Autonomous Vision-based Tethered-assisted Rover Docking

Dorian Tsai[1], Issa A.D. Nesnas[2], and Dimitri Zarzhitsky[3]

***Abstract*— Many intriguing science discoveries on planetary surfaces, such as the seasonal flows on crater walls and skylight entrances to lava tubes, are at sites that are currently inaccessible to state-of-the-art rovers. The *in situ* exploration of such sites is likely to require a tethered platform both for mechanical support and for providing power and communication. Mother/daughter architectures have been investigated where a mother deploys a tethered daughter into extreme terrains. Deploying and retracting a tethered daughter requires undocking and re-docking of the daughter to the mother, with the latter being the challenging part. In this paper, we describe a vision-based tether-assisted algorithm for the autonomous re-docking of a daughter to its mother following an extreme terrain excursion. The algorithm uses fiducials mounted on the mother to improve the reliability and accuracy of estimating the pose of the mother relative to the daughter. The tether that is anchored by the mother helps the docking process and increases the system's tolerance to pose uncertainties by mechanically aligning the mating parts in the final docking phase. A preliminary version of the algorithm was developed and field-tested on the Axel rover in the JPL Mars Yard. The algorithm achieved an 80% success rate in 40 experiments in both firm and loose soils and starting from up to 6 m away at up to 40° radial angle and 20° relative heading. The algorithm does not rely on an initial estimate of the relative pose. The preliminary results are promising and help retire the risk associated with the autonomous docking process enabling consideration in future martian and lunar missions.**

I. Introduction

A. Motivation and Background

Recent orbital observations have resulted in intriguing scientific discoveries on planetary surfaces. For example, the discovery of seasonal flows [1], such as those observed in the Newton crater, are on steep slopes (25°- 40°) that are hundreds of meters down from the crater rim. *In situ* analysis and sample capture of outflow deposits, that have interacted directly with water on Mars, makes their retrieval scientifically important for future return to Earth [2], [3]. The sampling of exposed strata on crater walls provides insight into the composition, structure, and history of Mars. Access to pit chains and collapsed lava tubes on Mars and the Moon would be of interest for human exploration as they could serve as potential habitats for astronauts and safe havens from solar radiation [4]. Other planetary bodies, such as Titan and Europa, offer similarly challenging terrains [5]. The Mars Exploration Rovers (MERs), and the Mars Science Laboratory (MSL) rover were designed to traverse slopes up to 20°and 30°respectively [6]. Access to much steeper terrains is likely to require tethered robots, in particular, when terrain properties are not known *a priori*. A tether/umbilical not only provides mechanical support for the rover, but also provides power and communication in the absence of direct sunlight for energy and line-of-sight for communication.

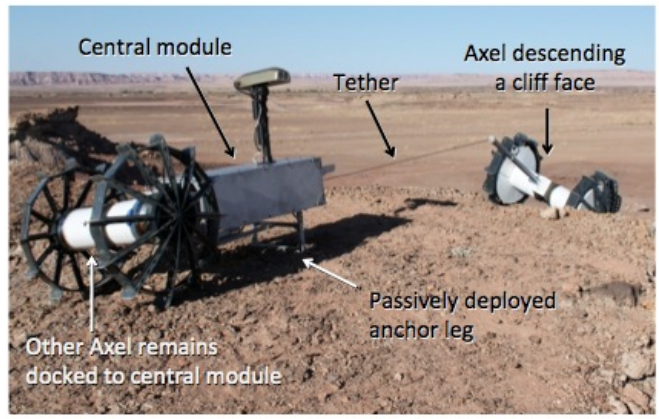


Fig. 1: The Axel rover (right) descends a cliff, while the central module and another Axel (left) acts as an anchor

One mother/daughter architecture to address extreme terrain access is the DuAxel/Axel platform (Fig. 1). The Axel rover (Fig. 2) is a two-wheeled, differentially-driven, tethered rover capable of rappelling steep slopes and traversing rocky terrain [7]. The DuAxel rover is formed by docking two Axel rovers to either side of a central module (CM). In a typical scenario, the four-wheeled DuAxel rover traverses untethered to an extreme terrain site such as a crater or a cliff face and anchors itself at a safe distance from the edge. The two-wheeled Axel rover then undocks from the central module and descends over the edge into the crater. Following its excursion, Axel re-docks to the central module and the now

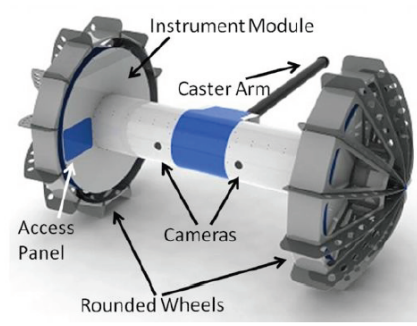


Fig. 2: The Axel rover with major components annotated

*This work was supported by NASA/JPL Research and Technology Development Program, the Keck Institute for Space Studies, and Erasmus Mundus Foundation

[1]D. Tsai carried out this research at the Jet Propulsion Laboratory. He is with the School of Electrical Engineering, Aalto University, Helsinki, Finland, dorian.tsai at utoronto.ca

[2]I.A. Nesnas is with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, {issa.nesnas} at jpl.nasa.gov

[3]D. Zarzhitsky was at the JPL. He is now with Amazon.com

reconstituted DuAxel rover drives to a new site.

