PARAMETRIC PERFORMANCE CHARACTERIZATION OF VISUAL ODOMETRY FOR THE SAMPLE FETCH ROVER

Virtual Conference 19–23 October 2020

Matteo De Benedetti[1], Martin Azkarate[2], Innocenti Mario[3]

[1]University of Pisa, Via Giunta Pisano 28 Pisa, Italy, E-mail: matteo.debenedetti91@gmail.com
[2]European Space Agency, Keplerlaan 1 Katwijk, The Netherlands, E-mail: martin.azkarate@esa.int
[3]University of Pisa, Via Giunta Pisano 28 Pisa, Italy, E-mail: mario.innocenti@unipi.it

ABSTRACT

ESA is working with NASA to plan and carry out an international Mars Samples Return (MSR) campaign between 2020 and 2030. A relevant part of the upcoming MSR mission is the Sample Fetch Rover (SFR), tasked to collect sample tubes of Martian soil prepared by Mars2020 rover Perseverance.

This work focuses on the localization capabilities of SFR and the potential reuse of the functionalities present in the ExoMars rover. Visual Odometry (VO), a vision-based localization algorithm, is often the main component of the localization process in planetary robotics.

The goal of this study is to investigate the possibility of transferring the ExoMars VO solution to a valid SFR implementation, compliant with mission requirements.

First, the main differences between the two missions, SFR and ExoMars have been studied, in order to identify the most critical parameters for the VO process. Then, using a testing rover available in the Planetary Robotics Lab (PRL) at European Space TEchnology and research Centre (ESTEC), the effect of the previously identified parameters on the VO performances was evaluated, identifying the most crucial ones and proposing some solutions to face them. This work could lead the way to future studies about the localization for the Sample Fetch Rover and what are the main and most critical factors that would have to be taken into account in order to achieve an accurate and reliable localization system.

1 INTRODUCTION

VO is the process of estimating the motion of a robot, usually represented as a 6-DOF pose, using pairs of stereo images as inputs. For autonomous vehicles knowing their position is extremely important, and while on Earth this problem is usually relatively easy to solve, in space exploration it becomes much more difficult.

On Earth, GPS is often used for localization and, if the terrain is not too rough and does not present high slip, then Wheel Odometry can be used. On Mars, all the previous problems are present at the same time: there is no GPS and the terrain is usually hard to navigate, making inertial odometry solutions drift quickly. In addition, due to the distance between Earth and Mars, a direct control is impossible. The Mars Rovers are usually commanded once a day, it is then imperative to have a well performing localization system onboard, since they will need to rely on this onboard capability to perform their tasks throughout the day without human intervention.

This is the scenario where VO came into play. It was first introduced in the two Mars Exploration Rover (MER), Spirit and Opportunity [1], as a ”bonus feature” but, as it proved essential in many critical situations, it quickly became the main mean of localization. From NASA’s first Mars rovers, VO has been regularly used on planetary rovers, and at this moment it is already implemented as the central component for the complex GNC architecture of the ExoMars Rover [2], the next ESA rover which will launch in 2020. The current state-of-the-art of Visual Odometry is able to reach very high levels of accuracy, achieving an error around 1-2% of the distance traversed [3], [4].

The paper first introduces the SFR mission and the advantages and reasons of using VO for the localization of planetary rovers. Then, Section 2 describes the many challenges posed on SFR and how they could affect the VO process. In Section 3 the reader is introduced, first, to the facilities available in the PRL and used for the testing activity, and then, to the plan of the tests aimed at investigating the differences for the VO between SFR and ExoMars, identified in the previous section. Next, Section 4 shows the results and preliminary conclusions of the tests. Lastly, Section 5 collects all the conclusions drawn from the previously shown tests, exposing the most critical aspects and how they could be mitigated. Additionally, some possible future and interesting activities are presented.

