Camera-Projector-Robot Calibration for Human-Robot Collaboration

Alexandre Angleraud[1], Quentin Houbre[1], Niko Siltala[1] and Roel Pieters[1]

***Abstract*— Effective and safe human-robot collaboration relies on the clear communication of information relevant to the shared task. Contextual information, such as work areas, object placement locations and safety zones can be projected on the work space, observable by the human operator, in an unobtrusive way. To facilitate this, a projector-camera system needs to be installed alongside a robotic arm as perception and communication tool. As the camera, projector and robot all operate in their own work space, calibration is needed to align the workspaces and compensate for distortion effects. This paper describes the calibration processes and all required tools for such system, including calibration between projector and camera, robot end-effector and camera, and camera intrinsics. As outcome the calibration should find the relationship between projected points and the world coordinate frame. The approach is evaluated with two workcells; a collaborative robot workcell and an industrial robot workcell, each with different configurations for the projector-camera system. The results demonstrate the procedures for all three calibration methods in detail, showing that an accurate calibration can be obtained for both workcells. The software is available at https://github.com/cogrob-tuni/projector-interface.**

I. INTRODUCTION

Industry is increasingly looking for higher customization and lower production times, thereby aiming to integrate more flexible systems into production lines. To keep up with this changing demand the manufacturing domain needs to adapt. Robots are thus transitioning from being dedicated to mass production to operating on smaller batches. Furthermore, collaboration between robots and human operators offers new possibilities for increased flexibility of manufacturing processes. As a result, technologies that support this collaboration have seen a growing interest. Augmented Reality (AR) technologies [1], [2] can facilitate interaction between robots and human operators by helping to gain insight into robot operations. Projector-based AR, by overlaying digital elements directly in the real world, has the advantage of augmenting the whole workspace, not just the user's field of view and allows for interaction without the addition of wearables (see Fig. 1). This concept is referred to as spatially augmented reality (SAR) [3]. Application areas for AR projects in manufacturing include robot programming, task guidance, safety awareness, etc.

This paper presents and discusses the different calibration processes required for camera, projector and robot in order to achieve a functional and accurate projector-based SAR system (see Fig. 1). Accurate calibration enables corresponding coordinate frames between camera, projector and robot such that projected digital information aligns with the real environment as accurately as possible without the need for visual detection of objects or markers during operation. The paper also presents and explains the practical steps needed for calibration and identifies several challenges that must be tackled. Our suggested calibration process includes four calibration steps:

1) Intrinsic camera calibration
2) Camera hand-eye calibration
3) Projector-camera calibration
4) Calibration refinement

First, intrinsic camera calibration aims to identify the intrinsic parameters of a camera, to compensate for disturbances such as lens distortions. Second, camera hand-eye calibration aims to find the correspondences between the camera and the origin of world coordinates, i.e., the robot base coordinate frame. Third, projector-camera calibration aims to find the transformation between the camera and the projector, in image coordinates. Finally, any residual misalignment between projection and world coordinates is captured by a calibration refinement step. The calibration procedures are evaluated and validated by two different setups; a collaborative robot setup and industrial robot setup, each with different camera-projector configurations. The motivation behind the developments of the projector-based SAR system is to enhance communication between human and robot in collaborative scenarios, ideally leading to more

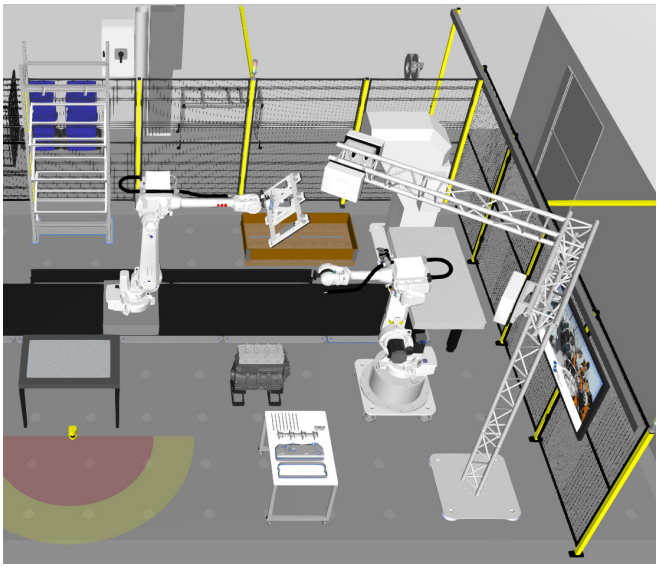


Fig. 1: Industrial robot workcell with cameras and projectors mounted on a truss, to enable projector-based Human-Robot Collaboration.

[1]Unit of Automation Technology and Mechanical Engineering, Tampere University, Tampere, Finland; firstname.surname@tuni.fi

intuitive and more explainable collaboration.

