Close-range real-time camera pose estimation and AR-guided alignment for large-scale industrial components

ISPRS Journal of Photogrammetry and Remote Sensing 235 (2026) 58–71

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Close-range real-time camera pose estimation and AR-guided alignment for large-scale industrial components

Francesco Messina [a], Alessio Martino [b], Francesca Matrone [b,*], Andrea Maria Lingua [b], Pablo Puerto [c]

[a] Politecnico di Torino, Department of Electronics and Telecommunications (DET), PoliTo Interdepartmental Center for Service Robotics (PIC4SeR), Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[b] Politecnico di Torino, Department of Environment, Land and Infrastructure Engineering (DIATI), Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[c] IDEKO, Basque Research and Technology Alliance (BRTA), Arriaga Kalea 2E, 20870 Elgoibar, Spain

ARTICLE INFO

Keywords:
Real-time camera pose estimation
Industrial metrology
Zero-Defect manufacturing
Augmented Reality

ABSTRACT

Precise alignment of large-scale industrial components on machine tools is essential to ensure machining accuracy, product quality, and process efficiency. Errors introduced during the setup phase can propagate throughout the manufacturing process, often resulting in costly rework. Conventional alignment methods rely on laser tracker systems, which, despite their high precision, require specialized equipment, skilled operators, and long setup times, making them expensive and operationally demanding. To overcome these limitations, this work presents a real-time collaborative camera pose estimation framework that simplifies and accelerates the alignment process. The proposed solution integrates predictive simulation, acquisition trajectory planning, and augmented reality (AR) to enable fast, accurate, and intuitive alignment, even for non-expert users. The system built upon the IDEKO VSET solution was further developed within the TACCO project, co-funded by EIT Manufacturing and the European Union. The framework starts from high-fidelity 3D models and employs a Monte Carlo-based optimization strategy to determine the optimal placement of auxiliary components, including calibrated scale bars, coded targets (igloos), and cross reference frame. Configurable image acquisition strategies allow users to balance accuracy and computational complexity. Candidate configurations are evaluated through a least-squares bundle block adjustment simulation, based on collinearity equations and a Structure from Motion approach, enabling uncertainty propagation analysis and the generation of 95% confidence error ellipsoids prior to data acquisition. A key innovation lies in the system's ability to translate optimized planning solutions into real-time immersive AR guidance. Operators are guided in the placement of auxiliary components and camera positioning through an AR head-mounted display, ensuring complete and accurate data capture. The proposed approach achieves alignment accuracies closer to laser tracker systems while significantly reducing setup time, cost, and dependence on specialized personnel, offering a scalable and cost-effective alternative for industrial component alignment.

1. Introduction

Metrology plays a fundamental role in the industrial sector, especially when working with large-scale components. Such components are commonly used in mission-critical sectors, including aerospace, shipbuilding, energy production, and transportation infrastructure. In such domains, raw parts often exhibit linear dimensions extending to tens of meters and demand exceptional machining precision to ensure functionality, safety, and regulatory compliance. Due to their role in primary and safety-critical systems, these components require rigorous control and planning throughout the entire manufacturing pipeline, encompassing initial machining, intermediate alignment procedures, and final assembly and finishing operations.

Large-volume metrology (LVM) is essential to ensure the correct positioning and accurate measurement of elements subject to milling or other machining operations. The current state of the art in LVM includes using expensive equipment such as Coordinate Measuring Machines (CMMs) or laser trackers. These systems provide high accuracy measures but require substantial investment, dedicated facilities, trained personnel, and its considerably time-expensive.

An effective approach to addressing these challenges lies in photogrammetry, which enables the acquisition of accurate and reliable

* Corresponding author.
E-mail address: francesca.matrone@polito.it (F. Matrone).

https://doi.org/10.1016/j.isprsjprs.2026.02.034
Received 8 September 2025; Received in revised form 26 January 2026; Accepted 27 February 2026
Available online 6 March 2026

measurements from images (Bösemann, 2005; Wang et al., 2013; Galantucci et al., 2016). By employing high-resolution cameras in combination with specialized algorithms (Girelli et al., 2022), it is possible to reconstruct detailed three-dimensional models of industrial components efficiently and without the need for highly trained personnel, through the rapid acquisition of multiple images of the object.

This contribution addresses a portion of the TACCO project, aimed at the rapid and accurate setup of defect-free parts. The objective is to design an intuitive and visually appealing Human–Machine Interface (HMI) to guide non-expert users in the LVM of industrial components up to 30 m in size. The system aims to achieve a level of relative accuracy comparable to that of the conventional high-precision methods previously discussed. The project, co-funded by the European Union under the European Institute of Innovation and Technology (EIT) Manufacturing initiative, represents an enhancement of the VSET system presented in Mendikute et al. (2017), which utilizes calibrated, coded, and retro-reflective auxiliary elements for the precise acquisition of machining point positions.

Building upon the previous work presented by Messina et al. (2024) which validated the proposed methodology through simulated scenarios, this study extends the validation to a real-world application involving a medium-sized industrial component. By comparing the photogrammetric results with high-precision measurements acquired using a Leica Absolute Tracker AT403, this study demonstrates the practical applicability, robustness, and accuracy of the TACCO system under operational conditions. The findings confirm that the methodology not only meets the target accuracy thresholds but also provides a reliable alternative to traditional large-volume metrology systems in complex industrial environments.

The main contributions of this work can be summarized as follows:

- definition of a simulation-based framework for the optimal placement of auxiliary photogrammetric components and camera poses in large-volume industrial alignment tasks;
- integration of a Monte Carlo least-squares simulation to analytically estimate uncertainty propagation and error ellipsoids before data acquisition;
- use of an augmented reality interface to guide non-expert operators during both auxiliary component placement and image acquisition, reducing setup time and operator dependency;
- experimental validation of the proposed methodology through comparison with high-precision laser tracker measurements in a real industrial scenario.

2. Related works

This section reviews the main state-of-the-art approaches related to industrial photogrammetry for large-volume metrology, uncertainty modeling, and the use of augmented reality for measurement and alignment tasks.

The accurate measurement of industrial components through photogrammetry has been studied for many years, Fraser (1988), even for large-scale elements, although traditionally limited by the need to use film and opto-mechanical systems. The impact of computer vision in recent decades has significantly reduced the time required for both camera calibration and data processing, as demonstrated by Remondino and Fraser (2006) and Balletti et al. (2014). This has been made possible by developing automatic self-calibration techniques, and improving the processes used in close-range photogrammetry (Luhmann et al., 2016). Simultaneous calibration of multi-camera measurement systems also uses auxiliary elements, such as scale bars placed around the volume, to obtain self-calibrated bundle adjustment solutions (Luhmann, 2010; Franceschini et al., 2014; Sun et al., 2018, 2019). The use of these auxiliary components, along with targets placed on the object, has played a crucial role in measuring large raw parts for high-precision machining tasks as reported in Puerto et al. (2022).

Through the calibration and orientation of multi-camera systems for LVM, it is possible to improve measurement accuracy significantly, achieving precision levels approaching those of a laser tracker (LT) but at a much lower cost (Leizea et al., 2023).

2.1. Photogrammetric solutions for industrial metrology

In the state-of-the-art, diverse photogrammetric methodologies are applied to industrial metrology, each addressing specific challenges associated with large-volume and high-accuracy measurements in complex industrial environments.

One of the most practical contributions is offered in the study on in-process portable photogrammetry with optical targets (Leizea et al., 2023), which proposes an efficient and cost-effective alternative to traditional pre-calibrated camera systems. By employing low-cost cameras and adhesive optical markers, combined with real-time bundle adjustment and self-calibration, the method achieves high measurement precision without relying on expensive metric equipment. This approach stands out for its adaptability in dynamic, large-scale scenarios, such as machining alignment or quality control of oversized components, with computational efficiency and ease of integration within existing industrial workflows. Nonetheless, a key limitation is the increasing computational burden associated with intermediate bundle adjustments as the number of images and markers grows, particularly in very large measurement volumes. This aspect is common with the first version of the TACCO system, the above-mentioned VSET system (Mendikute et al., 2017), and it has been faced in this contribution (Section 3) trying to minimize the number of images with comparable accuracy.

In contrast, the methodology proposed by Puerto et al. (2022) places stronger emphasis on standardization and traceability. Their framework evaluates photogrammetric accuracy based on the (*Verein Deutscher Ingenieure*/Association for Electrical, Electronic & Information Technologies) VDI/VDE 2634 Part 1 standard (a protocol guiding the evaluation of the accuracy of optical 3D measuring systems) using a highly controlled reference environment. This study provides a rigorous and reproducible method for assessing system performance, showing that careful planning of camera network geometry and the placement of auxiliary markers significantly impacts the quality of the measurement results. A primary advantage of this approach is its metrological robustness and predictive accuracy modeling, although it is highly sensitive to operator expertise and requires a carefully maintained experimental setup. This criticality led our proposal to approach a practical and handy solution as the AR (Augmented Reality) system, to reduce the impact of operator skill. Additionally, the use of Monte Carlo simulations to anticipate uncertainty levels introduces a powerful predictive capability that further strengthens the reliability of the method, providing the basis for our workflow.

An experimental investigation into markerless photogrammetry expands the discussion toward more flexible and user-friendly alternatives (Prieto et al., 2024). In this context, the feasibility of applying photogrammetry without artificial targets is assessed using textured and non-textured models. Markerless methods are notably advantageous when physical access to the object is restricted or when dense 3D reconstructions are prioritized over strict dimensional accuracy. They require less preparation and fewer physical components, thus reducing operational constraints and costs. However, the results consistently demonstrate that accuracy levels are significantly inferior to marker-based systems, particularly on reflective or texture-poor surfaces. The technique also demands substantial computational resources, and the variability introduced by automatic feature matching reduces repeatability and traceability, which are critical in industrial metrology. For these reasons, in this study the use of markerless photogrammetry has been discarded.

The broader technical and methodological landscape is effectively summarized in Luhmann's comprehensive review of close-range photogrammetry (Luhmann, 2010). This reference study categorizes systems as off-line or on-line, highlighting that off-line systems, typically using digital Single Lens Reflex cameras and self-calibration, offer the highest achievable accuracies for static scenes. On-line systems, although less accurate due to real-time constraints and sensor limitations, are more suited to dynamic applications such as robot tracking or assembly monitoring. Luhmann underscores that accuracy is not solely a function of hardware, but largely dependent on the network geometry, camera calibration models, and software capabilities, including bundle adjustment and error propagation analysis.

