

Autonomous Multilateral Debridement with the Raven Surgical Robot

Ben Kehoe¹, Gregory Kahn², Jeffrey Mahler², Jonathan Kim², Alex Lee², Anna Lee², Keisuke Nakagawa⁴, Sachin Patil², W. Douglas Boyd⁴, Pieter Abbeel², Ken Goldberg³

Abstract—Robotic surgical assistants (RSAs) enable surgeons to perform delicate and precise minimally invasive surgery. Currently these devices are primarily controlled by surgeons in a local tele-operation (master-slave) mode. Introducing autonomy of surgical sub-tasks has the potential to assist surgeons, reduce fatigue, and facilitate supervised autonomy for remote tele-surgery. This paper considers the sub-task of *surgical debridement*: removing dead or damaged tissue fragments to allow the remaining healthy tissue to heal. We present an implemented automated surgical debridement system that uses the Raven, an open-architecture surgical robot with two cable-driven 7 DOF arms. Our system combines stereo vision for 3D perception, trajopt, an optimization-based motion planner, and model predictive control (MPC). Experiments with autonomous sensing, grasping, and removal of over 100 fragments suggest that it is possible for an autonomous surgical robot to achieve robustness comparable to human levels for a surgically-relevant subtask, although for our current implementation, execution time is 2–3× slower than human levels, primarily due to replanning times for MPC. This paper provides three contributions: (i) introducing debridement as a surgically-relevant sub-task for robotics, (ii) designing and implementing an autonomous multilateral surgical debridement system that uses both arms of the Raven surgical robot, and (iii) providing experimental data that highlights the importance of accurate state estimation for future research.

I. INTRODUCTION

Robotic surgical assistants (RSAs), such as Intuitive Surgical’s da Vinci[®] system, have proven highly effective in facilitating precise minimally invasive surgery [9]. Currently these devices are primarily controlled by surgeons in a local tele-operation mode (master-slave with negligible time delays). Introducing autonomy of surgical sub-tasks has potential to assist surgeons, reduce fatigue, and facilitate supervised autonomy for remote tele-surgery.

In this paper, we introduce *surgical debridement* as a relevant subtask for autonomous surgical robotics. Surgical debridement is a tedious surgical sub-task in which dead or damaged tissue is removed from the body to allow the remaining healthy tissue to heal [2], [10]. Autonomous surgical debridement requires perception to locate fragments, grasp and motion planning to determine collision free trajectories

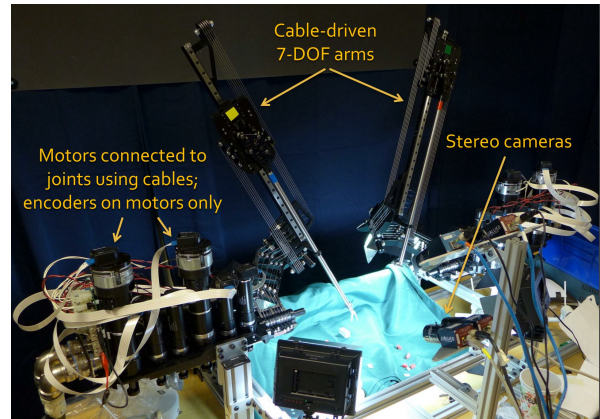


Fig. 1. Hardware for autonomous multilateral debridement. The Raven surgical robot has two cable-driven 7-DOF arms. The system uses a pair of digital cameras for stereo 3D vision.

for one or more arms and grippers to grasp them, and control to deposit them into a receptacle (see Fig. 2(b)). We study autonomous multilateral execution of this subtask.

We use the Raven surgical robot system [12], augmented with a stereo vision system for perception, as a testbed for evaluating the feasibility of autonomously performing surgical debridement with two arms (multilateral operation), as shown in Fig. 1.

Autonomous execution of surgical sub-tasks is challenging. Many surgical robots, such as the da Vinci and the Raven, have 6 DOF per arm (plus a grasp DOF), so there is no joint redundancy. Also, each arm must enter the body through a port that constrains the motion at that point akin to a spherical joint. These constraints, combined with a very limited workspace, make for a particularly challenging robotic manipulation setting. Further, the need for sterilization has resulted in the placement of actuators and encoders outside the body, where they connect inside using cable-driven mechanisms, which have much less precise control than stiff geared actuators.

For perception, the constrained and highly reflective surfaces of the actuators and body organs create difficulties for standard 3D point cloud methods. We use stereo vision, which is increasingly common in RSA systems. The challenges of autonomous computer vision are well known and include noise, calibration, segmentation, and occlusions.

Because of the uncertainty outlined above, replanning is required to prevent robot collisions with obstacles (the other arm, the worksurface, or other objects). Collisions are a familiar problem in robotics, but are exacerbated with surgical robots because a collision with the worksurface can

¹Department of Mechanical Engineering; benk@berkeley.edu

²Department of Electrical Engineering and Computer Sciences; {gkahn, jmahler, jonkim93, alexlee-gk, leeanna, pabbeel}@berkeley.edu

³Department of Industrial Engineering and Operations Research and Department of Electrical Engineering and Computer Sciences; goldberg@berkeley.edu

¹⁻³ University of California, Berkeley; Berkeley, CA 94720, USA

⁴Division of Cardiothoracic Surgery; University of California Davis Medical Center; Sacramento, CA 95817, USA; {keisuke.nakagawa, walter.boyd}@ucdmc.ucdavis.edu

