Super-Resolution-Enhanced Binocular Vision for Accurate Positioning of Prefabricated Wall Panels via Multi-Scale Ellipse Center Detection

Lizhi Long[1] Honghu Chu[3] Lu Deng[1,*] and Weiwei Chen[2]

[1] College of Civil Engineering, Hunan University, China
[2] Bartlett School of Sustainable Construction, University College London, UK
[3] College of Civil Engineering and Architecture, Zhejiang University of Water Resources and Electric Power, China
longlizhi2lxl@hnu.edu.cn, denglu@hnu.edu.cn, weiwei.chen@ucl.ac.uk, chuhh@zjweu.edu.cn

Abstract

Reliable on-site positioning is essential for robotic installation of autoclaved lightweight concrete (ALC) wall panels. A significant challenge arises from the use of close-range wide-angle binocular vision, which frequently produces marker images suffering from low spatial resolution and low contrast. These degraded marker images increase positioning errors, which propagate through ellipse-center positioning, stereo triangulation, and 6D pose recovery. This paper proposes an enhanced binocular vision-based positioning measurement framework that integrates super-resolution reconstruction (SRR) with a multi-scale ellipse center detection (MSECD) strategy. The method first restores marker edge details using an SRR model trained with recursive network, then detects marker to constrain positioning region, and finally estimates sub-pixel ellipse centers by cross-scale detection and ellipse merging. The recovered centers are triangulated to obtain 3D coordinates, enabling robust pose estimation and deviation reporting for alignment guidance. Laboratory experiments show millimeter-level positioning accuracy within a 3 mm tolerance, and the proposed framework reduces average absolute coordinate errors by up to 29.7%, 21.6%, and 8.5% along the x-, y-, and z-axes, respectively, compared with a bicubic upsampling baseline.

Keywords –

prefabricated wall panels; binocular stereo vision; super-resolution reconstruction; ellipse center detection; pose estimation; positioning measurement

1 Introduction

Autoclaved lightweight concrete (ALC) wall panels are increasingly used to accelerate prefabricated construction [1]. During installation, panels must be into a target pose under strict geometric tolerances to prevent connector interference and cumulative layout drift. Although semi-automated machines are deployed for panel grasping and transportation [2], the critical task of precise alignment and final pose adjustment still heavily relies on manual intervention. This human-in-the-loop positioning step is inherently operator-dependent, increasing the risk of misalignment, which compromises the placement precision of ALC panels and ultimately reduces overall construction efficiency.

Although laser-based metrology can deliver high accuracy, it is often limited by cost and deployment complexity constraints on construction sites [3–5]. Conversely, Marker-based binocular vision systems offer a cost-effective and easily integrated alternative for measuring the assembly pose of precast concrete components [6,7]. Given that the surfaces of precast concrete elements are typically texture-less, extracting reliable natural features for assembly guidance is inherently difficult. To address this, existing approaches employ artificial circular markers to create prominent reference features for component pose estimation. These markers are pre-installed on surface of both the component and the existing structure to define their respective coordinate systems. This setup simplifies the assembly pose to determining the rotation and translation between the two coordinate systems (see Figure 2) [7]. In this design, the large circle serves as the primary target for positioning, while the adjacent small circle encodes unique identification. While recent binocular-vision alignment systems, such as the trajectory monitoring method by Cheng et al. [6] and the deep learning-based pose estimation framework by Long et al. [7], significantly enhance automation and accuracy under standard viewing conditions, they face severe performance degradation when applied to large-scale ALC wall panels in confined spaces. Specifically, close-range assembly scenarios necessitate cameras with extremely wide fields of view (short-focal-length lenses). Consequently, markers suffer from severe optical degradation, appearing as small, blurry ellipses with weak edges and low contrast. Direct application of the single-scale ellipse-fitting algorithms utilized in prior works [8,9] to these degraded images results in highly unstable center localization and unacceptable 6D pose estimation errors.

To address this bottleneck, this paper proposes a super-resolution-enhanced positioning framework that strengthens the target detection-to-pose estimation chain for ALC wall panels positioning. Specifically, the conventional stereo pipeline is upgraded into four steps: (1) apply super-resolution reconstruction (SRR) to recover edge details of small fiducials; (2) detect markers to obtain regions of interest (ROIs) for constrained processing; (3) localize ellipse centers within ROIs using a multi-scale ellipse center detection (MSECD) strategy that is robust to small targets and degraded edges; (4) triangulate center points for 3D reconstruction and perform 6D pose estimation and deviation computation. The main contributions are: (i) a unified SRR–detection–MSECDs pipeline tailored to enhance the accuracy of positioning measurement accuracy rather than merely improving image quality; (ii) a cross-scale ellipse-center positioning scheme that provides sub-pixel centers under low-resolution wide-FoV imaging; and (iii) concise validations in laboratory demonstrating millimeter-level positioning, establishing a vital baseline accuracy requirement for future on-site ALC installation under complex field conditions.

