

# Caveats on the first-generation da Vinci Research Kit: latent technical constraints and essential calibrations

Zejian Cui<sup>\*,1,2</sup>, João Cartucho<sup>\*,1</sup>, Stamatia Giannarou<sup>1</sup>, and Ferdinando Rodriguez y Baena<sup>1,2</sup>

**Abstract**— Telesurgical robotic systems provide a well established form of assistance in the operating theater, with evidence of growing uptake in recent years. Until now, the da Vinci surgical system (Intuitive Surgical Inc, Sunnyvale, California) has been the most widely adopted robot of this kind, with more than 6,700 systems in current clinical use worldwide [1]. To accelerate research on robotic-assisted surgery, the retired first-generation da Vinci robots have been redeployed for research use as “da Vinci Research Kits” (dVRKs), which have been distributed to research institutions around the world to support both training and research in the sector. In the past ten years, a great amount of research on the dVRK has been carried out across a vast range of research topics. During this extensive and distributed process, common technical issues have been identified that are buried deep within the dVRK research and development architecture, and were found to be common among dVRK user feedback, regardless of the breadth and disparity of research directions identified. This paper gathers and analyzes the most significant of these, with a focus on the technical constraints of the first-generation dVRK, which both existing and prospective users should be aware of before embarking onto dVRK-related research. The hope is that this review will aid users in identifying and addressing common limitations of the systems promptly, thus helping to accelerate progress in the field.

## I. INTRODUCTION

Robotic-Assisted Minimally Invasive Surgery (RMIS) has gained significant popularity in recent years due to its advantage of causing less tissue trauma and reducing hospitalization time for patients. Of all surgical robotic platforms available in the market, the da Vinci robot (Intuitive Surgical Inc, Sunnyvale, California) has dominated RMIS, with more than 10M operations having been performed on this platform in the past twenty years [2]. To further propel RMIS research on the da Vinci surgical platform, in 2012, an initiative was launched to repurpose retired first-generation da Vinci robots by converting them into da Vinci Research Kits (dVRKs) [3], which consist of both a software package and hardware controllers. These dVRKs have been distributed to research institutions around the world. Until now, more than 40 research groups across 10 countries have benefited from the initiative, forming a thriving dVRK research community, with more than 250 peer-reviewed research papers on the dVRK having been published.

\*These authors contributed equally to the work.

Zejian Cui: [zejian.cui19@imperial.ac.uk](mailto:zejian.cui19@imperial.ac.uk)

João Cartucho: [j.cartucho19@imperial.ac.uk](mailto:j.cartucho19@imperial.ac.uk)

<sup>1</sup>The Hamlyn Centre for Robotic Surgery, Imperial College London, London SW7 2AZ, UK

<sup>2</sup>Mechatronics in Medicine Lab, Department of Mechanical Engineering, Imperial College London, London SW7 2AZ, UK

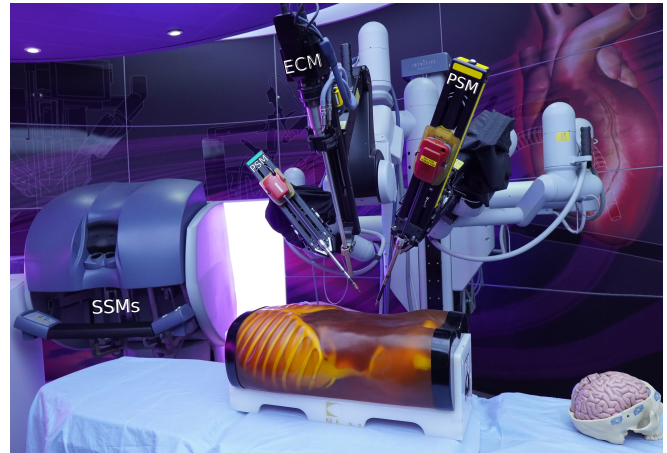


Fig. 1. The first-generation dVRK with two surgeon-side manipulators (SSMs), three patient-side manipulators (PSMs) and one endoscopic camera manipulator (ECM)

Although research conducted on the dVRK branches into different topics, almost all of these works deal with data directly provided by the dVRK, containing kinematics, image and system information, as described in [2]. These data are arguably the best source of information available to users who wish to keep abreast of the latest state of the robot as they develop novel research on the dVRK. However, blind reliance on the fidelity of these data can be risky. The inaccuracy of dVRK data partly arises from the intrinsic design of the telesurgical platform, and partly from limitations of the factory calibration of the hardware, which has only been improved upon in recent years thanks to a growing interest in the field and the many contributions of researchers working with the dVRK. Awareness of these seminal works should be heightened within the dVRK community, while thus far, there only exists one review paper about the dVRK [2], which focuses on categorizing research in terms of relevant high-level application areas.

To fill in this gap, in this paper we aim to analyze the technical constraints identified within the first-generation dVRK so that users can be aware of these shortcomings and take corrective actions to overcome these, in pursuit of better research outcomes. We have based this review paper on existing literature and the feedback from researchers of the dVRK community, as detailed in Sec. II.

According to our initial research, there is a general consensus among dVRK users about the inaccuracies within the dVRK kinematics and image data, which necessitates studies on essential kinematics and camera calibration. Allowing for

the fact that hand-eye calibration is an indispensable part of teleoperation, the accuracy of which can have a profound impact on subsequent task implementations, we also stress technical concerns about hand-eye calibration. Additionally, we address potentiometer calibration after weighing up its substantial impact on the performance of autonomous task execution, even though, thus far, to the best of our knowledge, there is no literature about this topic. In the following sections, kinematics calibration, hand-eye calibration, potentiometer calibration and camera calibration problems are detailed in Sec.III, VI, IV and V, respectively. For each of these, we first identify the nature of the problem, then distill prevailing schools of thought concerned on how to tackle them, and finally we point out potential paths for better addressing these problems, placed in the dVRK framework. A discussion is presented in Sec.VII, where we summarize the main findings, and other miscellaneous dVRK technical issues uncovered during the survey.

## II. SEARCH METHODOLOGY

### A. Literature review

To identify the most significant technical constraints that have hampered research progress with the dVRK, we first looked into all papers on the dVRK across different research application areas and gathered common issues that researchers claimed to have affected task performance, taking advantage of the latest review paper on the subject, which encompasses papers on the dVRK from year 2014 to 2021 [2]. We also employed the keywords “dVRK” paired with “error”, “calibration” and “task performance” individually on Google Scholar, ScienceDirect and RefWorks and included all relevant findings in this survey.

### B. Consultation with researchers

1) *Leading researchers:* Having gathered a list of technical issues reported by dVRK users in the literature, we consulted with leading researchers of the dVRK community (Mr. Anton Deguet and Dr. Simon DiMaio) about whether these issues have been discovered under the first-generation dVRK and thus can be regarded as universal.

2) *Peer researchers:* We designed a questionnaire featuring selected technical issues, which we deemed to be “universal”, and had it distributed within the dVRK community<sup>1</sup>. The design of the questionnaire and a summary of user responses are detailed in Appendix I.

## III. KINEMATICS CALIBRATION

A well-acknowledged technical issue within the dVRK is its time-variant inaccurate forward kinematics, which induces a poor surgical tool tip pose estimation, reflected by a discrepancy in readings of up to 1.02mm [4], [5] between the alleged tip pose streamed from the dVRK and the ground truth observed by an external means. This discrepancy, termed positioning error in this paper, has exceeded the sub-millimetre positioning accuracy required by

RMIS [6], having repercussions on diverse research projects implemented on the dVRK. Note that kinematics errors exist in all individual joints of the dVRK; however, it is in the end effector position that the discrepancy caused by the kinematics errors is most evident, because of the accumulation of errors across all the joints.

### A. Need for kinematics calibration

For research on surgical subtask automation, accurate positioning is required, such as in debridement [7], suturing [8], [9], needle extraction [10], and peg transfer [11]–[13], where a surgical instrument is required to move to the target grasping point. For research on the development of Active Constraints (AC, also known as virtual fixtures) via the dVRK, positioning accuracy will affect the proximity query result, and therefore the entire enforcement stage of the AC implementation [14]–[17]. Even for research on image-based hand-eye calibration [18], inaccurate surgical tool kinematic data leads to an inaccurate 3D pose estimation and thereby affects the accuracy of its back-projected 2D pixel position, which serves as a criterion for evaluating different hand-eye calibration strategies.

### B. Challenges in accurate positioning

1) *Cable-driven effects:* The dVRK surgical platform, which is a cable-driven robotic system [19], has an inherent issue in obtaining accurate joint readings from the built-in encoders. This is because all encoders are mounted adjacent to actuators but distant from joints, where the actual joint angles are estimated through transmission kinematics [20]. However, unknown parameters such as pulley-cable friction and nonlinearities [21] contribute to inaccuracies in joint angle estimation, which result in dVRK joint encoder readings not reflecting actual joint positions. These errors in joint space subsequently affect the estimated end effector tip pose through Denavit-Hartenberg (DH) parameters and forward kinematics. The DH parameters here only concern active joints.

2) *Inaccurate kinematic parameters:* The wear and tear of the dVRK not only induces cable slack [22], which affects the parameters in the transmission kinematics, resulting in inaccurate joint encoder readings, but also entails instrument damage which leads to additional inaccuracies in the DH parameters of the forward kinematics, causing the estimated tool-tip pose to stray from the actual one. Another significant source of error is due to the potentiometers, as explained in detail in IV.

3) *Other non-kinematic error sources:* Other than kinematic factors, non-kinematic factors such as external forces exerted onto the tool shaft and backlash between the tool shaft and cannula, illustrated in Fig. 2, would also contribute to an overall erroneous end-tip pose estimation [23], [24]. These external forces will affect the tension of cables associated to tool tips, leading to an error in transmission kinematics. In addition, they will also cause a displacement of the tool shaft because of compliance, as illustrated in Fig. 3; this displacement is not detectable by the encoders [23], and thus positioning error accumulates.

