www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

SEAMLESS AUGMENTED REALITY REGISTRATION SUPPORTING FACILITY MANAGEMENT OPERATIONS IN UNPREPARED ENVIRONMENTS

SUBMITTED: May 2024
REVISED: October 2024
PUBLISHED: December 2024
GUEST EDITORS: Vito Getuli, Farzad Rahimian, Nashwan Dawood, Pietro Capone, Alessandro Bruttini
DOI: 10.36680/j.itcon.2024.051

Leonardo Messi, Dr,
Department of Construction, Civil Engineering and Architecture, Università Politecnica delle Marche, Ancona, Italy
l.messi@staff.univpm.it

Francesco Spegni, Dr,
Department of Construction, Civil Engineering and Architecture, Università Politecnica delle Marche, Ancona, Italy
f.spegni@staff.univpm.it

Massimo Vaccarini, Dr,
Department of Construction, Civil Engineering and Architecture, Università Politecnica delle Marche, Ancona, Italy
m.vaccarini@staff.univpm.it

Alessandra Corneli, Dr
Department of Construction, Civil Engineering and Architecture, Università Politecnica delle Marche, Ancona, Italy
a.corneli@staff.univpm.it

Leonardo Binni, Eng
Department of Construction, Civil Engineering and Architecture, Università Politecnica delle Marche, Ancona, Italy
l.binni@pm.univpm.it

SUMMARY*: Despite its great potential, Augmented Reality (AR) still struggles to be widely used in real processes in the construction industry. This is mainly due to limitations associated with current AR registration methodologies, including the lack of continuity between different scenarios, drifts over distances, and the need to prepare the scene in advance with alignment infrastructures and to resort to manual registration procedures. In addition, users may not be skilled enough or allowed to prepare the considered environment, since this is a task that requires a significant amount of measurement and calibration. To promote the application of AR in Facility Management (FM), which typically involves both complex indoor and outdoor environments, an automatic localization method is needed, thus providing accurate AR registration in mixed and unprepared environments. In order to fill this gap, a system implementing a markerless approach for seamless indoor-outdoor AR registration has been developed by integrating multiple AR registration engines with a cloud platform. These engines, primarily relying on Global Navigation Satellite Systems Real-Time Kinematic (GNSS-RTK) and Computer Vision (CV) technologies, are managed by an additional system component that automatically assigns priority depending on the current scenario. The system proposed in this study is tested on site on an FM use case related to a university campus for qualitative assessment. Furthermore, a quantitative assessment of the system's accuracy is also performed. In both evaluations, the system shows very promising results in terms of (i) applicability for FM operations, and (ii) accuracy. Specifically, BIM holograms are automatically superimposed to their real counterparts seamlessly, in mixed indoor-outdoor scenarios, without manual procedures, demonstrating applicability of the proposed approach in unprepared environments. In terms of accuracy, the AR overlaying discrepancies in outdoor and indoor scenarios (D_{L-SQ}) result equals* $0.090\ m$ *and* $0.082\ m$ *with a maximum among all the analyzed scenes (D_M) of* $0.075\ m$ *and* $0.071\ m$*, respectively. Given these results, the work presented in this paper provides a substantial contribution to the dissemination of AR technologies in the FM field and to the support of FM activities in terms of better coordination, visualization, communication and on-field access to high-quality information. This paper extends the work presented at the 23rd International Conference on Construction Applications of Virtual Reality (CONVR 2023).*

KEYWORDS*: Augmented Reality, Seamless Registration, Feature Matching, 6-DoF Pose Estimation, Real-Time Kinematic, Building Information Modeling, Facility Management.*

REFERENCE*: Leonardo Messi, Francesco Spegni, Massimo Vaccarini, Alessandra Corneli & Leonardo Binni (2024). Seamless Augmented Reality Registration Supporting Facility Management Operations in Unprepared Environments. Journal of Information Technology in Construction (ITcon), Special issue: 'Managing the digital transformation of construction industry (CONVR 2023)', Vol. 29, pg. 1156-1180, DOI: 10.36680/j.itcon.2024.051*

1. INTRODUCTION

The AECO industry, although commonly recognized as one of the least digitized industries, is adopting computer-based technologies to provide better performance in various stages of the lifecycle of buildings (Albahbah et al., 2021). The Operation and Maintenance (O&M) phase of Facility Management (FM) accounts for the largest proportion of the whole life costs of the building process (Salman & Ahmad, 2023). The costs of the O&M phase represent 50–70% of the total annual facility operating costs and 85% of the entire lifecycle cost of a building. The complexity of facilities has grown constantly over the years, as well as the complexity of the daily management tasks they require. In fact, the increased need of the construction industry for visualization technologies arises from the complex nature of the industry and its high demand for information access for assessment, communication, and collaboration. Lack of coordination between facility managers and field workers results in delays and cost overruns which could easily be avoided with better coordination and visualization tools.

In this domain, Augmented Reality (AR) technologies can be used as visualization and interaction tools for facilities' O&M tasks and can provide significant advantages (Gómez Jáuregui et al., 2019; Salman & Ahmad, 2023). For example, current inspection and maintenance practices are characterized by scattered and disoriented facility information that the operators must retrieve through specifications, maintenance reports, and checklists. As a consequence, 50% of the on-site maintenance time is still spent on localizing and navigating targets inside a facility (Salman & Ahmad, 2023). Even after locating the target, operators must put additional effort into seeing the target as it could be concealed in the case of piping, overhead ducts or behind a wall.

AR, which can support similar tasks even by displaying concealed elements, poses several demanding technological requirements for its implementation, such as display and precise position tracking (Costanza et al., 2009). In fact, in order to give the illusion that virtual objects are located at fixed physical positions or aligned to their physical counterparts, the system must know the position of relevant physical objects relative to the display system. This problem, also known as spatial registration, has been considered since the earliest days as a core part of AR functionality (Albahbah et al., 2021; Salman & Ahmad, 2023). Spatial registration can combine virtual objects and the real environment with the correct spatial perspective relationship by calculating the corresponding relation of both the virtual world and the real-world coordinate systems (Cheng et al., 2020). More in detail, spatial registration is responsible for calculating the user's correct spatial position and orientation in accordance with the real-world coordinate systems (Albahbah et al., 2021). Although great efforts have been made so far by researchers in the spatial registration domain, further contributions are required. In fact, to promote the application of AR in the FM, which typically involves both indoor and outdoor environments, an advanced localization method that can provide a seamless registration in heterogeneous and vast scenarios is needed. To boost AR applicability to real-life scenarios, as in the case of FM, there is the need for an AR system not requiring any preliminary preparation of the environment. Scene preparation, which is typically required for those systems defined as marker-based, involves the deployment and/or use of alignment infrastructures within the real environment, such as distinctive patterns recognizable by cameras, beacons for sending signals, predefined models, and so on. The preparation process also includes further processing on the virtual model for reference placement. For those reasons, the environment preparation processes are tasks that imply a significant amount of measurement and calibration both in real and virtual scenes, hence limiting scalability. Furthermore, users may also not be skilled enough or allowed to prepare the considered environment (Azuma et al., 1999).

In order to fill these gaps, a localization system for seamless indoor-outdoor AR registration has been developed by defining a cloud platform that hosts two registration engines, one based on Global Navigation Satellite Systems Real-Time Kinematic (GNSS-RTK) and the other based on Computer Vision (CV). Between these two, there is an integrated engine switcher that manages the priority between them. The system proposed in this study is tested on site on an FM use case related to a university campus for qualitative assessment. A quantitative assessment of the system's accuracy is also performed while it is being used. In both assessments, the system shows very promising results in terms of (i) applicability for FM operations, and (ii) accuracy. This paper extends the work presented at the 23rd International Conference on Construction Applications of Virtual Reality (CONVR 2023). The remainder of this paper is structured as follows. In Section 2, a literature review is presented. Section 3 reports the methodology adopted for the development of the proposed system and its implementation. In Section 4, experiments design, execution, and results discussion are presented. Finally, Section 5 is dedicated to conclusions and limitations.

2. LITERATURE REVIEW

In this section, a literature review concerning existing AR registration methodologies is reported. A differentiation is made between technologies mainly applied to either indoor (Section 2.1) or outdoor (Section 2.2) environments for description purposes, as they face different challenges. Understanding strength and possible gaps in approaches proposed by past studies and commercially available solutions paves the way to the definition of the indoor-outdoor seamless registration system proposed by this study.

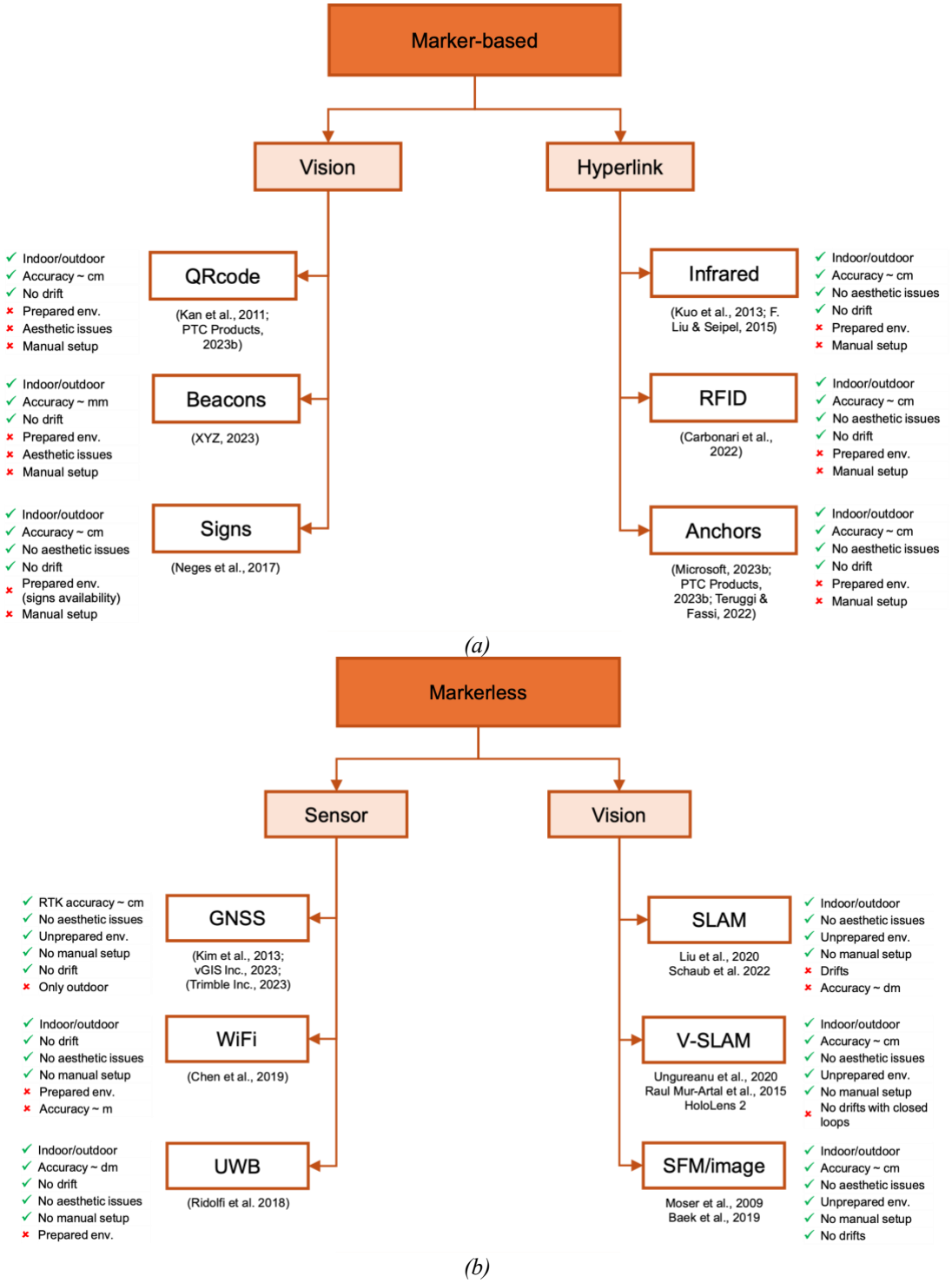


Figure 1. Taxonomy of AR registration technologies: (a) marker-based and (b) markerless.

It is common knowledge that AR can provide a relevant contribution to FM. For example, AR can support locating elements to be maintained in both indoor and outdoor environments and provide real-time instructions (Salman & Ahmad, 2023). To make AR fully applicable, some challenges must be addressed; in particular, spatial registration, which is the process responsible for calculating the user position and orientation to overlap virtual and real entities (Albahbah et al., 2021). In this regard, it is particularly useful to define a taxonomy of existing spatial registration technologies (localization technologies). As shown in Figure 1, localization technologies can be hierarchically divided into marker-based and markerless technologies.

