

# Reflective-AR Display: An Interaction Methodology for Virtual-to-Real Alignment in Medical Robotics

Javad Fotouhi , Tianyu Song , Arian Mehrfard, Giacomo Taylor , Qiaochu Wang, Fengfan Xian, Alejandro Martin-Gomez, Bernhard Fuerst, Mehran Armand, Mathias Unberath , and Nassir Navab

**Abstract**—Robot-assisted minimally invasive surgery has shown to improve patient outcomes, as well as reduce complications and recovery time for several clinical applications. While increasingly configurable robotic arms can maximize reach and avoid collisions in cluttered environments, positioning them appropriately during surgery is complicated because safety regulations prevent automatic driving. We propose a head-mounted display (HMD) based augmented reality (AR) system designed to guide optimal surgical arm set up. The staff equipped with HMD aligns the robot with its planned virtual counterpart. In this user-centric setting, the main challenge is the perspective ambiguities hindering such collaborative robotic solution. To overcome this challenge, we introduce a novel registration concept for intuitive alignment of AR content to its physical counterpart by providing a multi-view AR experience via reflective-AR displays that simultaneously show the augmentations from multiple viewpoints. Using this system, users can visualize different perspectives while actively adjusting the pose to determine the registration transformation that most closely superimposes the virtual onto the real. The experimental results demonstrate improvement in the interactive alignment of

a virtual and real robot when using a reflective-AR display. We also present measurements from configuring a robotic manipulator in a simulated trocar placement surgery using the AR guidance methodology.

**Index Terms**—Surgical robotics; laparoscopy, computer vision for medical robotics, augmented reality.

## I. INTRODUCTION

ROBOTIC-ASSISTED minimally invasive surgery is becoming increasingly common due to its associated benefits that include higher accuracy, and tremor and fatigue reduction. Robotic systems can augment the surgeon's abilities with stereo endoscopic imaging and intuitive control which help the surgeon's hand-eye coordination and reduce physical workload during surgery [1]. Furthermore, robotic surgery has benefits over traditional laparoscopic techniques with patients experiencing reduced blood loss and shorter post-operative hospital stays [2].

Quick and accurate set up of robotic systems leading up to surgery remains a major challenge in the endeavor of making robotic surgery the standard of care. After a patient has been positioned, anesthetized, and trocars inserted, the robotic arms must be positioned and docked before operation can begin. This procedure is a crucial step of workflow and grows more complex the more joints the robotic arms have. While many different configurations of the robot's joints may allow the robotic arm to dock, sub-optimal positioning increases the likelihood of collisions and inadequate reach during teleoperation. Any repositioning or undocking necessary to adjust the robot significantly decreases operating room efficiency [3]. For many procedures including minimally invasive gastrectomy, "junk time," the time taken to set up or reposition robotic arms, is often the sole reason for increased procedure time in robotic procedures over purely laparoscopic approach [4].

Optimal set up of robotic arms, consequently, is critical to increasing the efficiency of robotic surgery and foster acceptance. Due to safety and regulatory concerns, having a robot automatically drive itself to a pre-operative position is infeasible. In modern surgical robotic systems, the set up of the arms can be supported by lasers, as shown in Fig. 2. Though lasers assist the staff in aligning the robot, it can still exhibit challenges in a complex system with joint redundancies, as it does not directly show the desired configuration of all joints. As most manual methods are error prone and induce a steep learning curve to operating room staff unfamiliar with a system, we investigate an augmented reality (AR) solution for guidance during robotic

Manuscript received September 10, 2019; accepted January 17, 2020. Date of publication February 10, 2020; date of current version February 25, 2020. This letter was recommended for publication by Associate Editor M. Zecca and Editor P. Valdastri upon evaluation of the reviewers' comments. (Corresponding author: Javad Fotouhi.)

Javad Fotouhi is with the Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA, and also with the Laboratory for Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: javad.fotouhi@jhu.edu).

Tianyu Song, Arian Mehrfard, Qiaochu Wang, and Fengfan Xian are with the Laboratory for Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: tsong11@jhu.edu; amehrfal1@jhu.edu; qwang85@jhu.edu; fxian1@jhu.edu).

Giacomo Taylor is with the Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA and with the Laboratory for Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA, and also with the Verb Surgical Inc., Mountain View, CA 94043 USA (e-mail: giacomotaylor@verbsurgical.com).

Alejandro Martin-Gomez is with the Laboratory for Computer Aided Medical Procedures, Technical University of Munich, 80333 Munich, Germany (e-mail: alejandro.martin@tum.de).

Bernhard Fuerst is with Verb Surgical Inc., Mountain View, CA 94043 USA (e-mail: benfuerst@verbsurgical.com).

Mehran Armand is with Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 21218 USA (e-mail: mehran.armand@jhuapl.edu).

Mathias Unberath is with the Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA, and also with the Laboratory for Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: unberath@jhu.edu).

Nassir Navab is with the Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA and with the Laboratory for Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA, and also with the Laboratory for Computer Aided Medical Procedures, Technical University of Munich, 80333 Munich, Germany (e-mail: nassir.navab@jhu.edu).

Digital Object Identifier 10.1109/LRA.2020.2972831

set up. Using an optical see-through head-mounted display (OST-HMD), setup staff can be interactively guided through joint-by-joint steps to optimally position the robot in an efficient manner.

The works by Qian *et al.* are similar to our solution in the spirit of using AR for robotic surgery [5]–[7]; however, the focus of their works were on optimal instrument insertion and manipulation by showing the extension of the arms inside the abdomen using AR. Our methodology addresses the alignment of robot arms for optimal reach and minimum collision. Other early applications of AR in robot-assisted surgery focused on multi-modal registration of medical imaging data with the endoscopic view [8], [9].