2 MOTIVATIONS AND OBJECTIVES

As previously mentioned, Visual Odometry is the main component of the ExoMars localization and it is the baselined solution for the Sample Fetch Rover. Nevertheless, there are many differences between the two missions, the two rover designs and the way they will operate. The objective of the work is to investigate those differences, trying to replicate them as accurately as possible in the Mars Test Bed of the PRL, characterize their effect on the VO performances, and where possible suggest some actions to mitigate them.

The MSR mission architecture places indeed a number of challenging requirements [5] on the SFR, extending beyond those firstly defined in the assessment study [6]. The start of the SFR surface mission coincides with the Martian dust storm season, which means the SFR will need to be able to maintain the same performances across high optical depth. The rover is also required to traverse up to 15 km within 150 sols, this raised the required traverse speed of SFR, which at the moment is planned to be around 6.7 cm/s in contrast with the 1 cm/s designed for ExoMars. The higher speed alone could strongly affect the VO by amplifying the estimation errors and increasing the drift rate. It also affects many factors of the VO process. A higher speed would increase the spatial distance and reduce the amount of image overlap between two frames. This is likely to cause a decrease in the motion estimation accuracy, up to the point of making it impossible if the overlap becomes too small and so does the number of common features between two consecutive frames.
A high-speed traverse leads to a blurred image with very noisy features, making it hard to detect them, match them between images and also have a good precision in the position they are detected at. These factors would of course worsen the VO performances, but they could also be mitigated acting on the exposure time of the camera. A fast exposure time will result in darker but sharper images, which would reduce the effect of the motion blur. It is important to consider that shortening the exposure time comes with consequences that cannot be ignored, in particular for the SFR scenario, where the rover will have to function. If the ambient light and visibility are already low, then the exposure time cannot be reduced too much, otherwise it will result in very dark images, on which it would be almost impossible to run the VO.

At the same time, the spatial relationship between the two consecutive frames also depends on the frequency the VO runs at and the camera position and orientation. A camera that is higher or points directly forward will result in images with much more overlap and many far away features, while a camera aimed more at the ground will lead to a smaller overlap and considerably closer features. the best choice is to have both, because close features lead to a good position estimate and far away features will instead provide a better orientation. Additionally, the rover's velocity changes the blurriness of the images.

One more interesting aspect to be investigated for the SFR scenario is the terrain. Having to traverse a significantly longer distance is bound to bring the rover on new and different terrains very often. The localization must be robust to these changes. Also, reference images of the area near the SFR landing site have become recently available and they show a very rough ground, made of flat fractured plates surrounded and partially covered by fine sand. This kind of terrain will represent a challenge for the small rover that will have to climb over these plates, while trying to keep the planned high speed and good localization performances.

3 TESTS SETUP

The testing activity has been performed at the PRL at ESTEC. The laboratory offers the Planetary Utilization Test Bed (PUTB), a 9-by-9 meters terrain filled with different types of sand, pebbles, and rock, specifically chosen to mimic the Martian environment. It is also equipped with VICON, a precise motion capture system, used to acquire ground truth localization data of objects in the PUTB.

The tests were carried out on the ExoMars Testing Rover (ExoTeR), a half-scale laboratory prototype of the ExoMars rover that mimics the locomotion and navigation subsystems of the real rover. The rover mounts a multitude of external sensors used for localization, such as a Bumblebee BB2-08S2C stereo camera mounted in front of the rover (often referred as Localization Camera (LocCam)), two more Bumblebee stereo cameras, one of which is often referred to as Navigation Camera (NavCam), mounted on top of a mast actuated by a Pan and Tilt Unit (PTU), an Inertial Measurement Unit (IMU) and encoders for the motors and the passive joints of the locomotion system.

The VO solution implemented in ExoTeR, referred to as SpartanVO [7] in this paper, is developed within the SPAring Robotics Technologies for Autonomous Navigation (SPARTAN) project with the funding of the European Space Agency. The project aims at developing a robust hardware accelerated implementation of computer vision algorithms for planetary exploration rovers that offer limited computational capabilities. Particular emphasis is given to make it as reliable but also as fast as possible.

