Erciyes University Journal of Institue Of Science and Technology

Research Article
10.65520/erciyesfen.1968365

Imprint:
Volume: 42(3)
Year: 2026
Page: 249-271

Dursun Alp Kızılöz[a*]
Murat Peker[b]

[a] Res. Asst., Bursa Technical University, alp.kiziloz@btu.edu.tr
[b] Assoc. Prof., Bursa Technical University, murat.peker@btu.edu.tr

[*] Corresponding Author

Received: 6/10/2026
Accepted: 8/3/2026

Citation:
Dursun Alp Kızılöz, Murat Peker (2026). Design of Low-Cost High Accuracy Motion Capture System. *Erciyes University Journal of Institue Of Science and Technology*, 42(3), 249-271. https://doi.org/10.65520/erciyesfen.*1968365*

Screened by iThenticate for Authors & Researchers

Except where otherwise noted, content in this article is licensed under a Creative Commons 4.0 International license. Icons by Font Awesome.

Design of Low-Cost High Accuracy Motion Capture System

Abstract

In this study, a low-cost passive marker-based optical motion capture system was developed and experimentally validated as an alternative to high-cost commercial systems that are widely regarded as the industry standard. The proposed system consists of multiple IP cameras operating in IR mode, passive reflective markers, and computer vision-based image processing algorithms. Camera calibration was completed with reprojection errors ranging from 0.02 to 0.04 pixels. The markers were detected with sub-pixel accuracy, and their three-dimensional positions were calculated using triangulation. To evaluate the system, four passive markers with 10 cm spacing were mounted on a platform. Experiments were conducted while the platform remained stationary and while it was moved by hand along the X, Y and Z axes. The three-dimensional distances between the markers were analysed using RMSE, MAE, signed mean error (bias), and standard deviation metrics. Under static conditions, the frame-to-frame standard deviation of the calculated distances ranged from 0.086 to 0.204 mm, while the RMSE values ranged from 0.222 to 3.360 mm. The dominant source of error was found to be fixed offsets caused by the manual placement of the markers and the tolerance of the reference measurement device, rather than measurement noise. The findings demonstrate that, despite its low cost, the proposed system provides sub-millimeter repeatability and millimeter-level accuracy, making it a feasible alternative to industrial motion capture systems for static and low-speed applications.

Keywords: Optical Motion Capture, Image Processing, Computer Vision, Ground Truth Generation

Düşük Maliyetli Yüksek Doğruluklu Hareket Yakalama Sisteminin Tasarımı

Öz

Bu çalışmada, endüstri standardı kabul edilen yüksek maliyetli ticari sistemlere alternatif olarak, düşük maliyetli pasif işaretleyicili bir optik hareket yakalama sistemi geliştirilmiş ve deneysel olarak doğrulanmıştır Geliştirilen sistem, IR modunda çalışan çoklu IP kameralar, pasif yansıtıcı işaretleyiciler ve bilgisayarlı görü tabanlı görüntü işleme algoritmalarından oluşmaktadır. Kamera kalibrasyonu 0.02–0.04 piksel yeniden projeksiyon hatalarıyla tamamlanmış; işaretleyiciler alt piksel hassasiyetle tespit edilerek üç boyutlu konumları üçgenleme yöntemiyle hesaplanmıştır. Sistemi değerlendirmek amacıyla aralarında 10 cm mesafe bulunan dört pasif işaretleyici, bir platform üzerine yerleştirilmiştir. Platformun X, Y ve Z eksenleri boyunca sabit ve el ile hareket ettirilmesiyle deneyler gerçekleştirilmiştir. İşaretleyiciler arasındaki üç boyutlu mesafeler; RMSE, MAE, işaretli ortalama hata (bias) ve standart sapma metrikleri kullanılarak analiz edilmiştir. Statik durumda hesaplanan mesafelerin kareler arası standart sapması 0.086–0.204 mm aralığında kalmış; RMSE değerleri ise 0.222–3.360 mm arasında bulunmuştur. Hatanın baskın kaynağının ölçüm gürültüsü değil, işaretleyicilerin elle yerleştirilmesinden ve

referans cihazın toleransından kaynaklanan sabit ofsetler olduğu belirlenmiştir. Elde edilen bulgular, önerilen sistemin düşük maliyetine rağmen milimetre altı tekrarlanabilirlik ile milimetre seviyesinde doğruluk sunduğunu ve statik ile düşük hızlı uygulamalar için endüstriyel hareket yakalama sistemlerine uygulanabilir bir alternatif oluşturduğunu göstermektedir.

Anahtar kelimeler: Optik Hareket Yakalama, Görüntü İşleme, Bilgisayarlı Görü, Yer Gerçeği Üretimi

1.INTRODUCTION

Motion Capture is a technology used to track, digitize, and record the movements of objects or living beings in three-dimensional (3D) space [1]. Initially used as a photogrammetric analysis tool in biomechanical research in the 1970s and 1980s, these systems, which focus on transferring the movements and performance of living things and objects to the digital environment, have become a fundamental technological tool in many different sectors today thanks to their high sensitivity [2].

Motion capture systems are widely used across a broad and diverse range of fields, extending from biomechanics to the entertainment industry. In the entertainment and media sector, they are utilized in film, television, video games, 3D character animation, and virtual/augmented reality (VR/AR) applications [2-7]. In sports sciences, they are used for purposes such as maximizing performance by analysing the biomechanical data of athletes, identifying and preventing injury risks, improving techniques, and support the more efficient design of new sports equipment [8]. In the fields of clinical medicine and healthcare, these systems provide objective data for the diagnosis of neurological and orthopaedic disorders, 3D gait analysis, and the monitoring of rehabilitation processes [9-11]. Furthermore, in industrial and ergonomic applications, they are utilized to conduct risk assessments aimed at protecting the occupational health of workers in sectors such as manufacturing, construction, and logistics, preventing musculoskeletal disorders by analysing working postures, and optimizing workstations [12,13].

This technology plays a critical role, particularly in the field of robotics; it is used in collecting high-precision ground truth data needed to test the accuracy of SLAM (Simultaneous Localization and Mapping) algorithms widely used in robotic systems, as well as in advanced tasks such as developing intelligent autonomous robots and enabling human-robot interaction [14,15].

The techniques employed in motion capture systems are divided into various categories based on the principles of motion detection and the underlying technology. For instance, during the early years of motion capture development, mechanical and magnetic systems were utilized [2]. Mechanical systems are those in which angular changes are measured directly via electromechanical potentiometers located at the joints of exoskeleton-like structures worn directly on the user's body [16,17]. Magnetic systems, on the other hand, estimate position and orientation by measuring the electromagnetic fields between receivers placed on the user's body and a central transmitting antenna [16,17]. Another approach, inertial systems, involves placing wearable inertial measurement units such as accelerometers, gyroscopes, and magnetometers on specific points of the body to continuously measure orientation and acceleration [16,17].

Finally, optical systems, which operate with high precision and are the most widely utilized, are categorized into three subgroups. The first comprises passive marker-based optical motion capture systems, which rely on calculating the position of reflective markers via the triangulation method, using fixed, high-speed cameras positioned around the measurement volume [17]. Offering extremely high sub-millimeter precision, these systems are regarded as the gold standard [18]. Vicon and OptiTrack are examples of such systems [2,19]. The second category consists of active marker optical motion capture systems. In contrast to the purely reflective markers used in passive systems, the markers in these systems are composed of diodes that actively emit infrared light [16,17]. Lastly, markerless optical motion capture systems are those that eliminate the need for physical markers by leveraging image processing, deep learning, computer vision, and artificial intelligence (AI)-based algorithms [20,21].

CC: BY-NC-ND

This study develops an alternative system to passive marker optical motion capture systems. Passive marker-based optical motion capture systems transfer the movements of objects or living beings into digital environments through the use of multiple cameras and specialized markers placed on the subject. To create a passive marker optical motion capture system, a recording area is first defined. High-resolution, high-speed cameras are placed around the recording area. The object or living being, with passive markers positioned on it, is recorded by the cameras as it moves within the defined recording area. The markers are detected on the 2D images obtained from the cameras, and their position in 3D space is calculated using the triangulation method [16]. Table 1 lists the datasets generated for SLAM algorithms. It is observed that mocap systems, particularly those recorded indoors and triggered by hardware, have an accuracy of approximately 1 mm in the static state. The mocap system proposed in this study achieved the same accuracy with software triggering, reaching millimeter accuracy and sub-millimeter repeatability in the static state.

Companies such as OptiTrack or Vicon perform this process using their proprietary software and equipment, providing real-time 3D coordinate positioning and offering sub-millimeter precision [2, 19]. However, the primary disadvantage of these systems is the high cost of the advanced technological equipment utilized. For instance, in our study, a passive marker-based optical motion capture system consisting of 6 Dahua Wizmind IP cameras, capable of recording at 25 FPS with a resolution of 2688x1520 pixels, was built for approximately $1,622.50 [29]. In contrast, the approximate cost of a professional system, configured with 6 cameras capable of recording at 360 FPS with a 2048x1088 pixel resolution and calculating 3D positions with 0.2 mm precision, is stated to be $27,032.50 [19].

