

Article

Six-Degree-of-Freedom Freehand 3D Ultrasound: A Low-Cost Computer Vision-Based Approach for Orthopedic Applications

Lorenzo De Sanctis ^{1,2}, Arianna Carnevale ¹, Carla Antonacci ^{1,3}, Eliodoro Faiella ¹, Emiliano Schena ^{1,3}
and Umile Giuseppe Longo ^{1,2,*}

- ¹ Fondazione Policlinico Universitario Campus Bio-Medico, Via Álvaro del Portillo, 200, 00128 Rome, Italy; lorenzodesanctis1@gmail.com (L.D.S.); arianna.carnevale@policlinicocampus.it (A.C.); carla.antonacci@unicampus.it (C.A.); e.faiella@policlinicocampus.it (E.F.); e.schena@unicampus.it (E.S.)
² Research Unit of Orthopaedic and Trauma Surgery, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Via Álvaro del Portillo, 21, 00128 Rome, Italy
³ Laboratory of Measurement and Biomedical Instrumentation, Department of Engineering, Università Campus Bio-Medico di Roma, Via Álvaro del Portillo, 21, 00128 Rome, Italy
* Correspondence: g.longo@policlinicocampus.it

Abstract: In orthopedics, X-rays and computed tomography (CT) scans play pivotal roles in diagnosing and treating bone pathologies. Machine bulkiness and the emission of ionizing radiation remain the main problems associated with these techniques. The accessibility and low risks related to ultrasound handling make it a popular 2D imaging method. Indeed, 3D ultrasound assembles 2D slices into a 3D volume. This study aimed to implement a probe-tracking method for 6 DoF 3D ultrasound. The proposed method involves a dodecahedron with ArUco markers attached, enabling computer vision tracking of the ultrasound probe's position and orientation. The algorithm focuses on the data acquisition phase but covers the basic reconstruction required for data generation and analysis. In the best case, the analysis revealed an average error norm of 2.858 mm with a standard deviation norm of 5.534 mm compared to an infrared optical tracking system used as a reference. This study demonstrates the feasibility of performing volumetric imaging without ionizing radiation or bulky systems. This marker-based approach shows promise for enhancing orthopedic imaging, providing a more accessible imaging modality for helping clinicians to diagnose pathologies regarding complex joints, such as the shoulder, replacing standard infrared tracking systems known to suffer from marker occlusion problems.

Keywords: freehand 3D ultrasound; musculoskeletal ultrasound; probe tracking; ArUco markers; orthopedics



Citation: De Sanctis, L.; Carnevale, A.; Antonacci, C.; Faiella, E.; Schena, E.; Longo, U.G. Six-Degree-of-Freedom Freehand 3D Ultrasound: A Low-Cost Computer Vision-Based Approach for Orthopedic Applications. *Diagnostics* **2024**, *14*, 1501. <https://doi.org/10.3390/diagnostics14141501>

Academic Editor: Zhuhuang Zhou

Received: 19 May 2024

Revised: 30 June 2024

Accepted: 10 July 2024

Published: 12 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Musculoskeletal disorders are considered to be the leading worldwide cause of pain and physical impairment [1,2]. Among them, shoulder diseases (SDs) play a significant role, and patients who suffer from SDs may experience several consequences, such as reduced range of motion, worse quality of life, and severe restrictions in performing Activities of Daily Living [3–5]. The early diagnosis and treatment of musculoskeletal pathologies like osteoarthritis are essential to prevent additional joint damage and improve patients' quality of life [3–5].

Medical imaging plays a crucial role in the diagnosis and treatment of musculoskeletal diseases [6]. The primary imaging techniques used for pathology diagnosis are X-rays, computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound (US) imaging [7]. In X-ray and CT scans, the examination involves the patient being exposed to ionizing radiation, with a minimum of 0.02 mSv for a chest X-ray and 8.00 mSv for the corresponding CT scan [8]. The use of ionizing radiation requires the application of specific protocols and the execution of exams in shielded rooms to avoid exposing personnel,

patients, and visitors to unnecessary doses [9]. MRI, even if it does not involve the emission of ionizing radiation, requires bulky instruments in dedicated spaces. Not all patients can undergo exams involving the intense magnetic fields that this technique uses [10]. US uses mechanical (sound) waves and their interaction with the tissues to produce an image [11]. This imaging system does not require the emission of ionizing radiation; it is considered a relatively safe technique in most situations, and there are commercially available devices that can be transported easily [12]. Standard B-mode US can output 2D cross-sections of the anatomical region scanned [13].

Freehand 3D US is an imaging technique that consists of manually moving a conventional 2D US probe over the area of interest to acquire a series of 2D images, which are then reconstructed into a three-dimensional representation. This approach relies on tracking the position and orientation of the probe, often using a tracking device or external sensor, to accurately assemble 2D slices into a 3D volume [14].

The first US probe optical tracking attempts were detailed in the literature as early as 1990 [15]. Today, infrared (IR) stereophotogrammetric optical tracking systems vary based on the number of cameras and the application. For example, compact systems with only two cameras are suitable for use in the operating room, whereas entire laboratories are employed for kinematic gait analysis [16,17]. These systems recognize only the marker as a generic point, and six-degree-of-freedom (6 DoF) tracking is possible only with a cluster of at least three rigidly attached markers. Indeed, 6 DoF tracking of an object might require more than two cameras, as the object itself could obscure the marker's view from the camera, a phenomenon known as self-occlusion.

Several methods for probe-tracking are reported in the literature [18]. In particular, freehand 3D ultrasound localization systems may include optical tracking systems, mechanical devices, electromagnetic sensors, and acoustic systems [19–21]. The mentioned systems have some limitations: optical systems are prone to occlusion, mechanical systems are often bulky and heavy, electromagnetic systems have low positional accuracy and cannot be applied to all patients, and acoustic systems have non-negligible latency [22]. In the literature, an attempt to validate freehand 3D ultrasound with a low-cost approach is reported, using a single printable fiducial marker to capture the US probe trajectory, albeit without solving the occlusion problem and limiting the freedom of scanning [23].

This study aimed to implement a probe-tracking method utilizing a standard RGB camera, an identification-tracking algorithm based on computer vision (CV), and binary ArUco fiducial markers (consisting of a black square with a white pattern made of paper) [24]. The idea was to create a cluster of ArUco markers glued onto a dodecahedron as support and use an algorithm to assess the positions and orientations of the US images in the camera frame, applying the same principle used in [25] to track a ball pen. Attaching this cluster to the US probe makes it possible to calculate a rigid body transformation to obtain a volume filled with the acquired US images. This multiple passive marker method could help physicians to perform and interpret US scans of curved surfaces such as human joints by solving the occlusion problem of traditional optical systems.

