



Mobile projective augmented reality for collaborative robots in construction

Siyuan Xiang, Ruoyu Wang, Chen Feng^{*}

Tandon School of Engineering, New York University, Brooklyn, NY 11201, USA



ARTICLE INFO

Keywords:

Mobile projective AR (MPAR)
Camera-projector system
Pose estimation
Construction co-robots

ABSTRACT

Augmenting virtual construction information directly in a physical environment is promising to increase onsite productivity and safety. However, it has been found that using off-the-self augmented reality (AR) devices, such as goggles or helmets, could potentially cause more health, safety, and efficiency concerns in complex real-world construction projects, due to the restricted field of view and the non-negligible weight of those devices. To address these issues, we propose a mobile projective AR (MPAR) framework in which the AR device is detached from human workers and carried by one or more mobile collaborative robots (co-robots). MPAR achieves glassless AR that is visible to the naked eye using a camera-projector system to superimpose virtual 3D information onto planar or non-planar physical surfaces. Since co-robots often need to move during the operation, we design algorithms to ensure consistent mobile projection with two major components: projector pose estimation and projection image generation. For planar surfaces, MPAR is achieved by a homography-based pose estimation and image warping. For non-planar surfaces, MPAR uses iterative closest point (ICP) for pose estimation and common graphics pipelines to generate projection images. We conducted both qualitative and quantitative experiments to validate the feasibility of MPAR in a laboratory setting, by projecting 1) as-planned building information onto a planar surface, and 2) as-built 3D information onto a piece-wise planar surface. Our evaluation demonstrated centimeter-level projection accuracy of MPAR from different distances and angles to the two types of surfaces.

1. Introduction

The construction industry has been labeled as hazardous [1,2] and susceptible to casualties and economic losses [3,4]. McKinsey [5] reported that the average annual productivity growth increased only 1% over the past 20 years, lower than 2.8% for the total world economy. Occupational Safety and Health Administration(OSHA) reports that the fatal injury rate in the construction industry is higher than the national average in any other industries in the US [6]. Some researchers believe these issues might be caused by lack of job site information for on-site construction operators [7–9]. To address the low efficiency and safety issues, AR has been applied to the construction industry [10] in different phases.

An AR system can help the collaboration and discussion between different parties, as it provides an interface that allows people to retrieve information when they are making decisions. One potential application is for designers to better communicate with customers about the interior design [11–15]. By displaying the virtual facility or decoration

information generated by AR on a real site, clients can see exactly what a new piece of furniture or design will look like in their home before they purchase it. Lin et al. [16] designed a visualization system using AR technology to display the public building or facility information, facilitating the discussion process for different parties. To improve the interaction between field operators and the constructional information, Kim et al. [17] developed a system that can assist users in determining the optimum scenario for equipment operation by involving them in a shared AR environment.

Also, AR techniques have already been applied to visualize HVAC (heating, ventilation, and air conditioning) information [18–20]. Conventionally, when making HVAC renovation, construction workers take design papers to site and make a careful measurement to ensure they are making renovation in the correct place. With AR techniques, they can see directly what is behind the wall while working, e.g., drilling a hole according to the superimposed virtual information. AR can also assist the construction industry in facilitating employee training [21,22]. With the help of AR training platforms, construction workers

* Corresponding author.

E-mail addresses: siyuan@nyu.edu (S. Xiang), ruoyuwang@nyu.edu (R. Wang), cfeng@nyu.edu (C. Feng).

can be trained with virtual materials, tools, or instructions, without being exposed to some dangerous training scenarios. Therefore, in conjunction with some other applications, such as hazard recognition and avoidance, AR has the potential to improve the safety for the construction industry [10,23–25].

However, current AR technologies have several disadvantages that could be problematic for architectural, engineering, construction, and facility management (AEC/FM). For head-mounted display (HMD) AR like HoloLens, the restricted field of view (FOV) might impede construction worker from obtaining environmental information [26,27] and the caused situational unawareness might result in fatal injuries to field workers [28]. Also, wearing a heavy helmet or goggles for long hours might cause occupational diseases for construction workers. Tablet-based AR can assist engineering processes by collaboratively visualizing computer-generated models [29]. However, one drawback is that construction workers cannot work while holding the tablet, and the inconvenience might limit the application of AR in construction. Also, Yan et al. [30] pointed out that tablet-based AR provides less immersive user experience compared to other AR technologies.

Therefore, one potential solution is to design a mobile augmented reality (MPAR) system for AEC/FM, and Building Information Modeling (BIM) data retrieval, and construction education [31–34]. Besides avoiding these disadvantages caused by HMD AR or tablet-based AR, a portable and mobile AR system itself will not be a burden to construction workers' workload. Also, the user can take the advantage of the MPAR technique to assist construction collaborative robots in human robot interaction [35,36], by observing the context-related virtual objects augmented by the MPAR system.

In this paper, we focus on the design and systematic evaluation of an MPAR system that uses a mobile projector system to blend virtual contents into real scenes while avoiding the problems mentioned above. This work is a continuation of our preliminary research [37]. Our method realizes AR with a camera-projector system and ensures a consistent projection during the working process, even if the projector is moved for various reasons (e.g., to enlarge/move the field of view, to make room for co-workers). The MPAR system benefits users in several aspects: 1. Users can place the MPAR system at arbitrary poses as long as the projection still covers the desired area. This is important because unlike in the classroom or movie theatre where projectors are strategically placed and fixed during the whole projection period, it is often difficult to ensure that the projector could stay static on a construction site with complex terrain or environment. 2. MPAR system is cost-efficient since it is only composed of a camera, a portable projector, and a computational platform such as a laptop. 3. It is easy to mount the MPAR system on a robot, such as an iRobot, such as the example shown in Fig. 1 and our video demonstration at <https://youtu.be/rZE2RcdJ3E>. In this prototype, the camera-projector system serves as the perception and projection module. The pan-tilt is used to control the pose of the camera-projector system. The Nvidia Jetson TX2 is the computational platform that can provide the functionality of computation, rendering, and human-robot interaction. The iRobot is the mobile base for the MPAR system. Such an MPAR system enables human-robot collaboration on construction job sites.

This paper first presents comprehensive applications or research of AR in the construction field, and related methods might be useful for our MPAR work. It then introduces the MPAR system's method in the methodology section. Finally, two experiments using the MPAR system are presented, and the quantitative evaluation work is also implemented to give an analysis for using the MPAR system.

Our contributions are listed as follows:

- We propose MPAR as a method to generate a proper projection image that can ensure that the projection result stays on designed locations, which is of fundamental importance for informational/cognitive collaboration between a collaborative construction robot and its human colleagues.

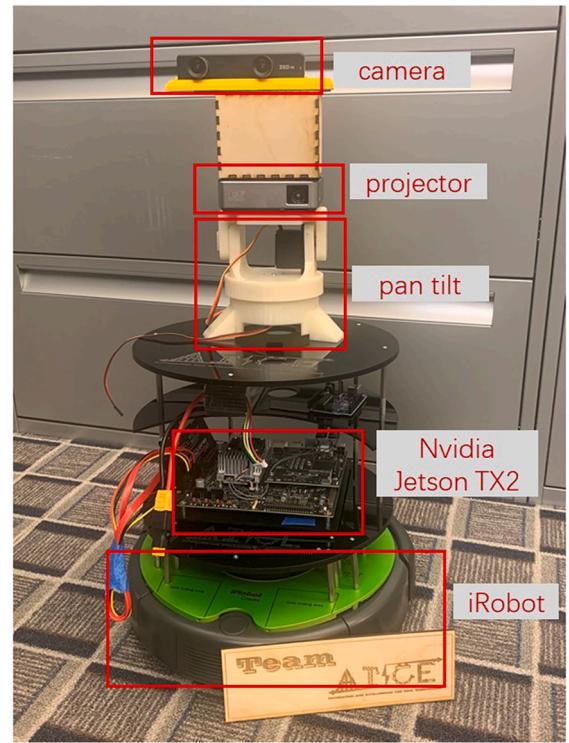


Fig. 1. MPAR prototype design.

- We develop a mobile camera-projector system with both hardware and software that can achieve consistent projection results on both planar and non-planar surfaces during movement.
- We design a systematic way to automatically measure the projection accuracy on both planar and non-planar surfaces, which enables efficient comprehensive evaluation when deploying MPAR in real-world applications.

2. Related work

2.1. BIM and AR

BIM is defined as an intelligent 3D model-based process that can provide professionals with insight and tools in different construction phases. Compared to traditional construction workflow, where various stakeholders need to retrieve building or facility information from scattered 2D drawings, documents, or reports, BIM technology can assist them in obtaining the data fast and accurately with high structured digital models. The advantages of using BIM in the construction industry have been widely acknowledged [38–43]. Since BIM is broadly used in AEC/FM and can provide data as the input of an AR system, some researchers have been focusing on integrating AR technology with BIM. These researches mainly focus on employing AR as a visualization tool to display as-planned or as-built BIM data in the real environment for architectural visualization [44], facility management [45], defect management [46], and other construction-related applications [47–49].

2.2. Projector-based AR

To avoid the limitations mentioned above of using conventional AR devices such as tablets or goggles, we propose using the MPAR system to achieve the virtual content blending in the real world. Projector-based AR is not an entirely new idea, and it has been introduced to many fields already. For example, augmenting details for cultural heritage artifacts [50], improving users' gaming experience [51], or combining head-mounted display (HMD) with view-dependent projection to

achieve immersive AR user experience [52].

Projector-based AR is also used for human-robot interaction. Tatsuno et al. [53] presented the framework that the robot can project the information onto surfaces to provide visual support for human activities. [54] proposed work aiming at improving the robot's ability to communicate with humans by projecting the robot's internal state information on the shared floor. [55] evaluated a mobile robot's communication capability with humans by projecting visual arrows and simplified maps, which indicates the robot's intended movement. For architecture design, projector-based AR has been used for assisting designers in exploring alternative designs and creating an immersive environment [56–58]. [59] proposed a remote-collaboration system that can project the virtual attribute information on real site and send the augmented on-site scene to the remote supporter. Therefore, they can provide useful instructions to the on-site worker, thus improving the on-site worker's work efficiency. However, some of these methods still require wearable devices, which might produce the same problems as using HMD. But our MPAR system only uses the projector as the device to augment the reality so these problems can be avoided. Moreover, to our best knowledge, these methods do not focus on a consistent projection with the mobile projector system, and it could be problematic when we want to achieve as-planned building or facility information augmentation for real construction applications. For example, if the projection content is changed during the movement of the projection system, users cannot retrieve or observe the virtual BIM data that should be projected onto its designed location.

Our work is related to [49] in that the authors also proposed a framework to retrieve building information by projecting 2D images onto the construction site. However, the MPAR system differs from the previous designed system in: 1) our MPAR system does not require users to specify the projector's location; 2) our MPAR system ensures consistent and correct building or facility information projection on their as-planned location; 3) our MPAR system can be used for projection on multiple planes or even non-planar surfaces.

2.3. Projection on non-planar surfaces

The real construction site is complex, and sometimes the projector-based AR applications will be required to project the virtual content onto multiple planes or even non-planar surfaces. In this paper, we treat the projection on a single 2D plane or non-planar surfaces with two different methods, so we will survey some related methods that can benefit the MPAR system regarding two kinds of projection technique. [60–66] focused on projection onto non-planar surfaces. Since the non-planar projection surface structure is irregular, these methods all focused on correcting the geometric distortions caused by the irregular attribute of the surface, to achieve the desired projection effect. These methods proposed different projection geometric correction schemes to compensate for the surface's structure. However, these methods suffer from high computational complexity, and [67] proposed a saliency-guided projection geometric correction method to lower the computation load. Another challenge for projector-based AR is that the projection surface might be changing during different phases. Lindlbauer et al. [68] developed a framework that combines spatial AR with shape-changed interfaces, with the ability to render arbitrary objects and shadows.

