

SEAMLESS INDOOR/OUTDOOR MARKER-LESS AUGMENTED REALITY REGISTRATION SUPPORTING FACILITY MANAGEMENT OPERATIONS

Leonardo Messi, Francesco Spegni, Massimo Vaccarini, Alessandra Corneli & Leonardo Binni

Università Politecnica delle Marche (UNIVPM), Faculty of Engineering, DICEA Department, Ancona, Italy

ABSTRACT: *Augmented reality (AR) still struggles to be widely used in real processes in the construction industry despite its great potential. This is partly due to the difficulties that exist in aligning holograms and maintaining their stability, especially for outdoor applications. In addition, being indoor-outdoor interactions crucial for built environment management, it would be important that AR apps can work seamlessly. Alignment in indoor environments cannot make use of methods such as GNSS, nor can all environments be assumed to have been previously initialized with AR tools. Thus, marker-less AR registration is crucial for indoor applications. This paper presents an approach for marker-less AR registration seamlessly in both outdoor and indoor environments. Real-time kinematic positioning (RTK) and Inertial Measurement Units (IMU) technologies have been chosen for outdoor registration, while image comparison based on convolutional neural networks (CNN) for indoor registration. In this research, the application of these two technologies and their integration have been studied and tested on site on a real Facility Management use case related to a university campus. The proposed approach has shown very promising results in displaying BIM elements of the electrical system seamlessly superimposed through AR to their physical counterparts in mixed indoor-outdoor environments.*

KEYWORDS: *Augmented Reality, Seamless Registration, Feature Matching, Pose Estimation, Real-Time Kinematic, Facility Management.*

1. INTRODUCTION

The AECO industry, although commonly recognized as one of the least digitized industries, is increasingly moving towards embracing more and more computer-based technologies to provide better performance in various stages of buildings lifecycle (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. Ever since the facilities started to become more complex, the day-to-day tasks have also become more difficult. 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 tools for facilities' O&M tasks and can provide significant advantages. AR impacts the mobile computing industry by radically changing the type of interaction between humans and computers. In fact, such technology creates direct, automatic, and practicable connections between the physical world and digital information by providing a simple and immediate user interface to a digitally enhanced physical world. Since AR allows virtual objects to be simultaneously superimposed on the real world, it helps to locate and view the occluded facilities and equipment and provides maintenance guiding instruction for the field workers (Salman & Ahmad, 2023). Through a hand-held device (HHD) or head mount device (HMD), it augments the real world by making implicit information apparent to the user when required. Many journal publications demonstrate that AR technology can be applied to various domains in AECO industry, especially in the O&M phase of the project lifecycle (Baek et al., 2019; Jurado et al., 2021; Naticchia et al., 2021; Vaccarini et al., 2022). For example, current maintenance practices are characterized by scattered and disoriented facility information that the maintenance staff must fetch through specifications, maintenance reports, and checklists. In fact, 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, maintenance staff 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 poses a number of demanding technological requirements for its implementation (Costanza et al., 2009). One challenge is related to display technology, which has registered remarkable breakthroughs in the last decades. Precise position tracking constitutes another significant challenge. In order to give the illusion that virtual objects

are located at fixed physical positions or attached to physical items, the system must know the position of relevant physical objects relative to the display system. Since the earliest days, spatial registration has been considered as one of the most important technical aspects of an AR systems and is considered 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).

The spatial registration methods are generally classified into two categories: “marker-based” and “marker-less” methods. The first one is considered the most widely used spatial registration method. Markers can be 2D images with visual features or natural 3D objects in the real environment (Cheng et al., 2020). The high use of this method may be returned to the simplicity, efficiency, and convenience 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 for calculating the relationship of relative positions. A similar approach can be applied with invisible markers, such as infrared and RFID ones (El Barhoumi et al., 2022). With marker-less approaches, instead of tracking features of markers, localization technologies are used to control the relative position between the real environment and virtual objects. GNSS is the most popular marker-less localization technology due to its suitability for use in a large open area such as a construction site and the ease of its signal receiving by common mobile devices (Cheng et al., 2020). According to the official US Government information on GNSS, the user range error (URE) for civil commitments cannot reach lower than 0.8 m (Cheng et al., 2020). The localization accuracy of GNSS is much worse in indoor environments, given that the buildings block the GNSS signal. The low accuracy of GNSS is not suitable for activities that require high accuracy or that mainly occur indoors. Compared with GNSS, some other marker-less localization technologies, such as Wi-Fi, Ultra-wideband (UWB), Inertial Measurement Unit (IMU), and Simultaneous Localization and Mapping (SLAM) can provide higher accuracy and can be applied to indoor activities. Another marker-less category is represented by vision-based methods using natural features for registration purposes (El Barhoumi et al., 2022). Each of these methods for spatial registration has its limitations in either accuracy or practicality. To promote the application of AR in the FM, which typically involves both indoor and outdoor environments, an advanced localization method that can provide an accurate and seamless registration in heterogeneous scenarios is needed. In order to cover this gap, a marker-less localization system for seamless indoor/outdoor AR registration has been developed by defining a cloud platform that hosts an indoor registration engine, an outdoor registration engine, plus a switch engine that manages the priority between the two. The developed system, tested on site on a real FM use case related to a university campus, has shown very promising results. 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. In Section 4, experiments design and execution on a FM use case are presented. Finally, Section 5 is devoted to results discussion and conclusions.