<sup>1</sup>More info at: <https://jhdvrk.slack.com>

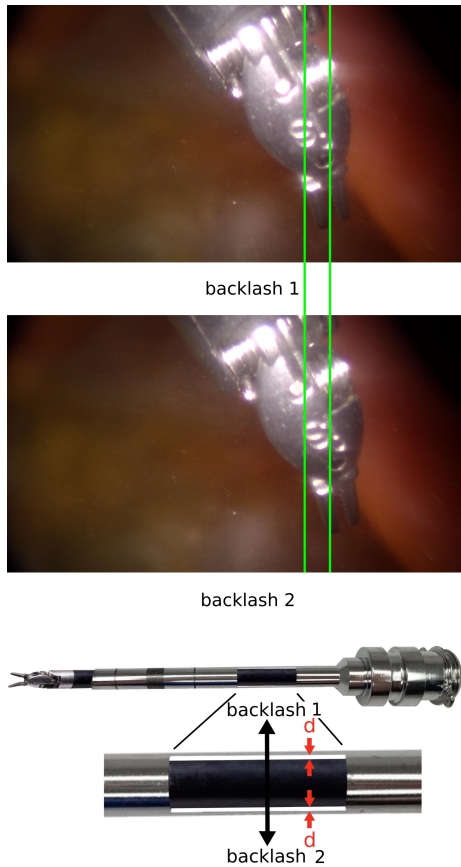


Fig. 2. As illustrated in red (d), there is a gap between the surgical instrument’s shaft and the cannulae. Therefore, the point of contact between this shaft and the cannulae can change. This phenomenon is known as backlash. Backlash causes the tool to move without being reflected by the encoder readings. Backlash is evident in the figure, illustrated in green, where the “S” character of Intuitive’s logo has moved to a different location, given that the camera is static.



Fig. 3. Compliance of the surgical instrument. Similarly to backlash, this compliance is not reflected on encoder readings, and hence not reflected in the joint position estimation.

### C. Methods for kinematics calibration

In this subsection, we gather methods that are available in the current literature to compensate for dVRK positioning errors, and we classify these in term of the specific error source they are aiming to address, namely encoder readings, kinematics parameters and other non-kinematic factors. Methods within the former category aim to estimate the current joint positions, whereas those within the latter category aim to force the system to reach the desired position.

1) *Compensation for encoder readings:* For methods that aim to compensate encoder reading inaccuracies, we further divide these into two classes: encoder reading calibration and tool-tip pose estimation. Both classes recognize the nature of the erroneous encoder readings; however, the former focuses

on finding actual joint positions in joint space, while the latter places an emphasis on finding the actual tool-tip pose.

a) *Encoder reading calibration:* Methods for calibrating encoder readings follow a standardized pipeline that involves modeling, measurement and parameter fitting [25]. To start with, a model parameterized by  $\eta$ , which maps encoder readings obtained from transducers to actual joint positions, is conceived. In the subsequent measurement stage, robot joints are commanded to move to designated positions while ground truth values of joint positions are gathered via an external measurement system. Finally, different fitting models, which can be either linear or nonlinear, are adopted to estimate the  $\eta$  that best fits designated joint positions to ground truth measurements.

Major differences between methods in this class are the selection of calibration models and approaches to measuring the ground truth. Following the calibration paradigm, Huang *et al.* [24] started off using a linear model to map encoder readings to ground truth values gathered by tracking infrared markers mounted on the end effector with an optical camera system, namely the NDI Polaris [26]. To better estimate and account for the nonlinear nature of cable stretch and friction, Hwang *et al.* adopted a data-driven approach to training deep learning models that map collected readings to actual joint positions gathered in a similar way [11].

With access to the forward kinematics of a system, which maps encoder readings to actual joint positions, one can find an inverse calibration model that generates encoder commands that drive a robot to reach desired joint positions. Such methods lend themselves to implementing automatic surgical tasks [7], [11], [13], where a surgical tool is requested to accurately move to a given pose. However, there are several drawbacks in model-based calibration methods. First, as pointed out in [27], model-based calibration methods are hard to implement outside of a lab setting because of the need for additional sensors, calibration objects, and tedious data collection procedures. Second, an invariant calibration model fails to accurately reflect a dynamic mapping relationship as, e.g., the cable wears and tears over time. Third, adopting model-based approaches requires greater efforts to find the ground truth values of the joint positions for the 7-DoF da Vinci robot. Although there exists numerical inverse kinematics solvers for the SSMs and the PSMs, running at 1.5 KHz<sup>2</sup>, there is a lack of kinematics parameters identification procedures. Hence the fidelity of the nominal kinematics parameters for the disposable surgical instruments relies on the quality of manufacture, which, however, can be imprecise. Furthermore, the accuracy of the ground truth measurement provided by an external sensor, a depth-sensing camera for example, is also inevitably affected by the inaccuracy in camera parameters and the hand-eye transformation matrix, as analyzed in V and VI, respectively.

b) *Tool-tip pose estimation:* In addition to generating joint commands that drive a robot to reach the desired state, there are times when the actual tool-tip pose in the

<sup>2</sup>One numerical solver is provided at <https://github.com/jhu-cisst/cisst>

Cartesian space is more important, not least in situations where the minimum distance between a surgical tool and its surrounding geometry needs to be precisely returned. Typically, pose estimation methods utilize both kinematic and image data. First, a tool-tip pose inferred from encoder readings and the forward kinematics of the system serves as an initial guess, then another tip-pose is obtained out of 2D image data, serving as an observation. By combining the initial guess with the observed result, a more accurate tip-pose can be estimated.

Literature on dVRK-based tip-pose estimation varies mainly in terms of the choice of computer vision-based strategy for 3D tool-tip pose reconstruction and the selection of 3D pose estimators. Moccia *et al.* [14] proposed a binary segmentation-based 3D pose reconstruction strategy, where the 3D pose is reconstructed by first finding the tool-tip on both binary-masked images captured by a stereoscopic endoscopic camera, followed by triangulation. Without relying on a pre-trained tool segmentation model, Richter *et al.* [27] simply attached markers on tool joints and used these to correct the discrepancy between detected markers and back-projected joint positions on the image plane. Ye *et al.* [28] leveraged the CAD model of surgical instruments and conducted keypoints matching between online generated part-based templates and 2D images to measure the accuracy of the tool pose estimated from the dVRK. For 3D pose estimators, Extended Kalman Filters (EKFs) and Particle Filters have also been proposed [14], [27]–[29].

Given that most pose estimation methods exploit image data, the accuracy of 3D tool tip pose reconstruction is highly reliant on the quality of gathered images and the robustness of the pose reconstruction strategies. Therefore, image-dependent pose estimation methods are likely to deliver subpar performances in situations when the camera fails to capture surgical tools in full view or when blood-stained surgical tools fail to be detected and segmented, which are likely to occur within a surgical setting. In addition, the accuracy of image-based methods is also dependent on the precision of camera calibration methods, which are elaborated in Sec. V

In theory, estimator-based pose estimation methods can also be utilized for estimating joint positions  $j_1$ – $j_6$ ; however, as suggested by the findings in [27], an increase in the number of unknown variables increases the system nonlinearity and yields less accurate estimation results. It is also worth mentioning that spurious feature detection and false feature correspondences could also impair the estimation accuracy. Finally, the computational efforts for feature detection in some pose estimation methods are not negligible, such as edge detection in [27], tool segmentation in [14], which may render these methods less suitable to be incorporated in a modern clinical setting.

2) *Compensation of kinematics parameters:* The dVRK returns the surgical tool-tip pose based on both joint readings and the system’s forward kinematics, which means that calibrated encoder readings are a necessary but not sufficient condition for accurate tool tracking. The nominal dVRK DH

kinematics parameters assume that adjacent joint axes are orthogonal and without intersections [30]; however, these assumptions fail to hold up in a real clinical setting. In [24], authors have shown that there are intersections between adjacent tool axes and additional calibration needs to be conducted on the dVRK kinematics parameters to account for these. A standardized approach to calibrating these kinematics parameters relies on external sensors attached onto the joints, which are used to track complex movements as the robot is driven along random trajectories. This process enables joint axes vectors to be estimated, from which corresponding kinematics parameters can be extracted using linear algebra.

3) *Compensation of external force factors:* Research on this topic stems from the observation that the dVRK surgical tools are intrinsically mildly compliant and can deflect under external environmental forces. However, such deflection is undetectable by joint encoder readings, and hence the end tip pose estimation can suffer further. One way to compensate for the external force factors is through estimation of the external forces/torques. In [23], it is assumed that the deflection on the slender dVRK tool shaft is the most salient one, and a compliance model is established by connecting the deflection of the surgical tool in the lateral direction with encoder readings for the first two joints. Alternatively, Yilmaz *et al.* [31] proposed a neural network-based method for estimating inverse dynamic torques such that the external forces/torques can be obtained by subtracting the inverse dynamic torques from the joint torque readings. In both cases, the estimation of the external force factors builds upon the knowledge of system dynamic equations, which are supported by studies on dVRK dynamics modeling, such as [32]. Another way to address the external force factors is to directly monitor the forces/torques by attaching force sensing devices without further modifications of the original instruments, as proposed in [33], although there is still great room for improvement in sensing accuracy.

Apart from joint-level calibration approaches, it is also possible to interpret and account for positioning errors through hand-eye calibration, as seen in [27], [28], [34], where the positioning error is compensated by conducting hand-eye calibration on the fly. Methods falling within this category are discussed in Sec. VI.

#### IV. POTENTIOMETER CALIBRATION

Potentiometer calibration is part of the kinematics calibration process since the potentiometers are used to calculate the joint encoder readings, which serve as input to the kinematic parameters. In this paper, we present it as a separate section since we explore potentiometer calibration in detail. To the best of our knowledge, so far there is no explicit mention of potentiometer calibration in the context of dVRK literature.

The dVRKs are conjured out of the retired first-generation da Vinci systems which were previously used for commercial purposes. The original potentiometer calibration values, for the commercial version of the system, are not always available, and even when available, the calibration values were

obtained using a different reference Voltage. This reference Voltage is required to get digital readings, using analogue-to-digital converters (ADCs), since the first-generation da Vinci systems use analogue potentiometers [3]. The ADCs and other electronic components in the dVRK are different from those employed on the commercial version of the system, and therefore, a different reference Voltage is required. Additionally, the potentiometer calibration values tend to change over time, and hence recalibration is necessary.

All manipulators of the dVRK robot consist of a series of individual joints that are controlled by actuators. Each actuator is coupled with an encoder (directly attached to the actuator) and a potentiometer (indirectly attached to the actuator), both of which can be used to measure the actuator position ( $\beta$ ), which can be further converted to the joint encoder reading by multiplying with given coupling matrices and gear ratios. According to Walker *et al.* [35], the actuator position ( $\beta$ ) can be returned from both the potentiometer and the encoder using the following equations:

$$\beta_P = k_P \times P + b_P \quad (1)$$

$$\beta_E = k_E \times E + b_E \quad (2)$$

$$\beta = \beta_E \approx \beta_P \quad (3)$$

where  $P$  is the Voltage reading from the ADC converter of the potentiometer;  $k_P$  is a scale factor;  $b_P$  is an offset for the Voltage readings;  $E$  represents the reading from the encoder counter;  $k_E$  is the radians per encoder count, which is known a priori; and  $b_E$  is an offset for the encoder.

