

# Hybrid electromagnetic-ArUco tracking of laparoscopic ultrasound transducer in laparoscopic video

Xinyang Liu<sup>1</sup>,<sup>a</sup> William Plishker<sup>2</sup>,<sup>b</sup> and Raj Shekhar<sup>a,b,\*</sup>

<sup>a</sup>Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Hospital,  
Washington, DC, United States

<sup>b</sup>IGI Technologies, Inc., Silver Spring, Maryland, United States

## Abstract

**Purpose:** The purpose of this work was to develop a new method of tracking a laparoscopic ultrasound (LUS) transducer in laparoscopic video by combining the hardware [e.g., electromagnetic (EM)] and the computer vision-based (e.g., ArUco) tracking methods.

**Approach:** We developed a special tracking mount for the imaging tip of the LUS transducer. The mount incorporated an EM sensor and an ArUco pattern registered to it. The hybrid method used ArUco tracking for ArUco-success frames (i.e., frames where ArUco succeeds in detecting the pattern) and used corrected EM tracking for the ArUco-failure frames. The corrected EM tracking result was obtained by applying correction matrices to the original EM tracking result. The correction matrices were calculated in previous ArUco-success frames by comparing the ArUco result and the original EM tracking result.

**Results:** We performed phantom and animal studies to evaluate the performance of our hybrid tracking method. The corrected EM tracking results showed significant improvements over the original EM tracking results. In the animal study, 59.2% frames were ArUco-success frames. For the ArUco-failure frames, mean reprojection errors for the original EM tracking method and for the corrected EM tracking method were 30.8 pixel and 10.3 pixel, respectively.

**Conclusions:** The new hybrid method is more reliable than using ArUco tracking alone and more accurate and practical than using EM tracking alone for tracking the LUS transducer in the laparoscope camera image. The proposed method has the potential to significantly improve tracking performance for LUS-based augmented reality applications.

© 2021 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.8.1.015001](https://doi.org/10.1117/1.JMI.8.1.015001)]

**Keywords:** ArUco; augmented reality; electromagnetic tracking; laparoscopic ultrasound.

Paper 20213R received Aug. 13, 2020; accepted for publication Jan. 12, 2021; published online Feb. 3, 2021.

## 1 Introduction

Laparoscopic surgery is a widely used alternative to conventional open surgery and is known to achieve improved outcomes, cause less scarring, and lead to significantly faster patient recovery.<sup>1,2</sup> Despite this success, surgeons cannot visualize anatomic structures and surgical targets below the exposed organ surfaces in standard laparoscopy. Laparoscopic ultrasound (LUS) imaging provides information on subsurface anatomy, but ultrasound images are presented separately and can only be integrated with the laparoscopic video in the surgeon's mind. Moreover, focus is distracted from the laparoscopy screen when viewing ultrasound images presented on a separate screen. To enhance intraoperative visualization, a number of groups have developed augmented reality (AR) systems that fuse live ultrasound images with laparoscopic video in real time.<sup>3-8</sup> Determining the pose (i.e., position and orientation) of the LUS transducer in the laparoscopic camera coordinate system is essential in these AR applications. Once the pose of the LUS transducer is determined, the coordinates of the ultrasound image in the camera coordinate system can be calculated using ultrasound calibration.<sup>9,10</sup> The ultrasound image can then

\*Address all correspondence to Raj Shekhar, [rshekhar@childrensnational.org](mailto:rshekhar@childrensnational.org)

be projected on the camera image using camera projection matrix obtained through camera calibration.<sup>11</sup> In addition to the AR applications, tracking the pose of the LUS transducer can help register intraoperative ultrasound data with preoperative imaging during laparoscopic liver surgery.<sup>12,13</sup>

Conventional methods to track an LUS transducer include approaches based on either tracking hardware or computer vision (CV). Among tracking hardware, optical tracking and electromagnetic (EM) tracking are the two established real-time tracking methods.<sup>14</sup> For surgical applications, an optical tracking system typically uses an infrared camera to track wireless passive markers, whereas an EM tracking system usually tracks small (~1-mm diameter) wired sensors inside a working volume of a magnetic field created by a field generator. For AR applications based on tracking hardware, the pose sensors are attached to the LUS transducer and the laparoscope such that the sensors maintain a fixed spatial relationship with the respective imaging tips.<sup>4,5,7,15</sup> In comparison, CV-based methods require no special tracking hardware and rely on detecting user-introduced patterns placed on the LUS transducer directly from the laparoscopic camera.<sup>3,6,8,16–18</sup> Another form of reported CV-based methods do not use custom patterns, and instead estimate the LUS probe's pose from the video image alone.<sup>19,20</sup> This “marker-less” approach has been applied to localize other surgical instruments as well.<sup>21</sup>

### 1.1 Hardware-Based Tracking

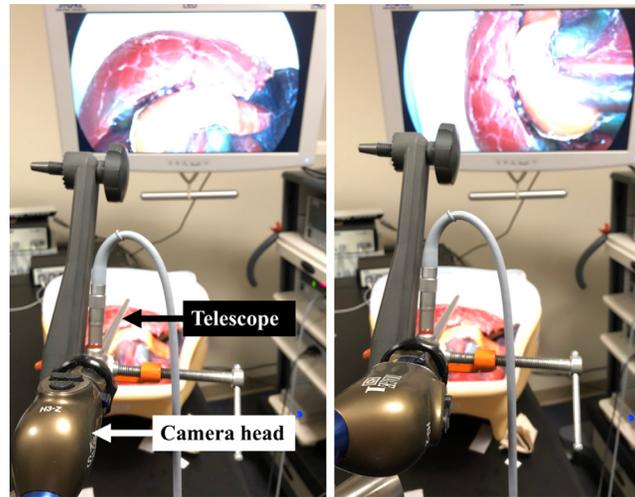
Compared with CV-based tracking methods, hardware-based tracking methods are robust to occlusion and low-quality video images. However, the hardware-based methods have their limitations. In image-guided surgery applications, typically, an object (e.g., ultrasound image) in the  $S_{LUS}$  (sensor attached on the LUS transducer) coordinate system is transformed to the  $S_{Lap}$  (sensor attached to the laparoscope) coordinate system via the tracking hardware, and then to the camera coordinate system through hand-eye calibration.<sup>22</sup> Therefore, tracking hardware error and hand-eye calibration error are inherent to this type of methods. In static, well-controlled, laboratory-based experimental settings, the ultrasound image-to-video target registration errors (TREs) have been reported to be  $3.34 \pm 0.59$  mm (left-eye) and  $2.76 \pm 0.68$  mm (right-eye) for an optical tracking-based stereoscopic AR system,<sup>5</sup> and  $2.59 \pm 0.58$  mm (left-eye) and  $2.43 \pm 0.48$  mm (right-eye) for an EM tracking-based stereoscopic AR system.<sup>7</sup> In a dynamic clinical setting, the TRE is expected to be larger than these numbers.

For AR applications, the ultrasound image is projected on the video image based on camera calibration. If the camera optics, such as the zoom, are changed during the procedure, the camera needs to be recalibrated. For hardware-based tracking, this often means withdrawing the laparoscope from the patient's body and performing camera calibration mid-surgery in the operating room (OR), which is not practical.

Another limitation of hardware-based tracking becomes apparent when it is applied to the commonly used oblique-viewing laparoscopes. Compared with forward-viewing (i.e., 0 deg) laparoscopes, the oblique-viewing laparoscopes have an angled (e.g., 30 deg) lens relative to the camera. During a laparoscopic procedure, the surgeon usually holds the camera head relatively steady and rotates the telescope to expand the surgical field of view. This relative rotation can be modeled equivalently by holding the telescope steady and rotating the camera head. As shown in Fig. 1, the camera image rotates about a rotation center in the video image plane in this case.