Marker-based methods, which are the most widely used spatial registration methods, register virtual holograms using some kind of visible or hyperlink markers (El Barhoumi et al., 2022). In the first group, markers can be 2D images with visual features or natural 3D objects in the real environment (Cheng et al., 2020). The extensive use of these methods may be owing to the simplicity, efficiency, and availability of image recognition for superimposing virtual objects to the real world. Image recognition methods rely on extracting features from images instead of using complicated algorithms to calculate the relationship between relative positions. A similar approach can be applied with invisible hyperlink markers, such as infrared and RFID ones (El Barhoumi et al., 2022). These solutions can be potentially applied in mixed environments, but they require a preliminary preparation of the environment that usually includes a survey, marker positioning both in the real environment and in the virtual scene, and a manual setup. Furthermore, markers often generate aesthetic issues that, in some cases (e.g., historic buildings), cannot be accepted (Baek et al., 2019).

On the other hand, markerless methods solve some of the marker-based issues. In fact, instead of tracking features of markers, localization technologies are used to check the relative position between the real environment and virtual objects. For example, GNSS-based methods are the most popular markerless localization technology due to their suitability for use in a large open area such as a construction site and the ease of signal reception by common mobile devices (Cheng et al., 2020). Such registration methods can cope with unprepared environments as they only rely on the GNSS tracking system for registering the position of the user, without the need for any kind of preparation (Azuma et al., 1999). These are usually for wide or complex environments, or those that the user cannot verify because they are not sufficiently skilled or they are not allowed to (Azuma et al., 1999). Preparing an environment for an AR system is hard work, requiring a significant amount of measurement and calibration that limits scalability. However, GNSS-based methods are limited to outdoor scenarios, since buildings block the GNSS signal (Salman & Ahmad, 2023). Other markerless localization technologies, such as Wi-Fi, Ultra-WideBand (UWB), Inertial Measurement Unit (IMU), Simultaneous Localization and Mapping (SLAM), and image comparison, can be applied indoors. Nevertheless, Wi-Fi and UWB require a preliminary preparation of the environment, whereas IMU and SLAM are affected by drifts and closed loops (El Barhoumi et al., 2022). Vision-based methods using natural features from images for registration purposes can solve most of the previous issues and work very well indoors (Schönberger & Frahm, 2016). Image comparison solutions require a preliminary survey for collecting reference images of the real scene only. This task is not considered as a preparation of the environment, as it requires no special efforts or skills. However, it is hard to collect reference images of large assets (such as infrastructure). Therefore, this technology has limitations in usability in large scenarios such as the outdoor ones.

2.1 Indoor AR registration methodologies

In the AECO industry, several AR registration methodologies for indoor applications have been developed and tested so far. Past studies have exhaustively tested marker-based approaches using vision-based and hyperlink-based markers (Figure 1 (a)). Vision-based markers are distinctive in the scene and, even though they are advantageous in terms of robustness in detection, they should be installed all over the facility before on-site activities commence (e.g., FM activities) (Lee & Akin, 2011; Park et al., 2013). Examples of vision-based markers are QRcode (Kan et al., 2011; PTC Products, 2023) and beacons (XYZ, 2023). Visual markers can trigger aesthetic issues because of their distinctive appearance. This issue can be overcome using natural markers (Koch et al., 2014; Neges et al., 2017), with the limitation of dependance on signs, including exit signs, fire extinguisher signs, and textual information signs. If the scene does not have such designated signs, localization can be restricted (Baek et al., 2019). Alternative solutions are represented by invisible markers, such as infrared (Kuo et al., 2013; F. Liu & Seipel, 2015) and Radio Frequency IDentification (RFID) (Carbonari et al., 2022; Naticchia et al., 2021) technologies, which do not aesthetically change the scene. However, even though invisible markers do not have aesthetic issues, they should be pre-installed. Another solution is spatial anchors provided by commercial AR

libraries, such as Vuforia (PTC Products, 2023) and World Locking Tools (WLTs) (Microsoft, 2023b). These are tested in indoor environments (Ashour et al., 2022; El Barhoumi et al., 2022) and comparative tests show better results of WLTs over Vuforia Image Target (Teruggi & Fassi, 2022). Marker-based methods, despite their applicability indoors, require a preparation of the considered environment.

Among markerless methods (Figure 1 (b)), GPS-based AR systems are studied widely (Kim et al., 2013). Although they work in unprepared environments, they are unsuitable for indoor applications because of their low accuracy (Chen et al., 2019). Therefore, many studies employ Wi-Fi fingerprinting technology for indoor localization purposes (Ahmad et al., 2020; Chen et al., 2019). This approach loses accuracy in the case of multiple mobile devices. In fact, the localization accuracy of the order of 1 m, ensured by Wi-Fi-based collaborative systems (Chen et al., 2019), can still be improved. UWB technologies ensure higher accuracy (Ridolfi et al., 2018) compared to Wi-Fi fingerprinting; despite this, it is not enough for accurate AR registration. SLAM is the computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it (J. Liu et al., 2020; Schaub et al., 2022). The process of using vision sensors to perform SLAM is called Visual Simultaneous Localization and Mapping (VSLAM) (Microsoft, 2023a; Mur-Artal et al., 2015; Ungureanu et al., 2020). Utilizing visual data in SLAM applications has the advantages of more cost-effective hardware requirements, more straightforward object detection and tracking, and the ability to provide rich visual and semantic information (Tourani et al., 2022). In order to limit drifts of such technologies, trajectories must follow closed loops. Another markerless methodology suitable for indoor localization, that does not require closed loops and works with different cameras, is based on image comparison through CV technologies. Image-based localization is classically tackled by estimating a camera pose from correspondences established between scattered local features (Ethan Rublee et al., 2011) and a 3D Structure-from-Motion (SfM) (Mooser et al., 2009; Schönberger & Frahm, 2016) map of the scene (Li et al., 2012). Image comparison methodologies are classified into direct-matching and image-retrieval methodologies (Baek et al., 2019). Direct-matching methodologies do not render images for dataset, but directly find correspondences between 3D structure and the queried image (Humenberger et al., 2020). This pipeline scales to large scenes using image retrieval (Cao et al., 2020). Image-retrieval methodologies attempt to find the closest image to the queried image among the preliminarily prepared dataset. The dataset images can be preliminarily collected by photographs or rendered from three-dimensional structure estimation. Recently, many of these steps or even the end-to-end pipeline have been successfully learned with neural networks (DeTone et al., 2017; Lindenberger et al., 2023; Sarlin, Cadena, et al., 2018). Although this approach may lose accuracy whenever there is lack of context or repetitive elements, it does show great potential to develop mainly indoor AR registration apps with applicability for non-expert users (Baek et al., 2019). In fact, image-based localization methods work in unprepared environments. Collecting images is not considered a preparation of the environment since it requires no special efforts or skills. In light of this, image comparison through CV technologies is the most promising technology for indoor AR registration. Despite its applicability outdoors being possible, managing large image dataset could be challenging, hence suggesting the need for it to be integrated with other solutions for outdoor AR registration towards a seamless solution.

2.2 Outdoor AR registration methodologies

AR registration in outdoor environments presents distinct challenges compared to those encountered in indoor environments, such as (i) the absence of readily available reference points, (ii) expanded spatial dimensions, (iii) computational costs in large environments, and (iv) the dynamic nature of external spaces that undergoes frequent changes. Therefore, the AR registration approach in outdoor environments cannot solely rely on local reference systems but must necessarily be based on an absolute reference system, enabling the determination of the geographic pose of both the user and virtual objects (Cyrus et al., 2019; Ling et al., 2019; Marchand et al., 2016). To this end, hybrid AR registration approaches, based on the combined use of IMU and high-precision GNSS such as the Real-Time Kinematics (RTK), are pursued for the visualization of underground pipelines and subsurface data (Hansen et al., 2021; He et al., 2006; Roberts et al., 2002), for urban navigation (Guarese & Maciel, 2019; Zhao et al., 2016), for agricultural vehicle navigation (Kaizu & Choi, 2012), and for the alignment of multiple smaller maps from an existing SLAM tracking system (Ling et al., 2019). Even from a commercial standpoint, there are currently not many available solutions that ensure the use of AR apps in outdoor environments either without relying on some additional infrastructures (e.g., QRcodes, beacons, RFIDs, etc.) or without the need for manual/semi-manual alignment procedures and scene preparation, with some exceptions. For instance, Trimble Site Vision (Trimble Inc., 2023) makes use of the built-in GNSS receiver to achieve centimeter-level horizontal

positioning accuracy under RTK coverage. Similarly, Engineering-grade AR for AEC (vGIS Inc., 2023) achieves the same centimeter-level positioning accuracy under RTK coverage. In this case, the RTK antenna is not directly integrated into the system but needs to be obtained from third-party vendors. In any case, relying on GNSS technologies only means that the system cannot cope with urban-canyon scenarios and indoor environments (Figure 1 (b)). Due to this limitation and the persistently high costs, these solutions have not yet experienced widespread adoption in the construction industry. Nevertheless, the GNSS-RTK, ensuring accurate outdoor localization without any preliminary preparation of the environment, appears to be one of the most promising technologies for applicability in outdoor unprepared environments. For this reason, it is the ideal solution to be integrated with CV technologies for developing an AR registration system that works seamlessly in mixed and unprepared environments.

2.3 Research questions

As demonstrated by the literature review, several AR registration methodologies exist. Limitations of existing indoor AR registration approaches must be considered. Marker-based approaches share the limitation of requiring a preliminary survey to install vision-based or hyperlink-based markers or the existence of signs to be used as natural markers (Baek et al., 2019). Among markerless sensor-based approaches, GNSS-based solutions are inappropriate for indoor applications because of the weakness of the signals (Chen et al., 2019), whereas Wi-Fi-based and UWB-based solutions do not ensure enough accuracy (Salman & Ahmad, 2023). On the other hand, although they may lose accuracy whenever there is lack of context or repetitive elements, vision-based solutions using images show great potential for development of indoor AR registration applications with applicability for non-expert users (Baek et al., 2019).

In outdoor scenarios, on the other hand, marker-based approaches can barely be applied due to the need for installation of several markers to cover large environments. Similarly, such large scenarios can hardly cope with markerless vision-based solutions due to possible difficulties in managing a large open space or the need to carry out a large survey (Azuma et al., 1999). Also, limitations of the markerless GNSS-based approaches must be considered. Firstly, the reduced reliability of GNSS in urban-canyon scenarios (e.g., proximity to urban elements, such as buildings, roofs, trees, and so on) limits applications possibilities (Cheng et al., 2020; Ling et al., 2019). In addition, GNSS-based solutions are currently expensive (especially for high-precision GNSS-RTK systems). Despite this, since single RTK receiver components are becoming available at affordable prices, the development of in-house devices is showing promising growth (Hansen et al., 2021).

Finally, outdoor AR registration approaches are affected by the lack of integration with indoor scenarios, except through the use of additional supporting infrastructure (such as beacons) that constrain the deployment area (Cheng et al., 2020). Considering that indoor-outdoor interactions are crucial for built environment management, it is important that AR applications can work seamlessly even during changes of environment. In addition, in order to ensure wide AR application, registration procedures should work in unprepared environments. If AR registration could be achieved without the need to carefully prepare the environment in advance, that would be a major step in reducing the difficulty of operating an AR system.

In order to fill these gaps, this study aims to answer the following research questions:

RQ1 What system architecture would enable seamless markerless AR indoor-outdoor registration?

RQ2 What technical solutions would make AR registration infrastructureless in unprepared environments for wider applicability even among non-expert users?

RQ3 How can a seamless indoor-outdoor AR registration methodology support facility management operations?

RQ4 What is the hologram registration accuracy of the proposed AR registration system?

3. METHODOLOGY

Given the gaps identified in the literature review (Sections 2.1 and 2.2), and in response to the RQs previously identified (Section 2.3), this paper proposes an approach to AR hologram registration of BIM models supporting FM operations such that: (i) no preparation of the environment is required; (ii) no manual alignment is required; (iii) adopted technologies are not subject to drift issues; and (iv) seamless use in mixed indoor-outdoor scenarios

is guaranteed. Based on this approach, a system architecture that meets the above requirements is thereby proposed and its accuracy assessed based on the hologram registration error.

The architecture of the proposed system and its components are described in Section 3.1. The implementation of the system is covered in Section 3.2. Finally, the methodology followed to assess both qualitative and quantitative performances of the proposed system is defined in Section 3.3.