2. Materials and Methods

The proposed approach's goal was to map the pixels of an US image into a volume. It was, thus, necessary to assign, to each pixel i , a set of three coordinates (x_i, y_i, z_i) to define its position $(\mathbf{u}_i)$ in space relative to the camera reference frame as follows:

$$\mathbf{u}_i = (x_i, y_i, z_i)^T$$

Pixel coordinates were calculated through a concatenation of homogeneous transformation matrices. A rigid transformation in $SE(3)$ is written in the form of a 4×4 matrix

$$\mathbf{T} \in SE(3) : \mathbf{T} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix}$$

where $\mathbf{R}$ is the rotation sub-matrix such that

$$\mathbf{R} \in \left\{ \mathbf{R} \in M_{3,3}(\mathbb{R}) : \mathbf{R}^\top = \mathbf{R}^{-1}, \det(\mathbf{R}) = +1 \right\}$$

and $\mathbf{t} \in \mathbb{R}^3$ is the translation vector [26,27]. To maintain dimensional correctness during multiplication, vectors $\mathbf{u}$ were expressed as points $\mathbf{p} = \begin{pmatrix} \mathbf{u} \\ 1 \end{pmatrix}$ so that $\mathbf{p}' = \mathbf{T} \cdot \mathbf{p}$. Specifically, the US image pixels were transformed from the image reference frame to the probe virtual frame, from the probe frame to the marker cluster frame, and from the marker cluster frame to the camera frame, as shown in Figure 1.

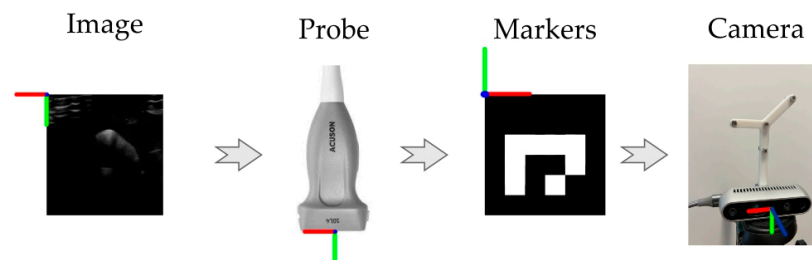


Figure 1. A representation of the frame transformations required to convert 2D image pixels in 3D camera coordinates. Reference axis colors: red (x-axis), green (y-axis), blue (z-axis).

For this study, a RealSense D435i (Intel Corporation Inc., Santa Clara, CA, USA) camera was used to obtain a video stream, but its depth capabilities have not been utilized in any way. The detection algorithm requires the marker to be attached to a planar surface. A regular dodecahedron was designed using a Computer-Aided Design (CAD) software (Fusion 360, version 2.0.18961, Autodesk Inc., San Francisco, CA, USA) and fabricated in white polylactic acid with the assistance of a 3D printer (Ender 3, Creality Co, Ltd., Shenzhen, China) (Figure 2a). A set of twelve 25 mm edge ArUco markers were printed on standard A4 paper using an office laser printer. Each of the twelve markers was cut out and glued using a dedicated mask (Figure 2b) onto one face of the dodecahedron with its side parallel to the dodecahedron edge at a 2.5 mm distance from it. A universal fixture was fabricated to attach the dodecahedron posteriorly to the US probe body (Figure 2c). At the meeting point between three edges, posteriorly, a cluster of IR markers was press fitted for tracking the probe with the reference technique (Figure 2d). To record the position of the IR markers, a set of 10 cameras was used (Qualisys, AB, Göteborg, Sweden).

Another IR marker cluster was placed on the camera using a dedicated fixture to calculate the homogeneous transformation between the motion capture system frame and the camera frame (Figure 3). The marker positions relative to the camera frame were extracted from the CAD software of the fixture and the technical documentation of the camera.

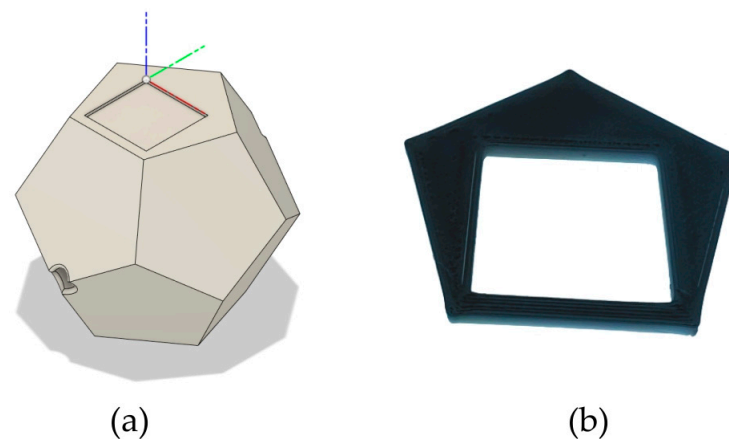


Figure 2. Cont.

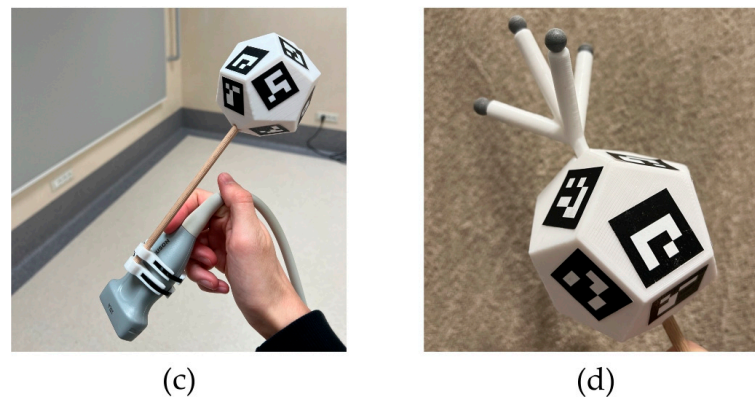


Figure 2. (a) Dodecahedron CAD design, (b) marker placement mask, (c) US probe fixture, and (d) removable IR cluster module.

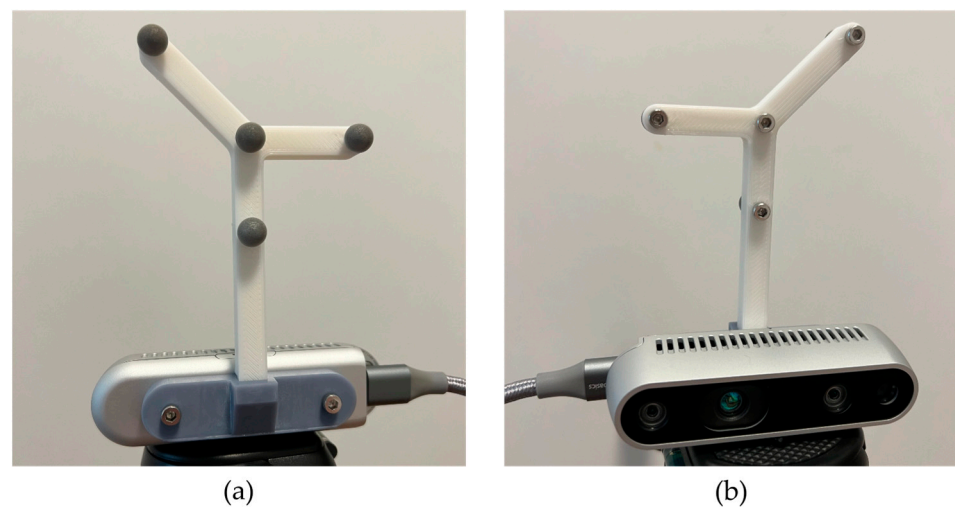


Figure 3. Camera IR marker cluster: (a) rear view and (b) front view.