The relative rotation between the telescope and the camera head changes the camera optics and the hand-eye calibration, creating a rotational offset to any virtual object overlaid on the video image. Current hardware-based solutions to correct this offset include attaching two pose sensors, one on the telescope and another on the camera head,<sup>23–25</sup> or using a rotary encoder<sup>26</sup> to track the relative rotation. Although demonstrated in the laboratory setting, these approaches are generally not practical for the OR use.

In addition to the above-mentioned common limitations, optical tracking is limited to tracking only a rigid LUS transducer because of the line-of-sight requirement, whereas EM tracking accuracy may be impaired by the distortion of the magnetic field created by ferrous metals or conductive materials inside the working volume.<sup>27</sup>



**Fig. 1** The telescope of a 30-deg-laparoscope was fixed by a clamp. The camera head was rotated  $\sim 90$  deg in the physical space. The camera image was observed to rotate by the same angle ( $\sim 90$  deg) about a rotation center in the image plane.

## 1.2 Computer Vision-Based Tracking

CV-based methods do not need tracking hardware and they can be more accurate in tracking the tool in the video image if the CV marker/feature is not occluded and detectable in the image. Compared with the marker-less approaches, CV pattern-based methods are in general more accurate and robust in extracting 3D poses in the camera coordinate system. For example, a CV marker-based AR system was reported to achieve a TRE of 1.1 to 1.3 mm for a monocular setup and 0.9 to 1.1 mm for a stereoscopic setup.<sup>8</sup> Some of these patterns, such as the checkerboard<sup>16,17</sup> and the circular dot pattern,<sup>6</sup> need to be on a flat surface. Some patterns, on the other hand, can work with a cylindrical surface.<sup>3,8,18</sup> Despite these advances, CV-based methods can still be unreliable as the sole tracking method in a complex, dynamic surgical environment. The patterns can be occluded by a variety of sources, such as the organ tissues, surgical tools, blood, and smoke. Lighting conditions and the specular reflection of light may also obscure pattern detection. The camera may also lose focus on the pattern if the laparoscope or the LUS probe is moved fast.

## 1.3 Contribution

We present a new method of tracking the LUS transducer in laparoscopic video by combining the hardware- and the CV-based tracking methods, and refer to it as hybrid tracking. Because AR is our motivating application, our purpose focused on tracking the LUS transducer in the laparoscope video image. Because we focused on camera space, CV-based methods are inherently more advantageous compared to hardware-based methods in terms of accuracy. For the tracking hardware, we chose EM tracking to track a common LUS transducer with an articulating imaging tip. For the CV-based tracking, we chose the ArUco marker<sup>28,29</sup> for its popularity within the general AR community and ease of implementation with OpenCV.<sup>30</sup> In addition to the ArUco pattern, our method can use other patterns as well, such as the ARToolKit pattern<sup>31</sup> or the ARTag pattern.<sup>32</sup>

The ArUco marker is a flat synthetic square composed of a wide black border and an inner binary matrix that determines its identifier (ID). The ArUco library first detects the corners of the markers in the camera image. If all four corners of a marker are detected, the marker identification is attempted to match it to a particular predetermined pattern. Once the marker is identified, its pose relative to the camera can be estimated by iteratively minimizing the reprojection error of the corners using the Levenberg–Marquardt algorithm.<sup>33</sup> To improve accuracy and robustness, multiple markers can be assembled to form an ArUco board (AB), which can be a single flat surface or a combination of multiple contiguous flat surfaces of known geometry.

The ArUco software estimates the pose of the AB using all identified markers. The more markers that are identified the more accurate the pose estimation of the board is. Reprojection error comparing the detected corners of the identified markers and the reprojected corners based on the estimated pose is given by the ArUco library.

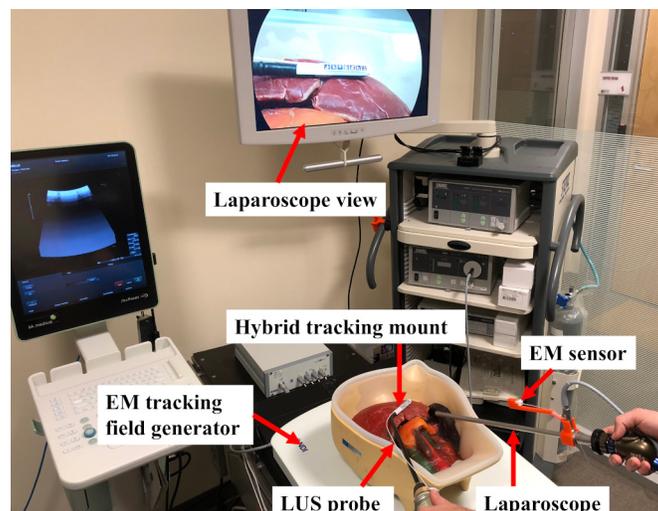
The proposed method was validated first on a visually realistic abdominal phantom and then on an *in-vivo* porcine model. Special cases such as introducing distortion to the EM field, changing the laparoscope camera zoom, and rotating the telescope relative to the camera head were considered during these experiments. Through these experiments, we demonstrated that our proposed hybrid tracking method is more accurate than using the hardware-based method alone and more reliable than using the CV-based method alone. Our hybrid method was inspired by some previous works. For example, Schneider et al.<sup>17</sup> compared the EM tracking method with the CV-based tracking method for a pick-up ultrasound transducer, but there was no discussion of combining the two tracking methods. Although Tella et al.<sup>34</sup> integrated EM tracking data with visual data from laparoscopic camera images, their application was image mosaicking but not surgical instrument tracking. Unlike this work, our preliminary idea of hybrid tracking was to use CV technique without any markers;<sup>20</sup> however, the resulting accuracy, robustness, and computational time were not acceptable for practical use.

## 2 Method

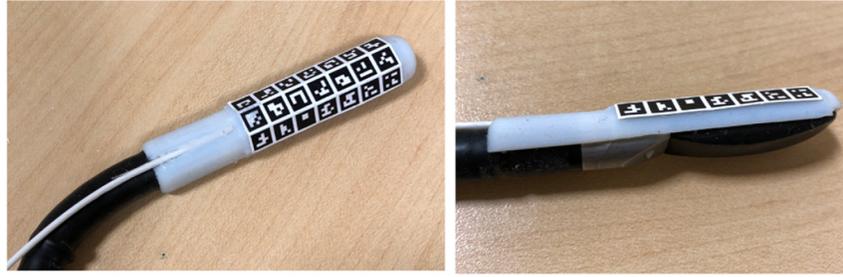
### 2.1 System Setup

As shown in Fig. 2, the study used a standard laparoscopic vision system (Image 1 Hub; KARL STORZ, Germany) with 0 deg and 30 deg 10-mm laparoscopes; an ultrasound scanner (Flex Focus 700; BK Ultrasound, Analogic, Peabody, Manchester) with a four-way articulating LUS transducer (8666-RF); and an EM tracking system with the Tabletop field generator (Aurora; Northern Digital, Waterloo, Ontario, Canada). To track the laparoscope using EM tracking, a custom-designed tracking mount, containing a six degrees-of-freedom (DOF) EM sensor, was fixed on the handle of the laparoscope.<sup>7</sup>

Because ArUco tracking requires the ArUco markers to be flat, we designed and 3D printed a hybrid tracking mount, which contains a six-DOF EM sensor and three flat surfaces for attaching ArUco markers (Fig. 3). The mount was designed to maximize the area of flat surfaces while keeping it as clinically feasible as possible. Specifically, the transducer with the mount can be introduced through a 12-mm trocar, the same size used for the original transducer without the mount. An AB with  $3 \times 7$ , 4.5-mm markers with different IDs were fixed on the hybrid tracking mount.



