

Article

Mixed Reality-Based Multi-Scenario Visualization and Control in Automated Terminals: A Middleware and Digital Twin Driven Approach

Yubo Wang ¹, Enyu Zhang ¹, Ang Yang ^{1,*} , Keshuang Du ¹ and Jing Gao ² 

¹ School of Maritime Economics and Management, Dalian Maritime University, Dalian 116026, China; yubo.wang@dlmu.edu.cn (Y.W.); zey87@dlmu.edu.cn (E.Z.); keshuang.du@argsen.com (K.D.)

² STEM Academic Unit, University of South Australia, Adelaide, SA 5095, Australia; jing.gao@unisa.edu.au

* Correspondence: ang.yang@dlmu.edu.cn

Abstract

This study presents a **Digital Twin–Mixed Reality (DT–MR)** framework for the immersive and interactive supervision of **automated container terminals (ACTs)**, addressing the fragmented data and limited situational awareness of conventional 2D monitoring systems. The framework employs a middleware-centric architecture that integrates heterogeneous subsystems—covering terminal operation, equipment control, and information management—through standardized industrial communication protocols. It ensures synchronized timestamps and delivers semantically aligned, low-latency data streams to a multi-scale Digital Twin developed in Unity. The twin applies level-of-detail modeling, spatial anchoring, and coordinate alignment (from **Industry Foundation Classes (IFCs)** to east–north–up (ENU) coordinates and Unity space) for accurate registration with physical assets, while a **Microsoft HoloLens 2** device provides an intuitive Mixed Reality interface that combines gaze, gesture, and voice commands with built-in safety interlocks for secure human–machine interaction. Quantitative performance benchmarks—latency ≤ 100 ms, status refresh ≤ 1 s, and throughput $\geq 10,000$ events/s—were met through targeted engineering and validated using representative scenarios of quay crane alignment and **automated guided vehicle (AGV)** rerouting, demonstrating improved anomaly detection, reduced decision latency, and enhanced operational resilience. The proposed DT–MR pipeline establishes a reproducible and extensible foundation for real-time, human-in-the-loop supervision across ports, airports, and other large-scale smart infrastructures.

Keywords: automated container terminal; digital twin; mixed reality; immersive visualization; real-time monitoring



Academic Editor: Nikos A. Salingeros

Received: 15 September 2025

Revised: 19 October 2025

Accepted: 20 October 2025

Published: 27 October 2025

Citation: Wang, Y.; Zhang, E.; Yang, A.; Du, K.; Gao, J. Mixed Reality-Based Multi-Scenario Visualization and Control in Automated Terminals: A Middleware and Digital Twin Driven Approach. *Buildings* **2025**, *15*, 3879. <https://doi.org/10.3390/buildings15213879>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Background and Significance

Smart building environments and large-scale facilities are increasingly exploring Digital Twin (DT) and Mixed Reality (MR) technologies to enhance operational monitoring and management [1,2]. Automated container terminals (ACTs), as a special class of complex built infrastructure, exemplify these challenges. An ACT is a wide-area cyber–physical system coordinating quay cranes, yard cranes, automated guided vehicles (AGVs), and control systems across expansive indoor–outdoor spaces [3]. This scale and heterogeneity make real-time supervision and decision-making highly non-trivial [4]. Conventional

supervisory control and data acquisition (SCADA) interfaces remain predominantly two-dimensional and equipment-centric, which limits the operator's ability to perceive spatial relationships and system-wide status. Knatz et al. (2022) [3] noted that traditional console displays in automated terminals lead to fragmented situational awareness and increased cognitive load, slowing down responses to operational disruptions. Similarly, Yang and Ergan (2014) [5] found that a lack of spatial context in current monitoring interfaces hinders operator accuracy and efficiency during facility management tasks.

These shortcomings underscore the need for more intuitive and integrated monitoring approaches in ACT operations, aligning with broader trends in smart infrastructure management [6]. More importantly, ACTs should also be recognized as specialized built environments that differ from generic industrial plants. They combine wide-area outdoor infrastructure, complex building-scale logistics, and BIM/IFC/CityGML integration requirements. In summary, ACTs represent an important yet particularly challenging built environment for which new human-in-the-loop technologies are needed to ensure transparent, safe, and efficient operations.

1.2. Challenges in Current Terminal Operations

Despite the rapid adoption of automation in container terminals, several fundamental problems continue to impede reliable and immersive supervision. Grounded in industry practice and recent studies, we identify five key challenges that motivate a DT–MR approach for ACTs:

- **Heterogeneous and siloed data environments:** Operational information across the Terminal Operating System (TOS), Equipment Control System (ECS), and other industrial subsystems remains fragmented, making it difficult to form a unified real-time view of terminal activities. Insufficient cross-system data fusion and time synchronization lead to gaps in situational understanding and slower decision-making [7,8].
- **DT modeling complexity:** Building a high-fidelity DT of a large, dynamic terminal is non-trivial. Excessive detail in the virtual model can impair runtime performance, whereas oversimplification compromises fidelity and usefulness [9,10]. Achieving an optimal balance between representational accuracy and computational efficiency remains an ongoing challenge for terminal-scale DT implementations [11].
- **Restricted situational awareness in conventional monitoring systems:** Traditional 2D dashboards and SCADA interfaces limit operators' ability to perceive multi-equipment coordination and process-level dynamics in real-time, motivating the need for spatially aligned visualization and interaction approaches [12]. In particular, conventional systems obscure interdependencies among multiple devices (e.g., AGVs and yard cranes), limiting awareness of coordination bottlenecks and reducing the ability to detect emergent conflicts.
- **User interface and interaction constraints:** Fixed control-room human–machine interfaces and limited input modalities do not translate well to large 3D scenes. In an MR environment, operators can easily experience information overload without multimodal interaction techniques and spatial navigation aids to rapidly reorient and target relevant assets [13,14]. The narrow field of view and positioning drift of head-mounted devices under varying conditions further challenge stable MR usage in expansive outdoor terminals. This directly connects to the operator's ability to navigate a wide-area facility efficiently—a key concern for ensuring practical adoption in real operations.
- **Performance and latency requirements:** A DT–MR system must remain highly responsive under live operational loads. End-to-end latency, frame rate stability, and data throughput scalability are critical for maintaining immersion and preventing

user discomfort. For instance, collision avoidance and equipment control in ACTs demand millisecond-level responsiveness [15]. Ensuring low-latency data updates and synchronization (on the order of only a few dozen milliseconds) across hundreds of devices is a major technical hurdle [16,17].

Beyond spatial tracking and interface design, visual comfort during prolonged use represents a critical human factor challenge in MR deployments. Conventional head-mounted displays often induce vergence–accommodation conflict (VAC), where mismatched cues between ocular convergence and focal adjustment cause visual fatigue, dizziness, and reduced operational endurance. To mitigate these effects, our framework adopts a fixed-focal-plane configuration (optimized at approximately 2.0 m), which maintains consistent visual depth across all holographic content. This design choice significantly reduces cognitive and physiological strain, enabling sustainable supervision sessions—an essential requirement for continuous port operations.

Collectively, these challenges—including visual ergonomics—define the problem space addressed in this study and outline the practical requirements for next-generation supervision systems in automated container terminals.

1.3. Emerging Solutions: Digital Twin and Mixed Reality

Recent advances suggest that DT and MR technologies could jointly tackle the above challenges. DT provides data-driven virtual representations of physical assets and processes, continuously synchronized with sensor and operational data for analysis and control [9]. In ACT contexts, DT approaches have demonstrated tangible benefits for optimizing operations. For example, Gao et al. (2023) [18] proposed a DT-based scheduling method that significantly reduced terminal energy consumption through dynamic equipment dispatching. Likewise, Li et al. (2021) [8] introduced an AdaBoost-driven DT framework for terminal operations, showing improved coordination between heterogeneous resources under dynamic conditions. These studies establish DT as a robust foundation for modeling complex port systems and enabling data-informed optimization. In the building domain, integrating DT models with Building Information Modeling (BIM) has also supported lifecycle monitoring and collaborative decision-making, illustrating DT's broader applicability to smart infrastructure management [19].

Complementing DT on the human interface side, MR offers an intuitive and immersive way to present live operational data. MR overlays digital information onto the physical environment via head-mounted displays, enabling in situ visualization and interaction with system data in a spatially registered manner [20]. Empirical evidence shows that such MR interfaces can enhance situational awareness and reduce cognitive burden in field operations by presenting context-rich, 3D information at the point of need. For instance, Choi et al. (2022) [21] developed an integrated DT–MR platform for human–robot collaboration, and demonstrated that holographic overlays improved operator awareness and reduced cognitive load during safety-critical tasks. Similarly, Teizer et al. (2013) [22] found that an MR-based training system for construction workers enhanced users' spatial understanding and helped reduce on-site safety risks, a result echoed in recent construction safety studies [2].

Driven by these parallel developments, researchers have started exploring the convergence of DT and MR in unified frameworks. Lee et al. (2024) [16] proposed a Mixed Reality Virtual Device architecture to seamlessly bridge MR front-end devices with IoT/DT back-end systems, emphasizing interoperability and real-time data consistency. Shin et al. (2023) [23] developed a DT pipeline integrated with MR visualization for interactive simulation monitoring, allowing users to intuitively observe and steer scientific workflows.

Nevertheless, these solutions focus on either optimization or visualization in isolation, and lack an integrated framework addressing data fusion, modeling fidelity, immersive interaction, and runtime safety together in large-scale ACT contexts. This gap provides the motivation for the present research, which aims to converge DT and MR into a coherent system for large-scale terminal operation monitoring and control. This convergence of DT and MR technologies provides the conceptual motivation for the present study. Section 2 further reviews related research efforts in greater depth, highlighting methodological limitations and open challenges that this work aims to address.

In summary, while DT and MR technologies have been extensively explored independently, their integrated implementation in automated container terminals remains limited. Existing approaches rarely establish a unified middleware architecture that ensures real-time data fusion, low-latency visualization, and safe human-in-the-loop control. To bridge this gap, this study develops a middleware-driven DT–MR framework that integrates heterogeneous terminal subsystems, optimizes model fidelity for Mixed Reality visualization, and quantitatively evaluates system responsiveness through controlled scenario-based testing. This design aims to demonstrate a scalable, reliable, and reproducible approach for coupling Digital Twin and Mixed Reality in complex industrial supervision environments.

1.4. Research Scope and Contributions

Based on the above objectives, the scope and key contributions of this study are as follows:

- **Unified data integration with real-time synchronization:** We design a middleware-based integration framework that unifies heterogeneous operational data from terminal subsystems into a coherent DT model, ensuring real-time synchronization and consistency across equipment and control layers.
- **Immersive DT-MR multi-scenario interface:** We implement an MR interface that connects the DT with real terminal environments, enabling intuitive, spatially aligned visualization and human–machine interaction across representative operational scenarios. This immersive interface enhances situational awareness and decision efficiency compared with traditional 2D control systems.
- **Operational safety and control mechanisms:** To ensure safe human-in-the-loop control, we incorporate multiple runtime safety interlocks and control policies. These include command permission hierarchies and lockout mechanisms, audit logging of user actions, and a one-touch emergency stop/fallback function that immediately freezes or rolls back MR-issued commands. We establish a secure feedback loop between the MR interface and the conventional control system, allowing operators to seamlessly revert to the traditional control console or halt equipment in case of any detected anomalies. This safety-centric design minimizes the risk of erroneous or unauthorized operations when introducing immersive controls.
- **Systematic evaluation of performance and human factors:** We conduct a rigorous evaluation in three stages: (1) performance benchmarking, which confirms that the integrated system meets real-time performance requirements (maintaining, for example, <100 ms latency, 60 fps rendering, and processing of over 10,000 events/s); (2) qualitative usability assessment, which conceptually contrasts the MR interface with a baseline 2D SCADA display to illustrate potential advantages in situational awareness and interaction efficiency; and (3) scenario-based case analysis, which demonstrates the system’s effectiveness in handling realistic operational events. The findings suggest that the proposed DT-MR framework can enhance anomaly detection speed and operational awareness compared with traditional 2D interfaces, while empirical user evaluation remains a direction for future work.

