Toward Practical and Accurate Touch-Based Image Guidance for Robotic Partial Nephrectomy

James M. Ferguson*, E. Bryn Pitt*, Andria A. Remirez, Michael A. Siebold, Alan Kuntz, Nicholas L. Kavoussi, Eric J. Barth, S. Duke Herrell III, and Robert J. Webster III

Abstract—Partial nephrectomy involves removing a tumor while sparing surrounding healthy kidney tissue. Compared to total kidney removal, partial nephrectomy improves outcomes for patients but is underutilized because it is challenging to accomplish minimally invasively, requiring accurate spatial awareness of unseen subsurface anatomy. Image guidance can enhance spatial awareness by displaying a 3D model of anatomical relationships derived from medical imaging information. It has been qualitatively suggested that the da Vinci robot is well suited to facilitate image guidance through touch-based registration. In this paper we validate and advance this concept toward real-world use in several important ways. First, we contribute the first quantitative accuracy evaluation of touch-based registration with the da Vinci. Next, we demonstrate real-time touch-based registration and display of medical images for the first time. Lastly, we perform the first experiments validating use of touch-based image guidance to improve a surgeon’s ability to localize subsurface anatomical features in a geometrically realistic phantom.

Index Terms—Robot-assisted surgery, image guidance, robot calibration, image registration, kidney surgery.

I. INTRODUCTION

TREATMENT of renal cell carcinoma typically requires surgically removing the tumor and surrounding kidney tissue. Some cases require radical nephrectomy, in which the entire kidney is removed. However, for patients with localized tumors, the American Urological Association and the European Association of Urology recommend nephron-sparing partial nephrectomy, in which only part of the kidney is removed [1], [2]. Compared to radical nephrectomy, partial nephrectomy leads to improved long-term patient outcomes by allowing the patient to retain some kidney function and reducing the risk of chronic kidney disease [3], [4]. Partial nephrectomy remains underutilized, however, likely due to the extreme technical challenges associated with the procedure, especially when performed minimally invasively [5], [6].

Robot-assisted partial nephrectomy (RAPN) performed using the da Vinci Surgical System (Intuitive Surgical, Inc., Sunnyvale, CA, USA) can help mitigate many challenges of minimally invasive partial nephrectomy [7], but RAPN does not inherently address the challenge of relying primarily on direct visualization via an endoscopic camera for surgical navigation. This results in a limited field of view that inhibits surgeons’ ability to intuit the anatomical context of the surgical environment, i.e., the location of surgical tools relative to critical, frequently subsurface, anatomical features. Locating these anatomical features, such as large blood vessels and the
tumor itself, is critical to safely and successfully performing RAPN. Surgical image guidance can help surgeons locate these features, providing additional anatomical context by accurately registering 3D anatomical volumes (typically generated by segmentation of preoperative computed tomography (CT) or magnetic resonance (MR) images) to the surgical environment and displaying this information to the surgeon during the procedure (see Fig. 1). Accurate image guidance has the potential to improve patient outcomes by making localization, dissection, and isolation of critical vascular and organ structures, as well as correct margin selection, easier for surgeons.

It has been suggested that the da Vinci robot’s kinematics could be used to achieve accurate registration for image guidance [8], [9]. In this work, we create such a system, quantify its performance, and demonstrate its ability to improve an experienced surgeon’s performance. Our image guidance system uses the instruments of the da Vinci robot itself as 3D localizers for digitizing anatomical surfaces. By lightly tracing an instrument tip over the surface of the target anatomy while recording the robot’s joint values, our system generates a set of points on the surface of the anatomy. Our system computes a surface-based registration between the preoperative images and the patient’s anatomy during the surgery. Using this registration, we then display a 3D model of the patient’s anatomy segmented from the preoperative imaging to the surgeon in the da Vinci’s surgeon console, enabling the surgeon to visualize the location of subsurface anatomy that is not visible via the endoscope (see Fig. 1). By using the inherent capabilities of the da Vinci for registration, our system provides an image guidance approach that is well suited to the clinical workflow.