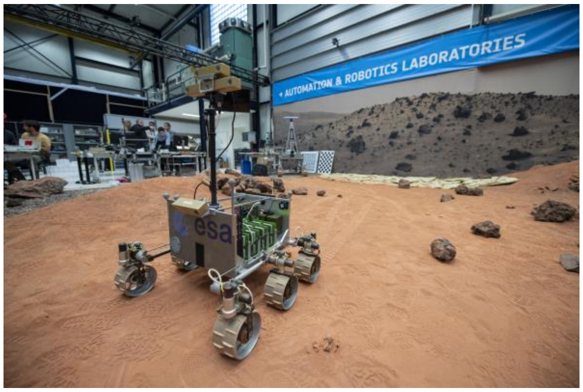

Figure 1: ExoTeR in the Test Bed of the PRL.

4 EXPERIMENTAL RESULTS

4.1 Motion Blur and Ambient Light

Since the higher translational velocity is one of the main differences of the SFR with respect to other rover missions, it was decided to start the tests activity by separately varying it and see how it would affect the VO performances. Then, the relationship between the rover's speed and the ambient light and Exposure Time (ET), which is expressed in milliseconds, will be investigated and finally a low ambient light scenario will be tested. Increasing the ET of the camera has two main effects:

- The image brightness will increase, which can be beneficial in case of a low ambient light.
- The motion blur will also increase, which is not beneficial for the VO since it adds uncertainty and noise in the features.

The ExoTeR camera (Bumblebee BB2-08S2C) also has an Auto Exposure mode (AE) which internally computes an optimal value for the ET. The first set of tests that has been run is made of sequences where the ExoTeR rover moved forward for 5 meters at increasing translational velocities. This particular range of velocities was chosen to include the speed designed for ExoMars (0.01 m/s) and the one planned for SFR (0.067 m/s), up to the maximum speed of ExoTeR (0.09 m/s). It has been deemed more relevant to the testing evaluation to look at the slopes of the error curve: a 1st degree polynomial has been fitted to the curve using Least Squares, and the following results, shown in Figure 2 and Table 1, were obtained.

Velocity	Slope of the error curve
0.01 m/s	0.0090
0.03 m/s	0.0076
0.05 m/s	0.0047
0.07 m/s	0.0055
0.09 m/s	0.0088

Table 1: Results of the velocity tests.

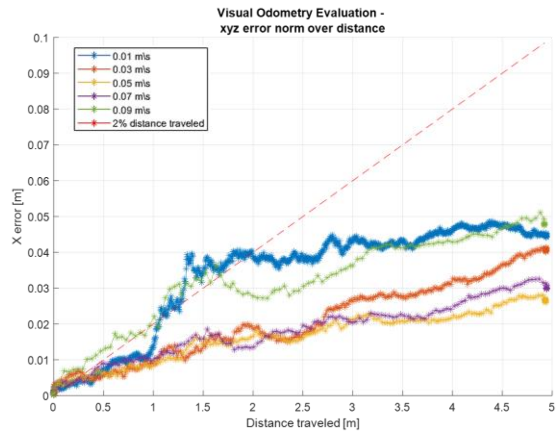

Figure 2: Velocity tests.

There is variation between the slopes of the tests, but it is not significant and more importantly does not look like it increases with the speed. This led to believe that only increasing the translational velocity, while all the other parameters (such as ambient light, VO frequency, exposure time of the camera etc.) remain the same, does not seem to have a negative effect on the VO performances. This conclusion is valid as long as the speed remains inside the tested range of 0.01 to 0.09 m/s and for the environmental conditions of the PRL, while it is likely that higher speeds would negatively affect the VO performances.

Different ETs have been tested against a constant velocity of 0.07 m/s and the following results, in Figure 3, have been obtained.

