

Article

Optimising Camera–ChArUco Geometry for Motion Compensation in Standing Equine CT: A CT-Motivated Benchtop Study

Cosimo Aliani ¹, Cosimo Lorenzetto Bologna ², Piergiorgio Francia ^{1,*} and Leonardo Bocchi ¹

¹ Department of Information Engineering, University of Florence, Via di Santa Marta 3, 50139 Florence, Italy; cosimo.aliانى@unifi.it (C.A.); leonardo.bocchi@unifi.it (L.B.)

² Epica Imaginalis, Via Rodolfo Morandi 13/15, 50019 Florence, Italy; c.lorenzetto@imaginalis.it

* Correspondence: piergiorgiofrancia@libero.it

Abstract

Standing equine computed tomography (CT) acquisitions are susceptible to residual postural sway, which can introduce view-inconsistent motion and degrade image quality. External optical tracking based on ChArUco fiducials is a promising, low-cost strategy to enable projection-wise motion compensation, yet quantitative guidance on how camera–marker geometry affects pose-estimation performance remains limited. This CT-motivated benchtop study characterizes how the relative camera–ChArUco configuration influences both the accuracy (bias with respect to ground truth) and the precision (repeatability) of pose estimates obtained from RGB images using OpenCV ChArUco detection and reprojection-error minimization to estimate the rigid camera-to-board transformation. Controlled experiments systematically varied acquisition protocol (continuous repeated estimates at fixed pose versus cyclic repositioning), viewing angle over a wide angular range at two working distances, and camera-to-board distance over multiple depth settings. Ground truth for angular configurations was defined by a stepper-motor rotation stage, while distance ground truth was obtained by ruler measurements. Performance was summarized via mean absolute error and standard deviation across repeated measurements, complemented by variance-based statistical testing with multiple-comparison correction. Cyclic repositioning did not yield evidence of increased variability relative to continuous acquisitions, supporting view-by-view sampling. Viewing angle induced a consistent accuracy–precision trade-off for rotations: frontal views minimized mean error but exhibited higher variability, whereas oblique views reduced jitter at the expense of increased bias. Increasing working distance reduced repeatability, most prominently for depth-related components. Overall, these findings provide pre-clinical guidance for selecting camera/marker placement (moderately oblique viewpoints, limited working distance, sufficient image footprint) before in-scanner and in-vivo validation for standing equine CT motion compensation.



Academic Editors: Sebastian Iwaszenko and Karolina Nurzynska

Received: 15 January 2026

Revised: 8 February 2026

Accepted: 12 February 2026

Published: 18 February 2026

Copyright: © 2026 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/).

Keywords: computed tomography; equine; ChArUco; camera–marker geometry; motion compensation

1. Introduction

Computer Vision (CV) focuses on the automatic extraction, analysis, and interpretation of meaningful information from single images or sequences of images, and relies on theoretical frameworks and algorithms for autonomous visual understanding [1]. A long-standing challenge concerns three-dimensional (3D) reconstruction from two-dimensional

(2D) images, which arises in multiple domains—robotics [2,3], medicine [4,5], and virtual reality [6,7]—and stems from the loss of depth information when projecting a 3D scene onto the image plane ($f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$). Approaches such as photogrammetry, stereo vision (SV), and structure-from-motion (SfM) address this task by detecting distinctive features (keypoints), establishing correspondences across views, and combining image evidence with camera parameters to infer 3D coordinates for scene points.

In many of these pipelines, accurate and stable camera pose estimation is a key determinant of reconstruction quality. Physical fiducials with known geometry are widely used to stabilize pose: ArUco markers [8,9] offer robust, uniquely identifiable square patterns but limited corner refinement; chessboards enable subpixel-accurate corner localization yet require full visibility and are less tolerant to occlusions [10,11]. ChArUco boards combine these strengths by embedding ArUco markers within a chessboard grid, achieving reliable detection under partial occlusion together with high-precision corner refinement via interpolated chessboard corners [11]. Thanks to the known layout, a single image suffices to estimate the rigid transformation between camera and board, typically decomposed into rotation (yaw, pitch, roll) and position components (T_x, T_y, T_z).

A setting where precise, projection-wise pose tracking is particularly valuable is computed tomography (CT) of standing horses. Dedicated systems for standing acquisitions of the head and distal limbs allow imaging of sedated equine patients in a weight-bearing position, avoiding the risks of general anaesthesia while preserving diagnostic performance [12–16]. However, residual postural sway during gantry rotation induces view-inconsistent motion that manifests as blur and streak artifacts and may reduce confidence in the assessment of fine osseous detail [14,15]. To mitigate such motion, previous work has mainly relied on mechanical restraints and acquisition protocols, while software-based motion compensation is still uncommon in routine equine imaging and lacks quantitative design guidelines.

As shown in Figure 1, a practical strategy is to rigidly mount a lightweight ChArUco board near the region of interest and observe it with an auxiliary RGB camera during acquisition. Each frame provides a board pose that, via a fixed extrinsic calibration to the scanner, can be mapped to object motion for projection-wise motion compensation. In this context, the camera–marker geometry—specifically, viewing angle and working distance—directly affects both pose accuracy (bias with respect to ground truth) and pose precision (variance/noise), and thus the effectiveness of motion correction.

By contrast, in human CT, several prototype systems already exploit external motion sensing for projection-wise motion correction. Optical trackers, inertial sensors, and RGB-D cameras have been used to monitor patient motion during helical or cone-beam scans and to warp the projections accordingly, recovering image quality that approaches that of motion-free acquisitions [17–19]. Low-cost Intel RealSense devices in particular have been characterized for clinical motion tracking and CT perfusion motion correction, showing millimetric pose accuracy in realistic scanner environments [20–22]. These developments motivate the use of compact RGB or RGB-D modules as external motion sensors in the standing equine CT workflow.

Recent work also highlights complementary directions to handle motion and tracking uncertainty. On the motion-compensation side, projection-driven and learning-based strategies have been proposed in CT to estimate and compensate motion directly from image/projection data, reducing reliance on external sensors in some settings [23–25]. On the tracking side, the fiducial-marker literature has progressed beyond standard OpenCV pipelines: deep-learning-based detection has been introduced to improve robustness under adverse illumination and imaging conditions [26], multi-marker fusion with spatial/temporal outlier rejection has been shown to improve both accuracy and precision

especially at large tilt angles [27], and recent benchmarking studies report that camera type and imaging geometry can materially affect fiducial-based pose stability [28]. These advances reinforce that both the tracking algorithm and the camera–marker geometry contribute to performance.



Figure 1. Representative photograph illustrating the intended standing equine cone beam computed tomography (CBCT) head-imaging workflow. A ChArUco board is rigidly attached to the horse’s head and observed by an RGB camera for external pose tracking in the target clinical context. This image is provided for illustrative purposes only; no data from the depicted animal/procedure were acquired or used in the present study.

Although ChArUco boards are already adopted across diverse applications (e.g., RGB-D fusion [29], omnidirectional camera calibration [30], and robust detection in challenging conditions [31]), the quantitative relationship between pose-estimation accuracy/precision and the relative camera–board configuration remains underexplored. In particular, there is limited evidence on how viewing angle and camera-to-plane distance shape the error with respect to a known ground truth and the noise of estimated rotations and positions, despite the practical impact on reconstruction quality in motion-aware workflows and on the design of standing equine CT setups.

The present study addresses this gap through a controlled experimental analysis that systematically varies camera-to-plane angles and camera-to-plane positions, acquiring repeated measurements to characterize how pose accuracy and precision change with viewpoint. The resulting trends are interpreted as CT-motivated, pre-clinical guidance to inform camera/marker placement prior to future in-scanner and in-vivo validation, thereby limiting iterative optimization on live animals.