The undocking and redocking of the tethered rover to its mother is a critical part of a mother/daughter architecture. Given the communication delays and limited bandwidths (up to 40 minutes round-trip between Earth and Mars and few communication windows), docking and undocking lend themselves to being done autonomously. Docking is defined as "moving from an initial, to a desired position and orientation, while following a safe trajectory" [8]. For the Axel rover, docking involves guiding the rover with its caster arm into a funnel-like docking port on the central module, while actively managing the tether that connects the two.

B. Related Work

Autonomous docking has been successfully demonstrated in a number of applications. Mobile service robots rely on docking stations to recharge their batteries [9], [10]. Docking mechanisms can extend functionality in a variety of ways, acting as communication waypoints [11], activating autonomous power lifts for wheelchairs [12], harvesting produce in horticultural environments [13], and enabling unattended vehicle parking [14]. In the aerospace industry, autonomous vision-based rendezvous and docking is used for connecting to the International Space Station [15] and is being developed for satellite servicing. The main differences between our application and some of these systems is that our docking strategy uses a combination of forces generated by wheel/soil interaction on uneven terrains, subject to different soil conditions, and forces generated by the pull of a tether. Moreover, our docking has to be robust to different outdoor lighting conditions.

Vision-based docking have either used visual servoing or vision-guided approaches. Both use optical feedback to control the robot's motion by first comparing the current system configuration to the desired state, generating an error term, and then regulating the error to zero over time. Visual servoing often requires efficient computation to handle frequent visual updates for better system response, which allows for small image change assumption in the algorithms. In contrast, a visual-guided approach typically involves the post-processing of a minimal number of image features to estimate the pose of the robot, planning a path based on this pose, followed by a step toward the docking station [16]. Because of the limited computational and power resources of flight missions, most space applications have adopted a vision-guided approach. We employ such an approach to enable the infusion of such algorithms into future flight missions.

The rest of this paper describes an approach for autonomous vision-based tether-assisted docking of a daughter rover to its mother. Section II describes the overall approach, while Section III details the approach used to identify the pose of the mother relative to the daughter in an outdoor environment. Section IV describes the motion planning approach while Section V presents the field experiments in the JPL Mars Yard and analyzes its overall performance. The summary of our results and lessons learned are outlined in Section VI.

II. Concept of Operations

In a mother/daughter architecture, there are two strategies for vision-based docking: eco-docking, where the perception sensors are mounted on the stationary (mother) platform and ego-docking, where the perception sensors are mounted on the moving (daughter) platform. While there are advantages and disadvantages to each approach, Santos-Victor and Sandini have shown that both share the same control [17]. For the DuAxel rover, eco-docking has advantage of greater computation in the stationary platform (central module), higher vantage point for the mast-mounted perception sensors, and the ability to use the same sensors for docking the two opposite Axels. The drawback is a more complex architecture that require hand-shaking between the central module and the rover. More importantly, the mast-mounted perception sensors would not have an optimal view of the mating parts.

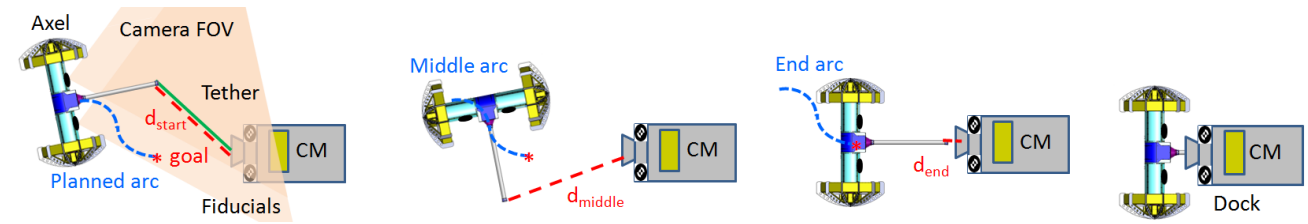
Given the drawbacks and the availability of sensors onboard the rover (daughter) for navigation, we selected an ego-docking approach. The current Axel rover is equipped with a stereo camera pair with a wide (35 cm) baseline to accommodate the tether spool. The rover can pitch its cameras independent of the rover's motion and uses an inertial measurement unit to measure the rover's roll and pitch and estimate yaw. The tether is managed by the Axel rover and is paid out as the rover traverses away from the central module and picked up as the rover returns. This is in lieu of reeling the tether from the central module, which increases abrasion of the tether on the terrain. The reeling of the tether plays an important part in the docking process.

The docking concept is shown in Fig. 3, and the software flow diagram is shown in Fig. 4. We use a three-dimensional vision-guided approach to docking that uses a sense-plan-act sequence that is typical for planetary applications. The algorithm running on the rover uses fiducial markers mounted on the central module to detect and compute its relative pose. As the rover approaches the central module, the relative pose uncertainty decreases. Then, it uses a motion planner to position and orient the rover such that it aligns its arm with the docking cone. If the misalignment error is outside acceptable bounds and the rover is too close, the rover reverses its direction to allow enough space to re-estimate the pose and retry the maneuver. Once the rover's arm contacts the docking cone (based on a depth estimate), the algorithm puts the wheels in free rotation mode and reels in the tether 25 cm inside the central module, aligning the rover as it enters the docking cone. In the final phases of this process, the fiducial markers are too close to be identified.