Broadly speaking, no single methodology universally outperforms the others, each serves a distinct niche depending on the specific constraints of the industrial application. In-process marker-based photogrammetry is optimal for real-time, large-scale measurements where speed and cost-efficiency are critical. Markerless photogrammetry, though less precise, is highly suited to preliminary inspections or non-contact scanning when measurement traceability is less critical. Finally, the principles laid out by Luhmann provide essential guidance for system design, calibration, and integration, framing photogrammetry as a versatile yet demanding technology whose successful deployment depends on aligning methodological strengths with operational objectives.

2.1.1. Network design

In this framework, the proposed methodology took inspiration from the photogrammetric network design. This is fundamental to achieving high accuracy and reliability in industrial and close-range applications, because the camera configuration and control layout directly condition the bundle adjustment. Fraser's classic work on network design considerations for non-topographic photogrammetry formalizes the problem in terms of zero-, first-, second- and third-order design (datum, configuration, weight and densification), and shows through simulations that strongly convergent imagery, well-distributed control, and adequate redundancy are essential to minimize parameter correlations and maximize precision (Fraser, 1984). These principles are summarized and extended in Fraser (1996), which provides practical guidelines for camera placement and control distribution in industrial measurement scenarios. Building on this foundation, Mason's works (Mason, 1995a,b) introduced expert-system approaches to network design, encoding expert rules for choosing camera stations and weights, ordering the design tasks according to Fraser's zero-/first-/second-order framework and automating the search for near-optimal configurations. Recent work in industrial photogrammetry has revisited these ideas in the context of automated systems and repeated deformation measurements. Shen et al. for example, analyze "network configuration and repeatability precision" and introduce the concept of network consistency to explain how changes in configuration between epochs affect repeatability, confirming that carefully controlled, Fraser-style network geometry remains critical in modern high-precision industrial metrology (Shen et al., 2023).

2.2. Augmented reality

AR systems enable the overlay of digital content (e.g., instructions, measurement data, or simulation results) onto the physical environment in real time, supporting tasks such as assembly guidance, inspection, and quality control.

Recent studies indicate that AR is evolving from a purely visual technology toward a metric-aware interface tightly coupled with photogrammetric and geospatial data. The work of Bolkas et al. (2024) highlights the growing maturity of AR and Mixed Reality (MR) systems as operational interfaces, demonstrating their effectiveness in integrating complex spatial data, instruments, and digital models into real-world workflows, an aspect that directly supports industrial and construction-related applications beyond education. For example, in MR construction applications, accurate spatial alignment is identified as a critical requirement: Einizinab et al. (2025) demonstrate that SLAM (Simultaneous Localization And Mapping) drift severely affects BIM-reality registration and propose camera pose refinement based on wireframe reconstruction to ensure stable AR visualization in inspection tasks. Immersive environments have also been explored for quantitative spatial analysis, as shown by Subramaniyam et al. (2024), who employ virtual reality to assess UAV-derived 3D model accuracy through interactive GCP-based (Ground Control Points) error evaluation, highlighting the benefits of immersive visualization for understanding metric uncertainty. Georeferenced AR systems integrating GNSS, inertial sensors, and visual–inertial odometry have been also proposed for on-site architectural visualization, though limitations in positioning accuracy and system robustness remain evident (Assumpcao and Cuperschmid, 2024). Beyond the built environment, AR has been even applied to geospatial data visualization, such as air quality mapping (Sanità et al., 2024), where GIS-derived thematic information is overlaid onto real-world scenes to enhance spatial awareness.

Collectively, these works confirm that AR effectiveness depends on precise georeferencing, reliable pose estimation, and integration with photogrammetric models, rather than on visualization alone.

AR can be thus applied differently in the domains, such as for training, inspection, quality control, or implementation strategy, and adopts different methodological approaches according to the main purpose. In this context, the training of non-expert users has been one of the main aspects taken into account in the definition of our methodology.

In detail, the research from Moghaddam et al. (2021) presents a controlled experimental study comparing AR-based and paper-based instruction for electro-mechanical assembly tasks. The use of HoloLens 2 and a mixed-mode content delivery system (text, images, videos, and 3D animations) allows for a detailed evaluation of AR's impact on performance, error rates, learning retention, and user autonomy. The authors show that although AR does not necessarily reduce task time, it significantly improves accuracy and knowledge transfer, suggesting its value for high-precision and knowledge-intensive tasks.

On the other side, although much AR research in photogrammetry and metrology has traditionally focused on foundational technologies (e.g., camera pose estimation and spatial localization), recent works (Yalcin et al., 2025) demonstrate the convergence of photogrammetric data products with AR interfaces in practical applications. For instance, mobile AR applications have been developed that overlay CityGML (City Geography Markup Language)-based photogrammetric models in real time, enabling interactive visualization of complex spatial data directly on physical scenes, with implications for planning, cadastral management, and stakeholder engagement in built environments.

Therefore, AR has shown to have diverse industrial applications and that it can improve productivity, efficiency, and work quality across manufacturing sectors, although challenges relating to software maturity and usability persist. Empirical evaluations of commercial AR tools in manufacturing contexts highlight the need for robust interaction design and domain-specific features to increase adoption in general industrial settings (Fernández-Moyano et al., 2025). Beyond specific applications, broader conceptual work positions AR and MR as foundational components of the industrial metaverse, where digital twins, Internet of Things (IoT), and eXtended Reality (XR) come together to deliver real-time interactive experiences for operators and planners in Industry 4.0 and beyond.

2.2.1. Augmented reality for Industry 4.0

In the context of Industry 4.0, XR tools are increasingly investigated as enabling technologies for the integration of digital models, real-time measurements, and human operators within industrial workflows, particularly in alignment with the broader aims of Industry 4.0 itself and

with support for even non-expert users or operators. Experimental studies in real and realistic industrial scenarios have showed that AR-based guidance significantly improves manual assembly performance, reducing task completion time, error rates, and perceived cognitive workload compared to traditional paper-based documentation, especially for non-expert operators (Marino et al., 2024). Beyond procedural guidance, AR is widely adopted for industrial inspection and quality control, where spatially registered overlays of digital models support the identification of assembly discrepancies. Recent comparative studies demonstrate that the choice of AR visualization and rendering techniques plays a decisive role in inspection accuracy and operator efficiency (Fortuna et al., 2024). At a higher level, a qualitative evaluation and meta-analysis (Morales Méndez and del Cerro Velázquez, 2024) confirm a statistically significant positive impact of AR on industrial assistance and training, highlighting its effectiveness in reducing cognitive load and improving skill acquisition.

In a more specialized direction, Ho et al. (2021) narrow their focus to the intersection of AR and metrology for quality control. They evaluate the growing, but still nascent, body of research on AR in Quality 4.0 contexts. Their analysis draws attention to the potential of AR to enhance visual inspection, overlay tolerances, and support operator judgment. Importantly, they emphasize the need for accurate tracking and real-time data integration, especially in applications requiring high precision, such as aerospace and automotive. These two works allowed the test and validation of the AR TACCO system for high precision tasks on the operators, dealing with the "workforce acceptance" and driving to opt for low-cost solutions.

The effectiveness of low-cost solutions has been proven by Marino et al. (2021) by implementing a complete AR inspection tool and evaluating it through field testing with factory workers and engineers. Their methodology combines software development with usability evaluation, cognitive load analysis, and user interviews. This mixed-methods approach enables them to assess not only task performance but also the user's subjective experience. The study confirms that AR can significantly support discrepancy detection in large, complex assemblies, while maintaining a low cognitive burden, grounding in real-world conditions, using off-the-shelf tablets and marker-based tracking to ensure system robustness. Yet, it is constrained by some technical limitations, such as lighting sensitivity and marker visibility, which affect tracking accuracy and thus the reliability of AR overlays.

Altogether, these studies illustrate the multifaceted potential of AR in industrial environments, while also underscoring critical limitations. Empirical studies show AR's effectiveness in specific tasks, particularly where accuracy is paramount. However, broader adoption continues to face challenges related to hardware usability, environmental variability, system integration, and organizational readiness. A comprehensive approach, one that combines user-centered design, technical robustness, and systemic alignment, is essential for unlocking AR's full value in Industry 4.0 contexts.

3. Methodology

Based on the analysis of the related works, the following section presents the proposed methodology, which aims to achieve high photogrammetric accuracy in large-volume industrial scenarios through a robust, simulation-based workflow designed to be usable even by non-expert operators.

To achieve photogrammetric measurement accuracies in the range of 1:100.000, a level typically associated with high-end metrology, it is essential to establish a robust and user-friendly workflow that can be executed even by untrained personnel. A key requirement of this workflow is the accurate placement of pre-calibrated scale bars and coded auxiliary components, which serve as spatial references and anchors for the bundle block adjustment step. These components are fundamental to reliably identifying the locations of critical machining points on the industrial part, marked by reflective targets referred to hereafter as Uncoded Targets (UTs).

The study does not aim to achieve a complete reconstruction of the entire workpiece. Instead, it focuses on accurately measuring a predefined set of critical points, namely the UTs, that are known in advance and correspond to machining locations. Accordingly, the acquisition strategy is designed to ensure sufficient accuracy at these specific points rather than full object coverage. This requirement is directly embedded in the optimization framework through a maximum allowable ellipsoidal error defined for each UT.

To support operators throughout the entire measurement process and to ensure the high accuracy and quality of photogrammetric results, a dedicated methodology has been developed. This approach enables the simulation and estimation of the error ellipsoids associated with each UT, depending on various acquisition strategies that the user can define or select. These ellipsoids provide an analytical estimate of spatial uncertainty, allowing for the comparison of different measurement setups before any physical acquisition takes place.

The simulation process is driven by three core inputs:

- a high-fidelity STL or STEP 3D model of the industrial component to be machined;
- the defined spatial positions of the Uncoded Targets, i.e., the machining points requiring high-accuracy localization;
- a configuration file that includes customizable parameters governing all remaining aspects of the simulation, such as camera specifications, auxiliary component settings and limits, threshold values, desired vertical and horizontal image overlap, and various other tunable options.

Using these inputs, the system performs an automated and iterative simulation that virtually replicates the steps an operator would undertake during a real-world measurement session. The goal of this simulation is to generate a sub-optimal yet highly effective configuration that minimizes uncertainty and maximizes visibility and measurement redundancy.

