EXECUTIVE SUMMARY REPORT

ATENA

Prepared by:	ATENA Team
Approved by:	Marcos Avilés
	Digitally signed by Marcos Avilés Rodrigálvarez Date: 2022.06.13 12:14:25 +02'00'
Authorized by:	Marcos Avilés

Code:	GMV-ATENA-EXR
Version:	1.0
Date:	24/05/2022
Internal code:	GMV 22087/22 V1/22

This study was commissioned by the European Defence Agency jointly with the European Space Agency in response to the invitation to tender No.18.ESI.OP.006. The study does not, however, express the Agencies' official views. The views expressed and all recommendations made are those of the authors.

DOCUMENT STATUS SHEET

Version	Date	Pages	Changes
1.0	24/05/2022	26	First release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1. INTRODUCTION

1.1. PURPOSE AND SCOPE

The purpose of this document is to summarise the work carried out during the course of the ATENA activity and summarise its results.

This work has been performed under the ESA contract No. 4000133392/21/F/CP. The consortium participating in this project was formed by GMV Aerospace and Defense as prime, and AIKO as subcontractors.

1.2. DEFINITIONS AND ACRONYMS

1.2.1. DEFINITIONS

Concepts and terms used in this document and needing a definition are included in the following table:

Table 1-1 Definitions

Concept / Term	Definition

1.2.2. ACRONYMS

Acronyms used in this document and needing a definition are included in the following table:

Table 1-2 Acronyms

Acronym	Definition
AFC	Asteroid Framing Camera
AG	Attitude Guidance
AIM	Asteroid Impact Mission
ADCS	Attitude Determination Control System
AMF	Actuator Management Function
AOCS	Attitude and Orbit Control Systems
APE	Absolute Pointing Error
ASW	Application Software
CNN	Convolutional Neural Network
COM	Center of Mass
COP	Close Observation Phase
COTS	Commercial Off-The-Self
DCP	Detailed Characterisation Phase
DKE	Dynamic-Kinetic Environment
DRM	Didymos Reference Model
DSA	Deep Space Antenna
ECP	Early Characterisation Phase
ECSS	European Cooperation for Space Standardization
ELP	End-of-Life Phase
EOP	Extended Operations Phase
ERN	Enhanced Relative Navigation
ESA	European Space Agency
EXP	Experimental Phase
FDIR	Fault Detection, Isolation, and Recovery

Acronym	Definition
FOV	Field Of View
GNC	Guidance, Navigation and Control
HGA	High Gain Antenna
HIL	Hardware In the Loop
LEOP	Launch and Early Orbit Phase
LPC	Launch Period Closure
LPO	Launch Period Opening
MC	Monte Carlo
MIL	Model In the Loop
MPE	Mean Performance Error
MVM	Mode Vehicle Management
OBC	On-Board Computer
OBSW	On-Board Software
PDP	Payload Deployment Phase
RDV	Rendezvous
RPE	Relative Pointing Error
SADM	Solar Array Drive Mechanisms
SOW	Statement Of Work
TBC	To Be Confirmed
TBD	To Be Defined
TBW	To Be Written
VCFB	Very Close Fly-By

2. REFERENCES

2.1. APPLICABLE DOCUMENTS

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority. They are referenced in this document in the form [AD.x]:

Table 2-1 Applicable Documents

Ref.	Title	Code	Version	Date
[AD.1]	Appendix 1- SOW AO/1-10356/20/F/CP	ESA-TECSAG-SOW-018784	1.0	11/06/2020

2.2. REFERENCE DOCUMENTS

The following documents, although not part of this document, amplify or clarify its contents. Reference documents are those not applicable and referenced within this document. They are referenced in this document in the form [RD.x]:

Table 2-2 Reference Documents

Ref.	Title	Code	Version	Date
[RD.1]	ATENA Assessment of AI Techniques	GMV-ATENA-TN-01	1.0	21/03/2021
[RD.2]	ATENA Algorithm Design and Analysis	GMV-ATENA-TN-02	1.0	17/06/2021
[RD.3]	ATENA: Performance Analysis Report	GMV-ATENA-TN-03	1.0	22/11/2021
[RD.4]	ATENA Validation Report	GMV-ATENA-TN-04	2.0	17/06/2021
[RD.5]	HERA: Didymos Reference Model	ESA-TECSP-AD-017258	Issue 5, Revision 4	12/03/2020
[RD.6]	M. Dusmanu.I. Rocco, T. Pajdla, M. Pollefeys, J. Sivic, A. Torii and T. Sattler: D2-Net: A Trainable CNN for Joint Description and Detection of Local Features. Proceedings Computer Vision and Pattern Recognition, 2019			2019
[RD.7]	D.G. Lowe: Object recognition from local scale-invariant features. International Conference on Computer Vision, Corfu, Greece, pp. 1150–1157			1999

3. INTRODUCTION

Space missions benefit greatly by the capability of the on-board GNC system to adapt rapidly to unknown environment. Autonomous vision-based navigation is a particular technology under implementation in several ESA missions. One of the most interesting applications in that field is the proximity operations around a small asteroid, like those in HERA.

The goal of this activity was to develop a navigation algorithm with the capability to fly over an unknown terrain and achieve better navigation performances than current vision-based techniques based on unknown feature tracking.

More specifically, the following objectives needed to be achieved:

- To design a navigation architecture that uses techniques for image processing and a navigation filter that performs data fusion of the IP output and the rest of the GNC sensors.
- To develop a training strategy that uses limited a priori knowledge of the terrain (only a coarse shape model is known at arrival to the target) to be able to learn keypoints and demonstrate generalization of the performances (different distances, surface characteristics, viewing orientation).
- To validate the performances of the autonomous navigation system in a high-fidelity MIL testbench considering the representative proximity operations of HERA during very low altitude fly-bys.
- To demonstrate the capability of the technique to achieve good navigation results in a reduced-scale mock-up of binary asteroid Didymos using real images acquired at a HW facility.

To further expand the field of options, different IP approaches were explored. The first one, using conventional IP techniques. The second one, using more recent Artificial Intelligence (AI) based techniques. Current AI techniques like Convolutional Neural Networks (CNN) open new possibilities which needed to be evaluated. The third one, a Hybrid algorithm, using conventional IP techniques fused with AI techniques.

The chosen project reference scenario has been the Very Close Fly-By phase of the HERA mission. It is however important to remark that none of the techniques studied during this activity were designed for this particular mission. HERA was considered as a very realistic test case.

3.1. HERA MISSION OVERVIEW

An overview of the mission is given in Figure 3-1. DART will be launched in 2021 and impact Dimorphos in September 2022. HERA will be launched in October 2024 and will arrive at Didymos on January 2027. The early characterisation phase of Didymos will start from distances between 20 to 30 km to determine the shape and the gravity field. The detailed characterisation phase will be conducted from about 10-20 km distance. During this phase the CubeSats will be released. Very close flybys of Dimorphos are envisioned at the end of the mission.