III. Target Identification

A. Visual Markers

The rover's docking algorithm requires reliable detection of the relative pose of the central module from five or so meters away to less than half a meter. It has to do so under varying lighting conditions, with large perspective distortion



(a) Vision locates the CM. Rover plans S-arc (2-arc) maneuver to align itself with CM

(b) Rover traverses first arc (showing predicted tether length at the midpoint)

(c) Rover aligned with the CM

(d) Rover reels itself in to complete the docking

Fig. 3: Concept of operations for the Axel rover docking

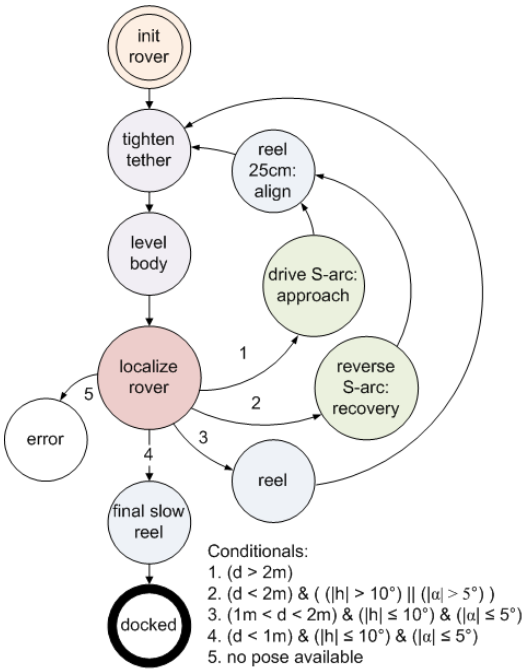


Fig. 4: Software flow of tether-assisted docking algorithm. Docking regions are shown in Fig. 7

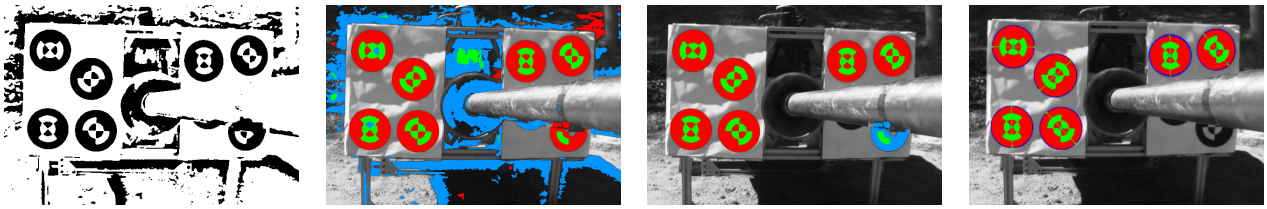
(oblique angle and orientation) and with occlusions from the arm/tether. We use visual markers (i.e., fiducials) to reduce computation and increase robustness. The algorithm not only detects the fiducials but also determines their correspondence to an *a priori* model to estimate pose.

1) Design: Circular fiducials are often used due to their compact shape, rotation invariance, and simple centroid location, which can be determined with sub-pixel accuracy [18]. We selected 10 cm diameter fiducials with four possible rotations in 45°increments (-45°, 0°,45°, and 90°) with easily detectable centroids that are perspective invariant (Fig. 5), similar to the ones used on the Curiosity rover [19]. We experimented with unique fiducial tags (such as ARTag [20] of similar size), but found that they were often indistinguishable with fish-eye (90°) lenses from more than 4 m away, and were more sensitive to partial occlusions. As a result, we chose an asymmetrical configuration of eight large fiducials to efficiently detect and compute the unique pose of the central module in light of partial occlusions. A homography transform requires detecting only four fiducials.

2) Detection: The fiducial detection algorithm uses a multi-step pipeline with the major steps shown in Fig. 5. The first step uses adaptive thresholding to handle lighting variations while still providing sufficient contrast for the blob detection [21]. The threshold levels were chosen empirically from a large data set. While generally edge detection may be less sensitive to lighting variations, limitations in depth of focus across a large distance range would necessitate more sophisticated filtering to detect the defocused edges at the distance extremes. The second step used a blob filter to identify connected black/white regions based on their size, eccentricity, Euler number, and coincident centroids of black and white blobs. To capture the large variation in appearance, thresholds on individual filters had wide ranges and hence were not very effective individually, but collectively they provided robust detection. The third was a refinement step based on the fiducial design. A finer, local thresholding was performed to recover the lost internal pattern of the candidate fiducial. The centroid of the blob was computed by fitting the perimeter of the blob to an ellipse using a least-squares fit [22]. Following the blob analysis, very few false positives remained, which were eliminated by the correspondence algorithm.

3) Correspondence: The correspondence problem has been extensively studied with a large body of published work [23], [24]. This algorithm has to compute the correspondence of a partial set of detected fiducials to an *a priori* known model. With eight fiducials, there are $8!$ (40320) possible permutations for correspondence. However, taking into account fiducials orientations and physical constraints (e.g. the central module cannot be upside down), the number of possible correspondences reduces to 16, which allows for a much simpler and computationally efficient algorithm. For each permutation, a set of initial correspondences is assumed, forming a hypothesis. The relationship between the hypothesis and the known model is determined using a homography. The inverse homography is then used to reproject the model onto the current image, which then allows the hypothesis to be verified by comparing the projected model to the remaining fiducials.