2 Related Works

To overcome the accuracy degradation of binocular stereo vision in close-range, wide-FoV, and low-resolution settings, this study builds on two key research streams: super-resolution reconstruction (SRR) and ellipse center detection. SRR is used to enhance image details and preserve marker-edge geometry without costly hardware upgrades, while robust ellipse center detection is essential for reliable marker-based 6D pose estimation under perspective distortion and construction-site disturbances.

2.1 Super-Resolution Reconstruction (SRR)

High-precision vision measurement traditionally relies on high-end, high-resolution cameras, which are expensive and difficult to deploy. SRR offers a software alternative by reconstructing high-resolution (HR) images from low-resolution (LR) inputs, improving edge clarity and feature fidelity [10]. Deep learning SRR methods are dominant today and are typically CNN-based or GAN-based [11]. However, measurement-oriented SRR requires not only perceptual sharpness but also geometric integrity and continuous, well-defined edges, as reconstruction artifacts can directly affect subsequent fitting and positioning [12]. Although SRR has been shown to improve deformation/displacement measurement and stereo-based metric estimation, its impact on ellipse-based marker recognition and center positioning, especially for small, low-contrast targets, remains insufficiently studied [13,14]. This motivates the adoption of deep SRR to improve marker visibility and boundary quality in wide-FoV, short-focal-length imaging.

2.2 Ellipse Center Detection

Circular artificial markers are widely used in the positioning of prefabricated components, but they commonly appear as ellipses in camera images due to perspective effects. In ALC wall panel installation, wide-angle close-range imaging often yields small, blurry, low-contrast ellipses, further challenged by occlusion, uneven lighting, and cluttered backgrounds, leading to missed detections and inaccurate centers positioning [7,15]. Existing methods are broadly voting-based (e.g., Hough variants) or contour-based (edge tracking and arc merging) [16–18]. Voting-based methods are sensitive to noise and small targets, whereas contour-based methods often require multiple hand-tuned parameters and generalize poorly in complex scenes [19]. To address these limitations, this study develops a multi-scale ellipse center detection (MSECD) strategy with cross-scale feature fusion and local structure awareness, enabling more reliable center positioning for high-precision pose estimation in controlled settings, thereby laying a solid groundwork for future applications in complex real-world installation environments.

3 Methodology

Our system takes synchronized stereo images captured by a binocular camera rig mounted on installation equipment. Circular targets attached to the panel serve as fiducials. The processing chain is designed to output two types of information: (i) the 3D positions of markers in the reference frame, and (ii) the panel pose and its deviation from the desired installation pose.

Figure 1 outlines the workflow. First, the raw stereo images are enhanced by super-resolution reconstruction to improve contour readability. Second, marker regions are localized, and each region is processed by MSECD to obtain a stable sub-pixel ellipse center. Third, matched left-right centers are triangulated to recover 3D marker coordinates, which are then used for rigid pose estimation and deviation reporting.

