EuroGNC Navigation Guidance Control CEAS Bristol, UK June 11th-13th 2024
University of BRISTOL ROYAL AERONAUTICAL SOCIETY AIAA

A small body open-source dataset for image processing algorithms

Mattia Pugliatti PhD Student, Department of Aerospace Science and Technology, Politecnico di Milano, 20156, Milan, Italy. mattia.pugliatti@polimi.it

Shubham Vyas Researcher, Robotics Innovation Center, DFKI GmbH, Robert-Hooke-Strasse 1, Bremen 28359, Germany. Shubham.Vyas@dfki.de

Marko Jankovic Researcher, Robotics Innovation Center, DFKI GmbH, Robert-Hooke-Strasse 1, Bremen 28359, Germany. Marko.Jankovic@dfki.de

Francesco Topputo Full professor, Department of Aerospace Science and Technology, Politecnico di Milano, 20156, Milan, Italy. francesco.topputo@polimi.it

ABSTRACT

The capability to accurately detect surface morphological features such as craters and boulders of multiple sizes and at different scales on the surface of small bodies is of paramount importance for various vision-based applications around small bodies. The development of this capability, however, is hindered by significant challenges: the environmental conditions due to the irregularity of the bodies, properties, and distribution of the features, rapidly changing illumination conditions, and most importantly, the lack of publicly available datasets for training, validation, or testing. In this work, the authors describe the methodology used to generate a general-purpose open-access dataset specifically designed to simplify access to labeled datasets about small bodies and to enable the design and application of advanced data-driven image processing algorithms.

Keywords: Dataset; Analog facility; Small bodies; Space exploration; Drone; Segmentation; Navigation

1 Introduction

Physical setups such as one that uses 3D printed shape models within robotic facilities or terrain analog resembling the actual environment are valuable for Hardware-in-the-Loop (HIL) experiments. A variety of these facilities exists [1–5] and are currently operated around the world at various capacities to validate Image Processing (IP) and vision-based algorithms. Their main advantages lie in the accurate photometric acquisitions made possible by the natural scattering of the light from an analog material stimulated by the artificial lamps and a real-time simulation framework (no waiting time for rendering the scene). Their main drawback is the high operational costs, the large resources needed to be invested in their design and maintenance, and the rigid setup, which poses geometrical constraints (self-occlusions, self-shadowing, range of the simulation limited to the size of the facility, lighting conditions).

For these motivations, these facilities are often used only in the final stages of an incremental validation campaign. De facto, data acquired with this setting is rarely available outside of the environment, with few notable exceptions of open-access datasets for space applications [6, 7].

The capability to accurately detect surface morphological features such as craters and boulders at different scales on the surface of small bodies is of paramount importance for a variety of vision-based applications around small bodies. The development of this capability, however, is hindered by significant challenges: the environmental conditions due to the irregularity of the bodies, properties, and distribution of the features, rapidly changing illumination conditions, and most importantly, the lack of publicly available datasets for training, validation, or testing.

This latter challenge is addressed with the detailed description of a setup used at Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) to create a general-purpose open-access dataset designed to simplify access to labeled data about small bodies. The dataset can be freely accessed at https://zenodo.org/records/11190629.

2 Facility

The Robotic Innovation Center (RIC), part of DFKI, operates a terrain analog facility to test free-climbing robotic systems and demonstrates their mobilities. The facility is roughly 15.0 $m \times 10.3$ m wide and 4.4 m high, with a total testing area of about 146.5 m^2. A curtain can be mechanically lowered from the ceiling to make a dark environment representative of space conditions. A studio lamp can then be used within the facility to simulate the Sun's illumination conditions. A sizeable unused volume exists on top of the facility, given that there is a distance of about 12 m between the top of the facility and the ceiling. The surface has been built using real data from the south polar craters on the Moon and images from the Apollo missions and is composed of three contiguous areas with 25°, 35°, and 45° slopes and two flat areas at the top and at the bottom. Obstacles such as boulders can be attached at predisposed screw points distributed at regular intervals on the terrain to provide challenging surface conditions for locomotion systems. Finally, a sandy substrate is also simulated with fine-grained material (<1 mm) in the lower part of the facility. A Vicon [1] tracking system is mounted surrounding the facility from above. Finally, in the lower section, a dedicated space serves as a control center containing Vicon's servers and a working station. A 2D map of the facility seen from above is visible in Figure 2 while some examples of the views of the terrain are visible in Figure 1.

