCHES demonstration mission on Mt. Etna. The team comprises, for this experiment, two LRU units, which have different high-level capabilities but share the same sensor modalities for the navigation and mapping. The third robot is the lander, which behaves here as a *static* robot and, through featuring a dedicated computer, solves the SLAM problem independently of the presence of the other robots. While the creation of the submaps is distributed between all robots [11, 12], the lander computer is also dedicated to the conver-

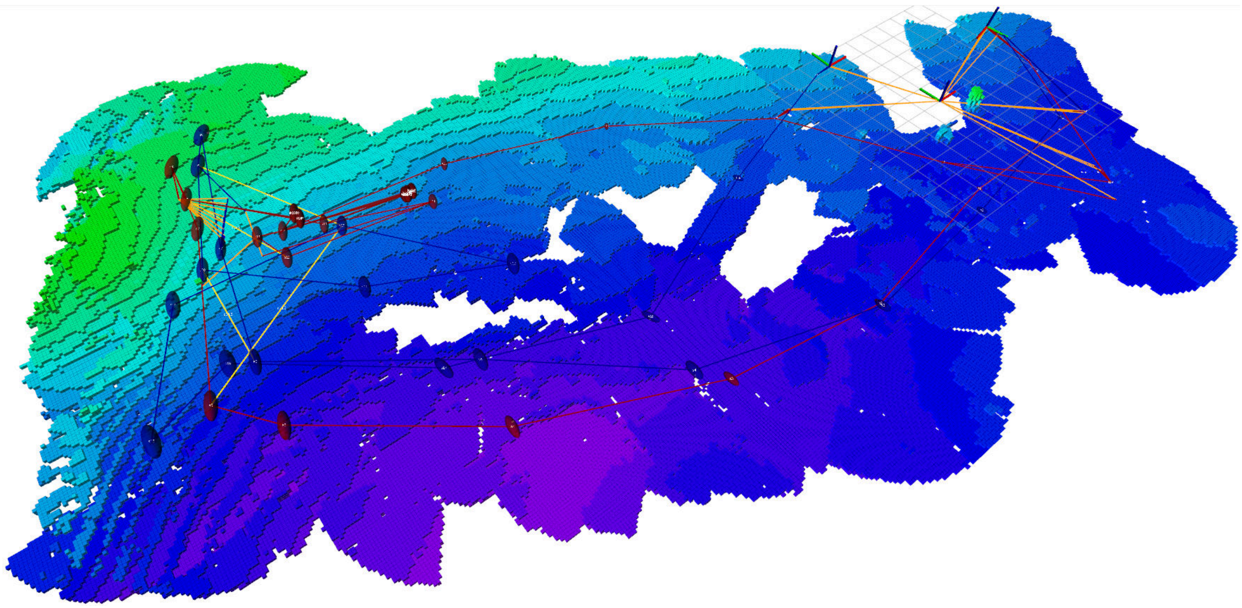

Fig. 3: Octomap and SLAM graph at the end of the multi-robot exploration session, as computed on the Lander and shown from its perspective. Blue and red ellipsoids are, respectively, submap origins and covariances for the LRU and LRU2 robots. Orange lines are multi-robot detections, and yellow lines are GPGMap matches that constrain the relative positions of their origins. The reference frames on the upper-right denote LRU, LRU2 and, at the center of the orange robot detection lines, the Lander (which is treated as a static robot). The mapped area has an approximate size of 40x25 meters.

sion from submaps to GPGMaps for all robots in the team, therefore alleviating the computational load on the robots computers. When GPGMaps are ready, they are dispatched to all robots within a ROS network, so that they can look for matches among the ones belonging to themselves and the other robots. Validated GPGMap matches are then used to establish 6D pose constraints between the submaps in the SLAM graph. The graph optimization itself is then performed in a fully decentralized fashion on all of the agents in the multi-robot team.