For the first-generation dVRK, actuator positions measured from the potentiometer  $\beta_P$  are usually noisier than the ones measured from the encoder  $\beta_E$ , and hence the robot selects  $\beta = \beta_E$ . Note that  $b_E$  needs to be accurately calibrated for each actuator, to ensure an accurate tool-tip pose estimation result. As explained next,  $b_E$  is only accurate if  $k_P$  and  $b_P$  are also accurate, hence the need for potentiometer calibration.

#### A. Need for potentiometer calibration

1) *Encoder calibration:* The first-generation dVRK robots use relative encoders, which require calibration of  $b_E$  every time the dVRK controllers are powered ON, via the equation below (resulting from Eqs. 1, 2, 3):

$$b_E = -k_E \times E + k_P \times P + b_P \quad (4)$$

It is shown that the accuracy of  $b_E$  depends on the accuracy of  $k_P$  and  $b_P$ , which are the outputs of the potentiometer calibration. This explains why appropriate potentiometer calibration is an important prerequisite towards obtaining an accurate value of  $b_E$ .

2) *Safety:* With the assumption that  $\beta_E \approx \beta_P$ , the dVRK potentiometers work as redundant sensors in tandem with their encoders to ensure task safety. In the unlikely event that the encoders fail to reflect the actual actuator positions, which may cause the manipulator to keep moving towards an unachievable target position, the discrepancy  $|\beta_E - \beta_P|$  can be used to monitor whether the encoders are functional and to decide whether the robot needs to be stopped in order to avoid potential damage. For this, accurate potentiometer calibration is needed.

#### B. Challenges in potentiometer calibration

The goal of potentiometer calibration is to accurately estimate  $k_P$  and  $b_P$ . The scale factor  $k_P$  of a potentiometer is calibrated by measuring  $E$  and  $P$  at different time steps ( $t_1$  and  $t_2$ ), during which a large actuator motion is enforced, as follows:

$$k_P = k_E \frac{E_{t_2} - E_{t_1}}{P_{t_2} - P_{t_1}} \quad (5)$$

However, it should be noted that the assumption of a constant value for  $k_P$  may not hold true in practice, as real-world potentiometers have inherent non-linearity (i.e., Eq. 1 assumes linearity). Potentiometer non-linearity refers to the deviation of the potentiometer's output from a linear relationship with the input rotation of the potentiometer's shaft, along its range of motion. This is due to manufacturing errors, physical wear etc.

Besides  $k_P$ , the offset  $b_P$  of a potentiometer is hard to estimate. It is required to place the actuator at the zero-position, where  $\beta_P = 0$  so that we have  $b_P = -k_P \times P$ , via substitution in Eq. 1. The challenge here is that it is hard to determine the exact zero-position of an actuator, especially in the context of this research system.

#### C. Methods for potentiometer calibration

The methods described in this section can be implemented preoperatively.

1) *Potentiometer nonlinearity calibration:* Common solutions for correcting potentiometer nonlinearity are based on Bezier or polynomial models [36]. These models are used to create a mathematical function that can correct for non-linearities in the potentiometer output. Another approach is to create a lookup table. This involves measuring the output of the potentiometer with its shaft rotated to various input angles and creating a lookup table that maps input angles to corrected output values. Different types of interpolation (e.g., linear, or spline) can be used to estimate corrected output values for input angles that were not measured directly. The lookup approach is also generally faster since it reduces the need for complex calculations.

2) *Calibrating potentiometer offsets ( $b_P$ ) on a PSM:* Each PSM consists of a total of seven actuators, hence there are seven potentiometer offsets that need to be calibrated. The offsets for the last four potentiometers ( $b_{P4}$ ,  $b_{P5}$ ,  $b_{P6}$ ,  $b_{P7}$ ) are usually estimated first with a simple process, whereas calibrations of the offsets  $b_{P1}$ ,  $b_{P2}$  and  $b_{P3}$  are harder

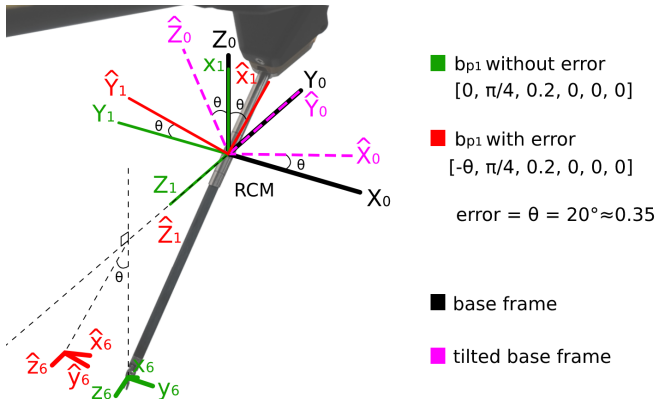


Fig. 4. Illustration of the impact of the offset error ( $b_{P1}$ ) in the potentiometer of the first joint of the PSM. For the sake of illustration, an error of  $20^\circ$ ,  $0.35 \text{ rad}$ , is added onto the first joint of the PSM (1), causing a rotation of  $20^\circ$  at the tool-tip position ( $\hat{\delta}$ ). With this error, hand-eye calibration results in a transformation from the camera frame to the tilted base frame illustrated in pink ( $\hat{0}$ ). Hand-eye calibration rectifies the error  $b_{P1}$  because the erroneous rotation is compensated by the tilted frame, which is rotated in the opposite direction of the error. Note that joint 3 is in  $m$  instead of  $rad$ .

to implement but have a more significant impact on the accuracy of tip-pose estimation, since the first three actuators are connected to the largest links of the PSM.

The solution for calibrating  $b_{P4}$  to  $b_{P7}$  is to attach a rigid holder to the four wheels that control the movement of a surgical tool. When the wheels are connected to the holder, the last four actuators are held at the zero-position, where  $b_P = -k_P \times P$ . The holder can either be 3D printed or laser cut, using models that are available on the dVRK's Wiki [37].

As illustrated in red in Figure 4, regarding the 1<sup>st</sup> potentiometer of the PSM, a constant error in the offset (a wrong  $b_{P1}$  value after powering ON the PSM), would cause a constant rotation error around the 1<sup>st</sup> joint. A mathematical description of this effect is the next sub-section (Sec. IV-D). The tool-tip pose would suffer from the same rotation error. Here, we assume that the scale ( $k_{P1}$ ) has been correctly calibrated as explained in Sec. IV-B. If the user's goal is to obtain the tool-tip pose relative to the camera, then the user usually needs to perform hand-eye calibration after powering ON the PSM. This hand-eye calibration, in the presence of an offset error in  $b_{P1}$ , would result in a transformation from the camera to a tilted base frame, shown in pink in Figure 4. Errors in calibration of offsets  $b_{P1}$  introduce a constant error in the joint positions (since  $b_{E1}$  in Equation 4 depend on  $b_{P1}$ ). Therefore, the tool-tip cartesian pose relative to the PSM's base will be erroneous. However, if hand-eye calibration is used, the tool-tip cartesian pose relative to the camera will be correct, since the hand-eye calibration has already compensated for the errors by outputting a transformation from the camera frame to a tilted PSM base frame.

The compensation approach mentioned above does not apply to the 2<sup>nd</sup> joint or the 3<sup>rd</sup> joint, as further detailed in the next sub-section (Sec. IV-D). The reason is that no

constant transformation matrix exists when the potentiometer offsets ( $b_{P2}$  and  $b_{P3}$ ) are not well calibrated. To quantify the impact of the offset error  $e_2$ , resulting from an uncalibrated second potentiometer, we generated a reference trajectory for the end-effector with a total of 1,000 points, by interpolation of joint positions ranging between  $[-\frac{\pi}{5}, -\frac{\pi}{6}, 0.2m, 0, 0, 0, 0]$  and  $[\frac{\pi}{5}, \frac{\pi}{6}, 0.2m, 0, 0, 0, 0]$ . We then repeated generating the same trajectory but added  $e_2$  to the second joint position to simulate an end-effector trajectory under the presence of an erroneous  $b_{P2}$ . We found that each  $1^\circ$  increment in  $e_2$  resulted in an additional RMSE of  $3.4 \pm 0.1 \text{ mm}$  between the reference and the actual trajectory. This error was observed on surgical instruments including Large Needle Driver (400006), ProGrasp Forceps (400093), Cadière Forceps (400049) and Maryland Bipolar Forceps (400172). If  $e_2$  reaches  $3^\circ$ , the discrepancy in the end effector trajectory will be slightly over  $1 \text{ cm}$ , which highlights the need for calibration of the second potentiometer offset. Currently, there is no existing technique for accurate calibration of the second potentiometer's offset ( $b_{P2}$ ).

Conversely, a recent technique has been proposed [38] to calibrate  $b_{P3}$ . The general idea of this technique is to place a reference point of the kinematic chain, that is easy to visually track, at the remote centre of motion (RCM). If  $b_{P3}$  is accurately calibrated, then this point should not move when located on the RCM, independently of joint 1 and joint 2 actuator positions. Specifically, this technique uses as a reference point the axis between the coordinate frame of joint 4 ( $O_4$ ) and joint 5 ( $O_5$ ), which is located at the beginning of a surgical instrument's wrist. If a camera is recording this reference point, then its pixel position should remain static. Therefore, if the reference point moves when the position of joint 1 or joint 2 changes, then it means that the offset ( $b_{P3}$ ) is still not accurately calibrated. This technique tries different insertion positions until the motion of the reference point is minimized when the reference point is closest to the RCM. Once the insertion position is known,  $b_{P3}$  can be directly measured. The results of our questionnaire indicate that 60% of the respondents were not aware of this technique used to calibrate the third potentiometer offset.

