

Absolute Visual Servoing for Precise Earth Target Pointing Onboard Small Satellites

Johannes Dauner^{a*}, Joshua Redelbach^{a,b}, Lisa Elsner^a, Ilham Mammadov^a, Klaus Schilling^a

^a Zentrum für Telematik e.V., Würzburg, Germany

^b University of Würzburg, Germany

* Corresponding Author, johannes.dauner@telematik-zentrum.de

Abstract

Earth observation and high-bandwidth communication applications impose high demands on the attitude and orbit control system (AOCS) of a satellite in order to achieve precise pointing to the observation target or the ground station on Earth. Since it is necessary not only to control the satellite's attitude, but also to know its exact position in orbit, these satellites mostly rely on star trackers for attitude determination and GNSS receivers for orbit position determination.

In the scope of the Telematics Earth Observation Mission (TOM), three formation-flying CubeSats are developed by the German research institute Zentrum für Telematik e.V. (ZfT) for photogrammetric observations of volcanic ash clouds. An alternative to achieve the required precise pointing is vision-based attitude control utilizing onboard cameras, also referred to as visual servoing. In previous work conducted at ZfT, a visual servoing approach for coordinated attitude control of all satellites in the TOM formation has already been developed and extensively tested. The goal of this approach is the joint tracking of the same target area to enable sufficient image overlap for photogrammetric postprocessing. Since the attitude of all satellites is controlled based on the pointing of a designated leader satellite, this method is referred to as *relative* visual servoing.

A new, modified *absolute* visual servoing approach enables the precise determination of the absolute attitude and position of a single satellite in the inertial frame with the help of the satellite's camera images. This approach will also be tested in the scope of the TOM mission. A database of ground control points (GCPs), i.e. points or structures on the surface of the Earth of known location and with unique appearance in satellite imagery, is created on ground and then stored onboard the satellite. By detecting the same points in the images taken by the satellite's camera, the relation between the camera frame and the Earth-fixed frame can be found. This information is used to calculate the absolute attitude and position of the satellite, which can for example be fed into a control law to calculate the control inputs for the satellite's reaction wheels. To optimize performance, reliability, database size and onboard runtime, different image processing algorithms to identify and track ground control points are investigated and evaluated. For verification of the onboard matching process, an Earth observation simulator generates realistic imagery of the Earth's surface.

Keywords: CubeSat, Attitude Determination, Attitude and Orbit Control System (AOCS), Visual Servoing, Formation Flying, Earth Observation

1. Introduction

As electronics, sensors, and actuators become more and more miniaturized and less expensive, many satellite missions can be realized by using multiple small and inexpensive satellites instead of one large expensive satellite. Certain applications can even only be realized by using formations or constellations of satellites, especially in the field of communications and Earth observation. An obvious

application of satellite formations is the observation of the Earth's surface from different perspectives in order to create three-dimensional images using photogrammetric methods. One such mission is the Telematics Earth Observation Mission (TOM), realized by the German research institute Zentrum für Telematik e.V. (ZfT), whose imagery is to enable the study and monitoring of volcanic ash clouds with the help of photogrammetric methods [1–3].

For the 3D reconstruction of the observed areas on the Earth's surface, sufficient overlap of the images acquired by the three TOM satellites is necessary, which in turn requires precise attitude control of all satellites in the formation. Therefore, an additional objective of TOM is the demonstration of a vision-based attitude control approach which uses the images taken by the onboard cameras to improve the pointing accuracy.

Vision-based control methods originate from the field of robotics and are referred to as *visual servoing*. Typical applications are positioning and tracking tasks of robots [4–6]. The usage of image-based methods has also been proposed for attitude determination [7–9] and/or attitude control [10–12] of satellites. Previous visual servoing developments regarding the TOM mission are described by Dauner et al. [13] and Elsner [14].

This publication is structured as follows: After the introduction, an overview of the TOM mission is given (section 2). The basic operating principles of the visual servoing approach are discussed in section 3. Additionally, algorithms and software as well as test environments are described. In section 4, the previously developed Relative Visual Servoing (RVS) method is briefly revisited. Section 5 outlines the general approach of Absolute Visual Servoing (AVS), and provides details on database creation, onboard matching, and attitude and position determination. In addition, the testing of the AVS approach is described in this section. The paper ends with a conclusion and outlook on next steps in section 6.

2. The TOM Mission

TOM is Bavaria's contribution to the international Telematics International Mission (TIM) project, which is founded by the Regional Leaders Summit (RLS). In the frame of this project, space research institutes from the different partner regions contribute their capabilities to form a Ground Station Network (GSNW) spanning five continents. [15]

The main objective of TOM is to create 3D imagery of the Earth by using photogrammetric methods. Primary observation targets are volcanic eruptions and the resulting ash clouds. Since the image data obtained by a single satellite does not allow to capture rapidly evolving structures and also suffers from the parallax effect phenomena, TOM uses a formation of three 3U+ CubeSats, which enables the simultaneous observation of a target area from different viewing angles. [3] The three satellites will form a triple pendulum formation with a baseline



Fig. 1: One of the three TOM satellites.

distance of about 100 km. Required formation flying technologies have already been investigated and demonstrated in the NetSat mission [16], and several approaches for formation control have been developed at Zft [1].