- **Value for built environment research:** Beyond port operations, this study highlights the applicability of DT-MR methods to complex built environments such as port terminals, positioning them as part of the broader discourse on smart building and facility management.

Sections 2 and 3 further elaborate on the theoretical background and current technological limitations of automated terminal systems, forming the conceptual foundation for the proposed DT-MR framework that is technically detailed in Section 3.2.

2. Related Work

2.1. MR/AR in AECO Operations

MR and Augmented Reality (AR) technologies have seen growing adoption in architecture, engineering, construction, and operations (AECO) to improve on-site safety, training, and monitoring. Recent studies confirm that AR/MR are beneficial for construction safety training and hazard visualization, allowing workers to experience dangerous scenarios virtually and thereby improving hazard awareness [22,24]. MR is also identified as a key enabling technology in the construction industry's Industry 4.0 transformation, often integrated with Building Information Modeling (BIM) and Digital Twin systems to enhance operational awareness and decision-making [25,26].

State-of-the-art implementations in AECO range from AR head-mounted displays for on-site guidance to MR applications that superimpose BIM models onto real structures for maintenance and inspection. Guo et al. provide a comprehensive survey of AR/MR head-mounted devices under various environmental conditions, reflecting improvements in tracking accuracy and display capabilities [13]. These advancements enable more reliable anchoring of virtual content with real-world infrastructure—a critical requirement for construction sites and facilities. A notable example is the use of MR for facility operations: Hidayat et al. developed a DT system for smart building management that integrates blockchain for data integrity and uses MR to visualize building performance in situ [27].

Despite these advances, most AECO-focused MR applications remain confined to pilot stages or controlled environments. Challenges persist in anchoring virtual models across extensive, dynamic sites like ports, where GPS inaccuracies and lack of reference points hinder robust registration. Moreover, most MR research emphasizes static visualization or training; few tackle real-time sensor integration or bidirectional control. Consequently, current MR/AR implementations do not fully address immersive interaction with live systems in a large-scale port setting.

2.2. Digital Twin in Ports and Industrial Facilities

DT technology—virtual replicas of physical systems with real-time data linkage—has gained traction in smart ports and industrial facilities. A recent bibliometric review indicates that “smart port” research is rapidly expanding, with DT emerging as a core theme for improving operational efficiency and automation [28]. Gao et al. propose a DT-based approach to optimize energy consumption at automated container terminals, demonstrating efficiency gains through improved scheduling and reduced idle times [18]. Yang et al. present a comprehensive DT framework for Qingdao Port, which includes a geo-spatial 3D model linked with real-time operational data for monitoring and anomaly detection [9].

Beyond monitoring, DTs are being leveraged for decision support and optimization. Zhang et al. integrate reinforcement learning with a DT to optimize scheduling of quay cranes, automated guided vehicles (AGVs), and yard cranes, showing how virtual environments can safely train scheduling policies before deployment [29]. Such approaches highlight the potential of DTs in enabling closed-loop optimization.

However, existing port DTs remain siloed, focusing on specific applications such as energy optimization [18] or equipment scheduling, and largely presenting results through 2D dashboards. Few efforts integrate DTs with immersive MR interfaces for real-time operator interaction. Current solutions emphasize analytics and planning rather than human-in-the-loop operations, leaving a gap for large-scale immersive monitoring.

2.3. Three-Dimensional Visualization vs. Two-Dimensional Dashboards

A central consideration for operational monitoring systems is the visualization modality: immersive 3D versus traditional 2D dashboards. Conventional port management systems rely on 2D interfaces, which are efficient for abstract data but lack spatial context. Research shows that 3D visualization significantly improves situational awareness by aligning data with the physical environment [5,30]. For instance, Shin et al. demonstrate an MR visualization pipeline where simulation results are interactively rendered in 3D, enabling richer comprehension of complex data [23].

Nevertheless, 3D/MR visualization brings performance challenges. Complex models and real-time overlays can overwhelm rendering on wearable MR devices. Techniques such as level-of-detail (LOD) rendering are essential to ensure responsiveness across large-scale environments [31]. Moreover, without careful design, 3D views may clutter or occlude key information, whereas 2D dashboards provide concise clarity. Overall, while 3D MR visualization enhances contextual awareness, research has yet to evaluate its trade-offs in real industrial settings like ports.

2.4. Middleware and Data Governance in IIoT

Scalable DT–MR integration requires robust middleware and effective data governance. DT ingest heterogeneous sensor data, telemetry, and operational logs, making data harmonization a major challenge. Correia et al. highlight issues such as interoperability, standardization, and data quality management across DT implementations [32]. Lee et al. propose a Mixed Reality Virtual Device (MRVD) framework, which abstracts IoT devices as virtual entities to simplify MR–DT integration and support scalable interaction [16].

Data governance also involves ensuring security and traceability. Hidayat et al. demonstrate how blockchain-enhanced DTs can safeguard sensor data integrity and support trustworthy audits [27]. Despite these advances, many DT–MR middleware solutions remain bespoke prototypes with limited scalability. High-frequency data streams require optimized architectures to achieve millisecond-level latency, yet performance benchmarks are rarely discussed. Governance policies—defining access rights and retention—are often overlooked, despite their importance in multi-stakeholder environments such as ports.

2.5. Closed-Loop Control and Safety in DT–MR Systems

An important promise of DT–MR systems is enabling closed-loop control, where real-time data flows from physical systems to virtual models and back, ensuring both efficiency and safety. Xie et al. present a closed-loop human–cyber–physical system in mining, where AR interfaces provide operators with real-time alerts and allow them to issue control commands [33]. This improves response time and integrates human judgment into system control.

Similarly, Choi et al. designed an MR-based safety framework for human–robot collaboration, where workers visualize dynamic safety zones and the DT triggers emergency stops if thresholds are violated [21]. These implementations highlight the potential of MR for safety-critical applications.

Yet, most remain confined to controlled testbeds. Reliability in outdoor port environments is rarely studied, and few systems address graceful degradation, advisory versus

automatic control modes, or long-term human factors. Ensuring robustness and fail-safe mechanisms in DT–MR closed-loop systems remains an open challenge.

3. Requirements for ACT Visualization and Control

3.1. Evolution and Current Limitations of ACT Operational Systems

While previous sections outlined the evolution and research context, this section focuses specifically on the operational constraints that motivate the proposed middleware and visualization strategy.

Automated container terminals (ACTs) have evolved from mechanized and semi-automated facilities toward fully automated systems, where quay cranes, yard cranes, and automated guided vehicles (AGVs) are integrated into centralized control platforms. While this evolution has reduced manual involvement and improved throughput, it has simultaneously amplified the scale and complexity of operations. The increased heterogeneity of systems demands higher levels of coordination and real-time oversight.

In modern ACTs, operational domains such as ship-to-shore handling, yard storage, and container transport are managed through layered subsystems, including the Terminal Operating System (TOS), Equipment Control System (ECS), and Terminal Management System (TMS). Despite their functional roles, these systems frequently operate in silos, generating fragmented and non-uniform data streams. The absence of seamless data integration creates barriers to holistic decision-making and reduces transparency across the terminal.

Operators remain constrained by two-dimensional SCADA-based monitoring interfaces. While effective for presenting isolated metrics, these displays lack spatial depth and contextual richness, which are essential for wide-area terminal environments. They increase cognitive workload, limit situational awareness, and delay responses during disruptions. Navigating complex layouts via multiple windows and menus further compounds operator fatigue and slows interventions. Together, these limitations restrict the capacity to manage highly dynamic and interdependent equipment fleets. As illustrated in Figure 1, a fully automated container terminal integrates quay cranes (QCs), yard cranes (YCs), and automated guided vehicles (AGVs) under centralized supervision, with operators overseeing operations through Mixed Reality interfaces such as the HoloLens. The figure conceptually highlights how spatially distributed assets and information layers are unified into an interactive control ecosystem, underscoring the importance of immersive situational awareness for coordinated terminal management.

The persistent challenges of fragmented data, 2D visualization, rigid interaction, and performance constraints underscore the urgent need for an integrated, immersive, and low-latency DT–MR system. Such a system must bridge data silos, enhance spatial awareness, reduce operator workload, and provide resilient support for safety-critical decision-making.

3.2. Data and Functional Requirements for MR-Based Solutions

Building on the limitations identified in Section 3.1, five categories of requirements are defined for MR-based solutions to ensure their effectiveness in ACT environments:

The first requirement is real-time integration of heterogeneous data sources to eliminate fragmentation across operational subsystems. Information from the TOS, ECS, and IMS must be consolidated through a unified middleware layer that aligns protocols and semantics while supporting sub-second synchronization. Such integration ensures consistency between physical operations and their digital representation, thereby enabling supervisors to make informed decisions based on coherent and up-to-date information.

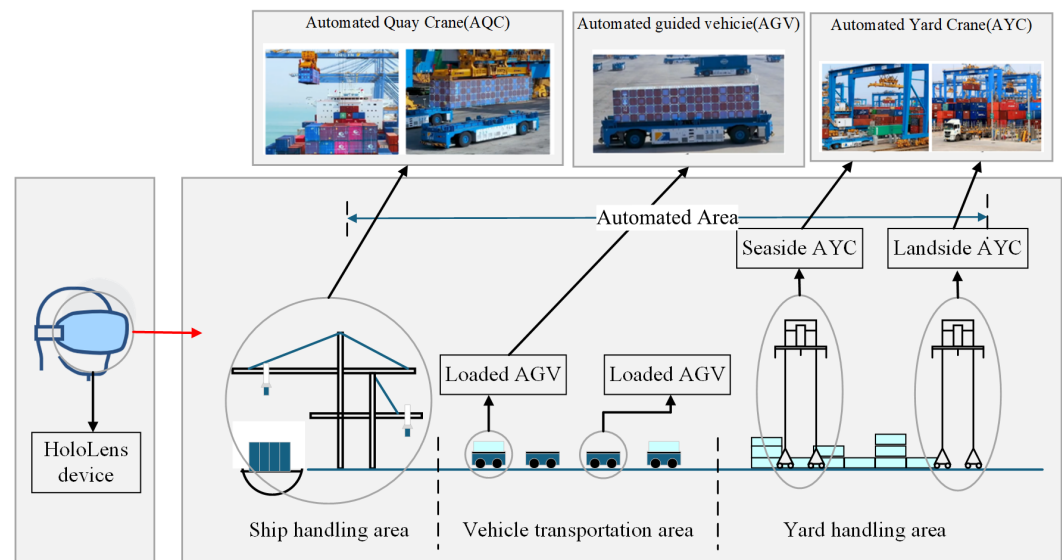


Figure 1. Operational modality of a fully automated container terminal highlighting the role of human operators in supervision and control.

A second requirement concerns visualization fidelity. Dynamic assets such as quay cranes and AGVs must be represented with high-fidelity three-dimensional models to capture operational details accurately. In contrast, static elements like container stacks can be rendered at lower levels of detail to reduce computational load. Precise spatial calibration between the digital and physical environments is essential to avoid distortions, ensuring that operators can rely on immersive views for wide-area monitoring.