#### D. Compensation for potentiometer offset errors using hand-eye calibration

As previously mentioned, potentiometer offset errors result in deviations between the actual joint frames ( $0$ ,  $1$ , and  $2$ ) and the virtual frames ( $\hat{0}$ ,  $\hat{1}$ , and  $\hat{2}$ ). It is posited that the accuracy will not be affected by the potentiometer offset error in a certain joint, provided that there exists a constant homogeneous transformation matrix ( ${}^{\hat{0}}T_0$ ) from the virtual base frame ( $\hat{0}$ ) to the actual base frame ( $0$ ), so that the virtual joint frame can be realigned to the actual joint frame despite the offset errors in all the preceding joints. The realignment is achieved by combining a constant transformation ( ${}^{\hat{0}}T_0$ ), with the result of hand-eye calibration ( ${}^cT_{\hat{0}}$ ) and any virtual transformations, *e.g.*  ${}^{\hat{0}}T_{\hat{1}}$ .

An example of how hand-eye calibration compensates for the offset error in the first joint is shown in Fig. 5. The

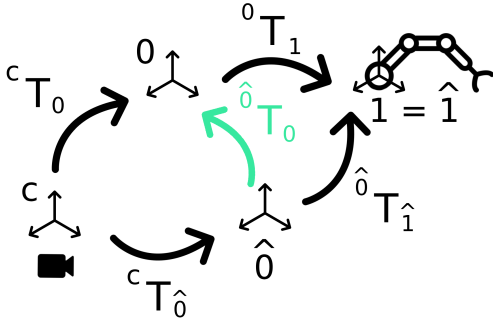


Fig. 5. This figure illustrates how hand-eye calibration compensates for the offset error in the first potentiometer with the transformation ( ${}^cT_{\hat{0}}$ ), from the camera frame (c) to the virtual base frame ( $\hat{0}$ ). The new hand-eye calibration matrix  ${}^cT_{\hat{0}}$  post-multiplied by  ${}^{\hat{0}}T_{\hat{1}}$  results in the realignment for joint frame 1 ( $1 = \hat{1}$ ), and hence the following joint frames will be aligned if no offset errors exist in the rest of the joints. Hand-eye calibration is capable of compensating for the potentiometer offset errors only when there exists a constant transformation matrix ( ${}^{\hat{0}}T_0$ ) from the virtual base frame ( $\hat{0}$ ) to the actual base frame (0). It is later proved that hand-eye calibration only addresses the offset error in joint 1, but not in joint 2

goal is to find a constant transformation ( ${}^{\hat{0}}T_0$ ), so that after post-multiplying  ${}^{\hat{0}}T_{\hat{1}}$  to the new hand-eye calibration matrix ( ${}^cT_{\hat{0}}$ ), realignment is achieved for joint 1 ( $1 = \hat{1}$ ). After that, all subsequent joint coordinate frames will be aligned if no potentiometer offset errors exist in the remaining joints. For the first joint, the following equation is obtained:

$${}^{\hat{0}}T_0 = {}^{\hat{0}}T_{\hat{1}=1}({}^0T_1)^{-1} \quad (6)$$

The same logic also applies when the offset errors in both joint 1 and joint 2 are considered:

$${}^{\hat{0}}T_0 = {}^{\hat{0}}T_{\hat{1}}{}^{\hat{1}}T_{\hat{2}=2}({}^1T_2)^{-1}({}^0T_1)^{-1} \quad (7)$$

in which case, realignment is achieved for the second joint frame ( $2 = \hat{2}$ ). Next, the analytical solution for  ${}^{\hat{0}}T_0$  is found for the above three cases, with the modified DH transformation matrix being:

$${}^{n-1}T_n = \begin{bmatrix} c(\theta_n) & -s(\theta_n) & 0 & a_{n-1} \\ s(\theta_n)c(\alpha_{n-1}) & c(\theta_n)c(\alpha_{n-1}) & -s(\alpha_{n-1}) & -d_n s(\alpha_{n-1}) \\ s(\theta_n)s(\alpha_{n-1}) & c(\theta_n)s(\alpha_{n-1}) & c(\alpha_{n-1}) & d_n c(\alpha_{n-1}) \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where  $c()$  and  $s()$  represent  $\sin()$  and  $\cos()$ , respectively, and the DH parameters for the dVRK PSMs being:

$n$	$\alpha_{n-1}$	$a_{n-1}$	$d_n$	$\theta_n$
1	$\frac{\pi}{2}$	0	0	$q_1 + \frac{\pi}{2}$
2	$-\frac{\pi}{2}$	0	0	$q_2 - \frac{\pi}{2}$

After solving Eq. 6 by substituting in the DH parameters above ( $q_1 = j_1$  into  ${}^0T_1$  and  $q_1 = j_1 + \Delta_1$  into  ${}^{\hat{0}}T_{\hat{1}=1}$ ), the following constant transformation is obtained:

$${}^{\hat{0}}T_0 = \begin{bmatrix} c(\Delta_1) & 0 & -s(\Delta_1) & 0 \\ 0 & 1 & 0 & 0 \\ s(\Delta_1) & 0 & c(\Delta_1) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where  $\Delta_1$  is the offset error in the first potentiometer (a constant value). This transformation represents a pure rotation around the y-axis of the actual base frame (0) by  $-\Delta_1$ .

Now, assume that  $\Delta_1 = 0$  to single out the contribution of the offset error in the second potentiometer to  ${}^{\hat{0}}T_0$  and solve Eq. 7 (substituting  $q_2 = j_2 + \Delta_2$  into  ${}^{\hat{1}}T_{\hat{2}=2}$  and  $q_2 = j_2$  into  ${}^1T_2$ ), the following transformation is obtained:

$${}^{\hat{0}}T_0 = \begin{bmatrix} c(j_1)^2 + c(\Delta_2)s(j_1)^2 & s(\Delta_2)s(j_1) & -((-1 + c(\Delta_2))c(j_1)s(j_1)) & 0 \\ -s(\Delta_2)s(j_1) & c(\Delta_2) & c(j_1)s(\Delta_2) & 0 \\ -((-1 + c(\Delta_2))c(j_1)s(j_1)) & -c(j_1)s(\Delta_2) & c(\Delta_2)c(j_1)^2 + s(j_1)^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Notice that, when there exists an offset error in the second potentiometer ( $\Delta_2$ ),  ${}^{\hat{0}}T_0$  is not constant because it depends on the value of the first joint ( $j_1$ ).

Therefore, it is concluded that a single hand-eye calibration only compensates for the offset error in the first potentiometer ( $\Delta_1$ ) but not for the errors in the second potentiometer ( $\Delta_2$ ). It also implies that, in the presence of potentiometer offset errors in joint 2, a non-constant hand-eye calibration matrix needs to be updated on the fly.

#### E. Evaluate potentiometer calibration accuracy

The accuracy of potentiometer calibration can be evaluated by configuring a manipulator at multiple positions and examining if  $\beta_E \approx \beta_P$  holds true for each joint across all joint configurations.

## V. CAMERA CALIBRATION

In this section, different aspects that need to be taken into consideration during the calibration of the endoscopic camera on the dVRK are discussed. It is assumed that a conventional computer vision method for camera calibration [39] is used, which relies on the pinhole camera model [40] and a planar checkerboard with a known geometrical pattern.

#### A. Need for accurate camera calibration

Accurate camera calibration ensures the fidelity of algorithms that take endoscopic images as input for estimating 3D spatial information. Therefore, camera calibration plays a crucial role in image-based clinical applications such as image-guided surgery and autonomous robotic surgery, since these applications rely on algorithms that process 3D data. Such algorithms include hand-eye calibration, 3D reconstruction, 3D tracking, camera or instrument pose estimation, constraining the motion of surgical instruments to avoid collisions, interacting with the tissue via visual servoing, and registration of 3D structures in the endoscopic scene using Augmented Reality (AR).

An image-based hand-eye calibration, for example, cannot be accurate unless the lens distortion and the camera intrinsic parameters are precisely calibrated. Similarly, a disparity-based 3D reconstruction cannot be accurate unless stereo images are accurately rectified, which requires not only accurate distortion and intrinsic parameters but also an accurate estimation of the geometric transformation from the left to the right stereo camera.

## B. Challenges in camera calibration

### 1) Camera calibration parameters are not constant:

One fact that needs to be taken into consideration is that for the cameras of the dVRK, the calibration parameters should not be considered constant. These parameters might change unintentionally (due to camera self-heating [41], vibrations or mechanical shocks [42]), or intentionally by manual adjustment of the camera's focus, for example, to improve the imaging quality of an anatomical target or to adjust the camera's field-of-view. When the focus of a stereo camera is adjusted, its left and right camera sensors move relative to the tip of the endoscope. This sensor motion would cause changes in camera intrinsic parameters, principal point and lens distortion parameters [43], which hence requires the camera calibration to be repeated. It has been proposed to calibrate the camera before surgery [44]; however, this is unrealistic because surgeons must first insert the endoscope inside the patient before adjusting the field-of-view or the focus to the target anatomy [43]. Changing camera parameters is also an open problem for deep-learning image-based algorithms. The reason behind it is that a model that is trained on a dataset captured by one camera cannot be directly applied to a different camera. Even with the same camera, when a surgeon changes its focus setting, the model may become inaccurate because of changes in the camera parameters.

2) *Human error during camera calibration:* Common sources of error include inaccuracies of the calibration target (e.g. due to the target being printed on the wrong scale or being attached to a non-flat surface), the use of a small number of images or corners, limited distribution of the corners over the field of view of the camera, inaccurate 2D pixel positioning of the detected corners, overexposed corners due to a powerful endoscopic light source, and the presence of vibrations/motion when pictures are taken [45], [46].

3) *The interlaced video problem:* Interlaced video is a technique that creates an illusion of having a camera with a higher frame rate by doubling the frame rate (here from  $25Hz$  to  $50Hz$ ) while maintaining the same bandwidth [47]. The downside of interlacing is that half of the image information at each timestamp is lost. The reason is that, in interlaced video, an image's pixels are captured at two sequential timestamps. At the first timestamp, the image's odd rows are captured and at the second timestamp, the even rows are captured, or vice-versa. Since the odd and even rows are captured at different timestamps, there can be significant artefacts when there is rapid movement from either tissue, the camera or surgical instruments. Figure 6, shows an example of this artefact in a video of a real *ex vivo* surgery. Interlaced video can have adverse effects on the accuracy of camera calibration. The reason is that, when the calibration target is in motion, it may be captured at two different poses, which are stored in the odd and even rows of a video frame. As a result, image keypoints can be distorted, which leads to inaccurate detection of keypoints in the calibration target.