Table 1. Comparison of SLAM datasets

Dataset	Year	Environment	Camera	Trigger	Ground Truth
KITTI [22]	2013	outdoors	1 stereo RGB, 2×1392×512 @10Hz; 1 stereo gray, 2×1392×512 @10Hz	sw	OXTS RT 3003 pose @10Hz, acc. <10 cm
Malaga Urban [23]	2014	outdoors	1 stereo RGB, 2×1024×768 @20Hz	sw	GPS @1Hz, low accuracy
UMich NCLT [24]	2015	Indoors /outdoors	6 RGB, 1600×1200 @5Hz	sw	GPS/IMU/laser pose @150Hz, acc≈10 cm
EuRoC MAV [25]	2016	indoors	1 stereo gray, 2×752×480 @20Hz	hw	laser tracker @20Hz, motion capture pose @100Hz, acc≈1 mm
PennCOSYVIO [26]	2017	Indoors /outdoors	4 RGB 1920×1080 @30Hz; 1 stereo gray 2×752×480 @20Hz; 1 fisheye gray 640×480 @30Hz	hw /sw	fiducial marker pose @30Hz, acc≈15 cm
TUM VI [27]	2018	Indoors /outdoors	1 stereo gray, 2×1024×1024 @20Hz	hw	partial motion capture pose @120Hz, marker pose acc≈1 mm
M2DGR [28]	2021	indoors	1 RGB 1280×1024 @15Hz; 1 IR 640×512 @25Hz; 1 event 640×480 @15Hz; 1 VI-sensor RGB 640×480 @15Hz; IMU @200Hz	hw	Vicon Vero 2.2 MoCap @50Hz, accuracy ≈1 mm

Consequently, in this study, a high-precision and low-cost passive marker-based optical motion capture system was developed by utilizing accessible IP cameras instead of these high-cost systems. The development, calibration, and experimental evaluation of the system are described in detail in the following section.

CC: BY-NC-ND

2. MATERIALS AND METHODS

This study develops an alternative system to passive marker-based optical motion capture systems. For this purpose, a study area was first defined, and 6 IP cameras were placed around this area. The cameras were calibrated to ensure high-accuracy measurements. Then, images were taken synchronously from all cameras to first detect the passive markers in 2D and then calculate their 3D position.

This section will first introduce the equipment used in the study. Following an explanation of the camera calibration processes, the software developed for detecting the passive reflective markers and calculating their 3D position will be presented. Finally, the experimental setup established for the study will be described.

2.1. Experimental infrastructure and equipment

In this study, 14 mm diameter reflective M4 markers and marker bases manufactured by OptiTrack were used for position detection [19,29]. An image of the product used in this study is presented in Figure 1. Since these markers have high retroreflective properties under infrared (IR) illumination, they appear with enhanced contrast when the cameras operate in IR mode.

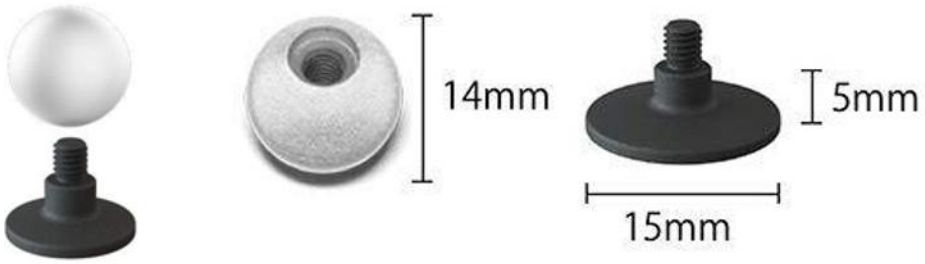


Figure 1. OptiTrack marker and marker base [29]

Six Dahua Wizmind IP cameras were used to detect the passive reflective markers. The cameras, used in IR mode, were triggered by software and recorded images at 2688x1520 pixel resolution with 25 FPS. The image of the camera used is given in Figure 2, and its technical specifications are given in Table 2.

Figure 2. Dahua Wizmind DH-IPC-HFW5442E-ZE IP camera [30]

Table 2. Dahua Wizmind DH-IPC-HFW5442E-ZE IP camera technical data [30]

Image Sensor	1/1.8" CMOS
Maximum Resolution	2688 (H) × 1520 (V)
Focal Length	2.7 mm–12 mm
FPS	2688 × 1520@ (1 fps–25/30 fps)
Network Port	RJ-45 (10/100 Enhancement Base-T)

CC: BY-NC-ND

In this study, a Bosch PLR 50 C laser distance meter was utilized to obtain the reference measurements for the actual positions of the reflective spheres [31]. The technical specifications regarding the measurement capabilities of the laser meter are presented in Table 3, while an image of the device is provided in Figure 3.

Figure 3. Bosch PLR 50 C laser distance meter [31]

Table 3. Bosch PLR 50 C laser distance meter technical data [31]

Measurement Range	0.05 m – 50 m
Measurement Accuracy	± 2.0 mm

For the power supply and data transmission of the cameras, a Dahua S4210-8GT-110 model high-performance 10-port managed Gigabit PoE switch, designed for security and enterprise network requirements, was utilized [32]. A workstation equipped with an Intel i9-14900K 32-core processor and an NVIDIA RTX 4000 graphics card was employed for image recording from cameras in IR mode and performing the detection of the reflective markers.

2.2. Camera calibration

Camera calibration is the process of determining the optical and geometric characteristics of cameras to extract metric information from 2D images and to obtain accurate, reliable measurements from imaging systems [33]. The calibration process consists of two distinct stages as intrinsic and extrinsic calibration. Both stages were performed in Python using the camera calibration module of the OpenCV library.

2.2.1. Intrinsic calibration

Camera lenses inherently cause certain distortions when capturing images, primarily radial and tangential distortions. Radial distortion is greater the further points are from the center of the image. Radial distortions typically manifest as a fisheye effect [33]. These distortions can be modelled as follows:

$$x_{distorted} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (1)$$

$$y_{distorted} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (2)$$

Tangential distortion occurs because the lenses capturing the images are not perfectly parallel to the image plane. This situation can be represented by the following formulas:

$$x_{distorted} = x + [2p_1 xy + p_2(r^2 + 2x^2)] \quad (3)$$

$$y_{distorted} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \quad (4)$$

In these equations, (x, y) represent the undistorted pixel coordinates. Consequently, there are five parameters used to define these distortions:

$$Distortion\ Coefficients = (k_1, k_2, p_1, p_2, k_3) \quad (5)$$

To perform intrinsic calibration, it is also necessary to calculate the camera's focal length (f_x, f_y) and optical center (c_x, c_y). The focal length and optical centers can be utilized to construct a camera matrix, which is used to correct the distortion originating from a specific camera's lenses. The camera matrix is unique to a specific camera; therefore, once calculated, it can be reused for subsequent images captured by the same camera [33]. It is expressed as a 3x3 matrix:

$$Camera\ Intrinsics\ Matrix = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

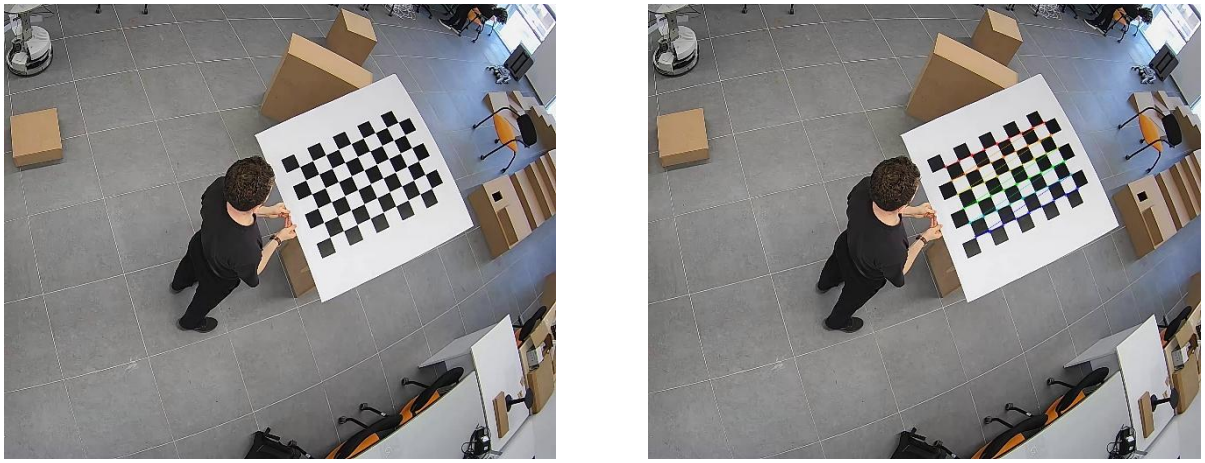
During the calibration process, a chessboard with a square size of 80 mm, containing 6 inner corners per row and 9 inner corners per column, was employed as the calibration pattern. Numerous images of the chessboard must be captured from various angles and elevations for each camera. For instance, in this project, between 105 and 145 images were recorded per camera to be utilized in the intrinsic calibration process. Example images captured for calibration purposes is presented in the Figure 4.

$$A = \begin{bmatrix} 1506.993 & 0 & 1341.067 \\ 0 & 1506.903 & 744.483 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

$$d = [-0.370\ 0.191\ 0.001\ 0\ -0.067\ ] \quad (8)$$

The camera matrix and distortion coefficients obtained as a result of intrinsic calibration are as shown above. Theoretically, the camera focal lengths are expected to be approximately 1. For instance, upon examining the output provided above, it can be observed that the camera focal lengths are indeed approximately 1. Similarly, the optical centers are expected to correspond to the midpoint of the image resolution. Given that the captured images frame a resolution of 2688x1520 pixels, their center coordinates are anticipated to be 1344x760. Examination of the calibration matrix reveals that the calculated midpoint values are quite close to the theoretical values. The fact that both (f_x, f_y) and (c_x, c_y) values are calculated very close to the theoretical values indicates that the calibration performed is robust. In addition, reprojection error was calculated to evaluate the difference between the estimated parameters and the true observed image points. With a reprojection error of 0.03 pixels, sub-pixel accuracy was achieved, providing strong evidence for the reliability of the measurements and constituting an important basis for millimeter-level precision. For all cameras, the reprojection error was found to lie within the range of 0.03–0.04 pixels. Reprojection error provides a reliable estimate of how accurate the obtained calibration parameters are. The closer the reprojection error is to zero, the more accurate the estimated parameters become.