This paper presents our touch-based image guidance system and analyzes its accuracy. We also describe practical steps for deploying it in the operating room using a clinical da Vinci Si system. We present a series of phantom experiments to provide a thorough accuracy analysis of a touch-based registration for image guidance. Finally, we present a phantom experiment demonstrating the utility of our image guidance system for improving an experienced surgeon’s ability to localize subsurface anatomical features important in partial nephrectomy. By providing practical and accurate image guidance, our method has the potential to improve surgeons’ ability to accurately accomplish partial nephrectomy. Success in achieving this has the potential to increase utilization of partial nephrectomy, thereby providing enhanced health outcomes to many more patients.

II. RELATED WORK

Image guidance has previously been recognized as potentially useful in facilitating partial nephrectomy, and numerous research groups have sought to implement such image guidance systems. One approach to image guidance in laparoscopic partial nephrectomy involved inserting fiducial markers on barbed needles directly into the kidney [10], [11]. The kidneys and fiducials were then imaged and segmented intraoperatively to enable registration by direct point-to-point correspondence between the fiducials in the segmented images and those same fiducials in the endoscopic video. While providing highly accurate, real-time guidance, these fiducial-based methods increase the risk and complexity of surgery by requiring foreign objects to be manually inserted into the kidney by the surgeon. Furthermore, the need for intraoperative imaging and segmentation represents a time-intensive interruption of the surgical workflow. Indeed, the robotic system is not compatible with intraoperative CT, and thus one would have to fully remove the robot to register the image set.

A less invasive approach to registration is fiducial-free manual registration. In manual registration, the surgeon is tasked with visually aligning 3D images or models to the surgical field. In [12] and [13], preoperative MR and CT images and 3D anatomical models were displayed alongside endoscopic video in the da Vinci’s surgeon console, and surgeons could manually adjust the orientation of the images to match the endoscopic view. Ukimura et. al. [14] and Nakamura et. al. [15] presented augmented reality systems in which surgeons manually aligned translucent 3D anatomical models directly overlaying the image feed from a laparoscopic endoscope. These studies found that surgeons benefited from having preoperative imaging information more readily available with respect to the live camera images. However, this approach increases cognitive burden on surgeons and provides no accuracy guarantees. Indeed, relying on human hand-eye coordination and spatial reasoning to perform registration makes accuracy highly dependent on an individual user’s skill, resulting in low registration precision, as evidenced by large variations in registration accuracy from trial-to-trial in these studies.

To enhance precision and facilitate objective accuracy, others have sought to employ stereo endoscopes for instrument tracking and registration to patient anatomy. Su et. al. [16] proposed a multi-step CT-to-endoscope registration method where the segmented kidney surface was first manually aligned with the stereoscopic video. Surface-based video tracking techniques were then used to refine and stabilize the registration during system operation. Pratt et. al. [17] utilized an augmented endoscope overlay by first identifying a matching feature in both of the stereo images and the preoperative scans to align the translational degrees of freedom and then using a rolling-ball interface to manually align the rotational degrees of freedom. We direct the reader to [18] for a thorough overview of research aimed at using computer vision algorithms to automatically detect and track the da Vinci instruments in the stereo endoscope video. These vision-based approaches are limited by a requirement for persistent, direct line of sight between the endoscope camera and either the anatomical surface or the surgical instruments. During surgery, line of sight is often obstructed by blood, smoke, and other surgical tools. Furthermore, endoscope-based methods typically require accurate tracking of the endoscope position itself to localize tracked objects in the surgical field. Accurate endoscope tracking usually requires an external tracking system and a calibration process to determine the rigid transformation from the tracked frame to the camera frame, such as the method presented in [19]. Some researchers have sought to augment camera-based tracking methods by combining them with either geometric or kinematic information to improve
accuracy [20]–[22]. These results show promise and warrant further investigation, but have yet to fully address the above limitations of endoscope-based methods.