Several studies have discussed the challenges of aligning virtual and real objects, and have emphasized the importance of this step for various room-scale and spatially-aware AR solutions [10]–[13]. Nuernberger *et al.* suggested a semi-automatic alignment strategy to register virtual and real spaces [14]. Their work relied on scene content and environment constraints such as edges and surfaces for snapping the virtual content to real. In a different study, various rendering and visualization techniques were compared for alignment of different virtual models in fully immersive environments [15]. Results indicated that static visualization techniques which exhibited lower occlusion in a single view yielded better alignment.

In order to properly augment virtual assistance on a physical robot, we require an intuitive and fast approach to align the AR environment provided by a OST-HMD to the robot. This registration must be robust to perceptual ambiguities that arise during AR alignment [16], [17]. To this end, we propose virtual-real active alignment (ViRAA) to register a virtual model of the robot to its real counterpart. Our method enables the user to create and view multiple AR mirrors which show the current 3D scene (including real and virtual robot) from different view-points. By providing this overlay from multiple perspectives simultaneously, users can actively adjust the 6 degree-of-freedom (DOF) transformation parameters that best align the virtual and real objects in all views.

We summarize the contributions of the present work as 1) reflective-AR displays as the multi-view and marker-free paradigm for co-registration between the virtual and real spaces, hence enabling spatially-aware AR, and 2) using AR for assistance during robot set up (Fig. 1).

## II. METHODOLOGY

An important step in many AR scenarios is to bring the virtual content that lives in a controlled environment into alignment with the physical reality that is present in the unmodeled environment [16]. In this work, to enable seamless interaction of a surgical robot manipulator and its virtual representation during an AR experience, we introduce reflective-AR displays that enable multi-view visualization and interactive alignment of virtual and real objects. In Sec. II-A, we present the problem formulation for registering virtual-to-real. A key contribution of this work, which is the AR reflectors, is presented in Sec. II-B. Finally, in Sec. II-C, we suggest AR guidance to facilitate robot



Fig. 1. AR-assisted robot arm positioning.

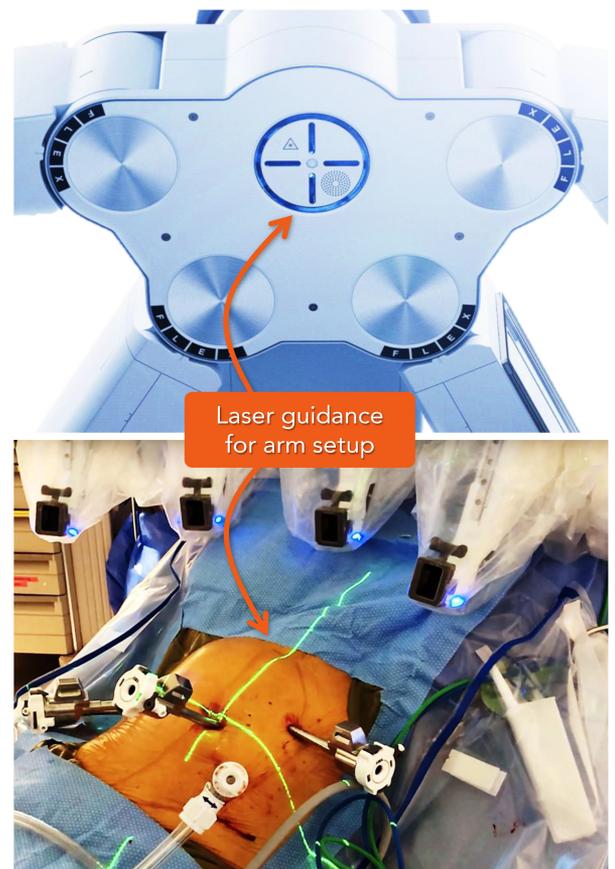


Fig. 2. Da Vinci Xi surgical robot uses multiple lasers to assist the positioning and docking of the robotic arms.

set up during surgical interventions. It is important to note that in Sec. II-A and II-B we discuss the problem of “virtual-to-real” alignment to enable spatially-aware AR, and in Sec. II-C we discuss “real-to-virtual” alignment to provide spatially-aware AR guidance.

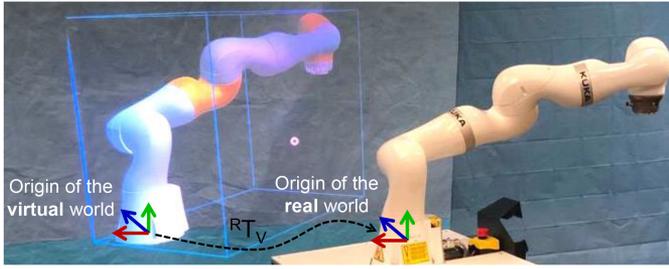


Fig. 3. ViRAAI strategy estimates the virtuality to reality transformation  ${}^R\mathbf{T}_V$ .

#### A. Virtual-Real Active Alignment (ViRAAI)

To estimate the virtual-to-real 6 DOF alignment shown in Fig. 3, we estimate the transformation  ${}^R\mathbf{T}_V = (\bar{R}, \bar{t})$  via interactively registering a robot with its virtual model at  $N$  pre-defined joint configurations. Each time a rigid-body transformation  $\{(R_i, t_i)\}_{i=1}^N$  is obtained, where  $(R_i, t_i) \in SE(3)$ , and  $SE$  is the Special Euclidean group.