2. LITERATURE REVIEW

In this section, a literature review concerning existing AR registration methodologies applied to both indoor (Section 2.1) and outdoor (Section 2.2) environments is reported. Understanding strength and eventual gaps of approaches proposed by past studies and commercially available solutions paved the way to the definition of the indoor/outdoor seamless registration system proposed by this study.

2.1 Indoor AR registration

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 visual markers distinctive in the scene (Lee & Akin, 2011; Park et al., 2013). Even though artificial markers are advantageous in terms of robustness in detection, they should be installed all over the facility before on-site activities, such as the actual FM, occur. In addition, visual markers can trigger aesthetic issues because of their distinctive appearance. Alternative solutions are represented by invisible markers, such as infrared (Kuo et al., 2013) and RFID (Carbonari et al., 2022; Naticchia et al., 2021), and natural markers (Koch et al., 2014) that do not aesthetically change the scene. However, even though invisible markers do not have aesthetic issues, they should be pre-installed. Natural markers, instead, have the limitation of depending on signs, including exit signs, fire extinguisher signs, and textual information signs. If the scene does not have such designated signs, the localization can be restricted (Baek et al., 2019). Commercial AR libraries, such as Vuforia (PTC Products, 2023), ARcore (Google LLC, 2023), and World Locking

Tools (WLTs) (Microsoft, 2023) have been tested in indoor environments (Ashour et al., 2022; El Barhoumi et al., 2022). Comparative tests in indoor environments of Vuforia Image Target and WLTs showed better results of the second ones (Teruggi & Fassi, 2022).

Among marker-less methods, GNSS-based AR systems have been largely studied (Kim et al., 2013). However, they are considered inappropriate for indoor applications because of their low accuracy (Chen et al., 2019). Therefore, many studies have employed the 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 be still improved. Another marker-less methodology for indoor localization is based on image comparison. Image-based localization is classically tackled by estimating a camera pose from correspondences established between sparse local features (Ethan Rublee et al., 2011) and a 3D Structure-from-Motion (SfM) (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). This approach, although may lose accuracy whenever there is lack of context or repetitive elements, has shown great potential to develop indoor AR registration apps with applicability for non-expert users (Baek et al., 2019).

2.2 Outdoor AR registration

AR registration in outdoor environments presents distinct challenges compared to those encountered in indoor environments. Tracking and alignment approaches such as SLAM enable the placement of virtual objects relative to a local reference frame. However, the reliability of such tracking approaches in open environments is compromised due to (i) expanded spatial dimensions, (ii) the absence of readily available reference points, (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 Global Navigation Satellite System (GNSS) such as the Real-Time Kinematics (RTK), have been 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 solutions available that ensure the use of AR apps in outdoor environments either without relying on some kind of additional infrastructures (e.g., markers, QRcode, RFID, beacons, etc.) or without the need of manual/semi-manual alignment procedures, with some exceptions. For instance, Trimble Site Vision (Trimble Inc., 2023) makes use of the built-in GNSS receiver to achieve 1 centimeter of horizontal accuracy under RTK coverage. Similarly, Engineering-grade AR for AEC (vGIS Inc., 2023) developed by vGIS achieves the same centimeter-level 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. However, relying on GNSS technology only means that the system cannot cope with urban-canyon scenarios and indoor environment. Due to these limitations and the persistently high costs, these solutions have not yet experienced widespread adoption in the construction industry. Delving deeper into this last notion, it is noteworthy that the individual components of RTK receivers are economically affordable, thereby fostering the proliferation of applications of this technology (Hansen et al., 2021).