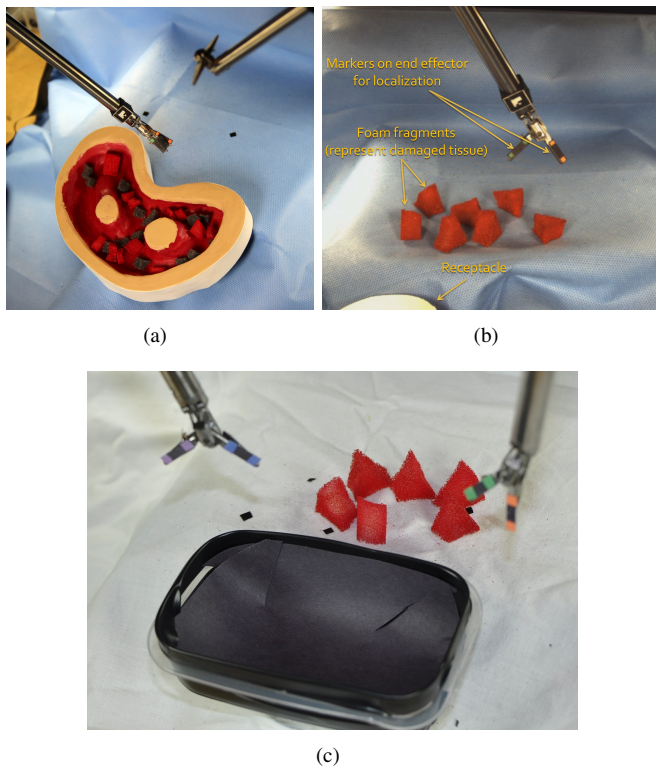


Fig. 2. Varying levels of realism in the surgical debridement task. (a) Two arm surgical debridement with simulated anatomical structures containing multiple foam colors. (b) Single arm surgical debridement. (c) Two arm surgical debridement with plain white background. This is the setup used in the experiment.

snap cables requiring extensive repair time in addition to the cost of collision with patient anatomy.

To the best of our knowledge, this is the first description of a Raven performing a manipulation task autonomously. For context, we report experiments performed by a medical student with training in laparoscopic surgery. The human executes surgical debridement using a 3D display fed by the stereo camera pair and local teleoperation devices. We compare one-arm and two-arm autonomous performance in terms of robustness (failure modes) and timing.

Our experiments indicate that grasping and manipulation tasks associated with the surgical debridement task can be performed reliably with a Raven surgical robot system. We evaluate performance with the human debriding 30 fragments, with one-arm debriding 60 fragments, and with two-arms debriding 60 fragments.

II. RELATED WORK

Existing robotic surgical systems can be categorized into a spectrum based on the modality of interaction with the surgeon [30], [37]. These systems range from pure teleoperated or master/slave systems that directly replicate the motions performed by the surgeon [11], [30], to supervisory or shared-control systems where the surgeon holds and remains in control of the medical instrument and the robot provides assistance [36], to purely autonomous systems where medical motions are planned off-line when detailed quantitative pre-operative plans of the surgical procedure

can be laid out and executed autonomously without intra-operative modification [38]. In addition, intelligent robotic assistants have also been proposed for rendering assistance in minimally invasive surgery [17], [20].

In this work, we focus on autonomous execution of a tedious surgical sub-task known as surgical debridement [2], [10], which involves removing damaged tissue from an affected area to allow the surrounding tissue to heal. We note that prior work has addressed the problem of designing planning and control algorithms for autonomous execution of other surgical sub-tasks such as knot tying or suturing [28], [39] and tissue retraction during surgery [22], [14].

Recent advances in motion planning, control, and perception have enabled robotic systems to perform complex manipulation tasks in real world domains [3], [8], [7], [31], [33]. These systems perform integrated task and motion planning (see e.g., [1], [6], [16], [40]) by using state machines or task graphs [4], [34] for high-level task specification and motion planning algorithms for realization of low-level sub-tasks. Extensions have been proposed to consider uncertainty in task execution [15], [35]. Our work uses a similar architecture for autonomy that integrates a high-level task specification in terms of a state machine [4] with low-level planning. However, instead of open-loop execution of motion plans for accomplishing low-level sub-tasks, we re-plan after every time-step in the spirit of model predictive control [27] to mitigate uncertainty.

There is extensive prior work on calibration of kinematic parameters of robotic manipulators [13]. Extensions have been proposed to simultaneously calibrate robot and sensor (e.g., camera) parameters [24], [41]. These methods cannot account for errors resulting from material non-linearities such as cable stretch, prevalent in cost-effective cable-driven actuation mechanisms. We follow the data-driven approach of Pastor et al. [21] to characterize the systematic error in camera calibration and registration using an optimization-based approach, and residual kinematic errors due to non-linearities using Gaussian process regression [26].

III. SURGICAL DEBRIDEMENT

Laparoscopic surgery requires the execution of many different sub-tasks, including incisions, suturing, clamping, retraction, etc. Not all of these are suited to autonomous operation; for example, cutting into organs requires very high precision and is high risk. Different sub-tasks require different instruments, when some surgical robots only have gripping tools available.

The Fundamentals of Laparoscopic Surgery [29], a skills training program for laparoscopic surgery, provides a pick-and-place task involving the transfer of triangular blocks between vertical pegs (see Fig. 3). However, this task is not well-suited to autonomous operation. The size of the blocks requires very high precision, and the rigid pegs can cause damage to the robot if it impacts them.

A. Task Definition

We propose *surgical debridement* as task for performing experiments with autonomous surgical robots. Surgical

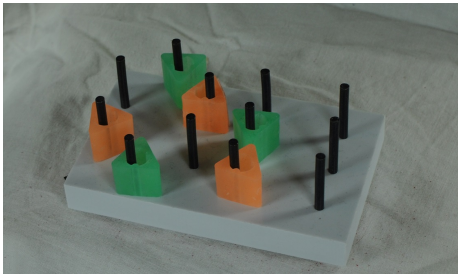


Fig. 3. Fundamentals of Laparoscopic Surgery pick-and-place task, used in a skills training program. The blocks must be grasped and transferred between pegs.

debridement is a tedious surgical sub-task in which dead or damaged tissue is removed from the body to allow the remaining healthy tissue to heal faster [2], [10].