2. Materials and Methods

This work adopts a controlled benchtop experimental design to quantify how camera–ChArUco geometry affects pose-estimation accuracy and precision. Repeated pose estimates were collected under identical conditions while systematically varying, separately, the camera-to-plane viewing angle and the camera-to-board working distance (and associated translations). Ground truth for angular configurations was defined by the commanded stepper-motor rotation stage (1° increments), whereas ground truth for distance changes was obtained via ruler measurements. For each geometric condition, pose was estimated from RGB images using OpenCV ChArUco detection and perspective- n -point

(PnP), and performance was summarized using mean absolute error with respect to ground truth (accuracy) and the standard deviation across repeated measurements (precision). An overview of the full implementation workflow and the three experimental branches is provided in Figure 2.

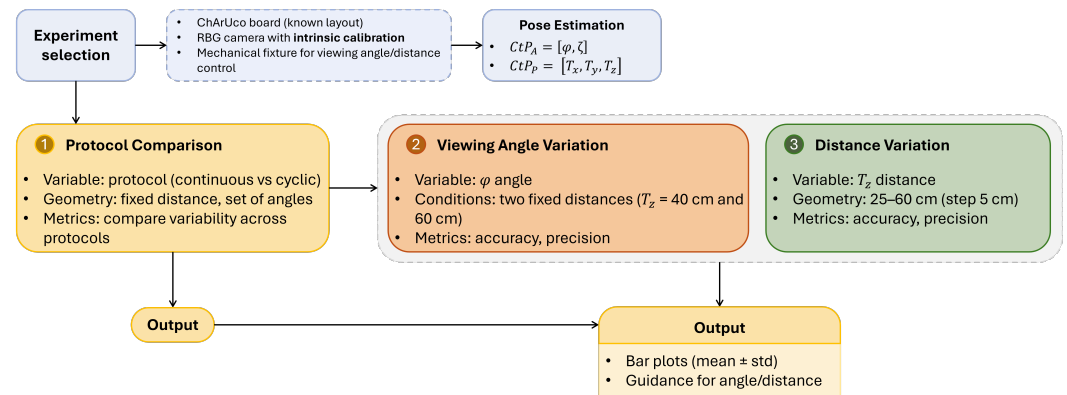


Figure 2. Flowchart of the experimental workflow and study design. A common pose-estimation pipeline (setup, acquisition, and ChArUco-based pose estimation) is applied across three experiments: protocol comparison (continuous vs. cyclic), viewing-angle variation at fixed working distances, and working-distance variation. Outputs include accuracy/precision metrics (mean \pm standard deviation) used to derive practical guidance on camera–board geometry.

2.1. ChArUco Board

The ChArUco board used in this study, shown in Figure 3a, consists of a 7×7 chessboard grid with 24 unique ArUco markers placed inside the white squares. Each chessboard square measures 15 mm per side, while each ArUco marker measures 10 mm per side. The relative pose between the camera and the board was parameterized by five scalar quantities:

- **Camera-to-plane angles (CtP_A):** $CtP_A = [\varphi, \zeta]$, where φ is the angle between the camera optical axis (\hat{z}_c) and the board normal vector \hat{n} , and ζ is the in-plane rotation around \hat{n} .
- **Camera-to-plane positions (CtP_P):** $CtP_P = [T_x, T_y, T_z]$, describing the position of the board origin expressed in the camera reference system.

To improve clarity, parameters defining both CtP_A and CtP_P are depicted in Figure 3b.

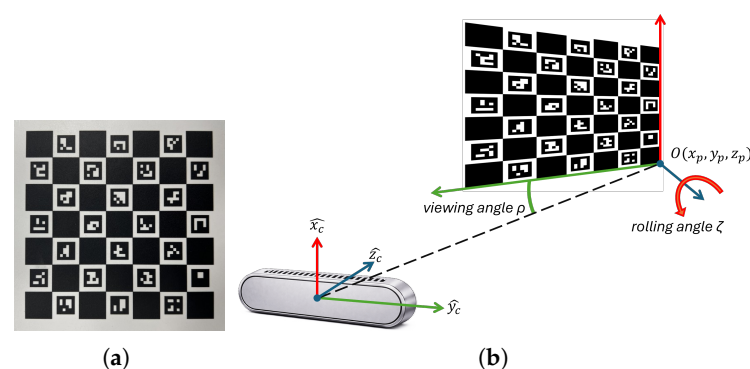


Figure 3. (a) ChArUco board used during the experiments. (b) Camera and ChArUco reference systems and definition of CtP_A and CtP_P . ChArUco's origin $O(x_p, y_p, z_p)$ is expressed in the camera reference system.

Both the detection of the ChArUco board and the camera pose estimation were performed using the open-source OpenCV library [32], in conjunction with the Qt Creator development environment (version 5.12.6) and the C++ programming language.

2.2. RGB Camera

The RGB camera employed in this study is the RGB module of the Intel RealSense D435 3D camera (Intel Corporation, Santa Clara, CA, USA). Its key specifications are summarized in Table 1.

Table 1. Main characteristics of the RGB module of the Intel D435 3D camera.

Frame resolution	1920 × 1080
Frame rate	30 fps
Sensor technology	Rolling shutter
Sensor FOV (H × V)	69° × 42°
Sensor resolution	2 MP

2.3. Mechanical Systems for Camera or ChArUco Handling

Once rigidly attached to the horse, the ChArUco target is expected to undergo predominantly translational excursion T_z (towards/away from the observing camera) and changes in viewing angle φ , rather than in-plane rotation ζ about the board normal. This reflects the residual postural sway of the sedated standing subject around its resting stance.

Thus, the experimental setups were designed to independently vary either the viewing-angle φ or the position along the z-axis T_z between the ChArUco board and the camera, while monitoring the remaining pose parameters to verify that off-axis positions and ζ remained negligible:

- **Viewing angle φ setup:** as shown in Figure 4a, this setup consists of a steel mechanical wheel that rotates precisely about its axis in 1° increments and is driven by a stepper motor. Two orthogonal grooves intersecting at the wheel's centre enable precise placement of the ChArUco board such that its coordinate origin coincides with the wheel's axis of rotation. This alignment is critical to minimize undesired translational components and rotation about ζ during wheel motion.
- **T_z position setup:** as shown in Figure 4b, this setup uses a 1-m-long steel rail with a central horizontal groove. The ChArUco board is fixed to an L-shaped bracket and mounted orthogonally at one end of the base. The camera is attached to a separate L-shaped bracket and positioned to face the ChArUco board. By sliding the bracket along the groove, the distance between the camera and the ChArUco board is adjusted along the z-axis while minimizing off-axis motion.

For the viewing-angle experiments, the wheel was controlled via its dedicated software, which allowed setting angular positions in 1° increments with respect to a reference configuration defined as $\varphi = 0^\circ$. Accordingly, the ground-truth rotation φ_{GT} was taken as the wheel angle set in the control software. To reduce the effect of mechanical backlash, target angles were approached consistently using the same rotation direction.

For the distance experiments, the ground-truth camera-to-plane distance $T_{z,GT}$ was measured using a ruler. Measurements were taken between the camera reference point and the ChArUco board plane (as defined in Figure 3b) at each tested position, and recorded with the ruler resolution.

Pose accuracy was quantified as the absolute error with respect to ground truth, e.g., $e_\varphi = |\varphi - \varphi_{GT}|$ and $e_{T_z} = |T_z - T_{z,GT}|$, while precision was quantified as the variability (e.g., standard deviation) across repeated measurements under identical conditions.