3.1 System architecture

In order to answer the research questions (Section 2.3), the system architecture reported in Figure 2 is defined. The proposed architecture is built on an AR cloud platform and on an AR client (the AR device), hosting the following four services: (i) the data storage, (ii) the GNSS-based AR registration engine, (iii) the image-based AR registration engine, and (iv) the engine switcher (Figure 2). The AR Cloud Platform serves as a decentralized resource and processes hub that facilitates the data processing, storage, and distribution tasks by means of a RESTful API. The latter represents a universal, web-oriented, communication interface between the aforementioned services and several client and user interfaces across the Internet. An important characteristic of the platform is its ability to host, localize, and align BIM models, images, and point clouds within a common geospatial context. This geolocation feature enables the precise mapping of virtual assets and features to their corresponding real-world locations (based on the WGS-84 standard), facilitating integration between the virtual and physical realms. One of the key responsibilities of the AR cloud platform is therefore to manage the alignment processes. Particularly, the correct positioning of images within the platform (also known as image registration) is achieved by referencing the absolute world coordinates of the acquisition point, along with the accurate rotations for a full 6-Degree-of-Freedom (6-DoF) pose estimation. This process ensures that images (and point clouds) are precisely georeferenced and aligned to their real-world locations. This approach lays the foundation for understanding the subsequent paragraphs, which delve into the concept of "images in the vicinity of the user's position".

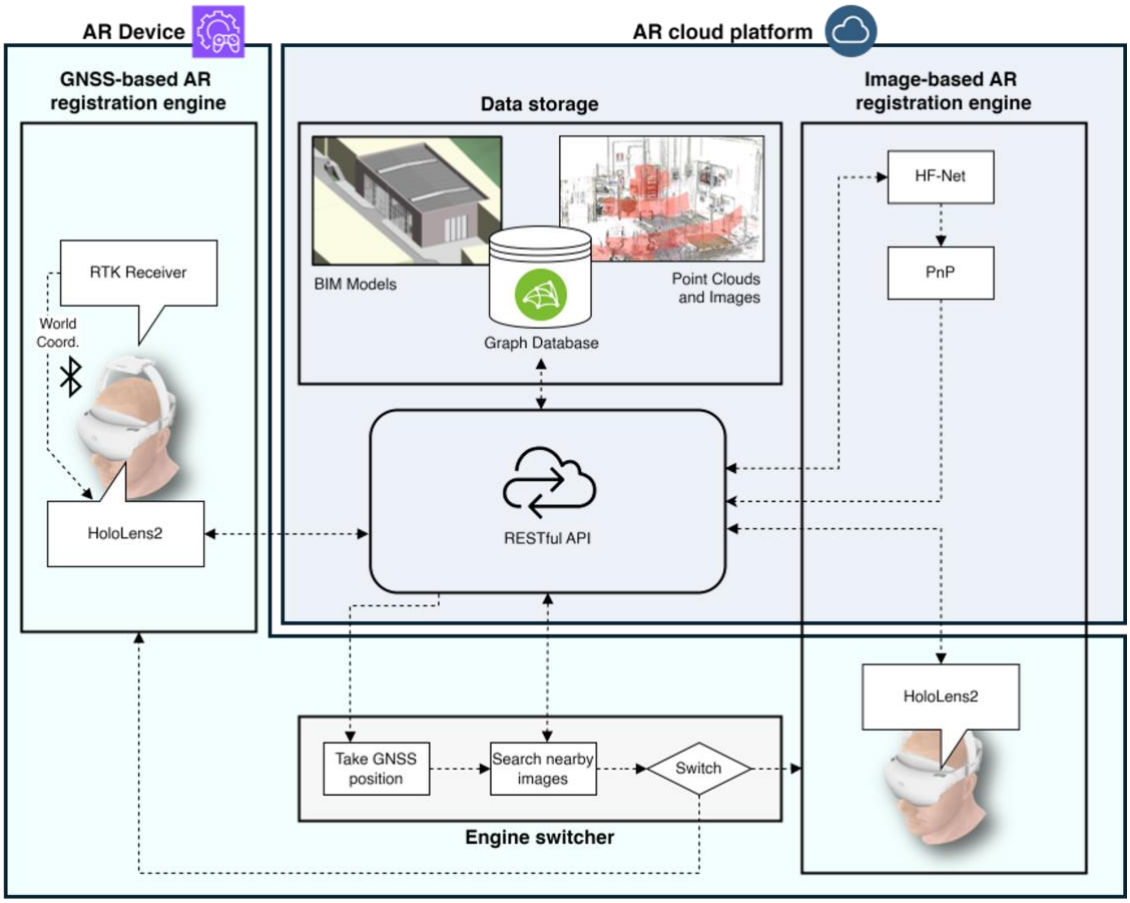


Figure 2: Architecture of the proposed system for seamless and markerless indoor-outdoor AR registration.

The data storage environment is responsible for storing and managing structured (e.g., .ifc files) and unstructured (e.g., images) data. It eases accessibility to AR applications through dedicated clients. A graph database provides a resilient backbone, offering efficient storage as well as a powerful query language for retrieval and traversal of interconnected heterogeneous data elements. Unstructured data, such as binary large objects (e.g., images, point clouds, etc.), are stored in a proper data-lake and they are linked to the structured information stored in the graph database so they can be easily accessed. The next integral components are the two distinct registration engines, specialized for outdoor and indoor environments, respectively. The GNSS-based AR registration engine, which relies on the combination of GNSS-RTK and the IMU system of the AR device, is tailored to tackle the unique challenges presented in open spaces, such as dealing with the absence of reliable reference points and coping with large and dynamic environments. On the other hand, the image-based AR registration engine is based on computer vision algorithms and designed to excel in environments characterized by restricted access to GNSS signals, leveraging features like point clouds and images to achieve accurate positioning. For this purpose, Convolutional Neural Networks (CNN) that simultaneously predict local features and global descriptors are applied for accurate 6-DoF localization (Sarlin, Cadena, et al., 2018; Sarlin, Debraine, et al., 2018). Finally, the pivotal feature of this system lies in the engine switcher, which effectively serves as an integrator between the GNSS-based and image-based AR registration engines. It is a rule-based engine that assesses the availability of either GNSS signals or features and it dynamically switches between the two registration approaches to maintain a consistent and uninterrupted AR experience. By synergistically combining the aforementioned four elements within the AR cloud platform, the system architecture (Figure 2) delivers a robust markerless AR system that can seamlessly adapt to both indoor and outdoor scenarios (i.e., answer to RQ1). Although the approach and methodology proposed in this paper are applicable to both AR Hand Held Devices (HHDs) and Head Held Devices (HMDs), the focus from this point forward will be specifically on the usage of the Microsoft AR tool, HoloLens 2. To this end, a novel addition to the tool has been introduced, developed, and physically manufactured, enabling a robust connection between the HoloLens 2 and an RTK receiver for precise calibration between the systems.

3.1.1 GNSS-based AR registration engine

Following the works available in literature (Hansen et al., 2021; Ling et al., 2019), the GNSS-based AR registration engine proposed in this paper relies on the combination of a GNSS-RTK tracking system and the HoloLens 2 built-in inertial tracking system. By aligning the local frame reference of the HoloLens 2 device to global coordinates, a geographical SLAM algorithm can be developed in order to have an absolute 6-DoF localization of the AR device and therefore accurately aligned virtual objects (Figure 3). Some considerations must be made:

- a general 3D object in the AR cloud platform has its own reference frame located in the world that is fully specified by its geographical coordinates and its orientation with respect to the North: Latitude (°), Longitude (°), Altitude (m), and Azimuth (°). The object coordinates refer to the WGS-84 standard;
- the coordinates retrieved from the RTK system refer to the WGS-84 standard;
- the RTK receiver is developed in-house, assembling electronic devices in a 3D printed case, with contained costs;
- the RTK receiver and the HoloLens 2 must be firmly connected to each other in order to have a fixed distance correction factor between the RTK receiver position and the centre of the AR device;
- the RTK system is reliable as long as the receiver is within 10 km of the RTK base station antenna;
- the local coordinate frame of HoloLens 2 originates at the point where the AR application is turned on.

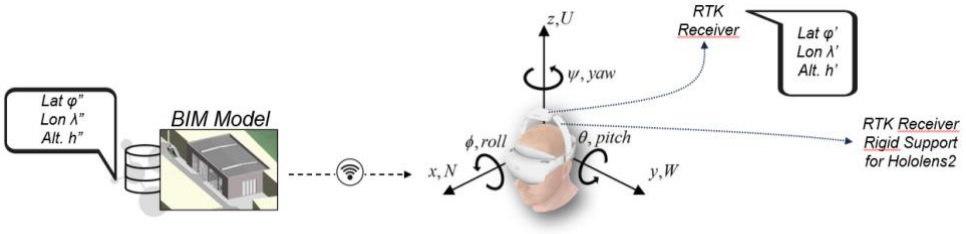


Figure 3: A diagram of the parameters involved in the computation of the geographical 6-DoF pose of the AR device.

The objective of placing world-referenced 3D BIM objects into the HoloLens 2 local frame can be achieved by carrying out the following steps: (i) aligning the local frame with the North direction, (ii) adjusting the object position, (iii) adjusting the object altitude, and (iv) placing objects based on the distance from the observer. This sequence of tasks constitutes the hologram registration process at the core of the GNSS-based AR registration engine, and it is executed automatically (i.e., answer to RQ2). Hence, this registration process, which does not require any particular actions from the user, may find applicability even among less experienced users in unprepared environments.

The first part of the tracking involves the initialization phase of the HoloLens 2 position. It includes acquiring initial samples from both the RTK and the IMU systems. Once initialized, the RTK system provides absolute 3D coordinates of the body frame (Latitude φ', Longitude λ', and Altitude h'), while the IMU provides for local 3D coordinates (x, y, z) and rotations (ϕ, θ, ψ) of the body frame. The equirectangular projection method is used to convert absolute global coordinates into planar coordinates, the horizontal x axis denotes longitude λ, and the vertical y axis generally denotes latitude φ. The forward projection (i.e., Equation 1) transforms spherical coordinates into planar coordinates. The reverse projection (i.e., Equation 2) transforms from the plane back onto the sphere.

$$\begin{cases} x = R(\lambda - \lambda_0)\cos\varphi_1 \\ y = R(\varphi - \varphi_0) \end{cases} \tag{1}$$

$$\begin{cases} \lambda = \dfrac{x}{R\cos\varphi_1} + \lambda_0 \\ \varphi = \dfrac{y}{R} + \varphi_0 \end{cases} \tag{2}$$

Where λ is the longitude of the location to project, φ is the latitude of the location to project, φ_1 is the standard parallel (North and South of the Equator) where projection is at true scale, φ_0 is the central parallel of the map, λ_0 is the central meridian of the map, x is the horizontal coordinate of the projected location on the map, y is the vertical coordinate of the projected location on the map, and R is the radius of the globe.

When trying to locally project geographical coordinates into planar ones, it is preferable to adopt the origin of the local frame as central point (φ_0, λ_0) and the corresponding latitude as reference φ_1, where no distortion is produced. Therefore, we will assume here that $\varphi_0 = \varphi_1$ and that the movement is made within a small surround of the central point (φ_0, λ_0).

The position of the local frame (φ_0', λ_0') is the first unknown to be solved for. When the user's GNSS coordinates (φ', λ') and local planar coordinates (x, y) are known, the geographical position of its reference frame can be computed by inverting the reverse projection as follows:

$$\lambda_0' = \lambda' - \frac{x}{R\cos\varphi_1} \tag{3}$$

$$\varphi_0' = \varphi' - \frac{y}{R} \tag{4}$$

The local equirectangular projection must have y axis directed toward the North. However, at the beginning, the HoloLens 2 local frame has an arbitrary unknown orientation with respect to the North. To solve this problem, it is necessary to consider the movement of the user. When it moves from the starting position, the path followed has an orientation β' with respect to the y' axis. The same movement defines a bearing angle β with respect to the North as shown in Figure 4.

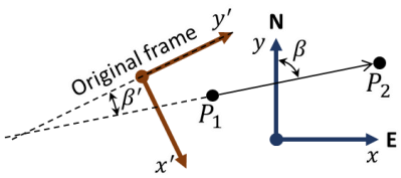


Figure 4: Bearing angles with respect to the North direction.

Given the latitude and longitude of the start point P_1 (φ_1, λ_1) and the latitude and longitude for the end point P_2 (φ_2, λ_2) of a straight line along a great-circle area, the initial bearing (sometimes referred to as forward azimuth)

can be computed as follows:

$$\beta = \operatorname{atan2}\left(\frac{\sin(\lambda_2' - \lambda_1')\cos(\varphi_2')}{\cos(\varphi_1')\sin(\varphi_2') - \sin(\varphi_1')\cos(\varphi_2')\cos(\lambda_2' - \lambda_1')}\right) \quad (5)$$

Once the bearing is known, the local frame can be adjusted to make the y axis look North by simply rotating (counterclockwise) the local frame by the vertical axis for an angle equal to $\beta - \beta'$.

On the other hand, when the origin of the local frame geographic position (φ_0', λ_0') and the object geographic coordinates (φ'', λ'') are known, the corresponding local planar coordinates can be computed by the forward projection:

$$x = R(\lambda'' - \lambda_0')\cos\varphi_1' \quad (6)$$

$$y = R(\varphi'' - \varphi_0') \quad (7)$$

When the observer moves too far from the origin, the local reference should be translated into the new position and the 3D objects must be positioned with respect to the new reference system by re-applying the previous equations.