Third, the system must strengthen situational awareness by embedding contextual information directly into the immersive environment. Holographic overlays, including equipment status panels, progress indicators, and visual alerts, should be integrated into the operator's field of view. By presenting such information in situ, the system allows supervisors to monitor concurrent workflows at a glance, improving anomaly detection and enhancing responsiveness to dynamic conditions.

The fourth requirement relates to interaction modalities. To overcome the rigidity of traditional 2D interfaces, the system should support multimodal interactions that combine gestures, gaze, and voice commands with intuitive navigation tools such as teleportation and dynamic scaling. These capabilities lower cognitive workload and enable operators to interact naturally with terminal-scale DT, thereby improving efficiency and reducing the risk of oversight during high-density operations.

Finally, robust system performance is indispensable for reliable deployment in automated terminals. The DT-MR framework must sustain end-to-end latencies below 100 ms, refresh operational states within one second, and process more than 10,000 events per second. Rendering quality must remain at or above 60 frames per second on head-mounted devices, while spatial registration errors should not exceed ± 5 cm. Meeting these quantitative benchmarks ensures the stability and usability of the immersive system under demanding, large-scale operational conditions. As illustrated in Figure 2, the proposed MR-augmented information architecture integrates the Terminal Operating System (TOS), Equipment Control System (ECS), and Terminal Management System (TMS) through a unified middleware layer. Each subsystem manages its functional domain—such as berth planning, vehicle scheduling, and business management—while the MR interface provides a real-time, immersive visualization layer connecting these traditionally siloed components. This architecture supports synchronized data exchange and human-machine interaction across operational, control, and strategic decision levels.

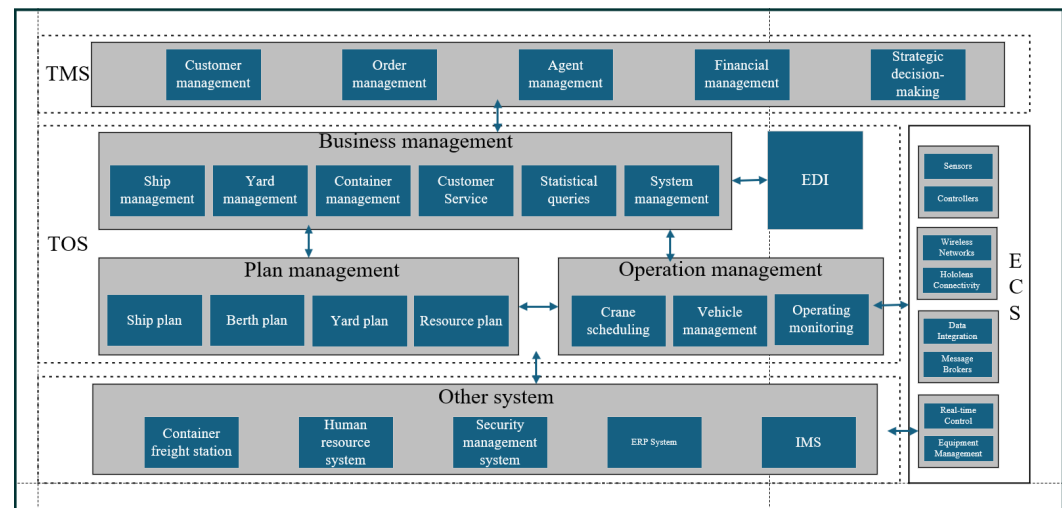


Figure 2. Architecture of the MR-augmented information system for ACTs.

In summary, these five requirements—data integration, visualization fidelity, situational awareness, interaction, and performance—collectively define the design objectives of the DT–MR framework. They establish both functional and performance baselines, guiding the architecture described in Section 4 and providing measurable benchmarks for subsequent validation.

4. Mixed Reality-Based Visualization Framework for Automated Container Terminals

4.1. Design Rationale and Overall Framework

The DT–MR system employs a layered architecture, as illustrated in Figure 3. This hierarchical structure consists of (1) a **Device Control and Acquisition Layer** for real-time data collection, (2) a **Task and Scheduling Layer** driven by TOS/ECS, (3) a **Maintenance and Management Layer** that integrates IMS and safety modules, and (4) a **Data Management and Twin Layer** dedicated to DT models. Such a modular layered design ensures that each subsystem can be independently optimized and scaled. Explicit interfaces between the layers enforce performance constraints—such as an end-to-end latency of ≤ 100 ms and a refresh rate of ≤ 1 s—while facilitating standardized data exchange. In practice, this allows each layer to be developed and maintained independently without impacting others. Developers can trace performance metrics—such as latency or throughput—through the pipeline to identify potential bottlenecks. As shown in Figure 3, the overall framework provides a structured foundation for system extensibility—such as incorporating new equipment or data sources—while guaranteeing the required responsiveness and reliability.

4.2. Digital Twin Modeling Pipeline

The DT models are constructed from heterogeneous data sources. We import detailed 3D geometry of cranes, vehicles, and yard infrastructure from design data (e.g., BIM/IFC models, laser scans, or CityGML maps) and synchronize dynamic state from sensors. To balance fidelity and performance, we employ a tiered level-of-detail (LOD) strategy. High-fidelity LOD500 meshes are used for dynamic equipment (quay cranes, yard cranes, AGVs) where precise kinematics (trajectories, hoist motion, collision geometry) matter. In contrast, static structures like container stacks, buildings, and roads use simplified LOD300–400 meshes, capturing overall layout with far fewer polygons. In deployment, these LODs can switch based on distance: for example, distant yard cranes may render a coarse LOD, while cranes within 50–100 m use a detailed model. Typical parameters might be, say, ≤ 50 k triangles for LOD400 stack models versus ~ 200 k for LOD500 crane models.

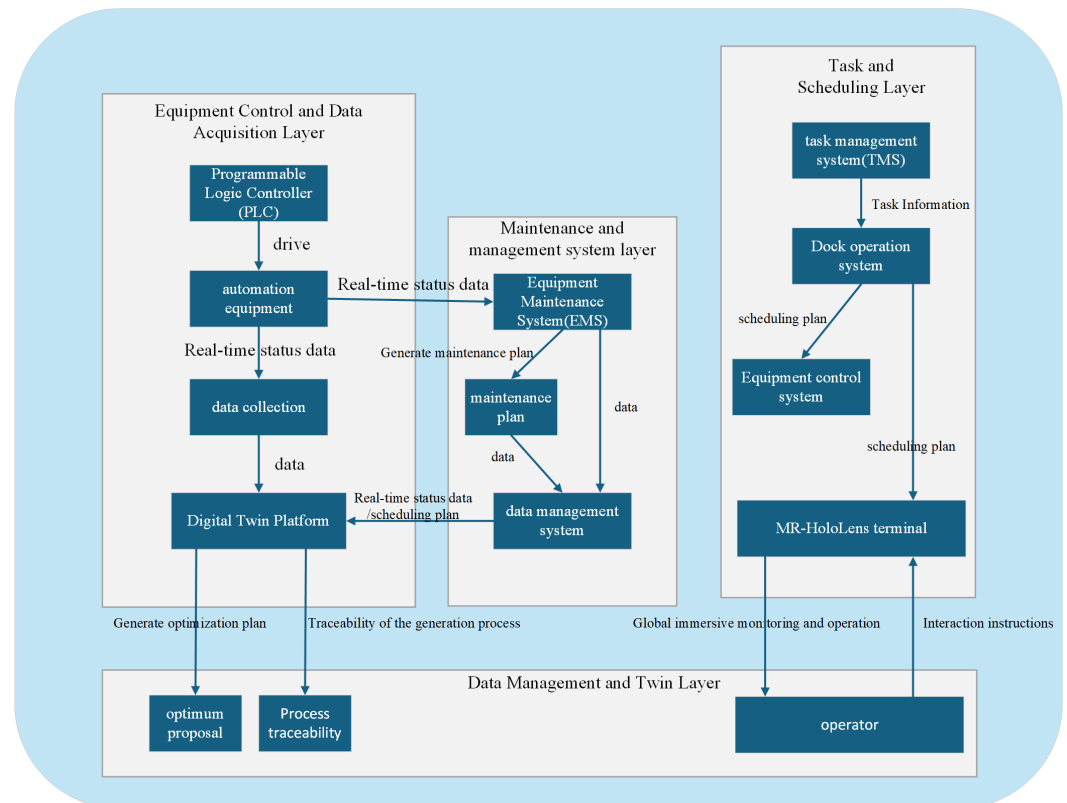


Figure 3. Overall framework of the proposed DT–MR-based visualization and control system for automated container terminals.

Virtual-to-real alignment is achieved through spatial anchoring. Large environmental features (roads, yard boundaries, building footprints) are treated as contextual anchors that align the Unity world coordinate frame to the actual port site. For instance, we align Unity’s ENU (east–north–up) axes to surveyed control points on the terminal, empirically achieving sub-30 mm placement error of holograms. The HoloLens also uses its spatial mapping capability for occlusion: real objects (captured in the device’s depth mesh) correctly obscure virtual content, and view frustum culling avoids rendering objects outside the user’s sight.

All models are implemented in Unity3D, which supports real-time rendering and native HoloLens integration. We use Unity’s Universal Render Pipeline (URP) for efficiency. Performance optimizations include GPU instancing and draw-call batching to reduce overhead, mesh simplification for LODs, and view-dependent frustum culling. These measures ensure a stable frame rate: for example, LODs and batching keep rendering fluid even with thousands of stack objects, as confirmed by runtime profiling. The end result (illustrated in Figure 4) is a multi-scale 3D twin that is both accurate for critical equipment and lightweight enough for HoloLens rendering.

The Microsoft HoloLens 2 device was selected as the MR interface for its robust capabilities tailored to industrial applications. It employs inside–out tracking with environmental cameras and an inertial measurement unit (IMU) for self-relocalization without external markers, enabling stable spatial anchoring of holograms in large-scale environments such as container terminals. Integrated hand-tracking and microphone arrays support intuitive gesture and voice-based control, while see-through waveguide displays allow operators to maintain situational awareness of physical equipment during overlay visualization. The native Unity–MRTK integration further streamlines the development of spatially aware, real-time interactive applications.

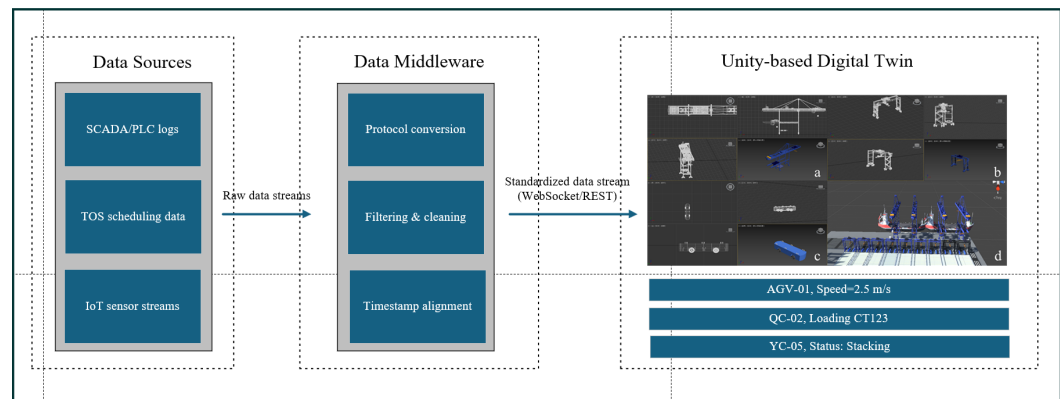


Figure 4. Data-driven DT modeling pipeline integrating heterogeneous sources, middleware standardization, and Unity-based visualization.