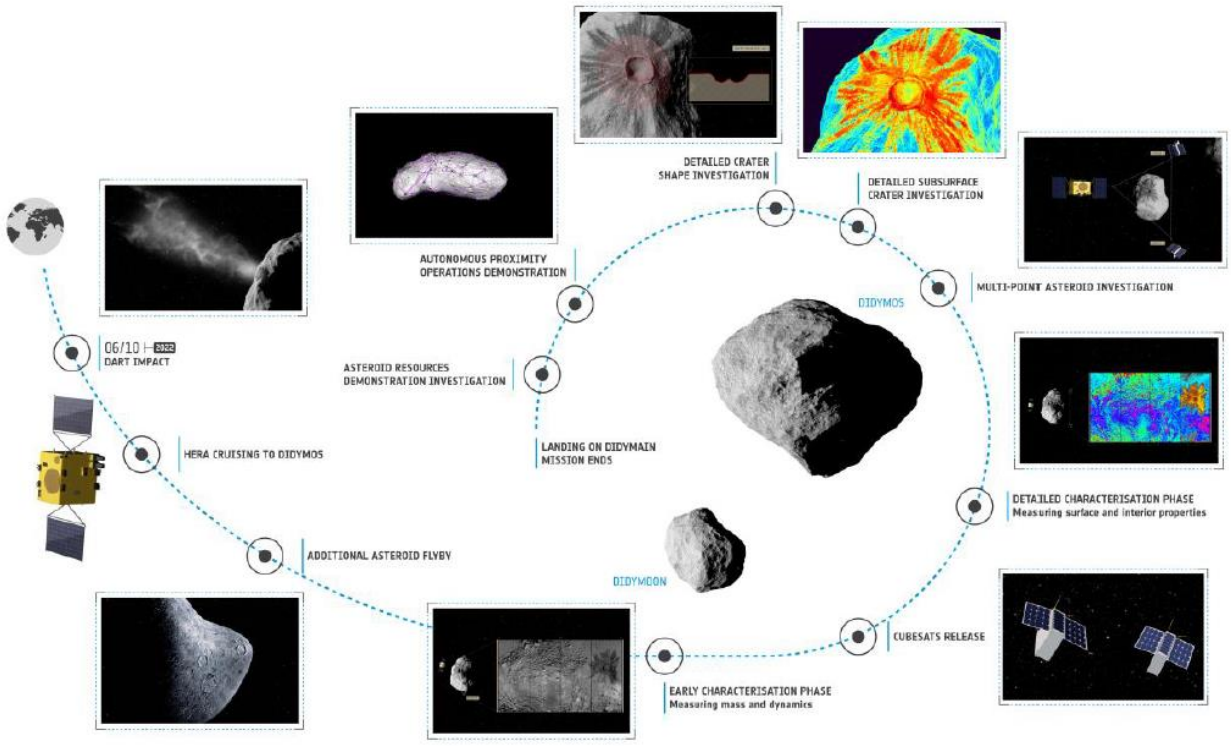


Figure 3-1: AIDA Architecture.

The end-of-mission currently foresees an attempt to land the HERA spacecraft in the polar region of Didymos. The CubeSats will try landing on Dimorphos at the end-of-mission.

The DART mission will demonstrate that the technology to deflect an asteroid by kinetic impact is available, in particular the terminal guidance system. HERA will characterize an asteroid of the size range most important for deflection. DART and HERA together will allow to quantify the deflection and to enable the application of the results to other asteroids, therefore fully validating the technique. This is critical information, a mandatory

4. ALGORITHM DESIGN

4.1. CONVENTIONAL IP/HYBRID ENHANCED RELATIVE NAVIGATION

A high-level diagram depicting the blocks of the Enhanced Relative Navigation (ERN) is provided in Figure 4-1. From the image acquired by the spacecraft camera a first stage of *Feature Extraction* is executed. The feature extraction comprises two steps: a feature detection step, where salient regions of the image are identified, and a feature description step, where a compact vector of values describing the area surrounding the feature is built. This vector, also called, descriptor, allows comparing and matching features in different images. The output of the Feature Extraction block is a list of features (together with their descriptors) that is compared with a reference landmark database within the *Landmark Matching* block. The landmark database stores the 3D coordinates of the landmarks together with a feature descriptor. The result of the matching process is a list of correspondences between the features detected in the camera image and features in the database. These 2D-3D correspondences are used in the *Pose Estimation* block to estimate the position and the orientation of the camera with respect to the asteroid.

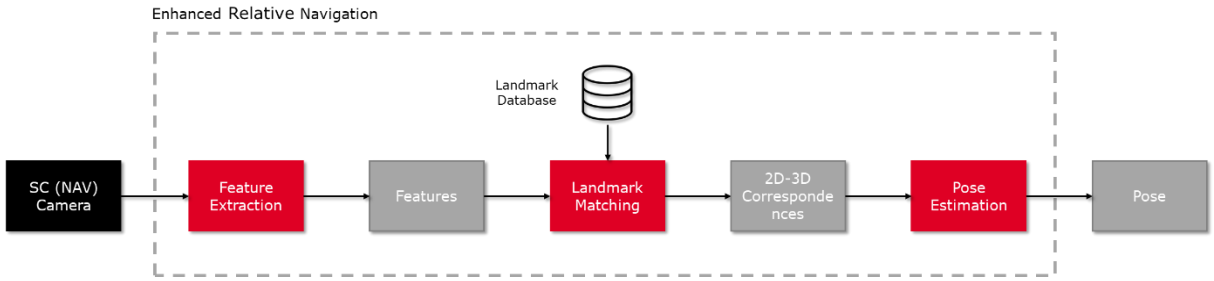


Figure 4-1: Enhanced Relative Navigation

As the reader can notice, a very important stage is the creation of the Landmark Database, which needs to cover the areas of the asteroid that are to be navigated and with a high accuracy so that the Navigation can use it to reset the accumulated error resulting from the propagation of relative measurements.

4.1.1. ERN DATABASE GENERATION

The generation of the landmark database is an important stage as it largely affects the accuracy of the ERN solution but also its feasibility (when the ERN can be called)

Two main strategies are considered, the first one using directly the images acquired on-board with the associated navigation, and the second using the images acquired for Ground navigation and orbit determination, together with their navigation.

4.1.1.1. DATABASE GENERATION USING ON-BOARD IMAGES

Given the relative dynamics between the spacecraft and the asteroid, in which the second is observed multiple times during different passes, the most straight forward approach is to generate the database directly in one pass and use this database in later passes.

This approach is simple to implement, as there are no dependencies with Ground. However, it is subject to the higher error of the on-board navigation (which will ultimately affects the accuracy of the ERN as it mostly recovers the error up to the point in which the database was created).

4.1.1.2. DATABASE GENERATION USING GROUND NAVIGATION IMAGES

The second approach is to use the navigation images acquired for orbit determination together with its navigation solution. The navigation images are acquired every hour during periods of 16 hours (the remaining 8 hours are the communication window used to download them to Ground).

The database produced using this method would have a considerably lower error (since the on-Ground navigation solution will be more accurate) and hence the performances of the ERN would be much better. On the other hand, the dependencies with Ground (data needs to be uploaded to the spacecraft) makes it more difficult to implement at mission level.

The number of points on the trajectory in which the ERN could be executed might also be lower as it would be on those previously observed in a navigation image and with comparable conditions. The comparable conditions refer to:

- Asteroid view angle: The area of the asteroid must be seen with a similar incidence angle.
- Ground sample distance: The area of the asteroid must be seen with a similar resolution.
- Illumination conditions: The area of the asteroid must be seen under similar illumination conditions.

4.1.2. CONVENTIONAL-IP VS HYBRID

Both the Conventional-IP and Hybrid Enhanced Relative Navigation implementation share most of the design described before. The difference is in the Feature Extraction stage (and associated Matching step, which takes into account the respective descriptor). While the Conventional-IP uses the widely used SIFT feature detector [RD.7], the Hybrid implementation uses an AI-based feature detector. The goal of the Hybrid solution is to evaluate novel solutions that could have a better performance under other challenging conditions where SIFT performs worse (such as with different illumination conditions).

SIFT algorithm detects features as the local maxima or minima across scales of the Difference of Gaussians (DoG), which is, coarsely speaking, the scale derivative of the Gaussian scale space. This allows detecting features which are invariant to scale change

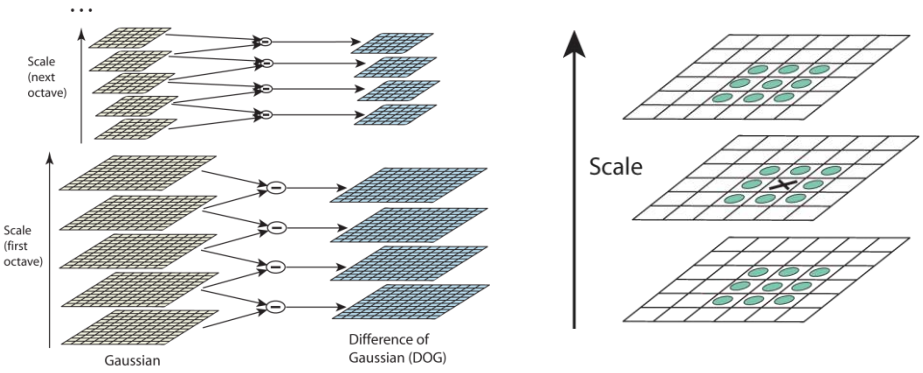
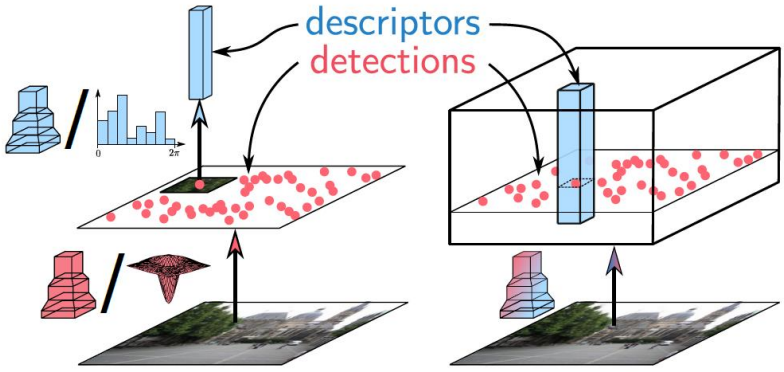