Fig. 6 shows a comparison of the model reprojection to measured fiducials using a closeness measure. This is used to detect occlusions and false positives. Then all the fiducials within the closeness threshold are used to compute an overall



(a) (b) (c) (d)

Fig. 5: Fiducial detection process *from left to right*: (1) thresholding and connected compoments, (2) area and eccentricity elimination (blue), (3) centroid pair elimination and B/W ratio elimination (blue), and (4) major-axis orientations and ellipses.

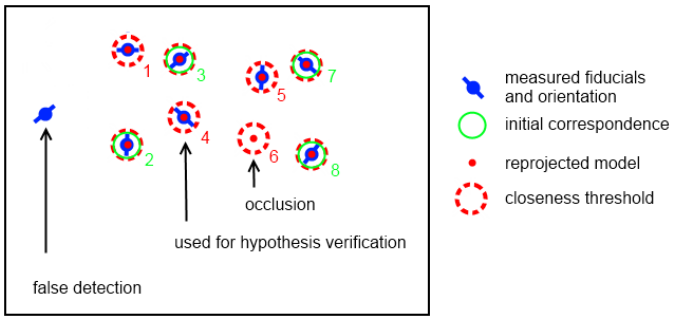
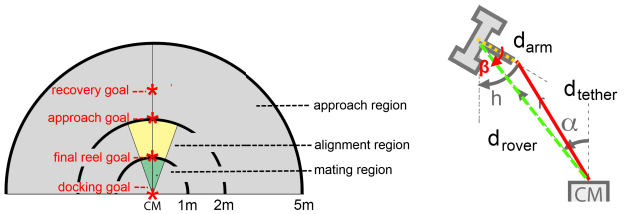


Fig. 6: Fiducial correspondence on the image plane, showing initial correspondence, model reprojection, an occlusion and a false detection.



(a) Docking regions and goal points (b) Taut tether distance

Fig. 7: Docking parameters

homography, and the model is again reprojected onto the image. The Euclidean distance between the non-occluded projections and the measured fiducials (excluding false detections) is finally calculated as the image reprojection error. Of the 16 permutations, the permutation with minimal error is selected as the correct correspondence.

It is important to note that even if not all 8 fiducials are detected (due to occlusions or false negatives), a successful correspondence allows the homographic projection of the model to approximate the missing (undetected) fiducial centroids. We independently apply this to each of the left and right images. The detected fiducials together with the projected fiducials from the model complete the set of 8 fiducials for each of the left and right images. The complete set is used by the point-stereo algorithm to estimate the pose, which allows for stereo-based pose estimates even if the algorithm does not detect the same fiducials in the left and right images. Combining detected and projected fiducials may introduce some inconsistencies in the pose, but the magnitude is small compared to noise and calibration errors.

B. Target Pose Estimation

Dense stereo can be used to compute the fiducials' three-dimensional points in a world frame. Alternatively, for uncalibrated cameras and to reduce computation, one can run the fiducial detection on both the left and right stereo images and only compute the point stereo of individual fiducials using triangulation.

To compute the pose of the central module relative to the rover, we used a least-squares minimization to match the 3D points of the model to the detected 3D points, similar to the Iterative Closest Point algorithm.

Monocular-based pose estimation is used to handle contingencies when the fiducials on the central module are in the field-of-view (FOV) of only one camera. This situation occurs when the rover takes a sharp turn. To estimate the pose, we use the Robust Planar Pose Estimation algorithm, specifically designed for planar objects, and commonly used in augmented reality [25]. Due to time constraints, this algorithm was only implemented using the CAHV model[1], which does not compensate for lens distortion, and was hence less accurate.

IV. Docking Framework

Given the relatively small terrain undulations near the docking site, we use flat terrain approximations for the motion planner to control the (x, y) position, pitch and heading of the rover relative to the docking station. For tethered-assisted docking, we use radial coordinates, where r denotes the radius and α denotes the angular position relative to the central module (Fig. 7). The rover's orientation is denoted by its heading h relative the central module. To dock the rover to the docking station, the algorithm basically positions and aligns the Axel rover's arm at the docking cone within a certain error bound and then reels itself in.

One can break down the docking maneuver into three regions: the approach region, the alignment region and the mating region (Fig. 7). The tether usually relaxes the stringent docking requirements on pose uncertainty and motor control accuracy. In the approach region, the rover approaches and aligns itself to the docking station (central module). Generally, in this region, there is a low risk of collision, so the rover can reverse and realign itself if the alignment error is too large as a result of slippage, poor perception or pose estimation. In the alignment region, the

[1]The linear version of the CAHVORE model.

rover proceeds toward the mating position while refining its alignment. In the mating region, precise positioning is required and speed is reduced to avoid collision or jamming. The rover rechecks its alignment then puts its wheels in free rotation mode. Then it uses the mating geometries and the tether to pull and align itself in the mating position. As of this writing, the rover did not have a sensing capability to confirm the completed docking. However, stereo vision pose estimates from the penultimate step proved sufficiently accurate and reliable to use as a condition for termination.

A. Path Planning

1) Arc Trajectories: Arc maneuvers are used on the Ackermann-steered planetary rovers, since they are often resource-constrained, which prevents simultaneous steering and driving. While in the case of the two-wheeled rovers, such a constraint is not applicable, we leveraged existing arc-based motion planning algorithms [26]. A single arc is required to move any point on the rover to an (x, y) goal. However, in order to control both position and final heading of the rover (x, y, h), a minimum of two arcs are needed. This underconstrained problem has an infinite number of two-arc maneuvers that can achieve that goal. To resolve the redundancy, we imposed an additional optimization constraint that minimizes both the distance traveled and the amount of turning to keep the docking station within the camera's FOV as much as possible. Minimizing turning also reduces risk of tether entanglement.