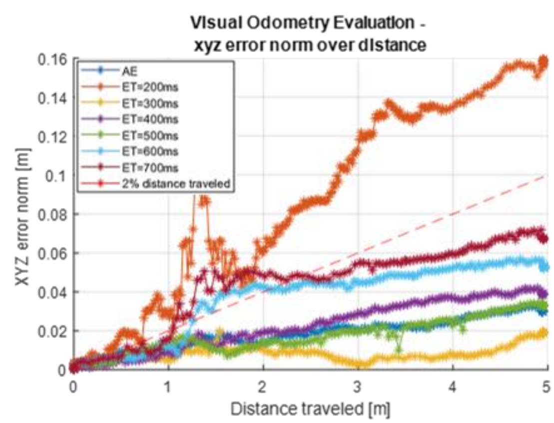

Figure 3: Position Error in the tests at different Exposure Times.

The position estimation with ET=300ms performed better than the AE (which in this case applied an ET~530ms). This means that lowering the exposure time can be beneficial at high speeds because it helps reducing the motion blur, but if lowered too much it starts to be counterproductive because it also strongly reduces the brightness and contrast of the image (ET= 200ms), while a higher ET (ET= 400ms, 500ms, 600ms, 700ms) leads to an even higher motion blur that will negatively affect the VO performances.

A low light/visibility scenario has been reproduced in the PRL and the ambient light has then been measured using a digital light meter. All the previous tests have been run with the lab fully lit at 422.24 lux, while the low light scenario is at 226.27 lux. The first test performed in the low light scenario had the rover moving at 0.07 m/s and showed that the SpartanVO in AE was struggling and reducing the ET did not improve the performances. Successively, the possibility of increasing the ET, which would have the positive effect of increasing the brightness but also the negative effect of increasing the motion blur, has been also investigated. The test showed that it actually decreases the VO performances with respect to the use of AE. Finally, a low light and slow speed (0.02 m/s) scenario has been tested. From the results it was clear that either increasing or decreasing the ET is not at all beneficial with respect to the use of AE, but the VO in AE at this lower speed was still able to achieve very good performances. All these tests in the low light scenario showed that the only way to maintain the same level of performances was to reduce the speed from 0.07 m/s to 0.02 m/s.

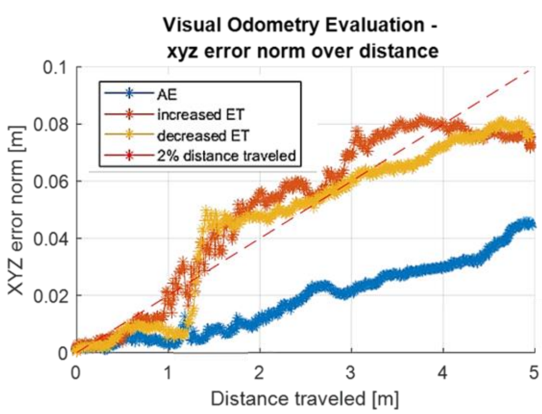

Figure 4: Low Ambient Light tests with low speed.

4.2 Terrain

In order to characterize the robustness of the SpartanVO to a variety of terrains, it has been tested on the three different type of terrains available in the PUTB of the PRL: sand, sand with some rocks, and a layer of small pebbles, as shown in Figure 5.

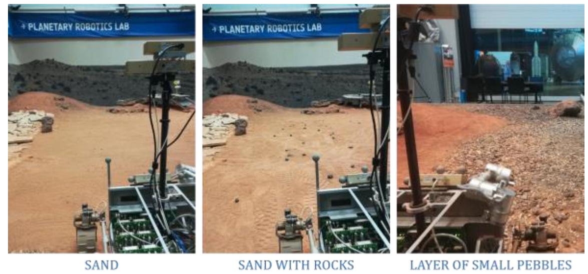

Figure 5: Different types of terrain of the PRL.

The different terrains were challenging in their own way for the VO: in the sand there are not many features to be extracted and the VO risked to struggle estimating an accurate motion; the occasional small rock would generate abnormal motions when climbed that the VO could struggle to pick; and the layer of pebbles would induce an almost constant vibration on the rover along the whole traverse. The rover was commanded a 3.5 meters traverse at the SFR velocity of 0.07 m/s and the results obtained on the three terrains are plotted in Figure 6. They clearly show a good robustness of the SpartanVO against different types of terrains, with a small increase in the slope of the error in the traverse with the rocks, but still below the 2% line.

