VISION-BASED ACCURATE PLANETARY ROBOTIC ARM PLACEMENT

Anton Donchev [(1)], Calum Murray [(1)], Yang Gao [(2)], Affan Shaukat [(2)], Wissam Albukhanajer [(2)], Daisy Lachat[(1)]

[(1)] Airbus Defence & Space Ltd, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2AS, United Kingdom, Email: anton.donchev@airbus.com, gnc.uk@astrium.eads.net

[(2)] University Of Surrey, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom, Email: yang.gao@surrey.ac.uk

ABSTRACT

The ability to accurately place robotic arm payloads on a target is critical to investigate scientifically interesting sites for sample handling missions such as Mars Sample Return (MSR), and its precursor missions (e.g. to Phobos or to the Moon).
A new innovative technique is presented within this paper which uses visual data to control the positioning of a robotic arm (i.e. visual servo-control) developed by Airbus Defence & Space ltd. This is a very promising technique which has been designed to enable the precise and repeatable placement of instrumentation while also automatically compensating for mechanical deflection of the arm under various levels of gravity.
In addition, the vision-based approach allows the same control system to be used for both sample collection and transfer tasks as required by missions such as MSR.

1. INTRODUCTION

Accurate placement of robotic arm payloads is essential for planetary missions aiming to probe and analyse sites of scientific interest. In particular, sample handling missions, such as Mars Sample Return, require the development of techniques to carry out the entire sample handling chain.

This paper presents the work produced within the scope of a UK Space Agency NSTP-2 project to develop a vision-based robotic arm control system which aims to address this challenge.

This study developed building blocks of a system capable of performing the complete sample handling chain via the processing of visual data acquired from a monocular camera attached to the robotic arm. This class of control technique (known as *Visual Servo Control*) allows the flexure or deflection of the arm joints or limbs to be compensated automatically, as well as removing the need to precisely locate the target point with respect to the fixed platform.

In addition, the use of vision data to control the positioning of the robotic arm allows the same control system to perform both sample collection as well as the sample transfer.

In order to achieve this task, novel image processing techniques are used in order to detect and track the location of the final target with respect to the mounted monocular camera. Two separate cases are considered for the target:

- firstly a rock from which the sample will be obtained, and
- secondly, the receptacle where the sample must be deposited.

During sample transfer, a dedicated fiducial marker is used in order to accurately determine the position of the sample receptacle, and hence help guide the robotic arm towards it.

The proposed system is illustrated in Fig. 1 below.

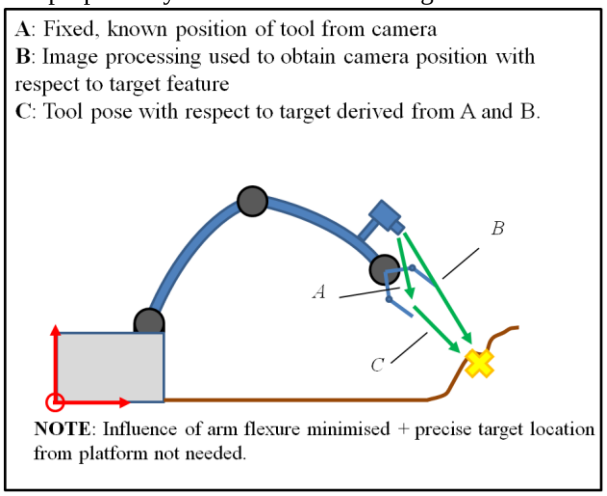

Figure 1. Monocular camera mechanical configuration

The techniques developed within the presented study are expected to enable the entire sample handling chain to be achieved for future planetary missions. They are also relevant for other applications, such as the accurate placement of scientific instruments or the exchange of end-effector tools. Beyond the space sector, it is anticipated that the use of some of these techniques will find applications in other industries that require robust autonomous operation of robotic elements, such as precision agriculture or nuclear decommissioning.

Section 1.1 provides an overview of the overall system

and Section 2 describes the simulated test environment. Section 3 describes the design of the control system – both for structured and unstructured environment cases, finally Section 4 presents the results, followed by a conclusion and description of future works in Sections 5 and 6.

1.1. OVERALL SYSTEM

The overall system considered for this project was a sampling platform which has two receptacles for sample collection. Sample collection and return is achieved via a 6-Degree of Freedom (DoF) robotic arm with an end-effector capable of rock sampling.

