



OPEN Mixed reality navigation system for ultrasound-guided and fusion biopsies using markerless instrument tracking

Michał Trojak^{1,2}, Maciej Stanuch² & Andrzej Skalski^{1,2}✉

Percutaneous needle biopsy is commonly performed under ultrasound guidance to visualize the target lesion and monitor needle position. This study aims to improve spatial information for the operator by accurately visualizing the ultrasound image at its true anatomical location and scale within a head-mounted display, co-registered with pre-procedural imaging data. Current mixed reality systems rely on fiducial markers to track the instruments, which limits their practical utility in clinical settings; therefore, this study evaluated a custom markerless method for ultrasound probe tracking that achieved accuracy suitable for clinical applications. By displaying the ultrasound image relative to the probe's spatial position, the system enables seamless fusion with pre-procedural imaging data. Combined with markerless needle tracking, this approach provides a complete navigation solution. Target registration error analysis between 2D ultrasound and 3D computed tomography demonstrated a 95% likelihood that the error remains within 3.41 mm. A biopsy navigation assessment test confirmed the system's strong potential to reduce the number of puncture attempts compared to conventional ultrasound guidance alone. By providing comprehensive spatial and imaging information, the navigation system may contribute to reduced operator error and improved procedural success rates.

Keywords Augmented Reality, Biopsy, Computer Vision, Locatable Ultrasound, Mixed Reality, Surgical Navigation

Biopsy is a medical procedure that involves removal of a tissue sample from a living individual for diagnostic examination¹. One of its most commonly performed type is percutaneous needle biopsy (PNB), which is a minimally invasive method that relies on inserting a needle through the skin². It is conducted with the assistance of imaging modalities like ultrasound (US), fluoroscopy, computed tomography (CT), magnetic resonance imaging (MRI), cone beam CT (CBCT) and positron emission tomography CT (PET-CT).

US is often preferred over other modalities because it lacks ionizing radiation and is portable, accessible, and fast, while providing real-time needle visualization and precise guidance in almost any anatomical plane^{3,4}. It can also adjust to tissue changes such as pressure, respiration, injection, and bladder expansion. US-guided PNB is widely used for diagnosing thoracic lesions⁵, peripheral pulmonary lesions⁶ or musculoskeletal soft tissue masses⁷. Reported diagnostic success rates in past studies include 90.5% for thoracic lesion (lesion sizes 1.5–16 cm, mean 7 cm)⁵, 81.8% for peripheral pulmonary lesions (72.0% ≤ 20 mm, 86.8% for 21–49 mm, 79.7% ≥ 50 mm; lower accuracy due to necrosis in 28.8% of cases)⁶, and 81.4% for peripheral pulmonary nodules ≤ 2 cm⁸.

Recent work has explored head-mounted display (HMD) applications for US-guided procedures. Nguyen et al.⁹ have streamed and rendered US images in an HMD with respect to the tracked probe by detecting specific markers printed on a custom probe case. The study revealed a 17% faster completion time for novices using the mixed reality system; however, for experts, the system resulted in a 5% slower performance, likely due to their familiarity with traditional methods. Costa et al.¹⁰ used QR codes to track both US probe and needle within an HMD, displaying live US video stream and trajectory from needle tip to the lesion. The study revealed longer procedure times but improved needle alignment and lesion targeting compared to a conventional approach. Li et al.¹¹ tracked retro-reflective spheres both on the US probe and needle with depth camera. An examination carried out on a phantom containing a tumor with a 5 mm radius showed that mixed reality guidance increased

¹AGH University of Krakow, Krakow PL30059, Poland. ²MedApp S.A., Krakow PL30037, Poland. ✉email: skalski@agh.edu.pl

the procedure success rate from 26% to 98% for out-of-plane punctures and from 45% to 93% for in-plane punctures.

Fusion biopsy, which combines real-time images provided by the US and high-resolution features of MRI, is now widely used for prostate cancer diagnosis^{12–14}. Studies report that fusion biopsy achieves higher overall diagnostic accuracy and sensitivity than conventional US-guided techniques, including detecting high-grade tumors (Gleason score ≥ 7)^{15,16}. The fusion of MRI and US is typically performed using three distinct methods. One approach is cognitive fusion, which relies on MRI to identify suspicious region and then target it using US without any specialized software^{17,18}. This method depends heavily on operator experience and does not outperform standard biopsy techniques¹⁹. Another techniques are rigid registration that is based on identifying corresponding anatomical landmarks and computing translation and rotation to bring one image set into alignment with the other^{20,21} and elastic registration that additionally tries to compensate for possible deformations, such as those caused by patient movements, pressure from surgical instruments, or injections²². Research conducted by Hale et al.²³ indicated that the rigid registration achieves a registration error of 4.87 ± 3.50 mm but elastic registration does not considerably improve the results, achieving error of 4.11 ± 2.09 mm.

In recent years, deep learning methods for estimating the six-degree-of-freedom pose of surgical instruments have gained considerable attention. Hein et al.²⁴ compared state-of-the-art single-view methods ZebraPose²⁵ and SurfEmb²⁶ with multi-view method EpiSurfEmb²⁷ for markerless surgical drill and screwdriver pose estimation. They report that monocular pose estimation do not satisfy the high accuracy requirements of clinical applications due to depth ambiguities. Moreover, the depth refinement approach introduced for SurfEmb does not reliably enhance pose accuracy and can, in fact, substantially degrade it. Similarly, the Iterative Closest Point refinement proposed for ZebraPose substantially reduces the overall pose estimation accuracy. The multi-view approach reached satisfactory results, resulting in an average positional error of 1.01 mm and an orientation error of 0.89° for the surgical drill, and 2.79 mm and 3.33° for the screwdriver under optimal conditions. Agethen et al.²⁸ proposed a deep learning method for markerless surgical tool tracking that combines pose estimation from multiple views with recurrent neural network. This approach has achieved an error below 3.0 mm on the dataset acquired in a simulated operating room environment, using a four-cameras configuration.

Our study relies solely on built-in RGB and depth cameras within an HMD for markerless tool tracking. It advances US applications in biopsy by building upon mixed reality navigation system introduced in²⁹. It contained markerless needle tracking and imaging data superimposition onto the patient's physical body. The present work extends it with markerless US probe tracking using single-view pose estimation and custom refinement method with depth data, enabling real-time overlay of US images in accurate scale and anatomical location being examined. It enables possibility of performing fusion of live US images with pre-procedural 3D CT/MRI data without any complex intensity-based registration algorithms, extending fusion biopsy beyond prostate cancer. The system also supports marking of target point on anchored US image and injection site using gestures and voice commands and guides the operator to precisely perform the procedure. This approach reduces human error, potentially shortens procedure time, and improve detection accuracy. The fully markerless setup provides a better user experience and greater flexibility by eliminating the need to maintain focus on markers to ensure they remain in the camera's field of view. In conclusion, the key contributions of this study are as follows:

- Extension of a mixed reality navigation system with markerless US probe tracking using single-view pose estimation and a custom depth-based refinement approach that improves the resulting accuracy;
- Simultaneous markerless tracking of biopsy needle and US probe within unified coordinate system;
- Fusion of real-time 2D freehand US images with pre-procedural 3D CT/MRI data by slicing the 3D volumes with a plane defined by the US probe pose;
- Assessment of the mixed reality-guided biopsy navigation with US imaging on a radiological phantom and an ex vivo tissue model.

Methods

The mixed reality biopsy navigation system consists of three major elements, presented in Fig. 1. It involves reconstructing CT/MRI data into a volumetric model visualized on an HMD, along with needle and US probe tracking to overlay US images and assist the user during needle insertion.

Mixed reality visualization and patient registration

Visualization of imaging data in mixed reality is done using CarnaLife Holo technology (MedApp S.A., Poland)³⁰. It processes DICOM data into a 3D volumetric model that can be visualized through an HMD. Users can interact with it including moving, rotating, scaling, slicing, or taking measurements³¹. The system can also superimpose the visualized data onto the patient's body, maintaining accurate real-world dimensions^{32,33}. This requires scanning the subject with radiological markers that serve as reference points and remain clearly visible in the imaging data across multiple slices. By matching corresponding points in the imaging data and physical sites within the world coordinate system of the headset as presented in Fig. 2, the transformation between these two point sets is computed. In a result, the imaging data is adjusted to real-world dimensions and perceived in an HMD as a volumetric model aligned with the subject²⁹.