In summary, this paper offers the following contributions:

- Detailed procedures for camera-projector-robot calibration including calibration between projector and camera, robot end-effector and camera, and camera intrinsics.
- An experimental analysis of our methods with two industrially relevant scenarios, including both collaborative and industrial robot

II. Literature review

A. Augmented Reality in manufacturing

Augmented Reality (AR) is commonly used to augment the environment in which robots and workers operate [2]. Operators can wear hardware, in the form of body-attached displays or rely on spatial displays, such as projectors. Fixed projection-based Spatial Augmented Reality (SAR) typically assumes a static distance between the projector and the projection surface. Projected elements can be purely visual or can be used to interact with the system, creating a Mixed Reality (MR) experience. Virtual elements are then not just information being displayed but coexist with real-world objects. In some cases, systems allow users to interact with digital information through real physical objects. This is referred to as a Tangible Graphical User interface (TUI) [4]. Frameworks such as reacTIVision [5] allow for the development of TUIs and some proprietary tools are commercially available [4]. However, most projection-based systems rely on custom-built software with specific hardware configurations, making quick deployment and reuse of the technology challenging.

Depth sensing has become a popular approach for monitoring shared workspaces [6], [7]. This combines a robot model with elements perceived during runtime, such as human hands or objects to manipulate, allowing to calculate the distance between the robot and humans or human body parts and projected elements. [6] shows a method to maintain safety and offer interactive buttons to control the flow of robot actions. The authors divide safety violations in two cases. One if something gets in the way of the projected image, the other if something get in the way of the camera. A mask is created corresponding to the expected position of projections without any obstruction, comparing it to the actual state of the image. In [8], wearable AR and projector-based AR were compared to a traditional monitor-based display. AR was found advantageous, however, compared to current state-of-art in wearables, the projector-based AR was found to be more realistic.

Mobile projection systems have been explored as a more flexible alternative to fixed setups. For instance, [7] describes a mobile interface to interact with the robot action flow. Such interface has the benefit of being flexible and being usable in different setups. In [9] a mobile projector attached to a robotic arm for collaborative assembly was utilized, enabling assembly operations without any prior instructions. In a context where small batch sizes forces manufacturing lines to change, [10] addresses the issue of projection uncertainty.

Finally, works also exist dedicated to specific tasks and specific application areas for SAR, such as mobile industrial robot teams [11] and robot grasping [12].

These related works describe significant developments for SAR, however, often the importance of calibration is given little attention, or assumed as a problem easy to solve. In this work, we present all the calibration steps needed for SAR, as well as their different practical tools, such as calibration patterns, software and data collection.

B. Calibration techniques

A taxonomy of calibration techniques is provided in [13], identifying four criteria to take into account when deciding which calibration technique to use:

- Projection surface type or geometry - Projection mapping is the process of turning non-planar surfaces into projection surfaces.
- Static or dynamic displays - Moving displays require geometric calibration, considering the projector as an inverse camera.
- Ease of use of the calibration procedure.
- Accuracy - Higher accuracy is needed in case of robot tasks while lower accuracy is sufficient for information overlay.

The two main ways of calibrating a projector SAR are pixel-by-pixel mapping, using structured light patterns such as in [14] to compute point correspondences and estimate depth or finding the linear correction between two planes by calculating a homography matrix. Recent works commonly employ the approach of [15] in which a checkerboard pattern is placed in different positions and orientations and calibration parameters are computed using homographies. In [16], this method is extended to use a marker from the ARToolkit and compute a set of homographies.

[17] proposes a method for calibrating a projector using an RGB-D camera. Some methods fix the location of a printed calibration pattern and require the projector to be moved [18]. Others try to iteratively fit a projected pattern onto a printed pattern [19]. To calibrate multi-projector setups, [20] identifies two types or corrections needed - *intra* and *inter-projectors*. In other words, the calibration of the projector itself with respect to the display surface and the alignment of all projectors with each other, when boundaries do not match. Different levels of self-calibrating setups have been investigated [21]. Self-calibrating systems have the advantage of not requiring human intervention, saving time and effort during the original installation. It also ensures that if the parameters being calibrated change over time, the system remains accurate. Common approaches involve embedding cameras or sensors to detect calibration patterns or features on the projection surface automatically [22]. Some systems use iterative feedback loops in which the projector adjusts its output based on sensor data to correct misalignment or distortions dynamically [23].

III. Calibration Methods

The calibration process involves the coordinate frames of the robot base, which is also the world frame W, the camera C and the projector P, as depicted in Fig. 2. The goal of the calibration is to find the relationships between all three frames, meaning that a projected point on the surface plane can be expressed in world frame coordinates. In addition, calibration should compensate for the inaccuracies that occur due to lens distortions. The transformation chains in the eye-to-hand camera-projector-robot setup illustrate the known (solid line) and unknown (dashed line) transformations between components. The robot's end effector pose E relative to the world W is denoted by T_E^W, known from forward kinematics. T_W^C, represents the hand-eye transformation from world W to camera C, and T_B^E, denotes the transformation between the calibration board B and the robot's end-effector pose E. T_P^C, denotes the transformation between the projector P and the camera C, and T_S^C, denotes the transformation between the surface projected image I_P and the camera image I_C.