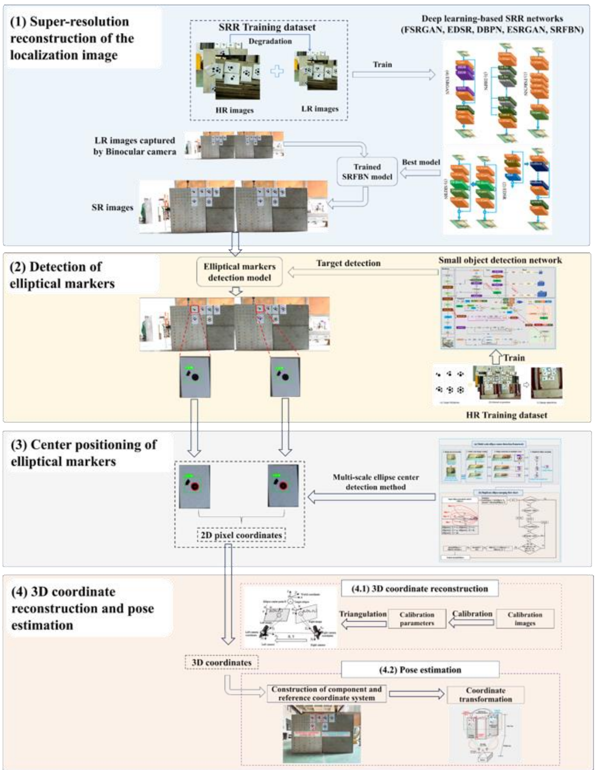


Figure 1. Flowchart of the research framework

3.1 Measurement-Oriented Super-Resolution Reconstruction

The goal of SRR module is to restore boundary details of small elliptical markers captured by short-focal-length lenses, thereby stabilizing downstream center fitting. To this end, representative CNN-based SRR models (FSRCNN [20], EDSR [21], DBPN [22], SRFBN [14]) and a GAN-based model (ESRGAN) are evaluated. Considering the trade-off between reconstruction quality and inference efficiency, the Super-Resolution Feedback Network (SRFBN) is selected as the default SRR backbone.

In practice, The SRR model is trained using paired high-and low-quality marker images synthesized via blurring, downsampling and noise injection. A feedback-based network is utilized due to its strong edge recovery and stable generalization. Ultimately, the SRR output is fed to the subsequent detection and ellipse center detection modules.

3.2 Marker Detection and ROI Extraction

Given the sensitivity of ellipse center detection algorithms to background edge clutter, the search space is constrained by initially detecting marker Regions of Interest (ROIs). To this end, a lightweight object detector named YOLO-ODCA [7], as shown in Figure 3, is trained on marker images collected under varying viewing distances, angles, and illumination conditions. In addition to standard convolutional blocks, the detector integrates two upgrades to improve small-target recognition: (i) the coordinate attention (CA) mechanism to enhance informative channels and suppress background noise, and (ii) an Omni-Dimensional Convolution (ODConv) design to better model edge patterns at different orientations. Subsequently, the detected bounding boxes are expanded by a small margin to retain complete marker boundaries, and each ROI is forwarded to the center-positioning module.

Detection performance is evaluated using precision (P), recall (R), and F1-score, where F1 = 2PR/(P+R), as well as mAP@0.5:0.95. These metrics serve to validate the robust extraction of ROIs, which is critical for minimizing background-induced errors during the subsequent ellipse center detection stage.

3.3 Multi-Scale Ellipse Center Detection (MSECD)

The goal of MSECD strategy is to derive a reliable ellipse center even when the original contour is incomplete or poorly contrasted. Instead of relying on a single fitting operation at the native scale, geometric consistency is enforced across a multi-resolution hierarchy.

Given a detected ROI, a pyramid of re-scaled patches (e.g., ×1, ×2, ×3) are generated. At each scale, edges are extracted and an ellipse is fitted using a direct least-squares (DLS) estimator. Each fitted center is then mapped back to the original image coordinates.

To finalize the estimation, a consensus fusion mechanism is applied: outliers defined as centers deviating from the median beyond a threshold proportional to the ROI size are rejected, and the remaining candidates are averaged. This cross-scale fusion effectively suppresses scale-specific quantization noise and stabilizes the final sub-pixel center estimate.

MSECD takes as input an ROI cropped around a target. Its output is the ellipse center in the original image coordinates.

(1) Multi-scale generation: create a small set of resized patches $\{I_k\}$ at predefined scales S_k (e.g., 1×, 2×, 3×).

(2) Edge extraction and ellipse center detection: for each I_k, detect edges and fit an ellipse using a direct least-squares estimator; obtain center C_k.

(3) Back-projection: map each C_k back to the original ROI using the inverse scaling transform.

(4) Cross-scale fusion: compute a robust consensus center by merging multi-scale candidates. Detections are grouped and averaged if their spatial Euclidean distance is strictly within a defined clustering threshold (empirically set to 10 pixels). Isolated candidates failing to cluster are treated as outliers and rejected.

The robustness of the MSECD module depends fundamentally on the scale factors and the spatial clustering threshold. Since the input ROI has already been enhanced by the SRR module, limiting the scale factors to {1×, 2×, 3×} prevents the exponential computational overhead associated with higher scales, which typically offer diminishing returns due to interpolation blur. Furthermore, sensitivity tests on the merging threshold indicate that a 10-pixel radius provides an optimal stability region. A threshold that is too strict (e.g., < 3 pixels) fails to absorb the natural sub-pixel coordinate shifts induced by multi-scale resizing, while a threshold that is too loose (e.g., > 20 pixels) introduces the risk of incorrectly merging adjacent background noise or distinct targets. This balanced thresholding ensures stable sub-pixel center localization even under degraded imaging conditions.

