Integrating Real-time Video View with Pre-operative Models for Image-guided Renal Navigation: An *in vitro* Evaluation Study

Peter Jackson[1], Richard Simon[2], and Cristian Linte[1,2]

Abstract— To provide a complete picture of a scene sufficient to conduct a minimally invasive, image-guided renal intervention, real-time laparoscopic video needs to be integrated with underlying anatomy information typically available from pre- or intra-operative images. Here we present a simple and efficient hand-eye calibration method for an optically tracked camera, which only requires the acquisition of several poses of a Polaris stylus featuring 4 markers automatically localized by both the camera and the optical tracker. We evaluate the calibration using both the Polaris stylus, as well as a patient-specific 3D printed kidney phantom in terms of the number of poses acquired, as well as the depth of the imaged scene into the field of view of the camera, by projecting the several landmarks on the imaged object at known location in the 3D world onto the camera image. The RMS projection error decreases with increasing distance from the camera to the imaged object from 7 pixels at 15-18 mm, to under 2 pixels at 28-30 mm, which corresponds to a 2 mm and 1 mm error, respectively, in 3D space.

I. INTRODUCTION

To reduce trauma and improve patient outcome, minimally invasive image-guided renal interventions rely on laparoscopic visualization of the surgical scene as a surrogate for direct vision. However, laparoscopic video provides only organ surface information, with no underlying anatomy. Such information is usually available from pre-operative images or anatomical models derived from computed tomography (CT) or magnetic resonance imaging (MRI), or intra-operative imaging modalities, such as ultrasound (US) or X-ray imaging. This information can be used to further enhance laparoscopic visualization during image-guided renal interventions. The pre-operative images or models derived from these images can be directly overlaid onto the laparoscopic images in a virtual reality (VR) or augmented reality (AR) scene to provided additional information, such as subsurface vasculature or internal surgical target location. However, to generate a spatially correct VR or AR environment representative of the surgical scene, the pre-operative images, virtual representations of tracked surgical instruments, and real-time, intra-operative video images need to be correctly registered to display the "real" and "virtual" surgical scene from the same perspective.

External tracking systems (e.g. NDI Polaris Spectra optical tracking system (OTS) and others) provide position and orientation of tracked dynamic reference frames (DRFs). A DRF can be rigidly attached to a camera and then tracked to determine its pose and position in space relative to the tracker. The images from the camera can then be overlaid with pre-procedural anatomical models to generate an augmented reality view. However, the generation of such overlays requires accurate calibration between the camera optical axis and the DRF. This calibration is referred to as a hand-eye calibration [1], which is not as straightforward to perform in a surgical environment, due to sterile and time constraints.

There have been several methods and techniques proposed in the literature to perform the hand-eye calibration [2]–[4] with the simplest method using a Procrustes analysis of paired-points in the image and reference frames. Some techniques to obtain the paired-points rely on using a tracked checkerboard to provide the 3D coordinates of the corners, which are also automatically identified in the camera image [5], [6]. Another technique to obtain point pairs is rotating a camera around a fixed point [2]. Once several poses featuring a set of landmarks and corresponding coordinates in 3D space and 2D image space are identified, the calibration transformation is solved for using some form of the Perspective-n-Point (PnP) problem or other formulation [7]. Morgan *et al.* used an optically tracked stylus with a calibrated tool-tip to obtain the coordinates of the tool-tip in 3D space and image space from several poses; these homologous coordinates were then used to solve for the camera calibration.

In this paper, we propose a hand-eye calibration method for a renal access multi-modality imaging guidance platform and assess its uncertainty. The proposed method build on work described by Morgan *et al.* [8], but instead of using the coordinates of the tip of the tracked stylus to determine the calibration transformation, we rely directly on the built-in reflective markers on a DRF. This approach eliminates any uncertainty associated with a tool tip calibration or associated with tracking a checkerboard pattern using an attached DRF. We assess the calibration accuracy in terms of the 2D error between the projected landmarks from the 3D world space onto the 2D image space for both a Polaris stylus and a 3D printed patient-specific kidney phantom and additionally, assess the 3D error between the corresponding landmarks. Finally, we demonstrate the generation of both real video views and virtual views for the surgical scene from the same view perspective, along with a corresponding overlay.