A. Intrinsic Camera Calibration

Intrinsic camera calibration is the process of determining a camera's internal parameters that affect how it projects 3D points onto a 2D image [15]. These parameters can be captured by a pinhole model [24] and include the focal length (f_x, f_y) and principal point (c_x, c_x), in matrix-form as:

$$K = \begin{bmatrix} f_x & 0/s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

Lens distortion can be taken into account to compensate for radial and tangential effects [25] . Radial distortion affects how points deviate from the optical center and can be modeled as:

$$r^2 = x^2 + y^2, \tag{2}$$
$$x' = x\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right), \tag{3}$$
$$y' = y\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right), \tag{4}$$

where (x, y) are the normalized (undistorted) coordinates, (x', y') are the radially distorted coordinates and k_1, k_2, k_3 are the radial distortion coefficients.

Tangential distortion occurs due to lens misalignment and can be modeled as:

$$x'' = x' + \left(2p_1 xy + p_2(r^2 + 2x^2)\right) \tag{5}$$
$$y'' = y' + \left(p_1(r^2 + 2y^2) + 2p_2 xy\right), \tag{6}$$

where: (x'', y'') are the tangentially distorted coordinates and p_1, p_2 are the tangential distortion coefficients.

B. Camera Hand-Eye Calibration

The goal of hand-eye calibration is to find the position of the camera with respect to the world or robot coordinate frame, denoted as the transformation T_W^C. Formally, the procedure aims to minimize a cost function [26], representing the discrepancy between the observed pose of the calibration target in the camera frame and the pose predicted via forward kinematics from the robot base through the estimated camera transform:

$$T_W^C = \min \sum_{i=1}^{N} \|B_i - T_W^C A_i\|^2, \tag{7}$$

where $B = T_E^W T_B^E$ is the pose of the calibration pattern expressed in robot frame and A is the pose of the same calibration object as observed by the camera.

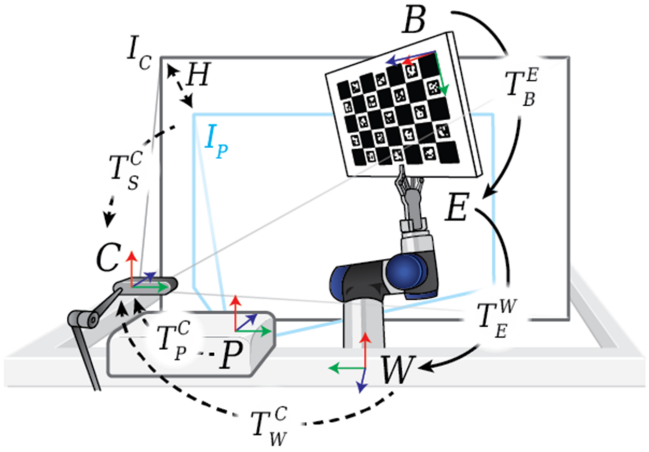


Fig. 2: Camera-Projector-Robot system depicting known (solid line) and unknown (dashed line) transformations between different coordinate frames.

C. Projector-Camera Calibration

Projector-camera calibration aims at finding the transformation T_P^C between points in the 2D image of the camera I_C and 3D points projected on a planar surface I_P. As T_P^C is difficult to obtain, T_S^C can be estimated by taking advantage of the planar relationship between the camera image I_C and the projector image I_P, which can be estimated by the homograpy H. As a transformation between two planes can be captured by a perspective projection, homography estimation was used to find this relationship:

$$x_p = H x_c, \tag{8}$$

where x_p is a point in the projector image and x_c is a point in the camera image. To estimate the homography matrix, there are different approaches [24]. Here, the Direct Linear Transform (DLT) is used, requiring at least 4 pairs of known correspondences. Using these point pairs, Eqn. (8) can be rearranged in the form $Ah = 0$, where Singular Value Decomposition (SVD) can be used to solve for h.

D. Calibration refinement

After completion of the different calibration steps for a robot–camera–projector system, residual systematic errors may persist due to imperfections in the calibration process. A compensation function can then be calculated to correct for such inaccuracies, taking into account the discrepancy between the actual observed positions of projected features