CC: BY-NC-ND



(a) (b)

Figure 4. Sample images from intrinsic calibration process. (a) Regular frame. (b) Corners detected frame

2.2.2. Extrinsic calibration

Extrinsic parameters correspond to the rotation and translation vectors that transform the coordinates of a 3D point into another coordinate system. The camera intrinsic matrix, shown by formula 6, projects 3D points given in the camera coordinate system to 2D pixel coordinates. When the rotation and translation vectors obtained from extrinsic calibration are integrated with the camera's intrinsic matrix, the distortion-free projective transformation formula modelled with the pinhole camera model, expressed by formula 9 below, is obtained. In the below formula s represents the scale factor, p represents the 2D pixel in the image plane, A denotes the camera intrinsic matrix, $[R \mid t]$ represents the rotation matrix and translation vector, namely the extrinsic camera matrix and P_w represents the 3D point in the world coordinate system.

$$s\, p = A\, [R \mid t]\, P_w \tag{9}$$

Prior to initiating the calibration process, a 9 m^2 workspace was designated for conducting the measurements. To perform the extrinsic calibration, a chessboard was placed at the midpoint of the designated 9 m^2 workspace, with the first corner of the chessboard positioned precisely at the center of the area. The purpose of this setup was to define the center of the working space as the origin of the global coordinate system. The origin was determined using a single image captured simultaneously by all six cameras, and the positions of the cameras were calculated relative to this origin. The extrinsic calibration file, containing the rotation and translation matrices that define the camera positions, was subsequently saved. In this stage, a reprojection error of 0.02 pixels was calculated for each camera. Images showing the virtual origin determined as a result of the calibration are presented in the Figure 5. Figure 5(a) illustrates the assignment of the first corner of the chessboard as the origin. Furthermore, Figure 5(b) confirms that the designated origin point is located precisely at the center of the workspace.

(a) (b)

Figure 5. Images depicting the virtual origin. (a) Determination of the origin using the corner of the chessboard. (b) The established origin visualized without the chessboard.

2.2.3. Software of recording frames, detecting markers and triangulation

The data acquisition software developed to record from six IP cameras at a resolution of 2688 × 1520 pixels and a frame rate of 25 FPS was built upon a producer–consumer architecture. In order to achieve synchronized image capture at 25 FPS through software triggering, each camera was triggered every 40 ms by separate parallel threads. To improve synchronization performance, the buffer size was set to 1, ensuring that only the most recent frame available at the moment of the trigger signal was retrieved from each camera. In addition, since performing frame acquisition and disk writing simultaneously could lead to frame drops and latency, the disk writing process was organized asynchronously to begin only after the recording stop command was issued. Consequently, the frames captured at 25 FPS were temporarily stored in RAM during acquisition. As the cameras were operated in IR mode to detect reflective markers, the captured frames were instantly converted into grayscale format and buffered in RAM. Once the recording session was terminated via an interrupt signal, the asynchronous writer module was activated and transferred the accumulated timestamped images from memory to the hard disk in PNG format, together with a structured log file. In this way, the temporal integrity of the dataset was preserved. The overall workflow of the developed recording algorithm is presented in the Figure 6.

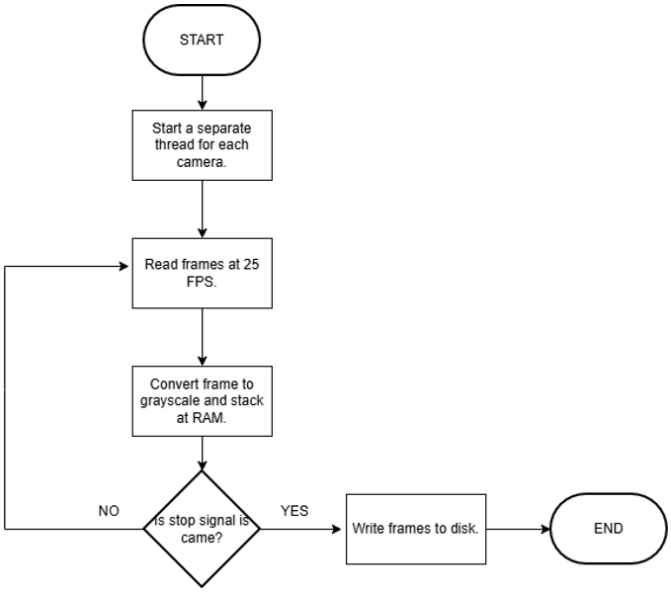


Figure 6. Flowchart of record algorithm

CC: BY-NC-ND

A computer vision script has been developed to detect and track the 2D pixel coordinates of the four passive reflective markers within the images recorded by the IP cameras with sub-pixel precision. The frames, having been converted to grayscale, are scanned across various threshold values to identify potential circle candidates. At the conclusion of this stage, the integer (rough) center coordinates (c_x, c_y) and radius (r) of the circular regions are calculated. However, the objective of the script is to calculate the center points with sub-pixel accuracy and to determine the circle that perfectly aligns with the reflective sphere. To achieve this, utilizing the initial rough centers, the Gaussian moments method is employed; by treating the pixel intensity values as mass, the brightness center of mass is calculated to a precision of three decimal places. Since this estimator uses the intensity distribution across the entire marker region rather than relying solely on edge sharpness, it remains relatively stable under moderate image blur. In particular, symmetric blur broadens the intensity profile without substantially shifting its centroid. Moreover, the high retro reflectivity of the markers under IR illumination preserves sufficient local contrast under relatively low ambient-light conditions. Subsequently, the Canny Edge Detection algorithm is applied around this bright center to extract the outer contours of the point. Then, a perfect circle was fitted to the detected edge pixels using the Least Squares mathematical method. This ensures that the center and radius are determined with maximum precision.

To guarantee that the identified passive reflective markers are tracked in a consistent order across every frame, the user manually assigns an identification number to each passive reflective marker in the initial frame. Subsequent detections are executed by taking these identification numbers into account. This procedure prevents the sequence of the passive reflective markers from becoming disorganized in subsequent frames and mitigates the risk of incorrect detections. Finally, combinations of four are generated from all the detected points. These combinations were scored according to circularity, size consistency, brightness, and proximity to their positions in the previous frame. The quadruple group attaining the highest score is accepted as the final result for that specific frame. The complete workflow of the marker detection and tracking procedure is summarized in Figure 7.

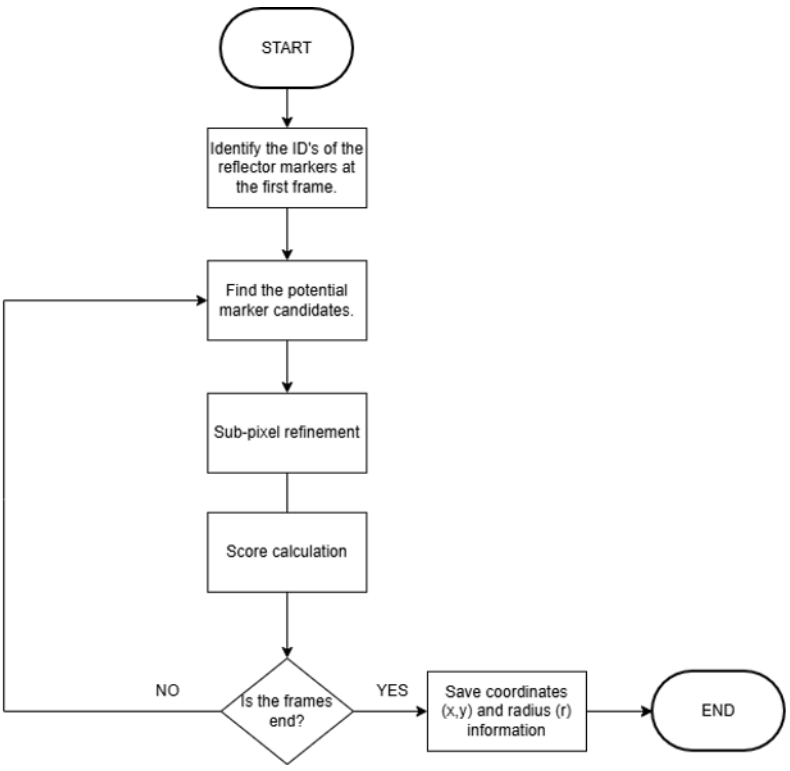


Figure 7. Flowchart of marker detection algorithm

Finally, the 2D image plane coordinates acquired from the multi-camera system were transformed into 3D spatial coordinates utilizing the intrinsic and extrinsic calibration parameters of the cameras. Following the correction of radial and tangential optical distortions on the sensor plane, the viewing rays from the six cameras were simultaneously resolved in 3D space for each reflective marker using the N-View Direct Linear Transformation (DLT) method. For this procedure, Singular Value Decomposition (SVD) was employed to calculate the point at which the rays most closely converge with sub-millimeter precision. Consequently, the 3D spatial positions of the four passive reflective markers were successfully determined. The flowchart of the triangulation procedure is presented in Figure 8.