[1] Center for Imaging Science, Rochester Institute of Technology Rochester, NY 14623, USA `pj7837@rit.edu`

[2] Department of Biomedical Engineering, Rochester Institute of Technology Rochester, NY 14623, USA `calbme@rit.edu`

II. METHODOLOGY

A. Instrumentation and Apparatus

We used a Logitech C910 webcam as our imaging device; the camera has a maximum resolution of 5 MP, but was used at 1.3 MP for a faster frame rate. The camera was instrumented with a DRF to be tracked with a NDI Polaris Spectra OTS. To simulate a renal intervention scene, we used a 3D printed patient-specific kidney phantom created by segmenting an abdominal MRI and generating a virtual model. The model was "instrumented" with 14 3-mm dia. hemispherical divots used as registration fiducials or targets, as well as an additional 10 landmarks that serve solely to emulate targeting.

The OTS was positioned such that its optimal tracking volume covered the working volume around the 3D printed kidney phantom. The phantom, tracked stylus, and the tracked camera were positioned such that the OTS could unobstructively "see" each component, while allowing the camera to "see' the tracked stylus and phantom. In addition, we also ensured the camera remained within the OTS field of view while being manipulated to image the phantom from various distances and viewing angles.

B. Hand-eye Camera Calibration Procedure

The transformation relationships between the respective coordinate frames of the components of a IGI system augmented with a tracked camera is shown in Fig. 1, where T_a^b is a 4x4 rotation and translation matrix that describes the transformation from coordinate space a to b. The transformations
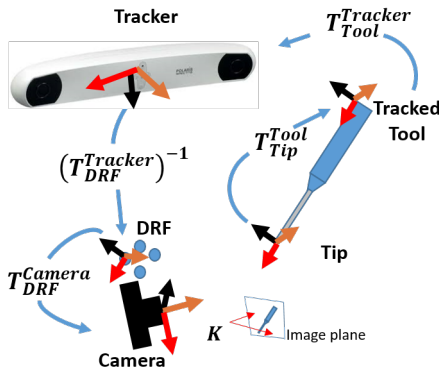

Fig. 1. Diagram representation for a fully calibrated system to project points that enables the projection of landmarks from the 3D world onto a 2D camera image. Each T represents a transformation matrix between coordinate systems.

labeled in Fig. 1 as $T_{DRF}^{Tracker}$ and $T_{Tool}^{Tracker}$, are given by the OTS. T_{Tip}^{Tool} is found using a pivot calibration [9], which is later used to register the 3D kidney phantom from the 3D world to its virtual counterpart using the tracking system. T_{DRF}^{Camera} is the unknown hand-eye calibration transformation matrix, while K is the camera intrinsic matrix which governs the perspective transformation of 3D camera space coordinates into 2D image space coordinates.

Assuming a pinhole camera model, the hand-eye calibration can be formulated as,

$$X_{Image} = K T_{World}^{Camera} X_{World}, \qquad (1)$$

where X_{Image} and X_{World} are the homogeneous 2D coordinates in the image and 3D coordinates in world space, respectively. T_{World}^{Camera} is a generic transformation that maps world points to 3D camera space and can be solved for using PnP algorithms, which requires a set of n corresponding landmarks in both the 3D world and 2D image.

To obtain a set of corresponding landmarks, while maintaining a simple hand-eye calibration solution, we used the retro-reflective markers on the tracked stylus and transformed them into the coordinate system of the DRF attached to the camera. This mapping was accomplished using Equation 2:

$$X_{DRF} = (T_{DRF}^{Tracker})^{-1} T_{Stylus}^{Tracker} X_{Stylus}, \qquad (2)$$

where X_{Stylus} and X_{DRF} represent the coordinates of the stylus markers in the 3D stylus coordinate system and the 3D DRF coordinate system attached to the camera, respectively. The coordinates of the stylus markers in image space were obtained by capturing an image of the stylus and manually identifying the center of the visible marker in the image, which served as the ground truth for our subsequent error analysis. The images were captured simultaneously with the 3D coordinates captured by the tracker. Each image, or pose, provides a set of 4 corresponding landmarks, one for each marker. We captured a total of 18 poses across the image viewing plane to use for calibration. The camera was calibrated using Zhang's [10] method as implemented in OpenCV to solve for the intrinsic camera parameters, K.