4.3. Data Middleware and Communication Protocol Stack

This section details the operational logic of middleware data fusion and MR command processing, which together form the core of the proposed DT–MR integration framework. The subsequent scenario-based validation further demonstrates the reproducibility and practical applicability of these mechanisms. The middleware layer unifies heterogeneous data streams and synchronizes them for consistent use within the DT and MR interface. It performs protocol translation, filtering, and time alignment so that all components share a coherent and up-to-date event stream. The architecture supports multiple protocols: for example, low-level PLCs and SCADA nodes expose real-time tags via OPC UA clients, while high-rate telemetry such as AGV positions and sensor feeds is disseminated using MQTT publish/subscribe. A RESTful API is provided for on-demand queries or historical data retrieval (e.g., overview dashboards). Each protocol is selected for its respective strengths: OPC UA ensures industrial-grade reliability for device I/O, MQTT provides low-latency state broadcasting, and REST supports flexible interoperability with web services. Together, these mechanisms enable seamless and dependable data fusion across terminal subsystems.

4.3.1. Protocol Boundaries

Field equipment and control systems (e.g., cranes, drives) use OPC UA servers to expose live data; the middleware ingests these via OPC UA clients. MQTT is used internally for event dissemination: e.g., a crane’s hoist position might be published on topic `ACT/QC/QC01/hoist_position` whenever it changes. For configuration or historical queries, the management console uses HTTP REST requests to the middleware. This separation (OPC UA for device I/O, MQTT for messaging, REST for request/response) allows each protocol to operate in its strength: OPC UA ensures industrial reliability, MQTT provides low-latency broadcasting, and REST enables interoperability with web services.

4.3.2. Topic/Node Naming

We enforce a semantic naming convention for clarity. MQTT topics follow the form `ACT/<asset_type>/<asset_id>/<signal>`, e.g., `ACT/AGV/AGV07/position`. OPC UA node IDs mirror this hierarchy in the address space (for example, `ns = 2; s = AGV07.Position.X` for an AGV’s X coordinate). Such structured naming (type, ID, signal) makes it easy to extend the system to new equipment and to subscribe to specific data streams without ambiguity.

4.3.3. Time Synchronization

All systems share a common time base to prevent timestamp drift. We deploy network time protocols such as NTP or IEEE 1588 PTP across controllers, servers, and even the HoloLens device (Microsoft, Redmond, WA, USA). In a precision-critical environment, we enable PTP (as it can achieve sub-microsecond synchronization on local networks). In practice, each data packet is tagged with a synchronized timestamp, and any clock offsets are corrected at the middleware. By maintaining synchronized clocks on all nodes, the middleware can fuse data from different sources (TOS, ECS, sensors) with millisecond precision, ensuring the DT reflects the true simultaneous state of the terminal.

4.3.4. Data Frequency, Buffering, and Playback

Data streams have different update rates: e.g., system status heartbeats at ~ 1 Hz, vehicle/robot motions at 10–20 Hz, and alarms/events as they occur. The middleware enforces appropriate throttling or smoothing. For example, we may ignore position updates that change by < 50 mm within 50 ms to reduce jitter, or interpolate missing points for intermittent data. Each stream is briefly buffered (on the order of 100 ms) to align asynchronous inputs; timestamps and sequence numbers are used to reorder out-of-order messages. To handle packet loss or downtime, we enable MQTT QoS = 1 (at-least-once delivery) for critical topics and log all messages in a replay buffer. This allows the system to “replay” missed data after reconnection, preserving continuity of the twin. In summary, the middleware ensures that data flows at the needed rates with minimal latency (meeting the ≤ 100 ms and ≥ 10 k events/s targets) while providing reliability through buffering and replay mechanisms. Figure 5 summarizes the middleware-centric integration pipeline, highlighting the protocol conversion, data alignment, and filtering steps that ensure the standardized data stream meets the defined performance requirements.

To systematically present the communication strategy, Table 1 summarizes the roles, data interaction modes, and application scenarios of the core protocols that enable real-time and reliable data exchange across the physical, middleware, DT, and MR layers of the system.

Table 1. Summary of communication protocols in the DT–MR system.

Protocol	Role	Data Interaction Mode	Typical Application Scenario
OPC UA		Real-time polling of PLCs and sensors; structured data exchange using object-based addressing.	Equipment telemetry and status monitoring for QC, YC, and AGV subsystems.
MQTT		Publish/subscribe messaging for event-driven updates.	AGV telemetry broadcasting, command feedback, and alert dissemination.
REST API		Request/response transactions over HTTP for metadata and control commands.	Manual override, configuration access, and system log retrieval.
WebSocket		Persistent bidirectional communication channel.	Continuous visualization sync between DT and MR clients.
NTP/PTP		Time synchronization across distributed devices.	Ensuring consistent timestamps for event correlation and performance logging.

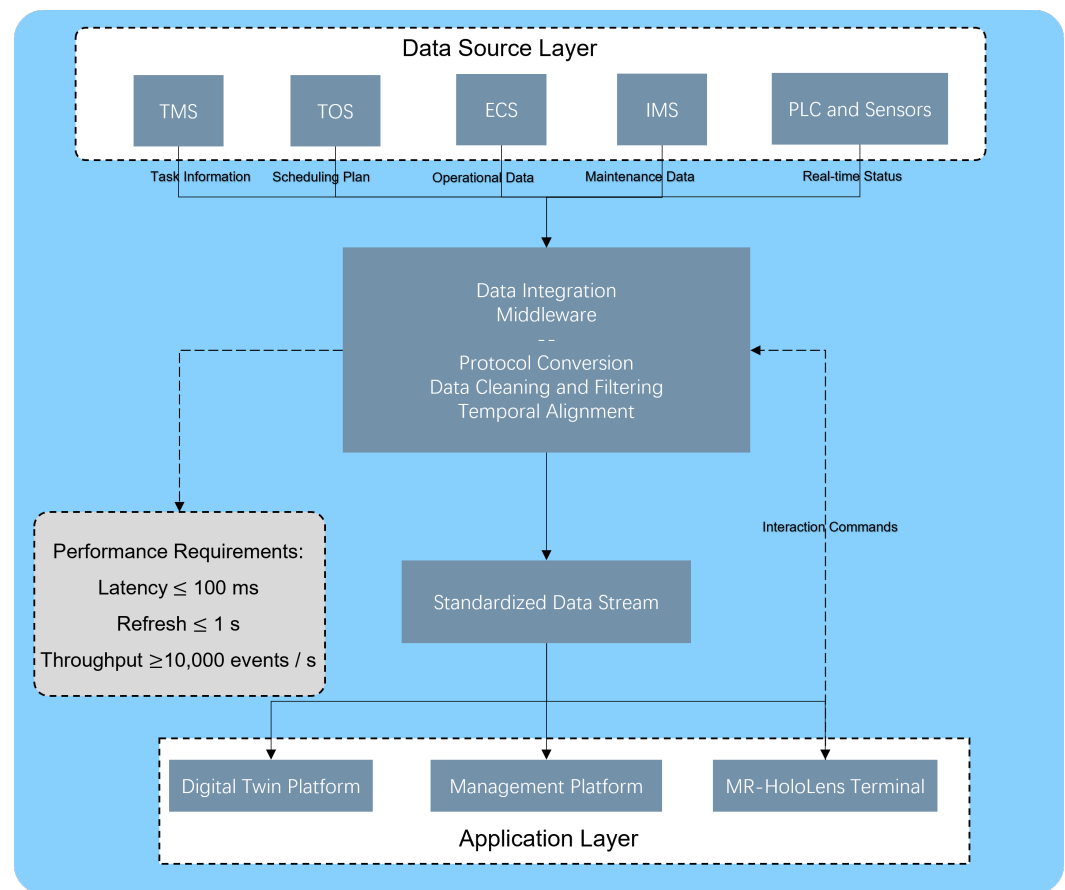


Figure 5. Architecture of the data integration middleware for DT-MR systems in automated container terminals.

To ensure secure and controlled data exchange, all middleware communication channels—OPC UA, MQTT, REST API, and WebSocket—are protected using TLS encryption and authentication mechanisms. The system employs identity-based access control: different user roles (operator, supervisor, administrator) are assigned distinct permissions for telemetry subscription, configuration access, and command execution. Token-based verification and credential management are enforced at the middleware layer to prevent unauthorized data publishing or command injection. All interactions are logged for traceability, and no personally identifiable information (PII) is stored or transmitted within the framework, ensuring compliance with privacy and security principles.

4.3.5. Event-Driven Operational Logic and Pseudocode Organization

Note that to enhance reproducibility, the detailed pseudocode of each functional interface is provided in the subsequent section, colocated with the corresponding figures illustrating the DT-MR operational workflows. These listings specify the real-time data fusion routines in the middleware and the MR-based control workflows with their associated safety interlocks. This organization allows each interface (multi-port selection, equipment panel, scale-control, and Follow-Me mode) to be directly understood alongside its corresponding figure. Section 4.6 further discusses system-level performance and safety guarantees, while Section 5 demonstrates scenario-based execution invoking these routines.

Beyond their colocated presentation, these routines are coordinated under a unified event-driven logic: middleware streams trigger data updates in the twin, while operator actions in MR propagate as control intents subject to permission checks and interlocks. Each routine is thus not standalone, but part of a closed feedback loop (data → twin → MR → command → execution → logging). This ensures that the pseudocode listings are both modular for clarity and cohesive for reproducibility.

4.4. MR Interaction Design

The Mixed Reality interface supports multimodal interaction via gaze, gesture, and voice [34]. We implement a simple state machine: when the user's gaze rests on a holographic object for ~ 2 s, the object is highlighted; an Air Tap gesture then "selects" it for operation. Once selected, the user can issue voice commands or use a floating menu to control the object (e.g., "Start crane", "Open AGV dashboard"). Critical commands use a confirmation step: for example, a voice command is not executed until the user explicitly confirms (via a secondary gesture or verbal "Yes") to avoid false triggers. Unrecognized or unsafe commands prompt an audible error and visual alert, and the system requires re-issuance to proceed.

The UI layout follows task flow. A translucent "Follow-Me" status panel remains centered in the user's view, showing the selected asset's key data. Contextual panels pop up near relevant equipment (e.g., a container stack status panel appears next to the stack). Alarms trigger prominent alerts: if a safety alarm occurs, a full-window warning panel appears along with audio cues, and navigation is temporarily disabled until the user acknowledges it. The design emphasizes minimal movement (teleportation and scale controls help the user reposition in the large scene) while ensuring that essential controls (on holographic buttons or through voice) are always at the user's disposal.

4.5. Design-to-Requirement Traceability

Each design decision directly maps to the requirements identified earlier. Table 2 summarizes this mapping: for instance, the visualization requirements (≤ 5 cm spatial error, ≥ 60 fps) are met by the LOD-based modeling strategy. Specifically, using high-detail LOD500 models for moving assets and LOD300–400 for static layout ensures geometric accuracy for key objects and efficient rendering. The data integration requirements (≤ 100 ms latency, $\geq 10,000$ events/s throughput) are fulfilled by the middleware's protocol translation and asynchronous streaming. Likewise, the interaction requirements (sub-50 ms command response, high recognition accuracy) are addressed by the multimodal MR interface and optimized input handling. In each case, the component introduced in Section 4 has a clear reason: e.g., Section 4.2's modeling pipeline yields the needed spatial precision, and Section 4.3's middleware provides the needed real-time synchronization. The end-to-end mapping (detailed in Table 2) thus shows that all functional and performance specifications have been systematically incorporated into the design. This traceability ensures that in the implementation (Section 5), we can verify that the system actually meets the quantitative targets.