Markerless surgical instruments tracking

The system is embedded with simultaneous tracking of biopsy needle and US probe without the need for physical markers, as presented in Fig. 3. Needle tracking enables guiding the operator to precisely inject the needle at the selected location, ensuring that it reaches the target area. Moreover, US probe tracking allows for real-time US images to be displayed relatively to the probe location, with accurate scale and orientation.

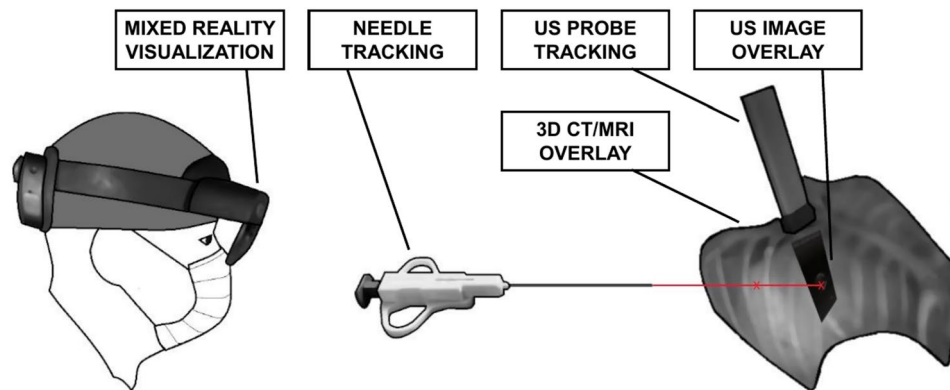


Figure 1. Scheme of the mixed reality biopsy navigation system with US imaging. Pre-procedural CT/MRI data is reconstructed into a volumetric model that is displayed on an HMD and aligned with patient's physical body. The US probe tracking enables the display of the US image at its real size and position, correlated with the probe's location. The solution also enables the selection of a needle trajectory and guides the operator along this path during needle placement.

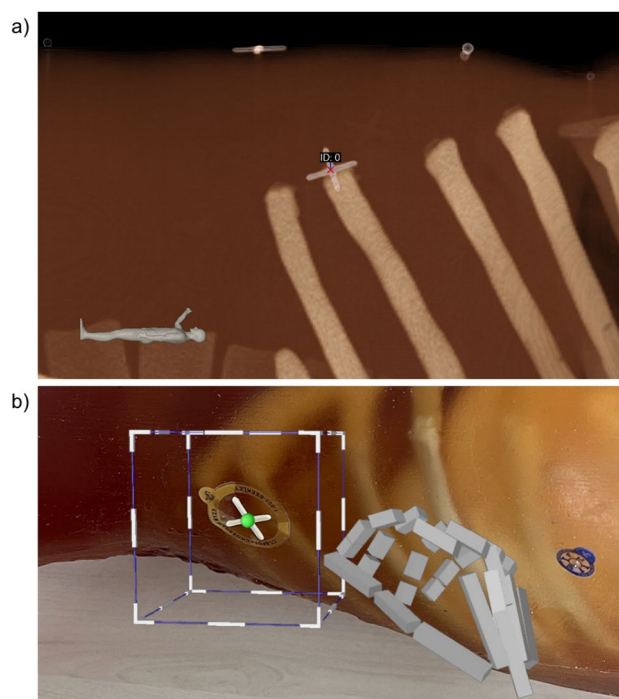


Figure 2. Selecting corresponding reference points in the imaging data and physical sites in world coordinate system of the headset to perform patient registration: (a) selected reference point in the imaging data; (b) mixed reality interface to select corresponding reference point within an HMD on a physical site.

The system utilizes the Microsoft HoloLens 2 headset that is equipped with a set of built-in cameras and sensors accessible by Research Mode³⁴. The process of needle tracking begins with segmenting the tool's handle on RGB image with FastInst³⁵ and identifying the proper needle line through the Hough transform. Subsequently, the tool's position within a mixed reality space is reconstructed by aligning points captured by depth camera with the predefined needle line²⁹.

US probe tracking begins by detecting the tool in RGB image with YoloX³⁶ and estimating its pose using ZebraPose²⁵. The Perspective-n-Point (PnP) problem is solved using the Progressive-X (Prog-X) algorithm³⁷, which iteratively generates pose hypotheses, quickly rejects redundant ones and integrates the new ones into a compound set using energy minimization, enforcing spatial coherence by penalizing inconsistent or non-adjacent 3D–2D projections. This method has shown high efficiency on the LM-O dataset³⁸ that contains multiple occluded and texture-less objects²⁵. The predicted pose is then refined with depth data by establishing correspondences between model point cloud and points captured by depth camera³⁹. These correspondences

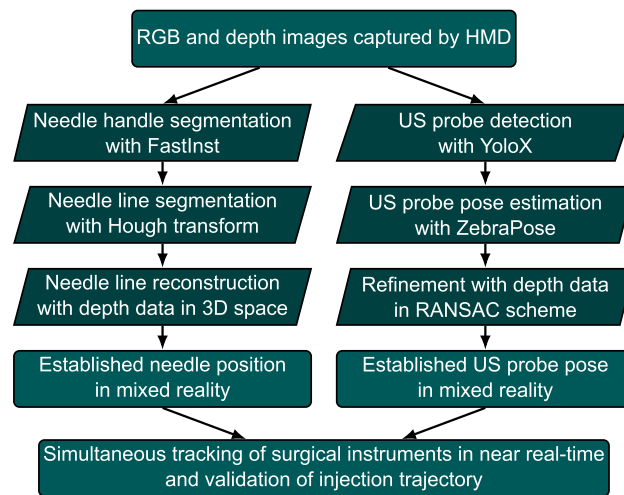


Figure 3. Scheme of markerless surgical instruments tracking pipeline. The inputs are RGB and depth images captured by the built-in HMD cameras and in result algorithm outputs estimated needle position and pose of the US probe. It enables possibility to validate correctness of injection in predefined trajectory.

are determined by projecting the model point cloud onto the RGB image plane, which is aligned with the depth camera image. Using the relationship between model points transformed by the predicted pose and their distance from camera we exclude those points that are not visible from the camera's view. Additionally, points not related to the object are excluded using segmented mask predicted by the ZebraPose which handles occlusions by other elements such as a hand. Finally, a rigid transformation between the two point sets is computed via least-squares fitting⁴⁰ in a RANSAC scheme, and pose stability is improved using Kalman filtering and Exponential Moving Average (EMA). The study utilized US Chison SonoAir 70 (Chison, China) system with linear probe.

Since the instrument tracking relies solely on the built-in sensors, no extra calibration step is required. The predicted pose is obtained within internal Photo-Video camera coordinate system and can be directly transformed to the world coordinate system of the headset. That is a major simplification over solutions that consists of external cameras or tracking systems which require calibration before each procedure.

Multi-modal image fusion

Unifying the coordinate systems of tracked surgical instruments, pre-procedural imaging data, and an HMD enables the fusion of 2D US images with 3D CT/MRI data. To achieve this, the reconstructed volumetric model from the pre-procedural scan must first be accurately aligned with the patient's physical body. Once the US probe is tracked, real-time US images can be also overlaid on the patient's body in their true spatial position and scale, allowing seamless integration with the pre-procedural imaging data. The estimated position and orientation of the US image determines the cutting plane through the superimposed volumetric data, creating a cross-section of the 3D CT/MRI. Given that obtained US probe pose and superimposed 3D CT/MRI volumetric model are in world coordinates of the headset, the transformation of the US probe pose to volumetric model coordinate system is calculated as follows:

$$(T_V^W)^{-1} T_P^W = \begin{bmatrix} R_P^V & t_P^V \\ 0 & 1 \end{bmatrix} \quad (1)$$

where T_V^W and T_P^W denote the transformations of the volumetric model and the US probe into world coordinates, respectively, while R_P^V and t_P^V represent the rotation and translation of the US probe relative to the volumetric model. The cutting plane through volumetric model can be then calculated as:

$$\Pi = \{ x^V \in \mathbb{R}^3 \mid n_V^\top (x^V - p_V) = 0 \}; \quad n_V = R_P^V n_P; \quad p_V = t_P^V + R_P^V p_P \quad (2)$$

where x^V denotes a point in the volumetric model, n_P and n_V are the normals to the US probe plane in world and volumetric model coordinates, respectively, and p_P and p_V are the positions of the US probe in world and volumetric model coordinates. Within this method it is possible to overlay 2D US onto 3D CT/MRI without any advanced intensity-based registration algorithms. As a result, the operator gains a comprehensive information about the patient's anatomy, enabling more precise and informed decision-making during the procedure.