These point-pairs, which consist of the 3D coordinates of the stylus markers transformed into the camera DRF space and their 2D coordinates in image space, allowed us to rewrite Equation 1 as follows:

$$X_{Image} = K T_{DRF}^{Camera} X_{DRF}, \qquad (3)$$

which we then solve for the calibration matrix, T_{DRF}^{Camera}, using OpenCV SolvePNP. We also computed the calibration using an implementation of the algorithm described in [7] as a comparison to the OpenCV algorithm. The method proposed a solution to determine the registration between homologous points and lines using an Iterative Closest Point model, which minimizes the distance between 3D points and lines.

Once the calibration matrix is determined, a known point in the OTS space can be projected onto the image plane by

$$X_{Image} = K T_{DRF}^{Camera} (T_{DRF}^{Tracker})^{-1} X_{Tracker}, \qquad (4)$$

where $X_{Tracker}$ represents the point's coordinates in 3D OTS space and X_{Image} represents a point's homogeneous coordinate in 2D image space. Fig. 2 shows an example of the stylus with the coordinates of the markers in 3D space projected onto the camera image.
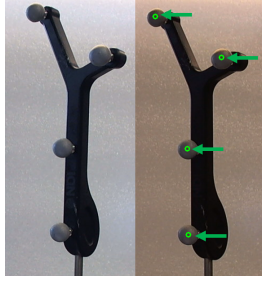
Fig. 2. Image showing a pose of the stylus featuring 4 markers (left panel) and their corresponding 3D coordinates projected onto the camera image (right panel), indicated by the arrows.

C. Calibration Validation with Stylus

Following hand-eye calibration, we collected an independent test set consisting of 39 poses to assess calibration uncertainty and analyze the errors. We positioned the stylus at multiple locations on a plane parallel to the image plane, and repeated the measurements at increasing distances. The stylus markers were transformed from their 3D world coordinate system into the 2D camera image coordinate system using Eq. 4. The calibration error was estimated in image space as the pixel distance between the manually identified and projected marker locations. We also calculated the calibration error in 3D space in terms of the distance between the 3D world coordinates of a point transformed into camera space and a line that connects the origin (i.e., camera) and the point in image space, which would ideally intersect the marker in 3D space. As such, if the hand-eye calibration is correct, and in the absence of noise, the distance between the 3D marker and this line should be zero. The distance is calculated using Equation 5:

$$d = \frac{|\,(q-p) \times (p-r)\,|}{|q-p|},\qquad (5)$$

where p is the camera origin, q is the point on the image plane, and r is the point in 3D space. The methods to assess 2D pixel error in image space and 3D mm error in 3D space are schematically illustrated in Fig. 3
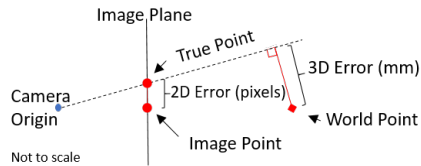

Fig. 3. The location of ground truth, image, and world points relative to the image plane and the associated error for the 2D and 3D cases. Points on the image plane have (x, y) coordinates with units in pixels or mm, while the world point has (x, y, z) coordinates with units in mm.

D. Calibration Validation with Phantom

To assess the calibration error in the context of the desired image-guided renal navigation procedure, we conducted a validation study using the 3D printed patient-specific kidney phantom. We registered the physical kidney phantom to its

corresponding virtual model by using a set of six (6) paired-point landmarks sampled with the optically tracked stylus. To ensure a sufficiently accurate registration, we used the fiducial configuration that yielded the best registration, as reported in our previous work [11], characterized by a RMS FRE on the order of 0.57-mm.

Following registration, we acquired several poses of the kidney phantom and its associated fiducial markers from two viewing angles, with the camera located at two different distances from the kidney: proximal (180 mm) and distal (290 mm). Similar to the stylus-based validation, we recorded the 3D landmarks projected onto the camera image of the phantom, along with the landmark locations manually identified in the camera image.

