

# Visual kinematic force estimation in robot-assisted surgery - application to knot tying

P.J. “Eddie” Edwards<sup>a</sup>, Emanuele Colleoni<sup>a</sup>, Aswhin Sridhar<sup>b</sup>, John D. Kelly<sup>b</sup> and Danail Stoyanov<sup>a</sup>

<sup>a</sup>Surgical Robot Vision Group, WEISS, UCL, Charles Bell House, 43-45 Foley St, London W1W 7TS, UK; <sup>b</sup> Urology Department, Westmoreland Street Hospital, UCLH, London, UK;

## ARTICLE HISTORY

Compiled Wednesday 26<sup>th</sup> August, 2020

## ABSTRACT

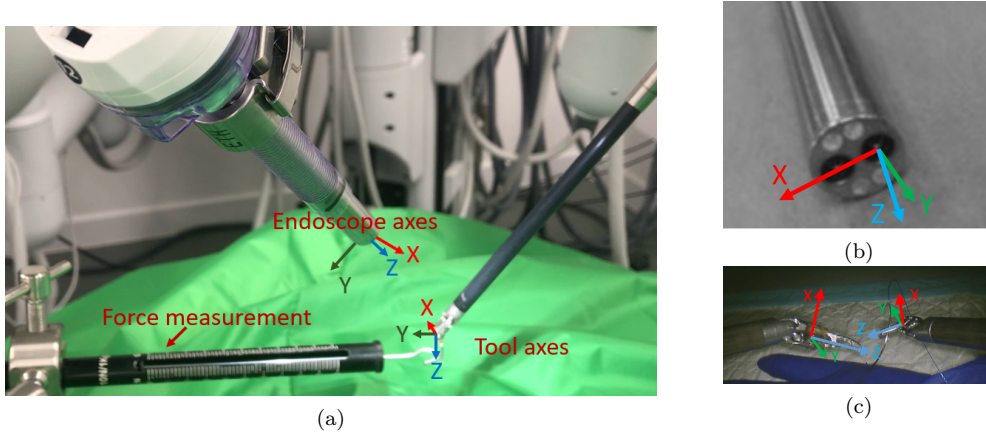
Robot-assisted surgery has potential advantages over more conventional methods but lacks any force feedback, which can lead to errors such as broken stitches or tissue damage. More experienced surgeons are able to judge the tool-tissue forces visually and finding an automated way of capturing this skill would seem desirable. Methods to measure force tend to involve complex measurement devices or visual tracking of tissue deformation. We investigate whether surgical forces can be estimated simply from the discrepancy between kinematic prediction and visual measurement of the tool position. We show that combined visual and kinematic force estimation can be achieved that does not rely on external measurements or modelling of tissue deformation. An initial alignment from kinematic pose to the visual scene is performed when no force is applied to the tool. When a subsequent force is applied the visual and kinematic estimates of tool position diverge. We plot visual/kinematic displacement using forces measured with a strain gauge with both vision and marker based tracking. Graphs of force vs displacement demonstrate the ability to discern the forces involved in knot tying. We also visualize our displacement force using the publicly available JIGSAWS dataset as well as clinical examples of knot tying with the da Vinci surgical system. We anticipate that the ability to visualize or feel forces using this method will offer a significant advantage to those learning robotic surgery as well as adding to the information available to more experienced surgeons.

## KEYWORDS

Force estimation; robotic surgery; visual tracking; kinematic tracking

## 1. Introduction

Robot-assisted endoscopic surgery has been rapidly gaining acceptance and has many advantages over conventional laparoscopic surgery. These include features that make the surgical task easier, including the enhanced dexterity of wrist motion, scaling of movements and reduction in tremor as well as better outcomes for the patient such as improved cosmesis and reduced blood loss and infection rates. Such collaborative robots lack haptic feedback, however, which can be a hindrance to both the learning and practice of robot-assisted surgery (Okamura 2009).