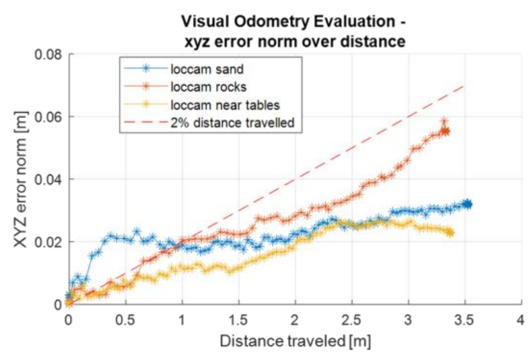

Figure 6: Test on different terrains in the PRL.

Recent reference images of the SFR landing site revealed that part of the terrain around it is expected to be made of solid fractured plates partially covered in sand, as shown in Figure 7.

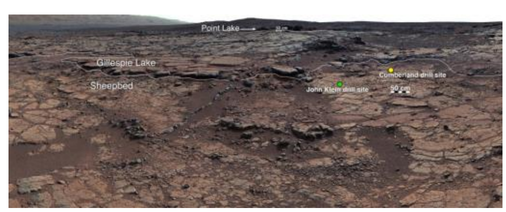

Figure 7: Fractured terrain in the SFR landing site.

This type of terrain is definitely different than the usual sand where most of the tests in the PUTB have been conducted so far.

It has then be decided to run a test on a part of the PUTB in the lab that resembles such fractured terrain, as shown in Figure 8, and compare it with a traverse at the same speed on the sand.

Figure 8: Fractured terrain in the PUTB of the PRL.

The fractured terrain led to a more irregular traverse, with sudden changes in position (both vertically and horizontally) when the rover climbed up or down the plates. The comparison between the VO performances on the fractured and sandy terrain, whose results are reported in Figure 9, proved that the VO is quite robust to the new terrain, with a slightly higher drift, mostly on the z component of the position estimate, which was expected from the changes in altitude from this new kind of terrain.

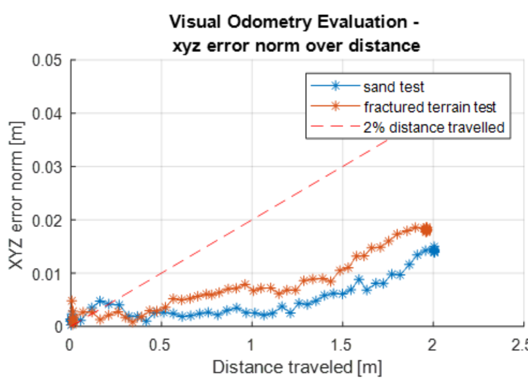

Figure 9: Comparison between Fractured and Sandy Terrain.

4.3 VO Frequency

The VO frequency has a great influence in the VO process, in particular on the relationship between the two consecutive stereo pairs that are used for the motion estimation. Therefore, to better understand and characterize it, two concepts are introduced:

- The Inter-Frame Distance (IFD), expressed in meters, is defined as the spatial distance in the world between the points of acquisition of two frames. The IFD also depends on the rover's speed and it is computed simply by multiplying the VO period with the rover's speed.
- The Image Overlap Percentage (IOP), is the percentage of the area of an image that is common between two frames at different time steps. It is conceptually similar to the IFD but also more general, since it encodes information about the content of the images and it also depends on the camera placement on the rover (height and orientation), but it loses information about the physical distance between frames.

The computation of the IOP is more complicated and an approximated model has been defined. Assuming the rover travels on perfectly flat ground with exactly the same motion as commanded, meaning no vibrations from rocks and slip of the terrain, the problem of computing the IOP becomes purely geometrical.