For this type of vision-based control system, a monocular camera is attached near the end effector of the robotic arm, after the last movable joint, which is commonly referred to as “eye-in-hand” configuration, as opposed to "end-point open-loop" where the camera is fixed elsewhere in the world. A monocular camera is used in order to reduce the mass and power requirements of the vision-based system with respect to a stereo-camera.

This camera output is then processed in order to provide the top level control inputs to the control system and navigate the arm’s end-effector towards the desired target.

In addition a stereo-camera system mounted on to the platform itself is used to select a sampling target in addition it is also used to provide the initial as depth to the control system.

A prototype system to test this concept was implemented by Airbus Defence and Space Ltd. and the University Of Surrey. It was tested in a simulated environment in order to preliminary assess the system performance.

2. SIMULATED ENVIRONMENT

2.1. PLATFORM

The platform which is considered for this study is a lander equipped with two sample receptacles, each one being identified with a different fiducial marker.

In addition, the platform supports a mast with a stereo camera which is used to select the initial sample target and determine an estimated depth to the target.

The platform includes a 6-degree of freedom robotic arm which is based on the LARAD robotic arm developed by Airbus Defence and Space Ltd. [2]. This arm has a monocular camera mounted near the end effector after the last joint.
Fig. 2 shows the simulated platform used to test the vision based control system.

This particular configuration of the platform has two receptacles and therefore two tracking markers. This was done in order to test the system under a more complex environment (e.g. when more than one visual marker was present).

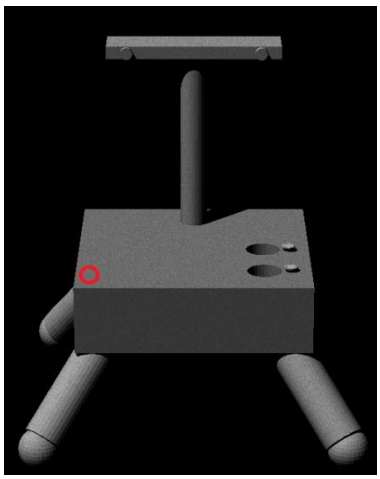

Figure 2. Simulated Sampling Platform

The robotic arm itself was not added to the visual simulation as it is not in the field of view of the monocular camera during nominal operation. It is also out of the field of view of the stereo cameras during the initial stereo image acquisition.

This simulated platform represents a use case similar to MSR. While only this platform configuration was used for testing within the scope of this study, the system design is not limited to working with such configuration type and will be capable of operating with other platforms as well.

2.2. SIMULATED VISUAL ENVIRONMENT

In order to have the vision-based simulations as close as possible to the real world testing, a camera model with several visual distortion effects was incorporated into the environment.

The vision simulation environment used an Airbus Defence and Space Ltd. proprietary software vision simulator, called SurRender [3] which is traditionally used space scene simulation. This was configured to provide the visual environment required for testing the control system.
In addition, a camera model, which simulates the following optical and sensory effects, was incorporated into the simulation:

- Lens Flare artefacts
- Modular Transfer Function and Point Spread functions

- Vignetting
- Camera thermal and electronic noise

These effects are introduced in order to provide confidence in the vision-based tracking algorithms, showing that they are capable of operating when the input images are distorted and contain error due to real camera system imperfections.

2.3. ROBOTIC ARM DYNAMICS

The Robotics-Toolbox [5] was used to provide the rigid-body kinematics simulation of the arm.

In addition, a deflection model was added, which off-sets the position of the camera in the direction of the gravity vector.

During testing, the parameters of the arm were varied around the baseline in order to test the robustness of the control algorithm with respect to different physical arm properties.

3. THE VISION-BASED CONTROL SYSTEM

The implemented system is split into two distinct operating modes.

The first mode, which will be referred to as the *structured tracking mode*, is used to move the robotic arm end effector towards the designated receptacle on the platform. This is the mode used for sample return and is provided in order to accurately guide the end effector towards the sample receptacle.

The second mode, which will be referred to as the *unstructured tracking mode*, allows the system to navigate to a user selected rock. This mode is used for the sample collection part of the acquisition chain of the system.

A third mode is required to move the robotic arm to a pre-defined position, so that either the structured target or the rock to be sampled is in the field of view of the monocular camera. This mode was not implemented within the scope of this study, however this does not require accurate positioning and a simple inverse kinematics calculation could be used. Note that for this third mode, a path planning with obstacle avoidance calculation may be desired.

3.1. ARCHITECTURE

Fig. 3 shows the overall architecture of the vision-based control system for both modes.