**Fig. 2** The imaging and tracking devices for demonstrating hybrid tracking.



**Fig. 3** The hybrid tracking mount for the LUS transducer. The mount contains a six-DOF EM sensor and an AB with 21 markers fixed on three flat surfaces.

### 2.2 EM Tracking Approach

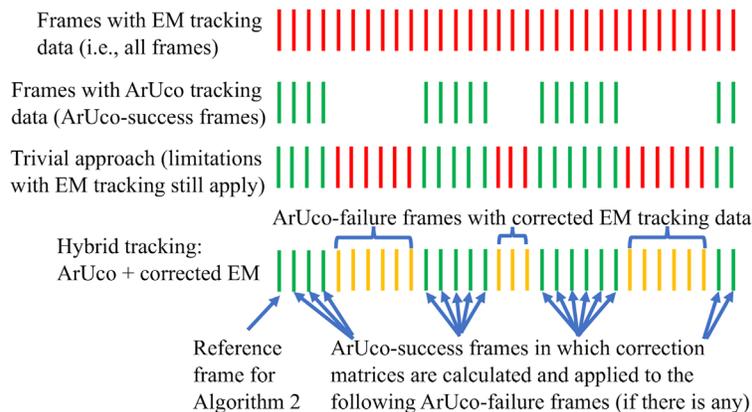
With the developed AB, tracking the LUS transducer in the laparoscope camera space becomes tracking the AB attached to the transducer. To use EM tracking to track the AB, we first acquired the coordinates of the outer corners of the AB in the EM sensor (i.e., the sensor in Fig. 3) coordinate system. This was accomplished using a tracked stylus (Aurora six-DOF Probe). The coordinates of the same corners in the AB coordinate system were known from the design of the AB. The transformation from the AB coordinate system to the EM sensor ( $S_{LUS}$ ) coordinate system  $S_{LUS}T_{AB}$  was determined with a root-mean-square error of 0.38 mm, by registering the two coordinate systems using a SlicerIGT module<sup>35,36</sup> implementing Horn’s algorithm.<sup>37</sup> The transformation from the AB coordinate system to the camera (C) coordinate system using the EM tracking approach ( $C T_{AB}^{(EM)}$ ) can be written as

$$C T_{AB}^{(EM)} = C T_{S_{Lap}} \cdot S_{Lap} T_{EMT} \cdot EMT T_{S_{LUS}} \cdot S_{LUS} T_{AB}, \quad (1)$$

where EMT denotes the EM tracker and  $S_{Lap}$  denotes the sensor attached on the laparoscope as shown in Fig. 2. Based on our previous work,<sup>25</sup>  $C T_{S_{Lap}}$  was obtained using OpenCV’s function of solving the perspective-n-point problem<sup>38</sup> with a special calibration plate. It can be determined using the standard hand-eye calibration as well.

### 2.3 Hybrid Tracking Framework

The general framework of the proposed method is shown in Fig. 4. The first consideration in hybrid tracking is that EM tracking is available at all times, whereas ArUco tracking can be intermittent. Second, we assume that ArUco tracking is more accurate than EM tracking in



**Fig. 4** Overview of the proposed hybrid tracking method.

estimating the pose of the AB in camera space if the ArUco pattern is not occluded and detectable in the video image. The primary idea behind hybrid tracking is to use ArUco tracking if the AB can be successfully recognized by the camera (called ArUco-success) and use what we call corrected EM tracking otherwise (i.e., ArUco-failure). We developed and tested two algorithms that calculate either a single correction matrix (Algorithm 1) or three correction matrices (Algorithm 2) to improve EM tracking results. For an ArUco-success video frame, a correction matrix  $T_{corr}$  is calculated to transform  ${}^C T_{AB}^{(EM)}$  in Eq. (1) to  ${}^C T_{AB}^{(ArUco)}$  (i.e., the transformation from the AB coordinate system to the camera coordinate system through the ArUco tracking approach):

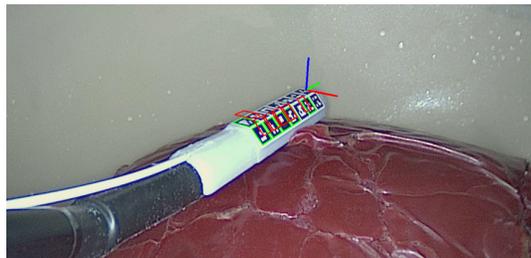
$$\text{Algorithm 1: } {}^C T_{AB}^{(ArUco)} = T_{corr} \cdot {}^C T_{AB}^{(EM)}. \quad (2)$$

Once calculated for the most recent ArUco-success frame,  $T_{corr}$  is applied to correct EM tracking for the following ArUco-failure frames until a new ArUco-success frame appears.

To develop the criteria for determining ArUco-success, we collected developmental data by scanning a tissue-mimicking laparoscopic abdominal phantom (IOUSFAN, Kyoto Kagaku Co. Ltd., Kyoto, Japan) as shown in Fig. 2. The 0-deg laparoscope was calibrated using our single-image calibration method.<sup>39</sup> The camera calibration result is used by the ArUco library to estimate the pose of the AB. Using a frame grabber, we recorded a laparoscopic video (968 frames at a 10-Hz frame rate) of the LUS sweeping the liver surface. After data collection, the ArUco library was used to detect ArUco markers and estimate the pose of the AB for each video frame. Figure 5 shows an example frame from the developmental data.

Based on the results of this experiment, we decided that the first criterion for determining ArUco-success was to have at least two (out of 21) detected ArUco markers. Although ArUco can estimate the entire board pose based on just one marker, such estimation could be susceptible to noise. The chance that two (or more) detected markers are both noise signals is much smaller. For the developmental data, the ArUco library was able to detect at least two markers in 81.7% of the total frames (791 out of 968). Of these qualified frames, the mean number of detected markers was  $5.4 \pm 2.4$  with the maximum being 12. The mean ArUco reprojection error was  $1.51$  ( $\sim 0.2$  mm)  $\pm 1.38$  pixel for full HD resolution ( $1920 \times 1080$  pixels). We can refer to Sec. 4 for how we correlate pixels to distance in the 3D space. Although the AB has three faces, the library usually detected markers on only one or two faces. This is to be expected because not all three faces can be visible to the camera for most LUS probe orientations as shown in Fig. 5.

A second criterion is to limit the reprojection error to a certain threshold  $\epsilon$  to exclude those frames having larger-than-normal reprojection error. Based on the developmental data, we chose  $\epsilon$  to be 2.89 pixel, which is the mean reprojection error plus one standard deviation. We defined reprojection error as the distance between the detected marker corners and their reprojections calculated from the estimated pose of the AB. For reference, the average marker edge length of the detected markers in Fig. 5 is  $\sim 40$  pixels. About 79.1% of the total frames in the developmental data satisfied both criteria.



**Fig. 5** Example frame from the developmental data. The green squares are the reprojected ArUco markers that were successfully detected. The red squares are the projections of the same detected markers using the EM tracking method. To show ArUco has successfully estimated the pose of the AB based on the detected markers, the estimated pose is illustrated as the blue-red-green axes.