3.4 Stereo Triangulation and Panel Pose Estimation

With ellipse centers extracted in both views, left-right correspondences are established based on epipolar constraints and marker IDs when available. 3D marker coordinates are obtained by stereo triangulation after camera calibration, which can be shown in Figure 2.

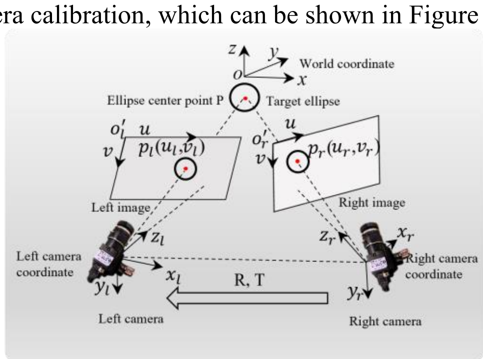


Figure 2. Binocular vision model

As shown in Figure 3, the panel pose is estimated by aligning the reconstructed 3D marker set to its as-designed coordinates using a rigid transformation solved in least squares. This yields the panel position and orientation in the reference coordinate system.

To support on-site adjustment, deviation metrics specifically translational offsets along and rotational errors around the installation axes are computed. These outputs can be directly used by operators or automated controller to guide fine alignment.

All geometric computations are performed in calibrated camera coordinate systems. Intrinsic parameters and distortion coefficients are obtained offline using a standard planar calibration procedure. For stable triangulation, stereo image pairs are rectified to align corresponding points along identical polar line.

After rectification, a 3D marker point can be recovered from its pixel correspondence (u_l, v) and (u_r, v) using the disparity $\delta = u_l - u_r$. With focal length f and baseline B, the depth is calculated as $Z = f\,\mathrm{B} / \delta$, while the lateral coordinates are derived via $X = (u_l - c_x) \cdot Z / f$ and $Y = (v - c_y) \cdot Z / f$, where (c_x, c_y) denotes the principal point. These closed-form relations provide an efficient implementation and facilitate error propagation analysis.

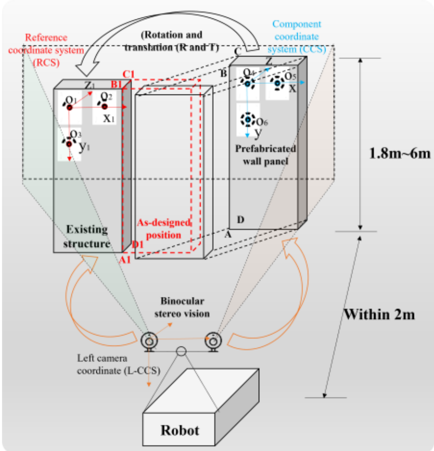


Figure 3. Pose estimation of the panel component

4 Experimental Validation

To demonstrate the effectiveness of the proposed framework, a comprehensive positioning experiments was conducted using precast concrete wall panels. The performance of the proposed ALC positioning method were evaluated at four levels: (1) Super-resolution reconstruction for elliptical marker; (2) Elliptical marker center detection; (3) Coordinate measurement of elliptical markers; (4) Pose estimation of ALC panels.

4.1 Experiment Setup

To rigorously validate the proposed framework, a measurement experiment was conducted using a precast concrete wall panel. As depicted in the experimental setup (see Figure 4), reference markers were affixed to the panel surface to establish the Component Coordinate System (CCS), while additional test markers served as validation targets. Ground-truth 3D coordinates for all markers were acquired using a high-precision Total Station to eliminate layout deviations. The vision system's measurements were transformed into the CCS, and absolute errors relative to the ground truth were computed along the x, y, and z axes.

The evaluation compared three image processing strategies: (1) Original High-Resolution (HR) images as the benchmark, (2) Bicubic interpolation (representing standard upsampling), and (3) the proposed SRFBN-based super-resolution. Tests were executed at four discrete distances (3 m, 4 m, 5 m, and 6 m) to simulate typical on-site installation ranges.