Surgical debridement has several features that make it attractive for autonomous operation. It is a tedious task, which means automation could reduce surgeon fatigue. A simulation of the task can be performed with only a gripper, and it involves detection, grasping, and motion planning components. Importantly, the task or any one of its components can be approximated with varying degrees of realism (see Fig. 2), allowing autonomous surgical robots to start with simple environments and build up to more and more realistic situations.

We consider an idealized environment in which fragments designated as damaged tissue are distributed throughout the environment. The robot must find the damaged tissue fragments, grasp them, and place them in a receptacle. The level of complexity for different aspects of the task can be varied, from providing fragments to be avoided, to situating the fragments in a tissue phantom, even to attaching the fragments to the work surface and requiring a cutting action for removal. Alternatively, any or all of these may be simplified to provide a low barrier to entry for autonomous surgical robotics research. The goal is to facilitate safe (collision-free) multilateral autonomous execution of the surgical debridement task using a surgical robot.

B. Failure Modes

We define several failure modes for this task:

Identification:

- 1) Fragment false negative: no detection of a fragment in the workspace.
- 2) Fragment false positive: detection of a fragment where none exists.
- 3) Pickup false negative: after grasping, no detection of a fragment in the gripper, causing an unnecessary regrasp.
- 4) Pickup false positive: after a pickup failure (see below), detection of a fragment in the gripper.

Grasping:

- 5) Grasp failure: the gripper is closed, but no part of the fragment is within the gripper.
- 6) Multiple grasp: the gripper unintentionally grasps multiple fragments. When targeting a single fragment for

pickup, any other fragments grasped could possibly be healthy tissue, even if they happen not to be.

- 7) Pickup failure: the gripper has closed on some part of the fragment, but the fragment falls out of the gripper on lifting.

Movement:

- 8) Drop en route: after lifting, the fragment falls out during the move to the receptacle.
- 9) Dropoff failure: the fragment is dropped from the gripper upon arrival to the receptacle, but the fragment lands outside the receptacle.

IV. HARDWARE

A. Raven Surgical Robot

We use a Raven surgical robot system (Fig. 1) for evaluating the feasibility of multilateral autonomous execution of the surgical debridement task.

The Raven is an open-architecture surgical robot for laparoscopic surgery research with two cable-driven 7 DOF arms. It is intended to facilitate collaborative research on advances in surgical robotics [12].

The primary difficulty in using the Raven for autonomous operation is state estimation. The joints do not have encoders on them; instead, measurements of the joint angles can only be taken from encoders mounted on the motors (i.e., the cable capstans) but not on the joints. The joints on the arms are connected to motors at the base using up to four meters of small-gauge cable, and the cables are coupled through multiple joints, so the motor measurements do not reflect the slack or stretch in cables that affect the joint angles. As a result, even a small amount of slack or stretch in the cables can greatly increase the uncertainty in gripper pose. State estimation has previously been explored in simulation [25], but not in physical experiments.

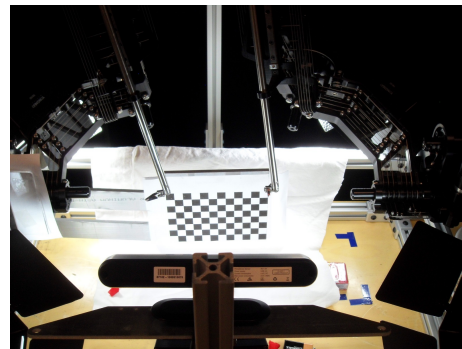


Fig. 4. Removable bracket for rigidly mounting a checkerboard in the workspace for registering the stereo cameras. We use routines from OpenCV for this purpose [5].

B. Vision Hardware

Since the kinematics introduce considerable uncertainty in the calculation of the gripper pose, we use a vision system to obtain direct measurements of the pose. The Raven presents challenges on this front as well. The size of the grippers is too small to use complex fiducial markers like those based on 2D bar codes. We were able to place a fiducial marker on the

wrist link of the robot, but the small size meant the cameras had trouble detecting the marker, and the measurement was highly noisy even when it was detected.

We use a stereo vision system to estimate the pose using colored dots mounted on the gripper (Fig. 2(b)). The stereo vision system is also used to construct a static 3D point cloud from the disparity image, which is used to localize the fragments. Off-the-shelf stereo cameras are usually built for larger workspaces, and thus the camera pair are too widely separated for our environment. We constructed a custom stereo camera using a pair of Prosilica GigE GC1290C cameras with 6 mm focal length lenses at a separation of 4.68 cm for this purpose.

We also experimented with a Primesense Carmine sensor for obtaining point clouds of the environment. However, the Carmine relies on a projected texture, which does not work on specular reflective surfaces like the stainless steel the Raven tool is constructed from. Therefore, the Carmine cannot be used for detecting the gripper.

The cameras must be registered to the robot frame to allow their detections to be used to direct the robot. However, the small size of the workspace prevents the camera field of view from including the robot base. To register the cameras, we fabricated a removable bracket for a checkerboard that could be mounted to the robot base (see Fig. 4), putting the checkerboard in the camera field of view with a known pose relative to the base. This also allows calculation of the transform between bases of the individual arms, which are not precision mounted relative to each other, by using the camera as an intermediate frame.