2) Tether Management: While the tether assists the docking process, it poses additional challenges for motion planning. When the tether is slack, Axel's arm trails the rover's traverse direction to provide the necessary reaction force for the rover's motion. However, for the docking process, the arm must lead (i.e. be in front of the rover) to prevent tether entanglement and enable mating. A taut tether can provide the necessary reaction force to overcome the wheel terrain friction in a leading arm configuration. One can envision an algorithm that servos on the tether tension during the docking process. Unfortunately, at the time of the experiments, the rover did not have an integrated means for measuring the magnitude nor direction of the tether tension, so we relied on the nosier motor current measurements to infer tether tension and used an approximate motion predictive model to manage the tether during the docking maneuver. To maintain proper tether tension during motion, we used a geometric approach to compute the instantaneous exposed tether length d_{tether} from the tip of its arm to the docking station (Fig. 7) using the cosine law as follows:

$$d_{tether} = \sqrt{d_{arm}^2 + d_{rvr}^2 - 2d_{arm}d_{rvr}\cos(\alpha - h)}, \quad (1)$$

For each arc in the parallel parking maneuver (Fig. 3), we computed the change in tether length and used this linear approximation to control the tether length. In future work, we will either use the instantaneous tether length or tether tension to control the reel motor.

Fig. 8: Autonomous docking experiments with the VICON ground truth system in the JPL Mars Yard.

V. Experimental Results

A. Field Tests

We conducted a total of 40 autonomous docking experiments in the JPL Mars Yard using two terrain types: (a) loose sand and (b) compact mixed sand/gravel terrain. We tested docking from a radius of up to 7 m away and an angle α of up to $40°$, which was the limit of what the cameras could see, where only part of the docking station would be visible at the boundary of one of the stereo images (Fig. 9). Cases where the docking station was entirely outside the camera's FOV were not tested. For ground truth measurements with millimeter accuracy, we used the VICON system that uses six tripod-mounted infrared cameras to track the 6-DOF pose of the rover using an asymmetric configuration of rover-mounted reflective balls (Fig. 8).

B. Overall Performance

Of the 40 autonomous docking tests, 29 were entirely successful. Fig. 9 shows the different tests and annotates some of the failures. The tests demonstrated successful docking from a maximum distance of 6 m, with a relative heading of $20°$, and up to $40°$ in radial angle. This distance limitation was expected since the fiducials were designed to be discernable from 5 m. Excluding experiments on the the workspace boundaries, 20 out of 24 tests (or 80%) were successful for this initial version of the docking algorithm.

Figures 10 and 11 show results from two experiments that show the points of image acquisition, planned S-arcs and actual path based on ground truth data from VICON, as well as the pose estimates and planned S-arcs based on vision measurements. The corresponding roll, pitch and yaw histories are also shown. From these figures, it is evident that the vision pose could exhibit inaccuracies as large as 30 cm from 4 m away; however, in this self-correcting algorithm, as the rover approaches the docking station, the vision estimates become more accurate. Fig. 10 shows a typical run, but Fig. 11 shows a situation where following the rover's first step, the arm's misalignment error is still large. The rover detects and corrects this misalignment by taking

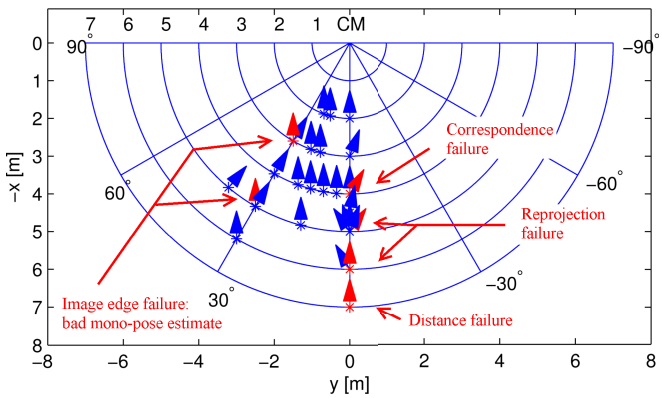


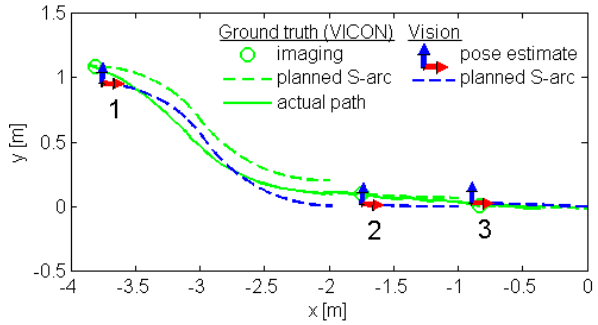
Fig. 9: Docking experiment grid. Arrows indicate starting position and orientation, with successful docks (blue), and failures (red).

a subsequent step in the reverse direction and retrying the approach. Oscillation between the forward and reverse steps is possible, though only observed once in the experiments. Imposing limit cycles for such recovery steps would address such oscillations.