Results

Ablation study

The first step in markerless US probe tracking pipeline is to determine its pose from RGB image. The selection of a model for this task was guided by the need to balance high accuracy with minimal inference time given real-time performance requirements. The study compared ZebraPose against PVNet⁴¹ in terms of US probe

pose estimation from monocular RGB images, considering also the difference in accuracy between RANSAC and Prog-X in solving the PnP problem. Moreover, the study examined the influence of incorporating depth data by evaluating output poses after the refinement step and those predicted by the HiPose model⁴². The refinement was performed without utilizing the Kalman filter and EMA, since they necessitate input data from consecutive time steps. The assessment was conducted using Maximum Symmetry-Aware Surface Distance (MSSD) and Maximum Symmetry-Aware Projection Distance (MSPD) metrics⁴³ that accounts for both the translation and rotation of the object. The MSSD measures the maximal surface distance between the estimated and ground-truth object poses, while the MSPD measures the maximal pixel displacement between their projections:

$$e_{\text{MSSD}}(\hat{P}, \bar{P}, \mathcal{S}_M, \mathcal{V}_M) = \min_{S \in \mathcal{S}_M} \max_{x \in \mathcal{V}_M} \|\hat{P}x - \bar{P}Sx\|_2 \quad (3)$$

$$e_{\text{MSPD}}(\hat{P}, \bar{P}, \mathcal{S}_M, \mathcal{V}_M) = \min_{S \in \mathcal{S}_M} \max_{x \in \mathcal{V}_M} \|\text{proj}(\hat{P}x) - \text{proj}(\bar{P}Sx)\|_2 \quad (4)$$

where M is the object's 3D model, $\hat{P}$ is the estimated pose, $\bar{P}$ is the ground-truth pose, $\mathcal{S}_M$ is a set of object symmetries and $\mathcal{V}_M$ is a set of points sampled from the object's 3D model. Moreover, errors in estimating the object's origin position and projection were also studied. Both models were trained on almost 16000 images and tested on 2200 images gathered using Intel Realsense D455 and D435i stereo cameras (Intel, USA) with annotated US probe poses. The test set included mainly hand-held and occlusion scenarios. The image annotation process involved reconstructing the point cloud of a given scene and manually aligning the 3D model of the object with its corresponding location. This approach made it possible to determine the object's translation and rotation relative to the camera coordinate system for each image. The recalls for accurate estimates at various thresholds are shown in Fig. 4. The superior outcomes of ZebraPose over PVNet demonstrate that the model excels at handling low-textured objects with limited number of distinctive features. The use of Prog-X instead of the standard RANSAC/PnP approach also increases the accuracy level. The custom refinement method with depth data significantly improves the results: 70.86% of cases fall below the 1 cm MSSD threshold and 80.09% below the 10 px MSPD threshold, compared to only 8.41% and 50.27%, respectively, when using ZebraPose with Prog-X alone. An exemplary predicted poses before and after refinement are presented in Fig. 5. Overall, HiPose generally achieves pose estimations with less error compared to ZebraPose alone, with 50.50% of cases below a 1 cm MSSD threshold and 56.14% below a 10 px MSPD threshold, though it does not reach the precision achieved by refining ZebraPose's results with the developed method. However, there are certain instances where HiPose performs worse than ZebraPose, with only 91.00% of cases falling below the 10 cm MSSD threshold, as opposed to 98.05% for ZebraPose. The achieved results demonstrate that appropriate utilization of depth data is crucial for accurate pose estimation, especially in determining the correct distance of the object from the camera.

The impact of each step in the custom refinement method using depth data was assessed by estimating object orientation with different versions of the algorithm. Each subsequent version activated the next element, which was then compared to the reference orientation determined by the surrounding ArUco markers on a prepared 3D-printed plate, presented in Fig. 6. To improve the reference orientation accuracy, the center of each of the

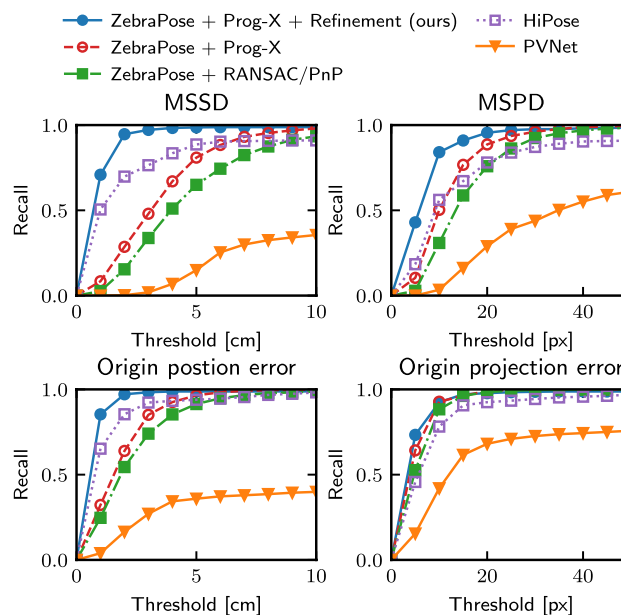


Figure 4. Recalls of US probe pose estimations in camera coordinates at various thresholds for evaluated models. Origin point is the geometric center of the object. ZebraPose with Prog-X solver excels with monocular input. Incorporating depth data significantly enhances accuracy, where refining ZebraPose output using our developed technique yields superior results compared to HiPose.

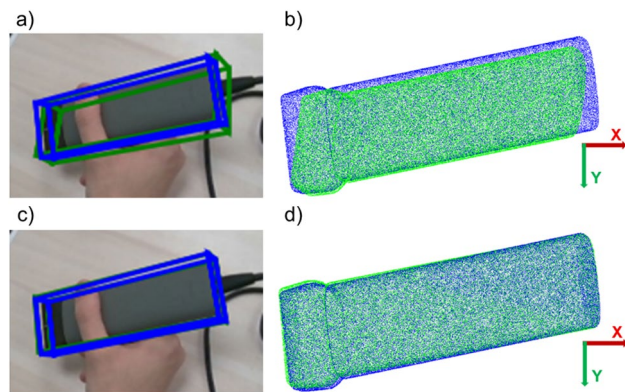


Figure 5. ZebraPose predictions using Prog-X solver before and after custom refinement method with depth data where blue boxes/points represents ground-truth poses, green represents predictions: (a) MSPD 9.79 px; (b) MSSD 5.14 cm; (c) MSPD 2.25 px; (d) MSSD 0.31 cm.



Figure 6. 3D-printed plate with ArUco markers to determine reference object orientation for ablation study.

Post-processing	Non-occluded	Hand-occluded
Baseline	$10.79 \pm 5.42^\circ$	$12.94 \pm 6.47^\circ$
+ RANSAC	$4.87 \pm 3.03^\circ$	$9.81^\circ \pm 3.89^\circ$
+ Mask	$4.20 \pm 2.77^\circ$	$4.86 \pm 3.03^\circ$
+ Kalman filter	$3.51 \pm 2.18^\circ$	$3.85 \pm 2.16^\circ$
+ EMA	$2.94 \pm 1.60^\circ$	$3.21 \pm 1.69^\circ$

Table 1. Impact of post-processing elements to refine the output pose in world coordinates. The error is calculated as a mean of angles between corresponding direction vectors of the estimated and ground-truth orientations. The baseline version of the algorithm performs least squares fitting across all model and depth points. The RANSAC-based approach selects the transformation achieving the highest number of inliers. The predicted object mask is then used to perform early outlier rejection. Finally, the pose is stabilized using a combination of a Kalman filter and EMA.