A photo of one of the three TOM satellites is shown in fig. 1. TOM accommodates an in-house developed, redundant camera system as payload. It consists of a global shutter sensor with a resolution of 2432 (H) × 2048 (V) and a ruggedized 100 mm focal lens. The camera system has a Field of View (FoV) of 4.8° and provides a Ground Sampling Distance (GSD) of 20.7 m and a swath of 50.3 km (at 600 km altitude). As payload computers for image capturing, storage, and onboard processing, two independent Linux single board computers (SBCs) are used. [1]

The Attitude and Orbit Control System (AOCS) used in the TOM satellites enables three-axis attitude control as well as orbit/formation control. It consists of the following components [1]:

- Four Sun sensors, integrated in the four large panels of the satellite
- Inertial Measurement Unit (IMU), providing three-axis gyro and magnetometer measurements
- Six miniature reaction wheels (two for each axis)
- Five Magnetorquers, distributed in the panels of the satellite
- Global Navigation Satellite System (GNSS)-based orbit determination
- Propulsion unit for orbit/formation control.

Since high-precision attitude sensors like star trackers are not used in TOM, precise pointing will be achieved by using visual servoing approaches, which are described in detail in the following sections.

3. Visual Servoing Fundamentals

3.1 Basic Operating Principles

Within the proposed work, the operating principles of visual servoing are based on the detection and matching of features in satellite imagery. Features can be considered as highly distinctive points in an image which differ from their immediate neighborhood and which can be detected repeatedly in different images showing the same scenery. To ensure this, in addition to uniqueness, the properties of a feature should be invariant to geometric changes (translation, scaling, or rotation), to photometric changes (changes in brightness or exposure), and to noise [17].

In order to detect and locate these points of interest in an image, feature detection algorithms search for areas with significant structural information and specific shape, such as edges, corners or blobs. This is done by defining a certain salience measure, such as pixel intensity gradient, and looking for local extrema across the image pixels. Fig. 2 exemplifies the detection of features in an image showing an area of southern Germany.

To be able to compare features detected in different images of the same scenery, each feature requires a compressed, unique signature. Therefore, feature description algorithms analyze the neighborhood of a feature and encode interesting information of these pixels into a descriptor vector acting as a numerical fingerprint. Just like the features, these descriptor vectors have to be robust to noise as well as to changes in viewpoint and in photometric imaging conditions [18].

In order to find corresponding features in two images, the two associated sets of descriptor vectors are compared during the feature matching process. After the matching, only features that occur in both sets should remain. This is done by calculating the Euclidean or Hamming distance between all feature descriptors of the two sets. Each vector from the first set is matched with its nearest neighbor of the second set regarding descriptor distance [18]. Fig. 3 shows the results of this matching process in which features of a small region (same as in fig. 2) are matched with features from a wider area. Since it is not guaranteed that every feature of the first set appears in the second set, mismatches with similar features which

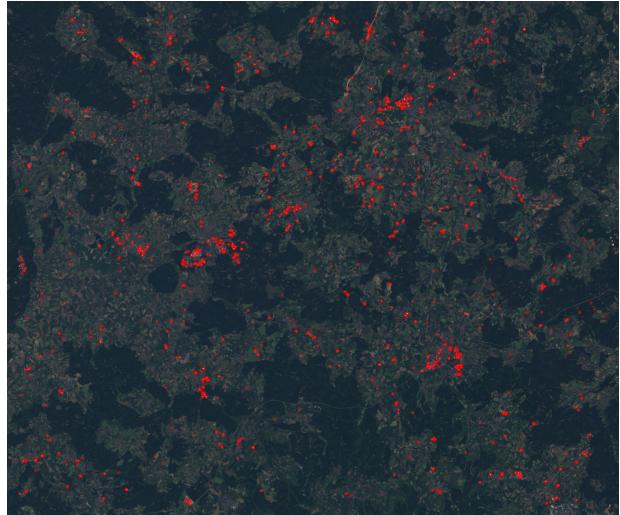


Fig. 2: Feature detection in an image showing parts of southern Germany. Earth imagery credit: Landsat 8 [20].

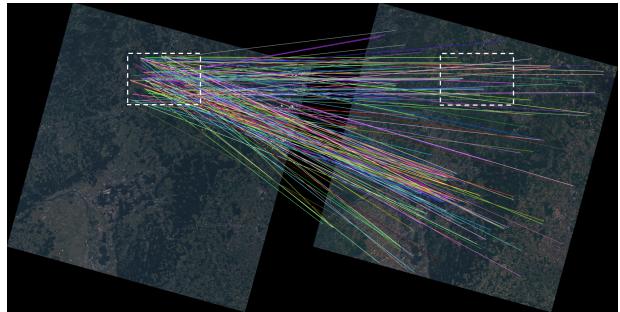


Fig. 3: Feature matching based on nearest neighbor regarding descriptor distance. Earth imagery credit: Landsat 8 [20].

are located at another position are unavoidable.

To discard these false matches and thus to provide a more reliable feature matching, additional checks need to be performed during feature validation. Therefore, the plausibility of all matches is evaluated based on different properties, e.g. by only selecting matches having a descriptor distance smaller than a certain threshold (fig. 4), and by checking the relative pixel positions of the matched features using the Random Sample Consensus (RANSAC) algorithm [19] (fig. 5).

3.2 Algorithms and Software

There are several algorithms for feature detection and description. The current state-of-the-art algorithms include SURF [21], ORB [22], BRISK [23]

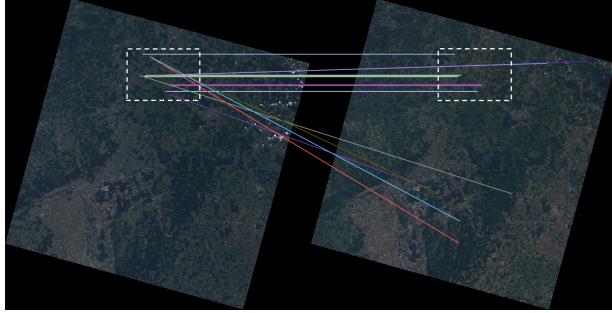


Fig. 4: Feature matching after applying threshold criterion. Earth imagery credit: Landsat 8 [20].