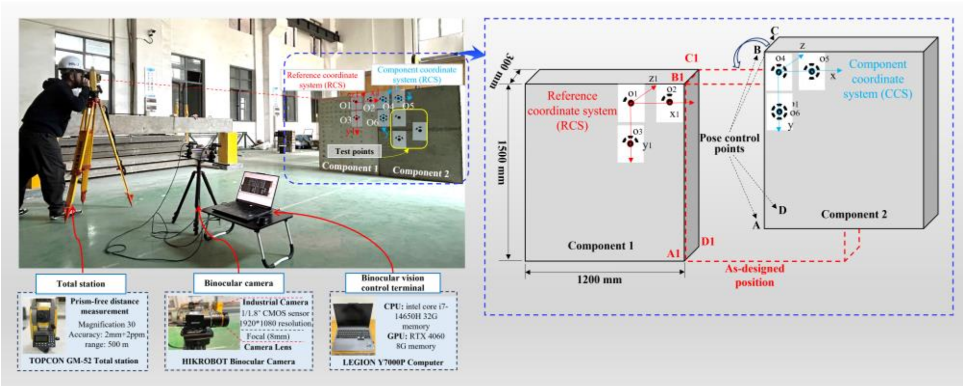


Figure 4. Binocular vision position measurement system

4.2 Datasets

To train the target detector and the SRR models, we constructed a comprehensive dataset with strict separation between training and testing scenes to accurately evaluate the method's generalization capability.

Marker Detection Dataset: We collected 1,100 high-resolution (1920 × 1080) images of artificial markers affixed to prefabricated wall panels under various viewing angles and distances. To ensure strict evaluation, the dataset was split into 800 images for training, 200 for validation, and a strictly separated set of 100 images for testing. These 100 test images were collected from independent prefabricated panel backgrounds to prevent data leakage. All images were cropped to 960 × 960 pixels and manually annotated. During training, the Mosaic data augmentation strategy was adopted to enrich dataset diversity and improve the model's robustness in detecting small elliptical markers under complex backgrounds.

SRR Dataset: To train the deep learning-based SRR models, we combined 1,000 natural scene images from the DIV2K dataset with 1,000 images from our HR-Marker dataset, yielding 2,000 high-resolution (HR) images. This combined dataset was randomly split at an 8:2 ratio, yielding 1,600 images for training and 400 for validation. Data augmentation techniques including horizontal and vertical flipping, as well as rotations of 90°, 180°, and 270°, were applied to expand the training set to 8000 images.

To evaluate the SRR performance under practical conditions, the strictly separated 100-image test set from the marker dataset was utilized. Due to the difficulty of acquiring paired real low-resolution (LR) images, LR inputs were synthetically generated from these HR test images using a standard image degradation model. The degradation process involves a 4× downsampling (using bicubic interpolation) combined with a Gaussian blur kernel (standard deviation of 1.6) and noise injection (intensity of 5) to simulate realistic on-site degradation such as motion blur and sensor noise.

4.3 Super-Resolution Reconstruction for Elliptical Marker

The SRR stage directly improves the observability of marker boundaries. Table 1 summarizes the quantitative performance of the six reconstruction methods in terms of PSNR, SSIM, and inference time. The results demonstrate that all deep learning-based architectures significantly outperform the traditional Bicubic interpolation baseline. The substantial improvements in PSNR and SSIM validate the superior capability of deep neural networks in restoring the structural fidelity and texture details essential for accurate marker positioning. Among these models, the feedback-based SRR model yield higher PSNR/SSIM on degraded marker patches.

Table 1. Representative SRR results for marker enhancement

Model	PSNR (dB)	SSIM (%)	Inference time (s/img)
Bicubic	27.94	76.13	87.21
FSRCNN	33.11	86.70	1.390
EDSR	33.45	87.19	1.310
ESRGAN	29.97	80.76	0.410
DBPN	33.70	87.48	1.510
SRFBN	33.72	87.53	0.610

4.4 Elliptical Marker Center Detection

The specific roles of SRR and MSECD were evaluated through an ablation study (i) baseline processing on low-quality images, (ii) bicubic upsampling with single-scale fitting, (iii) SRR with single-scale fitting, and (iv) SRR with MSECD. Detection performance was evaluated on a randomized subset of 100 low-resolution images by quantifying the false detection rate and miss detection rate across the four comparative methods. As detailed in Table 2, untreated low-resolution inputs exhibit the highest false and miss detection rates for artificial elliptical markers. While Bicubic interpolation yields a reduction in both false and miss detection rates, the improvement is marginal. In contrast, images enhanced by the SRFBN model demonstrate a significant increase in the accurate detection rates. However, the false detection rate remained elevated, the super-resolution process amplified edge details, causing nearby fan-ring structures to be misclassified as marker ellipses. This issue is effectively mitigated by the proposed MSECD strategy, which significantly reduces false detection rate through the application of multi-scale constraints and iterative ellipse merging.