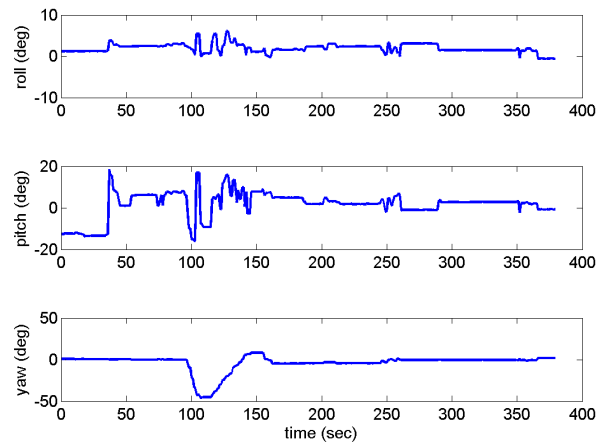
Comparing the rover's planned (desired) trajectory with the actual ground truth data, it can be seen that the rover was able to follow its desired path fairly well on different terrains (loose sand and compacted terrain) while managing the tether. Discrepancies are likely due to a combination of uneven wheel diameter from the wheel paddles, wheel slip, and coordination of the wheel's motion with the tether. A qualitative comparison on the effect of ground type was made during the experiments. Half of the tests were performed on compacted terrain, and the other half on soft sand. On compacted terrain, Axel's motion involved significant tumbling and rolling, owing to its large, discrete grousers. When reeled in, a "side-to-side" motion was often induced from alternating grousers contacting the ground. This caused Axel to occasionally exceed the $\pm 5°$ heading tolerances, and drive Axel back to the recovery goal point. As expected, this behavior was not observed on the soft sand, since the grousers sank into the sand, resulting in a smoother traverse. The corresponding two-arc maneuvers can be seen in the yaw measurements at time $t = 100$ seconds for both runs. The large pitch angles in Fig. 11, are attributed to the slack in the tether during a two-arc maneuver. Tether management is discussed in Sec. V-F.

C. Pure Reeling vs. Active Docking

Vision-based tethered-assisted docking was compared to the control experiment of (non-vision) tethered-based docking. With the tether-based, or perhaps more accurate "tethered-forced" docking, the rover puts its wheel in free rotation mode and uses only the spool motor to reel and align itself with the docking station. The tether-based docking experiments were tested from a distance of up to 4 m and from range of angles and orientations. Distance was not a significant factor for reeling. Pure reeling demonstrated successful docking for $\alpha < |5°|$, and $h < |90°|$. However,

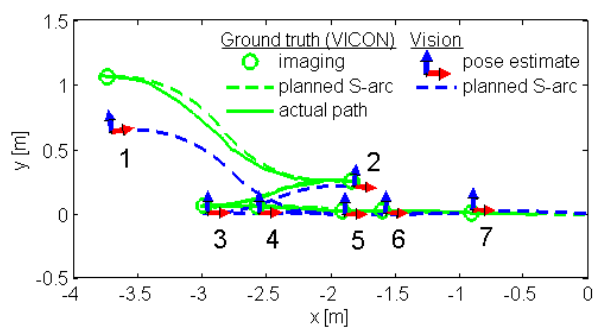


(a) X-Y Position

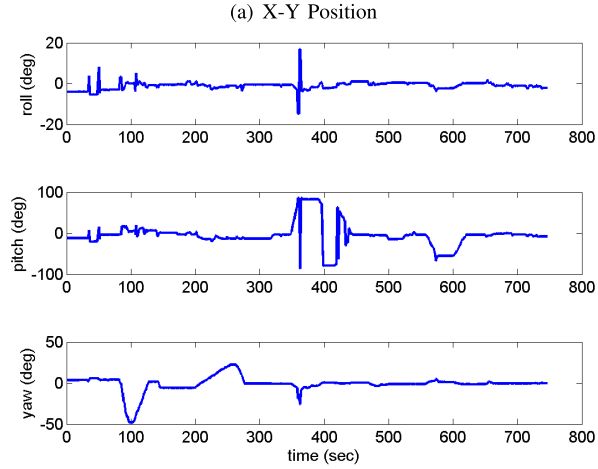


(b) Roll, pitch, yaw time history

Fig. 10: Typical dock from 4m distance, $15°\alpha$, and $0°$ h.



(a) X-Y Position



(b) Roll, pitch, yaw time history

Fig. 11: Typical dock illustrating a recovery maneuver.

for $\alpha \geq |5°|$, Axel pushed against the docking station cone during the final dock, which shifted the docking station horizontally to accommodate Axel's arm. This produced large internal loads on the structure. Additionally, higher currents of 1.5A–2.0A were recorded on the spool motor. These relatively high loads placed unnecessary additional stress on the spool motor.

Conversely, the vision-based tether-assisted docking was limited to docking with $h < |20-30°|$ due to the FOV of the lenses, and was only tested up to $\alpha \leq |40°|$. Lower currents of 0.5A–1.0A were observed on all motors. Distributing the motor load leads to less wear and tear on individual motors and would likely increase the operational lifetime of the robot. It is clear that the vision-based docking algorithm enabled wider α docking capability and better motor properties, but the tether-based docking was able to handle much larger heading angle h. Tether-assisted docking can serve as a backup approach in case of failure of the vision sensors. Moreover, a combination of tether-based and vision-based tether-assisted approaches could yield significantly improved docking range and heading capability.