**Figure 1.** Coordinate system definitions, showing the overall experimental setup (a), the endoscope axes with the origin on the left camera (b) and the tool axes centred on the last wheel with Z-axis in the direction of the tool (c)

Force torque measurements have been shown to distinguish expert from novice surgeons (Richards et al. 2000), which shows that force is a useful indicator of surgical skill and also suggests that communication of force might help shorten the learning curve for such procedures. Mechanical sensors placed at or near the robot end effector can provide accurate force measurement (Lee et al. 2016), but add significant complexity to the robot design and may not be cost effective (Okamura 2009). Forces on the robotic tool can be estimated using measurements of the currents in the motors and tool velocities in conjunction with a physical model of the robot. Most of the sensors are distant from the robot end effector, however, which can lead to inaccuracies and such methods have been largely restricted to the final grip force (Li et al. 2016).

Experienced surgeons are able to mitigate the effects of lack of haptic feedback using visual cues. One can emulate this perception by visual estimation of the distortion of the tissue, coupled with either a mechanical model of the tissue (Haouchine et al. 2018) or deep learning to provide force estimation (Noohi et al. 2014; Aviles et al. 2017). Being based solely on the endoscopic video, these methods add no complexity to the robot design and can potentially resolve very small forces. Most work to date is limited to a particular tissue type, such as liver (Marban et al. 2019) or brain (Giannarou et al. 2016), but this method shows some promise.

We recently made the observation that the predicted position from kinematic constraints and the measured position from vision will diverge when a force is applied to the robot. Any distortion of the robotic joints will result in inaccurate kinematic localisation, but this also gives an indication of the force acting upon the tool. In addition to established surgical cues, such as observation of tissue deformation and suture shape, observing a reduction of tool motion when moving the manipulator could be one of the ways experienced surgeons are able to perceive forces visually. This idea has been proposed in a patent (Zhao et al. US Patent 2010/0169815, Intuitive Surgical Inc, 2010), but to our knowledge has not been considered in the academic literature to date.

The purpose of this paper is to document this effect and assess its potential as a means of achieving force estimation that is independent of the tissue or procedure and requires no modification of the robotic tools. We concentrate on knot tying. Previous experiments using external force measurements have shown an improved learning

curve and a reduction in broken sutures when force is visually communicated to the trainee (Reiley et al. 2008), making this a good potential application.

## 2. Method

### 2.1. Equipment

We use two da Vinci™ surgical systems - a “classic” or original da Vinci with PAL resolution (576i) endoscopic video and a da Vinci Si providing either 720p or 1080i video. The dVLogger system from Intuitive Surgical, Inc. is used to record time-stamped stereoendoscopic video and kinematic information. The experimental setup and corresponding coordinate systems can be seen in Fig. 1.

### 2.2. Data Processing

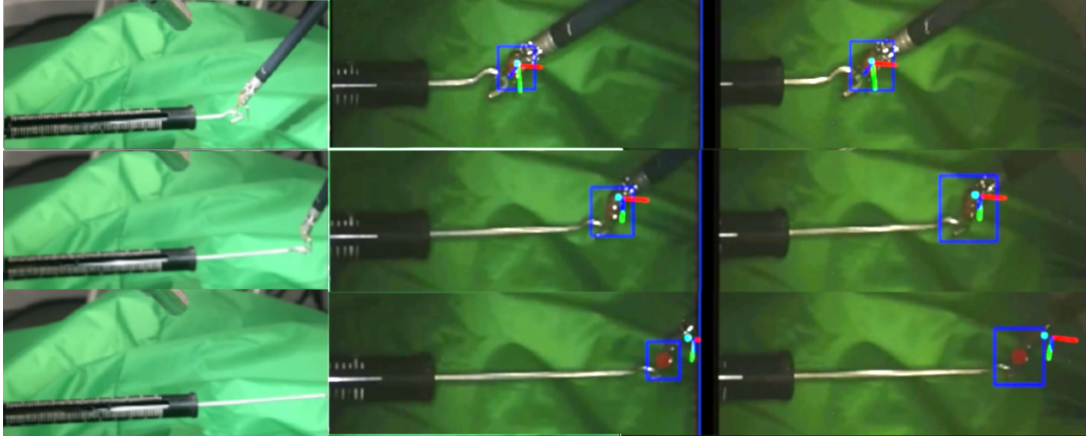
We investigate the force-displacement relationship of the da Vinci robot. For our initial experiments we approximate the tool position visually using the CSRT tracker from Lukežič et al. (2017). This is consistently one of the best performing patch-based object tracking methods and is available in OpenCV (Kristan et al. 2017). The CSRT tracker provides an initial match from the left to right frame as well as subsequent tracking for both frames. This gives 3D tool position tracking from stereo. The origin of the tool provided by the dVLogger is in the centre of the final wheel that enables jaw opening and closing (see Fig. 1(c)). Though this point is internal and not directly visible, it can be inferred from the position of the tool parts. The position of this origin is manually marked using a box in the first frame of the video segment. The CSRT tracker was able to track subsequent motion of this point with an accuracy of better than 5 pixels over short periods of a few tens of frames. To establish whether improved visual tracking would enable more accurate estimation of forces, we also attached Aruco markers to the Si robot to provide consistent tracking over longer periods.

