# Camera-Robot Calibration for the Da Vinci Robotic Surgery System

Orhan Özgüner [ID], *Member, IEEE*, Thomas Shkurti, Siqi Huang, *Member, IEEE*,

Ran Hao, *Graduate Student Member, IEEE*, Russell C. Jackson [ID], *Member, IEEE*,

Wyatt S. Newman, *Senior Member, IEEE*, and M. Cenk Çavuşoğlu [ID], *Senior Member, IEEE*

*Abstract*—The development of autonomous or semiautonomous surgical robots stands to improve the performance of existing teleoperated equipment but requires fine hand-eye calibration between the free-moving endoscopic camera and patient-side manipulator arms (PSMs). A novel method of solving this problem for the da Vinci robotic surgical system and kinematically similar systems is presented. First, a series of image-processing and optical-tracking operations are performed to compute the coordinate transformation between the endoscopic camera view frame and an optical-tracking marker permanently affixed to the camera body. Then, the kinematic properties of the PSM are exploited to compute the coordinate transformation between the kinematic base frame of the PSM and an optical marker permanently affixed thereto. Using these transformations, it is then possible to compute the spatial relationship between the PSM and the endoscopic camera using only one tracker snapshot of the two markers. The effectiveness of this calibration is demonstrated by successfully guiding the PSM end-effector to points of interest identified through the camera. Additional tests on a surgical task, namely, grasping a surgical needle, are also performed to validate the proposed method. The resulting visually guided robot positioning accuracy is better than the earlier hand-eye calibration results reported in the literature for the da Vinci system while supporting the intraoperative update of the calibration and requiring only devices that are already commonly used in the surgical environment.

*Note to Practitioners*—The problem of hand-eye calibration for the da Vinci robotic surgical system and kinematically similar systems is addressed in this article. Existing approaches have insufficient accuracy to automate low-level surgical subtasks and often require external patterns or subjective human intervention, none of which are applicable to practical robotic minimally invasive surgery (RMIS) scenarios. This article breaks down the calibration procedure into systematic steps to reduce error accumulation. Most of the time-consuming steps are performed offline, allowing them to be retained between movements. Each time the passive joints of the manipulator or the endoscope move, all that needs to be done is to refresh the transformation between the fixed markers. This key idea enables intraoperative updates of the hand-eye calibration to be performed online without sacrificing precision. The calibration method presented here demonstrates that the achieved accuracy is sufficient for

automating basic surgical manipulation tasks, such as grasping a suturing needle. The hand-eye calibration will be incorporated into a visually guided manipulation framework to perform high-precision autonomous surgical tasks.

*Index Terms*—da Vinci research kit (dVRK), medical robots and systems, surgical robotics: laparoscopy.

## I. INTRODUCTION

**D**UE to the nature of the master–slave teleoperation required to operate systems, such as the da Vinci robotic surgical system (Intuitive Surgical, Inc., Sunnyvale, CA, USA), robotic minimally invasive surgery (RMIS) usually results in longer operation times, and the learning curve is steeper than other minimally invasive surgical techniques [1]. In addition, the narrow laparoscopic camera view and limited workspace can make low-level tasks challenging and time-consuming. In order to reduce operation time and enhance surgeon performance, autonomous robotic surgical assistants [2] have been proposed to perform low-level surgical manipulation tasks, such as suturing [3]–[8], debridement [9], dissection, and retraction [10].

In order to perform these automated tasks with precision, it is necessary to know the transformation between the base frame of the robot manipulators ("hands") and the stereo endoscopic camera ("eye"). Once the hand-eye transformation is known, a detected object in the camera coordinate system can be located in the manipulator coordinate system, and the necessary motions to manipulate the object can be performed autonomously. Unfortunately, the surgical environment possesses several undesirable attributes that preclude traditional calibration strategies—most prominently, external objects, such as calibration grids, cannot be safely introduced into the patient, and the camera and patient-side manipulator arms (PSMs) are repositioned with respect to each other multiple times during an operation. This last condition further demands that calibration to be performed or updated "online" in a time-efficient manner.

This article focuses on computing the transformation between the endoscopic camera and the robotic manipulators in real time using an external tracking system. The presented work was performed using a Polaris Vicra optical tracking system (Northern Digital Inc., Waterloo, ON, Canada) and da Vinci robotic surgery system, but the methods developed can be used to calibrate any remote-center-of-motion (RCM)-based robot with any tracking system. In order to reduce the accumulation of error, this article proposes a procedure that is divided into systematic steps instead of finding a hand-eye transformation directly. Most of the time-consuming elements

of the calibration are performed offline in a manner that allows them to be retained between movements and even between surgical procedures, leaving the remainder able to be computed in real time without sacrificing precision. This key idea enables the intraoperative updates of the hand-eye transformation to be performed instantly.

This article is organized as follows. Section II discusses related studies on robot-camera (hand-eye) calibration. The experimental hardware is described in Section III. In Section IV, the problem formulation and proposed calibration methods are described. The details of the hardware-based validation tests and the validation results are presented in Section V, followed by the conclusions in Section VI.

## II. Related Studies

The problem of hand-eye calibration in robotic systems has been well studied by many researchers using various techniques. Early studies described the problem by separating the transformation into translation and rotation components [11]–[15]. Others then argued that the separation implies that the rotation component has nothing to do with the translation component, which is not a viable assumption, and attempted to find a solution simultaneously [16]–[19]. The problem with this method was that orientation errors were propagating into positional errors. To reduce the rotational error propagation, researchers then proposed to solve the problem iteratively [20]–[22]. The extension of the robot-world calibration problem to hand-eye calibration has also been studied [23]–[25].