For the unstructured control mode it can be seen from Fig. 3, as already mentioned the stereo camera is only used during the initial target selection to provide the selected target and estimated depth to the selected target, for the structured tracking mode the stereo camera is not required.

After this initial step the monocular camera is used in order to control the robotic arm to the final unstructured target.

The structured target mode architecture and system is very similar to the unstructured target system, only varying the target feature set, which is pre-defined.

In addition, the image processing is changed in order to accommodate the different targets which are to be tracked.

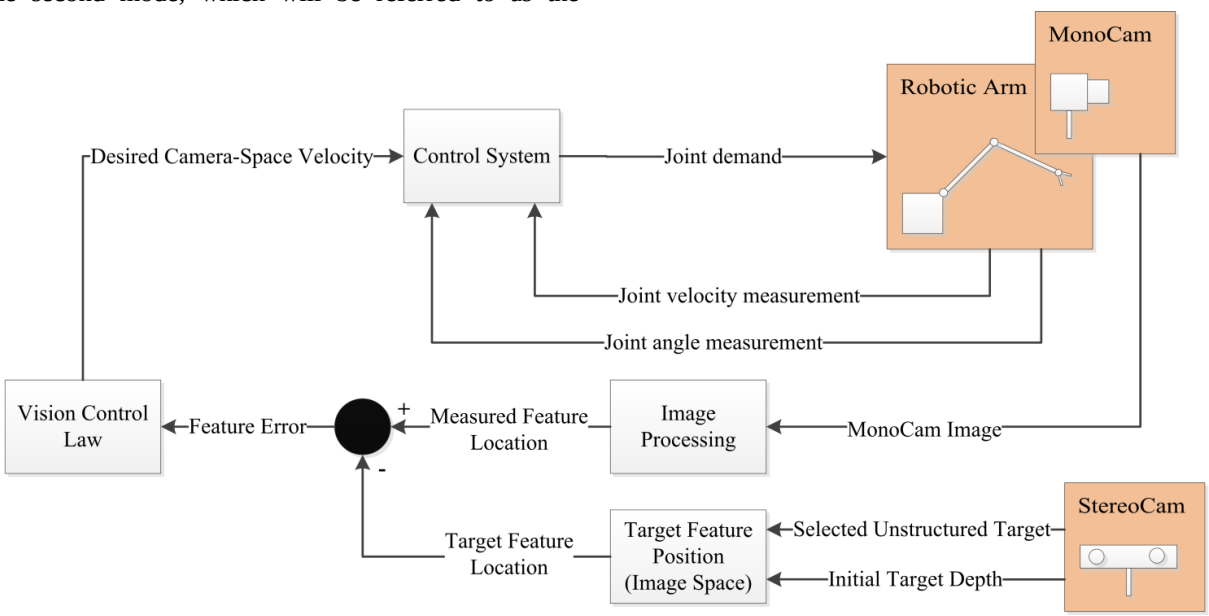

Figure 3. Unstructured Vision-based Control Architecture

3.2. UNSTRUCTURED TARGET TRACKING

The unstructured tracking algorithm was developed by the University Of Surrey.

A hyper-salient feature tracking algorithm was used to allow tracking of surface features of a rock in order to allow vision-based control of the robotic arm.

It also incorporates the target selection, as part of the algorithm. Some of the tracked features on a simulated rock surface can be seen in Fig. 4.

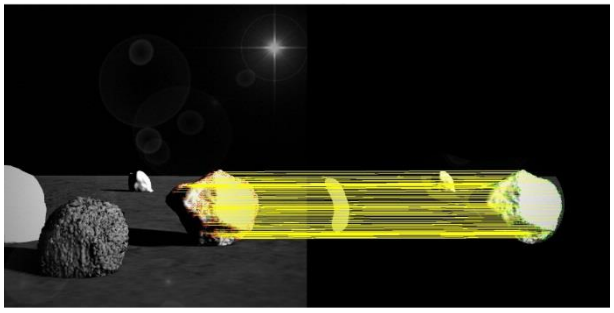

Figure 4. Rock face feature tracking.

These features are then provided to the vision-based control algorithm which uses them as control points in order to move the robotic arm towards the target point on the surface of the rock.

3.3. STRUCTURED TARGET TRACKING

The structured target tracking utilises a fiducial marker as the vision target, this is used in order to simplify and accelerate the retrieval part of the system and it only requires one marker per receptacle to be mounted.

For the testing performed as part of this study, two markers (and receptacles) were mounted on the platform in order to test the ability of the structured tracking to distinguish between different fiducial marker targets.