Table 2. Qualitative ablation summary of module contributions

Variant	SRR	MSE CD	False detection rate (%)	Miss detection rate (%)
Raw + single-scale fit	–	–	30.1	62.3
Bicubic + single-scale fit	–	–	35.4	56.2
SRR + single-scale fit	✓	–	24.2	14.5
SRR + MSECD (proposed)	✓	✓	2.3	1.6

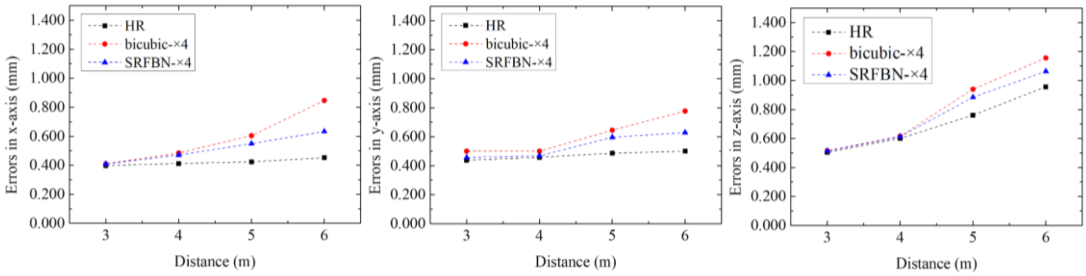
It is important to note that the high false detection rate (24.2%) observed in the "SRR + single-scale fit" method directly quantifies the negative propagation of SRR artifacts. While deep SRR significantly improves edge contrast, it occasionally introduces localized geometric distortions or 'hallucinated' edge pixels. When fed into a standard single-scale fitting algorithm, these artifacts inevitably cause sub-pixel shifts in the detected center. The proposed MSECD strategy effectively acts as a geometric low-pass filter by enforcing consensus across multiple image scales, it suppresses these localized hallucinated textures, preventing SRR-induced pixel shifts from propagating downstream.

4.5 Coordinate Measurement of Elliptical Markers

The quantitative results for coordinate measurement accuracy are summarized in Figure 5. The analysis reveals two critical findings:

First, a correlation exists between measurement distance and error magnitude, particularly along the z-axis (depth). This phenomenon is inherent to binocular vision, where depth is inferred from disparity; as distance increases, disparity decreases, amplifying the sensitivity of depth estimation to pixel-level localization errors.

Second, regarding image reconstruction strategies, the Bicubic method exhibits significant performance degradation at longer distances (e.g., 6 m). This is attributed to the blurring of marker edges in low-resolution images, which compromises the stability of ellipse center fitting. In contrast, the SRFBN-based approach, coupled with the MSECD strategy, effectively restores high-frequency edge details while actively suppressing hallucinated distortions. This prevents SRR artifacts from translating into disparity errors during stereo matching, thereby stabilizing the sub-pixel center estimation and keeping depth (z-axis) triangulation errors tightly bounded. Consequently, the proposed method maintains coordinate measurement errors within 1 mm across the full 6-meter range, achieving an accuracy profile closely approximating that of the HR benchmark and satisfying the requirements for precision assembly.



(a) (b) (c)

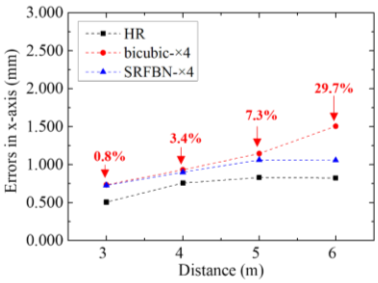
Figure 5. Correlation between coordinate measurement error and camera distance for the binocular cameras: (a) Measurement errors in x-axis, (b) Measurement errors in y-axis, (c) Measurement errors in z-axis

4.6 Pose Estimation of ALC Panels

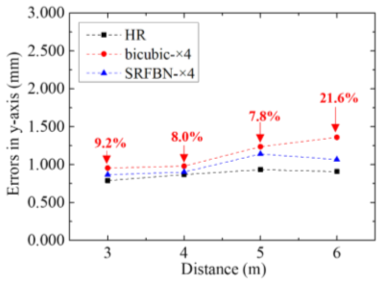
To comprehensively assess the final assembly precision, the experiment was extended to evaluate the 6-DoF pose estimation capability. This involved determining the deviation of the wall panel relative to its design pose by aligning the Component Coordinate System (CCS) with the Reference Coordinate System (RCS). As illustrated in Figure 7, the panel was manually positioned such that its four corner control points (A, B, C, D) aligned with their corresponding design references. Given the practical constraints of physical alignment, the actual spatial coordinates of the control points were measured using a high-precision Total Station to establish the ground truth. The system performance was then evaluated by localizing these points using High-Resolution (HR), Bicubic, and SRFBN-enhanced images across varying distances (3–6 m).