V. PROBLEM DEFINITION AND METHOD

The surgical debridement task environment considered for this work focuses primarily on the motion planning component of the task. The vision component is simplified through the use of a uniform white background. For the grasping components, the tissue fragments were modeled with small, irregular pieces of foam rubber.

This section covers the vision system for fragment detection and gripper pose estimation in Section V-A, the optimization-based MPC approach in Section V-B, and the multilateral coordination required by the task in Section V-C.

A. Vision System

We use the vision setup outlined in Section IV-B to detect and segment the fragments and for detecting the gripper pose.

1) *Fragment Segmentation*: In order to reliably retrieve the fragments, we must localize the fragments with respect to the robot using the vision system. To simplify the localization, we restricted all fragments to be a specific red hue with a known upper and lower bound of HSV (Hue, Saturation, Value) given the lighting conditions of the workspace. Furthermore, this HSV range was not present elsewhere in the workspace. Given this constraint, localizing the fragments was a three-step process. First, we threshold the image based on HSV values to identify the groups of pixels representing the fragments. Then, we find a reference

point for the fragments by tracing the contours and computing the region centroid. Finally, we use the disparity of the fragment centroid between the left and right images to calculate the position of the fragment centroid in 3D space. To help deal with partially occluded foam pieces, we use an alternative fragment reference point with a constant offset from the lower bound of the fragment in the image.

2) *Pose Estimation*: Reliable, autonomous execution requires precise determination of the gripper pose during execution. Since the forward kinematics produce an unreliable estimate of the gripper pose, we use the vision system to detect the gripper pose.

We estimate the gripper pose using color-based fiducial marks. For each gripper finger, we designate a specific color with a known range of HSV values, given the constrained lighting conditions of the workspace. Each gripper finger has exactly two such marks of the same color, one on the end closest to the joint, and one on the end closest to the tip (see Fig. 2(b)). Using a process similar to the fragment segmentation, we threshold incoming images from the stereo pair for each of the four known HSV values and use the centroids of the regions along with the disparity to find the points in 3D space.

To determine the origin of the gripper pose, we take the average of the position of upper left and upper right fiducial marks on each of the grippers. We calculate the orientation of the gripper pose by finding the vectors along each gripper finger using the fiducial marks. These vectors are coplanar, and the orientation is determined from the component-wise average of the vectors (the axis along the center of the gripper) and the normal (parallel to the gripper joint axis).

The detected pose is assumed to be the true pose. However, these updates occur at best at 10 Hz but are often slower, and may not happen for some time, for example while carrying a fragment, the markers on the gripper may be occluded. To allow for estimated poses in between these updates, we use a pose estimation algorithm using updates from the forward kinematics pose.

We denote the detected gripper pose at time t as $T_{c,t}$, and the pose calculated by the forward kinematics as $T_{FK,t}$.

Given the detected and calculated poses at two times t_0 and t_1 , we first calculate the delta-pose for each, $\Delta T_{c,t_0 \rightarrow t_1} = T_{c,t_0}^{-1} T_{c,t_1}$ and $\Delta T_{FK,t_0 \rightarrow t_1} = T_{FK,t_0}^{-1} T_{FK,t_1}$. As described below, we find two adjustment transforms $T_{A,L}$ and $T_{A,R}$ such that

$$\Delta T_{c,t_0 \rightarrow t_1} = T_{A,L} \Delta T_{FK,t_0 \rightarrow t_1} T_{A,R} \quad (1)$$

Then, given a new calculated pose at time t_2 , without having received a new detected pose, we estimate the true pose by finding $\Delta T_{FK,t_1 \rightarrow t_2} = T_{FK,t_1}^{-1} T_{FK,t_2}$, and applying the adjustment transforms:

$$\hat{T}_{c,t_2} = T_{A,L} \Delta T_{FK,t_1 \rightarrow t_2} T_{A,R} \quad (2)$$

The adjustment matrices in Eq. 1 are iteratively updated with each received detected pose, alternating between updating $T_{A,L}$ and $T_{A,R}$. They are both initialized to identity, $T_{A,L,0} = T_{A,R,0} = I_4$. Given the above update at t_1 , we

keep $T_{A,R,1} = T_{A,R,0}$ and update the left adjustment matrix as follows:

$$T_{A,L,1} = \text{interp}(T_{A,L,0}, \Delta T_{c,t_0 \rightarrow t_1} (\Delta T_{FK,t_0 \rightarrow t_1} T_{A,R,0})^{-1})$$

Then, given a second update of the detected pose at t_3 , we keep $T_{A,L,2} = T_{A,L,1}$ and update the right adjustment matrix:

$$T_{A,R,2} = \text{interp}(T_{A,R,1}, (T_{A,L,1} \Delta T_{c,t_1 \rightarrow t_3})^{-1} \Delta T_{FK,t_1 \rightarrow t_3})$$

where the *interp* function is linear interpolation of the position and spherical linear interpolation for the orientation.

B. Optimization-based Motion Planning with trajopt

Due of the large kinematic uncertainty, an arm may not closely follow the path it is given, which increases the chance of collisions. There are two options for dealing with this situation: (i) allow for this error with a safety margin around the path that the other arm must keep out of, or (ii) use a Model Predictive Control approach and replan frequently using updated pose estimates. Because of the small size of the workspace for the Raven, the first option is not feasible; the size of the safety margin would preclude the other arm from operating anywhere near it. Additionally, this means that both arms must plan together; a path planned for a single arm would have to include this safety margin if the other arm was independently planning its own path. For more details, see Section V-C.

Frequent replanning is also required to maneuver the arm onto the fragment for grasping. In the current system, each arm is permitted to move a maximum of 2.5cm before replanning. With this maximum distance, the safety margin can be set very small. During experiments, the safety margin was set to 1 mm with no collisions occurring.