For the fiducial marker, a RUNE-Tag [1] was selected as the fiducial marker which will be used by the tracking system. The RUNE-Tag consists of a set of circular dots which are arranged co-centrically around the middle of the marker, as seen in Fig. 5.

This tracker must be placed in such a way that it is always within the field of view of the monocular camera during navigation to the receptacle.

This is the case as the control system uses the features in order to control the position of the camera and by extending the end effector as this is assumed to be rigidly connected to the monocular camera.

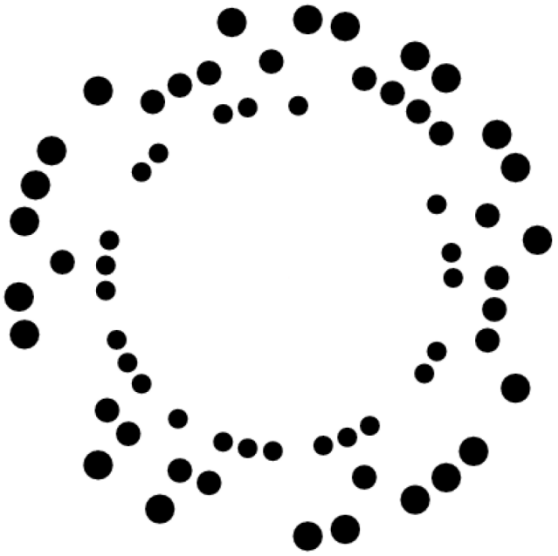

Figure 5. RUNE-tag fiducial marker example [1]

As the tracking algorithm does not require all points of the marker to be visible, meaning that the fiducial marker tracking is highly resilient to occlusion of the marker. RUNE-tag was shown to be detected without degradation up to 50% occlusion [1].

This is crucial for operations on Mars, as dust deposition on the surface of the fiducial marker will produce occlusions which the vision-based tracking system must be able to cope.

Another potential source of occlusion is shadowing, which may occur either from external sources or from the limbs of the robotic arm itself. This was not tested explicitly as it would have the same effect as the test already performed.

3.4. CONTROL SYSTEM

The control algorithm used for the system consists of two parts arranged in a nested control loop: the inner part is a controller which commands the robotic arm joints and the outer part is a navigation system which uses vision based algorithms in closed loop, in order to guide the robotic arm to the final position of the end effector.

The output of the vision outer loop of the control algorithm is camera velocities. The inner loop of the control then attempts to match the demanded velocities of the end effector. Fig. 6 shows the six Cartesian-space velocities demands of the camera (three for positional velocity and three for angular velocity).

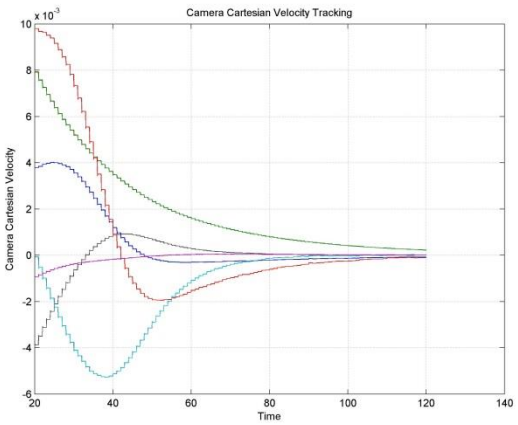

Figure 6. Convergence of Camera space velocities

The control system uses an iterative approach in order to solve the positioning problem.

By separating the control system into two distinct loops, it is possible to switch which vision-based control and tracking system is used for the outer loop. This enables the same control system to perform both the sample collection, where the robotic arm must reach the rock target, and the sample return to the receptacle, where the arm must accurately control to the deposition receptacle.

In order to achieve this, a third control mode is envisioned that moves the arm to a predetermined position. This mode, in the case of the unstructured tracking, will position the arm so that the field of view of the monocular camera intersects the field of view of the stereo cameras mounted on the platform. From this location the vision-based control mode can be activated. For the structured target case, a pre-determined position can be defined before flight which ensures that the fiducial marker is in the field of view of the monocular camera.

For both cases it is not required for the positioning of the robotic arm to be very accurate, as long as the target is within the field of view of the monocular camera.

4. RESULTS

4.1. STRUCTURED TARGET TRACKING RESULTS

The fiducial marker tracking algorithm was tested within the simulated environment. Both resilience to occlusion and ability of the system to detect the marker at different light levels were assessed.

Fig. 7 shows fiducial marker detection at the furthest possible distance considered for this project.