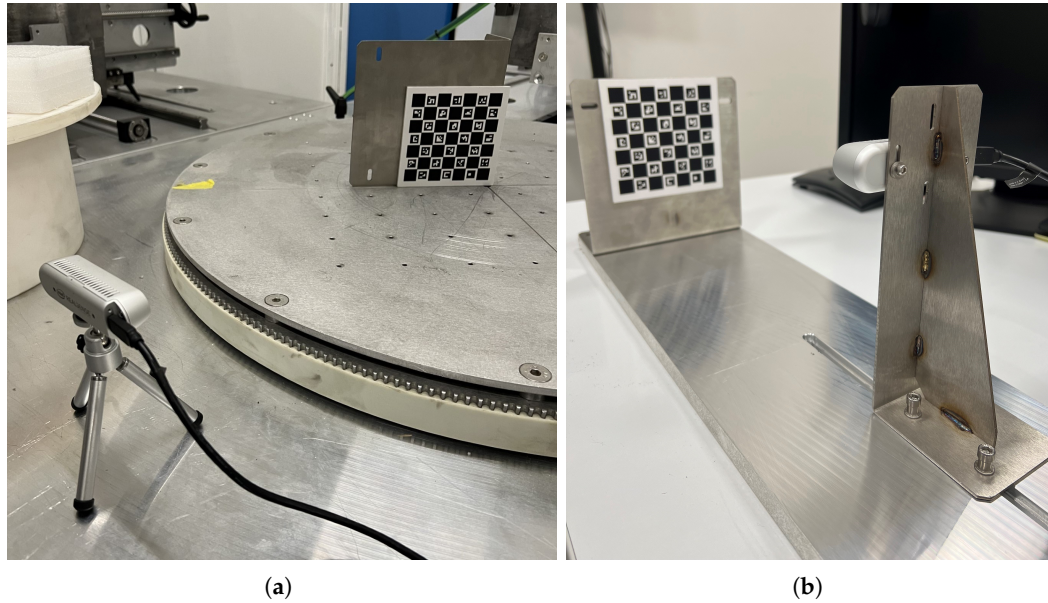


Figure 4. (a) Setup used during the test for the evaluation of the viewing angle φ parameter. The camera was placed on a tripod to ensure stability. (b) Setup used during the test for the evaluation of the T_z parameter.

2.4. Measurements Acquisition Procedure

Multiple tests were conducted to evaluate the accuracy of camera pose estimation relative to the ChArUco marker. In each test, the camera was initially positioned relative to the ChArUco board, and the pose parameters $CtP_A = [\varphi, \zeta]$ and $CtP_P = [T_x, T_y, T_z]$ were acquired.

Before each run, an initial reference pose was recorded at the start position: $CtP_{A_0} = [\varphi_0, \zeta_0]$ and $CtP_{T_0} = [T_{x_0}, T_{y_0}, T_{z_0}]$. These reference values were subtracted from each subsequent measurement to describe rotation/position increments rather than absolute pose. Repeatability and drift were quantified across repeated runs using statistics of these relative pose trajectories. Absolute accuracy was also evaluated as the error with respect to the known ground truth defined for each test.

2.4.1. Continuous and Cyclic Acquisitions

The first experiment aimed to assess the noise introduced by the camera and by the mechanical movement system. For this test, the camera was fixed at a distance of $T_z = 40$ cm from the ChArUco board, and the viewing angle φ was varied using the mechanical wheel. The tested angles were $[45^\circ, 25^\circ, 5^\circ, 0^\circ, -5^\circ, -25^\circ, -45^\circ]$, using two acquisition methods:

- **Continuous acquisitions:** To evaluate noise introduced solely by the camera, each pose was estimated 10 times at each angle before rotating the wheel to the next angle.
- **Cyclic acquisitions:** To assess the influence of the mechanical system, only one pose was estimated at each angle. After reaching -45° , the wheel was rotated back to 45° in a single motion, and the process was repeated 10 times in total.

After each acquisition, the standard deviation of all five pose parameters was computed to compare the consistency of the two acquisition approaches.

This comparison is motivated by projection-wise motion tracking in CT: during an acquisition, the tracker delivers one pose estimate per projection while the gantry (or the object) moves between successive views. “Continuous” acquisitions isolate the intrinsic measurement noise of the camera/estimator at a fixed viewpoint (no repositioning). “Cyclic” acquisitions mimic repeated view-by-view sampling across angles and re-positioning, there-

fore capturing additional variability introduced by the mechanical motion (e.g., backlash, settling, and repeatability of the rotation wheel), which is analogous to view-to-view perturbations during a CT sweep.

2.4.2. Viewing Angle Variation

The second experiment evaluated how viewing angle estimation accuracy varies with the viewing angle. Angles ranging from 60° to -60° were tested: $[60^\circ, 50^\circ, 45^\circ, 40^\circ, 30^\circ, 20^\circ, 10^\circ, 0^\circ, -10^\circ, -20^\circ, -30^\circ, -40^\circ, -45^\circ, -50^\circ, -60^\circ]$. The ChArUco board was aligned with the rotation axis of the mechanical wheel, and the camera was placed at fixed distances of $T_z = 40$ cm and $T_z = 60$ cm. At each angle and distance, 20 pose estimations were acquired to ensure statistical robustness.

2.4.3. Z-Axis Position Variation

The third experiment focused on how the position along the z-axis affects the estimation of the T_z parameter. Camera poses were estimated at distances ranging from $T_z = 25$ cm to $T_z = 60$ cm, in 5 cm increments ($\Delta T_z = 5$ cm). As in the previous test, each position was measured 20 times to support statistical analysis.

2.5. Pose Estimation Model and Performance Metrics

To provide theoretical support for the implementation, the pose-estimation model and the adopted performance metrics are formalized below.

Let $\mathbf{x}_i = [X_i, Y_i, Z_i]^\top$ be a 3D point on the ChArUco board expressed in the board frame. Under a pinhole camera model, its image projection $\mathbf{u}_i = [u_i, v_i, 1]^\top$ satisfies

$$s \mathbf{u}_i = \mathbf{K} (\mathbf{R} \mathbf{x}_i + \mathbf{t}), \quad (1)$$

where s is a projective scale factor (proportional to the point depth in the camera frame), \mathbf{K} is the intrinsic calibration matrix, and (\mathbf{R}, \mathbf{t}) is the rigid transform from board to camera coordinates, with \mathbf{R} mapping vectors from the board frame to the camera frame.

Given detected 2D corners $\tilde{\mathbf{u}}_i$, the pose is obtained by solving a perspective- n -point (PnP) problem, i.e.,

$$(\hat{\mathbf{R}}, \hat{\mathbf{t}}) = \arg \min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^N \|\tilde{\mathbf{u}}_i - \pi(\mathbf{K}(\mathbf{R} \mathbf{x}_i + \mathbf{t}))\|^2, \quad (2)$$

where $\pi(\cdot)$ denotes perspective division, and \mathbf{x}_i denotes the corresponding 3D points in the board frame.

The estimated translation is directly given by $\hat{\mathbf{t}} = [\hat{T}_x, \hat{T}_y, \hat{T}_z]^\top$ (board origin expressed in the camera frame). Let $\hat{\mathbf{n}}_c = \hat{\mathbf{R}} \mathbf{n}_p$ be the board normal expressed in the camera frame, with $\mathbf{n}_p = [0, 0, 1]^\top$. The viewing angle is then

$$\hat{\phi} = \arccos\left(\frac{\hat{\mathbf{n}}_c^\top \hat{\mathbf{z}}_c}{\|\hat{\mathbf{n}}_c\| \|\hat{\mathbf{z}}_c\|}\right), \quad (3)$$

where $\hat{\mathbf{z}}_c = [0, 0, 1]^\top$ is the camera optical axis. The in-plane rotation $\hat{\zeta}$ is obtained from the estimated rotation matrix $\hat{\mathbf{R}}$ as the rotation about the board normal $\hat{\mathbf{n}}_c$, i.e., the in-plane (roll) component of the board orientation.

3. Results

In the following subsections, the results for each of the three conducted tests will be presented.

In all bar plots, bar height represents the mean absolute error with respect to ground truth, while error bars indicate the standard deviation across repeated measurements under the same condition.

3.1. Continuous vs. Cyclic Acquisition

The results of this test are reported in Figures 5 and 6. Figure 5a,b show the mean absolute errors of the CtP_A parameters for the continuous and cyclic protocols, respectively. Figure 6a,b show the corresponding mean absolute errors for the CtP_P parameters.