IV. Results and Discussion

We present here and discuss the results of a mapping session performed *online* during the ARCHES demonstration mission in July 2022. The SLAM session involved LRU and LRU2 driving autonomously towards waypoints set remotely by human operators. The robots started their navigation in proximity of the lander, such that, through the detection of AprilTags [17], they could join a common reference frame and be aware of their respective position therein. The test environment was characterized by a stone field, where soft sand on a slight slope was often interrupted by small stones up to 30/40cm size, recognized as obstacles from our navigation pipeline, and offering useful details for the purpose of detecting loop closures. The mapping experiment lasted a duration of 43 minutes and covered, in a quite complete manner, an area of approximately 40 by 25 meters. Fig. 3 shows a top view of the resulting map and SLAM pose graph. The terrain is here represented as an Octomap, or an array of cube meshes showing occupied space that are color-coded by elevation. The SLAM graph shows the local reference frames, as ellipsoids, and the factors connecting their origins. Yellow lines represent GPGMap matches, connecting non-subsequent local reference frames. Orange lines connect robot poses at the time of visual robot detections.

Fig. 5 shows a collection of GPGMap matches, collected from LRU and LRU2, aligned and overlaid using the transformation determined after validating the matches. The gradients shown help to understand how little structure is provided by this type of terrain, which makes it especially challenging to establish matches between plain point clouds. In fact, while GPGM-SLAM validated 5 GPGMap matches during the sequence, over the 40 total GPGMaps generated by the robots, the structure-based matching from our baseline system [11, 12] was unable to establish any match.

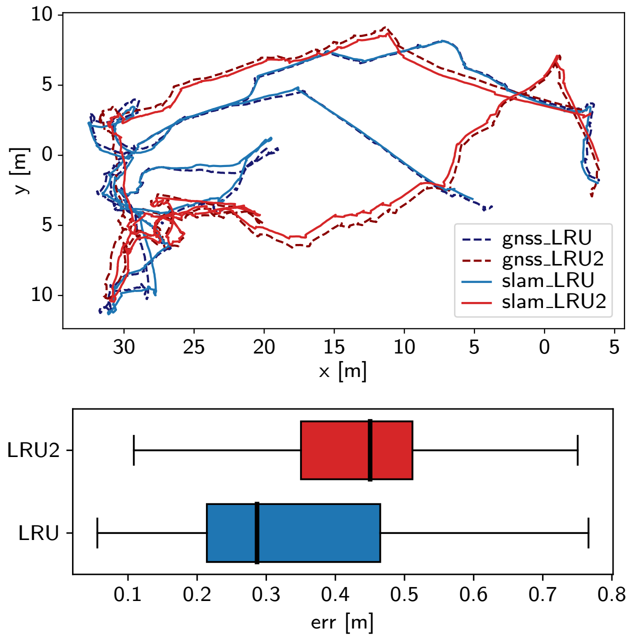

Fig. 4: **Top**: Trajectories of LRU and LRU2, according to the submap poses optimized by the multi-robot SLAM during the experiment on Mt. Etna, aligned to the differential GNSS solution. The coordinate system is centered on the Lander reference frame. **Bottom**: Absolute position errors for the trajectories of LRU and LRU2 aligned to the ground truth.

To accurately evaluate the pose estimation performances of our SLAM system, we equipped the LRU robots with a GPS antenna, and recorded a differential GPS solution for both. In Fig. 4, we plot the full trajectories of the LRUs using the SLAM solution for each local reference frame, which was estimated by the SLAM instance running onboard the lander. The fully optimized trajectories, at the end of the multi-robot SLAM session, are shown against the D-GPS tracks, after an alignment using Horn's method. An analysis of the absolute position errors between the estimated trajectory and ground truth, represented at the bottom of Fig. 4, shows that, at the end of the SLAM session, the medians of the errors result approximately at 30 cm for LRU and 45 cm for LRU2.