3. MPAR method

Mobile projective axis is not perpendicular to the projection surface, the projection result will have distortion; when the projector is moving back and forth towards the projection surface, the size of the projection result will change. So it is essential to generate an appropriate input image for the projector, which will counteract the distortion and scaling issues caused by the projector's inappropriate pose and location. Therefore, the technical challenge of MPAR is to find out such an input

image for a correct projection. In this paper, we describe the geometry of this projection process using the well-known camera projection matrix in 3D computer vision:

$$X_p = K_p [R_w^p, t_w^p; \mathbf{0}, 1] X_w \quad (1)$$

where X_w , X_p are the homogeneous representations of the object points in the world frame and image points in the projector frame respectively, K_p is the intrinsic matrix of the projector, and R_w^p , t_w^p represent the extrinsic matrix of the projector with respect to the world frame. The MPAR method contains two phases: 1) the pose estimation phase to obtain the extrinsic matrix of the projector; 2) the image generation phase to obtain the image points in the projector frame.

In this paper, the projection could be in two cases: 1) projection on a single 2D plane and 2) projection on non-planar surfaces. The first case can be considered as a special case of the second. However, to achieve fast and accurate projection, we used different tracking and image generation methods for the two cases. We will first discuss the MPAR workflow and then explain the details for the two cases, respectively.

The MPAR system workflow overview can be seen in Fig. 2. Once the onsite operators arrive at the construction site, they can decide the projection surface according to the location of the BIM data and set up the MPAR system working environment. The consistent projection can be achieved by two steps: 1) pose estimation, and 2) image generation. In the pose estimation phase, the MPAR system needs to perceive localization information from the environment and compute the projector's pose in the world frame. In the image generation phase, the MPAR system needs to first retrieve the BIM data from the database, then determine whether the projection is happening on a single 2D plane or a non-planar surface, and finally generate the image accordingly. Afterwards, by sending the generated image to the projector, the desired projection result can be achieved. Case 1 and case 2 in Fig. 2 illustrate the workflow of projection on planar surface and non-planar surface. The projection process is completely automatic and efficient: once an operator selects the projection area and set up the MPAR working environment, the system can estimate the projector's pose in the world frame and generate the projection image continuously and autonomously.

3.1. Projector pose estimation

In this section, we will introduce the method to obtain the projector's pose in the world frame. Since the projector itself cannot perceive localization information from the environment, we can use some intermediate devices, like cameras, to help estimate the projector's pose. We formulated our pose estimation module as passing sequential transformations from the world coordinate system to the projector coordinate system through the intermediate devices' poses. In this paper, the transformation from a coordinate system to b coordinate system is represented by a matrix T_a^b . Usually, the transform matrix is composed by a translational vector t and a rotational vector R . w represents the world coordinate system, p represents the projector coordinate system, and c represents the camera coordinate system.

In the MPAR system, we used cameras as intermediate devices to transform the pose. The pose estimation module aims to obtain: 1) calibration results of the camera-projector system, which is the extrinsic matrix between the projector and the camera; 2) the relative 6-DOF pose of the camera in the world frame.

3.1.1. Camera-projector system calibration

Calibration of the camera-projector system is essential for the applications using the MPAR system, and the projection results highly depend on the calibration accuracy. The intrinsic matrices of the camera and the projector K_c , K_p , and the extrinsic matrix T_c^p between them need to be obtained from the calibration. In this paper, we used a structured light based software provided by [69] to calibrate our camera-projector

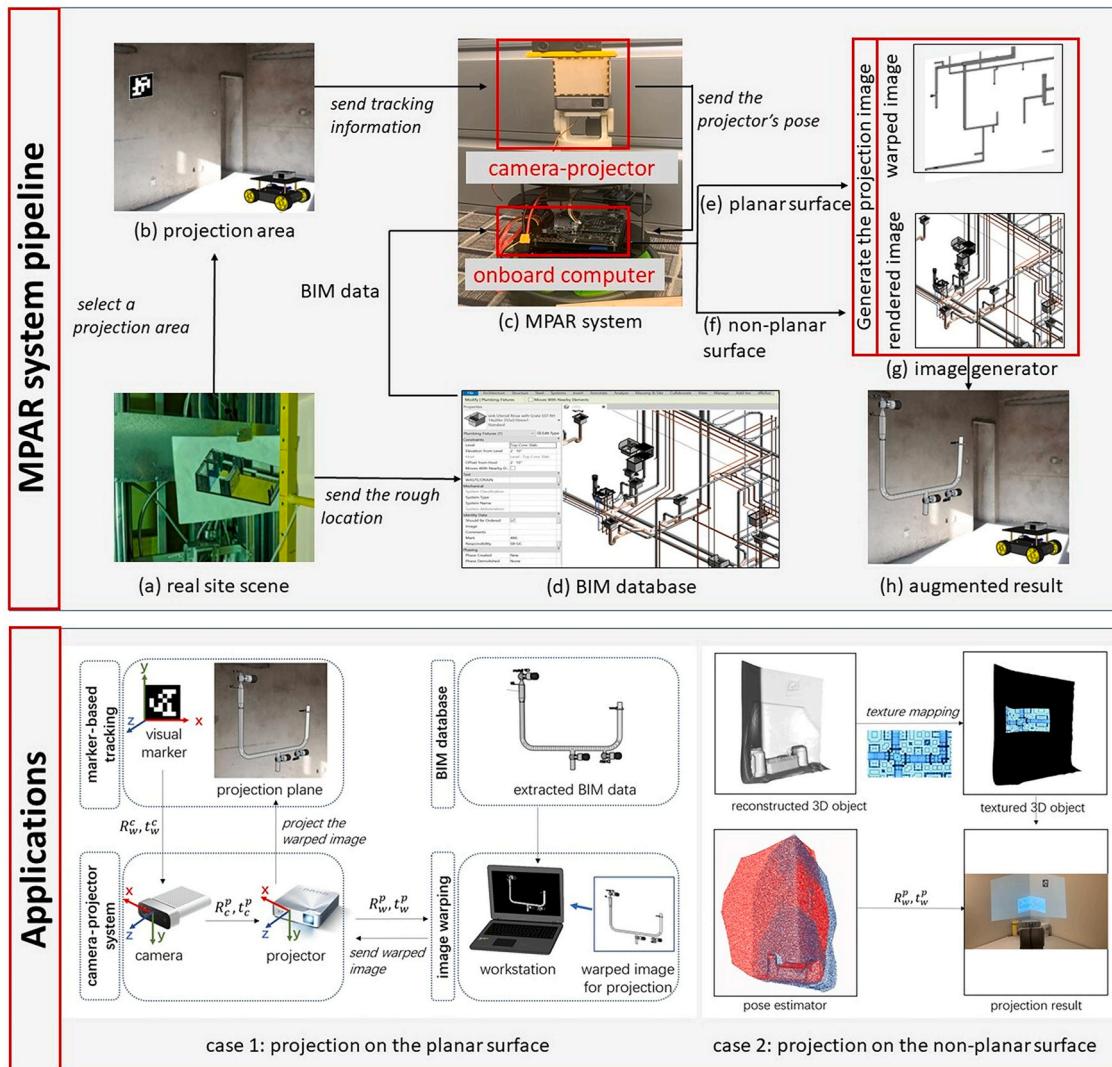


Fig. 2. MPAR system workflow overview.

system. The calibration process is automatic and user-friendly, and the calibration results can satisfy the specific projection accuracy requirement. Similar to most calibration methods, the user needs to print a checkerboard with known length and width for each grid. The user should then place the projector and camera in several positions and ensure: 1) the projection area can cover the checkerboard; 2) the frames of the projector and the camera are correctly overlapped, where it is explained in [69]. After setting up the devices, the user can start the work stream of the software. A series of patterns will be sent to the projector, then displayed on the checkerboard automatically. Meanwhile, the camera will capture an image when there is a new pattern projected onto the checkerboard. These captured images are used for computing the calibration parameters. Note that once the camera-projector system is rigidly-bounded and calibrated, we assume the intrinsic matrices and extrinsic matrix are also fixed. Although the focal lengths might vary slightly depending on different working distances.

3.1.2. Projector pose tracking

Although many research works currently focuses on tracking techniques of AR, it is still challenging to obtain a stable, accurate tracking result with low-latency due to the limitation of the computational power of the AR computing platform.

Kim et al. [70] concluded three tracking technologies, which are widely used in the AR tracking problem: 1) vision-based tracking methods; 2) sensor-based tracking methods; 3) hybrid tracking methods. In the vision-based tracking methods, visual features from captured images are used to determine the location. In contrast, in the sensor-based tracking methods, other features like the structure of the environment can provide unique location information. Hybrid tracking methods use both the visual signals and the sensor perceived features for localization.

In the MPAR system, we used vision-based tracking techniques for projection on a 2D plane. Specially, we used marker-based tracking techniques since 1) the markers can be installed on 2D plane; 2) the marker-based method can provide localization information even when the environment is texture-less. For the projection on non-planar surfaces, we selected the hybrid tracking methods, since the vision-based method can help provide the initial pose of the camera with respect to the world frame, while the sensor-based method can help avoid the “lose track” issue when the environment is texture-less.

3.1.3. Marker-based pose tracking

Vision-based tracking methods in the AR community can be categorized into two techniques: marker-based tracking and markerless tracking. Marker is a designed pattern and often used to provide unique

visual features for AR tracking. Though markerless AR is more flexible than marker-based AR for not requiring marker installation, there are several reasons that we favor the marker-based method in our 2D projection experiment. First, when the environment is texture-less, using the markerless tracking method might cause a lose track issue. Besides, markers can be used as landmarks to determine the world coordinate system. In this paper, we used AprilTag, a visual fiducial marker that can provide 3D position, orientation, and identity of the markers relative to the camera [71,72] for pose estimation. The pose estimation accuracy of using the AprilTag system has been studied in [73].

As aforementioned, we need to obtain the camera's pose in the world frame during the tracking phase. The pose $T_w^c = [R_w^c, t_w^c; \mathbf{0}, 1]$ can be estimated by L ($L \geq 1$) AprilTags with known size and spacing. And the process to obtain this relative pose are composed of the following steps:

- 1) Put up the AprilTag(s) on the projection plane, with their physical 3D coordinates recorded as X_w . Note in this paper, we used the four corners of an AprilTag to represent its pose.
- 2) Capture the frames that contain these AprilTags using the camera;
- 3) Extract the 2D coordinates of the corners for AprilTags in the camera frame, represented as X_c ;

For the sake of completeness, we introduced the basic knowledge for visual camera pose tracking techniques from the computer vision community [74]. With the known X_w and X_c , we can compute the homography H_w^c , an invertible non-singular 3×3 projective matrix, which maps the points from the world frame to the camera frame. Since all the AprilTags are attached to a 2D plane, for example, a wall plane, we can set all the Z coordinates of X_w to be zero, noted as \tilde{X}_w . The homography-based coordinate transformation is:

$$\lambda X_c = H_w^c \tilde{X}_w \quad (2)$$

Notice all points are represented in homogeneous coordinates.

To compute the homography matrix, we reorganized the Eq. (2) as the format of $Ah = 0$, where A represents the known coordinates of points, and h represents the unknown parameters in the homography matrix. By solving $Ah = 0$ using SVD (singular value decomposition) with the smallest singular value, we can obtain the desired homograph H_w^c . The homography matrix can only be computed when there are at least four sets of non-collinear points.