We collected another data set using a 3D kidney phantom by placing the camera at two distances and recording the locations of landmarks in 3D world space and 2D image space three times for each of the two camera perspectives: anterior and posterior. This experimental protocol yielded a total of 12 poses, each featuring 10 recorded landmarks.
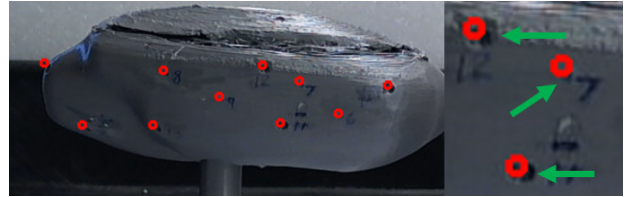

Fig. 4. Video view of the 3D printed kidney phantom showing a set of phantom landmarks projected from 3D world space onto the 2D camera image (left panel) and a close-up view of projected landmarks (circles) and actual landmarks (indicated by arrows in the right panel)

Fig. 4 shows a view of the 3D printed kidney phantom along with a set of 3D landmarks projected onto the camera image. The points on the phantom are projected by first registering the phantom to the OTS and then using this world registration transformation along with the other transformations to map 3D world coordinates to 2D camera image coordinates, as described by Equation 6:

$$X_{Image} = K T_{DRF}^{Camera} (T_{DRF}^{Tracker})^{-1} T_{Phantom}^{Tracker} X_{Phantom}, \qquad (6)$$

where $X_{Phantom}$ represents the 3D coordinates of the landmarks on the phantom and X_{Image} represents homogeneous coordinates of landmarks in the image plane.

E. Augmented Reality Overlay

With a fully calibrated system, we can use Equation 6 to overlay the virtual model onto the camera image. We used 3D Slicer [12] along with the IGT Extension [13] and PLUS [14] to capture a view of the virtual model from the same perspective as the real camera view, and either show these two views side by side or as a virtual on real view overlay. This allows the user to visually check the registration between the models and then use the overlay image to improve navigation. As a verification step of the correctness of the image overlay, we recorded the 2D image

space coordinates of the stylus markers in 10 images using manual identification, as well as the projection of the 3D coordinates of the markers onto the image using Eq. 4 and the virtual camera in 3D Slicer. We then calculated the RMS error between each projected set of coordinates and the manually identified locations.

III. RESULTS

A. Calibration Transformation Assessment

We first studied the error associated with the calibration transformation matrix and its convergence following the acquisition of several poses.

One way to visualize the stability of the calibration and verify that the calibration transformation is not changing is to examine the sum of the squared difference between the calculated transformation matrices as more features are added. Fig 5 indicates a large difference at first, when using only two calibration poses; the sum of the squared difference then drops sigficantly with the use of 4 or more calibration poses.
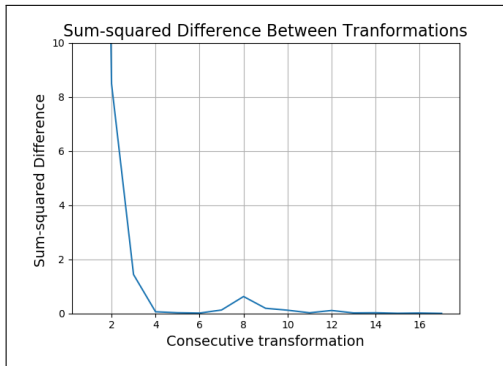

Fig. 5.   Sum-squared difference between calculated T_{DRF}^{Camera} transformation as a function of increasing number of poses used to determine the transformation. This shows the estimated transforms converge rapidly to a stable value.

In addition to assessing the convergence of the estimated calibration transformation, we also assessed the effect of the number of poses used to compute the calibration transformation in terms of the RMS distance error between the 3D projected landmarks onto the image space and the image space landmarks. We randomly selected n poses to calculate the calibration transformation and evaluated the fit of the calibration using the same pose, then repeated this five times for each n. Fig. 6 indicates that using at least three poses to perform the calibration, the error associated with the projection of the 3D space also stabilizes. These observations confirm previous studies [2], [7], [8], which indicated that twelve feature points are sufficient to generate a stable and sufficiently accurate calibration transformation.