12 ArUco markers was averaged over 100 successive detections. Two direction axes were obtained by averaging vectors between marker centers, and the third axis was defined as their cross product. This experiment did not involve external tracker to produce ground-truth as unifying its coordinate system with the headset would provide additional bias⁴⁴. During the test, the probe remained stationary while the headset moved around it, observing the object from various perspectives. The results of 1000 consecutive iterations of the algorithms are presented in Table 1. In both non-occluded and hand-occluded scenarios, the error after all steps is over one-third of a baseline. Without occlusion, RANSAC doubles the accuracy, though the mask does not substantially affect the results. Nevertheless, for occlusion scenarios, RANSAC substantially enhances its influence on the results after employing the mask to remove points not related to the tracked object.

Comparison to marker-based tracker

Pose obtained with ArUco markers used in ablation study has limited precision as a mean registration accuracy of the marker center is above 1 mm even in the best case scenario⁴⁵. To effectively validate positioning accuracy of the developed solution it was compared to a different navigation system. The study utilized ClaroNav MicronTracker H3-60 (ClaroNav, Canada) which is designed to detect and track objects by identifying passive, visible-light markers of unique black-and-white checkered patterns called Xpoints⁴⁶. It has root mean square calibration accuracy of 0.20 mm, measured for single Xpoint over 20000 averaged positions at depths in range 40–100 cm.

To avoid discrepancies caused by coordinate system alignment between both systems, the study examined relative motion, specifically observing the change in the estimated pose after moving or rotating the object⁴⁷. The HMD and MicronTracker were positioned statically in fixed locations, each observing the object from the opposing directions. For translation error measurements, the US probe was translated by 5 cm up to 30 cm, the HMD was placed 60 cm from the object center at the midpoint, angled at 50° towards the object while MicronTracker was placed 65 cm from the marker at the midpoint and angled at 15° towards it. In case of rotation error measurements, the US probe was rotated by 10° up to 90°, the HMD was placed 45 cm from the object, angled at 50° towards it while MicronTracker was placed 65 cm from the rotation center and angled at 35° towards it. Schematic illustrations of experimental setup are depicted on Fig. 7. For both tests the initial object pose estimated by the proposed solution was averaged over 20 consecutive iterations and for each step 20 consecutive estimations were compared with the readings from MicronTracker that served as the ground-truth. In case of rotation test, the delta was calculated along two axes perpendicular to the axis of rotation and taken the average of them. The translation error for each step was calculated as follows:

$$e_{\text{translation, step}} = \frac{1}{n} \sum_{i=1}^n \left| \|p_{\text{ref}}^H - p_{\text{step},i}^H\| - \|p_{\text{ref}}^T - p_{\text{step}}^T\| \right| \quad (5)$$

where n is number of measurements, p_{ref}^H and p_{ref}^T are the initial object positions estimated from the HMD and measured by the external tracker, respectively, $p_{\text{step},i}^H$ is the displaced position of the i -th measurement estimated from the HMD, and p_{step}^T is the displaced position measured by the external tracker. The rotation error for each step was calculated as follows:

$$e_{\text{rotation, step}} = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} \sum_{j \in \{a,b\}} \left| \angle(v_{j,\text{ref}}^H, v_{j,\text{step},i}^H) - \angle(v_{j,\text{ref}}^T, v_{j,\text{step}}^T) \right| \quad (6)$$

where n is number of measurements, $j \in \{a,b\}$ indexes two axes perpendicular to the axis of rotation, $v_{j,\text{ref}}^H$ and $v_{j,\text{ref}}^T$ are the initial orientation vectors estimated from the HMD and measured by the external tracker, respectively, $v_{j,\text{step},i}^H$ is the displaced orientation vector of the i -th measurement estimated from the HMD, and $v_{j,\text{step}}^T$ is the displaced orientation vector measured by the external tracker.

The results of the tests are presented in Fig. 8. They demonstrate a reliable positioning accuracy that aligns closely with the readings from the ground-truth.

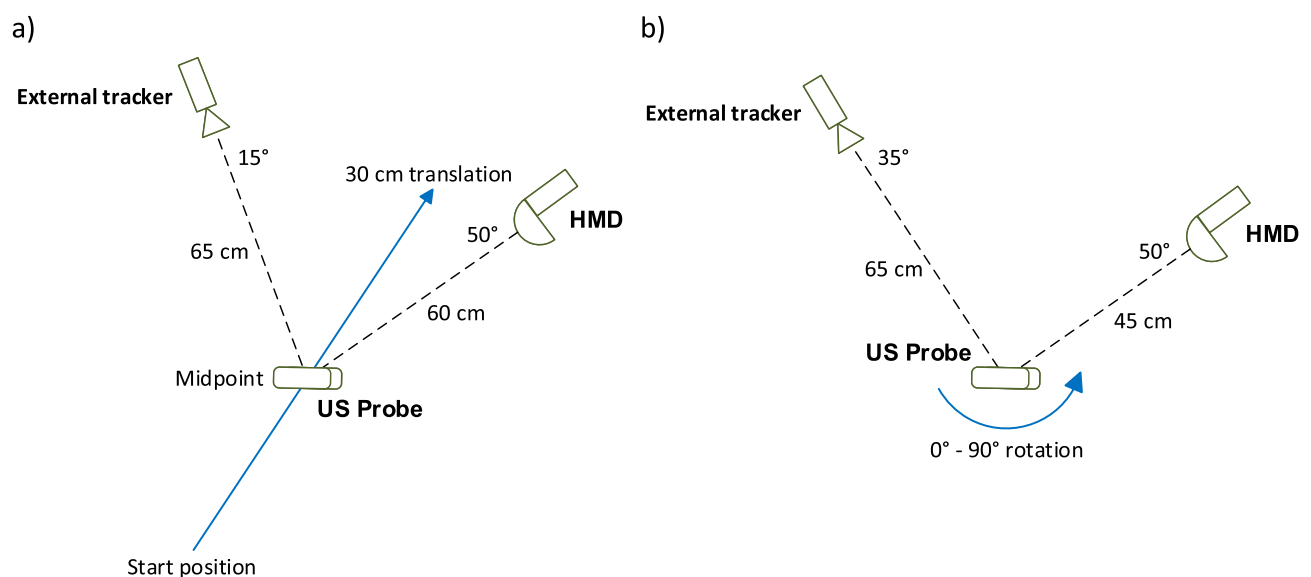


Figure 7. Setup for relative motion test with marker-based tracker: (a) translation error measurement; (b) rotation error measurement. Blue arrows represent object movement. The external camera and HMD were positioned to keep the object within the field of view throughout the entire range of movement.

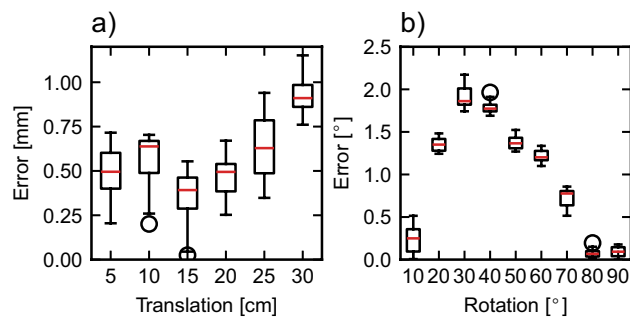


Figure 8. Results of the relative motion test: (a) translation error; (b) rotation error. Estimations of the markerless solution closely aligns with readings from the marker-based tracker.

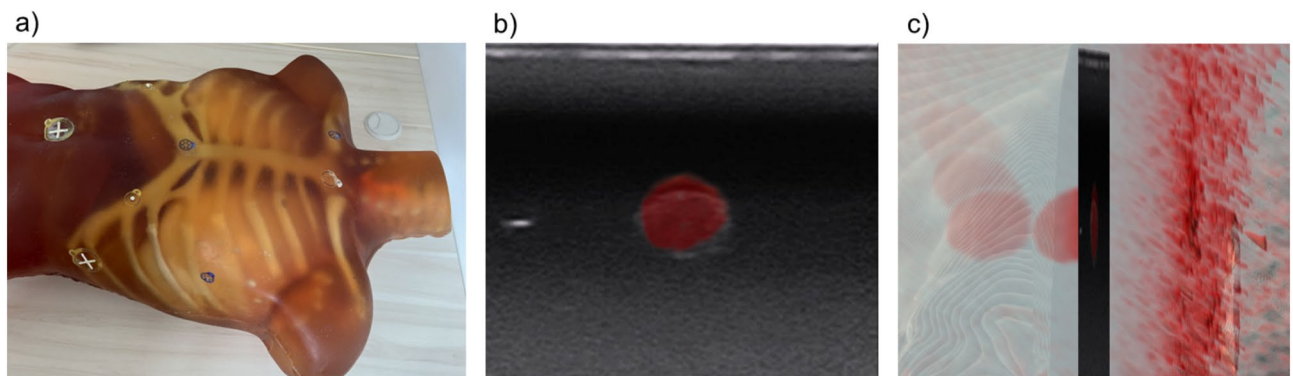


Figure 9. Rigid 2D US - 3D CT fusion (a) Rigid radiological phantom used for the evaluation (b) US image with red overlay representing lesion region in CT data of mean TRE 1.28 ± 0.39 mm; (c) the same fused data shown from a different viewing angle, with the 2D US image positioned within the 3D CT volume.