With the calculated homography matrix H_w^c , the T_w^c , which is composed of R_w^c and t_w^c , can be computed by decomposing the homography matrix with the calibrated camera intrinsic matrix K_c . The process can be divided into three steps:

- 1) Compute the \hat{R}_1 , \hat{R}_2 and \hat{t} in the following equation:

$$\begin{bmatrix} \hat{R}_1 & \hat{R}_2 & \hat{t} \end{bmatrix} = K_c^{-1} H_w^c \quad (3)$$

Note that \hat{R}_1 , \hat{R}_2 and \hat{t} represent the first, second and third column of the computed matrix $K_c^{-1} H_w^c$.

- 2) Obtain the U , V matrix in SVD:

$$USV^T = \begin{bmatrix} \hat{R}_1 & \hat{R}_2 & \hat{R}_1 \times \hat{R}_2 \end{bmatrix} \quad (4)$$

- 3) Calculate the refined R_w^c , t_w^c which satisfy the constraint of $R_1^T R_2 = 0$ and $\|R_1\| = \|R_2\| = 1$ (R_1 , R_2 represent the first and second column of R_w^c respectively):

$$R_w^c = U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(UV^T) \end{bmatrix} V^T, \quad t_w^c = \hat{t} / \|\hat{R}_1\| \quad (5)$$

Therefore, $T_w^c = [R_w^c, t_w^c; \mathbf{0}, 1]$ is obtained. With the pre-calibrated extrinsic matrix of T_c^p , we can compute the pose of the projector in the

world frame as $T_w^p = T_c^p T_w^c$.

3.1.4. Hybrid pose tracking

For projection onto a texture-less or non-planar surface, we used the Iterative Closest Points (ICP) [75] based method to realize the pose estimation for the camera, by implementing point cloud registration between the known BIM data in the world frame and the scanned point cloud in the RGB-D camera frame. Once the registration result is converged, the 6DOF transformation matrix T_w^c can be obtained. Though we can use pure ICP based method for our tracking system, we still use a marker for two reasons: 1) A marker with a unique ID can serve as a landmark; 2) the ICP method requires an initial pose, and a marker can be used to provide the initial pose.

Other techniques can also be used for tracking. It could be helpful to use Extended Kalman Filter (EKF) [76–78] based localization method in the tracking process. By fusing different signals from different sensors, EKF based method can provide an accurate and stable localization result when the tracking environment lacks visual features. Usually, an RGB camera is used to provide visual signals, and an IMU is used to provide linear acceleration and angular velocity signals.

3.2. Projection image generation

3.2.1. Projection on planar surface

If the projection surface is a single 2D plane, the image generation problem can be formulated as a 2D–2D image warping problem. As can be seen from Fig. 3, to achieve the desired projection result in (d), an input image in (c) to the projector is needed. The input image can be generated by warping an original image in (a) using a warping matrix. The warping matrix is the homography matrix H_p^l , which maps each point from the image frame to the projector frame. Since $H_p^l = H_w^p H_l^w$, and H_w^p can be obtained from pose estimation process using: $H_w^p = K_p[R_1 R_2 t_w^c]$, we only need to compute H_l^w . By finding the corresponding points P_1, P_2, P_3, P_4 in Fig. 3(a) and (b), we can compute H_l^w using Eq. (2). In practice, the four corners' physical coordinates are determined by the as-planned or as-built BIM object's location, which means the coordinate system established by the four corners is the same as the BIM object's coordinate system. For the four corners in the image frame, their coordinates are related to the width w and height h , and the unit is pixel.

3.2.2. Projection on non-planar surface

According to Eq. (1), the projection image can be generated by converting all object points X_w in the world frame to image points X_p in the projector frame. The generated image is a black background image with image points drawn on it. However, it is more efficient to use some existing rendering tools for the image generation process. These tools have optimized the computing process and considered the image interpolation problem when the image points are located in subpixels.

To use existing rendering tools, since the physical principle and mathematical model of a projector is the same as a camera, we can consider a projector as an inverse camera. Therefore, the generated image for the projection in a real site should be an image captured by a virtual camera in a simulated environment of the real site. The image generation process can be seen from Fig. 4. Note the illustration image is depicted on a wall, yet the method can be used for non-planar surfaces. Fig. 4(a) is a rendered picture depicting the real site scene, where the MPAR system can project consistent virtual content on the real site. Fig. 4(b) shows the image generation process that can provide the projection image for the MPAR system in Fig. 4(a). As can be seen from Fig. 4(b), we first established a simulated environment of the real-world environment. Therefore, after obtaining the projector's pose T_w^p in the real world from the tracking phase, we can use it for a virtual camera in the simulated environment and capture an image. Note the virtual camera's intrinsic matrix is the calibrated projector's intrinsic matrix K_p . The captured image can be used as the input for the real world projector, and the augmented BIM result can be seen in Fig. 4(a).

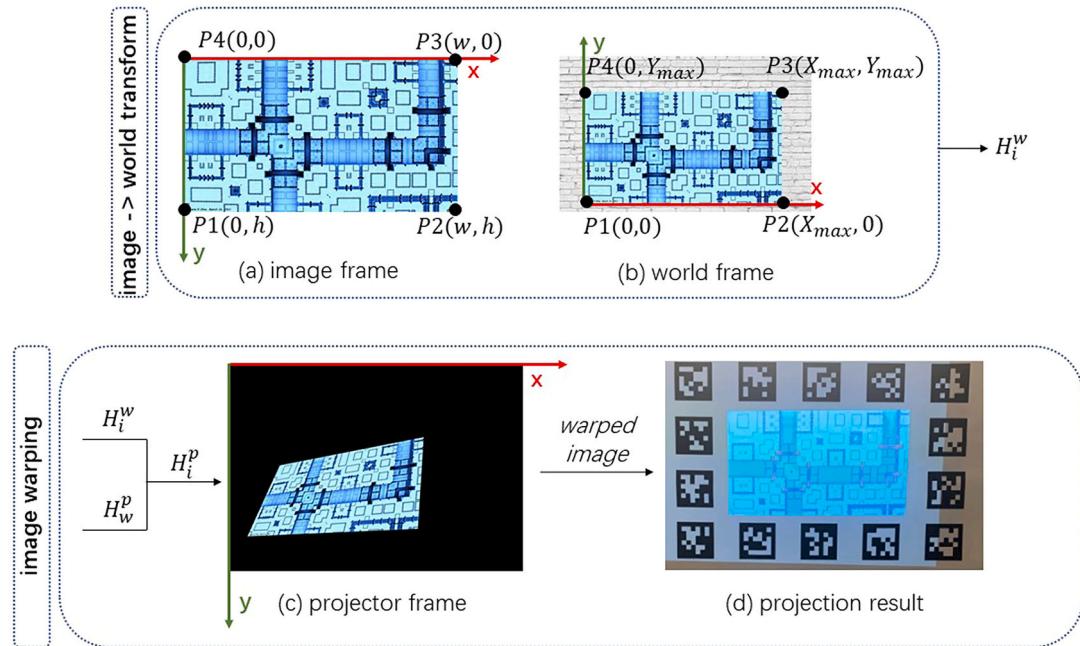


Fig. 3. Image generation workflow for projection on a single 2D plane.

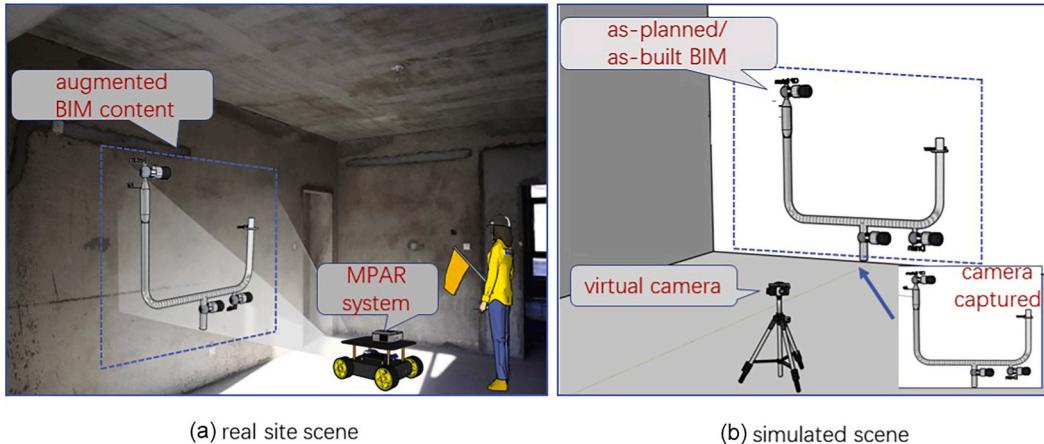


Fig. 4. Image generation for projection on non-planar surfaces.

4. Experiments

Based on the above methods, we designed two sets of experiments: qualitative evaluation (experiment 1, 2), and quantitative evaluation (experiment 3, 4). In experiment 1, as-planned building information was projected onto a single 2D plane; in experiment 2, as-built 3D model information was projected onto piecewise planar surface. In experiment 3 and 4, quantitative evaluation was conducted for projection on 1) a single 2D plane; 2) piecewise planar surface. Projection error is used as the metrics to evaluate the projection result.

These experiments can be readily extended to real construction scenario applications, since the environment setup is easy, and the projection process is highly automated. Some assumptions should be made before the implementation of these experiments. First, it is reasonable to assume that on-site workers or designers can have access to as-planned or as-built BIM data. Second, since the pose estimation method is based on detecting features from the environment, it is necessary to ensure that the MPAR system can always perceive visual features or structure features (3D depth information) from the

environment during the projection process.

4.1. Experiment 1: projection of As-planned building information on a single 2D plane

In this experiment, our MPAR system achieved a consistent and stable projection of as-planned BIM onto a single 2D plane. In most construction scenarios, walls are common surfaces we can project onto. Therefore, achieving a stable, accurate, and real-time projection is essential to promote our methods to many construction works. By designing the software in C++ with OpenCV library [79], the MPAR system achieved real-time projection with reasonable accuracy. The workflow of the experiment can be seen from Fig. 5. A water pipe image extracted from the BIM database in Fig. 5(b) is supposed to be projected onto the wall, as visualized in Fig. 5(a). Therefore, a warped image in Fig. 5(d) should be generated as the input image to the projector. Such an image generation process is explained in Figs. 3 and 5(c) is used to compute the projector's pose in the world frame. We first set up the working environment, then estimated the projector's pose and generate

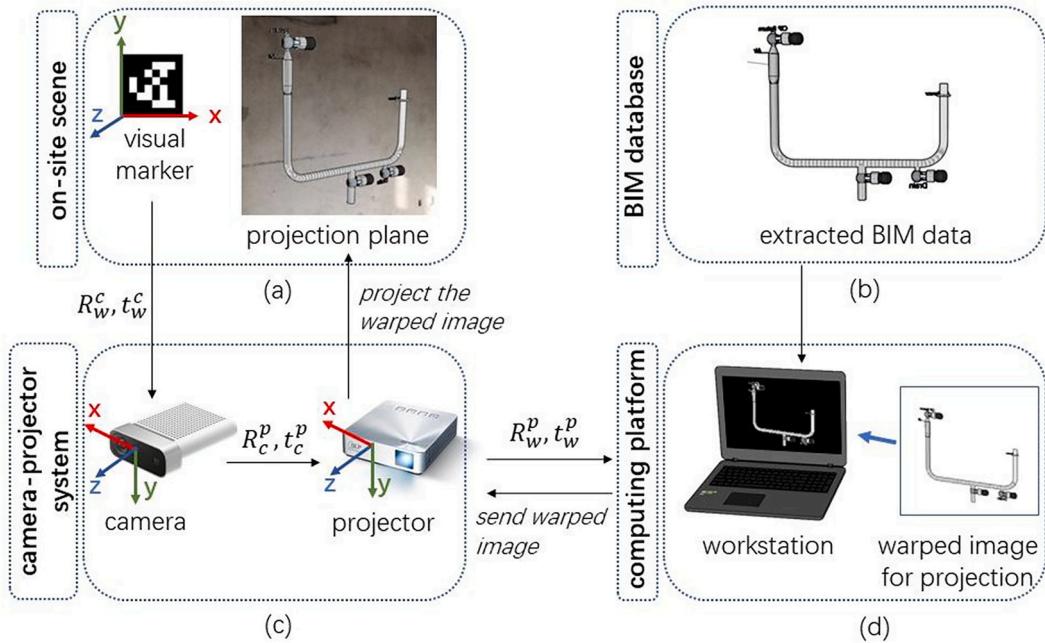


Fig. 5. The workflow of experiment 1.

the projection image.