To determine the required homogeneous transformation between the dodecahedron and the US probe and between the probe and the US video frame (i.e., probe calibration), a calibrator phantom was necessary [28]. Into a specifically designed container, two steel wires ($\varnothing$ 0.5 mm) were crossed at the first level at a depth of 8 mm from the surface (proximal layer in Figure 4a). Below this level, at a depth of 18 mm, a second one consisted of ten parallel wires ($\varnothing$ 0.5 mm) spaced 4 mm apart (distal layer in Figure 4a). Two tests were conducted to choose the following filling material (i.e., substrate): acetic silicone, shown in Figure 4b, and a 6.6% *w/w* agarose gel, as proposed by [29], mixed with benzalkonium chloride [30]. A mask that constrains the probe used to move perpendicular to the parallel wires was 3D printed and press-fitted onto the phantom.

The previously described calibrator was connected to a modular system consisting of two additional ArUco markers and an IR marker cluster. A linear transducer US probe (Acuson 10L4, Siemens Healthineers AG, Erlangen, Germany) was used.

Calibration was initiated when the operator, seeing only a single point (formed by the intersection of the two steel wires on the first layer) and the bottom of the container on the monitor, captured a snapshot of the tracking system marker trajectories and the US video stream. If multiple points or lines were present in the acquired US image, as shown in Figure 5, it would not be possible to reconstruct the calibration matrix because the calibrator landmarks (e.g., the intersection of the two wires on the first layer) would be lost. The calibration matrix was calculated as follows: first, the transformation from the dodecahedron frame to the centre of the two wires was found with the appropriate rotation, given by the formula

$$\mathbf{T}_{Cam2Dod}^{-1} \mathbf{T}_{Cam2Cal} = \mathbf{T}_{Probe}$$

Knowing that

$$\mathbf{T}_{Cam2Dod} \mathbf{T}_{Probe} \mathbf{T}_{Cam2Cal}^{-1} = \mathbf{I}_4$$

and then the pixel size was determined; subsequently, the transformation between the origin of the US frame and the origin of the calibrator was found, characterized by a simple two-dimensional translation.

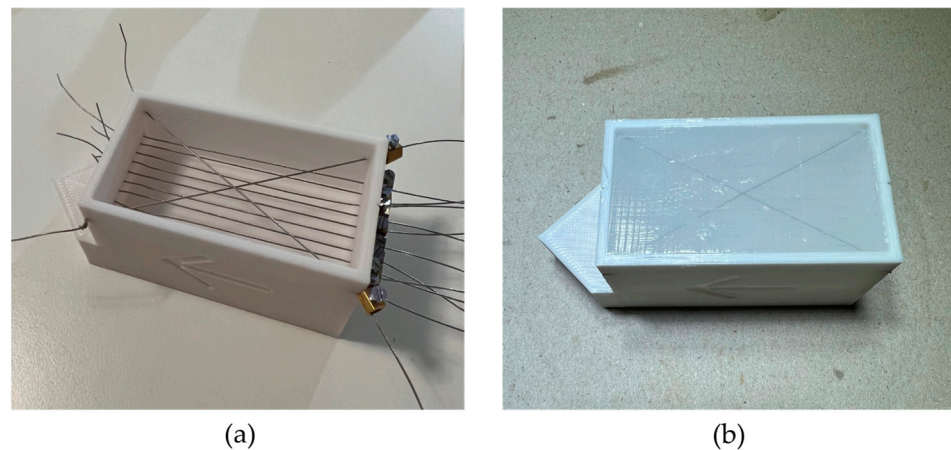


Figure 4. Calibrator phantom: (a) crossed-wire setup on multiple levels (b) phantom after silicone gel complete cross-linking.

To simulate a real US scan feeling, a phantom representing the humeral portion of the upper limb was made. The 3D model, segmented from an anonymized CT scan, is shown in Figure 6a. The bone model, made of PLA, was immersed in a 3% *w/v* agarose–water gel composed following the agar phantom operating standard procedure to mimic the soft tissue properties (Figure 6b) [31–33].

To capture and process data, a computer with a 6-core, 12-thread CPU (Ryzen 5 5600X, Advanced Micro Devices, Inc., Santa Clara, CA, USA), 16 GB of memory (Corsair Gaming, Inc., Fremont, CA, USA), and a 1TB NVMe solid-state storage was used. Factory camera calibration parameters were used to compute all data. After orienting the camera towards the scanning site and setting the acquisition parameters (capture duration, US machine’s region of interest, and starting delay), the software synchronized using Precision Time Protocol (PTPv2), the motion capture-dedicated workstation’s clock with the acquisition computer one, and started the IR recording [34,35]. The software then populated two arrays, respectively, with RGB frames from the camera, and their arrival timestamps until the capture interval condition was satisfied. Given the positions of each uniquely identified marker corner in the local frame (extracted from the CAD software) and having the coordinates of at least 4 points in the camera view, the dodecahedron pose was calculated by solving the Perspective-n-Point problem [36]. In this study, two tests were conducted: The first attempt consisted of computing the pose of the dodecahedron $\mathbf{T}_{Cam2Dod}^{RGB}$ and concatenating it with the probe matrix $\mathbf{T}_{Probe}$ obtained from the calibration phase (i.e., concatenation method), and the second one consisted of transforming all the corner coordinates in the probe frame and then directly computing the probe pose $\mathbf{T}_{Cam2Probe}^{RGB}$ (i.e., direct PnP method). To accomplish those tasks, an initial solution was obtained via the method proposed in [37] using the RANdom Sample And Consensus (RANSAC) algorithm [38]. The set of corner vectors was augmented with additional ones generated by averaging the identified pairs of points. Assuming k as the number of detected markers, the additional points were calculated as follows:

$$\mathbf{u}_{avg} = \frac{\mathbf{u}_i + \mathbf{u}_j}{2}, \quad \forall i, j = 1, \dots, k \mid i \neq j$$

The inlier input points selected by the RANSAC algorithm were used to refine the initial pose using the Levenberg–Marquardt optimization scheme [39,40]. Additionally, the individual components of the translation vector $\mathbf{t}$ of each transformation matrix were filtered through the Savitzky–Golay method [41]. For each set of coordinates obtained from the IR system’s software (Qualisys Track Manager, version 2024.1, Qualisys AB, Göteborg, Sweden), knowing the respective coordinates in the dodecahedron frame, the optimal transformation matrix between the motion capture system frame and the dodecahedron frame was calculated using the Kabsch–Umeyama algorithm [42,43]. The motion capture *World to Dodecahedron* transformation was

$$\mathbf{T}_{World2Dod} = \mathbf{T}_{World2Cam} \mathbf{T}_{Cam2Dod}^{IR}$$

where $\mathbf{T}_{World2Cam}$ is the same among all frames given the static pose of the camera, while $\mathbf{T}_{Cam2Dod}^{IR}$ varies.