As output, the simulation provides:

- A detailed list of optimal placement positions for all auxiliary components of the TACCO system, including reference frames, coded targets, and calibrated scale bars;
- A planned set of camera poses (positions and orientations) to ensure adequate image coverage and redundancy.

These results are then exported as 3D models in STL format and loaded into an AR visualization system. Through the AR headset, the operator can intuitively see where each auxiliary component must be physically placed on the real industrial part. This overlay of the digital plan onto the real geometry ensures correct placement even by non-expert users, significantly reducing setup time and potential errors (see Fig. 1).

Following the placement phase, the operator proceeds to the image acquisition step. Using a standard digital camera wirelessly connected to a PC, the user captures images according to the pre-planned trajectories. The AR system continues to provide real-time guidance, indicating the correct viewpoints and camera orientations to ensure comprehensive and consistent image coverage. This acquisition process enables the precise measurement of the relative positions of the UTs with respect to the industrial component.

The simulation methodological pipeline is summarized in the schematic diagram shown in Fig. 2. The mathematical formulation adopted in this work is based on standard photogrammetric models, including the collinearity equations and least-squares adjustment, which are widely documented in the literature. These formulations are intentionally adopted to provide a reliable foundation for the proposed simulation-based framework and to ensure methodological soundness and reproducibility. This research thus represents a significant advancement in photogrammetric process automation and accessibility, enabling high-precision industrial alignment with minimal need for expert intervention.

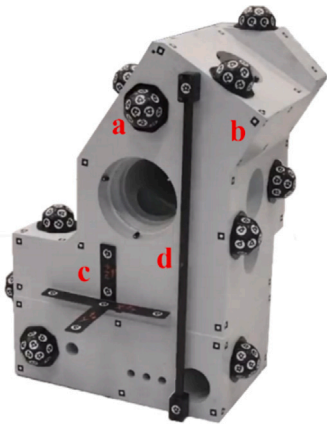


Fig. 1. Example of a raw component with auxiliary elements placed on its surface: (a) Coded targets (igloos); (b) UTs (Uncoded Target points where machining must be performed); (c) Cross reference system; (d) Calibrated scale bar.

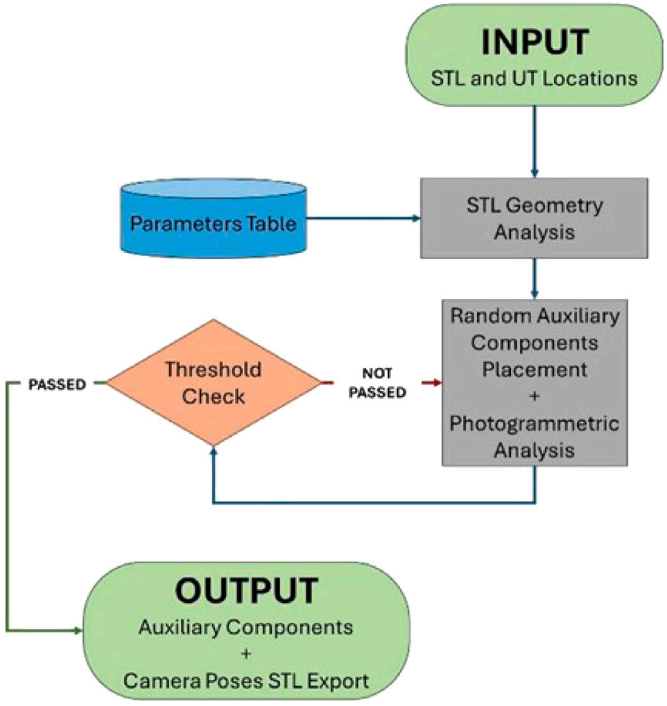
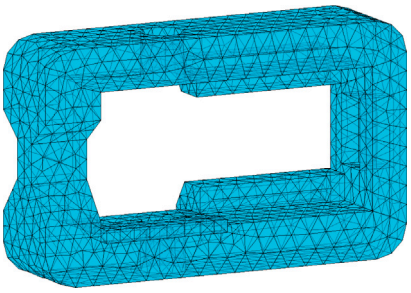
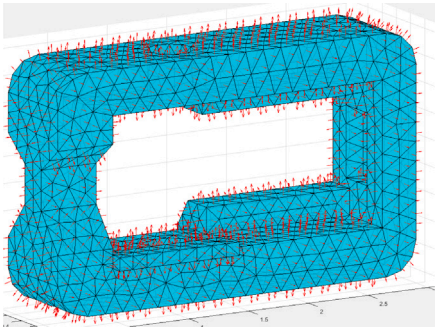


Fig. 2. Block diagram of the proposed methodology.



(a) Meshed STL object.



(b) Visualization of computed normal vectors.

Fig. 3. STL Geometry Analysis first step.

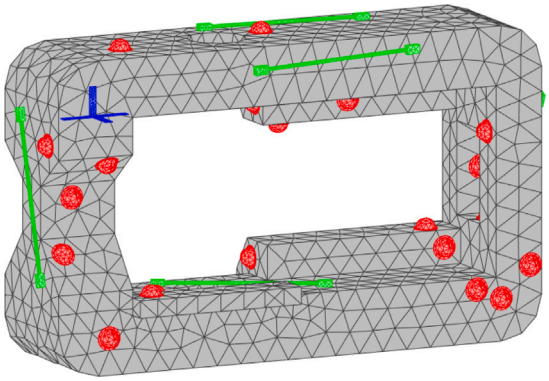


Fig. 4. Example of auxiliary components random placement on the model.

3.1. Simulation

The first step in the simulation is to identify the surfaces where auxiliary components can be placed. Once the object's model is loaded, it is meshed as finely as possible while keeping the computational load manageable, as shown in Fig. 3(a). The normal vectors of each node of the mesh are then calculated in order to be able to correctly orient and place each auxiliary component, as shown in Fig. 3(b).

The second step is the placement of the UTs, whose relative positions are known a priori and provided by the user on the surfaces to be machined. These positions are excluded from the list of possible nodes where the TACCO system's auxiliary components can be placed. Based on the previously described process, a list of available positions is generated for each planar face of the model, taking into account the geometric constraints of the different auxiliary components.

Given the maximum number of each type of auxiliary component available, some positions are randomly selected from random faces, and components are virtually placed there, checking for potential collisions at each iteration. The different types of components are placed in a predefined order: first the reference system, then the scale bars, and finally the coded targets according to their importance for the final calculation of the error ellipsoids.

Once all auxiliary components have been randomly positioned without collisions as shown in Fig. 4, the next step is the generation and definition of the camera poses, following the user-selected acquisition strategy. Three strategies, schematized in Fig. 5, have been defined: spherical, ellipsoidal, and buffer-based, which can be chosen by the user depending on the object geometry and the available space around it.

- The first strategy generates a series of circular trajectories forming a hemispherical pattern along the Z-axis centered at the centroid of the component.

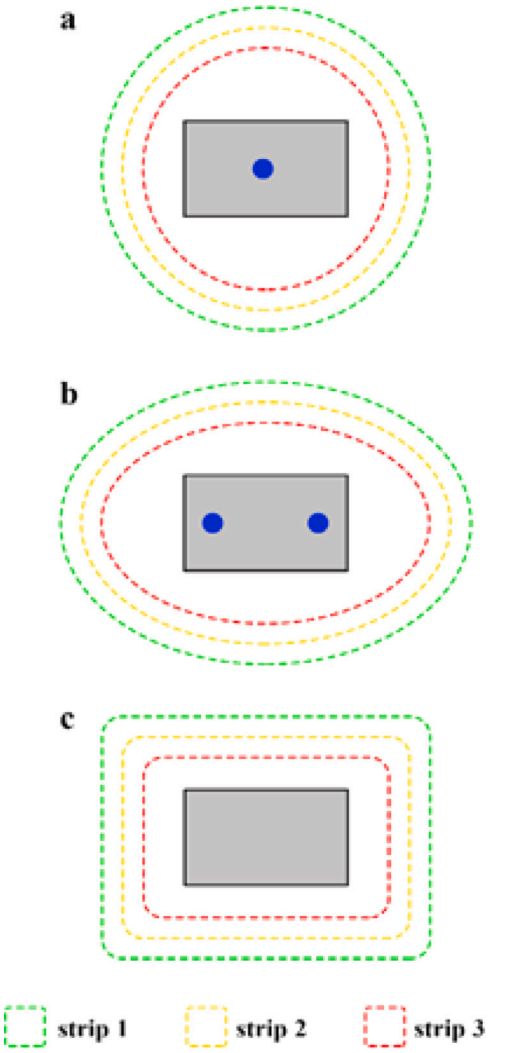


Fig. 5. A top-view schema of the three implemented strategies: spherical (a), ellipsoidal (b), and buffer (c).

- The second strategy defines ellipsoidal trajectories, with the two foci located along the major axis of the component, forming an ellipsoid extending along the Z-axis.
- The third strategy involves generating trajectories that follow the contour of the component at a fixed offset distance, created every few meters along the Z-axis depending on the component's height.

An example of camera pose generation using the ellipsoidal strategy is shown in Fig. 6.

The proposed shooting trajectory strategy is intentionally simple and heuristic-based, as the goal of this work is not to jointly optimize shooting paths and auxiliary target placement. Instead, the focus is on determining suitable auxiliary target positions along a predefined acquisition trajectory, supporting guided and supervised placement rather than arbitrary solutions. Shooting trajectories are selected according to established photogrammetric practices used in industrial inspection, and multiple trajectory configurations are analyzed to assess the impact of acquisition geometry. A sub-optimal but computationally efficient solution is deliberately adopted to validate the proposed methodology in a realistic and controlled setting. Moreover, the approach is designed to assist inexperienced operators: a structured and intuitive shooting strategy ensures practical usability in real-world industrial scenarios, balancing optimization performance with operational feasibility.

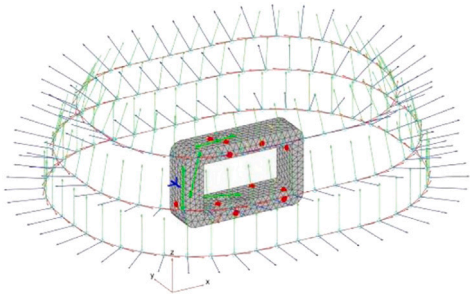


Fig. 6. Simulation of auxiliary component placement on the raw part and photogrammetric image acquisition using the ellipsoidal strategy.