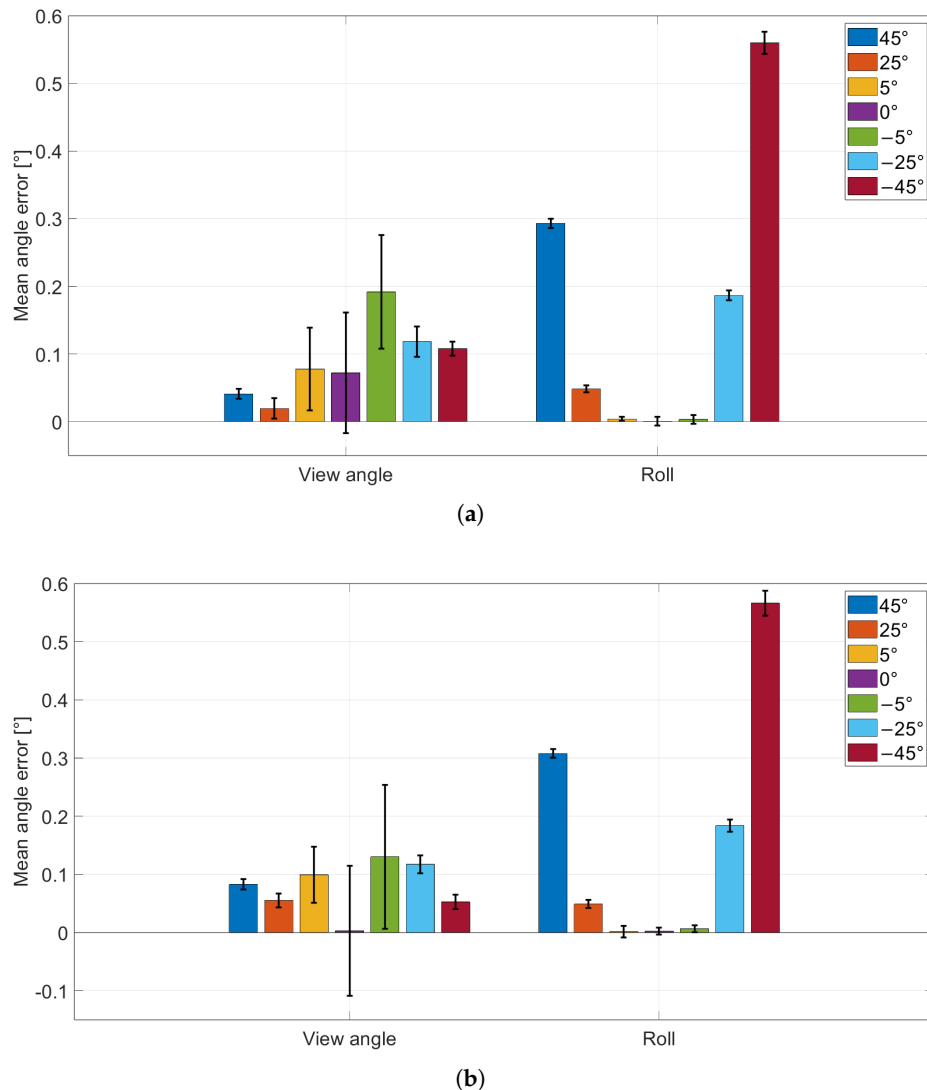


Figure 5. CtP_A parameters absolute error (bar height; viewing angle φ and roll ζ) with standard deviation (error bars) for each wheel position. (a) Continuous acquisition; (b) cyclic acquisition.

The mechanical wheel provided controlled viewpoint changes in 1° increments through stepper-motor actuation. The comparison between continuous and cyclic acquisitions was used to verify that repeated repositioning of the wheel does not introduce additional variability beyond the intrinsic pose-estimation noise.

To assess whether the two acquisition protocols differ in a statistically meaningful way, Levene's test for equality of variances was applied to the error distributions obtained with the continuous and cyclic protocols. For the formal statistical comparison, three representative wheel positions were selected (-45° , 0° , $+45^\circ$), corresponding to the two most oblique views and the frontal configuration. This choice captures the extremes of the investigated angular range (i.e., a worst-case condition for perspective effects) as well

as the frontal case, while limiting the number of multiple comparisons. Intermediate angles were analyzed descriptively in the plots and are expected to lie within the bounds defined by these representative configurations. For each selected wheel position, a one-to-one comparison was performed for each parameter (i.e., φ , ζ , T_x , T_y , T_z), resulting in 15 comparisons in total. Levene's test was selected because the mean absolute error can be influenced by systematic effects (e.g., small setup misalignments or mechanical offsets), whereas the variance (and thus the standard deviation) more directly reflects repeatability and intrinsic measurement noise, which is the key aspect when evaluating whether cyclic repositioning introduces additional variability. Across the 15 comparisons, only T_y at -45° showed a nominally significant result ($p = 0.03783$). A Bonferroni correction for multiple comparisons was applied, setting the corrected significance level to $\alpha_B = 0.05/15 \approx 0.0033$; under this threshold, no comparison was significant. Therefore, no evidence was found that cyclic repositioning systematically changes measurement variability compared to the continuous protocol, and subsequent experiments are reported using the continuous acquisition method only.

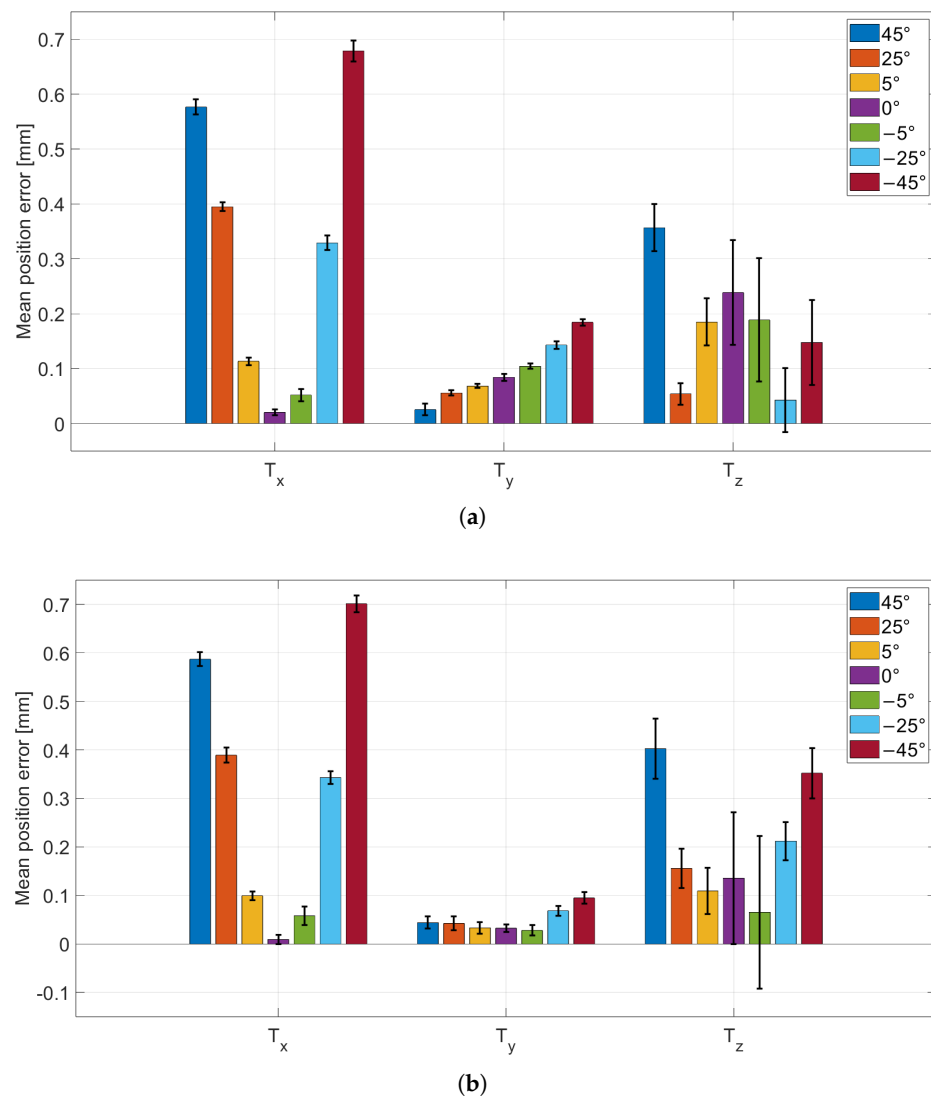


Figure 6. CtP_p parameters absolute error (bar height; T_x , T_y , T_z) with standard deviation (error bars) for each wheel position. (a) Continuous acquisition; (b) cyclic acquisition.