## 2.4 Modeling Zoom and Rotation

Algorithm 1 [Eq. (2)] considers no *a priori* information regarding camera zoom and relative rotation of the laparoscope. These parameters can be obtained from the video image and are specifically modeled in Algorithm 2. To model changes to the zoom and rotation parameters of the laparoscope, we used three correction matrices such that Eq. (2) becomes

$$\text{Algorithm 2: } {}^C T_{AB}^{(\text{ArUco})} = T_{corr}^{\text{Zoom}} \cdot T_{corr}^{\theta} \cdot {}^C T_{AB}^{(\text{EM})} \cdot T_{corr}^{\text{AB}}, \quad (3)$$

where  $T_{corr}^{\text{AB}}$  is the correction transformation in the AB coordinate system;  $T_{corr}^{\theta}$  and  $T_{corr}^{\text{Zoom}}$  are the correction transformations in the camera coordinate system to correct offsets introduced by performing the relative rotation and by changing the camera zoom, respectively. From an ArUco-success frame,  $T_{corr}^{\theta}$  and  $T_{corr}^{\text{Zoom}}$  can be obtained from the image features as described below. After  $T_{corr}^{\theta}$  and  $T_{corr}^{\text{Zoom}}$  are obtained, we used Eq. (3) to calculate  $T_{corr}^{\text{AB}}$ .

As shown in Fig. 1, when fixing the telescope while rotating the camera head of a 30-deg laparoscope, the camera image can be observed to rotate around a rotation center in the image plane. In the physical space, this camera head rotation can be modeled by rotating the camera lens coordinate system around a rotation axis.<sup>26</sup> This rotation axis can be approximated as the  $z$  axis of the lens coordinate system. Thus,  $T_{corr}^{\theta}$  can be modeled as a homogeneous rotation matrix

$$T_{corr}^{\theta} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (4)$$

where  $\theta$  is the relative rotation angle.

Because camera zoom is associated with the  $z$  axis of the camera coordinate system,  $T_{corr}^{\text{Zoom}}$  can be modeled to be

$$T_{corr}^{\text{Zoom}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5)$$

where  $\alpha$  is the zoom factor. The rotation angle and the zoom factor can be estimated from the camera image. Figure 6 shows an example frame after rotating the 30-deg laparoscope. The green squares are the ArUco reprojected markers for the detected markers. These align very well with the borders of the ArUco markers on the hybrid tracking mount. The red and black squares are the projections of the corresponding markers using the EM tracking and the reference-adjusted EM tracking methods, respectively.



**Fig. 6** Example frame after rotating the 30-deg laparoscope. The green squares are the ArUco reprojected markers for the detected markers. These align very well with the borders of the ArUco markers on the hybrid tracking mount. The red and black squares are the projections of the corresponding markers using the EM tracking and the reference-adjusted EM tracking methods, respectively.

The red squares are the EM tracking reprojection, which has no information regarding rotation change. The black squares are the reference-adjusted EM tracking reprojection, which will be explained next. The rotation angle was estimated by comparing the slopes of the corresponding line segments between the ArUco projection (green squares) and the reference-normalized EM projection (black squares). Similarly, the zoom factor was the ratio of the lengths of these corresponding line segments.

As shown in Fig. 4, we consider the first ArUco-success frame in the video sequence to be a reference frame. The zoom and rotation changes for the following frames are relative to this reference frame. For the reference frame, a reference AB correction  $T_{corr}^{AB_{ref}}$  can be calculated according to

$${}^C T_{AB}^{(ArUco)} = {}^C T_{AB}^{(EM)} \cdot T_{corr}^{AB_{ref}}. \quad (6)$$

Once calculated,  $T_{corr}^{AB_{ref}}$  is applied to the following ArUco-success frames to calculate the reference-adjusted EM tracking result  ${}^C T_{AB}^{(EM^*)}$  according to

$${}^C T_{AB}^{(EM^*)} = {}^C T_{AB}^{(EM)} \cdot T_{corr}^{AB_{ref}}. \quad (7)$$

The idea is to use  $T_{corr}^{AB_{ref}}$  to correct some tracking errors in the AB coordinate system before estimating  $\theta$  and  $\alpha$  from the video image. As shown in Fig. 6, the black squares were the reference-adjusted red squares before they were compared with the green squares to calculate  $\theta$  and  $\alpha$ . To summarize, for an ArUco-success frame other than the reference frame, we have the following algorithm to calculate the three correction matrices:

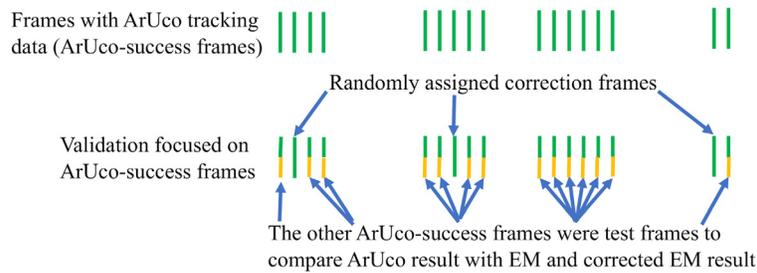
1. Calculate reference-adjusted EM tracking according to Eq. (7);
2. Use  ${}^C T_{AB}^{(EM^*)}$  to calculate  $T_{corr}^\theta$  and  $T_{corr}^{Zoom}$  from the camera image;
3. Calculate  $T_{corr}^{AB}$  using Eq. (3).

Once obtained for an ArUco-success frame, the three matrices will be used to correct EM tracking for the following ArUco-failure frames. The reason we used the first ArUco-success frame but not the most recent ArUco-success frame as the reference is that the errors in the estimated  $T_{corr}^{AB}$  of the previous ArUco-success frame would affect the estimation of  $T_{corr}^\theta$  and  $T_{corr}^{Zoom}$ , which, in turn, would affect the estimation of  $T_{corr}^{AB}$  in the current ArUco-success frame. This process will iterate and the errors could accumulate to become significant. On the other hand, the errors in estimating  $T_{corr}^{AB_{ref}}$  based on the first ArUco-success frame will be consistent in all following frames.

Although this study focuses on tracking the AB, it is straightforward to extend the hybrid tracking method to track the LUS image, i.e., by incorporating the ultrasound calibration result into the pipeline. Ultrasound calibration determines the transformation from the ultrasound image plane to the coordinate system of the sensor attached on the ultrasound probe. It can be performed using either the EM tracking approach<sup>7</sup> or the ArUco tracking approach.<sup>8,16</sup> Based on OpenCV and ArUco libraries, the hybrid tracking method was implemented using Python on a laptop computer with Intel Core i7 2.8 GHz quad-core processor and 32 GB RAM.

### 3 Result

We performed phantom and animal studies to evaluate the performance of our hybrid tracking method. Because hybrid tracking was designed to enhance the tracking performance in a complex, dynamic surgical environment, we therefore chose reprojection error, a metric that can be used to consistently evaluate framewise overlay accuracy for both phantom and animal studies. As used in most camera calibration works,<sup>11</sup> reprojection error is the average distance in the image space between the detected corners and the reprojected corners using the estimated pose. We can refer to Sec. 4 for more details on validity of using reprojection error to evaluate surgical AR systems and our potential future work to evaluate a more complete system using metrics in the 3D space.



**Fig. 7** Explanation of correction and test frames.

### 3.1 Phantom Study

Several video sequences were acquired using the same setup we used to acquire the developmental data. The video sequences included: a normal case; a distortion case in which an electronic device (a frame grabber) was repetitively brought in and taken out of the magnetic field; a zoom case in which the laparoscope's optical zoom was adjusted several times; a rotation case in which the 30-deg-telescope was rotated several times relative to the camera head; and a combination case that combined all the aforementioned situations. We placed a frame grabber close to the tip of the LUS transducer to generate significant distortion such that the overlay error caused by it was obvious. In practice, we do not anticipate such significant distortion during a normal laparoscopic procedure. It is worth noting that the zoom and rotation changes were made arbitrarily during the video acquisition, which meant we did not have the ground truth zoom factors and rotation angles. The acquired video sequences were post-processed by our developed software to generate EM and corrected EM tracking results. The detected corners and ArUco reprojected corners were obtained using the ArUco library. Although the video sequences were acquired at 10 frames per second (fps) that was limited by the frame grabber we used, our post-processing time was fast enough to keep up with the conventional 30-fps video frame rate. In other words, if the frame grabber could acquire images at 30 fps, our implementation is capable of real-time processing.