We hypothesize that the average transformation computed from these  $N$  estimates, can yield a closer approximation of the true virtual-to-real alignment compared to each individual  $N$  transformations. Hence, from these  $N$  estimates, we seek to compute the mean rotation and translation. The mean rotation matrix  $\bar{R}$  is computed on the Special Orthogonal group  $SO(3)$  by minimizing:

$$\arg \min_{\bar{R} \in SO(3)} \sum_{i=1}^N d(R_i, \bar{R})^2, \quad (1)$$

where  $d(\cdot)$  denotes a distance function on the Riemannian manifold. To establish  $d(\cdot)$ , the rotation matrix is expressed in the Lie algebra (tangent space) of the Lie group as  $R = e^{\hat{w}}$ . The tangent space  $\mathbf{w}$  is then obtained as  $\log(R) = \hat{w}$ , such that  $\hat{w}$  is the skew-symmetric matrix constructed from the vector  $\mathbf{w}$ . Consequently, the mean rotation is estimated as [18]:

$$\arg \min_{\bar{R} \in SO(3)} \sum_{i=1}^N \|\log(R_i^T \bar{R})\|_F^2, \quad (2)$$

where  $\|\cdot\|_F^2$  is the Frobenius norm. The mean translation  $\bar{t}$  is computed in Euclidean space as:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^N \mathbf{t}_i. \quad (3)$$

#### B. Reflective-AR Display

Due to the projective property of human visual system and the differences in perceptual cues in virtuality and reality, the scale and depth between real and virtual objects are easily misjudged [17]. To overcome depth ambiguities and enhance 3D perception during an AR experience, we introduce reflective-AR displays that allow simultaneous visualization of the scene from various viewpoints. The reflective-AR displays shown in Fig. 4 are constructed by displaying images from the integrated camera sensor of the OST-HMD as if the user observed the

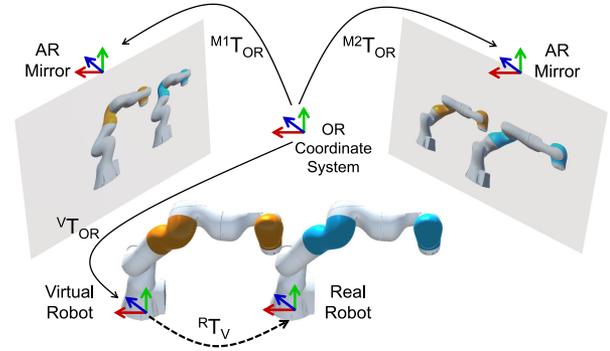


Fig. 4. Reflective-AR displays enable simultaneous alignment from multiple views.

real scene from different viewpoints simultaneously, and are augmented with the projections of the 3D virtual objects. To compute a geometrically relevant pose for displaying these images, we compute the associated observer poses to the coordinate frame of the AR scene in the operating room ( ${}^O R_{OR}, {}^O \mathbf{t}_{OR}$ ) via simultaneous-localization and mapping (SLAM).

The observer imaging geometry in Fig. 5 is formulated as:

$$P_o = K_o P \begin{bmatrix} {}^O R_{OR} & {}^O \mathbf{t}_{OR} \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (4)$$

where  $K_o$  is the matrix of intrinsic parameters and  $P$  is the projection operator. Next, to simulate a mirror-like view, we construct a reversed frustum as (Fig. 5):

$$P_m = K_m P \begin{bmatrix} {}^O R_{OR} & {}^O \mathbf{t}_{OR} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} R_x(\pi) & \begin{bmatrix} 0 \\ 0 \\ D \end{bmatrix} \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (5)$$

In Eq. 5, the optical center of the observer frustum is rotated by the amount  $\pi$  around the  $x$  axis, and translated by the amount  $D$  along the principle ray of the frustum. Distance  $D$  is approximated as the Euclidean distance between the camera center, and an arbitrary point on the surface of the robot that is acquired by colliding the gaze cursor with the spatial map of the AR scene. The distance is merely used as a reference to position the optical center of the reversed frustum, and does not affect the rendering content in the reflective-AR display. To compute the reversed frustum's intrinsic matrix  $K_m$ , the  $y$ -axis of the image plane is flipped according to Eq. 5. Lastly, to give rise to a mirror-like AR display, the 3D virtual structures are projected into the image plane with the imaging geometry  $P_m$ , thus, enabling joint visualization of real and virtual in the reflective display.

Since the reflective AR displays are constructed based on the imaging geometry of the observer frustum, we adopted the same

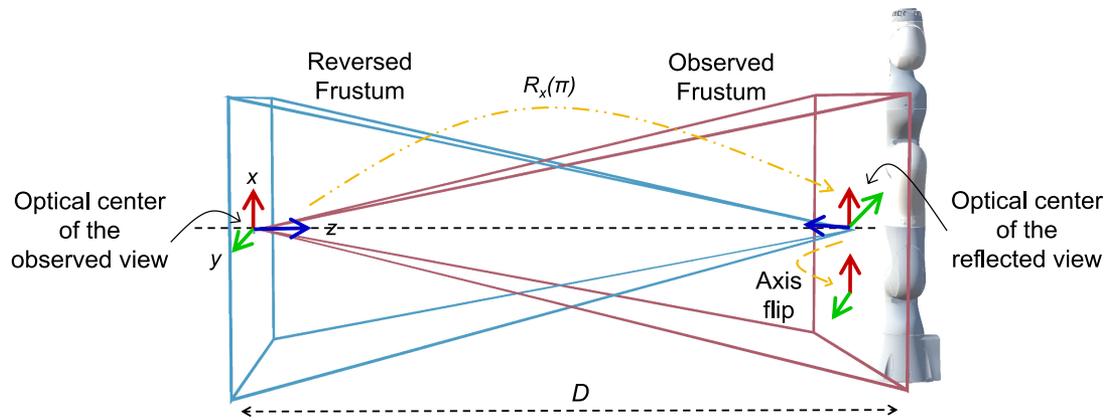


Fig. 5. Imaging geometries of the observed and reversed frustums in relation to the robotic manipulator.