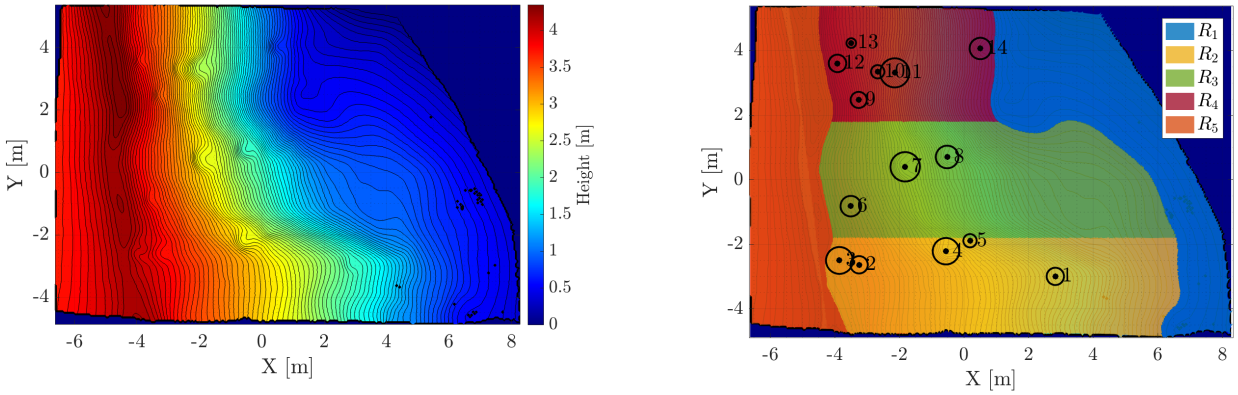
Usually, the facility simulates surface interactions to test the robotic systems developed at RIC. However, as part of a collaboration with DFKI, it has been exploited in a dedicated activity designed for visual-based applications.

As illustrated in Figure 2b, the facility is divided into 5 different regions whose boundaries are designed to separate the three sloped areas at the center, and the flat ones at the top and bottom. From R_1 to R_5 each region corresponds to a surface area of roughly 26.7 m^2, 34.0 m^2, 39.1 m^2, 22.5 m^2, and 25.9 m^2. Within these regions, 14 large craters have been manually identified and assigned an ID for referencing. Additionally, for this activity, a boulder field has been created with spare boulders on the lower portion of the terrain within R_1.

Since a digital terrain model of the facility exists, it is possible to develop a digital twin of the real analog terrain in Blender using Celestial Object Rendering TOol (CORTO)[8]. This is crucial since it will enable the creation of an artificial environment in which it is possible to generate synthetic labels that go beyond the reconstructed pose provided by the Vicon system. Some of the expected labels are illustrated in Figure 3. In order to link the real facility with the artificial one, several calibration procedures are necessary.

[1] `https://www.vicon.com/`, last accessed 12th of September 2023.

Fig. 1 Example of views of the facility with simulated lighting conditions.

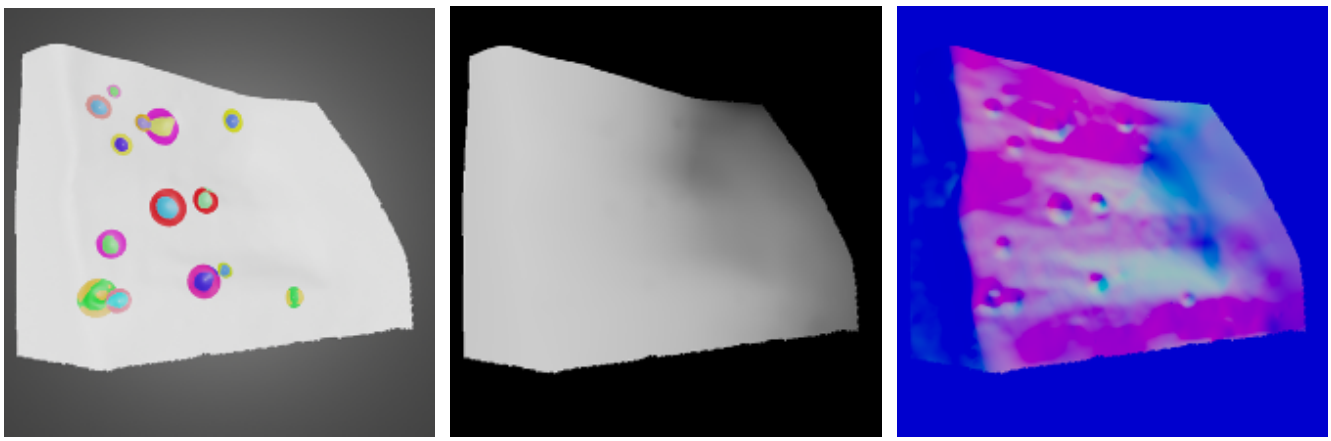


(a) **(b)**

Fig. 2 Height-map of the facility (a) with isolines (plotted at every 0.05 m intervals). The naming convention for the large craters and regions of the facility (b) used in this section.

2.1 Data collection setup

Exploiting the current design of the facility at the RIC and the free available space above the terrain, a drone is used as an image-collecting device to generate the dataset samples. A commercial drone is a cost-effective option for positioning a camera across the facility without installing other complex equipment. The complete setup used to generate the dataset is illustrated in Figure 4.



(a) (b) (c)

Fig. 3 Examples of labels obtained with the artificial environment of the terrain analog facility: (a) semantic segmentation labels for craters, (b) depth map, (b) slope map.