It is important to note that the iterative loop process can be efficiently parallelized, which significantly reduces the overall computation time required to perform multiple runs. This parallelization enables the system to explore a broader solution space more rapidly, increasing the likelihood of reaching a satisfactory or sub-optimal solution within a shorter timeframe.

3.2. Monte Carlo solution using Least square simulation

Analytically, the Monte Carlo solution is based on a Structure from Motion approach fully developed by the authors in the MATLAB environment, without relying on routines from the Computer Vision Toolbox or other external libraries, as shown in Fig. 7, where:

- X_j are the 3D coordinates in the object space (X, Y, Z) of the jth point expressed using homogeneous coordinates $(X_j, Y_j, Z_j, U_j)^T$;
- x_{ij} are the 2D image coordinates in the ith image (u, v) of the jth point expressed using homogeneous coordinates $(u_{ij}, v_{ij}, w_{ij})^T$;
- P_i are the camera pose matrices of ith image in homogeneous coordinates defined as in Eq. (1).

$$P_i = K\begin{bmatrix} R_i & t_i \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}_i = \begin{pmatrix} p_{1i} \\ p_{2i} \\ p_{3i} \end{pmatrix} \tag{1}$$

Where K is the Interior orientation matrix of used digital camera, R_i and t_i are the matrices of exterior orientation parameters: rotation angles ω_i, ϕ_i, κ_i and translation of projection center (X_{0i}, Y_{0i}, Z_{0i}) of ith image. In this work, the rotation matrix R_i follows the conventional photogrammetric formulation and is obtained from the successive rotations about the X, Y and Z axes defined by the angles ω_i, ϕ_i, κ_i, respectively (anticlockwise positive, Luhmann (2010), Luhmann et al. (2016)).

The main equations used to solve the problem are the collinearity equation expressed in homogeneous coordinates (image and object) (Eq. (2)):

$$\begin{pmatrix} u_{ij} \\ v_{ij} \\ w_{ij} \end{pmatrix} = P_i \begin{pmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{pmatrix}$$

$$\begin{cases} u_{ij} = \dfrac{w_{ij}u_{ij}}{w_{ij}} = \dfrac{p_1 X_j}{p_3 X_j} \\ v_{ij} = \dfrac{w_{ij}v_{ij}}{w_{ij}} = \dfrac{p_2 X_j}{p_3 X_j} \end{cases}$$

$$\begin{cases} u_{ij} = \dfrac{p_{11i}X_j + p_{12i}Y_j + p_{13i}Z_j + p_{14i}}{p_{31i}X_j + p_{32i}Y_j + p_{33i}Z_j + p_{34i}} \\ v_{ij} = \dfrac{p_{21i}X_j + p_{22i}Y_j + p_{23i}Z_j + p_{24i}}{p_{31i}X_j + p_{32i}Y_j + p_{33i}Z_j + p_{34i}} \end{cases} \tag{2}$$

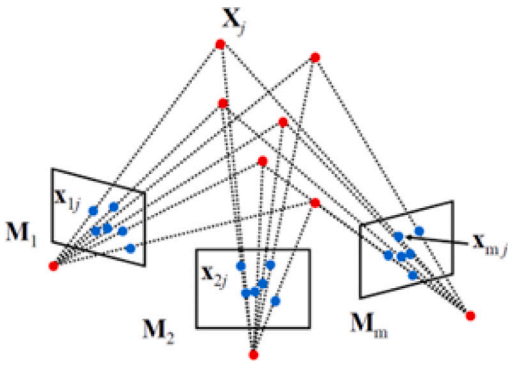


Fig. 7. Structure from Motion approach.

In this approach, the known parameters are:

- The 3D coordinates of the 6 coded targets on the reference system bar considered as constraints in the least square adjustment;
- The interior orientation parameters (fixed) of used digital camera in the K matrix (focal length, 2D coordinates of principal point, pixel size in u and in v, skew;
- The coefficients (fixed) of the polynomial representation of optics calibration (radial, decentring) of used digital camera $(k_1, k_2, k_3, k_4, p_1, p_2, p_3)$.
- The length L_{sb} of scale bars with a precision $\sigma_{sb} = 0.1$ mm;
- The measures of the 2D image coordinates of jth point in ith image $(u_{ij}, v_{ij})^T$ with a defined precision $\sigma_{uv} = kd_{pixel}$ with k = 0.1–0.5.

In the least square simulation, the unknown parameters are:

- The 3D coordinates of jth points $(X_j, Y_j, Z_j)^T$ as uncoded target, coded targets of auxiliary elements, start and end points of scale bars;
- The 6 exterior orientation parameters of ith images $(X_{0i}, Y_{0i}, Z_{0i}, \omega_i, \phi_i, \kappa_i)$;

We suppose that the digital camera has a calibration with an accuracy of less than 1 μm. The reason for this assumption resides in the fact that the camera interior orientation parameters are assumed to be known *a priori* from an off-line calibration procedure, and according to established close-range photogrammetry literature, such calibrations typically achieve sub-pixel residuals, corresponding to an uncertainty on the sensor plane of the order of 1 μm or better for high-resolution digital cameras (Fraser, 1988; Remondino and Fraser, 2006; Luhmann, 2010; Luhmann et al., 2016). This level of uncertainty is generally negligible, or at most of the same order of magnitude, with respect to the image coordinate measurement noise (σ_{uv} = 0.1–0.5 pixels) assumed in the simulations, and therefore it does not dominate the uncertainty propagation in the bundle adjustment.

The number of unknown parameters can be evaluated:

$$n_{Unknowns} = 6n_{Images} + 3n_{Uncoded\,Markers} + 6n_{ScaleBars} + 3n_{MarkerPerIgloos}n_{Igloos}.$$

The number of equations can be evaluated summarizing different parts. For fixing the constraint of reference system, we add 3 equations for the jth coded marker in the Reference System Bar:

$$\begin{cases} X_j = a_j\ (const) \\ Y_j = b_j\ (const) \\ Z_j = c_j\ (const) \end{cases} \tag{3}$$

for a total of 6*3=18 equations. For a scale bar, we add a weighted equation on the length:

$$\sqrt{(X_{ek} - X_{sk})^2 + (Y_{ek} - Y_{sk})^2 + (Z_{ek} - Z_{sk})^2} = L_{sb} \tag{4}$$

where (X_{sk}, Y_{sk}, Z_{sk}) and (X_{ek}, Y_{ek}, Z_{ek}) are the 3D coordinate of the starting and ending code markers of the jth scale bar. The weight of this equation is defined using the σ_{sb}. In particular, in the least-squares framework, constraints can be introduced as fictitious observations with assigned *a priori* weights, allowing them to act either as hard or soft constraints depending on their weighting. In this work, the constraints defined in Eq. (3), which fix the coordinates of the coded targets belonging to the reference system bar, are implemented as hard constraints by assigning them very high weights, thus defining a stable datum for the adjustment and ensuring numerical stability of the solution. Conversely, the scale bar constraints introduced in Eq. (4) are treated as soft constraints, whose weights are defined according to the measurement precision of the scale bar length (σ_{sb}). This formulation allows the scale to be enforced while properly accounting for its associated measurement uncertainty.

Starting and ending coded markers are measured in 2D image coordinates then 2 equations can be added for each image in which they are captured. The weights of these equations are defined using the σ_{uv}. For each coded marker in each auxiliary elements, the Matlab code checks, for every image, whether the point is "included and visible" and writes the 2 equations. The weights of these equations are defined using the σ_{uv}. We do not add constraints on the mutual distances between coded targets in an auxiliary element. In the same way, the Matlab code checks each uncoded marker, for each image, whether the point is "included and visible" and writes the 2 equations. The weights of these equations are defined using the σ_{uv}.

The linear system can be expressed using the design matrix A (also known as coefficient matrix, model matrix or regressor matrix):

$$AX = T + v \tag{5}$$

where T is the column vectors of the various measures (known terms), X is the column vector of unknowns, and v is the column vector of residuals. Generally, if the measures are available, the LSM estimates the solution minimizing the function $v^T v = min$ using the normalization of the system:

$$A^T AX = A^T T \tag{6}$$

Defining the normal matrix $N = A^T A$ and the normalized vector of known terms $TN = A^T T$, the normalized system can be solved with the inversion of the normal matrix:

$$NX = T_N \tag{7}$$

$$X = N^{-1} T_N \tag{8}$$

Using an example of a system with 3 unknowns and 4 equations, it is possible to write the coefficient matrix (normal matrix using Gaussian tiles):

$$\mathbf{A} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{bmatrix} \tag{9}$$

The subscripts contain the equation index. The normal matrix is shown in Fig. 8.

It can be noted that the different parts of a normal matrix can be evaluated with a cumulative approach, since the cyan terms can be calculated by only using Eq. (5), the pink ones by only using Eq. (6), and so on.

Therefore, the multiplication on the left of A^T is not required, and the N matrix can be written as a direct cumulation of the different contributions of each equation. From a computational perspective, the normal matrix resulting from the least-squares formulation exhibits a sparse structure. In the proposed implementation, this property

$$
\begin{aligned}
[aa] &= a_1a_1 + a_2a_2 + a_3a_3 + a_4a_4 \\
[ab] &= a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4 \\
[ac] &= a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4 \\
[bb] &= b_1b_1 + b_2b_2 + b_3b_3 + b_4b_4 \\
[bc] &= b_1c_1 + b_2c_2 + b_3c_3 + b_4c_4 \\
[cc] &= c_1c_1 + c_2c_2 + c_3c_3 + c_4c_4
\end{aligned}
$$

Fig. 8. Normal matrix.

is exploited by avoiding the explicit construction of the full design matrix. Instead, the normal matrix is built through an incremental accumulation of the individual contributions of each observation equation. Furthermore, the covariance matrix is not fully computed; only the submatrices corresponding to the uncoded target coordinates are extracted and used for error ellipsoid estimation. This strategy significantly reduces memory usage while ensuring numerical stability, which is prioritized in the Monte Carlo simulation framework adopted in this work. Although more advanced sparse linear algebra techniques could further improve computational performance, the current implementation was designed to balance efficiency, robustness, and clarity, given that the primary objective of the simulation is uncertainty assessment rather than real-time computation.