For each video sequence, ~80% of the frames met the ArUco-success criteria. We focused our validation on the ArUco-success frames. As shown in Fig. 7, the idea was to randomly assign a portion (called correction portion) of the ArUco-success frames as the correction frames, and the remaining ArUco-success frames as the test frames. The correction frames were used to calculate the correction matrices. For the test frames, the corrected EM and the original EM tracking results were compared with the ArUco tracking result. We experimented with three correction portions, which were 20%, 10%, and 5%. For each situation, the same video was processed 10 times with different random sets of correction frames.

Table 1 shows mean reprojection errors of the original EM, the corrected EM, and the ArUco tracking for different situations and different correction portions. Reprojection error was calculated using the corners of the markers detected by ArUco.

The corrected EM tracking results using both correction algorithms show significant improvements over the original EM tracking result, especially for the challenging situations (i.e., situations other than normal). The results of Algorithm 2 were better than those of Algorithm 1 in every situation. As anticipated, the greater the size of correction portion, the smaller was the reprojection error and the higher was the accuracy of hybrid tracking in all situations. It should also be noted that rotating the laparoscope led to larger errors compared with other challenging cases. We did not notice significant variation in results among the 10 runs with different random sets of correction frames. For example, the standard deviation of the 10 runs for the corrected EM tracking error in a normal situation (10% correction using Algorithm 2) is 1.3 pixel.

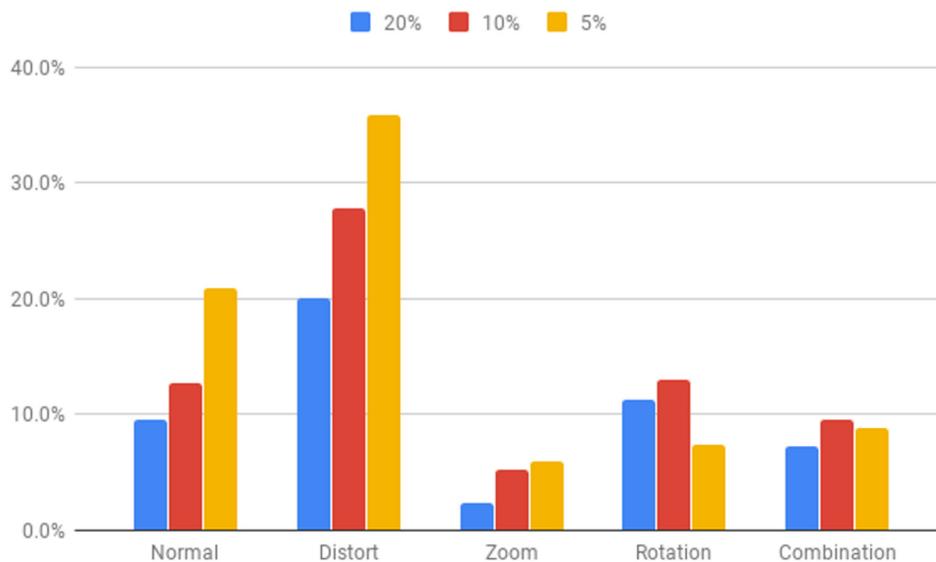
We believe the zoom and rotation cases of Table 1 warrant further explanation. A change in zoom will affect the parameters in the original calibrated camera matrix. For the EM tracking approach since hand-eye calibration does not change, the pose of the object (in our case the LUS transducer) in the camera space does not change either. The EM approach then projects the object with the original pose through an outdated camera matrix to the image space, which causes the

**Table 1** Mean reprojection error (in pixel) for different situations.

	EM	Corrected EM						ArUco
		20% correction, 80% test		10% correction, 90% test		5% correction, 95% test		
		Alg. 1	Alg. 2	Alg. 1	Alg. 2	Alg. 1	Alg. 2	
Normal	27.9	12.8	11.6	17.2	15.0	21.0	16.6	1.3
Distortion	60.3	12.4	10.0	19.1	13.8	27.5	17.7	1.4
Zoom	58.4	8.3	8.1	11.7	11.1	17.1	16.1	1.3
Rotation	366.7	24.0	21.3	39.3	34.3	58.4	54.1	1.4
Combination	181.0	10.9	10.1	17.0	15.4	24.9	22.7	1.9

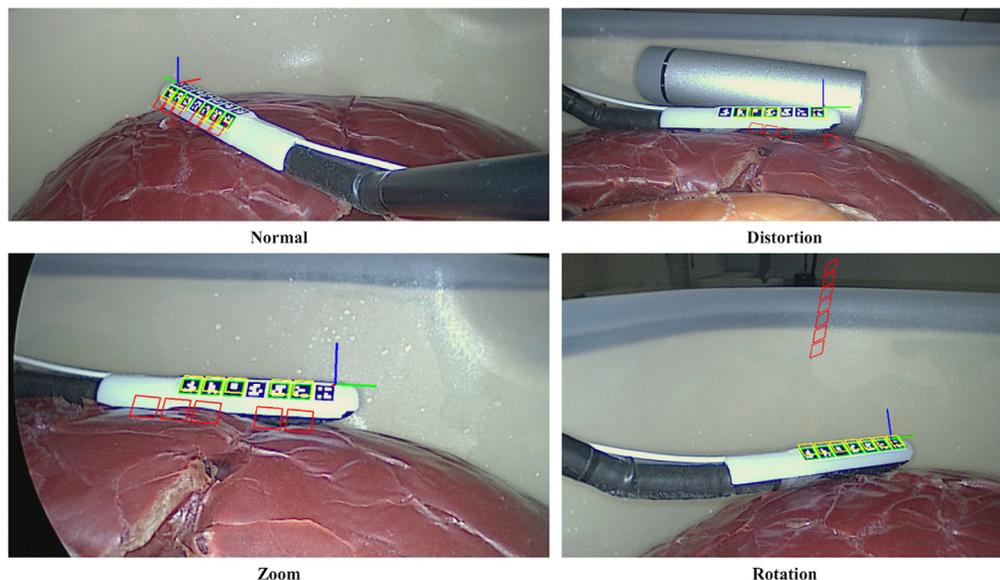
wrong overlay. On the contrary, after a zoom change, the ArUco method detects the object in a new location in the camera space despite the fact that neither camera nor the object has moved. For example, zooming in the camera is detected by the ArUco method as being closer to the camera, so it adapts the pose of the object in the camera space accordingly. Although the ArUco approach still uses the outdated camera matrix to project the object to the image space, based on the examples we have tried, errors caused by the outdated camera matrix have negligible impact on the ArUco approach in terms of overlay accuracy. As for the rotation case, it is worth noting that we used only a single sensor to track the 30-deg laparoscope. This seems unfair for the EM tracking approach because two sensors are needed to track the relative rotation without any assistance from an ArUco or other CV based technique. However, our purpose in this work was to show the proposed hybrid tracking method could work with a single sensor in which case the EM approach, as expected, would fail as also evident from the large reprojection error data in Table 1. Reducing the number of sensors from two to one carries significant benefits as it will greatly enhance the practicality of the resulting system.