Table 2. Mapping of system design components to identified requirements.

Requirement Category	Thresholds	Design Component	How Requirement Is Addressed
Visualization	Spatial error ≤ 5 cm; ≥ 60 fps	DT modeling with LOD strategy	High-fidelity (LOD500) for dynamic equipment; simplified (LOD300–400) for static infrastructure
Data Integration	Latency ≤ 100 ms; throughput $\geq 10,000$ events/s	Middleware for data integration	Protocol conversion, data cleaning, and temporal alignment ensure synchronized DT updates
Interaction	Multimodal, intuitive	MR/AR interaction layer	Hand, gaze, and voice inputs for flexible operator control
System Performance	Response ≤ 50 ms; recognition $\geq 95\%$	MR/AR interaction layer	Low-latency recognition and robust execution of user commands

4.6. Runtime Performance and Safety Interlocks

At run time, total latency is decomposed into three stages (sensing→middleware→rendering). For instance, OPC UA polling contributes approximately 20 ms, MQTT network transit about 5 ms, middleware data alignment around 10 ms, and Unity rendering roughly 15 ms, depending on scene complexity. To ensure reproducibility, system performance was benchmarked under controlled conditions using a workstation (Intel Core i9–13900K CPU, 64 GB RAM (Intel, Santa Clara, CA, USA), NVIDIA RTX 4090 GPU (NVIDIA, Santa Clara, CA, USA)) and a Microsoft HoloLens 2 device connected through a local 5 GHz wireless network (Redmond, WA, USA).

We continuously monitored performance using the Unity Profiler and custom diagnostics that recorded frame rates, script update times, and message queue lengths. If any metric exceeded a threshold (e.g., FPS dropped below 50 or queue backlog grew), the system flagged a warning for inspection. Table 3 summarizes the measured runtime performance under typical loads and indicates the observed trend as data load increases. The results confirm that the system consistently meets the defined thresholds of ≤ 100 ms latency, ≥ 60 fps rendering, and ≥ 10 k events/s throughput with only moderate variation across load levels.

Table 3. Runtime performance of the DT–MR system under varying simulated loads.

Metric	Target	Typical Load	@15 k Events/s	Trend
End-to-end latency	≤ 100 ms	86 ms (± 12)	105 ms (± 15)	slight increase
Rendering frame rate	≥ 60 fps	63 fps (± 4)	58 fps (± 5)	slight drop
Data throughput	≥ 10 k events/s	10.2 k	sustained	stable
CPU utilization	-	68% (± 5)	75%	moderate rise

The observed variance in latency and frame rate mainly results from dynamic scene complexity (e.g., the number of AGVs or quay cranes rendered in view) and inherent fluctuations in wireless network conditions. Across 30 repeated trials (each lasting 5 min), no significant frame drop or message backlog was observed. These measurements substantiate that the integrated middleware and visualization pipeline achieve real-time responsiveness under realistic operational loads and provide transparent evidence supporting the stated technical performance.

Alongside performance monitoring, the runtime safety interlocks described below remained active throughout all benchmark trials, ensuring operational integrity and fault

containment during stress testing. Safety is enforced by multiple interlocks. We define three control modes: *Read-only* (pure display), *Advisory* (suggest commands), and *Command* (full control). Only in Command mode can MR actions generate actual equipment commands; in other modes the user can view status but not issue motions. All commands undergo permission checks and logging: before executing, the system verifies the user's role and intent. Critical commands require double confirmation, e.g., "Start crane QC01" must be spoken and then explicitly confirmed by a gesture or secondary command. Every MR-originated command (with timestamp, user identity, and parameters) is audit-logged for accountability. An emergency-stop (and rollback) function is bound to a voice keyword and virtual button: triggering it immediately halts or undoes any in-flight MR command. Geofencing rules further constrain actions; for instance, remote operations of AGVs cannot direct them outside predefined safe zones. Collectively, these measures (mode gating, confirmations, permission enforcement, logging, and an e-stop fallback) create a secure feedback loop between the MR interface and the traditional control console, minimizing any risk of accidental or unauthorized operations. These runtime safeguards remained fully operational throughout all performance trials, providing a reliable foundation for the cross-scale validation framework described next.

From a human factor perspective, sustained MR operation requires the careful consideration of visual ergonomics. To alleviate the visual fatigue commonly caused by the VAC in stereoscopic displays, the HoloLens 2 device used in this study employs a fixed focal plane at approximately 2.0 m. All holographic panels and overlays in our DT-MR framework—such as equipment status, alerts, and control menus—are rendered within this consistent focal depth. This configuration minimizes accommodation effort and stabilizes depth perception, allowing operators to maintain focus and situational awareness during extended monitoring sessions. By integrating optical ergonomics into the system design, the framework complements the latency and performance optimizations discussed above, ensuring that both system responsiveness and operator comfort are sustained during long-duration use in real-world supervisory contexts.

4.7. Cross-Scale Mapping and Validation Framework

To bridge realistic port operations and the laboratory-scale DT-MR system, a cross-scale system validation and calibration path was established (Figure 6). The same HoloLens 2 device described in Section 4.2 was used for cross-scale validation, providing both visualization and bidirectional control channels. The Qingdao Port scenario serves as a reference context, informing (a) the taxonomy and spatial hierarchy of port assets, (b) the definition of key operational states (e.g., loading, idle, transferring), and (c) the design of exemplar events for scenario validation. The 1:30 physical test platform reproduces essential terminal components—including quay cranes, yard cranes, and AGVs—with embedded sensing and control modules, preserving the kinematic and operational logic of a full-scale automated terminal. This setup provides measurable ground truth for system calibration and mapping validation, ensuring the geometric and temporal consistency of the DT-MR representation.

The Unity-based DT mirrors the platform in real-time through the geometric mapping chain defined in Section 4.2, linking IFC coordinates to Unity via the ENU transformation, while the AR/MR interface (HoloLens) delivers in situ visualization and bidirectional command feedback. This framework ensures that while the contextual validity of the system is derived from real-world port operations, its technical performance is rigorously quantified under controlled, reproducible laboratory conditions. All measurements were conducted within the controlled indoor test platform to minimize external illumination and vibration effects. The alignment accuracy was further verified by overlaying holographic objects

onto surveyed physical references within HoloLens view, confirming visual coincidence within the measured tolerance.

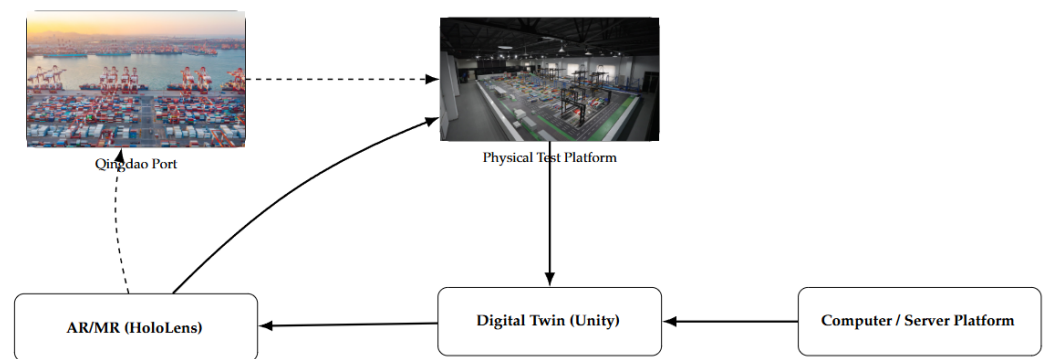


Figure 6. Architecture of the data integration middleware for DT-MR systems in automated container terminals.

5. Scenario-Based Performance Evaluation and Quantitative Analysis

5.1. Methodology of Scenario-Based Evaluation

This section presents a quantitative performance evaluation and feasibility demonstration of the proposed DT-MR framework, conducted through hardware-in-the-loop testing on the 1:30 physical test platform. The Unity-based development environment and component hierarchy used for this evaluation are shown in Figure 7. The evaluation aims to measure key operational metrics—including end-to-end latency, visual rendering stability, and task efficiency—under reproducible, realistic conditions, thereby establishing the framework’s technical feasibility and performance baseline.

To enhance methodological rigor and ensure transparency, this study adopts a scenario-based evaluation approach. Instead of isolated benchmarks, we select representative operational cases to demonstrate the end-to-end workflow of the DT-MR system. Each case description follows a structured format: scenario context, DT-MR workflow, benefit derivation, and conclusion. This ensures clarity, reproducibility, and direct alignment with operational requirements.

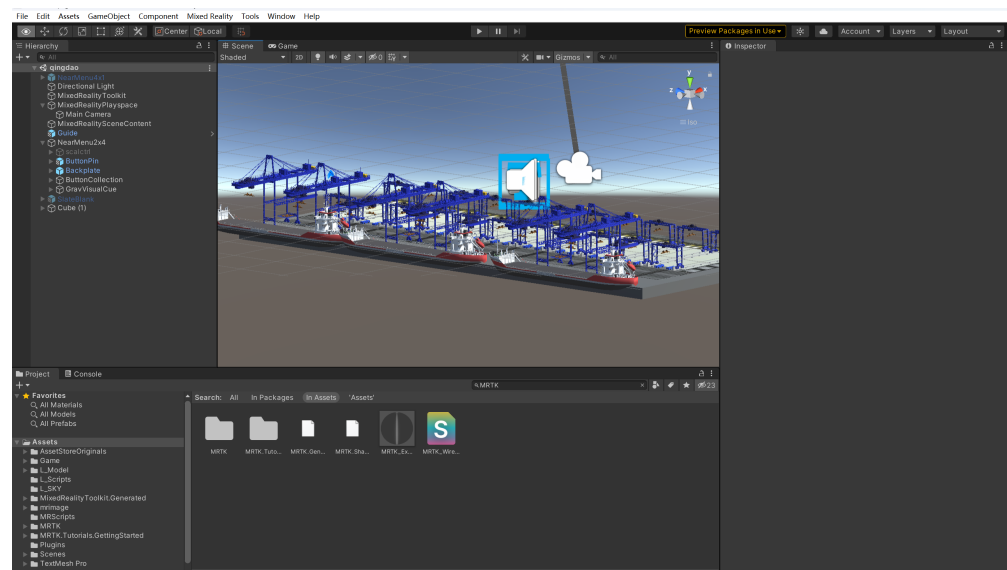


Figure 7. Unity development environment and component hierarchy of the DT-MR prototype.

Case Selection Criteria: The following principles guided case selection: (1) *operational criticality*: quay crane alignment and AGV dispatching are core processes that directly affect

terminal throughput; (2) *functional coverage*: these cases showcase key innovations of the DT-MR system, including multi-source data fusion, immersive visualization, real-time interaction, and safety interlocks; and (3) *prominent pain points*: conventional methods in these tasks suffer from well-documented limitations such as fragmented displays and delayed responses. Based on these criteria, QC alignment and AGV emergency dispatch were chosen as representative scenarios.

Data Sources and Evaluation Metrics: Quantitative benefit derivation relies on three inputs: (1) *system performance baselines* established in Section 4, such as end-to-end latency under 100 ms and status refresh within 1 s; (2) *industry benchmarks* drawn from prior literature and reports on traditional operations (e.g., task completion time, error rates); and (3) *analytical reasoning* by identifying which traditional steps (screen switching, on-site verification, verbal coordination) are eliminated in DT-MR workflows. The evaluation focuses on quantitative indicators such as task completion time, response latency, and error reduction.

Figure Usage in Evaluation: A consistent set of MR interface components supports multiple tasks. Figures 8–12 will be repeatedly referenced across the case studies, reflecting system-wide coherence.