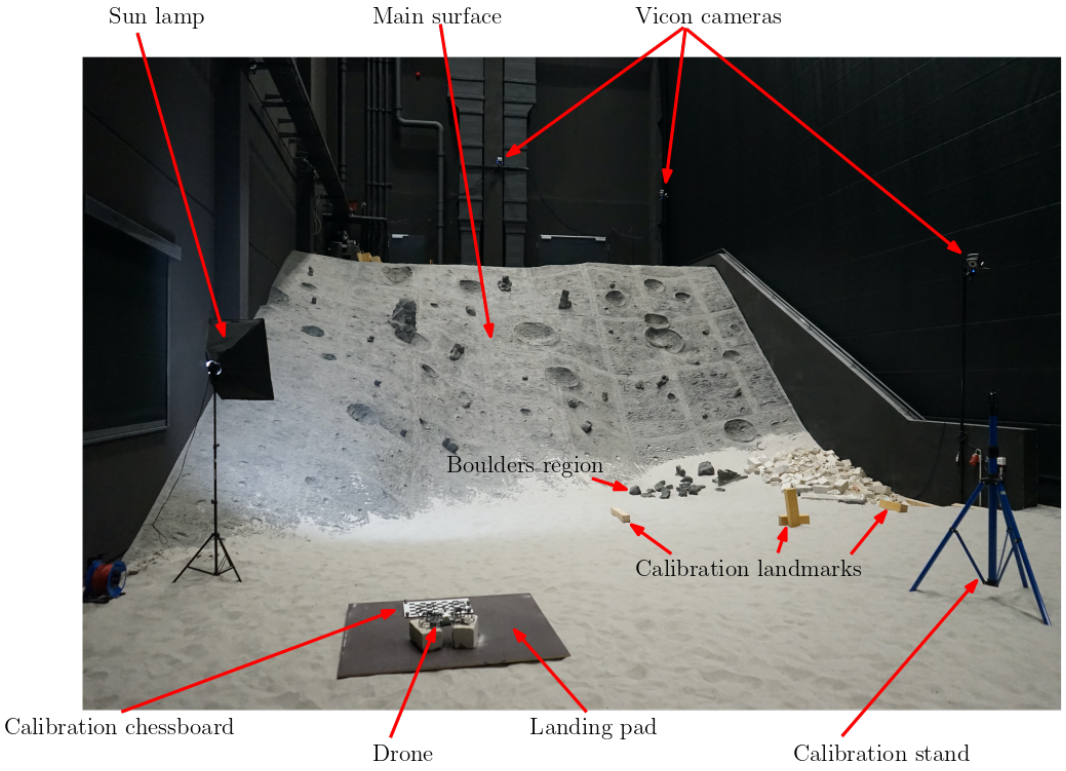


Fig. 4 Schematic of the setup used in the facility to generate the dataset.

To generate the dataset, the facility has been used toghether with additional components:

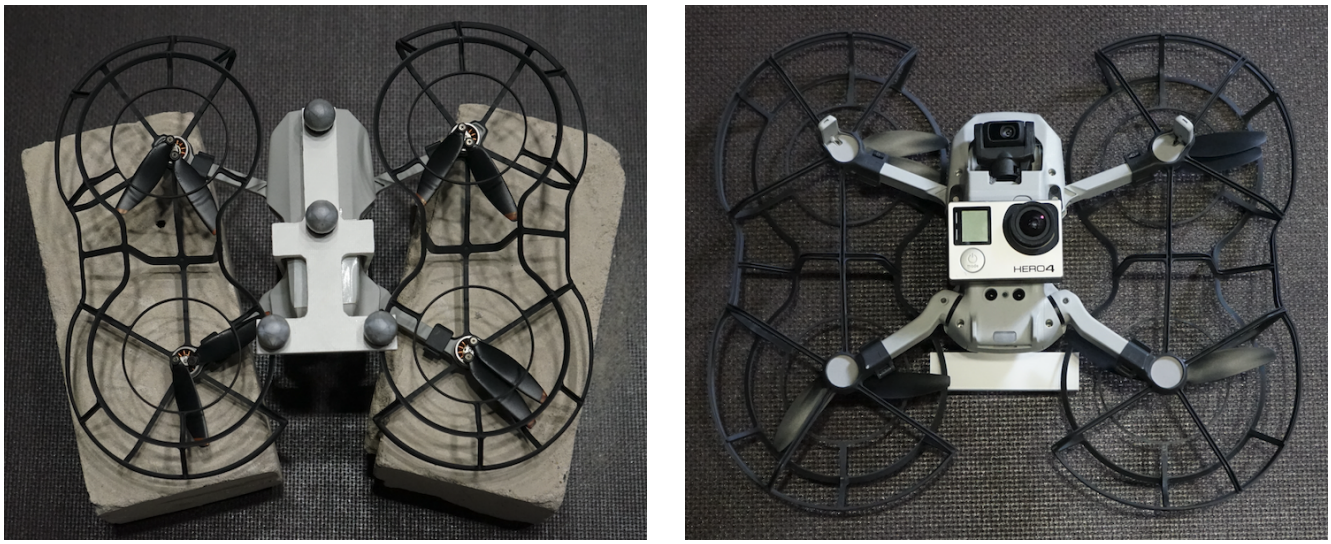
- **Drone**: A commercial drone (The DJI-mini SE [2] is used as a tool to position cameras around the facility. The drone is flight manually through a dedicated controller and is equipped with the following components, as illustrated in Figure 5:
  - **Propeller guards**: They guarantee safe operations and avoid damage to the curtain, the Vicon cameras, or any other element within the facility. The guards lower the drone performance in terms of battery life, shortening the usable flight time with each battery.
  - **3D-printed Vicon tracker stand**: A rigid 3D-printed support structure is attached to the drone for positioning 4 Vicon markers to allow the detection and tracking of the drone within

[2] `https://www.dji.com/id/mini-se/specs`, last accessed 11th of September, 2023.

the facility [3]. This 3D-printed structure demonstrated essential to increase the visibility of the trackers to the Vicon cameras, increasing the number of poses correctly generated within the facility during each flight.