The kinematic information inherently available in the da Vinci surgical system represents a means of 3D localization that relies neither on intraoperative use of external trackers nor on processing endoscopic video. Previous research found that the da Vinci’s active joints (which control motion of the laparoscopic instruments during operation) can be localized with sufficient accuracy for image guidance; however, the accuracy of the passive setup joints (used for gross manipulator positioning) was not suitable [23]–[25]. Kwartowitz et al. [25] proposed to address this shortcoming by using a “hybrid” tracking scheme that combines two tracking modalities (specifically kinematic tracking and optical tracking) to more accurately track the multiple manipulators of the da Vinci in a common coordinate system. In this hybrid tracking scheme, the base frames of the active kinematic chains are registered to external, optically tracked frames attached to the base of the da Vinci. Thus, all base frames can be localized within the coordinate system of the optical tracker and the manipulator tips can then be kinematically tracked relative to their respective base frames. Fiducial localization experiments in [26] later validated the accuracy of hybrid tracking with the da Vinci for image guidance applications. In this paper, we implement this hybrid tracking approach as part of a new calibration method that also simultaneously estimates kinematic parameters of the da Vinci system.

Kinematic tracking of the da Vinci instruments has also shown particular promise in combination with “drop-in” ultrasound probes. In [27], registering the image frame of the ultrasound to the kinematic frames of the robot enabled the ultrasound plane to be displayed in the live endoscopic video. Researchers also combined automatic detection of the robot instruments in ultrasound images [28] with kinematic tracking to produce semi-autonomous ultrasound guidance that tracked instrument motions [29]. Later work in [30] demonstrated registration between kinematically tracked ultrasound and preoperative CT images for application to partial nephrectomy. Researchers also combined automatic detection of the robot instruments in ultrasound images [28] with kinematic tracking to produce semi-autonomous ultrasound guidance that tracked instrument motions [29]. Later work in [30] demonstrated registration between kinematically tracked ultrasound and preoperative CT images for application to partial nephrectomy.

The feasibility of this ultrasound-based registration technique in the context of the operating room is, however, inherently coupled to the accuracy of intraoperative segmentation of the ultrasound images, and represents a skill- and time-intensive addition to the surgical workflow. As an alternative, the touch-based method examined in this work uses the da Vinci’s kinematically tracked instruments directly to digitize anatomical surfaces to enable registration.

The idea of a touch-based registration for image guidance with the da Vinci system was first introduced by Ong et al. [8]. During a partial nephrectomy case, the instrument tool tip was lightly traced over the kidney surface while recording the robotic joint values; the data was processed postoperatively to generate a sparse set of surface points that were used for a standard surface-based registration. The concept showed qualitative merit; however, the authors noted they were unable to perform quantitative analysis of the touch-based method due to the unavailability of a ground truth comparison during the human trial. Building upon this concept, we have further assessed surface-based registration with the da Vinci using rigid phantoms [31]; however, thorough analysis of registration error for this touch-based method using anatomically accurate phantoms has thus far remained unstudied. In this paper we take essential steps toward practical and accurate deployment of this touch-based registration concept by presenting a system that is suitable for deployment in a real-world operating room and accomplishes registration in real time. We also rigorously evaluate the accuracy of touch-based registration on anatomically accurate soft-tissue phantom models, and demonstrate its ability to improve the localization accuracy of an experienced surgeon.

III. SYSTEM OVERVIEW

A. Preoperative System Setup and Calibration

Preoperative calibration of the da Vinci Si system is necessary to achieve sufficient kinematic tracking accuracy to enable our touch-based registration. Figure 2 shows a clinical da Vinci Si deployed for preoperative calibration, which takes place as the da Vinci system is draped prior to surgery. Additively manufactured reference frames designed to interface with the da Vinci system are clamped rigidly to the distal ends of the setup arms (Fig. 2, upper right). These reference frames are equipped with reflective optical tracking markers. To maintain the sterile field, the reference frames are first clamped without reflective spheres before deploying the sterile drapes. After draping the robot, sterile, disposable, commercially available spheres are attached to mounting posts through the sterile plastic drapes. This process ensures a sterile barrier between the clamping system and the sterile surgical environment.

As shown in Fig. 2 (lower right), the da Vinci instruments grasp sterile, optically tracked calibration tools. Each calibration tool is previously pivot-calibrated so that the position of the interface with the instrument tip is accurately known relative to the optical markers. This enables measurement of the instrument position relative to the optically tracked reference frames at the base of each serial chain. Our system uses a Polaris Spectra (Northern Digital Inc., Waterloo, Ontario, Canada) optical tracking system, which is currently available for use in many operating rooms, including any operating room at our institution. The Polaris system has a reported tracking accuracy of 0.25 mm, and for this work, we consider measurements with the optical tracker to be ground truth [32].