Figure 8 shows the percentage of improvement when using Algorithm 2 over Algorithm 1 for varying correction portions. Algorithm 2 produced on average an  $\sim 21\%$  improvement over

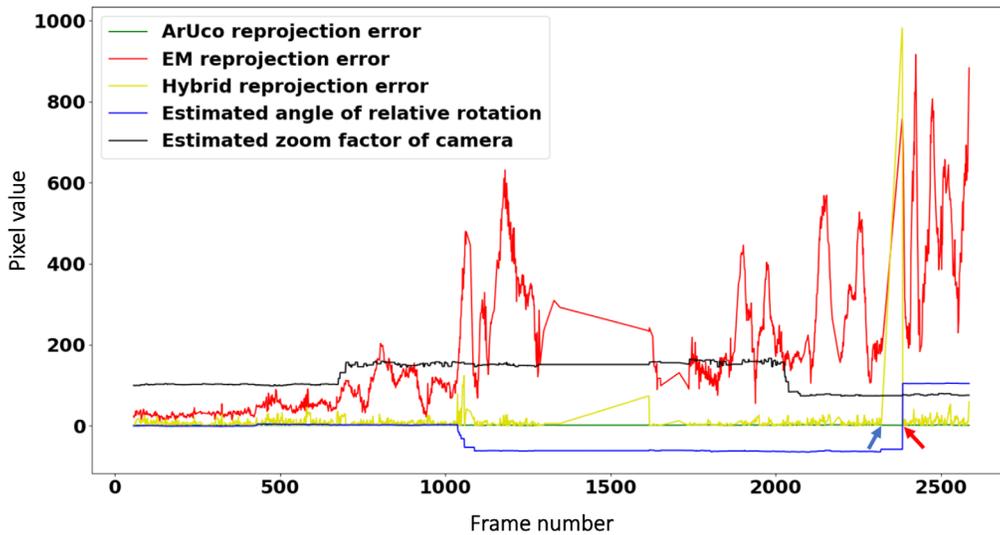
**Fig. 8** Percentage of improvement of Algorithm 2 over Algorithm 1 for different situations and for different correction portions.

Algorithm 1 for the normal and distortion cases. The improvement decreased to an average of  $\sim 8\%$  for the zoom, rotation, and combination cases. For the normal and distortion cases, there is a clear trend that the improvement increases as the correction portion decreases. However, this trend does not hold for the other cases involving zoom and rotation changes to the laparoscope. Although we modeled rotation and zoom changes in Algorithm 2 compared to Algorithm 1, one major difference between these two algorithms lies in the added  $T_{corr}^{AB}$  to correct errors in the AB coordinate system. This agrees with what we find in Fig. 8 in that the improvement of normal and distortion cases are greater than the other cases involving changes in the camera coordinate system. A video clip showing the tracking results, generated using Algorithm 2, during the combination case is provided as a multimedia material (Fig. 9).

Figure 10 shows plots of reprojection errors, the estimated relative rotation angle and the estimated camera zoom factor for one run of the combination case with 20% correction portion using Algorithm 2. The video started with the normal situation and the challenging events were then introduced over time and repeated. To generate the distortion of the magnetic field, an electronic device as shown in Video 1 was introduced and removed twice. As the reader may tell from Fig. 10, we also changed the camera zoom twice and rotated the telescope relative to the camera head twice. After changing the zoom, it may be necessary to adjust the camera focus. The original EM tracking result (red curve) became much worse after relative rotation was introduced. Note that the hybrid tracking result (yellow curve) includes both the ArUco tracking result and the corrected EM tracking result. As can be seen from the figure, the correction frame takes place when the yellow curve dips down to touch the green curve. In other words, the ArUco-failure frames take place when the yellow curve and the green curve do not overlap. In most frames, the hybrid tracking result remains close to the ArUco result and is significantly better than the original EM tracking result. One exception happens around frame number 2300, where the yellow curve has a spike. This is because a relative rotation takes place at this time (blue arrow), and we do not have a correction frame until a later time (red arrow). It should be noted that the algorithm calculates a new rotation angle only at a new correction frame, but not the frame where the actual rotation took place.



**Fig. 9** Example multimedia still images showing the results of the ArUco tracking (green), the corrected EM tracking using Algorithm 2 (yellow) and the original EM tracking (red). Only markers detected by ArUco were reprojected (Video 1, MP4, 11.1 MB [URL: <https://doi.org/10.1117/1.JMI.8.1.015001.1>]).



**Fig. 10** Reprojection errors, estimated relative rotation angle and estimated camera zoom factor for one run of the combination case with 20% of correction frames using Algorithm 2. The y axis is: pixel value for the reprojection error (green, red, and yellow curves); angle in degrees for the relative rotation angle (blue curve); and camera zoom factor times 100 (black curve).

### 3.2 Animal Study

An animal study on a 40-kg swine was performed to demonstrate the feasibility of using the hybrid tracking method. The study was approved by our Institutional Care and Animal Use Committee to ensure it was conducted in an acceptable ethical and humane fashion. In addition to the EM sensor on the hybrid tracking mount (Fig. 3), a second EM sensor was attached to the laparoscope (10 mm, 30 deg) in the same way as in the setup in Fig. 2. The EM tracking field generator, wrapped in a surgical cushion, was placed on the surgical table. The anesthetized swine was positioned supine on the field generator, with its liver at a desired location within the working volume of EM tracking. After insufflation, the laparoscope was introduced through a 12-mm trocar placed at the umbilicus (i.e., belly button). The LUS probe with the hybrid tracking mount was introduced through another 12-mm trocar placed at the left upper quadrant site. After the liver was examined with the LUS probe, the surgeon performed partial liver resection with the presence of the LUS probe in the surgical view.

We acquired two video recordings: one for the normal case and one for the challenging case, i.e., including rotation and zoom changes of the laparoscope. Table 2 has the ArUco tracking statistics for these animal study video recordings. The ArUco-success rates in the animal study were higher than our assumed correction portions (20%, 10%, and 5%) we studied earlier. The number of ArUco detected markers for the normal case was similar to what we obtained for the developmental data using the phantom. To be consistent with the phantom study, we assigned the same three portions of the ArUco-success frames to be correction frames, and the remaining ArUco-success frames to be test frames. Similar to the phantom study, the animal video was processed 10 times with different random sets of correction frames. Table 3 shows reprojection errors for the EM tracking, the corrected EM tracking using Algorithm 2, and the ArUco tracking. These errors were comparable to the errors obtained for the phantom study (Table 1).

**Table 2** Statistics for the two videos recorded during the animal study.

Case	Number of frames	ArUco-success rate	Number of markers detected (max)
Normal	3546	59.2%	$5.5 \pm 2.3$ (12)
Challenging	3599	31.8%	$3.9 \pm 1.7$ (11)

**Table 3** Mean and maximum (in parentheses) reprojection error (in pixel) for the animal data.

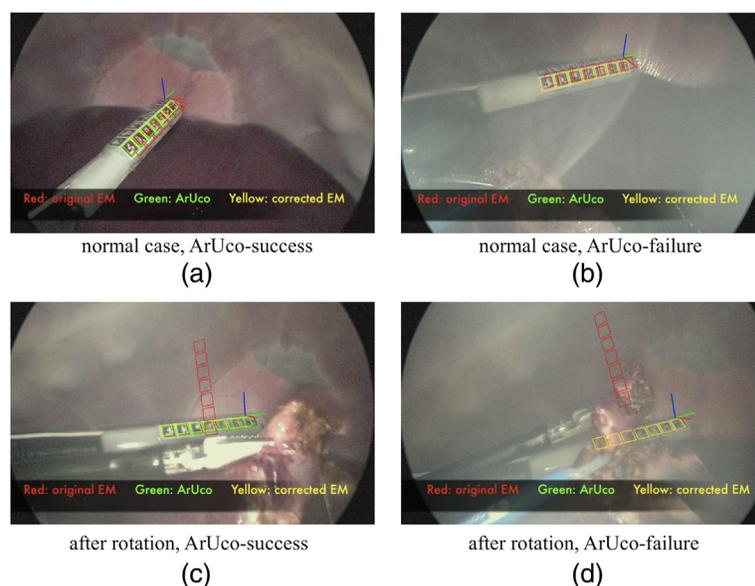
Case	EM	Corrected EM			ArUco
		20% correction, 80% test	10% correction, 90% test	5% correction, 95% test	
Normal	30.8 (123.8)	9.3 (151.2)	10.3 (151.2)	11.5 (151.2)	2.0 (2.9)
Challenging	149.0 (1079.4)	21.7 (911.9)	27.2 (935.6)	35.0(1059.9)	1.8 (2.9)