3.2. Influence of Viewing Angle on Pose Estimation

The results of this test are reported in Figures 7 and 8. Figure 7a,b show the mean absolute errors of the CtP_A parameters at fixed working distances $T_z = 40$ cm and $T_z = 60$ cm, respectively. Figure 8a,b show the corresponding mean absolute errors for the CtP_P parameters at the same distances.

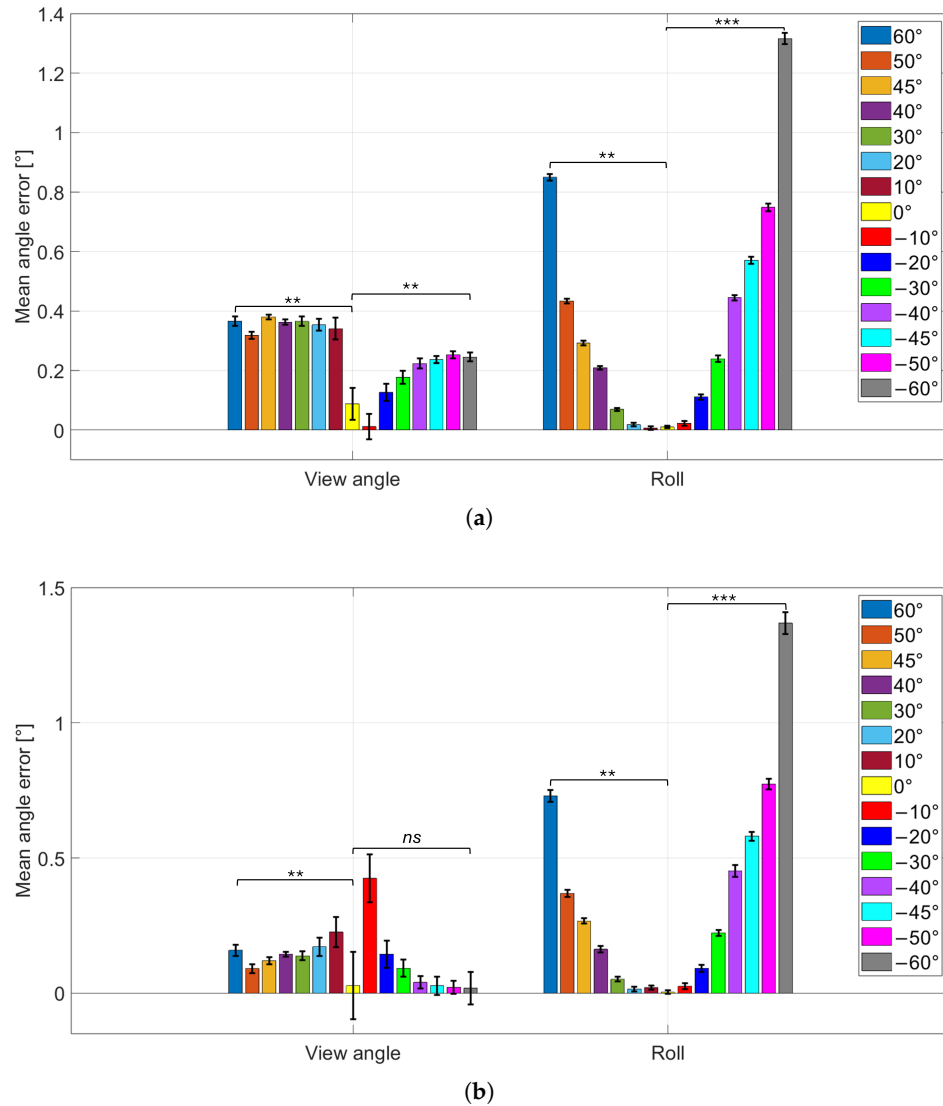


Figure 7. CtP_A parameters absolute error (viewing angle φ and roll ζ) for each wheel position. (a): T_z parameter fixed at 40 cm; (b): T_z parameter fixed at 60 cm. Statistical significance is indicated using symbols: ns for not statistically significant, ** for $p < 0.001$, and *** for $p < 0.0001$.

Additionally, to assess whether the observed changes in variability with viewing angle are statistically meaningful, Levene's test for equality of variances was applied to the error distributions. The statistical analysis was performed separately for the two working distances ($T_z = 40$ cm and $T_z = 60$ cm), since they represent distinct experimental conditions and were evaluated independently. Within each working distance, the variance at the frontal configuration (0°) was compared against the variance at the two extreme viewpoints (-60° and $+60^\circ$) for each parameter (φ , ζ , T_x , T_y , T_z). This resulted in 10 variance comparisons per working distance. Again, Levene's test was selected because systematic effects can influence the mean absolute error, whereas the variance directly reflects repeatability and intrinsic measurement noise. For each distance-specific family of tests, a Bonferroni cor-

rection was applied to control for multiple comparisons, setting the corrected significance level to $\alpha_B = 0.05/10 = 0.005$.