We use trajopt [32], a low-level motion planning algorithm based on sequential convex optimization to plan locally-optimal, collision-free trajectories simultaneously for both arms. An important feature of trajopt is the ability to check continuous collisions: the arm shafts are very narrow, which could allow them to pass through each other between points on the path.

Additionally, trajopt provides flexible facilities for integrating many different constraints, including collision constraints, pose constraints, and feasibility (e.g., joint limit) constraints. We use all three kinds of constraints. The pose constraint is used to ensure the orientation of the gripper keeps the colored markers towards the cameras so that pose estimation will continue receiving updates.

C. Multilateral Coordination

The surgical debridement task for two arms is very similar to two copies of the single-arm task. However, coordination is required between the two arms in several important places.

There is a single receptacle that is not big enough for both arms to drop fragments into at the same time. Therefore, the arms synchronize their access to the receptacle. The arms must also coordinate in allocating fragments so that they do not attempt to pick up the same fragment.

Besides the two above synchronization points, coordination must occur during path planning. If there was no uncertainty in kinematics, there would be no synchronization required: each arm could plan a path based on its knowledge of the other arm's future position. However, as noted above, the presence of large uncertainty in the Raven kinematics means that an arm planning around the path of another arm would have to use such an unworkably large safety margin. Therefore, the paths for both arms are always planned together.

Because the arms operate independently except for the cases listed above, the two-arm planning is done in a manner transparent to the control code. The control code for each arm proceeds independently in parallel, with a single separate planner also running in parallel. When an arm reaches a replanning point, it submits a request to the planner and waits for a response. Once the planner has received a planning request from both arms, it calculates a two-arm path and returns the individual paths to the respective arms, which then resume independent execution. The same system is used for one-arm operation; in this case, the planner can plan as soon as it receives a request from the arm.

VI. RESULTS AND DISCUSSION

A. Experimental Setup

The experiment was performed with six foam rubber fragments in a random configuration, as shown in Fig. 2(c). The receptacle, located at the front of the workspace, measured approximately 11×7 cm.

For teleoperation, the human operator viewed the workspace through the stereo pair using a 3D monitor, and controlled the Raven using the Razer Hydra controller.

We use the failure modes defined in Section III-B. We experienced occasional static cling in which the fragment would not fall out of an opened gripper; in these cases, failure was indicated if the fragment would have fallen outside the receptacle.

B. Autonomous Performance and Comparison with Teleoperation

The purpose of autonomous surgical robotics is to offload tedious tasks from the surgeon. For comparison, we had the task performed in teleoperation by a third-year medical student with experience on a laparoscopic simulator. To simulate surgical conditions, the teleoperation was performed by viewing the workspace on a 3D monitor using the same cameras used by the autonomous system.

Table I shows the comparison between single-arm autonomous, two-arm autonomous, and teleoperated execution. The autonomous system was executed ten times for each test, and the human operator executed the task five times.

The autonomous system in two-arm operation took on average $2.1 \times$ longer than in teleoperation. However, the amount of time spent in motion for the two arm system was actually slightly less than for the overall teleoperation execution time. This was despite the fact that, in the autonomous operation, the robot moved slowly due to the need

	Human	Autonomous	
	Local Teleoperation	1 arm	2 arm
Total number of fragments	30	60	60
Avg. time per fragment (s)	29.0	91.8	60.3
Avg. time for perception (%)	–	12.1	10.0
Avg. time for planning (%)	–	32.8	36.6
Avg. time for arm movement (%)	–	55.1	45.7
Avg. time waiting on other arm (%)	–	–	7.7
Avg. number of replanning steps	–	11.06	10.58
Fragment false negative (%)	0.0	1.9	0.0
Fragment false positive (%)	0.0	0.0	0.0
Pickup false negative (%)	0.0	0.0	0.0
Pickup false positive (%)	0.0	0.0	3.6
Grasp failure (%)	5.0	3.5	3.6
Grasp multiple fragments (%)	0.0	5.2	7.1
Pickup failure (%)	0.0	0.0	0.0
Drop en route (%)	0.0	0.0	1.8
Dropoff failure (%)	0.0	0.0	0.0

TABLE I

AVERAGE EXECUTION TIME AND OCCURRENCES FOR FAILURE MODES DEFINED IN SECTION III-B. THE NUMBER OF REPLANNING STEPS IS THE NUMBER OF TIMES DURING EXECUTION THAT THE SYSTEM ACCEPTS A NEW INPUT AND OUTPUT STATE AND GENERATES A NEW PLAN. THE TWO-ARM AUTONOMOUS SYSTEM PERFORMED APPROXIMATELY HALF AS FAST AS TELEOPERATION, BUT $1.5\times$ FASTER THAN THE ONE-ARM AUTONOMOUS SYSTEM. THE FAILURE RATES WERE SIMILAR BETWEEN AUTONOMOUS AND TELEOPERATION, WITH ALL BUT ONE FAILURE MODE OCCURRING LESS THAN 5% OF THE TIME. THE MULTIPLE-FRAGMENT GRASP FAILURE MODE OCCURRED DUE TO SEGMENTATION LUMPING CLOSE FRAGMENTS TOGETHER.

to obtain recent updates from the vision system. Although the teleoperator was permitted to use both arms simultaneously, we did not observe him using them in this manner. Each fragment was picked up sequentially. The autonomous system, however, was able to parallelize its arm movements. If the kinematics were reduced and the pose estimation improved, the camera updates could be less frequent and the speed of the robot higher.

The planning and perception together took nearly 50% of the time. The perception code was coded in Python and was not optimized to take advantage of available GPU hardware, which indicates that significant speedups can be made.