To collect calibration data prior to surgery, a robot operator simply moves the calibration tools throughout the robot workspace while recording data. In our current implementation, the operator momentarily pauses at discrete locations to ensure synchronization of optical tracking and robot encoder data streams. This is necessary only because we did not have direct access into the robot software to enable synchronization between the data streams of the optical tracker and robot. In future clinical implementations, the markers can simply be waved in front of the optical tracking system to collect calibration data. This data collection process is repeated for each da Vinci manipulator and each instrument to be used during surgery.

Our system employs a hybrid tracking technique [25] for calibration that enables the da Vinci’s manipulators to be
kinematically tracked relative to external, optically tracked reference frames (rather than the internal base frames of the da Vinci system). Using hybrid tracking bypasses the comparatively inaccurate setup joints of the da Vinci system, shortening the effective kinematic chain and improving tracking accuracy beyond the inherent capabilities of the da Vinci system. Our calibration process also simultaneously calibrates the parameters of the kinematic model using standard techniques [33]–[35]. The result of calibration is that each robotic instrument can be accurately tracked with respect to the location of the reference frames attached at the base of the active serial chain.

B. Touch-Based Registration

Our touch-based registration method aligns two sets of data: a densely sampled point set describing the organ surface in image space and a sparsely sampled point set of surface data describing the organ surface in physical space. The dense image space point set is obtained preoperatively from volumetric imaging. For the experiments presented in this paper, the kidney surface was manually segmented from CT images using 3D Slicer, an open-source medical image computing and visualization software platform [36]. In the future, however, when an image guidance system like ours is developed into a commercial product, it is likely that manual segmentation would be replaced by an automatic segmentation algorithm. Any existing or future segmentation algorithm would be straightforward to incorporate into the framework described in this paper, since our system assumes only the existence of a segmentation without regard for how the segmentation was accomplished.

The physical space point set is obtained intraoperatively by lightly tracing the surface of the patient’s organ with the tip of the da Vinci’s instrument. We track the instrument’s tip position in physical space during this process using the previously calibrated kinematics. Surface tracing is quick and non-disruptive to surgical workflow: acquiring a sufficient number of surface points for accurate registration requires only about 30 seconds. After tracing, the data are automatically downsampled to exclude data points within 2 mm of neighboring points to eliminate variations in point cloud density caused by variable tracing speed. This results in a set of points in physical space that lie on the surface of the patient’s kidney.

Previous work concluded that the physical-space data used for surface-based registration should include at least 28% of the anterior surface area of the kidney to ensure accurate registration [37]. Therefore, once surface tracing is complete, our system automatically analyzes the tracing data to verify that the tracing covers a sufficient area. Our system determines the surface area corresponding to a tracing by constructing a surface mesh from the tracing data using the ball-pivoting surface reconstruction algorithm [38] (illustrated in Fig. 3). The area of the reconstructed surface is compared to the kidney’s surface area, which is determined from the preoperative CT images.

Registration between the image space point set with the physical space point set is computed using the globally optimal iterative closest point (GoICP) algorithm [39]. GoICP does not require user initialization and as such is not subject to suboptimal initialization concerns associated with standard ICP algorithms. The resulting registration between the image space and the physical space relates knowledge of the patient’s anatomy present in the preoperative images to the current position of the robot with respect to the patient.

C. Real-Time Data Streaming and Visualization

Using the computed registration, we display the position of anatomical structures segmented from preoperative imaging to the surgeon in real time directly in the da Vinci surgeon console (see Fig. 1).

We built our image guidance system as a submodule of 3D Slicer, an open-source medical image computing and visualization platform that enables patient image segmentation, preoperative planning, and real-time model rendering for image guidance [36]. Our system interfaces with the clinical da Vinci Si application programming interface (API) [40] through a data acquisition module built with the open-source Plus Toolkit [41] that streams kinematic data output by the API to 3D Slicer using the standardized OpenGLESLink messaging protocol [42]. This enables the endoscope camera view and graphical models of the da Vinci’s manipulators in the image guidance display to track the movement of the physical instruments in real time. Our image guidance (see Fig. 1) is displayed directly in the da Vinci surgeon console through the console’s TilePro interface.