4.1.1. Experiment setup

The hardware design of the MPAR system is explained in Fig. 1. We used a different camera and computing platform to achieve the projection results, yet the functionality modules play the same role as described in the figure. To achieve accurate and robust detection results, we used Microsoft's Azure Kinect sensor, an RGB-D camera, to capture RGB images. It draws power and sends data through two USB cables. For projection, we used Asus S1 projector, which is portable and has 3 h cable-free projection time after fully charged. The camera and the projector are rigidly bound, and pre-calibrated using the abovementioned method. Thus, we knew the intrinsic matrix of the camera and the projector, and the extrinsic matrix between them. For data processing and rendering, we used a laptop with USB connections to the Kinect camera and the projector.

For the experimental environment setup, we used a smooth white panel as the projection plane, as can be seen from Fig. 6. We put up a poster with multiple AprilTags printed on it. The printed AprilTags' IDs are from 0 to 13, and their physical locations on the poster are pre-defined. These AprilTags can be used to determine the world coordinate system; they can also serve as visual features for pose estimation. After defining the projection location of the virtual BIM information on the panel, we should let the MPAR system face the projection plane and

ensure both the camera frame and the projector frame can cover the projection plane. Finally, we should start the software work stream while letting the MPAR system move during the projection process.

4.1.2. Projection results

The projection results can be seen from Fig. 6. We placed the MPAR system at random places with random poses, and the augmented BIM information remains at the same location on the projection plane.

4.2. Experiment 2: projection of As-built building information on piecewise planar surface

Since most common non-planar surfaces in construction are piecewise planar, we implemented the projection experiment on piecewise planar surfaces without loss of generality. The environment setup can be seen from Fig. 7(d). The environment is composed of three mutually perpendicular planes, with several cylinders and boxes put on the horizontal plane. Also, an AprilTag is attached to one of the planes. We used C++ language with OpenCV, Open3D [80], and PCL library [81] to implement the two applications.

4.2.1. Pose estimation

As mentioned in the method section, we used the ICP algorithm to localize the projector in the world frame. Therefore, two sets of point



Fig. 6. Results of experiment 1. Though the projector is placed at three different locations with different projection angles, the projection result remains consistent in the pre-defined area on the projection plane.

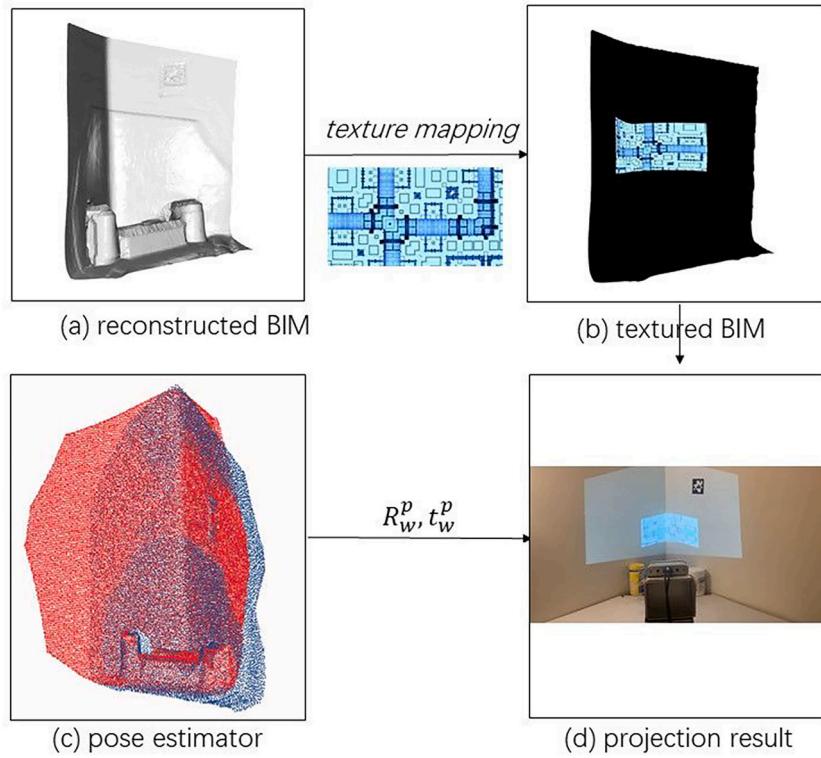


Fig. 7. The workflow of experiment 2.

cloud are required: the point cloud representing the BIM model and the point cloud scanned by the RGB-D Kinect camera in each frame. For the point cloud of the BIM model generation, we first used KinectFusion [82] method to generate a model; then, we registered the generated BIM model to the AprilTag's coordinate system. The point cloud of the BIM model is the blue set of the point cloud in Fig. 7(c). In each frame, we scanned the point cloud in the depth camera frame, which is visualized as red points in Fig. 7(c). By applying the ICP algorithm to these two sets of point cloud, the relative pose of the depth camera T_w^d with respect to the BIM model can be obtained, where w and d represents the world frame of the BIM system and the depth camera frame respectively. To obtain the projector's pose with respect to the BIM model, we can compute it as $T_w^p = T_c^p T_d^c T_w^d$. Here, the extrinsic matrix between the projector and the camera T_c^p , and that between the Kinect's RGB camera and the depth camera T_d^c are pre-calibrated.

4.2.2. Projection image generation

We used the two image generation ways mentioned in the method section to achieve the projection results, and we will discuss the two different implementations in this section, respectively.

4.2.2.1. Projection image generation using camera projection matrix. We have discussed the method to generate an input image for the projector using the projective projection matrix in the method section. In our experiment, we designed a grid pattern that we want to project onto the planes. The pattern is composed of several points on the piecewise planar surface and several green or red lines connecting these points. To generate the input image, we first need to define the corner coordinates of the pattern in the BIM model, and using Eq. (1) to compute the corresponding image points in the projector frame. We then drew these image points on a black image and connecting these points with the correct order to generate the lines in the pattern.

4.2.2.2. Projection image generation using rendering tools. As introduced in the method part, to use the existing rendering tools to generate the

input image, we need to simulate a virtual environment, and a virtual camera using the intrinsic and extrinsic matrix of the projector in the real world frame. For the simulated virtual environment, since we have already generated the point cloud BIM model, we need further to process this model to a textured mesh BIM model. In practice, we used MeshLab [83] to convert the point cloud to mesh, as in Fig. 5(a). Then we used Blender [84] to apply texture mapping, using a given pipe image as the texture, and the textured mesh can be seen from (b). For the virtual camera, the extrinsic matrix in the simulated environment equals to T_w^p , and the intrinsic matrix equals to K_p , which is pre-calibrated. Finally, we sent the textured BIM model to the simulated environment in Open3D and rendered an image using the parameters of the virtual camera.

4.2.3. Projection results

As we can see in Figs. 8 and 9, for both methods, the projection results are located at the designed place. Therefore, we believe the MPAR system can achieve reasonable projection results on non-planar surfaces. The projection inaccuracy might come from the system error, and we will conduct quantitative evaluation experiments in the following sections.

4.3. Experiment 3: quantitative evaluation for projection on a single 2D plane

One existing problem that might hinder AR applications in the construction industry is the lack of the accuracy of some AR methods. The inaccurate AR superimposition might be problematic and cause severe consequences for the construction industry. Therefore, before we widely apply these AR techniques to the construction industry, we need to do the quantitative evaluation work. In most of the AR methods, the inaccuracy is mainly from two aspects [85]: 1) pose estimation for the AR system; 2) image generation and projection. In our MPAR system, we believe that the projection results rely on: 1) the pose estimation result for the visual perception sensor, in our experiment is the Kinect camera; 2) the pre-calibration results for the intrinsic matrix of the projector and

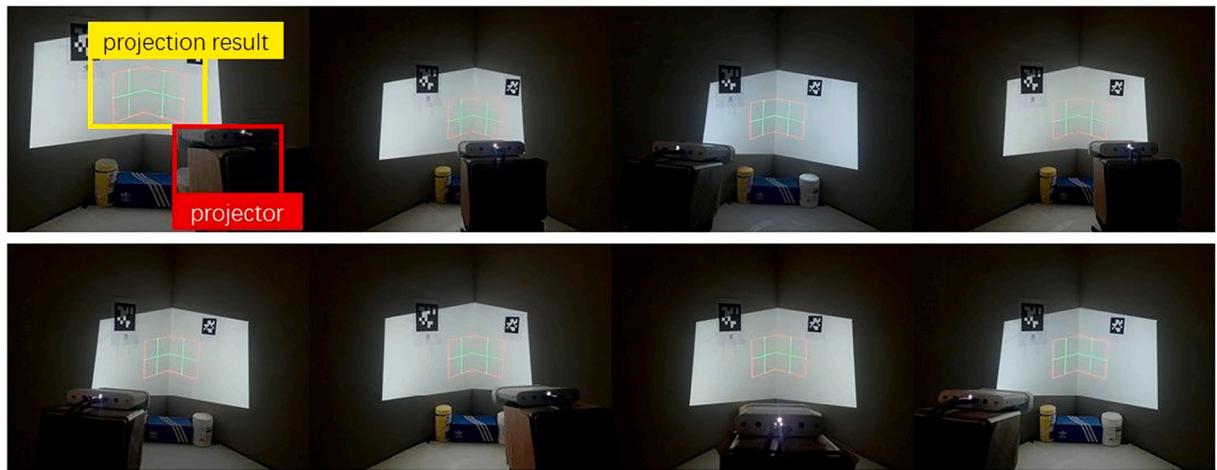


Fig. 8. Results of experiment 2 using projection matrix. Similar to experiment 1, the projection results remain consistent while the projector is placed at eight different locations with different projection angles.



Fig. 9. Results of experiment 2 using rendering tools.

the camera, and the extrinsic matrix between them; 3) the image generation accuracy. We will calculate the accuracy of our MPAR system and analyze the potential reasons for the projection results.