Wang *et al.* [26] proposed a method for hand-eye calibration of RCM-based robots, such as the da Vinci robotic surgery system. The translation part of the calibration matrix is computed by moving the endoscope to at least two different positions, and the orientation is obtained by constructing an optimization problem with different images of the robotic manipulator. The average translational error is reported as 16.08 mm.

Zhang *et al.* [27] proposed to solve the calibration problem for RMIS systems using an internal dot pattern attached to the tip of the manipulator. The hand-eye relationship was computed iteratively, and the convergence of this iterative computation was proven. While the resulting transforms were found to be self-consistent over multiple end-effector positions, they were not compared to ground truth.

D'Ettorre *et al.* [28] developed a vision-guided method for automatically grasping the suturing needle. The hand-eye calibration is identified by establishing 3-D point correspondences between the camera and the robot tooltip while positioning the tip at corners of a calibration grid and detecting these corners from the camera.

Pachtrachai *et al.* [29] developed a method without a calibration object for RMIS systems. The surgical manipulator's pose is estimated relative to the endoscopic camera using the 3-D instrument tracking method described in [30]. The minimum instrument tracking error was reported as about 20 mm, while the manipulator was in motion.

Seita *et al.* [31] studied the problem by employing a two-phase calibration procedure. In the first phase, the trajectories in the workspace are automatically explored with random targets, using red tape tied to the end-effector to visually track the location. A deep neural network (DNN) was trained in the obtained data. In the second phase of the procedure, the manipulator's end-effector was sent to a target on a printed calibration grid. The error of the end-effector was corrected by a human expert. A random forest (RF) was then trained on the collected data to predict the residual error.

These previous studies in the literature on hand-eye calibration showed great promise. However, achieved precision is either insufficient or requires an external pattern in order to perform low-level subtasks autonomously. Furthermore, some of these studies rely on simplifying assumptions, such as artificially coloring the end effectors; require vision-based tracking algorithms that are not computationally feasible; or employ subjective human expertise during algorithm training—none of which are applicable to practical RMIS scenarios.

Previous studies tried to solve the hand-eye calibration problem directly to find the transformation between the camera and the manipulator. In contrast, the presented study aims to perform the calibration procedure in a divide-and-conquer fashion, performing the majority of the calibration procedure preoperatively using an external optical sensor (which is already commonly used in operating rooms [32]). These components of the calibration are not affected by surgical instrument change or movement during the procedure. Once the relative transformations of the endoscopic camera and the manipulator RCM point to fixed passive markers are computed preoperatively, the remainder of the calibration procedure requires only a single snapshot of the relative marker transformations—this is very quick, operating-room friendly, and compatible with all RCM-based manipulators. This systematic breakdown also reduces errors in kinematic parameters, such as link and joint angle offsets.

To the best of our knowledge, this article is the only method currently available, which is applicable to a realistic laparoscopic surgical scenario, since it is capable of recovering hand-eye calibration in real time following the movement of the robot's passive joints. The presented method highlights and accommodates significant deviations from the idealized and experimentally identified Denavit–Hartenberg (DH) parameters for the physical da Vinci PSM. This kinematic intrinsic identification method can be used to characterize and correct any RCM-based robot. In addition, the open-loop point-reaching accuracy results achieved in this article are equivalent to or better than earlier results reported in the literature. Specifically, the achieved open-loop point-reaching accuracy is equivalent to the accuracy of earlier studies that rely on closed-loop visual servocontrol and significantly better than methods that reported open-loop reaching accuracy.

## III. Hardware Description

### A. Da Vinci Research Kit

In order to control the da Vinci surgical robotic system in an automated manner, we employed the da Vinci Research Kit (dVRK) developed by Johns Hopkins University and the Worcester Polytechnic Institute (WPI) [33] (see Fig. 1). The dVRK acts as a substitute for the teleoperation master
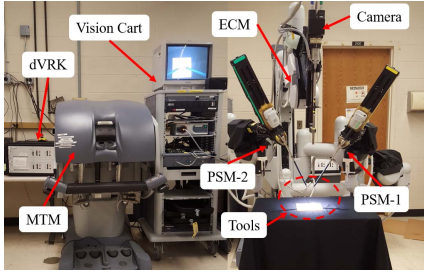
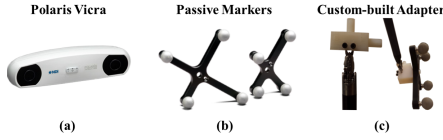Fig. 1. da Vinci surgical robotic system with the dVRK.



Fig. 2. (a) Polaris Vicra optical tracking system. (b) Passive markers. (c) Custom-built adapter.

station via a ROS interface that can be controlled from any desktop computer [33]. Forward and inverse kinematics allow this joint-level control to be leveraged into workspace (3-D Cartesian) control of the robot in the PSM base frame. Multiple PSMs can be managed in parallel with the device. In this article, only a single manipulator was employed; however, the procedure can be extended to calibrate multiple manipulators at the same time or in sequence.

### B. Polaris Optical Tracker

In order to identify kinematic calibration parameters, an optical tracker system—the NDI Polaris Vicra [34]—is employed as an external sensor. The Polaris Vicra can measure the position and orientation of passive markers composed of retroreflective spheres in specific and unique geometric shapes [34]. The unique marker geometry allows the system to track multiple markers simultaneously, and the ROS-compatible wrapper package enables a desktop computer to read the marker position and orientation at 20 Hz. The accuracy of the system is reported as 0.25-mm rms within its operating volume [34]. The system is already approved for and commonly used in surgical operations, most prominently in neurosurgery [32].