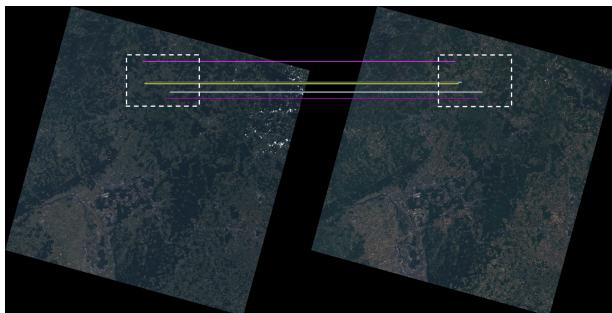


Fig. 5: Feature matching after validation based on RANSAC algorithm. Earth imagery credit: Landsat 8 [20].

and FREAK [24]. SURF, ORB and BRISK provide feature detection as well as description, whereas the FREAK algorithm only describes features.

Within the presented work, Open Source Computer Vision Library (OpenCV) is used for image processing. It is a cross-platform library which can be used to develop computer vision as well as machine learning applications. It supports Linux, Windows, Android and MacOS and provides interfaces for Python, C++, Java and Matlab. Furthermore, OpenCV offers implementations for all algorithms mentioned above. [25] Complementary to OpenCV, the cross-platform library Visual Servoing Platform (ViSP) provides visual feature tracking algorithms, e.g. the Lucas-Kanade-Tomasi (KLT) algorithm, which enables fast tracking of matched features for visual servoing applications [26].

3.3 Classification: Relative Visual Servoing (RVS) and Absolute Visual Servoing (AVS)

Visual servoing techniques for robots can be classified into Image-based Visual Servoing (IBVS) and Position-based Visual Servoing (PBVS) (and hybrid approaches) [4,5]. In IBVS, the control values are calculated directly based on the detected features. In contrast, in PBVS, a geometric model of the observed target or scenery, the camera model and the features found in the image are used to estimate the pose of the camera relative to the observation target.

In the presented work, a similar distinction is made. Relative Visual Servoing (RVS) is similar to IBVS: Tracked features in the current image frame are matched with given reference features, and the pixel offset between those is fed into the control law to calculate the torques to be applied in order to track the reference features with the camera.

On the other hand, Absolute Visual Servoing (AVS) is similar to PBVS: The detected features are matched with Ground Control Points (GCPs) with a known geographical location to calculate the pose of the camera and the resulting attitude and position of the satellite. These estimated values can then be used together with a reference attitude and/or position to calculate a attitude/position error, which is fed into the control law. In contrast to RVS, AVS can also be used as sensor for attitude and position determination without forwarding these values to a controller.

3.4 Test Environments

Three different test environments are used at ZfT to verify and improve the different visual servoing approaches.

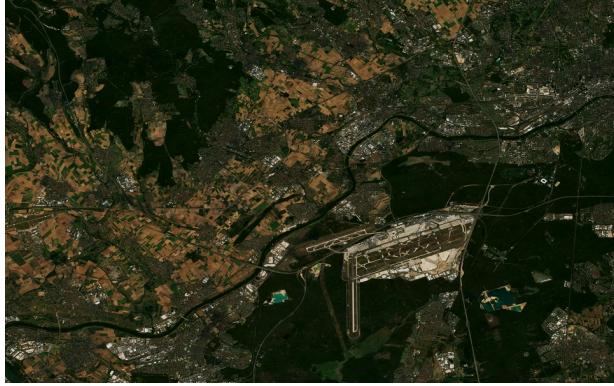


Fig. 6: Simulated camera view generated by the Earth Observation Simulator (EOS).



Fig. 7: Hardware-in-the-Loop (HiL) test setup using the Earth Observation Simulator (EOS).



Fig. 8: HiL turtable testbed, used for tests of Relative Visual Servoing (RVS) with two satellites.

The first test environment is solely software-based. Based on given orbit position and camera attitude, an EOS generates simulated imagery which represents the view of the satellite's camera. An example image is shown in fig. 6. The generated images are then fed into the visual servoing algorithms. The development of the EOS is described in detail by Friesen [27].

The second test environment is a HiL test setup which can even include the whole satellite. As depicted in fig. 7, the EOS is used in this setup to show simulated images on a screen. The satellite is positioned in front of the screen, and the camera is pointed towards the center of the screen.

The third test environment (see fig. 8) is a more complex HiL testbed: Two CubeSat models are mounted onto two high-precision motion simulators ("turntables"), which are rotated according to the attitude of the satellites. The observation target is moved by a mobile robot in order to replicate the relative orbit motion of the two satellites. The Inter Satellite Link (ISL) between the two satellites can be simulated via wi-fi or cable connection.

4. Relative Visual Servoing (RVS)

The Relative Visual Servoing (RVS) approach is already described in more detail by Dauner et al. [13] and Elsner [14], therefore only a brief summary is given here.

As already mentioned in section 3, the goal of RVS is to track given reference features. In TOM, the jointly observed target area of all satellites should be maximized. Therefore, one satellite is assigned as leader which defines which target area will be observed. The other satellites are designated as followers which have to track the area defined by the leader. The leader satellite sends the features it detected in the target area to the follower satellites as reference values for the tracking.