Figure 4-2: Computation of Difference of Gaussians (left); Maxima and minima are obtained by comparing with neighbors in at current and adjacent scales (right)

Apart from the detector, SIFT algorithm proposes its own strategy for matching the detected features in scenarios involving rotation, translation, scaling and, to a lower extend, to small skew or perspective transformations.

In the case of the Hybrid ERN, the feature extraction component is based on the D2-Net CNN network proposed by M. Dusmanu et. al. [RD.6]. This method proposes a *describe-and-detect* approach (rather than the classical *detect-then-describe* approach) to sparse local feature detection and description: Rather than performing feature detection early on based on low-level information, the detection stage is postponed. A set of feature maps is first computed via a Deep Convolutional Neural Network. These feature maps are then used to compute the descriptors (as slices through all maps at a specific pixel position) and to detect keypoints (as local maxima of the feature maps). As a result, the feature detector is tightly coupled with the feature descriptor. Detections thereby correspond to pixels with locally distinct descriptors that should be well-suited for matching. As a result, the proposed algorithm follows a *describe-and-detect* approach which is able to detect keypoints belonging to higher-level structures and locally unique descriptors.



(a) detect-then-describe (b) detect-and-describe

Figure 4-3: Comparison between different approaches for feature detection and description. Pipeline (a) corresponds to different variants of the two-stage detect-then-describe approach. In contrast, our proposed pipeline (b) uses a single CNN which extracts dense features that serve as both descriptors and detectors. Image courtesy of M. Dusmanu [RD.6].

4.1.3. NAVIGATION

The translational navigation is based on the hybrid EKF/UKF filter developed by GMV in several previous projects. The scheme block of the navigation algorithm is depicted in Figure 4-4. The propagation of the state is performed within the EKF part of the filter in order to reduce the computational load, while the measurement update is performed through an unscented transform to account for the highly nonlinear measurement equations. The filter also contains a state reconfiguration mechanism that is used to manage the addition and removal of features to the filter state. Furthermore, this mechanism is used in case an erroneous measurement is detected through a Mahalanobis check. The propagation step is executed at a frequency of 1 Hz and the measurement update is performed whenever a new image is available, currently assumed value for HERA is: every 48s. Since the measurement is provided with a delay (due to image transmission and image processing times), the navigation state is estimated N seconds behind the current time, and propagated forward to provide an estimate available for use by the guidance and control algorithms.

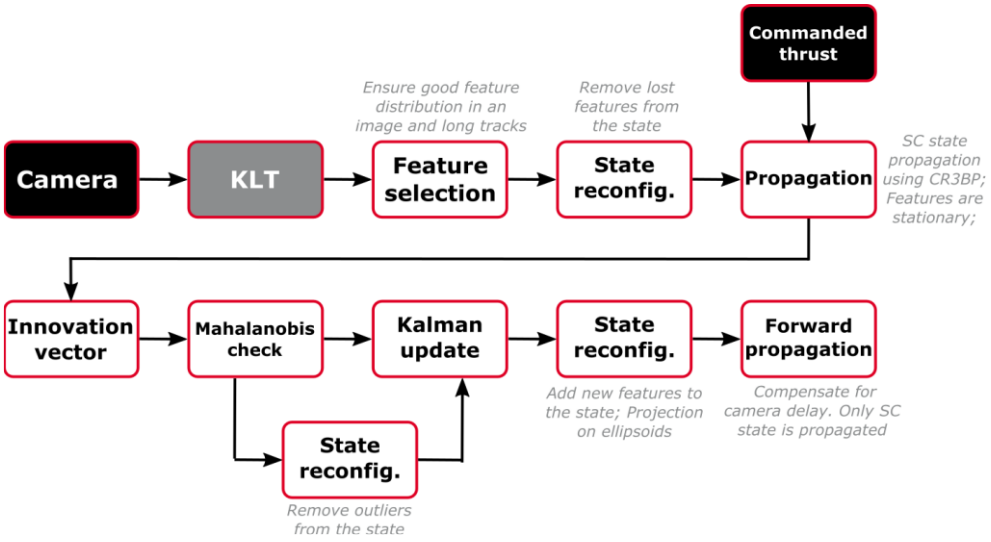


Figure 4-4 Translational navigation algorithm.

The navigation filter has also the ability to fuse the information coming from an altimeter. When available, the altimeter measurements can be used to estimate the bias of the ellipsoidal model of the asteroid with respect to the surface and provide a range estimate. In order to use the altimeter, the state is augmented with additional state – the bias of the altimeter.

Current implementation allows for the altimeter measurements to be processed at longer intervals, compared to the image measurements. For example, the altimeter measurement can be available every 10 images, as long as they are synchronized with the images. Processing of asynchronous measurements is currently not supported, however it can be implemented in future updates of the navigation system design.

The hybrid UKF filter used for the navigation filter with unknown features is justified due to the non-linearity of the measurements being fed to the filter. The shape model used to translate pixel coordinates into Cartesian coordinates and vice versa introduces non-linear derivatives that are complicated to handle by means of an EKF. However, the Enhanced Relative Navigation is directly providing a position vector with its covariance. Thus, the measurements are purely linear and the UKF scheme is not justified.

To ease the formulation, an EKF was proposed for the ERN. The position vector provided by the ERN is treated as a direct measurement to estimate the complete state vector (position and velocity) of the spacecraft. Thus, the velocity is also corrected during the measurement updates by means of the previously correlation introduced by the dynamics propagation.

4.2. AI-BASED SOLUTION

Among algorithms that were evaluated, SuperGlue and Patch2Pix algorithms have been selected as full learnable matching pipelines for four reasons:

1. These algorithms do not require annotation/labelling of input dataset. Therefore, they can be used/trained effortlessly using only raw input dataset.
2. It is proven in the literature that they show state-of-art results on benchmark datasets such as MegaDepth, ScanNet datasets. ...
3. It is also proven that they are embeddable thanks to their integration into modern SfM/SLAM systems for real-time scenarios without any problem.
4. Even though these two algorithms use full learnable matching pipelines, they follow different procedures when matching proposals: pixel-level and patch-level. Therefore, it is advantageous to understand which one is more suitable for an asteroid-based scenario.

As an initial step, all hyperparameters are chosen the same as the original ones for consistency. After this training step, it has been planned to change these hyperparameters according to our needs.

For the description of the AI algorithms, some preliminary definitions are required in order to differentiate the approach from the Hybrid methods that will be proposed in next sections.

In the case of an end-to-end AI algorithm, Feature Tracking (FT), recalled in this case as keypoint detection and feature matching between two input images is done as follows:

- Either SIFT or SuperPoint Network is used to extract keypoints, their descriptive features and generated ground truth (location info of these features - can be named as annotations) for the SuperGlue model. Then, a Graph Neural Network (GNN) is used to match all extracted features between input images.
- A Pre-trained ResNet backbone is used to extract keypoints, then extracted features are fed into a pre-trained NCNet matching layer to obtain/generate ground truth/annotations of extracted keypoints for Patch2Mix model. Then, the Patch2Pix network is used to match all extracted features between input images and refines the matches without requiring any extra step.