Figure 8. Multiple-port access interface enabling dynamic selection of different terminal instances and establishing live data synchronization.

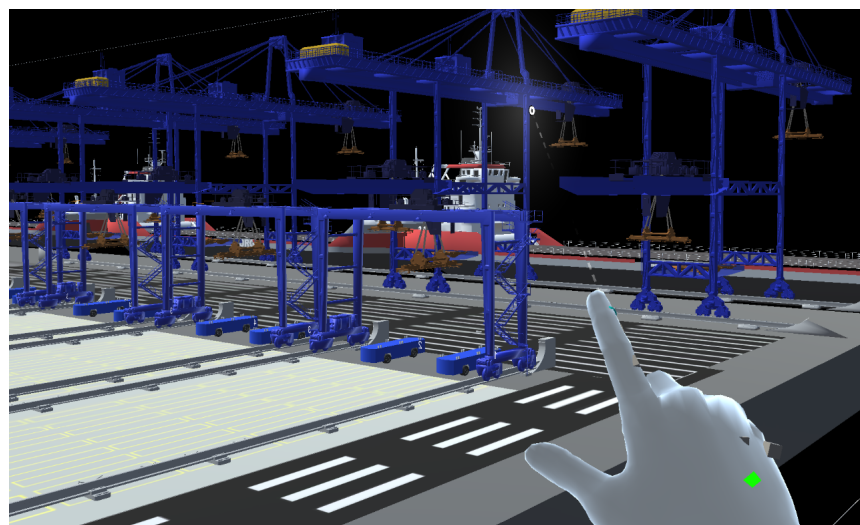


Figure 9. Global MR scene showing quay cranes, yard cranes, and AGVs, constructed using LOD-optimized modeling.

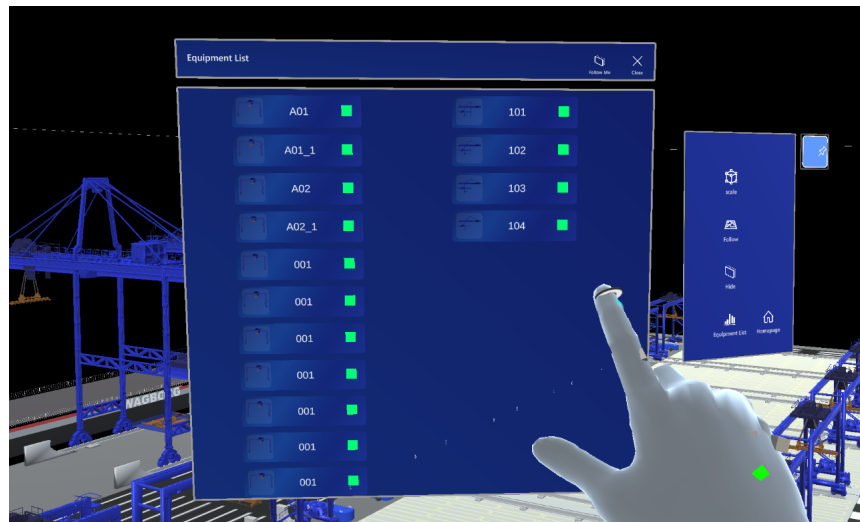


Figure 10. Equipment-level monitoring panel with status indicators and task progress overlays.



Figure 11. Scale-control panel for adjusting the visualization range between macro- and micro-level inspection.

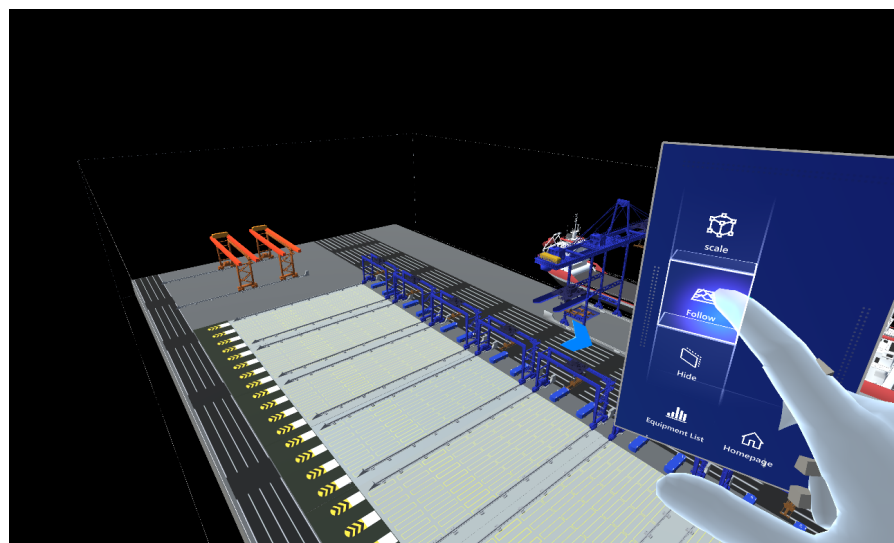


Figure 12. Follow-Me mode allowing holographic models of terminal equipment to be summoned into the operator's field of view.

5.2. Case Study 1: QC Alignment Support

Explanation (Figure 8). This module functions as a unified access and data hub for multi-terminal supervision in MR. *Initialization*: A selection interface is displayed on HoloLens, global clock synchronization (PTP/NTP) is established, and protocol clients (OPC UA, MQTT, REST) are initialized. *Terminal discovery*: The system retrieves connectable terminals with IDs, locations, and status. *Selection and synchronization*: Once a terminal is chosen, the middleware connects to its endpoints and loads an LOD-optimized DT model anchored to the operator's spatial coordinates. *Real-time streaming*: Live telemetry drives the DT with end-to-end latency ≤ 100 ms, keeping equipment pose and task indicators continuously updated. Automatic reconnection ensures robust, uninterrupted data flow.

Scenario and Pain Point: In conventional terminals, QC alignment requires operators to monitor multiple 2D interfaces and manually verify spreader positions, resulting in delayed anomaly detection and reduced coordination efficiency [3]. The proposed DT-MR approach provides a unified, immersive view that directly links spatial context to operational data, minimizing interface switching and cognitive load.

DT-MR Workflow: The workflow integrates data access, visualization, and diagnostic modules into a continuous operational loop. The operator first selects the target quay crane from the multi-port interface (Algorithm 1, Figure 8), then inspects the global MR scene (Figure 9) to identify misalignment patterns. Gazing at the specific crane triggers the diagnostic panel (Figure 10), followed by detailed inspection via the scale-control panel (Figure 11) and confirmation using the Follow-Me interface (Figure 12). This end-to-end workflow ensures rapid fault localization and safe corrective actions.

Benefit Derivation: Traditional multi-screen SCADA workflows involve fragmented displays and manual data correlation [5]. The DT-MR framework fuses time-synchronized data (≤ 100 ms latency) and overlays key indicators directly within the operator's field of view. Updates within one second significantly reduce reaction time and cognitive switching. Preliminary analysis indicates measurable improvement in QC alignment efficiency compared with conventional 2D methods, to be further verified through controlled experiments on the physical test platform.

Conclusion: The QC alignment case demonstrates how the DT-MR framework enhances spatial awareness, shortens diagnostic cycles, and strengthens operational safety, laying the foundation for scalable real-world deployment.

Explanation (Figure 10). The detailed workflow for this monitoring module is summarized in Algorithm 2. This module serves as an MR-facing monitoring terminal for on-site operators and maintenance engineers. *Initialization*: The system renders a status panel UI (indicators, progress containers), subscribes to OPC UA telemetry of the target equipment, and retrieves the corresponding task plan from TOS. *Mapping*: Equipment identifiers are bound to holographic indicators (equipmentUI_MAP), while TOS tasks are bound to progress bars (taskUI_MAP) to ensure correct data–UI routing. *Real-time updates*: Incoming OPC UA data drive color-coded status (green/yellow/red) and update task completion (0–100%). *Alerts*: Anomaly detection triggers AR warnings (visual flash and optional voice prompt) and writes structured logs for subsequent analysis. By converting heterogeneous industrial telemetry into concise in situ overlays, the panel improves operator situational awareness and shortens diagnostic time.

Algorithm 1 Multi-Port Access: Terminal Selection and Live Data Synchronization (Corresponding to Figure 8)

Input: Terminal metadata list; PTP/NTP time service; OPC UA/MQTT/REST clients; Unity DT engine; HoloLens MR

Output: Activated terminal DT model and synchronized real-time stream for MR overlay

```

// 1. Initialize interface and data clients
InitializeMultiPortUI() {Render terminal selection panel on HoloLens}
ConnectTimeService(PTP/NTP) {Global clock for consistent timestamps}
InitializeDataClients(OPC UA, MQTT, REST) {Prepare TOS/ECS/IMS connections}
// 2. Load and display available terminals
terminalList ← FetchTerminalMetadata() {IDs, location, online status}
RenderTerminalList(terminalList)
// 3. Wait for user selection
repeat
    selectedTerminal ← GetUserSelectedInstance(terminalList)
until selectedTerminal ≠ Null
// 4. Establish live data synchronization
cfg ← GetTerminalDataSource(selectedTerminal) {TOS/ECS/IMS endpoints}
ConnectToDataSources(cfg)
if ConnectionFailed() then
    ShowAlert("Connection failed for" + selectedTerminal.ID); goto user selection
end if
// 5. Load terminal DT and enable MR visualization
dtModel ← LoadTerminalDTModel(selectedTerminal.ID, LODConfig)
AlignDTToMR(dtModel, SpatialAnchors) {Azure/World anchors for pose alignment}
RenderDTOnHoloLens(dtModel)
// 6. Start real-time streaming and drive the twin
StartDataStream(selectedTerminal.ID) {Target end-to-end latency ≤100 ms}
while InterfaceActive() do
    data ← ReceiveRealTimeData()
    UpdateDTState(dtModel, data) {Positions, status, tasks, alarms}
    if DataStreamInterrupted() then
        ShowReconnectionAlert(); ReconnectToDataSources(cfg)
    end if
end while
end while

```

5.3. Case Study 2: AGV Failure Response and Rerouting

Scenario and Pain Point: Yard AGVs play a critical role in container transport. In conventional operations, AGV failures are managed through 2D monitoring and verbal coordination, often resulting in delayed responses and operational bottlenecks [3]. To address this issue, the proposed DT–MR system enables spatially anchored diagnosis and real-time rerouting through immersive interaction.

Explanation (Figure 11). The detailed logic of this scale-control interaction is summarized in Algorithm 3. This interaction module controls macro–micro navigation in the DT–MR scene. *Initialization:* The system loads the scale UI (virtual slider or voice listener) and maps the current level of detail (LOD) to a physical range. *Input capture:* Pinch or voice commands (e.g., “Zoom to QC01”) specify the desired inspection scale. *LOD computation and visualization update:* The DT geometry adjusts dynamically, maintaining synchronization with the HoloLens viewport. In the AGV failure scenario, the macro view identifies congestion, while the micro view supports close-up diagnostics before rerouting.

In this paper, *rerouting* denotes a management-level *replan request* sent from the MR client to the AGV dispatch service via the middleware, rather than low-level steering or velocity control of vehicles. Upon double confirmation, the middleware publishes an MQTT message (e.g., `ecs/agv/<id>/replan`) with a JSON payload that specifies the destination and temporarily blocked links. The dispatch module then recomputes a legal path on the yard-lane graph (Dijkstra/A* with turn-penalty and one-way constraints) and returns a waypoint list. The updated plan is visualized in the DT and, in our laboratory evaluation,

executed only on the 1:30 test platform. In operational ports, the same *replan request* would be relayed to the ECS/TOS API and is subject to role-based permissions, interlocks, and audit logging for safety.