Our calibration setup requires two permanent passive markers—one affixed to the base of the PSM and the other affixed to the camera, both external to the patient. In order to perform the offline stages of the calibration, we have also designed a removable adapter that is attached to the gripper and able to hold the third marker (see Fig. 2) and a machined calibration board with a fourth marker embedded in its surface (see Fig. 3). Neither of these devices needs to be present during the surgical procedure, as they are only used during the preoperative calibration phase.

## IV. CALIBRATION PROCEDURE

The hand-eye calibration procedure finds the transformation between the camera and the manipulators in three independent steps by introducing an external optical sensor. In the first step of the procedure, which is explained in detail in Section IV-B,

a fixed marker is attached to the camera, and the transformation between the camera and the marker is computed via a custom-built calibration board using the frame assignments shown in Section IV-A. The second step of the procedure is achieved by attaching an optical marker onto the robot manipulator base, and the transformation between the robot base and the marker is calculated using the method detailed in Section IV-C. In the third step, the transformation between the fixed markers is computed using the optical tracking system. The final hand-eye calibration is calculated by using these three intermediate transformations.

### A. Frames and Transformations of Interest

1) $C$ is the camera optical frame or "eye" frame. Pixel locations deprojected from image space to 3-D Cartesian space using the endoscope's binocular vision are natively expressed in this frame.
2) $M_c$ is the frame of a Polaris marker rigidly affixed to the endoscopic camera pole.
3) $P$ is the PSM base frame, "portal" frame, or "hand" frame. It is the origin of the workspace of the PSM, meaning that forward and inverse kinematics operate on points expressed in this frame.
4) $M_p$ is the frame of a Polaris marker rigidly affixed to the PSM base.
5) $D$ is the frame of a camera calibration grid freely movable within the robot workspace.
6) $M_d$ is the frame of a Polaris marker rigidly affixed to that calibration grid.
7) $M_t$ is the frame of a Polaris marker rigidly affixed to the PSM tooltip.

It should be noted that the Polaris sensor reports the 6-DOF location of a marker (position and orientation) in a Cartesian frame $N$ internal to the device—more precisely, a transform between the origin of $N$ and the origin of the marker. This location will change if the Polaris is rotated or moved and has no direct physical relevance to the da Vinci unit or objects within its workspace. Given any two reported locations $X$ and $Y$ from the same tracker position, it is possible to compute a transform $g_{XY}$ between them that is independent of the tracker and thus persistent following its removal:

$$X = g_{NX}, \quad Y = g_{NY}$$
$$\therefore \ [g_{NX}]^{-1} \cdot g_{NY} = g_{XN} \cdot g_{NY} = g_{XY}. \tag{1}$$

This method is used to compute several of the transforms (see Fig. 4) described in the following and can be performed using a single Polaris reading in real time.

1) $g_{M_c C}$ is the transform between the camera and the marker affixed to it, which also remains constant when the camera is moved. The determination of this transform is discussed in Section IV-B.
2) $g_{M_p P}$ is the transform between the PSM base and the marker affixed to it, which remains constant even when the PSM is moved. The determination of this transform is discussed in Section IV-C.
3) $g_{M_p M_c}$ is the transform between $M_p$ and $M_c$. It changes whenever the PSM or camera is moved but can be
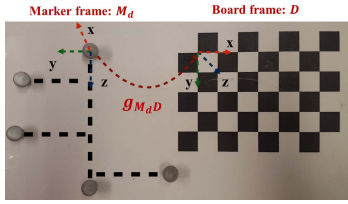
Fig. 3.    Custom-built board with known transformation between tracking marker and grid.
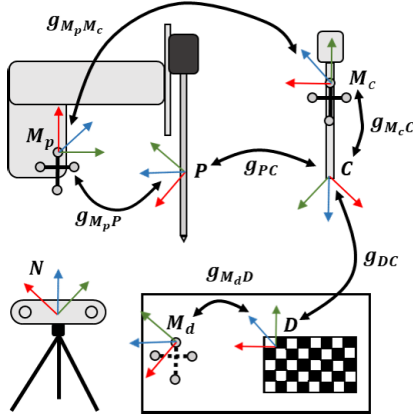


Fig. 4.    Relevant frames and transformations for the da Vinci surgical system.



Fig. 5.    Camera to Polaris calibration schematic diagram.
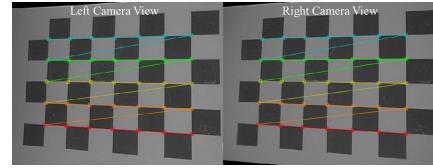


Fig. 6.    Corners of the board detected in the camera frame.

redetermined in real-time using only a single Polaris reading of both markers, as per (1).

4) $g_{dc}$ is the camera-to-calibration-grid transform. It changes based on the location of the calibration grid and does not exist at all when (as in surgery) the grid is not present. It is discussed further in Section IV-B.

5) $g_{M_dD}$ is the transform between the calibration grid and the marker affixed to it. The physical structure holding both components is CNC-machined so that this transform is constant and known to high precision

$$g_{M_dD} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & -80\,\text{mm} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}. \qquad (2)$$

6) $g_{M_dM_c}$ is the transform between the calibration grid and camera markers. It is discussed further in Section IV-B.