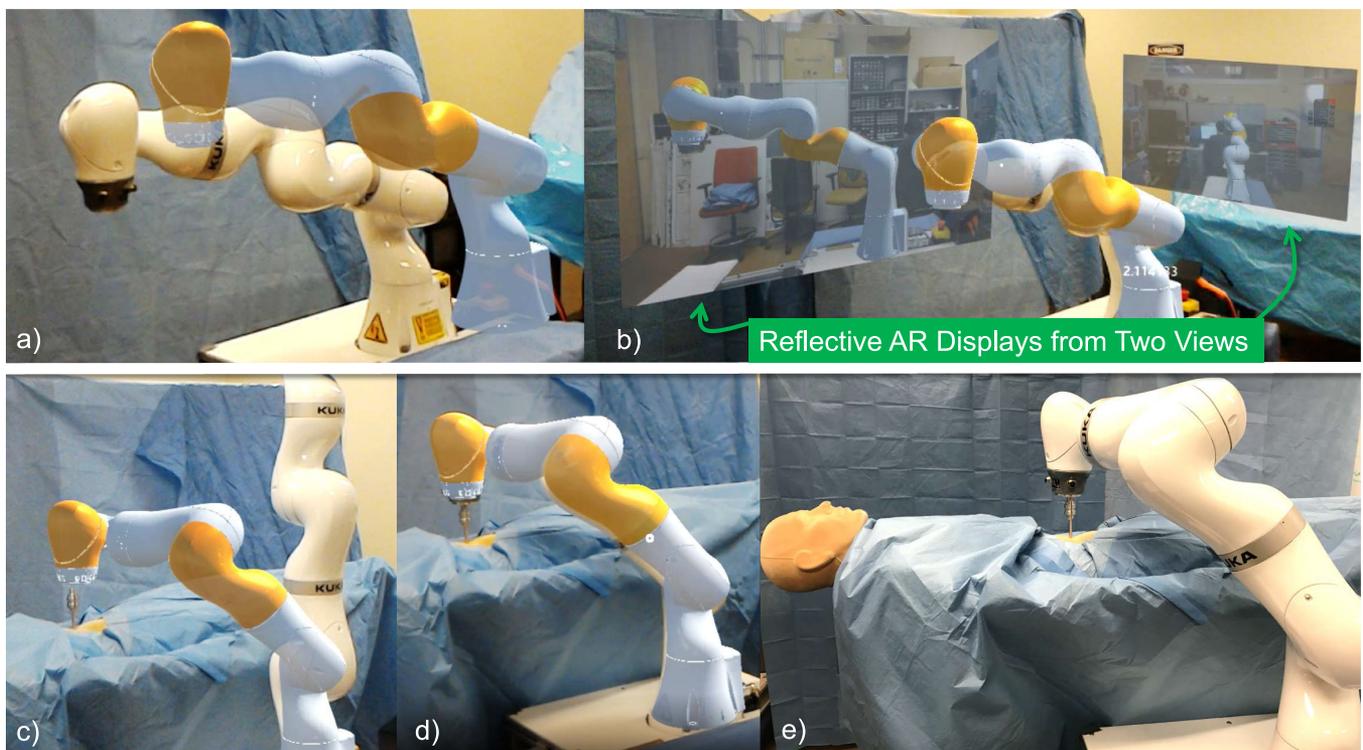


Fig. 6. During the surgical AR experience, the virtual model of the robot is first visualized at a known configuration (a). The alignment between the real and virtual is established in multiple views via reflective-AR displays (b). Once the 6 DOF rigid-body transformation is identified between the real and virtual content, a virtual robot is rendered into the scene at a safe surgical configuration (c). The robot assistant can then align the robot with the virtual counterpart (d), and dock it to the trocar (e).

axis convention used in the computer vision community. We set the z-axis in the direction pointing away from the camera, along the principal ray, connecting the origin to the principal point on the image plane.

### C. Augmented Reality Assistance for Robot Set Up

After the registration transformation is established between real and virtual worlds, a collision-free and safe virtual robot configuration can be presented to the medical assistant. The

desired configuration can either be estimated via the inverse kinematics of the robot, or can be adjusted interactively using the virtual robot and the patient position on the surgical bed. The robot set up is then performed in joint-by-joint steps, following the virtual planning.

## III. EXPERIMENTAL RESULTS

We first evaluate the ViRAAI strategy using a virtual and real robotic manipulator. Next, in a simulated surgical setup, we

TABLE I  
MEAN AND STANDARD DEVIATION OF MISALIGNMENT ERRORS IN MM

Alignment Method	$(t_x, \sigma_{t_x})$	$(t_y, \sigma_{t_y})$	$(t_z, \sigma_{t_z})$	$(\ \tilde{t}\ _2, \ \sigma_t\ _2)$
ViRAAI	(17.4, 16.1)	(11.9, 6.24)	(21.6, 16.5)	(30.2, 23.9)
ViRAAI + Reflective-AR Display	(9.00, 5.64)	(10.3, 7.45)	(9.18, 5.77)	(16.5, 11.0)

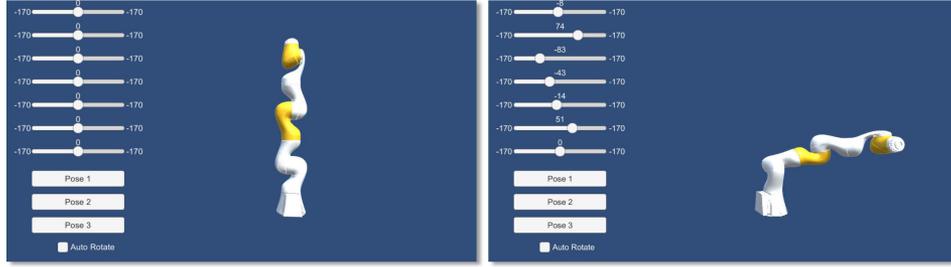


Fig. 7. Interactive fixed display demonstrating two different target joint configurations.