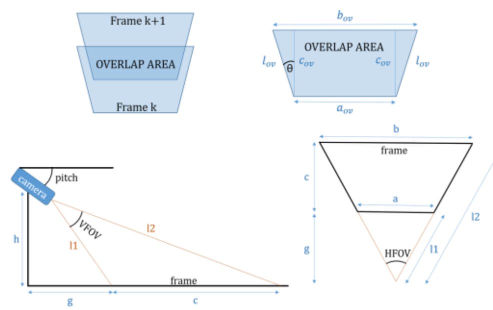

Figure 10: Image overlap Percentage scheme.

To investigate different values of Inter-Frame Distance (IFD) and Image Overlap Percentage (IOP), a vast number of tests with different combinations of rover velocity and VO frequency have been run and they all led to the following conclusion. The VO performances are affected by the IFD and IOP in a very similar way, whether the rover is moving at a high or low speed, as shown in Figure 11.

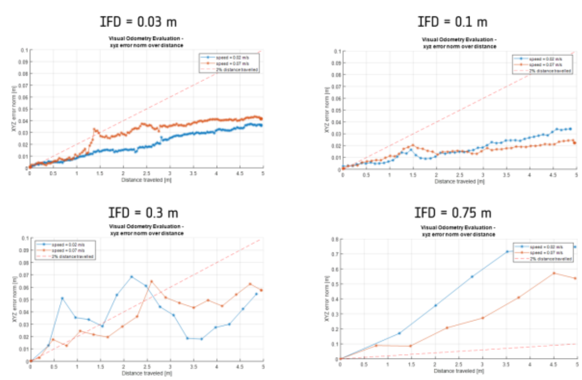

Figure 11: High and Low speed at different IFDs.

This suggested that both IFD and IOP are very good metrics to isolate the influence of the velocity and study the effect the VO frequency has on the frames and therefore the VO performances.

All the tests run suggested the existence of an upper limit for the IFD of 0.3 meters and a lower limit for the IOP of 75%, beyond which the VO accuracy starts to degrade, regardless of the speed. This was then validated with dedicated tests at both low and high speed, whose results are presented in Figure 12 and 13.

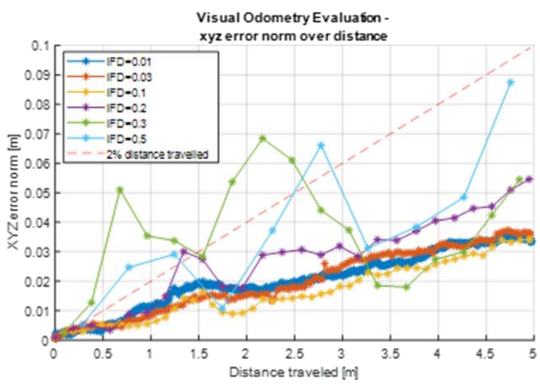

Figure 12: High speed tests at different IFDs.

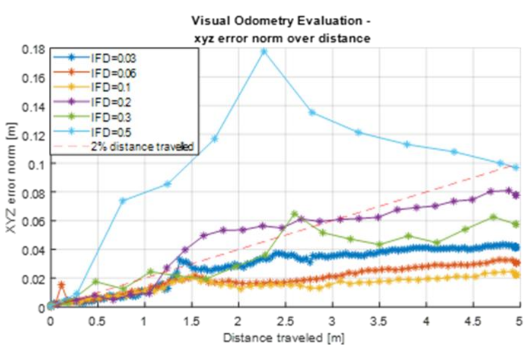

Figure 13: Low speed tests at different IFDs.

4.4 NavCam

All the tests described so far were run on the LocCam, but it is extremely interesting to also investigate different camera position and orientations. To do so the other available camera, the NavCam, has been used in the next tests. The NavCam is the same model as the LocCam (Bumblebee BB2-08S2C) but it is located on top of the mast on an actuated PTU, as shown in the schematic in Figure 14.

