SEAT TRACK LOCALIZATION AND TRACKING FOR ROBONAUT 2 MOBILITY ON THE INTERNATIONAL SPACE STATION

Barrett Ames\textsuperscript{1}, Seth Gee\textsuperscript{2}, Michael W. Lanighan\textsuperscript{3}, Robert R. Burridge\textsuperscript{2}

\textsuperscript{1}Duke University, Durham, NC, USA, E-mail: cbames@cs.duke.edu
\textsuperscript{2}TRACLabs Inc., Houston, TX, USA, E-mail: \{seth,burridge\}@traclabs.com
\textsuperscript{3}University of Massachusetts Amherst, Amherst, MA, USA, E-mail: lanighan@cs.umass.edu

ABSTRACT

Robonaut 2 (R2) was launched to the International Space Station (ISS) in 2011 as a torso with two arms and a head. In 2014, significant upgrades were performed to R2 on the ISS, including the addition of a pair of legs. Each leg has a specialized end-effector capable of both grasping handrails and docking with seat track. The seat track exists throughout ISS as an integral and predictable part of the structure. Handrails on the other hand are distributed less uniformly throughout the ISS due to crew member use. R2 lab personnel have previously developed the ability to localize and grasp handrails, enabling R2 mobility. In this work, R2's mobility options are extended by developing a complementary capability for seat track.

1 INTRODUCTION

In early 2011, NASA launched R2 on STS-133 to the ISS \cite{6}. This launch and the subsequent construction of the torso was part one of a three phase approach to developing a robot capable of assisting astronauts in Extra-Vehicular Activities (EVA). In phase 1, R2 was attached via a stanchion to the Destiny module of the ISS. During this first phase, experiments in robotic assistance under the conditions of microgravity were performed, with a specific focus on performing service and maintenance tasks. In 2014, as part of phase 2, a set of climbing legs (see Figure 1) was delivered and attached to the torso. The legs have grippers that were specifically designed for mobility grasping in the ISS (see Figure 2) \cite{1}\cite{5}. The work described henceforth is the first attempt at developing an automated procedure that would provide R2 the ability to grasp specific, but frequently-occurring, features of the space station.

2 TASK DESCRIPTION

In order to attain maximal mobility inside the International Space Station, R2's end effectors were designed to grasp two main features present throughout the space station. The first feature is the blue handrail. While being of a distinctive shape and color, the handrails are also used by the crew to move about the station, and as such are frequently re-positioned by the crew. This means that in some cases R2 might be unable to reach its destination if it relied solely on handrails. The other graspable feature is known as seat track. Seat track is prevalent...
throughout the station, and is rigidly attached to the structure of the station. Furthermore, the handrails used by the crew connect to seat track, so any location where the crew may wish for a handhold must already have seat track. An example of the material is in Figure 3. Thus, the task was to develop a procedure for safely and securely attaching the end effector of R2 to seat track, enabling mobility throughout the station.

2.1 Hardware Limitations

In order to perform the task safely, the gripper on the end of the leg is limited to applying a maximum force of 35N. If the gripper applies more than 35N, the gripper mechanism enters a fault mode which requires a physical reset. Furthermore, due to the tight mating tolerances of the seat track and gripper, any misalignment of the end effector with respect to the seat track may result in a failure to close.

3 METHODOLOGY

The approach taken to create the automated docking procedure follows a bottom-up development strategy. Each individual component was developed separately in order to make its performance as high as possible. Once the individual component’s performance was satisfactory, it was combined with other parts of the system, where it became part of an overall system test.

3.1 Sensor Preparation

The first components that were investigated were the sensors, as they provide the inputs for the control algorithms. The force torque sensors and the RGB cameras had previously been calibrated by the R2 lab, thus only the PMD Camboard Nanos required a thorough characterization and calibration. In addition to calibration, both the Camboard Nanos and the RGB cameras required significant filtering.

3.1.1 PMD Camboard Nano

The Camboard Nano is a reference design provided by PMD to demonstrate their sensor’s capabilities. The reference design was targeted towards short distance gesture recognition, not precise object recognition and localization. Since these devices were uncharacterized and poorly calibrated, a significant portion of time was devoted to understanding and calibrating the sensors. The cameras come from the factory with settings that are selected for general applicability, as opposed to high precision results. Using standard camera calibration procedures, the camera intrinsics were learned and then applied to produce a rectified point cloud. In Figure 4 the left most image is the initial point cloud received from the Nano. The cloud is warped, and it is very difficult to tell that there are distinct objects present in the point cloud. In the center image, however, the point cloud has been rectified. The rectification process makes edges more distinct and makes planar objects more closely resemble planar surfaces in the point cloud. After being rectified, the points are then passed through a depth discontinuity filter [15]. This filter attempts to remove points that lie along the boundary of objects that are neither part of the foreground nor the background. Figure 5 demonstrates output of the discontinuity filter.