The RVS algorithm for the cooperative target tracking in TOM has four main processing steps:

1. Feature initialisation: The leader identifies reference features in the first two images. Their descriptors are sent via the ISL to the followers. The feature initialisation step is only performed at the start of a new observation or if the camera perspective has changed significantly.
2. Feature tracking: The reference features are matched with the features identified in the current image frame.
3. Error calculation: Based on the difference between the initial positions of the reference

features in the first image frame and their tracked positions in the current image frame, the attitude error angles are computed.

4. Controller: The error angles are fed into the control law to calculate the reaction wheel torques to achieve precise coordinated tracking of the target area.

The presented RVS algorithm has been extensively tested in the available test environments [13]. The root-mean-square errors (RMSEs) of the angles measured during tests on the turntable testbed (test environment 3, shown in fig. 8) are summarized in table 1. As can be seen in this table, the RMSEs of the angles stay well below 0.05° . A video with impressions from the performed HiL tests is available here: <https://www.youtube.com/watch?v=S-KHxH6iLVQ>.

	$\Delta\theta$	$\Delta\phi$	$\Delta\psi$
Leader [°]	0.011	0.023	0.046
Follower [°]	0.021	0.036	0.048

Table 1: Root-mean-square errors (RMSEs) of the angles measured during tests of the visual servoing algorithms for TOM on the turntable testbed.

5. Absolute Visual Servoing (AVS)

To extend the applicability of image-based attitude control algorithms to further Earth observation missions, based on the existing Relative Visual Servoing (RVS) approach, an absolute one is developed: Absolute Visual Servoing (AVS). This way, the precise pointing of a single satellite towards a specific location on Earth's surface should be enabled by determining the absolute attitude and position of the satellite based on images taken by the onboard camera.

5.1 General Approach

In AVS, features are detected in the incoming image frames taken by the onboard camera. Then, they are compared and matched with GCPs stored in an onboard database.

In general, GCPs are structures on the Earth's surface with unique appearance in satellite imagery and with known geographic position. In the context of AVS, these GCPs are features extracted by feature detection algorithms applied to satellite images in advance. To be able to determine their exact geographic position, the GCPs are detected in

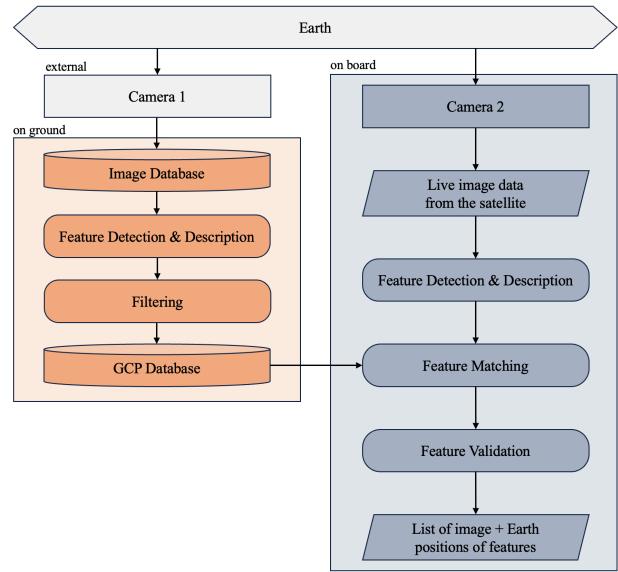


Fig. 9: Block diagram of the process of feature detection, description and matching in Absolute Visual Servoing (AVS).

satellite images with given metadata that provide useful information about the geographic position of the displayed area. Based on this information, the exact geographic position of the GCPs can be determined. Afterwards, each GCP gets a unique signature generated by a feature description algorithm. Subsequently, the set of GCPs is filtered, so that only the GCP features remain that are most likely to be matched onboard. The remaining descriptor vectors are stored in an onboard database mapped with their absolute geographic position on Earth.

After the feature matching and validation process onboard the satellite, the absolute attitude and position of the satellite can be determined based on the image positions and geographic positions of the matched GCP features in the current image frame. The process of feature detection, description and matching in AVS is illustrated in fig. 9. The three main steps of AVS (database creation, onboard matching and attitude and position determination) are described in detail in sections 5.2 to 5.4 respectively.

5.2 Database Creation

To create the GCP database, at first, a scene of the Worldwide Reference System-2 (WRS-2) is chosen that should be covered by the database. For this scene, two RGB-images generated based on the

Landsat 8 data [20] are selected. The Landsat 8 data packages are used due to the high quality, the free and easy access, and the detailed metadata provided. During selection, care is taken that the images have 0 % cloud coverage to ensure that the entire area on ground is visible.

Features are extracted from each of the two RGB-images and are described by the feature detection and description algorithm respectively. Next, the resulting features are filtered in order to only keep the ones that are most likely to be matched onboard as well as to reduce the total number of GCPs and thus the required memory space of the database. First, the descriptor vectors of the two images are matched. Second, for each match, the geographic positions of both features are determined and compared during the validation step. If their ground positions do not match, the matched features are discarded.

The remaining matches should only include features that are constant over time, can be detected in different images, and have a unique signature so they are recognized during the matching process. Therefore, these features are likely to also be detected and correctly matched onboard. The descriptor vectors and the geographic positions of the features are then stored in the database. The process is summarized in fig. 10.

5.3 Onboard Matching

As first step of onboard matching, features are detected in the RGB-images taken by the onboard camera and described by the same detection and description algorithms used to create the database. Next, these descriptor vectors are matched with the ones stored in the GCP database. Afterwards, the matches are validated based on whether the

descriptor distance exceeds a certain threshold or not. Subsequently, the remaining matches are validated by comparing the relative pixel positions of the matched features with their relative geographic positions on ground. This is done by applying a modified version of the RANSAC algorithm. These validation steps should guarantee that the resulting matches are correct and all mismatches are discarded. Fig. 11 illustrates this process. Afterwards, the image and geographic positions of the matched features are used to determine the attitude and position of the satellite.