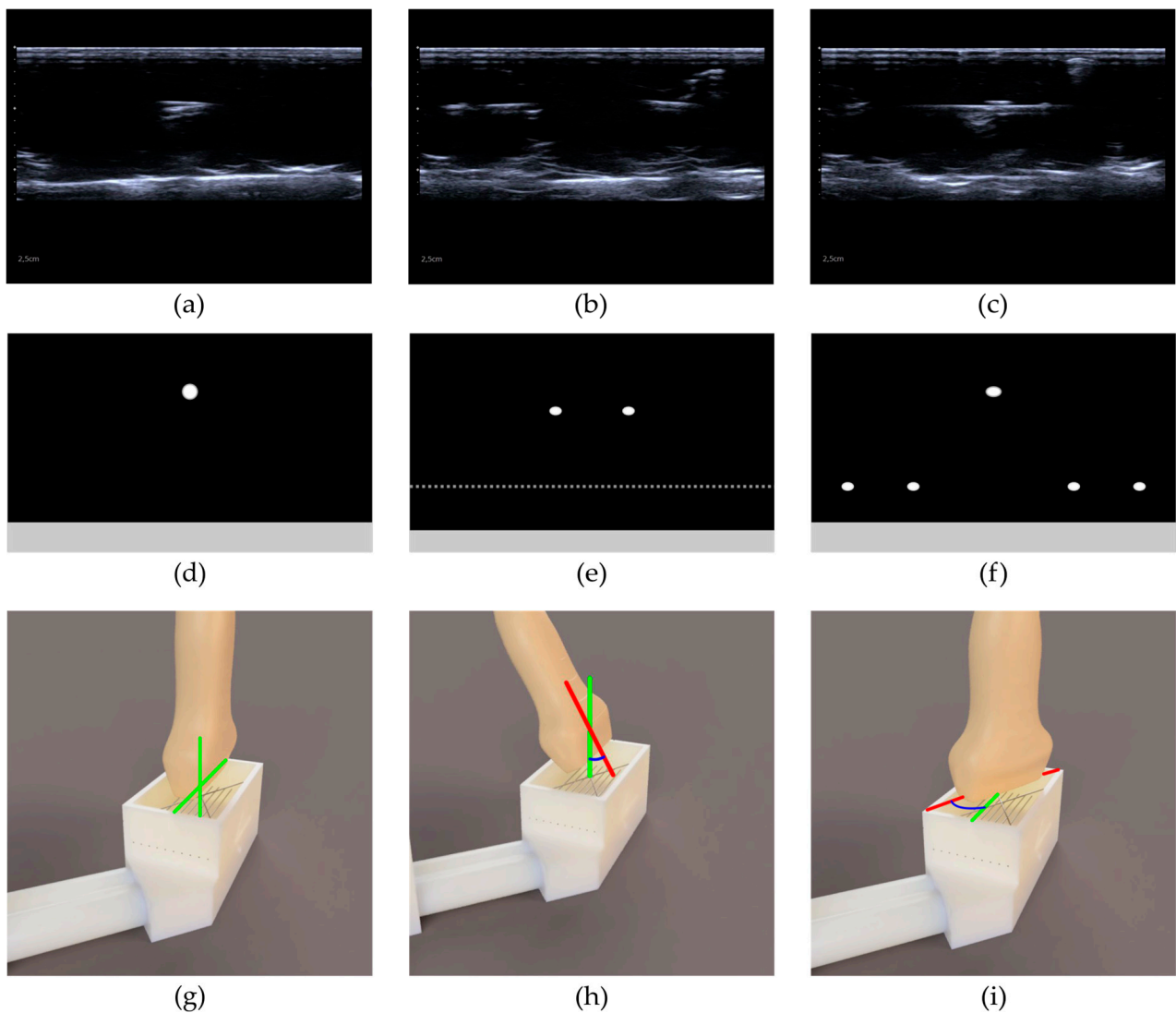


Figure 5. Probe calibration: (a) optimal case, (d) illustration, and (g) render; (b) pitch error, (e) illustration, and (h) render; and (c) roll error, (f) illustration, and (i) render. In the second row, slice thickness is assumed to be nearly zero to simplify the concept. In the third row, the angle in blue represents the rotation for which the above error is obtained.

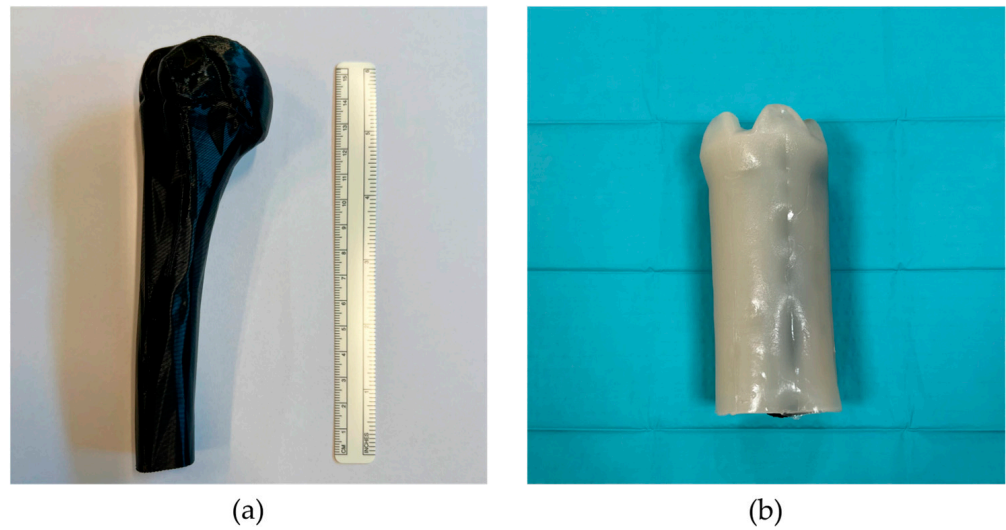


Figure 6. Phantom: (a) a 3D printed humerus model and (b) final agarose gel phantom.