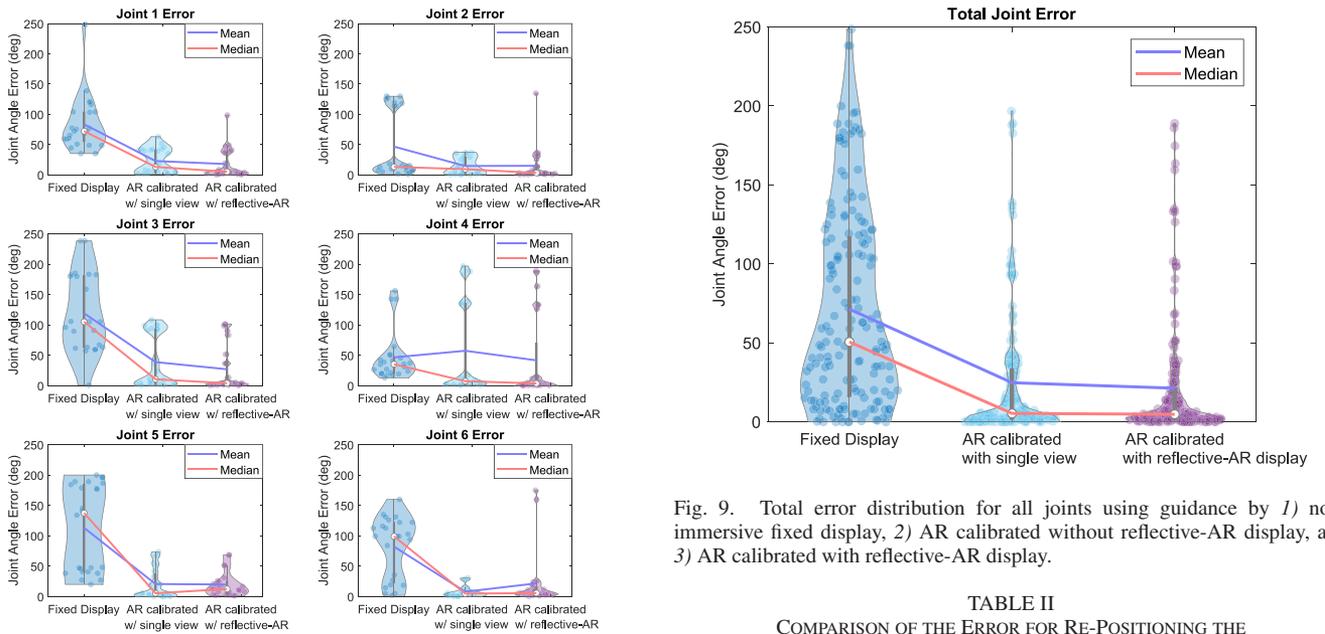


Fig. 8. Distribution of errors evaluated for each joint separately when guided by 1) non-immersive fixed display, 2) AR calibrated without reflective-AR display, and 3) AR calibrated with reflective-AR display.

assess the errors in moving the robot joints to achieve a desired joint configuration using AR guidance where a trocar must be inserted at a mannequin's umbilicus (Fig. 6). The umbilicus is commonly chosen as a robotic port and remote center of motion (RCM) for abdominal surgery. Training for port and trocar placement in umbilicus and optimal docking of the robot has a steep learning curve [19], [20].

### A. System and Design

For the experiments we used a 7 DOF KUKA LBR Intelligent Industrial Work Assistant (iiwa) 7 R800 redundant robot manipulator (KUKA AG, Augsburg, Germany). The joint configuration and end-effector pose of the robot was obtained

Fig. 9. Total error distribution for all joints using guidance by 1) non-immersive fixed display, 2) AR calibrated without reflective-AR display, and 3) AR calibrated with reflective-AR display.

TABLE II  
COMPARISON OF THE ERROR FOR RE-POSITIONING THE ROBOT JOINTS IN DEGREE UNITS

Joint Error	Mean	Median	Min	Max	Std
Reflective AR Display	23.7	4.93	0.02	188	41.1
AR	26.8	5.88	0.00	197	42.4
Fixed Display	71.4	50.6	0.64	249	61.8

through a ROS interface [21]. It is important to note that our solution is designed to address challenges in surgical settings, and the KUKA arm was merely used as an exemplary robot that was available for this research. The AR environment was delivered by a first-generation Microsoft HoloLens OST-HMD (Microsoft, Redmond, WA).

### B. Alignment of Virtual-to-Real

The ViRAAI strategy is evaluated by aligning the virtual and real robots with and without a reflective-AR display. Each experiment is repeated 10 times for 4 users. The error measurements

TABLE III  
P-VALUES FOR EACH INDIVIDUAL JOINT, AS WELL AS FOR ALL JOINTS COMBINED

P-value	Joint 1	Joint 2	Joint 3	Joint 4	Joint 5	Joint 6	Total
Reflective AR Display / AR	0.61	0.50	0.23	0.71	0.59	0.27	0.54
Reflective AR Display / Fixed Display	0.12e−6	0.15e−1	0.24e−6	0.77	0.22e−6	0.76e−4	0.12e−15
AR / Fixed Display	0.16e−6	0.39e−1	0.62e−4	0.87	0.12e−6	0.58e−7	0.20e−17

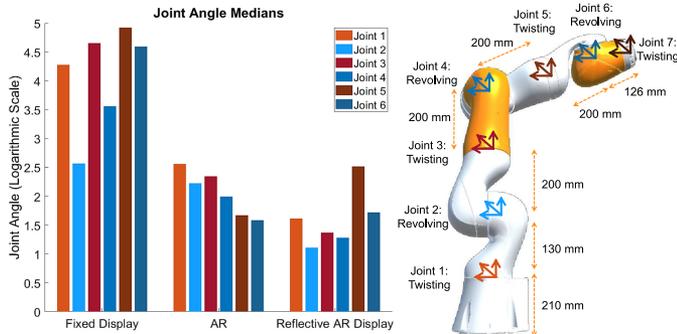


Fig. 10. The plot demonstrates an abstract comparison between the errors contributed by each joint. Since AR-based approaches yielded substantially smaller errors, we used Logarithmic scale for optimal visualization and comparison of errors with different orders of magnitude. Revolving joints with even indexes are shown in blue, and twisting joints with odd indexes are shown in red.

are presented in Table I. We did not incorporate an external marker-based tracking approach as the base-line since marker tracking exhibits high errors due to propagation, and does not include the user in the loop, i.e. it only determines the registration error and not the augmentation error in the user’s view. Instead, to quantify the amount of misalignment, for each iteration we located three pairs of distinct 3D landmarks on the surfaces of both real and virtual robots. These points were identified interactively by intersecting rays from the OST-HMD to the landmark using a gaze cursor.