5.4 Attitude and Position Determination

In order to calculate the attitude $(\theta, \phi, \psi)^{I \rightarrow B}$ and position $(\vec{x}, \vec{y}, \vec{z})^{I \rightarrow B}$ of the satellite and, thus, how the body-frame is shifted and rotated with respect

to the Earth Centered Inertial (ECI) frame, the following three steps need to be performed:

1. Determine position and attitude of the onboard camera with respect to the Earth Centered Earth Fixed (ECEF) frame ($E \rightarrow C$) based on the matched features.
2. Determine position and attitude of the satellite with respect to the ECEF frame ($E \rightarrow B$) based on the position and attitude of the camera.
3. Transform the position and attitude of the satellite from the ECEF frame to the ECI frame ($I \rightarrow B$).

The different coordinate frames are illustrated in fig. 12. The position and attitude of the onboard camera with respect to the ECEF frame can be computed using the collinearity equations [28]:

$$x^P = -k \frac{(x^E - x^{E \rightarrow C}) \cdot r_{11} + (y^E - y^{E \rightarrow C}) \cdot r_{12} + (z^E - z^{E \rightarrow C}) \cdot r_{13}}{(x^E - x^{E \rightarrow C}) \cdot r_{31} + (y^E - y^{E \rightarrow C}) \cdot r_{32} + (z^E - z^{E \rightarrow C}) \cdot r_{33}}, \quad [1]$$

$$y^P = -k \frac{(x^E - x^{E \rightarrow C}) \cdot r_{21} + (y^E - y^{E \rightarrow C}) \cdot r_{22} + (z^E - z^{E \rightarrow C}) \cdot r_{23}}{(x^E - x^{E \rightarrow C}) \cdot r_{31} + (y^E - y^{E \rightarrow C}) \cdot r_{32} + (z^E - z^{E \rightarrow C}) \cdot r_{33}} \quad [2]$$

with

$$R^{E \rightarrow C} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \quad [3]$$

$$\begin{bmatrix} \cos \phi \cos \psi & -\cos \phi \sin \psi & \sin \phi \\ \cos \theta \sin \psi + \sin \theta \sin \phi \cos \psi & \cos \theta \cos \psi - \sin \theta \sin \phi \sin \psi & -\sin \theta \cos \phi \\ \sin \theta \sin \psi - \cos \theta \sin \phi \cos \psi & \sin \theta \cos \psi + \cos \theta \sin \phi \sin \psi & \cos \theta \cos \phi \end{bmatrix}, \quad [4]$$

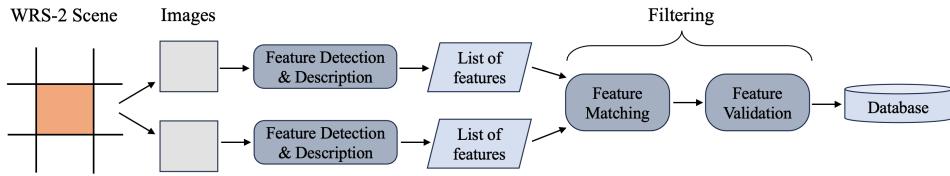


Fig. 10: Block diagram of creating the Ground Control Point (GCP) database.

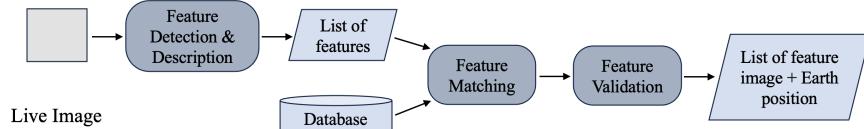


Fig. 11: Block diagram of onboard feature matching.

where k is a constant depending on the used camera. These equations link the pixel positions in the image (x^P, y^P) of the matched GCP features with their ground positions (ECEF frame; x^E, y^E, z^E). The components of the vector $[x^{E \rightarrow C}, y^{E \rightarrow C}, z^{E \rightarrow C}]^T$ represents the shift of the camera-frame with respect to the ECEF frame.

For each onboard matched feature, the two collinearity equations are set up to find the attitude and position of the camera $(x, y, z, \theta, \phi, \psi)^{E \rightarrow C}$. Six equations are required to enable the computation of all six parameters, leading to the need of at least three matched features. If more than three features are provided, the system is overdetermined and an approximate, best solution can be computed which results in a more accurate attitude and position.

The position and attitude of the satellite with respect to the ECEF frame $(x, y, z, \theta, \phi, \psi)^{E \rightarrow B}$ can be easily computed, as the relative attitude and position of the built in camera with respect to the body-frame is known.

The last step is to transform these parameters to the ECI frame which is a routine task for the Attitude Determination and Control System (ADCS) and thus is not described in detail.

5.5 Test Results

In order to achieve the goal of a reliable attitude and position determination by detecting and matching features stored in an onboard database, certain requirements must be met. First, the GCP features need to be equally distributed around the globe, so that the detection of a sufficient number of features within the FoV of the onboard camera is guaranteed. Second, the matched features must not contain any incorrect matches, in order to enable a

precise attitude and position determination. Third, due to the limited resources onboard the satellite, the required memory space of the database must be kept as low as possible. Fourth, since the determination of the satellite's attitude and position is a time-critical task, the onboard matching needs to be performed in real-time. The test results presented here are also described and discussed in more detail by Redelbach [29].