Since the video sampling frequency of the RGB camera (60 Hz) was lower than the IR system's one (200 Hz), data resampling (at 60 Hz) was needed: for the rotation sub-matrix, converted temporarily into quaternion form, the Spherical Linear intERPolation (SLERP) algorithm was used [44], while for translation vectors, linear interpolation was performed on each of the components. US reconstruction was beyond the scope of this article; thus, a set of 4 pixels representing the corners of a generic US image was considered. Assuming that the probe's reference frame was coincident with the transducer geometric centre orthogonal projection on the probe surface, with the y -axis pointing in the direction of the US beam and the x -axis parallel with the transducer array direction, while also considering the Field of View (FOV) of the US probe to be 38 mm and the general useful depth d to be 30 mm, the corners of the image had the following coordinates:

$$\begin{aligned} \mathbf{p}_{c1} &= \left(\frac{FOV}{2}, 0, 0, 1 \right)^T, & \mathbf{p}_{c2} &= \left(\frac{FOV}{2}, d, 0, 1 \right)^T \\ \mathbf{p}_{c3} &= \left(-\frac{FOV}{2}, d, 0, 1 \right)^T, & \mathbf{p}_{c4} &= \left(-\frac{FOV}{2}, 0, 0, 1 \right)^T \end{aligned}$$

For each acquired pose, the corners absolute coordinates were calculated using the reference matrix

$$\mathbf{p}_{ref} = \mathbf{T}_{Cam2Dod}^{IR} \mathbf{T}_{Probe} \mathbf{p}, \quad \mathbf{p} \in \{\mathbf{p}_{c1}, \mathbf{p}_{c2}, \mathbf{p}_{c3}, \mathbf{p}_{c4}\}$$

using the direct PnP method matrix

$$\mathbf{p}_{direct} = \mathbf{T}_{Cam2Probe}^{RGB} \mathbf{p}, \quad \mathbf{p} \in \{\mathbf{p}_{c1}, \mathbf{p}_{c2}, \mathbf{p}_{c3}, \mathbf{p}_{c4}\}$$

and using the dodecahedron origin matrix

$$\mathbf{p}_{concat} = \mathbf{T}_{Cam2Dod}^{RGB} \mathbf{T}_{Probe} \mathbf{p}, \quad \mathbf{p} \in \{\mathbf{p}_{c1}, \mathbf{p}_{c2}, \mathbf{p}_{c3}, \mathbf{p}_{c4}\}$$

To assess the accuracy and precision of the proposed system compared to the reference, the mean and standard deviation (STD) of the error were considered. Let N be the total number of corners analyzed (equivalent of the distributed pixels in the volume)

$$\boldsymbol{\varepsilon}_{direct} = \mathbf{p}_{direct} - \mathbf{p}_{ref}$$

be the error of the direct PnP method and

$$\boldsymbol{\varepsilon}_{concat} = \mathbf{p}_{concat} - \mathbf{p}_{ref}$$

be the error of the concatenation method, with $\bar{\varepsilon}$ as their relative mean and σ as their relative standard deviation. Thus, the following were considered:

$$\bar{\varepsilon} = \frac{1}{N} \sum \varepsilon, \quad \sigma = \sqrt{\frac{1}{N} \sum (\varepsilon - \bar{\varepsilon})^2}$$

A total of four consecutive tests were conducted to evaluate the accuracy and precision of the proposed system. The experiments were performed under consistent conditions throughout the tests: the dodecahedron was not repositioned on the US probe, and one probe calibration sequence was completed before conducting the tests. The same operator performed all experiments on the same phantom. The IR system was calibrated following the standard procedure. The scan on the model was performed longitudinally, crossing over the same spot twice.

3. Results

Numerical results of the four experiments are presented in Table 1. For easier data interpretation, the Euclidean norm of the mean error and the STD were also reported.

Table 1. The numerical results of the performed recordings.

Method	Test #	$\bar{\varepsilon}$ [mm]			$\ \bar{\varepsilon}\ $ [mm]	σ [mm]			$\ \sigma\ $ [mm]
		<i>x</i>	<i>y</i>	<i>z</i>		<i>x</i>	<i>y</i>	<i>z</i>	
Origin ¹	1	0.197	0.663	−0.045	0.693	2.655	2.686	3.312	5.024
	2	−0.135	1.734	−3.149	3.597	4.352	2.184	7.550	8.241
	3	−0.181	1.765	1.464	2.301	1.110	1.219	4.206	4.518
	4	−1.712	0.611	4.020	4.414	2.992	2.216	3.427	5.061
Direct	1	3.947	0.895	0.792	4.124	1.899	2.020	5.005	5.722
	2	0.928	0.599	−2.934	3.135	4.352	2.184	7.550	8.984
	3	−0.857	0.453	2.689	2.858	2.908	1.920	4.299	5.534
	4	0.675	3.860	0.373	3.937	2.206	1.773	5.817	6.469
Concatenation	1	3.893	0.928	0.629	4.051	2.578	2.478	5.965	6.954
	2	0.964	0.589	−2.797	3.017	4.860	2.678	7.643	9.445
	3	−0.846	0.505	2.596	2.777	3.539	2.496	4.561	6.290
	4	0.884	3.581	1.348	3.927	3.296	2.342	6.792	7.905

¹ For the dodecahedron origin, data are obtained by multiplying by the null element in SE(3), and it does not consider the rotational component of the last transformation matrix.

Test 3 yielded the lowest STD norm of 5.534 mm, with a mean error norm of 2.858 mm, using the direct method. The concatenation method, applied to the same data, produced an STD norm of 6.290 mm, with a mean error norm of 2.777 mm. Dodecahedron origin and probe trajectories were plotted and are shown in Figure 7a, whereas comparisons between direct and concatenation methods are presented in Figure 7b using the scatter plots of the errors.