The planning time was due in large part to the number of times the system must generate a new plan. Currently, the system must plan an average of 10.81 times during the move to, grasping, and dropoff of a single fragment. This is due to the 2.5 cm maximum distance that an arm is permitted to move before replanning. We found that increasing this distance caused the actual path to deviate too far from the planned path. Improved state estimation would reduce this deviation, allowing for longer distances between replanning.

As noted in Section V-B, the short replanning distance allowed for a very small safety margin to be used, 1 mm. This allowed the two arms to pick up closely-packed fragments more quickly, as the arms could pick up adjacent fragments without penetrating the safety margin.

The two-arm autonomous system was on average $1.5\times$ faster than the one-arm system. This is less than a $2\times$ speedup due in part to waiting time and to increased planning and perception times under the added complexity of two

arms.

Both autonomous and teleoperated systems were able to successfully complete all trials, recovering from grasp and motion failure modes. No false negatives were observed, though the vision system would occasionally lump two close fragments together as a single detection; once one of the fragments was picked up, the other would be correctly detected. The grasp failure rate was higher for teleoperation than for the autonomous system; we believe this is due to the 3D camera not being spaced optimally for human viewing, which led to the human operator reporting a lack of sufficient depth perception.

Further information on this research, including data, photos, and video, is available at:
<http://rll.berkeley.edu/raven>.

VII. CONCLUSION AND FUTURE WORK

We have demonstrated that it is possible for an autonomous surgical robot to achieve robustness comparable to human levels for a surgically-relevant subtask, yet many challenges remain. Execution time is $2\text{--}3\times$ slower than human levels, primarily due to replanning. Our current system essentially assumes the uncertainty grows at a rate such that, after 2.5 cm, the actual pose can be sufficiently far from the planned pose that replanning is required. Two factors could reduce replanning time: 1) faster processing of each replan step, which could be achieved by parallelization or by advances in Moore’s law, and 2) reducing the number of times the robot must replan: in our current system, the average is 10.81 planning calls per fragment. The latter could be achieved with improved state estimation for the following

reason: since the workspace for surgical robots is small and there are a number of obstacles, any decrease in uncertainty can play a crucial role. With more sophisticated state estimation and/or estimation that is state/trajectory/obstacle dependent, the state estimates used in a path would remain valid for longer, which reduces the number of replanning steps, in turn reducing overall execution time.

For surgical robots where space is limited and sterilization is essential, cable-driven actuators are often used and it is not feasible to install joint sensors at the distal ends of the devices. Such indirect control and sensing is inherently imprecise. One approach to improving state estimation is to learn and update systematic and residual error models over time, and to combine these into a dynamic state estimation filter. This can also be combined with a probabilistic state model to replace the worst-case error model.

In future work we will develop more sophisticated state estimation. In addition to reducing the amount of uncertainty (see next section for initial results), we will use tools to manage uncertainty in a more sophisticated way, including probabilistic state estimation using Belief Space [19], [23]. We will perform controlled experiments with humans at different skill levels (including at least one master surgeon). We will also explore increase task complexity, for example using a mix of fragments of different types and introducing other additional obstacles in the workspace, and a hybrid/supervisory system as in a remote tele-surgery scenario, where a human supervisor is in the loop to periodically confirm a set of detections and motion plans prior to execution.

A. Error Characterization

As we have demonstrated, state estimation is one of the main problems in performing autonomous operation with the Raven. In addition to improving pose estimation through machine vision, we also plan to improve the pose estimation through the forward kinematics.

Using the approach from Pastor et al. [21], we characterized the systematic error in the forward kinematics. We moved the gripper around a 3D grid in the workspace using a tele-operation interface and collected $n = 650$ data samples.

Each sample consisted of the gripper pose, denoted by the transformation matrix T_c , as estimated using the stereo vision system, and the gripper pose T_{FK} as computed using the forward kinematics of the robot [18]. The gripper pose T_c is in the camera frame and the gripper pose T_{FK} is in the world frame, and the transformation matrix between the camera frame and the world frame is given by $T_{c \rightarrow w}$. We use standard methods from the OpenCV library [5] to compute this registration of the stereo camera pair to the world frame. To minimize errors in computing $T_{c \rightarrow w}$, we rigidly mount a checkerboard to the robot frame, the dimensions of which are exactly known (Fig. 4). Ideally, the gripper poses estimated using vision and forward kinematics should be identical in the common world frame, i.e., $T_{FK} = T_{c \rightarrow w} T_c$, but this is not the case in practice. Then, we optimized the best transformation that minimizes the systematic pose

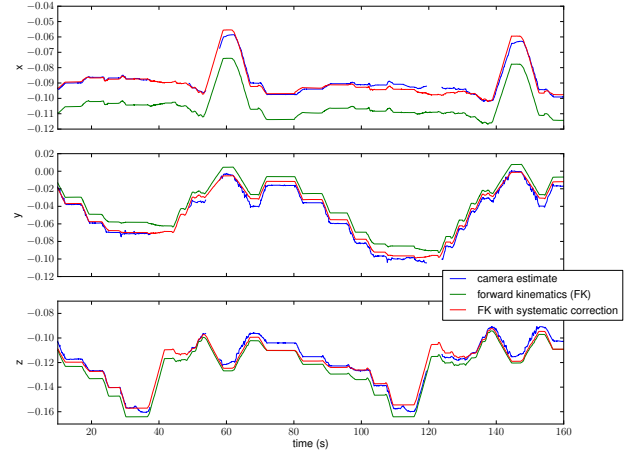


Fig. 5. Position estimates from machine vision, forward kinematics, and forward kinematics with systematic error correction. The position estimate from machine vision is assumed to be the true position (though it is noisy). The RMS error of the position (not per-axis) for the forward kinematics is 1.1 cm, and for the forward kinematics with systematic error correction it is 0.3 cm

error between the forward kinematics model and the pose determined using the stereo vision system in the world frame.