To analyze the applicability of the process of creating the GCP database as well as of the onboard matching with respect to these requirements, a GCP database is created covering approximately the area of Germany (see fig. 13; approx. 550 km \times 700 km). The onboard matching is performed using six test images, each displaying a different region on ground. The size of the test images is chosen based on the FoV of the camera integrated in TOM. The RGB-images for creating the database as well as for the onboard matching are generated based on the Landsat 8 data package [20].

Each test scenario of the onboard matching is evaluated based on the number of remaining matches, the percentage of correct matches and the runtime. Additionally, the required memory space of the database is determined for the covered area and estimated for the entire land surface of the Earth. Furthermore, to analyze the impact of the used algorithms, the entire process of database creation and onboard matching is performed using six different combinations of feature detection and description algorithm: ORB-ORB, ORB-FREAK, BRISK-BRISK, BRISK-FREAK, SURF-SURF and SURF-FREAK. The test scenarios are performed using an Apple MacBook Pro 14" 2021 with an Apple M1 Pro chip and 16 GB unified memory. The results

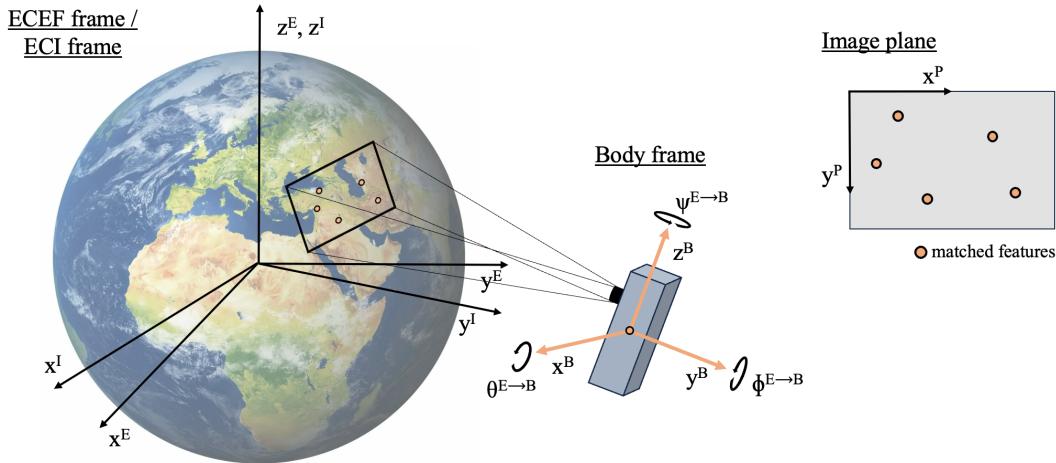


Fig. 12: Illustration of the different reference frames used to determine the attitude and position of the satellite based on the onboard matched features.

are summarized in table 2.

It gets clear that the used detection and description algorithms strongly influence the results of creating the GCP database as well as of the onboard matching. However, all combinations achieve on average a sufficient number of matches (see section 5.4; ≥ 3) enabling the attitude and position determination. Furthermore, besides the ones using the SURF detector, all combinations provide a reliable onboard feature matching by only selecting correct matches in all test scenarios. The required memory space of the database depends on the number of stored GCPs on the one hand, and on the required memory space of a single descriptor vector on the other hand. Both factors are mainly influenced by the used detection and description algorithm, explaining the great differences of the storage needed. However, it can be noted that a database covering the entire land surface of the Earth probably needs a high amount of memory space using the current setup (250-900 MB) independent of the chosen algorithms, which must be considered when designing the satellite. The total runtime of the onboard matching process depends on the runtime of the four performed steps: detection, description, matching and validation. Since the runtimes of these steps, again, are mainly influenced by the detection and description algorithm, the results regarding this requirement differ greatly as well. It can be seen that the onboard matching using the current setup has limited real-time capability for all combinations. In addition, matching the detected features with the stored GCPs would take even longer using a database

that covers the entire land surface of the Earth.

In the described test setup, two factors are not taken into account which could have an impact on the performance of the onboard matching. First, the RGB-images used to create the database and the ones used to test the onboard matching were recorded by the identical camera onboard the Landsat 8 satellite. This will not be the case in future missions. Second, the used Landsat 8 data was recorded when the satellite has been in nadir pointing mode. However, the attitude and position determination needs to be successfully done independent of the current attitude of the satellite.

Therefore, imagery generated by the EOS is used to test the onboard matching process. The test image shows the region around Frankfurt airport at an elevation angle of 60 deg. Furthermore, a small database using the ORB-ORB combination is created covering only one WRS-2 scene that displays an area of southern Germany including the region around Frankfurt airport. The generated database contains 3750 features. As can be seen in fig. 14 and fig. 11, the feature detection and the onboard matching are successfully performed resulting in 13 matches, which is a sufficient number to compute the satellite's attitude and position (as described in section 5.4).