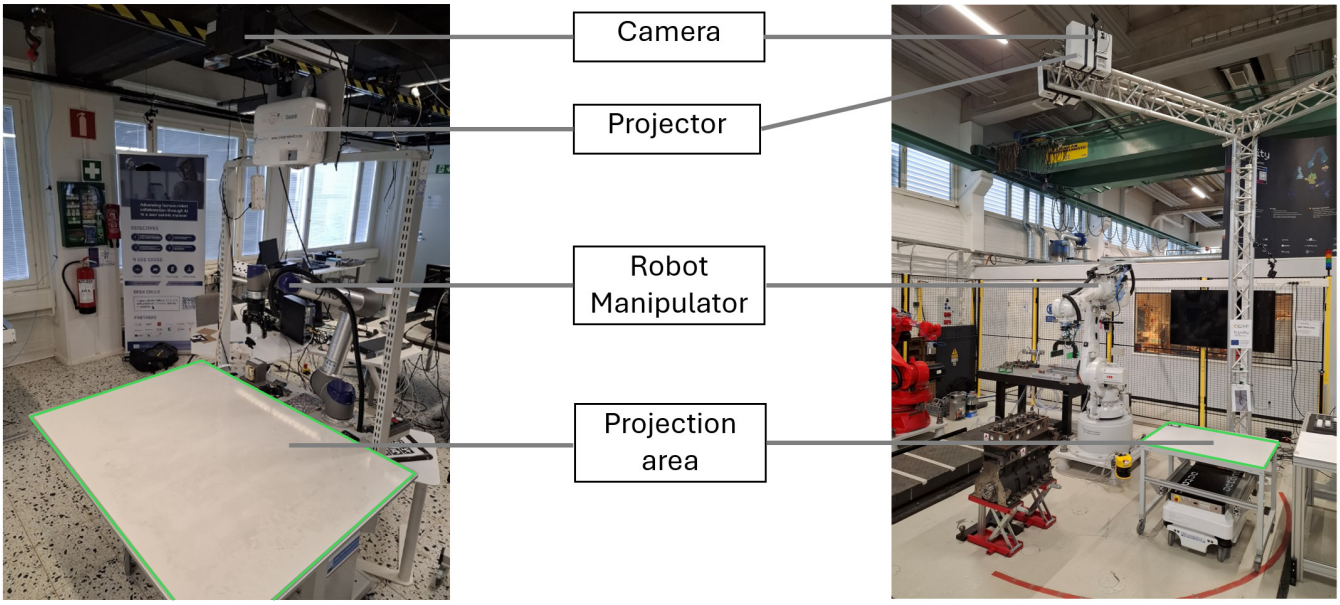


Fig. 3: Collaborative Robot (CR) workcell (left) and Industrial Robot (IR) workcell (right), annotated with relevant components and the projection area (green).

and their intended positions within the shared coordinate frame of the system. By analyzing the error vectors between corresponding points a compensation function can be derived as a transformation that minimizes these residuals, such as polynomials or Radial Basis Functions (RBF). This correction is then applied to subsequent projection operations to improve the overall accuracy of the system.

IV. Implementation

A. Experimental workcells

To evaluate our developments, two human-robot collaborative scenarios were prepared, one with a collaborative robot (UR5) and one with an industrial robot (ABB IRB 4600). Both scenarios use the same camera (Azure Kinect) but different projectors with same 1920 by 1080 pixels resolution; BenQ MW550 projector with a throw ratio of 1.55:1 to 1.7:1 (pixel size $= 0.39 - 0.44mm$), for the collaborative robot workcell and Optima DZ500 projector with a throw ratio of 1.4:1 to 2.4:1 (pixel size $= 0.72 - 1.23mm$), for the industrial robot workcell. The robot workcells can be seen in Fig. 3, which further emphasizes their differences. In particular, the collaborative robot workcell has relatively short distance between projector and surface ($\approx 1.3m$), while the industrial robot workcell has larger distance ($\approx 3.3m$). An additional difference between both workcells is their projection area. While for the collaborative robot workcell the projection area covers a large share of the image, for the industrial robot workcell the area is much smaller, owing to the height of the installed projector and camera.

Our developments are implemented through ROS noetic, with the transformation tree (tf) defining all relevant coordinate frames. The calibration results are exported with formats that can be loaded by the projector interface. Scikit-learn [27] was used for the calibration refinement. All software will be provided open-source after publication, including all required calibration tools, i.e., external libraries, calibration patterns and image processing software, and graphical user interfaces to streamline the different calibration steps.