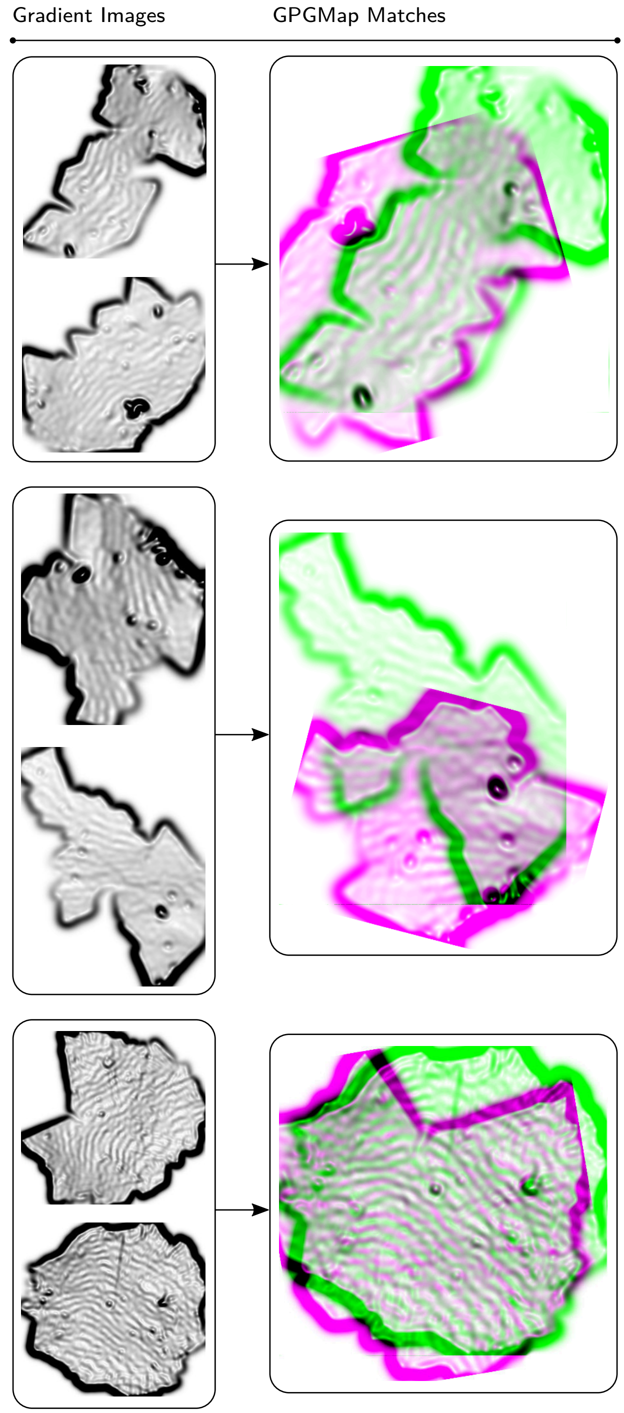

Fig. 5: Collection of GPGMap matches. **Left**: gradient images of the continuous inferred elevation. **Right**: aligned gradient images after match validation.

V. Conclusions

We presented here the results obtained with our novel SLAM pipeline, named GPGM-SLAM, which is targeted at extremely unstructured outdoor environments, during the ARCHES demonstration mission. It took place in the summer of 2022 on the volcanic slopes of Mt. Etna, Sicily, an environment designated as analogous to the lunar surface due to its appearance and geological properties. During this navigation experiment, which was performed online with a team of three robots, we demonstrated that our GPGMaps-based approach outperforms traditional structure-based place recognition, as it is able to match fine structural details that are otherwise impossible to be used in common point cloud-based matching approaches. Furthermore, we demonstrated how the computational load of Gaussian Process inference can be delegated to an computation node external to the robots in order to save resources for other robot-specific tasks, e.g., navigation and obstacle avoidance on the mobile systems.

VI. Acknowledgment

This work was supported by the Helmholtz Association, project ARCHES (contract number ZT-0033) and by the UA-DAAD Australia-Germany Joint Research Collaboration Scheme, project TARMA.

References

[1] G. De Cubber, D. Doroftei, D. Serrano, K. Chintamani, R. Sabino, and S. Ourevitch, "The euicarus project: developing assistive robotic tools for search and rescue operations," in *2013 IEEE international symposium on safety, security, and rescue robotics (SSRR)*. IEEE, 2013, pp. 1–4.

[2] J. Delmerico, S. Mintchev, A. Giusti, B. Gromov, K. Melo, T. Horvat, C. Cadena, M. Hutter, A. Ijspeert, D. Floreano, *et al.*, "The current state and future outlook of rescue robotics," *Journal of Field Robotics*, vol. 36, no. 7, pp. 1171–1191, 2019.

[3] A. Wedler, M. J. Schuster, M. G. Müller, B. Vodermayer, L. Meyer, R. Giubilato, M. Vayugundla, M. Smisek, A. Dömel, F. Steidle, P. Lehner, S. Schröder, E. Staudinger, B. Foing, and J. Reill, "German aerospace center's advanced robotic technology for future lunar scientific missions," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2188, p. 20190574, 2021.