To recap, feature tracking, named as keypoint extraction and feature matching in the Computer Vision literature, is completed using SIFT/SuperPoint and GNN for the SuperGlue model, Pre-trained ResNet, pre-trained NCNet and Patch2Pix network for the Patch2Pix model, respectively.

Concerning the Enhanced Relative Navigation (ERN) model, in this case the defining is referred to an enhanced feature matching and relative position system, capable mapping two input images. This is done as follows:

- Optimal Matching Layer is used to refine/optimize feature matches between two input images thanks to Sinkhorn Algorithm for the SuperGlue model. The final output is keypoint matches, relative poses based on iterative optimization algorithms (e.g.: RANSAC).

- Unlike SuperGlue, Patch2Pix does not require any enhanced feature matching since the matched features are already refined previously. The final output is keypoint matches, relative pose based on iterative optimization algorithms (e.g.: RANSAC).

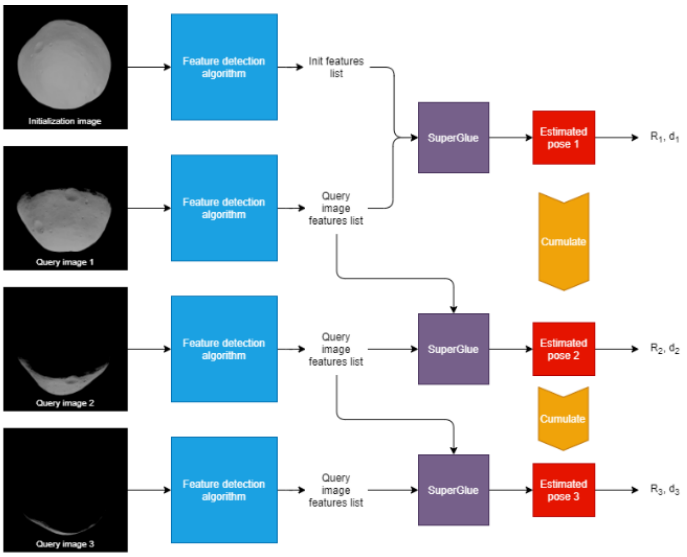


Figure 4-5: SuperGlue pipeline. The system is composed by two basic components: a feature detection (based on ORB/SIFT/Superpoints) and a feature matching algorithm (in the case of SuperGlue a Graph Neural Network is employed).

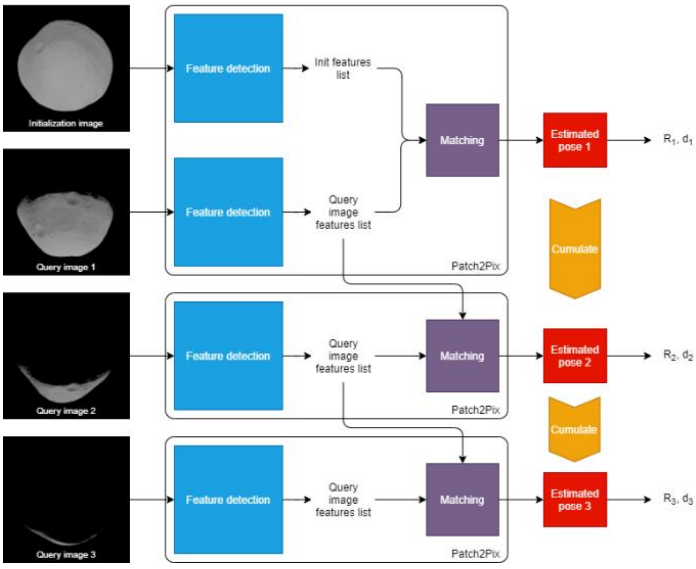


Figure 4-6: Patch2Pix pipeline. The approach is fully trainable and integrated. Thus, both feature detection and matching are part of the same component.

Because of this FT and ERN systems in this case are complementary and designed to work together in order to produce the expected measures required by the navigation systems. This process is detailed in Figure 4-5 and Figure 4-6.

Regarding SuperGlue (Figure 4-5) a set of two images is provided to the feature extraction algorithm. The output lists are then used by SuperGlue to estimate common features. An optimization routine is then used to estimate the pose difference between the reference image and the query image. The process is then repeated, cumulating the variation in rotation and translation.

On the other end, because Patch2Pix (Figure 4-6) is a fully trainable end-to-end approach both feature detection and matching is integrated within the same functional block.

5. ALGORITHM VALIDATION

5.1. CONVENTIONAL IP/HYBRID ERN

The following aspects were considered in the definition of the tests:

- Sensitivity to the illumination conditions. Although the HERA mission is defined so that the surface of the asteroid being captured by the camera is always illuminated, it is necessary to analyze the sensitivity to the choice of the tunable point in the VCFB to match conditions of the rehearsal and/or to different illumination angles.
- Robustness to the shape and appearance of the surface. Given that the ultimate appearance of the asteroid will not be known until the spacecraft reaches the body, it is important that the performance of the algorithms is independent of the shape and appearance of the asteroid.
- Robustness to the spacecraft pose knowledge as well as to asteroids attitude knowledge. As the reference landmarks used by the Enhanced Relative Navigation are both detected and later matched on-board, their estimated position is dependent on the available knowledge at the moment they are first computed. The algorithms must be compatible with the expected navigation error and the knowledge of the asteroid ephemerides.
- Importance of using real images to simulate the actual execution. Since no images of the asteroid will be available during the training phase, the algorithms need to be trained with synthetic images. It is thus necessary to ensure the algorithms work when real images acquired by a camera are used.

5.1.1. MIL VALIDATION

The MIL validation was performed using the HERA mission simulator developed in house (by GMV) under the HERA GNC activity and adapting it to the defined interfaces and with the mission scenario definition.

First, we report the position error of the HERA scenario, using a navigation based on Feature Tracking and Altimeter. This is useful to be taken as the baseline and to compare results with. Given that the the Navigation Filter uses the ERN only to correct the translation, all plots will refer to the position error.

Figure 5-1 shows the position error of the Navigation when using Feature Tracking (FT) and the Altimeter (ALT). The blue line shows the actual error at each instant of time whereas the red line shows the 3-sigma being considered by the filter. Each line corresponds to one sample of the Monte Carlo simulation. The navigation is performed with respect to Didymain up to t≈27000, and with respect Dimorphos, from that time onwards.

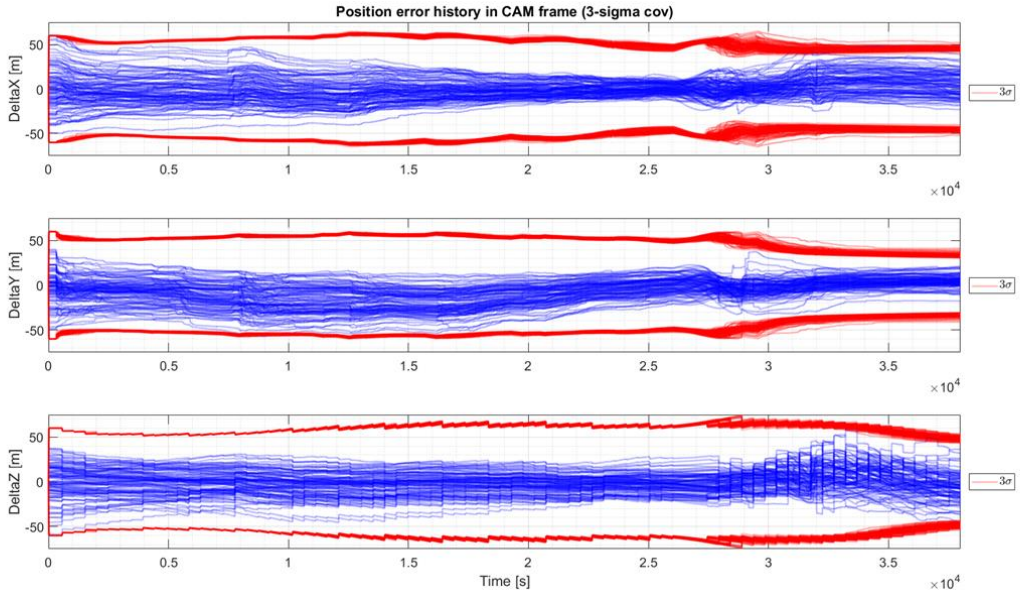


Figure 5-1. Position error FT+ALT (HERA reference navigation)

Figure 5-2 shows the position error when using Feature Tracking and ERN with eight activations while navigating with respect to Didymain and with two activations while navigating with respect Dimorphos. These eight activations correspond to three different areas of the asteroid observed while orbiting around the asteroid. Three activations are performed when the spacecraft sees the first and third areas of the asteroid depicted in Figure 5-3 (left), and the remaining two when spacecraft sees the second area. This distribution of the activations allow a better integration of the corrections within the filter, increasing the observability. In this case, given that the relative trajectory is mostly a linear approach, the two activations correspond to the same area of the asteroid (see Figure 5-3, right).