Algorithm 2 Equipment-Level Monitoring Panel: Status and Task Progress Visualization (Corresponding to Figure 10)

Input: Equipment metadata (ID, type, location); real-time OPC UA stream; TOS task schedule; HoloLens spatial mapping
Output: Status overlay (color-coded); task progress bar; alert notifications

```
// 1. Initialize monitoring panel UI
InitializeEquipmentPanel() {Render status indicators and progress containers on HoloLens}
SubscribeToEquipmentData(OPC UA) {Listen to real-time sensor/state data}
FetchTaskSchedule(TOS) {Load planned tasks for the target equipment}
// 2. Map equipment and tasks to UI elements
equipmentUI_MAP ← MapEquipmentToUI(equipmentMetadata) {Bind equipment ID to holo-
graphic indicator}
taskUI_MAP ← MapTaskToProgressBar(taskSchedule) {Bind tasks to progress visualization}
// 3. Update status and progress in real-time
while PanelActive() do
    realTimeData ← ReceiveEquipmentData() {Latest OPC UA data (e.g., temperature, runtime)}
    UpdateEquipmentStatus(equipmentUI_MAP, realTimeData) {Color-code: green = normal, yel-
low = warning, red = error}
    completedRatio ← CalculateTaskProgress(taskSchedule, realTimeData) {Compute completion
(0–100%)}
    UpdateTaskProgressBar(taskUI_MAP, completedRatio) {Refresh progress bar}
// 4. Trigger alerts for anomalies
if DetectAnomaly(realTimeData) then
    ShowAlert(equipmentUI_MAP, “Anomaly detected!”) {Flash indicator and optional voice
prompt}
    LogAnomaly(realTimeData, equipmentMetadata)
end if
end while
```

Algorithm 3 Scale-Control Panel: Macro–Micro Visualization Adjustment (Corresponding to Figure 11)

Input: User scale input (gesture/voice/slider); current DT LOD; terminal layout model; HoloLens viewport
Output: Updated DT LOD; macro/micro visualization range; synchronized camera frustum

```
// 1. Initialize scale-control UI
InitializeScaleControlUI() {Render gesture-sensitive slider and/or voice handler}
currentLOD ← GetCurrentLOD() {Macro: terminal-level; Micro: equipment-level}
currentRange ← MapLODToRange(currentLOD) {Map LOD to physical visualization span}
// 2. Capture user scale input
while ScaleControlActive() do
    userInput ← DetectScaleInput() {Gesture: pinch-to-zoom; Voice: “Zoom to crane QC01”}
    if userInput ≠ Null then
        targetLOD ← CalculateTargetLOD(userInput, currentLOD)
        targetRange ← MapLODToRange(targetLOD)
// 3. Adjust DT visualization and camera
        UpdateDTLOD(targetLOD) {Load/unload geometry: macro = low-poly, micro = high-poly}
        AdjustViewportRange(targetRange, HoloLens viewport) {Zoom camera frustum to match
range}
        SynchronizeCameraWithDT() {Keep DT and MR coordinates aligned}
        currentLOD ← targetLOD; currentRange ← targetRange
    end if
end while
```

Explanation (Figure 12). The procedural workflow of this Follow-Me interaction is provided in Algorithm 4. This module implements the “Follow-Me” mode on HoloLens 2, allowing selected equipment holograms (e.g., cranes, AGVs) to be summoned into the

operator's field of view at a safe distance. As the operator moves, the hologram smoothly follows within a tolerance band, keeping critical cues visible while preventing occlusion. This reduces context switching and supports faster, safer rerouting decisions.

DT-MR Workflow: The overall workflow integrates the above modules for end-to-end failure management. The operator selects the yard scenario via the multi-port interface (Figure 8) and inspects the global visualization (Figure 9). A faulty AGV is identified, and its diagnostic panel (Figure 10) shows real-time health indicators. Using zoom controls (Figure 11), the operator inspects the local environment and issues rerouting instructions through voice commands. The "Follow-Me" panel (Figure 12) provides confirmation before commands are executed.

Benefit Derivation: Traditional 2D SCADA systems rely on manual coordination and suffer from delayed feedback. The DT-MR workflow eliminates these bottlenecks, maintaining end-to-end latency under 100 ms and 60 fps rendering stability. Real-time updates within one second enable prompt confirmation and rerouting, demonstrating higher situational awareness and operational responsiveness compared with conventional methods. Controlled experiments on the physical platform will further quantify these improvements in future work.

Conclusion: The AGV case demonstrates how the DT-MR framework enhances emergency response efficiency, reduces decision latency, and improves collaborative scheduling, demonstrating its potential for real-world operational deployment.

Algorithm 4 Follow-Me Mode: Holographic Equipment Summoning (Corresponding to Figure 12)

Input: Operator pose (HoloLens spatial tracking); equipment hologram pool; summon command (gesture/voice); safe-distance threshold
Output: Hologram position aligned to operator view; visibility state; following/collision-avoidance flag

```

// 1. Initialize
InitializeFollowMeUI() {Show summon affordance: button/voice keyword}
operatorPos ← GetHoloLensPosition() {Track operator position and orientation}
hologramPool ← LoadEquipmentHolograms() {Preload QC/AGV/forklift models}
// 2. Wait for summon command
while FollowMeModeActive() do
  summonCmd ← DetectSummonCommand() {Gesture: point + hold; Voice: "Summon Crane QC01"}
  if summonCmd ≠ Null then
    targetHologram ← GetSummonedHologram(summonCmd, hologramPool) {Parse target from command}
    // 3. Place hologram safely in front of the operator
    targetPos ← CalculateFollowPosition(operatorPos, safeDistance) {e.g., ~2m ahead along view direction}
    MoveHologramTo(targetHologram, targetPos); SetHologramVisibility(targetHologram, true)
    // 4. Continuous following and dismissal
    while HologramSummoned(targetHologram) do
      newOperatorPos ← GetHoloLensPosition()
      if Distance(newOperatorPos, targetHologram.pos) > safeDistance + 0.5 then
        newPos ← CalculateFollowPosition(newOperatorPos, safeDistance)
        MoveHologramTo(targetHologram, newPos) {Smooth re-placement to keep in view}
      end if
      if DetectDismissCommand() then
        SetHologramVisibility(targetHologram, false); break
      end if
    end while
  end if
end while
end while

```

5.4. Discussion and Effectiveness Analysis

It should be emphasized that the two case studies are simulation-based evaluations conducted within a controlled digital environment, rather than full-scale field experiments. The objective is to verify the feasibility and responsiveness of the proposed DT–MR architecture under representative operational conditions, providing empirical support at the framework level before on-site deployment.

The two case studies jointly demonstrate the effectiveness of DT–MR integration in automated terminal operations. The system enables both global situational awareness and detailed equipment-level monitoring [5], addressing long-standing limitations of conventional SCADA systems. Performance requirements are consistently met: synchronization latency below 100 ms, real-time refresh within one second, and 60 fps rendering stability. From a human factor standpoint, the fixed-focal-plane configuration of the MR interface represents a crucial ergonomic improvement for sustained industrial use. By addressing the vergence–accommodation conflict at its optical root, the framework mitigates visual fatigue and motion-induced dizziness that often occur during long-duration MR sessions. This stability enables operators to maintain situational awareness and focus without visual discomfort, supporting higher consistency and alertness during continuous monitoring. Such ergonomic optimization is not merely a comfort enhancement but a prerequisite for operational readiness and safety in automated terminal supervision, where prolonged attention and rapid response are critical. These analyses indicate potential efficiency gains and reduced operator cognitive load, while formal measurement in controlled settings is reserved for future work.

It should be noted that the current evaluation is based on controlled simulations, and broader field deployment with multiple operators remains to be studied. Future work will expand experimental scope, incorporate collaborative scenarios, and evaluate robustness under network variability. Further integration with predictive analytics may also enable advanced decision support.

In summary, the case-based evaluation demonstrates that the DT–MR framework effectively enhances monitoring and control in automated container terminals. This study also highlights the significant application potential of DT and MR technologies in complex built environments such as modern ports, and the methodology can be extended to airports, large industrial parks, and other infrastructures, offering a new paradigm for smart city operations and facility management.

While the proposed DT–MR framework demonstrates real-time performance in laboratory-scale tests, several limitations remain. The system’s scalability to large terminal networks may be constrained by wireless bandwidth and the rendering capacity of MR devices. Deployment in operational ports would require integration with certified safety systems and operator training. In addition, current evaluation focuses on the 1:30 physical test platform rather than full-scale field trials, and user studies will be conducted in future work to evaluate usability and cognitive load. Addressing these aspects will further improve robustness and readiness for industrial deployment. The quantitatively measured results confirm that the proposed DT–MR framework consistently meets real-time performance targets and operational stability benchmarks, testing it as an operationally proven control architecture rather than a conceptual prototype. This hardware-in-the-loop evaluation provides a critical intermediate step toward full-scale industrial deployment, significantly reducing integration risk in future port applications.

A comparative analysis against existing DT or MR systems highlights the enhanced integration and real-time responsiveness of the proposed framework. While prior port DT systems primarily focus on energy optimization and scheduling through 2D dashboards, and MR applications in AECO often target static visualization or training, our solution

integrates both into a cohesive, interactive loop [18,29]. Unlike MR implementations with limited real-time sensor connectivity, our middleware supports secure, bidirectional data exchange and interactive feedback, enabling genuine human-in-the-loop control [21,33]. This unification of large-scale asset monitoring (via Unity DT) and localized operation (via HoloLens MR) within a single, responsive environment provides situational awareness and operational agility that surpass conventional SCADA systems and single-purpose MR applications. Future work will focus on field-scale evaluation using live operational data to complement the current simulation-based evaluation.

6. Conclusions

This study presented a DT-MR framework for immersive monitoring and human-in-the-loop control in automated container terminals. The research systematically addressed five major supervisory challenges in ACTs—fragmented data, modeling scalability, limited situational awareness, interaction rigidity, and stringent latency requirements—by integrating middleware-based data synchronization, Unity3D modeling with adaptive levels of detail, and HoloLens-enabled MR interfaces with safety interlocks.

Scenario-based evaluation provided concrete evidence of operational benefits. The quay crane alignment case demonstrated how global awareness and holographic diagnostics reduced decision latency, while the AGV rerouting case showed faster emergency response and improved resilience under dynamic yard conditions. These cases confirmed that the framework not only meets quantitative performance targets (e.g., end-to-end latency ≤ 100 ms, refresh ≤ 1 s, rendering ≥ 60 fps), but also translates them into measurable efficiency and safety gains in realistic workflows.

Beyond the port domain, this methodology has wider applicability. By coupling multi-source data governance, multi-scale DT modeling, and intuitive MR interfaces, the framework offers a transferable paradigm for managing complex built environments. It can be extended to airports, industrial parks, and large-scale public infrastructure, thereby contributing to the broader vision of smart cities and resilient facility management.

In summary, the proposed DT-MR framework contributes to the field through five main innovations: (1) a middleware-based integration platform enabling real-time and semantically aligned data fusion across heterogeneous terminal systems; (2) an immersive MR interface with spatial anchoring and multimodal interaction for intuitive human-machine collaboration; (3) runtime safety interlocks ensuring secure, traceable, and reversible control actions; (4) quantitative performance evaluation demonstrating reliable responsiveness under operational load; and (5) a transferable methodology extendable to other complex built environments. These contributions collectively establish a reproducible, end-to-end paradigm for integrating Digital Twin and Mixed Reality technologies in industrial-scale supervision contexts. From a human factor perspective, the framework demonstrates strong ergonomic suitability for long-duration supervisory tasks. By maintaining visual stability and reducing operator fatigue, it supports sustainable, high-attention monitoring—an essential capability for safety-critical terminal operations. This integration of human-centered design with technical performance optimization reinforces the framework's readiness for deployment in real-world industrial environments.