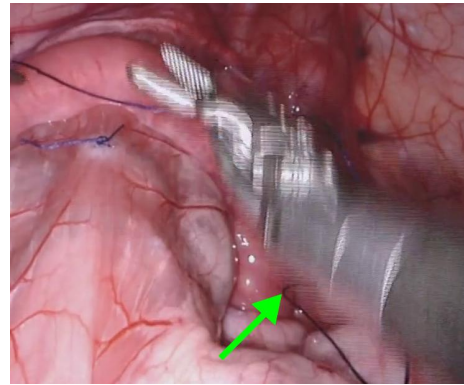


Fig. 6. A visual example of the interlaced video problem on an image from an *ex vivo* surgery. This image was taken from the SurgT challenge [48].

## C. Methods for camera calibration

### 1) Compensation for not constant calibration parameters:

In image-based robotic surgery, it is crucial that the robot is able to estimate the camera calibration parameters on-the-fly. Otherwise, the 3D information extracted from the endoscopic images will be inconsistent with the real dimensions of the surgical scene. Currently, there are three approaches to estimating the changeable camera calibration parameters during surgery, namely (a) using a calibration model, using (b) self-calibration, and (c) using a fixed-focus endoscope.

(a) *Using a calibration model:* Pratt *et al.* [43], created a calibration model which gathers the intrinsic and distortion parameters across the entire range of possible focal lengths of an endoscopic camera. The idea is that this model is first created preoperatively. Then during surgery, with the help of a single image of a calibration pattern attached to a surgical instrument, the focus setting can be estimated, and thereby its corresponding distortion and intrinsic parameters can be retrieved from the pre-created calibration model. The paper has also found that the geometric transformation from the left to the right camera is essentially constant for all the focus settings. One limitation of this paper is that it assumes that a calibration pattern is available intraoperatively. To address this drawback, Kalia *et al.* [49] made the assumption that given the fact that the focus is typically adjusted to the target anatomy, and that a surgical instrument is usually in contact with the target anatomy, the distance from the camera to the instrument can determine the focus setting. Then again, similar to Pratt *et al.* [43], this focus setting is used to extract the corresponding parameters from the pre-created calibration model. The second method presupposes that the surgical instrument's tip is in focus with the camera, which is not always the case. Moreover, it relies on accurate hand-eye calibration, as it involves estimating the instrument's pose relative to the camera. As a result, while it can be utilized in surgical procedures, the method's accuracy relies upon the soundness of these assumptions.

(b) *Using self-calibration:* In computer vision, self-calibration (or auto-calibration) estimates the camera calibration parameters using raw images without the requirement of a calibration pattern. Therefore, self-calibration methods can be used intraoperatively, as long as (1) these methods are able to estimate the calibration parameters on-the-fly, (2) are able to deal with non-rigid scenes, such as deforming soft-tissue. The degree of the deformation is proportional to the magnitude of the forces applied to the target anatomy. Self-calibration methods are mostly tailored for rigid world scenes. The earlier works solely focused on estimating the focal length [50], [51], but were later enhanced to estimate the focal length and principal point in addition to radial distortion [42], [52]. More recently, deep learning methods estimate all these parameters either via supervised learning [53], which trains with ground-truth calibration parameters, or via self-supervised learning [54], [55], which trains directly from raw videos. In addition to research on rigid world scenes, methods for self-calibration on non-rigid surgical scenes have also been developed. Some of these methods estimate focal length only [56]–[58]. A more recent method estimates focal length and principal point in addition to radial distortion for non-rigid scenes [59].

(c) *Using a fixed-focus endoscope:* Taking the changeable calibration parameters problem into consideration, and given that adjusting the focus during surgery is impractical, the newer da Vinci systems are now using fixed-focus cameras [60], also referred to as focus-free endoscopes. These new-generation endoscopes are still not available to dVRK systems and there is still little information about whether the calibration parameters on fixed-focus endoscopes are constant or not over time. Despite the lack of information regarding the stability of calibration parameters over time for these endoscopes, they remain a top choice for clinical translation, even if they change slightly over time. This is because any small changes in calibration parameters could be corrected either preoperatively or intraoperatively using self-calibration techniques.

2) *Compensation for human error during calibration:* If a checkerboard is used, for laboratory experiments, then it should be held at a static pose while capturing each of the camera calibration images. The reason is twofold. First, in the presence of checkerboard movement, inconsistent poses can be captured on the left and right stereo images if the cameras are not properly synchronized, and consequently, the estimation of the geometric transformation between two stereo cameras is inaccurate. The other reason is to avoid motion artefacts, which can significantly decrease the quality of camera calibration. Images containing motion artefacts should be pruned out of the calibration data, and fortunately, these motion artefacts can easily be identified thanks to the interlaced video technology that the dVRK endoscope camera is currently using. It is recommended to inspect the quality of the detected corners in the checkerboard images. This can be done by checking if the corners are correctly positioned between the squares of the checkerboard. It is also recommended to inspect the overall reprojection errors,

which can be used to identify possible outliers [61]. A common approach for verifying lens distortion parameters is to check if straight lines in the real-world are projected into straight lines in the 2D image [62]. Another technique for verifying stereo extrinsic and intrinsic parameters involves checking if keypoints are row-aligned in the rectified images when the left and right images are stacked horizontally [62].

3) *Compensation for interlaced video:* Many methods have been proposed for intra-frame deinterlacing. A popular solution for keeping the original resolution and perform real-time removal of this artefact is to use every other row (i.e., either the even rows or the odd rows) and interpolate the pixels for missing rows. There are other more complex algorithms for deinterlacing, such as recent deep learning-based approaches [63], [64]. Note that no matter how complex the algorithm is, the loss of image information at any given timestamp is inevitable. To avoid any possible impact of interlaced video on camera calibration, best practice is to maintain a static position for the calibration target during image capture. By doing so, the complete raw image from the dVRK can be utilized directly without requiring any deinterlacing. Moreover, holding the target stationary ensures that there is no motion blur, and any potential synchronization discrepancies between the left and right cameras do not affect the calibration precision. Nevertheless, it should be noted that maintaining a static position for the calibration target during image capture is only feasible in a controlled laboratory setting. In real-world scenarios, such as during intraoperative camera self-calibration, interlaced video remains a concern as it can potentially affect calibration precision. Therefore, deinterlacing may be necessary to utilize the complete raw image from the dVRK for self-calibration methods.

## VI. HAND-EYE CALIBRATION

Hand-eye calibration has been widely researched in the area of robotics; it aims to find the transformation from the camera frame to the robot frame. There exists some ambiguity in the definition of a hand-eye calibration problem in the dVRK community, with some researchers defining it as finding the transformation from the camera frame to the frame of the robot arm to which the camera is attached [65], [66], while most other researchers place their focus on finding the transformation from the camera frame to the base frame of another robot arm to which a surgical tool is attached. Although for a generalized hand-eye calibration problem, the robot frame is unspecified, in the remaining part of this section, we define the dVRK hand-eye calibration problem as per the latter school of thought. Hand-eye calibration is related to both kinematics calibration (including potentiometer calibration) and camera calibration. Therefore, it inherits the inaccuracy discussed in the previous sections.

### A. Need for hand-eye calibration

The importance of developing an accurate and time-efficient hand-eye calibration strategy is twofold. In a teleoperation scenario, an accurate hand-eye transformation matrix enables a better alignment between the surgeon's view and

the endoscopic camera view so that surgeons can see the surgical tools move in sync with their hand movements, making the teleoperation more intuitive. Through an operation, there are times when the robot arms need to be repositioned to adjust to different surgical tasks, and whenever its “set up joints” (SUJ) are relocated, the hand-eye calibration needs to be recomputed. Despite the fact that it is desirable to develop a hand-eye calibration method that is both accurate and time-efficient, in a teleoperation scenario, time-efficiency takes priority over accuracy because humans can compensate for a slight misalignment between the movements of their hands and the surgical tools. However, for automatic surgical tasks, the accuracy of hand-eye calibration is of paramount importance because all robot motion commands are generated out of visual information captured by the endoscopic camera, and these visual cues cannot be correctly perceived by the robot arm unless an accurate transformation matrix is provided.

### B. Challenges in dVRK hand-eye calibration

Although there is a rich literature on hand-eye calibration problems in a generalized robotic framework, the development of calibration strategies for the da Vinci robot, which is customized for MIS, is still challenging due to several intrinsic constraints of the system. As observed in [67]–[70], most widely-adopted hand-eye calibration methods require an external calibration object such as a planar checkerboard and external hardware. However, these methods are not applicable in a MIS environment where the working space is narrow and introducing external calibration objects is impractical. In addition, the Remote Center of Motion (RCM) constraint [71] of the patient-side robots limit the number of possible joint configurations available, which renders general calibration methods yet more challenging to implement in this context.

### C. dVRK hand-eye calibration methods

Most hand-eye calibration problems pivot on solving  $AX = XB$  [72] or its derivative  $AX = YB$  [73], where  $X$  represents the transformation matrix of interest, and  $A$  and  $B$  are matrices that can be deduced from kinematic readings and extrinsic calibration, such as the checkerboard calibration method [74]. To allow for the specific constraints of the da Vinci robot, researchers in the dVRK community build upon traditional hand-eye calibration methods to produce bespoke approaches that are better suited for the task. Although there are a handful of research works that rely on classical calibration objects (e.g. [75], [76]), most recent approaches have opted for more nuanced implementations. Pachtrachai *et al.* [66] proposed an approach which utilizes a tool-tracking algorithm so that the surgical tool itself functions as a calibration object *per se*. Cartucho *et al.* [77], [78] designed a cylindrical marker that can be attached around the shaft of a surgical tool, serving as a compatible calibration object. And Wang *et al.* [71] developed an algorithm which purely depends on the dVRK kinematic readings and image information captured by the endoscopic camera. Much as great progress has been made, the benefit of avoiding the

use of external calibration objects has come at the price of worse calibration accuracy.

To improve the calibration accuracy and the robustness of calibration algorithms against noise, research also focuses on designing advanced  $AX = XB$  solvers [34], [65], [66]. In cases where the hand-eye calibration needs to be performed in real-time to compensate for the dVRK kinematic positioning error, computational efficiency must also be considered [34]. Zhong *et al.* [18] also proposed an Interactive Manipulation-based (IM) computation method that enables fast extraction of the position and orientation term of hand-eye calibration, although they did not regard hand-eye calibration as  $AX = XB$  or  $AX = YB$  problem in their work.

## VII. DISCUSSION

The data information provided by the first generation dVRK is not flawless. Hence, without a proper understanding of the system’s constraints and calibration processes, the performance of surgical tasks can fall short of expectations. Due to page limits, this paper did not cover other technical limitations, such as inappropriate controller gains etc., which could also serve as valuable research directions for the dVRK community.