  - **GoPro Hero-4**: A GoPro camera is rigidly attached to the structure of the drone (on the bottom part) to acquire videos with nadir pointing during each flight. By default, the drone is also equipped with a gimbal-stabilized camera. Both cameras can record RGB videos at different framerates, resolutions, and Field Of View (FOV).

- **Calibration chessboard**: A rigid 7×10 calibration chessboard with 35 mm squares and Vicon markers attached to it.
- **Sun lamp**: A studio lamp to qualitatively simulate illumination conditions from the Sun.
- **Vicon system**: A Vicon motion tracking system with cameras surrounding the facility from above.
- **Calibration stand**: A tripod supporting the drone during the calibration procedures.
- **Calibration landmarks**: Visual landmarks used to perform manual calibration in the facility with the Vicon system, the drone, the calibration stand.
- **Boulders region**: portion of the facility where a boulder field is artificially simulated specifically for this activity. The boulder field is created with spare boulders not currently mounted within the terrain.
- **Main surface**: the main portion of the facility with craters, boulders, slopes, and regions with sand.
- **Take-off pad**: safe region for drone take-off and landing during each flight.



(a) **(b)**

Fig. 5 Top (a) and bottom (b) view of the drone used to generate the dataset.

2.2 Data generation

Given the setup illustrated in the previous section, the data is collected by 18 manual flights of the drone within the facility. The dataset is generated with varying camera properties and illumination conditions. Each flight is executed manually following an approximate predefined flight path. First, the drone traverses the R_2 region across the X axis until arriving at the R_5 region. Then, the drone is moved across the R_5 region, spanning the $R_4 - R_3 - R_2$ areas from top to bottom until the boulder field in R_1 is reached. The drone is then moved in a horizontal-vertical cross pattern to mimic a landing trajectory. After this passage, the drone is moved once again over the $R_2 - R_3 - R_4$ region with a random motion.

[3]The author would like to personally thank Houssemeddine Jebali at DFKI for having modeled and printed the support for the markers.

The duration of each flight is approximately 5 minutes from take-off to landing. The maximum duration allowable by each battery is greatly reduced by the installation of the propeller guards and by the extra weight of the GoPro and Vicon tracker stand attached respectively on the bottom and on top of the drone. Figure 6 and Figure 7 illustrate all the reconstructed positions recorded by the Vicon system during each flight. Note that the drone is flown visually from the control center within the facility during each flight. This is due to the absence of a GPS signal within the building, making it impossible with the current commercial setting to program any predefined path planning algorithm to automatize data collection in a sistematic way.

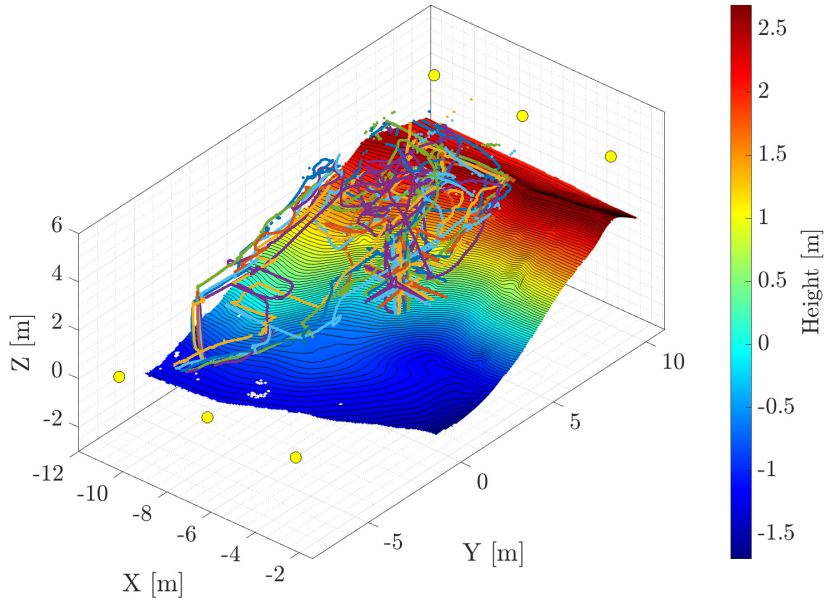


Fig. 6 3D view of the final 18 flights used to generate the dataset. The flights are displayed together with the facility elevation model and with the approximate location of the Sun's lamp (yellow dots).

During each flight, video streams, log files from the DJI onboard software, and Vicon data are simultaneously recorded as raw data. Note that in practice, both GoPro and DJI imaging sensors are simultaneously activated. However, the GoPro is considered as the primary data-collection sensor, while the DJI imaging data is collected on an opportunistic basis. This is due to an issue with the self-stabilyzing gimbal mechanism of the DJI camera, which invalidate a rigid hand-eye extrinsic calibration, as disussed in the next section.

