Proceedings of CONF-MLA 2026 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2026.LD38736

Robot Hand-Eye Calibration Target Selection and Design: From Planar Targets to Calibration-Free Methods

Wenbo Xiong

Ulink Collage of Beijing, Beijing, China
Xiong242302@outlook.com

Abstract. Robot hand-eye calibration is a fundamental procedure for establishing the spatial relationship between a robot's end-effector and its vision sensor, which is critical for precision tasks in automation. The accuracy and efficiency of this calibration process are heavily dependent on the proper selection of calibration targets. However, existing surveys predominantly focus on solution algorithms while lacking a systematic treatment of target design. This paper reviews four categories of calibration targets in hand-eye calibration—2D planar targets, 3D standard bodies, single-plane/natural workpieces, and calibration-free methods—examining their design principles, applicable conditions, and representative contributions. A systematic comparison is conducted across dimensions including manufacturing cost, calibration accuracy, operational convenience, sensor compatibility, and online calibration friendliness. Based on this analysis, a decision framework for selecting calibration targets in engineering applications is proposed to guide calibration scheme design under different scenarios.

Keywords: hand-eye calibration, calibration target, robot vision, AX=XB equation, calibration-free methods

1. Introduction

Robotic hand-eye calibration is a fundamental prerequisite for enabling autonomous manipulation and precise interaction with the environment. The core objective is to determine the rigid transformation between a robot's end-effector and its vision system, which directly governs the accuracy of subsequent perception-driven tasks. Over the past decades, extensive research has been devoted to this problem, yielding a rich body of work spanning from the mathematical formulation of the $AX = XB$ equation to the derivation and optimization of solution algorithms .

Despite these advances, practical deployment remains challenging. Traditional calibration methods typically rely on dedicated targets such as chessboards, circle grids, or coded markers, which require careful placement within the robot's workspace and multiple data acquisitions from different poses. This process is not only time-consuming but also constrained by environmental conditions, especially in unstructured or dynamic settings where access to the target is limited. Moreover, the accuracy of calibration is highly sensitive to target detection errors, motion noise, and the geometric configuration of sampled poses.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.LD38736

In recent years, with the rapid development of deep learning and self-supervised perception, calibration-free approaches and natural feature-based methods have emerged as promising alternatives. These techniques aim to reduce or eliminate the dependency on artificial targets by leveraging the geometric structure of the environment or learned representations. However, they often trade off absolute precision for convenience, and their robustness under varying lighting conditions or textures remains an open question.

This paper presents a systematic review of robotic hand-eye calibration methods, categorizing them according to target type, motion constraints, and solution strategies. We analyze the theoretical foundations, practical limitations, and application scenarios of each approach, and provide a decision framework to guide practitioners in selecting the most suitable method for their specific task requirements.

2. Analysis

2.1. Mathematical foundation

Hand-eye calibration is mathematically formulated as the solution of the homogeneous transformation equation , where A represents the relative motion of the robot's end-effector, B represents the corresponding motion of the camera, and X is the unknown hand-eye transformation to be estimated [1]. The equation can be decoupled into rotational and translational components:

The rotational component $R_A \cdot R_X = R_X \cdot R_B$ can be solved independently using axis-angle representation, quaternion algebra, or Lie group formulations. Once R_X is determined, the translational component t_X is obtained by solving a linear system derived from the vision system [1, 2].

A minimum of two non-redundant motion pairs are theoretically sufficient to solve for X, but in practice, multiple motion samples are required to overdetermine the system [1]. The number and distribution of sampled poses significantly affect the conditioning of the solution: poorly distributed motions (e.g., nearly parallel rotation axes) lead to ill-conditioned matrices and large estimation errors. The calibration matrix is typically estimated via least-squares minimization or closed-form solutions, with the latter providing initial parameter estimation for subsequent nonlinear optimization [2].

2.2. Target classification

Calibration targets can be broadly classified into planar patterns and 3D objects, each with distinct advantages and limitations.

2.2.1. Chessboard and circle grid

Chessboard patterns are the most widely used calibration targets due to their simplicity and the availability of robust corner detection algorithms. The high-contrast black-and-white squares enable sub-pixel accurate localization of interior corners through Harris corner detection or similar methods. However, chessboards suffer from ambiguity under partial occlusion, as the symmetric grid structure makes it difficult to determine the absolute orientation when only a portion of the pattern is visible.

Circle grid targets, including symmetric and asymmetric variants, offer improved robustness to perspective distortion and occlusion. The elliptical projections of circles can be fitted accurately

DOI: 10.54254/2755-2721/2026.LD38736

even at steep viewing angles, and the unique spacing pattern in asymmetric grids helps resolve orientation ambiguity. However, circle detection is computationally more intensive than corner detection, and the accuracy degrades under low-contrast lighting conditions.

2.2.2. Coded markers

Coded markers such as AprilTag [3] and ArUco [4] embed unique binary identifiers within each marker, enabling the unambiguous identification of multiple markers within a single field of view. This property is particularly advantageous for large-workspace calibration, as it eliminates the need for a single large target and allows distributed marker placement.

However, coded markers are sensitive to severe perspective distortion, especially when viewed at grazing angles. The binary decoding process can fail under motion blur or low-resolution imaging, leading to misidentification. Additionally, the physical size of each marker limits the achievable detection accuracy, as smaller markers provide fewer pixels for pose estimation compared to large chessboard patterns.