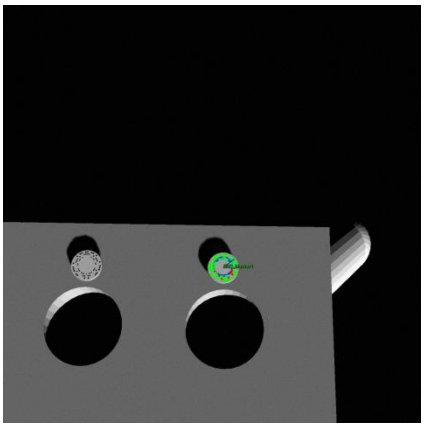

Figure 7. Fiducial Marker tracking at a far distance

In Fig. 7 green points signify detected ellipses of the marker and red points signify points which were not found. This is determined by comparing a mathematical model of the fiducial marker versus the detected points in the image.

The same test was repeated in a low light environment as presented in Fig. 8.

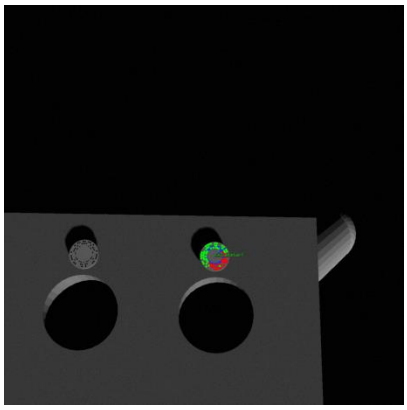

Figure 8. Fiducial Marker tracking at a far distance in low light conditions

Fewer points of the fiducial marker are detected in the low light environment, as can be seen from Fig. 8 which is expected, however the tracking algorithm is still capable of tracking the fiducial marker.

The tracking algorithm detection improves as the robotic arm end effector approaches to the marker. This signifies that the control system will be able to work under both the low light and bright light conditions from the far distance.

For both tests a secondary marker was included, as shown in Fig. 7 & Fig. 8. This was done in order to test if the algorithm is capable of detecting only the selected marker, which would allow more than one marker target to be added to the sampling platform. In both cases, the second marker was correctly ignored by the system.

In addition, several tests were performed in order to measure the occlusion resilience of the algorithm. This can be seen in Fig. 9, the marker is detected correctly when it is partially outside of the field of view of the camera.

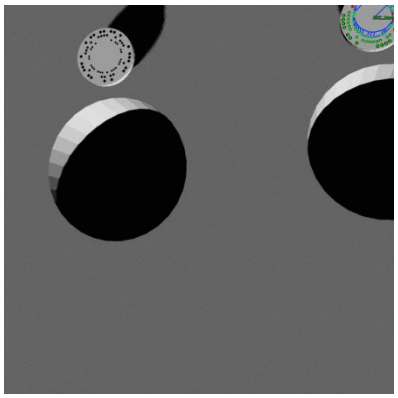

Figure 9. Fiducial Marker tracking when the marker is occluded

All marker tracking tests were performed with the camera model effects turned on as described in Section 2.2.

4.2. UNSTRUCTURED TARGET TRACKING RESULTS

This tracking algorithm was tested with two data-sets: an in-lab physical setup, and a simulated environment with a variety of rocks. The algorithm was capable of tracking a preselected rock at different distances and angles. A sample from both data-sets is shown in Fig. 10.

Figure 10. A sample of images from the Airbus Defence & Space Ltd simulated (left) and STAR Lab at the University Of Surrey (right) data

In order to test the performance of the algorithm, the bounding box of the selected target rock was compared to the human annotated target box. This can be seen in Fig. 11 where the blue outline box is the hybrid-salient feature tracking and the red outline box is the human annotated target.

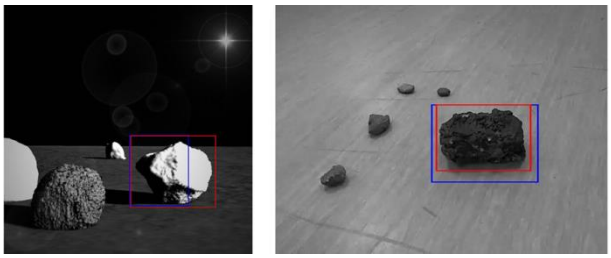

Figure 11. Comparison of the bounding box output from the hybrid-salient feature tracker (blue) against human annotated ground truth (red) using the Airbus Defence & Space Ltd (left) and STAR Lab at the University Of Surrey (right) data sets

Performance was measured in the simulated case with the results shown in Fig. 12 below. This ratio is the ratio between the overlapping areas of the selected and measured target.