The timestamps from the dVLogger should provide good synchronisation but there can still be residual mismatches in lag between video and kinematic data. We match the kinematic motion to the visually tracked tool using Procrustes alignment. The relative lag between visual and kinematic tracking is optimized so as to minimize the residual error to provide accurate temporal alignment.

As a simple indicator of the force we use the displacement between the kinematic and visual positions of the tool origin in pixels. This is a relative measure and gives an indication of the force rather than an absolute value. To measure forces in the lab setting we use a spring strain gauge. The gauge measurement is recorded using an external video camera which is synchronized to the endoscopic video using observations of the robot motion. The force measurement is then made by observing the position of the force indicator in the video.

### 2.3. Experiments

To assess the feasibility and accuracy of force measurement, we have performed four experimental evaluations. As the force on the end effector increases we expect the discrepancy between the visual and kinematic positions to increase. It is known that the kinematics of cable driven robots can have inaccuracies, causing drift or hysteresis in the predicted position. To overcome this, we perform experiments over a short time



**Figure 2.** Spring forces and visual discrepancy for the da Vinci Classic robot. External camera view of the spring gauge and scene is on the left. The endoscopic view is on the right with OpenCV tracked position in red for the right and left views, and predicted position and orientation from kinematics shown as axes. The brightness of the red dot provides an indication of force. Forces are 0N (top), 3N (middle) and 5N (bottom)

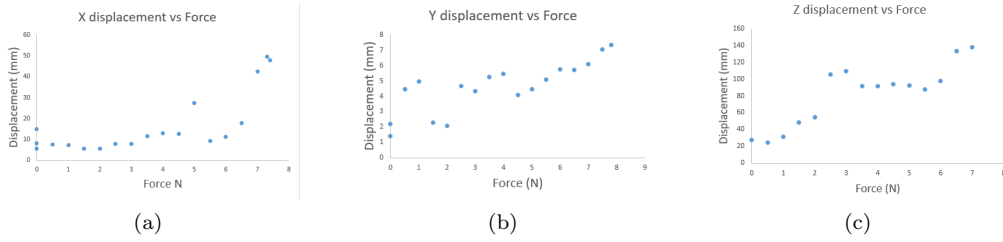
period and calculate an approximate hand-eye calibration by matching the visual and kinematic tool position when no force is present. Subsequent motion with force should deviate from the visually tracked position. Spring gauges measure the absolute force on the robot in the lab setting.

### *2.3.1. OpenCV visual force estimation with the da Vinci Classic*

In this experiment we gathered video and kinematic data from the da Vinci Classic. Stereo calibration is performed using a standard chessboard and OpenCV. In a first pass, the tool is moved without any force acting up it. The positions of the tool from vision and kinematics are recorded and registered by Procrustes analysis to orient and scale the 3D visual position from stereo to provide a best match. This effectively provides a hand-eye calibration that matches kinematic motion to the visual scene. After this initial hand-eye calibration, the tool is tracked and the visual and kinematic positions should remain well aligned. The strain gauge hook is then grabbed by the tool and a the spring is extended. The distance between kinematic and visual position is measured as a displacement in millimetres. This is displayed visually as a red dot of increasing intensity with increasing force. Experiments are performed in all 3 axes showing that lateral, vertical and depth tracking are possible using the stereo-endoscopic video. A separate camera captures the overall scene and is able to see the measurement from the strain gauge. Several time points are chosen and both the measured and visual force estimations are recorded.