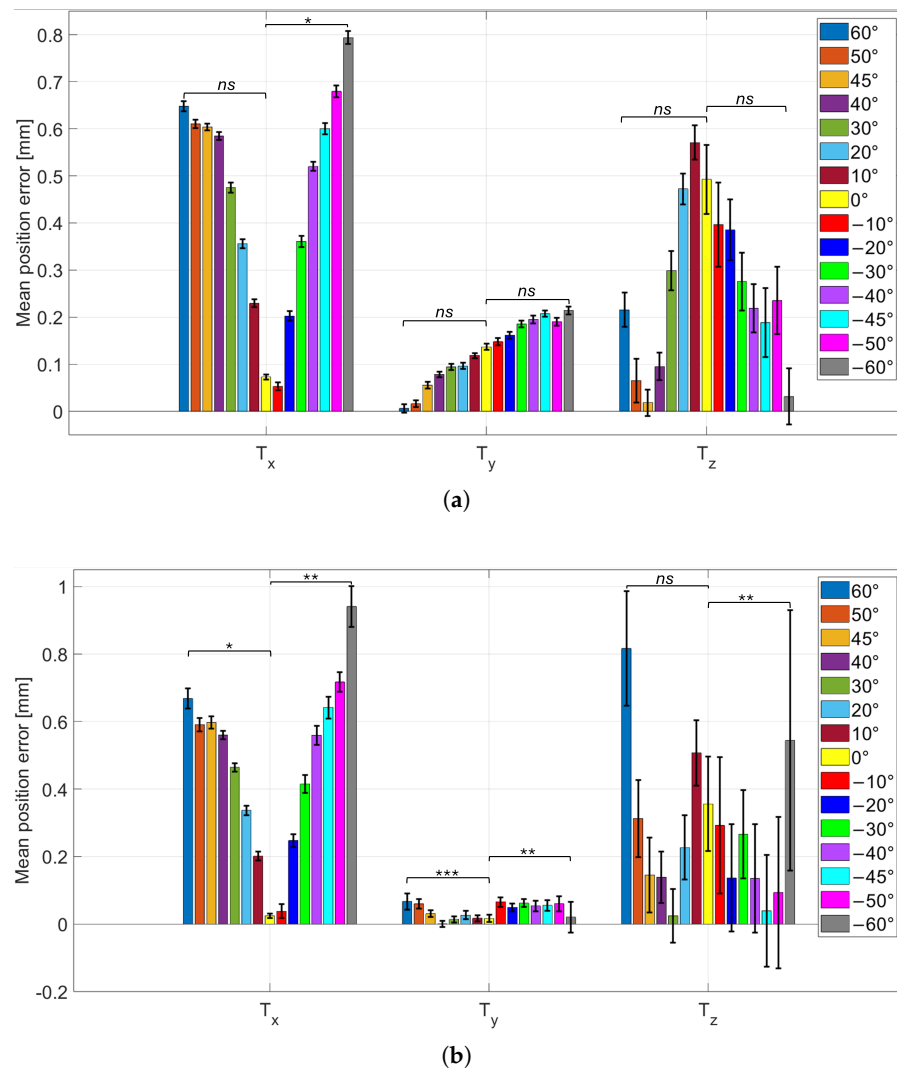


Figure 8. $CtPp$ parameters absolute error (T_x , T_y , T_z) for each wheel position. (a): T_z parameter fixed at 40 cm; (b): T_z parameter fixed at 60 cm. Statistical significance is indicated using symbols: ns for not statistically significant, * for $p < 0.005$, ** for $p < 0.001$, and *** for $p < 0.0001$.

Across viewing angles, the mean absolute error on φ is minimal in the frontal configuration ($\varphi_{GT} = 0^\circ$) and increases progressively at more oblique views. Conversely, the standard deviation (error bars) is highest at $\varphi_{GT} = 0^\circ$ and decreases as the viewing angle departs from frontal, indicating a trade-off between accuracy (lower mean absolute error) and precision (lower variability). A similar behavior is observed for the in-plane rotation ζ . Comparing working distances, moving from $T_z = 40$ cm to $T_z = 60$ cm yields lower mean absolute errors (lower bars) but larger standard deviations (higher error bars) for both φ and ζ . Notably, the angle-dependent trend in variability is preserved at both distances, with the largest standard deviations near the frontal configuration and decreasing variability at increasing obliquity. To complement these qualitative trends, Levene's test consistently indicated significant variance differences for the rotation parameters: the variability at 0° was significantly different from that at both -60° and $+60^\circ$ for φ and ζ . This statistically supports the observed decrease in standard deviation as the viewing angle departs from the frontal configuration. In contrast, translation parameters showed no significant variance differences in almost all comparisons, with the only exception of T_x

in the 0° vs. -60° comparison. Overall, this suggests that at $T_z = 40$ cm the viewing-angle dependence primarily affects the repeatability of rotation estimation, while translation variability remains largely stable across angles. For $T_z = 60$ cm, a different pattern emerged. Variance differences remained significant for most rotation comparisons, with the exception of the φ comparison between 0° and -60° , which was not significant. Despite this isolated non-significant case, the overall trend still supports reduced variability at oblique angles. For translation parameters, variance differences were significant in most comparisons, except for the T_z comparison between 0° and $+60^\circ$. This broader occurrence of significant variance changes in CtP_P at the larger working distance may indicate that increased measurement noise at 60 cm makes translation repeatability more sensitive to viewpoint, compared to the 40 cm condition.

3.3. Influence of Distance on Pose Estimation

The results of this test are presented in Figure 9. In Figure 9a, the average errors related to CtP_A parameters are presented, while in Figure 9b, the average errors related to CtP_P parameters are presented.

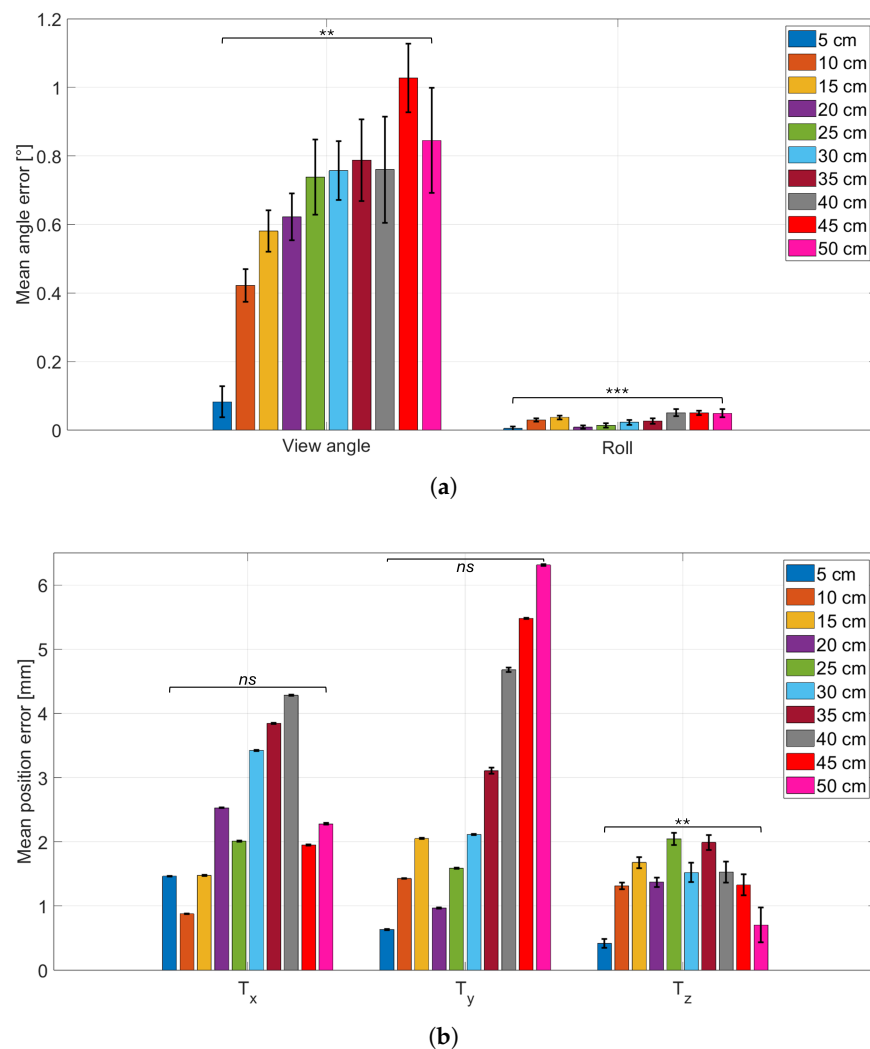


Figure 9. (a): CtP_A parameters absolute error (bar height; viewing angle φ and roll ζ) with standard deviation (error bars) for each camera-to-ChArUco distance. (b): CtP_P parameters mean absolute error (bar height; T_x , T_y , T_z) with standard deviation (error bars) across repeated measurements, for each camera-to-ChArUco working distance. Statistical significance is indicated using symbols: *ns* for not statistically significant, ** for $p < 0.0025$, and *** for $p < 0.00025$.

To assess whether the observed changes in variability with distance are statistically meaningful, Levene's test for equality of variances was applied to the error distributions. The variance at the lower working distance ($T_z = 5$ cm) was compared against the variance at the higher working distance ($T_z = 50$ cm) for each parameter (φ , ζ , T_x , T_y , T_z). Since this resulted in five variance comparisons, a Bonferroni correction was applied, setting the corrected significance level to $\alpha_B = 0.05/5 = 0.01$.