We define each ray  $i$  using the position of the user’s head  $\mathbf{h}_i$ , and the unit direction vector  $\mathbf{u}_i$  from the head to the annotated landmark on the spatial map of the environment. The landmark  $\mathbf{x}_i^*$  is estimated in a least-squares fashion from the intersection of two rays as:

$$\mathbf{x}_i^* = \arg \min_{\mathbf{x} \in \mathbb{R}^3} \sum_{i=1}^2 \|(I_3 - \mathbf{u}_i \mathbf{u}_i^T)(\mathbf{x} - \mathbf{h}_i)\|^2. \quad (6)$$

To quantify the error of our ground-truth measurement mechanism using 3D landmarks, we computed the Euclidean distance between different sets of targets on an optical table for a total of 12 times. We selected four combinations of landmarks which were 5 cm, 10 cm, 15 cm, and 20 cm apart. The average error for measuring distances using AR annotations was 3.6 mm.

For virtual-to-real object alignment, the results in Table I indicate a total error of  $16.5 \pm 11.0$  mm when using the reflective-AR display, and  $30.2 \pm 23.9$  mm when using AR without the additional mirror view. To demonstrate the change in alignment error when averaging multiple alignment transformations on the  $SE(3)$  manifold as presented in Sec. II-A, we computed the average transformation given Eq. 2 and Eq. 3 when using the AR reflective display. This experiment yielded a total error of

$11.3 \pm 1.01$  mm, which is lower than each individual alignment trial.

### C. Augmented Reality for Robot Set Up: Accuracy Analysis

During a simulated robot-assisted trocar placement,  $U = 8$  users moved the robot joints to achieve different target joint configurations. Each user performed this task with the guidance from 1) AR, 2) AR with reflective display, and 3) interactive fixed display, all in randomized orders. The fixed external monitor was used as a non-immersive baseline, displaying the desired robot configurations to the user at all time during the execution, and allowing the users to interact with the visualization by rotating or scaling (Fig. 7). The users performed each test three times, and each time with a different target joint. To set up AR we used the average Euclidean transformation over  $N = 3$  trials as described in Eq. 2 and Eq. 3. Errors in joint angles are demonstrated in Fig. 8. The total errors are shown in Fig. 9. The violin plots show the distribution of error within each guidance method.

The statistics for the overall error in joint positioning using each guidance techniques is compared in Table II. Statistical significance measures are shown in Table III. In Fig. 10 we present a comparison between the errors from each joint, and highlight the errors for twisting and revolving joints. A joint is characterized as twisting if the axis of rotation is parallel to the robot link, and revolving if the axis of rotation is orthogonal to the robot link.

### D. Augmented Reality for Robot Set Up: Time Analysis

Table IV presents the observed time for all eight users. Guidance using AR and AR reflector both required registration between the virtual and real content. Therefore, each time the users were given a unique joint configuration target, prior to moving the real robot, the user aligned the virtual model of the robot with its real counterpart three times  $\{\mathbf{T}_V^R(i)\}_{i=1}^3$ . Using the ViRAAI approach presented in Sec. II-A, the transformation  $\overline{\mathbf{T}_V^R}$  which expressed the geometric mean in  $SE(3)$  was computed. The average transformation was applied to the AR scene to register the origins of the real and virtual environments, and enable AR guidance. Table IV also presents the time required for ViRAAI with and without AR reflectors for all users.

## IV. DISCUSSION

The experimental results indicated an improved alignment when using a reflective-AR display. The L2-norm average misalignment error in this case was  $16.5 \pm 11.0$  mm, and showed improvement compared to  $30.2 \pm 23.9$  mm error when no reflective displays were used. Averaging the transformations on  $SE(3)$  manifold yielded an even lower error of  $11.3 \pm 1.01$  mm. This alignment error does not seem sufficient for tasks that

TABLE IV  
TIME REQUIRED FOR ViRAAL AND RE-POSITIONING THE ROBOT JOINTS IN MINUTE:SECOND UNITS

	Execution Time					Alignment Time				
	Mean	Median	Min	Max	Std	Mean	Median	Min	Max	Std
<b>Reflective AR Display</b>	2 : 00	1 : 50	0 : 39	3 : 47	0 : 50	4 : 32	4 : 42	2 : 06	6 : 59	1 : 41
<b>AR</b>	1 : 31	1 : 23	0 : 22	5 : 06	0 : 59	2 : 29	1 : 52	1 : 11	4 : 26	1 : 12
<b>Fixed Display</b>	1 : 34	1 : 35	0 : 27	3 : 20	0 : 48	-	-	-	-	-

require high accuracy such as defining biopsy targets in AR, however seems acceptable for providing intuition during robotic arm set up. Using more than one reflective-AR display did not improve the alignment due to two main reasons. First, the limited field of view of Microsoft HoloLens prohibited optimal view of multiple mirrors in their frustums simultaneously when standing in close proximity to the robot. Second, the poor quality of SLAM-based tracking and the unreliable spatial map of the HMD resulted in drifts, hence achieving alignment consensus in all views became challenging. Larger field of view, reliable head tracking, enhanced form factor, enhanced gesture input, and eye tracking capabilities can greatly improve the current limitations.