4.3.1. Experiment setup

The purpose of the quantitative evaluation experiment is to measure the discrepancy between the as-planned building information and the

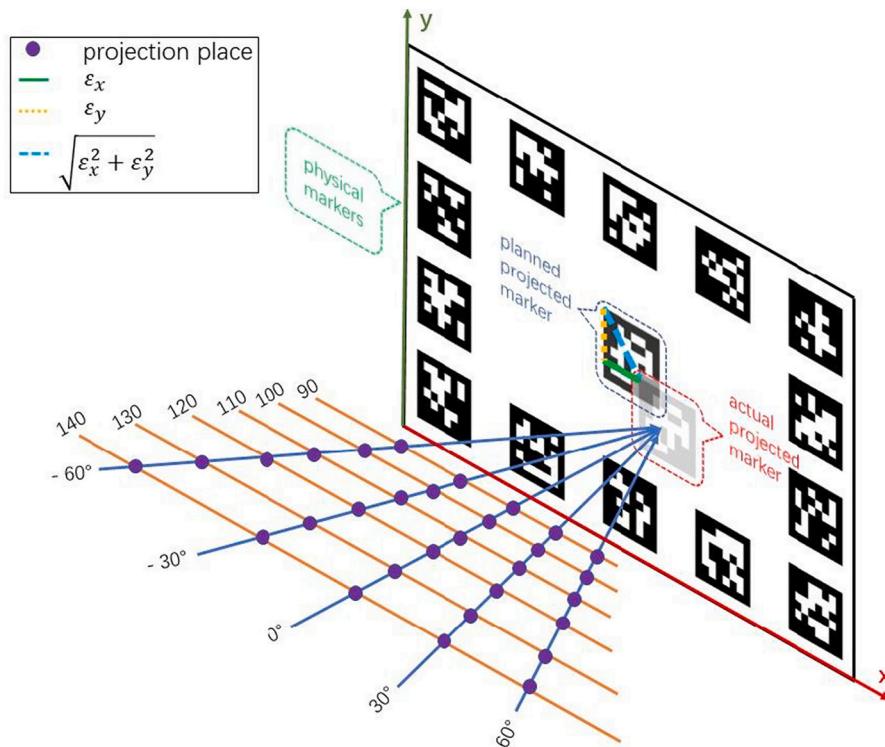


Fig. 10. Quantitative evaluation experiment setup for experiment 3.

superimposed virtual information. We designed an experiment to automatically compute the discrepancies between the planned projection locations and the actual projection locations. As can be seen from Fig. 10, we used the same poster as in experiment 1. The data is captured from six distances and five angles, and each projection place is visualized as the purple spot in the figure. Distances are measured as the length between the principal point of the projector and the projection plane, and the unit is cm. Angles are defined as negative and positive degrees, which are the angles between the projection ray and the perpendicular ray to the projection plane, for left and right projection, respectively. ε_x and ε_y represent the projection error along x -axis and y -axis respectively; $\sqrt{\varepsilon_x^2 + \varepsilon_y^2}$ is the translational error. We planned to project an AprilTag with ID 15 in the center of the poster, with 10cm for side length. Once the MPAR system starts to work, the AprilTag with ID 15 will be projected onto the poster, and we can compute its actual projection location. To efficiently compute the projected April Tag's location, we used an RGB camera to capture each frame during the projection process. In each frame, the printed AprilTags and the projected AprilTag can be detected. The printed AprilTags are used as control points to determine the projector's pose in each frame, since their physical locations are known ahead, and their coordinates in the captured camera frame can be detected. Therefore, the homography matrix H_w^c can be computed using corresponding points in the world frame and the camera frame. Finally, the actual projection location of AprilTag 15 can be calculated by converting its detected location in the camera frame to the world frame using H_w^c . By calculating the difference between the planned projection location and the actual projection location, we can obtain four projection errors (four corners of the AprilTag) from one frame.

We want to study how the projection distance and projection angle affect the projection results; therefore, we located the MPAR system in different distances, with different projection angles and computed the projection error for each combination of distance and angle. In practice, we captured 30 videos from six distances and five angles. The configuration is explained in Fig. 10. When computing the projection error from different projection distances, to reduce the random projection error for each distance, we used a large number of projection frames to compute the projection error. We took the average value as the projection error. To better illustrate the computation process, we give an example of how to compute the projection error in distance of 90cm. Note the projection error contains the error on x -axis, error on y -axis, and the translation error. We captured 5 videos from five different projection angles on distance 90cm, and for each video, we extracted 600 frames. Therefore,

five videos can produce 3,000 frames. In each frame, we can detect four corners of the projected AprilTag. So in all we have 12,000 points on distance 90cm. These 12,000 points are all used as the data to compute the average projection error. For the remaining locations, we followed the same rules for the projection error computation.

To compute the projection error from different projection angles, we used the same way as for the error computation from different distances. In this section, we used all data collected from this angle at different distances for a specific angle. The data contain 14,400 points: 6 distances for each angle; 600 frames extracted from each video; 4 corners detected for projection error computation.

4.3.2. Error analysis for projection from different distances

We found several error sources jointly affecting the accuracy of MPAR projection result:

- 1) Using a focal length parameter calibrated in a specific distance for a non fixed-focus projector might cause inaccurate projection results. As can be seen from Fig. 11, the data are collected from six distances, from 90cm to 140cm, with 10cm as the interval. The blue box, the yellow box represents the error on x -axis, y -axis, respectively; the green box represents the translational error. Note the red symbols in this figure are outliers in the data. We find projection error is highly related to distances. In a specific distance (100cm in our experiment), the absolute translational projection error is within 0.5cm, and the relative projection error is within 0.05. We also noticed that the translational projection error in distance 90cm and 110cm is also relatively small, compared to other distances. We think this is because the projector is a short-throw projector, and its focal length will change when the projection distance changes. Since we calibrated the camera-projector system around the distance of 100cm, the calibration parameters are likely to fit that distance, but not fit for other distances. If we still use that set of calibration parameters for other distances, they are probably not the actual calibration parameters. Using these calibration parameters will lead to inaccurate projection results. The inaccuracy caused by the focal length changes can also be seen from Fig. 12. The top line figures are the projection error visualization for four corners of the projected AprilTag. The bottom line figures are the error ellipse of the projection error data. Each ellipse's width and height are two times of the standard deviation of the data along the corresponding axes. We found the projection errors visualized in this figure are moving to the left bottom direction and becoming larger when the distance becomes larger,

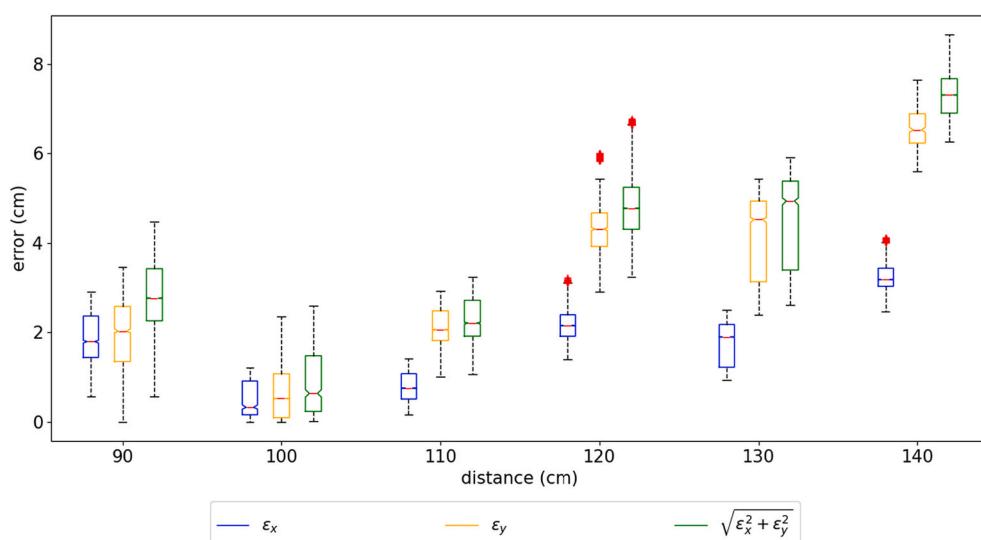


Fig. 11. Box plot for experiment 3 from different projection distances.

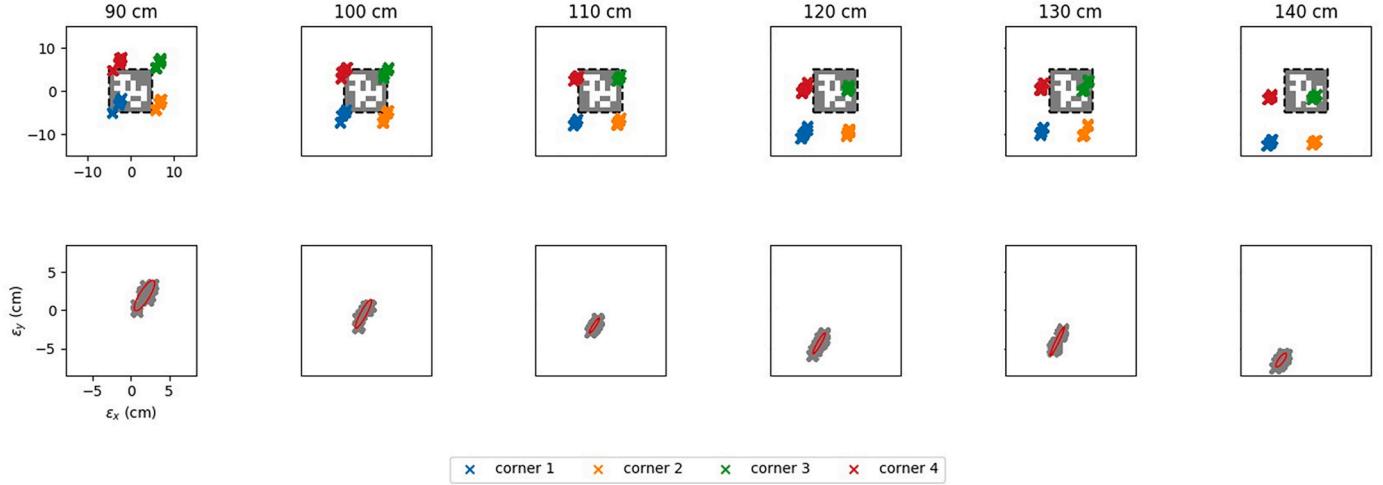


Fig. 12. Visualization for experiment 3 from different projection distances.

which might be the consequence of using an inappropriate foal length. As we can see in Eq. (3), the focal length of the camera will affect the calculation result. Therefore, using an inaccurate focal length will lead to inaccurate projection. Therefore, if people want to use such a short-throw projector on the working site, the range of projection distance should be considered ahead, and the calibration should be conducted on the most used distance.

- 2) Denser visual information can benefit pose estimation, and contribute to more accurate projection results. The density of the control points can be measured as the ratio of the number of control points to the distance. The error along $y - axis$ is more significant than that along $x - axis$. As we can see from the Fig. 10, the control points along the $x - axis$ are denser than that along the $y - axis$. From [86], it is proved that the denser control points can benefit the pose estimation result.. Therefore, the localization result along $x - axis$ will be more accurate, which can explain why the projection results along $x - axis$ are more accurate. If users are using visual features to do localization, then dense visual information could be beneficial.

4.3.3. Error analysis for projection from different angles

Though the difference of the projection accuracy from different angles are not as obvious as from different distances, there still exist several factors that could cause potential systematic errors for the projection results.

The MPAR system is not sensitive to projection angles. Compared to the quantitative evaluation results from different projection distances, the change on angles does not cause a noticeable difference for

projection results, as can be seen from Fig. 13. This conclusion can also be drawn from Fig. 14, where the error ellipse does not change much when the angle changes. However, we assume that most projector-based AR applications should still be constrained within a reasonable range of angles. Otherwise, it will fail due to the physical projection limitation, e. g. when the projector's principal axis is almost parallel to the projection plane.

4.4. Experiment 4: quantitative evaluation for projection on piecewise planar surface

Besides the factors that can cause inaccuracy discussed in the 2D single plane projection quantitative evaluation, we think several aspects can lead to inaccuracy in the piecewise planar surface projection experiment. First, since KinectFusion reconstructs the as-built BIM model, it might have some inaccuracy in the reconstructed model. Therefore, if we used the ICP algorithm to obtain the MPAR system's pose, it will lead to an inaccurate result and affect the final projection results. More potential factors that affect the projection accuracy will be discussed later.