7) $g_{PC}$ is the hand-eye transform of interest. Using this transform, a physical point identified in $C$ can be transformed into $P$ and inverse kinematics employed to make the da Vinci tool travel to that point. Once $g_{M_pM_c}$, $g_{M_pP}$, and $g_{M_cC}$ are known

$$g_{PC} = [g_{M_pP}]^{-1} \cdot g_{M_pM_c} \cdot g_{M_cC}. \qquad (3)$$

### B. Camera to Polaris Registration

One of the preoperative calibration procedures involves finding the transformation between the camera frame $C$ and the marker frame $M_c$ attached to the camera (see Fig. 5). The transformation $g_{M_cC}$ is computed offline using a custom-built visual calibration board that includes a tracking marker (see Fig. 3), which serves as a connection between the camera and the Polaris tracker. Before hand-eye calibration, the endoscopic
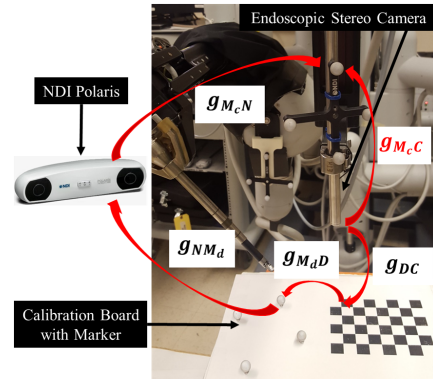
camera system optical intrinsics are calculated by the ROS industrial-calibration package [35]. The resulting intrinsics show a 0.37-pixel reprojection error, which contributes (along with detection errors impossible to directly measure) to a total error in locating the positions of detected objects, which averages 0.49 mm and has a maximum of 2.8 mm [36].

First, the board is positioned such that it is visible to both the Polaris tracker and the endoscopic camera. Using the images from the endoscope, the grid corner locations are then detected in the camera frame (see Fig. 6) using a corner detection algorithm [37]. The corner locations can be used as inputs to the OpenCV solvePnP optimization algorithm [37] to compute $g_{CD}$ using the camera's intrinsic parameters, the corner locations of the grid in camera frame ($C$), and corresponding corner locations in the grid frame ($D$). Since $g_{M_dD}$ is known and a single Polaris reading gives $g_{M_dM_c}$ with (4), $g_{M_cC}$ can be computed as follows:

$$g_{M_cC} = [g_{M_dM_c}]^{-1} \cdot g_{M_dD} \cdot g_{DC}. \qquad (4)$$

Since the corner detection and the Polaris readings might be affected by the board position and lighting conditions, 500 data samples using different positions and orientations of the board are collected. The developed software has the ability to automatically collect data, while the board is moved under the camera and tracker. After collecting the data, a constrained nonlinear optimization problem is constructed as

$$g_{M_cC} = \text{argmin}_{g_{M_cC}} \sum_{i=0}^{n} \text{error}\left(g_{M_cC}, g_{DC}^i, g_{M_dM_c}^i\right)$$

$$g_x^i = g_{M_cC} \cdot [g_{DC}^i]^{-1} \cdot [g_{M_dD}]^{-1} \cdot g_{M_dM_c}^i = \begin{pmatrix} R_x & p_x \\ 000 & 1 \end{pmatrix}$$

$$\text{error} = ||p_x|| + \alpha \cdot \left| \arccos\left(\frac{\text{Trace}(R_x) - 1}{2}\right) \right| \qquad (5)$$
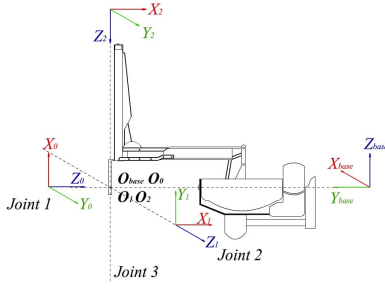
Fig. 7. Robot illustration with base frame and DH 0, 1, and 2 frames. The kinematics of the PSM is represented by the DH convention. The first three frames are ideally intersecting at the origin of the base frame and are mutually perpendicular.
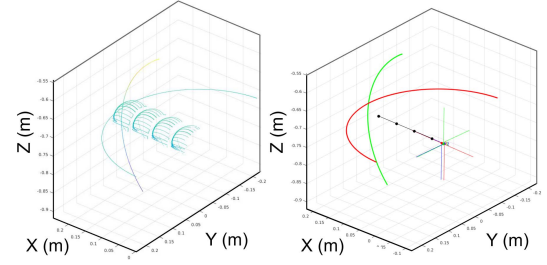


Fig. 8. Left: point cloud containing marker positions obtained by "circle-drawing" and "sphere-drawing" motions. Right: fitted arcs and sphere centers.

where $n$ is the size of the data pool (500 data sets were enough to cover the region in our case) and $\alpha$ is a scale factor identified empirically. The error is the weighted sum of the translational and rotational error, where the rotational error is given by the angle of the axis-angle representation. This optimization problem is solved using MATLAB's *fmincon* utility using the default (interior-point) optimization algorithm with no linear constraints or region.

$g_{M_cC}$ does not depend on the specific location of the calibration board. Once it is established, it remains valid even if the joints of the endoscope-holding arm are moved. Therefore, the procedures described earlier can be done offline and preserved across multiple camera positions and multiple surgeries.

### C. PSM Kinematic Model and PSM to Polaris Registration

Joint axes 1–3 ($\omega_1$–$\omega_3$) of the da Vinci PSM are designed to intersect at a single point (namely, RCM) and be mutually perpendicular, defining the PSM base frame $P$'s position and orientation (see Fig. 7). On a physical robot, this may not be the case because of imperfections in the mechanism, making the definition of $P$ and the subsequent kinematic calculations inaccurate. In order to improve the accuracy of the PSM kinematics and the PSM to Polaris registration, an appropriately defined DH parameterization of the PSM kinematics, taking into account these imperfections, is identified as part of the process.