IV. SYSTEM VALIDATION EXPERIMENTS

We first evaluate the efficacy of our calibration method to improve the overall kinematic accuracy of the da Vinci robot. We then evaluate the accuracy of our touch-based registration method.

A. Calibration Accuracy

We wish to determine the number of measurements that must be collected during preoperative setup to ensure good calibration results. In the context of robot calibration, this number is generally difficult to predict, as it varies from system to system and also depends on the measurement method used [43]. We performed a series of trials to determine the relationship between tracking accuracy and the number of measurements used in calibration, as described below.

For our touch-based application, the da Vinci instrument tip serves as the localizer. Given that surface-based registration relies only on discrete points of position data, only the positional (not rotational) accuracy of the localizer needs to be considered. Therefore, the accuracy of our system can be quantified by the fiducial localization error (FLE) of the da Vinci instruments, i.e., the distance between the model-predicted tip location and the true tip location:

$$\text{FLE} = \| \mathbf{p}_{\text{true}} - \mathbf{p}_{\text{model}} \|. \quad (1)$$

In practice, the FLE cannot be directly measured because our model-predicted position is measured in a different coordinate frame from our “true” position (as measured by the
Fig. 2. Hybrid tracking implemented with the da Vinci Si in the operating room. Optically tracked markers (top right) are rigidly clamped to the base of each manipulator, and sterile tracking spheres are attached to the markers over the robot drapes ensuring sterility. Calibration is achieved by gripping sterile calibration objects (bottom right) in the manipulators (or pressing them onto the endoscope) and waving them in front of the tracker preoperatively.

optical tracker). However, it is possible to indirectly estimate the expected value of the FLE from these data, as described below. The transformation between the two coordinate frames (the robot and the optical tracker) can be estimated from a standard, rigid, point-based registration between the model-predicted positions and the true measured positions [44]. The error associated with such a registration can be quantified by the fiducial registration error (FRE), which is the root-mean-square error in the alignment of the registered points. Performing numerous registrations using different sets of point samples provides a good estimate for the expected value of the FRE for registrations between the two frames. The expected value of the FRE can be used to estimate the expected value of the FLE, according to the following relationship derived in [44]:

$$\langle \text{FLE}^2 \rangle = \frac{\langle \text{FRE}^2 \rangle}{(1 - 2/N)}, \quad (2)$$

where $N$ is the number of points used in the registration and the angle bracket operator denotes the expected value of a random variable. This formulation relies on standard assumptions that the components of FRE are independent, isotropic, 3D normal random variables.

Our evaluation data set comprised 130 calibration measurements, collected at distinct poses representing a sparse sampling of the entire da Vinci Si active workspace. Each calibration measurement consists of a set of robot joint values and a corresponding Cartesian position, measured in the optical tracker’s workspace. All data was collected using the da Vinci’s EndoWrist Large Needle Driver instrument.

To determine the relationship between the model-predicted position accuracy and the number of data points used in model calibration, we performed a Monte Carlo cross-validation analysis of the evaluation data set. Each iteration of the cross-validation was performed as follows:

- A number $M \in \{10, 15, 20, \ldots, 95, 100\}$ of “training points” were selected uniformly at random from the complete set of 130 points.
- The kinematic model was calibrated using the training points.
- $K = 30$ “validation points” were selected uniformly at random from the remaining $130 - M$ points.
- A number $N \in \{5, 6, 7, \ldots, 19, 20\}$ of “registration points” were selected uniformly at random from the validation points. This process was repeated 1000 times for each value of $N$, resulting in a total of 16,000 distinct sets of registration points per iteration.
- Using each set of registration points, a rigid point-based registration between the (calibrated) model-predicted positions and the measured “true” positions was performed.
- The mean value of FLE for the calibrated system was computed from the average FRE of each registration according to Eq. (2).

This process was repeated a total of 1000 times for each
Fig. 4. Fiducial localization error (FLE) of the da Vinci Si vs. the number of measurements used for calibration of the hybrid tracking model. The red area indicates the standard deviation for each respective trial. Only marginal improvements to accuracy can be seen past \( M = 60 \).