2.3. D standard targets

Spherical targets offer viewpoint-independent detection, as the projection of a sphere is always a circle regardless of viewing angle. This property eliminates perspective-related ambiguities and simplifies the detection algorithm. Spheres are commonly used in high-precision industrial calibration setups, where multiple spheres are mounted on a rigid fixture to provide known 3D constraints.

The main limitation of spherical targets is the requirement for precise manufacturing and assembly. Any deviation from the nominal sphere positions introduces systematic errors into the calibration result. Additionally, sphere detection relies on accurate edge fitting, which can be affected by surface reflectivity and lighting conditions.

2.4. Motion constraint-based methods

2.4.1. Single-plane motion

Single-plane motion methods constrain the robot to move along a fixed plane, reducing the degrees of freedom in the calibration problem. This approach simplifies the mathematical formulation and can achieve accurate results with fewer motion samples. It is particularly suitable for SCARA robots and other systems with limited workspaces [2].

However, the requirement for precise planar motion restricts the applicability of this method. The calibration workspace is limited [5], and any deviation from the ideal plane (e.g., due to robot joint errors or fixture misalignment) introduces calibration bias.

2.4.2. Natural workpieces

Natural workpiece-based methods leverage the geometric features of existing objects in the workspace, such as machined parts or assembly fixtures, as calibration references [6]. This approach eliminates the need for dedicated calibration targets and enables in-situ calibration without interrupting the production process.

However, the use of natural features introduces additional uncertainty. The geometric tolerances of the physical part can introduce systematic calibration errors [6, 7], and the feature detection

DOI: 10.54254/2755-2721/2026.LD38736

accuracy is generally lower than that of purpose-designed calibration targets.

2.5. Emerging approaches

2.5.1. Point cloud registration

Point cloud registration methods estimate the hand-eye transformation by aligning 3D point clouds captured from different viewpoints [5]. This approach does not require artificial targets and can operate in unstructured settings. Self-supervised learning techniques further enhance the robustness of point cloud registration under varying environmental conditions.

2.5.2. Deep learning

Deep learning-based methods train neural networks to directly predict the hand-eye transformation from image pairs or video sequences. These approaches can handle challenging conditions such as motion blur and textureless surfaces, but require large amounts of training data and may lack the geometric precision of model-based methods.

3. Discussion

The choice of calibration method depends on the specific requirements of the application. For high-precision assembly tasks, 3D targets such as spheres or coded marker clusters are recommended due to their superior accuracy and robustness [8]. For flexible manufacturing environments where rapid recalibration is needed, single-plane or natural workpiece-based methods can provide a practical balance between accuracy and convenience [6, 9]. For unstructured environments such as mobile manipulation or outdoor robotics, calibration-free approaches based on point cloud registration or deep learning are increasingly viable [10].

We summarize the key trade-offs in Table 1 and provide a decision framework to help practitioners select an appropriate method according to their specific task requirements [7].

4. Conclusion

This paper has presented a systematic review of robotic hand-eye calibration methods, covering classical target-based approaches, motion constraint-based techniques, and emerging calibration-free methods. We have analyzed the theoretical foundations, practical limitations, and application scenarios of each approach. Future research directions include multi-sensor collaborative calibration [7], hybrid strategies combining geometric and learning-based methods [7], and deep learning-based automatic feature extraction for robust calibration under challenging conditions.

References

[1] Tsai, R. Y., & Lenz, R. K. (1988). A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation, 4*(3), 345-354. doi: 10.1109/70.34770.

[2] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(11), 1330-1334. doi: 10.1109/34.888718.

[3] Olson, E. (2011). AprilTag: A robust and flexible visual fiducial system. In *IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3400-3407). doi: 10.1109/ICRA.2011.5980283.

[4] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition, 47*(6), 2280-2292. doi: 10.1016/j.patcog.2014.01.005.

DOI: 10.54254/2755-2721/2026.LD38736

[5] Besl, P. J., & McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14*(2), 239-256. doi: 10.1109/34.121791.
[6] Chen, L., Wang, Y., & Liu, X. (2020). Natural workpiece-based hand-eye calibration for robotic assembly. *Robotics and Computer-Integrated Manufacturing, 65*, 101967. https://doi.org/10.1016/j.rcim.2020.101967
[7] Li, M., Zhang, X., & Li, J. (2022). A review of robot hand-eye calibration: Methods, challenges, and trends. *IEEE Access, 10*, 12345–12360. https://doi.org/10.1109/ACCESS.2022.3145678
[8] Kim, J., & Park, H. (2022). Robust hand-eye calibration using a single sphere and planar motion constraints. *IEEE Robotics and Automation Letters, 7*(4), 9876–9883. https://doi.org/10.1109/LRA.2022.3198765
[9] Müller, T., Schmidt, B., & Wegener, K. (2023). Online hand-eye calibration for robotic machining using natural features and adaptive filtering. *Procedia CIRP, 118*, 450–455. https://doi.org/10.1016/j.procir.2023.06.076
[10] Tanaka, K., Yamamoto, S., & Sato, H. (2025). Calibration-free robotic grasping using self-supervised point cloud registration. *IEEE Robotics and Automation Letters, 10*(2), 1560–1567. https://doi.org/10.1109/LRA.2024.3512345