Runtime analysis

Sensor data is transmitted from the HMD to an external computer, where all processing and rendering are performed, and the resulting frames are streamed back to the HMD for display. The runtime test was conducted on a machine equipped with an AMD Ryzen 9 5900HX CPU and an Nvidia RTX3080 GPU. The algorithm was implemented in C++ with Prog-X optimized to enable real-time performance and models accelerated with TensorRT. The measured algorithm processing runtimes are averaged over 1000 iterations. For US probe pose estimation, detection with YoloX takes 18.73 ± 0.67 ms, building 2D-3D correspondences by ZebraPose takes 11.46 ± 0.78 ms, Prog-X runs at 28.04 ± 6.25 ms, and refinement of the predicted pose with depth data takes 12.61 ± 0.85 ms, resulting in total processing of 72.61 ± 6.46 ms. For needle position estimation, segmenting the tool handle with FastInst and identifying the needle line through the Hough transform run mostly in parallel, together taking 19.07 ± 2.12 ms, while reconstructing the needle position in 3D space from depth data takes an additional 5.96 ± 0.87 ms, which results in a total processing time of 25.03 ± 2.41 ms. Simultaneous US probe pose and needle position estimation takes around 98 ms on average if both tools are within the camera's field of view; however, the transmission of the sensor data from HMD to the computer requires an extra 19.47 ± 2.44 ms. The new frame is rendered, transmitted and displayed on the HMD within around 113 ms. Although average processing refresh rate for tracking US probe alone is 10.86Hz, for tracking needle position alone is 22.47Hz, and for simultaneous tracking of both instruments is 8.54Hz, the overall end-to-end latency remains within 197–230 ms, as tool pose estimation runs asynchronously with the rendering loop.

CT-US target fusion error

Evaluation of the target registration error (TRE) between CT and US was carried out using a radiological phantom with embedded lesion imitations that was subjected to CT scanning. Twenty samples of US images with estimated positions and orientations within 3D CT volumetric model were collected from various locations. Each sample underwent 20 measurements between corresponding points of the lesion contours on the US image and 3D CT cross-section. An example of performed fusion is depicted in Fig. 9. The mean TRE across all measurements was found to be 1.97 ± 0.89 mm, and the complete distribution of errors is illustrated in Fig. 10. The survival function analysis revealed a 0.95 likelihood that the error remains within 3.41 mm. It should be noted that although the test was performed using CT imaging, the procedure and the user interface remain the same in the case of MRI.

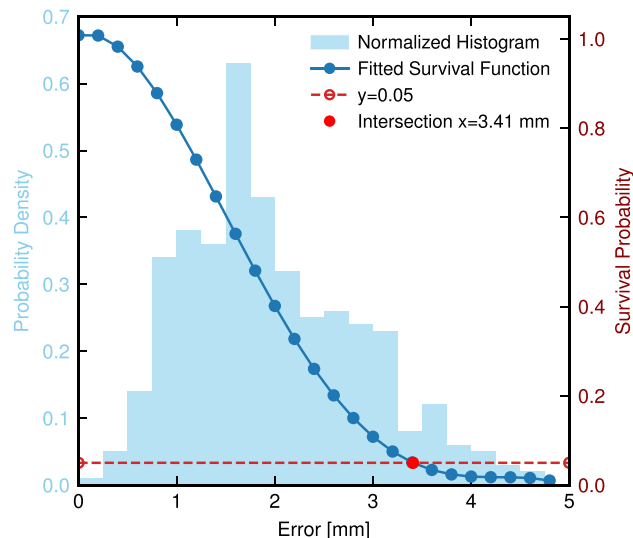


Figure 10. Distribution of target registration errors for 400 measurements. Survival function was fitted with 4th degree polynomial. The analysis indicates 0.95 likelihood that the error lies within 3.41 mm.

Biopsy navigation assessment

The impact of utilizing the developed mixed reality navigation system with markerless surgical instruments tracking was validated by simulating biopsy procedure on an ex vivo tissue model (a pork knee). The test did not incorporate fusion of US with CT/MRI but evaluated in practice suitability of both needle and US probe tracking accuracy. For the experiment, one wooden ball of a given size (diameter 8–20 mm, in 2 mm steps) was inserted into the tissue at a time. For each size, 10 injections were performed using both mixed reality navigation and US only. The US was used to visualize the ball position in the tissue and to confirm whether the needle successfully hit the target. To mitigate potential bias from cumulative operator learning, each puncture was performed from a different insertion site, with injection depths varying between 7 and 11 cm. Additionally to increase difficulty, each time the target ball was replaced with smaller size and repositioned to a new location within the subject.

In order to guide the user to perform a precise needle injection, the target and puncture sites had to be first selected. To select the target site, the user positioned both index finger tips to define a line that intersects the US image plane at the desired location. The puncture site was then selected with a single index finger tip. The selection and confirmation of points began with the use of appropriate voice commands. The two chosen points defined a trajectory that the needle should follow during insertion. Through needle tracking, the operator was guided to maintain the designated path²⁹. Mixed reality navigation interface is presented in Fig. 11.

The test was performed by a person without prior experience in performing medical procedures. The results are presented in Fig. 12. The target hit rate was calculated as the ratio of the number of successful punctures when needle has reached the ball to the total number of attempts for each ball size. The experiment demonstrated that utilizing mixed reality navigation significantly improves procedural accuracy, as the operator successfully reached the target in 87.14% of cases compared to only 32.86% when using US alone.

Discussion

The study validated the use of markerless US probe tracking in a mixed reality system for biopsy support. Applying a deep learning model to this task is challenging, as it must balance high accuracy and low inference time. According to the authors of ZebraPose, recognized as one of the fastest models for 6D pose estimation of seen objects with reasonably high precision, the model still operates in 250 ms²⁵. Performed optimizations enabled real-time operation and the applied post-processing aligned the model output with real-world data captured by depth camera, leading to increase pose estimation accuracy in comparison to alternative depth-based refinement methods²⁴. Compared to a well-established surgical tracking technology ClaroNav MicronTracker, the achieved accuracy proved suitable for clinical use.

Some recent studies have explored the markerless object pose estimation for HMD applications using only built-in sensors. Pöllabauer et al.⁴⁸ used a deep learning model to estimate 6D poses only from RGB images and anchor virtual overlays, reporting MSSD and MSPD values suitable for assistance but insufficient for high-precision alignment. Ali et al.⁴⁹ used Azure Object Anchors, trained on a 3D object model and relying solely on depth data, but noted limitations for similarly shaped or planar objects and for targets smaller than 70 cm. In contrast, our study combines RGB and depth data to take advantage of both modalities. RGB images provide color and texture information and are analyzed by deep learning models while depth camera data provides information about actual geometry and localization of the object. The employed framework can be applied to any object, although the resolution and quality of the built-in Time-of-Flight (ToF) camera might limit its applicability, which requires additional research.