The systematic error is characterized by the rigid transformation T_{sys} , which transforms the gripper pose estimated from the forward kinematics T_{FK} to better match the gripper pose estimated from the vision system in the world frame $T_{c \rightarrow w} T_c$. That is, in the absence of state-dependent errors, $T_{sys} T_{FK} = T_{c \rightarrow w} T_c$ should hold. We compute the best transformation $T_{sys} = \begin{bmatrix} R_{sys} & P_{sys} \\ \mathbf{0}^\top & 1 \end{bmatrix}$ by solving the the following constrained optimization:

$$\min_{T_{sys}} \sum_{i=0}^n \|T_{sys} T_{FK,i} - T_{c \rightarrow w} T_{c,i}\|_F^2, \quad \text{s. t. } R_{sys} R_{sys}^\top = I, \quad (3)$$

where $\|A\|_F$ denotes the Frobenius norm of a matrix, $T_{c,i}$ and $T_{FK,i}$ are the gripper poses obtained from vision and forward kinematics, respectively, for the i^{th} data sample, and $R_{sys} R_{sys}^\top = I$, or equivalently, $\text{vec}[R_{sys} R_{sys}^\top - I] = \mathbf{0}$, enforces the orthonormality constraint. We initialize the optimization with the identity transformation I .

The results are shown in Fig. 5. The detected pose, though noisy, is assumed to be the actual pose. The RMS position error for the forward kinematics is 1.2 cm, and for the forward kinematics with systematic error correction the RMS error is 0.3 cm. Though not shown on the plot, the RMS error for the pose estimator outlined in Section V-A.2 is 0.4 cm. The results suggest that even after correcting for systematic error such as camera registration, there remains a smaller but significant residual error that may be state dependent or non-affine. In future work, we will apply the the method of Pastor et al. to correct for residual errors using Gaussian Process Regression to further improve state estimation.

VIII. ACKNOWLEDGMENTS

We thank our many collaborators on this project, in particular PI Allison Okamura and co-PIs Greg Hager, Blake Hannaford, and Jacob Rosen. This work is supported in part by a seed grant from the UC Berkeley Center for Information Technology in the Interest of Science (CITRIS) and by the U.S. National Science Foundation under Award IIS-1227536: Multilateral Manipulation by Human-Robot Collaborative Systems.