Each of the flights is executed with different camera settings (resolution, Frame Per Second (FPS), and FOV) and illumination conditions. The camera settings used to generate the videos are summarized in Table 1. The first three (A, B, and C) represent the settings used for the GoPro, while the latter two (D and E) are the settings used for the DJI sensor. For each of the GoPro settings, six flights are executed, varying the illumination conditions by positioning the Sun lamp in different locations around the facility, as illustrated by the yellow points in Figure 6 and Figure 7. The lamp thus resulted positioned three times in R_1 with a left-center-right configuration pointing towards the center of the facility and then three times in R_5 with the same configuration. Repeating this light pattern for each of the A-B-C settings results in the 18 flights.

2.3 Calibrations

For this activity, several calibration proceudres are necessary. The Vicon system is assumed to be calibrated; thus, its calibration procedure is not presented.

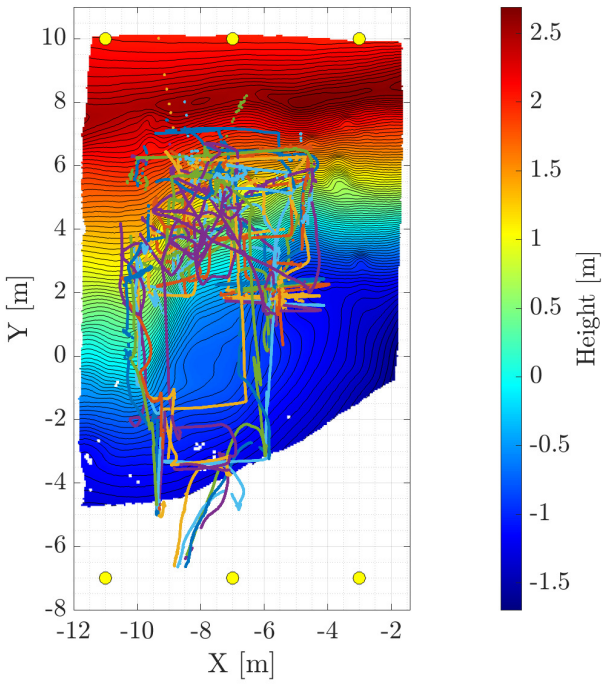


Fig. 7 Top view of the final 18 flights used to generate the dataset.

Table 1 Characteristics of the different camera settings used to generate the dataset.

ID	Resolution [px]	FPS	FOV	Sensor
A	1920×1080	120	Wide	GoPro
B	2704×1520	50	Medium	GoPro
C	1920×1080	120	Narrow	GoPro
D	1920×1080	50	Wide	DJI
E	2720×1530	50	Wide	DJI

To link the digital model of the facility with the real one, the reference frame of the Vicon system is exploited as an intermediary. Three markers are positioned at prominent features in the real facility while the Vicon system records their positions. The positions of the three trackers are then matched in the digital model of the facility. The transformation from the facility and Vicon reference frames can then be easily reconstructed. Since this is a rigid transformation, the calibration is performed only once. This procedure is referred to as the facility-facility calibration since it allows the transformation of a position vector from the real to the digital facility and vice-versa.

Several data streams are collected from different systems and sensors during each flight. Collection alone is not enough, as video streams, Vicon data, and flight log data internally recorded by the DJI software must be synchronized together in order to be useful. This procedure is referred to as time synchronization and is a laborious task that is performed in a semi-automatic way. To achieve time synchronization, a particular abrupt relative motion needs to be performed during each data-collection event (typically at the beginning) between the drone, the chessboard pattern, and the Vicon system. This is necessary to generate a noticeable and time-limited event that can be simultaneously registered in each data stream. During each flight, this is performed by a sudden sharp movement across the X direction in the facility above the calibration chessboard shortly after take off. On the other hand, during

calibration, when the drone is kept stationary, this is achieved by manually moving the chessboard in a clearly detectable pattern.

Currently, the time synchronization is performed only between the video streams and the Vicon system. Also, since these operate at different FPS (100 for the Vicon system, 50 or 120 for the GoPro, and 50 for the DJI camera), the time synchronization accuracy varies depending on the combinations of the three elements and the recording settings. The accuracy of the calibration have a lower limit given by the lowest FPS used (e.g. 20 ms). However, since the low-velocity of each flight and given that images are recorded as static data-points for the dataset, the time synchronization error is expected to be far lower than the error introduced by other sources (e.g. the right-hand extrinsic calibration). The time synchronization is performed in 2 steps: 1) The shared intervals in which all data streams exhibit the abrupt motion at the beginning of each flight are manually selected and considered as candidate solutions for time-synchronization. 2) The time-synchronization achieving the smallest extrinsic calibration error is selected as the correct one.

Intrinsic calibration is performed to estimate camera parameter and coefficients for tangential and radial distortion models. This procedures allow for the estimation of the intrinsic camera matrix K.

The extrinsic calibration, also referred to as hand-eye calibration, is necessary to find two unknown transformations in the chain of relationships between the elements of the facility: 1) the transformation matrix between the drone reference frame (D) (detected by Vicon markers) and the camera reference frame (C), namely T_C^D, and 2) the transformation matrix between the physical chessboard reference frame H (detected by Vicon markers) and the ones associated with the chessboard pattern J, namely T_J^H. Since both the Vicon markers and the GoPro camera are rigidly mounted with respect to each other (as it is possible to see in 5) and the same holds true between the markers on the physical chessboard and its pattern, both transformations are constant during the entire data-collection procedure. Note that because of the time-varying gimbal mechanism of the Drone's camera, the transformation between such camera and the Vicon trackers is time-varying, thus cannot be reliably estimated during calibration and assumed constant during each flight. This issue motivated the use of the GoPro camera as main data-collection sensor.