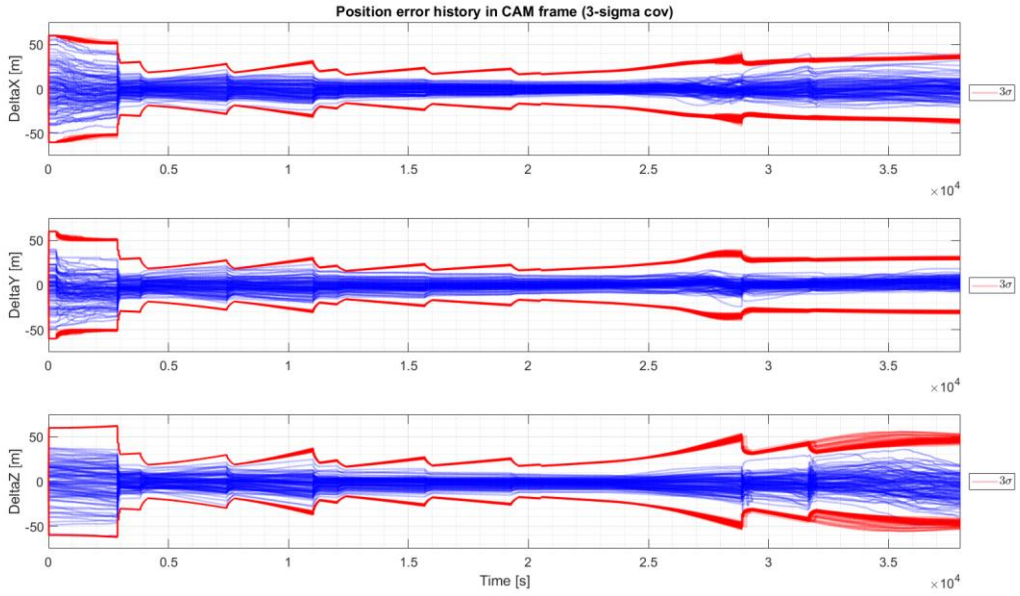


Figure 5-2. Position error FT+ERN_REHE8_8_2

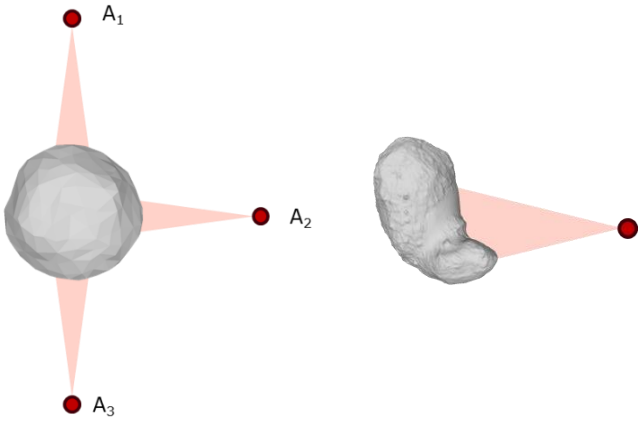


Figure 5-3. Left, distribution of the areas for the 8 ERN activations on Didymain; right, area of the two ERN activations on Dimorphos.

We can clearly notice how the error is quite constrained and how each activation of the ERN, especially the first ones, reduces the error. This can also be observed in the covariance, which decreases in each of the activations.

Table 5-1 summarizes the results of these tests and observe numerically the improvements provided by the ERN with respect the nominal FT+ALT navigation in HERA. The errors are provided at times t=25000, before switching from primary to secondary, and t=38000, the end of the trajectory.

Table 5-1: Error statistics

Test case	Statistic	t=25000				t=38000			
		x [m]	y [m]	z [m]	total [m]	x [m]	y [m]	z [m]	total [m]
FT_ALT	**min_error**	0.04	0.04	0.33	1.42	0.55	0.01	0.21	5.00
	max_error	15.65	24.88	28.06	35.66	45.99	20.91	44.82	59.64
	mean_error	0.56	-1.7	-2.24	14.41	5.23	4.55	-5.92	20.71
	std. dev.	6.37	9.57	10.51	6.41	13.9	7.02	14.81	10.61
	95% cases below	23.98				39.72			
FT_ERN_REHE8_8_2	**min_error**	0.07	0.01	0.01	1.39	0.06	0.02	0.15	3.27
	max_error	10.87	11.18	14.57	16.59	36.77	18.26	47.35	55.71
	mean_error	0.68	0.8	-2.68	7.6	0.27	2.9	-8.05	17.09
	std. dev.	3.81	4.37	5.24	3.31	10.19	5.08	13.55	9.6
	95% cases below	13.15				34.26			
FT_HyERN_REHE8_8_2	**min_error**	0.04	0.02	0.06	1	0.03	0.03	0.28	3.59
	max_error	11.43	14.43	22.64	23.43	57.39	25.23	74.23	93.72
	mean_error	1.81	0.71	-4.56	9.43	-8.14	6.77	-16.75	26.89
	std. dev.	4.16	5.08	6.73	4.85	15.35	6.54	17.34	15.66
	95% cases below	18.68				51.23			

As it can be noticed, the error during the first part of the trajectory is always considerably better for the FT+ERN combination. When observing the error at the end of the trajectory, we can see that the error of the FT+ERN combination, when using two activations in Dimorphos is slightly better than the one obtained with the nominal FT+ALT navigation. Not only that, even if no activations were performed in the second part, still the performance will not degrade significantly, thanks to the notably smaller error provided by the ERN during the first part.

We can also observe that the FT and Hybrid ERN (HyERN) combination performs worse than FT+ERN, but also worse than FT+ALT. This was due to the following reasons:

- Features were worse distributed (tent to concentrate always in higher gradients, such as in the asteroid limbo or shadows)
- Less matches were typically found, providing worse pose estimates (especially in more challenging situations)

- More sensitive to scale.

On-Ground vs On-Board Database

The benefits from creating the landmark database on-Ground using the images acquired for Orbit Determination with respect creating the database on-board using the on-board navigation solution were demonstrated already in the Performance Analysis Report [RD.3], where it was shown that the ERN allow recovering the error up to the accuracy of the navigation solution used to generate the database. Given that the on-Ground navigation solution will be considerably better than the on-board solution, so will be the accuracy of the ERN.

Table 5-2 summarizes the error numerically and compared with the results obtained when using the on-Ground database.

Table 5-2: Error statistics on-board database

Test case	Statistic	t=25000				t=38000			
		x [m]	y [m]	z [m]	total [m]	x [m]	y [m]	z [m]	total [m]
FT_ERN_OB_2_0	**min_error**	1.24	1.02	4.58	7.08	2.07	0.01	1.43	15.13
	max_error	7.42	30.89	78.34	81.72	32.02	20.01	50.15	56.09
	mean_error	1.06	-4.24	-18.1	33.6	-10.20	4.66	-20.06	30.95
	std. dev.	5.11	17.25	31.33	20.88	17.25	8.34	18.40	15.09
	95% cases below	81.72				56.09			
FT_ALT_ERN_OB_2_0	**min_error**	0.09	3.53	3.62	5.55	4.28	0.55	2.37	9.51
	max_error	6.3	32.4	54.48	56.8	30.99	14.89	47.42	56.94
	mean_error	1.31	-1.57	-17.7	31.01	-2.66	5.1	-9.53	22.86
	std. dev.	3.11	16.65	28.03	18.89	14.44	6.73	19.15	13.33
	95% cases below	56.80				56.94			

Robustness to Asteroid Shape and Appearance

Due to the uncertainty on the asteroid shape and appearance, dedicated tests on evaluating the performance of the ERN under different shapes (more concave asteroid vs convex one) and appearances (cratered, rocky). The shape of the asteroid did not have any relevant impact on the performances. Conventional IP ERN provided comparable performances independently of the appearance of the asteroid.