We evaluated the results in Table I based on a novel user-in-the-loop concept. Alternatively, these errors can be evaluated by using an external navigation system and fiducials. However, the latter will exclude the user and will only evaluate the registration error, instead of the augmentation error. Considering the large improvement (>45%) reported across multiple trials in Table I, we expect the conclusion will not change when using optical navigation.

For modern surgical platforms, the first assistants are trained to set up, dock, tear down, and re-configure the robot using extensive pre-operative E-learning or instructor-led training. In addition to the training that they receive by the manufacturer, general guidelines, and demonstrations in the form of text or visualizations are available in the operating room. In our experiments, to exclude the bias of training, we substituted the conventional training-based approach by recruiting users with no background in setting up surgical robots and focused on demonstrating the effectiveness of intra-operative AR guidance. Comparison with base-line training is a subject of future work, which requires randomized studies in clinical settings with larger populations.

The alignment between the real and virtual, which establishes the registration, is an entirely user-dependent step as the registration chain implicitly takes into account the internal relations between the users eyes, AR camera on the headset, and the AR displays. Since human visual systems are different, these relations differ, and consequently cannot be done in a user-out-the-loop setting. Therefore, automatic approaches will not generalize.

Fig. 9 presented the misalignment errors for bringing the real robot to a desired configuration. The expected accuracy for the robotic set up depends on the design, number of arms, and number of joints of the particular surgical system. These parameters all vary depending on the surgical use case. The higher the number of joints and arms, the higher the chance for collision; therefore, higher accuracy is demanded for a more complex system. This step during the AR-assisted workflow,

which involves the alignment of a real object to virtual, can in future leverage from multi-view strategies by using collaborative AR devices or external cameras.

Several AR publications have shown that time-saving of AR cannot be quantified immediately with dedicated user studies [22], partly because of the unfamiliar interface and exposure to additional information. Time-saving only manifests after the user is proficient with the system. We hypothesize that while there is overhead in setting up AR, the rate of failure/collision would drop leading to a net reduction of overall junk-time. The Riemannian averaging in Eq. 2 and Eq. 3 can increase this set up time, but is a one-time process which takes place before intervention and can result in a more accurate fusion of information and improved AR experience. We hope AR assistance minimizes the training time and allows operators to verify and inspect the proper alignment of robot arms, both quantitatively and visually.

HMD-based AR is challenged by recurrent estimation of transformations and corrections such as user-to-display and drift, respectively. There is a great wealth of literature directed towards addressing these issues [23], which our solution benefits from as new hardware becomes available. It should be noted that drift will be quite limited because the working volume near the patient bed and the robot is restricted.

Our proposed approach enables the co-registration between the real and virtual spaces and delivers spatially-aware AR. We also demonstrated the application of ViRAAL for AR guidance during minimally-invasive robotic surgery. The estimation of the overall registration greatly benefits from the averaging strategy presented in Eqs. 1–3, which are suggested to compute the mean transformation that satisfies the properties of  $SE(3)$  manifold. This average estimate is a rigid transformation that has the shortest distance to all other estimates around the true pose.

The range of errors exhibited by all intra-operative guidance methods, particularly by the non-immersive fixed display, prove the complexity and importance of this problem for robot manipulation. We computed p-values and compared all pairs of methods in Table III to identify the most effective assistive approach. Statistical significance was considered if  $p < 0.05$ . Results suggested that guidance using AR with and without reflective display yielded significantly lower errors compared to non-immersive fixed display. The AR guidance approach using reflective displays outperformed the AR system with no mirrors, however in this comparison statistical significance was not achieved.

In Fig. 10, we compared the error contributed by revolving and twisting joints separately. Results indicated that re-positioning of revolving joints in all three guidance methods are consistently more accurate than the twisting joints. We hypothesize that the higher error is the result of the inherent symmetry in twisting motion that may lead to ambiguities.

## V. CONCLUSION

In this manuscript we presented a novel multi-view strategy to align virtual and real content, and demonstrated an application of it for improving surgical robotic workflows. The reflective-AR displays were introduced to eliminate the 3D scale ambiguities and improve the AR scene realism. We have demonstrated an AR interface that accommodates multiple reflective displays, and allows the users to scale the images within their viewing frustum [24].

The virtual-to-real registration approach, ViRAAI, is an interactive and user-specific method that calibrates the real and virtual worlds directly to the user's display. No external camera or tracking system, other than the HMD itself, is used in order to keep maximum flexibility and transferability of the system into different surgical environments.

Seamless overlays of virtual content onto the reflective AR displays are achieved by placing virtual cameras at the optical centers where the images were acquired, hence allowing to render virtual and real from an identical imaging geometry. The reflective displays require a static scene, therefore, are suited for aligning virtual-to-real, and not vice-versa.

The focus of this work is beyond the KUKA robot and its redundant design; it is instead on complex surgical platforms with multiple arms and various joints. The task of alignment is expected to be more difficult for redundant manipulators with more joints. Nonetheless, we expect AR to provide an effective guidance mechanism to reconfigure complex redundant arms at the bedside.

Surgical robots are only certified to be controlled by the surgeon in its active mode, and due to safety reasons, their set up by the surgical staff are performed entirely manually. Our solution is designed around this concept of full manual interaction. It should be noted that our contribution is not on computing joint configurations that minimize collision, but instead we show that if such configuration exists, then with the support of AR it can be manually achieved by the first assistant.

A user interface such as the AR reflectors, can accelerate interaction during surgery. By measuring the exact time of the staff this could be validated in future work. We believe that the proposed alignment strategy can extend to other realms of computer-assisted surgery, namely for surgical training and AR guidance during image-guided therapies.