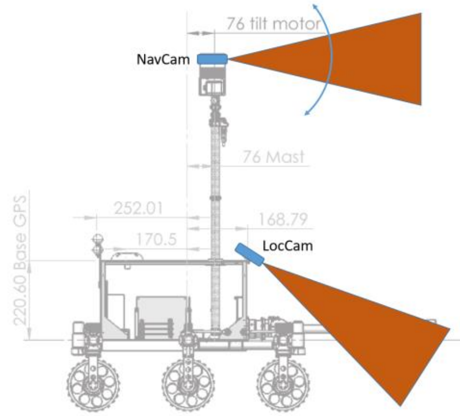

Figure 14: NavCam on the ExoTeR rover.

The major differences between the NavCam and the LocCam are that the former is in a higher position, approximately 1 meter above the ground, and has the possibility of changing the pitch of the camera, while to LocCam was fixed on the rover's body with a pitch of 30 degrees with respect to the rover body frame. Changing the pitch of the camera has the effect of decoupling the IFD and IOP, since different pitch values will still lead to the same IFD (which is only function of the rover's speed and VO frequency) but will now mean different IOPs. For example, with the same speed and VO frequency, a camera pointed to the ground below the rover (pitch=90deg) will lead to a considerably smaller IOP than a camera pointed forward, like in the case of the LocCam.

Many tests, with different combinations of pitch values, speed and VO frequency, were executed. It was concluded that the NavCam can provide acceptable VO estimation only with a pitch between 30-40 degrees from the horizontal, and even in that window it performs generally worse than the LocCam. In addition to the worse performances, the VO run on the NavCam did not change no matter the IFD and IOP. An example of the results from different pitch values is shown in Figure 15 and clearly shows how the pitch of 30and 40 degrees stays below the 2% error line, along with the LocCam, while 20 and 50 degrees diverge rapidly.

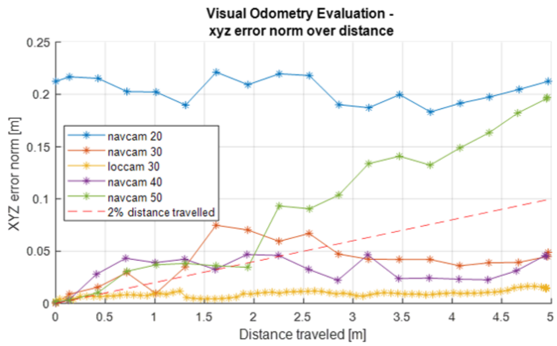

Figure 15: NavCam tests results with different pitch.

Considering the results obtained so far, the NavCam showed in general inferior performances compared to the LocCam, no matter the IFD and IOP. This could then be caused by two factors:

- The higher position of the camera does not allow close features to be used in the motion estimation, even with a high pitch, which will then lead to a very small IOP and the loss of far features and therefore low VO accuracy.
- The camera mounted on the PTU is considerably more sensible to vibrations of the mast when the rover moves, compared to the LocCam bracket which is better fixed to the ExoTeR body.

To further investigate these two possible reasons, additional tests were planned and executed. For both cameras a 3.5 m long traverse at 0.07 m/s on the sandy terrain in the center of the Mars Test Bed was compared with two other interesting sequences:

- Buried Rocks: small pebbles were placed in the rover's path to evaluate the effect of the mast vibrations when the rover moves over them. In order to avoid the positive effect of adding more features, the rocks were buried in the sand, as shown in Figure 16.

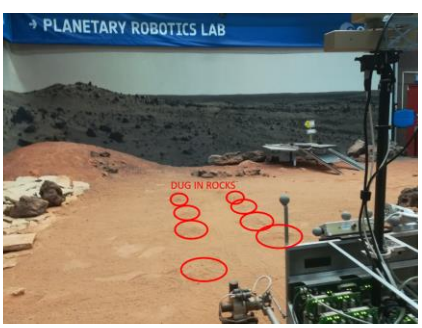

Figure 16: Terrain for the buried rocks test.