Considering the diverse research interests in the dVRK community, we aim to provide general guidance for dVRK users about the implementation of these calibrations based on the repercussions of these implementations on their research outcomes, as shown in Table.I. First, the need to conduct specific calibrations should pivot on the type of data being used, these data spanning image data, kinematics data, dynamics data etc. From this perspective, it is the researchers who need to be aware of the defect in these data and conduct corresponding calibrations. For example, in automating surgical subtasks, it is a common practice to use both image and kinematics data for localising the target position and commanding a surgical instrument to move towards the target, and therefore all of the four calibrations are needed to achieve an optimal performance. In studies involving AC, where accurate tip position is needed for enquiring the minimum distance between a surgical instrument and a moving tissue surface, it is necessary to perform kinematics calibration. Second, we suggest that the existence of a human element in a study largely diminishes the importance of conducting these calibrations, because a human can compensate for inaccuracies. Therefore, for researches that feature surgeon training, teleoperation assistance, where the accuracy requirement is less stringent, the implementation of these calibrations are left to the discretion of the researchers themselves.

Kinematics calibration aims to reduce the dVRK positioning error. We have narrowed error factors down to three categories, namely “cable-driven effects”, “inaccurate kinematics parameters” and “other non-kinematic sources”. In this paper, we evaluate these calibration methods based on their application scenarios, given that some are designed for laboratory trials while others are for clinical use. Overall,

TABLE I

THIS TABLE PROVIDES GUIDANCE ON THE IMPLEMENTATION OF EACH CALIBRATION BASED ON DIFFERENT RESEARCH OBJECTIVES. ‘✓’ REPRESENTS “USUALLY REQUIRED” AND ‘-’ REPRESENTS “USUALLY NOT REQUIRED”.

Requirements	<i>Kinematics calibration</i>	<i>Potentiometer calibration</i>	<i>Camera calibration</i>	<i>Hand-eye calibration</i>
Human teleoperation? (e.g. surgeon training, skill assessment and gesture recognition)	-	-	-	-
Requires accurate encoder position readings? (e.g. System simulation and modelling)	✓	✓	-	-
Requires extracting 3D information from images? (e.g. Augmented Reality, Depth Estimation, 3D Tracking)	-	-	✓	-
Requires accurate end-effector pose relative to the camera? (e.g. Autonomous Control)	✓	✓	✓	✓

the methods that rely on an external marker or a calibration object are not suitable for clinical application. Although the calibration methods that build upon a pre-training model have the potential to be translated into clinical practice, as they are implemented preoperatively, whether the effectiveness of these calibration methods will diminish over time remains to be assessed. Therefore, it is preferable to develop intraoperative calibration methods that take real-time information into consideration.

Potentiometer calibration is crucial to obtain accurate joint encoder readings, thereby indirectly impacting the accuracy of tip-pose estimation. Despite this, until now, the dVRK community had not yet investigated the inherent nonlinearities of potentiometers. One of the challenges that have been addressed is the calibration of the potentiometer offsets. These offsets are difficult to calibrate as their calibration requires users to position the actuators at their zero-position. For the PSMs, techniques have been developed to calibrate the offset of the last five potentiometers. However, calibrating the offset for the first two potentiometers remains an open problem. It is discussed in this paper that the first potentiometer offset error can be compensated for using hand-eye calibration, but this does not apply to the second potentiometer. Additionally, it has been identified that each 1° error on the second joint encoder reading results in a mean error of 3.4 mm in the end-effector position. Therefore, it is essential to develop a technique for calibrating the offset of the second potentiometer, which is still an open problem for the community. Furthermore, for the ECM and SSMs, there is still no technique to determine their exact zero-position for offset calibration.

Regarding camera calibration, obtaining accurate values of lens distortion, camera intrinsic parameters and the transformation between two stereo cameras is crucial for tasks that extract 3D information out of endoscopic images. With current endoscopic technology employed with the dVRK, calibrated parameters are not constant, mainly due to the dynamic adjustment of focus settings, which happens frequently in real surgeries. In the authors’ opinion, it would be preferable if dVRK researchers focus on developing methods to continuously update current camera parameters using natural surgical video without the aid of a calibration target that is unusable in real surgery. However, for laboratory experiments, a calibration target such as a checkerboard

can be used. Furthermore, another interesting observation is that instead of trying to deinterlace input images for deep learning models, one can choose to use only every other row (e.g., even rows). This is because most deep learning models currently require lower resolution images, and therefore one can use only half of the rows (without interlaced video issues) as a natural down-scaling technique.

Hand-eye calibration has been widely researched in the robotics area, but most of these developed methods are not suitable for an RMIS setting, where the task space is narrowly confined, and the accuracy requirement can be much higher. In a clinical setting, the da Vinci robots are already equipped with sensorized SUJs, which enable automatic updates of hand-eye transform. Although the inaccuracy is up to around 5 cm [79], it suffices for teleoperation scenarios due to human compensation. From this perspective, all current works focusing on the development of hand-eye calibration with high accuracy requirements are to be applied for laboratory trials. In theory, hand-eye calibration only needs to be performed once before experiments. However, due to the nature of hand-eye calibration resolving the inaccuracy in encoder readings, it would be more efficient to conduct on-the-fly hand-eye calibration for applications where accurate kinematics data is also required. Furthermore, considering that hand-eye calibration mostly benefits surgical applications where visual guidance plays a crucial role, we suspect that more emphasis will be placed on the development of clinically adaptable hand-eye calibration methods with high accuracy requirements once the need for greater uptake in surgical subtasks automation will have materialized.

We believe that, as newer versions of the dVRKs gradually become available, many of the technical issues under the first-generation dVRKs discussed in this paper can be alleviated to a large extent, if not fully eradicated. For example, the latest da Vinci XI robots adopt digital potentiometers, as opposed to analogue ones installed on the da Vinci Classic, and hence the accuracy and reliability of joint encoder readings could be potentially enhanced for future versions of the dVRK. Also, since the advent of the da Vinci Si, the endoscope cameras installed on the robots produce images with better quality, which would inherently improve the quality of visual feedback for future dVRK researchers.

Nevertheless, it is likely that future da Vinci robots will remain tendon-driven despite having shorter kinematics

chains; therefore, essential kinematics calibration will still be required. In addition, the practice of moving the SUJs and zooming in/out the endoscopic camera during operation will still require repeating hand-eye calibration and camera calibration. To streamline the overall surgical workflow, we suggest that one avenue for future research should be the development of a unified approach that could merge the calibration procedures and be implemented on the fly. We have stressed the interrelationship of kinematics calibration, hand-eye calibration and potentiometer calibration, which supports the benefit of a unified approach. We envision that a single clinically adaptable marker, attached around the tool shaft, could serve as a mutual calibration object, with great potential to unify the calibration procedures. The unified approach would help reduce potential error factors from individual research efforts by future dVRK users.

### VIII. CONCLUSION

In this paper we covered the most widely-reported dVRK technical limitations discovered in the currently available literature with a view to aiding future research on the subject. However, there are other underlying technical constraints that may have equal, if not more of a severe impact on surgical task performance. These underlying constraints, including but not limited to misalignment between surgeon-side and patient-side manipulators and naive inverse kinematics solvers, have already been discovered by leading researchers of the dVRK community. This paper not only aims to provide guidance for dVRK users such that they are aware of these potential factors, but also to encourage future works on the comparative assessment of different calibration and performance accuracy assessment studies. The constraints identified here are rooted in the first-generation dVRK teleoperational surgical platform, though are likely to exist in similar form as new versions are made available. Therefore, the development of unified solutions that can be incorporated within future versions of da Vinci robots are also recommended. Finally, given that the cable-driven nature and imaging system of the da Vinci system are mirrored in other surgical robotic implementations (e.g. RAVEN), the hope is that our paper can also be of use by researchers outside the dVRK community.

### IX. ACKNOWLEDGMENTS

We would like to express our appreciation towards three leading researchers in the dVRK community: Mr. Anton Deguet, Dr. Simon DiMaio and Prof. Peter Kazanzides for their valuable advice during the course of completing this work, especially to Anton, who kindly agreed to review the potentiometer calibration section of this paper, and Peter, who provided valuable feedback towards our revised manuscript. Also, we thank each and every one of the peer researchers within the dVRK community who completed the questionnaire and provided us with invaluable feedback based on their unique research experience.

### APPENDIX I QUESTIONNAIRE

The questionnaire consists of three sections, corresponding to kinematic/dynamics data (data from the PSMs, ECM, and SSMs), imaging data (images captured by the dVRK's endoscope), and system data (e.g., signals from the foot pedals for teleoperation), respectively. For each section, users first need to answer a yes-or-no question to determine if they have used the type of data before. If not, they could skip to the next section; otherwise, they are instructed to point out the sources of error in that data type which they reckon may have affected their research results. Users are prompted to identify the sources of error from a list of selected options that were gathered from literature review and consultation with leading researchers. In addition, the participants are also available to put in other potential sources of error by themselves. A total of 35 responses have been collected from 20 different institutions, summarized as follows:

#### A. Kinematics and dynamics data

Approximately 95% of the participants have experience in dealing with kinematics and dynamics data. The most commonly identified problems are "inaccurate forward-kinematics" (66%) and "inaccurate hand-eye calibration" (66%), followed by issues with the surgical instruments (56%), and "inaccurate potentiometer calibration" (50%), as shown in Figure 7.

#### B. Imaging data

Approximately 70% of the participants have experience in dealing with imaging data streamed from the endoscope of the dVRK. The most commonly identified problem with imaging data is "inaccurate camera calibration" (83%), followed by "low pixel resolution" (65%), as shown in Figure 7.

In addition to the sources of error listed from the selection box, 20% of the participants have reflected that there is a significant difference between the left and right image, mainly perceived as colour mismatch. Also, 13% of the participants have pointed out the presence of noise in the images.

#### C. System state data

Approximately 30% of the participants have experience in dealing with system state data. No significant sources of error have been identified here.