The solution is thus the following: the Eqs. (2) and Eq. (4) need a linearization in the unknown parameters to be solved, defining a linear system which can be solved using the least squares algorithm. The simulation does not aim to solve the system, but to evaluate the following parts:

- Normal matrix and inverse normal matrix using the Moore–Penrose pseudo inverse based on singular value decomposition (SVD) to prevent ill-conditioned problems of the equation system.
- Covariance matrix of unknown parameters is evaluated using inverse normal matrix and weight matrix; the submatrices devoted to the description of precision of uncoded targets are extracted and used for standard ellipsoid generation (95% of probability).
- The redundant matrix is estimated to prevent solution with too small redundancy.

3.3. Least square simulation of bundle block adjustment

Once all camera positions have been calculated according to the chosen strategy, a least-squares simulation of a conventional photogrammetric bundle adjustment is performed. The adjustment is formulated using the collinearity equations in a projective framework and is employed here to analyze uncertainty propagation rather than to solve an actual image-based reconstruction. The image coordinates of various features and the lengths of the scale bars are appropriately weighted to estimate the variance–covariance matrix of the coordinates of the uncoded targets. This matrix is then used to compute 95% confidence level error ellipsoids.

The maximum semi-axis of each error ellipsoid is compared against a user-defined threshold, which sets the maximum allowable error for that specific measurement. The entire process is carried out in an external loop, where the random positions of the auxiliary elements are recalculated until the threshold constraint is satisfied.

From the variance–covariance matrix of the coordinate estimates, the 3 × 3 submatrix corresponding to the precision parameters of the ith point is extracted:

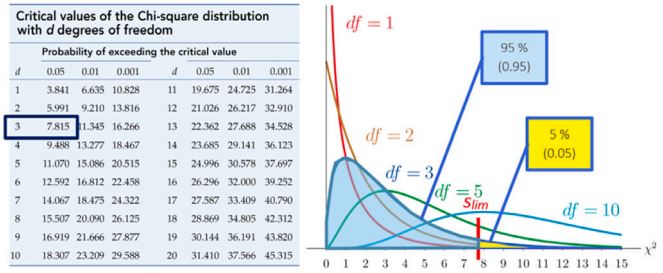


Fig. 9. Critical values of the Chi-square distribution for various degrees of freedom.

$$
\boldsymbol{C_{XX}} = \begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \sigma^2_{X_i} & \sigma_{X_iY_i} & \sigma_{X_iZ_i} & \cdots \\ \cdots & \cdots & \sigma^2_{Y_i} & \sigma_{Y_iZ_i} & \cdots \\ \cdots & \cdots & \cdots & \sigma^2_{Z_i} & \cdots \\ simm & \cdots & \cdots & \cdots & \cdots \end{vmatrix}
$$

Eigenvalues and eigenvectors are then computed to rigorously determine the error ellipsoid, which represents the isosurface of the 3D Gaussian distribution at a specified confidence level. A 95% confidence level is commonly adopted, meaning that the error ellipsoid defines the region containing 95% of all possible point estimates drawn from the underlying Gaussian distribution.

The three eigenvectors define the directions of the ellipsoid's semi-axes, while the corresponding eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ represent their lengths. To compute the lengths of the three semi-axes at the 95% confidence level (a, b, c), the eigenvalues are scaled using the critical value s_{lim} from a Chi-square distribution (Fig. 9) with 3 degrees of freedom:

$$
\begin{aligned}
\mathrm{P}(\mathrm{s} < \mathrm{s}_{lim}) &= 1 - 0.05 = 0.95 \\
\mathrm{s}_{lim} &= 7.815
\end{aligned}
$$

Given the eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$, the lengths of the 95% confidence error ellipsoid semi-axes (a, b, c) are:

$$
\begin{aligned}
\mathrm{a} &= 2\lambda_1\sqrt{7.815} \\
\mathrm{b} &= 2\lambda_2\sqrt{7.815} \\
\mathrm{c} &= 2\lambda_3\sqrt{7.815}
\end{aligned}
$$

The orientation of the error ellipsoid's semi-axes is determined by the directions of its three eigenvectors. If the maximum error remains within the defined threshold, the positions of auxiliary elements and the corresponding camera poses are saved and exported for integration with the VR system. The objective is to minimize the ellipsoid error while also reducing the number of auxiliary components and images needed to reach the desired accuracy. This is crucial because the complexity of bundle block adjustment increases significantly with a large number of images. These parameters are defined according to the accuracy and precision requirements of the TACCO system (a minimum admissible semi-axis of the error ellipsoid, λ_{adm}), set by the user to develop the specific industrial application, aiming to reduce the number of required components and images.

3.4. AR headset integration

To assist inexperienced operators in placing the auxiliary elements and acquiring images, a system was designed to guide them to the correct positions virtually. Through an Augmented Reality (AR) headset, the PICO 4 (Fig. 10), the operator can simultaneously view the real world with the object to be processed and the virtual models derived

Fig. 10. PICO 4 AR headset.

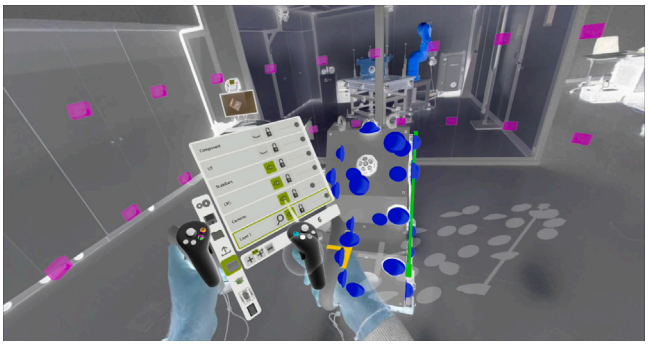


Fig. 11. AR visualization of a demo workpiece with the possibility to select the different layers: component, uncoded targets (UT), scalebars, cross reference and cameras.

from the simulation, along with the optimal placement of the auxiliary elements and camera poses for the TACCO system's acquisition process.

Using a free digital 3D design application, Gravity Sketch, it is possible to measure, group elements (Fig. 11), and manage several aspects such as color, transparency, and visibility. Most importantly, it allows the user to seamlessly transition from a VR environment, a blank canvas with STL models, to an AR environment, perfectly merging the virtual and real-world models.

In the proposed workflow, once the models are exported from MATLAB and loaded into the AR device via PC through the app's management platform, the user can freely adjust the appearance and visibility of each element group.

To merge real and virtual environments accurately, the headset processor performs complex spatial calibration by establishing a persistent World Coordinate System. Using the four tracking cameras, unique environmental features are identified to create a point cloud of the room. Virtual objects are then "anchored" to these specific coordinates. This calibration ensures that when a virtual asset is placed on a physical surface, the software understands its position in 3D space relative to the headset's origin, maintaining the illusion of physical presence through continuous coordinate transformation.

As the user moves, the inertial measurement units (IMUs) and tracking cameras work to update the headset's position and orientation at high frequencies. This allows virtual objects to remain "pinned" to their real-world locations even as the user's perspective changes.

Fig. 12 shows the virtual model that can be manually rotated and translated until it perfectly overlaps with its real-world counterpart, allowing the operator to accurately position the auxiliary components on the object and subsequently acquire the images from correct camera poses. The final positioning of the elements on the test piece and the relative visualization during the acquisition phase in a real case scenario is also depicted.

4. First results and validation

To assess the feasibility and relevance of the proposed methodology, we initially conducted a series of simulations on a variety of industrial components, encompassing a wide range of physical scales and structural complexities.

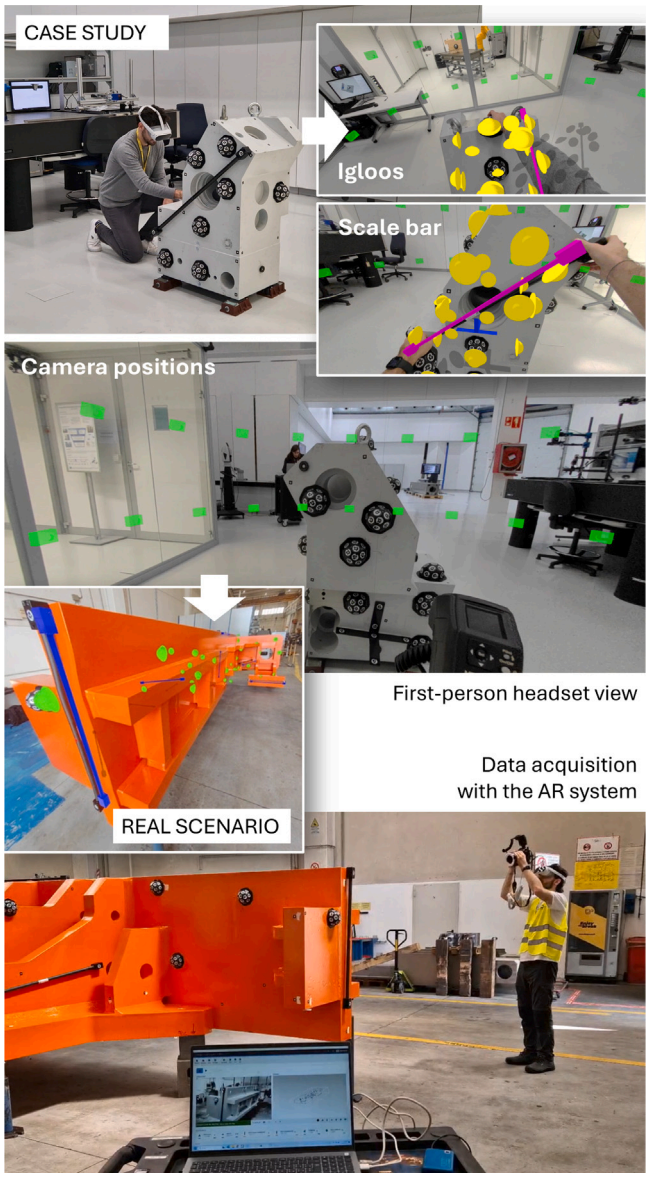


Fig. 12. AR visualization of a demo workpiece with the possibility to select the different layers: component, uncoded targets (UT), scalebars, cross reference and cameras.

These simulations aimed to replicate realistic acquisition conditions and geometrical challenges typical of industrial components. The outcomes of the simulated reconstructions were then systematically compared with analytical estimates derived a posteriori, enabling us to validate the accuracy and consistency of the method under different scenarios. Beyond validation, this comparison also guided the fine-tuning of key simulation parameters, helping to identify the most effective trade-off between computational complexity, scalability, and the desired level of relative accuracy.