Multi-modal fusion of US and CT/MRI is crucial in various clinical contexts, as the advantages of each modality enhance diagnosis and procedure planning. A substantial amount of previous research focused on

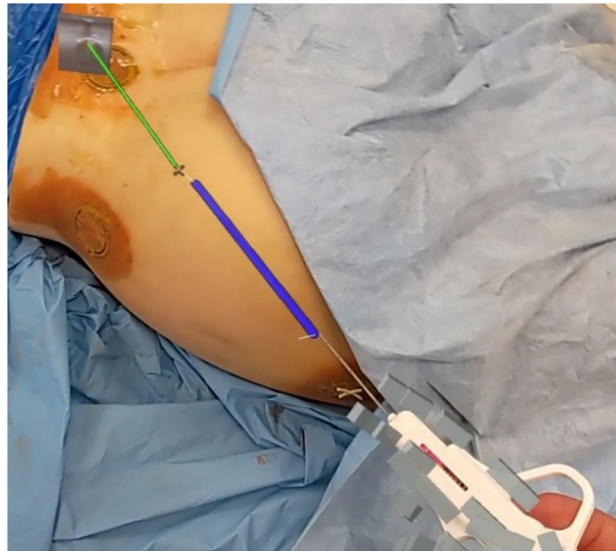


Figure 11. Interface of mixed reality navigation with US-guidance. The overlaid US image contains lesion region, allowing the system to designate target and injection points, guiding the user to align the estimated needle position (blue line) with the chosen trajectory (green line). The user task is to minimize the length of leading lines (white lines) that indicate the needle deviation from the trajectory.

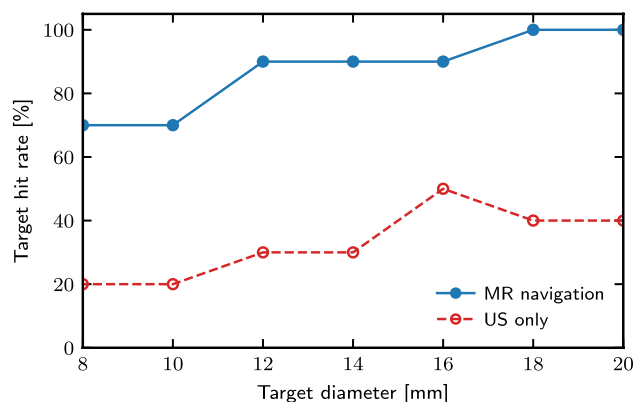


Figure 12. Mixed reality navigation assessment performed on the tissue for injections at depths in range 7–11 cm. The use of developed solution outperforms standard US-guidance in terms of target hit rate.

improving this process and handling its major difficulties. Gueziri et al.⁵⁰ proposed a framework for rigid registration between CT and intraoperative US of a single vertebra, achieving a median accuracy of 1.48 mm and requiring 11 seconds of computation time. Wei et al.⁵¹ employed a classification network to estimate the US probe orientation, then refined through segmentation and achieved a mean registration error of 4.7 ± 1.7 mm. Chi et al.⁵² proposed a method based on deep learning for registering 2D US with 3D CT for kidney scans of free breathing, achieved a mean contour distance of 0.79 ± 0.22 mm. Last but not least Wang et al.⁵³ proposed a solution to register 2D US with 3D MRI of liver images, obtaining mean TRE 4.95 ± 2.23 mm. The literature often highlights that a major challenge in registration involves US noise. The presented solution relies only on estimated US probe pose instead of complex intensity-based registration algorithms, yet it still delivers satisfactory performance, making it simple and adaptable across many anatomies and procedures. The user can perform real-time fusion of images as the results can be observed immediately. However, one limitation is the requirement for accurate 3D CT/MRI alignment with the patient, which is currently performed by manually identifying landmarks. Nonetheless, recent research in this field may lead to automate this process without the need for markers⁵⁴. Another study constraint is performing the fusion on a rigid, non-deformable phantom, which limits the ability to simulate US examination on human participants. In clinical practice, misalignment between live US and pre-procedural CT or MRI data may result from tissue deformation induced by patient repositioning, respiratory and cardiac movements, or surgical instruments pressure. Recent studies have explored non-rigid image registration methods⁵⁵; however, this topic lies beyond the scope of the present study.

Further work should also consider the potential misalignment between the 2D US plane and the 3D CT/MRI cross-section, although obtaining ground-truth data remains a challenge.

The standard US-guided biopsy lack of spatial information to coordinate needle movement in relation to the lesion position, especially for punctures out of the image plane. The developed system addresses this issue as the user perceive anchored US image with marked target and is guided to adjust proper needle placement. The conducted biopsy navigation assessment test proved usability of the solution as the target hit rate increased by over 2.5 times compared to using US alone. This directly affects the reduction of required punctures, thereby decreasing complication rates, blood loss, and patient discomfort. Mixed reality navigation may be particularly useful for targets near critical structures, improving spatial awareness and reducing cognitive load compared to 2D imaging alone. The image fusion with CT/MRI can be essential in targeting lesions that are poorly visible in the US. The system can also serve as a training tool, helping users attain improved performance when carrying out more complex biopsy procedures. Challenges of deploying the system for real clinical scenarios include maintaining sterile environment, ensuring quick setup time and remaining stable tracking despite occlusions, variable lightning, and long procedure times. The system must not cause user discomfort, fatigue or distraction which can lead clinicians to choose conventional imaging despite the theoretical navigation advantages.

One of the study limitations is relying only on the built-in ToF camera for depth measurements that is prone to errors caused by distractions such as lighting, reflections, colors and materials^{56,57}. The further work should explore the use of HMDs equipped with stereo cameras as they provide higher resolution and a limited amount of noise⁵⁸. The evaluation of the developed markerless US probe tracking pipeline on data captured with an Intel RealSense camera, as reported in the ablation study, demonstrated the portability of the solution. Nevertheless, generating precise disparity maps without compromising real-time performance continues to be a demanding task⁵⁹. Another technical constraint arises from using an HMD that relies on the IEEE 802.11ac wireless standard, as data transmission is a key factor. Employing a device that supports the newer IEEE 802.11ax standard could improve transfer speeds⁶⁰, and adopting IEEE 802.11be could further reduce latency⁶¹. It should also be emphasized that the reported end-to-end latency corresponds to the scenario in which both tools are used at the same time, which is unlikely to be common in routine practice. Lastly, although low latency is most critical when instruments are moved rapidly, precise surgical maneuvers are typically slow and deliberate, so this limitation is less significant in those contexts. The study also employed a single type of needle and US probe. The proposed method can be adapted to other tool types and is expected to maintain comparable accuracy, as clinical instruments generally share similar geometric characteristics and the estimation precision is mainly limited by accuracy of the depth map. The US probe used in this study features a symmetrical, low-texture design which are factors typically associated with reduced tracking performance, yet the method still yielded satisfactory results. However, smaller tools or those covered with sterile sheaths may pose additional challenges, which requires further investigation.

Conclusion

The study presented and evaluated mixed reality navigation system for biopsy that features fusion between 2D US and 3D CT/MRI. The developed markerless surgical instruments tracking method was validated against state-of-the-art RGB-only and RGB-D methods and achieved the best outcomes in case of US probe pose estimation. The method was also evaluated alongside a different navigation system, and the results were closely matched. The CT-US target registration error test indicated a 0.95 likelihood that the error remains within 3.41 mm which is suitable for clinical use. The biopsy navigation assessment test demonstrated that mixed reality navigation significantly enhances procedural success while minimizing human errors.

Data Availability

The datasets created and analyzed during the current study are not publicly available due to privacy and confidentiality restrictions but are available from the corresponding author on reasonable request.