Fig 6 also shows that as the number of poses used to compute the transformation increases, the RMS error converges to approximately 2 pixels. This error consists of both the calibration error , as well as the uncertainty associated with the user selection of the landmarks in the camera image. This user selection uncertainty in localizing the markers in
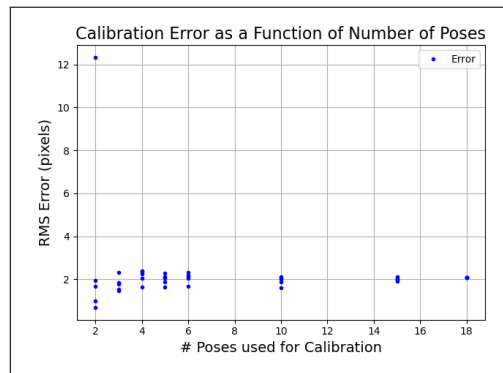

Fig. 6.   Calibration error as a function of number of poses. The calibration was calculated using n random poses which was repeated five times for each n. As n increases, the error converges to 2 pixels.

the camera image was measured and showed a 1.2 pixel variability.

We also compared the calibration transform obtained using OpenCV to the calibration transform calculated using the method described by Chen *et al*. [7]. The sum-squared difference between our reconstructed calibration transformation matrices and those calculated using method of Chen *et al*.was 0.07, indicating that both methods yield similar transformations.

B. Stylus Calibration Validation

To ensure the hand-eye calibration was accurate, we computed the error associated with the projection of 3D landmarks from an independent set of poses (i.e., not used to determine the hand-eye calibration) onto the camera image. In 2D image space, the RMS error is the distance (in pixels)
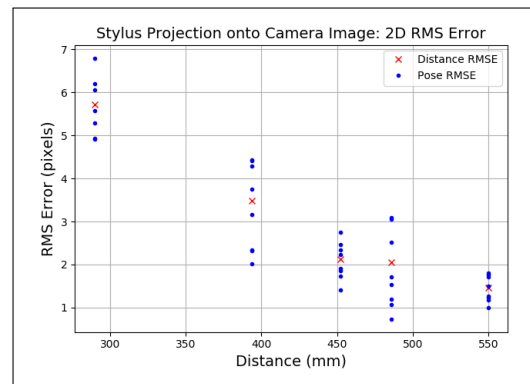

Fig. 7.   Projection error in 2D space (pixels) for poses collected at varying distances. The RMS error for each pose is marked using a 'dot', while RMS error computed across all points from all poses collected at a single distance from the camera is marked with an 'x'. The distance is measured using the average z distance from camera origin to the 3D world coordinates of the pose.

between the manually identified landmarks and the projected 3D landmarks in each pose and computed across all the the points collected a specific distance from the camera. Fig. 7 illustrates the the RMS error (in pixels), which decreases with increasing distance between the camera and imaged

object. The plot shows the RMS error associated with each pose (dot), as well as the RMS error for all the points collected at a given distance (x). The error was plotted as a function of distance from the camera, which was determined using the average z coordinate value of the location in camera space.
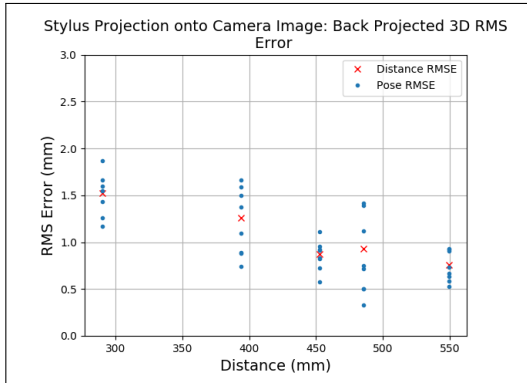

Fig. 8. Estimated RMS distance between "true" and projected points in 3D world space. The RMS error for each pose is marked using a 'dot', while the RMS error computed across all points from all poses collected at a single distance from the camera is marked with an 'x'. The distance is measured using the average z distance from camera origin to the 3D world coordinates of the pose

While the 2D error is important, the projection error in 3D space is more important to the user, as it indicates the uncertainty in the system in physical units. Fig. 8 shows the error of points projected into 3D space, which quantifies the uncertainty in terms of the physical distance between the projected 3D landmarks and the same landmarks captured in the camera image.