For each component, 100 independent iterations were performed following the Monte Carlo approach. Each iteration introduced randomized placement positions for auxiliary components that affect the final accuracy results in terms of error of the ellipsoids on uncoded targets. Each simulation setup included 6 scale bars, 10 coded targets, and 20 uncoded targets. Among the 100 runs, the one with the lowest maximum error on UTs was selected as the representative result for that component. Additionally, the average maximum error across all runs was reported.

These simulations were executed on an Alienware desktop computer equipped with an Intel Core i7-9700K processor running at 3.60 GHz, 32 GB of RAM. A summary of the best-case results across all components, together with other key parameters, is provided in Table 1.

The simulations in Table 1 were conducted using the intrinsic parameters of the Nikon Z6 II digital camera. Specifically, the configuration included a resolution of 3936 × 2624 pixels, a focal length of 35 mm, and a pixel size of 5.94 μm. These parameters were selected to closely mirror realistic data acquisition conditions for the basic version of the TACCO system. The simulations in Table 2 were carried out using the intrinsic parameters of the Fujifilm GFX 100 II digital camera that is the suggested camera for the Pro version of the TACCO system. Specifically, the configuration included a resolution of 11648 x 8736 pixels, a focal length of 30 mm, and a pixel size of 3.76 μm.

Two distinct types of simulation tests were performed using the Fujifilm GFX 100 II. Accordingly, each component in Table 2 is represented by two rows:

- The first row corresponds to tests where only the camera parameters were changed, allowing the simulation to autonomously determine the optimal number of stripes, images, and camera poses following the same approach used with the Nikon Z6 II simulations.
- The second row represents a test in which the simulation was constrained to use the same number of images and shooting positions as in the Nikon-based simulations, so that the only thing that change is the camera parameters.

This setup enables a direct comparison of performance and results from two different perspectives: one based on fully optimized settings and the other based on matched acquisition conditions.

Based on the simulation results, several key observations can be made regarding the performance and robustness of the proposed photogrammetric methodology:

- Minimum scale bar requirement: a minimum of four scale bars is necessary to achieve a stable and reliable solution, even when dealing with relatively small-scale industrial components. Using fewer scale bars can lead to an unstable scale definition and degraded accuracy; according to Blaha and Sandwith (2001), each calibrated scale bar provides an independent estimate of the scale factor, and these estimates are combined through a weighted mean. Therefore, additional well-calibrated scale bars reduce the uncertainty of the scale factor.
- Effect of auxiliary coded targets: The simulations indicate that the total number of auxiliary targets has a negligible impact on the final solution quality. This suggests that, beyond a minimal threshold, the redundancy provided by additional coded targets does not significantly improve accuracy.
- Scalability of image acquisition: The number of required images increases proportionally with the physical dimensions of the component. This ensures that sufficient image overlap and geometric coverage are maintained to support high-precision and robust bundle adjustment. Unfortunately, for very large components, this implies the need for a substantial number of images, which can result in longer acquisition procedures, extended simulation times, and potentially an increased number of required scale bars. To address this simulation limitation, we increased the computer's RAM dedicated to this process and parallelized the iterations, running up to four tasks simultaneously. This led to a reduction in overall processing time by approximately 40% compared to the data presented in the results tables.
- The geometric complexity of some components significantly complicates the task of identifying appropriate positions for the placement of all the auxiliary elements. This challenge can hinder the efficiency of the simulation. As a result, the overall processing time for each measurement run may increase substantially, limiting the system's scalability and responsiveness even for medium and small size components.
- Another significant factor contributing to higher processing time is the total number of images involved in the simulation. For each image, the algorithm must perform several computationally intensive tasks: checking for occlusions along the line of sight, determining which targets are visible, and generating the corresponding collinearity equations. As the number of images increases, these operations scale non-linearly, leading to a noticeable increase in overall computation time. The number of images is primarily driven by the number of strips required to cover the component. This, in turn, depends on factors such as the height of the component and the desired vertical overlap between images. Taller structures or stricter overlap requirements demand more vertical image rows, thereby increasing the number of strips and, consequently, the number of images. Each additional strip compounds the computational burden, especially in dense or geometrically complex scenarios where occlusion checks become more frequent and costly.
- Recommended imaging distances: The most accurate results are obtained when the camera to object distance and corresponding image capture range are maintained between 1.5 to 3 m. This configuration ensures optimal image resolution and depth variation, while keeping the field of view wide enough to capture necessary surface detail.
- By comparing Tables 1 and 2, two key observations can be made. First, when keeping the number of images constant, using the Fujifilm camera results in an improvement in terms of the uncertainty ellipsoids for the uncoded targets. However, these improvements come at the cost of increased processing time, likely due to the higher image resolution, which results in more occlusion checks and a greater number of collinearity equations being processed per image. Second, when the Fujifilm camera is used and the simulation is allowed to automatically optimize the number and placement of images, the resulting accuracy is

Table 1
Simulations results using Nikon Z6 II.

Components dimensions (m)	Processing Time per run (s)	Maximum error UTs (mm)	Mean error UTs over 100 runs	Number of images
0.8 × 1.21 × 2.44	91	0.08	0.091	80
3 × 0.92 × 1.8	185	0.083	0.124	132
1.79 × 1.4 × 1.16	178.5	0.096	0.141	106
5.37 × 4.2 × 3.48	281	0.055	0.071	259
5.78 × 0.94 × 3.34	537	0.05	0.179	261
7.5 × 2.3 × 4.5	123	0.053	0.265	166

Table 2
Simulations results using Fujifilm GFX 100 II.

Components dimensions (m)	Processing time per run (s)	Maximum error UTs (mm)	Mean error UTs over 100 runs	Number of images
0.8 × 1.21 × 2.44	74.5	0.077	0.091	25
0.8 × 1.21 × 2.44	158	0.069	0.09	80
3 × 0.92 × 1.8	109	0.084	0.126	44
3 × 0.92 × 1.8	383	0.045	0.111	132
1.79 × 1.4 × 1.16	125	0.104	0.152	42
1.79 × 1.4 × 1.16	312	0.044	0.121	106
5.37 × 4.2 × 3.48	162	0.045	0.052	57
5.37 × 4.2 × 3.48	812	0.045	0.049	259
5.78 × 0.94 × 3.34	330.5	0.052	0.079	84
5.78 × 0.94 × 3.34	857	0.045	0.052	261
7.5 × 2.3 × 4.5	104	0.047	0.102	58
7.5 × 2.3 × 4.5	299	0.046	0.050	166

comparable to that achieved with the Nikon camera. Nevertheless, this optimization leads to a reduced number of required images, offering a significant advantage in terms of computational processing time. Additionally, fewer images translate into faster and more efficient real-world data acquisition, making this approach not only computationally advantageous but also practically beneficial in field operations.

To verify the reliability of the above-mentioned methodology, a comparison of the estimated accuracy was made using Fraser's empirical formula (Fraser, 1996):

$\sigma_{XYZ} = \frac{qD}{c\sqrt{k}}\sigma_i$

Where:

- k is the average number of images capturing the same point (multiplicity);
- q is the configuration factor (0.4–0.8), which describes the spatial resection geometry quality and is generally related to the angles between intersecting homologous rays;
- D is the average distance between the images and the object points;
- c is the focal length;
- σ_i is the precision of the measured image coordinates, assumed to be 0.1 pixels in this case.

The resulting precision estimates from the simulation closely align with those predicted by Fraser's empirical formula, thereby validating the accuracy and reliability of the adopted simulation methodology. This consistency supports the reliability of the simulated photogrammetric results for subsequent analysis and system evaluation.

4.1. Real case scenario validation

The proposed methodology was subsequently tested and validated on a real-world case involving a medium-sized (6.5 × 2.5 × 1.5 m) industrial component. For reference measurements, a Leica Absolute Tracker AT403 (Fig. 13) was employed, a laser tracker widely used in industrial settings for large-scale metrology and precision quality control. The AT403 features a maximum measurement range of 320 m and delivers an accuracy of up to 10 μm. It is also engineered to maintain high performance under demanding environmental conditions, including temperature fluctuations and mechanical vibrations, without compromising measurement precision.

The positions of components defined in the simulation have been compared with the real location of the elements in the physical environment, defined using the TACCO System. During this comparison (Table 3), some minor deviations were identified, although none were critical to the system's proper functioning. The discrepancies observed were within the established tolerance range and did not affect the overall functionality of the system.

Fig. 13. Leica AT403 Measurement to validate the simulation on a real-world component.

Table 3
Comparison between the simulated location and the real location of placed components.

Parameters	3D discrepancies [mm]
mean	20
std	100
min	5
max	149

The following phase of the validation process is to compare the simulated location of the images with the acquired one by using the mixed reality system. These tests (Table 4) lead to the conclusion that the module developed is highly effective for guiding and assisting users during the measurement process.

The next step is to compare the location of uncoded targets (UTs) with the measurements taken using the laser tracker. In this comparison, the normal distance to the surface was used as the primary metric, since the laser tracker cannot measure the center of the marker directly. This methodology enabled a precise evaluation of the proposed system's performance under real operating conditions. To perform this operation, a preliminary, approximated 3D rigid roto-translation (Fig. 14, top) has been estimated using the reference coordinate system bar using the target coordinates defined with TACCO system and surveyed with laser tracker. Then, starting from this approximated solution, a new 3D rigid roto-translation (Fig. 14, bottom) has been evaluated using only the coordinates of surveyed points normal to the raw component surface.

An approach suggested by the solution of the absolute model orientation in photogrammetry has been used to solve this problem. Analytically, this is the equation of the absolute orientation,

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_u \\ Y_u \\ Z_u \end{pmatrix} + R_{\Omega\Phi K} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where:

$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ are the 3D coordinates surveyed by Leica Laser Tracker;

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ are the 3D coordinates acquired by the proposed system;

$\begin{pmatrix} X_u \\ Y_u \\ Z_u \end{pmatrix}$ are the 3D translation parameters;

$R_{\omega\phi\kappa}$ is the 3D rotation matrix depending on the 3 rotation angles ω, ϕ, κ. The equation regarding the coordinate in the direction of the normal to the object surface can be used, with one equation for each point in the X, Y, or Z direction. The first evaluation of the approximate solutions using all the coordinates allows for the transformation from the TACCO system's coordinates into a 3D coordinate system similar to the one of the Laser Tracker measurements:

$$\mathbf{X}_0 = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} \approx \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{X}$$

Table 4
Comparison between the simulated location and the real location of images.