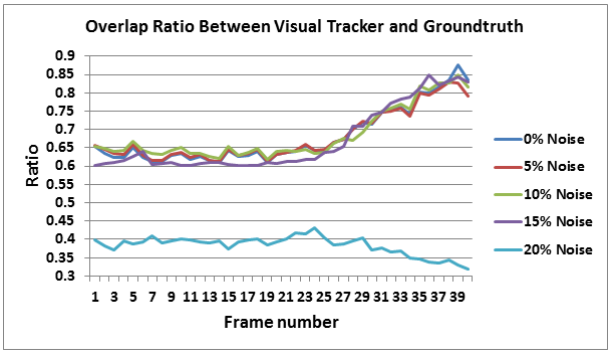

Figure 12. Ratio of the spatial overlap between the bounding box outputs from the vision system and the ground-truth human annotations

Mars can have extreme weather conditions where noise due to planetary dust cannot be ignored. It is therefore an important requirement for the vision systems to be robust against this type of noise that can severely degrade the quality of input images. A common factor for noisy visual inputs is granulated dust particles. Such type of noisy conditions can be simulated by adding "salt-and-pepper" noise to images Each input image is corrupted by adding "salt-and-pepper" noise over a varying scale of 0 – 20% (with 5% increment), resulting in 4 distinct samples for each individual frame in terms of varying noise levels. This allows checking the performance in noisy conditions, and it is in addition to normal camera noise.

Referring to Fig. 12, it is worth noting that the vision system seems to outperform the minimum spatial overlap threshold requirement for noise level as high as 15%. Beyond 15% the overlap drop below 0.5, however it is very much expected that the mission activity will be halted at such high degree of adverse weather conditions. The images at 0% noise and the two higher noise levels are shown in Fig. 13 below.

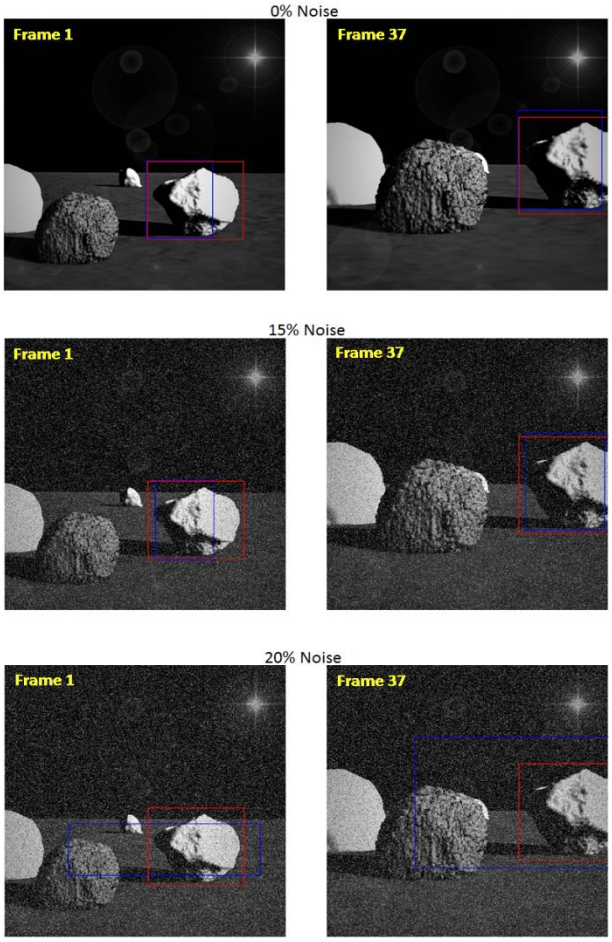

Figure 13. Comparison of the vision system output against ground-truth at different noise levels and different instants of the data sequence

The images used to test the vision tracking algorithm, were taken as the camera approached the target rock, which can be seen as an increase of ratio as the camera gets closer to the object in Fig. 12. At the end of the frames, the target rock begins to be partially outside of the view frame of the camera. This can be seen on Fig.8 in frame 37. This is also replicated in Fig. 12, as the ratio begins to dip towards the end of the curve.

4.3. INTEGRATION AND CONTROL RESULTS

The vision-based control system was integrated with both the structured and unstructured tracking systems and tested in simulation.