Levene's test indicates significantly different variances for the rotation parameters (φ and ζ) when comparing $T_z = 5$ cm to $T_z = 50$ cm, consistent with a distance-dependent change in repeatability. This behavior may reflect increased sensitivity to pixel-level noise at larger working distances, where the board occupies fewer pixels and corner localization becomes less stable; alternatively, it may also be influenced by uncertainty in the distance-setting procedure. For the translation parameters, T_x and T_y do not show significant variance differences across the two distances, suggesting that their repeatability remains largely stable in the tested range. By contrast, T_z exhibits a significant variance difference, supporting the observation that depth estimation becomes less repeatable at larger camera-to-board distances. Overall, these results indicate that increasing working distance primarily reduces the repeatability of T_z estimation, even when the average agreement with ground truth remains comparable.

4. Discussion

The present CT-motivated benchtop study shows that camera–ChArUco geometry has distinct and sometimes opposing effects on pose-estimation accuracy (bias with respect to ground truth) and precision (repeatability). In particular, cyclic repositioning does not measurably increase variability compared to repeated estimates at fixed pose, viewing angle induces an accuracy–precision trade-off in rotation estimation, and increasing working distance reduces repeatability, especially for depth-related components.

The first objective was to verify whether repeated repositioning introduces additional variability beyond intrinsic pose-estimation noise. This hypothesis was not supported: variance-based testing did not provide evidence of increased variability under cyclic acquisitions after multiple-comparison correction. This result is relevant for CT-motivated, projection-wise tracking, where one estimate per view is delivered while the relative geometry changes between successive projections. In standing equine CT, motion mitigation has been mainly pursued via mechanical restraint and protocol choices rather than software-based projection-wise correction [12–16]. At the same time, recent CT research has proposed image-driven and learning-based motion compensation approaches that estimate motion directly from projections or reconstructed volumes [23–25]. The present finding complements these directions by supporting the feasibility of stable, view-by-view external sampling as an additional information source in motion-aware workflows.

The second objective was to characterize how viewpoint affects pose-estimation performance. Rotation estimation exhibits a consistent trade-off: frontal views yield lower mean absolute error but higher variability, whereas more oblique views reduce variability at the cost of larger mean error. Recent fiducial-marker studies confirm that challenging viewpoints and adverse imaging conditions are key failure modes: deep-learning-based detectors have been introduced to improve robustness under difficult illumination [26], and multi-marker strategies with outlier rejection have been shown to reduce both mean errors and variability, especially at large tilt angles [27]. Benchmarking across different cameras and setups further reports that pose stability can vary materially with sensing hardware and geometry [28]. In the standing equine CT context, where residual postural sway can lead to view-inconsistent motion and image artefacts, precision (low frame-to-frame jitter) can be a critical limiting factor for motion compensation pipelines. The present

results highlight that a geometry optimized for minimal bias (strictly frontal view) may not be optimal when repeatability is prioritized.

The third objective was to evaluate the effect of working distance. Increasing the camera-to-board distance is associated with reduced repeatability, most markedly affecting depth-related components. A plausible mechanism is the reduced board footprint in the image at larger distances, which makes corner localization more sensitive to pixel-level noise. This observation is consistent with prior characterizations of low-cost RGB/RGB-D sensing for motion tracking [20–22] and aligns with recent fiducial-based camera-positioning analyses emphasizing the role of marker image coverage and viewing geometry in pose quality [33]. For CT-motivated deployments, this implies that distance should be selected to keep the target sufficiently resolved in the image while satisfying practical constraints of the scanner environment.

Although the present work is not a clinical validation study, the observed trends can inform pre-clinical configuration of an external tracking setup before in-scanner and in-vivo evaluation. In standing equine CT, where external tracking may support projection-wise motion compensation, the following implications arise:

- A strictly frontal configuration may minimize bias but can maximize jitter; therefore, a moderately oblique nominal viewpoint may be preferable when repeatability is prioritized.
- Excessive working distances should be avoided to preserve repeatable depth estimation.
- The board should occupy a sufficiently large image region to stabilize corner localization.
- Rigid mounting and minimization of occlusions/illumination changes remain practical requirements in the clinical workflow.

Several limitations should be acknowledged. Experiments were conducted under controlled laboratory conditions using a single board layout and a single RGB sensor; real deployments may involve occlusions, specularities, illumination changes, and multi-axis motion. Only limited degrees of freedom were explored (single-axis rotation and axial distance changes). In addition, repeated CT acquisitions solely for marker-placement optimization would increase radiation exposure without clinical benefit; therefore, the present study focuses on benchtop characterization to reduce the need for iterative animal experiments. Finally, the analysis focuses on pose-estimation metrics rather than end-to-end image-quality outcomes.

Future work should extend the study to multi-axis trajectories and validate the proposed setup choices in realistic scanner environments. When ethically justified (e.g., clinically indicated scans), in-vivo validation should quantify the impact of the chosen geometry on motion-compensated reconstructions and image quality (e.g., sharpness and streak reduction). Additional extensions include evaluating different ChArUco scales/layouts, multi-camera fusion, and the integration of RGB-D information where available.

5. Conclusions

This CT-motivated benchtop study provides pre-clinical guidance on how camera-ChArUco geometry impacts pose-estimation accuracy and repeatability in external tracking setups relevant to standing equine CT. The results highlight that viewpoint and working distance can affect bias and jitter differently, motivating geometry choices based on the relative importance of accuracy versus repeatability in motion-compensation pipelines. Future work will validate these setup choices in realistic scanner environments and, when ethically justified, in vivo using clinically indicated scans, including end-to-end evaluation on motion-compensated reconstructions.

Author Contributions: Conceptualization, C.A. and L.B.; methodology, C.A. and L.B.; software, C.A.; validation, C.L.B., P.F. and L.B.; investigation, C.A.; resources, C.L.B. and L.B.; data curation, C.A.; writing—original draft preparation, C.A.; writing—review and editing, C.A., C.L.B., P.F. and L.B.; visualization, C.A.; supervision, C.L.B., P.F. and L.B.; project administration, L.B. 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. This study did not involve experiments on live animals, did not collect animal-derived data, and did not use clinical/in vivo acquisitions. All measurements were performed in a controlled laboratory setup using a ChArUco board, an RGB camera (Intel RealSense D435), and mechanical fixtures; ground truth was obtained from the mechanical positioning system and ruler measurements.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to legal reasons.