Received: 28 January 2026; Accepted: 14 April 2026

Published online: 20 April 2026

References

1. Anjum, R. et al. Biopsy techniques: A review. *Journal of Oral Medicine, Oral Surgery, Oral Pathology and Oral Radiology* **10**, 181–184. <https://doi.org/10.18231/j.jooo.2024.034> (2024).
2. Veltri, A., Bargellini, I., Giorgi, L., Almeida, P. A. M. S. & Akhan, O. CIRSE Guidelines on Percutaneous Needle Biopsy (PNB). *Cardiovasc. Intervent. Radiol.* **40**, 1501–1513. <https://doi.org/10.1007/s00270-017-1658-5> (2017).
3. Serati, L. et al. Diagnostic yield and accuracy of image-guided percutaneous core needle biopsy of paediatric solid tumours: An experience from Italy. *Pediatric Hematology Oncology Journal* **6**, 12–17. <https://doi.org/10.1016/j.phoj.2020.10.003> (2021).
4. Kim, J. W. & Shin, S. S. Ultrasound-Guided Percutaneous Core Needle Biopsy of Abdominal Viscera: Tips to Ensure Safe and Effective Biopsy. *Korean J. Radiol.* **18**, 309–322. <https://doi.org/10.3348/kjr.2017.18.2.309> (2017).
5. Portela-Oliveira, E. et al. Ultrasound-guided percutaneous biopsy of thoracic lesions: High diagnostic yield and low complication rate. *Clin. Radiol.* **76**, 281–286. <https://doi.org/10.1016/j.crad.2020.12.004> (2021).
6. Guo, Y.-Q. et al. Ultrasound-Guided Percutaneous Needle Biopsy for Peripheral Pulmonary Lesions: Diagnostic Accuracy and Influencing Factors. *Ultrasound Med. Biol.* **44**, 1003–1011. <https://doi.org/10.1016/j.ultrasmedbio.2018.01.016> (2018).
7. Szaro, P., Wong, A., Blain, E., Gataa, K. G. & Geijer, M. The diagnostic value of ultrasound-guided percutaneous core needle biopsy of musculoskeletal soft tissue lesions. *Journal of Ultrasonography* **21**, 22–33. <https://doi.org/10.15557/JoU.2021.0004> (2021).
8. Huang, W. et al. Ultrasound-Guided Percutaneous Core Needle Biopsy of Peripheral Pulmonary Nodules ≤ 2 cm: Diagnostic Performance, Safety and Influence Factors. *Front. Oncol.* **11**, 671884. <https://doi.org/10.3389/fonc.2021.671884> (2021).
9. Nguyen, T., Plishker, W., Matisoff, A., Sharma, K. & Shekhar, R. HoloUS: Augmented reality visualization of live ultrasound images using HoloLens for ultrasound-guided procedures. *Int. J. Comput. Assist. Radiol. Surg.* **17**, 385–391. <https://doi.org/10.1007/s11548-021-02526-7> (2022).

10. Costa, N. et al. Augmented Reality-Assisted Ultrasound Breast Biopsy. *Sensors* **23**, 1838. <https://doi.org/10.3390/s23041838> (2023).
11. Li, H. et al. Navigate biopsy with ultrasound under augmented reality device: Towards higher system performance. *Comput. Biol. Med.* **174**, 108453. <https://doi.org/10.1016/j.combiomed.2024.108453> (2024).
12. Siddiqui, M. M. et al. Comparison of MR/Ultrasound Fusion-Guided Biopsy With Ultrasound-Guided Biopsy for the Diagnosis of Prostate Cancer. *JAMA* **313**, 390–397. <https://doi.org/10.1001/jama.2014.17942> (2015).
13. Bak, S., Budzen, K., Wesolowski, P., Stachura, T. & Gladysz, K. Review of diagnostic methods for prostate cancer with consideration of MRI-TRUS fusion biopsy. *Journal of Education, Health and Sport* **12**, 720–728. <https://doi.org/10.12775/JEHS.2022.12.09.085> (2022).
14. Musaddaq, T. & Musaddaq, B. Recent Advances in Image-Guided Tissue Sampling. *Cureus* **16**, e71613. <https://doi.org/10.7759/cureus.71613> (2024).
15. Fang, J.-H. et al. Comparative diagnostic accuracy of multiparametric magnetic resonance imaging-ultrasound fusion-guided biopsy versus systematic biopsy for clinically significant prostate cancer. *PeerJ* **11**, e16614. <https://doi.org/10.7717/peerj.16614> (2023).
16. Zhu, G. & Wang, Q. Comparisons between magnetic resonance/ultrasound fusion-guided biopsy and standard biopsy in the diagnosis of prostate cancer: A prospective cohort study. *Medicine* **97**, e11962. <https://doi.org/10.1097/MD.00000000000011962> (2018).
17. Chang, S. D. et al. MRI Targeted Prostate Biopsy Techniques: AJR Expert Panel Narrative Review. *Am. J. Roentgenol.* **217**, 1263–1281. <https://doi.org/10.2214/AJR.21.26154> (2021).
18. Marks, L., Young, S. & Natarajan, S. MRI-ultrasound fusion for guidance of targeted prostate biopsy. *Curr. Opin. Urol.* **23**, 43–50. <https://doi.org/10.1097/MOU.0b013e32835ad3ee> (2013).
19. Şahin, B., Dörücü, D., Tinay, İ., Filinte, D. & Çam, H. K. Is Cognitive MR Fusion Biopsy Superior to Standard TRUS Guided Prostate Biopsy?. *Our Clinical Experience. Bulletin of Urooncology* **21**, 35–39. <https://doi.org/10.4274/uob.galenos.2021.2021.2.4> (2022).
20. Naselli, A. et al. Fusion prostate biopsy: Tips and tricks to improve rigid registration. *Journal of Men's Health* **18**, 1, <https://doi.org/10.31083/j.jomh.1805120> (2022).
21. Cornud, F. et al. TRUS-MRI image registration: A paradigm shift in the diagnosis of significant prostate cancer. *Abdom. Imaging* **38**, 1447–1463. <https://doi.org/10.1007/s00261-013-0018-4> (2013).
22. Venderink, W., de Rooij, M., Sedelaar, J. P. M., Huisman, H. J. & Fütterer, J. J. Elastic Versus Rigid Image Registration in Magnetic Resonance Imaging-transrectal Ultrasound Fusion Prostate Biopsy: A Systematic Review and Meta-analysis. *Eur. Urol. Focus* **4**, 219–227. <https://doi.org/10.1016/j.euf.2016.07.003> (2018).
23. Hale, G. R. et al. Comparison of Elastic and Rigid Registration during Magnetic Resonance Imaging/Ultrasound Fusion-Guided Prostate Biopsy: A Multi-Operator Phantom Study. *J. Urol.* **200**, 1114–1121. <https://doi.org/10.1016/j.juro.2018.06.028> (2018).
24. Hein, J. et al. Next-generation surgical navigation: Marker-less multi-view 6DoF pose estimation of surgical instruments. *Med. Image Anal.* **103**, 103613. <https://doi.org/10.1016/j.media.2025.103613> (2025).
25. Su, Y. et al. ZebraPose: Coarse To Fine Surface Encoding for 6DoF Object Pose Estimation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 6738–6748, <https://doi.org/10.1109/CVPR52688.2022.00662> (2022).
26. Haugaard, R. L. & Buch, A. G. SurfEmb: Dense and Continuous Correspondence Distributions for Object Pose Estimation with L learnt Surface Embeddings, <https://doi.org/10.48550/arXiv.2111.13489> (2022).
27. Haugaard, R. L. & Iversen, T. M. Multi-view object pose estimation from correspondence distributions and epipolar geometry, <https://doi.org/10.48550/arXiv.2210.00924> (2023).
28. Agethen, N., Rosskamp, J., Koller, T. L., Klein, J. & Zachmann, G. Recurrent multi-view 6DoF pose estimation for marker-less surgical tool tracking. *Int. J. Comput. Assist. Radiol. Surg.* **20**, 1589–1599. <https://doi.org/10.1007/s11548-025-03436-8> (2025).
29. Trojak, M., Stanuch, M., Kurzyna, M., Darocha, S. & Skalski, A. Mixed Reality Biopsy Navigation System Utilizing Markerless Needle Tracking and Imaging Data Superimposition. *Cancers* **16**, 1894. <https://doi.org/10.3390/cancers16101894> (2024).
30. Wierzbicki, R. et al. 3D mixed-reality visualization of medical imaging data as a supporting tool for innovative, minimally invasive surgery for gastrointestinal tumors and systemic treatment as a new path in personalized treatment of advanced cancer diseases. *J. Cancer Res. Clin. Oncol.* **148**, 237–243. <https://doi.org/10.1007/s00432-021-03680-w> (2022).
31. Tsai, T.-Y. et al. Feasibility and accuracy of real-time 3D-holographic graft length measurements. *European Heart Journal - Digital Health* **5**, 101–104. <https://doi.org/10.1093/ehjdh/ztad071> (2024).
32. Bronowicki, K. et al. An attempt to evaluate the use of mixed reality in surgically treated pediatric oncology patients. *Npj Digital Medicine* **8**, 262. <https://doi.org/10.1038/s41746-025-01638-7> (2025).
33. Galyfos, G. et al. Using a novel three-dimensional holographic technology to perform open vascular surgery procedures. *Journal of Vascular Surgery Cases, Innovations and Techniques* **10**, 101440. <https://doi.org/10.1016/j.jvscit.2024.101440> (2024).
34. Ungureanu, D. et al. HoloLens 2 Research Mode as a Tool for Computer Vision Research. *arXiv preprint arXiv:2008.11239* <https://doi.org/10.48550/arXiv.2008.11239> (2020).
35. He, J., Li, P., Geng, Y. & Xie, X. FastInst: A Simple Query-Based Model for Real-Time Instance Segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 23663–23672, <https://doi.org/10.1109/CVPR52729.2023.02266> (2023).
36. Ge, Z., Liu, S., Wang, F., Li, Z. & Sun, J. YOLOX: Exceeding YOLO Series in 2021. *arXiv preprint arXiv:2107.08430* <https://doi.org/10.48550/arXiv.2107.08430> (2021).
37. Barath, D. & Matas, J. Progressive-x: Efficient, anytime, multi-model fitting algorithm. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2019).
38. Brachmann, E. et al. Learning 6D Object Pose Estimation Using 3D Object Coordinates. *European conference on computer vision* 536–551, https://doi.org/10.1007/978-3-319-10605-2_35 (2014).
39. Trojak, M., Jurgas, A., Stanuch, M., Kownacki, L. & Skalski, A. Mixed Reality 6D Object Pose Estimation using Deep Learning for Visual Markerless Surgical Navigation. *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* 947–952, <https://doi.org/10.1109/VRW66409.2025.00193> (2025).
40. Arun, K. S., Huang, T. S. & Blostein, S. D. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-9**, 698–700, <https://doi.org/10.1109/TPAMI.1987.4767965> (1987).
41. Peng, S., Liu, Y., Huang, Q., Zhou, X. & Bao, H. PVNet: Pixel-wise voting network for 6dof pose estimation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 4556–4565, <https://doi.org/10.1109/CVPR.2019.00469> (2019).
42. Lin, Y. et al. HiPose: Hierarchical Binary Surface Encoding and Correspondence Pruning for RGB-D 6DoF Object Pose Estimation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 10148–10158, <https://doi.org/10.1109/CVPR52733.2024.00967> (2024).
43. Burde, V., Benbihi, A., Burget, P. & Sattler, T. Comparative Evaluation of 3D Reconstruction Methods for Object Pose Estimation. *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* 7669–7681, <https://doi.org/10.1109/WACV61041.2025.00745> (2025).
44. Guo, N. et al. An Online Calibration Method for Microsoft HoloLens. *IEEE Access* **7**, 101795–101803. <https://doi.org/10.1109/ACCESS.2019.2930701> (2019).
45. Florkowska, A. et al. Metrological Analysis of HoloLens 2 for Visual Marker-Based Surgical Navigation. *2023 IEEE International Conference on Imaging Systems and Techniques (IST)* 1–6, <https://doi.org/10.1109/IST59124.2023.10355714> (2023).