Fig. 5. Distribution of RMS errors between the model-predicted robot tip position using hybrid tracking and the ground truth, optically tracked tip position. A significant decrease in error is seen when using our calibration method (red) over using the nominal robot parameters (blue). Results are for 1000 calibration trials with \( M = 60 \) measurements per trial.

value of \( M \). Figure 4 shows the results of this analysis, which indicates that using more than 60 data points to compute the hybrid tracking calibration offers only marginal improvements to localization accuracy.

Figure 5 shows the accuracy improvement of the calibrated da Vinci model compared to the nominal model from [25]. The mean and standard deviation of the calibrated system’s FLE are 0.95 mm and 0.14 mm, respectively, representing a 34% reduction in mean localization error. While no well-defined localization accuracy threshold exists for image guidance applications, it is clear that increased accuracy is always desired. Our calibrated system accuracy is comparable to prior methods used to track the absolute position of the da Vinci’s instrument tips (1.31 mm in [24] and 1.39/1.95 mm for PSM1/PSM2 in [25]).

We wish to emphasize that the accuracy values reported here reflect only the deviation between the model-predicted position of a robot manipulator and the measured position (i.e. where the robot “thinks” the manipulator is versus where it truly is). These accuracy results do not describe the accuracy with which a surgeon can direct the da Vinci manipulators during teleoperation (i.e. where the surgeon wants the manipulator to be versus where it truly is). While touch-based image guidance relies on a highly accurate model-predicted position, teleoperation with visual feedback and a human in the loop does not.

B. Registration Accuracy

We evaluated the accuracy of our touch-based registration method in a series of experiments using a commercially available synthetic kidney model (SynDaver Labs, Tampa, FL, USA) that accurately reflects the geometry and mechanical soft-tissue properties of a human kidney. For our experiments, the model was fixed to an optically tracked platform, as shown in Fig. 6.

Prior to experiments, the entire platform was CT scanned using an xCAT ENT Scanner (Xoran Technologies LLC, Ann Arbor, MI, USA) using a section thickness of 0.3 mm. The kidney surface as well as the optical tracking markers were manually segmented from the CT images using 3D Slicer [36]. The kidney surface segmentation was used to produce the required image space point set for registration. The segmentation of the optical tracking markers was used to compute the ground truth pose of the kidney model relative to the makers, which were optically tracked in the operating room.

To acquire data in physical space, an experienced urologic surgeon thoroughly traced the entire anterior surface of the kidney phantom using a Large Needle Driver instrument in a calibrated da Vinci Si while our system recorded tool tip position data. The resulting data set comprised 1241 evenly
spaced position measurements. From this large data set, we randomly generated 700 smaller continuous tracing intervals. The size of each sub-interval was randomly chosen from a discrete, uniform distribution to be between 200 and 400 position measurements. All of the 700 tracings met or exceeded the minimum surface area threshold.

For each of the 700 tracings, we performed registration using GoICP, mapping the segmented kidney model into the robot’s workspace. We then compared the registration to the ground truth pose of the kidney phantom, as measured using the optically tracked experiment platform. To evaluate the quality of each computed registration, we compared the vertex positions of the registered kidney model to the corresponding vertex positions in the tracked, ground truth model. The target registration error (TRE) at each vertex of the registered model was computed as the distance between that vertex and the corresponding vertex of the ground truth model.

Figure 7 shows an example registration with the TRE visualized as a heat map over the entire surface. The red lines shown in the figure represent the tracing (comprising 276 points in this example) of the physical kidney surface used for registration. In the region of the kidney where surface data was collected with the robotic instrument tip, TRE is approximately 2 mm while RMS TRE over the entire kidney surface is 2.75 mm. The TRE tends to increase as the distance from the data collection area increases, as should be expected.

To evaluate the overall consistency and reliability of our registration technique, we computed the RMS TRE over the entire kidney surface for all 700 tracings, as shown in Fig. 8. The average RMS TRE over all of the 700 registrations was 3.69 mm with standard deviation 0.61 mm. Note that while we performed registration using only a small number (200–400) of data points collected on the anterior kidney surface, we have reported RMS TRE over the entire kidney surface (at ∼175,000 mesh vertices). As opposed to considering TRE at a few points in the vicinity of the surface tracings, these results more realistically depict errors that can be expected over the surgical work volume when using touch-based registration.