The PSM to the Polaris registration and the identification of the PSM kinematic parameters are performed with the help of a passive marker $M_t$ affixed to the PSM end-effector (see Fig. 2) and tracked by the Polaris sensor.

In order to determine $\omega_1$, joint $J_1$ is swept from its minimum to maximum value while freezing joints $J_2$–$J_7$ at 0 displacement. The passive marker affixed to the PSM end-effector will describe an arc in three-space, which can be fit by a least-squares method to a full circle [38]. $\omega_1$ passes through the center of the circle and is parallel to the normal vector of the plane containing it, allowing its position and orientation to be calculated in $N$. Analogously, sweeping $J_2$ while holding $J_1$ and the other joints at 0 displacement allows us to determine $\omega_2$ (see Fig. 8).

To determine $\omega_3$, $J_1$ and $J_2$ are held at 0 displacement, and $J_3$ is moved to four different positions. At each, wrist joints

TABLE I
IDEAL AND REALISTIC DH FRAME PARAMETERS OF THE PSM

| Joint | $\theta_{ideal}$ (rad) | $\alpha_{ideal}$ (rad) | $a_{ideal}$ (mm) | $d_{ideal}$ (mm) |
|---|---|---|---|---|
| $J_1$ | $\theta_1$ | $\frac{\pi}{2}$ | 0 | 0 |
| $J_2$ | $\theta_2 + \frac{\pi}{2}$ | $\frac{\pi}{2}$ | 0 | 0 |
| $J_3$ | 0 | 0 | 0 | $\theta_3$ |
| Joint | $\theta_{real}$ **(rad)** | $\alpha_{real}$ **(rad)** | $a_{real}$ **(mm)** | $d_{real}$ **(mm)** |
| $J_1$ | $\theta_1 + \pi$ | $\frac{\pi}{2} - 0.0062$ | 5.74 | 0 |
| $J_2$ | $\theta_2 - \frac{\pi}{2} - 0.022$ | $\frac{\pi}{2} - 0.010$ | 1.12 | $-2.97$ |
| $J_3$ | 0 | 0 | 0 | $\theta_3 - 7.41$ |

$J_4$ and $J_5$ are swept from their minimum to maximum values, causing the tip of the gripper and the marker affixed thereto to describe a small ($\approx$40-mm diameter) sphere. The origin of each sphere is then identified using a similar least-squares fitting algorithm [38]. A line passing through all four sphere center points can then be determined by another least-squares fit [38], which describes $\omega_3$ in the Polaris frame (see Fig. 8).

Based on the aforementioned calculations, it was confirmed that $\omega_1$, $\omega_2$, and $\omega_3$, indeed, are not perpendicular and do not intersect. The difference between the idealized and experimentally identified DH parameters for the physical da Vinci PSM used in the experiment is shown pictorially in Fig. 9 and numerically in Table I.

Once the position and orientation of the PSM base frame relative to the Polaris frame, $g_{NP}$, are known, (1) can be used to determine $g_{M_pP}$ using $g_{NM_p}$. Similar to $g_{M_cC}$, the end result of this process does not depend on the position of the Polaris unit and remains valid even if the joints of the PSM-holding arm are moved, or the tip marker $M_t$ removed. Therefore, the procedures described can be done offline and preserved across multiple PSM positions and multiple surgeries.

### V. EXPERIMENTAL VALIDATION

#### A. Experiment Design

The accuracy of the calibration method can be evaluated by locating a point-of-interest (POI) or a set of POIs through the endoscopic camera and then commanding the da Vinci robot to move its PSM to those points (experimental setup is shown in Fig. 10). A consistent offset between the physical POIs $R_i$ and the end-effector point-of-arrival $Q_i$ represents an error in the calibration. $R_i$ is generated from the inner intersections of a $5 \times 7$ gray-scale checkerboard with 15-mm squares, which are located in the camera space using [37]. As these points correspond to the physical features of an object, it is possible
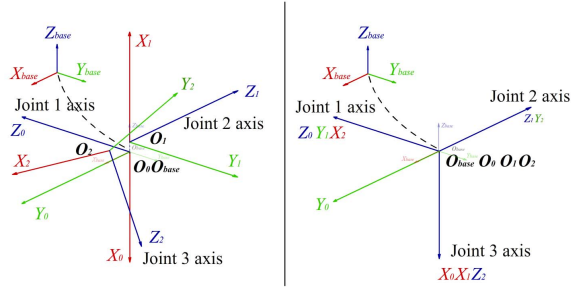
Fig. 9. Idealized (right) and real (left) axes of joints 1–3. Frames 0, 1, and 2 are intersecting at the origin of base frame under the ideal assumption, which conflicts with the reality. The difference increases the cumulative positioning error significantly.
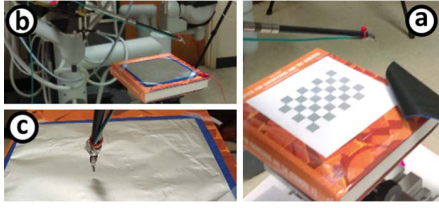


Fig. 10. Experimental test setup. (a) Board is placed under the endoscope to detect the corners. Carbon paper is placed on the board to record the X- and Y-error components. (b) Conductive foil is then placed on top to record the Z-axis error. (c) Manipulator is set to touch to the board corners.

to precisely measure the distance $\varepsilon_i$ between each $R_i$ and the resulting $Q_i = g_{PC} \cdot R_i$ as $X$, $Y$, and $Z$ components in an arbitrary 3-D frame (in this case, the board frame $D$)

$$\varepsilon_i = ||R_i - Q_i|| = \sqrt{\varepsilon_{Xi}^2 + \varepsilon_{Yi}^2 + \varepsilon_{Zi}^2}. \qquad (6)$$

In order to determine these components, after the POIs are recorded, but before the robot moves, the calibration grid is covered with a sheet of carbon paper backed by conductive foil (with a total thickness less than 0.1 mm). The foil is connected to the positive terminal of a computerized voltmeter, the conductive metal tip of the PSM is wired to a +5-V dc power source, and the voltmeter and power source share a common ground. Thus, when the PSM tip makes physical contact with the foil (and, by extension, the checkerboard intersection beneath), the voltage reported by the meter software jumps from 0 V to approximately +5 V.