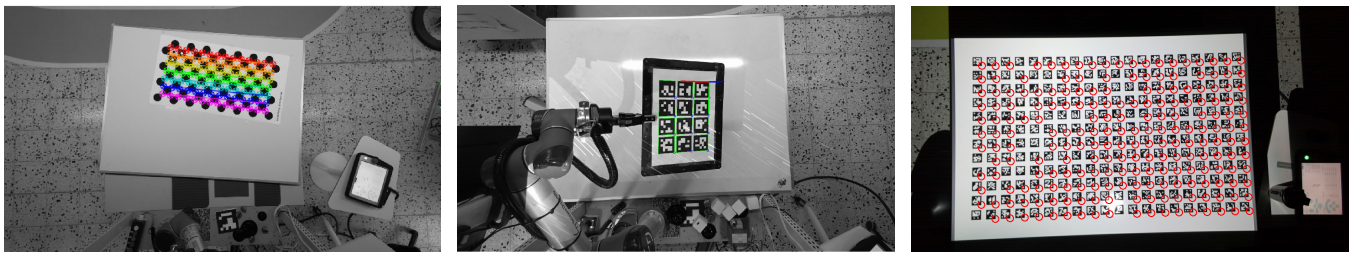
B. Data collection methodology

A checkerboard pattern was used for intrinsic camera calibration as depicted in Fig. 4a and 4d. Aruco [28] markers were used for hand-eye calibration, as depicted in Fig. 4b and 4e. Both were mounted on the robot's end effector and moved around to collect a suitable number of calibration samples. To determine the homography matrix, visual markers were projected onto the projection surface and detected as 2D coordinates with respect to the camera coordinate frame, as depicted in Fig. 4c and 4f. The correspondences between individual 2D marker points are maintained such that two arrays are obtained, one with points from the camera image and one with points from the projector image, respectively.

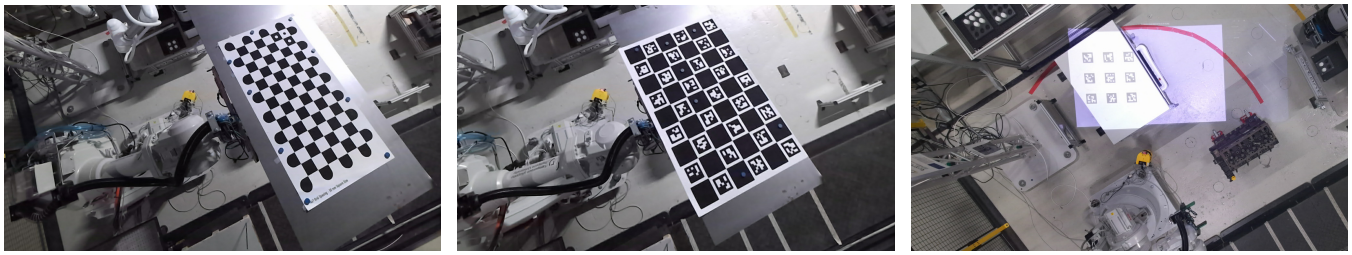
To evaluate the calibration steps, a set of circles were projected onto the projection surface and the robot was sent to these same locations to draw a circle with a pen. For each circle, the distance between the center point of the drawn circle and the projected circle is measured. From these values, the Root Mean Square Error (RMSE) between corresponding circles in I_c on the workspace can be calculated:

$$e_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|\mathbf{p}_i^{(p)} - \mathbf{p}_i^{(r)}\right|^2}, \tag{9}$$

where $p_i^{(p)}$ and $p_i^{(r)}$ are the coordinates of the $i-th$ circle in the projected coordinate frame and robot frame, respectively, $|\cdot|$ denotes the Euclidean norm, and N is the total number of circle correspondences.



(a) (b) (c)



(d) (e) (f)

Fig. 4: Calibration images. Top row depicts the intrinsic camera calibration (a), the hand-eye calibration (b) and the projector-camera calibration (c) for the collaborative robot (CR) workcell. Bottom row depicts the intrinsic camera calibration (d), the hand-eye calibration (e) and the projector-camera calibration (f) for the industrial robot (IR) workcell.

TABLE I: Calibration data collection for industrial (IC) and collaborative robot workcells (CC).

Calibration	**Intrinsic camera calibration**	**Hand-eye calibration**	**Projector-camera calibration**	**Calibration refinement**
Method	Reprojection error minimization	Reprojection error minimization	Homography estimation	compensation function
Pattern	Checkerboard	Aruco/Charuco markers 6D poses	Aruco markers 2D points	Projected and robot drawn circles
Samples CR	30 images	10 images, 3×4 marker pattern	1 image, 12×20 markers	12 circles 5 image
Samples IR	30 images	10 images, 3×5 marker pattern	1 image 12×20 markers	12 circles 5 images

The coordinates of each circle in robot space are transformed using the results from the calibration steps to obtain the center point in projector space. First, the 3D pose in robot space is transformed to 3D pose in camera frame. Then, that pose is projected onto 2D camera space before using the homography matrix to obtain its corresponding pixel value in the projector image. Note that in practice, for the robot to grasp the marker, a special end effector was designed and 3D printed. Therefore, an extra calibration step was necessary to obtain the tool center point (TCP). For the collaborative robot workcell, this is done through the teaching pendant, which requires moving the TCP around the same 3D point and recording four different poses. A similar approach was utilized for the industrial workcell. Image processing was finally used to process the images and calculate the center point of both sets of circles in the projected frame and robot frame.

The calibration data collection details are summarized in Table I.