As shown in Figure 6, the proposed framework demonstrates robust performance. While errors naturally increase with distance, the SRFBN model significantly mitigates the precision loss observed in standard upsampling. Specifically, at the maximum tested distance of 6 m, the SRFBN reconstruction reduces average absolute errors by 29.7%, 21.6%, and 8.5% along the x, y, and z axes, respectively, compared to the Bicubic baseline. Ultimately, the maximum positioning error remains below 2 mm, well within the stringent 3 mm tolerance required for prefabricated ALC wall panel installation.

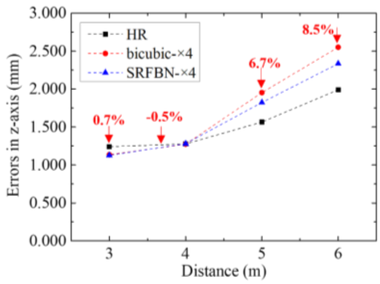
Beyond positioning accuracy, the computational efficiency of the complete pipeline was evaluated on a desktop workstation equipped with a GeForce RTX 3090 (24 GB VRAM). Processing a single pair of stereo images sequentially involves: SRR enhancement (0.610 seconds), YOLO-ODCA marker detection and ROI extraction (0.029 seconds), MSECD center localization (0.040 seconds), and 3D triangulation with pose estimation (0.005 seconds). The total end-to-end processing time is approximately 0.684 seconds per frame. Given that the final fine-alignment of heavy ALC wall panels is inherently a low-speed process to ensure construction safety, this processing latency (operating at roughly 1.5 Hz) fully satisfies the requirements for quasi-real-time positioning feedback during on-site robotic installation.



(a)



(b)



(c)

Figure 6. Correlation between coordinate measurement error and camera distance for the binocular cameras: (a) Measurement errors in x-axis, (b) Measurement errors in y-axis, (c) Measurement errors in z-axis

5 Conclusions

This paper presented a measurement-oriented enhancement framework for marker-based binocular positioning in close-range wide-FoV installation scenarios. By combining SRR for edge recovery, ROI-based marker detection, and a multi-scale ellipse center detection strategy, the method stabilizes the entire measurement pipeline from ellipse-center positioning and stereo triangulation to final 6D pose estimation. Experimental results in laboratory demonstrate millimeter-level positioning accuracy at an acceptable end-to-end processing latency of approximately 0.684 seconds, verifying its suitability for the fine-alignment stage of ALC wall panel installation.

While the current laboratory experiments demonstrate the baseline millimeter-level accuracy of the framework, its robustness against complex field challenges (such as dynamic occlusion, heavy dust, uneven lighting, and severe equipment vibration) remains untested. Future work will prioritize rigorous on-site trials to evaluate and enhance the system's performance in real-world construction environments, alongside investigating tighter integration of SRR and center positioning to further reduce computation and to improve robustness under severe motion blur.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Grant Number 2023YFC3806800), the Zhejiang Provincial Natural Science Foundation (Grant No. QN26E080027), the Hangzhou Municipal Natural Science Foundation (Grant No. 2025SZRJJ1193), the Horizon Europe Project, D-HYDROFLEX (Grant Number 101122357), the Horizon Europe Project, INHERIT (Grant Number 101123326).

References

[1] M. Kalpana, S. Mohith, Study on autoclaved aerated concrete: Review, Mater Today Proc 22 (2020) 894–896. https://doi.org/10.1016/j.matpr.2019.11.099.

[2] S. Li, X. Zhou, G. Cheng, W. Teng, Y. Zeng, G. Wei, Y. Chen, A scheme of installing ALC wall panels based on autonomous mobile robot, Journal of Building Engineering 94 (2024) 109991. https://doi.org/10.1016/j.jobe.2024.109991.

[3] D. Kim, Y. Kwak, H. Sohn, Accelerated cable-stayed bridge construction using terrestrial laser scanning, Autom Constr 117 (2020) 103269. https://doi.org/10.1016/j.autcon.2020.103269.