For each $Q_i$, the system is commanded to move first to a position 5 mm above the calculated $Z$-value and then descend with ±0.02-mm increments until an increase in voltage is registered—this means that the PSM tip is in contact with the grid and, therefore, at the true physical $Z$-coordinate of $R_i$. $R_{Zi} - Q_{Zi} = \varepsilon_{Zi}$. In addition, when the tip applies pressure to the foil, a dark mark is transferred from the carbon paper beneath it to the checkerboard. The displacement from this mark to the corresponding corner of the board is measured using calipers to determine $\varepsilon_{Xi}$ and $\varepsilon_{Yi}$. Each test involved 15 points, and the tests were repeated 13 times with the board at different locations under the camera field of view, for a total data set of 195 unique points.

In order to measure calibration performance in terms of rotation, the robot was commanded to approach identified

TABLE II
DESCRIPTIVE STATISTICS OF COMPONENT AND TOTAL
ERRORS (mm AND °)

|  | $\varepsilon_X$ | $\varepsilon_Y$ | $\varepsilon_Z$ | $\varepsilon$ | $\varepsilon_\theta$ |
|---|---|---|---|---|---|
| Mean | 1.5 | 1.1 | 0.5 | **2.1** | **3.2°** |
| Std. Dev. | 1.8 | 1.2 | 0.3 | **0.9** | **3.4°** |
| Max. | 3.8 | 3.0 | 1.4 | **4.3** | **6.0°** |

checkerboard points with the gripper held at various angles (90°, 60°, and 30°). Then the angle between the robot grippers and the board surface was measured manually with a digital angle gauge (Pittsburgh Digital Angle Gauge, with accuracy = ±0.3° as reported in the tool user manual). The difference between the commanded angle and the physically measured angle was recorded as the rotational error. Each angle was tested with 100 trials (total of 300 unique positions). The experimental procedure is summarized in the video attachments.

In addition to the quantitative measurements, the quality of the calibration was evaluated in an application-relevant task, namely, by grasping a surgical needle. The needle was placed randomly into the camera view (ensuring that the needle is reachable by the robot end-effector). A flat white surface, a flat printout of a photographed surgical procedure, a soft flat pad, and a suturing training pad were used as backgrounds. The location of the needle is determined by a needle tracking algorithm [3] in the camera frame. By using the calibration, coordinates of the needle are then transformed into the robot base frame, and the robot is commanded to grasp the needle. A total of 20 needle-grasping trials were performed.

The quality of the intraoperative updates of the hand-eye calibration after the motion of the camera, and the robot base was also experimentally evaluated. Specifically, after the calibration procedure is performed and the camera-robot calibration was obtained, the passive joints of the robot arm were manually moved to several new configurations. At each configuration, the optical tracker was used to obtain the transformation between the fixed markers attached to the camera and the robot base frames, and the new camera-robot calibration was obtained by updating only the affixed marker transformation. A total of five different configurations obtained by manually moving the passive joints were tested, where each configuration was evaluated by performing five sets of POI positioning tests (total of 375 unique points) and five visually guided needle grasps (totaling 25 trials).

### B. Results

Table II shows a statistical breakdown of the error across all 195 points. $\varepsilon_{Xi}$ and $\varepsilon_{Yi}$ are measured to a resolution of ±0.1 mm using calipers; $\varepsilon_{Zi}$ was measured using the descent method described earlier to a resolution of ±0.02 mm. Figs. 11 and 12 show the distribution of the error values. Table II also shows the rotational error breakdown $\varepsilon_\theta$. Out of 300 trials, the mean rotational error is calculated as 3.2° with the standard deviation of 3.4° and a maximum of 6°.

The needle-grasping experiments were performed a total of 20 times. The robot successfully grasped the needle
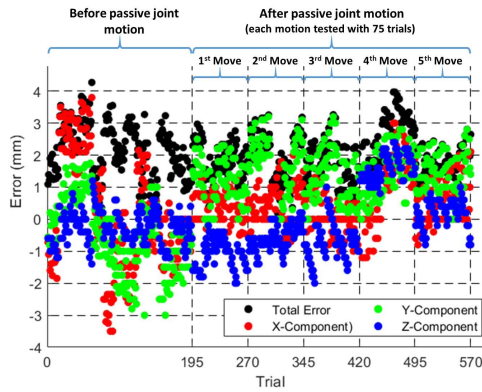
Fig. 11. Euclidean distance and component-axis errors plotted with respect to point number.
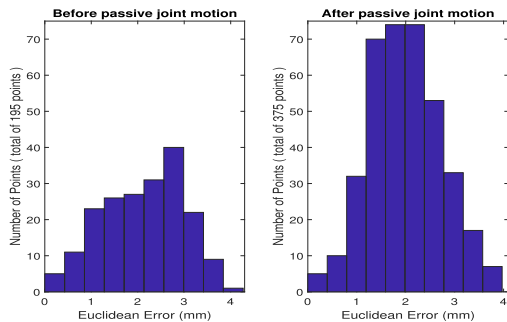


Fig. 12. Histogram of the Euclidean errors before and after passive joint motions.

in 18 cases and failed in two cases. In the failure cases, the robot grippers were very close to the needle. In both cases, the gripper made contact with the needle. The video attachment contains several successful needle grasps under varying background surfaces as well as the failure cases.