On the other hand, the Hybrid ERN was indeed impacted by this appearance, especially for the rocky version. There were considerably lower detection and matching rates, which translated into less effective and accurate corrections. Specific trainings of the CNN with only cratered or only rocky asteroids behaved slightly better than the combined trainings (using both images with craters and images with rocks). The same problems already identified for the default scenario also happened in the appearance variations: worse distributed features, less matches and sensitivity to scale.

Sensitivity Analyses

Given the dependence of having a database created under similar conditions to the moment when ERN is actually called, it was important to understand how close these conditions need to be. For this reason, we performed different sensitivity tests on the three factors that have the major influence: illumination conditions, viewing direction and viewing scale.

The algorithm shown to be robust to differences in illumination ranging from -15 to -60, where the accuracy was fairly constant in all cases. The covariance, nevertheless, captured the increasing uncertainties and clearly indicated when the measurements were not reliable. For the positive values, the range is more limited and the ERN provides accurate measurements up to +15 degrees.

The reason of this asymmetric range is that the nominal phase angle is approximately -30 degrees. The algorithm is robust to large differences in the phase angle (up to 60 degrees) as long as the difference only translates into projected shadows more or less elongated, but in the same direction. If the difference in the phase angle translates into shadows almost disappearing (due to a phase angle of 0 degrees) or projecting in the opposite direction, the algorithm is not capable of finding matches between the reference database and the current image.

This is depicted in Figure 5-4, where images with a difference in phase angle of -30, 0, and +30 are shown. The offset of -30 mostly translates into more elongated shadows (apart from the displacement of the terminator line). On the other side, the offset of +30 directly makes an absolute phase angle of 0 degrees, resulting in the shadows disappearance.

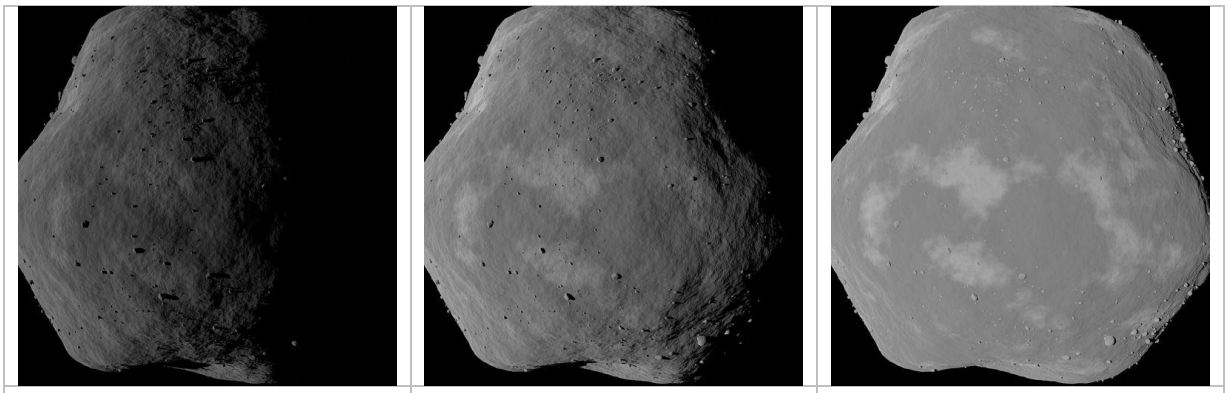


Figure 5-4. Images with different illumination offset angles.

Regarding to the viewing direction, up to a difference of ±10 degrees in the viewing direction between the database and the on-board image, the ERN performed without significate degradation. Starting on ±20 degrees, the performance was degraded but the covariance managed to capture this decrease in the performance.

With regards to the difference in scale, up to a scale factor of 2, the ERN performed without significate degradation (with increasing covariance, showing the lower confidence) but starting on 2.5, the performance was slightly degraded. For a scale factor of 3, the algorithm hardly provides good measurements. Again, the covariance captured the degradation.

5.1.2. HIL VALIDATION

The hardware in the loop tests must be understood as inserting in the V&V tests 'real' images captured by a real camera, in this case HIL tests do not refer to actually using space rated avionics in the loop (as for instance a LEON processor or a space rated camera). The HIL tests must be understood as MIL tests but using a real camera, even if it is a COTS camera, in a real robotic environment in order to provide the images to digest instead of using a synthetic image generator in the loop as in previous steps of the activity.

In any case the camera was representative of the proposed camera for the HERA mission. In that sense the camera and lens chosen for the HIL Validation Plan provided the same Field of View as the HERA camera, and the same resolution as well.

platform-art© was agreed with ESA to be used for the HIL tests of the HERA autonomous GNC, including the use of the AFC camera. These tests with the AFC camera cover the DCP phase and also the close fly-by scenario (with the main goal of testing the autonomous retargeting manoeuvres). For the close fly-by some differences with respect to the nominal scenario due to practical limitations in the set up of the robotic lab and to the focusing capability of the qualification model of the AFC were implemented.

Taking advantage of GMV's experience and activities being performed in the frame of the HERA autonomous GNC development, ***platform-art©*** was also proposed as baseline facility to support the HIL validation campaign here in ATENA.

Previous HERA mock-ups had a scale factor of about 2000, which means a Didymain mock-up of 40cm diameter and Dimorphos mock-up of 10cm diameter This was considered the best trade-off value for HERA because they were originally oriented at simulation of phases at farther distances to validate the centroiding algorithm and they lack lower-level detail.

Figure 5-5: 1:2000 HERA mock-ups

For ATENA, considering that the focus is on simulating the VCFB and the final part with Dimorphos, the proposed solution was thus to simulate the first part of VCFB with the previous 1:2000 mock-up of Didymain and simulate the final part of VCFB with a new 1:300 mock-up of Dimorphos. The creation of a larger size mock-up also allowed adding small sized details missing in the existing model. The VCFB trajectory became two separate trajectories (one for Didymain and one for Dimorphos), where we maximized the coverage that was so constrained due to the use of a narrow FOV at short distances.Figure 5-6 shows the new manufactured model and a close-up showing the fine detail added to simulate the asteroid relief at closer distances.

Figure 5-6. New model of Dimorphos.

The same VFCB trajectory used in the MIL simulations was implemented in the laboratory. Figure 5-7 shows a comparison between the reference image generated with PANGU and the one captured at the laboratory. Note that due to physical constrains, exactly the same illumination conditions could not be recreated.

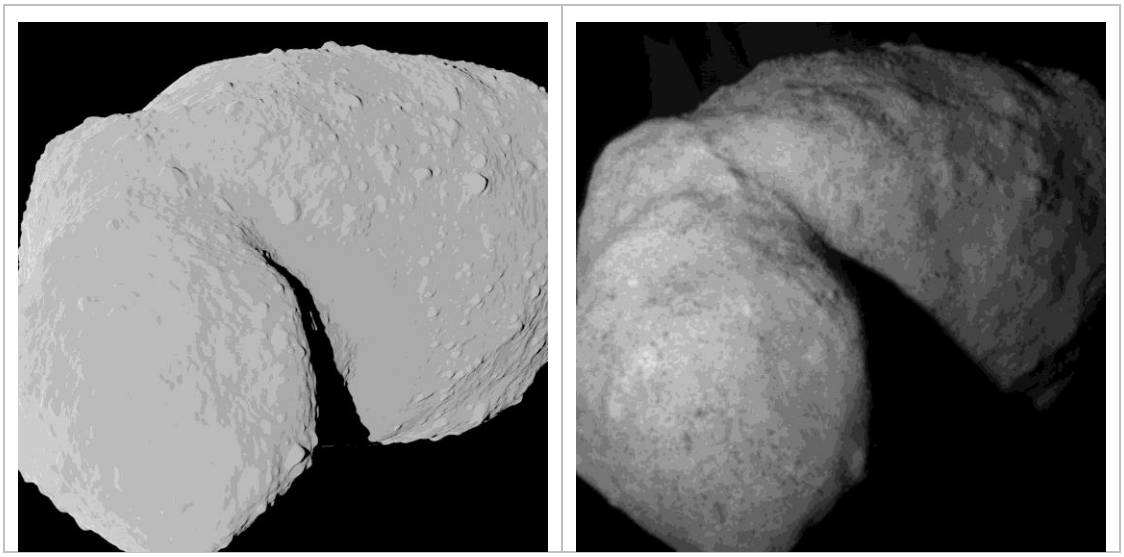


Figure 5-7. Reference (PANGU) image vs Laboratory image

The ERN performance was evaluated at each activation with a small Monte Carlo. For each of these frames, 20 executions were performed, each of them with a different database extracted by sampling the navigation error of the on-Ground solution. Therefore, at each image where the ERN is called there are 20 solutions, for which we computed the mean and the standard deviation. The mean of the errors for each image is depicted with the blue line whereas the mean±the standard deviation is depicted in blue.