**Table 4** Mean and maximum (in parentheses) number of frames since last correction frame.

Case	Corrected EM		
	20% correction, 80% test	10% correction, 90% test	5% correction, 95% test
Normal	9 (788)	18 (818)	33 (836)
Challenging	15 (773)	30 (858)	62 (943)

Compared with EM tracking, the corrected EM tracking consistently yielded better results. Table 4 shows mean and maximum number of frames since the last correction frame was also given. As expected, large intervals without ArUco correction increased errors. It is worth noting that our evaluation method as explained in Fig. 7 generates larger-than-actual intervals without correction. This is because we only assigned a portion of the ArUco-success frames as the correction frames, and the other ArUco-success frames for testing contributed to the intervals without correction.

A video clip overlaid with tracking results is provided as a multimedia material (Fig. 11). To visually assess the results of hybrid tracking, we reprojected all markers (no matter they were detected or not) of the entire one face of the AB based on the estimated pose. This is for easier visually comparison with the original ArUco pattern in the blurred situation found in the animal study. As shown in Fig. 11(d), corrected EM tracking performed well even when the ArUco pattern was entirely occluded.

**Fig. 11** Example multimedia still images of the submitted video clip showing tracking results during the animal study (Video 2, MP4, 15.6 MB [URL: <https://doi.org/10.1117/1.JMI.8.1.015001.2>]).

## 4 Discussion

Our contribution in this work is a new hybrid tracking framework that combines hardware (i.e., EM)- and CV (i.e., ArUco)-based tracking to improve the overall tracking performance. We proposed two algorithms to calculate correction matrices applied to the original EM tracking. The proposed method was evaluated using an abdominal phantom first, followed by a feasibility study using a porcine model. We discuss below the results of the study and insights gained.

In both phantom and animal studies, ArUco tracking was very accurate ( $\leq 2$ -pixel reprojection error) if the ArUco library could successfully identify the pattern. We carried out simulations of hybrid tracking assuming 5%, 10%, and 20% ArUco-success frames in a given recording. The ArUco-success frame rates were higher than the upper threshold (20%) in the animal study, with the smallest rate being 31.8% for the challenging case. This rate was most likely smaller than typical because we deliberately tried to occlude the ArUco pattern in the animal study to challenge the hybrid tracking algorithm, as illustrated in [Video 2](#). Therefore, we believe the rates of ArUco-success frames we found in our animal study are representative of the rates that one would expect during an actual laparoscopic procedure, proving the feasibility of hybrid tracking.

For the video frames in which ArUco tracking fails, our proposed correction methods corrected the original EM tracking to improve the overall tracking performance. Of the two correction algorithms, Algorithm 2 outperformed Algorithm 1, and will be the preferred choice for future use. With corrected EM tracking, our hybrid tracking method not only increases tracking accuracy for a normal case, but also improves system practicality, i.e., allowing tracking of zoom and rotation changes of the laparoscope without adding an additional EM sensor as discussed in the Introduction section.

One potential limitation of the proposed method could be long dropouts in ArUco tracking and consequently large intervals of correction as reported in [Table 4](#). Such situations may result in large reprojection errors comparable to the original EM tracking ([Table 3](#)). In these situations, the proposed method loses its advantage over the original EM tracking method. The error can increase if a challenging event, such as zoom or rotation change of the laparoscope, occurs during such interval. This will cause a high peak of reprojection error as illustrated in [Fig. 10](#). When using the system clinically, the surgeon will be advised to expose the pattern to the camera if there is an extended period of occlusion of the pattern, or if a challenging event has happened.

Although metrics in the 3D space are ideal to determine the AR overlay accuracy, obtaining such measurements usually requires a static, well-planned experimental setup. For an initial demonstration that a surgical AR system can work in practical situations such as during animal or human procedures, reprojection error has been used in many previous works to compare among different methods. For example, Espinel et al.<sup>40</sup> used reprojection error to compare between the manual and the automated methods to register a preoperative 3D liver model to 2D laparoscopy image. They achieved 20 to 30 pixels of reprojection errors with  $1920 \times 1080$  image resolution (same as ours) for multiple *in-vivo* human data. To give the readers an approximate idea of what the pixel values mean in this paper, we approximately correlate the pixel value to the distance in the 3D space as follows. In selected video frames with typical distances from the camera to the target such as those shown in [Video 1](#), we manually quantified the pixel distances of edges of the AR markers. Since the actual distance of each edge is known to be 4.5 mm, the pixel distance can be correlated to the actual distance. This results in an average of 7.4 pixel per mm. It should be noted that this is an estimation because the pixel values depend on the distance from the camera to the target. In the future, we will integrate ultrasound calibration into the hybrid tracking pipeline, and evaluate the more complete system using metrics in the 3D space such as the TRE<sup>7</sup> and the vessel reconstruction error.<sup>17</sup>

The hybrid tracking software was implemented using Python on a computer with Linux operating system. It was independent from our AR software, which was implemented using C++ on a computer with Windows operating system. Because the data was acquired using the AR system, we evaluated the hybrid tracking method offline and retrospectively. When the hybrid tracking software is incorporated into the AR software, we anticipate hybrid tracking will work in real time. When using hybrid tracking during the procedure, the overlay accuracy can be visually evaluated by checking if the projected virtual ArUco pattern aligns with the physical pattern on the transducer. For clinical implementation, the EM sensor in the hybrid tracking mount ([Fig. 3](#))

can be embedded inside the LUS transducer,<sup>41</sup> and a sterilizable and biocompatible hybrid tracking mount need to be developed. Our hybrid tracking framework is flexible to employ new CV-based tracking technologies. For example, it is worth investigating using cylindrical patterns<sup>8,18</sup> instead of the planar ArUco pattern as most laparoscopic tools tend to have rounded surfaces. Ideally, the cylindrical pattern could be laser marked on the transducer surface. New developments in marker-less CV-based tracking could also be incorporated.

## 5 Conclusions

Combining EM tracking and ArUco tracking, we developed a hybrid tracking method to track the imaging tip of an LUS transducer in the laparoscope video image. Through phantom and animal studies, we showed that the new method is more reliable than using ArUco tracking alone and more accurate and practical than using EM tracking alone. The new hybrid method has the potential to significantly improve tracking performance for LUS-based AR applications. The hybrid tracking framework can be extended to track other surgical instruments.

## Disclosures

Raj Shekhar and William Plishker are cofounders of IGI Technologies, Inc.

## Acknowledgments

This work was supported by the National Institutes of Health under grant No. 2R42CA192504.