D. Fiducial Performance

Fiducial detection proved reliable under good lighting, but some sensitivity to uneven lighting conditions was observed. For the first two days of testing, lighting conditions were fairly uniform or lightly speckled across the fiducials. Under these conditions, the docking algorithm was 84% successful (21/25 runs), only 2 of which were attributed to failed detection. However, on the third day, the sun cast an uneven and intense shadow over the front of the central module. This reduced the contrast of the fiducials and resulted in more failures due to fixed thresholds. The docking performance dropped to 53% (8/15 runs), which could have been addressed by adjusting the adaptive threshold value; but for consistency, it was held constant during the experiments. In future work, we will incorporate pre-processing steps to handle intense partial shadows, such as using a variable adaptive threshold, an initial ellipse edge detection, or acquiring and merging of images with multiple exposures.

As expected, fiducial detection at the image boundaries was less reliable due to more compressed fiducials with higher lens distortion. Fewer blobs were detected on the boundaries, which limited the heading range h. The eccentricity constraint could have been increased to handle the extreme perspective, but for consistency, detection parameters and software were held constant throughout the experiments. The parameters were determined empirically from several image data-sets, which did not have a full representation of the extreme cases performed during the field tests.

Fiducial correspondence was able to handle most occlusions and occasional false positives; however, 5 of the 11 failures were attributed to a known bug in the correspondence module for handling false positives. Despite typically having 6-7 detected fiducials, only 4 were input to the consensus-based homography algorithm. But a minimum of 4 points are required to define a homography, so if one of the points was invalid, which was then forced to fit with the remaining points, the data-set became degenerate and the homography failed. To handle this problem, at least 5 points should be used to determine the homography, allowing the algorithm to reject outliers, and still maintain a reasonable population for homography fitting.

E. Pose Estimation

Stereo-based pose estimation was reliable given correctly measured and corresponded fiducials. On average, the reprojection error was 8.21 ± 2.4 pixels, approximately 1 pixel/fiducial. At 5 m, this amounted to 20 cm in position error when using a 3.5 mm focal length lens, but the error decreased as the rover approached the docking station. At 2 m, the mean error was 4.12 cm for r and and 4.58° for h and 2.56 cm and 2.5° at 1 m, which were sufficiently accurate for reeling and docking.

The performance degraded when the rover was initially pointed away from the docking station. The perception system suffered from partial visibility of the central module, significant lens distortion at the image boundaries, and a very wide stereo camera baseline (35 cm). This caused fiducial detection or correspondence to fail, and forced the algorithm to rely on monocular pose estimation. The latter did not account for the lens distortion, and in turn, produced unreliable pose estimates and ultimately resulted in docking failure. With proper lens distortion models, monocular methods can be improved to serve as a backup for stereo failures.

F. Tether Observations

For motion planning, the most challenging issue was perhaps tether management. The triangular tether approximation was sufficient for most docking maneuvers; particularly those involving turns of less than 30°. For tighter arcs, the shortcomings of the linear tether approximation became apparent. An excessively taut tether prevents Axel from driving very wide arcs, which results in significant slippage and high torque on the spool motor. On the other hand, a loose tether results in changing the rover's pitch without making forward progress as the tether did not provide the needed reaction force. This caused the large pitch in Fig. 11. Fortunately, even if significant slip occurred, as long as the rover could detect the docking station, it iterated over the docking procedure. Over a few iterations, Axel was able to compensate for the slip or pitch change as it incrementally reeled in the tether and moved towards the docking station. Many tests suffered from the slack tether during a wide arc, but were still able to dock successfully. This showed that the tether was able to compensate for some pose inaccuracies and locomotion challenges leading to successful docking. Future work would use finer trajectory control either by using spline-based motor profiles in lieu of the currently used trapezoidal profiles or by using frequent trajectory updates during the spool motor's motion. With the availability of tether tension and direction sensors, force-based control algorithms can be used for smoother and visually optimal trajectories (*i.e.* maximizing the time the docking station stays in the rover's FOV).

During the initial tests of the passive reeling experiments, the tip of arm got occasionally caught against the flat face of the docking cone, which prevented further reeling. Future designs of the docking station and tether arm must account for this interaction. But this docking failure was circumvented by tipping the docking station forward by $5°$, which angles the docking cone toward the ground. Following that, the docking cone interaction caused only one failure out of 40 field tests, proving passive reeling to be reliable for docking.

VI. Conclusion

In conclusion, autonomous vision-based tethered-assisted mating of a rover with a docking station from a range of initial poses can be made reliable in an outdoor environment under varying lighting and terrain conditions without the need for high precision perception, motion estimation or control. We have demonstrated a vision-based algorithm that detects fiducials mounted on a docking station, matches the 3D fiducial constellation to an *a priori* model of the docking station to compute its pose, and generates a sequence of visual pose updates as the rover approaches the docking station using parallel parking maneuvers to position and orient the rover. The algorithm was designed to handle partial occlusions that result from the tether arm and to use monocular pose estimation as a fall-back to stereo-based pose estimation when the docking station is visible only in one camera. Moreover, the algorithm used a linear approximation for managing the tether length, which was sufficient for docking in most cases. The tether plays a significant role in guiding the rover's motion thus increasing the algorithm's robustness and reducing requirements on precise estimation and control. This algorithm was adapted to and field-tested on the Axel rovers in the JPL Mars Yard to understand its operational limits. Axel demonstrated reliable docking from up to 6 m away with up to $40°$ in radial angle, and $20°$ in relative heading from the docking station. The tests were carried out under a range of lighting conditions and on two terrain types: loose sand and a compacted mixture. The success rate of the algorithm under these constraints was 80%. All docking failures were anticipated and can be addressed in future work. The algorithm developed in this paper could serve as a baseline for docking of extreme terrain tethered rovers. Furture work will improve its robust for consideration in extreme terrain exploration missions.