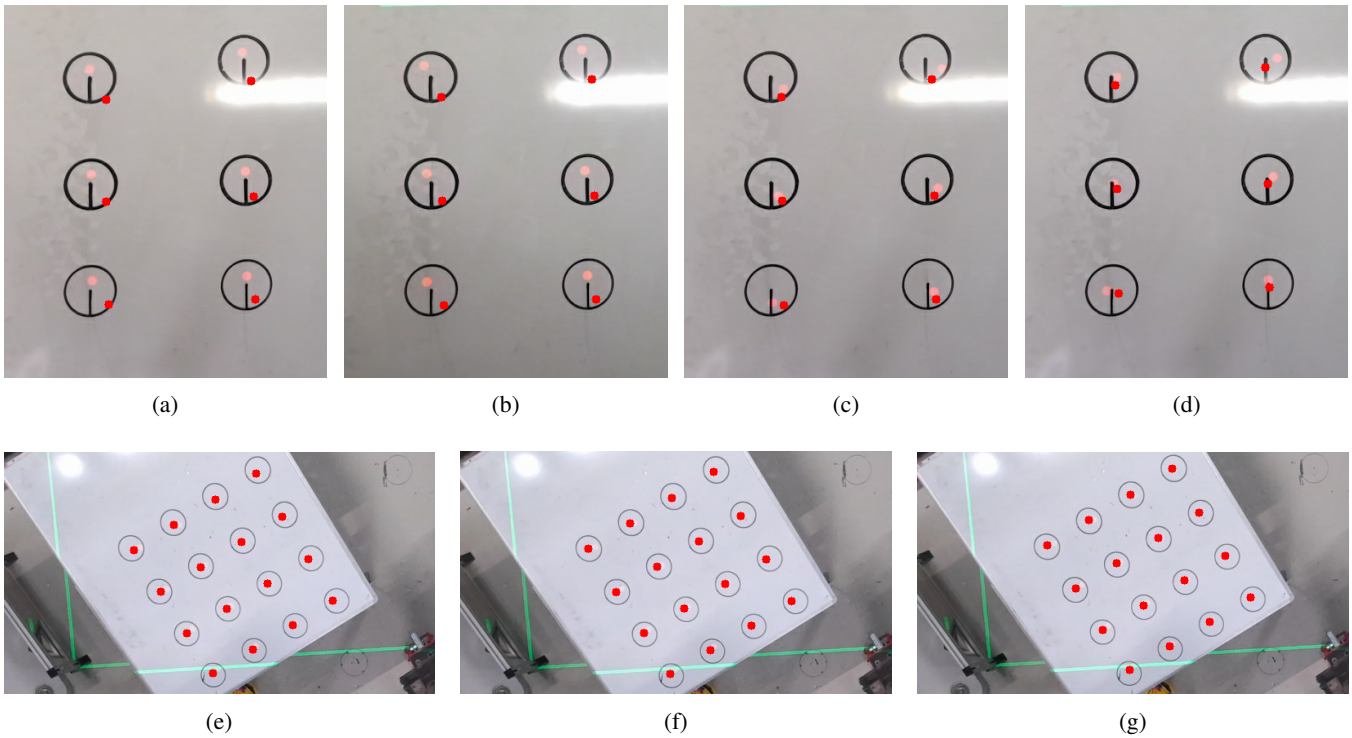
V. Results

A. Individual calibration results

The results for each of the calibration steps are reported in Table II and visualized in Fig. 5, indicating how each step influences the final projection accuracy. In the figure, e_{RMSE} was calculated according to Eqn. (9) with the circle center and the projected red point. Overall, each calibration step (i.e., intrinsic, hand-eye and projector-camera calibration) leads to an error of roughly < 10 mm for the collaborative workcell and roughly < 5 mm for the industrial workcell, indicating a well executed calibration procedure.

As expected, intrinsic camera calibration effects the accuracy in a non-uniform way, as shown by the non-uniform distortion effect for the set of project circles (see Fig. 5a).

Despite the accuracy of each of the three calibration steps, still an error remains (e_{RMSE}) before the calibration refinement of 22 px ($\approx 9mm$) for the collaborative workcell and 3.4 px ($\approx 3mm$) for the industrial workcell. Calibration refinement thus aims to capture this disturbance by finding a



(a) (b) (c) (d)

(e) (f) (g)

Fig. 5: Calibration results for the CR workcell (top row) and the IR workcell (bottom row). The black circle is drawn by the robot and the red point indicates the projected point. The influence on accuracy is depicted for intrinsic calibration (a), hand-eye calibration (b, e), projector-camera calibration (c, f) and calibration refinement (d, g).

function and its parameters that disturb the projected image. For both workcells the functions that returned the best results were polynomial models. As final results, the calibration steps for both workcells achieved an error e_{RMSE} of 8.2 px ($\approx 3mm$) for the collaborative workcell and 3.4 px ($\approx 4mm$) for the industrial workcell. Considering that the size of a single pixel can be as small as 0.4mm for the collaborative workcell and 0.7mm for the industrial workcell, we conclude that these are satisfactory results.