The distance between the observer at a certain geographic position (λ', φ') w.r.t. the object in geographic position (λ'', φ'') must be computed by using the forward projection formula (i.e., Equation 1). The corresponding distance is then:

$$d = R\sqrt{\left((\lambda' - \lambda'')\cos\varphi'\right)^2 + (\varphi' - \varphi'')^2} \quad (8)$$

To track the true value of the observer altitude, each time a GNSS measurement is acquired for the observer's altitude h', its height in the local coordinate system can be stored for later use as height of the origin of the local frame. Consequently, given the altitude of the local frame (h_0'), its height in local coordinates z_0 and the altitude of an object (h''), the corresponding vertical coordinate z of an object can be computed by:

$$z = h'' - h_0' + z_0 \quad (9)$$

If the observer vertically moves the objects at z' to match their true height from the ground, the resulting true altitude is computed and stored. When the observer moves too far from the origin, the local reference should be vertically translated into the new altitude and the 3D objects must be positioned with respect to the new reference system by re-applying the previous equations.

3.1.2 Image-based AR registration engine

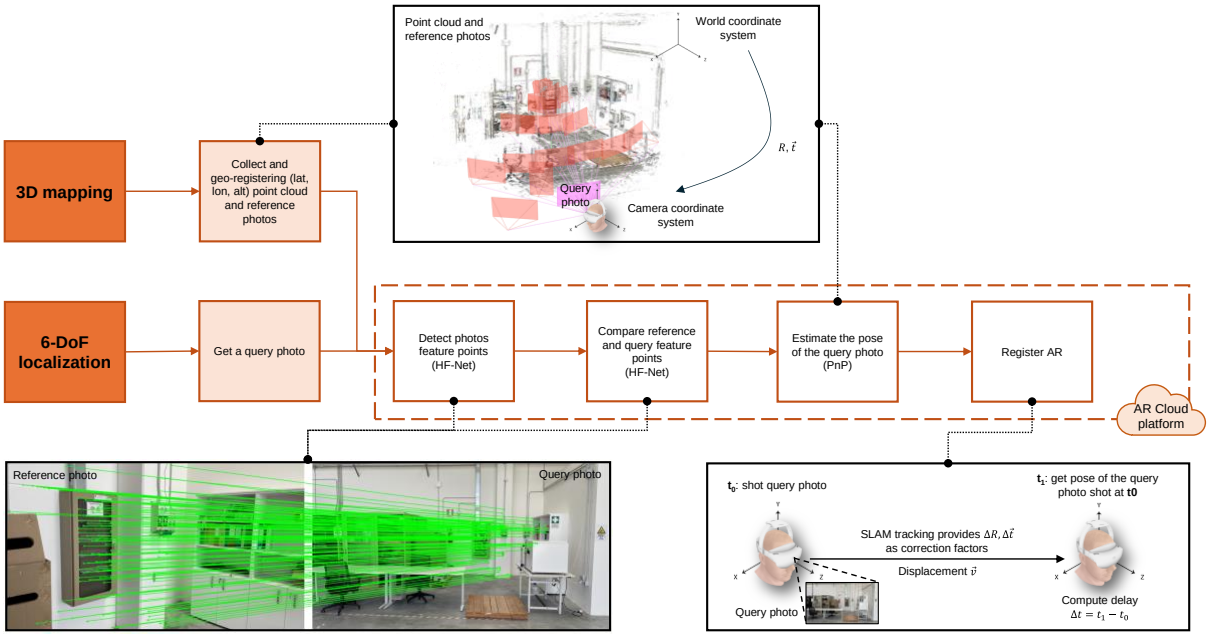


Figure 5: A diagram of the image-based AR registration engine's processes.

The image-based AR registration engine is designed to achieve accurate positioning and cover environments characterized by restricted access to GNSS signals by leveraging visual features extracted from and matched between images through CV algorithms. As depicted by Figure 5, the functioning of this engine can be described by the following two macro-phases: 3D mapping and 6-DoF localization.

The 3D mapping consists of carrying out a survey for the environment in question, aiming to collect point clouds and photos. Such a survey can be either executed using a combination of cameras and a LiDAR scanner (e.g., GeoSLAM ZEB Horizon), or simply with a camera (e.g., smartphone). In the first case, a survey directly provides a georeferenced 3D mapping as the point cloud and photos with absolute coordinates. In the second case, a 3D reconstructed model of the environment, to be used as a reference for localization, must be generated from images in post-processing. In fact, point clouds can be generated from image collections using the incremental SfM methodology implemented by the COLMAP library (Schönberger & Frahm, 2016). In this case, an additional step is required to geo-register and scale the 3D reconstructed model by providing absolute coordinates of at least three features or shot points.

Afterwards, the 6-DoF localization consists of estimating the pose of the HoloLens 2 by comparing a frame from its current view (i.e., query image) with the 3D mapping of the analyzed environment (i.e., reference images and point cloud), stored in the AR cloud platform. In this study, the Hierarchical Feature Network (HF-Net) technology is implemented for image comparison (Sarlin, Cadena, et al., 2018; Sarlin, Debraine, et al., 2018). The HF-Net consists of a CNN able to extract from and match visual features between reference and query images by computing local and global descriptors. The "hierarchical" attribute refers to the HF-Net feature close to the human aptitude of naturally localizing, in a previously visited environment, with a "from-coarse-to-fine" approach. In other words, humans first localize themselves by looking at the global scene appearance and subsequently inferring an accurate location from a set of likely places using local visual clues. This means that for each HoloLens 2 registration call, a coarse search, consisting of global-descriptor matching between the query image and the reference images, is performed. Afterwards, a finer search based on local-descriptor matching between the query image's 2D keypoints and the point cloud's 3D points co-visible in reference images is executed. Given a set of point cloud's 3D points and their corresponding 2D projections in the image, the pose of the query image can be estimated by solving the Perspective-n-Point (PnP) pose problem (Kneip et al., 2011). The estimated pose is described, as reported as follows, by a position vector $\vec{t}$ and a quaternion $\vec{q}$, referred to the camera coordinate system:

$$\vec{t}_{\text{ccs}} = (t_x,\, t_y,\, t_z) \tag{10}$$

$$\vec{q} = (q_w, q_x,\, q_y,\, q_z) \tag{11}$$

Once the quaternion $\vec{q}$ is transformed into the corresponding 3x3 rotation matrix $R(\vec{q})$ (Kuipers, 1999), the coordinates of the HoloLens 2 camera can be converted into the world coordinate system of 3D mapping. The transformation matrix is given by the negative transpose rotation matrix $-R(\vec{q})^T$ (Schönberger & Frahm, 2016), as reported as follows:

$$R(\vec{q}) = \begin{bmatrix} 2(q_w^2 + q_x^2) - 1 & 2(q_x q_y - q_w q_z) & 2(q_x q_z - q_w q_y) \\ 2(q_x q_y - q_w q_z) & 2(q_w^2 + q_y^2) - 1 & 2(q_y q_z - q_w q_x) \\ 2(q_x q_z - q_w q_y) & 2(q_y q_z - q_w q_x) & 2(q_w^2 + q_z^2) - 1 \end{bmatrix} \tag{12}$$

$$\vec{t}_{\text{wcs}} = -R(\vec{q})^T \cdot \vec{t}_{\text{ccs}} \tag{13}$$

This step is required to express all the $\vec{t}_{\text{ccs}}$ in the same coordinate system and hence compare them. At this point, since the 3D mapping world coordinate system is referenced to the global coordinate system, the user is localized and holograms are projected. Finally, it must be noted that the AR registration must consider that the user may move while the pose is being computed. So, delay correction factors Δt and ΔR can be retrieved by the tracking module of the HoloLens 2 and applied to the AR registration.

Once the survey of the interested area is completed, the presented image-based AR registration engine process (Figure 5) can be automatically executed and thus enables automatic AR registration. Hence, the image-based AR registration process, not requiring the user to carry out any particular actions, can be applied even in unprepared environments (i.e., answer to RQ2).

3.1.3 The registration engine switcher

As illustrated in Figure 2, the engine switcher serves as a seamless integrator between the two types of registration,

contributing to answer both RQ1 and RQ2. This component is a rule-based engine that enables seamless AR registration.

There are 4 possible context scenarios in which the operators may find themselves while using the system:

1. the first scenario outlines situations where the RTK service is available, yet the database lacks geo-located images and/or point clouds of the user's surroundings. This scenario is most frequently encountered in outdoor environments;
2. the second scenario outlines situations where geo-located images and/or point clouds of the user's surroundings are available, but the RTK service is not accessible. Such scenarios are typically encountered in indoor environments;
3. the third scenario represents situations where both the RTK service and images and/or point clouds of the environment are available. Such circumstances may arise especially in the proximity of externally mapped buildings.
4. the last scenario outlines situations where the database lacks images and/or point clouds, and the RTK service remains unavailable. Such circumstances primarily arise within indoor environments. Given how important it is to digitize existing assets and the simplicity of environment mapping, also facilitated by smartphone photography (as described in Section 3.1.2), please refer to scenario 3;

According to the described scenarios, the engine switcher is responsible for changing the AR registration method depending on the presence of images and/or point clouds in the user's vicinity.

3.2 System implementation

In this section, the implementation of the system architecture presented in Section 3.1 is described. Section 3.2.1 provides details about the implementation of the GNSS-based AR registration engine, whereas Section 3.2.2 describes the implementation of the image-based AR registration engine. Finally, Section 3.2.3 is devoted to the engine switcher.

The system architecture is implemented as a set of micro-services deployed using state-of-the-art containerization technologies (i.e., Docker). The services are integrated using RESTful APIs and together form the so-called backend of the AR cloud platform.

3.2.1 Implementation of the GNSS-based AR registration engine

The GNSS-based AR registration engine is implemented in Unity. It is worth noting that, given the micro-services architecture of the system, the GNSS-based AR registration engine is also a client of the AR cloud platform and is directly deployed on the AR device. Basically, as shown in Figure 6, this engine makes use of both the RTK-GNSS and HoloLens 2 tracking systems to compute the distances traveled in global (geographic) and local coordinates. The engine is initialized at position T_0 and obtains the first global and local position samples. While moving, the position changes to T_n. The distance $\overrightarrow{T_0T_n}$ is computed using both global and local samples. If the distances are comparable, the engine proceeds computing the Azimuth (i.e., the bearing angle β w.r.t. the North direction, clockwise), otherwise it sets the initial position as the current one and waits for new samples acquired during the next iteration. The Azimuth is then filtered and used to rotate the local coordinate system, aligning the y axis to the North direction. Then, if the accuracy h_{Acc} of the RTK sample is sufficient (i.e., horizontal accuracy lower than a given threshold corresponding to the GNSS receiver in RTK Fix status), the system translates the origin of the local coordinate system to the last computed position and consequently places the asset's holograms in the space.

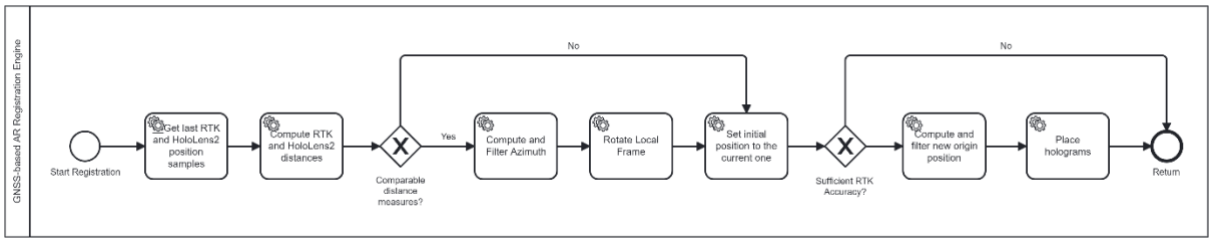


Figure 6. BPMN describing the implementation of the GNSS-based AR registration engine.

3.2.2 Implementation of the image-based AR registration engine

The image-based AR registration engine is implemented as a set of micro-services deployed using state-of-the-art containerization technologies (viz. Docker). Figure 7 depicts as a BPMN the implementation of the image-based AR registration engine, assuming 3D mapping based on a camera (e.g., smartphone) and SfM (e.g., COLMAP library (Schönberger & Frahm, 2016)). At the core, there is a backend web-service exposing a RESTful API which allows:

- initializing, deleting, and renaming a project;
- uploading of more photos into the 3D mapping dataset of the project or deletion of existing photos;
- commencement of a new 3D mapping task, identifying feature points of the dataset photos, identifying the pose of each photo, and assigning (x, y, z) coordinates to each identified feature point, even generating a 3D point cloud from the dataset photos;
- uploading and 6-DoF localizing of a query photo w.r.t. the 3D point cloud generated during the 3D mapping task;
- return of the pose of the query photo to HoloLens 2 to perform AR registration.