This was tested by setting the initial position of the arm such that the target rock or fiducial marker (depending on which tracking system was in the loop for the particular case) is in the field of view of the monocular camera.

The control algorithm was run until the least absolute deviation error (otherwise known as the L1 Norm) between the target feature location and the actual measure features approached zero.

This error is calculated as per Eq. 1 below.

$$S = \sum_{i=1}^{n} |\overrightarrow{y_i} - f(\overrightarrow{x_i})| \quad (1)$$

Where $\overrightarrow{y_i}$ is the current position of the feature point,
$\overrightarrow{x_i}$ is the calculated target position of the feature point

The measured feature points are the points provided by each separate tracking algorithm. In the structured case, these are six of the fiducial marker points detected by the tracking algorithm. The target point is pre-determined. All feature points are calculated in image space.

For the unstructured tracking algorithm, the target point is calculated based on the output of the vision-based algorithm. This is compared to the measured feature points and used to determine the overall positioning error.

Fig. 14 and Fig. 16 show this feature error reducing as the simulation progresses. This demonstrates that the system is correctly controlling the arm towards the final desired position of the end effector.

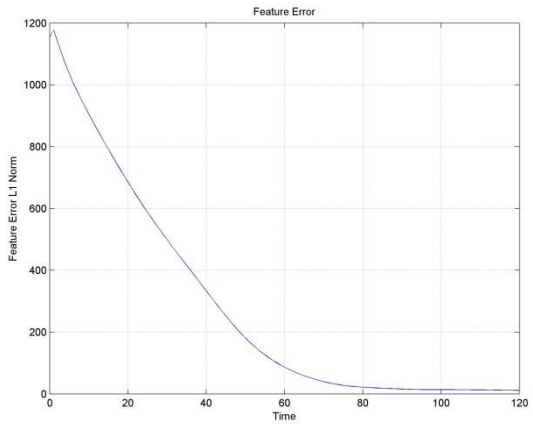

Figure 14. Structured tracking case feature error with versus time

Fig. 15 shows the initial position (on the left) as seen from the monocular camera when the test begins. On the right on Fig. 15 is shown the image at the final position of the monocular camera.

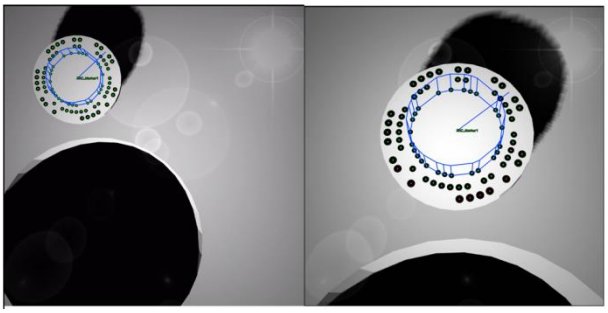

Figure 15. (left) Monocular camera position at beginning of test, (right) Monocular camera position at end of test

From Fig. 15 it can be seen that the camera has moved towards the fiducial marker. As the camera is slightly offset from the centre of the robotic arm end effector, the end effector itself is positioned directly over the receptacle.

For the unstructured tracking and control case the Feature error is shown in Fig. 16.

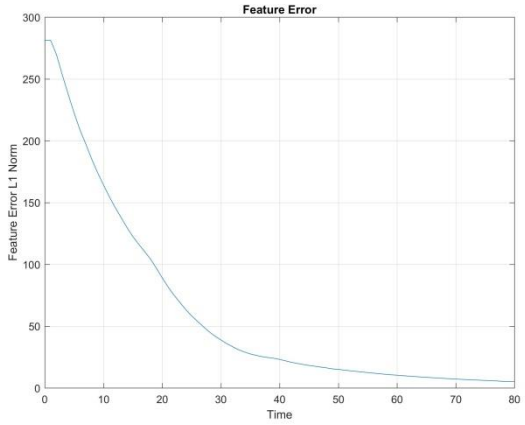

Figure 16. Unstructured tracking case feature error with versus time

In addition, images of the initial and final position of the monocular camera are shown in Fig. 17 below.

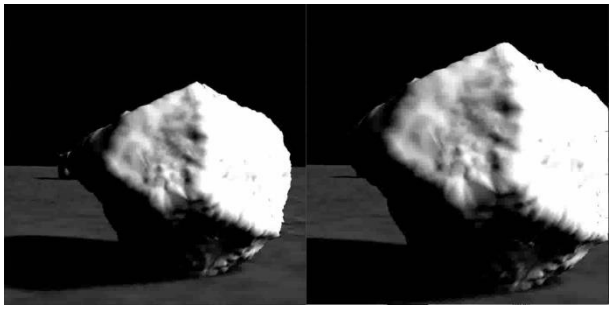

Figure 17. (left) Monocular camera position at beginning of test, (right) Monocular camera position at end of test

As can be seen from Fig. 17, the robotic arm end-effector approaches the rock target.