4.4.1. Experiment setup

In this section, we used the same devices and environment as the projection experiment on piecewise planar surface. The illustration figure can be seen from Fig. 8. Similar to the quantitative evaluation on a single plane, which uses the homography transformation between the world frame and the projector frame to help compute the projection

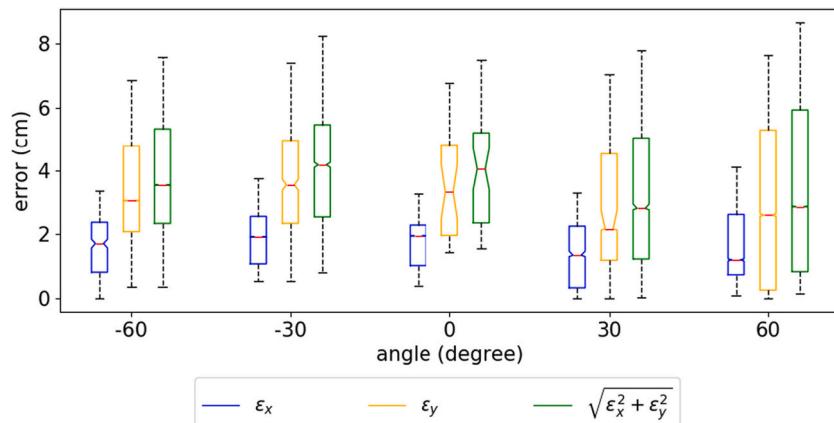


Fig. 13. Box plot for experiment 3 from different projection angles. The legend is the same as in Fig. 11.

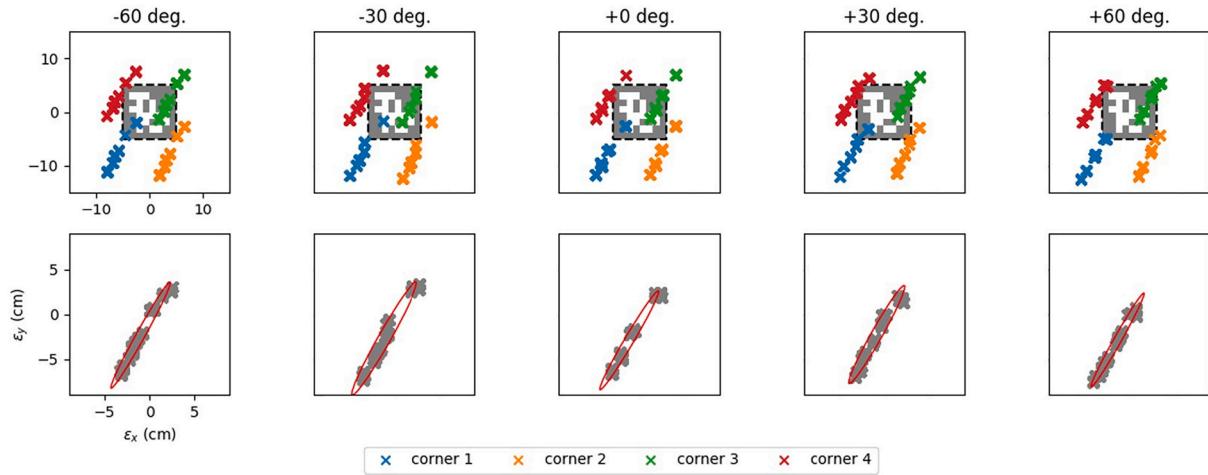


Fig. 14. Projection error visualization and error ellipse for experiment 3 from different projection angles. The legend is the same as in Fig. 12.

error, we treated the piecewise planar surface projection quantitative evaluation as uses two homography transformations to help compute projection error. For each plane, we specified two points to be projected onto the specified location. Then we detected the projected points in the captured camera frame, using fast feature detectors [87]. We followed the method used in experiment 3 obtain the projected points' coordinates in the world frame, and calculate the projection error between the specified projection location and actual projected location. The two AprilTags attached on two planes are used to serve as control points, and their coordinates in the world are known.

As we have already explored the factors that might affect projection error according to different distances, ranging from 90cm to 140cm in the previous quantitative evaluation part, we focused on the projection error from a relatively small distance range, from 90cm to 110cm this time. It has the best performance in the previous quantitative evaluation experiment from these three distances. We believe it is more worthy of exploring the projection results after filtering out some extreme distances. For the selection of angles, we selected angles -30° , 0° and 30° since our experiment environment consists of three mutually perpendicular planes, and the corner is 90° , therefore study on these three angles can represent the typical angle range that the MPAR system works.

4.4.2. Error analysis for projection from different distances

The projection error on 90cm, 100cm, and 110cm can be seen from Fig. 15. From the error ellipse in Fig. 16 regarding different projection

distances, the projection errors are mainly random errors, instead of system errors, since: 1) the center of the error ellipse is around the original point, which means the bias is close to 0; 2) the correlation between ε_x and ε_y is not obvious.

4.4.3. Error analysis for projection from different angles

The projection error from -30° , 0° and 30° can be seen from Fig. 15. See from the error ellipse from different projection angles in Fig. 16, random errors are still the main component of the projection errors.

4.5. Limitation and discussion

Admittedly, there are several limitations to MPAR. One limitation is the projection speed of the current implementation when the tracking method uses the ICP algorithm for projection on non-planar surfaces. In our experiment, it takes 4 to 5 seconds to obtain a converged pose. We believe the speed can be improved in the future by using: 1) GPU-based ICP; or 2) other SLAM-based pose tracking methods. Another limitation is the image brightness of the projector could be easily affected by other light sources such as the sunlight, leading to less visible AR in some situations (e.g., outdoor). This can be alleviated by using high-end projectors or laser-based ones.

Some may find that our MPAR system does not provide an AR environment as immersive as see-through based AR using helmets or goggles. We stress that this is a design choice we made as a trade-off between the immersive user experience, which is less desirable for

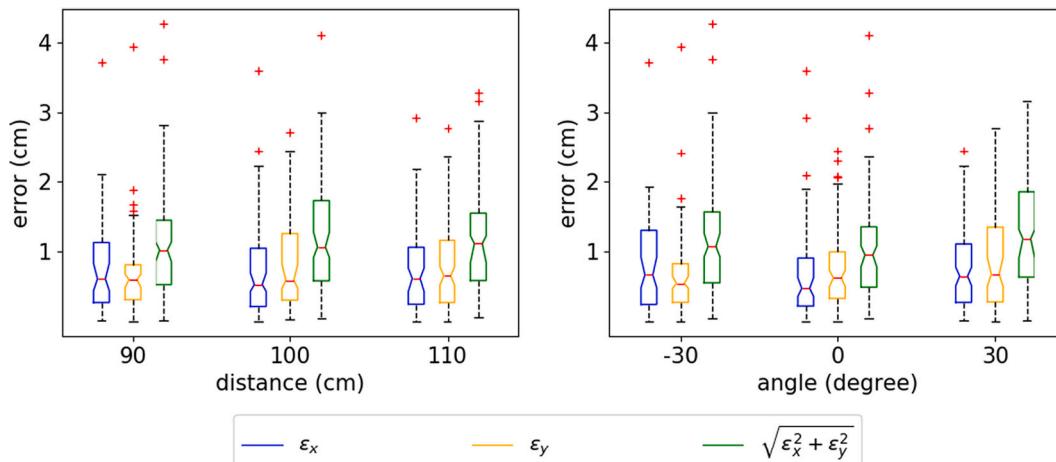


Fig. 15. Box plot for experiment 4 from different projection distances and angles. The legend is the same as in Fig. 11.

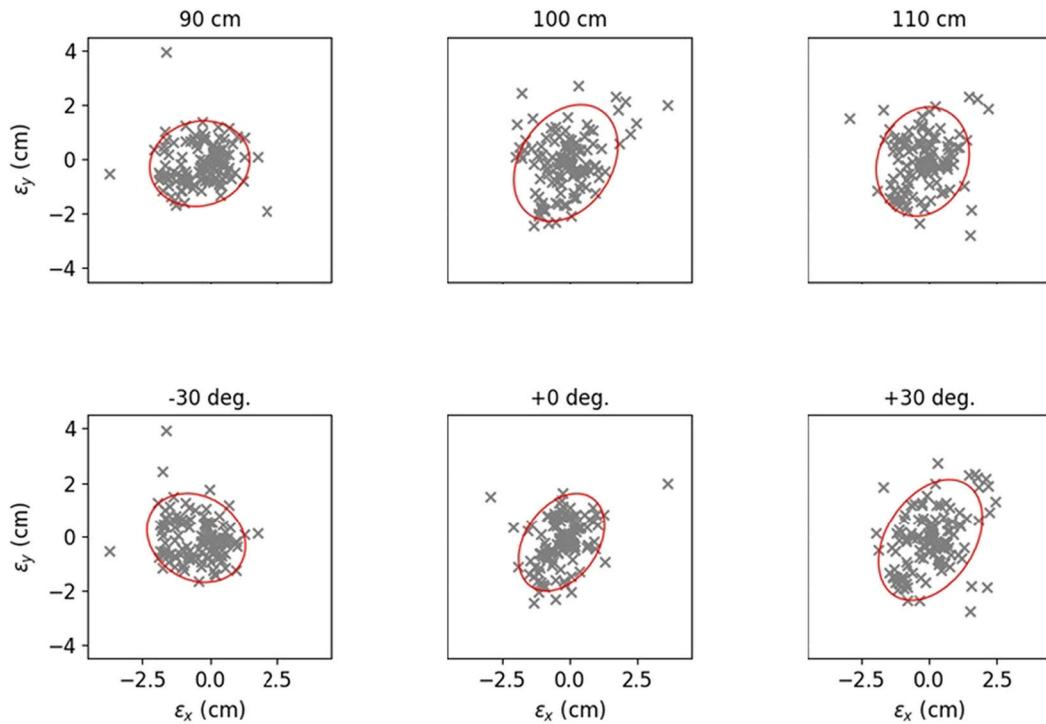


Fig. 16. Error ellipse for experiment 4 from different projection distances and angles. Each ellipse's width and height are two times of the standard deviation of the data along the corresponding axes.

onsite operations, and the safety and efficiency concerns of helmets or goggles in practice, which is due to the restricted field of view and non-negligible weight of those existing off-the-self AR devices. Similarly, although our current implementation uses fiducial markers which might imply extra installation workload, it should be noted that it is not a core requirement of MPAR method and can be easily updated to newer place recognition and pose estimation methods.

Additionally, the quantitative evaluation part shows that there is still room for projection accuracy improvement. Using better camera-projector calibration and pose estimation method could further improve the projection accuracy. However those methods are not the main contribution of this paper and can be plug-and-play in our MPAR framework when available. Lastly, the MPAR system's capability relies on the frame rate and resolution of the camera and the projector: using a more precise industrial camera and projector might also improve the projection accuracy.

5. Conclusions

Motivated by previous researchers' safety and efficiency concerns of applying existing AR techniques in construction scenarios, we proposed MPAR, a camera-projector based AR system, that ensures a consistent projection even when the device might move during the projection process. Our MPAR system can project virtual construction information on a single 2D plane or non-planar surfaces. We demonstrated that our MPAR system is feasible to be applied to some construction scenarios and it can achieve centimeter-level projection accuracy. We also analyzed the potential factors that might affect the projection accuracy from the quantitative evaluation experiments on projection and provided insight for how to further improve the projection accuracy.