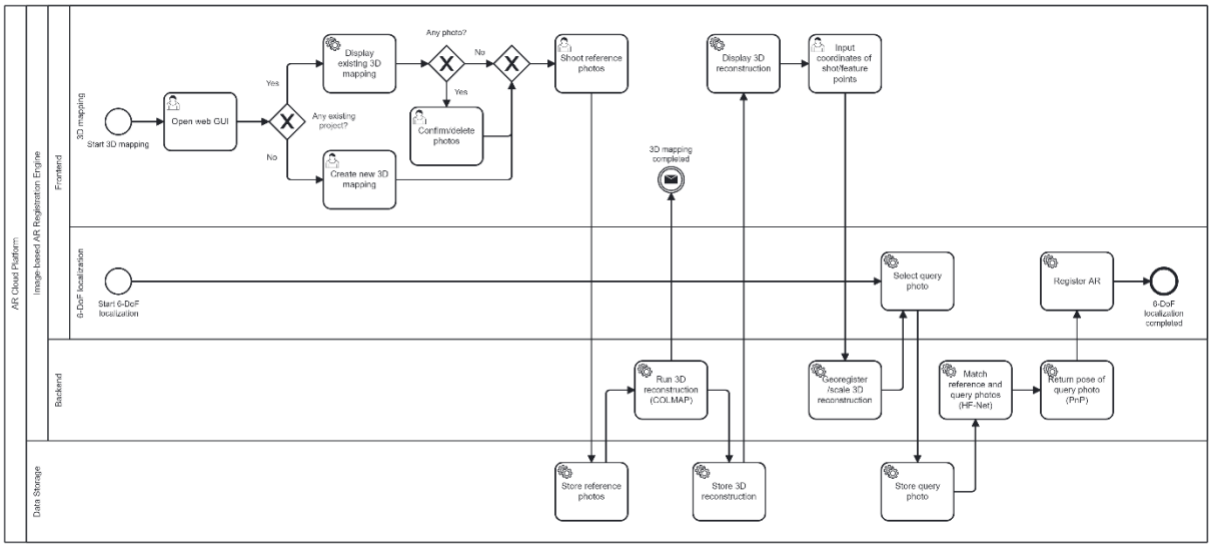


Figure 7. BPMN describing the implementation of the image-based AR registration engine.

The backend web service stores permanent information using volumes provided by a Network-Attached Storage (NAS): while the original pictures are stored as files, the extracted information (e.g., feature points and 3D point clouds) is stored using one database for each project. In addition to the backend web service, a minimal responsive frontend web GUI allows to test the implemented architectures from HoloLens 2 (Figure 8).

3.2.3 Implementation of the engine switcher

Following the methodology presented in Section 3.1.3, the implementation of the engine switcher is based on the acquisition of the user's current geographical position and on the search for images and/or point clouds in the vicinity of the user (Figure 9). Thus, the user's location, acquired by the RTK receiver, although it may not be accurate (as it is probably in an indoor environment as described by scenario 2 in Section 3.1.3), is sent seamlessly to the engine switcher. The current geographic location of the user is used to query the database for images in the vicinity of the location itself. Note that the threshold distance by which the word "vicinity" is understood in this text is discretionary. A higher threshold potentially includes a higher number of features and slows down the computation. On the other hand, an excessively low threshold may result in too narrow a search and therefore in uncovered areas in cases where the retrieved position is not very accurate (mainly indoors). If features are found in the database within the threshold distance, the engine switcher activates the image-based AR registration engine. Otherwise, if no features are found by the query, the system activates the GNSS-based AR registration engine. The procedure is looped in order to achieve seamless switching.

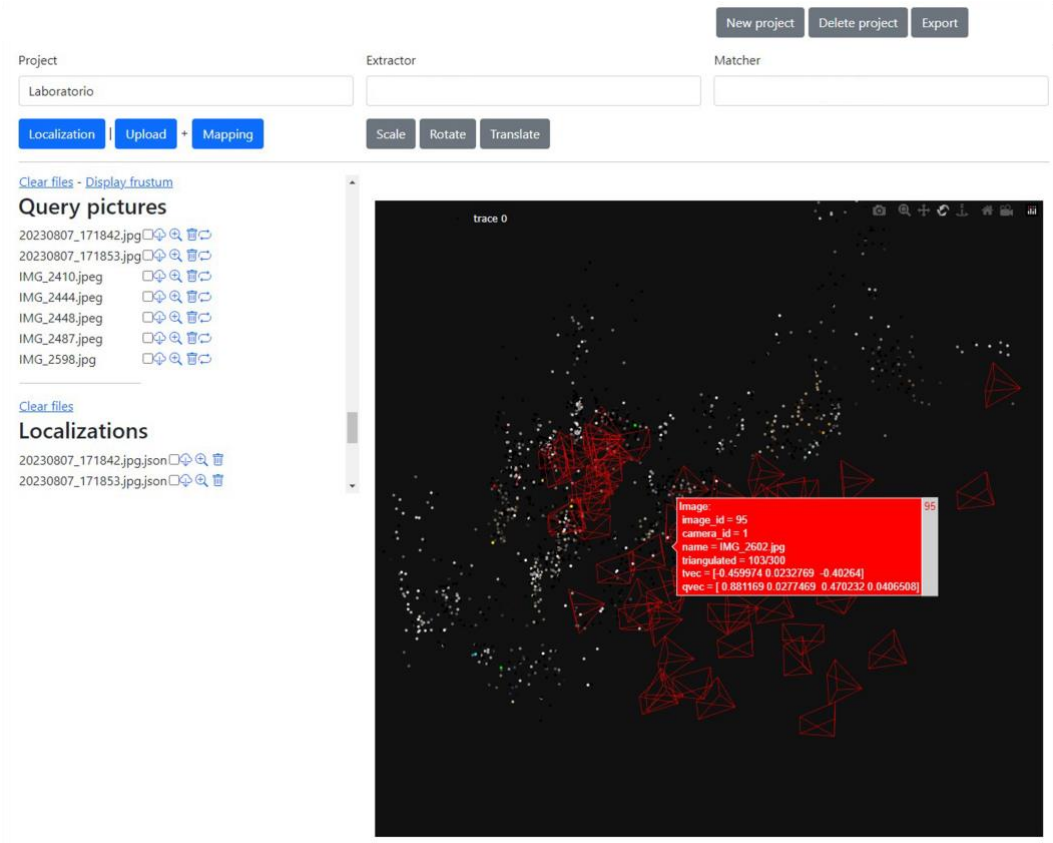


Figure 8. Frontend GUI of the image-based AR registration engine.

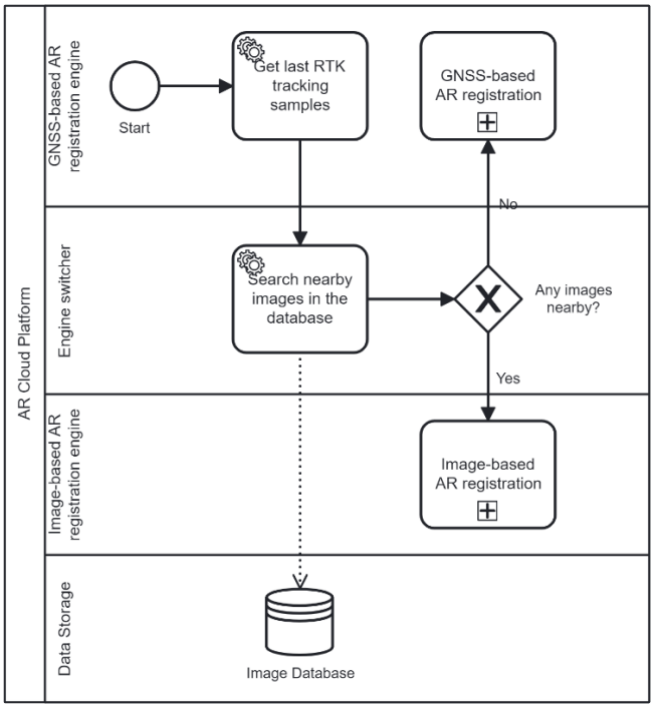


Figure 9. BPMN describing the implementation of the engine switcher.

3.3 Evaluation methodology

The seamless indoor-outdoor AR registration solution presented in this paper has undergone both qualitative and quantitative assessments. The former aims to prove the relevance of the proposed seamless AR registration solution for applicability in unprepared environments during FM operations (i.e., answer to RQ2 and RQ3). The latter, on the other hand, consists of carrying out an accuracy assessment of the proposed GNSS-based and image-based AR registration engines (i.e., answer to RQ4).

The qualitative assessment of the proposed system is based on features that an AR system must exhibit for easy and scalable use in FM scenarios. Thus, the hypotheses underlying the qualitative assessment of the proposed system are the following:

- during FM activities, the operator must not resort to manual procedures for registering the device pose and holograms. Any manual procedures waste time, disrupt activities and may require expert skills;
- the system must be deployable in unprepared scenarios. The need to prepare real and virtual environments using markers and/or other infrastructures prior to the use of the system narrows its area of use, drastically increases operation times, may require expert skills, and finally, limits scalability;
- FM scenarios usually include mixed indoor and outdoor environments. The system must be seamless. Any interruptions in the service and/or misalignments that occur as a result of the change of environment would lead to a limitation in the use of the system;
- the system must not be subject to drift, especially for medium to long distances, as activities may extend over large spaces depending on the type of asset (e.g., road infrastructure, large building sites, and so on). Drift issues restrict the system's area of use and limit scalability;

On the other hand, the quantitative evaluation methodology proposed for the AEC/FM domain is applied in this study to measure hologram registration errors (Gómez Jáuregui et al., 2019). The so-called registration error occurs when the virtual objects displayed in the AR device appear in the wrong position in relation to the real environment. It must be noted that the deviation shown in pixels in a 2D screenshot cannot be used to measure the real displacement of the virtual points with respect to the real points, since the sightline between the observer and the virtual point holds infinite positions in the 3D space. It would be necessary to combine two or more different perspectives to calculate the actual position of the virtual point. However, in practice, it may not be possible for two reasons: (i) it is very unlikely that those sightlines intersect in a single point and (ii) each one of those virtual points in the different pictures would be affected differently by the lens (i.e., they would be placed in different positions of the distortion map). Therefore, the deviation between virtual and real points must be assessed in real scale, and not in screen pixels, as the deviation of the sightline $P_{prp} - P_v$ of a virtual point P_v with respect to the real position of that point taken by the camera P_r (Figure 10). Hence, $P_{v_{UL}}$, measured on the photo w.r.t. the upper-left-screen coordinate system, must be converted into world-coordinates. For this purpose, the following transformation matrix M from world coordinates to upper-left-screen pixels is computed and inverted:

$$M = S_{UL} \cdot S_{CS} \cdot M_P \cdot R \cdot T \rightarrow M^{-1} \tag{14}$$

where R and T are the rotation and translation matrixes from the UTM world (X_w, Y_w, Z_w) to the viewing (X_v, Y_v, Z_v) coordinate system respectively, M_P is the transformation matrix from the viewing to the perspective-projection coordinate system, S_{CS} is the transformation matrix from the perspective-projection to the center-screen (X_{cs}, Y_{cs}, Z_{cs}) coordinate system, and S_{UL} is the transformation matrix from the center-screen (X_{cs}, Y_{cs}, Z_{cs}) to the upper-left-screen (X_{ul}, Y_{ul}, Z_{ul}) coordinate system (Figure 10). At this point, $P_{v_{UL}}$ can be converted into world coordinates as follows:

$$P_{v_{WC}} = M^{-1} \cdot P_{v_{UL}} \tag{15}$$

Considering k to be different scenes where point P is observable, the shortest distance in world coordinates between the sightline $P_{prp} - P_v$ and the position of the real element P is calculated for each scene k by finding the length of the perpendicular drawn from the point to the line:

$$dist(P, P'_v) = \min dist\ (P, \overline{P_{prp} P_v}) \tag{16}$$

For the purpose of this study, a survey carried out with GeoSLAM Backpack Vision provides ground truth UTM world coordinates of P points. Furthermore, the least-squares intersection of the sightline $P_{prp} - P_v$ of each scene

is calculated to find the point $\dot{P}$ that best fits the intersection of those sightlines of a same point from different points of view (one for each scene). The point $\dot{P}$, which minimizes the sum of perpendicular distances to the sightlines of the k scenes, can be computed by the following linear system of equations:

$$R \cdot \dot{P} = q \tag{17}$$

$$R = \sum_{J=1}^{k} \left(I - h_j \cdot h_j^T\right), q = \sum_{J=1}^{k} \left(I - h_j \cdot h_j^T\right) \cdot P_{prp_j} \tag{18}$$

Where h_j is the unit vector of the j-th sightline $P_{prp} - P_v$, and I is the identity matrix. Finally, according to (Gómez Jáuregui et al., 2019), two values are considered to estimate the AR superimposition accuracy:

- D_M: maximum distance between the sightlines $P_{prp} - P_v$ and the real position of point P;
- D_{L-SQ}: distance between the optimum point $\dot{P}$ achieved at the least-squares solution and the real position of point P.