## REFERENCES

- [1] R. v. van der Schatte Olivier, C. vant Hullenaar, J. Ruurda, and I. Broeders, "Ergonomics, user comfort, and performance in standard and robot-assisted laparoscopic surgery," *Surgical Endoscopy*, vol. 23, no. 6, 2009, Art. no. 1365.
- [2] J. Jayakumar, S. D. Patel, B. K. Gangrade, D. M. Narasimulu, S. R. Pandian, and C. Silva, "Robotic-assisted laparoscopy in reproductive surgery: A contemporary review," *J. Robot. Surgery*, vol. 11, no. 2, pp. 97–109, 2017.
- [3] R. M. Juza, J. R. Lyn-Sue, and E. M. Pauli, "Intraoperative considerations for robotic repair," in *Laparoscopic and Robotic Incisional Hernia Repair*. Berlin, Germany: Springer, 2018, pp. 103–115.
- [4] H. Liu, T. Kinoshita, A. Tonouchi, A. Kaito, and M. Tokunaga, "What are the reasons for a longer operation time in robotic gastrectomy than in laparoscopic gastrectomy for stomach cancer?" *Surgical Endoscopy*, vol. 33, no. 1, pp. 192–198, 2019.
- [5] L. Qian, A. Deguet, Z. Wang, Y.-H. Liu, and P. Kazanzides, "Augmented reality assisted instrument insertion and tool manipulation for the first assistant in robotic surgery," in *Proc. Int. Conf. Robot. Autom.*, 2019, pp. 5173–5179.
- [6] L. Qian, X. Zhang, A. Deguet, and P. Kazanzides, "Aramis: Augmented reality assistance for minimally invasive surgery using a head-mounted display," in *Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2019, pp. 74–82.
- [7] L. Qian, A. Deguet, and P. Kazanzides, "Arssist: Augmented reality on a head-mounted display for the first assistant in robotic surgery," *Healthcare Technol. Lett.*, vol. 5, no. 5, pp. 194–200, 2018.
- [8] O. Mohareri *et al.*, "Ultrasound-based image guidance for robot-assisted laparoscopic radical prostatectomy: Initial in-vivo results," in *Int. Conf. Inf. Process. Comput.-Assisted Interventions*, 2013, pp. 40–50.
- [9] P. Pessaux, M. Diana, L. Soler, T. Piardi, D. Mutter, and J. Marescaux, "Towards cybernetic surgery: Robotic and augmented reality-assisted liver segmentectomy," *Langenbeck's Archives Surgery*, vol. 400, no. 3, pp. 381–385, 2015.
- [10] D. Reiners, D. Stricker, G. Klinker, and S. Müller, "Augmented reality for construction tasks: Doorlock assembly," in *Proc. IEEE ACM Int. Workshop Augmented Reality*, 1998, vol. 98, no. 1, pp. 31–46.
- [11] S. J. Henderson and S. K. Feiner, "Augmented reality in the psychomotor phase of a procedural task," in *Proc. 10th IEEE Int. Symp. Mixed Augmented Reality*, 2011, pp. 191–200.
- [12] S. Feiner, B. Macintyre, and D. Seligmann, "Knowledge-based augmented reality," *Commun. ACM*, vol. 36, no. 7, pp. 53–62, 1993.
- [13] T. P. Caudell and D. W. Mizell, "Augmented reality: An application of heads-up display technology to manual manufacturing processes," in *Proc. 25th Hawaii Int. Conf. Syst. Sci.*, 1992, vol. 2, pp. 659–669.
- [14] B. Nuernberger, E. Ofek, H. Benko, and A. D. Wilson, "Snaptoreality: Aligning augmented reality to the real world," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2016, pp. 1233–1244.
- [15] A. Martin-Gomez, U. Eck, and N. Navab, "Visualization Techniques for Precise Alignment in VR: A Comparative Study," in *Proc. IEEE Virtual Reality 3D User Interfaces*, 2019, pp. 735–741.
- [16] P. Milgram and D. Drascic, "Perceptual effects in aligning virtual and real objects in augmented reality displays," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, 1997, vol. 41, no. 2, pp. 1239–1243.
- [17] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr, "Effects of stereo viewing conditions on distance perception in virtual environments," *Presence: Teleoperators Virtual Environ.*, vol. 17, no. 1, pp. 91–101, 2008.
- [18] M. Moakher, "Means and averaging in the group of rotations," *SIAM J. Matrix Anal. Appl.*, vol. 24, no. 1, pp. 1–16, 2002.
- [19] C. Chang, Z. Steinberg, A. Shah, and M. S. Gundeti, "Patient positioning and port placement for robot-assisted surgery," *J. Endourology*, vol. 28, no. 6, pp. 631–638, 2014.
- [20] P. Iranmanesh, P. Morel, O. J. Wagner, I. Inan, F. Pugin, and M. E. Hagen, "Set-up and docking of the da vinci surgical system: Prospective analysis of initial experience," *International J. Med. Robot. Comput. Assisted Surgery*, vol. 6, no. 1, pp. 57–60, 2010.
- [21] M. Safeea and P. Neto, "KUKA sunrise toolbox: Interfacing collaborative robots with MATLAB," *IEEE Robot. Autom. Mag.*, vol. 26, no. 1, pp. 91–96, Mar. 2019.
- [22] A. Dünser and M. Billinghurst, "Evaluating augmented reality systems," in *Handbook of Augmented Reality*. Berlin, Germany: Springer, 2011, pp. 289–307.
- [23] M. Tuceryan, Y. Genc, and N. Navab, "Single-point active alignment method (SPAAM) for optical see-through HMD calibration for augmented reality," *Presence: Teleoperators Virtual Environ.*, vol. 11, no. 3, pp. 259–276, 2002.
- [24] J. Fotouhi *et al.*, "Interactive flying frustums (IFFs): spatially aware surgical data visualization," *Int. J. Comput. Assisted Radiol. Surgery*, vol. 14, no. 6, pp. 913–922, 2019.