In addition to the software steps taken to rectify and filter, several hardware factors were identified and normalized. In Figure 2 the camboard Nano is the small lens and LED array that are on the bottom left of the foot. Directly in front of them is the anodized aluminum gripper and the interior gripping surface. These surfaces are highly IR reflective. This reflective surface proves to be a problem for Time-of-Flight devices such as the Nano. The light reflected by these nearby structures hits much of the lower half of the sensor array and causes the readings to be heavily biased. The bias makes everything appear closer than they are in reality [11]. This was remedied via application of IR absorbent material along the gripper. In addition to the errors caused by the environment in which the sensor was used, there was at least one source of error from the particular implementation of the sensor. Specifically, the depth error was dependent on the temperature of the device [18]. In order to avoid this error the device was allowed to come to a steady state temperature before perform-
ing any tasks. Lastly, the kinematic calibration of the sensor was fine tuned by hand so that depth pixels aligned with the color pixels of the RGB camera.

Figure 5: The green points have passed through the discontinuity filter

3.1.2 Point Grey Research Flea 3 RGB Cameras

Although the RGB cameras were calibrated they still required significant filtering before they could be used successfully. The filtering process was designed specifically to highlight the features that were believed to be important for tracking and localization. The features selected for tracking were the screw heads that are located in the valley of the seat track. They were chosen for their distinct, closed geometry, as well as their different albedo and geometric relationship to each other. The first step in making the screw heads more apparent was an “unsharp” mask, as it increases contrast along boundaries. This was specifically selected to help compensate for any blurring that may occur due to the motion of the arm during image capture. Next, a homomorphic filter was applied to the image. The homomorphic filter removes low frequency sources of light, thus compensating for the washout created by the foot light source. The different stages of filtering can be seen in Figure 6.

3.2 Model Based Sensor Fusion

Unfortunately, once all of the sensors had been well calibrated and filtered, they were still unable to provide the necessary information to perform adequately. Thus, a model based fusion approach was performed to attain the highest level of performance.

3.2.1 Optical Localization

Optical localization proceeds in two stages, the first is a rough pass that finds a portion of the image to focus in on. The second stage finds the fine localization of the seat track. The filtered image developed in Section 3.1.2 is used as the basis for the model-based operations that follow. First, it was passed to an ensemble edge detector. One part of the ensemble was Canny’s edge detector. The other part of the ensemble was based on combining Sobel derivatives with an adaptive threshold. This method is referred
to as adaptive edge detection in the remainder of this work. The two edge images were then logically OR’ed together. The resulting image was pruned by thresholding on the length of each contour, if the contour was under a specific length, then it was ignored as noise. The different components of the edge detection ensemble can be seen in Figure 7. The resulting edges were used to find the outer edge of the seat track. This was accomplished by finding two long edges that were very close to parallel. The precision of this step is not particularly important, as it is used to create a sub-image for further processing. Thus, as long as the sub-image contains the portions needed for the next stage of the pipeline the actual size of the sub-image has little effect. A typical sub-image found by this process is depicted in Figure 8. The next step in the pipeline is to find the screw heads in the sub-image provided by the previous step. The screw heads are located by first applying a contrast limited adaptive histogram equalization (CLAHE). The benefit of CLAHE is to increase contrast in a local way, thus further fending off the washout caused by the LED in the leg. Following the CLAHE, a bilateral filter is applied to sharpen the edges in the image while blurring the spaces between the edges. The resulting image is Figure 8, which demonstrates a fairly clear delineation between the screw heads and the seat track. Images of the type depicted in Figure 8 are then passed into the fine localization portion of the pipeline. The fine localization is based on the centroids of screwhead-blobs that are detected in the sub-image. The blobs are detected by a rolling threshold approach. A threshold of a small range is applied to the sub-image, and then is OR’ed with all the other threshold ranges that are applied to the sub-image. The resulting blobs are not a 1 to 1 mapping of blobs to screw heads as can be seen in Figure 9. In many cases the false positives are of an odd shape, and thus success was found by filtering the blobs based on eccentricity and convexity. The resulting blobs are then fit to a line using RANSAC. Figure 10 demonstrates a typical linear fit that is output by the optical localization routine.
3.2.2 Optical and Depth Fusion

The next step in the seat track localization process combines the information from the optical and depth sensors. The camera works well for identification and 2D localization, but structure from motion 3D localization is quite difficult. Furthermore, the 3D sensor does not contain the necessary resolution for object recognition. Combining the two sensors allows for weaknesses to be compensated, and strengths to be exploited. In addition, including model information compensates for noisy or missing data.