As a reference, we first present the results using the same approach but for the synthetic images. Figure 5-8 shows the range error (which is the most representative of the accuracy of the ERN and the component that is most affected by the navigation error in the database generation) but also the corresponding component of the returned covariance matrix. In total, eight activation regions are considered (corresponding to the Didymain), each of them composed of 5-10 images. For each of these frames, 20 executions were performed, each of them with a different database extracted by sampling the navigation error of the on-Ground solution. Therefore, at each image where the ERN is called there are 20 solutions, for which we computed the mean and the standard deviation. The mean of the errors for each image is depicted with the blue line whereas the mean±the standard deviation is depicted in blue (to better understand the envelope of the error). It is also important to remark that the database with the camera located in the position of the central frame of each activation window. Taking into account that the spacecraft is relatively orbiting around the asteroid, in the borders of the activation region the camera during the execution of the ERN will be observing the asteroid from a slightly different point of view and therefore two "error" contributions are actually being evaluated: the difference in illumination angle and the difference in viewpoint. This can be observed in Figure 5-8, where the error with offset=0 is depicted (this means each database was created at the same location and with the same illumination of the central frame of each activation region). In this case, the error of the ERN basically corresponds to the error of the database (the one of the on-Ground navigation solution). The covariance plot (right) shows as that the in the central frame the ERN is fully confident of its solution whereas in the borders of each activation region, the covariance increases slightly, showing that the slight difference in the viewing angle increases slightly the uncertainty (even if the actual error is still good).

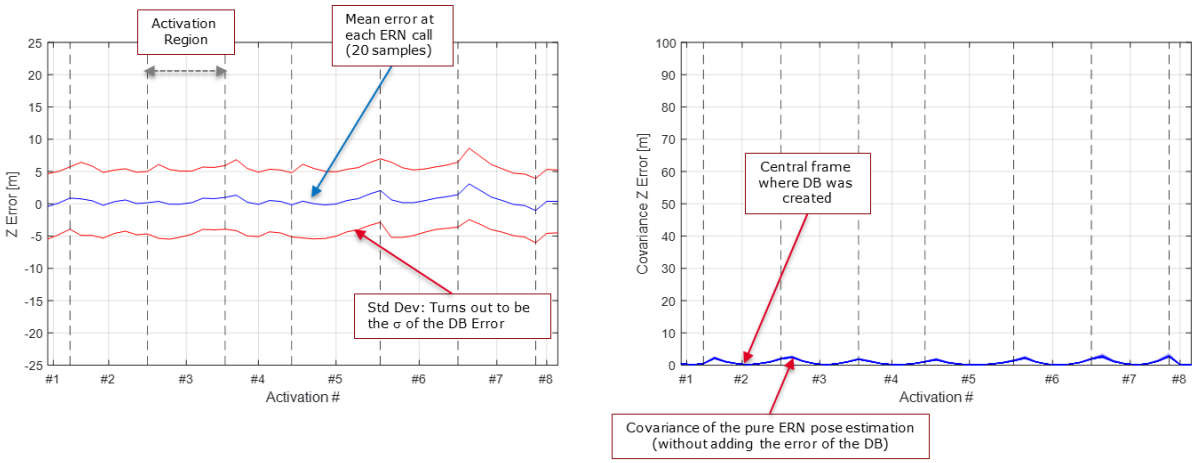


Figure 5-8. Interpretation of the sensitivity analyses error plots

Figure 5-9 shows the range error (left) and its associated covariance (right). In this case, a total of 9 activation regions were considered, 8 in the Didymain and 1 in Dimorphos.

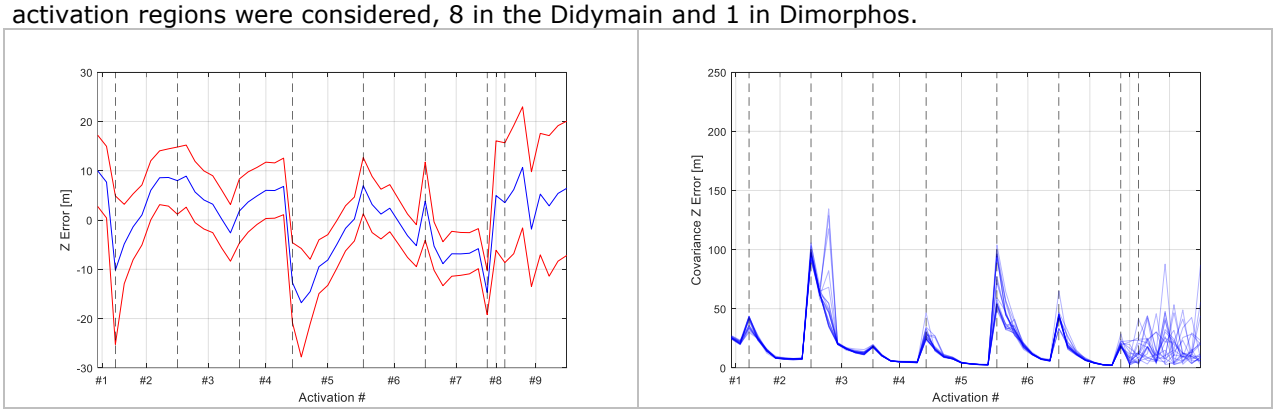


Figure 5-9. Range error (left) and associated covariance (right) for the HIL test

We can observe that the performances are slightly lower than those obtained in the synthetic images. In addition, the covariance show less confidence on the measurements. Nevertheless, it is important to note that these performances should not be compared directly with those obtained in MIL tests due to the additional sources of error:

- Ground-truth obtained in the laboratory is not perfect because of the calibration residuals and errors in the robots telemetries.
- Scaling factors. The primary mock-up has a scale of 1:2000, whereas the mock-up of the secondary has a scale of 1:300. This does not mean obtained in the HIL tests should be divided by 1:2000, because the navigation errors used to create the database where scaled accordingly, but inaccuracies both of the algorithm and in the ground-truth are scaled by a large factor.
- Blurring of the images was higher than what should be expected from the real camera.

Even with these issues and also with an excessively simple mock-up of Didymain (this model was manufactured within HERA for farther distances), the algorithm performed well and no significant degradation of the performances has been found.

5.2. AI-BASED SOLUTION VALIDATION

To validate the performances of the algorithm, is necessary to test it in different conditions. For this reason, GMV provided three different synthetic datasets. Each dataset has features to use for navigation. In particular:

- Default dataset: mixed set of features, both craters and boulders.
- Cratered dataset: only craters are present on both Didymain and Dimorphos.
- Rocky dataset: only boulders are present on both Didymain and Dimorphos.

The tests were performed on the Very Close Flight-By (VCFB) trajectory.

GMV provided a set of 50 trajectories plus a reference trajectory that represents the ground-truth. Normal noise is injected into the 50 testing trajectories analysed. This will represent possible measurement noise. The robustness of the algorithm is tested against these conditions.

A subset of the trajectory is used for the evaluation. It means that, considering the whole dataset provided, the images are used as follows:

- From frame 0 to frame 400, are used for recreating the point-cloud, then used for validation.
- From frame 400 to frame 800, are used for the navigation and pose estimation validation.
- From frame 800 to the end, are discarded from the analysis.