- Near Features: the rover moves parallel to the tables and various equipment in the lab, this has the effect of adding a considerable amount of close features, though only on one side, that the NavCam was missing before, as shown in Figure 17.

Figure 17: Terrain for the near features test.

The following results, Figure 18, were obtained. The LocCam performed very similarly in all 3 sequences, with a small decrease in accuracy in the rocks traverse compared to the normal sand traverse. The NavCam instead showed considerably lower performances with a steeper error curve in the buried rocks sequence, while the near features sequence gave a small increase in accuracy. These tests confirm that the worse performance in the NavCam VO is indeed induced by the vibration of the mast while the rover moves and the lack of close features due to the NavCam position on top of the mast.

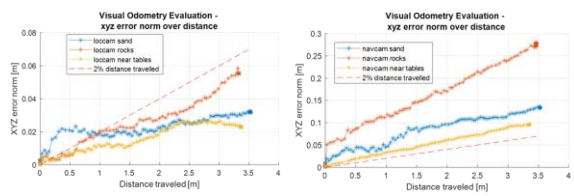

Figure 18: LocCam (left) and NavCam (right) in the buried rocks and near tables tests.

5 CONCLUSIONS

The testing activity was focused on the aspects that will change from the ExoMars mission to the SFR of the MSR mission. Increasing the velocity separately did not lead to any considerable loss in performances, and it could be compensated by lowering the Exposure Time in order to reduce the motion blur effect. In the case of a low ambient light scenario, the best solution that was found was to lower the speed of the rover to recover some of the lost performances with respect to a fully lit scenario.

To better describe the effect of the VO frequency, two concepts were defined and introduced: the IFD and IOP. As thoroughly explained in the Subsection 4.3 regarding the VO frequency tests, a region of these two values, out of which the accuracy of the VO starts to degrade, was experimentally found and validated in successive tests.

An additional aspect which was investigated is the kind of terrain the rover is going to traverse. The VO showed good performances in all the differet terrains available in the PRL, keeping the error around 1-1.5%. Another interesting test was performed based on recently acquired reference images for the SFR landing site. The SpartanVO was tested in a very similar terrain available in the PRL and the results showed a very good robustness of the VO with respect to this new challenging terrain.

Then, the option of using one of the other available cameras was investigated. It was concluded that the NavCam is not able to perform as well as the LocCam, possibly due to the two following reasons: the lack of close feature-points and, mostly, the vibrations of the mast when the rover moves.

All in all, the SpartanVO proved to be very robust for every new condition it was tested in and it was able to achieve very good accuracy, well below the 1-2% of the state-of-the-art.

5.1 Future Work

It would be interesting to focus even more in detail on peculiar aspects of the MSR mission and SFR. One possibility would be to study longer traverses, which are very interesting for the MSR mission, and could not be possible during this thesis due to the limited 9 by 9 meters Mars Test Bed of the PRL. Another interesting development could be to make the VO more adaptive. In this work it has been experimentally demonstrated how the rover speed and VO frequency are crucial to obtain good localization performances, and their relationship is well characterized by the IFD and IOP. Instead of running at a fixed frequency, the VO could then estimate online the appropriate frequency according to the current motion of the rover.

References

[1] Maimone M, Cheng Y (2007) Two Years of Visual Odometry on the Mars Exploration Rovers.

[2] Bora L and Nye B, (2017) Exomars rover control, localisation and path planning in a hazardous and high disturbance environment.

[3] Shaw A and Woods M (2013) Robust visual odometry for space exploration.

[4] Kostavelis I, Boukas E, Nalpantidis L, Gasteratos A (2013) Visual odometry for autonomous robot navigation through efficient outlier rejection.

[5] Wayman A and Meacham P (2018) Engineering Challenges or a Sample Fetch Rover.

[6] Merlo A and Larranaga J (2018) Characterization of a Visual Localization Solution for an ESA Mars Rover.

[7] Kostavelis I, Boukas E, Nalpantidis L, Gasteratos A (2016) Stereo-based visual odometry for autonomous robot navigation.