The reliability of the calibration after the movement of the passive joints was tested with a total of five different configurations with 25 needle-grasping experiments (five experiments for each configuration). After the manual motion on the passive joints, using the updated calibration, the robot successfully grasped the needle in 21 cases and failed in four cases (the gripper hit the needle in two cases and missed the needle in two cases). The measured POI positioning errors (cumulative for the five configurations) were 2.0 mm (rms), 0.7 mm (std. dev.), and 4.0 mm (max).

The calibration procedure described in this article takes about 40 min. Specifically, the part described in Section IV-B takes ~5 min (data collection 4 min and processing 1 min) and in Section IV-C takes ~35 min (data collection 34 min and processing 1 min). A video of the robot moving to different POIs after the movement of the passive joints and calibration update is included in the video attachment.

## VI. CONCLUSION

In this article, a method for hand-eye calibration of the da Vinci robotic surgery system is presented. Although this article specifically focuses on the calibration of the da Vinci system, the proposed method is broadly applicable to any RCM-based robotic mechanism, such as the Raven surgical robotic system [39].

Instead of trying to directly find a transformation between the camera and the manipulators, the calibration procedure is performed in three independent steps to reduce the cumulative error and to allow the calibration procedure to be partitioned into preoperative (offline) and intraoperative (online) parts. An external optical sensor—the NDI Polaris Vicra—is used to identify the kinematic calibration parameters during the calibration procedure. In the first step of the procedure, a fixed marker is attached to the camera, and the transformation between the camera and the marker via a custom-built calibration board is computed. In the second step of the procedure, an optical marker is attached to the robot manipulator base, and the transformation between the robot base and the marker is calculated. These two steps can be preoperatively performed offline, and the transforms discovered persist between surgical operations. In the third step, the Polaris Vicra is used to identify the transformation between the affixed markers to find the final calibration. Each time the passive joints of the manipulator or the endoscope move, all that needs to be done is to refresh the transformation between the fixed markers, enabling the quick, online intraoperative update of the hand-eye transformation.

The resulting calibration methodology produced visually guided end-effector motions with an rms positioning error of 2.1 mm and a maximum error of 4.3 mm. The rotation accuracy of the method is calculated as 3.2° with the standard deviation of 3.4° and a maximum of 6°. The resulting accuracy of the system after calibration is better than the earlier hand-eye calibration results reported in the literature for the dVRK. The quality of the calibration is evaluated in an application relevant task, namely, by grasping a surgical needle (45 experiments and 39 successful needle grasps). This functional validation results also demonstrated that the achieved calibration accuracy is sufficient for basic surgical manipulation tasks, such as needle grasping. In addition, following a one-time offline calibration step, the intraoperative update of the calibration can be performed in the order of a few seconds in real time without human intervention, excessive computational resources, or the introduction of nonsurgical equipment into the operating area. Although the relatively small visual range of the Polaris Vicra limits the workspace in which calibration is possible, this limitation could easily be rectified by substituting another optical tracker, such as the Polaris Spectra or Vega, with a larger workspace. Nonoptical trackers can also be employed [40]. Since the calibration procedure depends on the external tracking equipment, the cost should be taken into account although it is typically a small fraction of the overall system cost.

## REFERENCES

[1] P. Yohannes, P. Rotariu, P. Pinto, A. D. Smith, and B. R. Lee, "Comparison of robotic versus laparoscopic skills: Is there a difference in the learning curve?" *Urology*, vol. 60, no. 1, pp. 39–45, Jul. 2002.

[2] G. P. Moustris, S. C. Hiridis, K. M. Deliparaschos, and K. M. Konstantinidis, "Evolution of autonomous and semi-autonomous robotic surgical systems: A review of the literature," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 7, no. 4, pp. 375–392, Dec. 2011.

[3] O. Ozguner, R. Hao, R. C. Jackson, T. Shkurti, W. Newman, and M. C. Cavusoglu, "Three-dimensional surgical needle localization and tracking using stereo endoscopic image streams," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 6617–6624.

[4] R. Hao, O. Ozguner, and M. C. Cavusoglu, "Vision-based surgical tool pose estimation for the da Vinci robotic surgical system," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 1298–1305.

[5] R. C. Jackson and M. C. Cavusoglu, "Needle path planning for autonomous robotic surgical suturing," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 1669–1675.

[6] T. Liu and M. C. Cavusoglu, "Needle grasp and entry port selection for automatic execution of suturing tasks in robotic minimally invasive surgery," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 552–563, Apr. 2016.

[7] R. C. Jackson, V. Desai, J. P. Castillo, and M. C. Cavusoglu, "Needle-tissue interaction force state estimation for robotic surgical suturing," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 3659–3664.

[8] R. C. Jackson, R. Yuan, D.-L. Chow, W. S. Newman, and M. C. Cavusoglu, "Real-time visual tracking of dynamic surgical suture threads," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 3, pp. 1078–1090, Jul. 2018.

[9] B. Kehoe *et al.*, "Autonomous multilateral debridement with the Raven surgical robot," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 1432–1439.

[10] B. Kehoe *et al.*, "Autonomous tumor localization and extraction: Palpation, incision, debridement and adhesive closure with the da Vinci research kit," in *Proc. Hamlyn Surgical Robot. Conf.*, 2015, pp. 1–27.