B. Workcell specific calibration results

The baseline error for the collaborative workcell is significantly higher than for the industrial workcell. The industrial workcell presents errors that are more localized and therefore easier to capture. After hand–eye calibration, the error is already reduced to below 5 mm. In this case, the homography calibration almost perfectly captured the transformation so the polynomial refinement did not lead to any real improvement. In contrast, for the collaborative workcell the polynomial correction step was impactful, reducing the reprojection error by nearly a factor of three. Despite these differences in sources of inaccuracy, the final errors are comparable across both workcells. This shows the importance of understanding which calibration steps address which sources of inaccuracy in order to achieve robust results.

C. Projection use cases

Contextual information projected by spatial augmented reality serves to provide contextual information to the operator, such as work areas, object placement locations, and safety zones, and offer collaborative functionalities, such as virtual buttons and task instructions. Accurate calibration thus ensures that projected information aligns with correct work areas and placed objects, without requiring sophisticated perception tools.

Fig. 6 depicts different use cases for the collaborative robot workcell (a) and the industrial workcell (b), where virtual buttons are projected on a table within the projection area (green square). In addition, Fig. 6a shows projected areas (small green square) in which objects can be placed, highlighting that the projected information and the world coordinate frame are aligned. The difference in complexity between the use cases can be seen by the difference in distance for projection; for the CR workcell the projection area covers a large part of the image, as the distance from projector to table is around $\approx 1.3m$, while for the IR workcell the distance is around $\approx 3.3m$, resulting in a small projection area and a configuration where the (movable) table can partly leave the projection area.

VI. Discussion

Calibration is a crucial and well-studied topic in robotics and vision-related fields, as demonstrated by the many solutions available in literature. Despite these, with several calibration steps in sequence, identifying where an inaccuracy or error has slipped in, or how much data is needed to capture all disturbances, is not an easy task. In this work, we aimed to provide an overview of all calibration steps needed, reported

TABLE II: Root-mean-square error (RMSE) after calibration steps for industrial (IC) and collaborative robot workcells (CC).

	Root Mean Square Error (RMSE)															
Calibration	Original camera calibration				Intrinsic camera calibration				Hand-eye calibration				Calibration refinement			
	IC		**CC**		**IC**		**CC**		**IC**		**CC**		**IC**		**CC**	
Method	**[px]**	**[mm]**	**[px]**	**[mm]**	**[px]**	**[mm]**	**[px]**	**[mm]**	**[px]**	**[mm]**	**[px]**	**[mm]**	**[px]**	**[mm]**	**[px]**	**[mm]**
Error minimization	4.26	5.2	23.0	10.1												
Error minimization					4.1	5.0	22.5	9.9								
Error minimization									3.4	4.2	22.0	9.7				
Polynomial correction													**3.4**	**4.2**	**8.2**	**3.6**

(a) (b)

Fig. 6: Spatial Augmented Reality (SAR) for the Collaborative Robot (CR) workcell (a) and the Industrial Robot (IR) workcell (b). The largest green square indicates the projection area, while the circular buttons indicate the virtual user interface buttons for human-robot collaboration.

the data captured and analyzed how each step influences the final accuracy.

The main limitation of our work is that there is no single solution that works for all camera-projector-robot setups. Each calibration step might need a different algorithm to achieve optimal accuracy as different components might be affected by different disturbances.

The presented approach relies on a planar surface for projection, making it suitable for workspaces such as tabletops. Projection on non-planar surfaces can lead to blur or misalignment, if the depth variation of the surface is too large. The solution to this would be to map the surface of the projection area and finding a suitable pixel-wise transformation that can be applied to compensate for the surface variations.

The calibration procedures presented in this paper are primarily for static setups, i.e, the camera, projector, and robot maintain fixed relative positions during operation. Expanding the system to handle dynamic projection areas in the environment would increase its applicability but would require recalibration of certain steps (e.g., camera-projector calibration). Alternatively, self-calibration could be a potential solution that, at regular intervals, performs re-calibration, allowing for changes in the position of the camera, projector and projection area. As can be seen in Fig. 6b, small variations in table positions (linear translation and rotation) can already be captured by the detection of a marker attached to the tabletop surface.

Finally, as depicted in Fig. 6, projection of information and virtual buttons is possible without active perception tools. Future work will enable collaboration between the human and the projector system by perception of the human such that projection acts as an active user interface.

VII. CONCLUSIONS

This paper presented the different calibration routines needed for a camera-projector-robot system, including calibration between projector and camera, robot end-effector and camera, and camera intrinsics. An additional refinement step to compensate for misalignment between projection and world coordinates is then applied by measuring the error between projected and robotically drawn circles. The calibration steps are evaluated with two workcells, one with a collaborative robot and one with an industrial robot.

Results demonstrate that a projector-based Spatial Augmented Reality system can be successfully used to overlay

contextual information in shared workspaces and assist human operators in manufacturing tasks.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101017141 and 101135708.

REFERENCES

[1] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Transactions on Information and Systems*, vol. 77, no. 12, pp. 1321–1329, 1994.