Parameters	X [mm]	Y [mm]	Z [mm]	O [gon]	F [gon]	K [gon]
mean	30	−26	40	1.5	2.6	−1.8
std	85	95	77	5.3	5.9	7.2
min	−150	−191	−180	−12.3	−13.5	−11.6
max	172	185	197	13.9	15.2	10.6

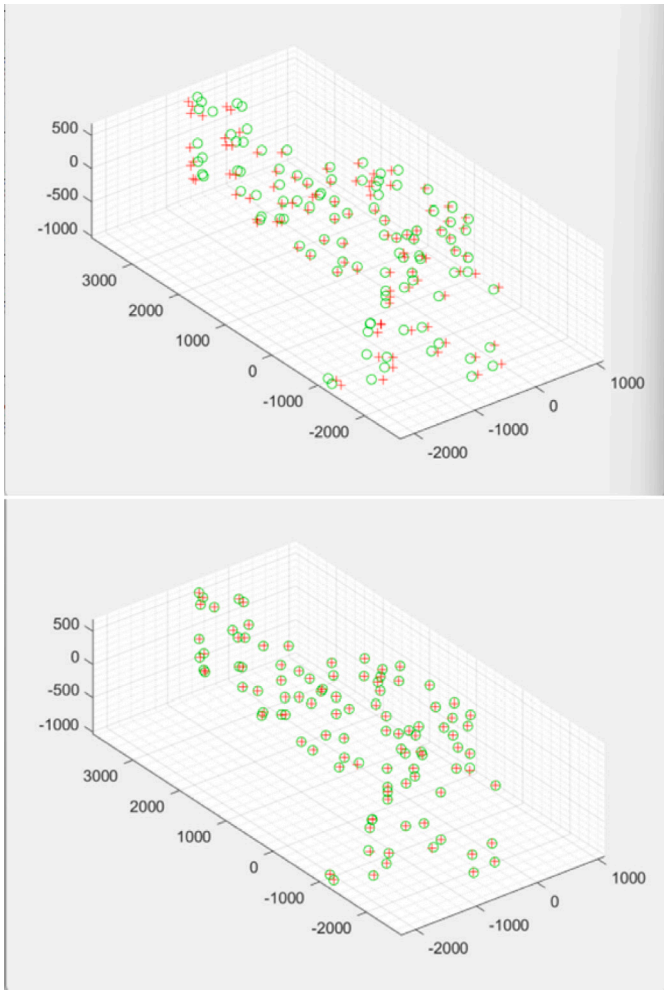


Fig. 14. TACCO system point cloud (in green) and AT403 point cloud (in red), before and after the 3D rigid rototranslation.

The roto-translation parameters are now really small, making it possible to write:

$$\mathbf{X} \approx \mathbf{x} = \mathbf{X}^0$$
$$\downarrow$$
$$\Omega = d\Omega, \quad \Phi = d\Phi, \quad K = dK,$$
$$X_u = dX_u, \quad Y_u = dY_u, \quad Z_u = dZ_u$$

$$\mathbf{R} = \begin{pmatrix} \cos\varphi\cos\kappa & -\cos\varphi\,sen\kappa & sen\varphi \\ \cos\omega\,sen\kappa + sen\omega sen\varphi\cos\kappa & \cos\omega\cos\kappa - sen\omega sen\varphi sen\kappa & -sen\omega\cos\varphi \\ sen\omega sen\kappa - \cos\omega\,sen\varphi\cos\kappa & sen\omega\cos\kappa + \cos\omega\,sen\varphi sen\kappa & \cos\omega\cos\varphi \end{pmatrix}$$

With small value angles, the rotation matrix can be simplified:

$$dR = \begin{pmatrix} 1 & -dK & d\Phi \\ dK & 1 & -d\Omega \\ -d\Phi & d\Omega & 1 \end{pmatrix}$$

and the 3D roto-translation equation can be converted into linear equations:

$$X = dX_u + Z^0 d - Y^0 dK + X^0$$
$$Y = dY_u - Z^0 d\omega + X^0 dK + Y^0$$
$$Z = dZ_u + Y^0 d\omega - X^0 d\phi + Z^0$$

The problem can now be solved with an iterative approach. In the practical case, 127 equations are written for the 127 surveyed points, estimating the parameters of the final 3D roto-translation used to convert the coordinates of TACCO system into laser tracker ones.

These two sets of coordinates have been compared to evaluate the final precision, and the results are summarized in Fig. 15 and Table 5, where it is possible to highlight that the mean value is close to 0, meaning there are no significative biases; the standard deviation of discrepancies is 0.254 mm, close to the value of 1/50000; the distribution of discrepancies is not Gaussian, but there are about 30% of points with very small discrepancies (close to 0), denoting the very high precision of the instrument.

Additionally, this procedure confirmed that the system can measure large components, even in workshop environments where conditions may be less controlled. This capability for industrial-scale measurements adds significant value to the system, facilitating its implementation across a wide range of practical applications and ensuring reliable and accurate measurements. The images acquired by the Fujifilm GFX 100 II have been processed using a Structure from Motion approach in the software Agisoft Metashape Professional, simulating the proposed system with manual plotting of all points (reference system bar, scale bars, igloos, uncoded target) defining a new simulated solution for TACCO system. The same procedure has been applied to obtain two sets of coordinates, laser tracker and our improved proposed system, to be compared for evaluating the final precision. The results are summarized in Fig. 16 and Table 6 where it is possible to highlight that the mean value is close to 0, meaning there are no significative biases; the standard deviation of discrepancies is 0.197 mm, close to the value of 1/61000; the distribution of discrepancies is not Gaussian, but there are about 45% of points with very small discrepancies (close to 0), denoting a higher precision of the PRO version of the system. Therefore, it was confirmed that the system can be improved using a very high resolution camera up to a 22%. This result demonstrates that the TACCO system can be adapted in a very wide range of applications, using specific modules.

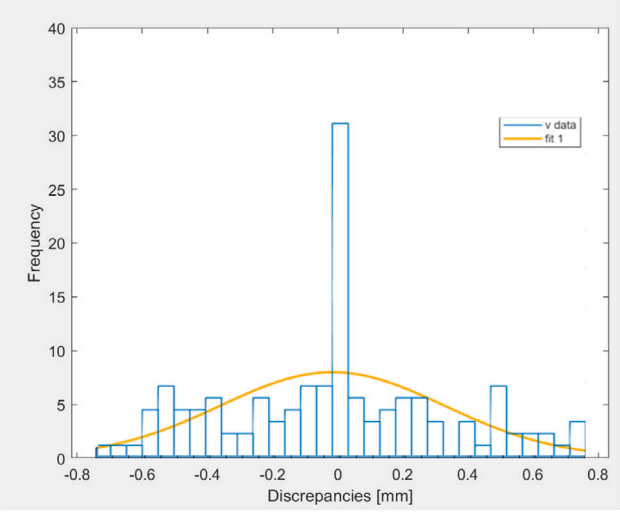


Fig. 15. Histogram of the discrepancies among AT403 laser tracker and TACCO system measurements of uncoded targets.

Table 5
Results of the comparison between the AT403 and the TACCO system.

Parameters	3D discrepancies [mm]
mean	−0.000
std	0.254
min	−0.742
max	0.796

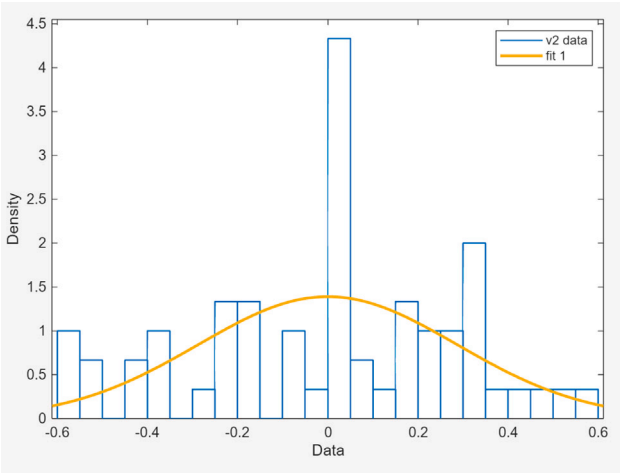


Fig. 16. Histogram of the discrepancies among AT403 laser tracker and the improvement of TACCO system using digital camera Fujifilm GFX 100 II.

Table 6
Results of the comparison between the AT403 and the improvement of TACCO system using digital camera Fujifilm GFX 100 II.

Parameters	3D discrepancies [mm]
mean	−0.000
std	0.197
min	−0.585
max	0.559

The pronounced peak around zero observed in the discrepancy histograms (about 32% in Fig. 15 and about 43% in Fig. 16) is an expected outcome of the adopted validation strategy. After rigid roto-translation between the photogrammetric and laser tracker coordinate systems, discrepancies are evaluated along the surface normal direction, which reduces systematic effects and emphasizes random measurement errors. Points characterized by favorable visibility conditions, strong ray intersection geometry, and high observation redundancy exhibit very small residuals, resulting in a concentration of discrepancies close to zero. This behavior indicates a stable adjustment and confirms the high internal consistency of the proposed photogrammetric solution.

5. Conclusions

This study presents and validates a novel photogrammetric methodology developed within the TACCO project to address the challenges of large-volume industrial alignment. By integrating simulation-based design, Monte Carlo optimization, and AR-assisted user guidance, the proposed system enables the accurate placement of auxiliary components and the acquisition of photogrammetric data, even by non-expert operators.

Extensive simulation tests demonstrated the ability of the system to achieve submillimetric accuracy, with results closely aligned with both analytical predictions and reference measurements obtained from a Leica Absolute Tracker. Notably, the consistency between simulation output and real-world measurements confirms that the target relative accuracy of 1:100,000 is not only theoretically desirable but realistically achievable under operational conditions. The presence of a pronounced peak around zero in the discrepancy histograms further confirms the robustness of the proposed methodology, indicating that a significant portion of the measured points exhibits very small residuals due to favorable measurement geometry and high observation redundancy.