## References

1. M. Rosen and J. Ponsky, "Minimally invasive surgery," *Endoscopy* **33**(4), 358–366 (2001).
2. H. S. Himal, "Minimally invasive (laparoscopic) surgery," *Surg. Endosc.* **16**(12), 1647–1652 (2002).
3. J. Leven et al., "DaVinci canvas: a telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability," *Lect. Notes Comput. Sci.* **3749**, 811–818 (2005).
4. C. L. Cheung et al., "Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study," in *Lect. Notes Comput. Sci.*, **6363**, 408–415 (2010).
5. X. Kang et al., "Stereoscopic augmented reality for laparoscopic surgery," *Surg. Endosc.* **28**(7), 2227–2235 (2014).
6. P. Pratt et al., "Robust ultrasound probe tracking: initial clinical experience during robot-assisted partial nephrectomy," *Int. J. Comput. Assist. Radiol. Surg.* **10**(12), 1905–1913 (2015).
7. X. Liu et al., "Laparoscopic stereoscopic augmented reality: toward a clinically viable electromagnetic tracking solution," *J. Med. Imaging (Bellingham)* **3**(4), 045001 (2016).
8. U. L. Jayarathne et al., "Robust, intrinsic tracking of a laparoscopic ultrasound probe for ultrasound-augmented laparoscopy," *IEEE Trans. Med. Imaging* **38**(2), 460–469 (2018).
9. P. W. Hsu et al., "Comparison of freehand 3-D ultrasound calibration techniques using a stylus," *Ultrasound Med. Biol.* **34**(10), 1610–1621 (2008).
10. T. K. Chen et al., "A real-time freehand ultrasound calibration system with automatic accuracy feedback and control," *Ultrasound Med. Biol.* **35**(1), 79–93 (2009).
11. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000).
12. Y. Song et al., "Locally rigid, vessel-based registration for laparoscopic liver surgery," *Int. J. Comput. Assist. Radiol. Surg.* **10**(12), 1951–1961 (2015).
13. J. Ramalhinho et al., "Breathing motion compensated registration of laparoscopic liver ultrasound to CT," *Proc. SPIE* **10135**, 101352V (2017).

14. G. Xiao et al., "Electromagnetic tracking in image-guided laparoscopic surgery: comparison with optical tracking and feasibility study of a combined laparoscope and laparoscopic ultrasound system," *Med. Phys.* **45**(11), 5094–5104 (2018).
15. M. Feuerstein et al., "Magneto-optical tracking of flexible laparoscopic ultrasound: model-based online detection and correction of magnetic tracking errors," *IEEE Trans. Med. Imaging* **28**(6), 951–967 (2009).
16. P. Pratt et al., "Intraoperative ultrasound guidance for transanal endoscopic microsurgery," *Int. J. Comput. Assist. Radiol. Surg.* **15**(1), 463–470 (2012).
17. C. Schneider et al., "Tracked 'pick-up' ultrasound for robot-assisted minimally invasive surgery," *IEEE Trans. Biomed. Eng.* **63**(2), 260–268 (2016).
18. L. Zhang et al., "Real-time surgical tool tracking and pose estimation using a hybrid cylindrical marker," *Int. J. Comput. Assist. Radiol. Surg.* **12**(6), 921–930 (2017).
19. M. Feuerstein et al., "New approaches to online estimation of electromagnetic tracking errors for laparoscopic ultrasonography," *Comput. Assist. Radiol. Surg.* **13**(5), 311–323 (2008).
20. W. Plishker et al., "Hybrid tracking for improved registration of laparoscopic ultrasound and laparoscopic video for augmented reality," *Lect. Notes Comput. Sci.* **10550**, 170–179 (2017).
21. M. Allan et al., "Toward detection and localization of instruments in minimally invasive surgery," *IEEE Trans. Biomed. Eng.* **60**(4), 1050–1058 (2013).
22. Y. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form  $AX=XB$ ," *IEEE Trans. Rob. Autom.* **5**(1), 16–29 (1989).
23. S. De Buck et al., "Evaluation of a novel calibration technique for optically tracked oblique laparoscopes," *Lect. Notes Comput. Sci.* **4791**, 467–474 (2007).
24. C. Wu et al., "A full geometric and photometric calibration method for oblique-viewing endoscopes," *Comput. Aided Surg.* **15**(1–3), 19–31 (2010).
25. X. Liu et al., "Fast calibration of electromagnetically tracked oblique-viewing rigid endoscopes," *Int. J. Comput. Assist. Radiol. Surg.* **12**(10), 1685–1695 (2017).
26. T. Yamaguchi et al., "Development of a camera model and calibration procedure for oblique-viewing endoscopes," *Comput. Aided Surg.* **9**(5), 203–214 (2004).
27. A. M. Franz et al., "Electromagnetic tracking in medicine—a review of technology, validation, and applications," *IEEE Trans. Med. Imaging* **33**(8), 1702–1725 (2014).
28. S. Garrido-Jurado et al., "Generation of fiducial marker dictionaries using mixed integer linear programming," *Pattern Recogn.* **51**, 481–491 (2016).
29. F. J. Romero-Ramirez et al., "Speeded up detection of squared fiducial markers," *Image Vision Comput.* **76**, 38–47 (2018).
30. G. Bradski, "The OpenCV library," *Dr. Dobb's J. Software Tools* **25**(11), 120, 122–125 (2000).
31. H. Kato and M. Billingham, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," in *Proc. IWAR*, pp. 85–94 (1999).
32. M. Fiala, "ARTag, a fiducial marker system using digital techniques," in *Proc. CVPR*, pp. 590–596 (2005).
33. D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM J. Appl. Math.* **11**(2), 431–441 (1963).
34. M. Tella et al., "A combined EM and visual tracking probabilistic model for robust mosaicking: application to fetoscopy," in *Proc. CVPRW*, pp. 524–532 (2016).
35. T. Ungi et al., "Open-source platforms for navigated image-guided interventions," *Med. Image Anal.* **33**, 181–186 (2016).
36. A. Fedorov et al., "3D Slicer as an image computing platform for quantitative imaging network," *Magn. Reson. Imaging* **30**(9), 1323–1341 (2012).
37. B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Am. A* **4**, 629–642 (1987).
38. V. Lepetit et al., "EPnP: an accurate  $O(n)$  solution to the PnP problem," *Int. J. Comput Vision* **81**, 155 (2009).
39. X. Liu et al., "On-demand calibration and evaluation for electromagnetically tracked laparoscope in augmented reality visualization," *Int. J. Comput. Assist. Radiol. Surg.* **11**(6), 1163–1171 (2016).

40. Y. Espinel et al., "Combining visual cues with interactions for 3D–2D registration in liver laparoscopy," *Ann. Biomed. Eng.* **48**(6), 1712–1727 (2020).
41. X. Liu et al., "GPS laparoscopic ultrasound: embedding an electromagnetic sensor in a laparoscopic ultrasound transducer," *Ultrasound Med. Biol.* **45**(4), 989–997 (2019).

**Xinyang Liu** received his BS degree in electrical engineering from Beijing Institute of Technology in 2003 and his PhD in biomathematics from Florida State University in 2010. He was previously with Johns Hopkins Hospital and Brigham and Women's Hospital. He is currently a staff scientist in the Sheikh Zayed Institute for Pediatric Surgical Innovation at the Children's National Hospital. His research interests include medical imaging, computer-assisted surgery, AR, and machine learning.

**William Plishker** received the BS degree in computer engineering from Georgia Tech, Atlanta, Georgia and a PhD in electrical engineering from the University of California, Berkeley. His PhD research centered on application acceleration on network processors. His postdoctoral work at the University of Maryland included new dataflow models and application acceleration on graphics processing units. He is currently the CEO of IGI Technologies, Inc., which focuses on medical image processing.

**Raj Shekhar** is a principal investigator in the Sheikh Zayed Institute for Pediatric Surgical Innovation at the Children's National Hospital and a professor of radiology and pediatrics in the George Washington School of Medicine and Health Sciences. He leads research and development in surgical visualization, AR, signal and image processing, machine learning, and mobile applications. He was previously with the Cleveland Clinic and the University of Maryland and has founded two medical technology startups to commercialize his academic research.