5. CONCLUSION & FUTURE WORK

The algorithms have been successfully developed and integrated. The resulting system is capable of controlling a robotic arm using vision-based feedback to a desired position, for both sample collection and transfer. Preliminary positional performance assessment showed the system to be fully functional and highly promising to tackle the tasks at hand. Furthermore, the study has demonstrated the feasibility of applying cognitive vision techniques (i.e. visual saliency) in addressing visual servo challenges using unstructured features seen on planetary rocks.

During the initial study, it was identified that the system needed specific control modes to complement the implemented vision-based control system. Notably the critical need to incorporate a recovery action in case of image tracking failure and the limitations of the vision-based control at very close range calling for the design of a specific final approach mode. This presented a problem as the overall system is not robust to failures and does not reach the target upon contact.

In addition the current unstructured tracking is only used to control the position of the camera, the orientation of the camera is a pre-set vector used by the control system, in the future this system will be extended so that the orientation of the camera and by extension the end effector is in line with the normal of the surface of the rock.

Another limitation which is currently imposed on the system but will be updated in future iterations is the ability to select a specific point of interest on the surface of a rock, currently the tracking and control algorithm allows the selection of a specific rock and then controls to the centre of the selected rock target.

The final improvement to be added in the future is an improvement in the depth estimation from the monocular camera which would make this system more applicable to non-contacting scientific instruments.

6. ACKNOWLEDGMENT

This research work has been funded by Airbus Defence and Space Ltd. the University of Surrey and the UK Space Agency.

7. REFERENCES

[1] F. Bergamasco, A. Albarelli, E. Rodola, and A. Torsello. *Rune-tag: A high accuracy fiducial marker with strong occlusion resilience*. In Computer Vision

and Pattern Recognition (CVPR), 2011 IEEE Conference on, pages 113-120, 2011.

[2] R. Marc et al. *Design Concepts and Implementation of the Lightweight Advanced Robotic Arm Demonstrator (LARAD).* 6th International Astronautical Congress (IAC), 2016.

[3] R. Brochard, N. Despre, *SurRender: Image Generation software for scientific-level space scene simulation.* IPPW-13 06_PO_05 Poster Presentation.

[4] Abhinav Bajpai, Affan Shaukat, Guy Burroughes, Yang Gao, *Planetary Monocular Simultaneous Localization and Mapping (PM-SLAM)*, Journal of Field Robotics, Volume 33, Issue 2, pages 229–242, 2015, DOI:10.1002/rob.21608.

[5] Corke, Peter. "A robotics toolbox for MATLAB." *Robotics & Automation Magazine, IEEE* 3.1 (1996): 24-32

[6] Yang Gao, (Ed.), Contemporary Planetary Robotics – An Approach to Autonomous Systems, pp. 1-450, Berlin: Wiley-VCH, ISBN-10: 3527413251, ISBN-13: 978-3527413256, August 2016.

[7] Affan Shaukat, Yang Gao, Jeffrey A. Kuo, Bob A. Bowen and Paul E. Mort, "Visual Classification of Waste Material for Nuclear Decommissioning," Robotics and Autonomous Systems, Volume 75, Part B, Pages 365-378, 2016, doi: 10.1016/j.robot.2015.09.005.

[8] Abhinav Bajpai, Affan Shaukat, Guy Burroughes, Yang Gao, Planetary Monocular Simultaneous Localization and Mapping (PM-SLAM), Journal of Field Robotics, Volume 33, Issue 2, pages 229–242, 2015, DOI:10.1002/rob.21608.

[9] Conrad Spiteri, Yang Gao, Said Al-Milli, and Aridane Sarrionandia de León, Real-time Visual Sinkage Detection for Planetary Rovers, Robotics and Autonomous Systems, Vol 72, pp. 307–317, 2015.

[10] Guy Burroughes and Yang Gao, "Ontology-Based Self-Reconfiguring Guidance, Navigation, and Control for Planetary Rovers". AIAA Journal of Aerospace Information Systems, 2016, doi:10.2514/1.I010378.

[11] Collewet, Christophe, Chaumette, *Visual servoing on non-planar objects from active vision,* 2007, IEEE