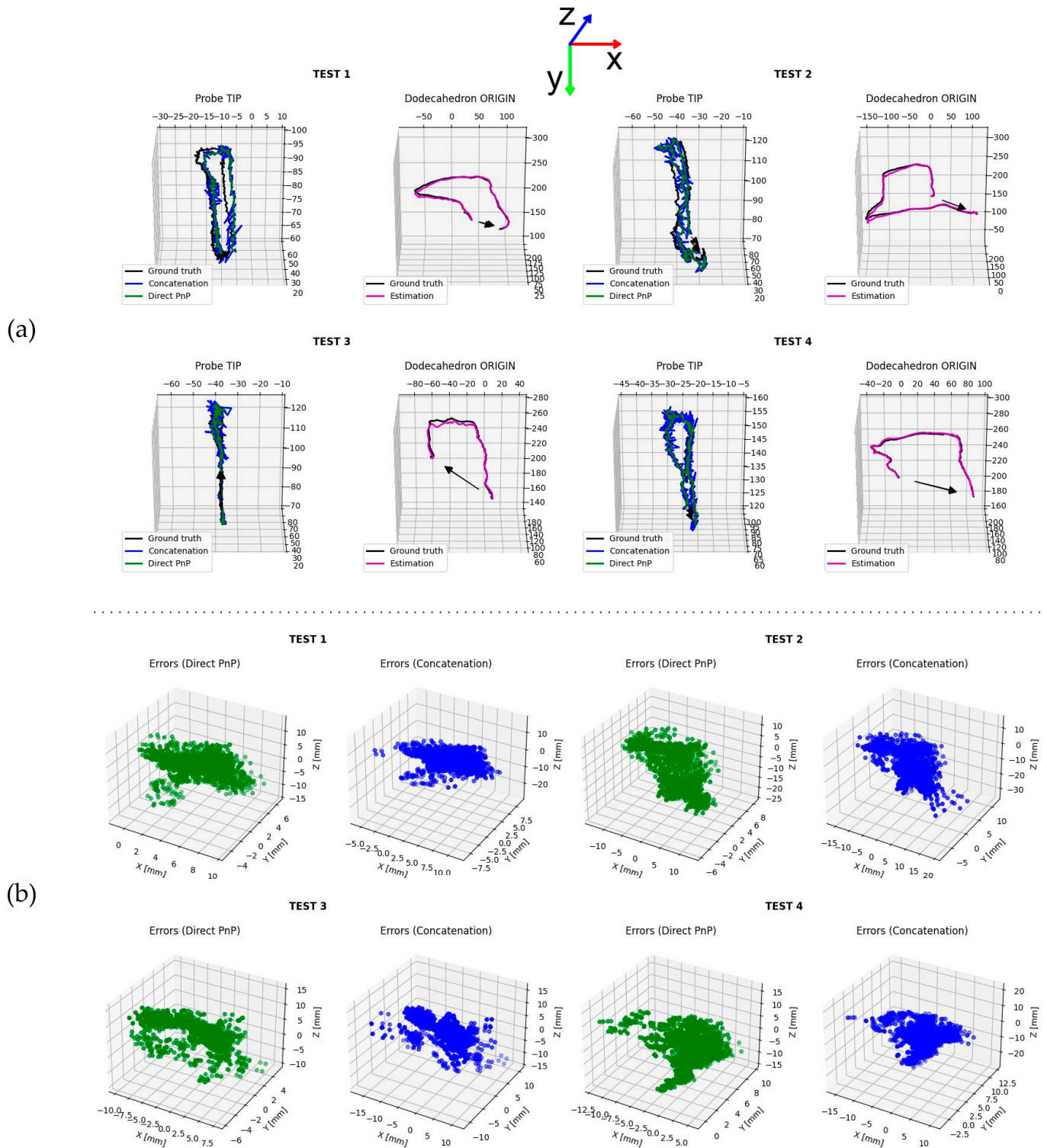


Figure 7. Plots: (a) probe tip and dodecahedron origin trajectories during all acquisitions and (b) probe tip trajectory error scatter plots. Axes are not equally scaled for visualization purposes.

4. Discussion

Starting from a video camera and attaching ArUco markers to a dodecahedron, fixed to an US probe, it was possible to reconstruct the trajectory of the instrument using a low-cost system. The results were compared to an infrared optical tracking system considered to be the reference. The choice of using the RealSense D435i stereo camera was exclusively dictated by the possibility of having extensive technical documentation, a factory camera

calibration, and mechanical drawings of the device. The regular dodecahedron, among all platonic solids, has a high number of faces, allowing the simultaneous detection of multiple ArUco markers, and it has a regular pentagon as its faces, permitting a maximum marker size equal to the edge length. Multiple-marker detection was needed to discern eventual pose ambiguities generated during the PnP problem resolution. The marker size choice was a compromise between visibility and the compactness of the dodecahedron: a larger marker makes detection simpler but also increases the dodecahedron's dimensions, making it bulkier and uncomfortable to manipulate. Regarding the calibrator, the 4 mm spacing between the wires in the second layer is greater than the average slice thickness described for a linear probe but adequate to prevent artifacts caused by the beam shape [45,46]. The agarose substrate of the calibrator allows a sound velocity of 1571 ± 12 m/s compared to that of water, which is 1496 m/s at 25 °C [47]. The sound velocity in silicone corresponds to 1030 ± 60 m/s, which is significantly lower than that of water. This makes it not an ideal material for the described application [48]. The low deformation of both materials allowed us to maintain the probe at a constant distance of 8 mm from the first level and 18 mm from the second. The silicone phantom demonstrated much more solidity than the agarose one.

In this type of application, the main causes of artifacts are the distance between the camera sensor and the dodecahedron, camera calibration, and detection algorithm. The results in Figure 7a highlighted an inverse correlation between the z-axis distance of the markers from the camera sensor plane and the trajectory accuracy. Even with sub-pixel marker detection, the features are shrunk in the image as the object moves away from the sensor depending on the focal lengths of the camera, thus making the pose estimation less accurate. Furthermore, camera calibration, not mentioned in this work given the factory calibration used, is crucial; indeed, the lens manifests optical distortion affecting pixels that are further away from the optical centre of the camera. This phenomenon can be seen especially in Test 2 (Figure 7a), where the dodecahedron origin trajectory gradually shifts from the ground truth as the origin moves away from the origin in the $-x$ direction. Fine-tuning the marker detection and pose refining algorithms can improve the accuracy and precision of the system. The parameters were adjusted empirically by running the elaboration script on Test 1 data until the optimal configuration was found (i.e., local minima). Considering the direct PnP and the matrix concatenation method, the two approaches are comparable, if the error mean norm is considered. The direct method features a lower error standard deviation norm because of the Savitzky–Golay translation vector filtering. US imaging provides only a limited view of the cortical bone. Therefore, a wide range of trabecular bone pathologies, especially osteoporosis, cannot be diagnosed through US [49]. It is necessary to specify that the US technician plays a crucial role in the scan execution: moving the probe slowly with a constant velocity and avoiding rapid changes in orientation drastically changes the reconstruction result [50]. In this work, the phantom was static in the camera frame, but in a real-world application, the patient could move, shifting pixel positions, thus making the scan uninterpretable. The static model used in the experiments did not fully simulate the dynamic changes in clinical environments. Future studies may take into account dynamic changes as these could affect image quality and the following interpretation. Robotic manipulators that have the potential to improve scan quality and repeatability by accommodating patients' unique anatomical differences could replace the operator in the future [51]. Finally, it should be considered that the error measure defined to study the accuracy and precision of the proposed system is relative to the IR system taken as a reference. IR measurements also have inherent uncertainty, with respect to the actual position value, that could affect the obtained results. Compared with IR optical systems, although less accurate, the proposed system is not subject to self-occlusion since the ArUco marker is uniquely identified. Compared with electromagnetic and acoustic systems, the markers in the proposed system are passive and do not emit waves of any kind, and the latency problem is not present. Compared with mechanical systems, the proposed approach is lightweight since the dodecahedron fill percentage can be adjusted and does not require the use of bulky arms or mechanical components [22].