During the hand-eye calibration, the drone is positioned vertically on a calibration stand within the facility, clearly visible by the Vicon cameras. The chessboard calibration pattern is then moved manually to be captured by the drone-mounted camera. During the procedure, n acquisitions of the chessboard pattern are performed. In each of these acquisitions, m corner points embedded by the chessboard pattern are extracted. During this procedure the Vicon system (V) simultaneously tracks the drone (D) and the chessboard markers (H). If at any given point of the calibration data is missing about the drone, chessboard, or the chessboard pattern is not correctly identified in the image, that data is discarded from the procedure.

The purpose of the calibration is to solve for the two unknown rigid transformation matrices T_C^D and T_J^H using the error on the retro-projection error of the chessboard pattern. To do so, during calibration the homogeneous coordinates of a single chessboard pattern point p_i are transformed in physical world coordinates as X_i with the following chain of transformations:

$$X_i = T_D^C \cdot T_V^D \cdot T_C^V(t) \cdot \left(T_J^H \cdot p_i\right) \tag{1}$$

Note that T_C^V is the only time-varying transformation, given by the manual movement of the camera within the facility during flights. Keeping the drone stationary during calibration, makes it constant. The two unknown transformations are expressed for simplicity as seven-element vectors made by origin coordinates and quaternions:

$$\{T_C^V\}_v = \left[x_0,\ y_0,\ z_0,\ q_0,\ q_1,\ q_2,\ q_3\right] \tag{2}$$

$$\{T_J^H\}_v = \left[x_0,\ y_0,\ z_0,\ q_0,\ q_1,\ q_2,\ q_3\right] \tag{3}$$

Using the distortion coefficients extracted from the intrinsic calibration, and the intrinsic calibration matrix K, the point X_i gets projected on the image plane as a point p_k via a pinhole camera model as:

$$p_j = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot X_i \tag{4}$$

$$x = \frac{p_{j,1}}{p_{j,3}} \quad y = \frac{p_{j,2}}{p_{j,3}} \quad r^2 = x^2 + y^2 \tag{5}$$

$$x^{'} = 2p_1xy + p_2\left(r_2 + 2x^2\right) \tag{6}$$

$$y^{'} = p_1\left(r_2 + 2y^2\right) + 2p_2xy \tag{7}$$

$$x^{''} = x\left(1 + k_1r_2 + k_2r_2^2 + k_3r_2^3\right) \tag{8}$$

$$y^{''} = y\left(1 + k_1r_2 + k_2r_2^2 + k_3r_2^3\right) \tag{9}$$

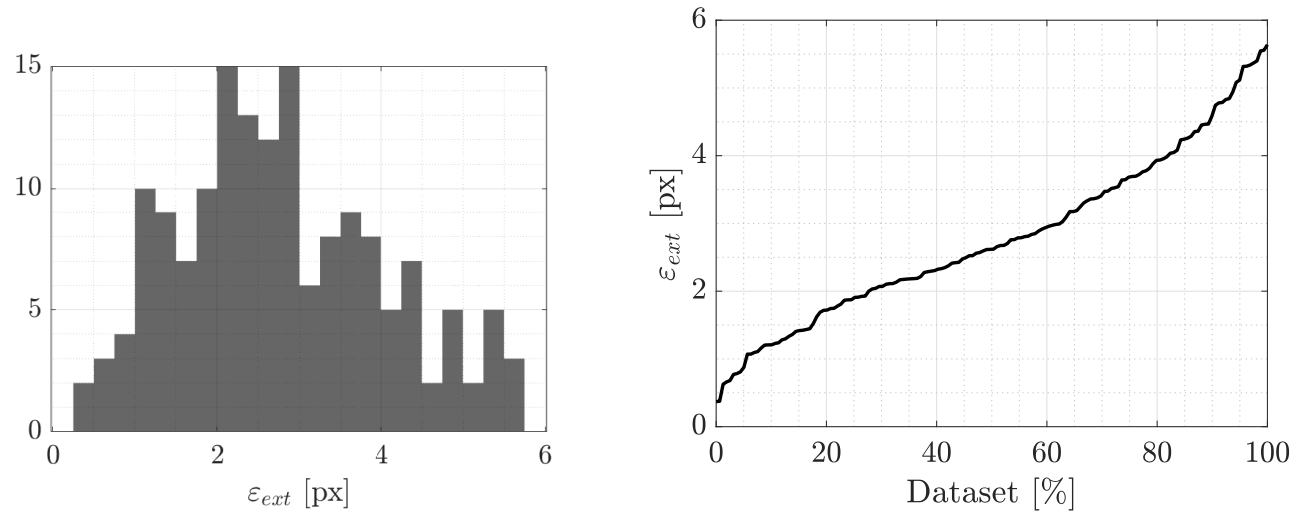
$$p_k = K\begin{bmatrix} x^{'} + x^{''} \\ y^{'} + y^{''} \\ 1 \end{bmatrix} \tag{10}$$

where $x^{'}$ and $y^{'}$ represent points coordinate after applying tangential distortion and $x^{''}$ and $y^{''}$ those after applying radial distortion, and K is the intrinsic camera calibration matrix.