REFERENCES

- [1] R. Alami, R. Chatila, S. Fleury, M. Ghallab, and F. Ingrand, "An architecture for autonomy," *Int. Journal of Robotics Research*, vol. 17, no. 4, pp. 315–337, 1998.
- [2] C. E. Attinger, E. Bulan, and P. A. Blume, "Surgical debridement: The key to successful wound healing and reconstruction," *Clinics in podiatric medicine and surgery*, vol. 17, no. 4, p. 599, 2000.
- [3] J. A. Bagnell, F. Cavalcanti, L. Cui, T. Galluzzo, M. Hebert, M. Kazemi, M. Klingensmith, J. Libby, T. Y. Liu, N. Pollard, *et al.*, "An integrated system for autonomous robotics manipulation," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2012, pp. 2955–2962.
- [4] J. Bohren and S. Cousins, "The SMACH high-level executive [ROS news]," *IEEE Robotics & Automation Magazine*, vol. 17, no. 4, pp. 18–20, 2010.
- [5] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly, 2008.
- [6] S. Cambon, R. Alami, and F. Gravot, "A hybrid approach to intricate motion, manipulation and task planning," *Int. Journal of Robotics Research*, vol. 28, no. 1, pp. 104–126, 2009.
- [7] S. Chitta, E. G. Jones, M. Ciocarlie, and K. Hsiao, "Perception, planning, and execution for mobile manipulation in unstructured environments," *IEEE Robotics and Automation Magazine*, vol. 19, no. 2, pp. 58–71, 2012.
- [8] A. Cowley, B. Cohen, W. Marshall, C. Taylor, and M. Likhachev, "Perception and motion planning for pick-and-place of dynamic objects," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS) (to appear)*, 2013.
- [9] S. A. Darzi and Y. Munz, "The impact of minimally invasive surgical techniques," in *Annu Rev Med.*, vol. 55, 2004, pp. 223–237.
- [10] M. Granick, J. Boykin, R. Gamelli, G. Schultz, and M. Tenenhaus, "Toward a common language: Surgical wound bed preparation and debridement," *Wound repair and regeneration*, vol. 14, no. s1, pp. 1–10, 2006.
- [11] G. Guthart and J. Salisbury Jr, "The Intuitive telesurgery system: Overview and application," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, vol. 1, 2000, pp. 618–621.
- [12] B. Hannaford, J. Rosen, D. C. Friedman, H. King, P. Roan, L. Cheng, D. Glozman, J. Ma, S. Kosari, and L. White, "Raven-II: AN open platform for surgical robotics research," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 954–959, Apr. 2013.
- [13] J. Hollerbach, W. Khalil, and M. Gautier, "Model identification," in *Springer Handbook of Robotics*. Springer, 2008, ch. 14, pp. 321–344.
- [14] R. Jansen, K. Hauser, N. Chentanez, F. van der Stappen, and K. Goldberg, "Surgical retraction of non-uniform deformable layers of tissue: 2d robot grasping and path planning," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2009, pp. 4092–4097.
- [15] L. P. Kaelbling and T. Lozano-Pérez, "Integrated task and motion planning in belief space," *Int. Journal of Robotics Research*, 2013.
- [16] —, "Hierarchical task and motion planning in the now," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2011, pp. 1470–1477.
- [17] H. Kang and J. T. Wen, "Robotic assistants aid surgeons during minimally invasive procedures," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 1, pp. 94–104, 2001.
- [18] H. King, S. Kosari, B. Hannaford, and J. Ma, "Kinematic analysis of the Raven-II research surgical robot platform," University of Washington, Tech. Rep. UWEEETR-2013-0006, 2012, available at: <https://www.ee.washington.edu/techsite/papers/refer/UWEEETR-2012-0006.html>.
- [19] A. Lee, S. Patil, J. Schulman, Z. McCarthy, J. van den Berg, K. Goldberg, and P. Abbeel, "Gaussian belief space planning for imprecise articulated robots," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS) (to appear)*, 2013.
- [20] V. Munoz, C. Vara-Thorbeck, J. DeGabriel, J. Lozano, E. Sanchez-Badajoz, A. Garcia-Cerezo, R. Toscano, and A. Jimenez-Garrido, "A medical robotic assistant for minimally invasive surgery," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, vol. 3, 2000, pp. 2901–2906.
- [21] P. Pastor, M. Kalakrishnan, J. Binney, J. Kelly, L. Righetti, G. Sukhatme, and S. Schaal, "Learning task error models for manipulation," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2013.
- [22] S. Patil and R. Alterovitz, "Toward automated tissue retraction in robot-assisted surgery," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2010, pp. 2088–2094.
- [23] S. Patil, Y. Duan, J. Schulman, K. Goldberg, and P. Abbeel, "Gaussian belief space planning with discontinuities in sensing domains," in *Int. Symp. on Robotics Research (ISRR) (in review)*, 2013.
- [24] V. Pradeep, K. Konolige, and E. Berger, "Calibrating a multi-arm multi-sensor robot: A bundle adjustment approach," in *Int. Symp. on Experimental Robotics (ISER)*, 2010.
- [25] S. Ramadurai, S. Kosari, H. H. King, H. Chizeck, and B. Hannaford, "Application of unscented kalman filter to a cable driven surgical robot: A simulation study," in *2012 IEEE International Conference on Robotics and Automation, St. Paul-Minneapolis*, May 2012.
- [26] C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [27] J. Rawlings, "Tutorial overview of Model Predictive Control," *IEEE Control Systems Magazine*, vol. 20, no. 3, pp. 38–52, 2000.
- [28] C. E. Reiley, E. Plaku, and G. D. Hager, "Motion generation of robotic surgical tasks: Learning from expert demonstrations," in *Int. Conf. on Engg. in Medicine and Biology Society (EMBC)*, 2010, pp. 967–970.
- [29] E. M. Ritter and D. J. Scott, "Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic surgery," *Surgical Innovation*, vol. 14, no. 2, pp. 107–112, 2007. [Online]. Available: <http://sri.sagepub.com/content/14/2/107.abstract>
- [30] J. Rosen, B. Hannaford, and R. M. Satava, *Surgical robotics: Systems, applications, and visions*. Springer, 2011.
- [31] R. B. Rusu, I. A. Sucan, B. Gerkey, S. Chitta, M. Beetz, and L. E. Kavraki, "Real-time perception-guided motion planning for a personal robot," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2009, pp. 4245–4252.
- [32] J. Schulman, J. Ho, A. Lee, H. Bradlow, I. Awwal, and P. Abbeel, "Finding locally optimal, collision-free trajectories with sequential convex optimization," in *Robotics: Science and Systems (RSS)*, 2013.
- [33] S. S. Srinivasa, D. Ferguson, C. J. Helfrich, D. Berenson, A. Collet, R. Diankov, G. Gallagher, G. Hollinger, J. Kuffner, and M. V. Weghe, "HERB: A home exploring robotic butler," *Autonomous Robots*, vol. 28, no. 1, pp. 5–20, 2010.
- [34] I. A. Sucan and L. E. Kavraki, "Mobile manipulation: Encoding motion planning options using task motion multigraphs," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2011, pp. 5492–5498.
- [35] —, "Accounting for uncertainty in simultaneous task and motion planning using task motion multigraphs," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2012, pp. 4822–4828.
- [36] R. Taylor, P. Jensen, L. Whitcomb, A. Barnes, R. Kumar, D. Stoianovici, P. Gupta, Z. Wang, E. Dejuan, and L. Kavoussi, "A steady-hand robotic system for microsurgical augmentation," *Int. Journal of Robotics Research*, vol. 18, no. 12, pp. 1201–1210, 1999.
- [37] R. Taylor, A. Menciassi, G. Fichtinger, and P. Dario, "Medical robotics and computer-integrated surgery," in *Springer Handbook of Robotics*. Springer, 2008, pp. 1199–1222.
- [38] R. H. Taylor, B. D. Mittelstadt, H. A. Paul, W. Hanson, P. Kazanzides, J. F. Zuhars, B. Williamson, B. L. Muisits, E. Glassman, and W. L. Bargar, "An image-directed robotic system for precise orthopaedic surgery," *IEEE Trans. on Robotics and Automation*, vol. 10, no. 3, pp. 261–275, 1994.
- [39] J. Van Den Berg, S. Miller, D. Duckworth, H. Hu, A. Wan, X.-Y. Fu, K. Goldberg, and P. Abbeel, "Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2010, pp. 2074–2081.
- [40] J. Wolfe, B. Marthi, and S. J. Russell, "Combined task and motion planning for mobile manipulation," in *Int. Conf. on Automated Planning and Scheduling (ICAPS)*, 2010, pp. 254–258.
- [41] H. Zhuang, K. Wang, and Z. S. Roth, "Simultaneous calibration of a robot and a hand-mounted camera," *IEEE Trans. on Robotics and Automation*, vol. 11, no. 5, pp. 649–660, 1995.