Furthermore, the methodology supports scalability across a wide range of component sizes and geometries, with processing efficiencies enhanced through parallel computation. The integration of augmented reality (AR) plays a key role in improving system accessibility: by providing real-time, spatially aligned visual guidance, the AR interface enables even unexperienced operators to correctly locate auxiliary elements (igloos, scale bars and cross reference) and acquire the required image sets with minimal training. Validation tests also confirmed the system's practical applicability, showing excellent correspondence between simulated and real-world data. Additionally, results show that the use of high-resolution cameras (e.g., Fujifilm GFX 100 II) can improve accuracy by up to 22% while reducing the number of required images, significantly streamlining the workflow.

Overall, the TACCO system represents a significant advancement in industrial photogrammetry by providing a cost-effective, precise, and operator-friendly alternative to conventional laser-based alignment methods. Its combination of real-time visualization, robust simulation, and metrological accuracy positions it as a scalable and practical solution for zero-defect manufacturing in complex industrial environments.

As a future enhancement aimed at further improving accuracy, the human involvement in image acquisition could be minimized. By exporting the simulated capture geometry as a flight plan and uploading it to a UAV (Unmanned Aerial Vehicle) equipped with a stabilized gimbal, images could be captured automatically at precisely defined positions and orientations. This drone-based workflow would reduce field time, ensure repeatable camera geometry, and significantly lower the current residual error from 0.3 mm to well below 0.1 mm. An additional improvement can be related to the code development environment since the current MATLAB implementation is mainly intended to verify the validity and effectiveness of the proposed methodology. Moving the implementation to C++ would significantly reduce the computational time. The reported computation times are thus included only to provide an indicative measure of the method's potential, rather than to represent a fully optimized final system.

CRediT authorship contribution statement

Francesco Messina: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Alessio Martino:** Writing – original draft, Visualization, Validation, Software, Investigation, Data curation. **Francesca Matrone:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Investigation, Data curation, Conceptualization, Methodology. **Andrea Maria Lingua:** Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Pablo Puerto:** Writing – review & editing, Supervision, Resources.

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

The TACCO Project (tacco-project.eu) is co-funded by the EIT Manufacturing and the European Commission. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the EIT Manufactuing nor the European Union or the European Health and Digital Executive Agency (HADEA). Among the partners are Ideko, Mondragon, Soraluce, Politecnico di Torino, Renishaw Iberica, Officine Meccaniche B.B.M., Maschinenfabrik Wüstwillenroth GmbH, and Dr. Matzat. The contribution of the research team at Politecnico di

Torino focused on improving the optimal placement of auxiliary components, enhancing the overall accuracy and user experience, building upon the work carried out by the partners on VSET to achieve new milestones in the application of photogrammetry for industrial alignment. During the preparation of this work, the author(s) used ChatGPT in order to assist with proofreading and improve the clarity and structure of the English language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

Assumpcao, Joao Paulo Franco, Cuperschmid, Ana Regina Mizrahy, 2024. Exploring georeferenced augmented reality for architectural visualization with unmanned aerial vehicles. ISPRS Int. J. Geo-Information 13 (11), 389.

Balletti, Caterina, Guerra, Francesco, Tsioukas, Vassilios, Vernier, Paolo, 2014. Calibration of action cameras for photogrammetric purposes. Sensors 14 (9), 17471–17490.

Blaha, George, Sandwith, Scott C., 2001. Methodology for accuracy assessments and verifications in digital photogrammetry. In: Machine Vision and Three-Dimensional Imaging Systems for Inspection and Metrology, vol. 4189, SPIE, pp. 218–228.

Bolkas, Dimitrios, O'Banion, Matthew, Parisi, Erica Isabella, Bauer, Peter, Detchev, Ivan, 2024. Demonstration of immersive technologies for geospatial learning. ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 10, 1–7.

Bösemann, Werner, 2005. Advances in photogrammetric measurement solutions. Comput. Ind. 56 (8–9), 886–893.

Einizinab, Sajjad, Khoshelham, Kourosh, Winter, Stephan, Christopher, Philip, 2025. Camera pose refinement for precise BIM alignment in mixed reality visualization. J. Comput. Civ. Eng. 39 (5), 04025072.

Fernández-Moyano, Jon A, Remolar, Inmaculada, Gómez-Cambronero, Águeda, 2025. Augmented reality's impact in industry—A scoping review. Appl. Sci. 15 (5), 2415.

Fortuna, Santina, Barbieri, Loris, Marino, Emanuele, Bruno, Fabio, 2024. A comparative study of augmented reality rendering techniques for industrial assembly inspection. Comput. Ind. 155, 104057.

Franceschini, Fiorenzo, Galetto, Maurizio, Maisano, Domenico, Mastrogiacomo, Luca, 2014. Large-scale dimensional metrology (LSDM): from tapes and theodolites to multi-sensor systems. Int. J. Precis. Eng. Manuf. 15, 1739–1758.

Fraser, Clive, 1984. Network design considerations for nontopographic photogrammetry. Photogramm. Eng. Remote Sens. 256–280.

Fraser, Clive S., 1988. State of the art in industrial photogrammetry. IAPRS, Kyoto 27, 166–181.

Fraser, Clive, 1996. Close range photogrammetry and machine vision. Netw. Des. 1115–1126.

Galantucci, L.M., Pesce, M., Lavecchia, F., 2016. A powerful scanning methodology for 3D measurements of small parts with complex surfaces and sub millimeter-sized features, based on close range photogrammetry. Precis. Eng. 43, 211–219.

Girelli, V.A., Tini, M.A., Bitelli, G., et al., 2022. Very high-resolution 3d surveying and modelling experiences in civil engineering applications. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 43 (2–2022), 673–678.

Ho, P.T., Albajez, J.A., Yagüe, J.A., Santolaria, J., 2021. Preliminary study of augmented reality based manufacturing for further integration of quality control 4.0 supported by metrology. In: IOP Conference Series: Materials Science and Engineering, vol. 1193, (1), IOP Publishing, 012105.

Leizea, Ibai, Herrera, Imanol, Puerto, Pablo, 2023. Calibration procedure of a multi-camera system: Process uncertainty budget. Sensors 23 (2), 589.

Luhmann, Thomas, 2010. Close range photogrammetry for industrial applications. ISPRS J. Photogramm. Remote Sens. 65 (6), 558–569.

Luhmann, Thomas, Fraser, Clive, Maas, Hans-Gerd, 2016. Sensor modelling and camera calibration for close-range photogrammetry. ISPRS J. Photogramm. Remote Sens. 115, 37–46.

Marino, Emanuele, Barbieri, Loris, Bruno, Fabio, Muzzupappa, Maurizio, 2024. Assessing user performance in augmented reality assembly guidance for industry 4.0 operators. Comput. Ind. 157, 104085.

Marino, Emanuele, Barbieri, Loris, Colacino, Biagio, Fleri, Anna Kum, Bruno, Fabio, 2021. An augmented reality inspection tool to support workers in industry 4.0 environments. Comput. Ind. 127, 103412.

Mason, S.O., 1995a. Conceptual model of the convergent multistation network configuration task. Photogramm. Rec. 15 (86), 277–299.

Mason, Scott, 1995b. Expert system-based design of close-range photogrammetric networks. ISPRS J. Photogramm. Remote Sens. 50 (5), 13–24.

Mendikute, Alberto, Yagüe-Fabra, José A, Zatarain, Mikel, Bertelsen, Álvaro, Leizea, Ibai, 2017. Self-calibrated in-process photogrammetry for large raw part measurement and alignment before machining. Sensors 17 (9), 2066.

Messina, Francesco, Lingua, Andrea Maria, Martino, Alessio, Matrone, Francesca, Maschio, Paolo, 2024. Fast and accurate zero-defect manufacturing using collaborative real-time photogrammetry. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 48, 495–502.

Moghaddam, Mohsen, Wilson, Nicholas C, Modestino, Alicia Sasser, Jona, Kemi, Marsella, Stacy C, 2021. Exploring augmented reality for worker assistance versus training. Adv. Eng. Informatics 50, 101410.

Morales Méndez, Ginés, del Cerro Velázquez, Francisco, 2024. Impact of augmented reality on assistance and training in industry 4.0: Qualitative evaluation and meta-analysis. Appl. Sci. 14 (11), 4564.

Prieto, Ines, Puerto, Pablo, Leizea, Ibai, Herrera, Imanol, 2024. Assessing the feasibility of markerless close range photogrammetry in industrial inspection: A foundational study. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 48, 121–128.

Puerto, Pablo, Heißelmann, Daniel, Müller, Simon, Mendikute, Alberto, 2022. Methodology to evaluate the performance of portable photogrammetry for large-volume metrology. Metrology 2 (3), 320–334.

Remondino, Fabio, Fraser, Clive, 2006. Digital camera calibration methods: considerations and comparisons. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 36 (5), 266–272.

Sanità, Marsia, Fratini, Jonathan, Muralikrishna, Nikhil, Pierdicca, Roberto, Malinverni, Eva Savina, 2024. Augmented reality for air quality monitoring: Case study in the marche region (Italy). Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 48, 389–395.

Shen, Xu, Yue, Jianping, Dou, Shihao, Huang, Guiping, 2023. Network configuration and repeatability precision of industrial photogrammetry and its application in high precision deformation measurement. Measurement 211, 112589.

Subramaniyam, Bala, Shylesh, Dharshan, Ramasamy, Jaganathan, Kumar, Navin, 2024. A virtual reality tool for accuracy assessment of 3D models in an immersive virtual environment. ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 10, 333–340.

Sun, Peng, Lu, Naiguang, Dong, Mingli, Wang, Jun, Yan, Bixi, 2019. Calibration and orientation of industrial online photogrammetry systems in situ. J. Eng. 2019 (23), 9137–9142.

Sun, Peng, Lu, Nai-Guang, Dong, Ming-Li, Yan, Bi-Xi, Wang, Jun, 2018. Simultaneous all-parameters calibration and assessment of a stereo camera pair using a scale bar. Sensors 18 (11), 3964.

Wang, Qing, Zissler, Nick, Holden, Roger, 2013. Evaluate error sources and uncertainty in large scale measurement systems. Robot. Comput.-Integr. Manuf. 29 (1), 1–11.

Yalcin, Salih, Erkek, Bilal, Ayyildiz, Ekrem, 2025. Integrating photogrammetric 3D city models and citygml data into augmented reality for enhanced urban planning and cadastre management. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 48, 381–386.