Acknowledgment

This work was performed at JPL under contract to the National Aeronautics and Space Administration. We grateful for the sponsorship of JPL's Research and Technology Development Program, the Keck Institute of Space Studies, and the Erasmus Mundus Foundation. We would also like to thank Prof. Joel Burdick of the California Institute of Technology and Adnan Ansar of JPL, for their advice and support.

References

[1] L. Ojha, A. McEwen, C. Dundas, S. Mattson, S. Byrne, E. Schaefer, and M. Masse, "Recurring slope lineae on mars: Updated global survey results," in *43rd Lunar and Planetary Science Conference*, 2012.

[2] S. Squyres, *Vision and Voyages for Planetary Science in the Decade 2013-2022*. The National Academies Press, 2011. Committee on the Planetary Science Decadal Survey.

[3] M. E. P. A. G. (MEPAG), *Mars Science Goals, Objectives, Investigations, and Priorities*. 2010.

[4] F. Horz, "Lava tubes - potential shelters for habitats," in *Lunar bases and space activities of the 21st century*, pp. 405–412, Lunar and Planetary Institute, 1985.

[5] G. E. Cushing, T. N. Titus, J. J. Wynne, and P. R. Christensen, "Themis observes possible cave skylights on mars," *Geophysical Research Letters*, vol. 34, p. L17201, September 2007.

[6] J. D. Mitchell *et al.*, "Mars exploration rover mobility assembly design, test and performance," *Systems, Man and Cybernetics*, vol. 1, pp. 450–455, October 2005.

[7] I. A. D. Nesnas *et al.*, "Axel and duaxel rovers for the sustainable exploration of extreme terrain," *Journal of Field Robotics*, vol. 29, pp. 663–685, June 2012.

[8] N. M. Barnes and Z. Q. Liu, "Vision guided circumnavigating autonomous robots," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, pp. 689–714, September 2000.

[9] U. Kartoun *et al.*, "Vision-base autonomous robot self-docking and recharging," in *World Automation Congress*, July 2006.

[10] P. Petrov, C. Boussard, S. Ammoun, and F. Nashashibi, "A hybrid control for automatic docking of electric vehicles for recharging," *IEEE Int. Conf. on Robotics & Automation*, pp. 2966–2971, May 2012.

[11] F. Ferreira and R. Ventura, "Autonomous docking of a tracked wheels robot to its tether cable using a vision-based algorithm," *Workshop on Robotics for Disaster Response, IEEE Int. Conf. on Robotics & Automation*, 2009.

[12] H. Sermeno-Villalta and J. Spletzer, "Vision-based control of a smart wheelchair for the automated transport and retrieval system (atrs)," in *IEEE Int. Conf on Robotics & Automation*, pp. 3423–3428, February 2006.

[13] E. J. VanHenten *et al.*, "An autonomous robot for harvesting cucumbers in greenhouses," *Autonomous Robots*, vol. 13, pp. 241–258, 2002.

[14] M. Kazemi, K. Gupta, and M. Mehrandezh, "Path-planning for visual servoing: A review and issues," *Visual Servoing via Advanced Numerical Methods*, no. 2, pp. 189–207, 2010.

[15] A. Petit, E. Marchand, and K. Kanani, "Vision-based space autonomous rendezvous: A case study," *Intelligent Robots and Systems*, pp. 619–624, September 2011.

[16] E. M. P. Low, I. R. Machester, and A. V. Savkin, "A biologically inspired method for vision-based docking of wheeled mobile robots," *Robotics and Autonomous Systems*, vol. 55, pp. 769–784, 2007.

[17] J. Santos-Victor and G. Sandini, "Visual behaviors for docking," *Computer Vision and Image Understanding*, vol. 67, pp. 223–238, September 1997.

[18] V. Lepetit and P. Fua, "Monocular model-based 3d tracking of rigid objects: A survey," *Foundations and Trends in Computer Graphics and Vision*, vol. 1, no. 1, pp. 1–89, 2005.

[19] L. Jandura, "Mars science laboratory sample acquisition, sample processing and handling: subsystem design and test challenges," in *40th Aerospace Mechanisms Symposium*, pp. 233–248, May 2010.

[20] M. Fiala, "Designing highly reliable fiducial markers," *IEEE Trans. on Pattern Analysis & Machine Intelligence*, vol. 32, pp. 1317–1324, July 2010.

[21] P. Kovesi, "Matlab and octave functions for computer vision and image processing," 2012.

[22] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, pp. 476–480, May 1999.

[23] F. Ababsa and M. Mallem, "A robust circular fiducial detection technique and real-time 3d camera tracking," *Journal of Multimedia*, vol. 3, pp. 34–41, October 2008.

[24] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE A&E Systems Magazine*, vol. 19, pp. 5–18, January 2004.

[25] G. Schweighofer and A. Pinz, "Robust pose estimation from a planar target," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 2024–2030, December 2006.

[26] I. A. D. Nesnas, M. W. Maimone, and H. Das, "Rover maneuvering for autonomous vision-based dexterous manipulation," *IEEE Int. Conf. on Robotics & Automation*, March 2000.