The prediction of the coordinates of the points p_k for the chessboard pattern can then be confronted with their actual detection from images p_d, acquired during calibration procedure. The difference between the two, averaged for the entire pattern, is used to compute a global error metric:

$$\varepsilon_{ext} = \frac{\sum_n \sqrt{\left(\sum_m \left(p_k - p_d\right)^2\right)}}{n} \tag{11}$$

Using ε_{ext} as loss function and the vector representation of the two unknown transformations $\{T_C^V\}_v$ and $\{T_J^H\}_v$, an optimization problem is solved in Matlab using the *fmincon* function with a sequential quadratic programming solver while imposing the quaternion equality constraint. At convergence, the optimization routine generates the two estimates for the transformations with an error of ε_{ext} of just few pixels. The results of one of the extrinsic calibration procedures are illustrated in 8 and 9, which reports a calibration error of 2.79 px on a 1080×1920 px sensor.



(a) **(b)**

Fig. 8 Examples of histogram (a) and cumulative distribution function (b) of the extrinsic calibration error.

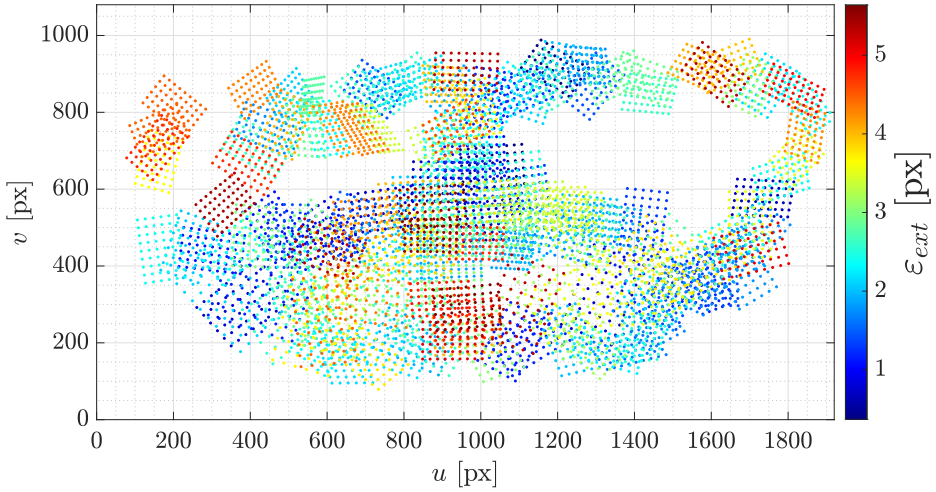


Fig. 9 Example of calibration patterns in camera image plane with their corresponding retro-projection errors.

Finally, data is also recorded to perform radiometric calibration. Black and white images are recorded with different camera settings in representative illumination conditions encountered during flights. This data can be optionally used to calibrate the level of dark noise and other camera effects in the two sensors used for data collection.

It is remarked that the dataset will be provided with raw calibration data, so that the final user can experiment with different calibration strategies than the ones performed by the authors.

3 Results

The total size of the raw data (composed of videos, DJI log files, and Vicon data) by the 18 flights and the calibration procedures is about 105 Gb. The final dataset will be further processed from this raw data and subdivided into three levels of increasing complexity:

- **Level 0 (Images):** Single images alone or in sequence are extracted from the raw videos. These can be used to qualitately assess the domain gap of image processing algorithm between synthetic and real scenarios.

- **Level 1 (Images + Poses):** Images are accompanied by full poses reconstructed via the Vicon system and/or using the transformations estimated using the extrinsic calibration procedure. At this level, the dataset can serve to perform quantitative assessment for visual-based navigation algorithms. Applying extrinsic calibration procedure illustrated in this work, errors in the order of few pixels are expected (as illustrated before). Using the raw Vicon data, the error is limited by the time synchronization procedure, given that the Vicon system can generate position labels with mm-level accuracy for the camera.
- **Level 2 (Images + Poses + Image Labels):** Images are accompanied by fully reconstructed poses and label masks such as the ones illustrated in Figure 3. The labels are generated using the digital twin of the facility.

A random sample of images from one of the flights is illustrated in Figure 10. All raw and calibration data has been already collected. However, at the current stage, the dataset is made publicly available for Level 0 and Level 1 data (only with Vicon position labels), while other levels will need to be processed in future stages. The dataset can be freely accessed on Zenodo https://zenodo.org/records/11190629, with the doi: 10.5281/zenodo.11190628.