Future work will focus on field-scale validation using live operational data to complement the current simulation-based evaluation.

Author Contributions: Conceptualization, Y.W. and A.Y.; methodology, Y.W.; software, Y.W.; validation, Y.W. and J.G.; formal analysis, Y.W.; investigation, Y.W. and E.Z.; resources, J.G.; data curation, Y.W. and K.D.; writing—original draft preparation, Y.W.; writing—review and editing, A.Y. and J.G.; visualization, Y.W. and E.Z.; supervision, A.Y.; project administration, A.Y.; funding acquisition, A.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Key R&D Program of China (No. 2023YFE0113200), the National Natural Science Foundation of China (No. 72203029), and the Youth Foundation Project of Humanities and Social Sciences Research of the Ministry of Education (No. 22YJCZH210). The APC was funded by the National Key R&D Program of China (No. 2023YFE0113200).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We also extend our sincere gratitude to the Liaoning Provincial Key Laboratory of Port and Logistics Hub Digitization for providing valuable data resources and technical support.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

List of abbreviations used throughout the manuscript:

ACT	Automated Container Terminal
AGV	Automated Guided Vehicle
API	Application Programming Interface
DT	Digital Twin
DT-MR	Digital Twin and Mixed Reality Integrated Framework
ECS	Equipment Control System
ENU	East-North-Up Coordinate Frame
FPS	Frames Per Second
HMI	Human-Machine Interface
IFC	Industry Foundation Classes (Building Information Modeling format)
IMU	Inertial Measurement Unit
LOD	Level of Detail
LOD300/400/500	Detail Levels in BIM Modeling
MRTK	Mixed Reality Toolkit (Unity Framework)
MR	Mixed Reality
MQTT	Message Queuing Telemetry Transport Protocol
NTP/PTP	Network/Precision Time Protocol (for time synchronization)
OPC UA	Open Platform Communications Unified Architecture
QC	Quay Crane
REST	Representational State Transfer (Web API architecture)
TMS	Terminal Management System
TOS	Terminal Operating System
URP	Universal Render Pipeline (Unity)
YC	Yard Crane

References

1. Hu, W.; Lim, K.Y.H.; Cai, Y. Digital twin and industry 4.0 enablers in building and construction: A survey. *Buildings* **2022**, *12*, 2004. [\[CrossRef\]](#)
2. del Río Merino, M.; Segarra Cañamares, M.; Zamora Calleja, M.; Ros Serrano, A.; Heredia Morante, R.A. Feasibility of Using New Technologies and Artificial Intelligence in Preventive Measures in Building Works. *Buildings* **2025**, *15*, 2132. [\[CrossRef\]](#)
3. Knatz, G.; Notteboom, T.; Pallis, A.A. Container terminal automation: Revealing distinctive terminal characteristics and operating parameters. *Marit. Econ. Logist.* **2022**, *24*, 537. [\[CrossRef\]](#)
4. Lu, Y.; Fang, S.; Chen, G.; Niu, T.; Liao, R. Cyber-physical integration for future green seaports: Challenges, state of the art and future prospects. *IEEE Trans. Ind.-Cyber-Phys. Syst.* **2023**, *1*, 21–43. [\[CrossRef\]](#)
5. Yang, X.; Ergon, S. Evaluation of visualization techniques for use by facility operators during monitoring tasks. *Autom. Constr.* **2014**, *44*, 103–118. [\[CrossRef\]](#)

6. Tai, J.L.; Hameed Sultan, M.T.; Grzejda, R.; Shahar, F.S. Remote Non-Destructive Testing of Port Cranes: A Review of Vibration and Acoustic Sensors with IoT Integration. *J. Mar. Sci. Eng.* **2025**, *13*, 1338. [\[CrossRef\]](#)
7. Wang, R.; Li, J.; Bai, R. Prediction and Analysis of Container Terminal Logistics Arrival Time Based on Simulation Interactive Modeling: A Case Study of Ningbo Port. *Mathematics* **2023**, *11*, 3271. [\[CrossRef\]](#)
8. Li, Y.; Chang, D.; Gao, Y.; Zou, Y.; Bao, C. Automated Container Terminal Production Operation and Optimization via an AdaBoost-Based Digital Twin Framework. *J. Adv. Transp.* **2021**, *2021*, 1936764. [\[CrossRef\]](#)
9. Yang, W.; Bao, X.; Zheng, Y.; Zhang, L.; Zhang, Z.; Zhang, Z.; Li, L. A digital twin framework for large comprehensive ports and a case study of Qingdao Port. *Int. J. Adv. Manuf. Technol.* **2024**, *131*, 5571–5588. [\[CrossRef\]](#)
10. Aheleroff, S.; Xu, X.; Zhong, R.Y.; Lu, Y. Digital twin as a service (DTaaS) in industry 4.0: An architecture reference model. *Adv. Eng. Inform.* **2021**, *47*, 101225. [\[CrossRef\]](#)
11. Yang, Z.; Aihemaiti, M.; Abudurehman, B.; Tao, H. High-Precision Optimization of BIM-3D GIS Models for Digital Twins: A Case Study of Santun River Basin. *Sensors* **2025**, *25*, 4630. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Khallaf, R.; Khallaf, L.; Anumba, C.J.; Madubuike, O.C. Review of digital twins for constructed facilities. *Buildings* **2022**, *12*, 2029. [\[CrossRef\]](#)
13. Guo, H.J.; Bakdash, J.Z.; Marusich, L.R.; Prabhakaran, B. Augmented reality and mixed reality measurement under different environments: A survey on head-mounted devices. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–15. [\[CrossRef\]](#)
14. Rokhsaritalemi, S.; Sadeghi-Niaraki, A.; Choi, S.M. A review on mixed reality: Current trends, challenges and prospects. *Appl. Sci.* **2020**, *10*, 636. [\[CrossRef\]](#)
15. Xin, J.; Negenborn, R.R.; Corman, F.; Lodewijks, G. Control of interacting machines in automated container terminals using a sequential planning approach for collision avoidance. *Transp. Res. Part Emerg. Technol.* **2015**, *60*, 377–396. [\[CrossRef\]](#)
16. Lee, S.; Kim, S.; Roh, B.h. Mixed reality virtual device (MRVD) for seamless MR-IoT-digital twin convergence. *Internet Things* **2024**, *26*, 101155. [\[CrossRef\]](#)
17. Yang, S.W.; Lee, Y.; Kim, S.A. Design and Validation of a Real-Time Maintenance Monitoring System Using BIM and Digital Twin Integration. *Buildings* **2025**, *15*, 1312. [\[CrossRef\]](#)
18. Gao, Y.; Chang, D.; Chen, C.H. A digital twin-based approach for optimizing operation energy consumption at automated container terminals. *J. Clean. Prod.* **2023**, *385*, 135782. [\[CrossRef\]](#)
19. Chatsuwan, M.; Ichinose, M.; Alkhalaf, H. Enhancing Facility Management with a BIM and IoT Integration Tool and Framework in an Open Standard Environment. *Buildings* **2025**, *15*, 1928. [\[CrossRef\]](#)
20. Chen, H.; Li, S.; Fan, J.; Duan, A.; Yang, C.; Navarro-Alarcon, D.; Zheng, P. Human-in-the-loop robot learning for smart manufacturing: A human-centric perspective. *IEEE Trans. Autom. Sci. Eng.* **2025**. [\[CrossRef\]](#)
21. Choi, S.H.; Park, K.B.; Roh, D.H.; Lee, J.Y.; Mohammed, M.; Ghasemi, Y.; Jeong, H. An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation. *Robot.-Comput.-Integr. Manuf.* **2022**, *73*, 102258. [\[CrossRef\]](#)
22. Teizer, J.; Cheng, T.; Fang, Y. Location tracking and data visualization technology to advance construction ironworkers' education and training in safety and productivity. *Autom. Constr.* **2013**, *35*, 53–68. [\[CrossRef\]](#)
23. Shin, J.H.; Park, S.J.; Kim, M.A.; Lee, M.J.; Lim, S.C.; Cho, K.W. Development of a digital twin pipeline for interactive scientific simulation and mixed reality visualization. *IEEE Access* **2023**, *11*, 100907–100918. [\[CrossRef\]](#)
24. Papadopoulos, T.; Evangelidis, K.; Kaskalis, T.H.; Evangelidis, G.; Sylaiou, S. Interactions in augmented and mixed reality: An overview. *Appl. Sci.* **2021**, *11*, 8752. [\[CrossRef\]](#)
25. Tuhaise, V.V.; Tah, J.H.M.; Abanda, F.H. Technologies for digital twin applications in construction. *Autom. Constr.* **2023**, *152*, 104931. [\[CrossRef\]](#)
26. Yang, A.; Han, M.; Zeng, Q.; Sun, Y. Adopting building information modeling (BIM) for the development of smart buildings: A review of enabling applications and challenges. *Adv. Civ. Eng.* **2021**, *2021*, 8811476. [\[CrossRef\]](#)
27. Hidayat, W.N.; Hakim, O.S.; Sukaridhoto, S.; Zainuddin, M.A.; Prayudi, A.; Arissabarno, C.; Achmad, Z.M.; Budiarti, R.P.N. Digital Twin System for Smart Buildings Integrated with Blockchain and Mixed Reality Technology. In Proceedings of the 2024 IEEE International Symposium on Consumer Technology (ISCT), Kuta, Bali, Indonesia, 13–16 August 2024; IEEE: New York, NY, USA, 2024; pp. 339–345.
28. Li, K.X.; Li, M.; Zhu, Y.; Yuen, K.F.; Tong, H.; Zhou, H. Smart port: A bibliometric review and future research directions. *Transp. Res. Part Logist. Transp. Rev.* **2023**, *174*, 103098. [\[CrossRef\]](#)
29. Zhang, Y.; Bao, X.; Zhang, L.; Chen, L.; Tang, X.; Zhang, Z.; Zheng, Y. Digital Twin enhanced reinforcement learning for integrated scheduling in automated container terminals. In Proceedings of the 2023 IEEE 19th International Conference on Automation Science and Engineering (CASE), Auckland, New Zealand, 26–30 August 2023; IEEE: New York, NY, USA, 2023; pp. 1–6.
30. Tu, X.; Autiosalo, J.; Jadid, A.; Tammi, K.; Klinker, G. A mixed reality interface for a digital twin based crane. *Appl. Sci.* **2021**, *11*, 9480. [\[CrossRef\]](#)

31. Wang, J.; Ma, Q.; Wei, X. The application of extended reality technology in architectural design education: A review. *Buildings* **2023**, *13*, 2931. [[CrossRef](#)]
32. Correia, J.B.; Abel, M.; Becker, K. Data management in digital twins: A systematic literature review. *Knowl. Inf. Syst.* **2023**, *65*, 3165–3196. [[CrossRef](#)]
33. Xie, J.; Liu, S.; Wang, X. Framework for a closed-loop cooperative human Cyber-Physical System for the mining industry driven by VR and AR: MHCPS. *Comput. Ind. Eng.* **2022**, *168*, 108050. [[CrossRef](#)]
34. Park, K.B.; Choi, S.H.; Lee, J.Y.; Ghasemi, Y.; Mohammed, M.; Jeong, H. Hands-free human–robot interaction using multimodal gestures and deep learning in wearable mixed reality. *IEEE Access* **2021**, *9*, 55448–55464. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.