### *2.3.2. Aruco marker force feedback with the da Vinci Si*

To assess the effect of tracking accuracy on the force-displacement relationship we performed a second experiment using the da Vinci Si. This experiment has a very similar setup to 2.3.1, but an Aruco marker is attached to the relevant section of the robot. Aruco markers provide automated marker tracking and identification in OpenCV (Romero-Ramirez et al. 2018). The da Vinci Si provides images at 1080i which are then deinterlaced to provide 1080 resolution video data without interlacing issues such as comb artefact. OpenCV stereo calibration is performed as before. Accurate



**Figure 3.** Graphs of displacement vs force for the da Vinci classic showing horizontal motion (a), vertical motion (b) and depth motion (c). The increasing trend is apparent in all directions.

positional tracking using the markers along with higher resolution video from the da Vinci Si should enable smaller discrepancies to be found and hence allow more accurate force estimation. We also display tool orientation as well as position to assess how the kinematic tracking may be affected by torque forces.

### 2.3.3. Visual force estimation with the JIGSAWS dataset

We also estimate the forces in surgical knot-tying in training sessions. The JIGSAWS dataset from Johns Hopkins University provides a standard for surgical gesture recognition and skills assessment (Gao et al. 2014). This dataset consists of videos and corresponding kinematic poses for the two tools. After an initial synchronisation of the video and kinematic information, we perform visual tracking using the OpenCV CSRT tracker previously described. We perform analysis on the left eye. The kinematic location of the tool should be in endoscope coordinates (see Fig. 1(b)), which are centred on the left camera facing along the optical axis. In this coordinate system, the values  $x/z$  and  $y/z$  should be proportional the pixel coordinates  $u, v$ . The scaling factor is estimated from an initial calibration with no force present. Subsequent discrepancy between visual and kinematic location represents the estimated force. Without stereo calibration the analysis is 2D and force is measured in pixels.

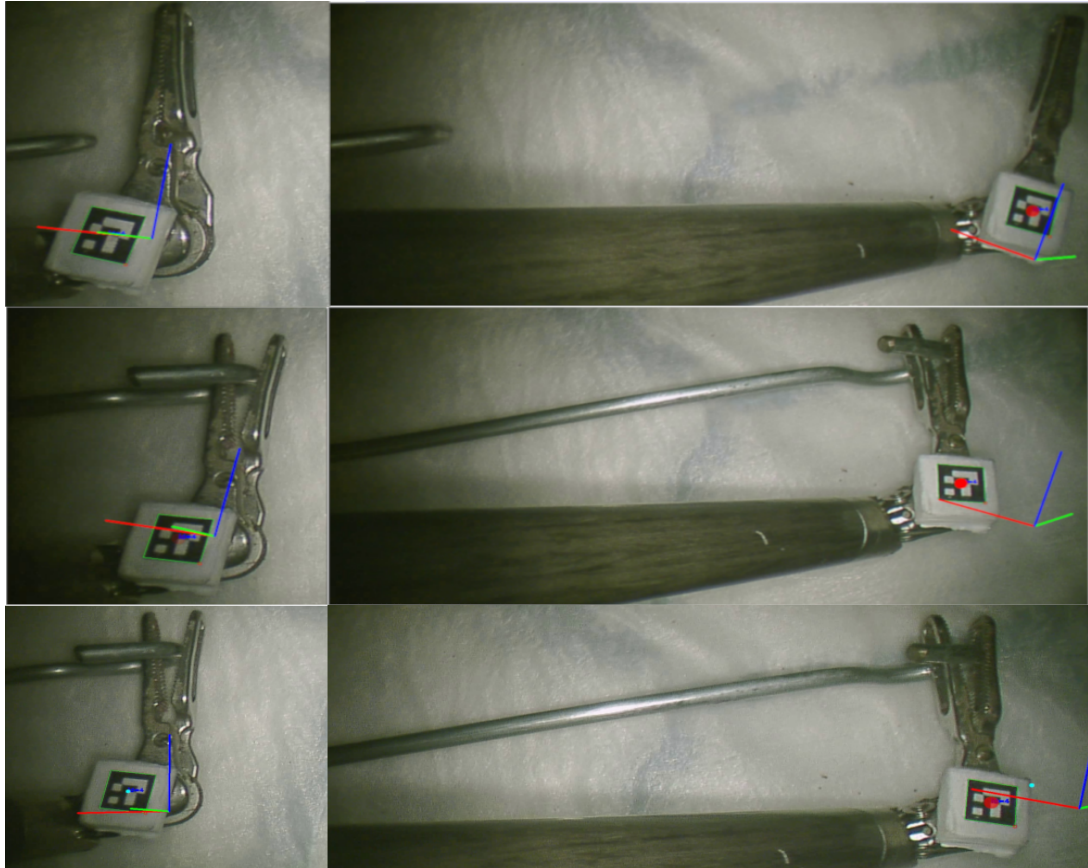
### 2.3.4. Visual force feedback in a clinical dataset