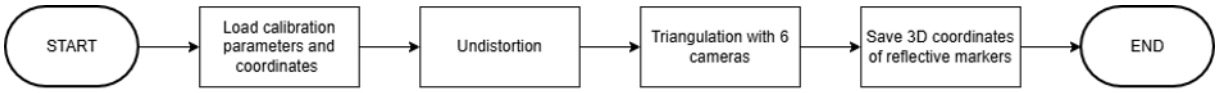


Figure 8. Flowchart of triangulation algorithm

All of the software described in this section, comprising the recording, marker detection, and triangulation modules, was developed in Python. The image processing operations and the undistortion of the 2D observations were performed with the OpenCV library, whereas the numerical computations, including the SVD based solution of the DLT system, were carried out with NumPy.

2.2.4. Experimental setup

The objective of this study is to experimentally validate the developed object detection and 3D position calculation software, and to determine its accuracy and precision, by calculating the spatial coordinates and inter-distances of four reflective markers placed at fixed locations with a constant spacing of 10 cm between them. In accordance with this objective, the reflective markers were sequentially and linearly arranged on a platform with dimensions of 30 cm x 80 cm x 80 cm (width x length x height), as illustrated in Figure 9. The levelness of this platform was verified utilizing a spirit level. Initially, the fourth reflective marker was positioned at a distance of 60 cm from the edge of the platform, and its exact distance was confirmed using a laser distance meter. Subsequently, the remaining reflective markers were linearly positioned at respective distances of 50 cm, 40 cm, and 30 cm from the starting point of the platform, with their distances similarly verified via the laser distance meter. Double-sided adhesive tape was employed to firmly secure the reflective markers. Images depicting the position measurements of the four reflective markers are provided in Figure 10.

Figure 9. Verification of the platform's levelness via a spirit level on the rigid platform (30 cm x 80 cm x 80 cm)

CC: BY-NC-ND

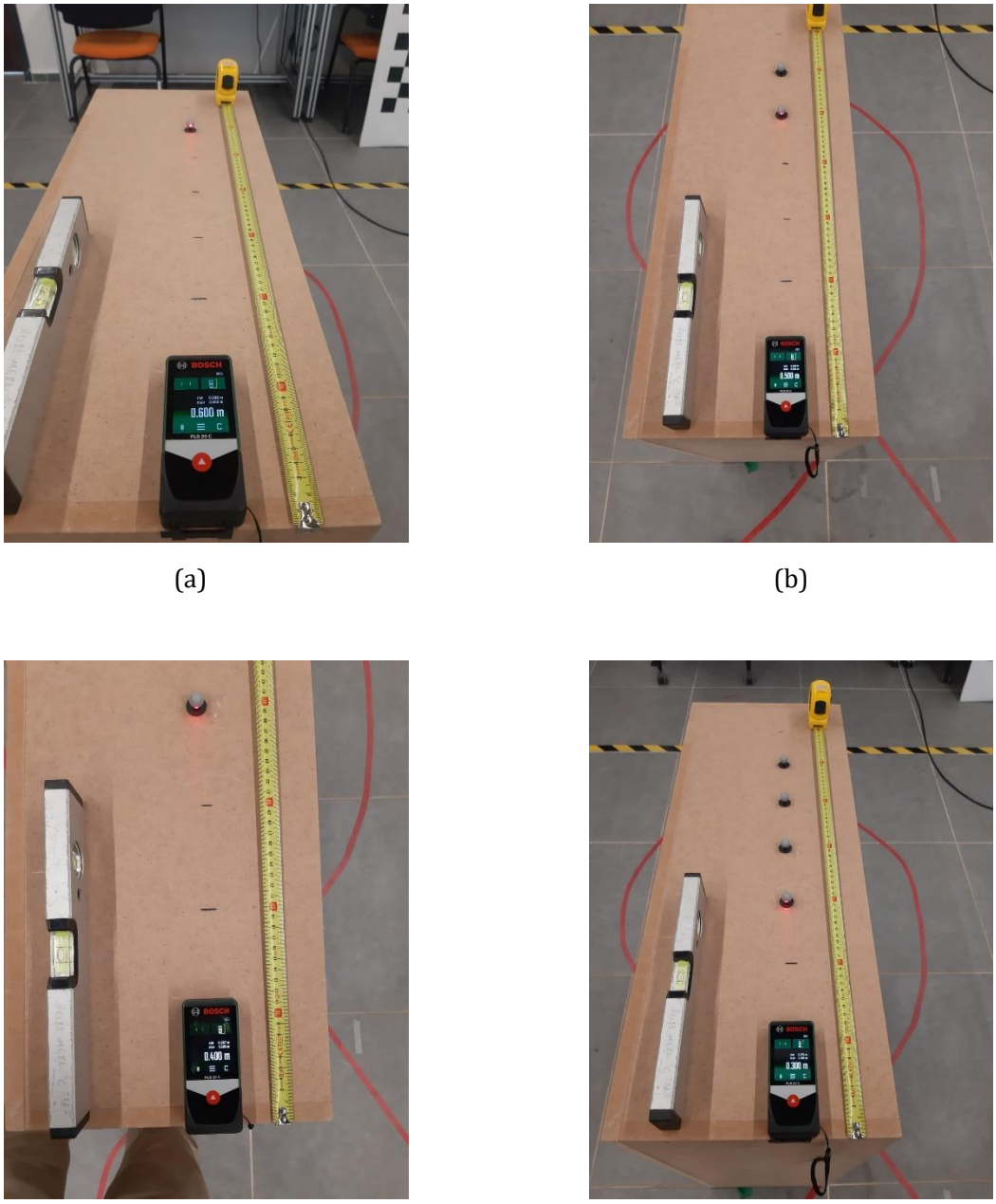


Figure 10. Reflector markers fixed on a solid platform. (a) 60 cm. (b) 50 cm. (c) 40 cm. (d) 30cm

Since the markers are rigidly attached to the platform, the distances between them remain physically constant regardless of how the platform moves; the laser-verified inter-marker distances can therefore serve as a reference not only while the platform is stationary but also while it is in motion. Accordingly, three separate recordings were acquired with the platform oriented along the X, Y, and Z axes of the workspace. In each recording, the platform was first held stationary, then moved by hand along the X axis in the X and Z axes variations and along the Y axis in the Z axis variation within the capture volume, and held stationary again, so that every recording contains both static and dynamic segments. The frames were classified as static and dynamic with a threshold of 0.05 m/s based on the velocity calculated from the marker trajectories; static evaluation was performed on 100 frames sampled at equal intervals on each axis, and dynamic evaluation was performed on all moving frames (139, 132, and 273 frames for X, Y, and Z respectively; peak velocities 2.82, 0.82, and 0.84 m/s). All experiments were conducted indoors under constant artificial lighting. Ambient light is not a

decisive factor in this configuration, as the cameras operate in IR mode and the detection of the retroreflective markers relies on IR contrast rather than on the visible illumination of the scene.

2.2.5. Analysis metrics

The experimental results presented in Section 3 are evaluated by means of five complementary statistical metrics computed over the distances between the consecutive marker pairs. In the following equations, d_i denotes the inter-marker distance calculated at frame i, d_{ref} denotes the corresponding reference distance verified by the laser distance meter, and N denotes the number of frames included in the evaluation. The signed distance error at frame i and its relative counterpart are defined as:

$$e_i = d_i - d_{ref} \tag{10}$$

$$e_i(\%) = \frac{|e_i|}{d_{ref}} x\ 100 \tag{11}$$

The mean signed error, referred to as the bias, quantifies the systematic component of the measurement error and is defined as:

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(d_i - d_{ref}) \tag{12}$$

Since positive and negative deviations cancel in Equation 12, the bias reveals whether a distance is systematically overestimated or underestimated, but it does not reflect the typical magnitude of the error. This magnitude is captured by the Mean Absolute Error (MAE), which averages the absolute deviations defined as follows [34]:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|d_i - d_{ref}| \tag{13}$$

The Root Mean Square Error (RMSE) likewise measures the overall deviation from the reference; however, as the deviations are squared before averaging, larger errors are penalized more heavily, which renders the RMSE more sensitive to outliers than the MAE [34, 35]:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - d_{ref})^2} \tag{14}$$

Finally, the standard deviation (σ) of the calculated distances characterizes the frame-to-frame repeatability of the system independently of the reference value, where $\overline{d}$ denotes the mean of the calculated distances:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - \overline{d})^2} \tag{15}$$

Bias, MAE, and RMSE characterize the accuracy of the system by quantifying how closely the measurements agree with the reference value, whereas standard deviation reflects precision by indicating the consistency of repeated measurements. By construction, these quantities satisfy the

CC: BY-NC-ND

decomposition $RMSE^2 = bias^2 + \sigma^2$ which enables the systematic and random components of the total measurement error to be assessed separately. The relationship between MAE and the absolute bias also provides insight into the distribution of the errors. When all errors have the same sign, the MAE approaches the absolute value of the bias; as the measurements begin to fluctuate on both sides of the reference value, the difference between these two metrics becomes more pronounced.