2.3 Research questions

As demonstrated by the literature review reported in Sections 2.1 and 2.2, 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 markers or the existence of signs to be used as natural markers (Baek et al., 2019). Among marker-less approaches, GNSS-based solutions are inappropriate for indoor applications because of the weakness of GNSS signals (Chen et al., 2019), whereas Wi-Fi-based approaches lose localization accuracy in case of multiple mobile devices (Salman & Ahmad, 2023). On the other hand, image-

based systems, although may lose accuracy whenever there is lack of context or repetitive elements, show great potential to develop indoor AR registration applications with applicability for non-expert users (Baek et al., 2019). With reference to outdoor AR registration approaches, limitations of marker-less GNSS-based approaches must be considered. First of all, the reduced reliability of GNSS in urban-canyon scenarios (e.g., proximity to urban elements, such as buildings, roofs, trees, and so on) limits possibilities of applications (Cheng et al., 2020; Ling et al., 2019). In addition, GNSS-based solutions are currently expensive (especially for high-precision RTK GNSS systems). Despite this, since single RTK receivers' 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 to 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 managing the built environment, it would be important that AR applications can work seamlessly even during changes in environment. In addition, in order to ensure a wider applicability of AR, preliminary set up procedures for registration should be as simple as possible. In order to cover these gaps, this study aims to answer the following research questions:

RQ1 What system architecture would ensure a seamless AR indoor/outdoor registration?

RQ2 What technical solutions would make AR registration “plug-and-play” for wider applicability even among non-expert users?

3. METHODOLOGY

3.1 System architecture

In order to answer the research questions (Section 2.3), the system architecture, reported in Fig. 1, has been defined. The proposed architecture is built on top of a BIM Cloud Platform which hosts the following four elements, each playing a crucial role in the overall functionality: (i) the Common Data Environment (CDE), (ii) the Outdoor AR Registration Engine, (iii) the Indoor AR Registration Engine, and (iv) the Switch Engine (Fig. 1). The BIM Cloud Platform serves as a centralized resource and processes hub that facilitates data processing, storage, and distribution. An important characteristic of the platform is its ability to host, localize, and align BIM models, images, and point clouds within a geospatial context. This geolocation feature enables the precise mapping of virtual assets and features to their corresponding real-world locations, facilitating the integration between the virtual and physical realms. One of the key responsibilities of the BIM Cloud Platform is therefore to manage the alignment processes. Particularly, the positioning of images within the platform is achieved by referencing the absolute world coordinates of the acquisition point, along with the accurate rotations. 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”.

The CDE is responsible for structured (e.g., .ifc files) and unstructured (e.g., images) data storage. It facilitates accessibility to AR applications through dedicated clients. In this work, the CDE of the DICEA Department of the Università Politecnica delle Marche has been used. At its core, a graph database provides a resilient backbone, offering efficient storage, retrieval, and traversal of interconnected data elements. The next integral components are the two distinct registration engines, specialized for outdoor and indoor environments, respectively. The Outdoor AR Registration Engine, which relies on the combination of RTK GNSS and IMU systems, 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 Indoor AR Registration Engine is designed to excel in environments characterized by restricted access to GNSS signals, leveraging features like point clouds and aligned images to achieve accurate positioning. To this purpose, convolutional neural networks (CNN) (Sarlin, Cadena, et al., 2018; Sarlin, Debraine, et al., 2018) that simultaneously predict local features and global descriptors have been applied for accurate 6-DoF localization. Finally, the pivotal feature of this system lies in the Switch Engine, which effectively serves as an integrator between the outdoor and indoor registration engines. It is a rule-based engine that assesses the availability of either GNSS signals, or features, or both, and dynamically switches between the two registration approaches to maintain a consistent and uninterrupted AR experience. By synergistically combining the aforementioned four elements within the BIM Cloud Platform, the system architecture (Fig. 1) delivers a robust and marker-less 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 head-mounted and hand-held AR devices, the focus from this point forward will be specifically on the usage of Microsoft's AR tool, HoloLens2. To this end, a novel addition to the tool is introduced, developed,

and physically realized, enabling a robust connection between the HoloLens2 device and the RTK receiver for precise calibration between the systems. The presented system architecture has been implemented in an AR application for HoloLens2 developed using the C# programming language and the serious game engine Unity3D.

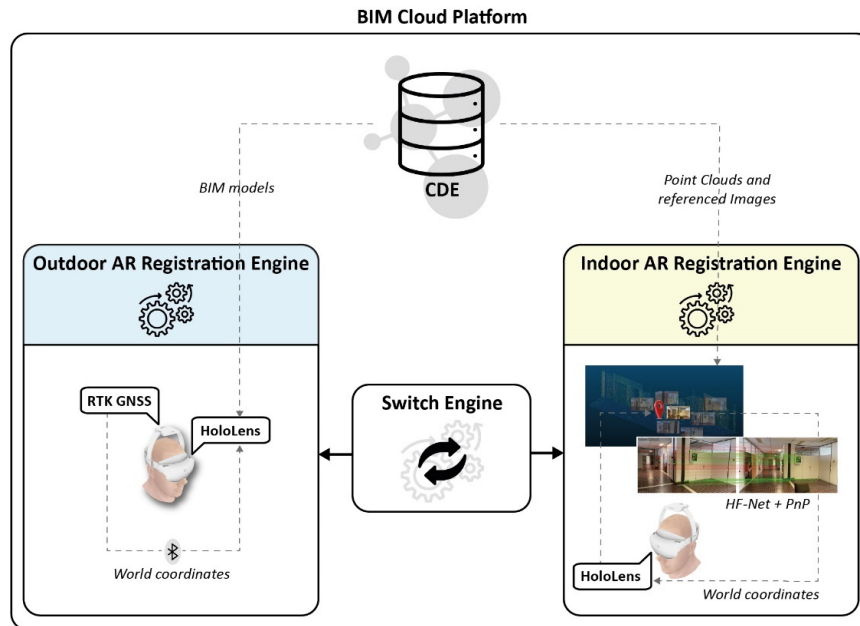


Fig. 1: Architecture of the proposed system for seamless outdoor and indoor AR registration.