Finally, we have access to clinical data from robot-assisted radical prostatectomies performed at a collaborating hospital. Though visual tracking of the tool in the real surgical scene is difficult due to the presence of blood and smoke, we use manual alignment to provide initial registration and overlay only the kinematic position of the tool. The surgeon can use their visual perception of the tool to provide an estimation of force.

## 3. Results

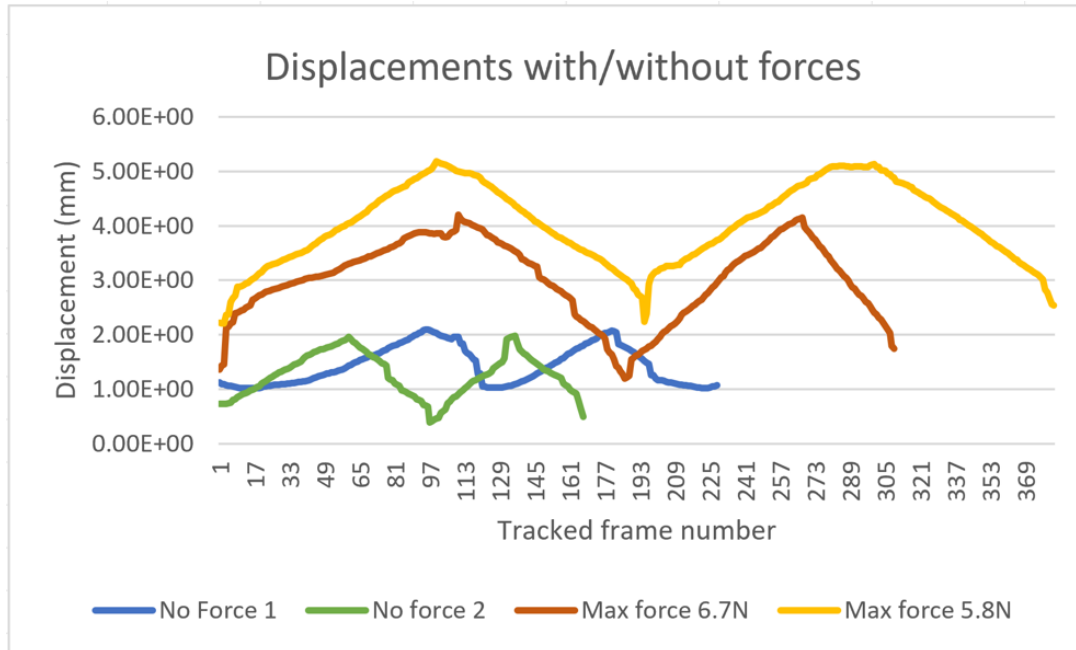
### 3.1. OpenCV visual force estimation with the da Vinci Classic

Some example frames from this experiment can be seen in Fig. 2. The visualisation of increasing discrepancy with increasing force is observed. Fig. 3 shows the force-displacement relationship when the spring gauge is extended multiple times. A clear correlation is visible, but there are also significant errors visible. This may be due to visual tracking errors or hysteresis in the force-displacement relationship for the robot. Such issues are perhaps to be expected as the robot is not designed exhibit



**Figure 4.** The Aruco marker position is shown in red and predicted position and orientation from kinematics is shown as axes. Above we can see good tracking when no force is applied (top image). In both the examples in the lower images, the kinematic position shows inaccuracy in both position and orientation as forces increase.