Conflicts of Interest: Author Cosimo Lorenzetto Bologna was employed by the company Epica Imaginalis. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Klette, R. *Concise Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 233.
2. Fazakas, T.; Fekete, R.T. 3D reconstruction system for autonomous robot navigation. In *Proceedings of the 2010 11th International Symposium on Computational Intelligence and Informatics (CINTI), Budapest, Hungary, 18–20 November 2010*; IEEE: New York, NY, USA, 2010; pp. 267–270.
3. Tang, Y.; Chen, M.; Wang, C.; Luo, L.; Li, J.; Lian, G.; Zou, X. Recognition and localization methods for vision-based fruit picking robots: A review. *Front. Plant Sci.* **2020**, *11*, 510. [[CrossRef](#)] [[PubMed](#)]
4. Huang, Y.; Qiu, Z.; Song, Z. 3D reconstruction and visualization from 2D CT images. In *Proceedings of the 2011 International Symposium on IT in Medicine and Education (ITME 2011), Guangzhou, China, 9–11 December 2011*; IEEE: New York, NY, USA, 2011; Volume 2, pp. 153–157.
5. Rengier, F.; Mehndiratta, A.; Von Tengg-Kobligk, H.; Zechmann, C.M.; Unterhinninghofen, R.; Kauczor, H.U.; Giesel, F.L. 3D printing based on imaging data: Review of medical applications. *Int. J. Comput. Assist. Radiol. Surg.* **2010**, *5*, 335–341. [[CrossRef](#)]
6. Ebner, T.; Feldmann, I.; Renault, S.; Schreer, O.; Eisert, P. Multi-view reconstruction of dynamic real-world objects and their integration in augmented and virtual reality applications. *J. Soc. Inf. Disp.* **2017**, *25*, 151–157. [[CrossRef](#)]
7. Tang, F.; Wu, Y.; Hou, X.; Ling, H. 3D mapping and 6D pose computation for real time augmented reality on cylindrical objects. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *30*, 2887–2899. [[CrossRef](#)]
8. Romero-Ramirez, F.J.; Muñoz-Salinas, R.; Medina-Carnicer, R. Speeded up detection of squared fiducial markers. *Image Vis. Comput.* **2018**, *76*, 38–47. [[CrossRef](#)]
9. Garrido-Jurado, S.; Muñoz-Salinas, R.; Madrid-Cuevas, F.J.; Medina-Carnicer, R. Generation of fiducial marker dictionaries using mixed integer linear programming. *Pattern Recognit.* **2016**, *51*, 481–491. [[CrossRef](#)]
10. De la Escalera, A.; Armingol, J.M. Automatic chessboard detection for intrinsic and extrinsic camera parameter calibration. *Sensors* **2010**, *10*, 2027–2044. [[CrossRef](#)]
11. ChArUco Marker Detection. Available online: https://docs.opencv.org/3.4/df/d4a/tutorial_charuco_detection.html (accessed on 8 February 2026).
12. Dakin, S.; Lam, R.; Rees, E.; Mumby, C.; West, C.; Weller, R. Technical set-up and radiation exposure for standing computed tomography of the equine head. *Equine Vet. Educ.* **2014**, *26*, 208–215. [[CrossRef](#)]
13. Davies, T.; Skelly, C.; Puggioni, A.; D’Helft, C.; Connolly, S.; Hoey, S. Standing CT of the equine head: Reducing radiation dose maintains image quality. *Vet. Radiol. Ultrasound* **2020**, *61*, 137–146. [[CrossRef](#)] [[PubMed](#)]
14. Mathee, N.; Robert, M.; Higgerty, S.M.; Fosgate, G.T.; Rogers, A.L.; d’Ablon, X.; Carstens, A. Computed tomographic evaluation of the distal limb in the standing sedated horse: Technique, imaging diagnoses, feasibility, and artifacts. *Vet. Radiol. Ultrasound* **2023**, *64*, 243–252. [[CrossRef](#)]
15. Pauwels, F.E.; Van der Vekens, E.; Christan, Y.; Koch, C.; Schweizer, D. Feasibility, indications, and radiographically confirmed diagnoses of standing extremity cone beam computed tomography in the horse. *Vet. Surg.* **2021**, *50*, 365–374. [[CrossRef](#)] [[PubMed](#)]

16. Gaida, J.L.; Steinberg, T.; Stieger-Vanegas, S.M.; Merle, R.; Lischer, C.J. Equine Standing Multidetector Computed Tomography of the Distal Thoracic Limb and Tarsus Has a Lower Cumulative Radiation Dose than Digital Radiography. *Vet. Radiol. Ultrasound* **2025**, *66*, e70049. [[CrossRef](#)] [[PubMed](#)]
17. Kim, J.; Nuyts, J.; Kyme, A.; Kuncic, Z.; Fulton, R. A rigid motion correction method for helical computed tomography (CT). *Phys. Med. Biol.* **2015**, *60*, 2047. [[CrossRef](#)]
18. Brombal, L.; Arana Peña, L.M.; Arfelli, F.; Longo, R.; Brun, F.; Contillo, A.; Di Lillo, F.; Tromba, G.; Di Trapani, V.; Donato, S.; et al. Motion artifacts assessment and correction using optical tracking in synchrotron radiation breast CT. *Med Phys.* **2021**, *48*, 5343–5355. [[CrossRef](#)]
19. Maier, J.; Nitschke, M.; Choi, J.H.; Gold, G.; Fahrig, R.; Eskofier, B.M.; Maier, A. Rigid and non-rigid motion compensation in weight-bearing CBCT of the knee using simulated inertial measurements. *IEEE Trans. Biomed. Eng.* **2021**, *69*, 1608–1619. [[CrossRef](#)]
20. Siena, F.L.; Byrom, B.; Watts, P.; Breedon, P. Utilising the intel realsense camera for measuring health outcomes in clinical research. *J. Med Syst.* **2018**, *42*, 53. [[CrossRef](#)]
21. House, R.; Lasso, A.; Harish, V.; Baum, Z.; Fichtinger, G. Evaluation of the Intel RealSense SR300 camera for image-guided interventions and application in vertebral level localization. In *Proceedings of the Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling, SPIE, Orlando, FL, USA, 14–16 February 2017*; SPIE: Washington, DC, USA, 2017; Volume 10135, pp. 771–777.
22. Moghari, M.D.; Noonan, P.; Henry, D.; Fulton, R.R.; Young, N.; Moore, K.; Evanns, A.; Kyme, A. Characterisation of the Intel RealSense D415 Stereo Depth Camera for Motion-Corrected CT Perfusion Imaging. *arXiv* **2024**, arXiv:2403.16490.
23. Thies, M.; Wagner, F.; Maul, N.; Yu, H.; Goldmann, M.; Schneider, L.S.; Gu, M.; Mei, S.; Folle, L.; Preuhs, A.; et al. A gradient-based approach to fast and accurate head motion compensation in cone-beam CT. *IEEE Trans. Med Imaging* **2024**, *44*, 1098–1109. [[CrossRef](#)]
24. Lu, A.; Huang, H.; Hu, Y.; Zbijewski, W.; Unberath, M.; Siewerdsen, J.H.; Weiss, C.R.; Sisniega, A. Vessel-targeted compensation of deformable motion in interventional cone-beam CT. *Med Image Anal.* **2024**, *97*, 103254. [[CrossRef](#)]
25. Maier, J.; Sawall, S.; Arheit, M.; Paysan, P.; Kachelrieß, M. Deep learning-based cone-beam CT motion compensation with single-view temporal resolution. *Med. Phys.* **2025**, *52*, e17911. [[CrossRef](#)]
26. Berral-Soler, R.; Muñoz-Salinas, R.; Medina-Carnicer, R.; Marín-Jiménez, M.J. DeepArUco++: Improved detection of square fiducial markers in challenging lighting conditions. *Image Vis. Comput.* **2024**, *152*, 105313. [[CrossRef](#)]
27. Kim, D.; Bong, J.H.; Jeong, S. Enhancing pose estimation using multiple graphical markers with spatial and temporal outlier detection. *Appl. Sci.* **2024**, *14*, 10225. [[CrossRef](#)]
28. Krumstroh, F.; Eisenmann, A.; Franssen, J.; Rindermann, M.; Pfeiffer, T. Comparing Fiducial Marker Tracking Across Cameras in Virtual and Physical Environments. In *Proceedings of the GI Workshop 2025 on Virtual and Augmented Reality*; Gesellschaft für Informatik: Bonn, Germany, 2025; pp. 10–18420.
29. de Medeiros Esper, I.; Smolkin, O.; Manko, M.; Popov, A.; From, P.J.; Mason, A. Evaluation of RGB-D Multi-Camera Pose Estimation for 3D Reconstruction. *Appl. Sci.* **2022**, *12*, 4134. [[CrossRef](#)]
30. An, G.H.; Lee, S.; Seo, M.W.; Yun, K.; Cheong, W.S.; Kang, S.J. Charuco Board-Based Omnidirectional Camera Calibration Method. *Electronics* **2018**, *7*, 421. [[CrossRef](#)]
31. Hu, D.; DeTone, D.; Malisiewicz, T. Deep charuco: Dark charuco marker pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019*; IEEE: New York, NY, USA, 2019; pp. 8436–8444.
32. OpenCV: Open Source Computer Vision Library. Available online: <https://opencv.org> (accessed on 8 February 2026).
33. García-Ruiz, P.; Romero-Ramirez, F.J.; Muñoz-Salinas, R.; Marín-Jiménez, M.J.; Medina-Carnicer, R. Sparse Indoor Camera Positioning with Fiducial Markers. *Appl. Sci.* **2025**, *15*, 1855. [[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.