The last part of the dataset is discarded because the modelling of Dimorphos cannot be performed due to the linear approach trajectory and thus the error in estimated poses is outside the requirements.

To reduce possible stereoscopic degeneration of the pose computation, the evaluation of images is computed on pairs selected using a skipping step. In the final evaluation a skipping step of 8 is considered.

Figure 5-10 is summarizing the results obtained for the three datasets. The most relevant metrics that identifies the performances of the algorithm is the total non-dimensional error and that is why this is the only reported in this figure, together with the number of matches.

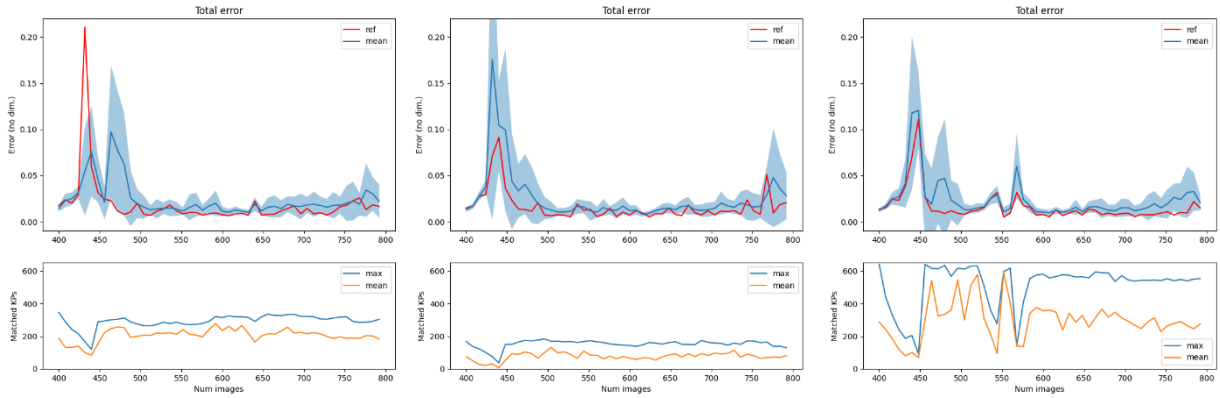


Figure 5-10. Compared total error score between all the datasets (respectively: cratered, default rocky)

For all the datasets there is a clear pattern in the values of the errors obtained. Is possible to identify three regions:

- From frame 0 to frame 500. In this region the error is fluctuating towards higher values, for all the scenarios considered. This is due to a bad coverage of the generated point-cloud, in this particular region. A further refinement of the feature extraction algorithm is required, in order to mitigate this effect
- From frame 0 to frame 700. The error in this region is stable and sufficiently low, to be considered for navigation.
- From frame 700 to the end. The error starts increasing again because the features detected in this area are changing, due to the changed scale of the problem. SLAM/VO approaches with progressive refinement of the reference point-cloud, should be considered, in order to reduce this effect.

The error reported in the previous graphs is also correlated with the number of matched features. In this sense, there is a clear pattern in the measured errors and as expected, a lower number of matched features equals to lower accuracy in pose estimation and vice-versa.

This aspect is important for the information to be fed the navigation system. The number of matched features offers a good parameter to consider for considering or discarding a particular measurement obtained by the AI-based pose estimation algorithm.

Analysing now the number of matched features between pairs of images is possible to conclude that the rocky dataset provides to Superpoint + Superglue the most reliable set of features to use for pose-estimation. This pattern is clearer considering the summarized performances reported in Table 5-3. In this case the mean number of matched features is of 514 in the case of rocky dataset, 293 for the default dataset and 153 for the default dataset.

This behaviour is probably due to the lack of sufficient features in the case of the Default dataset. To mitigate this effect, a complete model retrain of the feature extraction model should be performed.

The number of matched features directly affects the performances of the pose estimation algorithm, in fact, considering again Table 5-3 all the metrics considered are worst for the Default dataset, if compared with the other scenarios considered.

Rocky dataset once again offers the best performances and thus the best set of features to use for navigation.

Table 5-3: Mean values computed for all the testing datasets

	Cratered	Default	Rocky
Translation error (no-dim)	**0.0109**	0.0119	0.0112
Quaternion error (no-dim)	0.0129	0.0133	**0.0128**
Rotation error (deg.)	0.7416	0.7626	**0.7331**
Total error (no-dim)	0.0186	0.0162	**0.0161**
Translation error magnitude (m)	**73.3764**	80.2447	75.5064
Transl. x-axis error (m)	10.8568	4.7651	13.9877
Transl. y-axis error (m)	20.1382	30.3418	19.8682
Transl. z-axis error (m)	0.5471	0.2053	0.6038
Rotation alpha error (deg.)	-0.1848	-0.1725	-0.1296
Rotation beta error (deg.)	0.4571	0.4362	0.0423
Rotation gamma error (deg.)	0.1725	0.1517	0.1844
Number of matches	293	153	**514**

Considering now the decomposed translations, Table 5-3 shows that the y-axis contributes the most to the overall translation vector estimation.

Similar considerations are possible also for the decomposed rotations. The beta-angle (associated to the y-axis) shows again the highest contribution to the overall error (except for the rocky dataset).

6. CONCLUSIONS

A navigation architecture has been proposed using image processing techniques integrated within a navigation filter performing the data fusion of the UP and the rest of GNC sensors.

Different proposals have been considered following conventional IP algorithms but also new AI-based approaches, together with a hybrid version trying to combine benefits from both worlds.

The exhaustive testing and validation campaign has allowed to evaluate the performances of the Enhanced Relative Navigation under a wide range of situation and confirm its suitability for a mission such as HERA. The following list summarizes the main conclusions obtained:

The combination of Feature Tracking and ERN (without altimeter) provides comparable and even slightly improved performances to the nominal Feature Tracking plus Altimeter navigation of HERA if the database is created with the on-Ground navigation solution.

- Activations in Dimorphos are needed to achieve this higher performance.
- Creating a database for Dimorphos with images sent to Ground can be more challenging from an operations point of view as it involves observing it first under similar conditions to those to be faced during the actual ERN execution (this was shown to be fairly reasonable to achieve for Didymain).
- Nevertheless, sensitivity analyses have shown a reasonable tolerance to illumination conditions, viewing angle and scaling, which should facilitate this problem.
- Still, even if no ERN is used in Dimorphos, and just Feature Tracking is used for this part, the navigation solution does not degrade significantly (only the range accuracy is slightly reduced).

The database creation with the on-board navigation solution does not bring benefit except as a contingency in case of an (unexpected) filter divergence. This is due to the good design of the HERA navigation filter, which uses a SLAM-based approach and therefore no direct integration of the error in the relative measurements of the Feature Tracking is taking place. In case of a simpler filter design, where current pose is achieved as an integration of relative measurements, the ERN would allow resetting the overall accumulated error to a value comparable to the one when the database was created.

The ERN provides comparable performances independently of the appearance of the asteroid, and no additional training or tuning is needed.

HIL tests have confirmed the solution is not affected when switching to real images acquired by a real camera, especially considering some issues where affecting the quality of such images (defocusing, low detail on Didymain model)

Hybrid ERN has also a worse performance than the (conventional IP) ERN.

- Features are worse distributed (tend to concentrate always in higher gradients, such as in the asteroid limbo or shadows)
- Less matches are typically found, providing worse pose estimates (especially in more challenging situations)
- More sensitive to scale.
- More sensitive to the asteroid appearance. Needs a dedicated training for the asteroid appearance to improve its performance (the combined training with both craters and rocks performed worse).

AI-based solutions have also shown its suitability and potential to be used for navigation purposes, but still need to be consolidated, especially considering the additional implications they impose (training, execution time…)

As future work, a breadboarding of the algorithms in representative HW should be performed to characterize the execution time (although given that no continuous ERN corrections are to be performed, the requirement on the execution time could be less demanding than for the FT which is executed continuously). Given the dependencies with Ground operations, a more comprehensive study of the required interfaces and the definition of the Concept of Operations, especially for the on-ground database generation would also be required.

END OF DOCUMENT