[4] Y. Xu, Y. Luo, J. Zhang, Laser-scan based pose monitoring for guiding erection of precast concrete bridge piers, Autom Constr 140 (2022) 104347. https://doi.org/10.1016/j.autcon.2022.104347.

[5] S. Yoon, Q. Wang, H. Sohn, Optimal placement of precast bridge deck slabs with respect to precast girders using 3D laser scanning, Autom Constr 86 (2018) 81–98.

[6] Y. Cheng, F. Lin, W. Wang, J. Zhang, Vision-based trajectory monitoring for assembly alignment of precast concrete bridge components, Autom Constr 140 (2022) 104350. https://doi.org/10.1016/j.autcon.2022.104350.

[7] L. Long, J. Guo, H. Chu, S. Wang, S. Xu, L. Deng, Binocular vision-based pose monitoring technique for assembly alignment of precast concrete components, Advanced Engineering Informatics 65 (2025) 103205. https://doi.org/10.1016/j.aei.2025.103205.

[8] G. Carbone, A.S. Gurtatta, D. Malyshev, Development and implementation of a hybrid visual prediction algorithm for robotic smart tomato harvesting, Eng Appl Artif Intell 161 (2025) 112261. https://doi.org/10.1016/j.engappai.2025.112261.

[9] Q. Chen, J. Wen, T. Lin, H. Ren, Target recognition and localization of environmental perception system based on binocular cameras for unmanned cleaning vehicle, Eng Appl Artif Intell 143 (2025) 109994. https://doi.org/10.1016/j.engappai.2024.109994.

[10] H. Chu, J. Gai, W. Chen, J. Ma, CBRFormer: rendering technology-based transformer for refinement segmentation of bridge crack images, Advanced Engineering Informatics 69 (2026) 103868. https://doi.org/10.1016/j.aei.2025.103868.

[11] D.C. Lepcha, B. Goyal, A. Dogra, V. Goyal, Image super-resolution: A comprehensive review, recent trends, challenges and applications, Information Fusion 91 (2023) 230–260. https://doi.org/10.1016/j.inffus.2022.10.007.

[12] C. Xiang, W. Wang, L. Deng, P. Shi, X. Kong, Crack detection algorithm for concrete structures based on super-resolution reconstruction and segmentation network, Autom Constr 140 (2022) 104346. https://doi.org/10.1016/j.autcon.2022.104346.

[13] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C.C. Loy, Y. Qiao, X. Tang, ESRGAN: Enhanced super-resolution generative adversarial networks, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018: pp. 0–0. http://arxiv.org/abs/1809.00219.

[14] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, W. Wu, Feedback network for image super-resolution, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2019: pp. 3862–3871. https://doi.org/10.1109/CVPR.2019.00399.

[15] M. Fornaciari, A. Prati, R. Cucchiara, A fast and effective ellipse detector for embedded vision applications, Pattern Recognit 47 (2014) 3693–3708. https://doi.org/10.1016/j.patcog.2014.05.012.

[16] D.K. Prasad, M.K.H. Leung, S.-Y. Cho, Edge curvature and convexity based ellipse detection method, Pattern Recognit 45 (2012) 3204–3221. https://doi.org/10.1016/j.patcog.2012.02.014.

[17] R.A. McLaughlin, Randomized Hough Transform: Improved ellipse detection with comparison, Pattern Recognit Lett 19 (1998) 299–305. https://doi.org/10.1016/S0167-8655(98)00010-5.

[18] J. Wu, Robust real-time ellipse detection by direct least-square-fitting, in: 2008 International Conference on Computer Science and Software Engineering, IEEE, 2008: pp. 923–927. https://doi.org/10.1109/CSSE.2008.789.

[19] C. Lu, S. Xia, M. Shao, Y. Fu, Arc-support line segments revisited: An efficient high-quality ellipse detection, IEEE Transactions on Image Processing 29 (2019) 768–781.

[20] C. Dong, C.C. Loy, X. Tang, Accelerating the super-resolution convolutional neural network, in: European Conference on Computer Vision, Springer International Publishing, 2016: pp. 391–407. https://doi.org/10.1007/978-3-319-46475-6_25.

[21] B. Lim, S. Son, H. Kim, S. Nah, K.M. Lee, Enhanced deep residual networks for single image super-resolution, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, IEEE, 2017: pp. 1132–1140. https://doi.org/10.1109/CVPRW.2017.151.

[22] M. Haris, G. Shakhnarovich, N. Ukita, Deep back-projection networks for super-resolution, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018) 1664–1673. http://arxiv.org/abs/1803.02735.