**Figure 5.** Scatter plots showing the motion variation as the tool is moved throughout the visual field. The green and blue lines correspond to no applied force. The residual variation could be further reduced with improved hand-eye calibration. With maximum spring forces of 6.7N and 5.8N applied (red and orange lines) the displacement is greater. Roughly linear increase with applied force is seen, though the maximum displacement varies in each case.

linear elasticity.

### 3.2. Aruco marker force feedback with the da Vinci Si

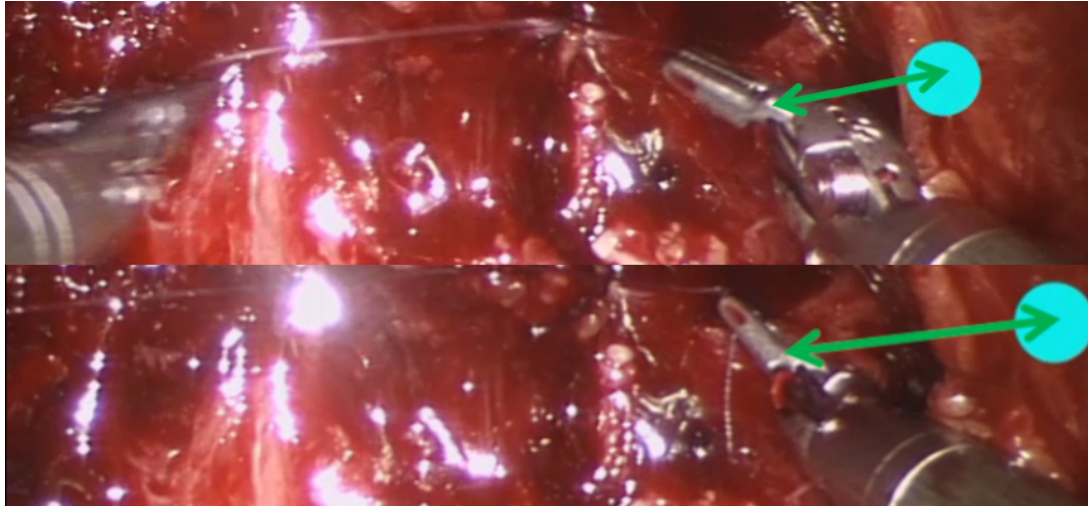
Some frames from this experiment can be seen in Fig. 4. It is clear that there is not only a displacement of the tool from its kinematically predicted position, but also a rotational inaccuracy. Since the spring is acting towards the end of the tool at a distance from the origin in the centre of the wheel this is not surprising, but this observation underlines the need for a more complex model of the robot forces than mere displacement. When a force is applied the discrepancy increases in a roughly linear fashion with improved visual tracking. The maximum force as measured by the spring gauge is 6.7N and 5.8N in the two cases shown. The corresponding maximum displacement is 4mm and 5mm, so is lower for the higher applied force. This may be due to hysteresis or changes in the configuration of the robot arms. But overall it seems that forces greater than 4-5N will produce displacements that can be reliably differentiated from zero force.

### 3.3. Visual force estimation with the JIGSAWS dataset

In Fig. 6 we can see captured frames from the JIGSAWS dataset. When a knot is tightened there is motion of the kinematic position away from the visual position, resulting in a clearly visualized force. This experiment shows that our method can be applied to historical data to give an estimate of surgical forces where no force measurements were made. Endoscopic video and corresponding kinematics are all that



**Figure 6.** JIGSAWS knot tying examples. No force on the left, knot tightening on the right. The brightness of the red dot gives a clear indication of force.



**Figure 7.** Clinical example of knot tightening during robot-assisted radical prostatectomy. As the knot is tightened (bottom image) the predicted positions of the tools from kinematics diverge (green arrow)

is required.

### *3.4. Visual force feedback in a clinical dataset*

In clinical examples it has not yet been possible to provide an accurate visual tool tracker due to the presence of blood. However, manual delineation of a few points is enough to provide a decent overlay of the kinematic position. It is then up to the surgeon to provide their own visual tracking of the actual tool position. This can still be an effective communication of force, as shown in Fig. 7, where the kinematic tool positions diverge as the knot is tightened.