V. IMAGE GUIDANCE PHANTOM EXPERIMENT

To demonstrate the utility of our system in the operating room, we performed a phantom experiment comparing the surgeon’s accuracy in localizing subsurface features both with and without the image guidance provided by our system. Figure 9 shows the setup for this experiment.

A challenge of many robotic surgical procedures is the localization of subsurface anatomy, making it difficult for the surgeon to know where to cut to remove lesions or avoid vessels. This is especially true in robot-assisted partial nephrectomy. During removal of perirenal fat surrounding the patient’s kidney, the surgeon must identify the locations of the renal artery, the renal vein, the ureter, and the tumor. All of these features are hidden beneath the fat layer. The fat must be carefully dissected, and the anatomical features must be uncovered while avoiding unnecessary damage which could result in blood loss or positive tumor margins.

We manufactured a soft, realistic silicone kidney phantom based on patient CT imaging. Eight acrylic spheres approximately 12 mm in diameter were set in the silicone material as it cured to serve as localization targets for experiments. Four targets were completely endophytic while the remaining 4 targets were at least partially exophytic. The model was fixed to the same optically tracked platform used in our registration experiments. The entire platform was CT scanned and segmented as before in Sec. IV-B. Optical tracking of the phantom platform was used solely to determine the ground truth positions of the embedded targets in the operating room to enable post hoc analysis of the surgeon’s localization accuracy.

During the experiment, the phantom was partially covered in a 10 mm thick layer of SynDaver synthetic fat. The fat completely covered all exophytic targets such that none of the localization targets were directly visible to the participating surgeon. Approximately 40% of the anterior surface of the
phantom was left uncovered to simulate the results of fat dissection required at the beginning of a partial nephrectomy procedure.

Prior to the experiment, the participating surgeon reviewed CT images of the phantom to develop a “mental map” of the locations of subsurface targets. The surgeon was allowed to reference the CT images throughout each phase of the experiment.

In the first phase of the experiment, the surgeon attempted to localize the subsurface targets based solely on his interpretation of the CT images. To localize a target, the surgeon was instructed to point (using the calibrated da Vinci instrument) to the perceived location of the center of each of the 8 targets. Each time the surgeon pointed to a target, our system recorded the transformation from the instrument tip to the system’s world coordinate frame to use for analysis. The surgeon repeated this pointing task 5 times for a total of 40 measurements.

In the next phase of the experiment, the surgeon repeated the subsurface localization task using our image guidance system. The surgeon first lightly traced the exposed portion of the kidney surface to collect surface data for our touch-based registration protocol. With the image guidance enabled by the registration, the surgeon repeated the above process of identifying target locations with the robotic instrument.

The final phase of the experiment served to establish a baseline accuracy for the target localization task. The image guidance was disabled, and the fat layer was completely removed from the phantom. With direct visualization of the entire kidney surface through the da Vinci’s endoscope, the surgeon then repeated the target localization process as before. Note that with direct visualization of the kidney surface, 4 of the target spheres were at least partially visible while the remaining 4 targets were completely concealed beneath the surface.

For all phases of the experiment, localization accuracy was evaluated by comparing target locations identified by the surgeon to the corresponding optically tracked, ground
truth locations. For a particular surgeon-identified location and ground truth pair, the localization error was taken as the minimum distance between the Z axis extracted from the end effector tip transform (the pointing direction of the instrument tip) and the ground truth target position (see the inset in Fig. 10). Figure 10 summarizes the localization accuracy results for each phase of the experiment.

Performing a pairwise t test on the set of all errors with image guidance and the set of all errors without image guidance showed a significant reduction in error. Mean error was reduced by 67% (from 9.2 mm to 3.0 mm) by using image guidance ($p < 0.001$).

For brevity, references to measurements made “with image guidance” or “without image guidance” in this discussion will refer specifically to those measurements made with fat partially obscuring the kidney surface. The term “direct visualization” refers to measurements made with all fat removed from the kidney surface (and no image guidance).