6. Conclusion and Outlook

The present work describes the initial developments of an AVS approach for satellite attitude and position determination. The first tests show promising results. Nevertheless, additional

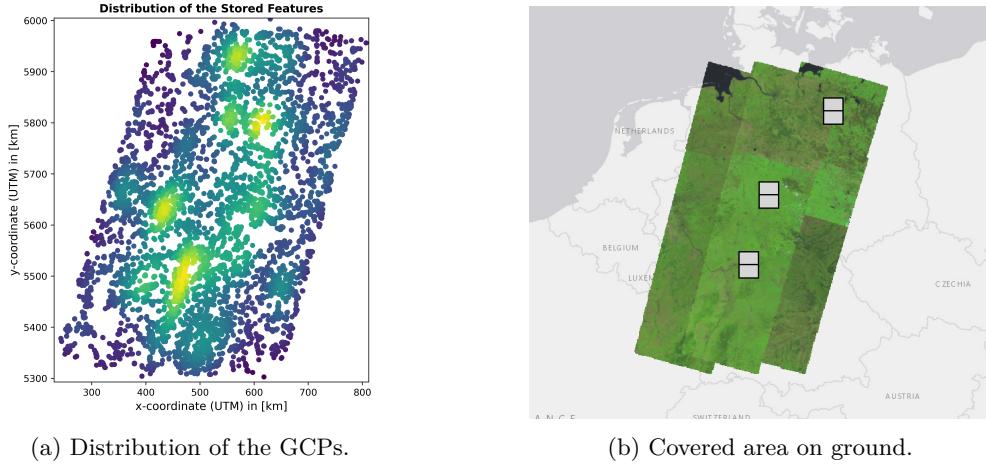


Fig. 13: Illustration of the Ground Control Points (GCPs) stored in the database covering approximately the area of Germany. In (a), each datapoint illustrates a GCP. Its color represents the density of the GCPs: The brighter the color, the more GCPs are located within this region. Image in (b) is taken from Earth Explorer [30] representing the covered area of the database. The marked regions illustrate the area displayed in the six test images.

Property	Detector-Descriptor					
	ORB-ORB	ORB-FREAK	BRISK-BRISK	BRISK-FREAK	SURF-SURF	SURF-FREAK
Average number of matches	28.5	22.3	21.3	31.3	10.5	11.0
Reliability	100 %	100 %	100 %	100 %	83.3 %	83.3 %
Descriptor size [B]	32	64	64	64	256	64
Number of stored GCPs (Germany)	27740	24372	28162	27309	8771	9405
Database size (Germany) [MB]	1.1	1.8	2.0	2.0	2.3	0.7
Projected number of stored GCPs (Earth's land surface)	$10.8 \cdot 10^6$	$9.5 \cdot 10^6$	$10.9 \cdot 10^6$	$10.6 \cdot 10^6$	$3.4 \cdot 10^6$	$3.7 \cdot 10^6$
Projected database size (Earth's land surface) [MB]	432.0	684.0	784.8	763.2	897.6	266.4
Runtime [ms]	559.96 ± 456.13	584.81 ± 457.30	197.39 ± 189.03	744.79 ± 444.94	131.85 ± 168.23	391.76 ± 404.04

Table 2: Summary of the test results using different combinations of detection and description algorithms. „Average number of matches“ indicates the average number of remaining matches after onboard matching over the six test scenarios. „Reliability“ indicates the percentage of scenarios during which all remaining matches are correct. „Descriptor size“ and „Database size“ means the required memory space of one descriptor vector and the onboard database respectively. „Number of stored GCPs“ indicates the number entries in the database. „Runtime“ specifies how long the onboard matching takes on average over the six scenarios.



Fig. 14: Feature detection in a test image generated by the Earth Observation Simulator (EOS).

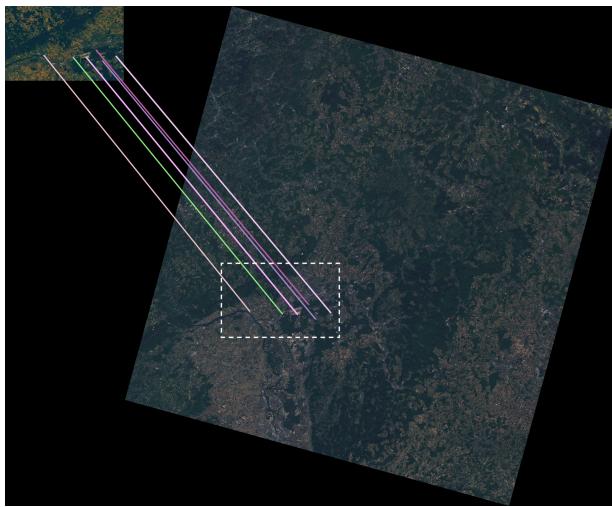


Fig. 15: Feature matching after validation using a test image generated by the Earth Observation Simulator (EOS) (top left). Earth imagery credit (right image): Landsat 8 [20].

simulations and tests need to be performed to evaluate the method in more detail. For example, the influences of cloud coverage and atmospheric distortions are challenges which have yet to be solved. Especially dense cloud coverage of the target observation area poses a big challenge to the absolute visual servoing approach, since the number of detected features could then not be sufficient to calculate attitude and position of the satellite. In addition, more tests with images originating from other sources have yet to be executed to optimize the different steps of the AVS algorithm.

The presented approach is planned to be demonstrated using onboard images taken by the TOM satellites. The TOM CubeSat formation is to be launched in 2024. Therefore, the AVS approach will be further developed and tested in the upcoming months.

Acknowledgments

The authors thank all the collaborators for their contributions within the Telematics Earth Observation Mission (TOM), supported by the Bavarian Ministry of Economics and the Regional Leaders Summit (RLS) for the cooperation in the Telematics International Mission (TIM). In addition, we gratefully acknowledge the support of the Elite Network Bavaria (ENB) for providing funds for the academic degree program “Satellite Technology”.