3.1.1 Outdoor AR Registration Engine

Following the works made by Hansen et al. (2021) and Ling et al. (2019), the Outdoor AR Registration Engine proposed in this paper relies on the combination of a RTK GNSS tracking system and the HoloLens2' built-in inertial tracking system. By aligning the local frame reference of the HoloLens2 device to global coordinates exploiting the RTK measurements and using the HoloLens2 capability to localize itself in the environment through a real time mapping service, a geographical SLAM algorithm can be developed in order to have an absolute 6-DoF localization of the AR device and therefore aligned virtual objects. Some considerations must be made:

- a general 3D object in the BIM 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 ($^{\circ}$), Longitude ($^{\circ}$), Altitude (m), and Azimuth ($^{\circ}$). Object's coordinates refer to the WGS-84 standard;
- the coordinates retrieved from the RTK system are based on the WGS-84 standard;
- the RTK receiver has been developed in-house with contained costs in the perspective of widespread adoption;
- the RTK receiver and the HoloLens2 must be solidly connected to each other;
- the RTK system is reliable as long as the receiver is within 10 km of the RTK base station antenna;
- the HoloLens2' local coordinate frame originates at the point where the AR application is turned on.

The problem of placing world-referenced 3D BIM objects into the HoloLens2 local frame can be achieved by fulfilling the following steps (Fig. 2): (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. The resulting Outdoor AR Registration Engine's process (Fig. 2) is automatically executed (i.e., answer to RQ2). Hence, the outdoor registration process, not requiring any particular action from the user, can find applicability even among non-expert users.

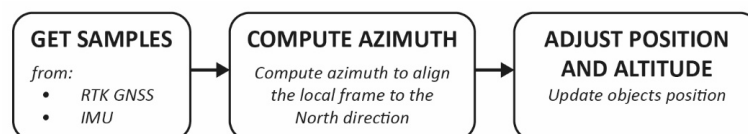


Fig. 2: A schematization of the outdoor AR registration engine's processes.

The first part of the outdoor engine involves the initialization phase of the HoloLens2 position. It includes

acquiring initial samples from both the RTK and IMU systems. Once initialized, the RTK system provides absolute 3D coordinates of the body frame (Latitude φ_1 , Longitude λ_1 , and Altitude h_1), while the IMU provides for local 3D coordinates (x, y, z) and rotations of the body frame. The local equirectangular projection must have y axis directed toward the North. However, at the beginning, the HoloLens2' local frame has an arbitrary unknown orientation with respect to the North. To solve this, when moving between two positions, the performed segment has an orientation β' with respect to the y' axis but the same movement forms a bearing angle β with respect to North as shown in Fig. 3.

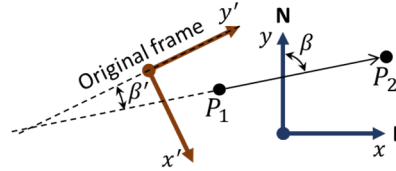


Fig. 3: Bearing angles with respect to North direction.

Given latitude and longitude of the start point (φ_1, λ_1) and latitude and longitude for the end point 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 follow:

$$\beta = \text{atan2} \left(\frac{\sin(\lambda_2 - \lambda_1) \cos(\varphi_2)}{\cos(\varphi_1) \sin(\varphi_2) - \sin(\varphi_1) \cos(\varphi_2) \cos(\lambda_2 - \lambda_1)} \right) \quad (1)$$

The geographical position of its reference frame (φ_0, λ_0) can be computed by inverting the reverse projection:

$$\lambda_0 = \lambda' - \frac{x}{R \cos \varphi_1} \quad (2)$$

$$\varphi_0 = \varphi' - \frac{y}{R} \quad (3)$$

assuming that $\varphi_0 = \varphi_1$ and where (φ', λ') are the GNSS coordinates and R is the radius of the globe. When the local frame position (φ_0, λ_0) and the object's geographic coordinates (φ'', λ'') are known, the corresponding local planar coordinates can be computed by the forward projection:

$$x = R(\lambda'' - \lambda_0) \cos \varphi_1 \quad (4)$$

$$y = R(\varphi'' - \varphi_0) \quad (5)$$

When the observer moves too far from the origin, the local reference should be translated in the new position and the 3D objects must be positioned with respect to the new reference system by re-applying the previous equations. To track the true value of the observer altitude, each time a GNSS measure 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:

$$z = h'' - h_0 + z_0 \quad (6)$$

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 in 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 Indoor AR Registration Engine