#### D. Questionnaire Analysis

After further analysis, it is discovered that 80% of dVRK users did not conduct kinematics calibration when conducting hand-eye calibration despite they are in the knowledge that the accuracy of joint encoder readings could have been improved. We hypothesise that this results from a lack of readily available resources for kinematics calibration, which could be leveraged by dVRK users when undertaking their individual research. Also, 90% of researchers adopted

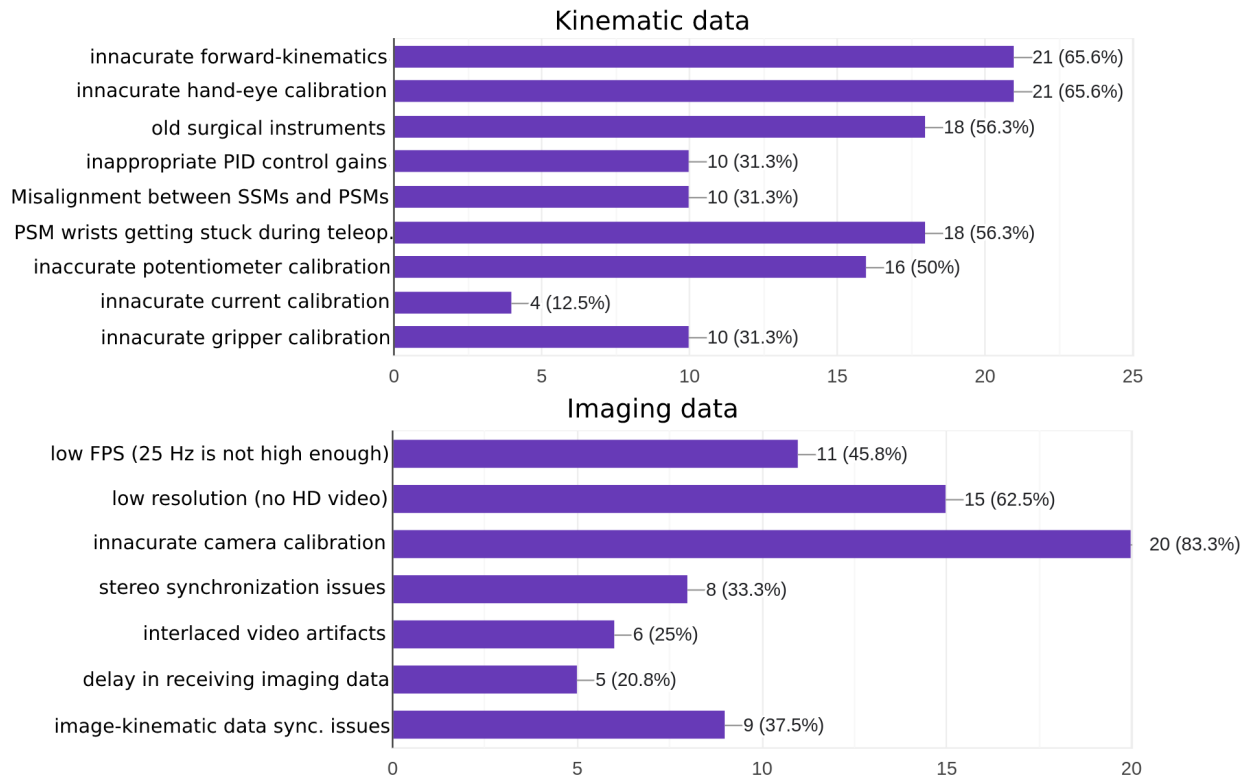


Fig. 7. Answers to the questionnaire identifying sources of error both on kinematic and imaging data captured by the first-generation dVRK.

marker-based approaches for hand-eye calibration, instead of marker-free ones.

It is also discovered that a significant percentage of dVRK users are either not aware of the necessity for certain calibrations, or not familiar with the methods for certain calibrations. Specifically, 40% of the users are not aware that the camera intrinsic parameters are not constant; 20% are not aware of the methods for kinematics calibration, and 30% are not aware of potentiometer calibration. In addition, 25% of the users reported that they bypassed hand-eye calibration in their research, which can be explained by the fact that some institutes mainly use the dVRKs for the purpose of training or conducting user studies, where the human element makes redundant the hand-eye calibration.

After discussions with the participants, it is gathered that 63% of the dVRK users are dissatisfied with the current standard definition (SD) cameras on the first generation dVRK. It is suggested from our analysis that the incorporation of cameras with high-definition (HD) and a faster frame rate into the dVRK system would be preferable, such that the image data would be better synchronized with the kinematics data, thereby contributing towards better system functionality and usability.

#### REFERENCES

- [1] "Intuitive corp. 2021. intuitive sustainability report." <https://www.intuitive.com/en-us/-/media/ISI/Intuitive/Pdf/2021-intuitive-sustainability-report.pdf>, 2022.
- [2] C. D'Ettore, A. Mariani, A. Stilli, F. Rodriguez y Baena, P. Valdastrì, A. Deguet, P. Kazanzides, R. H. Taylor, G. S. Fischer, S. P. DiMaio, A. Menciassi, and D. Stoyanov, "Accelerating surgical robotics research: A review of 10 years with the da vinci research kit," *IEEE Robotics & Automation Magazine*, vol. 28, no. 4, pp. 56–78, 2021.
- [3] 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 *2014 IEEE international conference on robotics and automation (ICRA)*, pp. 6434–6439, IEEE, 2014.
- [4] D. M. Kwartowitz, S. D. Herrell, and R. L. Galloway, "Toward image-guided robotic surgery: determining intrinsic accuracy of the da vinci robot," *International Journal of Computer Assisted Radiology and Surgery*, vol. 1, pp. 157–165, 2006.
- [5] J. M. Ferguson, L. Y. Cai, A. Reed, M. Siebold, S. De, S. D. Herrell, and R. J. Webster, "Toward image-guided partial nephrectomy with the da vinci robot: exploring surface acquisition methods for intraoperative re-registration," in *Medical imaging 2018: Image-guided procedures, robotic interventions, and modeling*, vol. 10576, pp. 49–59, SPIE, 2018.
- [6] A. Baumkircher, K. Seme, M. Munih, and M. Mihelj, "Collaborative robot precision task in medical microbiology laboratory," *Sensors*, vol. 22, no. 8, p. 2862, 2022.
- [7] D. Seita, S. Krishnan, R. Fox, S. McKinley, J. Canny, and K. Goldberg, "Fast and reliable autonomous surgical debridement with cable-driven robots using a two-phase calibration procedure," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6651–6658, 2018.
- [8] A. Wilcox, J. Kerr, B. Thananjeyan, J. Ichnowski, M. Hwang, S. Paradis, D. Fer, and K. Goldberg, "Learning to localize, grasp, and hand over unmodified surgical needles," in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 9637–9643, 2022.
- [9] F. Zhong, Y. Wang, Z. Wang, and Y.-H. Liu, "Dual-arm robotic needle insertion with active tissue deformation for autonomous suturing," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2669–2676, 2019.
- [10] P. Sundaresan, B. Thananjeyan, J. Chiu, D. Fer, and K. Goldberg, "Automated extraction of surgical needles from tissue phantoms," in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pp. 170–177, 2019.
- [11] M. Hwang, B. Thananjeyan, S. Paradis, D. Seita, J. Ichnowski, D. Fer, T. Low, and K. Goldberg, "Efficiently calibrating cable-driven surgical