References

- [1] A. Kleinschrodt, I. Mammadov, E. Jäger, J. Dauner, J. Scharnagl, and K. Schilling, “TOM/TIM Realization-A satellite Earth Observation (EO) formation flying mission of three nano satellites for retrieving multi view stereoscopic data,” in *Proceedings of the 73rd International Astronautical Congress*. Paris, France: IAF, 2022.
- [2] K. Schilling, T. Tzschichholz, I. Motroniuk, A. Aumann, I. Mammadov, O. Ruf, C. Schmidt, N. Appel, A. Kleinschrodt, S. Montenegro *et al.*, “Tom: A formation for photogrammetric earth observation by three cubesats,” in *4th IAA Conference on University Satellite Missions*, Roma, 2017.
- [3] K. Zakšek, M. R. James, M. Hort, T. Nogueira, and K. Schilling, “Using picosatellites for 4-d imaging of volcanic clouds: Proof of concept using iss photography of the 2009 sarychev peak

- eruption,” *Remote Sensing of Environment*, vol. 210, pp. 519–530, 2018.
- [4] S. Hutchinson, G. D. Hager, and P. I. Corke, “A tutorial on visual servo control,” *IEEE transactions on robotics and automation*, vol. 12, no. 5, pp. 651–670, 1996.
- [5] F. Chaumette and S. Hutchinson, “Visual servo control. i. basic approaches,” *IEEE Robotics & Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.
- [6] C. Cai, “6d visual servoing for industrial manipulators applied to human-robot interaction scenarios,” 2017.
- [7] T. Kouyama, A. Kanemura, S. Kato, N. Imamoglu, T. Fukuhara, and R. Nakamura, “Satellite attitude determination and map projection based on robust image matching,” *Remote Sensing*, vol. 9, no. 1, 2017. [Online]. Available: <https://www.mdpi.com/2072-4292/9/1/90>
- [8] S. Koizumi, Y. Kikuya, K. Sasaki, Y. Masuda, Y. Iwasaki, K. Watanabe, Y. Yatsu, and S. Matsunaga, “Development of attitude sensor using deep learning,” 2018.
- [9] A. Dagvasumberel and K. Asami, “A visual-inertial attitude propagation for resource constrained small satellites,” *Journal of Aeronautics and Space Technologies*, vol. 12, no. 1, pp. 65–74, 2019.
- [10] G. Klančar, S. Blažič, D. Matko, and G. Mušič, “Image-based attitude control of a remote sensing satellite,” *Journal of intelligent & robotic systems*, vol. 66, no. 3, pp. 343–357, 2012.
- [11] M. Hladký, “Vision based attitude control,” Master’s thesis, 2018.
- [12] M. Robic, K. Lagadec, R. Fraisse, E. Marchand, and F. Chaumette, “Visual servoing of an earth observation satellite of the lion constellation,” 09 2022.
- [13] J. Dauner, L. Elsner, O. Ruf, D. Borrmann, J. Scharnagl, and K. Schilling, “Visual servoing for coordinated precise attitude control in the TOM small satellite formation,” *Acta Astronautica*, vol. 202, pp. 760–771, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0094576522005409>
- [14] L. Elsner, “Development and integration of cooperative vision-based attitude control for the TOM mission,” Master’s thesis, University of Würzburg, 2022.
- [15] K. Schilling, G. Loureiro, Y. Zhang, A. Nüchter, J. Scharnagl, I. Motroniuk, and A. Aumann, “TIM: An international nano-satellite formation for photogrammetric earth observation,” in *Proceedings of the 70th International Astronautical Congress*, 2019.
- [16] J. Scharnagl, R. Haber, V. Dombrovski, and K. Schilling, “Netsat—challenges and lessons learned of a formation of 4 nano-satellites,” *Acta Astronautica*, vol. 201, pp. 580–591, 2022.
- [17] T. Tuytelaars, K. Mikolajczyk *et al.*, “Local invariant feature detectors: a survey,” *Foundations and trends® in computer graphics and vision*, vol. 3, no. 3, pp. 177–280, 2008.
- [18] F. Alhwarin, “Fast and robust image feature matching methods for computer vision applications,” Ph.D. dissertation, Bremen, Universität Bremen, Diss., 2011, 2011.
- [19] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [20] U. S. G. Survey, “Landsat 8,” (online; accessed 12-September-2023). [Online]. Available: <https://www.usgs.gov/landsat-missions/landsat-8>
- [21] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” *Lecture notes in computer science*, vol. 3951, pp. 404–417, 2006.
- [22] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “Orb: An efficient alternative to sift or surf,” in *2011 International conference on computer vision*. Ieee, 2011, pp. 2564–2571.
- [23] S. Leutenegger, M. Chli, and R. Y. Siegwart, “Brisk: Binary robust invariant scalable keypoints,” in *2011 International conference on computer vision*. Ieee, 2011, pp. 2548–2555.
- [24] A. Alahi, R. Ortiz, and P. Vandergheynst, “Freak: Fast retina keypoint,” in *2012 IEEE conference on computer vision and pattern recognition*. Ieee, 2012, pp. 510–517.

- [25] OpenCV-Team. (2020) About-opencv. [Online]. Available: <https://opencv.org/about/>
- [26] ViSP-Team, “Open source visual servoing platform library,” 2020. [Online]. Available: <https://visp.inria.fr/>
- [27] D. Friesen, “Development of an earth observation simulator,” Master’s thesis, University of Würzburg, 2023.
- [28] T. Schenk, “Introduction to photogrammetry,” *The Ohio State University, Columbus*, vol. 106, no. 1, 2005.
- [29] J. Redelbach, “Analysis of feature detection and matching algorithms for absolute visual servoing,” Bachelor’s thesis, University of Würzburg, 2023.
- [30] U. S. G. Survey, “Earth explorer,” (online; accessed 14-September-2023). [Online]. Available: <https://earthexplorer.usgs.gov/>