3. RESULTS AND DISCUSSION

This section presents the experimental results of the low-cost, high-precision passive-marker optical motion capture system proposed in this study and analyses the acquired findings. To verify the 3D positioning accuracy and reliability of the developed system, the margins of error between the actual distances measured by a reference laser distance meter and the spatial coordinates calculated by the software that detailed in Section 2.2.3 were analysed. For this purpose, utilizing the experimental setup described in the previous section, the rigid platform was positioned such that the spatial coordinates of the reflective markers varied sequentially across the X, Y, and Z axes. The platform was initially positioned along the X-axis within the area designated as the virtual origin. An image of the platform illustrating the established virtual origin is provided in Figure 11. Utilizing the images acquired in IR mode, the reflective markers were detected with sub-pixel precision, and their respective center points were determined. The outputs derived from this stage are presented in Figure 12.

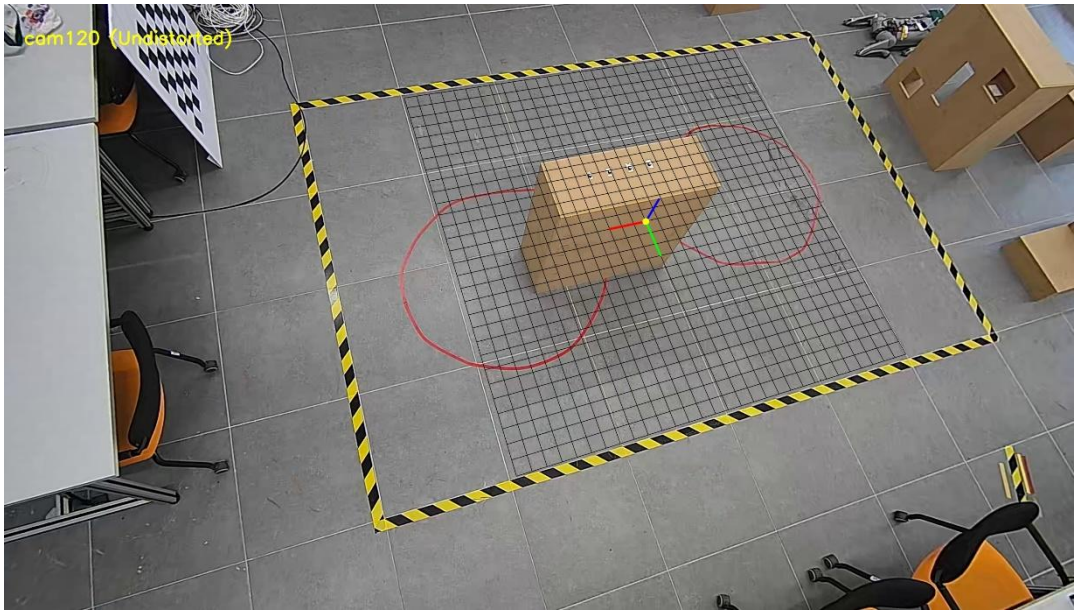


Figure 11. Image depicting the virtual origin (rigid platform positioned along the X-axis)

CC: BY-NC-ND

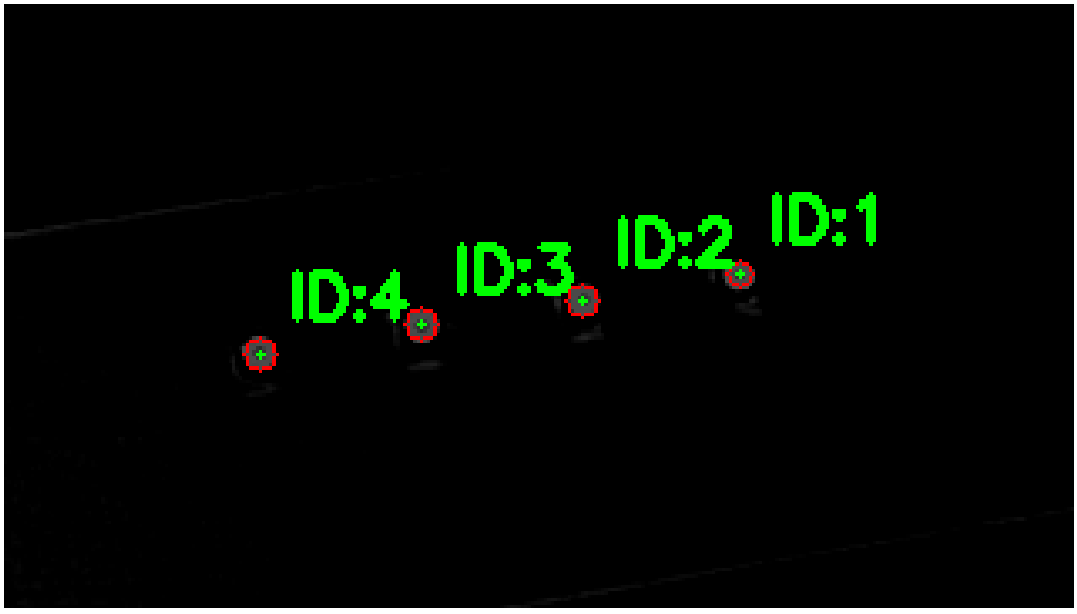


Figure 12. Reflector marker detections in the image acquired in IR mode (zoomed-in view)

Utilizing the detections acquired in IR mode, triangulation was performed to determine the 3D positions of the passive reflective markers relative to the established origin. These results are presented in Table 4. Furthermore, the error calculations that computed under the assumption that the laser distance meter yielded flawless measurements and that the linear positioning of the reflective markers was executed perfectly are detailed in Table 5.

A holistic examination of the findings presented in Table 4 and Table 5 reveals that the developed motion capture system exhibits high precision and stability in 3D space. Table 4 details the X, Y, and Z spatial coordinates of the fixed markers, calculated relative to the virtual origin across three consecutive frames, when the rigid platform was positioned along the X-axis. The fact that the deviations in the calculated X, Y, and Z coordinates of the four distinct passive reflective markers are confined to sub-millimeter levels (0.0001 m) not only verifies the physical levelness of the experimental setup ensured by the spirit level but also substantiates the spatial consistency of the marker detection and 3D position calculation software. An analysis of the four distinct Y and Z coordinates calculated for the respective markers indicates that the detections were achieved with millimeter-level precision.

Furthermore, the error calculations provided in Table 5 were performed under the premise that the laser distance meter yielded flawless measurements and that the linear positioning of the passive reflective markers was executed with absolute perfection. For instance, it is assumed that the initial passive reflective marker was placed precisely at a distance of 0.30000 m from the edge of rigid platform and the subsequent passive reflective marker was positioned with perfect linearity at an exact distance of 0.40000 m. Additionally, the intrinsic ±2.0 mm precision tolerance of the utilized laser distance meter is disregarded in this context. Under these assumptions, it was determined that the calculated absolute error ranged between 0.326 mm and 2.377 mm with the maximum relative error constrained to 2.377%.

Using the evaluation protocol described in the previous section, the three-dimensional distances between each pair of consecutive markers were compared with the laser-verified reference distance of 100 mm. System performance was evaluated using the RMSE, the MAE, the signed bias, and the standard deviation of the calculated distances. The first three quantify the accuracy of the system, whereas σ characterizes its frame-to-frame precision independently of the reference; by construction, these metrics satisfy the decomposition $RMSE^2 = bias^2 + \sigma^2$, which allows the systematic and random components of the total error to be assessed separately. Table 6 extends the three-frame examination given in Table 5 to the complete X-axis recording and separates the static and dynamic behaviour of

CC: BY-NC-ND

the system. In the static segment, the RMSE of each pair is almost entirely accounted for by its bias (e.g., RMSE 3.360 mm versus bias +3.357 mm for pair 2-3), while the standard deviations remain confined to 0.127–0.133 mm; in other words, the deviations from the 100 mm reference do not fluctuate from frame to frame but persist as constant offsets, which is the signature of marker-placement uncertainty rather than of measurement noise. The dynamic segment preserves this structure on the accuracy side, so the bias values remain within approximately 0.2 mm of their static counterparts (+3.357 mm versus +3.444 mm for pair 2-3), and the RMSE values change only marginally. The effect of motion is instead concentrated in the precision component, where σ increases to 0.598–0.954 mm as the platform reaches a peak speed of 2.82 m/s. This behaviour is consistent with the software-triggered acquisition of the cameras: the residual inter-camera time offset introduces a speed-dependent scatter into the triangulated positions, degrading the frame-to-frame repeatability while leaving the systematic accuracy of the reconstructed geometry unaffected.

Table 4. Reflective marker positions (on the X-axis)

	Reflective Marker 1			Reflective Marker 2		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
Frame 1	0.300991	0.104196	0.803230	0.402707	0.103878	0.802849
Frame 2	0.300989	0.104235	0.803217	0.402704	0.103858	0.802852
Frame 3	0.300986	0.104229	0.803250	0.402701	0.103847	0.802854
	Reflective Marker 3			Reflective Marker 4		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
Frame 1	0.505083	0.103557	0.802784	0.604757	0.103253	0.802828
Frame 2	0.505022	0.103587	0.802833	0.605063	0.103317	0.802695
Frame 3	0.505081	0.103579	0.802806	0.605208	0.103278	0.802722

Table 5. Reflective marker error calculation

Distance Between Reflective Markers	Actual Distance (cm)	Calculated Distance (cm)	Error (mm)	Error (%)
1-2	10	10.1717	1.717	%1.717
2-3	10	10.2377	2.377	%2.377
3-4	10	9.9674	0.326	%0.326

Table 6. Reflective marker error calculation for static and dynamic frames (on the X-axis)

Static 100 Frames				
Distance Between Reflective Markers	RMSE of Calculated Distance (mm)	MAE of Calculated Distance (mm)	Bias (mm)	Standard Deviation (mm)
1-2	1.256	1.249	+1.249	0.127
2-3	3.360	3.357	+3.357	0.133
3-4	1.002	0.994	-0.994	0.128
Dynamic 139 Frames				
1-2	1.426	1.172	+1.060	0.954
2-3	3.496	3.484	+3.444	0.598
3-4	1.286	1.060	-0.942	0.876

It is imperative to note that the observed minor deviations do not solely represent software-induced error tolerances; they inherently encompass physical measurement uncertainties stemming from the manual placement of the reflective spheres using double-sided adhesive tape. In conclusion, these data show that the calibration and sub-pixel detection algorithms of the proposed low-cost system perform consistently, delivering 3D positioning with millimeter-level accuracy that can serve as a viable alternative to industrial-standard systems.