C. Phantom Calibration Validation

We conducted a similar analysis using the 3D patient-specific kidney phantom as the object of interest We collected points across the surface of the kidney phantom, while positioned at a 170-185 mm range from the camera for the proximal measurement and 270-300 mm range for the distal measurement. Fig. 9 shows the RMS error associated with
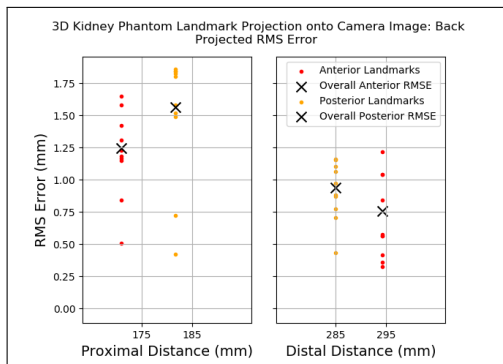

Fig. 9. Estimated RMS distance between "true" and projected points in 3D world space (mm). It shows the RMS error for each point on the anterior and posterior side of the phantom with a '.' and the overall error with a 'x'. Data points were collected at 170-185 mm for the proximal distance and 270-300 mm for the distal distance.

projecting 3D landmarks onto the 2D camera image of the anterior and posterior views of the kidney. Note the slightly higher error on the posterior view compared to the anterior view, which may be, in part, due to the greater spread in distance from the camera of target fiducials on the posterior side of the phantom, as well as perhaps sub-optimal tracking of the camera by the OTS when acquiring the posterior view of the phantom.

D. Same Perspective Real and Virtual Views with Overlay

Fig. 10 shows a real and virtual view of the *in vitro* image-guided renal intervention scene, from the same perspective, along with an overlay showing the augmentation of the real video view with the virtual view. Fig. 10 also shows the
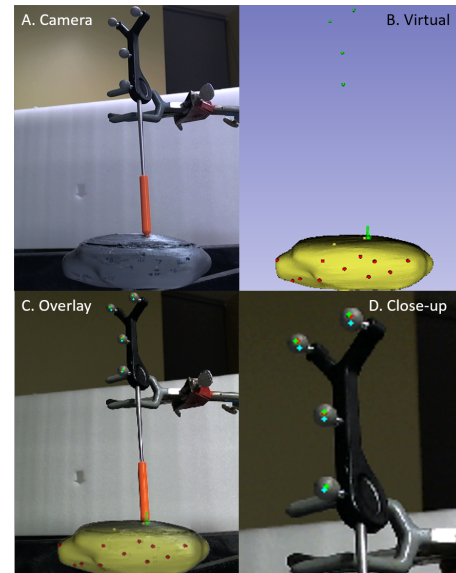

Fig. 10. Real (A) and virtual (B) views of the *in vitro* image-guided renal intervention scene from the same perspective as well the corresponding overlay (C) showing the augmentation of the real video view with its virtual counterpart as well as a close-up, detailed view of stylus marker locations projected onto the image (D). The center of the marker is indicated with a red dot, the 3D Slicer model projection is a green dot, and calibration re-projection is a blue dot

locations of the stylus markers projected onto the camera image of the stylus. The points indicate the center of the markers identified from the camera image (red), the location projected using a model in 3D Slicer (green) and overlaid to create the augmented reality view, and also the location of the marker projected using Eq. 4 (blue). We calculated the RMS error between the 3D Slicer projected points and the manually identified marker center as 1.0 pixels. Similarly, we computed a 1.8 pixel RMS distance error between the points projected using Eq. 4 and the manually identified marker centers. Ideally, these three locations should coincide, but the small error (on the order of 1-2 pixels) given the inherent uncertainty, indicates the position of the 3D Slicer virtual camera and the real camera are sufficiently identical.

IV. DISCUSSION

Our study suggests that the proposed and implemented hand-eye calibration method stabilizes to an acceptable error

on the order of two pixels upon the acquisition of only three poses, well within the tolerance for the proposed application. Our calibration results mirror those reported by Morgan *et al*. [8] in that a minimum of three (3) poses, or 12 features, are needed to perform a consistent and accurate hand-eye calibration and error decreases as a function of increasing distance.

While the 2D image projection error at the closest distance was on the order of 7 pixels, when converted to 3D space, it amounted to 2 mm. This error dropped to approximately 2 pixels, which corresponds to 1 mm in 3D space, when the imaged object is further from the camera. The counterintuitive result of error decreasing with increasing distance away from the camera is primarily a function of the angular size of objects.

In the future, we intend to more fully analyze the uncertainty associated with overlaying the virtual model on the camera image. We will then use the system to investigate the benefits of augmented reality for renal navigation and the possibility of improving registration using information from the camera image.