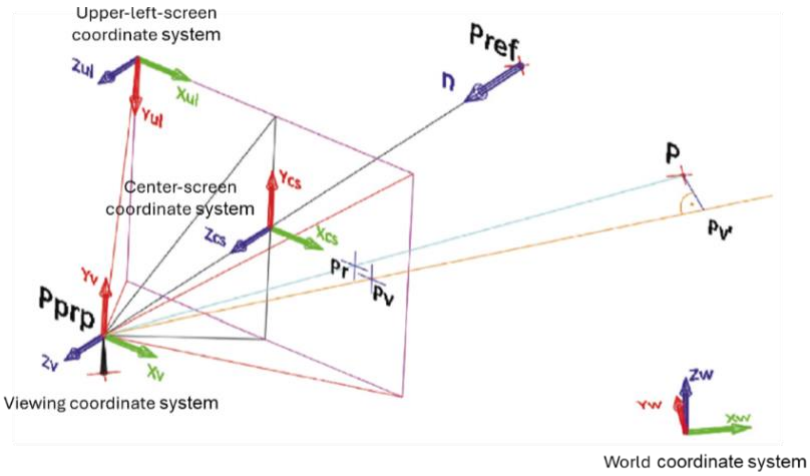


Figure 10: Scheme of the registration error assessment adapted from (Gómez Jáuregui et al., 2019).

4. EXPERIMENTS AND RESULTS

In this section, experiments and results are reported. Two kinds of experiments have been conducted. The first one, reported in Section 4.1 and referred to as the "experiment 1", aims to provide a qualitative assessment of the developed system during facility management operations. The second one, on the other hand, reported in Section 4.2 and referred to as "experiment 2", aims to provide a quantitative assessment of the developed system in terms of overlaying discrepancies between holograms and reality.

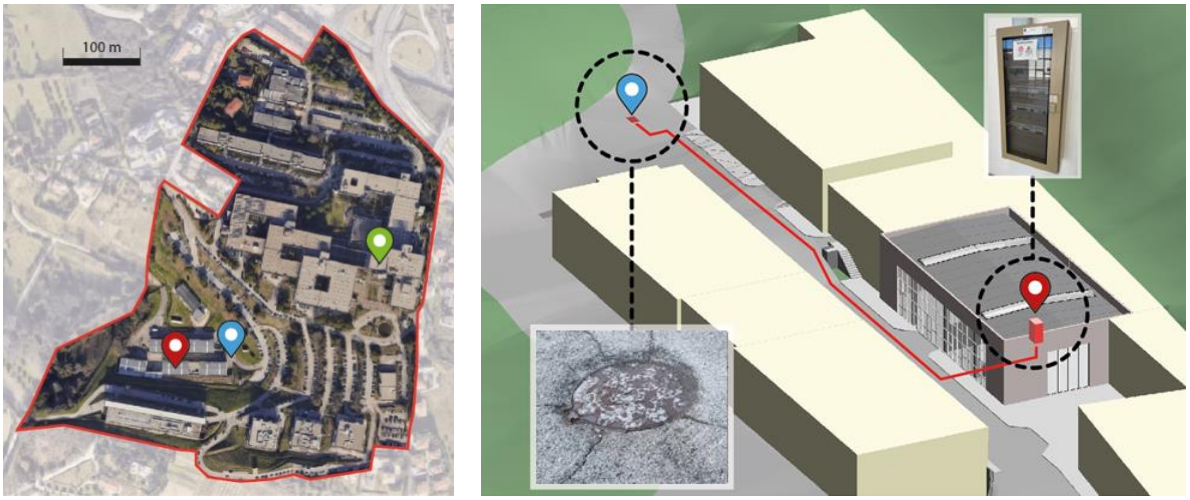
4.1 Qualitative assessment

The system implemented in this study is tested on an FM use case based on a university campus, assumed as a case study. This experiment will be referred to as experiment 1. It aims to prove the relevance of the proposed seamless AR registration solution during facility management operations (i.e., answer to RQ2 and RQ3). Specifically, the study focused on the FM of the Digital Construction Capability Centre (DC3) Lab at the Università Politecnica delle Marche. The DC3 Lab, which covers an area as large as $240\ m^2$, is composed of a main open space, a changing room, an office, and the restroom. Within this context, the management of the electrical system, and in particular of the internal electrical panel of the DC3 Lab, is considered. During this activity, the operator in charge of FM operations spends most of the time first locating the electrical panel. Then, in order to find the root cause of the problem, the technician may be asked to locate the panel's associated cabling, which extends outside the building. These cables can be accessed through manholes located on the road in front of the building, connecting it to the rest of the campus (Figure 11 (a)). Having located all the elements of interest to FM operations, the technician may need technical information about the electrical system. For this purpose, he/she needs to access the as-built BIM model (Figure 11 (b)). By testing the proposed system on the presented case, its applicability in real-life situations will be assessed. The experimentation primarily revolves around the

utilization of an AR application implemented on the HoloLens 2 device.

The developed system is tested on the selected use case in order to provide a qualitative assessment of the developed system during FM operations. First of all, a preliminary set-up phase consisting of collecting and initializing input data is carried out. This phase must only be done once since related input and settings are maintained. It includes:

1. collecting point clouds and referenced images by carrying out a survey of the DC3 Lab. In this study, the survey is conducted using GeoSLAM Backpack Vision (Figure 12 (a)), which simultaneously collects both point clouds and referenced photos (Figure 12 (b)). Alternatively, the survey can be carried out by collecting photos (e.g., with a smartphone) and then generating a point cloud through the SfM methodology implemented by the COLMAP library (Schönberger & Frahm, 2016);
2. collecting the BIM model of the DC3 Lab (Figure 11 (b));
3. uploading the BIM model, point cloud, and images related to the selected use case, onto the BIM cloud platform. It must be noted that point cloud and images are aligned directly from the survey. The BIM model must be aligned to the previous dataset by selecting reference points; it must be noted that this alignment is only done once since it is maintained.



(a) (b)

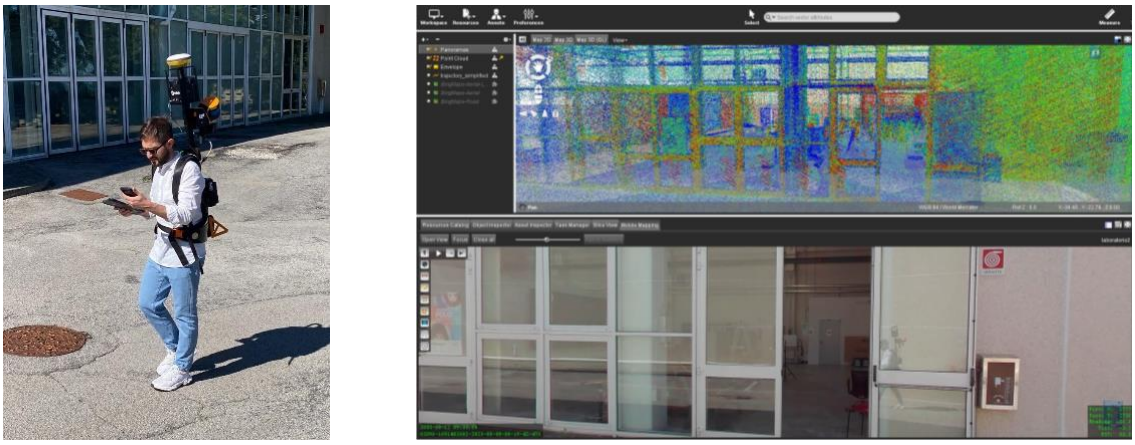
Figure 11: (a) Aerial view of the university campus, adopted as a case study, identifying the positions of the DC3 Lab (i.e., red placemark), the manhole cover (i.e., blue placemark), and the RTK base antenna (i.e., green placemark); (b) view of the BIM model of the DC3 Lab identifying the positions of the indoor electrical panel (i.e., red placemark) and the outdoor manhole cover (i.e., blue placemark).

Once the preliminary steps above are completed, the AR-based inspection of the electrical system related to the selected use case can start. In this study, the head-mounted AR device Microsoft HoloLens 2 is used (Figure 13). The AR application, based on the system architecture reported in Figure 1, is used to support the operators inspecting the electrical system distributed in a heterogeneous indoor/outdoor scenario. The main contribution of the proposed system is the markerless AR registration, seamlessly displaying BIM models superimposed to the whole inspected environment. The following steps are executed on-site:

4. putting on the HoloLens 2 with RTK receiver set up (Figure 13) and launching the AR application;
5. setting the search distance threshold. During these tests, the threshold is set at 20 meters (which corresponds to the average dimension of the case study building) to balance query time and number of features found;
6. moving outdoors around the campus of the Faculty of Engineering at Università Politecnica delle Marche. During this preliminary step outside the DC3 Lab, the GNSS-based AR registration engine is automatically triggered due to the absence of outdoor images;
7. heading to the internal electrical infrastructure to inspect the electrical panel (Figure 11). As the user moves

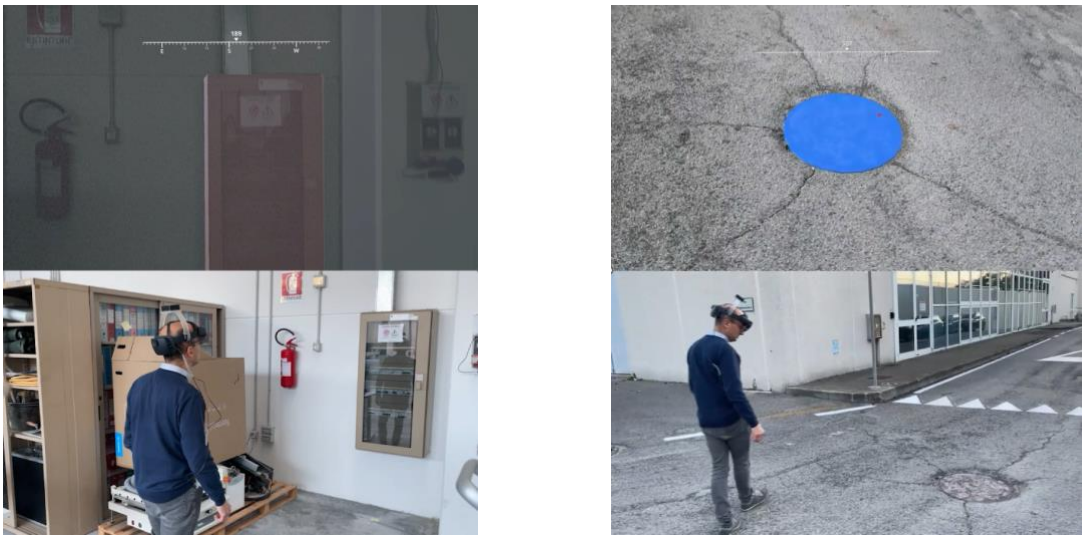
from the outdoor to the indoor environement, the GNSS coverage decreases and the collected dataset (i.e., point clouds and images) is found in the surroundings of the user position. Hence, the system automatically switches to the image-based AR registration engine. This transition occurs without interruption as the system switches between registration modes through the engine switcher's algorithm;

8. inspecting the internal electrical infrastructure, specifically focusing on the indoor electrical panel (Figure 13 (a));
9. heading to the external electrical infrastructure to inspect the cabling associated with the internal electrical panel;
10. inspecting the external electrical infrastructure, specifically focusing on the manhole cover located on the street facing the building (Figure 11 (b)). During the outdoor phase, the AR application relies on the GNSS-based AR registration engine, leveraging geospatial data retrieved from the RTK receiver to overlay BIM data on the real asset (Figure 13 (b)).



(a) *(b)*

Figure 12: (a) Survey phase of the DC3 Lab using GeoSLAM Backpack Vision for collecting (b) point clouds and geo-located photos.



(a) (b)

Figure 13: Visualization of the aligned holograms during the test inspection within an unprepared indoor-outdoor scenario: (a) the indoor electrical panel and (b) the outdoor manhole cover. The test is performed using the Microsoft HoloLens 2 integrated with an RTK receiver by means of a 3D printed support.

To sum up, tasks 1-3 are preliminary tasks that only need to be done once, since related input and settings are maintained. Tasks 4-9, on the other hand, describe FM operations related to the adopted case study and prove how a seamless indoor-outdoor AR registration can smoothen such activities.