## **4. Discussion**

This paper highlights an effect that has received little attention in the literature so far, namely that a discrepancy between the kinematic and visual position of a robotic tool can provide an estimate of the force applied. The accuracy with which the force can be detected depends on the flexibility of the robot, which is generally minimized by design, and the accuracy of visual tool tracking.

The methods used here are simple and used to demonstrate and evaluate the effect.



Kinematic tracking uses the dVLogger provided by Intuitive Surgical, Inc., which provides the tool position within the endoscope frame. A more complex kinematic model showing the position of all parts of the surgical tool may be more illuminating.

Accurate tool tracking is difficult, particularly in real surgical scenes in the presence of blood contamination and smoke. We used simple block matching from OpenCV which is effective only over short time periods of a few tens of frames. Tool tracking is a highly active area of research, however, and the ability to include a force dimension adds to the motivation for this work. One of the few works to use both kinematic and visual data is the da Vinci tool tracker from Ye et al. (2016). Our future work will build on this work and other tool tracking methodologies, such as Du et al. (2016), to incorporate kinematic information and a virtual CAD model to provide more accurate tracking and improved visualisation. The analysis here incorporates both stereo views, but not the full visual instrument pose.

The force estimation described in this paper is based purely on displacement. The absolute force is not estimated. Accurate force estimation would require a physical model of the robot coupled with 3D visual estimation of the robotic tool. This will form part of our future research. Such a model may provide a better relationship between force and displacement. The distortion of both orientation and position that is apparent in Fig. 4 due to the force being applied near the tip of the instrument suggests that such a model may provide better force estimation.

However, a full physical model of the robot will be difficult to generate. There may be differences in tool shaft bending depending on how far the tool extends beyond the trocar and whether the force applied is along or perpendicular to this shaft, for example. But external observation of the robot suggests that distortion also occurs in the set up joints that bear the weight of the robot arms. A combined model taking into account all of these effects will not be easily achieved.

Perhaps a more feasible solution would be to train a neural network to recognize this effect and relate it to the force applied at the tool. A regression network may be able to predict the forces given kinematic and visual data and could learn when effects such as hysteresis occur. To train such a network would require a wide range of forces in many positions measured with a more accurate force-torque sensor rather than spring gauges. The force estimation as it stands may feed into other learning systems seeking to establish the skill level of the surgeon, for example.

The clinical examples here are limited to knot tying. This requires significant force in the horizontal plane so is well suited to this implementation. The force estimation we propose will provide extra and complementary information that may be useful when assessing surgical skill. In order to limit damage to tissue and provide palpation, the ability to measure force in the Z direction to a higher accuracy will be necessary. A machine learning approach that incorporates the information described in this paper as well as visual cues of surface deformation may be capable this level of accuracy.

Despite the limitations of the experiments described in this paper, we have shown that forces of 4-5N can be detected and shown to the surgeon. Abiri et al. (2017) measured the forces required to break sutures used in robotic surgery. Even the finest 5-0 gauge silk sutures had breaking forces of over 8N. This suggests that the effect observed in this paper could be used as an early warning system for suture failure, which is a common problem observed when learning robotic surgery (Abiri et al. 2018).

There are many avenues for future development of this method, but the effect demonstrated in this paper can already provide an indication of robotic surgical forces, which may be beneficial to trainee and expert surgeons alike.

## Acknowledgements

This work was supported by the Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS) at UCL (203145Z/16/Z), EPSRC (EP/P027938/1, EP/R004080/1) and the H2020 FET (GA 863146). Danail Stoyanov is supported by a Royal Academy of Engineering Chair in Emerging Technologies (CiET1819/2/36) and an EPSRC Early Career Research Fellowship (EP/P012841/1). We are also grateful to technical support provided by Intuitive Surgical, Inc. under a research agreement.