5. Conclusions

Both elaboration methods have demonstrated comparability, both among themselves and with respect to the reference, giving validity to the proposed capture system. Without the need to perform the dual recording (both with infrared and RGB system) carried out in this work for validation purposes, it could be possible to watch in real time the scanned region. Implementing deep learning algorithms, which have the potential to entirely replace tracking systems, including the one proposed in this study, could help to align captured US frames, thus reducing the error [52]. In the future, a refinement of the algorithm could further improve the accuracy and precision of the system, making it more reliable and useful. Given the results produced by the proposed approach, its effectiveness will be further investigated by involving real patients in a clinical study.

Author Contributions: Conceptualization, L.D.S.; Formal analysis, A.C.; Investigation, L.D.S. and C.A.; Methodology, L.D.S. and A.C.; Project administration, U.G.L.; Resources, C.A. and E.F.; Software, L.D.S.; Supervision, E.S. and U.G.L.; Validation, A.C. and E.S.; Visualization, C.A.; Writing—original draft, L.D.S.; Writing—review and editing, L.D.S. and C.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sets generated during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Combes, D.; Lancigu, R.; de Cepoy, P.; Caporilli-Razza, F.; Hubert, L.; Rony, L.; Aubé, C. Imaging of Shoulder Arthroplasties and Their Complications: A Pictorial Review. *Insights Imaging* **2019**, *10*, 90. [\[CrossRef\]](#) [\[PubMed\]](#)
- Maffulli, N.; Longo, U.G.; Berton, A.; Loppini, M.; Denaro, V. Biological Factors in the Pathogenesis of Rotator Cuff Tears. *Sports Med. Arthrosc. Rev.* **2011**, *19*, 194–201. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kolk, A.; Henseler, J.F.; de Witte, P.B.; van Zwet, E.W.; van der Zwaal, P.; Visser, C.P.J.; Nagels, J.; Nelissen, R.G.H.H.; de Groot, J.H. The Effect of a Rotator Cuff Tear and Its Size on Three-Dimensional Shoulder Motion. *Clin. Biomech.* **2017**, *45*, 43–51. [\[CrossRef\]](#) [\[PubMed\]](#)
- Neer, C.S.; Craig, E.V.; Fukuda, H. Cuff-Tear Arthropathy. *J. Bone Jt. Surg. Am.* **1983**, *65*, 1232–1244. [\[CrossRef\]](#) [\[PubMed\]](#)
- Longo, U.G.; Facchinetti, G.; Marchetti, A.; Candela, V.; Risi Ambrogioni, L.; Faldetta, A.; De Marinis, M.G.; Denaro, V. Sleep Disturbance and Rotator Cuff Tears: A Systematic Review. *Medicina* **2019**, *55*, 453. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pope, T.; Bloem, J.L.; Morrison, W.B.; Wilson, D.J.; White, L. *Musculoskeletal Imaging*; Elsevier Health Sciences: Amsterdam, The Netherlands, 2020; ISBN 978-3-030-57376-8.
- Greenspan, A.; Beltran, J. *Orthopaedic Imaging: A Practical Approach*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2020; ISBN 1975136497.
- Valentin, J. 2. How High Are the Doses? *Ann. ICRP* **2000**, *30*, 19–24. [\[CrossRef\]](#)
- Martin, C.J. Radiation Shielding for Diagnostic Radiology. *Radiat. Prot. Dosim.* **2015**, *165*, 376–381. [\[CrossRef\]](#)
- Sammet, S. Magnetic Resonance Safety. *Abdom. Radiol.* **2016**, *41*, 444–451. [\[CrossRef\]](#)
- Aldrich, J.E. Basic Physics of Ultrasound Imaging. *Crit. Care Med.* **2007**, *35*, S131–S137. [\[CrossRef\]](#)
- ter Haar, G. Ultrasound Bioeffects and Safety. *Proc. Inst. Mech. Eng. Part H* **2009**, *224*, 363–373. [\[CrossRef\]](#)
- Martin, K. Introduction to B-Mode Imaging. In *Diagnostic Ultrasound*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2019; pp. 1–5. ISBN 9781138893603.
- Gee, A.; Prager, R.; Treece, G.; Berman, L. Engineering a Freehand 3D Ultrasound System. *Pattern Recognit. Lett.* **2003**, *24*, 757–777. [\[CrossRef\]](#)
- Mills, P.H.; Fuchs, H. 3D Ultrasound Display Using Optical Tracking. In Proceedings of the First Conference on Visualization in Biomedical Computing, Atlanta, GA, USA, 22–25 May 1990; pp. 490–497.
- Wu, H.; Lin, Q.; Yang, R.; Zhou, Y.; Zheng, L.; Huang, Y.; Wang, Z.; Lao, Y.; Huang, J. An Accurate Recognition of Infrared Retro-Reflective Markers in Surgical Navigation. *J. Med. Syst.* **2019**, *43*, 153. [\[CrossRef\]](#)
- Longo, U.G.; De Salvatore, S.; Carnevale, A.; Tecce, S.M.; Bandini, B.; Lalli, A.; Schena, E.; Denaro, V. Optical Motion Capture Systems for 3D Kinematic Analysis in Patients with Shoulder Disorders. *Int. J. Environ. Res. Public Health* **2022**, *19*, 12033. [\[CrossRef\]](#)