In the future, we plan to further extend the current MPAR method and explore its potential in complex job sites. First, it is necessary to enlarge the current projection area, which may require the simultaneous localization and mapping (SLAM) based localization method. Therefore, we can study how to combine the large-scale SLAM method with the

pose estimation method in our MPAR system. Second, it could be beneficial if we could link the building-level BIM database with the MPAR system and assist the construction workers in different rooms in the building. Therefore, we can investigate the feasibility to efficiently and accurately retrieve the required BIM data according to the MPAR system's location and pose. Lastly, as we mentioned before, introducing a laser-based projector to replace the current projector could be helpful if we want the MPAR system to work in the outdoor environment, or even deploy the MPAR system on the drones for outdoor projection.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by NSF Future Manufacturing program under EEC-2036870. Siyuan Xiang gratefully thanks the IDC Foundation for its scholarship. The authors also thank Yuhui Fu for refining the software, Harish Kuppam and Akshay Kumar V Kutty for building the hardware, when they worked with the authors as a team on prototyping MPAR in the North America Final Trail of an International Construction Innovation Competition held by the VINCI Construction company.

References

- [1] K.M. Lundein, V.R. Kamat, C.C. Menassa, W. McGee, Scene understanding for adaptive manipulation in robotized construction work, *Autom. Constr.* 82 (2017) 16–30, <https://doi.org/10.1016/j.autcon.2017.06.022>.
- [2] E. Nadhim, C. Hon, B. Xia, I. Stewart, D. Fang, Falls from height in the construction industry: a critical review of the scientific literature, *Int. J. Environ. Res. Public Health* 13 (7) (2016), <https://doi.org/10.3390/ijerph13070638>, p. 638.
- [3] L. Ding, C. Zhou, Development of web-based system for safety risk early warning in urban metro construction, *Autom. Constr.* 34 (2013) 45–55, <https://doi.org/10.1016/j.autcon.2012.11.001>.

- [4] Y. Zhou, L. Ding, L. Chen, Application of 4d visualization technology for safety management in metro construction, *Autom. Constr.* 34 (2013) 25–36, <https://doi.org/10.1016/j.autcon.2012.10.011>.
- [5] F. Barbosa, J. Woetzel, J. Mischke, M.J. Ribeirinho, M. Sridhar, M. Parsons, N. Bertram, S. Brown, *Reinventing Construction: A Route to Higher Productivity*, McKinsey Global Institute, 2017.
- [6] OSHA, Worker Safety Series [Online; accessed 1-Jan-2019]. [Online]. Available, <https://www.osha.gov/Publications/OSHA3252/3252.html>, 2018.
- [7] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, D. Schmalstieg, Handheld augmented reality for underground infrastructure visualization, *Pers. Ubiquit. Comput.* 13 (4) (2009) 281–291, <https://doi.org/10.3390/ijerph1307063810.1007/s00779-008-0204-5>.
- [8] C. Woodward, M. Hakkarainen, O. Korkalo, T. Kantonen, M. Aittala, K. Rainio, K. Kähkönen, Mixed reality for mobile construction site visualization and communication, in: *10th International Conference on Construction Applications of Virtual Reality*, 2010, pp. 4–5.
- [9] H.-L. Chi, Y.-C. Chen, S.-C. Kang, S.-H. Hsieh, Development of user interface for tele-operated cranes, *Adv. Eng. Inform.* 26 (3) (2012) 641–652, <https://doi.org/10.1016/j.aei.2012.05.001>.
- [10] X. Li, W. Yi, H.-L. Chi, X. Wang, A.P. Chan, A critical review of virtual and augmented reality (vr/ar) applications in construction safety, *Autom. Constr.* 86 (2018) 150–162, <https://doi.org/10.1016/j.autcon.2017.11.003>.
- [11] S. Siltanen, Diminished reality for augmented reality interior design, *Vis. Comput.* 33 (2) (2017) 193–208, <https://doi.org/10.1007/s00371-015-1174-z>.
- [12] V.T. Phan, S.Y. Choo, Interior design in augmented reality environment, *Int. J. Comput. Appl.* 5 (5) (2010) 16–21, <https://doi.org/10.5120/912-1290>.
- [13] J. Cook, S. Gibson, T. Howard, R. Hubbolt, Real-time photo-realistic augmented reality for interior design, in: *ACM SIGGRAPH 2003 Sketches & Applications*, ACM, 2003, <https://doi.org/10.1145/965400.965427>, pp. 1.
- [14] J. Hui, Approach to the interior design using augmented reality technology, in: *2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications*, IEEE, 2015, pp. 163–166, <https://doi.org/10.1109/ISDEA.2015.50>.
- [15] S. Siltanen, C. Woodward, Augmented interiors with digital camera images, in: *Proceedings of the 7th Australasian User Interface Conference Vol. 50*, Australian Computer Society, Inc., 2006, pp. 33–36, <https://doi.org/10.1145/1151758.1151761>.
- [16] T.-H. Lin, C.-H. Liu, M.-H. Tsai, S.-C. Kang, Using augmented reality in a multiscreen environment for construction discussion, *J. Comput. Civ. Eng.* 29 (6) (2014), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000420](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000420), p. 04014088.
- [17] B. Kim, C. Kim, H. Kim, Interactive modeler for construction equipment operation using augmented reality, *J. Comput. Civ. Eng.* 26 (3) (2011) 331–341, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000137](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000137).
- [18] R. Ghimire, K.R. Pattipati, P.B. Luh, Fault diagnosis and augmented reality-based troubleshooting of hvac systems, in: *2016 IEEE Autotestcon*, IEEE, 2016, pp. 1–10, <https://doi.org/10.1109/AUTEST.2016.7589590>.
- [19] M. Aftab, S.C.-K. Chau, M. Khanji, Enabling self-aware smart buildings by augmented reality, in: *Proceedings of the Ninth International Conference on Future Energy Systems*, 2018, pp. 261–265, <https://doi.org/10.1145/3208903.3208943>.
- [20] S. Lee, Ö. Akin, Augmented reality-based computational fieldwork support for equipment operations and maintenance, *Autom. Constr.* 20 (4) (2011) 338–352, <https://doi.org/10.1016/j.autcon.2010.11.004>.
- [21] X. Wang, P.S. Dunston, Design, strategies, and issues towards an augmented reality-based construction training platform, *J. Inform. Technol. Constr.* 12 (25) (2007) 363–380.
- [22] R.E. Pereira, M. Gheisari, B. Esmaeili, Using panoramic augmented reality to develop a virtual safety training environment, *Constr. Res. Congr.* (2018) 29–39, <https://doi.org/10.1061/9780784481288.004>.
- [23] Q.T. Le, A. Pedro, C.R. Lim, H.T. Park, C.S. Park, H.K. Kim, A framework for using mobile based virtual reality and augmented reality for experiential construction safety education, *Int. J. Eng. Educ.* 31 (3) (2015) 713–725.
- [24] K. Kim, H. Kim, H. Kim, Image-based construction hazard avoidance system using augmented reality in wearable device, *Autom. Constr.* 83 (2017) 390–403, <https://doi.org/10.1016/j.autcon.2017.06.014>.
- [25] A. Albert, M.R. Hallowell, B. Kleiner, A. Chen, M. Golparvar-Fard, Enhancing construction hazard recognition with high-fidelity augmented virtuality, *J. Constr. Eng. Manag.* 140 (7) (2014), [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000860](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000860), p. 04014024.
- [26] R. Xiao, H. Benko, Augmenting the field-of-view of head-mounted displays with sparse peripheral displays, in: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 1221–1232, <https://doi.org/10.1145/2858036.2858212>.
- [27] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, H. Fuchs, Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources, in: *ACM SIGGRAPH 2014 Emerging Technologies*, ACM, 2014, <https://doi.org/10.1145/2601097.2601141>, p. 20.
- [28] S. Kim, M.A. Nussbaum, J.L. Gabbard, Augmented reality “smart glasses” in the workplace: industry perspectives and challenges for worker safety and health, *IIE Trans. Occupat. Ergon. Hum. Factors* 4 (4) (2016) 253–258, <https://doi.org/10.1080/21577323.2016.1214635>.
- [29] S. Dong, A.H. Behzadan, C. Feng, V.R. Kamat, Collaborative visualization of engineering processes using tabletop augmented reality, *Adv. Eng. Softw.* 55 (2013) 45–55, <https://doi.org/10.1016/j.adengsoft.2012.09.001>.
- [30] D. Yan, H. Hu, Application of augmented reality and robotic technology in broadcasting: a survey, *Robotics* 6 (3) (2017), <https://doi.org/10.3390/robotics630018>, p. 18.
- [31] H. Bae, M. Golparvar-Fard, J. White, High-precision vision-based mobile augmented reality system for context-aware architectural, engineering, construction and facility management (aec/fm) applications, *Vis. Eng.* 1 (1) (2013), <https://doi.org/10.1186/2213-7459-1-3>, p. 3.
- [32] J. Irizarry, M. Gheisari, G. Williams, B.N. Walker, Infospot: a mobile augmented reality method for accessing building information through a situation awareness approach, *Autom. Constr.* 33 (2013) 11–23, <https://doi.org/10.1016/j.autcon.2012.09.002>.
- [33] A. Shirazi, A.H. Behzadan, Design and assessment of a mobile augmented reality-based information delivery tool for construction and civil engineering curriculum, *J. Prof. Issues Eng. Educ. Pract.* 141 (3) (2014), [https://doi.org/10.1061/\(ASCE\)EL.1943-5541.0000229](https://doi.org/10.1061/(ASCE)EL.1943-5541.0000229), p. 04014012.
- [34] M. Kopsida, I. Brilakis, Markerless bim registration for mobile augmented reality based inspection, in: *Proceedings of the 11th European Conference on Product and Process Modeling*, 2016, pp. 6–8.
- [35] H. Hedayati, M. Walker, D. Szafir, Improving collocated robot teleoperation with augmented reality, in: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 78–86, <https://doi.org/10.1145/3171221.3171251>.
- [36] O. Kyjanek, B. Al Bahar, L. Vasey, B. Wannemacher, A. Menges, Implementation of an augmented reality ar workflow for human robot collaboration in timber prefabrication, in: *Proceedings of the 36th International Symposium on Automation and Robotics in Construction*, 2019, pp. 1223–1230, <https://doi.org/10.22260/ISARC2019/0164>.
- [37] S. Xiang, R. Wang, C. Feng, Towards mobile projective ar for construction co-robots, in: *Proceedings of the International Symposium on Automation and Robotics in Construction*, Vol. 36, IAARC Publications, 2019, pp. 1106–1113, <https://doi.org/10.22260/ISARC2019/0147>.
- [38] S. Azhar, Building information modeling (bim): trends, benefits, risks, and challenges for the aec industry, *Leadersh. Manag. Eng.* 11 (3) (2011) 241–252, [https://doi.org/10.1061/\(ASCE\)LM.1943-5630.0000127](https://doi.org/10.1061/(ASCE)LM.1943-5630.0000127).
- [39] C.M. Eastman, C. Eastman, P. Teicholz, R. Sacks, K. Liston, *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*, John Wiley & Sons, 2011.
- [40] R. Volk, J. Stengel, F. Schultmann, Building information modeling (bim) for existing buildings—literature review and future needs, *Autom. Constr.* 38 (2014) 109–127, <https://doi.org/10.1016/j.autcon.2013.10.023>.
- [41] S. Azhar, M. Hein, B. Sketo, Building information modeling (bim): benefits, risks and challenges, in: *Proceedings of the 44th ASC Annual Conference*, 2008, pp. 2–5, [https://doi.org/10.1061/\(ASCE\)LM.1943-5630.0000127](https://doi.org/10.1061/(ASCE)LM.1943-5630.0000127).
- [42] J.D. Goedert, P. Meadati, Integrating construction process documentation into building information modeling, *J. Constr. Eng. Manag.* 134 (7) (2008) 509–516, [https://doi.org/10.1061/\(ASCE\)0733-9364\(2008\)134:7\(509\)](https://doi.org/10.1061/(ASCE)0733-9364(2008)134:7(509)).
- [43] B. Becerik-Gerber, S. Rice, The perceived value of building information modeling in the u building industry, *J. Inform. Technol. Constr.* 15 (15) (2010) 185–201.
- [44] Y. Jiao, S. Zhang, Y. Li, Y. Wang, B. Yang, Towards cloud augmented reality for construction application by bim and sns integration, *Autom. Constr.* 33 (2013) 37–47, <https://doi.org/10.1016/j.autcon.2012.09.018>.
- [45] M. Gheisari, J. Irizarry, Investigating human and technological requirements for successful implementation of a bim-based mobile augmented reality environment in facility management practices, *Facilities* 34 (1/2) (2016) 69–84, <https://doi.org/10.1108/F-04-2014-0040>.
- [46] C.-S. Park, D.-Y. Lee, O.-S. Kwon, X. Wang, A framework for proactive construction defect management using bim, augmented reality and ontology-based data collection template, *Autom. Constr.* 33 (2013) 61–71, <https://doi.org/10.1016/j.autcon.2012.09.010>.
- [47] X. Wang, P.E. Love, M.J. Kim, C.-S. Park, C.-P. Sing, L. Hou, A conceptual framework for integrating building information modeling with augmented reality, *Autom. Constr.* 34 (2013) 37–44, <https://doi.org/10.1016/j.autcon.2012.10.012>.
- [48] X. Wang, M. Truijens, L. Hou, Y. Wang, Y. Zhou, Integrating augmented reality with building information modeling: onsite construction process controlling for liquefied natural gas industry, *Autom. Constr.* 40 (2014) 96–105, <https://doi.org/10.1016/j.autcon.2013.12.003>.
- [49] K.-C. Yeh, M.-H. Tsai, S.-C. Kang, On-site building information retrieval by using projection-based augmented reality, *J. Comput. Civ. Eng.* 26 (3) (2012) 342–355, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000156](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000156).
- [50] B. Ridel, P. Reuter, J. Laviole, N. Mellado, N. Couture, X. Granier, The revealing flashlight: interactive spatial augmented reality for detail exploration of cultural heritage artifacts, *J. Comput. Cult. Herit.* 7 (2) (2014), <https://doi.org/10.1145/2611376>, p. 6.
- [51] B. Jones, R. Sodhi, M. Murdock, R. Mehra, H. Benko, A. Wilson, E. Ofek, B. MacIntyre, N. Raghuvanshi, L. Shapira, Roomalive: magical experiences enabled by scalable, adaptive projector-camera units, in: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, ACM, 2014, pp. 637–644, <https://doi.org/10.1145/2642918.2647383>.
- [52] H. Benko, E. Ofek, F. Zheng, A.D. Wilson, Fovear: Combining an optically see-through near-eye display with projector-based spatial augmented reality, in: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, ACM, 2015, pp. 129–135, <https://doi.org/10.1145/2807442.2807493>.
- [53] K. Tatsumoto, S. Iwaki, T. Ikeda, Tracking projection method for 3d space by a mobile robot with camera and projector based on a structured-environment approach, *Artif. Life Robot.* 22 (1) (2017) 90–101, <https://doi.org/10.1007/s10005-016-0312-7>.
- [54] R.T. Chadalavada, H. Andreasson, R. Krug, A.J. Lilenthal, That's on my mind! robot to human intention communication through on-board projection on shared