[2] R. Suzuki, A. Karim, T. Xia, H. Hedayati, and N. Marquardt, "Augmented Reality and robotics: A survey and taxonomy for AR-enhanced human-robot interaction and robotic interfaces," in *CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–33.

[3] R. Raskar, G. Welch, and H. Fuchs, "Spatially augmented reality," in *International workshop on Augmented Reality: placing artificial objects in real scenes*, 1999, pp. 64–71.

[4] J. Spadaccini and H. McDonald, "The evolution of tangible user interfaces on touch tables: New frontiers in UI & UX design," Ideum, Tech. Rep., 2017.

[5] M. Kaltenbrunner and R. Bencina, "reacTIVision: a computer-vision framework for table-based tangible interaction," in *International conference on Tangible and embedded interaction*, 2007, pp. 69–74.

[6] C. Vogel, C. Walter, and N. Elkmann, "Safeguarding and supporting future human-robot cooperative manufacturing processes by a projection-and camera-based technology," *Procedia Manufacturing*, vol. 11, pp. 39–46, 2017.

[7] D. Monakhov, J. Latokartano, M. Lanz, R. Pieters, and J.-K. Kämäräinen, "Mobile and adaptive user interface for human robot collaboration in assembly tasks," in *International Conference on Advanced Robotics (ICAR)*, 2021, pp. 812–817.

[8] A. Hietanen, R. Pieters, M. Lanz, J. Latokartano, and J.-K. Kämäräinen, "AR-based interaction for human-robot collaborative manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 63, p. 101891, 2020.

[9] C. Lengenfelder, C. Frese, A. Zube, M. Voit, and J. Beyerer, "A cooperative HCI assembly station with dynamic projections," in *International Symposium on Robotics*, 2020, pp. 1–8.

[10] J. Pang and P. Zheng, "ProjecTwin: A digital twin-based projection framework for flexible spatial augmented reality in adaptive assistance," *Journal of Manufacturing Systems*, vol. 78, pp. 213–225, 2025.

[11] J. Berg, A. Lottermoser, C. Richter, and G. Reinhart, "Human-Robot interaction for mobile industrial robot teams," *Procedia CIRP*, vol. 79, pp. 614–619, 2019.

[12] L. L. Gong, S. K. Ong, and A. Y. C. Nee, "Projection-based Augmented Reality interface for robot grasping tasks," in *International Conference on Robotics, Control and Automation*, 2019, p. 100–104.

[13] R. B. Madhkour, M. Mancas, and T. Dutoit, "A taxonomy of camera calibration and video projection correction methods," *EAI Endorsed Transactions on Creative Technologies*, vol. 2, no. 2, 2 2015.

[14] Z. Li, Y. Shi, C. Wang, and Y. Wang, "Accurate calibration method for a structured light system," *Optical Engineering*, vol. 47, no. 5, pp. 053 604–053 604, 2008.

[15] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[16] S. Audet and M. Okutomi, "A user-friendly method to geometrically calibrate projector-camera systems," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2009, pp. 47–54.

[17] R. B. Madhkour, M. Mancas, and B. Gosselin, "Automatic geometric projector calibration - application to a 3D real-time visual feedback," in *International Conference on Computer Vision Theory and Applications*, 2013, pp. 420–424.

[18] H. Anwar, I. Din, and K. Park, "Projector calibration for 3D scanning using virtual target images," *International Journal of Precision Engineering and Manufacturing*, vol. 13, pp. 125–131, 2012.

[19] I. Martynov, J.-K. Kamarainen, and L. Lensu, "Projector calibration by inverse camera calibration," in *Scandinavian Conference on Image Analysis*, 2011, pp. 536–544.

[20] M. Brown, A. Majumder, and R. Yang, "Camera-based calibration techniques for seamless multiprojector displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 2, pp. 193–206, 2005.

[21] F. Li, H. Sekkati, J. Deglint, C. Scharfenberger, M. Lamm, and D. a. Clausi, "Simultaneous projector-camera self-calibration for three-dimensional reconstruction and projection mapping," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 74–83, 2017.

[22] M. Son and K. Ko, "Multiple projector camera calibration by fiducial marker detection," *IEEE Access*, vol. 11, pp. 78 945–78 955, 2023.

[23] M. T. Ibrahim, M. Gopi, and A. Majumder, "Self-calibrating dynamic projection mapping system for dynamic, deformable surfaces with jitter correction and occlusion handling," in *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2023, pp. 293–302.

[24] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.

[25] D. C. Brown, "Decentering distortion of lenses," *Photogrammetric Engineering*, vol. 32, pp. 444–462, 1966.

[26] D. Allegro, M. Terreran, and S. Ghidoni, "Multi-camera hand-eye calibration for human-robot collaboration in industrial robotic workcells," *IEEE Robotics and Automation Letters*, vol. 9, no. 11, pp. 9852–9859, 2024.

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[28] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.