18. Mozaffari, M.H.; Lee, W.-S. Freehand 3-D Ultrasound Imaging: A Systematic Review. *Ultrasound Med. Biol.* **2017**, *43*, 2099–2124. [CrossRef]
19. Huang, Q.; Zeng, Z. A Review on Real-Time 3D Ultrasound Imaging Technology. *Biomed. Res. Int.* **2017**, *2017*, 6027029. [CrossRef]
20. Huang, Q.; Xie, B.; Ye, P.; Chen, Z. Correspondence-3-D Ultrasonic Strain Imaging Based on a Linear Scanning System. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2015**, *62*, 392–400. [CrossRef]
21. Huang, Q.-H.; Yang, Z.; Hu, W.; Jin, L.-W.; Wei, G.; Li, X. Linear Tracking for 3-D Medical Ultrasound Imaging. *IEEE Trans. Cybern.* **2013**, *43*, 1747–1754. [CrossRef]
22. Peng, C.; Cai, Q.; Chen, M.; Jiang, X. Recent Advances in Tracking Devices for Biomedical Ultrasound Imaging Applications. *Micromachines* **2022**, *13*, 1855. [CrossRef]
23. Léger, É.; Gueziri, H.E.; Collins, D.L.; Popa, T.; Kersten-Oertel, M. Evaluation of Low-Cost Hardware Alternatives for 3D Freehand Ultrasound Reconstruction in Image-Guided Neurosurgery. In Proceedings of the Simplifying Medical Ultrasound, Strasbourg, France, 27 September 2021; Noble, J.A., Aylward, S., Grimwood, A., Min, Z., Lee, S.-L., Hu, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 106–115.
24. Kalaitzakis, M.; Cain, B.; Carroll, S.; Ambrosi, A.; Whitehead, C.; Vitzilaios, N. Fiducial Markers for Pose Estimation. *J. Intell. Robot. Syst.* **2021**, *101*, 71. [CrossRef]
25. Wu, P.-C.; Wang, R.; Kin, K.; Twigg, C.; Han, S.; Yang, M.-H.; Chien, S.-Y. DodecaPen: Accurate 6DoF Tracking of a Passive Stylus. In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, Quebec, QC, Canada, 22–25 October 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 365–374.
26. Eade, E. Lie Groups for 2d and 3d Transformations. 2013; Volume 117. Available online: <http://ethaneade.com/lie.pdf> (accessed on 18 May 2024).
27. Siciliano, B.; Khatib, O.; Kröger, T. *Springer Handbook of Robotics*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 200.
28. Prager, R.W.; Rohling, R.N.; Gee, A.H.; Berman, L. Rapid Calibration for 3-D Freehand Ultrasound. *Ultrasound Med. Biol.* **1998**, *24*, 855–869. [CrossRef]
29. Manickam, K.; Machireddy, R.R.; Seshadri, S. Characterization of Biomechanical Properties of Agar Based Tissue Mimicking Phantoms for Ultrasound Stiffness Imaging Techniques. *J. Mech. Behav. Biomed. Mater.* **2014**, *35*, 132–143. [CrossRef]
30. Merchel Piovesan Pereira, B.; Tagkopoulos, I. Benzalkonium Chlorides: Uses, Regulatory Status, and Microbial Resistance. *Appl. Environ. Microbiol.* **2019**, *85*, e00377–19. [CrossRef]
31. Souza, R.M.; Santos, T.Q.; Oliveira, D.P.; Alvarenga, A.V.; Costa-Felix, R.P.B. Standard Operating Procedure to Prepare Agar Phantoms. *Proc. J. Phys. Conf. Ser.* **2016**, *733*, 12044.
32. de Assis, M.K.M.; Souza, R.M.; Costa-Félix, R.P.B.; Alvarenga, A.V. Assessment of Ultrasonic Properties of an Agarose Phantom at the Frequency Range 2.25 MHz to 10 MHz. *J. Phys. Conf. Ser.* **2021**, *1826*, 012005. [CrossRef]
33. Madsen, E.L.; Sathoff, H.J.; Zagzebski, J.A. Ultrasonic Shear Wave Properties of Soft Tissues and Tissue-like Materials. *J. Acoust. Soc. Am.* **1983**, *74*, 1346–1355. [CrossRef]
34. *IEEE Std 1588–2008 (Revision of IEEE Std 1588–2002)*; IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems. IEEE: Piscataway, NJ, USA, 2008; pp. 1–269.
35. Scheiterer, R.L.; Na, C.; Obradovic, D.; Steindl, G. Synchronization Performance of the Precision Time Protocol in Industrial Automation Networks. *IEEE Trans. Instrum. Meas.* **2009**, *58*, 1849–1857. [CrossRef]
36. Marchand, E.; Uchiyama, H.; Spindler, F. Pose Estimation for Augmented Reality: A Hands-On Survey. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 2633–2651. [CrossRef]
37. Terzakis George and Lourakis, M. A Consistently Fast and Globally Optimal Solution to the Perspective-n-Point Problem. In Proceedings of the Computer Vision—ECCV 2020, Online, 23–28 August 2020; Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 478–494.
38. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
39. Eade, E. Gauss-Newton/Levenberg-Marquardt Optimization. *Tech. Rep.* **2013**. Available online: <https://www.ethaneade.org/optimization.pdf> (accessed on 18 May 2024).
40. Madsen, K.; Nielsen, H.B.; Tingleff, O. *Methods for Non-Linear Least Squares Problems*; Technical University of Denmark: Copenhagen, Denmark, 2004.
41. Gorry, P.A. General Least-Squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method. *Anal. Chem.* **1990**, *62*, 570–573. [CrossRef]
42. Kabsch, W. A Solution for the Best Rotation to Relate Two Sets of Vectors. *Acta Crystallogr. Sect. A* **1976**, *32*, 922–923. [CrossRef]
43. Umeyama, S. Least-Squares Estimation of Transformation Parameters between Two Point Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 376–380. [CrossRef]
44. Shoemake, K. Animating Rotation with Quaternion Curves. *SIGGRAPH Comput. Graph.* **1985**, *19*, 245–254. [CrossRef]
45. Scholten, H.J.; Weijers, G.; de Wild, M.; Korsten, H.H.M.; de Korte, C.L.; Bouwman, R.A. Differences in Ultrasound Elevational Beam Width (Slice Thickness) between Popular Handheld Devices. *WFUMB Ultrasound Open* **2023**, *1*, 100009. [CrossRef]
46. Goldstein, A. Slice Thickness Measurements. *J. Ultrasound Med.* **1988**, *7*, 487–498. [CrossRef] [PubMed]
47. Del Grosso, V.A.; Mader, C.W. Speed of Sound in Pure Water. *J. Acoust. Soc. Am.* **1972**, *52*, 1442–1446. [CrossRef]

48. Zell, K.; Sperl, J.I.; Vogel, M.W.; Niessner, R.; Haisch, C. Acoustical Properties of Selected Tissue Phantom Materials for Ultrasound Imaging. *Phys. Med. Biol.* **2007**, *52*, N475–N484. [[CrossRef](#)] [[PubMed](#)]
49. Lane, J.M.; Russell, L.; Khan, S.N. Osteoporosis. *Clin. Orthop. Relat. Res.* **2000**, *372*, 139–150. [[CrossRef](#)]
50. Huang, Q.; Lu, M.; Zheng, Y.; Chi, Z. Speckle Suppression and Contrast Enhancement in Reconstruction of Freehand 3D Ultrasound Images Using an Adaptive Distance-Weighted Method. *Appl. Acoust.* **2009**, *70*, 21–30. [[CrossRef](#)]
51. Huang, Q.; Gao, B.; Wang, M. Robot-Assisted Autonomous Ultrasound Imaging for Carotid Artery. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–9. [[CrossRef](#)]
52. Prevost, R.; Salehi, M.; Jagoda, S.; Kumar, N.; Sprung, J.; Ladikos, A.; Bauer, R.; Zettinig, O.; Wein, W. 3D Freehand Ultrasound without External Tracking Using Deep Learning. *Med. Image Anal.* **2018**, *48*, 187–202. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.