The AR registration in indoor environments, which contrarily to the outdoor ones are affected by restricted GNSS signals, required a dedicated solution different from the one presented in Section 3.1.1. For this reason, an Indoor AR Registration Engine, based on image comparison with survey data of the analyzed environment, has been developed (Fig. 4). A preliminary on-site survey with a camera and LiDAR scanner (e.g., GeoSLAM ZEB Horizon) must be carried out in order to collect point cloud and aligned photos of the analyzed environment. Alternatively, point clouds can be generated from a photos collection using the incremental Structure-from-Motion methodology implemented by the COLMAP library (Schönberger & Frahm, 2016). The basic idea is 6-DoF localizing the HoloLens2 by comparing a frame from its current view (i.e., query image) with images referenced

to the point cloud of the analyzed environment (i.e., reference images). Both reference images and the point cloud are stored in the BIM Cloud Platform. In this study, the Hierarchical Feature Network (HF-Net) technology has been implemented for image comparison (Sarlin, Cadena, et al., 2018; Sarlin, Debraine, et al., 2018). HF-Net consists of a CNN able to simultaneously detect feature keypoints and compute local and global descriptors for accurate 6-DoF localization. The “hierarchical” attribute refers to the HF-Net feature close to the humans’ attitude 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 HoloLens2 registration call, a coarse search, consisting in a global-descriptors matching between the query image and the reference images, is performed. Afterwards, a finer search based on a local-descriptors matching between the query image’s 2D keypoints and the point cloud’s 3D points covisible in reference images is executed. Finally, a 6 DoF pose estimation of the query image is carried out by solving the Perspective-n-Point (PnP) problem (Kneip et al., 2011). The estimated pose is thus the rotation and the translation vectors that allow transforming 3D points expressed in the world coordinate system into the camera coordinate system. These parameters enable the indoor AR registration of HoloLens2’s gaze.

The presented Indoor AR Registration Engine’s process (Fig. 4), once the survey of the interested area is completed, can be automatically executed (i.e., answer to RQ2). Hence, the indoor registration process, not requiring the user to do any particular actions, can find applicability even among non-expert users.



Fig. 4: A schematization of the Indoor AR Registration Engine’s processes.

3.1.3 The Switch Engine

As illustrated in Fig. 1, the Switch Engine 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 indoor/outdoor AR registration. The Switch Engine acts differently according to 3 possible scenarios: (i) RTK only (outdoor), (ii) RTK plus images/point clouds (outdoor), and (iii) images/point clouds only (indoor). In the first scenario, the Switch Engine identifies the availability of a stable and reliable RTK connection in outdoor scenarios. Consequently, the Switch Engine triggers the outdoor AR Registration Engine. The second scenario occurs when, in outdoor environments, images and point clouds are available simultaneously with RTK connection. The third scenario, instead, refers to indoor scenarios in which only images and point clouds are available. In both the second and third scenarios, as the Switch Engine identifies the presence of images and point clouds in the vicinity of the user’s real-world position (e.g., when approaching a previously surveyed building or asset), the system triggers the Indoor AR Registration Engine. At that point, the system entirely relies on images and point clouds for AR registration.

4. EXPERIMENTS

4.1 FM use case

The methodology proposed in this study has been tested on a FM use case based on a university campus, assumed as case study. Specifically, the study focused on the FM of the Digital Construction Capability Centre (DC3) Lab at the Università Politecnica delle Marche (Fig. 5 (a)). The DC3 Lab, which covers an area as large as 240 m², 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, has been considered. During this activity, the technician in charge of FM operations spends 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 externally to the building. These cables can be accessed through manholes located on the road in front of the building (Fig. 5 (a)). Once located all the elements interested by FM operations, the technician may need technical information about the electrical system. To this purpose, he/she needs to access the as-built BIM model (Fig. 5 (b)). The implementation of the proposed methodology (Section 3) enables the seamless AR registration in heterogeneous indoor/outdoor scenarios. Testing the proposed system to the presented use case, its

applicability in real-life situations will be assessed. As described more in detail in Section 4.2, the experimentation primarily revolves around the utilization of an AR application implemented on the HoloLens2 device.