4.2 Quantitative assessment of overlaying discrepancies

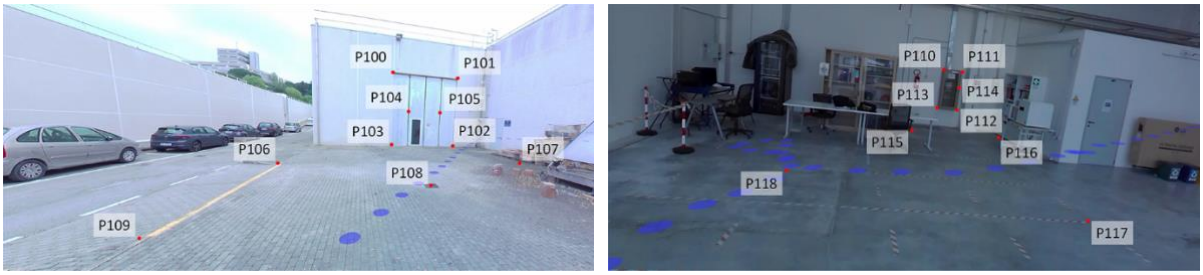
The system implemented in this study is tested in order to perform a quantitative assessment of overlaying discrepancies between holograms and reality. This experiment will be referred to as experiment 2. Essentially, an outdoor and indoor scenario related to the FM use presented in Section 4.1 are selected in order to test the proposed GNSS-based and image-based AR registration engines respectively (i.e., answer to RQ4).

For this purpose, the evaluation methodology described in Section 3.3 is considered to finalize the experiments and process collected data.

In the outdoor scenario, overlaying discrepancies between the real door and its hologram are assessed (Figure 14 (a)). In the indoor scenario, on the other hand, overlaying discrepancies between the real electrical panel and its hologram are assessed (Figure 14 (b)). In both scenarios experiments are carried out assuming that the preliminary actions as described by tasks 1-5 in the numbered list reported in Section 4.1, are already fulfilled. The quantitative assessment includes the following additional tasks:

1. placing the HoloLens 2 on a tripod near a projection reference point P_{prp} and directing the HoloLens 2 so that its central point view is aligned with a reference point P_{ref}. It must be noted that:
    a. the $P_{prp} - P_{ref}$ combination identifies a specific scene as it represents the position and orientation of the device;
    b. a setup offset must be applied to P_{prp} to retrieve the real position of the HoloLens 2 RGB camera taking the tripod height into consideration;
2. saving a frame of the HoloLens 2 view;
3. using Equation 14, computing the transformation matrix M^{-1} from upper-left-screen pixels to world coordinates;
4. w.r.t. the selected frame of the HoloLens 2, retrieving the virtual point $P_{v_{UL}}$ in the upper-left-screen coordinate system corresponding to real point P assumed as control point;
5. using Equation 15, converting $P_{v_{UL}}$ into the virtual point $P_{v_{WC}}$ in world coordinates;
6. using Equation 16, computing the shortest distance in world coordinates between the sightline $P_{prp} - P_v$ and the position of the real point P assumed as a control point;
7. moving the HoloLens 2 to the next scene $P_{prp} - P_{ref}$ combination;
8. using Equations 17 and 18, applying the least-squared method to all the shortest point-line distances computed through point 6 (i.e., one for each scene) to compute D_{L-SQ};
9. identifying D_M as the maximum distance of the scenes between the sightlines $P_{prp} - P_v$ and the real position of point P.

The experiment setup illustrating the position of the considered points in the real world for both outdoor and indoor scenarios is depicted in Figure 14. Three scenes for each of the two scenarios, defined by $P_{prp} - P_{ref}$ combinations, are summarized respectively in Table 1 and Table 2.



(a) *(b)*

Figure 14. Overview of the control point selected in both the (a) outdoor and (b) indoor scenarios.

The same tables report the accuracy performance of the proposed AR registration system for each of the two

scenarios expressed in terms of D_{L-SQ} and D_M. In the outdoor scenario, AR overlaying discrepancies between the virtual point estimated through the least-square method and the real control point is equal to $0.090\ m$ (i.e., D_{L-SQ}) with a maximum of $0.075\ m$ (i.e., D_M) in the single analyzed scenes (Table 1). In the indoor scenario, however, AR overlaying discrepancies are equal to $0.082\ m$ (i.e., D_{L-SQ}) with a maximum of $0.071\ m$ (i.e., D_M) in the single analyzed scenes (Table 2).

To sum up, tasks 1, 2, and 7 describe the part of the experimental campaign carried out on-site. Collected photos are processed in the back-office according to tasks 3-6, 8, and 9 to compute overlaying discrepancies between holograms and their physical counterparts.

Table 1. Analysis of results for the "P103" control point for three different outdoor scenes for the assessment of the GNSS-based AR registration system.

Scene	P_{prp}	P_{ref}	$P_{v_{UL}}$	$P_{v_{WC}}$	$Dist(P, P_{v'})\ [m]$	$D_{L-SQ}\ [m]$
1	P106 (379993.741, 4826946.908, 131.539,1)	P105 (380000.121, 4826937.583, 134.275,1)	(1196, 1661)	(382313.405, 4824864.476, -388.306)	0.073	
2	P107 (379993.547, 4826936.063, 131.528,1)	P104 (380000.197, 4826940.187, 134.285,1)	(1517, 1801)	(382414.942, 4828845.782, -390.340)	0.031	**0.090**
3	P109 (379988.289, 4826946.995, 131.522,1)	P105 (380000.121, 4826937.583, 134.275,1)	(1246, 1387)	(382772.413, 4825597.540, -199.011)	**0.075**	

Table 2. Analysis of results for the "P113" control point for three different indoor scenes to assess of the image-based AR registration system.

Scene	P_{prp}	P_{ref}	$P_{v_{UL}}$	$P_{v_{WC}}$	$Dist(P, P_{v'})\ [m]$	$D_{L-SQ}\ [m]$
1	P115 (380010.091, 4826932.621, 131.578, 1)		(426, 774)	(379035.559, 4825880.093, -179.343)	0.047	
2	P116 (380007.253, 4826932.751, 131.571, 1)	P114 (380008.166, 4826931.083, 132.887, 1)	(390, 765)	(380910.671, 4825802.936, -134.860)	**0.071**	**0.082**
3	P117 (380008.019, 4826937.851, 131.575, 1)		(553, 484)	(380125.287, 4825545.223, 46.161)	0.068	

4.3 Discussion

This section discusses how this paper contributes to the body of knowledge by answering the research questions formulated in Section 2.3. Since existing AR solutions do not ensure a seamless registration method in mixed indoor and outdoor scenarios, this study tries to fill this gap by presenting a novel system architecture (Section 3.1). The latter, built on an AR cloud platform and on an AR client, hosts a GNSS-based AR registration engine and an image-based AR registration engine. The integration between the two engines, ensuring a seamless AR registration in mixed environments, is provided by an engine switcher managing priorities between them. This answers RQ1.

Furthermore, two markerless technologies, namely GNSS-RTK (Section 3.1.1) and image comparison based on CV (Section 3.1.2), are adopted to ensure complementarily AR registration in all circumstances without the need for reference infrastructure to be installed prior, hence answering RQ2. In fact, GNSS-RTK best performs in large outdoor environments where carrying out complete 3D mapping would be too laborious. Image comparison based on CV, however, generally excels where the GNSS-RTK signal is weak (e.g., indoors and urban canyons). Such

technologies ensure an easy use of the proposed AR registration system emerging from experiment 1 (Section 4.1). This experiment, carrying out a qualitative assessment of the developed system during facility management operations, proves that it is possible to apply it to unprepared environments even among non-expert users. In fact, in experiment 1, tasks 1-3 (Section 4.1) are preliminary tasks that only need to be done once since related input and settings are maintained. Tasks 4-9 (Section 4.1), instead, prove that FM operations related to the adopted case study can be executed in unprepared environments and with no setup efforts by the user. This further answers RQ2.

In addition, experiment 1 demonstrates that efficiency of FM activities is expected to be considerably improved by the proposed AR registration system, ensuring automatic and seamless access to holograms of BIM models directly superimposed onto their physical counterparts. First, this would facilitate the time-consuming task of locating elements to be maintained and accessing their information in complex and mixed environments, including concealed elements. Second, given the possibility of linking BIM holograms to technical specifications of elements, maintenance reports, and checklists, the contribution of this study would support FM activities in terms of better coordination, visualization and communication. This answers RQ3.

The promising implications discussed so far are encouraged by the robustness of the proposed AR registration system that emerging from experiment 2 (Section 4.2). This experiment carries out a quantitative assessment of the developed system in terms of overlaying discrepancies between holograms and reality in terms of D_{L-SQ} and D_M (Section 3.3). The first one provides an overall measure of the holograms' registration error based on three perspectives (i.e., scenes). D_{L-SQ} is quite similar in both outdoor (Table 1) and indoor (Table 2) scenarios with values which do not exceed $0.090\ m$. Similarly, discrepancies in the single scenes registered maximum values D_M which do not exceed $0.075\ m$ in both outdoor (Table 1) and indoor (Table 2) scenarios. This analytically proves the robustness of the proposed AR registration system in mixed and unprepared environments. In fact, it ensures a stable alignment of holograms to their physical counterparts with centimeter-level discrepancies, enabling the user to freely explore even heterogeneous unprepared environments where FM operations take place. This answers RQ4. In addition, the results obtained in the current study confirm comparable results to the ones provided by previous studies for other markerless technologies (El Barhoumi et al., 2022).

5. CONCLUSIONS

This study focuses on AR registration in mixed indoor-outdoor and unprepared environments by firstly reviewing the state-of-the-art to identify existing gaps and then developing and implementing a seamless and infrastructureless AR registration system. To the best of the authors' knowledge, the one proposed in the current study is the first all-in-one system based on a seamless and infrastructureless methodology. In this respect, GNSS and image-based AR registration engines, combined with a priority switch, enable automatic alignment of holograms to their physical counterparts in mixed and unprepared environments, often subject to changes, without suffering from the drift issue. The AR registration system, described in detail in Section 3.1, has been designed to operate smoothly during changes in the environment.

Implemented using Microsoft HoloLens 2 as the AR interface, the system is tested on a university campus, adopting the maintenance of the electrical systems of a laboratory as a use case. This specific use case was selected based on the need for technicians to locate and inspect multiple mutually interconnected elements which are usually spread across diverse unprepared environments of the campus. Both qualitative and quantitative assessments are conducted for evaluation purposes. Qualitative evaluation shows how seamless and infrastructureless AR registration enhances facility management (FM) activities. The quantitative analyses, on the other hand, measure the discrepancies in hologram-reality overlay accuracy.

The findings demonstrated that the proposed system can successfully register holograms of BIM models automatically in mixed unprepared environments, without resorting to any manual procedure or suffering from drift issues, thereby simplifying operations and being accessible even to inexperienced users. Holograms of interior and exterior elements, such as an electrical panel and manhole cover respectively, are seamlessly aligned with their physical counterparts with high accuracy. Specifically, in the outdoor scenario, the measured discrepancy between the virtual point estimated through the least-square method and the real control point was equal to $D_{L-SQ} = 0.090$ m with a maximum of $D_M = 0.075$ m in the single analyzed scenes. In the indoor scenario, the measured discrepancy was $D_{L-SQ} = 0.082$ m with a maximum $D_M = 0.071$ m in the single analyzed scenes.

The efficiency of FM activities is expected to be considerably improved by automatic and seamless access to

holograms of BIM models directly superimposed onto their physical counterparts with centimeter-level accuracy. First, this would facilitate the time-consuming task of locating elements to be maintained and accessing their information in complex and mixed environments, including concealed elements. Secondly, given the possibility of linking BIM holograms to technical specifications of elements, maintenance reports, and checklists, the contribution of this study would support FM activities in terms of better coordination, visualization and communication directly on the field.

Current limitations of the proposed methodology can be reconducted to technologies adopted by the system's engines. Accuracy loss may affect image comparison, adopted by the image-based AR registration engine, in case of reference and/or query images with a lack of context or repetitive elements. Follow-up studies may quantify such limitation. On the other hand, the GNSS-based AR registration engine strongly relies on the availability of RTK coverage. Despite the fact that such technology is currently expensive, it should be noted that since the individual components of RTK systems are affordable, it is foreseeable that applications of this technology will proliferate.

Further studies will be carried out in order to optimize switching thresholds based on GNSS coverage and datasets (i.e., point clouds and images) availability in the surroundings of the user position. Future developments will also focus also on the definition of a graphical user interface to better manage the entire AR registration workflow. Finally, the proposed system will be provided to non-expert users in order to quantify its contribution in terms of time saving for completion of an FM task.

ACKNOWLEDGMENTS

This paper is an extended version of our previous work, which was presented at the 23rd International Conference on Construction Applications of Virtual Reality (CONVR 2023), in Florence, Italy. The authors acknowledge the support and feedback from Prof. Farzad Rahimian and Dr. Vito Getuli as the Chairs of the International Scientific Committee of CONVR 2023.