- robots with rgbd fiducial sensing and recurrent neural networks,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5937–5944, 2020.
- [12] M. Hwang, B. Thananjeyan, D. Seita, J. Ichnowski, S. Paradis, D. Fer, T. Low, and K. Goldberg, “Superhuman surgical peg transfer using depth-sensing and deep recurrent neural networks,” *arXiv preprint arXiv:2012.12844*, 2020.
- [13] M. Hwang, D. Seita, B. Thananjeyan, J. Ichnowski, S. Paradis, D. Fer, T. Low, and K. Goldberg, “Applying depth-sensing to automated surgical manipulation with a da vinci robot,” in *2020 International Symposium on Medical Robotics (ISMR)*, pp. 22–29, 2020.
- [14] R. Moccia, C. Iacono, B. Siciliano, and F. Ficuciello, “Vision-based dynamic virtual fixtures for tools collision avoidance in robotic surgery,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1650–1655, 2020.
- [15] A. Banach, K. Leibrandt, M. Grammatikopoulou, and G.-Z. Yang, “Active constraints for tool-shaft collision avoidance in minimally invasive surgery,” in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 1556–1562, 2019.
- [16] M. Grammatikopoulou, K. Leibrandt, and G.-Z. Yang, “Motor channelling for safe and effective dynamic constraints in minimally invasive surgery,” in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4317–4323, 2016.
- [17] Y. Cai, P. Choi, C.-W. V. Hui, R. H. Taylor, and K. W. S. Au, “A task space virtual fixture architecture for teleoperated surgical system with slave joint limit constraints,” *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 1, pp. 69–80, 2022.
- [18] F. Zhong, Z. Wang, W. Chen, K. He, Y. Wang, and Y.-H. Liu, “Hand-eye calibration of surgical instrument for robotic surgery using interactive manipulation,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1540–1547, 2020.
- [19] M. Morris, M. Shoham, and M. Nahon, “Applications and theoretical issues of cable-driven robots,” in *Proceedings of the Florida Conference on Recent Advances in Robotics, Boca Raton, FL, USA*, pp. 21–22, 2009.
- [20] M. Haghhighipناه, M. Miyasaka, Y. Li, and B. Hannaford, “Unscented kalman filter and 3d vision to improve cable driven surgical robot joint angle estimation,” in *2016 IEEE international conference on robotics and automation (ICRA)*, pp. 4135–4142, IEEE, 2016.
- [21] M. Miyasaka, J. Matheson, A. Lewis, and B. Hannaford, “Measurement of the cable-pulley coulomb and viscous friction for a cable-driven surgical robotic system,” in *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 804–810, IEEE, 2015.
- [22] M. Haghhighipناه, M. Miyasaka, Y. Li, and B. Hannaford, “Unscented kalman filter and 3d vision to improve cable driven surgical robot joint angle estimation,” in *2016 IEEE international conference on robotics and automation (ICRA)*, pp. 4135–4142, IEEE, 2016.
- [23] G. Chryssilla, N. Eusman, A. Deguet, and P. Kazanzides, “A compliance model to improve the accuracy of the da vinci research kit (dvrk),” *Acta Polytechnica Hungarica*, vol. 16, no. 8, 2019.
- [24] S. Huang, *Kinematic Calibration for da Vinci Surgical Robot*. PhD thesis, Case Western Reserve University, 2019.
- [25] Z. Roth, B. Mooring, and B. Ravani, “An overview of robot calibration,” *IEEE Journal on Robotics and Automation*, vol. 3, no. 5, pp. 377–385, 1987.
- [26] G. Fattori, A. J. Lomax, D. C. Weber, and S. Safai, “Technical assessment of the ndi polaris vega optical tracking system,” *Radiation oncology*, vol. 16, no. 1, pp. 1–4, 2021.
- [27] F. Richter, J. Lu, R. K. Orosco, and M. C. Yip, “Robotic tool tracking under partially visible kinematic chain: A unified approach,” *IEEE Transactions on Robotics*, vol. 38, no. 3, pp. 1653–1670, 2022.
- [28] M. Ye, L. Zhang, S. Giannarou, and G.-Z. Yang, “Real-time 3d tracking of articulated tools for robotic surgery,” in *International conference on medical image computing and computer-assisted intervention*, pp. 386–394, Springer, 2016.
- [29] F. Richter, Y. Zhang, Y. Zhi, R. K. Orosco, and M. C. Yip, “Augmented reality predictive displays to help mitigate the effects of delayed telesurgery,” in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 444–450, IEEE, 2019.
- [30] G. A. Fontanelli, F. Ficuciello, L. Villani, and B. Siciliano, “Modelling and identification of the da vinci research kit robotic arms,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1464–1469, 2017.
- [31] N. Yilmaz, J. Y. Wu, P. Kazanzides, and U. Tumerdem, “Neural network based inverse dynamics identification and external force estimation on the da vinci research kit,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1387–1393, IEEE, 2020.
- [32] Y. Wang, R. Gondokaryono, A. Munawar, and G. S. Fischer, “A convex optimization-based dynamic model identification package for the da vinci research kit,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3657–3664, 2019.
- [33] A. Novoseltseva, *Force Feedback for the Patient Side Manipulator of the daVinci Research Kit*. PhD thesis, Master’s Thesis, Worcester Polytechnic Institute, 2018.
- [34] Z. Zhang, L. Zhang, and G.-Z. Yang, “A computationally efficient method for hand–eye calibration,” *International journal of computer assisted radiology and surgery*, vol. 12, no. 10, pp. 1775–1787, 2017.
- [35] M. W. Walker, A. Dobryden, and J. Dionise, “Getting to know your robot,” in *Proceedings., IEEE International Conference on Robotics and Automation*, pp. 1245–1250, IEEE, 1990.
- [36] J. Song, P. Si, H. Hua, and M. Qiu, “Research on the inherent nonlinearity calibration of the potentiometer of a miniature linear series elastic actuator,” in *Actuators*, vol. 11, p. 207, MDPI, 2022.
- [37] “dvrk wiki on github regarding potentiometer calibration.” <https://github.com/jhu-dvrk/sawIntuitiveResearchKit/wiki/Calibration#potentiometers>. Accessed: 2022-08-10.
- [38] “dvrk wiki on github regarding potentiometer calibration of psm’s 3rd joint.” <https://github.com/jhu-dvrk/sawIntuitiveResearchKit/wiki/Potentiometer-calibration-for-PSM-insertion>. Accessed: 2022-08-10.
- [39] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [40] D. Forsyth and J. Ponce, *Computer vision: A modern approach*. Prentice hall, 2011.
- [41] L. Yu and G. Lubineau, “Modeling of systematic errors in stereo-digital image correlation due to camera self-heating,” *Scientific reports*, vol. 9, no. 1, pp. 1–15, 2019.
- [42] J. Civera, D. R. Bueno, A. J. Davison, and J. Montiel, “Camera self-calibration for sequential bayesian structure from motion,” in *2009 IEEE International Conference on Robotics and Automation*, pp. 403–408, IEEE, 2009.
- [43] P. Pratt, C. Bergeles, A. Darzi, and G.-Z. Yang, “Practical intraoperative stereo camera calibration,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 667–675, Springer, 2014.
- [44] F. Mourgues, E. Coste-Maniere, C. Team, et al., “Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 25–34, Springer, 2002.
- [45] J.-Y. Bouguet, “Camera calibration toolbox for matlab,” [http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html), 2004.
- [46] W. Qi, F. Li, and L. Zhenzhong, “Review on camera calibration,” in *2010 Chinese control and decision conference*, pp. 3354–3358, IEEE, 2010.
- [47] Y. Zhao, W. Jia, and R. Wang, “Rethinking deinterlacing for early interlaced videos,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 7, pp. 4872–4878, 2021.
- [48] J. Cartucho, A. Weld, S. Tukra, H. Xu, H. Matsuzaki, T. Ishikawa, M. Kwon, Y. Jang, K.-J. Kim, G. Lee, B. Bai, L. Kahrs, L. Boecking, S. Allmendinger, L. Muller, Y. Zhang, Y. Jin, B. Sophia, F. Vasconcelos, W. Reiter, J. Hajek, B. Silva, L. R. Buschle, E. Lima, J. L. Vilaca, S. Queiros, and S. Giannarou, “Surgt: Soft-tissue tracking for robotic surgery, benchmark and challenge,” 2023.
- [49] M. Kalia, P. Mathur, N. Navab, and S. E. Salcudean, “Marker-less real-time intra-operative camera and hand-eye calibration procedure for surgical augmented reality,” *Healthcare Technology Letters*, vol. 6, no. 6, pp. 255–260, 2019.
- [50] A. Azarbayejani and A. P. Pentland, “Recursive estimation of motion, structure, and focal length,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 6, pp. 562–575, 1995.
- [51] G. Qian and R. Chellappa, “Bayesian self-calibration of a moving camera,” *Computer Vision and Image Understanding*, vol. 95, no. 3, pp. 287–316, 2004.
- [52] N. Keivan and G. Sibley, “Online slam with any-time self-calibration and automatic change detection,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5775–5782, IEEE, 2015.
- [53] O. Bogdan, V. Eckstein, F. Rameau, and J.-C. Bazin, “Deepcalib: A deep learning approach for automatic intrinsic calibration of wide

- field-of-view cameras,” in *Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production*, pp. 1–10, 2018.
- [54] A. Gordon, H. Li, R. Jonschkowski, and A. Angelova, “Depth from videos in the wild: Unsupervised monocular depth learning from unknown cameras,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8977–8986, 2019.
- [55] J. Fang, I. Vasiljevic, V. Guizilini, R. Ambrus, G. Shakhnarovich, A. Gaidon, and M. R. Walter, “Self-supervised camera self-calibration from video,” in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 8468–8475, IEEE, 2022.
- [56] D. Stoyanov, A. Darzi, and G.-Z. Yang, “Laparoscope self-calibration for robotic assisted minimally invasive surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 114–121, Springer, 2005.
- [57] A. Bartoli, D. Pizarro, and T. Collins, “A robust analytical solution to isometric shape-from-template with focal length calibration,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 961–968, 2013.
- [58] S. Parashar, A. Bartoli, and D. Pizarro, “Self-calibrating isometric non-rigid structure-from-motion,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 252–267, 2018.
- [59] A. Agudo, “Total estimation from rgb video: on-line camera self-calibration, non-rigid shape and motion,” in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 8140–8147, IEEE, 2021.
- [60] “The da vinci vision.” <https://www.intuitive.com/en-us/products-and-services/da-vinci/vision>. Accessed: 2022-07-13.
- [61] B. Li, L. Heng, K. Koser, and M. Pollefeys, “A multiple-camera system calibration toolbox using a feature descriptor-based calibration pattern,” in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1301–1307, IEEE, 2013.
- [62] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [63] Y. Liu, X. Zhang, S. Wang, S. Ma, and W. Gao, “Spatial-temporal correlation learning for real-time video deinterlacing,” in *2021 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, IEEE, 2021.
- [64] Y.-C. Yeh, J. Dy, T.-M. Huang, Y.-Y. Chen, and K.-L. Hua, “Vd-net: video deinterlacing network based on coarse adaptive module and deformable recurrent residual network,” *Neural Computing and Applications*, pp. 1–14, 2022.
- [65] K. Pachtrachai, F. Vasconcelos, F. Chadebecq, M. Allan, S. Hailes, V. Pawar, and D. Stoyanov, “Adjoint transformation algorithm for hand-eye calibration with applications in robotic assisted surgery,” *Annals of biomedical engineering*, vol. 46, no. 10, pp. 1606–1620, 2018.
- [66] 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 *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2485–2491, 2016.
- [67] R. Horaud and F. Dornaika, “Hand-eye calibration,” *The international journal of robotics research*, vol. 14, no. 3, pp. 195–210, 1995.
- [68] J. Jiang, X. Luo, Q. Luo, L. Qiao, and M. Li, “An overview of hand-eye calibration,” *The International Journal of Advanced Manufacturing Technology*, pp. 1–21, 2021.
- [69] K. Daniilidis, “Hand-eye calibration using dual quaternions,” *The International Journal of Robotics Research*, vol. 18, no. 3, pp. 286–298, 1999.
- [70] R. Y. Tsai, R. K. Lenz, *et al.*, “A new technique for fully autonomous and efficient 3 d robotics hand/eye calibration,” *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 345–358, 1989.
- [71] Z. Wang, Z. Liu, Q. Ma, A. Cheng, Y.-h. Liu, S. Kim, A. Deguet, A. Reiter, P. Kazanzides, and R. H. Taylor, “Vision-based calibration of dual rcm-based robot arms in human-robot collaborative minimally invasive surgery,” *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 672–679, 2017.
- [72] Y. C. Shiu and S. Ahmad, “Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form  $ax=xb$ ,” 1987.
- [73] 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 Transactions on Robotics and Automation*, vol. 10, no. 4, pp. 549–554, 1994.
- [74] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [75] A. Roberti, N. Piccinelli, D. Meli, R. Muradore, and P. Fiorini, “Improving rigid 3-d calibration for robotic surgery,” *IEEE Transactions on Medical Robotics and Bionics*, vol. 2, no. 4, pp. 569–573, 2020.
- [76] O. Özgüner, T. Shkurti, S. Huang, R. Hao, R. C. Jackson, W. S. Newman, and M. C. Çavuşoğlu, “Camera-robot calibration for the da vinci robotic surgery system,” *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 4, pp. 2154–2161, 2020.
- [77] C. Wang, J. Cartucho, D. S. Elson, A. Darzi, and S. Giannarou, “Towards autonomous control of surgical instruments using adaptive-fusion tracking and robot self-calibration,” in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022.
- [78] J. Cartucho, C. Wang, B. Huang, D. S. Elson, A. Darzi, and S. Giannarou, “An enhanced marker pattern that achieves improved accuracy in surgical tool tracking,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, no. 4, pp. 400–408, 2022.
- [79] “dvrk wiki on github regarding calibration.” <https://github.com/jhudvrk/sawIntuitiveResearchKit/wiki/Calibration>. Accessed: 2022-08-10.