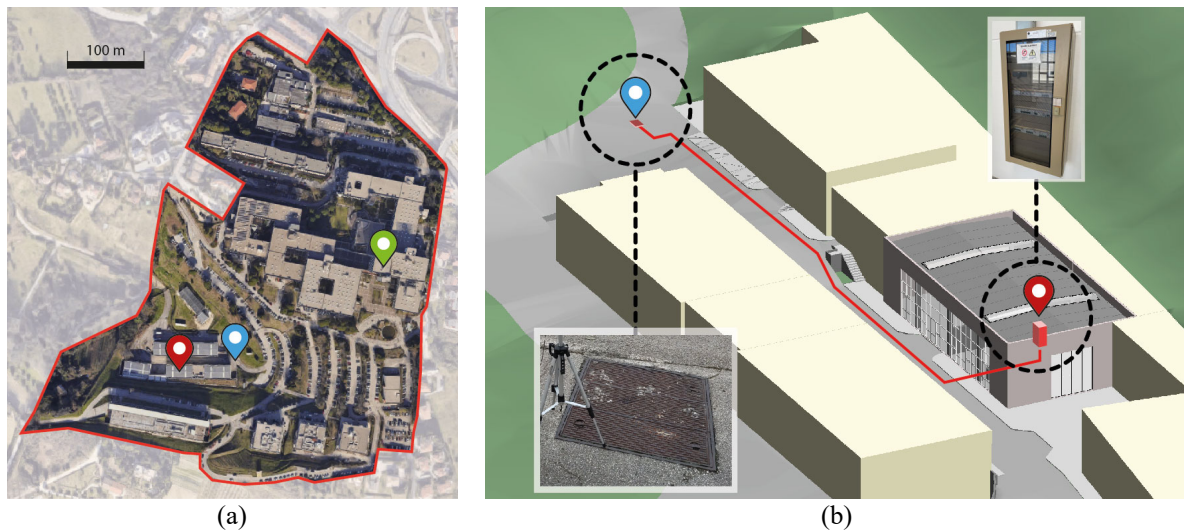


Fig. 5: (a) Aerial view of the university campus, assumed as case study, identifying the positions of the DC3 Lab (i.e., red placemark), the manhole cover (i.e., blue placemark), and the RTK base antenna (i.e., green placemark); (b) view of the BIM model of the DC3 Lab identifying the positions of the indoor electrical panel (i.e., red placemark) and the outdoor manhole cover (i.e., blue placemark).

4.2 Experiments design and execution

The developed system has been tested on the selected use case (Section 4.1) following the steps summarized in this section. First of all, a preliminary set-up phase consisting in collecting and initializing input data has been executed. This phase must be executed only once since related input and settings are maintained. It includes:

1. collecting point clouds and aligned photos by carrying out a survey of the DC3 Lab. In this study, the survey has been carried out by using GeoSLAM Backpack Vision (Fig. 6 (a)), which collects simultaneously both point clouds and aligned photos with a single scanning (Fig. 6 (b)). Alternatively, the survey can be carried out by collecting photos (e.g., with a smartphone) and then generating a point cloud through the Structure-from-Motion methodology implemented by the COLMAP library (Schönberger & Frahm, 2016);
2. collecting the BIM model of the DC3 Lab (Fig. 5 (b));
3. uploading the BIM model, point cloud, and images, related to the selected use case, on the BIM cloud platform. It must be noted that point cloud and images result aligned directly from the survey. The BIM model must be aligned to the previous dataset by selected reference points; it must be noted that this alignment is executed only once since it is maintained.

Once the previous preliminary steps 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 HoloLens2 has been used (Fig. 6 (c)). The AR application, based on the system architecture reported in Fig. 1, has been developed to support the technician inspecting the electrical system distributed in a heterogeneous indoor/outdoor scenario. The main contribution of the proposed system is the marker-less AR registration for displaying BIM models seamlessly superimposed to the whole inspected environment. The following steps have been executed on-site:

4. having on the HoloLens2 with installed RTK receiver (Fig. 6 (c)) and launching the AR application;
5. moving around the campus of the Faculty of Engineering at Università Politecnica delle Marche. During this preliminary step outside the DC3 Lab, the Outdoor AR Registration Engine is triggered in order to localize the user and drive him/her to the DC3 Lab;
6. heading to the internal electrical infrastructure to inspect the electrical panel located inside the DC3 Lab (Fig. 5 (a)). As the user moves from the outdoor to the indoor, the GNSS coverage decreases and the collected dataset (i.e., point clouds and aligned images) is found in the surrounding of the user position. Hence, the system seamlessly switches to the Indoor AR Registration Engine. This transition occurs without interruption as the system switches between registration modes through the Switch Engine's algorithm;
7. inspecting the internal electrical infrastructure, specifically focusing on the indoor electrical panel. During the

indoor phase, the AR application relies on the Indoor AR Registration Engine. It superimposes the digital model of the electrical panel on the real asset to let the facility managers have all the required information from the BIM model for the inspection (Fig. 7 (a));

8. heading to the external electrical infrastructure to inspect the cablings associated to the internal electrical panel. As the user moves from the indoor to the outdoor, the GNSS coverage rises again and the system seamlessly switches to the Outdoor AR Registration Engine. This transition occurs without interruption as the system switches between registration modes through the Switch Engine's algorithm;
9. inspecting the external electrical infrastructure, specifically focusing on the manhole covers located on the street facing the building (Fig. 5 (b)). During the outdoor phase, the AR application relies on the Outdoor AR Registration Engine, leveraging geospatial data retrieved from the RTK receiver to overlay BIM data on the real asset (Fig. 7 (b)).