Since our method currently relies on manually selecting the center of the stylus landmarks, it has limitations on the workflow and uncertainty. The user uncertainty is relatively small at approximately 1 pixel, but an automated method would be more consistent. The time required to determine the hand-eye calibration is less than 2 minutes to collect the minimum number of poses to ensure a consistent transformation. A future development would be to implement a method to automatically determine the location of landmarks in images, which would enable the rapid estimation of the hand-eye calibration by tracking a tool within an operating theater if needed. This approach eliminates the need for any special tools or instruments needed to perform the hand-eye calibration.

The proposed method can be used in an operating theater that utilizes an optical tracking system (such as the NDI Polaris Spectra or similar) for surgical instrument and patient localization in conjunction with an endoscopic guided intervention. This method enables on-the-fly calibration of the tracked video instrument and co-registration between the patient, tracked surgical instruments, pre-procedural images, and the intra-operative video. The co-registration enables the augmentation of the real-time video of the patient with virtual representations of surgical instruments and pre-procedural images for enhanced visualization and navigation.

V. CONCLUSION

We presented a method to perform hand-eye calibration for use with a renal intervention system. The proposed method provides a method which could be used in an environment such as an operating theater while avoiding the extra step of registering a checkerboard to a DRF. We assessed the 2D and 3D error in projecting model landmarks to the expected location in an image. The uncertainty was found to be adequate for the proposed purpose of augmented reality

for renal navigation and provided similar accuracy to other methods. Finally, we showed that the hand-eye calibration can be used for our goal of creating a multi-modal image-guidance platform by augmenting a real-time video view with virtual information extracted from pre-operative images.

REFERENCES

[1] R. Horaud and F. Dornaika, "Hand-Eye Calibration," *The International Journal of Robotics Research*, vol. 14, no. 3, pp. 195–210, 1995.

[2] S. Thompson, D. Stoyanov, C. Schneider, K. Gurusamy, S. Ourselin, B. Davidson, D. Hawkes, and M. J. Clarkson, "Hand–eye calibration for rigid laparoscopes using an invariant point," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 6, pp. 1071–1080, 2016.

[3] J. Marescaux, F. Rubino, M. Arenas, D. Mutter, and L. Soler, "Augmented-Reality–Assisted Laparoscopic Adrenalectomy," *JAMA*, vol. 292, no. 18, p. 2211, 2004.

[4] S. A. Nicolau, L. Goffin, and L. Soler, "A low cost and accurate guidance system for laparoscopic surgery," in *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '05*, (New York, New York, USA), p. 124, ACM Press, 2005.

[5] A. K. R. Voruganti and D. Bartz, "Alternative online extrinsic calibration techniques for minimally invasive surgery," in *Proceedings of the 2008 ACM symposium on Virtual reality software and technology - VRST '08*, (New York, New York, USA), p. 291, ACM Press, 2008.

[6] E. C. S. Chen, K. Sarkar, J. S. H. Baxter, J. Moore, C. Wedlake, and T. M. Peters, "An augmented reality platform for planning of minimally invasive cardiac surgeries," in *Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling* (D. R. Holmes III and K. H. Wong, eds.), vol. 8316, p. 831617, SPIE, 2012.

[7] E. C. S. Chen, T. M. Peters, and B. Ma, "Which point-line registration?," *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10135, no. March 2017, p. 1013509, 2017.

[8] I. Morgan, U. Jayarathne, A. Rankin, T. M. Peters, and E. C. Chen, "Hand-eye calibration for surgical cameras: a Procrustean Perspective-n-Point solution," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 7, pp. 1141–1149, 2017.

[9] Z. Yaniv, "Which pivot calibration?," in *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling* (R. J. Webster and Z. R. Yaniv, eds.), vol. 9415, p. 941527, SPIE, 2015.

[10] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 1, pp. 666–673, IEEE, 1999.

[11] P. Jackson, R. Simon, and C. Linte, "Effect of uncertainty on target registration error in image-guided renal interventions: from simulation to in-vitro assessment," in *Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling* (C. A. Linte and J. H. Siewerdsen, eds.), p. 15, SPIE, 2021.

[12] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper, and R. Kikinis, "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[13] T. Ungi, A. Lasso, and G. Fichtinger, "Open-source platforms for navigated image-guided interventions," 2016.

[14] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, "PLUS: Open-source toolkit for ultrasound-guided intervention systems," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2527–2537, 2014.