- floor space, in: European Conference on Mobile Robots, 2015, pp. 1–6, <https://doi.org/10.1109/ECMR.2015.7403771>.
- [55] M.D. Coover, T. Lee, I. Shindev, Y. Sun, Spatial augmented reality as a method for a mobile robot to communicate intended movement, *Comput. Hum. Behav.* 34 (2014) 241–248, <https://doi.org/10.1016/j.chb.2014.02.001>.
- [56] Y. Sheng, T.C. Yapo, C. Young, B. Cutler, A spatially augmented reality sketching interface for architectural daylighting design, *IEEE Trans. Vis. Comput. Graph.* 17 (1) (2009) 38–50, <https://doi.org/10.1109/TVCG.2009.209>.
- [57] Y. Sheng, T.C. Yapo, C. Young, B. Cutler, Virtual heliodon: Spatially augmented reality for architectural daylighting design, in: 2009 IEEE Virtual Reality Conference, IEEE, 2009, pp. 63–70, <https://doi.org/10.1109/VR.2009.4811000>.
- [58] J. Nasman, B. Cutler, Physical avatars in a projector-camera tangible user interface enhance quantitative simulation analysis and engagement, in: 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013, pp. 930–936, <https://doi.org/10.1109/CVPRW.2013.137>.
- [59] T. Machino, S. Iwaki, H. Kawata, Y. Yanagihara, Y. Nanjo, K.-i. Shimokura, Remote-collaboration system using mobile robot with camera and projector, in: Proceedings 2006 IEEE International Conference on Robotics and Automation, IEEE, 2006, pp. 4063–4068, <https://doi.org/10.1109/ROBOT.2006.1642326>.
- [60] S. Hirooka, H. Saito, Virtual display system using video projector onto real object surface, in: Proceedings of the 14th International Conference on Artificial Reality and Telexistence, 2004, pp. 305–310.
- [61] J.-P. Tardif, S. Roy, M. Trudeau, Multi-projectors for arbitrary surfaces without explicit calibration nor reconstruction, in: Fourth International Conference on 3-D Digital Imaging and Modeling, IEEE, 2003, pp. 217–224, <https://doi.org/10.1109/IM.2003.1240253>.
- [62] M. Ashdown, M. Flagg, R. Sukthankar, J.M. Rehg, A flexible projector-camera system for multi-planar displays, in: 2004 IEEE Conference on Computer Vision and Pattern Recognition Vol. 2, IEEE, 2004, <https://doi.org/10.1109/CVPR.2004.12> pp. II.
- [63] T. Takahashi, T. Kawano, K. Ito, T. Aoki, S. Kondo, Performance evaluation of a geometric correction method for multi-projector display using sift and phase-only correlation, in: 2010 IEEE International Conference on Image Processing, IEEE, 2010, pp. 1189–1192, <https://doi.org/10.1109/ICIP.2010.5651839>.
- [64] T. Yamanaka, F. Sakae, J. Sato, Adaptive image projection onto non-planar screen using projector-camera systems, in: 20th International Conference on Pattern Recognition, IEEE, 2010, pp. 307–310, <https://doi.org/10.1109/ICPR.2010.84>.
- [65] S. Jordan, M. Greenspan, Projector optical distortion calibration using gray code patterns, in: 2010 IEEE Conference on Computer Vision and Pattern Recognition Workshops, IEEE, 2010, pp. 72–79, <https://doi.org/10.1109/CVPRW.2010.5543487>.
- [66] J.-i. Jung, J. Cho, Screen adaptive geometric image calibration method for handheld video projector, in: 2010 Digest of Technical Papers International Conference on Consumer Electronics, IEEE, 2010, pp. 505–506, <https://doi.org/10.1109/ICCE.2010.5418843>.
- [67] A. Boroomand, H. Sekkati, M. Lamm, D.A. Clausi, A. Wong, Saliency-guided projection geometric correction using a projector-camera system, in: 2016 IEEE International Conference on Image Processing, 2016, pp. 2951–2955, <https://doi.org/10.1109/ICIP.2016.7532900>.
- [68] D. Lindlbauer, J.E. Grønbæk, M. Birk, K. Halskov, M. Alexa, J. Müller, Combining shape-changing interfaces and spatial augmented reality enables extended object appearance, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM, 2016, pp. 791–802, <https://doi.org/10.1145/2858036.2858457>.
- [69] D. Moreno, G. Taubin, Simple, accurate, and robust projector-camera calibration, in: 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, IEEE, 2012, pp. 464–471, <https://doi.org/10.1109/3DIMPVT.2012.77>.
- [70] K. Kim, M. Billingshurst, G. Bruder, H.B.-L. Duh, G.F. Welch, Revisiting trends in augmented reality research: a review of the 2nd decade of ismar (2008–2017), *IEEE Trans. Vis. Comput. Graph.* 24 (11) (2018) 2947–2962, <https://doi.org/10.1109/TVCG.2018.2868591>.
- [71] E. Olson, Apriltag: a robust and flexible visual fiducial system, in: 2011 IEEE International Conference on Robotics and Automation, IEEE, 2011, pp. 3400–3407, <https://doi.org/10.1109/ICRA.2011.5979561>.
- [72] C. Feng, V.R. Kamat, H. Cai, Camera marker networks for articulated machine pose estimation, *Autom. Constr.* 96 (2018) 148–160, <https://doi.org/10.1016/j.autcon.2018.09.004>.
- [73] C. Feng, V.R. Kamat, C.C. Menassa, Marker-assisted structure from motion for 3d environment modeling and object pose estimation, in: Construction Research Congress 2016, 2016, pp. 2604–2613, <https://doi.org/10.1061/9780784479827.259>.
- [74] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.
- [75] P.J. Besl, N.D. McKay, Method for registration of 3-d shapes, in: Sensor Fusion IV: Control Paradigms and Data Structures Vol. 1611, International Society for Optics and Photonics, 1992, pp. 586–606, <https://doi.org/10.1109/34.121791>.
- [76] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Courier Corporation, 2007.
- [77] R. Li, S. He, B. Skopljak, J. Jiang, P. Tang, A. Yilmaz, M. Banks, C. Oman, Development of a lunar astronaut spatial orientation and information system (laosios), in: Proceedings of the ASPRS 2010 Annual Conference, 2010, pp. 26–30.
- [78] R. Li, B. Skopljak, S. He, P. Tang, A. Yilmaz, J. Jiang, A spatial orientation and information system for indoor spatial awareness, in: Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, 2010, pp. 10–15, <https://doi.org/10.1145/1865885.1865889>.
- [79] G. Bradski, *The OpenCV library*, Dr. Dobb's J. Softw. Tools (2000).
- [80] Q.-Y. Zhou, J. Park, V. Koltun, Open3D: a modern library for 3D data processing, *arXiv* (2018), 1801.09847.
- [81] R.B. Rusu, S. Cousins, 3d is here: point cloud library (pcl), in: 2011 IEEE International Conference on Robotics and Automation, IEEE, 2011, pp. 1–4, <https://doi.org/10.1109/ICRA.2011.5980567>.
- [82] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, et al., Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera, in: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, 2011, pp. 559–568, <https://doi.org/10.1145/2047196.2047270>.
- [83] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, G. Ranzuglia, Meshlab: an open-source mesh processing tool, in: Eurographics Italian Chapter Conference Vol. 2008, Salerno, 2008, pp. 129–136, <https://doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136>.
- [84] B.O. Community, Blender - a 3D Modelling and Rendering Package, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018 [Online]. Available: <http://www.blender.org>.
- [85] V. Gomez-Jauregui, C. Manchado, D. Jesús, C. Otero, Quantitative evaluation of overlaying discrepancies in mobile augmented reality applications for aec/fm, *Adv. Eng. Softw.* 127 (2019) 124–140, <https://doi.org/10.1109/CVPR.2013.137>.
- [86] F. Herranz, K. Muthukrishnan, K. Langendoen, Camera pose estimation using particle filters, in: 2011 International Conference on Indoor Positioning and Indoor Navigation, IEEE, 2011, pp. 1–8, <https://doi.org/10.1109/IPIN.2011.6071919>.
- [87] P. Dollar, R. Appel, S. Belongie, P. Perona, Fast feature pyramids for object detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (8) (2014) 1532–1545, <https://doi.org/10.1109/TPAMI.2014.2300479>.