[11] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX=XB," *IEEE Trans. Robot. Autom.*, vol. 5, no. 1, pp. 16–29, Feb. 1989.

[12] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Trans. Robot. Autom.*, vol. 5, no. 3, pp. 345–358, Jun. 1989.

[13] C.-C. Wang, "Extrinsic calibration of a vision sensor mounted on a robot," *IEEE Trans. Robot. Autom.*, vol. 8, no. 2, pp. 161–175, Apr. 1992.

[14] F. C. Park and B. J. Martin, "Robot sensor calibration: Solving AX=XB on the Euclidean group," *IEEE Trans. Robot. Autom.*, vol. 10, no. 5, pp. 717–721, Oct. 1994.

[15] R. Horaud and F. Dornaika, "Hand-eye calibration," *Int. J. Robot. Res.*, vol. 14, no. 3, pp. 195–210, 1995.

[16] H. H. Chen, "A screw motion approach to uniqueness analysis of head-eye geometry," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1991, p. 145–151

[17] K. Daniilidis, "Hand-eye calibration using dual quaternions," *Int. J. Robot. Res.*, vol. 18, no. 3, pp. 286–298, Mar. 1999.

[18] E. Bayro-Corrochano, K. Daniilidis, and G. Sommer, "Motor algebra for 3D kinematics: The case of the hand-eye calibration," *J. Math. Imag. Vis.*, vol. 13, no. 2, pp. 79–100, 2000.

[19] Z. Zhao and Y. Liu, "Hand-eye calibration based on screw motions," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, 2006, pp. 1022–1026.

[20] Z. Zhao, "Hand-eye calibration using convex optimization," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2011, pp. 2947–2952.

[21] J. Angeles, G. Soucy, and F. P. Ferrie, "The online solution of the hand-eye problem," *IEEE Trans. Robot. Autom.*, vol. 16, no. 6, pp. 720–731, Dec. 2000.

[22] R. L. Hirsh, G. N. DeSouza, and A. C. Kak, "An iterative approach to the hand-eye and base-world calibration problem," in *Proc. ICRA. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 2171–2176.

[23] H. Zhuang, Z. S. Roth, and R. Sudhakar, "Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form AX=YB," *IEEE Trans. Robot. Autom.*, vol. 10, no. 4, pp. 549–554, Aug. 1994.

[24] M. Shah, "Solving the robot-world/hand-eye calibration problem using the Kronecker product," *J. Mech. Robot.*, vol. 5, no. 3, pp. 549–554, Aug. 2013.

[25] H. Li, Q. Ma, T. Wang, and G. S. Chirikjian, "Simultaneous hand-eye and robot-world calibration by solving the $AX = YB$ problem without correspondence," *IEEE Robot. Autom. Lett.*, vol. 1, no. 1, pp. 145–152, Jan. 2016.

[26] Z. Wang *et al.*, "Vision-based calibration of dual RCM-based robot arms in human-robot collaborative minimally invasive surgery," *IEEE Robot. Autom. Lett.*, vol. 3, no. 2, pp. 672–679, Apr. 2018.

[27] Z. Zhang, L. Zhang, and G.-Z. Yang, "A computationally efficient method for hand–eye calibration," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 10, pp. 1775–1787, Oct. 2017.

[28] C. D'Ettorre, G. Dwyer, X. Du, F. Chadebecq, F. Vasconcelos, E. DeMomi, and D. Stoyanov, "Automated pick-up of suturing needles forrobotic surgical assistance," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2018, pp. 1370–1377.

[29] K. Pachtrachai, M. Allan, V. Pawar, S. Hailes, and D. Stoyanov, "Hand-eye calibration for robotic assisted minimally invasive surgery without a calibration object," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 2485–2491.

[30] M. Allan, P.-L. Chang, S. Ourselin, D. J. Hawkes, A. Sridhar, J. Kelly, and D. Stoyanov, "Image based surgical instrumen pose estimation with multi-class labelling and optical flow," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, N. Navab, J. Hornegger, W. M. Wells, and A. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 331–338.

[31] D. Seita, S. Krishnan, R. Fox, S. McKinley, J. Canny, and K. Goldberg, "Fast and reliable autonomous surgical debridement with cable-drivenrobots using a two-phase calibration procedure," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, 2018, pp. 6651–6658.

[32] R. Khadem *et al.*, "Comparative tracking error analysis of five different optical tracking systems," *Comput. Aided Surg.*, vol. 5, no. 2, pp. 98–107, Jan. 2000.

[33] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da Vinci surgical system," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 6434–6439.

[34] *Polaris Vicra User Guide Rev. 6*, Northern Digital Inc., Waterloo, ON, Canada, 2012.

[35] *Opencv*. Accessed: Jan. 26, 2020. [Online]. Available: https://opencv.org/

[36] T. Shkurti, "Simulation and control enhancements for the da vinci surgical robot," M.S. thesis, Dept. Elect. Eng. Comput. Sci., Case Western Reserve Univ., Cleveland, OH, USA, Jan. 2019.

[37] *Rosindustrial*. Accessed: Jan. 26, 2020. [Online]. Available: https://github.com/ros-industrial/

[38] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, Apr. 2000.

[39] B. Hannaford *et al.*, "Raven-II: An open platform for surgical robotics research," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 4, pp. 954–959, Apr. 2013.

[40] X. Liu, W. Plishker, G. Zaki, S. Kang, T. D. Kane, and R. Shekhar, "On-demand calibration and evaluation for electromagnetically tracked laparoscope in augmented reality visualization," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, no. 6, pp. 1163–1171, Jun. 2016.