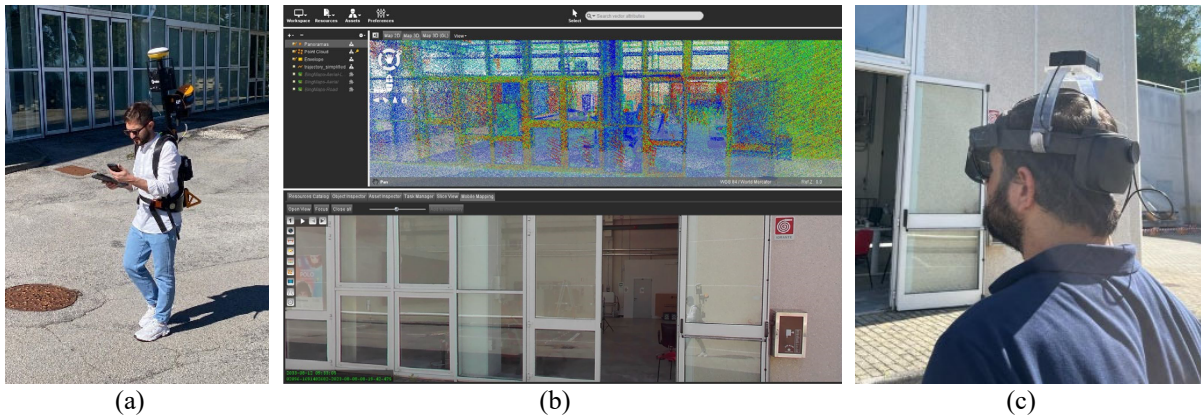


Fig. 6: (a) Survey phase of the DC3 Lab by using GeoSLAM Backpack Vision for collecting (b) point clouds and aligned photos; (c) inspection phase with the Microsoft HoloLens2 and the RTK receiver integrated by a 3D printed add-on.

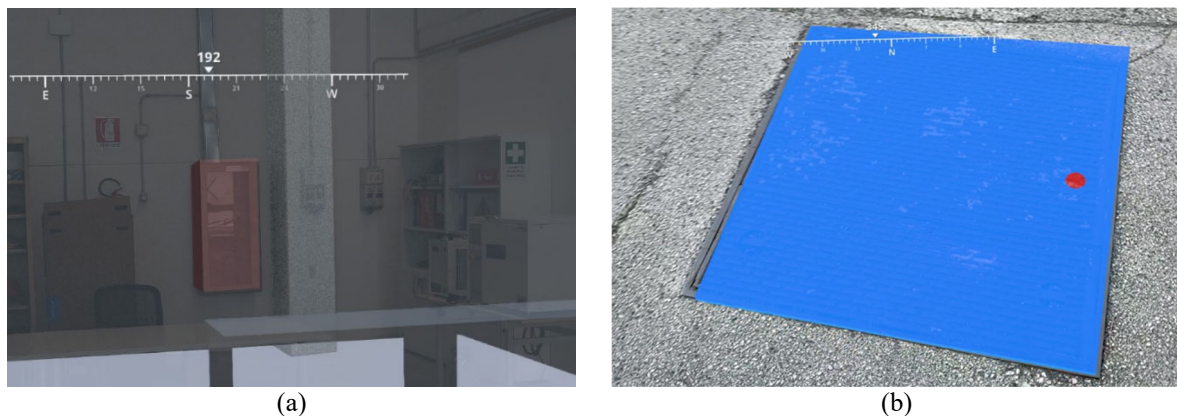


Fig. 7: Visualization of the aligned holograms of (a) the indoor electrical panel and (b) the outdoor manhole cover through the AR application deployed on HoloLens2.