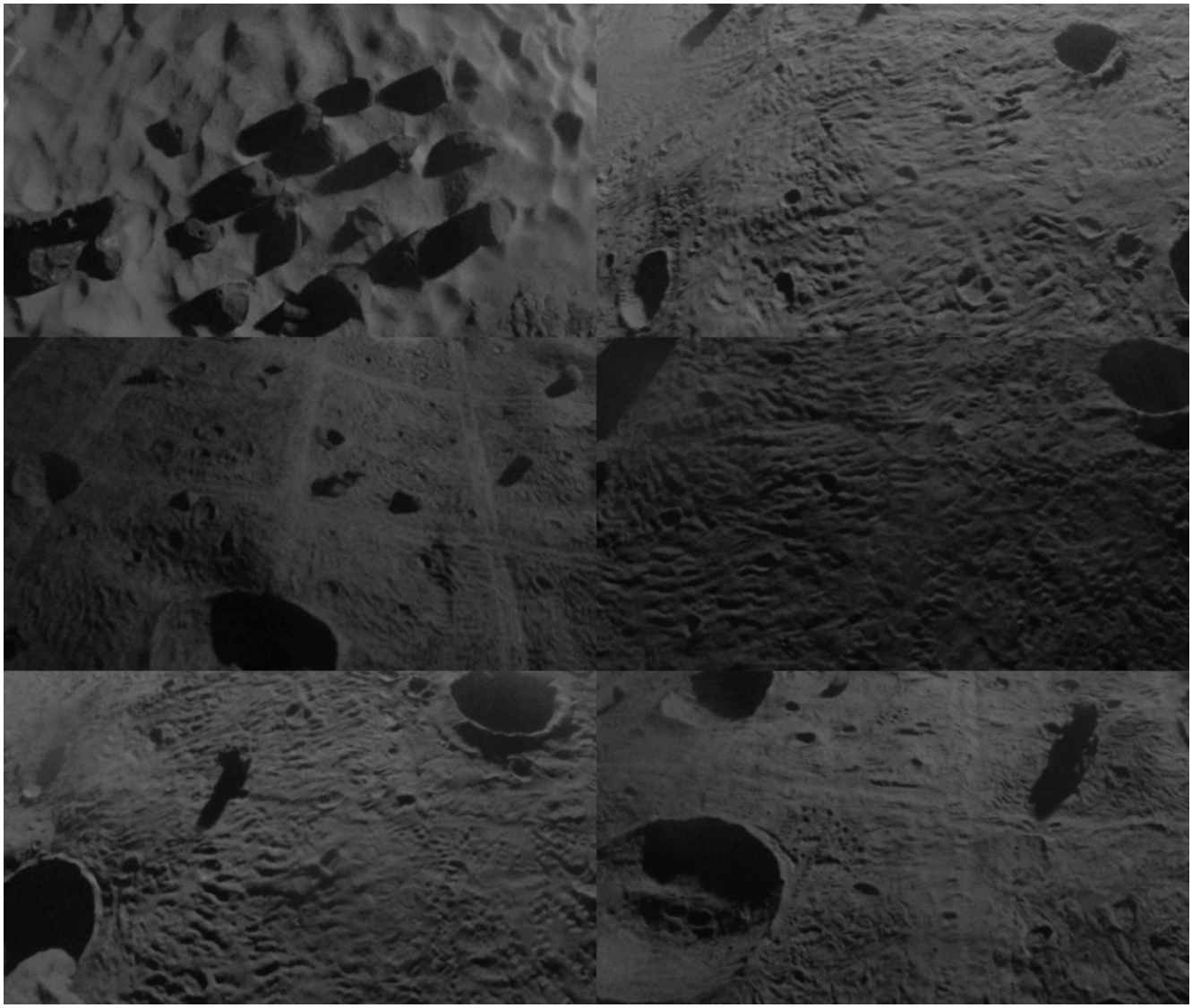


Fig. 10 Six samples of images captured by the GoPro during one of the flights with ID = B. The brightness of the images is artificially adjusted with γ = 0 .7 for visibility in the manuscript.

Acknowledgments

The authors would like to acknowledge the funding received from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813644.

References

[1] Hans Krüger and Stephan Theil. TRON - hardware-in-the-loop test facility for lunar descent and landing optical navigation. *IFAC Proceedings Volumes*, 43(15):265–270, 2010. DOI: 10.3182/20100906-5-jp-2022.00046.

[2] Martin Zwick, Irene Huertas, Levin Gerdes, and Guillermo Ortega. Orgl–esa's test facility for approach and contact operations in orbital and planetary environments. In *Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS), Madrid, Spain*, volume 6, 2018.

[3] Tae Ha Park, Juergen Bosse, and Simone D'Amico. Robotic testbed for rendezvous and optical navigation: Multi-source calibration and machine learning use cases. *arXiv preprint arXiv:2108.05529*, 2021.

[4] A Pellacani, M Graziano, and M Suatoni. Design, development, validation and verification of gnc technologies. In *8th European Conference for Aeronautics and Sciences (EUCASS)*, 2019.

[5] Mehregan Dor, Travis Driver, Kenneth Getzandanner, and Panagiotis Tsiotras. Astroslam: Autonomous monocular navigation in the vicinity of a celestial small body – theory and experiments, 2022.

[6] Travis Driver, Katherine A. Skinner, Mehregan Dor, and Panagiotis Tsiotras. AstroVision: Towards autonomous feature detection and description for missions to small bodies using deep learning. *Acta Astronautica*, 210:393–410, 9 2023. DOI: 10.1016/j.actaastro.2023.01.009.

[7] Tae Ha Park, Marcus Märtens, Gurvan Lecuyer, Dario Izzo, and Simone D'Amico. Speed+: Next-generation dataset for spacecraft pose estimation across domain gap. In *2022 IEEE Aerospace Conference (AERO)*, pages 1–15. IEEE, 2022.

[8] Mattia Pugliatti, Carmine Buonagura, and Francesco Topputo. Corto: The celestial object rendering tool at dart lab. *Sensors*, 23:9595, 2023. ISSN: 1424-8220. DOI: 10.3390/s23239595.