[4] S. Chiodini, M. Pertile, S. Debei, L. Bramante, E. Ferrentino, A. G. Villa, I. Musso, and M. Barrera, "Mars rovers localization by matching local horizon to surface digital elevation models," in *2017 IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)*. IEEE, 2017, pp. 374–379.

[5] P. J. Carle, P. T. Furgale, and T. D. Barfoot, "Long-range rover localization by matching lidar scans to orbital elevation maps," *Journal of Field Robotics*, vol. 27, no. 3, pp. 344–370, 2010.

[6] B. Wu, P. R. WK, P. Ludivig, A. S. Chung, and T. Seabrook, "Absolute localization through orbital maps and surface perspective imagery: A synthetic lunar dataset and neural network approach," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 3262–3267.

[7] Y. Zheng, T. Birdal, F. Xia, Y. Yang, Y. Duan, and L. J. Guibas, "6d camera relocalization in visually ambiguous extreme environments," 2022. [Online]. Available: https://arxiv.org/abs/2207.06333

[8] R. Giubilato, W. Stürzl, A. Wedler, and R. Triebel, "Challenges of slam in extremely unstructured environments: The dlr planetary stereo, solid-state lidar, inertial dataset," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8721–8728, 2022.

[9] R. Giubilato, M. Vayugundla, M. J. Schuster, W. Stürzl, A. Wedler, R. Triebel, and S. Debei, "Relocalization with submaps: Multisession mapping for planetary rovers equipped with stereo cameras," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 580–587, 2020.

[10] R. Giubilato, C. L. Gentil, M. Vayugundla, M. J. Schuster, T. Vidal-Calleja, and R. Triebel, "GPGM-SLAM: a Robust SLAM System for Unstructured Planetary Environments with Gaussian Process Gradient Maps," *Field Robotics*, vol. 2, 2022.

[11] M. J. Schuster, K. Schmid, C. Brand, and M. Beetz, "Distributed stereo vision-based 6d localization and mapping for multi-robot teams," *Journal of Field Robotics*, vol. 36, no. 2, pp. 305–332, 2019.

[12] M. J. Schuster, "Collaborative Localization and Mapping for Autonomous Planetary Exploration: Distributed Stereo Vision-Based 6D SLAM in GNSS-Denied Environments," Ph.D. dissertation, University of Bremen, 2019. [Online]. Available: http://nbn-resolving.de/urn:nbn:de:gbv:46-00107650-19

[13] C. Le Gentil, M. Vayugundla, R. Giubilato, W. Stürzl, T. Vidal-Calleja, and R. Triebel, "Gaussian process gradient maps for loop-closure detection in unstructured planetary environments," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1895–1902.

[14] R. Giubilato, M. Vayugundla, W. Stürzl, M. J. Schuster, A. Wedler, and R. Triebel, "Multi-modal loop closing in unstructured planetary environments with visually enriched submaps," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021.

[15] M. J. Schuster, M. G. Müller, S. G. Brunner, H. Lehner, P. Lehner, R. Sakagami, A. Dömel, L. Meyer, B. Vodermayer, R. Giubilato, M. Vayugundla, J. Reill, F. Steidle, I. von Bargen, K. Bussmann, R. Belder, P. Lutz, W. Stürzl, M. Smíšek, M. Maier, S. Stoneman, A. F. Prince, B. Rebele, M. Durner, E. Staudinger, S. Zhang, R. Pöhlmann, E. Bischoff, C. Braun, S. Schröder, E. Dietz, S. Frohmann, A. Börner, H. Hübers, B. Foing, R. Triebel, A. O. Albu-Schäffer, and A. Wedler, "The ARCHES Space-Analogue Demonstration Mission: Towards Heterogeneous Teams of Autonomous Robots for Collaborative Scientific Sampling in Planetary Exploration," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5315–5322, 2020.

[16] D. Eriksson, K. Dong, E. Lee, D. Bindel, and A. G. Wilson, "Scaling gaussian process regression with derivatives," in *Advances in Neural Information Processing Systems*, 2018, pp. 6867–6877.

[17] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE international conference on robotics and automation*. IEEE, 2011, pp. 3400–3407.