5. DISCUSSION AND CONCLUSIONS

This paper addresses the open issue of AR registration in mixed scenarios considering that indoor-outdoor interactions are crucial for built environment management. An extended literature review has found limitations of existing indoor and outdoor AR registration approaches, drawing the conclusion that an all-in-one solution conceived for heterogeneous scenarios do not exist yet. Hence, this work focuses on defining and implementing a system architecture for seamless AR registration even during changes in environment (i.e., RQ1). In doing this, technical solutions that would make AR registration applicable even among non-expert users have been considered (i.e., RQ2). In order to answer RQ1, a system architecture (Section 3.1), which delivers a robust and marker-less AR registration system for both indoor and outdoor scenarios, has been defined and implemented. The resulting system has been tested on-site on a FM use case (Section 4.1). The proposed marker-less localization system has been put in place considering the FM of a university laboratory's electrical system. It has been selected because in-charge technicians and facility managers are continuously asked to locate, inspect, and repair interrelated system elements distributed in a heterogeneous indoor/outdoor environment. Efficiency of such activities is expected to be

considerably improved by accessing BIM models directly superimposed to their physical counterparts. In this study, the head-mounted AR device Microsoft HoloLens2 has been used for 6DoF localization and seamless gaze registration in heterogeneous environments. Indoor inspections of the electrical system have specifically focused on the laboratory electrical panel, whereas the outdoor ones on the related cablings that can be inspected through a dedicated manhole cover. The proposed system has shown promising results in registering the BIM model on-site. In fact, the holograms of the electrical panel and the manhole cover resulted superimposed to their physical counterpart even if they are located respectively in indoor and outdoor environments. This has confirmed one of the main contributions of the proposed system, that is a marker-less AR registration for displaying BIM models seamlessly superimposed to the whole inspected environment. In order to ensure a user-friendly AR registration process in mixed environments, hence providing an answer to RQ2, the proposed system requires only a preliminary set-up phase (i.e., steps 1-3 in Section 4.2) whose settings are then maintained. In fact, as confirmed by experiments, the AR experience is supported by the Switch Engine that manages priorities of the indoor and outdoor AR registration engines. In areas with GNSS coverage (i.e., generally outdoor environment), the Switch Engine delegates holograms superimpositions to the Outdoor AR Registration Engine. This is the logic that regulates the superimposition of the external manhole cover hologram to its physical counterpart. As the GNSS coverage is lost and collected datasets (i.e., point clouds and aligned images) are found in the surrounding of the user position (i.e., generally indoor environments), the Switch Engine delegates holograms superimpositions to the Indoor AR Registration Engine. This is the logic behind the superimposition of the internal electrical panel. As a result, since AR registration is automatically fulfilled in heterogeneous environments, the proposed solution can find applicability even among non-expert users.

Current limitations of the proposed methodology can be traced back to technologies adopted by system's engines. Accuracy loss may affect image comparison, adopted by the Indoor AR Registration Engine, in case of reference and/or query images with lack of context or repetitive elements. Follow-up studies may quantify such limitation. On the other hand, the Outdoor AR Registration Engine strongly relies on the availability of RTK coverage. Despite such technology is currently economically expensive, it is noteworthy that since the individual components of RTK systems are economically affordable, the proliferation of applications of this technology is highly expected. Further studies will be carried out in order to assess registration accuracy of both the Indoor and Outdoor AR Registration Engine. More tests must be carried out in order to optimize switching thresholds based on GNSS coverage and datasets (i.e., point clouds and aligned images) availability in the surrounding of the user position. Future developments will focus also on the definition of a graphical user interface for better managing 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 saved time for completing a task.

REFERENCES

- Ahmad, A., Claudio, P., Alizadeh Naeini, A., & Sohn, G. (2020). Wi-fi rssi 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., Kivrak, 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>
- 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](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>

Google LLC. (2023). *ARCore*. <https://developers.google.com/ar?hl=it>

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>

Jurado, D., Jurado, J. M., Ortega, L., & Feito, F. R. (2021). Geuinf: Real-time visualization of indoor facilities using mixed reality. *Sensors (Switzerland)*, 21(4), 1–21. <https://doi.org/10.3390/s21041123>

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](https://doi.org/10.1016/S1881-8366(12)80021-8)

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>

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., Snively, 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>
- 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. (2023). *World Locking Tools documentation*. <https://learn.microsoft.com/en-us/mixed-reality/world-locking-tools/>
- 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>
- 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>
- 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). *From Coarse to Fine: Robust Hierarchical Localization at Large Scale*. <http://arxiv.org/abs/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>
- Schönberger, J. L., & Frahm, J.-M. (2016). *Structure-from-Motion Revisited*. <https://github.com/colmap/colmap>.
- 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>
- Trimble Inc. (2023). *Site Vision*. <https://sitevision.trimble.com/>
- Vaccarini, M., Carbonari, A., Spegni, F., & Giretti, A. (2022). Enhancing BIM through Mixed Reality for Facility Management. In *Building Information Modeling - A Sustainable Approach and Emerging Technologies [Working Title]*. IntechOpen. <https://doi.org/10.5772/intechopen.106186>
- vGIS Inc. (2023). *Engineering-grade AR for AEC*. <https://www.vgis.io/augmented-reality-bim-gis-ar-aec-civil-construction-engineering-bentley-autodesk-esri/>
- 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>