The first step in the combination is to find the 3D pose of the 2D location of the screw head. This is accomplished by upsampling the depth image using [13]. This produces a high resolution depth image by combining information from both the low resolution depth image, and a high resolution grayscale image. This same step is applied to the normals of the point cloud as well. The use of point normals allows for an entire 6-DOF pose to be found for each point in the image. Once each point has been localized, a 3D line is fit to all of the points using PCA. The position of each screw head is then projected onto the 3D line. This projection encodes the domain knowledge that the seat track will always be a flat bar and enforces that constraint on the localization. Next, the orientations of all the points are averaged and ensured to be pointing perpendicular to the 3D line. This model information is used because it is known that the seat track is not twisted and all mating holes are on the same side, so all will have the same orientation. The end result is a 6-DOF pose for each screw head that is along a line and oriented in the same direction.

3.3 Docking Procedure

The docking procedure is a set of movements that minimize perception error as well as perform the actual task of grasping the seat track. Minimizing perception error is necessary because when an object is viewed from a point that does not fall along the objects normal vector, the normal vectors will be biased toward the camera. In addition, because of the many distance dependent errors, it makes sense to move the Nano close to the target to minimize these errors. The actual task of grasping is rather straightforward once the gripper nubs have been inserted into the seat track, however significant finesse is required to get a positive lock from the gripper state machine. Lastly, the entire docking procedure was implemented in Robot Task Commander (RTC) (see below). The perception error minimization takes place in the first three movements that the leg takes. The first movement orients the leg so that the x axis of the camera is aligned with the centerline of the seat track. The second movement moves the leg so that the gripper is only 0.25m away from the seat track. This move is repeated, because due to perception errors, the relative motion that is first estimated is never quite right. Next, one seat track hole is selected and the leg is moved so as to align for insertion. This move is also repeated. The first motion ignores the orientation estimation for reasons explained above. Next, the leg attempts to dock; first it selects the same seat track mating hole, and then using the latest 6-DOF estimation it attempts to insert the pegs into the mating holes. The goal in this particular insertion movement is to be close to the seat track, but not entirely in the seat track. The motivation for this stems from the high gains and torque limits that are required for high precision position movements. Once the robot is hovering just inside the seat track, the gains and torque limits are immediately lowered. This allows for the robot to perform a soft dock. The benefit of the soft dock is three fold. First, if the gripper was positioned poorly than the low gains will help to wiggle the pegs into the mating hole. Second, the low gains prevent the robot from pressing too hard against the back of the seat track with excessive amounts of force, causing a shutdown. Finally, the low gains allow the gripper to properly align during closing if the misalignment is small.
3.4 High Level Organization

All of the procedures described previously were organized and connected using the RTC Framework [9]. The entire structure for the seat track docking task is in Figure 12. One of the benefits of using a tool like RTC is that once a low level behavior has been developed, it can be combined with other behaviors in a multitude of ways to create arbitrarily complex behavior. In this case, the richness of RTC was used to encode error checking and automated restart behavior. In addition to providing the necessary connections and organization for robust behavior, it also provides additional state information to the operator who is monitoring robot execution. Lastly, RTC was used to develop the tests that will be discussed in the following section.

4 RESULTS

This pipeline was tested on R2 at Johnson Space Center using the same seat track on which the pipeline was developed. The testing procedure consisted of 5 different initial configurations of the robot leg. The 5 initial conditions consist of the leg being offset in all 4 quadrants from the seat track, and one where it was directly above the seat track. In the offset cases, the end of the leg is oriented towards the seat track so that the seat track is visible to the sensors.

R2 was able to successfully dock with the seat track in all 5 cases. One starting configuration resulted in multiple failed attempts, however due to the high-level error catching it was able to reattempt each time, and successfully docked on third attempt.
5 CONCLUSIONS

This work demonstrated a wide variety of the skills and techniques necessary for performing tasks on R2. In particular, several different sensor error modes were highlighted, as well as calibration routines to handle those errors. In addition, a model based sensor fusion approach demonstrated how relatively poor sensor information, combined with a simple but clear model, provided adequate information to perform autonomous docking. At a higher level, this work demonstrates the ability of R2 to perform automated seat track engagement in a robust way from multiple initial conditions. This indicates that the sensors, computation, and actuation on R2 are adequate to perform the docking task. In the future, it would be of great use to consider obstacles on the path to the seat track. This might require detecting the seat track in spite of occlusions, as well as selecting a docking location so as to avoid any obstacles.

References