46. Liu, H.-G. et al. A Bulk Retrospective Study of Robot-Assisted Stereotactic Biopsies of Intracranial Lesions Guided by Videometric Tracker. *Front. Neurol.* **12**, 682733. <https://doi.org/10.3389/fneur.2021.682733> (2021).
47. Eastman, R. D., Marvel, J., Falco, J. & Hong, T. Measurement science for 6DOF object pose ground truth. 2013 IEEE International Symposium on Robotic and Sensors Environments (ROSE) 196–201, <https://doi.org/10.1109/ROSE.2013.6698442> (2013).
48. Pöllabauer, T., Rücker, F., Franek, A. & Gorschlüter, F. Detection and Pose Estimation of Flat, Texture-Less Industry Objects on HoloLens Using Synthetic Training. *Image Analysis* 569–585, https://doi.org/10.1007/978-3-031-31438-4_37 (2023).
49. Ali, M. H., Damouni, R., Nassar, A., Shimshoni, I. & Szpiro, S. F. A. Feasibility of Real Time 3D Object Detection for Accessibility Use Cases. 2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) 1178–1179, <https://doi.org/10.1109/VRW66409.2025.00235> (2025).
50. Gueziri, H.-E., Drouin, S., Yan, C. X. B. & Collins, D. L. Toward real-time rigid registration of intra-operative ultrasound with preoperative CT images for lumbar spinal fusion surgery. *Int. J. Comput. Assist. Radiol. Surg.* **14**, 1933–1943. <https://doi.org/10.1007/s11548-019-02020-1> (2019).
51. Wei, W., Haishan, X., Alpers, J., Rak, M. & Hansen, C. A deep learning approach for 2D ultrasound and 3D CT/MR image registration in liver tumor ablation. *Comput. Methods Programs Biomed.* **206**, 106117. <https://doi.org/10.1016/j.cmpb.2021.106117> (2021).
52. Chi, Y. et al. A two-step deep learning method for 3DCT-2DUS kidney registration during breathing. *Sci. Rep.* **13**, 12846. <https://doi.org/10.1038/s41598-023-40133-5> (2023).
53. Wang, Y. et al. Multimodal registration of ultrasound and MR images using weighted self-similarity structure vector. *Comput. Biol. Med.* **155**, 106661. <https://doi.org/10.1016/j.combiomed.2023.106661> (2023).
54. Jurgas, A., Trojak, M., Stanuch, M. & Skalski, A. Markerless Registration and Visualization of Volumetric Medical Data for Enhanced Surgical Precision Using Head-Mounted Displays. 2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) 925–930, <https://doi.org/10.1109/VRW66409.2025.00189> (2025).
55. Mitra, J. et al. A hybrid deformable registration method to generate motion-compensated 3D virtual MRI for fusion with interventional real-time 3D ultrasound. *Int. J. Comput. Assist. Radiol. Surg.* **18**, 1501–1509. <https://doi.org/10.1007/s11548-023-02833-1> (2023).
56. He, Y., Liang, B., Zou, Y., He, J. & Yang, J. Depth Errors Analysis and Correction for Time-of-Flight (ToF) Cameras. *Sensors* **17**, 92. <https://doi.org/10.3390/s17010092> (2017).
57. Schuon, S., Theobalt, C., Davis, J. & Thrun, S. High-quality scanning using time-of-flight depth superresolution. 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops 1–7, <https://doi.org/10.1109/CVPRW.2008.4563171> (2008).
58. Agresti, G., Minto, L., Marin, G. & Zanuttigh, P. Deep learning for confidence information in stereo and tof data fusion. Proceedings of the IEEE International Conference on Computer Vision Workshops 697–705, <https://doi.org/10.1109/ICCVW.2017.88> (2017).
59. Tosi, F., Bartolomei, L. & Poggi, M. A Survey on Deep Stereo Matching in the Twenties. *Int. J. Comput. Vision* **133**, 4245–4276. <https://doi.org/10.1007/s11263-024-02331-0> (2025).
60. Bandyopadhyay, D. et al. Network Throughput Improvement in Wi-Fi 6 over Wi-Fi 5: A Comparative Performance Analysis. 2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE) 1–6, <https://doi.org/10.1109/ICCECE.2023.10085684> (2023).
61. Liu, X., Dong, Y., Li, Y., Lin, Y. & Gan, M. IEEE 802.11be Wi-Fi 7: Feature Summary and Performance Evaluation. *IEEE Communications Standards Magazine* 1–7, <https://doi.org/10.1109/MCOMSTD.2025.3640281> (2025).

Acknowledgements

We would like to express our gratitude to Michal Kluska for his valuable assistance with the runtime analysis.

Author contributions

Conceptualization: M.T., M.S. and A.S.; methodology: M.T., M.S. and A.S.; software: M.T.; formal analysis and investigation: M.T.; resources: A.S.; writing—original draft preparation: M.T.; writing—review and editing: M.T., M.S. and A.S.; visualization: M.T.; supervision: A.S.; funding acquisition: A.S.

Funding

This work was funded by the Ministry of Science and Higher Education in Poland under the Implementation Doctorate Programme, project number DWD/6/0235/2022.

Declarations

Competing interests

The authors are MedApp S.A. employees and MedApp S.A. is the company that produces the CarnaLife Holo solution which was used for data visualization.

Additional information

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026