The measurements made with direct visualization provide helpful context for interpreting the localization accuracy results with and without image guidance. Direct visualization did not improve the surgeon’s localization accuracy of endophytic tumors when compared to the scenario without image guidance. This result indicates that increasing the visible surface area of the kidney did not help the surgeon form a more accurate mental registration between the CT images and the surgical scene. Using image guidance, on the other hand, enabled a substantial improvement of localization accuracy for endophytic targets when compared to both the scenario without image guidance and that with direct visualization. Taken together, these results demonstrate the utility of our image guidance system for enabling more accurate localization of subsurface features, specifically indicating that image guidance improved the surgeon’s localization ability beyond natural human ability.

With regards to the exophytic targets, direct visualization unsurprisingly resulted in substantially improved localization accuracy when compared to localization without image guidance (when the exophytic targets were completely obscured by fat cover), but direct visualization offered only slightly better accuracy than using image guidance. It is crucial to note, however, that direct visualization of exophytic tumors is very unlikely in a true clinical scenario. Surgeons rarely remove fat covering a tumor during partial nephrectomy for two reasons: (i) the fat provides a safe grasping point for manipulating and removing the tumor, and (ii) dissecting fat attached directly to the tumor dramatically increases the risk of unintentionally puncturing the tumor. Thus, image guidance was nearly as accurate as direct visualization while potentially being much safer.

While these results show great promise for our image guidance system, the current study was limited to one surgeon subject. Future work will be needed to explore whether the ability to accurately trace the kidney surface or to accurately identify subsurface features varies from subject to subject. Anecdotally, we have not observed large differences in these skills among either engineer or physician co-authors of this paper, but quantification of this anecdotal observation will be necessary in the future.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a practical and easily implemented image guidance system for the da Vinci family of robotic systems. Specifically, we proposed and tested a new method of accurately estimating required tracking parameters through calibration, achieving submillimetric absolute kinematic tracking accuracy for the first time with any da Vinci robot. We then presented the first quantitative accuracy evaluation of touch-based registration with the da Vinci, using an anatomically accurate SynDaver kidney model. We presented a system that incorporates these advancements to bring this concept toward real-world use. Finally, we demonstrated the utility of our system in the operating room in the first validation of touch-based image guidance to improve a surgeon’s ability during a subsurface target localization experiment.

The results of this work indicate a promising application for robotic partial nephrectomy which may increase the adoption of this underutilized alternative to total kidney removal; however, several significant challenges remain to be addressed before touch-based image guidance with the da Vinci can be fully realized in the operating room. The ex vivo experiments in this work provide a valuable proof of concept, but moving forward, it is crucial to perform in vivo evaluation of our image guidance system. In particular, we believe it may be useful to examine the effect of external forces applied to the da Vinci manipulators by the patient body wall and the insufflation system on kinematic tracking accuracy. Additionally, as prior works have noted, factors including patient positioning [45], arterial clamping [48], and kidney dissection [8], potentially cause organ deformation that can negatively affect registration accuracy when using preoperative images. Future work will be needed to address the effects of organ deformation that occurs throughout the surgical procedure. Substantial progress on these topics has been made (e.g., [8], [9], [49]) and can potentially be incorporated into our system in the future, particularly when accurate algorithms become computationally efficient enough for real-time use in a system like ours. Nonetheless, our results show that even rigid registration alone improves the ability of a surgeon to localize unseen subsurface objects in partial nephrectomy. Localizing these objects with confidence before cutting may, in the future, help to shift clinical decision making so that many more patients can receive the lifelong benefits of partial nephrectomy.

ACKNOWLEDGMENT

The authors would like to thank Robert Galloway and J. Michael Fitzpatrick for technical guidance, Simon Dimaio, Omid Mohareri, Cortney Jansen, and Intuitive Surgical for providing the API library necessary to interface with the clinical da Vinci system, and Gabor Fichtinger, Tamas Ungi, Andras Lasso, Kyle Sunderland, Csaba Pinter, Sal Choueib, Mark Asselin, and The Perk Lab at Queen's University in Kingston, Ontario for assistance with software. Note also that a preliminary (non-peer reviewed) version of some results in
this manuscript was presented in conference form in [31]; the current paper extends upon those results by more thoroughly describing the system as well as many improvements needed to facilitate eventual clinical translation, describing the calibration procedure rigorously, more thoroughly analyzing accuracy, and adding new experimental results demonstrating that the system can improve the performance of a human surgeon.

### References