This research has received funding from the Vitality project – Project Code ECS00000041, CUP I33C22001330007 - funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 - 'Creation and strengthening of innovation ecosystems,' construction of 'territorial leaders in R&D' – Innovation Ecosystems - Project 'Innovation, digitalization and sustainability for the diffused economy in Central Italy – VITALITY' Call for tender No. 3277 dated 30/12/2021, and Concession Decree No. 0001057.23-06-2022 of the Italian Ministry of University funded by the European Union – NextGenerationEU.

REFERENCES

Ahmad, A., Claudio, P., Alizadeh Naeini, A., & Sohn, G. (2020). Wi-fi rss fingerprinting for indoor localization using augmented reality. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 5(4), 57–64. https://doi.org/10.5194/isprs-Annals-V-4-2020-57-2020

Albahbah, M., Kıvrak, S., & Arslan, G. (2021). Application areas of augmented reality and virtual reality in construction project management: A scoping review. Journal of Construction Engineering, 4, 151–172. https://doi.org/10.31462/jcemi.2021.03151172

Ashour, Z., Shaghaghian, Z., & Yan, W. (2022). BIMxAR: BIM-Empowered Augmented Reality for Learning Architectural Representations. https://doi.org/https://doi.org/10.48550/arXiv.2204.03207

Azuma, R., Lee, J. W., Jiang, B., Park, J., You, S., & Neumann, U. (1999). Augmented Reality Tracking in unprepared environments for augmented reality systems. Computers & Graphics, 23, 787–793. https://doi.org/10.1016/S0097-8493(99)00104-1

Baek, F., Ha, I., & Kim, H. (2019). Augmented reality system for facility management using image-based indoor localization. Automation in Construction, 99, 18–26. https://doi.org/10.1016/j.autcon.2018.11.034

Cao, B., Araujo, A., & Sim, J. (2020). Unifying Deep Local and Global Features for Image Search. http://arxiv.org/abs/2001.05027

Carbonari, A., Franco, C., Naticchia, B., Spegni, F., & Vaccarini, M. (2022). A Mixed Reality Application for the On-Site Assessment of Building Renovation: Development and Testing. Sustainability (Switzerland), 14(20). https://doi.org/10.3390/su142013239

Chen, K., Chen, W., Fellow, P., Li, C. T., Student, M., & Cheng, J. C. P. (2019). A BIM-based location aware AR collaborative framework for facility maintenance management. In Journal of Information Technology in Construction (ITcon) (Vol. 24). http://www.itcon.org/2019/19

Cheng, J. C. P., Chen, K., & Chen, W. (2020). State-of-the-Art Review on Mixed Reality Applications in the AECO Industry. Journal of Construction Engineering and Management, 146(2). https://doi.org/10.1061/(asce)co.1943-7862.0001749

Costanza, E., Kunz, A., & Fjeld, M. (2009). Mixed Reality: A Survey. Lecture Notes in Computer Science, 5440. https://doi.org/https://doi.org/10.1007/978-3-642-00437-7_3

Cyrus, J., Krcmarik, D., Moezzi, R., Koci, J., & Petru, M. (2019). Hololens used for precise position tracking of the third party devices - Autonomous vehicles. Communications - Scientific Letters of the University of Žilina, 21(2), 18–23. https://doi.org/10.26552/com.c.2019.2.18-23

DeTone, D., Malisiewicz, T., & Rabinovich, A. (2017). SuperPoint: Self-Supervised Interest Point Detection and Description. http://arxiv.org/abs/1712.07629

El Barhoumi, N., Hajji, R., Bouali, Z., Ben Brahim, Y., & Kharroubi, A. (2022). Assessment of 3D Models Placement Methods in Augmented Reality. Applied Sciences (Switzerland), 12(20). https://doi.org/10.3390/app122010620

Ethan Rublee, Vincent Rabaud, Kurt Konolige, & Gary Bradski. (2011). ORB: An efficient alternative to SIFT or SURF. 2011 International Conference on Computer Vision. https://doi.org/10.1109/ICCV.2011.6126544

Gómez Jáuregui, V., Manchado del Val, C., Castillo Igareda, J. Á. del, & Otero González, C. A. (2019). Quantitative evaluation of overlaying discrepancies in mobile augmented reality applications for AEC/FM. Advances in Engineering Software, 127, 124–140. https://doi.org/10.1016/j.advengsoft.2018.11.002

Guarese, R. L. M., & Maciel, A. (2019). Development and Usability Analysis of a Mixed Reality GPS Navigation Application for the Microsoft HoloLens. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11542 LNCS, 431–437. https://doi.org/10.1007/978-3-030-22514-8_41

Hansen, L. H., Fleck, P., Stranner, M., Schmalstieg, D., & Arth, C. (2021). Augmented Reality for Subsurface Utility Engineering, Revisited. IEEE Transactions on Visualization and Computer Graphics, 27(11), 4119–4128. https://doi.org/10.1109/TVCG.2021.3106479

He, Z., Xia, Z., Chang, Y., Chen, W., Hu, J., & Wei, X. (2006). Research on underground pipeline augmented reality system based on ARToolKit. In H. Wu & Q. Zhu (Eds.), Proc. SPIE 6421, Geoinformatics 2006: Geospatial Information Technology (p. 642112). https://doi.org/10.1117/12.713123

Humenberger, M., Cabon, Y., Guerin, N., Morat, J., Leroy, V., Revaud, J., Rerole, P., Pion, N., de Souza, C., & Csurka, G. (2020). Robust Image Retrieval-based Visual Localization using Kapture. http://arxiv.org/abs/2007.13867

Kaizu, Y., & Choi, J. (2012). Development of a tractor navigation system using augmented reality. Engineering in Agriculture, Environment and Food, 5(3), 96–101. https://doi.org/10.1016/S1881-8366(12)80021-8

Kan, T.-W., Teng, C.-H., & Chen, M. Y. (2011). QR Code Based Augmented Reality Applications. In Handbook of Augmented Reality. Springer New York. https://doi.org/10.1007/978-1-4614-0064-6

Kim, C., Park, T., Lim, H., & Kim, H. (2013). On-site construction management using mobile computing technology. Automation in Construction, 35, 415–423. https://doi.org/10.1016/j.autcon.2013.05.027

Kneip, L., Scaramuzza, D., & Siegwart, R. (2011). A Novel Parametrization of the Perspective-Three-Point Problem for a Direct Computation of Absolute Camera Position and Orientation. https://doi.org/10.1109/CVPR.2011.5995464

Koch, C., Neges, M., König, M., & Abramovici, M. (2014). Natural markers for augmented reality-based indoor navigation and facility maintenance. Automation in Construction, 48, 18–30. https://doi.org/10.1016/j.autcon.2014.08.009

Kuipers, J. B. (1999). Quaternions and rotation sequences: a primer with applications to orbits. Princeton University Press.

Kuo, C., Jeng, T., & Yang, I. (2013). An invisible head marker tracking system for indoor mobile augmented reality. Automation in Construction, 33, 104–115. https://doi.org/10.1016/j.autcon.2012.09.011

Lee, S., & Akin, Ö. (2011). Augmented reality-based computational fieldwork support for equipment operations and maintenance. Automation in Construction, 20(4), 338–352. https://doi.org/10.1016/j.autcon.2010.11.004

Li, Y., Snavely, N., Huttenlocher, D., & Fua, P. (2012). Worldwide Pose Estimation using 3D Point Clouds. https://doi.org/https://doi.org/10.1007/978-3-642-33718-5_2

Lindenberger, P., Sarlin, P.-E., & Pollefeys, M. (2023). LightGlue: Local Feature Matching at Light Speed. http://arxiv.org/abs/2306.13643

Ling, F. F., Elvezio, C., Bullock, J., Henderson, S., & Feiner, S. (2019). A Hybrid RTK GNSS and SLAM Outdoor Augmented Reality System. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 1044–1045. https://doi.org/10.1109/VR.2019.8798315

Liu, F., & Seipel, S. (2015). Infrared-visible image registration for augmented reality-based thermographic building diagnostics. Visualization in Engineering, 3(1). https://doi.org/10.1186/s40327-015-0028-0

Liu, J., Xie, Y., Gu, S., & Chen, X. (2020). A SLAM-Based Mobile Augmented Reality Tracking Registration Algorithm. International Journal of Pattern Recognition and Artificial Intelligence, 34(1). https://doi.org/10.1142/S0218001420540051

Marchand, E., Uchiyama, H., & Spindler, F. (2016). Pose estimation for augmented reality: a hands-on survey. https://doi.org/10.1109/TVCG.2015.2513408

Microsoft. (2023a). HoloLens2. https://www.microsoft.com/en-us/hololens/hardware#document-experiences

Microsoft. (2023b). World Locking Tools documentation. https://learn.microsoft.com/en-us/mixed-reality/world-locking-tools/

Mooser, J., You, S., Neaumann, U., & Wang, Q. (2009). Applying robust Structure from Motion to markerless augmented reality. 2009 Workshop on Applications of Computer Vision (WACV). https://doi.org/10.1109/WACV.2009.5403038

Mur-Artal, R., Montiel, J. M. M., & Tardos, J. D. (2015). ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Transactions on Robotics, 31(5), 1147–1163. https://doi.org/10.1109/TRO.2015.2463671

Naticchia, B., Vaccarini, M., Corneli, A., Messi, L., & Carbonari, A. (2021). Leveraging Extended Reality technologies with RFID to enhance on field maintenance of buildings. http://itc.scix.net/paper/w78-2021-paper-038

Neges, M., Koch, C., König, M., & Abramovici, M. (2017). Combining visual natural markers and IMU for improved AR based indoor navigation. Advanced Engineering Informatics, 31, 18–31. https://doi.org/10.1016/j.aei.2015.10.005

Park, C. S., Lee, D. Y., Kwon, O. S., & Wang, X. (2013). A framework for proactive construction defect management using BIM, augmented reality and ontology-based data collection template. Automation in Construction, 33, 61–71. https://doi.org/10.1016/j.autcon.2012.09.010

PTC Products. (2023). Vuforia Enterprise AR Software. https://www.ptc.com/en/products/vuforia

Ridolfi, M., Vandermeeren, S., Defraye, J., Steendam, H., Gerlo, J., Clercq, D. De, Hoebeke, J., & Poorter, E. De. (2018). Experimental evaluation of uwb indoor positioning for sport postures. Sensors (Switzerland), 18(1). https://doi.org/10.3390/s18010168

Roberts, G. W., Evans, A., Dodson, A. H., Denby, B., Cooper, S., & Hollands, R. (2002). The Use of Augmented Reality, GPS and INS for Subsurface Data Visualisation. https://doi.org/https://api.semanticscholar.org/CorpusID:112931894

Salman, A., & Ahmad, W. (2023). Implementation of augmented reality and mixed reality applications for smart facilities management: a systematic review. Smart and Sustainable Built Environment. https://doi.org/10.1108/SASBE-11-2022-0254

Sarlin, P.-E., Cadena, C., Siegwart, R., & Dymczyk, M. (2018, December 9). From Coarse to Fine: Robust Hierarchical Localization at Large Scale. 2019 IEEE CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.48550/arXiv.1812.03506

Sarlin, P.-E., Debraine, F., Dymczyk, M., Siegwart, R., & Cadena, C. (2018). Leveraging Deep Visual Descriptors for Hierarchical Efficient Localization. http://arxiv.org/abs/1809.01019

Schaub, L., Podkosova, I., Schonauer, C., & Kaufmann, H. (2022). Point cloud to BIM registration for robot localization and Augmented Reality. Proceedings - 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct, ISMAR-Adjunct 2022, 77–84. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00025

Schönberger, J. L., & Frahm, J.-M. (2016). Structure-from-Motion Revisited. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/CVPR.2016.445

Teruggi, S., & Fassi, F. (2022). Mixed Reality Content Alignment in Monumental Environments. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 43(B2-2022), 901–908. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-901-2022

Tourani, A., Bavle, H., Sanchez-Lopez, J. L., & Voos, H. (2022). Visual SLAM: What Are the Current Trends and What to Expect? In Sensors (Vol. 22, Issue 23). MDPI. https://doi.org/10.3390/s22239297

Trimble Inc. (2023). Site Vision. https://sitevision.trimble.com/

Ungureanu, D., Bogo, F., Galliani, S., Sama, P., Duan, X., Meekhof, C., Stühmer, J., Cashman, T. J., Tekin, B., Schönberger, J. L., Olszta, P., & Pollefeys, M. (2020). HoloLens 2 Research Mode as a Tool for Computer Vision Research. http://arxiv.org/abs/2008.11239

vGIS Inc. (2023). Engineering-grade AR for AEC. https://www.vgis.io/augmented-reality-bim-gis-ar-aec-civil-construction-engineering-bentley-autodesk-esri/

XYZ. (2023). Engineering Grade AR.

Zhao, S., Chen, Y., & Farrell, J. A. (2016). High-Precision Vehicle Navigation in Urban Environments Using an MEM's IMU and Single-Frequency GPS Receiver. IEEE Transactions on Intelligent Transportation Systems, 17(10), 2854–2867. https://doi.org/10.1109/TITS.2016.2529000