CC: BY-NC-ND

The outputs of the calculations performed with the platform positioned along the Y-axis are presented in Table 7 and Table 8. Figure 13 illustrates the platform and the reflective spheres positioned along the Y-axis. Table 7 presents the calculated X, Y, and Z positional data of the fixed markers relative to the virtual origin across three consecutive frames when the rigid platform was positioned along the Y-axis. Similar to the results obtained when the rigid platform was positioned along the X-axis, the fact that the deviations in the calculated X, Y, and Z coordinates of the four distinct passive reflective markers are confined to sub-millimeter levels (0.0001 m) demonstrates that the system maintains its stability independent of the axis. According to the data in Table 7, when the markers are sequentially positioned along the Y-axis, the fluctuations in the X and Z coordinates which are theoretically expected to remain constant are again observed to occur at the sub-millimeter (0.0001 m) scale. This substantiates that the marker detection and 3D position calculation software does not lose its spatial accuracy across different orientations. The error analyses, conducted on the premise that the reference equipment and manual procedures are entirely flawless, are presented in Table 8. Examining the distance error analysis provided in Table 8, it was determined that, compared to the reference intervals verified as 10 cm by the laser distance meter, the absolute error in the values calculated by the system ranges between 0.567 mm and 2.069 mm, with the maximum relative error constrained to 2.069%. In parallel with the X-axis findings, it is evaluated that this calculated maximum deviation of approximately 2 mm is not solely an algorithmic tolerance but inherently encompasses physical setup uncertainties arising from the manual positioning of the reflective spheres using tape.

An examination of the full-recording statistics presented in Table 9 reveals that the pattern established for the X-axis is reproduced along the Y-axis. In the static segment, the standard deviations of the calculated inter-marker distances remain confined to the range of 0.086–0.139 mm, and the RMSE of each pair is once again governed almost entirely by its bias component, as exemplified by the 2-3 pair, for which an RMSE of 3.253 mm corresponds to a bias of +3.252 mm Because the peak platform velocity in this recording was limited to 0.82 m/s, the dynamic results remain closely aligned with the static measurements. The dynamic bias values differ from their static counterparts by less than 0.1 mm, while the RMSE values show only minor changes. Although the standard deviations increase moderately to between 0.196 and 0.231 mm, the overall measurement stability is largely preserved during motion. The bias values also retain the same sign pattern observed in the X-axis recording. It is furthermore noteworthy that the sign structure of the bias values coincides with that observed in the X-axis recording. This consistency substantiates the interpretation that the observed offsets originate from the physical placement of the reflective markers rather than from any axis-dependent characteristic of the measurement system.

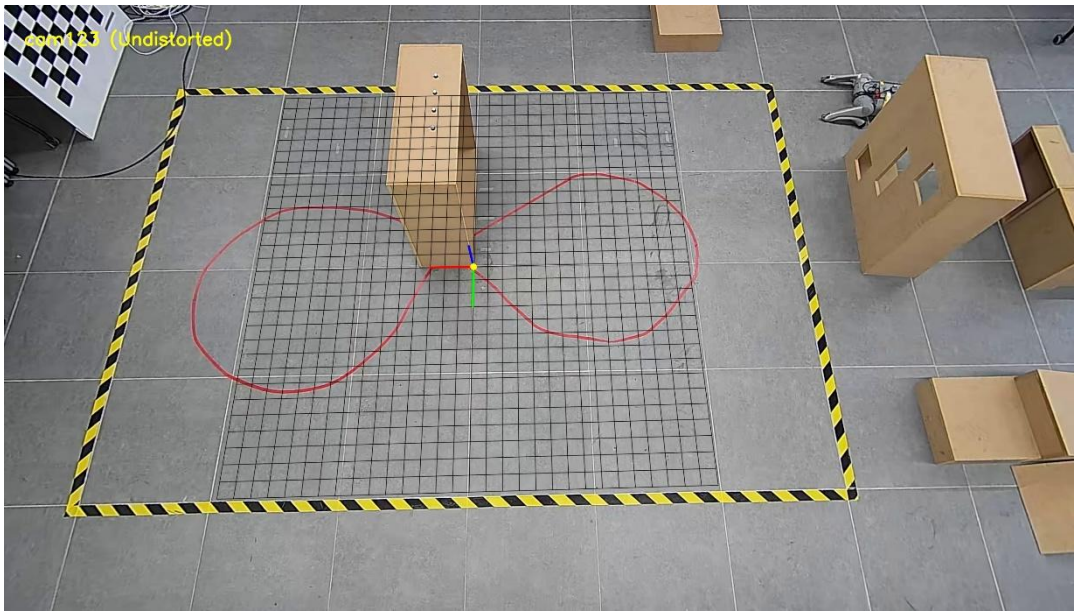


Figure 13. Image depicting the virtual origin (rigid platform positioned along the Y-axis)

Table 7. Reflective marker positions (on the Y-axis)

	Reflective Marker 1			Reflective Marker 2		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
Frame 1	0.112614	-0.301970	0.804430	0.112383	-0.403360	0.804812
Frame 2	0.112599	-0.302021	0.804449	0.112403	-0.403351	0.804769
Frame 3	0.112624	-0.301983	0.804460	0.112392	-0.403378	0.804799
	Reflective Marker 3			Reflective Marker 4		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
Frame 1	0.111634	-0.505426	0.804871	0.111408	-0.604859	0.804952
Frame 2	0.111583	-0.505419	0.804811	0.111404	-0.604848	0.804941
Frame 3	0.111592	-0.505423	0.804863	0.111436	-0.604862	0.804970

Table 8. Reflective marker error calculation

Distance Between Reflective Markers	Actual Distance (cm)	Calculated Distance (cm)	Error (mm)	Error (%)
1-2	10	10.1391	1.391	%1.391
2-3	10	10.2069	2.069	%2.069
3-4	10	9.9433	0.567	%0.567

Table 9. Reflective marker error calculation for static and dynamic frames (on the Y-axis)

Static 100 Frames				
Distance Between Reflective Markers	RMSE of Calculated Distance (mm)	MAE of Calculated Distance (mm)	Bias (mm)	Standard Deviation (mm)
1-2	0.793	0.784	+0.784	0.121
2-3	3.253	3.252	+3.252	0.086
3-4	1.276	1.269	-1.269	0.139
Dynamic 132 Frames				
1-2	0.793	0.768	+0.768	0.196
2-3	3.174	3.168	+3.168	0.202
3-4	1.339	1.319	-1.319	0.231

Finally, the platform was positioned such that the distances of the reflective spheres varied along the Z-axis, and the subsequent calculations were performed. The results are presented in Table 10 and Table 11. An image depicting the platform positioned along the Z-axis, in relation to the virtual origin, is provided in Figure 14. An examination of the data from Table 10 and Table 11, obtained by positioning the platform along the Z-axis (depth), confirms that the system exhibits a similarly high performance in depth perception. Evaluating Table 10 and the consecutive time-stamped log records, it is observed that when the passive reflective markers are aligned along the Z-axis, the deviations in the X and Y coordinates which theoretically should remain constant are once again maintained at sub-millimeter (0.0001 m) levels. This finding clearly demonstrates that the optical tracking software successfully sustains its spatial accuracy along the depth axis and does not compromise its stability across different orientations.

Furthermore, the error analysis in Table 11, predicated on the assumption of flawless reference measurement devices and perfectly linear passive reflective marker placement on the platform, numerically supports the system's reliability. In this analysis, which neglects the inherent error margin of the laser distance meter (±2.0 mm), the absolute deviations between the consecutive distances calculated by the system and the ideal true distances of 10 cm (100 mm) were found to fall within the range of 0.557 mm to 1.552 mm. The maximum relative error for this axis was calculated at 1.552%. As emphasized in the previous axis tests, these measured discrepancies of approximately 1.5

CC: BY-NC-ND

millimeters do not consist solely of software computational tolerances; they also inherently encompass physical setup misalignments that occur during the manual attachment of the markers to the platform. Ultimately, the Z-axis results support the conclusion that the proposed low-cost system maintains millimeter-level accuracy in all orientations of 3D space and can serve as a viable alternative to high-budget commercial systems in static and low-speed applications.