## References

- Abiri A, Askari SJ, Tao A, Juo YY, Dai Y, Pensa J, Candler R, Dutson EP, Grundfest WS. 2018. Suture breakage warning system for robotic surgery. *IEEE Transactions on Biomedical Engineering*. 66(4):1165–1171.
- Abiri A, Paydar O, Tao A, LaRocca M, Liu K, Genovese B, Candler R, Grundfest WS, Dutson EP. 2017. Tensile strength and failure load of sutures for robotic surgery. *Surgical endoscopy*. 31(8):3258–3270.
- Aviles AI, Alsaleh SM, Hahn JK, Casals A. 2017. Towards retrieving force feedback in robotic-assisted surgery: A supervised neuro-recurrent-vision approach. *IEEE Transactions on Haptics*. 10(3):431–443.
- Du X, Allan M, Dore A, Ourselin S, Hawkes D, Kelly JD, Stoyanov D. 2016. Combined 2d and 3d tracking of surgical instruments for minimally invasive and robotic-assisted surgery. *International Journal of Computer Assisted Radiology and Surgery*. 11(6):1109–1119.
- Gao Y, Vedula SS, Reiley CE, Ahmidi N, Varadarajan B, Lin HC, Tao L, Zappella L, in Béjar B, Yuh DD, et al. 2014. JHU-ISI gesture and skill assessment working set (JIGSAWS): A surgical activity dataset for human motion modeling. In: *Modeling and Monitoring of Computer Assisted Interventions (M2CAI) – MICCAI Workshop*; vol. 3. p. 3.
- Giannarou S, Ye M, Gras G, Leibrandt K, Marcus HJ, Yang GZ. 2016. Vision-based deformation recovery for intraoperative force estimation of tool–tissue interaction for neurosurgery. *International Journal of Computer Assisted Radiology and Surgery*. 11(6):929–936.
- Haouchine N, Kuang W, Cotin S, Yip M. 2018. Vision-based force feedback estimation for robot-assisted surgery using instrument-constrained biomechanical three-dimensional maps. *IEEE Robotics and Automation Letters*. 3(3):2160–2165.
- Kristan M, Leonardis A, Matas J, Felsberg M, Pflugfelder R, Cehovin Zajc L, Vojir T, Hager G, Lukežić A, Eldesokey A, et al. 2017. The visual object tracking vot2017 challenge results. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. p. 1949–1972.
- Lee D, Kim U, Gulrez T, Yoon WJ, Hannaford B, Choi HR. 2016. A laparoscopic grasping tool with force sensing capability. *IEEE/ASME Transactions on Mechatronics*. 21(1):130–141.
- Li Y, Miyasaka M, Haghhighipanah M, Lei Cheng, Hannaford B. 2016. Dynamic modeling of cable driven elongated surgical instruments for sensorless grip force estimation. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*; May. p. 4128–4134.
- Lukežić A, Vojir T, Čehovin Zajc L, Matas J, Kristan M. 2017. Discriminative correlation filter with channel and spatial reliability. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. p. 6309–6318.
- Marban A, Srinivasan V, Samek W, Fernández J, Casals A. 2019. A recurrent convolutional neural network approach for sensorless force estimation in robotic surgery. *Biomedical Signal Processing and Control*. 50:134 – 150.
- Noohi E, Parastegari S, Žefran M. 2014. Using monocular images to estimate interaction forces during minimally invasive surgery. In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*; Sep. p. 4297–4302.
- Okamura A. 2009. Haptic feedback in robot-assisted minimally invasive surgery. *Current Opin-*

- ion in Urology. 19(1):102–107.
- Reiley CE, Akinbiyi T, Burschka D, Chang DC, Okamura AM, Yuh DD. 2008. Effects of visual force feedback on robot-assisted surgical task performance. *The Journal of Thoracic and Cardiovascular Surgery*. 135(1):196 – 202.
- Richards C, Rosen J, Hannaford B, Pellegrini C, Sinanan M. 2000. Skills evaluation in minimally invasive surgery using force/torque signatures. *Surgical Endoscopy*. 14(9):791–798.
- Romero-Ramirez FJ, Muñoz-Salinas R, Medina-Carnicer R. 2018. Speeded up detection of squared fiducial markers. *Image and Vision Computing*. 76:38 – 47.
- Ye M, Zhang L, Giannarou S, Yang GZ. 2016. Real-time 3d tracking of articulated tools for robotic surgery. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Springer. p. 386–394.
- Zhao W, Zhao T, Larkin DQ. US Patent 2010/0169815, Intuitive Surgical Inc, 2010. Visual force feedback in a minimally invasive surgical procedure.