The Z-axis results presented in Table 12 complete the full-recording analysis. Under static conditions, the standard deviations range from 0.181 to 0.204 mm, indicating that the system maintains sub-millimeter repeatability in the depth direction. The dynamic segment was recorded at a peak velocity of 0.84 m/s, comparable to that of the Y-axis experiment, and shows only a limited increase in standard deviation to 0.303–0.344 mm. At the same time, the bias values remain largely unchanged, suggesting that platform motion has little effect on the systematic component of the measurement error. The 3–4 marker pair is particularly informative because its placement offset is close to zero, with a bias of −0.096 mm. For this pair, the static RMSE decreases to 0.222 mm, providing a clearer indication of the accuracy achievable by the system when the influence of marker-placement uncertainty is minimal.

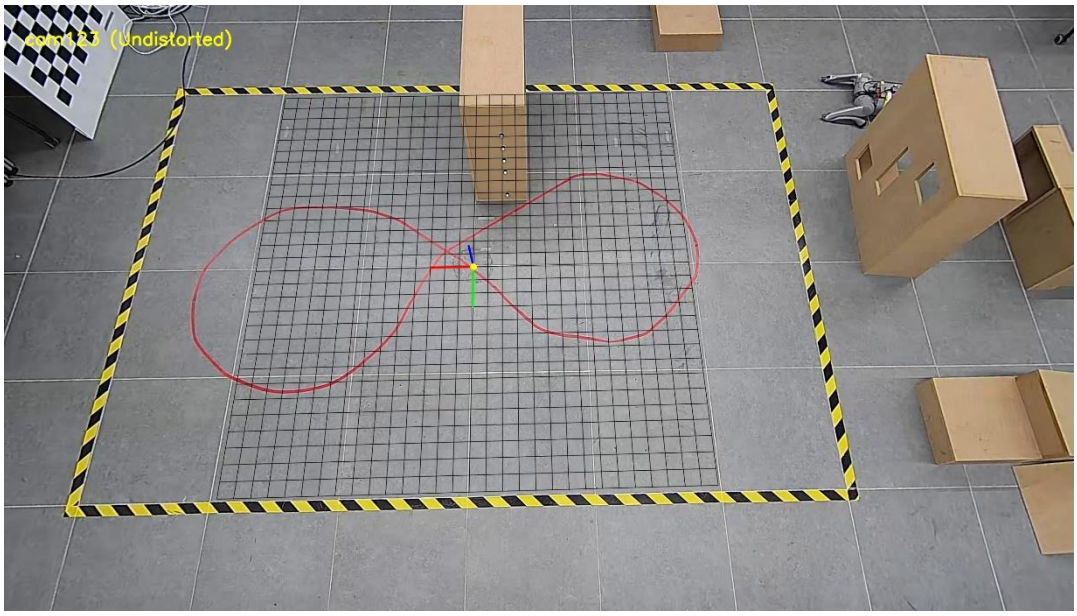


Figure 14. Image depicting the virtual origin (rigid platform positioned along the Z-axis)

Table 10. Reflective marker positions (on the Z-axis)

	Reflective Marker 1			Reflective Marker 2		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
Frame 1	-0.188725	-0.412876	0.289353	-0.188126	-0.412253	0.390747
Frame 2	-0.188777	-0.412929	0.289286	-0.188148	-0.412285	0.390726
Frame 3	-0.188777	-0.412929	0.289286	-0.188148	-0.412285	0.390726
	Reflective Marker 3			Reflective Marker 4		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
Frame 1	-0.187157	-0.412117	0.492294	-0.186360	-0.411735	0.591733
Frame 2	-0.187215	-0.412158	0.492211	-0.186393	-0.411753	0.591683
Frame 3	-0.187215	-0.412158	0.492211	-0.186393	-0.411753	0.591683

Table 11. Reflective marker error calculation

Distance Between Reflective Markers	Actual Distance (cm)	Calculated Distance (cm)	Error (mm)	Error (%)
1-2	10	10.1398	1.398	%1.398
2-3	10	10.1552	1.552	%1.552
3-4	10	9.9443	0.557	%0.557

Table 12. Reflective marker error calculation for static and dynamic frames (on the Z-axis)

Static 100 Frames				
Distance Between Reflective Markers	RMSE of Calculated Distance (mm)	MAE of Calculated Distance (mm)	Bias (mm)	Standard Deviation (mm)
1-2	0.971	0.950	-0.950	0.204
2-3	2.180	2.172	+2.172	0.181
3-4	0.222	0.156	-0.096	0.200
Dynamic 273 Frames				
1-2	1.250	1.202	-1.202	0.344
2-3	2.495	2.476	+2.476	0.303
3-4	0.375	0.307	-0.210	0.310

A holistic examination of Tables 6, 9, and 12 yields a consistent overall picture. The decomposition $RMSE^2 = bias^2 + \sigma^2$ is satisfied for every marker pair, with the frame-to-frame precision confined to 0.086–0.204 mm in all static cases. The bias values remain stable in sign and magnitude both across the three axes and across the static and dynamic separation, which indicates that they stem from the manual placement of the reflective markers and the ±2.0 mm tolerance of the reference instrument rather than from the measurement pipeline itself; consequently, the reported RMSE values should be interpreted as an upper bound on the accuracy of the system. Motion, in turn, affects only the random component of the error, which scales with the platform speed as a consequence of the software-triggered acquisition. Finally, it should be noted that throughout all frames of the three recordings, the four reflective markers were detected in a sufficient number of cameras for the triangulation to proceed without interruption, and no occlusion-related data loss was encountered.

4. DISCUSSION and CONCLUSION

In this study, as an alternative to high-cost commercial systems considered industry standards, a low-cost and high-accuracy passive marker optical motion capture system was developed and experimentally validated. To test the accuracy and algorithmic stability of the system in three-dimensional space, passive reflective markers placed at constant 10 cm intervals on a rigid platform were recorded along the X, Y and Z axes, and the recordings, each containing both stationary and moving segments, were evaluated statistically over their entire duration.

The error calculations conducted with reference to the laser distance meter indicate that the system achieves millimeter-level accuracy under the tested conditions. In the static segments, the RMSE of the inter-marker distances remained below 3.360 mm, 3.253 mm, and 2.180 mm for the X, Y and Z axes recordings, respectively, while the frame-to-frame standard deviation of the calculated distances did not exceed 0.204 mm in any case. The decomposition of the total error further showed that these RMSE values are governed by constant offsets that remain stable across the axes and across the static and dynamic segments. It is therefore evaluated that the millimetric deviations do not originate from software computation or algorithmic limits; rather, they encompass the tolerance of the reference measurement device and the physical setup uncertainties arising from the manual positioning of the markers with tape, so that the reported RMSE values constitute an upper bound on the accuracy of the system itself. The dynamic segments showed a similar pattern. The bias values remained at their static levels, and only the random component of the error grew with the platform

speed, reaching a standard deviation of 0.954 mm at a peak speed of 2.82 m/s, consistent with the software-triggered acquisition of the cameras.

In conclusion, thanks to its calibration techniques and sub-pixel level detection algorithms, the proposed low-cost system offers millimeter-level accuracy together with sub-millimeter repeatability in 3D positioning. In the future development phases of the system, it is aimed to analyse the dynamic trajectory tracking performance on mobile platforms, and to transform the system into a 'ground truth' reference to conduct accuracy tests and performance analyses of SLAM algorithms. However, certain limitations should be considered when using the system in dynamic applications. The fact that the cameras used in the system can provide image streaming at a maximum rate of 25 FPS directly limits the maximum speed at which a moving mobile robot can be tracked. Therefore, to ensure reliable and continuous position tracking, the motion speed of the mobile robot should be limited in accordance with the sampling capacity of the cameras. The passive reflective markers should be positioned on the robot with sufficient spacing between them and in a geometrically distinguishable configuration (for example, an asymmetric “L” configuration can be formed using three passive reflective markers with different inter-marker distances). In this way, the overlap of markers in the camera images can be minimized, allowing each marker to be detected individually and enabling the robot’s orientation information to be calculated reliably. Another limitation regarding the statistical generalizability of the experimental results is that each experimental condition was evaluated in a single trial.

Peer-review: External, Independent.

Acknowledgements:

This study was supported under the TÜBİTAK BİDEB 2210-C Domestic Priority Areas Master's Scholarship Program. We extend our sincerest gratitude to TÜBİTAK and BİDEB.

Declarations:

1. Statement of Originality:

This work is original.

2. Author Contributions:

Concept: DAK,MP; ***Conceptualization:*** DAK,MP; ***Literature Search:*** DAK; ***Data Collection:*** DAK; ***Data Processing:*** DAK; ***Analysis:*** DAK,MP; ***Writing – original draft:*** DAK,MP; ***Writing – review & editing:*** DAK,MP.

3. Ethics approval:

Not applicable.

4. Funding/Support:

This study was supported under the TÜBİTAK BİDEB 2210-C Domestic Priority Areas Master's Scholarship Program.

5. Competing Interests:

The authors declare no competing interests.

6. GenAI Usage Statement:

The authors declare that generative artificial intelligence (GenAI) tools were used only for language editing and grammatical improvement during manuscript preparation. The authors take full responsibility for the content of the manuscript.

7. Sustainable Development Goals:

REFERENCES

[1] Zhong, W., Zhang, L., Sun, Z., Dong, M., Zhang, M. 2024. UI-MoCap: An Integrated UWB-IMU Circuit Enables 3D Positioning and Enhances IMU Data Transmission. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 32, 1034-1044. DOI: https://doi.org/10.1109/TNSRE.2024.3369647

[2] Vicon, n.d. What Is Motion Capture, How It Works, and What It's Used For. Available at: https://www.vicon.com/about-us/what-is-motion-capture/ (Accessed: 14.05.2026).

[3] Wibowo, M. C., Nugroho, S., Wibowo, A. 2024. The Use of Motion Capture Technology in 3D Animation. International Journal of Computing and Digital Systems, 15, 975-987. DOI: https://doi.org/10.12785/ijcds//150169

[4] Damian, I., Bühling, R., Obaid, M., Billinghurst, M., André, E. 2013. Motion Capturing Empowered Interaction with a Virtual Agent in an Augmented Reality Environment. 2013 IEEE International Symposium on Mixed and Augmented Reality, Adelaide, SA, Australia, 1-6. DOI: https://doi.org/10.1109/ISMAR.2013.6671830

[5] Zordan, V. B., Hodgins, J. K. 2002. Motion Capture-Driven Simulations that Hit and React. Proceedings of the 2002 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, New York, NY, USA, 89-96. DOI: https://doi.org/10.1145/545261.545276

[6] Yamane, K., Hodgins, J. 2009. Simultaneous Tracking and Balancing of Humanoid Robots for Imitating Human Motion Capture Data. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, 2510-2517. DOI: https://doi.org/10.1109/IROS.2009.5354750

[7] Zou, Y., Wang, P., Tang, Q., Sun, Y. 2019. Design of Motion Capture System Based on Virtual Reality Scene. 2019 2nd International Conference on Safety Produce Informatization, Chongqing, China, 526-529. DOI: https://doi.org/10.1109/IICSPI48186.2019.9096046

[8] Du, Y. 2025. Research on Police Physical Training Based on Virtual Reality Technology. Quality in Sport, 38, 58534. DOI: https://doi.org/10.12775/QS.2025.38.58534

[9] Liparoti, M., Troisi Lopez, E., Agosti, V. 2020. Motion Capture System: A Useful Tool to Study Cyclist's Posture. Journal of Physical Education and Sport, 20, 2364-2367. DOI: https://doi.org/10.7752/jpes.2020.s4320

[10] Debarba, H. G., De Oliveira, M. E., Ladermann, A., Chague, S., Charbonnier, C. 2018. Augmented Reality Visualization of Joint Movements for Physical Examination and Rehabilitation. 2018 IEEE Conference on Virtual Reality and 3D User Interfaces, Reutlingen, Germany, 537-538. DOI: https://doi.org/10.1109/VR.2018.8446368

[11] Gan, Q., Li, Y., Wang, G., Zhang, Y. 2020. Application Research of Optical Tracking Point Layout in Computer Motion Capture Technology: Take Walking Motion as an Example. 2020 International Conference on Innovation Design and Digital Technology, Zhenjing, China, 548-552. DOI: https://doi.org/10.1109/ICIDDT52279.2020.00109

[12] Jalil, N. A. A., Amdan, S. M., Jorkasi, Z., Hussein, K., Jamal, N. 2025. A Systematic Review of Motion Capture Technologies Applied to Ergonomic Assessment. International Journal of Research and Innovation in Social Science, 9, 10380-10390. DOI:

https://doi.org/10.47772/IJRISS.2025.910000844

[13] Rybnikár, F., Kačerová, I., Hořejší, P., Šimon, M. 2022. Ergonomics Evaluation Using Motion Capture Technology—Literature Review. Applied Sciences, 13, 162. DOI: https://doi.org/10.3390/app13010162

[14] Shu, Z., Bei, S., Dai, J., Li, L., Chen, Z. 2025. MoCap2GT: A High-Precision Ground Truth Estimator for SLAM Benchmarking Based on Motion Capture and IMU Fusion. arXiv. DOI: https://doi.org/10.48550/arXiv.2507.12920

[15] Rodríguez-Ramirez, J. O., Castillo-Sánchez, J. P., Gonzalez-Hernandez, H. G., Reyes-Avendaño, J. A. 2025. Validation of LIDAR-Based Odometry Using a Motion Capture System on a Scaled Autonomous Platform. 2025 IEEE Mexican Humanitarian Technology Conference, Puebla, Mexico, 55-60. DOI: https://doi.org/10.1109/MHTC65808.2025.11086459

[16] Godil, A., Eastman, R., Hong, T. H. 2013. Ground Truth Systems for Object Recognition and Tracking. National Institute of Standards and Technology, NIST IR 7923. DOI: https://doi.org/10.6028/NIST.IR.7923

[17] Field, M., Stirling, D., Naghdy, F., Pan, Z. 2009. Motion Capture in Robotics Review. 2009 IEEE International Conference on Control and Automation, Christchurch, New Zealand, 1697-1702. DOI: https://doi.org/10.1109/ICCA.2009.5410185

[18] Ray, L. S. S., Zhou, B., Suh, S., Lukowicz, P. 2024. A Comprehensive Evaluation of Marker-Based, Markerless Methods for Loose Garment Scenarios in Varying Camera Configurations. Frontiers in Computer Science, 6, 1379925. DOI: https://doi.org/10.3389/fcomp.2024.1379925

[19] OptiTrack, n.d. Motion Capture Systems. Available at: https://optitrack.com (Accessed: 14.05.2026).

[20] Unger, T., Sal Moslehian, A., Peiffer, J. D., Ullrich, J., Gassert, R., Lambercy, O., Cotton, J., Awai Easthope, C. 2026. Differentiable Biomechanics for Markerless Motion Capture in Upper Limb Stroke Rehabilitation: A Comparison with Optical Motion Capture. IEEE Transactions on Medical Robotics and Bionics, 8, 90-97. DOI: https://doi.org/10.1109/TMRB.2025.3605962

[21] Maggioni, V., Azevedo-Coste, C., Durand, S., Bailly, F. 2025. Optimisation and Comparison of Markerless and Marker-Based Motion Capture Methods for Hand and Finger Movement Analysis. Sensors, 25, 1079. DOI: https://doi.org/10.3390/s25041079

[22] Geiger, A., Lenz, P., Stiller, C., Urtasun, R. 2013. Vision Meets Robotics: The KITTI Dataset. The International Journal of Robotics Research, 32, 1231-1237. DOI: https://doi.org/10.1177/0278364913491297

[23] Blanco-Claraco, J. L., Moreno-Dueñas, F. A., Gonzalez-Jimenez, J. 2014. The Málaga Urban Dataset: High-Rate Stereo and LiDAR in a Realistic Urban Scenario. The International Journal of Robotics Research, 33, 207-214. DOI: https://doi.org/10.1177/0278364913507326

[24] Carlevaris-Bianco, N., Ushani, A. K., Eustice, R. M. 2016. University of Michigan North Campus Long-Term Vision and Lidar Dataset. The International Journal of Robotics Research, 35, 1023-1035. DOI: https://doi.org/10.1177/0278364915614638

[25] Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W., Siegwart, R. 2016. The EuRoC Micro Aerial Vehicle Datasets. The International Journal of Robotics Research, 35, 1157-1163. DOI: https://doi.org/10.1177/0278364915620033

[26] Pfrommer, B., Sanket, N., Daniilidis, K., Cleveland, J. 2017. PennCOSYVIO: A Challenging Visual Inertial Odometry Benchmark. IEEE International Conference on Robotics and Automation, Singapore, 3847-3854. DOI: https://doi.org/10.1109/ICRA.2017.7989443

CC: BY-NC-ND

[27] Schubert, D., Goll, T., Demmel, N., Usenko, V., Stückler, J., Cremers, D. 2018. The TUM VI Benchmark for Evaluating Visual-Inertial Odometry. IEEE/RSJ International Conference on Intelligent Robots and Systems, Madrid, Spain, 1680-1687. DOI: https://doi.org/10.1109/IROS.2018.8593419

[28] Yin, J., Li, A., Li, T., Yu, W., Zou, D. 2021. M2DGR: A Multi-Sensor and Multi-Scenario SLAM Dataset for Ground Robots. arXiv. DOI: https://doi.org/10.48550/arXiv.2112.13659

[29] Amazon.co.jp, n.d. Reflective Markers + Marker Base Set, 10 pcs, Markers 0.6 inches (14 mm) × M4 + Base, W15 × H5. Available at: https://www.amazon.co.jp/dp/B08FWH2V8X (Accessed: 14.05.2026).

[30] Dahua Technology, n.d. IPC-HFW5442E-ZE: 4MP IR Vari-Focal Bullet WizMind Network Camera. Available at: https://www.dahuasecurity.com/tr/products/All-Products/Network-Cameras/WizMind-8-Series/5-Series/4MP/IPC-HFW5442E-ZE (Accessed: 14.05.2026)

[31] Bosch DIY, n.d. PLR 50 C. Available at: https://www.bosch-diy.com/storage/diy/zip-manuals/plr-50-c-100037744-zip-389489/tr-TR/index.html (Accessed: 14.05.2026).

[32] Dahua Transmission Technology, n.d. S4210-8GT-110: 10-Port Managed Desktop Gigabit Switch with 8-Port PoE. Available at: https://transmission.dahuasecurity.com/en/product-list/Business/CloudManagementSolution/NetworkSwitch/S4210-8GT-110 (Accessed: 14.05.2026).

[33] Shih, S. W. 1993. Accurate Linear Technique for Camera Calibration Considering Lens Distortion by Solving an Eigenvalue Problem. Optical Engineering, 32, 138. DOI: https://doi.org/10.1117/12.60087

[34] Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research 30, 79–82. DOI: https://doi.org/10.3354/cr030079

[35] Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscientific Model Development 7, 1247–1250. DOI: https://doi.org/10.5194/gmd-7-1247-2014.

CC: BY-NC-ND